# 5. Relativistic Wave Equations and their Derivation

## 5.1 Introduction

Quantum theory is based on the following axioms[1]:

1. The state of a system is described by a state vector $|\psi\rangle$ in a linear space.
2. The observables are represented by hermitian operators $A$..., and functions of observables by the corresponding functions of the operators.
3. The mean (expectation) value of an observable in the state $|\psi\rangle$ is given by $\langle A\rangle = \langle\psi| A |\psi\rangle$.
4. The time evolution is determined by the Schrödinger equation involving the Hamiltonian $H$

$$\mathrm{i}\hbar\frac{\partial |\psi\rangle}{\partial t} = H |\psi\rangle \ . \tag{5.1.1}$$

5. If, in a measurement of the observable $A$, the value $a_n$ is found, then the original state changes to the corresponding eigenstate $|n\rangle$ of $A$.

We consider the Schrödinger equation for a free particle in the coordinate representation

$$\mathrm{i}\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\boldsymbol{\nabla}^2\psi \ . \tag{5.1.2}$$

It is evident from the differing orders of the time and the space derivatives that this equation is not Lorentz covariant, i.e., that it changes its structure under a transition from one inertial system to another.

Efforts to formulate a relativistic quantum mechanics began with attempts to use the correspondence principle in order to derive a relativistic wave equation intended to replace the Schrödinger equation. The first such equation was due to Schrödinger (1926)[2], Gordon (1926)[3], and Klein (1927)[4]. This scalar wave equation of second order, which is now known as the Klein–Gordon equation, was initially dismissed, since it led to negative

[1] See QM I, Sect. 8.3.
[2] E. Schrödinger, Ann. Physik **81**, 109 (1926)
[3] W. Gordon, Z. Physik **40**, 117 (1926)
[4] O. Klein, Z. Physik **41**, 407 (1927)

probability densities. The year 1928 saw the publication of the Dirac equation[5]. This equation pertains to particles with spin 1/2 and is able to describe many of the single-particle properties of fermions. The Dirac equation, like the Klein–Gordon equation, possesses solutions with negative energy, which, in the framework of wave mechanics, leads to difficulties (see below). To prevent transitions of an electron into lower lying states of negative energy, in 1930[6] Dirac postulated that the states of negative energy should all be occupied. Missing particles in these otherwise occupied states represent particles with opposite charge (antiparticles). This necessarily leads to a many-particle theory, or to a quantum field theory. By reinterpreting the Klein–Gordon equation as the basis of a field theory, Pauli and Weisskopf[7] showed that this could describe mesons with spin zero, e.g., $\pi$ mesons. The field theories based upon the Dirac and Klein–Gordon equations correspond to the Maxwell equations for the electromagnetic field, and the d'Alembert equation for the four-potential.

The Schrödinger equation, as well as the other axioms of quantum theory, remain unchanged. Only the Hamiltonian is changed and now represents a quantized field. The elementary particles are excitations of the fields (mesons, electrons, photons, etc.).

It will be instructive to now follow the historical development rather than begin immediately with quantum field theory. For one thing, it is conceptually easier to investigate the properties of the Dirac equation in its interpretation as a single-particle wave equation. Furthermore, it is exactly these single-particle solutions that are needed as basis states for expanding the field operators. At low energies one can neglect decay processes and thus, here, the quantum field theory gives the same physical predictions as the elementary single-particle theory.

## 5.2 The Klein–Gordon Equation

### 5.2.1 Derivation by Means of the Correspondence Principle

In order to derive relativistic wave equations, we first recall the correspondence principle[8]. When classical quantities were replaced by the operators

$$\text{energy} \qquad E \longrightarrow \mathrm{i}\hbar\frac{\partial}{\partial t}$$

and

[5] P.A.M. Dirac, Proc. Roy. Soc. (London) **A117**, 610 (1928); ibid. **A118**, 351 (1928)
[6] P.A.M. Dirac, Proc. Roy. Soc. (London) **A126**, 360 (1930)
[7] W. Pauli and V. Weisskopf, Helv. Phys. Acta **7**, 709 (1934)
[8] See, e.g., QM I, Sect. 2.5.1

$$\text{momentum} \qquad \mathbf{p} \longrightarrow \frac{\hbar}{\mathrm{i}}\boldsymbol{\nabla} \,, \tag{5.2.1}$$

we obtained from the nonrelativistic energy of a free particle

$$E = \frac{\mathbf{p}^2}{2m} \,, \tag{5.2.2}$$

the free time-dependent *Schrödinger equation*

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\psi = -\frac{\hbar^2\boldsymbol{\nabla}^2}{2m}\psi \,. \tag{5.2.3}$$

This equation is obviously not Lorentz covariant due to the different orders of the time and space derivatives.

We now recall some relevant features of the special theory of relativity.[9] We will use the following conventions: The components of the space–time four-vectors will be denoted by Greek indices, and the components of spatial three-vectors by Latin indices or the cartesian coordinates $x,\ y,\ z$. In addition, we will use Einstein's summation convention: Greek indices that appear twice, one contravariant and one covariant, are summed over, the same applying to corresponding Latin indices.

Starting from $x^\mu(s) = (ct, \mathbf{x})$, the contravariant four-vector representation of the world line as a function of the proper time $s$, one obtains the four-velocity $\dot{x}^\mu(s)$. The differential of the proper time is related to $dx^0$ via $ds = \sqrt{1-(v/c)^2}\, dx^0$, where

$$\mathbf{v} = c\,(d\mathbf{x}/dx^0) \tag{5.2.4a}$$

is the velocity. For the four-momentum this yields:

$$p^\mu = mc\dot{x}^\mu(s) = \frac{1}{\sqrt{1-(v/c)^2}}\begin{pmatrix} mc \\ m\mathbf{v}\end{pmatrix} = \text{four-momentum} = \begin{pmatrix} E/c \\ \mathbf{p}\end{pmatrix} . \tag{5.2.4b}$$

In the last expression we have used the fact that, according to relativistic dynamics, $p^0 = mc/\sqrt{1-(v/c)^2}$ represents the kinetic energy of the particle. Therefore, according to the special theory of relativity, the energy $E$ and the momentum $p_x,\ p_y,\ p_z$ transform as the components of a contravariant four-vector

$$p^\mu = \left(p^0, p^1, p^2, p^3\right) = \left(\frac{E}{c}, p_x, p_y, p_z\right) . \tag{5.2.5a}$$

[9] The most important properties of the Lorentz group will be summarized in Sect. 6.1.

The metric tensor

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{5.2.6}$$

yields the covariant components

$$p_\mu = g_{\mu\nu} p^\nu = \left( \frac{E}{c}, -\mathbf{p} \right) \; . \tag{5.2.5b}$$

According to Eq. (5.2.4b), the invariant scalar product of the four-momentum is given by

$$p_\mu p^\mu = \frac{E^2}{c^2} - \mathbf{p}^2 = m^2 c^2 \; , \tag{5.2.7}$$

with the rest mass $m$ and the velocity of light $c$.

From the energy–momentum relation following from (5.2.7),

$$E = \sqrt{\mathbf{p}^2 c^2 + m^2 c^4} \; , \tag{5.2.8}$$

one would, according to the *correspondence principle* (5.2.1), initially arrive at the following *wave equation*:

$$\mathrm{i}\hbar \frac{\partial}{\partial t} \psi = \sqrt{-\hbar^2 c^2 \boldsymbol{\nabla}^2 + m^2 c^4} \; \psi \; . \tag{5.2.9}$$

An obvious difficulty with this equation lies in the square root of the spatial derivative; its Taylor expansion leads to infinitely high derivatives. Time and space do not occur symmetrically.

Instead, we start from the squared relation:

$$E^2 = \mathbf{p}^2 c^2 + m^2 c^4 \tag{5.2.10}$$

and obtain

$$-\hbar^2 \frac{\partial^2}{\partial t^2} \psi = (-\hbar^2 c^2 \boldsymbol{\nabla}^2 + m^2 c^4) \psi \; . \tag{5.2.11}$$

This equation can be written in the even more compact and clearly Lorentz-covariant form

$$\left( \partial_\mu \partial^\mu + \left( \frac{mc}{\hbar} \right)^2 \right) \psi = 0 \; . \tag{5.2.11$'$}$$

Here $x^\mu$ is the space–time position vector

$$x^\mu = (x^0 = ct, \mathbf{x})$$

and the covariant vector

$$\partial_\mu = \frac{\partial}{\partial x^\mu}$$

is the four-dimensional generalization of the gradient vector. As is known from electrodynamics, the d'Alembert operator $\Box \equiv \partial_\mu \partial^\mu$ is invariant under Lorentz transformations. Also appearing here is the Compton wavelength $\hbar/mc$ of a particle with mass $m$. Equation (5.2.11′) is known as the *Klein–Gordon* equation. It was originally introduced and studied by Schrödinger, and by Gordon and Klein.

We will now investigate the most important properties of the Klein–Gordon equation.

### 5.2.2 The Continuity Equation

To derive a continuity equation one takes $\psi^*$ times (5.2.11′)

$$\psi^* \left( \partial_\mu \partial^\mu + \left( \frac{mc}{\hbar} \right)^2 \right) \psi = 0$$

and subtracts the complex conjugate of this equation

$$\psi \left( \partial_\mu \partial^\mu + \left( \frac{mc}{\hbar} \right)^2 \right) \psi^* = 0 \, .$$

This yields

$$\begin{aligned} \psi^* \partial_\mu \partial^\mu \psi - \psi \partial_\mu \partial^\mu \psi^* &= 0 \\ \partial_\mu (\psi^* \partial^\mu \psi - \psi \partial^\mu \psi^*) &= 0 \, . \end{aligned}$$

Multiplying by $\frac{\hbar}{2m\mathrm{i}}$, so that the current density is equal to that in the non-relativistic case, one obtains

$$\frac{\partial}{\partial t} \left( \frac{\mathrm{i}\hbar}{2mc^2} \left( \psi^* \frac{\partial \psi}{\partial t} - \psi \frac{\partial \psi^*}{\partial t} \right) \right) + \boldsymbol{\nabla} \cdot \frac{\hbar}{2m\mathrm{i}} \left[ \psi^* \boldsymbol{\nabla} \psi - \psi \boldsymbol{\nabla} \psi^* \right] = 0 \, . \tag{5.2.12}$$

This has the form of a *continuity equation*

$$\dot{\rho} + \operatorname{div} \mathbf{j} = 0 \, , \tag{5.2.12′}$$

with density

$$\rho = \frac{\mathrm{i}\hbar}{2mc^2} \left( \psi^* \frac{\partial \psi}{\partial t} - \psi \frac{\partial \psi^*}{\partial t} \right) \tag{5.2.13a}$$

and current density

$$\mathbf{j} = \frac{\hbar}{2m\mathrm{i}} \left( \psi^* \boldsymbol{\nabla} \psi - \psi \boldsymbol{\nabla} \psi^* \right) \, . \tag{5.2.13b}$$

Here, $\rho$ is not positive definite and thus cannot be directly interpreted as a probability density, although $e\rho(\mathbf{x}, t)$ can possibly be conceived as the corresponding charge density. The Klein–Gordon equation is a second-order differential equation in $t$ and thus the initial values of $\psi$ and $\frac{\partial \psi}{\partial t}$ can be chosen independently, so that $\rho$ as a function of $\mathbf{x}$ can be both positive and negative.

### 5.2.3 Free Solutions of the Klein–Gordon Equation

Equation (5.2.11) is known as the free Klein–Gordon equation in order to distinguish it from generalizations that additionally contain external potentials or electromagnetic fields (see Sect. 5.3.5). There are two free solutions in the form of plane waves:

$$\psi(\mathbf{x}, t) = \mathrm{e}^{\mathrm{i}(Et - \mathbf{p}\cdot\mathbf{x})/\hbar} \tag{5.2.14}$$

with

$$E = \pm\sqrt{\mathbf{p}^2 c^2 + m^2 c^4}\ .$$

Both positive and negative energies occur here and the energy is not bounded from below. This scalar theory does not contain spin and could only describe particles with zero spin.

Hence, the Klein–Gordon equation was rejected initially because the primary aim was a theory for the electron. Dirac[5] had instead introduced a first-order differential equation with positive density, as already mentioned at the beginning of this chapter. It will later emerge that this, too, has solutions with negative energies. The unoccupied states of negative energy describe antiparticles. As a quantized field theory, the Klein–Gordon equation describes mesons[7]. The hermitian scalar Klein–Gordon field describes neutral mesons with spin 0. The nonhermitian pseudoscalar Klein–Gordon field describes charged mesons with spin 0 and their antiparticles.

We shall therefore proceed by constructing a wave equation for spin-1/2 fermions and only return to the Klein–Gordon equation in connection with motion in a Coulomb potential ($\pi^-$-mesons).

## 5.3 Dirac Equation

### 5.3.1 Derivation of the Dirac Equation

We will now attempt to find a wave equation of the form

$$\mathrm{i}\hbar\frac{\partial \psi}{\partial t} = \left(\frac{\hbar c}{\mathrm{i}}\alpha^k \partial_k + \beta m c^2\right)\psi \equiv H\psi\ . \tag{5.3.1}$$

Spatial components will be denoted by Latin indices, where repeated indices are to be summed over. The second derivative $\frac{\partial^2}{\partial t^2}$ in the Klein–Gordon equation leads to a density $\rho = \left(\psi^* \frac{\partial}{\partial t}\psi - \text{c.c.}\right)$. In order that the density be positive, we postulate a differential equation of first order. The requirement of relativistic covariance demands that the spatial derivatives may only be of first order, too. The Dirac Hamiltonian $H$ is linear in the momentum operator and in the rest energy. The coefficients in (5.3.1) cannot simply be numbers: if they were, the equation would not even be form invariant (having the same coefficients) with respect to spatial rotations. $\alpha^k$ and $\beta$ must be hermitian matrices in order for $H$ to be hermitian, which is in turn necessary for a positive, conserved probability density to exist. Thus $\alpha^k$ and $\beta$ are $N \times N$ matrices and

$$\psi = \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_N \end{pmatrix} \quad \text{an } N\text{-component column vector .}$$

We shall impose the following requirements on equation (5.3.1):

(i) The components of $\psi$ must satisfy the Klein–Gordon equation so that plane waves fulfil the relativistic energy–momentum relation $E^2 = p^2c^2 + m^2c^4$.

(ii) There exists a conserved four-current whose zeroth component is a positive density.

(iii) The equation must be Lorentz covariant. This means that it has the same form in all reference frames that are connencted by a Poincaré transformation.

The resulting equation (5.3.1) is named, after its discoverer, the *Dirac equation*. We must now look at the consequences that arise from the conditions (i)–(iii). Let us first consider condition (i). The two-fold application of $H$ yields

$$\begin{aligned} -\hbar^2 \frac{\partial^2}{\partial t^2}\psi = -\hbar^2 c^2 \sum_{ij} \frac{1}{2}\left(\alpha^i\alpha^j + \alpha^j\alpha^i\right)\partial_i\partial_j\psi \\ + \frac{\hbar m c^3}{\mathrm{i}} \sum_{i=1}^{3} \left(\alpha^i\beta + \beta\alpha^i\right)\partial_i\psi + \beta^2 m^2 c^4 \psi \,. \end{aligned} \tag{5.3.2}$$

Here, we have made use of $\partial_i\partial_j = \partial_j\partial_i$ to symmetrize the first term on the right-hand side. *Comparison with the Klein–Gordon equation* (5.2.11′) leads to the three conditions

$$\alpha^i\alpha^j + \alpha^j\alpha^i = 2\delta^{ij}\,\mathbb{1}\,, \tag{5.3.3a}$$

$$\alpha^i\beta + \beta\alpha^i = 0\,, \tag{5.3.3b}$$

$$\alpha^{i\,2} = \beta^2 = \mathbb{1}\,. \tag{5.3.3c}$$

### 5.3.2 The Continuity Equation

The row vectors adjoint to $\psi$ are defined by

$$\psi^\dagger = (\psi_1^*, \dots, \psi_N^*) .$$

Multiplying the Dirac equation from the left by $\psi^\dagger$, we obtain

$$\mathrm{i}\hbar\psi^\dagger \frac{\partial\psi}{\partial t} = \frac{\hbar c}{\mathrm{i}}\psi^\dagger\alpha^i\partial_i\psi + mc^2\psi^\dagger\beta\psi . \tag{5.3.4a}$$

The complex conjugate relation reads:

$$-\mathrm{i}\hbar\frac{\partial\psi^\dagger}{\partial t}\psi = -\frac{\hbar c}{\mathrm{i}}\left(\partial_i\psi^\dagger\right)\alpha^{i\dagger}\psi + mc^2\psi^\dagger\beta^\dagger\psi . \tag{5.3.4b}$$

The difference of these two equations yields:

$$\frac{\partial}{\partial t}\left(\psi^\dagger\psi\right) = -c\left(\left(\partial_i\psi^\dagger\right)\alpha^{i\dagger}\psi + \psi^\dagger\alpha^i\partial_i\psi\right) + \frac{imc^2}{\hbar}\left(\psi^\dagger\beta^\dagger\psi - \psi^\dagger\beta\psi\right) . \tag{5.3.5}$$

In order for this to take the form of a continuity equation, the matrices $\alpha$ and $\beta$ must be hermitian, i.e.,

$$\alpha^{i\dagger} = \alpha^i , \quad \beta^\dagger = \beta . \tag{5.3.6}$$

Then the *density*

$$\rho \equiv \psi^\dagger\psi = \sum_{\alpha=1}^{N}\psi_\alpha^*\psi_\alpha \tag{5.3.7a}$$

and the *current density*

$$j^k \equiv c\psi^\dagger\alpha^k\psi \tag{5.3.7b}$$

satisfy the *continuity equation*

$$\frac{\partial}{\partial t}\rho + \operatorname{div}\mathbf{j} = 0 . \tag{5.3.8}$$

With the zeroth component of $j^\mu$,

$$j^0 \equiv c\rho , \tag{5.3.9}$$

we may define a four-current-density

$$j^\mu \equiv (j^0, j^k) \tag{5.3.9$'$}$$

and write the continuity equation in the form

$$\partial_\mu j^\mu = \frac{1}{c}\frac{\partial}{\partial t}j^0 + \frac{\partial}{\partial x^k}j^k = 0 . \tag{5.3.10}$$

The density defined in (5.3.7a) is positive definite and, within the framework of the single particle theory, can be given the preliminary interpretation of a probability density.

### 5.3.3 Properties of the Dirac Matrices

The matrices $\alpha^k$, $\beta$ anticommute and their square is equal to 1; see Eq. (5.3.3a–c). From $(\alpha^k)^2 = \beta^2 = \mathbb{1}$, it follows that the matrices $\alpha^k$ and $\beta$ possess only the eigenvalues $\pm 1$.

We may now write (5.3.3b) in the form

$$\alpha^k = -\beta\alpha^k\beta \, .$$

Using the cyclic invariance of the trace, we obtain

$$\operatorname{Tr}\alpha^k = -\operatorname{Tr}\beta\alpha^k\beta = -\operatorname{Tr}\alpha^k\beta^2 = -\operatorname{Tr}\alpha^k \, .$$

From this, and from an equivalent calculation for $\beta$, one obtains

$$\operatorname{Tr}\alpha^k = \operatorname{Tr}\beta = 0 \, . \tag{5.3.11}$$

Hence, the number of positive and negative eigenvalues must be equal and, therefore, $N$ is even. $N = 2$ is not sufficient since the $2 \times 2$ matrices $\mathbb{1}$, $\sigma_x$, $\sigma_y$, $\sigma_z$ contain only 3 mutually anticommuting matrices. $N = 4$ is the smallest dimension in which it is possible to realize the algebraic structure (5.3.3a–c).

A particular representation of the matrices is

$$\alpha^i = \begin{pmatrix} 0 & \sigma^i \\ \sigma^i & 0 \end{pmatrix} \, , \quad \beta = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \, , \tag{5.3.12}$$

where the $4 \times 4$ matrices are constructed from the Pauli matrices

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \, , \quad \sigma^2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \, , \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{5.3.13}$$

and the two-dimensional unit matrix. It is easy to see that the matrices (5.3.12) satisfy the conditions (5.3.3a–c):

$$\text{e.g.,} \quad \alpha^i\beta + \beta\alpha^i = \begin{pmatrix} 0 & -\sigma^i \\ \sigma^i & 0 \end{pmatrix} + \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} = 0 \, .$$

The Dirac equation (5.3.1), in combination with the matrices (5.3.12), is referred to as the "*standard representation*" of the Dirac equation. One calls $\psi$ a four-spinor or spinor for short (or sometimes a bispinor, in particular when $\psi$ is represented by two two-component spinors). $\psi^\dagger$ is called the *hermitian adjoint* spinor. It will be shown in Sect. 6.2.1 that under Lorentz transformations spinors possess specific transformation properties.

### 5.3.4 The Dirac Equation in Covariant Form

In order to ensure that time and space derivatives are multiplied by matrices with similar algebraic properties, we multiply the Dirac equation (5.3.1) by $\beta/c$ to obtain

$$-\mathrm{i}\hbar\beta\partial_0\psi - \mathrm{i}\hbar\beta\alpha^i\partial_i\psi + mc\psi = 0 \ . \tag{5.3.14}$$

We now define new Dirac matrices

$$\begin{aligned} &\gamma^0 \equiv \beta \\ &\gamma^i \equiv \beta\alpha^i \ . \end{aligned} \tag{5.3.15}$$

These possess the following properties:

$\gamma^0$ is hermitian and $(\gamma^0)^2 = \mathbb{1}$. However, $\gamma^k$ is antihermitian.
$(\gamma^k)^\dagger = -\gamma^k$ and $(\gamma^k)^2 = -\mathbb{1}$.

Proof:

$$\left(\gamma^k\right)^\dagger = \alpha^k\beta = -\beta\alpha^k = -\gamma^k \ ,$$

$$\left(\gamma^k\right)^2 = \beta\alpha^k\beta\alpha^k = -\mathbb{1} \ .$$

These relations, together with

$$\gamma^0\gamma^k + \gamma^k\gamma^0 = \beta\beta\alpha^k + \beta\alpha^k\beta = 0 \qquad \text{and}$$

$$\gamma^k\gamma^l + \gamma^l\gamma^k = \beta\alpha^k\beta\alpha^l + \beta\alpha^l\beta\alpha^k = 0 \quad \text{for } k \neq l$$

lead to the fundamental algebraic structure of the Dirac matrices

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}\mathbb{1} \ . \tag{5.3.16}$$

The *Dirac equation* (5.3.14) now assumes the form

$$\left(-\mathrm{i}\gamma^\mu\partial_\mu + \frac{mc}{\hbar}\right)\psi = 0 \ . \tag{5.3.17}$$

It will be convenient to use the shorthand notation originally introduced by Feynman:

$$\not{v} \equiv \gamma \cdot v \equiv \gamma^\mu v_\mu = \gamma_\mu v^\mu = \gamma^0 v^0 - \boldsymbol{\gamma}\mathbf{v} \ . \tag{5.3.18}$$

Here, $v^\mu$ stands for any vector. The Feynman slash implies scalar multiplication by $\gamma_\mu$. In the fourth term we have introduced the covariant components of the $\gamma$ matrices

$$\gamma_\mu = g_{\mu\nu}\gamma^\nu \ . \tag{5.3.19}$$

In this notation the Dirac equation may be written in the compact form

$$\left(-\mathrm{i}\not{\partial} + \frac{mc}{\hbar}\right)\psi = 0 \ . \tag{5.3.20}$$

Finally, we also give the $\gamma$ matrices in the particular representation (5.3.12). From (5.3.12) and (5.3.15) it follows that

$$\gamma^0 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \; , \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \; . \tag{5.3.21}$$

**Remark.** A representation of the $\gamma$ matrices that is equivalent to (5.3.21) and which also satisfies the algebraic relations (5.3.16) is obtained by replacing

$$\gamma \to M\gamma M^{-1} \; ,$$

where $M$ is an arbitrary nonsingular matrix. Other frequently encountered representations are the Majorana representation and the chiral representation (see Sect. 11.3, Remark (ii) and Eq. (11.6.12a–c)).

### 5.3.5 Nonrelativistic Limit and Coupling to the Electromagnetic Field

#### 5.3.5.1 Particles at Rest

The form (5.3.1) is a particularly suitable starting point when dealing with the nonrelativistic limit. We first consider a free particle *at rest*, i.e., with wave vector $\mathbf{k} = 0$. The spatial derivatives in the Dirac equation then vanish and the equation then simplifies to

$$\mathrm{i}\hbar \frac{\partial \psi}{\partial t} = \beta m c^2 \psi \; . \tag{5.3.17$'$}$$

This equation possesses the following four solutions

$$\begin{aligned} \psi_1^{(+)} &= \mathrm{e}^{-\frac{\mathrm{i}mc^2}{\hbar}t} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \; , \quad \psi_2^{(+)} = \mathrm{e}^{-\frac{\mathrm{i}mc^2}{\hbar}t} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \; , \\ \psi_1^{(-)} &= \mathrm{e}^{\frac{\mathrm{i}mc^2}{\hbar}t} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \; , \quad \psi_2^{(-)} = \mathrm{e}^{\frac{\mathrm{i}mc^2}{\hbar}t} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \; . \end{aligned} \tag{5.3.22}$$

The $\psi_1^{(+)}$, $\psi_2^{(+)}$ and $\psi_1^{(-)}$, $\psi_2^{(-)}$ correspond to positive- and negative-energy solutions, respectively. The interpretation of the negative-energy solutions must be postponed until later. For the moment we will confine ourselves to the positive-energy solutions.

#### 5.3.5.2 Coupling to the Electromagnetic Field

We shall immediately proceed one step further and consider the coupling to an *electromagnetic field*, which will allow us to derive the Pauli equation.

In analogy with the nonrelativistic theory, the canonical momentum $\mathbf{p}$ is replaced by the kinetic momentum $\left(\mathbf{p}-\frac{e}{c}\mathbf{A}\right)$, and the rest energy in the Dirac Hamiltonian is augmented by the scalar electrical potential $e\Phi$,

$$\mathrm{i}\hbar\frac{\partial\psi}{\partial t}=\left(c\boldsymbol{\alpha}\cdot\left(\mathbf{p}-\frac{e}{c}\mathbf{A}\right)+\beta mc^2+e\Phi\right)\psi\,. \tag{5.3.23}$$

Here, $e$ is the charge of the particle, i.e., $e=-e_0$ for the electron. At the end of this section we will arrive at (5.3.23), starting from (5.3.17).

### 5.3.5.3 Nonrelativistic Limit. The Pauli Equation

In order to discuss the *nonrelativistic limit*, we use the explicit representation (5.3.12) of the Dirac matrices and decompose the four-spinors into two two-component column vectors $\tilde{\varphi}$ and $\tilde{\chi}$

$$\psi\equiv\begin{pmatrix}\tilde{\varphi}\\ \tilde{\chi}\end{pmatrix}, \tag{5.3.24}$$

with

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\begin{pmatrix}\tilde{\varphi}\\ \tilde{\chi}\end{pmatrix}=c\begin{pmatrix}\boldsymbol{\sigma}\cdot\boldsymbol{\pi}\,\tilde{\chi}\\ \boldsymbol{\sigma}\cdot\boldsymbol{\pi}\,\tilde{\varphi}\end{pmatrix}+e\Phi\begin{pmatrix}\tilde{\varphi}\\ \tilde{\chi}\end{pmatrix}+mc^2\begin{pmatrix}\tilde{\varphi}\\ -\tilde{\chi}\end{pmatrix}, \tag{5.3.25}$$

where

$$\boldsymbol{\pi}=\mathbf{p}-\frac{e}{c}\mathbf{A} \tag{5.3.26}$$

is the operator of the kinetic momentum.

In the nonrelativistic limit, the rest energy $mc^2$ is the largest energy involved. Thus, to find solutions with positive energy, we write

$$\begin{pmatrix}\tilde{\varphi}\\ \tilde{\chi}\end{pmatrix}=\mathrm{e}^{-\frac{\mathrm{i}mc^2}{\hbar}t}\begin{pmatrix}\varphi\\ \chi\end{pmatrix}, \tag{5.3.27}$$

where $\binom{\varphi}{\chi}$ are considered to vary slowly with time and satisfy the equation

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\begin{pmatrix}\varphi\\ \chi\end{pmatrix}=c\begin{pmatrix}\boldsymbol{\sigma}\cdot\boldsymbol{\pi}\,\chi\\ \boldsymbol{\sigma}\cdot\boldsymbol{\pi}\,\varphi\end{pmatrix}+e\Phi\begin{pmatrix}\varphi\\ \chi\end{pmatrix}-2mc^2\begin{pmatrix}0\\ \chi\end{pmatrix}. \tag{5.3.25$'$}$$

In the second equation, $\hbar\dot{\chi}$ and $e\Phi\chi$ may be neglected in comparison to $2mc^2\chi$, and the latter then solved approximately as

$$\chi=\frac{\boldsymbol{\sigma}\cdot\boldsymbol{\pi}}{2mc}\varphi\,. \tag{5.3.28}$$

From this one sees that, in the nonrelativistic limit, $\chi$ is a factor of order $\sim v/c$ smaller than $\varphi$. One thus refers to $\varphi$ as the large, and $\chi$ as the small, component of the spinor.

Inserting (5.3.28) into the first of the two equations (5.3.25′) yields

$$\mathrm{i}\hbar\frac{\partial\varphi}{\partial t} = \left(\frac{1}{2m}(\boldsymbol{\sigma}\cdot\boldsymbol{\pi})(\boldsymbol{\sigma}\cdot\boldsymbol{\pi}) + e\Phi\right)\varphi\ . \tag{5.3.29}$$

To proceed further we use the identity

$$\boldsymbol{\sigma}\cdot\mathbf{a}\,\boldsymbol{\sigma}\cdot\mathbf{b} = \mathbf{a}\cdot\mathbf{b} + \mathrm{i}\boldsymbol{\sigma}\cdot(\mathbf{a}\times\mathbf{b})\ ,$$

which follows from[10,11] $\sigma^i\sigma^j = \delta_{ij} + \mathrm{i}\epsilon^{ijk}\sigma^k$, which in turn yields:

$$\boldsymbol{\sigma}\cdot\boldsymbol{\pi}\,\boldsymbol{\sigma}\cdot\boldsymbol{\pi} = \boldsymbol{\pi}^2 + \mathrm{i}\boldsymbol{\sigma}\cdot\boldsymbol{\pi}\times\boldsymbol{\pi} = \boldsymbol{\pi}^2 - \frac{e\hbar}{c}\boldsymbol{\sigma}\cdot\mathbf{B}\ .$$

Here, we have used[12]

$$\begin{aligned}(\boldsymbol{\pi}\times\boldsymbol{\pi})^i\varphi &= -\mathrm{i}\hbar\left(\frac{-e}{c}\right)\varepsilon^{ijk}\left(\partial_j A^k - A^k\partial_j\right)\varphi \\ &= \mathrm{i}\frac{\hbar e}{c}\varepsilon^{ijk}\left(\partial_j A^k\right)\varphi = \mathrm{i}\frac{\hbar e}{c}B^i\varphi\end{aligned}$$

with $B^i = \varepsilon^{ijk}\partial_j A^k$. This rearrangement can also be very easily carried out by application of the expression

$$\boldsymbol{\nabla}\times\mathbf{A}\varphi + \mathbf{A}\times\boldsymbol{\nabla}\varphi = \boldsymbol{\nabla}\times\mathbf{A}\,\varphi - \boldsymbol{\nabla}\varphi\times\mathbf{A} = (\boldsymbol{\nabla}\times\mathbf{A})\,\varphi\ .$$

We thus finally obtain

$$\mathrm{i}\hbar\frac{\partial\varphi}{\partial t} = \left[\frac{1}{2m}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)^2 - \frac{e\hbar}{2mc}\boldsymbol{\sigma}\cdot\mathbf{B} + e\Phi\right]\varphi\ . \tag{5.3.29′}$$

This result is identical to the *Pauli equation* for the Pauli spinor $\varphi$, as is known from nonrelativistic quantum mechanics[13]. The two components of $\varphi$ describe the spin of the electron. In addition, one automatically obtains the correct gyromagnetic ratio $g = 2$ for the electron. In order to see this, we simply need to repeat the steps familiar to us from nonrelativistic wave mechanics. We assume a homogeneous magnetic field $\mathbf{B}$ that can be represented by the vector potential $\mathbf{A}$:

---

[10] Here, $\varepsilon^{ijk}$ is the totally antisymmetric tensor of third rank

$$\varepsilon^{ijk} = \begin{cases} 1 \text{ for even permutations of } (123) \\ -1 \text{ for odd permutations of } (123) \\ 0 \text{ otherwise} \end{cases} .$$

[11] QM I, Eq.(9.18a)

[12] Vectors such as $\mathbf{E}$, $\mathbf{B}$ and vector products that are only defined as three-vectors are always written in component form with upper indices; likewise the $\varepsilon$ tensor. Here, too, we sum over repeated indices.

[13] See, e.g., QM I, Chap. 9.

$$\mathbf{B} = \operatorname{curl} \mathbf{A} , \quad \mathbf{A} = \frac{1}{2} \mathbf{B} \times \mathbf{x} . \tag{5.3.30a}$$

Introducing the orbital angular momentum $\mathbf{L}$ and the spin $\mathbf{S}$ as

$$\mathbf{L} = \mathbf{x} \times \mathbf{p} , \quad \mathbf{S} = \frac{1}{2} \hbar \boldsymbol{\sigma} , \tag{5.3.30b}$$

then, for (5.3.30a), it follows[14,15] that

$$\mathrm{i}\hbar \frac{\partial \varphi}{\partial t} = \left( \frac{\mathbf{p}^2}{2m} - \frac{e}{2mc} (\mathbf{L} + 2\mathbf{S}) \cdot \mathbf{B} + \frac{\mathrm{e}^2}{2mc^2} \mathbf{A}^2 + e\Phi \right) \varphi . \tag{5.3.31}$$

The eigenvalues of the projection of the spin operator $\mathbf{S}\hat{\mathbf{e}}$ onto an arbitrary unit vector $\hat{\mathbf{e}}$ are $\pm\hbar/2$. According to (5.3.31), the interaction with the electromagnetic field is of the form

$$H_{\text{int}} = -\boldsymbol{\mu} \cdot \mathbf{B} + \frac{\mathrm{e}^2}{2mc^2} \mathbf{A}^2 + e\Phi , \tag{5.3.32}$$

in which the magnetic moment

$$\boldsymbol{\mu} = \boldsymbol{\mu}_{\text{orbit}} + \boldsymbol{\mu}_{\text{spin}} = \frac{e}{2mc} (\mathbf{L} + 2\mathbf{S}) \tag{5.3.33}$$

is a combination of orbital and spin contributions. The spin moment is of magnitude

$$\boldsymbol{\mu}_{\text{spin}} = g \frac{e}{2mc} \mathbf{S} , \tag{5.3.34}$$

with the gyromagnetic ratio (or Landé factor)

$$g = 2 . \tag{5.3.35}$$

For the electron, $\frac{e}{2mc} = -\frac{\mu_B}{\hbar}$ can be expressed in terms of the Bohr magneton $\mu_B = \frac{e_0 \hbar}{2mc} = 0.927 \times 10^{-20}$erg/G.

We are now in a position to justify the approximations made in this section. The solution $\varphi$ of (5.3.31) has a time behavior that is characterized by the Larmor frequency or, for $e\Phi = \frac{-Ze_0^2}{r}$, by the Rydberg energy (Ry $\propto mc^2\alpha^2$, with the fine structure constant $\alpha = e_0^2/\hbar c$). For the hydrogen and other nonrelativistic atoms (small atomic numbers $Z$), $mc^2$ is very much larger than either of these two energies, thus justifying for such atoms the approximation introduced previously in the equation of motion for $\chi$.

---

[14] See, e.g., QM I, Chap. 9.

[15] One finds $-\mathbf{p} \cdot \mathbf{A} - \mathbf{A} \cdot \mathbf{p} = -2\mathbf{A} \cdot \mathbf{p} = -2\frac{1}{2} (\mathbf{B} \times \mathbf{x}) \cdot \mathbf{p} = -(\mathbf{x} \times \mathbf{p}) \cdot \mathbf{B} = -\mathbf{L} \cdot \mathbf{B}$, since $(\mathbf{p} \cdot \mathbf{A}) = \frac{\hbar}{\mathrm{i}} (\boldsymbol{\nabla} \cdot \mathbf{A}) = 0$.

#### 5.3.5.4 Supplement Concerning Coupling to an Electromagnetic Field

We wish now to use a different approach to derive the Dirac equation in an external field and, to facilitate this, we begin with a few remarks on relativistic notation. The *momentum operator* in covariant and contravariant form reads:

$$p_\mu = \mathrm{i}\hbar\partial_\mu \quad \text{and} \quad p^\mu = \mathrm{i}\hbar\partial^\mu \, . \tag{5.3.36}$$

Here, $\partial_\mu = \frac{\partial}{\partial x^\mu}$ and $\partial^\mu = \frac{\partial}{\partial x_\mu}$. For the time and space components, this implies

$$p^0 = p_0 = \mathrm{i}\hbar\frac{\partial}{\partial ct} \, , \quad p^1 = -p_1 = \mathrm{i}\hbar\frac{\partial}{\partial x_1} = \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x^1} \, . \tag{5.3.37}$$

The coupling to the electromagnetic field is achieved by making the replacement

$$p_\mu \to p_\mu - \frac{e}{c}A_\mu \, , \tag{5.3.38}$$

where $A^\mu = (\Phi, \mathbf{A})$ is the four-potential. The structure which arises here is well known from electrodynamics and, since its generalization to other gauge theories, is termed *minimal coupling*.

This implies

$$\mathrm{i}\hbar\frac{\partial}{\partial x^\mu} \to \mathrm{i}\hbar\frac{\partial}{\partial x^\mu} - \frac{e}{c}A_\mu \tag{5.3.39}$$

which explicitly written in components reads:

$$\begin{cases} \mathrm{i}\hbar\dfrac{\partial}{\partial t} \to \mathrm{i}\hbar\dfrac{\partial}{\partial t} - e\Phi \\ \dfrac{\hbar}{\mathrm{i}}\dfrac{\partial}{\partial x^i} \to \dfrac{\hbar}{\mathrm{i}}\dfrac{\partial}{\partial x^i} + \dfrac{e}{c}A_i = \dfrac{\hbar}{\mathrm{i}}\dfrac{\partial}{\partial x^i} - \dfrac{e}{c}A^i \, . \end{cases} \tag{5.3.39$'$}$$

For the spatial components this is identical to the replacement $\frac{\hbar}{\mathrm{i}}\boldsymbol{\nabla} \to \frac{\hbar}{\mathrm{i}}\boldsymbol{\nabla} - \frac{e}{c}\mathbf{A}$ or $\mathbf{p} \to \mathbf{p} - \frac{e}{c}\mathbf{A}$. In the noncovariant representation of the Dirac equation, the substitution (5.3.39′) immediately leads once again to (5.3.23).

If one inserts (5.3.39) into the Dirac equation (5.3.17), one obtains

$$\left(-\gamma^\mu\left(\mathrm{i}\hbar\partial_\mu - \frac{e}{c}A_\mu\right) + mc\right)\psi = 0 \, , \tag{5.3.40}$$

which is the Dirac equation in relativistic covariant form in the presence of an electromagnetic field.

**Remarks:**

(i) Equation (5.3.23) follows directly when one multiplies (5.3.40), i.e.

$$\gamma^0 \left(\mathrm{i}\hbar\partial_0 - \frac{e}{c}A_0\right)\psi = -\gamma^i \left(\mathrm{i}\hbar\partial_i - \frac{e}{c}A_i\right)\psi + mc\psi$$

by $\gamma^0$:

$$\begin{aligned}
\mathrm{i}\hbar\partial_0\psi &= \alpha^i\left(-\mathrm{i}\hbar\partial_i - \frac{e}{c}A^i\right)\psi + \frac{e}{c}A_0\psi + mc\beta\psi \\
\mathrm{i}\hbar\frac{\partial}{\partial t}\psi &= c\,\boldsymbol{\alpha}\cdot\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)\psi + e\Phi\psi + mc^2\beta\psi\ .
\end{aligned}$$

(ii) The minimal coupling, i.e., the replacement of derivatives by derivatives minus four-potentials, has as a consequence the invariance of the Dirac equation (5.3.40) with respect to gauge transformations (of the first kind):

$$\psi(x) \to \mathrm{e}^{-\mathrm{i}\frac{e}{\hbar c}\alpha(x)}\psi(x)\ , \qquad A_\mu(x) \to A_\mu(x) + \partial_\mu\alpha(x)\ .$$

(iii) For electrons, $m = m_{\mathrm{e}}$, and the characteristic length in the Dirac equation equals the Compton wavelength of the electron

$$\lambda\!\!\!^{-}_c = \frac{\hbar}{m_{\mathrm{e}}c} = 3.8\times 10^{-11}\mathrm{cm}\ .$$

## Problems

**5.1** Show that the matrices (5.3.12) obey the algebraic relations (5.3.3a–c).

**5.2** Show that the representation (5.3.21) follows from (5.3.12).

**5.3** Particles in a homogeneous magnetic field.
Determine the energy levels that result from the Dirac equation for a (relativistic) particle of mass $m$ and charge $e$ in a homogeneous magnetic field $\mathbf{B}$. Use the gauge $A^0 = A^1 = A^3 = 0,\ A^2 = Bx$.

# 6. Lorentz Transformations and Covariance of the Dirac Equation

In this chapter, we shall investigate how the Lorentz covariance of the Dirac equation determines the transformation properties of spinors under Lorentz transformations. We begin by summarizing a few properties of Lorentz transformations, with which the reader is assumed to be familiar. The reader who is principally interested in the solution of specific problems may wish to omit the next sections and proceed directly to Sect. 6.3 and the subsequent chapters.

## 6.1 Lorentz Transformations

The contravariant and covariant components of the position vector read:

$$\begin{array}{lllllll} x^\mu & : & x^0 = ct\ , & x^1 = x\ , & x^2 = y\ , & x^3 = z & \text{contravariant} \\ x_\mu & : & x_0 = ct\ , & x_1 = -x\ , & x_2 = -y\ , & x_3 = -z & \text{covariant}\ . \end{array} \tag{6.1.1}$$

The metric tensor is defined by

$$g = (g_{\mu\nu}) = (g^{\mu\nu}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{6.1.2a}$$

and relates covariant and contravariant components

$$x_\mu = g_{\mu\nu} x^\nu\ , \quad x^\mu = g^{\mu\nu} x_\nu\ . \tag{6.1.3}$$

Furthermore, we note that

$$g^\mu{}_\nu = g^{\mu\sigma} g_{\sigma\nu} \equiv \delta^\mu{}_\nu\ , \tag{6.1.2b}$$

i.e.,

$$(g^\mu{}_\nu) = (\delta^\mu{}_\nu) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\ .$$

The d'Alembert operator is defined by

$$\Box = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \sum_{i=1}^{3} \frac{\partial^2}{\partial x^{i\,2}} = \partial_\mu \partial^\mu = g_{\mu\nu}\partial^\mu\partial^\nu \ . \tag{6.1.4}$$

*Inertial frames* are frames of reference in which, in the absence of forces, particles move uniformly. Lorentz transformations tell us how the coordinates of two inertial frames transform into one another.

The coordinates of two reference systems in uniform motion must be related to one another by a linear transformation. Thus, the inhomogeneous Lorentz transformations (also known as Poincaré transformations) possess the form

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu \ , \tag{6.1.5}$$

where $\Lambda^\mu{}_\nu$ and $a^\mu$ are real.

Remarks:

(i) *On the linearity of the Lorentz transformation*:
Suppose that $x'$ and $x$ are the coordinates of an event in the inertial frames $I'$ and $I$, respectively. For the transformation one could write

$$x' = f(x) \ .$$

In the absence of forces, particles in $I$ and $I'$ move uniformly, i.e., their world lines are straight lines (this is actually the definition of an inertial frame). Transformations under which straight lines are mapped onto straight lines are affinities, and thus of the form (6.1.5). The parametric representation of the equation of a straight line $x^\mu = e^\mu s + d^\mu$ is mapped by such an affine transformation onto another equation for a straight line.

(ii) *Principle of relativity*: The laws of nature are the same in all inertial frames. There is no such thing as an "absolute" frame of reference. The *requirement that the d'Alembert operator be invariant* (6.1.4) yields

$$\Lambda^\lambda{}_\mu g^{\mu\nu} \Lambda^\rho{}_\nu = g^{\lambda\rho} \ , \tag{6.1.6a}$$

or, in matrix form,

$$\Lambda g \Lambda^T = g \ . \tag{6.1.6b}$$

*Proof:* $\partial_\mu \equiv \dfrac{\partial}{\partial x^\mu} = \dfrac{\partial x'^\lambda}{\partial x^\mu}\dfrac{\partial}{\partial x'^\lambda} = \Lambda^\lambda{}_\mu \partial'_\lambda$

$$\begin{aligned} \partial_\mu g^{\mu\nu}\partial_\nu &= \Lambda^\lambda{}_\mu \partial'_\lambda g^{\mu\nu} \Lambda^\rho{}_\nu \partial'_\rho \stackrel{!}{=} \partial'_\lambda g^{\lambda\rho}\partial'_\rho \\ &\Rightarrow \Lambda^\lambda{}_\mu g^{\mu\nu} \Lambda^\rho{}_\nu = g^{\lambda\rho} \ . \end{aligned}$$

The relations (6.1.6a,b) define the Lorentz transformations.
Definition: Poincaré group $\equiv$ {inhomogeneous Lorentz transformation, $a^\mu \neq 0$}
The group of homogeneous Lorentz transformations contains all elements with $a^\mu = 0$.
A homogeneous Lorentz transformation can be denoted by the shorthand form $(\Lambda, a)$, e.g.,

| | |
|---|---|
| translation group | $(1, a)$ |
| rotation group | $(D, 0)$ . |

From the defining equation (6.1.6a,b) follow two important characteristics of Lorentz transformations:
(i) From the definition (6.1.6a), it follows that $(\det \Lambda)^2 = 1$, thus

$$\det \Lambda = \pm 1 \ . \tag{6.1.7}$$

(ii) Consider now the matrix element $\lambda = 0,\ \rho = 0$ of the defining equation (6.1.6a)

$$\Lambda^0{}_\mu g^{\mu\nu} \Lambda^0{}_\nu = 1 = (\Lambda^0{}_0)^2 - \sum_k (\Lambda^0{}_k)^2 = 1 \ .$$

This leads to

$$\Lambda^0{}_0 \geq 1 \qquad \text{or} \qquad \Lambda^0{}_0 \leq -1 \ . \tag{6.1.8}$$

The sign of the determinant of $\Lambda$ and the sign of $\Lambda^0{}_0$ can be used to classify the elements of the Lorentz group (Table 6.1). The Lorentz transformations can be combined as follows into the Lorentz group $\mathcal{L}$, and its subgroups or subsets (e.g., $\mathcal{L}^\downarrow_+$ means the set of all elements $L^\downarrow_+$):

**Table 6.1.** Classification of the elements of the Lorentz group

| | | sgn $\Lambda^0{}_0$ | det $\Lambda$ |
|---|---|---|---|
| proper orthochronous | $L^\uparrow_+$ | 1 | 1 |
| improper orthochronous* | $L^\uparrow_-$ | 1 | $-1$ |
| time-reflection type** | $L^\downarrow_-$ | $-1$ | $-1$ |
| space–time inversion type*** | $L^\downarrow_+$ | $-1$ | 1 |

* spatial reflection ** time reflection *** space–time inversion

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad PT = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{6.1.9}$$

| | |
|---|---|
| $\mathcal{L}$ | Lorentz group (L.G.) |
| $\mathcal{L}_+^\uparrow$ | restricted L.G. (is an invariant subgroup) |
| $\mathcal{L}^\uparrow = \mathcal{L}_+^\uparrow \cup \mathcal{L}_-^\uparrow$ | orthochronous L.G. |
| $\mathcal{L}_+ = \mathcal{L}_+^\uparrow \cup \mathcal{L}_+^\downarrow$ | proper L.G. |
| $\mathcal{L}_0 = \mathcal{L}_+^\uparrow \cup \mathcal{L}_-^\downarrow$ | orthochronous L.G. |

$$\mathcal{L}_-^\uparrow = P \cdot \mathcal{L}_+^\uparrow$$
$$\mathcal{L}_-^\downarrow = T \cdot \mathcal{L}_+^\uparrow$$
$$\mathcal{L}_+^\downarrow = P \cdot T \cdot \mathcal{L}_+^\uparrow$$

The last three subsets of $\mathcal{L}$ do not constitute subgroups.

$$\mathcal{L} = \mathcal{L}^\uparrow \cup T\mathcal{L}^\uparrow = \mathcal{L}_+^\uparrow \cup P\mathcal{L}_+^\uparrow \cup T\mathcal{L}_+^\uparrow \cup PT\mathcal{L}_+^\uparrow \tag{6.1.10}$$

$\mathcal{L}^\uparrow$ is an invariant subgroup of $\mathcal{L}$; $T\mathcal{L}^\uparrow$ is a coset to $\mathcal{L}^\uparrow$.
$\mathcal{L}_+^\uparrow$ is an invariant subgroup of $\mathcal{L}$; $P\mathcal{L}_+^\uparrow$, $T\mathcal{L}_+^\uparrow$, $PT\mathcal{L}_+^\uparrow$ are cosets of $\mathcal{L}$ with respect to $\mathcal{L}_+^\uparrow$. Furthermore, $\mathcal{L}^\uparrow$, $\mathcal{L}_+$, and $\mathcal{L}_0$ are invariant subgroups of $\mathcal{L}$ with the factor groups $(E, P)$, $(E, P, T, PT)$, and $(E, T)$.
Every Lorentz transformation is either proper and orthochronous or can be written as the product of an element of the proper-orthochronous Lorentz group with one of the discrete transformations $P$, $T$, or $PT$.

$\mathcal{L}_+^\uparrow$, the *restricted Lorentz group* = the *proper orthochronous L.G.* consists of all elements with $\det \Lambda = 1$ and $\Lambda^0{}_0 \geq 1$; this includes:

(a) Rotations
(b) Pure Lorentz transformations (= transformations under which space and time are transformed). The prototype is a Lorentz transformation in the $x^1$ direction

$$L_1(\eta) = \begin{pmatrix} L^0{}_0 & L^0{}_1 & 0 & 0 \\ L^1{}_0 & L^1{}_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cosh\eta & -\sinh\eta & 0 & 0 \\ -\sinh\eta & \cosh\eta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{\sqrt{1-\beta^2}} & -\frac{\beta}{\sqrt{1-\beta^2}} & 0 & 0 \\ -\frac{\beta}{\sqrt{1-\beta^2}} & \frac{1}{\sqrt{1-\beta^2}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{6.1.11}$$

with $\tanh\eta = \beta$. For this Lorentz transformation the inertial frame $I'$ moves with respect to $I$ with a velocity $v = c\beta$ in the $x^1$ direction.

## 6.2 Lorentz Covariance of the Dirac Equation

### 6.2.1 Lorentz Covariance and Transformation of Spinors

The *principle of relativity* states that the laws of nature are identical in every inertial reference frame.

We consider two inertial frames $I$ and $I'$ with the space–time coordinates $x$ and $x'$. Let the wave function of a particle in these two frames be $\psi$ and $\psi'$, respectively. We write the Poincaré transformation between $I$ and $I'$ as

$$x' = \Lambda x + a \; . \tag{6.2.1}$$

It must be possible to construct the wave function $\psi'$ from $\psi$. This means that there must be a local relationship between $\psi'$ and $\psi$:

$$\psi'(x') = F(\psi(x)) = F(\psi(\Lambda^{-1}(x' - a)) \; . \tag{6.2.2}$$

The principle of relativity together with the functional relation (6.2.2) necessarily leads to the requirement of *Lorentz covariance*: The Dirac equation in $I$ is transformed by (6.2.1) and (6.2.2) into a Dirac equation in $I'$. (The Dirac equation is form invariant with respect to Poincaré transformations.) In order that both $\psi$ and $\psi'$ may satisfy the linear Dirac equation, their functional relationship must be linear, i.e.,

$$\psi'(x') = S(\Lambda)\psi(x) = S(\Lambda)\psi(\Lambda^{-1}(x' - a)) \; . \tag{6.2.3}$$

Here, $S(\Lambda)$ is a $4 \times 4$ matrix, with which the spinor $\psi$ is to be multiplied. We will determine $S(\Lambda)$ below. In components, the transformation reads:

$$\psi'_\alpha(x') = \sum_{\beta=1}^{4} S_{\alpha\beta}(\Lambda)\psi_\beta(\Lambda^{-1}(x' - a)) \; . \tag{6.2.3$'$}$$

The Lorentz covariance of the Dirac equation requires that $\psi'$ obey the equation

$$(-\mathrm{i}\gamma^\mu \partial'_\mu + m)\psi'(x') = 0 \; , \qquad (c = 1,\ \hbar = 1) \tag{6.2.4}$$

where

$$\partial'_\mu = \frac{\partial}{\partial x'^\mu} \; .$$

The $\gamma$ matrices are unchanged under the Lorentz transformation. In order to determine $S$, we need to convert the Dirac equation in the primed and unprimed coordinate systems into one another. The Dirac equation in the unprimed coordinate system

$$(-\mathrm{i}\gamma^\mu \partial_\mu + m)\psi(x) = 0 \tag{6.2.5}$$

can, by means of the relation

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

and

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

be brought into the form

$$(-\mathrm{i}\gamma^\mu \Lambda^\nu{}_\mu \partial'_\nu + m) S^{-1}(\Lambda)\psi'(x') = 0 \; . \tag{6.2.6}$$

After multiplying from the left by $S$, one obtains[1]

$$-\mathrm{i}S\Lambda^\nu{}_\mu \gamma^\mu S^{-1}\partial'_\nu \psi'(x') + m\psi'(x') = 0 \; . \tag{6.2.6$'$}$$

From a comparison of (6.2.6′) with (6.2.4), it follows that the Dirac equation is form invariant under Lorentz transformations, provided $S(\Lambda)$ satisfies the following condition:

$$S(\Lambda)^{-1}\gamma^\nu S(\Lambda) = \Lambda^\nu{}_\mu \gamma^\mu \; . \tag{6.2.7}$$

It is possible to show (see next section) that this equation has nonsingular solutions for $S(\Lambda)$.[2] A wave function that transforms under a Lorentz transformation according to $\psi' = S\psi$ is known as a *four-component Lorentz spinor*.

### 6.2.2 Determination of the Representation $S(\Lambda)$

#### 6.2.2.1 Infinitesimal Lorentz Transformations

We first consider *infinitesimal (proper, orthochronous) Lorentz transformations*

$$\Lambda^\nu{}_\mu = g^\nu{}_\mu + \Delta\omega^\nu{}_\mu \tag{6.2.8a}$$

with infinitesimal and antisymmetric $\Delta\omega^{\nu\mu}$

$$\Delta\omega^{\nu\mu} = -\Delta\omega^{\mu\nu} \; . \tag{6.2.8b}$$

This equation implies that $\Delta\omega^{\nu\mu}$ can have only 6 independent nonvanishing elements.

---

[1] We recall here that the $\Lambda^\nu{}_\mu$ are matrix elements that, of course, commute with the $\gamma$ matrices.

[2] The existence of such an $S(\Lambda)$ follows from the fact that the matrices $\Lambda^\mu_\nu \gamma^\nu$ obey the same anticommutation rules (5.3.16) as the $\gamma^\mu$ by virtue of (6.1.6a), and from Pauli's fundamental theorem (property 7 on page 146). These transformations will be determined explicitly below.

These transformations satisfy the defining relation for Lorentz transformations

$$\Lambda^{\lambda}{}_{\mu} g^{\mu\nu} \Lambda^{\rho}{}_{\nu} = g^{\lambda\rho} \,, \tag{6.1.6a}$$

as can be seen by inserting (6.2.8) into this equation:

$$g^{\lambda}{}_{\mu} g^{\mu\nu} g^{\rho}{}_{\nu} + \Delta\omega^{\lambda\rho} + \Delta\omega^{\rho\lambda} + O\left((\Delta\omega)^2\right) = g^{\lambda\rho} \,. \tag{6.2.9}$$

Each of the 6 independent elements of $\Delta\omega^{\mu\nu}$ generates an infinitesimal Lorentz transformation. We consider some typical special cases:

$\Delta\omega^{0}{}_{1} = -\Delta\omega^{01} = -\Delta\beta$ : Transformation onto a coordinate system moving with velocity $c\Delta\beta$ in the $x$ direction (6.2.10)

$\Delta\omega^{1}{}_{2} = -\Delta\omega^{12} = \Delta\varphi$ : Transformation onto a coordinate system that is rotated by an angle $\Delta\varphi$ about the $z$ axis. (See Fig. 6.1) (6.2.11)

The spatial components are transformed under this *passive* transformation as follows:

$$\begin{aligned} x'^1 &= x^1 + \Delta\varphi x^2 \\ x'^2 &= -\Delta\varphi x^1 + x^2 \\ x'^3 &= x^3 \end{aligned} \quad \text{or} \quad \mathbf{x}' = \mathbf{x} + \begin{pmatrix} 0 \\ 0 \\ -\Delta\varphi \end{pmatrix} \times \mathbf{x} = \mathbf{x} + \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ 0 & 0 & -\Delta\varphi \\ x^1 & x^2 & x^3 \end{vmatrix} \tag{6.2.12}$$

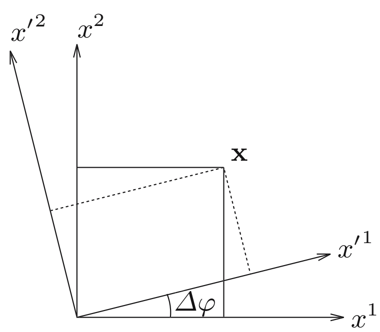



**Fig. 6.1.** Infinitesimal rotation, passive transformation

It must be possible to expand $S$ as a power series in $\Delta\omega^{\nu\mu}$. We write

$$S = 1\!\!1 + \tau \,, \quad S^{-1} = 1\!\!1 - \tau \,, \tag{6.2.13}$$

where $\tau$ is likewise infinitesimal i.e. of order $O(\Delta\omega^{\nu\mu})$. We insert (6.2.13) into the equation for $S$, namely $S^{-1}\gamma^{\mu}S = \Lambda^{\mu}{}_{\nu}\gamma^{\nu}$, and get

$$(1\!\!1 - \tau)\gamma^{\mu}(1\!\!1 + \tau) = \gamma^{\mu} + \gamma^{\mu}\tau - \tau\gamma^{\mu} = \gamma^{\mu} + \Delta\omega^{\mu}{}_{\nu}\gamma^{\nu} \,, \tag{6.2.14}$$

from which the equation determining $\tau$ follows as

$$\gamma^\mu \tau - \tau \gamma^\mu = \Delta\omega^\mu{}_\nu \gamma^\nu \ . \tag{6.2.14$'$}$$

To within an additive multiple of $\mathbb{1}$, this unambiguously determines $\tau$. Given two solutions of (6.2.14$'$), the difference between them has to commute with all $\gamma^\mu$ and thus is proportional to $\mathbb{1}$ (see Sect. 6.2.5, Property 6). The normalization condition $\det S = 1$ removes this ambiquity, since it implies to first order in $\Delta\omega^{\mu\nu}$ that

$$\det S = \det(\mathbb{1} + \tau) = \det \mathbb{1} + \operatorname{Tr} \tau = 1 + \operatorname{Tr} \tau = 1 \ . \tag{6.2.15}$$

It thus follows that

$$\operatorname{Tr} \tau = 0 \ . \tag{6.2.16}$$

Equations (6.2.14$'$) and (6.2.16) have the solution

$$\tau = \frac{1}{8} \Delta\omega^{\mu\nu} (\gamma_\mu \gamma_\nu - \gamma_\nu \gamma_\mu) = -\frac{\mathrm{i}}{4} \Delta\omega^{\mu\nu} \sigma_{\mu\nu} \ , \tag{6.2.17}$$

where we have introduced the definition

$$\sigma_{\mu\nu} = \frac{\mathrm{i}}{2} [\gamma_\mu, \gamma_\nu] \ . \tag{6.2.18}$$

Equation (6.2.17) can be derived by calculating the commutator of $\tau$ with $\gamma^\mu$; the vanishing of the trace is guaranteed by the general properties of the $\gamma$ matrices (Property 3, Sect. 6.2.5).

#### 6.2.2.2 Rotation About the z Axis

We first consider the rotation $R_3$ about the $z$ axis as given by (6.2.11). According to (6.2.11) and (6.2.17),

$$\tau(R_3) = \frac{\mathrm{i}}{2} \Delta\varphi \sigma_{12} \quad ,$$

and with

$$\sigma^{12} = \sigma_{12} = \frac{\mathrm{i}}{2} [\gamma_1, \gamma_2] = \mathrm{i} \gamma_1 \gamma_2 = \mathrm{i} \begin{pmatrix} 0 & \sigma^1 \\ -\sigma^1 & 0 \end{pmatrix} \begin{pmatrix} 0 & \sigma^2 \\ -\sigma^2 & 0 \end{pmatrix} = \begin{pmatrix} \sigma^3 & 0 \\ 0 & \sigma^3 \end{pmatrix} \tag{6.2.19}$$

it follows that

$$S = 1 + \frac{\mathrm{i}}{2} \Delta\varphi \sigma^{12} = 1 + \frac{\mathrm{i}}{2} \Delta\varphi \begin{pmatrix} \sigma^3 & 0 \\ 0 & \sigma^3 \end{pmatrix} \ . \tag{6.2.20}$$

By a succession of infinitestimal rotations we can construct the transformation matrix $S$ for a *finite rotation* through an angle $\vartheta$. This is achieved by decomposing the finite rotation into a sequence of $N$ steps $\vartheta/N$

$$\begin{aligned}\psi'(x') = S\psi(x) &= \lim_{N\to\infty}\left(1+\frac{\mathrm{i}}{2N}\vartheta\sigma^{12}\right)^N \psi(x)\\ &= \mathrm{e}^{\frac{\mathrm{i}}{2}\vartheta\sigma^{12}}\psi\\ &= \left(\cos\frac{\vartheta}{2}+\mathrm{i}\sigma^{12}\sin\frac{\vartheta}{2}\right)\psi(x)\ .\end{aligned} \tag{6.2.21}$$

For the coordinates and other four-vectors, this succession of transformations implies that

$$\begin{aligned}x' &= \lim_{N\to\infty}\left(\mathbb{1}+\frac{\vartheta}{N}\begin{pmatrix}0&0&0&0\\0&0&1&0\\0&-1&0&0\\0&0&0&0\end{pmatrix}\right)\cdots\left(\mathbb{1}+\frac{\vartheta}{N}\begin{pmatrix}0&0&0&0\\0&0&1&0\\0&-1&0&0\\0&0&0&0\end{pmatrix}\right)x\\ &= \exp\left\{\vartheta\begin{pmatrix}0&0&0&0\\0&0&1&0\\0&-1&0&0\\0&0&0&0\end{pmatrix}\right\}x = \begin{pmatrix}1&0&0&0\\0&\cos\vartheta&\sin\vartheta&0\\0&-\sin\vartheta&\cos\vartheta&0\\0&0&0&1\end{pmatrix}x\ ,\end{aligned} \tag{6.2.22}$$

and is thus identical to the usual rotation matrix for rotation through an angle $\vartheta$. The transformation $S$ for rotations (6.2.21) is *unitary* ($S^{-1}=S^{\dagger}$). From (6.2.21), one sees that

$$S(2\pi) = -\mathbb{1} \tag{6.2.23a}$$

$$S(4\pi) = \mathbb{1}\ . \tag{6.2.23b}$$

This means that spinors do not regain their initial value after a rotation through $2\pi$, but only after a rotation through $4\pi$, a fact that is also confirmed by neutron scattering experiments[3]. We draw attention here to the analogy with the transformation of Pauli spinors with respect to rotations:

$$\varphi'(x') = \mathrm{e}^{\frac{\mathrm{i}}{2}\boldsymbol{\omega}\cdot\boldsymbol{\sigma}}\varphi(x)\ . \tag{6.2.24}$$

#### 6.2.2.3 Lorentz Transformation Along the $\mathbf{x^1}$ Direction

According to (6.2.10),

$$\Delta\omega^{01} = \Delta\beta \tag{6.2.25}$$

and (6.2.17) becomes

$$\tau(L_1) = \frac{1}{2}\Delta\beta\gamma_0\gamma_1 = \frac{1}{2}\Delta\beta\alpha_1\ . \tag{6.2.26}$$

We may now determine $S$ for a finite Lorentz transformation along the $x^1$ axis. For the velocity $\frac{v}{c}$, we have $\tanh\eta = \frac{v}{c}$.

---

[3] H. Rauch et al., Phys. Lett. **54A**, 425 (1975); S.A. Werner et al., Phys. Rev. Lett. **35**, 1053 (1975); also described in J.J. Sakurai, *Modern Quantum Mechanics*, p.162, Addison-Wesley, Red Wood City (1985).

The decomposition of $\eta$ into $N$ steps of $\frac{\eta}{N}$ leads to the following transformation of the coordinates and other four-vectors:

$$
\begin{aligned}
x'^{\mu} &= \lim_{N\to\infty}\left(g+\frac{\eta}{N}I\right)^{\mu}{}_{\nu_1}\left(g+\frac{\eta}{N}I\right)^{\nu_1}{}_{\nu_2}\cdots\left(g+\frac{\eta}{N}I\right)^{\nu_{N-1}}{}_{\nu}x^{\nu}\\
g^{\mu}{}_{\nu} &= \delta^{\mu}{}_{\nu}\ ,\\
I^{\nu}{}_{\mu} &= \begin{pmatrix} 0 & -1 & 0 & 0\\ -1 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}\ ,\quad I^2=\begin{pmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}\ ,\quad I^3=I\\
x' &= \mathrm{e}^{\eta I}x=\left(1+\eta I+\frac{1}{2!}\eta^2 I^2+\frac{1}{3!}\eta^3 I+\frac{1}{4!}I^2\dots\right)x\\
x'^{\mu} &= \left(1-I^2+I^2\cosh\eta+I\sinh\eta\right)^{\mu}{}_{\nu}x^{\nu}\\
&= \begin{pmatrix} \cosh\eta & -\sinh\eta & 0 & 0\\ -\sinh\eta & \cosh\eta & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix} x^0\\ x^1\\ x^2\\ x^3\end{pmatrix}\ .
\end{aligned}
\tag{6.2.27}
$$

The $N$-fold application of the infinitesimal Lorentz transformation

$$L_1\left(\frac{\eta}{N}\right)=\mathbb{1}+\frac{\eta}{N}I$$

then leads, in the limit of large $N$, to the Lorentz transformation (6.1.11)

$$L_1(\eta)=\mathrm{e}^{\eta I}=\begin{pmatrix} \cosh\eta & -\sinh\eta & 0 & 0\\ -\sinh\eta & \cosh\eta & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\ . \tag{6.2.27$'$}$$

We note that the $N$ infinitesimal steps of $\frac{\eta}{N}$ add up to $\eta$. However, this does not imply a simple addition of velocities.

We now calculate the corresponding spinor transformation

$$
\begin{aligned}
S(L_1) &= \lim_{N\to\infty}\left(1+\frac{1}{2}\frac{\eta}{N}\alpha_1\right)^N=\mathrm{e}^{\frac{\eta}{2}\alpha_1}\\
&= \mathbb{1}\cosh\frac{\eta}{2}+\alpha_1\sinh\frac{\eta}{2}\quad .
\end{aligned}
\tag{6.2.28}
$$

For homogenous restricted Lorentz transformations, $S$ is *hermitian* ($S(L_1)^{\dagger}=S(L_1)$).

For *general infinitesimal transformations*, characterized by infinitesimal antisymmetric $\Delta\omega^{\mu\nu}$, equation (6.2.17) implies that

$$S(\Lambda)=\mathbb{1}-\frac{\mathrm{i}}{4}\sigma_{\mu\nu}\Delta\omega^{\mu\nu}\ . \tag{6.2.29a}$$

This yields the finite transformation

$$S(\Lambda) = \mathrm{e}^{-\frac{\mathrm{i}}{4}\sigma_{\mu\nu}\omega^{\mu\nu}} \tag{6.2.29b}$$

with $\omega^{\mu\nu} = -\omega^{\nu\mu}$ and the Lorentz transformation reads $\Lambda = \mathrm{e}^{\omega}$, where the matrix elements of $\omega$ are equal to $\omega^{\mu}{}_{\nu}$. For example, one can represent a rotation through an angle $\vartheta$ about an arbitrary axis $\hat{\mathbf{n}}$ as

$$S = \mathrm{e}^{\frac{\mathrm{i}}{2}\vartheta\hat{\mathbf{n}}\cdot\boldsymbol{\Sigma}} , \tag{6.2.29c}$$

where

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix} . \tag{6.2.29d}$$

#### 6.2.2.4 Spatial Reflection, Parity

The Lorentz transformation corresponding to a spatial reflection is represented by

$$\Lambda^{\mu}{}_{\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} . \tag{6.2.30}$$

The associated $S$ is determined, according to (6.2.7), from

$$S^{-1}\gamma^{\mu}S = \Lambda^{\mu}{}_{\nu}\gamma^{\nu} = \sum_{\nu=1}^{4} g^{\mu\nu}\gamma^{\nu} = g^{\mu\mu}\gamma^{\mu} , \tag{6.2.31}$$

where no summation over $\mu$ is implied. One immediately sees that the solution of (6.2.31), which we shall denote in this case by $P$, is given by

$$S = P \equiv \mathrm{e}^{\mathrm{i}\varphi}\gamma^{0} . \tag{6.2.32}$$

Here, $\mathrm{e}^{\mathrm{i}\varphi}$ is an unobservable phase factor. This is conventionally taken to have one of the four values $\pm 1$, $\pm \mathrm{i}$; four reflections then yield the identity $\mathbb{1}$. The spinors transform under a spatial reflection according to

$$\psi'(x') \equiv \psi'(\mathbf{x}', t) = \psi'(-\mathbf{x}, t) = \mathrm{e}^{\mathrm{i}\varphi}\gamma^{0}\psi(x) = \mathrm{e}^{\mathrm{i}\varphi}\gamma^{0}\psi(-\mathbf{x}', t) . \tag{6.2.33}$$

The complete spatial reflection (parity) transformation for spinors is denoted by

$$\mathcal{P} = \mathrm{e}^{i\varphi}\gamma^{0}\mathcal{P}^{(0)} , \tag{6.2.33$'$}$$

where $\mathcal{P}^{(0)}$ causes the spatial reflection $\mathbf{x} \to -\mathbf{x}$.

From the relationship $\gamma^{0} \equiv \beta = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}$ one sees in the rest frame of the particle, spinors of positive and negative energy (Eq. (5.3.22)) that are eigenstates of $P$ – with opposite eigenvalues, i.e., opposite parity. *This means that the intrinsic parities of particles and antiparticles are opposite.*

### 6.2.3 Further Properties of $S$

For the calculation of the transformation of bilinear forms such as $j^\mu(x)$, we need to establish a relationship between the adjoint transformations $S^\dagger$ and $S^{-1}$.

**Assertion:**

$$S^\dagger\gamma^0 = b\gamma^0 S^{-1} \quad , \tag{6.2.34a}$$

where

$$b = \pm 1 \quad \text{for} \quad \Lambda^{00}\begin{cases} \geq +1 \\ \leq -1 \end{cases} . \tag{6.2.34b}$$

*Proof:* We take as our starting point Eq. (6.2.7)

$$S^{-1}\gamma^\mu S = \Lambda^\mu{}_\nu\gamma^\nu \ , \qquad \Lambda^\mu{}_\nu \text{ real,} \tag{6.2.35}$$

and write down the adjoint relation

$$(\Lambda^\mu{}_\nu\gamma^\nu)^\dagger = S^\dagger\gamma^{\mu\dagger}S^{\dagger-1} \ . \tag{6.2.36}$$

The hermitian adjoint matrix can be expressed most concisely as

$$\gamma^{\mu\dagger} = \gamma^0\gamma^\mu\gamma^0 \ . \tag{6.2.37}$$

By means of the anticommutation relations, one easily checks that (6.2.37) is in accord with $\gamma^{0\dagger} = \gamma^0$, $\gamma^{k\dagger} = -\gamma^k$. We insert this into the left- and the right-hand sides of (6.2.36) and then multiply by $\gamma^0$ from the left- and right-hand side to gain

$$\gamma^0\Lambda^\mu{}_\nu\gamma^0\gamma^\nu\gamma^0\gamma^0 = \gamma^0 S^\dagger\gamma^0\gamma^\mu\gamma^0 S^{\dagger-1}\gamma^0$$

$$\Lambda^\mu{}_\nu\gamma^\nu = S^{-1}\gamma^\mu S = \gamma^0 S^\dagger\gamma^0\gamma^\mu(\gamma^0 S^\dagger\gamma^0)^{-1} \ ,$$

since $(\gamma^0)^{-1} = \gamma^0$. Furthermore, on the left-hand side we have made the substitution $\Lambda^\mu{}_\nu\gamma^\nu = S^{-1}\gamma^\mu S$. We now multiply by $S$ and $S^{-1}$:

$$\gamma^\mu = S\gamma^0 S^\dagger\gamma^0\gamma^\mu(\gamma^0 S^\dagger\gamma^0)^{-1}S^{-1} \equiv (S\gamma^0 S^\dagger\gamma^0)\gamma^\mu(S\gamma^0 S^\dagger\gamma^0)^{-1} \ .$$

Thus, $S\gamma^0 S^\dagger\gamma^0$ commutes with all $\gamma^\mu$ and is therefore a multiple of the unit matrix

$$S\gamma^0 S^\dagger\gamma^0 = b\,\mathbb{1} \quad , \tag{6.2.38}$$

which also implies that

$$S\gamma^0 S^\dagger = b\gamma^0 \tag{6.2.39}$$

and yields the relation we are seeking[4]

$$S^\dagger \gamma^0 = b(S\gamma^0)^{-1} = b\gamma^0 S^{-1} \, . \tag{6.2.34a}$$

Since $(\gamma^0)^\dagger = \gamma^0$ and $S\gamma^0 S^\dagger$ are hermitian, by taking the adjoint of (6.2.39) one obtains $S\gamma^0 S^\dagger = b^* \gamma^0$, from which it follows that

$$b^* = b \tag{6.2.40}$$

and thus $b$ is real. Making use of the fact that the normalization of $S$ is fixed by $\det S = 1$, on calculating the determinant of (6.2.39), one obtains $b^4 = 1$. This, together with (6.2.40), yields:

$$b = \pm 1 \, . \tag{6.2.41}$$

The significance of the sign in (6.2.41) becomes apparent when one considers

$$\begin{aligned} S^\dagger S &= S^\dagger \gamma^0 \gamma^0 S = b\gamma^0 S^{-1} \gamma^0 S = b\gamma^0 \Lambda^0{}_\nu \gamma^\nu \\ &= b\Lambda^0{}_0 \, \mathbb{1} + \sum_{k=1}^{3} b\Lambda^0{}_k \underbrace{\gamma^0 \gamma^k}_{\alpha^k} \, . \end{aligned} \tag{6.2.42}$$

$S^\dagger S$ has positive definite eigenvalues, as can be seen from the following. Firstly, $\det S^\dagger S = 1$ is equal to the product of all the eigenvalues, and these must therefore all be nonzero. Furthermore, $S^\dagger S$ is hermitian and its eigenfunctions satisfy $S^\dagger S \psi_a = a\psi_a$, whence

$$a\psi_a^\dagger \psi_a = \psi_a^\dagger S^\dagger S \psi_a = (S\psi_a)^\dagger S\psi_a > 0$$

and thus $a > 0$. Since the trace of $S^\dagger S$ is equal to the sum of all the eigenvalues, we have, in view of (6.2.42) and using $\mathrm{Tr}\, \alpha^k = 0$,

$$0 < \mathrm{Tr}\,(S^\dagger S) = 4b\Lambda^0{}_0 \, .$$

Thus, $b\Lambda^0{}_0 > 0$. Hence, we have the following relationship between the signs of $\Lambda^{00}$ and $b$:

$$\begin{aligned} \Lambda^{00} &\geq \phantom{-}1 \quad \text{for} \quad b = 1 \\ \Lambda^{00} &\leq -1 \quad \text{for} \quad b = -1 \, . \end{aligned} \tag{6.2.34b}$$

For Lorentz transformations that do not change the direction of time, we have $b = 1$; while those that do cause time reversal have $b = -1$.

[4] Note: For the Lorentz transformation $L_+^\uparrow$ (restricted L.T. and rotations) and for spatial reflections, one can derive this relation with $b = 1$ from the explicit representations.

### 6.2.4 Transformation of Bilinear Forms

The *adjoint* spinor is defined by

$$\bar{\psi} = \psi^\dagger \gamma^0 \ . \tag{6.2.43}$$

We recall that $\psi^\dagger$ is referred to as a hermitian adjoint spinor. The additional introduction of $\bar{\psi}$ is useful because it allows quantities such as the current density to be written in a concise form. We obtain the following transformation behavior under a Lorentz transformation:

$$\psi' = S\psi \Longrightarrow \psi'^\dagger = \psi^\dagger S^\dagger \Longrightarrow \bar{\psi}' = \psi^\dagger S^\dagger \gamma^0 = b\,\psi^\dagger \gamma^0 S^{-1} \ ,$$

thus,

$$\bar{\psi}' = b\,\bar{\psi} S^{-1} \ . \tag{6.2.44}$$

Given the above definition, the current density (5.3.7) reads:

$$j^\mu = c\,\psi^\dagger \gamma^0 \gamma^\mu \psi = c\,\bar{\psi}\gamma^\mu \psi \tag{6.2.45}$$

and thus transforms as

$$j^{\mu\prime} = c\,b\,\bar{\psi} S^{-1} \gamma^\mu S \psi = \Lambda^\mu{}_\nu c\,b\,\bar{\psi}\gamma^\nu\psi = b\Lambda^\mu{}_\nu j^\nu \ . \tag{6.2.46}$$

Hence, $j^\mu$ transforms in the same way as a vector for Lorentz transformations without time reflection. In the same way one immediately sees, using (6.2.3) and (6.2.44), that $\bar{\psi}(x)\psi(x)$ transforms as a scalar:

$$\begin{aligned} \bar{\psi}'(x')\psi'(x') &= b\bar{\psi}(x')S^{-1}S\psi(x') \\ &= b\,\bar{\psi}(x)\psi(x) \ . \end{aligned} \tag{6.2.47a}$$

We now summarize the transformation behavior of the most important bilinear quantities under *orthochronous Lorentz transformations* , i.e., transformations that *do not reverse the direction of time*:

$$\bar{\psi}'(x')\psi'(x') = \bar{\psi}(x)\psi(x) \qquad \text{scalar} \tag{6.2.47a}$$
$$\bar{\psi}'(x')\gamma^\mu\psi'(x') = \Lambda^\mu{}_\nu\bar{\psi}(x)\gamma^\nu\psi(x) \qquad \text{vector} \tag{6.2.47b}$$
$$\bar{\psi}'(x')\sigma^{\mu\nu}\psi'(x') = \Lambda^\mu{}_\rho\Lambda^\nu{}_\sigma\bar{\psi}(x)\sigma^{\rho\sigma}\psi(x) \qquad \text{antisymmetric tensor} \tag{6.2.47c}$$
$$\bar{\psi}'(x')\gamma_5\gamma^\mu\psi'(x') = (\det\Lambda)\Lambda^\mu{}_\nu\bar{\psi}(x)\gamma_5\gamma^\nu\psi(x) \qquad \text{pseudovector} \tag{6.2.47d}$$
$$\bar{\psi}'(x')\gamma_5\psi'(x') = (\det\Lambda)\bar{\psi}(x)\gamma_5\psi(x) \qquad \text{pseudoscalar,} \tag{6.2.47e}$$

where $\gamma_5 = \mathrm{i}\gamma^0\gamma^1\gamma^2\gamma^3$. We recall that $\det\Lambda = \pm 1$; for spatial reflections the sign is $-1$.

### 6.2.5 Properties of the $\gamma$ Matrices

We remind the reader of the definition of $\gamma^5$ from the previous section:

$$\gamma_5 \equiv \gamma^5 \equiv \mathrm{i}\gamma^0\gamma^1\gamma^2\gamma^3 \tag{6.2.48}$$

and draw the reader's attention to the fact that somewhat different definitions may also be encountered in the literature. In the standard representation (5.3.21) of the Dirac matrices, $\gamma^5$ has the form

$$\gamma^5 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} . \tag{6.2.48$'$}$$

The matrix $\gamma^5$ satisfies the relations

$$\{\gamma^5, \gamma^\mu\} = 0 \tag{6.2.49a}$$

and

$$(\gamma^5)^2 = \mathbb{1} . \tag{6.2.49b}$$

By forming products of $\gamma^\mu$, one can construct 16 linearly independent $4 \times 4$ matrices. These are

$$\Gamma^S = \mathbb{1} \tag{6.2.50a}$$
$$\Gamma^V_\mu = \gamma_\mu \tag{6.2.50b}$$
$$\Gamma^T_{\mu\nu} = \sigma_{\mu\nu} = \frac{\mathrm{i}}{2}[\gamma_\mu, \gamma_\nu] \tag{6.2.50c}$$
$$\Gamma^A_\mu = \gamma_5\gamma_\mu \tag{6.2.50d}$$
$$\Gamma^P = \gamma_5 . \tag{6.2.50e}$$

The upper indices indicate scalar, vector, tensor, axial vector (= pseudovector), and pseudoscalar. These matrices have the following *properties*[5]:

1. $(\Gamma^a)^2 = \pm\mathbb{1}$ (6.2.51a)

2. For every $\Gamma^a$ except $\Gamma^S \equiv \mathbb{1}$, there exists a $\Gamma^b$, such that

$$\Gamma^a\Gamma^b = -\Gamma^b\Gamma^a . \tag{6.2.51b}$$

3. For $a \neq S$ we have $\mathrm{Tr}\,\Gamma^a = 0$. (6.2.51c)
*Proof:* $\mathrm{Tr}\,\Gamma^a(\Gamma^b)^2 = -\mathrm{Tr}\,\Gamma^b\Gamma^a\Gamma^b = -\mathrm{Tr}\,\Gamma^a(\Gamma^b)^2$
Since $(\Gamma^b)^2 = \pm 1$, it follows that $\mathrm{Tr}\,\Gamma^a = -\mathrm{Tr}\,\Gamma^a$, thus proving the assertion.

[5] Only some of these properties will be proved here; other proofs are included as problems.

4. For every pair $\Gamma^a,\ \Gamma^b\ a \neq b$ there is a $\Gamma^c \neq \mathbb{1}$, such that $\Gamma_a \Gamma_b = \beta \Gamma_c$, $\beta = \pm 1,\ \pm \mathrm{i}$.
   Proof follows by considering the $\Gamma$.
5. The matrices $\Gamma^a$ are linearly independent.
   Suppose that $\sum_a x_a \Gamma^a = 0$ with complex coefficients $x_a$. From property 3 above one then has
   $$\mathrm{Tr} \sum_a x_a \Gamma^a = x_S = 0\ .$$
   Multiplication by $\Gamma_a$ and use of the properties 1 and 4 shows that subsequent formation of the trace leads to $x_a = 0$.
6. If a $4 \times 4$ matrix $X$ commutes with every $\gamma^\mu$, then $X \propto \mathbb{1}$.
7. Given two sets of $\gamma$ matrices, $\gamma$ and $\gamma'$, both of which satisfy
   $$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}\ ,$$
   there must exist a nonsingular $M$
   $$\gamma'^\mu = M \gamma^\mu M^{-1}\ . \tag{6.2.51d}$$
   This $M$ is unique to within a constant factor (Pauli's fundamental theorem).

## 6.3 Solutions of the Dirac Equation for Free Particles

### 6.3.1 Spinors with Finite Momentum

We now seek solutions of the free Dirac equation (5.3.1) or (5.3.17)

$$(-\mathrm{i}\partial\!\!\!/ + m)\psi(x) = 0\ . \tag{6.3.1}$$

Here, and below, we will set $\hbar = c = 1$.

For particles at rest, these solutions [see (5.3.22)] read:

$$\begin{aligned} \psi^{(+)}(x) &= u_r(m, \mathbf{0})\, \mathrm{e}^{-\mathrm{i}mt} \qquad\qquad r = 1, 2 \\ \psi^{(-)}(x) &= v_r(m, \mathbf{0})\, \mathrm{e}^{\mathrm{i}mt}\ , \end{aligned} \tag{6.3.2}$$

for the positive and negative energy solutions respectively, with

$$\begin{aligned} u_1(m, \mathbf{0}) &= \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix},\ u_2(m, \mathbf{0}) = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \\ v_1(m, \mathbf{0}) &= \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix},\ v_2(m, \mathbf{0}) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \end{aligned} \tag{6.3.3}$$

and are normalized to unity. These solutions of the Dirac equation are eigenfunctions of the Dirac Hamiltonian $H$ with eigenvalues $\pm m$, and also of the operator (the matrix already introduced in (6.2.19))

$$\sigma^{12} = \frac{\mathrm{i}}{2}[\gamma^1, \gamma^2] = \begin{pmatrix} \sigma^3 & 0 \\ 0 & \sigma^3 \end{pmatrix} \tag{6.3.4}$$

with eigenvalues $+1$ (for $r = 1$) and $-1$ (for $r = 2$). Later we will show that $\sigma^{12}$ is related to the spin.

We now seek solutions of the Dirac equation for finite momentum in the form[6]

$$\psi^{(+)}(x) = u_r(k)\,\mathrm{e}^{-\mathrm{i}k\cdot x} \qquad \text{positive energy} \tag{6.3.5a}$$
$$\psi^{(-)}(x) = v_r(k)\,\mathrm{e}^{\mathrm{i}k\cdot x} \qquad \text{negative energy} \tag{6.3.5b}$$

with $k^0 > 0$. Since (6.3.5a,b) must also satisfy the Klein–Gordon equation, we know from (5.2.14) that

$$k_\mu k^\mu = m^2 \ , \tag{6.3.6}$$

or

$$E \equiv k^0 = \left(\mathbf{k}^2 + m^2\right)^{1/2} \ , \tag{6.3.7}$$

where $k^0$ is also written as $E$; i.e., $k$ is the four-momentum of a particle with mass $m$.

The spinors $u_r(k)$ and $v_r(k)$ can be found by Lorentz transformation of the spinors (6.3.3) for particles at rest: We transform into a coordinate system that is moving with velocity $-\mathbf{v}$ with respect to the rest frame and then, from the rest-state solutions, we obtain the free wave functions for electrons with velocity $\mathbf{v}$. However, a more straightforward approach is to determine the solutions directly from the Dirac equation. Inserting (6.3.5a,b) into the Dirac equation (6.3.1) yields:

$$(\not{k} - m)u_r(k) = 0 \quad \text{and} \quad (\not{k} + m)v_r(k) = 0 \ . \tag{6.3.8}$$

Furthermore, we have

$$\not{k}\not{k} = k_\mu\gamma^\mu k_\nu\gamma^\nu = k_\mu k_\nu \frac{1}{2}\{\gamma^\mu, \gamma^\nu\} = k_\mu k_\nu g^{\mu\nu} \ . \tag{6.3.9}$$

Thus, from (6.3.6), one obtains

$$(\not{k} - m)(\not{k} + m) = k^2 - m^2 = 0 \ . \tag{6.3.10}$$

Hence one simply needs to apply $(\not{k} + m)$ to the $u_r(m, \mathbf{0})$ and $(\not{k} - m)$ to the $v_r(m, \mathbf{0})$ in order to obtain the solutions $u_r(k)$ and $v_r(k)$ of (6.3.8). The

[6] We write the four-momentum as $k$, the four-coordinates as $x$, and their scalar product as $k \cdot x$.

normalization remains as yet unspecified; it must be chosen such that it is compatible with the solution (6.3.3), and such that $\bar{\psi}\psi$ transforms as a scalar (Eq. (6.2.47a)). As we will see below, this is achieved by means of the factor $1/\sqrt{2m(m+E)}$:

$$u_r(k) = \frac{\not{k} + m}{\sqrt{2m(m+E)}} u_r(m, \mathbf{0}) = \begin{pmatrix} \left(\frac{E+m}{2m}\right)^{1/2} \chi_r \\ \\ \frac{\boldsymbol{\sigma}\cdot\mathbf{k}}{(2m(m+E))^{1/2}} \chi_r \end{pmatrix} \tag{6.3.11a}$$

$$v_r(k) = \frac{-\not{k} + m}{\sqrt{2m(m+E)}} v_r(m, \mathbf{0}) = \begin{pmatrix} \frac{\boldsymbol{\sigma}\cdot\mathbf{k}}{(2m(m+E))^{1/2}} \chi_r \\ \\ \left(\frac{E+m}{2m}\right)^{1/2} \chi_r \end{pmatrix} . \tag{6.3.11b}$$

Here, the solutions are represented by $u_r(m, \mathbf{0}) = \binom{\chi_r}{0}$ and $v_r(m, \mathbf{0}) = \binom{0}{\chi_r}$ with $\chi_1 = \binom{1}{0}$ and $\chi_2 = \binom{0}{1}$.

In this calculculation we have made use of

$$\begin{aligned} \not{k}\begin{pmatrix} \chi_r \\ 0 \end{pmatrix} &= \left[k^0 \begin{pmatrix} 1\!\!1 & 0 \\ 0 & -1\!\!1 \end{pmatrix} - k^i \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}\right] \begin{pmatrix} \chi_r \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} k^0 \chi_r \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ k^i \sigma^i \chi_r \end{pmatrix} = \begin{pmatrix} E\chi_r \\ \mathbf{k}\cdot\boldsymbol{\sigma}\chi_r \end{pmatrix} \end{aligned}$$

and

$$-\not{k}\begin{pmatrix} 0 \\ \chi_r \end{pmatrix} = \begin{pmatrix} 0 \\ k^0 \chi_r \end{pmatrix} + \begin{pmatrix} k^i \sigma^i \chi_r \\ 0 \end{pmatrix} , \; r = 1, 2 \; .$$

From (6.3.11a,b) one finds for the adjoint spinors defined in (6.2.43)

$$\bar{u}_r(k) = \bar{u}_r(m, \mathbf{0}) \frac{\not{k} + m}{\sqrt{2m(m+E)}} \tag{6.3.12a}$$

$$\bar{v}_r(k) = \bar{v}_r(m, \mathbf{0}) \frac{-\not{k} + m}{\sqrt{2m(m+E)}} . \tag{6.3.12b}$$

*Proof:* $\bar{u}_r(k) = u_r^\dagger(k)\gamma^0 = u_r^\dagger(m, \mathbf{0}) \frac{(\gamma^{\mu\dagger} k_\mu + m)\gamma^0}{\sqrt{2m(m+E)}} = u_r^\dagger(m, \mathbf{0}) \frac{\gamma^0(\gamma^\mu k_\mu + m)}{\sqrt{2m(m+E)}}$,

since $\gamma^{\mu\dagger} = \gamma^0\gamma^\mu\gamma^0$ and $(\gamma^0)^2 = 1\!\!1$

Furthermore, the adjoint spinors satisfy the equations

$$\bar{u}_r(k)\,(\not{k} - m) = 0 \tag{6.3.13a}$$

and

$$\bar{v}_r(k)\,(\not{k}+m)=0\;, \tag{6.3.13b}$$

as can be seen from (6.3.10) and (6.3.12a,b) or (6.3.8).

### 6.3.2 Orthogonality Relations and Density

We shall need to know a number of formal properties of the solutions found above for later use. From (6.3.11) and (6.2.37) it follows that:

$$\bar{u}_r(k)u_s(k)=\bar{u}_r(m,\mathbf{0})\frac{(\not{k}+m)^2}{2m(m+E)}u_s(m,\mathbf{0})\;. \tag{6.3.14a}$$

With

$$\begin{aligned}\bar{u}_r(m,\mathbf{0})(\not{k}+m)^2u_s(m,\mathbf{0}) &= \bar{u}_r(m,\mathbf{0})(\not{k}^2+2m\not{k}+m^2)u_s(m,\mathbf{0})\\ &= \bar{u}_r(m,\mathbf{0})(2m^2+2m\not{k})u_s(m,\mathbf{0})\\ &= \bar{u}_r(m,\mathbf{0})(2m^2+2mk^0\gamma^0)u_s(m,\mathbf{0})\\ &= 2m(m+E)\bar{u}_r(m,\mathbf{0})u_s(m,\mathbf{0})\\ &= 2m(m+E)\delta_{rs}\;,\end{aligned} \tag{6.3.14b}$$

$$\begin{aligned}\bar{u}_r(k)v_s(k) &= \bar{u}_r(m,\mathbf{0})\frac{\not{k}^2-m^2}{2m(m+E)}v_s(m,\mathbf{0})\\ &= \bar{u}_r(m,\mathbf{0})\;0\;v_s(m,\mathbf{0})=0\end{aligned} \tag{6.3.14c}$$

and a similar calculation for $v_r(k)$, equations (6.3.14a,b) yield the

**orthogonality relations**

$$\begin{array}{rclrcl}\bar{u}_r(k)\,u_s(k) &=& \delta_{rs} & \bar{u}_r(k)\,v_s(k) &=& 0\\ \bar{v}_r(k)\,v_s(k) &=& -\delta_{rs} & \bar{v}_r(k)\,u_s(k) &=& 0.\end{array} \tag{6.3.15}$$

Remarks:

(i) This normalization remains invariant under orthochronous Lorentz transformations:

$$\bar{u}'_r\,u'_s=u_r^\dagger\,S^\dagger\,\gamma^0\,S\,u_s=u_r^\dagger\,\gamma^0\,S^{-1}\,S\,u_s=\bar{u}_r\,u_s=\delta_{rs}\;. \tag{6.3.16}$$

(ii) For these spinors, $\bar{\psi}(x)\psi(x)$ is a scalar,

$$\bar{\psi}^{(+)}(x)\psi^{(+)}(x)=\mathrm{e}^{\mathrm{i}k\cdot x}\bar{u}_r(k)u_r(k)\mathrm{e}^{-\mathrm{i}kx}=1\;, \tag{6.3.17}$$

is independent of $k$, and thus independent of the reference frame.
In general, for a superposition of positive energy solutions, i.e., for

$$\psi^{(+)}(x) = \sum_{r=1}^{2} c_r u_r \ , \text{ with } \sum_{r=1}^{2} |c_r|^2 = 1 \ , \tag{6.3.18a}$$

one has the relation

$$\bar{\psi}^{(+)}(x)\psi^{(+)}(x) = \sum_{r,s} \bar{u}_r(k) u_s(k) c_r^* \, c_s = \sum_{r=1}^{2} |c_r|^2 = 1 \ . \tag{6.3.18b}$$

Analogous relationships hold for $\psi^{(-)}$.

(iii) If one determines $u_r(k)$ through a Lorentz transformation corresponding to $-\mathbf{v}$, this yields exactly the above spinors. Viewed as an active transformation, this amounts to transforming $u_r(m, \mathbf{0})$ to the velocity $\mathbf{v}$. Such a transformation is known as a "boost".

The *density* for a plane wave ($c = 1$) is $\rho = j^0 = \bar{\psi}\gamma^0\psi$. This is not a Lorentz-invariant quantity since it is the zero-component of a four-vector:

$$\begin{aligned}
\bar{\psi}_r^{(+)}(x)\gamma^0 \, \psi_s^{(+)}(x) &= \bar{u}_r(k)\gamma^0 \, u_s(k) \\
&= \bar{u}_r(k) \frac{\{\not{k}, \gamma^0\}}{2m} \, u_s(k) = \frac{E}{m} \, \delta_{rs} && (6.3.19a) \\
\bar{\psi}_r^{(-)}(x)\gamma^0 \, \psi_s^{(-)}(x) &= \bar{v}_r(k)\gamma^0 \, v_s(k) \\
&= -\bar{v}_r(k) \frac{\{\not{k}, \gamma^0\}}{2m} \, v_s(k) = \frac{E}{m} \, \delta_{rs} \, . && (6.3.19b)
\end{aligned}$$

In the intermediate steps here, we have used $u_s(k) = (\not{k}/m)u_s(k)$, $\bar{u}_s(k) = \bar{u}_s(k)(\not{k}/m)$ (Eqs. (6.3.8) and (6.3.13)) etc.

**Note.** The spinors are normalized such that the density in the rest frame is unity. Under a Lorentz transformation, the density times the volume must remain constant. The volume is reduced by a factor $\sqrt{1-\beta^2}$ and thus the density must increase by the reciprocal factor $\frac{1}{\sqrt{1-\beta^2}} = \frac{E}{m}$.

We now extend the sequence of equations (6.3.19).

$$\begin{aligned}
\text{For} \quad \psi_r^{(+)}(x) &= \mathrm{e}^{-\mathrm{i}(k^0 x^0 - \mathbf{k}\cdot\mathbf{x})} u_r(k) \\
\text{and} \quad \psi_s^{(-)}(x) &= \mathrm{e}^{\mathrm{i}(k^0 x^0 + \mathbf{k}\cdot\mathbf{x})} v_s(\tilde{k})
\end{aligned} \tag{6.3.20}$$

with the four-momentum $\tilde{k} = (k^0, -\mathbf{k})$, one obtains

$$\begin{aligned}
\bar{\psi}_r^{(-)}(x)\gamma^0 \, \psi_s^{(+)}(x) &= \mathrm{e}^{-2\mathrm{i}k^0 x^0} \bar{v}_r(\tilde{k})\gamma^0 \, u_s(k) \\
&= \frac{1}{2}\mathrm{e}^{-2\mathrm{i}k^0 x^0} \, \bar{v}_r(\tilde{k}) \left( -\frac{\tilde{\not{k}}}{m}\gamma^0 + \gamma^0 \frac{\not{k}}{m} \right) u_s(k) && (6.3.19c) \\
&= 0
\end{aligned}$$

since the terms proportional to $k_0$ cancel and since $\left\{k_i\gamma^i, \gamma^0\right\} = 0$. In this sense, positive and negative energy states are orthogonal for opposite energies and equal momenta.

### 6.3.3 Projection Operators

The operators

$$\Lambda_{\pm}(k) = \frac{\pm \not{k} + m}{2m} \tag{6.3.21}$$

project onto the spinors of positive and negative energy, respectively:

$$\begin{array}{ll} \Lambda_+ u_r(k) = u_r(k) & \Lambda_- v_r(k) = v_r(k) \\ \Lambda_+ v_r(k) = 0 & \Lambda_- u_r(k) = 0 \quad . \end{array} \tag{6.3.22}$$

Thus, the projection operators $\Lambda_{\pm}(k)$ can also be represented in the form

$$\begin{aligned} \Lambda_+(k) &= \sum_{r=1,2} u_r(k) \otimes \bar{u}_r(k) \\ \Lambda_-(k) &= -\sum_{r=1,2} v_r(k) \otimes \bar{v}_r(k) \ . \end{aligned} \tag{6.3.23}$$

The tensor product $\otimes$ is defined by

$$(a \otimes \bar{b})_{\alpha\beta} = a_\alpha \bar{b}_\beta \ . \tag{6.3.24}$$

In matrix form, the tensor product of a spinor $a$ and an adjoint spinor $\bar{b}$ reads:

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} (\bar{b}_1, \bar{b}_2, \bar{b}_3, \bar{b}_4) = \begin{pmatrix} a_1\bar{b}_1 & a_1\bar{b}_2 & a_1\bar{b}_3 & a_1\bar{b}_4 \\ a_2\bar{b}_1 & a_2\bar{b}_2 & a_2\bar{b}_3 & a_2\bar{b}_4 \\ a_3\bar{b}_1 & a_3\bar{b}_2 & a_3\bar{b}_3 & a_3\bar{b}_4 \\ a_4\bar{b}_1 & a_4\bar{b}_2 & a_4\bar{b}_3 & a_4\bar{b}_4 \end{pmatrix} \quad .$$

The projection operators have the following properties:

$$\Lambda_{\pm}^2(k) = \Lambda_{\pm}(k) \tag{6.3.25a}$$

$$\mathrm{Tr}\, \Lambda_{\pm}(k) = 2 \tag{6.3.25b}$$

$$\Lambda_+(k) + \Lambda_-(k) = 1 \ . \tag{6.3.25c}$$

*Proof:*

$$\begin{aligned} \Lambda_{\pm}(k)^2 &= \frac{(\pm\not{k} + m)^2}{4m^2} = \frac{\not{k}^2 \pm 2\not{k}m + m^2}{4m^2} = \frac{m^2 \pm 2\not{k}m + m^2}{4m^2} \\ &= \frac{2m(\pm\not{k} + m)}{4m^2} = \Lambda_{\pm}(k) \end{aligned}$$

$$\mathrm{Tr}\, \Lambda_{\pm}(k) = \frac{4m}{2m} = 2$$

The validity of the assertion that $\Lambda_{\pm}$ projects onto positive and negative energy states can be seen in both of the representations, (6.3.21) and (6.3.22), by applying them to the states $u_r(k)$ and $v_r(k)$. A further important projection operator, $P(n)$, which, in the rest frame projects onto the spin orientation $n$, will be discussed in Problem 6.15.

## Problems

**6.1** Prove the group property of the Poincaré group.

**6.2** Show, by using the transformation properties of $x_\mu$, that $\partial^\mu \equiv \partial/\partial x_\mu$ ($\partial_\mu \equiv \partial/\partial x^\mu$) transforms as a contravariant (covariant) vector.

**6.3** Show that the $N$-fold application of the infinitesimal rotation in Minkowski space (Eq. (6.2.22))

$$\Lambda = 1 + \frac{\vartheta}{N}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

leads, in the limit $N \to \infty$, to a rotation about the $z$ axis through an angle $\vartheta$ (the last step in (6.2.22)).

**6.4** Derive the quadratic form of the Dirac equation

$$\left[\left(\mathrm{i}\hbar\partial\!\!\!/ - \frac{e}{c}A\!\!\!/\right)^2 - \frac{\mathrm{i}\hbar e}{c}\left(\boldsymbol{\alpha E} + \mathrm{i}\boldsymbol{\Sigma B}\right) - m^2c^2\right]\psi = 0$$

for the case of external electromagnetic fields. Write the result using the electromagnetic field tensor $F_{\mu\nu} = A_{\mu,\nu} - A_{\nu,\mu}$, and also in a form explicitly dependent on $\boldsymbol{E}$ and $\boldsymbol{B}$.

*Hint:* Multiply the Dirac equation from the left by $\gamma^\nu\left(\mathrm{i}\hbar\partial_\nu - \frac{e}{c}A_\nu\right) + mc$ and, by using the commutation relations for the $\gamma$ matrices, bring the expression obtained into quadratic form in terms of the field tensor

$$\left[\left(\mathrm{i}\hbar\partial\!\!\!/ - \frac{e}{c}A\!\!\!/\right)^2 - \frac{\hbar e}{2c}\sigma^{\mu\nu}F_{\mu\nu} - m^2c^2\right]\psi = 0\,.$$

The assertion follows by evaluating the expression $\sigma^{\mu\nu}F_{\mu\nu}$ using the explicit form of the field tensor as a function of the fields $\boldsymbol{E}$ and $\boldsymbol{B}$.

**6.5** Consider the quadratic form of the Dirac equation from Problem 6.4 with the fields $\mathbf{E} = E_0\,(1,0,0)$ and $\mathbf{B} = B\,(0,0,1)$, where it is assumed that $E_0/Bc \le 1$. Choose the gauge $\mathbf{A} = B\,(0,x,0)$ and solve the equation with the ansatz

$$\psi(x) = \mathrm{e}^{-\mathrm{i}Et/\hbar}\mathrm{e}^{\mathrm{i}(k_y y + k_z z)}\varphi(x)\Phi\,,$$

where $\Phi$ is a four-spinor that is independent of time and space coordinates. Calculate the energy eigenvalues for an electron. Show that the solution agrees with that obtained from Problem 5.3 when one considers the nonrelativistic limit, i.e., $E_0/Bc \ll 1$.

*Hint:* Given the above ansatz for $\psi$, one obtains the following form for the quadratic Dirac equation:

$$\left[K(x,\partial_x)\mathbb{1} + M\right]\varphi(x)\Phi = 0\,,$$

where $K(x, \partial_x)$ is an operator that contains constant contributions, $\partial_x$ and $x$. The matrix $M$ is independent of $\partial_x$ and $x$; it has the property $M^2 \propto \mathbb{1}$. This suggests that the bispinor $\Phi$ has the form $\Phi = (\mathbb{1} + \lambda M)\Phi_0$. Determine $\lambda$ and the eigenvalues of $M$. With these eigenvalues, the matrix differential equation reverts into an ordinary differential equation of the oscillator type.

**6.6** Show that equation (6.2.14′)

$$[\gamma^\mu, \tau] = \Delta\omega^{\mu\nu}\gamma_\nu$$

is satisfied by

$$\tau = \frac{1}{8}\Delta\omega^{\mu\nu}(\gamma_\mu\gamma_\nu - \gamma_\nu\gamma_\mu) \ .$$

**6.7** Prove that $\gamma^{\mu\dagger} = \gamma^0\gamma^\mu\gamma^0$.

**6.8** Show that the relation

$$S^\dagger\gamma^0 = b\gamma^0 S^{-1}$$

is satisfied with $b = 1$ by the explicit representations of the elements of the Poincaré group found in the main text (rotation, pure Lorentz transformation, spatial reflection).

**6.9** Show that $\bar{\psi}(x)\gamma_5\psi(x)$ is a pseudoscalar, $\bar{\psi}(x)\gamma_5\gamma^\mu\psi(x)$ a pseudovector, and $\bar{\psi}(x)\sigma^{\mu\nu}\psi(x)$ a tensor.

**6.10** Properties of the matrices $\Gamma^a$.
Taking as your starting point the definitions (6.2.50a–e), derive the following properties of these matrices:
(i) For every $\Gamma^a$ (except $\Gamma^S$) there exists a $\Gamma^b$ such that $\Gamma^a\Gamma^b = -\Gamma^b\Gamma^a$.
(ii) For every pair $\Gamma^a$, $\Gamma^b$, $(a \neq b)$ there exists a $\Gamma^c \neq \mathbb{1}$ such that $\Gamma^a\Gamma^b = \beta\Gamma^c$ with $\beta = \pm 1, \pm \mathrm{i}$.

**6.11** Show that if a $4 \times 4$ matrix $X$ commutes with all $\gamma^\mu$, then this matrix $X$ is proportional to the unit matrix.
*Hint*: Every $4 \times 4$ matrix can, according to Problem 6.1, be written as a linear combination of the 16 matrices $\Gamma^a$ (basis!).

**6.12** Prove Pauli's fundamental theorem for Dirac matrices: For any two four-dimensional representations $\gamma_\mu$ and $\gamma'_\mu$ of the Dirac algebra both of which satisfy the relation

$$\{\gamma_\mu, \gamma_\nu\} = 2g_{\mu\nu}$$

there exists a nonsingular transformation $M$ such that

$$\gamma'_\mu = M\gamma_\mu M^{-1} \ .$$

$M$ is uniquely determined to within a constant prefactor.

**6.13** From the solution of the field-free Dirac equation in the rest frame, determine the four-spinors $\psi^{\pm}(x)$ of a particle moving with the velocity $\mathbf{v}$. Do this by applying a Lorentz transformation (into a coordinate system moving with the velocity $-\mathbf{v}$) to the solutions in the rest frame.

**6.14** Starting from

$$\Lambda_+(k) = \sum_{r=1,2} u_r(k) \otimes \bar{u}_r(k) \ , \qquad \Lambda_-(k) = -\sum_{r=1,2} v_r(k) \otimes \bar{v}_r(k) \ ,$$

prove the validity of the representations for $\Lambda_\pm(k)$ given in (6.3.22).

**6.15** (i) Given the definition $P(n) = \frac{1}{2}(1 + \gamma_5 \not{n})$, show that, under the assumptions $n^2 = -1$ and $n_\mu k^\mu = 0$, the following relations are satisfied

(a) $[\Lambda_\pm(k), P(n)] = 0 \quad ,$

(b) $\Lambda_+(k)P(n) + \Lambda_-(k)P(n) + \Lambda_+(k)P(-n) + \Lambda_-(k)P(-n) = 1 \quad ,$

(c) $\mathrm{Tr}\,[\Lambda_\pm(\mathrm{k})\mathrm{P}(\pm\mathrm{n})] = 1 \quad ,$

(d) $P(n)^2 = P(n)$

(ii) Consider the special case $n = (0, \hat{e}_z)$ where $P(n) = \frac{1}{2}\begin{pmatrix} 1+\sigma^3 & 0 \\ 0 & 1-\sigma^3 \end{pmatrix}$.

# 7. Orbital Angular Momentum and Spin

We have seen that, in nonrelativistic quantum mechanics, the angular momentum operator is the generator of rotations and commutes with the Hamiltonians of rotationally invariant (i.e., spherically symmetric) systems[1]. It thus plays a special role for such systems. For this reason, as a preliminary to the next topic – the Coulomb potential – we present here a detailed investigation of angular momentum in relativistic quantum mechanics.

## 7.1 Passive and Active Transformations

For positive energy states, in the non-relativistic limit we derived the Pauli equation with the Landé factor $g = 2$ (Sect. 5.3.5). From this, we concluded that the Dirac equation describes particles with spin $S = 1/2$. Following on from the transformation behavior of spinors, we shall now investigate angular momentum in general.

In order to give the reader useful background information, we will start with some remarks concerning active and passive transformations. Consider a given state $Z$, which in the reference frame $I$ is described by the spinor $\psi(x)$. When regarded from the reference frame $I'$, which results from $I$ through the Lorentz transformation

$$x' = \Lambda x \; , \tag{7.1.1}$$

the spinor takes the form,

$$\psi'(x') = S\psi(\Lambda^{-1}x') \; , \qquad \text{passive with } \Lambda \; . \tag{7.1.2a}$$

A transformation of this type is known as a *passive transformation*. One and the same state is viewed from two different coordinate systems, which is indicated in Fig. 7.1 by $\psi(x) \mathrel{\hat{=}} \psi'(x')$.

On the other hand, one can also transform the state and then view the resulting state $Z'$ exactly as the starting state $Z$ from one and the same reference frame $I$. In this case one speaks of an *active transformation*. For vectors and scalars, it is clear what is meant by their active transformation

[1] See QM I, Sect. 5.1

(rotation, Lorentz transformation). The active transformation of a vector by the transformation $\Lambda$ corresponds to the passive transformation of the coordinate system by $\Lambda^{-1}$. For spinors, the active transformation is defined in exactly this way (see Fig. 7.1).

The state $Z'$, which arises through the transformation $\Lambda^{-1}$, appears in $I$ exactly as $Z$ in $I'$, i.e.,

$$\psi'(x) = S\psi(\Lambda^{-1}x) \qquad \text{active with } \Lambda^{-1} \tag{7.1.2b}$$

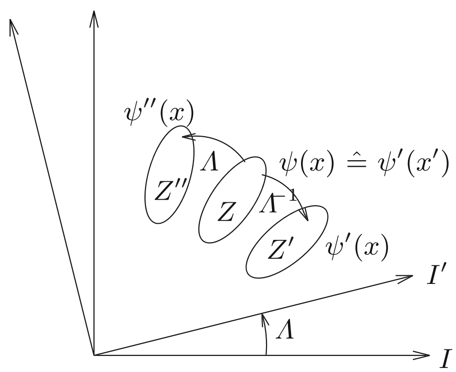



**Fig. 7.1.** Schematic representation of the passive and active transformation of a spinor; the enclosed area is intended to indicate the region in which the spinor is finite

The state $Z''$, which results from $Z$ through the active transformation $\Lambda$, by definition appears the same in $I'$ as does $Z$ in $I$, i.e., it takes the form $\psi(x')$. Since $I$ is obtained from $I'$ by the Lorentz transformation $\Lambda^{-1}$, in $I$ the spinor $Z''$ has the form

$$\psi''(x) = S^{-1}\psi(\Lambda x) , \qquad \text{active with } \Lambda . \tag{7.1.2c}$$

## 7.2 Rotations and Angular Momentum

Under the infinitesimal Lorentz transformation

$$\Lambda^{\mu}{}_{\nu} = g^{\mu}{}_{\nu} + \Delta\omega^{\mu}{}_{\nu} , \quad (\Lambda^{-1})^{\mu}{}_{\nu} = g^{\mu}{}_{\nu} - \Delta\omega^{\mu}{}_{\nu} , \tag{7.2.1}$$

a spinor $\psi(x)$ transforms as

$$\psi'(x') = S\psi(\Lambda^{-1}x') \qquad \text{passive with } \Lambda \tag{7.2.2a}$$

or

$$\psi'(x) = S\psi(\Lambda^{-1}x) \qquad \text{active with } \Lambda^{-1} . \tag{7.2.2b}$$

We now use the results gained in Sect. 6.2.2.1 (Eqs. (6.2.8) and (6.2.13)) to obtain

$$\psi'(x) = (\mathbb{1} - \frac{\mathrm{i}}{4}\,\Delta\omega^{\mu\nu}\sigma_{\mu\nu})\psi(x^\rho - \Delta\omega^\rho{}_\nu x^\nu)\ . \tag{7.2.3}$$

Taylor expansion of the spinor yields $(1 - \Delta\omega^\mu{}_\nu x^\nu \partial_\mu)\,\psi(x)$ , so that

$$\psi'(x) = (\mathbb{1} + \Delta\omega^{\mu\nu}(-\frac{\mathrm{i}}{4}\,\sigma_{\mu\nu} + x_\mu\partial_\nu))\,\psi(x)\ . \tag{7.2.3$'$}$$

We now consider the special case of rotations through $\Delta\boldsymbol{\varphi}$, which are represented by

$$\Delta\omega^{ij} = -\epsilon^{ijk}\Delta\varphi^k \tag{7.2.4}$$

(the direction of $\Delta\boldsymbol{\varphi}$ specifies the rotation axis and $|\Delta\boldsymbol{\varphi}|$ the angle of rotation). If one also uses

$$\sigma^{ij} = \sigma_{ij} = \epsilon^{ijk}\Sigma^k \quad , \quad \Sigma^k = \begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix}\ , \tag{7.2.5}$$

(see Eq. (6.2.19)) one obtains for (7.2.3′)

$$\begin{aligned}
\psi'(x) &= \left(1 + \Delta\omega^{ij}\left(-\frac{\mathrm{i}}{4}\,\epsilon^{ijk}\Sigma^k + x_i\partial_j\right)\right)\psi(x) \\
&= \left(1 - \epsilon^{ij\bar{k}}\Delta\varphi^{\bar{k}}\left(-\frac{\mathrm{i}}{4}\,\epsilon^{ijk}\Sigma^k - x^i\partial_j\right)\right)\psi(x) \\
&= \left(1 - \Delta\varphi^{\bar{k}}\left(-\frac{\mathrm{i}}{4}\,2\delta_{k\bar{k}}\Sigma^k - \epsilon^{ij\bar{k}}x^i\partial_j\right)\right)\psi(x) \\
&= \left(1 + \mathrm{i}\Delta\varphi^k\left(\frac{1}{2}\,\Sigma^k + \epsilon^{kij}x^i\frac{1}{\mathrm{i}}\,\partial_j\right)\right)\psi(x) \\
&\equiv \left(1 + \mathrm{i}\Delta\varphi^k J^k\right)\psi(x)\ .
\end{aligned} \tag{7.2.6}$$

Here, we have defined the total angular momentum

$$J^k = \epsilon^{kij}x^i\frac{1}{\mathrm{i}}\,\partial_j + \frac{1}{2}\,\Sigma^k\ . \tag{7.2.7}$$

With the inclusion of $\hbar$, this operator reads:

$$\mathbf{J} = \mathbf{x} \times \frac{\hbar}{\mathrm{i}}\,\boldsymbol{\nabla}\mathbb{1} + \frac{\hbar}{2}\,\boldsymbol{\Sigma}\,, \tag{7.2.7$'$}$$

and is thus the sum of the orbital angular momentum $\mathbf{L} = \mathbf{x} \times \mathbf{p}$ and the spin $\frac{\hbar}{2}\,\boldsymbol{\Sigma}$.

The total angular momentum (= orbital angular momentum + spin) is the generator of rotations: For a finite angle $\varphi^k$ one obtains, by combining a succession of infinitesimal rotations,

$$\psi'(x) = \mathrm{e}^{\mathrm{i}\varphi^k J^k}\, \psi(x)\,. \tag{7.2.8}$$

The operator $J^k$ commutes with the Hamiltonian of the Dirac equation containing a spherically symmetric potential $\Phi(\mathbf{x}) = \Phi(|\mathbf{x}|)$

$$[H, J^i] = 0\,. \tag{7.2.9}$$

A straightforward way to verify (7.2.9) is by an explicit calculation of the commutator (see Problem 7.1). Here, we consider general consequences resulting from the behavior, under rotation, of the structure of commutators of the angular momentum with other operators; Eq. (7.2.9) results as a special case. We consider an operator $A$, and let the result of its action on $\psi_1$ be the spinor $\psi_2$:

$$A\psi_1(x) = \psi_2(x)\ .$$

It follows that

$$\mathrm{e}^{\mathrm{i}\varphi^k J^k}\, A\, \mathrm{e}^{-\mathrm{i}\varphi^k J^k} \left(\mathrm{e}^{\mathrm{i}\varphi^k J^k}\ \psi_1(x)\right) = \left(\mathrm{e}^{\mathrm{i}\varphi^k J^k}\ \psi_2(x)\right)$$

or, alternatively,

$$\mathrm{e}^{\mathrm{i}\varphi^k J^k}\, A\, \mathrm{e}^{-\mathrm{i}\varphi^k J^k}\, \psi_1'(x) = \psi_2'(x)\ .$$

Thus, in the rotated frame of reference the operator is

$$A' = \mathrm{e}^{\mathrm{i}\varphi^k J^k}\, A\, \mathrm{e}^{-\mathrm{i}\varphi^k J^k}\ . \tag{7.2.10}$$

Expanding this for infinitesimal rotations ($\varphi^k \to \Delta\varphi^k$) yields:

$$A' = A - \mathrm{i}\Delta\varphi^k [A, J^k] \qquad . \tag{7.2.11}$$

The following special cases are of particular interest:

(i) $A$ is a scalar (rotationally invariant) operator. Then, $A' = A$ and from (7.2.11) it follows that

$$[A, J^k] = 0\,. \tag{7.2.12}$$

The Hamiltonian of a rotationally invariant system (including a spherically symmetric $\Phi(\mathbf{x}) = \Phi(|\mathbf{x}|)$) is a scalar; this leads to (7.2.9). Hence, in spherically symmetric problems the angular momentum is conserved.

(ii) For the operator $A$ we take the components of a three-vector $\mathbf{v}$ . As a vector, $\mathbf{v}$ transforms according to $v'^i = v^i + \epsilon^{ijk}\, \Delta\varphi^j\, v^k$. Equating this, component by component, with (7.2.11), $v^i + \epsilon^{ijk} \Delta\varphi^j v^k = v^i + \frac{\mathrm{i}}{\hbar}\Delta\varphi^j \left[J^j, v^i\right]$ which shows that

$$[J^i, v^j] = \mathrm{i}\hbar\, \epsilon^{ijk}\, v^k\ . \tag{7.2.13}$$

The commutation relation (7.2.13) implies, among other things,

$$\left[J^i, J^j\right] = \mathrm{i}\hbar\epsilon^{ijk} J^k \tag{7.2.14a}$$

$$\left[J^i, L^j\right] = \mathrm{i}\hbar\epsilon^{ijk} L^k \quad . \tag{7.2.14b}$$

It is clear from the explicit representation $\Sigma^k = \begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix}$ that the eigenvalues of the $4 \times 4$ matrices $\Sigma^k$ are doubly degenerate and take the values $\pm 1$. The angular momentum $\mathbf{J}$ is the sum of the orbital angular momentum $\mathbf{L}$ and the intrinsic angular momentum or spin $\mathbf{S}$, the components of which have the eigenvalues $\pm\frac{1}{2}$. Thus, particles that obey the Dirac equation have spin $S = 1/2$. The operator $\left(\frac{\hbar}{2}\boldsymbol{\Sigma}\right)^2 = \frac{3}{4}\hbar^2 1\!\!1$ has the eigenvalue $\frac{3\hbar^2}{4}$. The eigenvalues of $\mathbf{L}^2$ and $L^3$ are $\hbar^2 l(l+1)$ and $\hbar m_l$, where $l = 0, 1, 2, \ldots$ and $m_l$ takes the values $-l, -l+1, \ldots, l-1, l$. The eigenvalues of $\mathbf{J}^2$ are thus $\hbar^2 j(j+1)$, where $j = l \pm \frac{1}{2}$ for $l \neq 0$ and $j = \frac{1}{2}$ for $l = 0$. The eigenvalues of $J^3$ are $\hbar m_j$, where $m_j$ ranges in integer steps between $-j$ and $j$. The operators $\mathbf{J}^2$, $\mathbf{L}^2$, $\boldsymbol{\Sigma}^2$, and $J^3$ can be simultaneously diagonalized. The orbital angular momentum operators $L^i$ and the spin operators $\Sigma^i$ themselves fulfill the angular momentum commutation relations.

**Note:** One is tempted to ask how it is that the Dirac Hamiltonian, a $4 \times 4$ matrix, can be a scalar. In order to see this, one has to return to the transformation (6.2.6′). The transformed Hamiltonian including a central potential $\Phi(|\mathbf{x}|)$

$$(-\mathrm{i}\gamma^\nu\, \partial'_\nu + m + e\Phi(|\mathbf{x}'|)) = S(-\mathrm{i}\gamma^\nu\, \partial_\nu + m + e\Phi(|\mathbf{x}|))S^{-1}$$

has, under rotations, the same form in both systems. The property "scalar" means invariance under rotations, but is not necessarily limited to one-component spherically symmetric functions.

## Problems

**7.1** Show, by explicit calculation of the commutator, that the total angular momentum

$$\mathbf{J} = \mathbf{x} \times \mathbf{p}\, 1\!\!1 + \frac{\hbar}{2}\boldsymbol{\Sigma}$$

commutes with the Dirac Hamiltonian for a central potential

$$H = c\left(\sum_{k=1}^{3} \alpha^k p^k + \beta mc\right) + e\Phi(|\mathbf{x}|)\ .$$

CHAPTER 4

# Boson Wave Equations

The dynamical quantum-mechanical wave equations of spin-0 pions, spin-1 vector particles, and massless spin-1 photons are formulated in a consistent one-particle fashion. For the spin-0 Klein–Gordon equation, the interpretation of negative-energy states as describing antiparticles is stressed. The relativistic bound-state Coulomb problem is then solved for $\pi$-mesic atoms. The parallel is made between the massive spin-1 and photon wave equations. The notion of currents, current conservation and gauge invariance for photon amplitudes is discussed in detail and linked to the principle of minimal replacement. Minimal coupling of photons to charged particles will be the basis of the general electromagnetic interaction to be considered in later chapters. Second-quantized field theories are briefly described, and an analogy is made between (relativistic) photons and nonrelativistic phonons.

## 4.A Spin-0 Klein–Gordon Equation

**Derivation.** For a particle moving at relativistic velocities (i.e., having a kinetic energy that is a substantial fraction of its rest mass), the nonrelativistic approximation for energy, $E = m + \mathbf{p}^2/2m$, is no longer valid and one must use instead the exact relation $E = (\mathbf{p}^2 + m^2)^{\frac{1}{2}}$. The formal quantum-mechanical replacement $\mathbf{p} \rightarrow -i\nabla$ would then result in the Hamiltonian

$$H = (-\nabla^2 + m^2)^{\frac{1}{2}} \tag{4.1}$$

and a free particle Schrödinger-type equation

$$i\partial_t \psi = (-\nabla^2 + m^2)^{\frac{1}{2}}\psi. \tag{4.2}$$

Since the square-root operation in (4.1) and (4.2) is difficult to interpret, it would seem more reasonable to construct a relativistic wave equation associated with the square of the Hamiltonian operator, $(i\partial_t)^2\phi = H^2\phi$. Then defining the D'Alembertian operator as

$$\Box \equiv \partial_\mu \partial^\mu = \partial_t^2 - \nabla^2 = -p^2, \tag{4.3}$$

where $p_\mu = i\partial_\mu$, one is naturally led to the free-particle Klein–Gordon equation

$$(\Box + m^2)\phi(x) = 0, \qquad (p^2 - m^2)\phi(x) = 0. \tag{4.4}$$

Such a particle is said to be "on its mass shell", $p^2 = m^2$.

**Covariance.** The manifest covariance of (4.4) for a quantum-mechanical (Lorentz) scalar wave function, $\phi'(x') = \phi(x)$, i.e., for

$$|\phi'\rangle = U_\Lambda |\phi\rangle, \tag{4.5a}$$

$$\phi'(x) = \langle x|\phi'\rangle = \langle x|U_\Lambda|\phi\rangle = \langle \Lambda^{-1}x|\phi\rangle = \phi(\Lambda^{-1}x), \tag{4.5b}$$

implies that

$$(\partial'_\mu \partial'^\mu + m^2)\phi'(x') = (\partial_\mu \partial^\mu + m^2)\phi(x) = 0. \tag{4.6}$$

Due to these transformation properties, the wave function $\phi(x)$ must describe a spin-0 particle (e.g., a pion). Spin-0 solutions of the free-particle Klein–Gordon equation (4.4) are proportional to the invariant plane-wave functions

$$\begin{aligned} \phi_+(x) &\propto e^{-ip\cdot x} = e^{i\mathbf{p}\cdot\mathbf{x}}e^{-iEt}, \\ \phi_-(x) &\propto e^{ip\cdot x} = e^{-i\mathbf{p}\cdot\mathbf{x}}e^{iEt}. \end{aligned} \tag{4.7}$$

These equations are special cases of the general solutions of the Klein–Gordon wave equation containing a possible interaction term,

$$\phi_\pm(x) = \phi_\pm(\mathbf{x})e^{\mp iEt}. \tag{4.8}$$

It is therefore clear that the use of the operator $H^2$ in forming a wave equation leads to seemingly unphysical negative-energy solutions $e^{+iEt}$ as well as physical positive-energy solutions $e^{-iEt}$ for the quantum mechanical state of the particle. While this problem led to a temporary discarding of the Klein–Gordon equation in the late 1920s and early 1930s, we have since learned to live with it, as will be discussed shortly.

**Probability Current.** Another problem which arises with solutions of (4.4) is the construction of a positive definite probability density. Paralleling postulate $v$ in Section 1.A, one searches for a covariant probability current density

$j_\mu = (\rho, \mathbf{j})$ which obeys a continuity equation

$$\partial^\mu j_\mu = \partial^0 j_0 + \partial^i j_i = \partial_t j_0 + \nabla \cdot \mathbf{j} = 0, \tag{4.9}$$

where we have used (3.60), $\partial^\mu = (\partial_t, \nabla)$. The obvious candidate for $j_\mu$ is the hermitian form

$$j_\mu(x) = \phi^* i\overleftrightarrow{\partial}_\mu \phi \equiv \phi^* i\partial_\mu \phi - i(\partial_\mu \phi^*)\phi, \tag{4.10}$$

because this current is conserved ($\partial \cdot j = 0$) for states $\phi$ and $\phi^*$ obeying the Klein–Gordon equation (4.4):

$$\partial \cdot j = i\partial^\mu(\phi^* \partial_\mu \phi) - i\partial^\mu(\phi \partial_\mu \phi^*) = i\phi^* \Box \phi - i\phi \Box \phi^* = 0. \tag{4.11}$$

Furthermore, since the spatial part of (4.10), $\mathbf{j}$, is identical in form to the nonrelativistic current density (1.6) except for normalization, we are obliged by covariance arguments to accept the timelike component of (4.10), $\rho$, as the probability density. When combined with the general solutions (4.8), this probability density becomes

$$\rho_+(x) = \phi_+^* i\overleftrightarrow{\partial}_0 \phi_+ = 2E|\phi_+(\mathbf{x})|^2 \geq 0 \tag{4.12a}$$

for positive-energy solutions, and

$$\rho_-(x) = \phi_-^* i\overleftrightarrow{\partial}_0 \phi_- = -2E|\phi_-(\mathbf{x})|^2 \leq 0 \tag{4.12b}$$

for negative-energy solutions, with $E > 0$. Clearly a negative probability density is unacceptable; we shall contend with (4.12b) shortly.

**Wave Packets.** Construct general spin-0, free-particle wave packets in the Hilbert space of positive and negative energy states as (normalized in a box—see Section 1.B)

$$\phi_+(x) = \int \frac{d^3p}{2EV^{\frac{1}{2}}} a_\mathbf{p} e^{-ip\cdot x}, \qquad \phi_-(x) = \int \frac{d^3p}{2EV^{\frac{1}{2}}} b_\mathbf{p}^* e^{ip\cdot x}, \tag{4.13}$$

where the complex conjugation of $b_\mathbf{p}$ follows the usual convention. The factor of $2E$ ($E > 0$) in (4.13) is a manifestation of our covariant normalization of states. From (4.13) it is clear that the evolution of these packets does not alter their positive- and negative-energy character. Consider then a scalar product defined over the positive-energy states (4.13a) as (see Problem 4.1)

$$\langle \phi'_+, \phi_+ \rangle = \int d^3x\, \phi'^*_+(x) i\overleftrightarrow{\partial}_0 \phi_+(x) \tag{4.14a}$$

$$= \int \frac{d^3p}{2EV} a'^*_\mathbf{p} a_\mathbf{p} \tag{4.14b}$$

$$= \int d^4p\, \delta(p^2 - m^2)\theta(p_0) a'^*_\mathbf{p} a_\mathbf{p}/V, \tag{4.14c}$$

where $\delta(p^2 - m^2)\theta(p_0)$ in (4.14c) indicates that only positive-energy states with $p^2 = m^2$ are allowed. Note that this norm is time independent [differentiate (4.14a) with respect to time and use the Klein–Gordon equation; the result is obvious for (4.14b) and (4.14c)]. Note too that the norm (4.14) is a Lorentz invariant. Also, the covariant normalization of states, $\langle \mathbf{p}' | \mathbf{p} \rangle = 2E\delta^3(\mathbf{p}' - \mathbf{p})$, follows from (4.14b) with the replacement

$$a_{\mathbf{p}} V^{-\frac{1}{2}} \to 2E\delta^3(\mathbf{p} - \mathbf{p}_i) \tag{4.15}$$

for a plane wave of momentum $\mathbf{p}_i$.

**Interpretation of Negative-Energy States.** Historically, the resolution of the negative-energy and negative-probability-density problems led to a reformulation of the Klein–Gordon theory in a many-body context. It is possible, however, to stay within a single-particle framework (Stückelberg 1941; Feynman 1949) by interpreting (4.12a) as the charge density of a positively charged, positive-energy particle "propagating" forward in time ($t > 0$, $E > 0$) via the plane-wave phase $e^{-iEt}$. Similarly, one interprets (4.12b) as the charge density of a positively charged, negative-energy state propagating backward in time ($t = -|t| < 0$) via $e^{iEt} = e^{-iE|t|}$. Alternatively (4.12b) is the charge density of a negatively charged, positive-energy particle propagating forward in time via the complex conjugate of the phase, $e^{iEt} = (e^{-iEt})^*$. For neutral particles which are their own antiparticles (i.e., $\pi_c^0 = \pi^0$, where $c$ refers to the "charge conjugate" antiparticle—see Section 6.A), one can choose the wave function to be purely real or imaginary. In this case the probability density (4.12) vanishes, consistent with treating $\rho$ as a charge density. Unfortunately, a thorough understanding of this interpretation must await a discussion of charge conjugation in Chapter 6 and "backward propagation in time" in Chapters 7 and 10.

The Stückelberg–Feynman interpretation is ideally suited for scattering processes, where the particle is free and unlocalized before and after the collision. For bound-state wave packets, however, a particle constrained to $\Delta x \lesssim m^{-1}$ and $\Delta p \sim 1/\Delta x \sim m$ demands a superposition of all Fourier components, negative as well as positive energy:

$$\phi(x) = \int \frac{d^3p}{2EV^{\frac{1}{2}}} \left(a_{\mathbf{p}} e^{-ip\cdot x} + b_{\mathbf{p}}^* e^{ip\cdot x}\right). \tag{4.16}$$

Then $\phi^*\phi$ contains interference terms $e^{\pm 2iEt}$ which produce violent oscillations, referred to as "*Zitterbewegung*". Since such "jittery-motion" plays a more significant role for spin-$\frac{1}{2}$ particles, we postpone a detailed discussion of it until Chapter 5. Suffice it to say that as $E \to m$, such interference between positive and negative energy components could have physical consequences.

**Feshbach–Villars Formulation.** It is possible to circumvent these interference terms by constructing a Klein–Gordon bound-state wave function which has no negative-energy component in the nonrelativistic limit (Feshbach and Villars 1958). For $\phi$ satisfying the Klein–Gordon equation with time derivative $\dot{\phi}$, define

$$\varphi = \frac{1}{2}\left(\phi + \frac{i}{m}\dot{\phi}\right), \qquad \chi = \frac{1}{2}\left(\phi - \frac{i}{m}\dot{\phi}\right). \tag{4.17}$$

Because $\phi = \phi_+ \to e^{-imt}$ as $E \to m$ implies $\varphi \to e^{-imt}$ and $\chi \to 0$ by (4.17), $\varphi$ is called the "large" and $\chi$ the "small" positive energy component of $\phi$. Moreover $\varphi$ and $\chi$ satisfy coupled first-order equations in time,

$$i\dot{\varphi} = -\frac{\nabla^2}{2m}(\varphi + \chi) + m\varphi, \qquad i\dot{\chi} = \frac{\nabla^2}{2m}(\varphi + \chi) - m\chi. \tag{4.18}$$

These equations can be unified into one matrix equation satisfied by the column vector

$$\Phi = \begin{pmatrix} \varphi \\ \chi \end{pmatrix}, \tag{4.19}$$

and resembling the Schrödinger form

$$i\partial_t \Phi = H_0 \Phi, \tag{4.20}$$

where $H_0$ is the $2 \times 2$ matrix "hamiltonian"

$$H_0 = \beta(1 + \alpha)\mathbf{p}^2/2m + \beta m, \tag{4.21}$$

with

$$\beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \alpha = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{4.22}$$

While the positive- and negative-energy states are coupled in this scheme, the problems associated with negative energy are not completely eliminated, because $H_0$ as given by (4.21) is not hermitian in that

$$\beta(1 + \alpha) = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}.$$

This links the large and small components $\varphi$ and $\chi$ together, implying as in (4.12) that the probability density $\rho$ is not positive definite. If nothing else, however, this approach indicates that a second-order equation in time can be reduced to a first-order equation by doubling the Hilbert space of states via the column vector (4.19). [From an historical standpoint, Dirac learned this fact 30 years earlier when he discovered the first-order Dirac equation for spin-$\frac{1}{2}$ particles (Dirac 1928).] Furthermore this formalism is ideally suited for nonrelativistic reductions. We will exploit a similar pattern in the case of the Dirac equation in Chapter 5.

**External Fields.** Next we consider the modification of the Klein–Gordon equation for spin-0 particles in the presence of an electromagnetic field, specified by the vector and scalar potentials $\mathbf{A}$ and $A_0$ as the four-vector $A_\mu = (A_0, \mathbf{A})$. Following the minimal substitution procedure of classical physics, we take

$$i\partial_\mu = p_\mu \to p_\mu - eA_\mu, \tag{4.23}$$

where $e$ is the electric charge of the particle, taken as positive unless otherwise specified. The Klein–Gordon equation then becomes

$$[(i\partial - eA)^2 - m^2]\phi(x) = 0, \tag{4.24a}$$

or

$$(\Box + m^2)\phi(x) = -J(x), \tag{4.24b}$$

where $J(x)$ is a scalar current "source" density

$$J(x) = 2ieA \cdot \partial\phi + ie\phi\partial \cdot A - e^2A^2\phi, \tag{4.25}$$

or $J = V\phi$, $V$ being an effective potential operator acting on $\phi$.

This scalar current density should be contrasted with the vector current density $j_\mu$, altered from the form (4.10) by the minimal substitution law (4.23):

$$j_\mu(x) = \phi^*\overleftrightarrow{(i\partial_\mu - eA_\mu)}\phi = i\phi^*\overleftrightarrow{\partial}_\mu\phi - 2e\phi^*\phi A_\mu. \tag{4.26}$$

The doubling of the $A_\mu$ term in (4.26) is a consequence of $A_\mu$ being real. Following the procedure (4.11), one can demonstrate that (4.26) is also conserved, since in this case

$$\partial \cdot j = i\phi^*\Box\phi - i\phi\Box\phi^* - 2e\phi^*\phi\partial\cdot A - 2eA \cdot \partial(\phi^*\phi) = 0 \tag{4.27}$$

by use of (4.24b) and (4.25).

**Bound-State Coulomb Atom.** Finally we consider the specific bound-state Coulomb problem of a spin-0 $\pi^-$ particle with charge $-e$ bound to a heavy nucleus with charge $Ze$ ($\pi$-mesic atom). The static potential for this configuration is ($\alpha = e^2/4\pi$)

$$eA_\mu = V\delta_{\mu 0} = -\frac{Z\alpha}{r}\,\delta_{\mu 0}. \tag{4.28}$$

Writing the positive-energy wave function as $\phi(\mathbf{r}, t) = \phi(\mathbf{r})e^{-iEt}$, the spatial part obeys a Klein–Gordon equation obtained from (4.24):

$$\left[\left(E + \frac{Z\alpha}{r}\right)^2 + \nabla^2 - m^2\right]\phi(\mathbf{r}) = 0. \tag{4.29}$$

This latter equation is solved by the standard method of separation of variables with $\phi(\mathbf{r}) = \phi(r)Y_l^m(\hat{\mathbf{r}})$ and $\nabla^2 = r^{-1}(\partial^2/\partial r^2)r - \mathbf{L}^2/r^2$, leading to the one-dimensional radial equation

$$\left[\frac{1}{r}\frac{d^2}{dr^2}r - \frac{l(l+1) - (Z\alpha)^2}{r^2} + \frac{2EZ\alpha}{r}\right]\phi(r) = -(E^2 - m^2)\phi(r). \tag{4.30a}$$

In the limit $E \to m$, $E^2 - m^2 \to 2mE_{NR}$, and $Z\alpha \ll 1$ (i.e., for small-$Z$ atoms, $Z \ll 137$), (4.30a) becomes the usual nonrelativistic Schrödinger radial equation

$$\left[\frac{1}{r}\frac{d^2}{dr^2}r - \frac{l(l+1)}{r^2} + \frac{2mZ\alpha}{r}\right]\phi(r) = -2mE_{NR}\phi(r). \tag{4.30b}$$

We are familiar with the solutions of (4.30b) for integer values of $n - l = 1, 2, \ldots$, corresponding to the Bohr energy levels for $n = 1, 2, \ldots$,

$$E - m = E_{NR} = -\frac{m(Z\alpha)^2}{2n^2}. \tag{4.31}$$

Given (4.31), we can infer the Klein–Gordon energy levels by replacing $m \to E$, $E_{NR} \to (E^2 - m^2)/2E$ [by inspection of (4.30a)], and $n \to n'$ [see Schiff (1968)]:

$$E^2 - m^2 = -E^2(Z\alpha)^2/n'^2. \tag{4.32a}$$

This relation can be solved explicitly for the relativistic energy as

$$E = m\left[1 + \frac{(Z\alpha)^2}{n'^2}\right]^{-\frac{1}{2}}. \tag{4.32b}$$

Here $n'$ is a new relativistic principal quantum number with $n' - l'$ equal to the usual nonrelativistic quantum-number difference $n - l$, assuming only the integer values $n' - l' = 1, 2, 3, \ldots$; and $l'$ is the effective relativistic orbital angular momentum, inferred from (4.30a) to be

$$l'(l' + 1) = l(l + 1) - (Z\alpha)^2. \tag{4.33}$$

Solving (4.33) for $l'$ gives

$$l' = -\tfrac{1}{2} + [(l + \tfrac{1}{2})^2 - (Z\alpha)^2]^{\frac{1}{2}}, \tag{4.34a}$$

$$n' = n - l + l' = n - (l + \tfrac{1}{2}) + [(l + \tfrac{1}{2})^2 - (Z\alpha)^2]^{\frac{1}{2}}, \tag{4.34b}$$

where the positive sign of the square root has been chosen in order that $l'$ may be nonnegative as $Z\alpha \to 0$, corresponding to bound radial solutions $r^{l'}$ regular at the origin. This form of $n'$, (4.34b), is to be applied to the energy levels (4.32).

For $Z\alpha \ll 1$, both (4.32b) and (4.34) can be expanded in the form of (4.31) with $n = 1, 2, \ldots$ (see problem 4.2),

$$E - m = E_{NR} = -\frac{m(Z\alpha)^2}{2n^2}\left[1 + \frac{(Z\alpha)^2}{n}\left(\frac{1}{l + \frac{1}{2}} - \frac{3}{4n}\right)\right]. \tag{4.35}$$

This removes the $l$ degeneracy of $E$ in the $O(\alpha^4)$ relativistic fine-structure term. For large $Z$, (4.34) provides the constraint that for $l'$ and $n'$ real, the discriminant of the square root must be positive (of course, it cannot vanish) and for $s$-waves ($l = 0$) this means that

$$Z < \frac{1}{2\alpha} = \frac{137}{2}. \tag{4.36}$$

If instead $Z > 1/2\alpha$, the centrifugal barrier term in (4.30) becomes attractive for $s$-waves and the energy becomes imaginary, so that the wave function has a damped exponential part and the particle orbits become unstable. The situation is similar to the classical relativistic situation for the Coloumb potential; for $L^2 < (Z\alpha)^2$ the centrifugal barrier also becomes attractive and the particle spirals in to the origin. Multiparticle quantum states then presumably play a role, with short-distance corrections such as vacuum polarization by pair creation (see Chapter 15) modifying the single-particle wave function. In the next chapter we shall again return to this strong-field limit for the case of bound electrons.

## 4.B Spin-1 Wave Equation

**Derivation.** Starting with a three-component wave function $\phi_i$ describing a massive spin-1 free particle in its rest frame, two possible rest-frame covariant forms exist: a covariant four-vector $\phi_\mu = (0, \phi_i)$ and a rank-two antisymmetric (field) tensor $\phi_{\mu\nu}$ given by (recall the angular-momentum tensor operators $L_{\mu\nu}$ and $J_{\mu\nu}$ in Chapter 3) $\phi_{0i} = -\phi_{i0} = \partial_t \phi_i$ and $\phi_{00} = \phi_{ij} = 0$. In a general frame, the boosted form of $\phi_{\mu\nu}$ can be obtained from $\phi_\mu$ as

$$\phi_{\mu\nu} = \partial_\mu \phi_\nu - \partial_\nu \phi_\mu. \tag{4.37}$$

The free-particle dynamical relation between $\phi_\mu$ and $\phi_{\mu\nu}$ is called the Proca equation:

$$\partial^\nu \phi_{\mu\nu} = m^2 \phi_\mu. \tag{4.38}$$

Owing to the antisymmetric structure of $\phi_{\mu\nu}$, the derivative of (4.38) implies the subsidiary condition

$$\partial^\mu \phi_\mu = 0. \tag{4.39}$$

Since $\phi_\mu$ must transform according to the $(\frac{1}{2}, \frac{1}{2})$ representation of $\mathscr{L}$, we know that (4.39) is required to rule out the spin-0 component in $\phi_\mu$ (see Section 3.B). This in turn provides a group-theoretical justification for the dynamical equation (4.38). Moreover, using (4.37) to eliminate $\phi_{\mu\nu}$ in (4.38) and applying (4.39), we are led to a Klein–Gordon equation for the vector wave function,

$$(\Box + m^2)\phi_\mu(x) = 0, \tag{4.40}$$

which, as in the spin-0 case, guarantees the correct dynamical relation between energy and momentum for a free particle.

**Current Densities.** Given the wave functions $\phi_{\mu\nu}$ and $\phi_\mu$, it is possible to construct a conserved, hermitian current density analogous to (4.10) but describing spin-1 probabilities:

$$j_\mu(x) = i(\phi_{\mu\nu}\phi^{\nu *} - \phi^*_{\mu\nu}\phi^\nu). \tag{4.41}$$

The field equation (4.38) along with (4.37) then manifests $\partial \cdot j = 0$. As was the case for the spin-0 probability density, $j_0 = \rho$ obtained from (4.41) is not a positive definite quantity. Likewise, negative energy and *Zitterbewegung* problems again arise, to be handled in a manner similar to spin-0 particles. We therefore proceed to indicate the new aspects associated with spin-1 wave functions.

In the presence of an electromagnetic field, the minimal replacement (4.23) converts the field equation (4.38) from the spin-1 free-particle Klein–Gordon equation (4.40) to a similar equation containing a vector source current (see Problem 4.3),

$$(\Box + m^2)\phi_\mu(x) = J_\mu(x). \tag{4.42}$$

The actual structure of $J_\mu$ is not revealing, but if the subsidiary condition (4.39) is satisfied, then this source current is conserved. Consider too the modification of the spin one probability current density (4.41) in the presence of external electromagnetic fields. It is a straightforward task to show that the minimal replacement $i\partial_\mu \to i\partial_\mu - eA_\mu$ converts (4.41) to

$$\begin{aligned} j_\mu(x) = \phi^{*\alpha}[i\overleftrightarrow{\partial}_\mu g_{\alpha\beta} + i\overleftarrow{\partial}_\beta g_{\mu\alpha} - i\overrightarrow{\partial}_\alpha g_{\mu\beta} \\ - e(2A_\mu g_{\alpha\beta} - A_\beta g_{\mu\alpha} - A_\alpha g_{\mu\beta})]\phi^\beta. \end{aligned} \tag{4.43}$$

It can be demonstrated that this current is conserved ($\partial \cdot j = 0$) in much the same manner as for the spin-0 case (see Problem 4.4). It turns out, however, that unlike the spin-0 and spin-$\frac{1}{2}$ currents (the latter to be discussed in Chapter 5), the minimal spin-1 current is not unique (Lee 1965).

**Free Particle Solutions.** Finally we display the covariantly normalized free-particle, plane-wave solution of (4.40),

$$\phi_\mu(x) = \varepsilon_\mu(\mathbf{p})e^{-ip\cdot x}, \tag{4.44}$$

with the polarization vector further specified by the helicity eigenvalues, $\varepsilon_\mu^{(\lambda)}(\mathbf{p})$ and $\lambda = \pm 1, 0$. The subsidiary condition $\partial \cdot \phi = 0$ is therefore equivalent to $p \cdot \varepsilon = 0$, as derived earlier in (3.102). As in the spin-0 case, a factor of $(2E)^{-\frac{1}{2}}$ in (4.44) has been absorbed into the covariant normalization of the states, and a one-particle volume normalization factor of $V^{-\frac{1}{2}}$ has been set equal to unity. The orthogonality of the wave functions (4.44) then demands

$$\varepsilon_\mu^{(\lambda')*}(\mathbf{p})\varepsilon^{\mu(\lambda)}(\mathbf{p}) = -\delta_{\lambda'\lambda}, \tag{4.45}$$

a result which can be verified from the specific forms (3.98) and (3.101). Likewise, in the rest frame of the particle, the completeness property of the polarization vectors implies

$$\sum_\lambda^3 \varepsilon_i^{(\lambda)}(\hat{\mathbf{p}})\varepsilon_j^{(\lambda)*}(\hat{\mathbf{p}}) = \delta_{ij}, \tag{4.46a}$$

and since $\delta_{ij}$ can be expressed in four-dimensional language as $-(g_{\mu\nu} - m_\mu m_\nu/m^2)$, the boosted form of (4.46a) is

$$\sum_\lambda^3 \varepsilon_\mu^{(\lambda)}(\mathbf{p})\varepsilon_\nu^{(\lambda)*}(\mathbf{p}) = -\left(g_{\mu\nu} - \frac{p_\mu p_\nu}{m^2}\right). \tag{4.46b}$$

The $\rho$-meson ($m_\rho \approx 776$ MeV) and the $\omega$-meson ($m_\omega \approx 783$ MeV) are examples of spin-1 particles occurring in nature. They are significantly heavier than the spin-0 $\pi$-meson ($m_\pi \approx 140$ MeV).

## 4.C Spin-1 Maxwell Equation

A massless spin-1 photon obeys a quantum-mechanical wave equation somewhat similar in structure to the massive one, with $\phi_\mu$ and $\phi_{\mu\nu}$ respectively replaced by the four-vector potential $A_\mu$ and the antisymmetric field tensor $F_{\mu\nu}$, where

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{4.47}$$

The Maxwell free-field equation is then [setting $m = 0$ in (4.38)]

$$-\partial^\nu F_{\mu\nu} = \Box A_\mu - \partial_\mu(\partial \cdot A) = 0. \tag{4.48}$$

We take this as the dynamical wave equation for noninteracting photons.

**Existence of Gauges.** Note now that a subsidiary condition $\partial \cdot A = 0$ is no longer a direct consequence of the field equation itself, as was (4.39). Rather, an ambiguity now exists, and the value of $\partial \cdot A$ is correlated with the way we restrict the number of spin states of the photon to two, as required by the helicity constraint $\lambda = \pm 1$.

For plane waves, the representation for $A_\mu$ satisfying (4.48) for a noninteracting photon with $k^2 = \omega^2 - \mathbf{k}^2 = 0$ has the general form

$$A_\mu(x) = \varepsilon_\mu(\mathbf{k}) e^{-ik \cdot x}. \tag{4.49}$$

In Section 3.D we chose the "transverse gauge" for $\varepsilon_\mu^{(\pm 1)}(\mathbf{k})$ in order to insure only two independent components in $\varepsilon_\mu$:

$$\varepsilon_0^{(\pm 1)}(\mathbf{k}) = 0, \qquad \mathbf{k} \cdot \boldsymbol{\varepsilon}^{(\pm 1)}(\mathbf{k}) = 0. \tag{4.50}$$

But we cannot separately choose both conditions (4.50) in a frame-independent manner. A Lorentz-invariant choice is the combination $k \cdot \varepsilon(k) = 0$, and (4.49) then corresponds to the Lorentz gauge for (4.48):

$$\partial \cdot A = 0, \qquad \Box A_\mu = 0. \tag{4.51}$$

While (4.51) is not the only gauge decomposition of (4.48), it is the natural choice for free photons, paralleling massive spin-1 particles. The transverse or radiation gauge is $\boldsymbol{\partial} \cdot \mathbf{A} = 0$. In any case, the orthogonality and completeness relations for photon polarization vectors are analogous to (4.45) and (4.46):

$$\varepsilon^{*(\lambda')}(\mathbf{k}) \cdot \varepsilon^{(\lambda)}(\mathbf{k}) = -\delta_{\lambda'\lambda}, \tag{4.52}$$

valid in any gauge, while

$$\sum_{\lambda = \pm 1} \varepsilon_i^{(\lambda)}(\mathbf{k}) \varepsilon_j^{(\lambda)*}(\mathbf{k}) = \delta_{ij} - \hat{k}_i \hat{k}_j \tag{4.53}$$

is valid only in the transverse gauge (4.50).

**Gauge Invariance.** It is possible to transform $A_\mu$ to a different gauge so that (4.50) and (4.51) no longer hold but (4.48) remains an identity. This is achieved by a gauge transformation

$$A_\mu \to A_\mu + \partial_\mu \xi(x), \qquad \varepsilon_\mu \to \varepsilon_\mu + \xi k_\mu, \tag{4.54}$$

with $\xi(x)$ an arbitrary scalar function becoming $i\xi e^{-ik\cdot x}$ for a plane wave; thus $\partial \cdot A$ and $k \cdot \varepsilon$ are transformed to

$$\partial \cdot A \to \Box \xi(x), \qquad k \cdot \varepsilon \to 0. \tag{4.55}$$

Consequently, while $k \cdot \varepsilon$ must always vanish for $k^2 = 0$, $\partial \cdot A$ vanishes only in the Lorentz gauge (4.51). The connection between the minimal coupling replacement $i\partial_\mu \to i\partial_\mu - eA_\mu$ and the gauge transformation (4.54) is the principle of "gauge invariance of the second kind". This is the coordinate-dependent phase transformation of a charged-particle wave function,

$$\phi(x) \to e^{-ie\xi(x)}\phi(x), \tag{4.56a}$$

where $\xi(x)$ generates the gauge transformation (4.54), under which the canonical momentum transforms simply:

$$(i\partial_\mu - eA_\mu)\phi(x) \to e^{-ie\xi(x)}(i\partial_\mu - eA_\mu)\phi(x). \tag{4.56b}$$

Observables such as the current densities (4.26) and (4.43) are built up from bilinear products as $\phi^*\phi$ and $\phi^*(i\partial_\mu - eA_\mu)\phi$, manifestly invariant under the phase transformations (4.56); thus current conservation is naturally extended by this principle to include interactions in the presence of electromagnetic fields. From our viewpoint this is further justification for considering minimal replacement as a fundamental principle which generates the *only* interaction between charged particles and photons. In the context of lagrangian field theory, the principle of gauge invariance of the second kind [the first kind corresponding to a constant phase in (4.56) and linked to charge conservation] plays the central role and is sometimes considered the *raison d'être* for the existence of $A_\mu$ itself and its interaction with charged matter. Gauge principles recently have been used to generate other fundamental interactions (strong and weak), but such topics are beyond the scope of this book.

One final fact about gauge invariance of significant import for us later will be the manner in which the two physical spin states of the photon are realized for a general interaction with matter. This is most conveniently stated by the Lorentz-invariant $S$-matrix element, itself expressed in terms of the $M$-function of (3.87). Accordingly, we may write

$$S^{(\lambda)}(\mathbf{k}) = \varepsilon_\mu^{(\lambda)}(\mathbf{k})M^\mu(\mathbf{k}), \tag{4.57}$$

where $S^{(\lambda)}(\mathbf{k})$ and the polarization vector $\varepsilon_\mu^{(\lambda)}(\mathbf{k})$ represent a photon of momentum $\mathbf{k}$ and helicity $\lambda$. While $M_\mu$ transforms like a simple four-vector under $\mathscr{L}$, we recall that $\varepsilon_\mu$ has a slightly modified transformation law (3.109) because a photon really behaves like an E(2) [and not an O(3)] object under a little-group transformation. Applying (3.109) to (4.57) and using

$\Lambda_\mu{}^\nu = (\Lambda^{-1})^\nu{}_\mu$ then leads to (see Problem 4.5)

$$S^{(\lambda)}(\mathbf{k}) = e^{i\lambda\Theta}\left[\Lambda_\mu{}^\nu - \Lambda_0{}^\nu \frac{k_\mu}{\omega}\right] \varepsilon_\nu^{(\lambda)}(\Lambda^{-1}\mathbf{k})M^\mu(\mathbf{k})$$
$$= e^{i\lambda\Theta}\left[\varepsilon_\mu^{(\lambda)}(\Lambda^{-1}\mathbf{k})M^\mu(\Lambda^{-1}\mathbf{k}) - \Lambda_0{}^\nu\varepsilon_\nu^{(\lambda)}(\Lambda^{-1}\mathbf{k})\frac{k_\mu}{\omega}M^\mu(\mathbf{k})\right]. \tag{4.58}$$

Aside from an overall phase factor (of no consequence for physical probabilities), the Lorentz-invariant $S$-matrix transforms as a simple scalar (Weinberg 1964c)

$$S^{(\lambda)}(\mathbf{k}) = e^{i\lambda\Theta}S^{(\lambda)}(\Lambda^{-1}\mathbf{k}). \tag{4.59}$$

Then comparing (4.59) with (4.57) and (4.58), we see that $M_\mu$ must satisfy an additional constraint,

$$k^\mu M_\mu(\mathbf{k}) = 0. \tag{4.60}$$

This condition is sometimes referred to as "on-shell" gauge invariance (on-shell refers to physical, as opposed to virtual or off-shell, energy or momentum—more about this later). It guarantees that $\varepsilon_\mu^{(\pm 1)}$ transforms properly under little group transformations. The reference to gauge invariance means that (4.57) is invariant under the gauge transformation $\varepsilon_\mu \to \varepsilon_\mu + \xi k_\mu$, $\varepsilon \cdot M \to \varepsilon \cdot M + \xi k \cdot M = \varepsilon \cdot M$ by (4.60). We will take note of this fact many times in our later work.

**Charged-Matter Currents.** The probability current density $j_\mu$ for charged particles plays a dual role in that $ej_\mu$ is the charged source current density for the electromagnetic field. That is, the modification of the Maxwell field equation (4.48) in the presence of charged matter has the classical form

$$\partial^\nu F_{\mu\nu}(x) = -ej_\mu(x), \tag{4.61a}$$

or in terms of the dynamical (Klein–Gordon-type) equation for $A_\mu$ in the Lorentz gauge ($\partial \cdot A = 0$),

$$\Box A_\mu(x) = ej_\mu(x) \equiv j_\mu^{\text{em}}(x). \tag{4.61b}$$

Note that the sign of the spin-1 source current in (4.61b) is opposite to that of the spin-0 source current in (4.24b). As a consequence we shall show later that this sign change is linked to a fundamentally attractive spin-0 force (e.g., the pion-exchange strong force) and a fundamentally repulsive spin-1 force (e.g., the photon-exchange electromagnetic force between particles of like charge).

It is clear from the divergenceless nature of the left-hand side of (4.61) that this charged-matter current must be conserved: $\partial \cdot j(x) = 0$. As noted after (4.56), this result is also a consequence of the quantum-mechanical structure of the probability current density, as given for charged spinless particles in (4.26), coupled with the principle of gauge invariance of the second kind as applied to such particles via (4.56). The connection between current conservation and gauge invariance takes on further significance when formulated in momentum space. For a free, spinless, charged particle absorbing the four-momentum $q = p' - p$ from a photon, the current density constructed from positive-energy packets (4.13a) becomes, via (4.15),

$$\begin{aligned} j_\mu^+(x) &= \phi_+^*(x) i\overleftrightarrow{\partial}_\mu \phi_+(x) \\ &= \int \frac{d^3p}{2E} \int \frac{d^3p'}{2E'} a_{\mathbf{p}'}^* a_{\mathbf{p}} \langle \mathbf{p}' | j_\mu^+ | \mathbf{p} \rangle e^{iq\cdot x} \to \langle \mathbf{p}' | j_\mu^+ | \mathbf{p} \rangle e^{iq\cdot x}, \end{aligned} \tag{4.62}$$

and the differential operator $i\partial_\mu$ in (4.62) requires the off-diagonal momentum-space current to be

$$\langle \mathbf{p}' | j_\mu^+ | \mathbf{p} \rangle = (p' + p)_\mu \tag{4.63a}$$

for a free photon, $q^2 = (p' - p)^2 = 0$. On the other hand, in the presence of an external electromagnetic field, $q^2 = (p' - p)^2 \neq 0$; and on grounds of Lorentz covariance and current conservation, the off-diagonal momentum-space electromagnetic current must have the form

$$e\langle \mathbf{p}' | j_\mu^+ | \mathbf{p} \rangle = eF(q^2)(p' + p)_\mu. \tag{4.63b}$$

Here the charged particle is still considered as free, as in initial- and final-state scattering configurations. Since $\partial \cdot j \propto q^\mu \langle \mathbf{p}' | j_\mu | \mathbf{p} \rangle$, current conservation implies, for $p'^2 = p^2 = m^2$,

$$\begin{aligned} q^\mu \langle \mathbf{p}' | j_\mu^+ | \mathbf{p} \rangle &= F(q^2)(p' - p) \cdot (p' + p) \\ &= F(q^2)(m^2 - m^2) = 0, \end{aligned} \tag{4.64}$$

which demonstrates the absence of the only other possible covariant in (4.63b), $q_\mu = (p' - p)_\mu$. The dimensionless Lorentz-invariant function $F(q^2)$ is called a *form factor*. It represents all possible interactions between the charged particle and photons, altered only by "strongly interacting" particles also interacting with the spinless particle in question. At $q^2 = 0$, (4.63a) demands that $F(0) = 1$, regardless of the type of interacting particles present.

The conversion of the free-particle current [(4.10) in coordinate space or (4.63a) in momentum space] to the interacting forms (4.26) or (4.63b) for $q^2 \neq 0$ is a complicated dynamical process which we shall attempt to explain in the latter half of this book. At the present level of discussion the interesting connection is between the *dynamical* current-conservation statement (4.64) and the *kinematical* on-shell gauge-invariance constraint (4.60). For a free photon with momentum $k$ and $k^2 = 0$, (4.64) is a special case of (4.60) (with $q$ replaced by $k$). An analogous situation exists for spin-1 matter currents based upon (4.43) and also spin-$\frac{1}{2}$ matter currents to be described in the next chapter.

## 4.D Second Quantization: Photons and Phonons

Thus far we have treated a photon as a "particle" having a quantum-mechanical wave function obeying the dynamical wave equation (4.48) for a free photon and (4.61) for a photon interacting with charged matter. Needless to say, these quantum field equations are identical in form to Maxwell's equations for the classical electromagnetic field tensor, related to physical electric and magnetic fields as

$$-E_i = F_{0i} = \partial_t A_i - \partial_i A_0, \tag{4.65a}$$

$$B_i = \tfrac{1}{2}\varepsilon_{ijk} F_{jk} = \varepsilon_{ijk}\partial_j A_k. \tag{4.65b}$$

This complementarity relation between a quantum particle and a classical wave or field can be extended one step further, to a quantum field, then referred to as *second quantization.*

**Noninteracting Photons.** In such a quantum field theory, the vector potential of the radiation field is scaled to the space-time-averaged energy flux (Poynting vector), which in rationalized units (restoring $\hbar$ and $c$ here) is

$$\langle \mathbf{S} \rangle = c\langle \mathbf{E} \times \mathbf{B} \rangle = \sum_{\mathbf{k}} \frac{2\omega^2}{c} |\mathbf{A}_{\mathbf{k}}|^2 \hat{\mathbf{k}}, \tag{4.66}$$

where $\mathbf{A}_{\mathbf{k}}$ is the Fourier component of the vector potential

$$\mathbf{A}(\mathbf{r}, t) = \sum_{\mathbf{k}} (\mathbf{A}_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} e^{-i\omega t} + \mathbf{A}_{\mathbf{k}}^{\dagger} e^{-i\mathbf{k}\cdot\mathbf{r}} e^{i\omega t}). \tag{4.67}$$

[Note that we have anticipated that $\mathbf{A}$ will be treated as an operator and have used $\mathbf{A}^{\dagger}$ instead of $\mathbf{A}^*$ in (4.67).] If (4.66) is to represent a flux of photons of density $N/V$ with energy $\hbar\omega$ and angular momentum $\hbar$, then $\langle S \rangle = \hbar\omega c N/V$ implies

$$|\mathbf{A}_{\mathbf{k}}|^2 = \frac{\hbar c^2 N_{\mathbf{k}}}{2\omega V}. \tag{4.68}$$

In a second-quantized theory, (4.68) is interpreted as a matrix element $\langle N_{\mathbf{k}} | \mathbf{A}_{\mathbf{k}}^{\dagger} \cdot \mathbf{A}_{\mathbf{k}} | N_{\mathbf{k}} \rangle$ in a particle number "Fock space", with (restoring the polarization vector but suppressing its helicity components)

$$\mathbf{A}_{\mathbf{k}} = \sqrt{\frac{\hbar c^2}{2\omega V}}\, a_{\mathbf{k}} \boldsymbol{\varepsilon}(\mathbf{k}) \tag{4.69}$$

and

$$\langle N_{\mathbf{k}} | a_{\mathbf{k}}^{\dagger} a_{\mathbf{k}} | N_{\mathbf{k}} \rangle = N_{\mathbf{k}}. \tag{4.70}$$

We see that the second-quantized field operator $\mathbf{A}_{\mathbf{k}}$ has the same structure as the single-photon wave function (4.49), now noncovariantly normalized, except for the second-quantized operator $a_{\mathbf{k}}$, where $a_{\mathbf{k}}^{\dagger} a_{\mathbf{k}}$ is a number operator according to (4.70).

To proceed more rigorously, one treats (4.67) as a normal-mode harmonic-oscillator expansion and expresses the field energy as

$$E = \frac{1}{2} \int d^3x \, \langle \mathbf{E}^2 + \mathbf{B}^2 \rangle = \sum_k \hbar\omega(N_\mathbf{k} + \tfrac{1}{2}), \tag{4.71}$$

where use of (4.65), (4.67), and (4.69) leads to (4.71) with the identification (now being careful to respect the order of the operators $\mathbf{A}$ and $\mathbf{A}^\dagger$)

$$\langle N_\mathbf{k} | a_\mathbf{k}^\dagger a_\mathbf{k} + a_\mathbf{k} a_\mathbf{k}^\dagger | N_\mathbf{k} \rangle = 2N_\mathbf{k} + 1. \tag{4.72}$$

Then abstracting (4.71) to a hamiltonian operator expressed in terms of canonical coordinates and momenta, the commutation relation $[x_\mathbf{k}, p_k] = i\hbar$ leads to the second-quantized commutation relations (see Problem 4.6)

$$[a_\mathbf{k}, a_{\mathbf{k}'}^\dagger] = \delta_{\mathbf{k},\mathbf{k}'}, \qquad [a_\mathbf{k}, a_{\mathbf{k}'}] = 0. \tag{4.73}$$

It is therefore clear from (4.72) and (4.73) that $a_\mathbf{k}$ is an annihilation operator and $a_\mathbf{k}^\dagger$ a creation operator in Fock space satisfying

$$\langle N_\mathbf{k} - 1 | a_\mathbf{k} | N_\mathbf{k} \rangle = \langle N_\mathbf{k} | a_\mathbf{k}^\dagger | N_\mathbf{k} - 1 \rangle = \sqrt{N_\mathbf{k}} \tag{4.74}$$

and consistent with $a_\mathbf{k}^\dagger a_\mathbf{k}$ being the number operator as expected in (4.70).

**Spin and Statistics.** Second quantization is a natural way to build in the Bose statistics of the photon field from the outset via the commutation relations (4.73). Likewise the Bose statistics of spin-0 or spin-1 massive particles can be built in by commutation relations such as (4.73) to form a similar free quantum field theory. A field theory of fermions also can be constructed in this way, with the commutation relations (4.73) replaced by anticommutation relations in order to manifest the Fermi statistics. This connection between spin and statistics follows in a natural way from the requirement of positive definiteness for the free-field hamiltonian (Pauli 1940). Interacting quantum fields are described in terms of coordinate-space lagrangian densities. While the methods of quantum field theory are elegant and powerful, we shall follow the simpler and intuitive (one-particle) methods of Feyman for the greater part of this book. This means we will be investigating the structure of (one-particle) current and hamiltonian densities (usually in momentum space) rather than lagrangian densities (in coordinate space). The connection between spin and statistics will be invoked as a postulate, as elsewhere in quantum mechanics (postulate vi, Section 1.A); no reference will be made to the second-quantized commutation relations (4.73).

**Phonons.** Before leaving this subject, it will prove useful to consider a nonrelativistic quantum field construct, that of *phonons*, corresponding to boson quanta but associated with lattice vibrations in a solid. Like photons, phonons have no mass—but this is not dynamically relevant. Instead the relation between phonon energy (frequency) and momentum (wave number) depends upon a "dispersion law" resulting in $\omega_\mathbf{q} \approx$ constant at small $\mathbf{q}$ for "optical phonons" and $\omega_\mathbf{q} = c_s|\mathbf{q}|$ at small $|\mathbf{q}|$ for "acoustical phonons", where $c_s$ is the velocity of sound in a solid, $c_s \sim 10^5$ cm/sec. Since phonons are induced by displacements of lattice-site ions, they can be associated with a vector "spin", having longitudinal as well as transverse components. The scale of the phonon displacement amplitude $\mathbf{A}$ with Fourier coefficients $\mathbf{A}_\mathbf{q}$ and expansion (4.67) is set by the time-averaged displacement energy of a spring with spring constant $k = M\omega^2$, where $M$ is the lattice ion mass,

$$\langle E \rangle = 2 \times \tfrac{1}{2} M\omega_\mathbf{q}^2 |\mathbf{A}_\mathbf{q}|^2. \tag{4.75}$$

Since this classical energy represents only half the energy of the spring (the kinetic energy accounts for the other half, by the virial theorem), the quantum analog of (4.75) is $\frac{1}{2}\hbar\omega_\mathbf{q} N_\mathbf{q}$ for $N_\mathbf{q}$ phonons of angular frequency $\omega_\mathbf{q}$, giving

$$|\mathbf{A}_\mathbf{q}|^2 = \frac{\hbar N_\mathbf{q}}{2M\omega_\mathbf{q}}. \tag{4.76}$$

Paralleling the photon case (4.68) and (4.69), it would appear that the second-quantized phonon displacement field operator is

$$\mathbf{A}_\mathbf{q} = \sqrt{\frac{\hbar}{2M\omega_\mathbf{q}}}\, a_\mathbf{q} \boldsymbol{\varepsilon}(\mathbf{q}), \tag{4.77}$$

with $a_\mathbf{q}$ and $a_\mathbf{q}^\dagger$ satisfying the commutation relations (4.73) and $a_\mathbf{q}^\dagger a_\mathbf{q}$ is the number operator similar to (4.70). In Chapter 9 we shall employ the one-particle phonon wave function (noncovariantly normalized)

$$\mathbf{A}(\mathbf{x}, t) = \sqrt{\frac{\hbar}{2\rho\omega_\mathbf{q} V}}\, \boldsymbol{\varepsilon}(\mathbf{q}) e^{i\mathbf{q}\cdot\mathbf{x}} e^{-i\omega_\mathbf{q} t}, \tag{4.78}$$

where the ion mass $M$ in (4.77) is expressed as $\rho V$ in (4.78) so that the amplitude displays the usual one-particle box normalization. The ion density $\rho$ varies from 2 to 20 g/cm$^3$ in a solid, a typical value being 5 g/cm$^3$.

General references for boson wave equations are Hamilton (1959), Roman (1960), Schweber (1961), Bjorken and Drell (1964), Bethe and Jackiw (1968), Schiff (1968), Baym (1969), and Berestetskiĭ et al. (1971).

CHAPTER 5

# Spin-$\frac{1}{2}$ Dirac Equation

What the Schrödinger equation is to nonrelativistic physics, the Dirac equation is to relativistic physics. We begin this chapter by describing three alternative ways of deriving this spin-$\frac{1}{2}$ wave equation—the more the merrier, in order to develop as much intuition as possible about this fundamental dynamical tool. Next we formulate the Dirac equation in a manifest covariant manner and emphasize the structure of $\gamma$-matrix algebra and the positive and negative free-particle solutions. The Dirac equation in the presence of external fields is then generated by minimal replacement, and the resulting electron bound-state energies are obtained for the one-particle Coulomb atom and for a constant external magnetic field. We pay particular attention to the difference between the Dirac atom and the fine-structure level shifts in the Schrödinger atom. Finally, we develop free-particle Dirac equations for spin-$\frac{1}{2}$ massless neutrinos and spin-$\frac{3}{2}$ massive particles.

## 5.A Derivations of the Dirac Equation

The Dirac equation plays a fundamental role in any relativistic quantum theory, not only because it circumvents many of the problems arising from an unphysical interpretation of the Klein–Gordon equation, but also because it naturally describes the basic spin-$\frac{1}{2}$ constituents of matter at the atomic and nuclear level—electrons and nucleons (protons and neutrons). To appreciate the significance and beauty of the Dirac equation, it is well to describe three alternative derivations: the original relativistic approach of

Dirac, the *ad hoc* but elegant derivation via a Pauli spin-matrix replacement, and finally the group-theoretic derivation.

**Dirac's Derivation.** In order to obtain a positive definite probability density for spin-$\frac{1}{2}$ particles, Dirac searched for a relativistic differential equation which was first order in time (Dirac 1928, 1958). To this end he interpreted the Schrödinger equation, $(i\partial_t - H_0)\psi = 0$, as generated for spin-$\frac{1}{2}$ particles by a free-particle matrix hamiltonian $H_0$, first order in momentum for relativistic considerations:

$$H_0 = \boldsymbol{\alpha} \cdot \mathbf{p} + \beta m. \tag{5.1}$$

Here $\alpha_i$ ($i = 1, 2, 3$) and $\beta$ are four matrices to be determined, and $\mathbf{p}$ is the usual quantum-mechanical momentum operator $-i\nabla$ for a particle of mass $m$. Since the single-component spin-0 Klein–Gordon equation leads to a *two*-component first order matrix (Feshbach–Villars) equation (4.20), it is natural to suppose that a two-component spin-$\frac{1}{2}$ equation, second order in time, should be linked to a four-component matrix equation, first order in time. Such a "guess" was all the more impressive at the time because Dirac did not have the hindsight of the Feshbach–Villars equation. Thus one assumes that $\alpha_i$ and $\beta$ are $4 \times 4$ matrices and that the Dirac wave function $\psi \to \psi_\sigma$ is a four-component (Dirac) bispinor. The second-order equation in time for a free particle must, of course, be of the Klein–Gordon form, which in momentum space is obtained from the first-order equation via multiplication by $E + H_0$:

$$(E - H_0)\psi = 0 \to (E^2 - H_0^2)\psi = 0. \tag{5.2}$$

Dirac then demanded that the square of the $4 \times 4$ matrix hamiltonian (5.1) be constrained to $E^2 = \mathbf{p}^2 + m^2$ by (5.2), i.e.,

$$H_0^2 = \tfrac{1}{2}(\alpha_i\alpha_j + \alpha_j\alpha_i)p_i p_j + (\beta\alpha_i + \alpha_i\beta)p_i + \beta^2 m^2 = \mathbf{p}^2 + m^2. \tag{5.3}$$

This leads to the defining properties of the matrices $\alpha_i$ and $\beta$:

$$\{\alpha_i, \alpha_j\} = 2\delta_{ij}, \quad \{\beta, \alpha_i\} = 0, \quad \alpha_i^2 = \beta^2 = 1. \tag{5.4}$$

Furthermore, since $H_0$ must be an observable hermitian operator, so must $\alpha_i$ and $\beta$ be hermitian:

$$\alpha_i^\dagger = \alpha_i, \qquad \beta^\dagger = \beta. \tag{5.5}$$

The adjoint row bispinor $\psi^\dagger$ can be combined with the column bispinor $\psi$ to form a positive definite probability density

$$\rho = \psi^\dagger\psi = \sum_{\sigma=1}^{4} \psi_\sigma^*\psi_\sigma, \tag{5.6}$$

now naturally linked with the hermitian probability current density

$$\mathbf{j} = \psi^\dagger\boldsymbol{\alpha}\psi. \tag{5.7}$$

A justification for the form (5.7) is the resulting continuity equation $\partial_t \rho + \nabla \cdot \mathbf{j} = 0$, which follows from (5.6), (5.7), and the Dirac equation

$$(i\partial_t - H_0)\psi = 0, \tag{5.8}$$

where $H_0$ obeys (5.1) and (5.3) (see Problem 5.1).

From the properties (5.4), it is clear that $\alpha_i = -\beta\alpha_i\beta$ and $\beta = -\alpha_1\beta\alpha_1$, which means that these matrices are traceless (Tr $\beta = \sum_{\sigma=1}^{4} \beta_{\sigma\sigma}$):

$$\text{Tr}\ \alpha_i = \text{Tr}\ \beta = 0, \tag{5.9}$$

since Tr $BA$ = Tr $AB$. It is then useful to describe $\alpha_i$ and $\beta$ by a specific representation, such as the Dirac–Pauli representation,

$$\boldsymbol{\alpha} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix}, \qquad \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{5.10}$$

[which, of course, satisfies the general properties (5.4), (5.5), and (5.9)], where the elements, $\boldsymbol{\sigma}$ and 1, are themselves the $2 \times 2$ Pauli and identity matrices, obeying

$$\sigma_i \sigma_j = \delta_{ij} 1 + i\varepsilon_{ijk}\sigma_k. \tag{5.11}$$

Dirac used this representation (5.10) to obtain the first (but by no means the last) profound prediction of the Dirac equation. Making a two-component reduction of (5.8) in the presence of an electromagnetic field via the Dirac–Pauli representation (5.10), he discovered that the electron must have the unique magnetic (dipole) moment

$$\boldsymbol{\mu}_e = 2 \times \frac{e}{2m_e} \times \frac{\boldsymbol{\sigma}}{2} = \frac{e}{2m_e}\boldsymbol{\sigma}, \tag{5.12}$$

i.e., the Landé $g$-factor is $g_e = 2$, a result in almost perfect agreement with experiment and assumed *ad hoc* up to that time. We shall return to Dirac's method of predicting $g_e = 2$ in Section 5.D and to the small but important corrections to this result in Chapter 15.

**Derivation via Pauli-Spin-Matrix Replacement.** This leads naturally to the second derivation of the Dirac equation with $g_e = 2$ as input [see, e.g., van der Waerden (1932), Sakurai (1967)]. Recall from (5.11) that the square of any vector such as $\mathbf{p}^2$ is equivalent to the Pauli-spin-matrix replacement $(\boldsymbol{\sigma} \cdot \mathbf{p})^2$. The question then arises as to when this substitution is required, i.e., for all or only a few problems involving spin-$\frac{1}{2}$ particles. Furthermore, while $(\boldsymbol{\sigma} \cdot \mathbf{p})^2 = \mathbf{p}^2$, it is clear that in the presence of an electromagnetic field, the minimal replacement $\mathbf{p} \to \boldsymbol{\pi} = \mathbf{p} - e\mathbf{A}$ means $(\boldsymbol{\sigma} \cdot \boldsymbol{\pi})^2 \neq \boldsymbol{\pi}^2$, but instead (5.11) implies, along with $\mathbf{p} \times \mathbf{A} = -i\nabla \times \mathbf{A} - \mathbf{A} \times \mathbf{p}$, that

$$(\boldsymbol{\sigma} \cdot \boldsymbol{\pi})^2 - \boldsymbol{\pi}^2 = i\boldsymbol{\sigma} \cdot \boldsymbol{\pi} \times \boldsymbol{\pi} = -e\boldsymbol{\sigma} \cdot \mathbf{B}. \tag{5.13}$$

Need nonrelativistic or relativistic momentum operators be subject to this Pauli-spinor replacement? Consider first a nonrelativistic free-particle hamiltonian with $\mathbf{p}^2/2m \to (\boldsymbol{\sigma} \cdot \mathbf{p})^2/2m$. In the presence of an electromagnetic field, (5.13) then leads to the Pauli hamiltonian

$$H = \frac{1}{2m}(\boldsymbol{\sigma} \cdot \mathbf{p})^2 \to \frac{1}{2m}\boldsymbol{\pi}^2 - \frac{e}{2m}\boldsymbol{\sigma} \cdot \mathbf{B}. \tag{5.14}$$

The spin-dependent term in (5.14) has the form $-\boldsymbol{\mu} \cdot \mathbf{B}$, with $\boldsymbol{\mu} = g\mu_B \mathbf{J}$ where $\mathbf{J} = \boldsymbol{\sigma}/2$ and $g = 2$, a very desirable result.

Given this success as input, we are encouraged to apply this mnemonic to all problems involving $\mathbf{p}^2$ and spin $\frac{1}{2}$. In particular we modify the spin-$\frac{1}{2}$ Klein–Gordon equation to

$$[E^2 - (\boldsymbol{\sigma} \cdot \mathbf{p})^2]\phi = m^2\phi, \tag{5.15a}$$

which then has the time-dependent factored form

$$(i\partial_t - \boldsymbol{\sigma} \cdot \mathbf{p})(i\partial_t + \boldsymbol{\sigma} \cdot \mathbf{p})\phi = m^2\phi. \tag{5.15b}$$

Now we parallel the discussion of the Feshbach–Villars version of the Klein–Gordon equation (Section 4.A) by defining

$$\varphi_L \equiv \phi, \qquad \varphi_R \equiv \frac{1}{m}(i\partial_t + \boldsymbol{\sigma} \cdot \mathbf{p})\phi \tag{5.16}$$

and expressing the second-order equation (5.15) as two coupled first-order equations,

$$(i\partial_t + \boldsymbol{\sigma} \cdot \mathbf{p})\varphi_L = m\varphi_R, \qquad (i\partial_t - \boldsymbol{\sigma} \cdot \mathbf{p})\varphi_R = m\varphi_L. \tag{5.17}$$

Further defining

$$\varphi = \varphi_R + \varphi_L, \qquad \chi = \varphi_R - \varphi_L, \tag{5.18}$$

and using the Dirac–Pauli representation (5.10), the first-order equations (5.17) for these two-component spinors can be put into the four-component Dirac-equation form $(i\partial_t - H_0)\psi = 0$, with $\psi = \binom{\varphi}{\chi}$ and

$$H_0 = \begin{pmatrix} m & \boldsymbol{\sigma} \cdot \mathbf{p} \\ \boldsymbol{\sigma} \cdot \mathbf{p} & -m \end{pmatrix} = \boldsymbol{\alpha} \cdot \mathbf{p} + \beta m, \tag{5.19}$$

again the Dirac free-particle hamiltonian. Setting $\mathbf{p} = 0$ in (5.17) means that $\varphi_L(0) = \varphi_R(0)$ and $\chi(0) = 0$ for positive-energy states $\varphi_{L,R} \propto e^{-iEt}$, whereas $\varphi_L(0) = -\varphi_R(0)$ and $\varphi(0) = 0$ for negative-energy states $\varphi_{L,R} \propto e^{iEt}$. In four-component language with $\psi_\pm \propto e^{\mp iEt}$, these $\mathbf{p} = 0$ constraints can be expressed as

$$\beta\psi_\pm(0) = \pm\psi_\pm(0), \tag{5.20}$$

stating that the Dirac–Pauli representation diagonalizes the bispinor energy states in the extreme nonrelativistic limit.

Alternatively we may use another representation for $\alpha_i$ and $\beta$, called the Weyl representation

$$\boldsymbol{\alpha} = \begin{pmatrix} -\boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix}, \qquad \beta = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{5.21}$$

also satisfying the defining relations (5.4), (5.5), and (5.9). Given (5.21), the two-component equations (5.17) again can be combined into the four-component Dirac form (5.8) provided that

$$\psi = \begin{pmatrix} \varphi_L \\ \varphi_R \end{pmatrix}$$

in this representation (Problem 5.1). In Section 5.E we shall show that for $m = 0$, (5.17) indicates that $\varphi_R$ is polarized right-handed, corresponding to helicity $\lambda = +\frac{1}{2}$, and $\varphi_L$ is polarized left-handed, with $\lambda = -\frac{1}{2}$ [see also (3.94)]. Thus we see that the Weyl representation diagonalizes bispinor helicity states in the extreme relativistic ($m = 0$) limit.

**Group-Theory Derivation.** The third derivation of the Dirac equation combines the Weyl representation for $\alpha_i$ and $\beta$ with the group-theoretical formulation of relativistic spin-$\frac{1}{2}$ states in Section 3.B. The motivation is to combine the $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$ irreducible representations of the homogeneous Lorentz group into a larger matrix $(\frac{1}{2}, 0) + (0, \frac{1}{2})$ representation $\mathscr{D}$, which is then equivalent to the adjoint representation $\bar{\mathscr{D}}(\Lambda) = \beta\mathscr{D}^\dagger(\Lambda^{-1})\beta$. Not by accident, this $\beta$ corresponds to the Dirac $\beta$ in the Weyl representation (5.21).

For the two-component $(\frac{1}{2}, 0)$ boost $D(L_{\mathbf{p}})$ and the $(0, \frac{1}{2})$ boost $\bar{D}(L_{\mathbf{p}})$, the corresponding wave functions $\varphi_R$ and $\varphi_L$ satisfy the uncoupled equations

$$\begin{aligned} \varphi_L(\mathbf{p}) &= D(L_{\mathbf{p}})\varphi_L(0) \\ &= [2m(E+m)]^{-\frac{1}{2}}(E + m - \boldsymbol{\sigma}\cdot\mathbf{p})\varphi_L(0), \end{aligned} \tag{5.22a}$$

$$\begin{aligned} \varphi_R(\mathbf{p}) &= \bar{D}(L_{\mathbf{p}})\varphi_R(0) \\ &= [2m(E+m)]^{-\frac{1}{2}}(E + m + \boldsymbol{\sigma}\cdot\mathbf{p})\varphi_R(0), \end{aligned} \tag{5.22b}$$

from which one can derive the coupled relations (5.17) (see Problem 5.1). Now use (5.22) along with the matrix boost obtained from (3.51) in the Weyl representation as

$$\mathscr{D}(L_{\mathbf{p}}) = \begin{pmatrix} D(L_{\mathbf{p}}) & 0 \\ 0 & \bar{D}(L_{\mathbf{p}}) \end{pmatrix} = [2m(E+m)]^{-\frac{1}{2}}[(E+m)1 + \boldsymbol{\alpha}\cdot\mathbf{p}], \tag{5.23}$$

where 1 denotes the $4 \times 4$ unit matrix (sometimes deleted). Again expressing the Dirac bispinor in the Weyl representation as $\psi = \binom{\varphi_L}{\varphi_R}$, (5.22) and (5.23) boost $\psi(0)$ from rest to

$$\psi(\mathbf{p}) = \mathscr{D}(L_{\mathbf{p}})\psi(0). \tag{5.24}$$

Finally, to recover the Dirac equation, operate on (5.24) with $i\partial_t - H_0$, obtaining

$$\begin{aligned} (i\partial_t - H_0)\psi_\pm(\mathbf{p}) &\propto [\pm E - (\boldsymbol{\alpha}\cdot\mathbf{p} + \beta m)][\pm E + \boldsymbol{\alpha}\cdot\mathbf{p} + m]\psi_\pm(0) \\ &= (E^2 - \mathbf{p}^2 - m^2)\psi_\pm(0) \\ &\quad + m(\pm E + m - \boldsymbol{\alpha}\cdot\mathbf{p})(1 \mp \beta)\psi_\pm(0). \end{aligned} \tag{5.25}$$

The first term on the right-hand side of (5.25) vanishes by the energy–momentum relation, and the second term is zero by (5.20). Thus the Dirac equation $(i\partial_t - H_0)\psi = 0$ is equivalent to the Klein–Gordon constraint *plus* (5.20), which projects out the positive- or negative-energy parts of $\psi$. Put another way, the Dirac equation is the boosted form of the rest-frame energy-eigenvalue relation (5.20) subject to the energy–momentum (mass-shell) constraint $E^2 - \mathbf{p}^2 = m^2$ for $\mathbf{p}$ in an arbitrary direction.

In Chapter 6 we shall learn that (5.20) has invariant meaning as the spatial-inversion (parity) eigenvalue relation. Combining two-component spinors which transform into one another under spatial inversion as in (5.23) means that $\mathscr{D}$ is again equivalent to the parity transform via (3.47), $\beta\mathscr{D}(L_{\mathbf{p}})\beta = \mathscr{D}(L_{-\mathbf{p}})$. This is sometimes considered the main motivation for the four-component Dirac formalism. Another is the fermion mass; if it is zero then the two-component equations (5.17) suffice.

## 5.B Covariant Formulation

**Covariant γ-Matrices.** It is possible to formulate this relativistic four-component formalism in manifestly covariant language by expressing the Dirac matrices $\beta$ and $\boldsymbol{\alpha}$ in terms of covariant $\gamma$-matrices $\gamma_\mu = (\gamma_0, \boldsymbol{\gamma})$, defined as

$$\gamma_0 \equiv \beta, \qquad \boldsymbol{\gamma} \equiv \beta\boldsymbol{\alpha}. \tag{5.26}$$

These matrices have the "lengths" $(\gamma_i^2 = \gamma_1^2 = \gamma_2^2 = \gamma_3^2)$

$$\gamma_0^2 = -\gamma_i^2 = 1, \qquad \gamma_\mu\gamma^\mu = \gamma \cdot \gamma = 4 \tag{5.27}$$

and satisfy fundamental anticommutation relations following from (5.4):

$$\{\gamma_\mu, \gamma_\nu\} = \gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu = 2g_{\mu\nu}. \tag{5.28}$$

Note that (5.28) has the form of a covariant symmetric tensor equation; this implication will be discussed shortly. Next combine the hermitian and antihermitian relations $\gamma_0^\dagger = \gamma_0$ and $\gamma_i^\dagger = -\gamma_i$ into a "Dirac adjoint" matrix operation

$$\bar{\gamma}_\mu \equiv \gamma_0\gamma_\mu^\dagger\gamma_0 = \gamma_\mu, \tag{5.29}$$

where in general the Dirac adjoint (barred operation) of any $4 \times 4$ matrix is $\bar{A} = \gamma_0 A^\dagger \gamma_0$ with $\overline{AB} = \bar{B}\bar{A}$. That is, (5.29) implies that $\gamma_\mu$ is "self-barred". Then define the new $\gamma$-matrices

$$\gamma_5 \equiv \gamma_0\gamma_1\gamma_2\gamma_3 \tag{5.30a}$$

satisfying

$$\bar{\gamma}_5 = \gamma_5, \qquad \gamma_5^2 = -1, \qquad \gamma_5\gamma_\mu = -\gamma_\mu\gamma_5, \tag{5.30b}$$

$$\sigma_{\mu\nu} \equiv \tfrac{1}{2}i[\gamma_\mu, \gamma_\nu] = i(\gamma_\mu\gamma_\nu - g_{\mu\nu}) = -\sigma_{\nu\mu} \tag{5.31a}$$

satisfying $\bar{\sigma}_{\mu\nu} = \sigma_{\mu\nu}$, since $\bar{i} = -i$ and

$$\sigma_{ij} = \varepsilon_{ijk}\sigma_k, \qquad \sigma_i = \tfrac{1}{2}\varepsilon_{ijk}\sigma_{jk}, \tag{5.31b}$$

$$\sigma_{0i} = i\gamma_0\gamma_i = i\alpha_i = -\gamma_5\sigma_i, \tag{5.31c}$$

where $\sigma_i$ are now $4 \times 4$ matrices, built up from the $2 \times 2$ Pauli matrices as $\begin{pmatrix}\sigma & \\ & \sigma\end{pmatrix}$. In the latter case, $\boldsymbol{\alpha} = i\gamma_5\boldsymbol{\sigma}$ follows from $\alpha_1 = i\gamma_0\gamma_1\gamma_2\gamma_3 i\gamma_2\gamma_3 = \gamma_0\gamma_1$, etc.

Since $4 \times 4$ matrices have 16 possible elements, there must be 16 independent $\gamma$-matrices. We may choose these independent matrices as the self-barred set

$$\Gamma_i = 1,\ \gamma_\mu,\ \sigma_{\mu\nu},\ i\gamma_\mu\gamma_5,\ \gamma_5 = S(1),\ V(4),\ T(6),\ A(4),\ P(1), \tag{5.32}$$

where 1 is the one scalar ($S$) unit matrix, $\gamma_\mu$ are the four vector ($V$) matrices, $\sigma_{\mu\nu}$ the six antisymmetric tensor ($T$) matrices, $i\gamma_\mu\gamma_5$ the four axial-vector ($A$) matrices, and $\gamma_5$ the one pseudoscalar ($P$) matrix, so that in all there are $1 + 4 + 6 + 4 + 1 = 16$ independent matrices. The transformation laws for the various $\Gamma_i$ verify their independence; this will be demonstrated later along with the implications of the "pseudo" property of $A$ and $P$. Note the explicit factor of $i$ in $i\gamma_\mu\gamma_5$. It preserves the relation that all 16 $\Gamma_i$ are self-barred, i.e., $\overline{i\gamma_\mu\gamma_5} = \bar{\gamma}_5\bar{\gamma}_\mu\bar{i} = -i\gamma_5\gamma_\mu = i\gamma_\mu\gamma_5$. Thus $\bar{\Gamma}_i = \Gamma_i$. Beware, however, of the other possible (non-self-barred) choices for $\gamma_5$ appearing in the literature, $\pm i\gamma_0\gamma_1\gamma_2\gamma_3$. Finally, any general $4 \times 4$ matrix can be expanded in terms of these 16 $\Gamma_i$. In particular, one can verify that (see Problem 5.2)

$$\gamma_\mu\gamma_\nu = g_{\mu\nu} - i\sigma_{\mu\nu}, \tag{5.33a}$$

$$\gamma_5\sigma_{\mu\nu} = \tfrac{1}{2}\varepsilon_{\mu\nu\alpha\beta}\sigma^{\alpha\beta}, \tag{5.33b}$$

$$\gamma_\mu\gamma_\nu\gamma_\rho = g_{\mu\nu}\gamma_\rho - g_{\mu\rho}\gamma_\nu + g_{\nu\rho}\gamma_\mu - \varepsilon_{\mu\nu\rho\sigma}\gamma^\sigma\gamma_5, \tag{5.33c}$$

$$\varepsilon_{\mu\nu\rho\sigma}\gamma^\nu\gamma^\rho\gamma^\sigma = -3!\ \gamma_\mu\gamma_5, \tag{5.33d}$$

$$\varepsilon_{\mu\nu\rho\sigma}\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma = -4!\ \gamma_5, \tag{5.33e}$$

$$\begin{aligned}\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma = {} & g_{\mu\nu}\gamma_\rho\gamma_\sigma - g_{\mu\rho}\gamma_\nu\gamma_\sigma + g_{\nu\rho}\gamma_\mu\gamma_\sigma + g_{\rho\sigma}\gamma_\mu\gamma_\nu - g_{\nu\sigma}\gamma_\mu\gamma_\rho \\ & + g_{\mu\sigma}\gamma_\nu\gamma_\rho - g_{\mu\nu}g_{\rho\sigma} + g_{\mu\rho}g_{\nu\sigma} - g_{\mu\sigma}g_{\nu\rho} + \gamma_5\varepsilon_{\mu\nu\rho\sigma},\end{aligned} \tag{5.33f}$$

etc., where the $\gamma$-matrix products on the left-hand sides have been expressed as linear combinations of the 16 $\Gamma_i$ on the right-hand sides of (5.33), and we have used the fundamental properties of the Levi–Civita pseudotensor ($\varepsilon_{0123} = 1$),

$$\varepsilon_{\mu\nu\rho\sigma}\varepsilon^{\mu\nu\rho\sigma} = -4!, \tag{5.34a}$$

$$\varepsilon_{\mu\nu\rho\sigma}\varepsilon_\alpha{}^{\nu\rho\sigma} = -3!\ g_{\mu\alpha}, \tag{5.34b}$$

$$\varepsilon_{\mu\nu\rho\sigma}\varepsilon_{\alpha\beta}{}^{\rho\sigma} = -2!\begin{vmatrix} g_{\mu\alpha} & g_{\nu\alpha} \\ g_{\mu\beta} & g_{\nu\beta}\end{vmatrix}, \tag{5.34c}$$

etc.

As regards the matrices $\beta$ and $\boldsymbol{\alpha}$, the $\gamma$-matrices satisfying only (5.27) and (5.28) then have many possible representations. The representation-independence (Pauli–Good) theorem states that all representations of $\gamma$-matrices are equivalent up to a similarity transformation, $\gamma'_\mu = S^{-1}\gamma_\mu S$. The *Dirac–Pauli* representation, diagonalizing energy via $\gamma_0$ in the extreme non-relativistic limit, is

$$\begin{aligned} \gamma_0 &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad i\gamma_5 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ \boldsymbol{\gamma} &= \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & 0 \end{pmatrix}, \qquad \boldsymbol{\sigma} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix}, \end{aligned} \tag{5.35}$$

whereas the *Weyl* representation, diagonalizing helicity via $\gamma_5$ in the extreme relativistic limit, is

$$\begin{aligned} \gamma_0 &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad i\gamma_5 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \\ \boldsymbol{\gamma} &= \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & 0 \end{pmatrix}, \qquad \boldsymbol{\sigma} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix}. \end{aligned} \tag{5.36}$$

Other representations, including the Majorana and the light-plane representation, are worked out in Problem 5.3.

**Covariant Transformation Laws.** To put the Dirac equation $(i\partial_t - H_0)\psi = 0$ on a covariant footing so as not to single out time via $\partial_t$, multiply the equation on the left by $\gamma_0$—obtaining $(i\gamma_0\partial_0 - \boldsymbol{\gamma}\cdot\mathbf{p} - m)\psi = 0$—and use $\mathbf{p} = -i\boldsymbol{\nabla}$. This results in the covariant form of the Dirac equation,

$$(i\gamma\cdot\partial - m)\psi(x) = 0. \tag{5.37}$$

Henceforth it will be convenient to define a "slash" operation $\not{A} \equiv \gamma\cdot A$, so that (5.37) takes on the compact form $(i\not{\partial} - m)\psi = 0$. To verify that (5.37) is indeed covariant under homogeneous Lorentz transformations $x' = \Lambda x$, we rely upon the discussion in Chapters 2 and 3 to obtain a $4\times 4$ spinor matrix $\mathscr{S}(\Lambda)$ which transforms the bispinor wave functions in a fashion analogous to (2.1c),

$$\psi'(x') = \mathscr{S}(\Lambda)\psi(x). \tag{5.38}$$

Since the Hilbert-space operator is $U_\Lambda = \exp(-i\omega^{\mu\nu}J_{\mu\nu}/2)$, the obvious identification $J_{\mu\nu} \to \sigma_{\mu\nu}/2$ in the spin-$\frac{1}{2}$ Dirac space means we can write

$$\mathscr{S}(\Lambda) = \exp(-i\omega^{\mu\nu}\sigma_{\mu\nu}/4). \tag{5.39}$$

An infinitesimal transformation $\Lambda_{\mu\nu} \to g_{\mu\nu} + \omega_{\mu\nu}$ induces (5.39) to become $\mathscr{S}(\Lambda) \to 1 - i\omega^{\mu\nu}\sigma_{\mu\nu}/4$, and the $\gamma$-matrix identity (see Problem 5.2)

$$[\sigma_{\mu\nu}, \gamma_\rho] = 2i(g_{\rho\nu}\gamma_\mu - g_{\rho\mu}\gamma_\nu) \tag{5.40}$$

then leads to the transformation law for $\gamma$-matrices,

$$\mathscr{S}^{-1}(\Lambda)\gamma_\mu\mathscr{S}(\Lambda) = \Lambda_{\mu\nu}\gamma^\nu. \tag{5.41}$$

[Note that (5.41) bears the same relation to the Lorentz group as (2.15) and (2.52) do to the rotation group.] It is in the sense of (5.41) that $\gamma_\mu$ transforms like a four-vector; since $\gamma_\mu$ is a constant matrix, $\gamma'_\mu$ has no meaning other than (5.41). Then we establish the covariance of (5.37) by multiplying it on the left with $\mathscr{S}(\Lambda)$:

$$\begin{aligned} 0 &= \mathscr{S}(\Lambda)(i\gamma \cdot \partial - m)\psi(x) \\ &= [\mathscr{S}(\Lambda)\gamma_\mu \mathscr{S}^{-1}(\Lambda) i\partial^\mu - m]\mathscr{S}(\Lambda)\psi(x) \\ &= (i\Lambda^{-1}_{\mu\nu}\gamma^\nu\partial^\mu - m)\psi'(x') = (i\gamma \cdot \partial' - m)\psi'(x'), \end{aligned} \tag{5.42}$$

where we have used $\partial'_\nu = (\partial x_\mu / \partial x'^\nu)\partial^\mu = \Lambda^{-1}_{\mu\nu}\partial^\mu$. Thus, (5.42) implies that the Dirac equation (5.37) is valid in any frame $x' = \Lambda x$, an explicit demonstration of the meaning of covariance.

Now we extend the Dirac adjoint or "bar" operation to bispinor column vectors. Define the row bispinor

$$\bar{\psi} \equiv \psi^\dagger \gamma_0. \tag{5.43}$$

If $\psi$ satisfies (5.37), then $\bar{\psi}$ satisfies the adjoint Dirac equation

$$\overline{(i\not{\partial} - m)\psi} = \bar{\psi}(-i\overleftarrow{\not{\partial}} - m) = 0, \tag{5.44}$$

since $\bar{\gamma}_\mu = \gamma_\mu$. Then using $\bar{\mathscr{S}}(\Lambda) = \mathscr{S}^{-1}(\Lambda)$, which follows from (5.39) and $\bar{\sigma}_{\mu\nu} = \sigma_{\mu\nu}$, we have from (5.38)

$$\bar{\psi}'(x') = \bar{\psi}(x)\bar{\mathscr{S}}(\Lambda) = \bar{\psi}(x)\mathscr{S}^{-1}(\Lambda). \tag{5.45}$$

Combining (5.38), (5.41), and (5.45), we can obtain the transformation laws for the bilinear covariants $\bar{\psi}\Gamma_i\psi$ (now "$c$-numbers" in the Dirac space, since a row times a column matrix is a pure number—see Problem 5.4),

$$S: \quad \bar{\psi}'(x')\psi'(x') = \bar{\psi}(x)\psi(x) \tag{5.46a}$$

$$V: \quad \bar{\psi}'(x')\gamma_\mu\psi'(x') = \Lambda_\mu{}^\nu\bar{\psi}(x)\gamma_\nu\psi(x) \tag{5.46b}$$

$$T: \quad \bar{\psi}'(x')\sigma_{\mu\nu}\psi'(x') = \Lambda_\mu{}^\alpha\Lambda_\nu{}^\beta\bar{\psi}(x)\sigma_{\alpha\beta}\psi(x) \tag{5.46c}$$

$$A: \quad \bar{\psi}'(x')i\gamma_\mu\gamma_5\psi'(x') = (\det \Lambda)\Lambda_\mu{}^\nu\bar{\psi}(x)i\gamma_\nu\gamma_5\psi(x) \tag{5.46d}$$

$$P: \quad \bar{\psi}'(x')\gamma_5\psi'(x') = (\det \Lambda)\bar{\psi}(x)\gamma_5\psi(x). \tag{5.46e}$$

It is in this sense that the $\Gamma_i$ transform like integral representations of $\mathscr{L}$, indicating, for example, that the Dirac probability current (5.6) and (5.7), expressed in covariant form as

$$j_\mu = \psi^\dagger(1, \boldsymbol{\alpha})\psi = \bar{\psi}(\gamma_0, \boldsymbol{\gamma})\psi = \bar{\psi}\gamma_\mu\psi, \tag{5.47}$$

transforms as a four-vector, $j'_\mu(x') = \Lambda_\mu{}^\nu j_\nu(x)$, with a continuity equation $\partial \cdot j = 0$ which is invariant from frame to frame.

**Dirac Trace Algebra.** The bilinear covariants just discussed can be used not only to construct the probability current $j_\mu$, but also to form the $S$-matrix or transition-probability matrix elements for a process involving a spin-$\frac{1}{2}$ particle. In analogy with (3.87) we write

$$\langle \lambda' | S | \lambda \rangle = \bar{\psi}^{(\lambda')} M \psi^{(\lambda)}, \tag{5.48}$$

where $\psi^{(\lambda)}$ and $\bar{\psi}^{(\lambda')}$ are column and row bispinors which are eigenstates of helicity, and $M$ is a scalar $4 \times 4$ matrix composed of invariant products of momenta and $\gamma$-matrices. The complex conjugate (*) of (5.48) can be expressed in terms of the Dirac adjoint matrix, $\bar{M} = \gamma_0 M^\dagger \gamma_0$, according to

$$\begin{aligned} \langle \lambda' | S | \lambda \rangle^* &= (\bar{\psi}^{(\lambda')} M \psi^{(\lambda)})^\dagger = \psi^{\dagger(\lambda)} M^\dagger \bar{\psi}^{\dagger(\lambda')} \\ &= \psi^{\dagger(\lambda)} \gamma_0 \gamma_0 M^\dagger \gamma_0 \psi^{(\lambda')} = \bar{\psi}^{(\lambda)} \bar{M} \psi^{(\lambda')}. \end{aligned} \tag{5.49}$$

The physical quantity of interest is the transition probability, which corresponds to the product of (5.48) and (5.49); summing over all possible helicity states (unpolarized spin sum), we write

$$\sum_{\lambda', \lambda} |\langle \lambda' | S | \lambda \rangle|^2 = \sum_{\lambda'} \bar{\psi}^{(\lambda')} M \mathscr{P} \bar{M} \psi^{(\lambda')}, \tag{5.50}$$

where $\mathscr{P}$ is a Dirac "projection" operator

$$\mathscr{P} = \sum_{\lambda} \psi^{(\lambda)} \bar{\psi}^{(\lambda)} \tag{5.51}$$

obeying $\mathscr{P}^2 = \mathscr{P}$ for orthonormal bispinors $\bar{\psi}^{(\lambda')} \psi^{(\lambda)} = \delta_{\lambda' \lambda}$. Rearranging the bispinors in (5.50) to the form $(M \mathscr{P} \bar{M})_{\sigma' \sigma} \psi_\sigma^{(\lambda')} \bar{\psi}_{\sigma'}^{(\lambda')}$, a trace operation in the Dirac space ($\mathrm{Tr}\, A = \sum_\sigma A_{\sigma\sigma}$), and using (5.51) again, we obtain finally the unpolarized spin sum

$$\sum_{\lambda', \lambda} |\langle \lambda' | S | \lambda \rangle|^2 = \mathrm{Tr}\, M \mathscr{P} \bar{M} \mathscr{P}'. \tag{5.52}$$

Now we shall show shortly that the projection operators $\mathscr{P}$ and $\mathscr{P}'$ can be expressed in terms of $\gamma$-matrices, so (5.52) represents in principle a trace over the product of many $\gamma$-matrices. Consequently it will prove useful to describe the trace algebra of $\gamma$-matrices. To begin with, in the Dirac space

$$\mathrm{Tr}\ 1 = 4. \tag{5.53a}$$

Inspection of the Dirac–Pauli or Weyl representation [(5.35) and (5.36)] or application of the defining relations (5.28) and (5.4) along with the property $\mathrm{Tr}\, AB = \mathrm{Tr}\, BA$ gives

$$\mathrm{Tr}\ \gamma_\mu = \mathrm{Tr}\ \gamma_5 = 0, \tag{5.53b}$$

$$\mathrm{Tr}\ \gamma_\mu \gamma_\nu = \tfrac{1}{2}\, \mathrm{Tr}\{\gamma_\mu, \gamma_\nu\} = g_{\mu\nu}\ \mathrm{Tr}\ 1 = 4 g_{\mu\nu}, \tag{5.53c}$$

and from (5.30b) and (5.33b),

$$\mathrm{Tr}\ \gamma_5 \gamma_\mu = 0, \qquad \mathrm{Tr}\ \gamma_5 \gamma_\mu \gamma_\nu = -\tfrac{1}{2} i \varepsilon_{\mu\nu\alpha\beta}\ \mathrm{Tr}\ \sigma^{\alpha\beta} = 0. \tag{5.53d}$$

Next, note that (5.33c) then implies $\mathrm{Tr}\ \gamma_\mu \gamma_\nu \gamma_\rho = 0$, which can be extended to a product of any odd number of $\gamma$-matrices $\gamma_{\text{odd}}$ (excluding $\gamma_5$, which is an even product $\gamma_0 \gamma_1 \gamma_2 \gamma_3$) via multiplication of (5.33c) by successive pairs of $\gamma$-matrices, implying that

$$\mathrm{Tr}\ \gamma_{\text{odd}} = 0. \tag{5.53e}$$

For even numbers of $\gamma$-matrices, (5.53a) and (5.53c) are nonvanishing; for four $\gamma$'s, use of (5.33c) or (5.33f) leads to

$$\mathrm{Tr}\ \gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma = 4(g_{\mu\nu}g_{\rho\sigma} - g_{\mu\rho}g_{\nu\sigma} + g_{\mu\sigma}g_{\nu\rho}), \tag{5.53f}$$

$$\mathrm{Tr}\ \gamma_5\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma = -4\varepsilon_{\mu\nu\rho\sigma}, \tag{5.53g}$$

with $\varepsilon_{0123} = 1$. The last identity may be verified by choosing $\mu = 0$, $\nu = 1$, $\rho = 2$, $\sigma = 3$, and using (5.30a). Multiplication of (5.33f) by $\gamma_\sigma\gamma_\lambda$ leads to the trace of six $\gamma$-matrices, but at this point and thereafter it is easier to apply repeatedly the anticommutation relation (5.28) to permute $\gamma_\mu$ from one side of an even number of $\gamma$-matrices to the other and then use $\mathrm{Tr}\ A\gamma_\mu = \mathrm{Tr}\ \gamma_\mu A$ to find in general (see Problem 5.5)

$$\begin{aligned}\mathrm{Tr}\ \gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma \cdots &= g_{\mu\nu}\ \mathrm{Tr}\ \gamma_\rho\gamma_\sigma \cdots - g_{\mu\rho}\ \mathrm{Tr}\ \gamma_\nu\gamma_\sigma \cdots \\ &\quad + g_{\mu\sigma}\ \mathrm{Tr}\ \gamma_\nu\gamma_\rho \cdots - \text{etc.}\end{aligned} \tag{5.53h}$$

It is clear from the algebra (5.53) that including two additional $\gamma$-matrices in a trace can lead to a considerable complication. Quite often, however, the indices of two of the $\gamma$'s are summed (contracted), in which case repeated applications of (5.28) results in (see Problem 5.5)

$$\gamma_\alpha\gamma_\mu\gamma^\alpha = -2\gamma_\mu, \tag{5.54a}$$

$$\gamma_\alpha\gamma_\mu\gamma_\nu\gamma^\alpha = 4g_{\mu\nu}, \tag{5.54b}$$

$$\gamma_\alpha\gamma_\mu\gamma_\nu\gamma_\rho\gamma^\alpha = -2\gamma_\rho\gamma_\nu\gamma_\mu, \tag{5.54c}$$

$$\gamma_\alpha\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma\gamma^\alpha = 2(\gamma_\sigma\gamma_\mu\gamma_\nu\gamma_\rho + \gamma_\rho\gamma_\nu\gamma_\mu\gamma_\sigma), \tag{5.54d}$$

etc. The identities (5.54) should be applied before computing $\gamma$-traces according to (5.53).

The $\gamma$-matrix traces can be used to verify the completeness of the 16 Dirac covariants $\Gamma_i$ of (5.32), with

$$\mathrm{Tr}\ \Gamma_i\Gamma_j = 4\eta_i\delta_{ij}, \tag{5.55}$$

where $\eta_i = +1$ for $\Gamma_i = 1, \gamma_0, i\gamma_i\gamma_5, \sigma_i$ and $\eta_i = -1$ for $\Gamma_i = \gamma_5, \gamma_i, i\gamma_0\gamma_5, \sigma_{0i}$. Then any $4 \times 4$ matrix can be expanded as $M = \sum a_i\Gamma_i$ with the coefficients from (5.55) found to be $4a_i = \eta_i\ \mathrm{Tr}\ M\Gamma_i$. Also the completeness relation (5.55) is needed to develop Fierz "reshuffling" matrices (see Problem 5.6).

## 5.C Free-Particle Solutions of the Dirac Equation

**Positive- and Negative-Energy Spinors.** In the spirit of the free-particle spin-0 analysis, we now examine the positive- and negative-energy solutions of the covariant free-particle Dirac equation

$$(i\not{\partial} - m)\psi_\pm(x) = 0. \tag{5.56}$$

For the positive-energy solution corresponding to a particle of momentum $\mathbf{p}$ (with $V = 1$), we write

$$\psi_+(x) = u(\mathbf{p})e^{i\mathbf{p}\cdot\mathbf{x}}e^{-iEt}, \tag{5.57a}$$

while for the negative-energy solution with energy $-E$ ($E = \sqrt{\mathbf{p}^2 + m^2} > 0$) associated with momentum $-\mathbf{p}$,

$$\psi_-(x) = v(\mathbf{p})e^{-i\mathbf{p}\cdot\mathbf{x}}e^{iEt}. \tag{5.57b}$$

In both cases we require the spin-$\frac{1}{2}$ particle of mass $m$ to obey $E^2 - \mathbf{p}^2 = p^2 = m^2$. Substituting (5.57) into (5.56) then leads to the momentum-space Dirac equations

$$(\not{p} - m)u(\mathbf{p}) = 0, \qquad \bar{u}(\mathbf{p})(\not{p} - m) = 0, \tag{5.58a}$$

$$(\not{p} + m)v(\mathbf{p}) = 0, \qquad \bar{v}(\mathbf{p})(\not{p} + m) = 0. \tag{5.58b}$$

It is convenient to normalize the positive-energy spinors $u(\mathbf{p})$ and the negative-energy spinors $v(\mathbf{p})$, choosing the signs according to (5.20), so that

$$\bar{u}(\mathbf{p})u(\mathbf{p}) = 2m, \qquad \bar{v}(\mathbf{p})v(\mathbf{p}) = -2m, \tag{5.59}$$

while $\psi_\pm$ are normalized covariantly in analogy with the spin-0 Klein–Gordon solutions (4.7).

To obtain the free-particle solutions in the rest frame, we evaluate (5.58) in the limit $p_0 = m$, $\mathbf{p} = 0$, obtaining

$$\gamma_0 u(0) = u(0), \qquad \gamma_0 v(0) = -v(0). \tag{5.60}$$

[Recall that the notation $u(0)$ emphasizes the $\mathbf{p} = 0$ aspect of the rest-frame wave functions; rest-frame two-component spinors will be written as $\phi(\hat{\mathbf{p}})$ as a reference for rotations.] In the Dirac–Pauli representation with $\gamma_0$ diagonal, (5.60) diagonalizes the energy (mass, in the rest frame) with the spinors expressed as

$$u(0) = \sqrt{2m}\begin{pmatrix}\varphi\\0\end{pmatrix}, \qquad v(0) = \sqrt{2m}\begin{pmatrix}0\\\chi\end{pmatrix}, \tag{5.61}$$

where $\varphi$ and $\chi$ are two-component spinors, both normalized to unity in order that (5.59) may remain valid. In particular, we further specify these two component spinors as helicity eigenstates $\varphi^{(\lambda)}(\hat{\mathbf{p}})$ and $\chi^{(\lambda)}(\hat{\mathbf{p}}) = e^{\mp i\phi}\varphi^{(\lambda)}(-\hat{\mathbf{p}})$ [so that $v(\mathbf{p})$ represents a state with momentum $-\mathbf{p} = \mathbf{p}_{\pi-\theta,\pi+\phi}$ in the latter case—see (6.46)], obeying the eigenvalue equation (3.89), i.e.,

$$\tfrac{1}{2}\boldsymbol{\sigma}\cdot\hat{\mathbf{p}}\varphi^{(\lambda)}(\hat{\mathbf{p}}) = \lambda\varphi^{(\lambda)}(\hat{\mathbf{p}}) \tag{5.62}$$

and represented by the rotated spinors (3.91). The reason for the choice of phase $e^{\mp i\phi}$ for $\lambda = \pm\frac{1}{2}$ in $\chi(\hat{\mathbf{p}})$ will be explained in Chapter 6.

We may also recognize (5.60) as equivalent to the boost constraint (5.20) in the Lorentz-group derivation of the Dirac equation. In this case, the covariant version of the Dirac boost matrix (5.23) is, since $E + \boldsymbol{\alpha} \cdot \mathbf{p} = \not{p}\gamma_0$,

$$\begin{aligned} \mathscr{D}(L_{\mathbf{p}}) &= [2m(E+m)]^{-\frac{1}{2}}(\not{p}\gamma_0 + m) \\ &= \exp\left(\frac{-i\omega_p^{\mu\nu}\sigma_{\mu\nu}}{4}\right), \end{aligned} \tag{5.63}$$

so that $\mathscr{D}(L_{\mathbf{p}}) = \mathscr{S}(L_{\mathbf{p}})$. Application of this boost to (5.61) and using (5.60) with $\lambda = \pm\frac{1}{2}$ then leads to

$$u^{(\lambda)}(\mathbf{p}) = \mathscr{D}(L_{\mathbf{p}})u^{(\lambda)}(0) = \frac{\not{p} + m}{\sqrt{E+m}}\begin{pmatrix}\varphi^{(\lambda)}(\hat{\mathbf{p}}) \\ 0\end{pmatrix}, \tag{5.64a}$$

$$v^{(\lambda)}(\mathbf{p}) = \mathscr{D}(L_{\mathbf{p}})v^{(\lambda)}(0) = \frac{-\not{p} + m}{\sqrt{E+m}}\begin{pmatrix}0 \\ e^{\mp i\phi}\varphi^{(\lambda)}(-\hat{\mathbf{p}})\end{pmatrix}. \tag{5.64b}$$

From the form of (5.64) it is clear that these bispinors satisfy the free-particle Dirac equations (5.58), since $\not{p}\not{p} = \frac{1}{2}\{\not{p}, \not{p}\} = p^2 = m^2$ implies that

$$(\not{p} - m)u^{(\lambda)}(\mathbf{p}) \propto (\not{p} - m)(\not{p} + m) = p^2 - m^2 = 0,$$

$$(\not{p} + m)v^{(\lambda)}(\mathbf{p}) \propto (\not{p} + m)(\not{p} - m) = p^2 - m^2 = 0.$$

Further, the bispinors (5.64) are indeed helicity eigenstates, obeying by virtue of (5.62) and $[\boldsymbol{\sigma} \cdot \hat{\mathbf{p}}, \not{p}] = 0$

$$\tfrac{1}{2}\boldsymbol{\sigma} \cdot \hat{\mathbf{p}}u^{(\lambda)}(\mathbf{p}) = \lambda u^{(\lambda)}(\mathbf{p}), \tag{5.65a}$$

$$-\tfrac{1}{2}\boldsymbol{\sigma} \cdot \hat{\mathbf{p}}v^{(\lambda)}(\mathbf{p}) = \lambda v^{(\lambda)}(\mathbf{p}) \tag{5.65b}$$

for $\lambda = \pm\frac{1}{2}$, where $\boldsymbol{\sigma}$ in (5.65) is now the four-component Dirac spin matrix.

Thus, the advantage of the explicit boost construction in the Dirac–Pauli representation (5.64) is that it manifests both the Dirac equation and helicity constraints satisfied by positive- and negative-energy bispinors. Moreover, the connection between $u(\mathbf{p})$ and $v(\mathbf{p})$ for momentum states $-\mathbf{p}$ in the latter case is made readily apparent by (5.64). There is another use for the boost operator, that of "de-boosting" $u(\mathbf{p})$ and $v(\mathbf{p})$ back to the rest frame, sometimes referred to as the "nonrelativistic reduction" procedure. Recalling that $-\boldsymbol{\gamma}\gamma_0 = i\gamma_5\boldsymbol{\sigma}$, we express the boost (5.63) and its Dirac adjoint in terms of the even (diagonal) operator 1 and the odd (off-diagonal) operator $\gamma_5\boldsymbol{\sigma} \cdot \mathbf{p}$ in the Dirac–Pauli representation, converting the bispinors to the form

$$u^{(\lambda)}(\mathbf{p}) = \frac{E + m + i\gamma_5\boldsymbol{\sigma} \cdot \mathbf{p}}{\sqrt{E+m}}\begin{pmatrix}\varphi^{(\lambda)}(\hat{\mathbf{p}}) \\ 0\end{pmatrix}, \tag{5.66a}$$

$$\bar{u}^{(\lambda)}(\mathbf{p}) = (\varphi^{\dagger(\lambda)}(\hat{\mathbf{p}}), 0)\,\frac{E + m - i\gamma_5\boldsymbol{\sigma} \cdot \mathbf{p}}{\sqrt{E+m}}, \tag{5.66b}$$

$$v^{(\lambda)}(\mathbf{p}) = \frac{E + m + i\gamma_5\boldsymbol{\sigma} \cdot \mathbf{p}}{\sqrt{E+m}}\begin{pmatrix}0 \\ e^{\mp i\phi}\varphi^{(\lambda)}(-\hat{\mathbf{p}})\end{pmatrix}, \tag{5.66c}$$

$$\bar{v}^{(\lambda)}(\mathbf{p}) = (0, e^{\pm i\phi}\varphi^{\dagger(\lambda)}(-\hat{\mathbf{p}}))\,\frac{E + m - i\gamma_5\boldsymbol{\sigma} \cdot \mathbf{p}}{\sqrt{E+m}}. \tag{5.66d}$$

The nonrelativistic reduction of bilinear covariants such as $\bar{u}\Gamma_i u$ can be easily obtained using (5.66) to identify the net even operators (in the Dirac–Pauli representation) of $\bar{\mathscr{D}}\Gamma_i\mathscr{D}$. For $\bar{u}\Gamma_i u$ and $\bar{v}\Gamma_i v$, only net even operators (two-component diagonal) contribute, while for $\bar{u}\Gamma_i v$ and $\bar{v}\Gamma_i u$ only net odd operators (two-component off-diagonal) do (see Problem 5.7).

**Free-Particle Projection Operators.** Recalling that the two-component spinors satisfy the completeness relation

$$\sum_{\lambda}^{2} \varphi^{(\lambda)}(\hat{\mathbf{p}})\varphi^{\dagger(\lambda)}(\hat{\mathbf{p}}) = 1, \tag{5.67}$$

we expect the four-component bispinors to obey some sort of analogous relation. Given (5.64) and (5.67), we may compute

$$\begin{aligned}(E+m)\sum_{\lambda}^{2} u^{(\lambda)}(\mathbf{p})\bar{u}^{(\lambda)}(\mathbf{p}) &= (\not{p}+m)\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}(\not{p}+m)\\ &= \tfrac{1}{2}(\not{p}+m)(1+\gamma_0)(\not{p}+m).\end{aligned} \tag{5.68}$$

With the help of the anticommutation relations (5.28) and $p^2 = m^2$, we see that $\not{p}\gamma_0\not{p} + m^2\gamma_0 = 2E\not{p}$, $(\gamma_0\not{p} + \not{p}\gamma_0) = 2E$, and $(\not{p}+m)^2 = 2m(\not{p}+m)$. Then the factor of $(E+m)$ cancels from both sides of (5.68) and we obtain

$$\sum_{\lambda}^{2} u^{(\lambda)}(\mathbf{p})\bar{u}^{(\lambda)}(\mathbf{p}) = \not{p} + m, \tag{5.69a}$$

and similarly

$$\sum_{\lambda}^{2} v^{(\lambda)}(\mathbf{p})\bar{v}^{(\lambda)}(\mathbf{p}) = \not{p} - m. \tag{5.69b}$$

These useful relations can be verified by applying the Dirac operator $(\not{p} - m)$ to (5.69a) and $(\not{p} + m)$ to (5.69b); the normalizations are easily confirmed in the rest frame via (5.61). The identities (5.69) define respectively positive and negative projection operators in the sense of (5.51). That is,

$$\Lambda_{\pm}(p) = \frac{\pm\not{p} + m}{2m} \tag{5.70}$$

are normalized energy projection operators obeying $\Lambda_{\pm}^2 = \Lambda_{\pm}$ and $\Lambda_+\Lambda_- = 0$ for $p^2 = m^2$. The additional identity $\Lambda_+ + \Lambda_- = 1$, or equivalently from (5.69),

$$\sum_{\lambda} \left(u^{(\lambda)}(\mathbf{p})\bar{u}^{(\lambda)}(p) - v^{(\lambda)}(\mathbf{p})\bar{v}^{(\lambda)}(\mathbf{p})\right) = 2m \tag{5.71}$$

is the Dirac analog of the two-component completeness relation (5.67).

The usefulness of the energy projection operators (5.69) is that, for an $S$-matrix specified by positive- or negative-energy free-particle bilinear covariants, the general unpolarized-spin sum (5.52) becomes in particular

$$\sum_{\lambda',\lambda} |\bar{u}^{(\lambda')}(\mathbf{p}')Mu^{(\lambda)}(\mathbf{p})|^2 = \mathrm{Tr}\ M(\not{p} + m)\bar{M}(\not{p}' + m), \tag{5.71a}$$

$$\sum_{\lambda',\lambda} |\bar{v}^{(\lambda')}(\mathbf{p}')Mv^{(\lambda)}(\mathbf{p})|^2 = \mathrm{Tr}\ M(\not{p} - m)\bar{M}(\not{p}' - m), \tag{5.71b}$$

and similarly for $\bar{u}Mv$ and $\bar{v}Mu$.

It is also possible to construct a covariant spin projection operator. In the rest frame of the spin-$\frac{1}{2}$ particle, such a spin projection operator is $\frac{1}{2}(1 + \boldsymbol{\sigma} \cdot \hat{\mathbf{s}})$, where $\hat{\mathbf{s}}$ is the direction of spin polarization ($\hat{\mathbf{s}} \to \hat{\mathbf{p}}$ corresponds to a helicity projection operator). Defining $s_\mu = (0, \hat{\mathbf{s}})$ and $m_\mu = (m, \mathbf{0})$ in the rest frame, so that $m \cdot s = 0$, it is clear that in a boosted frame $p \cdot s = 0$. Then, replacing $\boldsymbol{\sigma}$ by $-i\gamma_5\boldsymbol{\gamma}\gamma_0$ and dropping the noncovariant factor $\gamma_0$ (since $\hat{\mathbf{p}}\gamma_0$ does not change sign when applied to either $u$ or $v$), we obtain the covariant spin projection operator

$$\Sigma(s) = \tfrac{1}{2}(1 + i\gamma_5\not{s}). \tag{5.72}$$

The invariant length $s^2 = -1$ means that $\Sigma^2(s) = \Sigma(s)$, characteristic of any projection operator. We may then compute positive- or negative-energy *polarized* spin sums using the Dirac trace algebra by simply replacing, say, $\not{p} \pm m$ in (5.71) with $\Sigma(s)(\not{p} \pm m)$ (or $(\not{p} \pm m)\Sigma(s)$, since $[\gamma_5\not{s}, \not{p}] = 0$) for a particle whose spin is not unobserved, but pointing in the "direction" $s$. Specific helicity states can then be generated using $s_\mu = \pm(p, E\hat{\mathbf{p}})/m$ in (5.72), whereas spin representations other than helicity follow from a different choice for $s_\mu$.

Finally, a third set of Dirac projection operators

$$P_\pm = \tfrac{1}{2}(1 \pm i\gamma_5) \tag{5.73}$$

will prove useful for our later work. They too obey $P_\pm^2 = P_\pm$, $P_+P_- = 0$, and $P_+ + P_- = 1$. One immediate application of (5.73) stems from $P_+\gamma_\mu = \gamma_\mu P_-$, which means that for free particle bispinors $\psi = m^{-1}i\not{\partial}\psi$ we have $P_\pm\psi = m^{-1}i\not{\partial}P_\mp\psi$, so that

$$\psi = P_+\psi + P_-\psi = m^{-1}(i\not{\partial} + m)P_+\psi. \tag{5.74}$$

Thus for $\phi = P_+\psi$,

$$(i\not{\partial} - m)\psi = -m^{-1}(\Box + m^2)\phi = 0, \tag{5.75}$$

which says that, given $\phi$ obeying the *less* restrictive Klein–Gordon equation (for bispinor wave functions), one may obtain $\psi$ via (5.74) which satisfies the *more* restrictive Dirac equation provided $\phi = P_\pm\psi$. This is reminiscent of our second derivation of the Dirac equation built up from two-component spinors related by the differential operator in (5.16), here equivalent to (5.74) in the Weyl representation. In the above approach, the Dirac–Pauli representation reduces (5.75) to two identical two-component spinor equations. This is true even in the presence of electromagnetic fields, and we shall take advantage of this fact shortly.

**Free-Particle Probability Current.** We have noted in (5.6), (5.7), and (5.47) that the bilinear covariant $\bar{\psi}\gamma_\mu\psi$ is the natural choice to represent a conserved probability current density. Consider the current density constructed from the free-particle positive-energy wave packet similar in structure to the Klein–Gordon form (4.13a), normalized in a box:

$$\psi_+(x) = \int \frac{d^3p}{2EV^{\frac{1}{2}}}\, a_{\mathbf{p}} u(\mathbf{p}) e^{-ip\cdot x} \tag{5.76}$$

(where we have deleted the helicity summation for clarity), giving

$$j_\mu^+(x) = \bar{\psi}_+(x)\gamma_\mu\psi_+(x) = \frac{1}{V}\int\frac{d^3p}{2E}\int\frac{d^3p'}{2E'}\, a_{\mathbf{p}'}^* a_{\mathbf{p}}\langle \mathbf{p}'\,|\,j_\mu^+\,|\,\mathbf{p}\rangle e^{iq\cdot x}, \tag{5.77}$$

with $q = p' - p$. The off-diagonal momentum-space current is then

$$\langle \mathbf{p}'\,|\,j_\mu^+\,|\,\mathbf{p}\rangle = \bar{u}(\mathbf{p}')\gamma_\mu u(\mathbf{p}). \tag{5.78}$$

In this language, current conservation $\partial\cdot j(x) = 0$ is equivalent to

$$q^\mu\langle \mathbf{p}'\,|\,j_\mu^+\,|\,\mathbf{p}\rangle = \bar{u}(\mathbf{p}')\not{q}u(\mathbf{p}) = \bar{u}(\mathbf{p}')(\not{p}' - \not{p})u(\mathbf{p}) = (m - m)\bar{u}u = 0, \tag{5.79}$$

by the free-particle Dirac equations (5.58a). An analogous negative-energy packet with $E = \sqrt{\mathbf{p}^2 + m^2} > 0$,

$$\psi_-(x) = \int\frac{d^3p}{2EV^{\frac{1}{2}}}\, b_{\mathbf{p}}^* v(\mathbf{p}) e^{ip\cdot x}, \tag{5.80}$$

leads to a $\langle \mathbf{p}'\,|\,j_\mu^-\,|\,\mathbf{p}\rangle$ which is also conserved in the sense of (5.79).

As we have already noted in Section 4.C, such off-diagonal momentum-space currents play an important role in the theory in their own right. Consider first a useful identity called the *Gordon reduction.* Defining $P = \frac{1}{2}(p' + p)$, $q = p' - p$ and using the $\gamma$-matrix property (5.31a), it is easy to show (Problem 5.8) that

$$i\sigma_{\mu\nu}q^\nu = \not{p}'\gamma_\mu + \gamma_\mu\not{p} - 2P_\mu. \tag{5.81}$$

Sandwiching this identity (5.81) between free-particle positive-energy bispinors, we note that $\not{p}'$ operating to the left and $\not{p}$ to the right both become $m$ in (5.81), again by the free-particle Dirac equations (5.58a). Rearranging this result, we obtain a form of the Gordon reduction,

$$\bar{u}(\mathbf{p}')\gamma_\mu u(\mathbf{p}) = \bar{u}(\mathbf{p}')\left(\frac{P_\mu}{m} + \frac{i\sigma_{\mu\nu}q^\nu}{2m}\right)u(\mathbf{p}), \tag{5.82}$$

where the $P_\mu$ and $\sigma_{\mu\nu}q^\nu$ terms in (5.82) are called the convection and spin currents, respectively. Clearly, relations similar to (5.82) can be generated by sandwiching (5.81) between $\bar{v}$ and $v$, or $\bar{u}$ and $v$, or $\bar{v}$ and $u$.

As a by-product of (5.82), the diagonal matrix element can be obtained with $q = 0$ and $P = p$:

$$\bar{u}(\mathbf{p})\gamma_\mu u(\mathbf{p}) = (p_\mu/m)\bar{u}(\mathbf{p})u(\mathbf{p}) = 2p_\mu. \tag{5.83a}$$

This result is "obvious" from covariance considerations alone, because the left-hand side of (5.83a) must be a simple four-vector and the only candidate depending upon $\mathbf{p}$ is $p_\mu$, with a normalization factor following from the contraction with $p^\mu$, using $p^2 = m^2$ and the Dirac equations (5.58). In a similar manner we deduce that

$$\bar{v}(\mathbf{p})\gamma_\mu v(\mathbf{p}) = (-p_\mu/m)\bar{v}(\mathbf{p})v(\mathbf{p}) = 2p_\mu. \tag{5.83b}$$

We may now calculate the expectation value of the Dirac velocity operator $\boldsymbol{\alpha}$. Noting that the spatial integral of (5.77) generates $\delta^3(\mathbf{p}' - \mathbf{p})$, which in turn picks out the diagonal matrix elements of $j_\mu^+$, we use (5.83a) to write [a similar relation holds for the Klein–Gordon current (4.62)]

$$\int d^3x\, j_\mu^+(\mathbf{x}, t) = \int \frac{d^3p}{2EV}\, |a_\mathbf{p}|^2 p_\mu/E, \tag{5.84}$$

with an analogous relation for $j_\mu^-$, where $a_\mathbf{p}$ is replaced by $b_\mathbf{p}$. For $\mu = 0$ we see that $p_0/E = 1$ in (5.84) as well as for the $j_0^-$ integral. Since these $j_0^\pm$ integrals are $\int d^3x\, \psi_\pm^\dagger(x)\psi_\pm(x)$, the latter integrals normalized to unity as the total probability, we learn that

$$\frac{1}{V}\int \frac{d^3p}{2E}\, |a_\mathbf{p}|^2 = \frac{1}{V}\int \frac{d^3p}{2E}\, |b_\mathbf{p}|^2 = 1. \tag{5.85}$$

Next we evaluate

$$\int d^3x\, \mathbf{j}^\pm(\mathbf{x}, t) = \int d^3x\, \psi_\pm^\dagger(x)\boldsymbol{\alpha}\psi_\pm(x) = \langle\boldsymbol{\alpha}\rangle_\pm, \tag{5.86}$$

and from (5.84) conclude that

$$\langle\boldsymbol{\alpha}\rangle_\pm = \left\langle\frac{\mathbf{p}}{E}\right\rangle_\pm = \langle\mathbf{v}_g\rangle_\pm, \tag{5.87}$$

where the latter two averages in (5.87) are weighted over the momentum-space probabilities (5.85). The important result (5.87) simply states that the positive- or negative-energy wave-packet expectation value of the current is just the relativistic group velocity $v_g$, a conclusion paralleling the nonrelativistic situation (1.6), and expected in this case because in the Heisenberg picture,

$$\frac{dr_j}{dt} = i[H, r_j] = i\alpha_k[p_k, r_j] = \alpha_j. \tag{5.88}$$

**Zitterbewegung.** It turns out, however, that both positive- and negative-energy states must be included in the construction of the eigenfunctions of $\boldsymbol{\alpha}$.

Consider then the general packet (again suppressing the helicity summation)

$$\psi(x) = \int \frac{d^3p}{2EV^{\frac{1}{2}}} [a_{\mathbf{p}} u(\mathbf{p}) e^{-ip\cdot x} + b_{\mathbf{p}}^* v(\mathbf{p}) e^{ip\cdot x}], \tag{5.89}$$

and computing $\langle \boldsymbol{\alpha} \rangle$ as before leads to

$$\langle \boldsymbol{\alpha} \rangle = \left\langle \frac{\mathbf{p}}{E} \right\rangle + \int \frac{d^3p}{2E} [a_{\mathbf{p}}^* b_{-\mathbf{p}}^* \bar{u}(\mathbf{p}) \boldsymbol{\alpha} v(-\mathbf{p}) e^{2iEt} + a_{\mathbf{p}} b_{-\mathbf{p}} \bar{v}(-\mathbf{p}) \boldsymbol{\alpha} u(\mathbf{p}) e^{-2iEt}]. \tag{5.90}$$

As was the case with the Klein–Gordon equation, the cross terms between the positive- and negative-energy states oscillate violently in time. Unlike the Klein–Gordon situation, however, the free-particle amplitudes multiplying the oscillating phases in (5.90) need not vanish as $\mathbf{p} \to 0$. To see this, use $\boldsymbol{\alpha} = i\gamma_5 \boldsymbol{\sigma}$ and the forms (5.66) to obtain the nonrelativistic reductions (Problem 5.7)

$$\bar{u}(\mathbf{p}) \boldsymbol{\alpha} v(-\mathbf{p}) \to 2m\varphi^\dagger(\hat{\mathbf{p}}) \boldsymbol{\sigma} \varphi(\hat{\mathbf{p}}), \qquad \bar{v}(-\mathbf{p}) \boldsymbol{\alpha} u(\mathbf{p}) \to 2m\varphi^\dagger(\hat{\mathbf{p}}) \boldsymbol{\sigma} \varphi(\hat{\mathbf{p}}). \tag{5.91}$$

Thus Dirac *Zitterbewegung* is nonvanishing to zeroth order in momentum. While unphysical consequences can be avoided for free particles by choosing either a positive- or a negative-energy packet as in (5.76) or (5.80), for bound states this cannot be done. We return to this latter problem shortly.

## 5.D Dirac Equation in an External Field

**Covariant Electromagnetic Interactions.** Following the procedure used in the Klein–Gordon analysis of a particle in the presence of an electromagnetic field, we consider for spin-$\frac{1}{2}$ particles the minimal replacement of $i\partial_\mu$ by $i\partial_\mu - eA_\mu$ in the Dirac equation (5.37), giving

$$(i\not{\partial} - e\not{A} - m)\psi(x) = 0, \quad \text{or } (i\not{\partial} - m)\psi(x) = e\not{A}\psi(x). \tag{5.92}$$

We will assume that this minimal coupling $e\not{A}$ represents the fundamental interaction between the spin-$\frac{1}{2}$ particles and photons. As in Dirac's original approach, the bispinor equation (5.92) obeys a Klein–Gordon constraint, obtained by multiplying (5.92) on the left by the operator $i\not{\partial} - e\not{A} + m$:

$$[(i\partial - eA)^2 - m^2 - e\tfrac{1}{2}\sigma_{\mu\nu} F^{\mu\nu}]\psi(x) = 0, \tag{5.93}$$

where we have used the anticommutation relations of the $\gamma$-matrices to write

$$\not{\partial}\not{A}\psi + \not{A}\not{\partial}\psi = \tfrac{1}{2}[\gamma_\mu, \gamma_\nu] F^{\mu\nu}\psi + \partial \cdot (A\psi).$$

We see from (5.93) that the Dirac wave function also obeys the Klein–Gordon equation in the presence of an external field (4.24a) apart from the spin-dependent factor

$$\tfrac{1}{2}\sigma_{\mu\nu} F^{\mu\nu} = i\boldsymbol{\alpha} \cdot \mathbf{E} - \boldsymbol{\sigma} \cdot \mathbf{B}, \tag{5.94}$$

which couples **E** and **B** respectively to the electric and magnetic dipole moments of the spin-$\frac{1}{2}$ particle. In particular, an energy eigenstate of $\psi(x)$ in (5.93) obeys, for $E^2 - m^2 = 2mE_{\mathrm{NR}} + \cdots$,

$$E_{\mathrm{NR}}\psi = -\frac{e}{2m}\boldsymbol{\sigma}\cdot\mathbf{B}\psi + \cdots, \tag{5.95}$$

showing once again that $\boldsymbol{\mu} = e\boldsymbol{\sigma}/2m$ and $g = 2$ are natural consequences of the Dirac equation.

It is also interesting to investigate the Dirac probability current, $j_\mu(x)$, in the presence of an external field. In spite of the absence of a derivative term in $\bar{\psi}\gamma_\mu\psi$, one may consider the Gordon decomposition in coordinate space analogous to (5.82) and then make a minimal replacement to find (see Problem 5.8)

$$j_\mu(x) = \frac{i}{2m}\bar{\psi}(x)\overleftrightarrow{\partial}_\mu\psi(x) - \frac{e}{m}\bar{\psi}(x)\psi(x)A_\mu(x) + \frac{1}{2m}\partial^\nu(\bar{\psi}(x)\sigma_{\mu\nu}\psi(x)). \tag{5.96}$$

It is clear that the first two terms in (5.96) have the same form as the nonrelativistic and Klein–Gordon current density (4.26), covariantly normalized in the latter case. The first and third terms in (5.96) correspond to the Gordon decomposition for $A_\mu = 0$. As in the spinless case (4.27), current conservation $\partial\cdot j(x) = 0$ is an obvious consequence of (5.96).

**Dirac Form Factors.** Following now the discussion in Section 4.C, the probability current density $j_\mu(x)$, whether due to nonrelativistic, Klein–Gordon, or Dirac particles, plays a dual role because $ej_\mu(x)$ represents a charged-matter current density which is a source for the electromagnetic field, $\Box A_\mu = ej_\mu$ (in the Lorentz gauge). Thus the spin-$\frac{1}{2}$ analog of the momentum-space current (4.63b) in the presence of an electromagnetic field for positive-energy scattered particles, obtained from (5.96), is

$$e\langle\mathbf{p}'|j_\mu^+|\mathbf{p}\rangle = e\bar{u}(\mathbf{p}')\left[F_1(q^2)\gamma_\mu + F_2(q^2)\frac{i\sigma_{\mu\nu}q^\nu}{2m}\right]u(\mathbf{p}). \tag{5.97}$$

This is the most general form consistent with Lorentz covariance and the Dirac equation. The four-momentum transferred from the photon to the spin-$\frac{1}{2}$ particles is $q = p' - p$, and the dimensionless invariant form factors $F_{1,2}(q^2)$ parallel $F(q^2)$ in (4.63b) for spinless particles. For free photons, $q^2 = 0$ and we must have $F_1(0) = 1$ in order that the coefficient of $\bar{u}\gamma_\mu u$ in (5.97) may be simply the charge $e$, consistent with the free-particle probability current density (5.78). Thus $F_1(q^2)$ is called the charge form factor. For $q^2 \neq 0$, the photon is not free, and the second term of (5.96) contributes to both $F_1(q^2)$ and $F_2(q^2)$ of (5.97) in a complicated dynamical manner. We shall return to the $q^2$ dependence of these form factors in later chapters.

Concerning the magnetic form factor $F_2(q^2)$ in (5.97), since the $\gamma_\mu$ current already consists of a convection and magnetic moment part as given by

(5.82), with $g = 2$ in the latter case, the additional magnetic-moment term in (5.97) described by $F_2(q^2)$ represents an *anomalous* correction to $g = 2$. That is, at $q^2 = 0$,

$$F_2(0) \equiv \kappa = \tfrac{1}{2}(g - 2) \tag{5.98}$$

is the dimensionless, anomalous magnetic moment of a spin-$\frac{1}{2}$ particle. For the electron $(e < 0)$,

$$\mu_e = \frac{e}{2m_e}(1 + \kappa_e) \approx (1.001)\frac{e}{2m_e}, \tag{5.99a}$$

or $\kappa_e \approx 0.001$, due totally to electromagnetic interactions. For the proton $(e > 0)$,

$$\mu_p = \frac{e}{2m_p}(1 + \kappa_p) \approx (2.79)\frac{e}{2m_p}, \tag{5.99b}$$

or $\kappa_p \approx 1.79$, primarily due to strong interactions—as is the anomalous magnetic moment of the chargeless neutron,

$$\mu_n = \frac{e}{2m_n}(0 + \kappa_n) \approx (-1.91)\frac{e}{2m_n}, \tag{5.99c}$$

or $\kappa_n \approx -1.91$. In Chapter 15 we shall show that the minimal (QED) coupling in (5.92) ultimately generates $\kappa_e = \alpha/2\pi + O(\alpha^2)$, in perfect agreement with experiment. It will also turn out that we can estimate $\kappa_p$ and $\kappa_n$ better than we have a right to expect. Lastly, the form factors $F_{1,2}(q^2)$ are not a unique description of the dynamics of spin-$\frac{1}{2}$ particles interacting with an external electromagnetic field. Two different but equivalent descriptions in terms of the *Sachs* form factors $G_E(q^2)$ and $G_M(q^2)$ are worked out in Problem 5.8.

**Constants of the Motion.** For a free-particle hamiltonian $\boldsymbol{\alpha} \cdot \mathbf{p} + \beta m$, the fundamental commutation relations (5.4) lead to

$$[H, \mathbf{r} \times \mathbf{p}] = -i\boldsymbol{\alpha} \times \mathbf{p}, \qquad [H, \boldsymbol{\sigma}] = 2i\boldsymbol{\alpha} \times \mathbf{p}, \tag{5.100}$$

and it is therefore clear that the total angular momentum $\mathbf{J} = \mathbf{r} \times \mathbf{p} + \frac{1}{2}\boldsymbol{\sigma}$ and helicity $\frac{1}{2}\boldsymbol{\sigma} \cdot \hat{\mathbf{p}}$ are conserved operators. In a central, spherically symmetric field $V = V(r) = eA_0$ with $\mathbf{A} = 0$, the Dirac hamiltonian becomes

$$H = \boldsymbol{\alpha} \cdot \mathbf{p} + \beta m + V(r), \tag{5.101}$$

and while commutation relations similar to (5.100) become more cumbersome to apply (see, e.g., Sakurai 1967), group-theory arguments alone tell us that the total-angular-momentum operator must remain conserved.

To simplify the search for constants of the motion and their eigenvalues for spin-$\frac{1}{2}$ particles, it will prove convenient to reduce the four-component Dirac analysis back to two-component form. We consider then the less restrictive second-order equation (5.93) satisfied by a Dirac bispinor $\psi$ for positive-energy stationary states with $\pi = \mathbf{p} - e\mathbf{A}$ and $e\mathbf{E} = -\nabla V$, $\mathbf{B} = 0$,

$$(\not{\pi} + m)(\not{\pi} - m)\psi = [(E - V)^2 + \nabla^2 - m^2 + i\boldsymbol{\alpha} \cdot \nabla V]\psi = 0. \quad (5.102)$$

Multiplying (5.102) on the left by the projection operators $P_\pm = \frac{1}{2}(1 \pm i\gamma_5)$ of (5.73), we see that the four-component wave function $\phi \equiv P_+ \psi$ satisfies (5.102), and the more restrictive Dirac wave function, obeying $H\psi = E\psi$, is then recovered from $\phi$ by

$$\psi = m^{-1}(\not{\pi} + m)\phi = m^{-1}[\gamma_0(E - V) - \boldsymbol{\gamma} \cdot \mathbf{p} + m]\phi, \qquad (5.103)$$

a situation similar to the free particle case (5.74). Next we work in the Dirac–Pauli representation with $P_+ = \frac{1}{2}\left(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right)$, giving

$$\phi = P_+\psi = \begin{pmatrix} \varphi \\ \varphi \end{pmatrix}. \qquad (5.104)$$

Then the four-component equation (5.102) reduces to two identical two-component equations [for $V = V(r)$],

$$\left[(E - V(r))^2 - m^2 + \frac{1}{r}\frac{\partial^2}{\partial r^2} r - \frac{\mathbf{L}^2}{r^2} + i\boldsymbol{\sigma} \cdot \hat{\mathbf{r}} \frac{dV}{dr}\right]\varphi = 0, \qquad (5.105)$$

equivalent to the Klein–Gordon form (4.29) except for the last spin-dependent term. If all we are interested in is the quantized energy levels, then we need only deal with (5.105), converting it to the form of the Schrödinger equation as in the Klein–Gordon case. We need not bother with recovering $\psi$ from $\phi$ via (5.103) and (5.104).

In order to separate variables in (5.105), we specialize further to a single-electron atom with $V(r) = -Z\alpha/r$ and $dV/dr = Z\alpha/r^2$. It is then natural to incorporate both $V^2$ and the spin-dependent term in the angular part of (5.105) along with $\mathbf{L}^2/r^2$, since these terms all fall off like $r^{-2}$. In particular we make use of the two-component decomposition $\sigma_i\sigma_j = \delta_{ij} + i\varepsilon_{ijk}\sigma_k$ to write

$$\mathbf{L}^2 = \boldsymbol{\sigma} \cdot \mathbf{L}(1 + \boldsymbol{\sigma} \cdot \mathbf{L}) = (1 + \boldsymbol{\sigma} \cdot \mathbf{L})^2 - (1 + \boldsymbol{\sigma} \cdot \mathbf{L}). \qquad (5.106)$$

This suggests that we define the two-component spin operator

$$-\Lambda = (1 + \boldsymbol{\sigma} \cdot \mathbf{L}) + (Z\alpha)i\boldsymbol{\sigma} \cdot \hat{\mathbf{r}}, \qquad (5.107)$$

which has the square

$$\Lambda^2 = (1 + \boldsymbol{\sigma} \cdot \mathbf{L})^2 - (Z\alpha)^2, \qquad (5.108)$$

the cross terms in (5.108) dropping out because of the identity (see Problem 5.9)

$$\{(1 + \boldsymbol{\sigma} \cdot \mathbf{L}), \boldsymbol{\sigma} \cdot \hat{\mathbf{r}}\} = 0. \qquad (5.109)$$

Combining (5.106)–(5.108), we see that the coefficient of $-r^{-2}$ in (5.105) has the simple form of an angular-momentum operator

$$\mathbf{L}^2 - (Z\alpha)i\boldsymbol{\sigma} \cdot \hat{\mathbf{r}} - (Z\alpha)^2 = \Lambda^2 + \Lambda = \Lambda(\Lambda + 1). \qquad (5.110)$$

Next we obtain the eigenvalues of $\Lambda^2$ and $\Lambda$ by squaring $\mathbf{J} = \mathbf{L} + \frac{1}{2}\boldsymbol{\sigma}$, to write for states $j = l \pm \frac{1}{2} \equiv j_\pm$

$$(\boldsymbol{\sigma} \cdot \mathbf{L})_{j_\pm} = (\mathbf{J}^2 - \mathbf{L}^2 - \tfrac{3}{4})_{j_\pm} = \begin{Bmatrix} l \\ -(l+1) \end{Bmatrix}, \tag{5.111a}$$

$$(1 + \boldsymbol{\sigma} \cdot \mathbf{L})_{j_\pm} = \begin{Bmatrix} l+1 \\ -l \end{Bmatrix} = \pm(j + \tfrac{1}{2}), \tag{5.111b}$$

where the upper (lower) cases in (5.111) correspond to $j = l + \frac{1}{2}$ ($j = l - \frac{1}{2}$). Then we have

$$(\Lambda^2)_{j_\pm} = (j + \tfrac{1}{2})^2 - (Z\alpha)^2, \tag{5.112a}$$

$$(\Lambda)_{j_\pm} = \mp[(j + \tfrac{1}{2})^2 - (Z\alpha)^2]^{\frac{1}{2}} \equiv \mp\lambda, \tag{5.112b}$$

where the signs in (5.112b) are determined by (5.107) and (5.111b) with $-\Lambda$ intrinsically positive for $j = l + \frac{1}{2}$ and $Z\alpha \to 0$. Following the Klein–Gordon analysis, we may now express the radial equation for the Dirac atom in the form of (4.30a),

$$\left[\frac{1}{r}\frac{d^2}{dr^2}r - \frac{l'(l'+1)}{r^2} + \frac{2EZ\alpha}{r}\right]\varphi(r) = -(E^2 - m^2)\varphi(r), \tag{5.113}$$

with (5.110) and (5.112) giving, for $j = l \pm \frac{1}{2}$,

$$l'_\pm(l'_\pm + 1) = [\Lambda(\Lambda + 1)]_{j_\pm} = \lambda(\lambda \mp 1). \tag{5.114}$$

Solving (5.114) for the effective orbital-angular-momentum eigenvalue, we see that

$$l'_\pm = \begin{Bmatrix} \lambda - 1 \\ \lambda \end{Bmatrix}. \tag{5.115}$$

As before, we have chosen the nonnegative solutions of $l'$ in (5.114) as $Z\alpha \to 0$, for otherwise the radial solution $r^{l'}$ will not be regular at the origin.

The orbital constants of the motion now being understood for the Dirac atom, we proceed as in the Klein–Gordon atom to find the energy eigenvalues. Since the radial equations are identical in the two cases except for different values of $l'$, we may conclude that the energy levels of the Dirac atom also are of the form (4.32b),

$$E = m\left[1 + \frac{(Z\alpha)^2}{n'^2}\right]^{-\frac{1}{2}}, \tag{5.116}$$

with $n' - l' = n - l$ constrained to the integer values 1, 2, ... as before. Substituting $l'$ given by (5.114) then leads to

$$n' = n - l + l'_\pm = n - (j + \tfrac{1}{2}) + [(j + \tfrac{1}{2})^2 - (Z\alpha)^2]^{\frac{1}{2}} \tag{5.117}$$

for both $j = l \pm \frac{1}{2}$. [Note that this two-component analysis eliminates the need for an auxiliary quantum number $k$, appearing in the four-component analysis. See e.g. Sakurai (1967).] We see that the energy levels of a Dirac atom are identical in form to the Klein–Gordon levels, but with $l$ in (4.34b) replaced by $j$ in (5.117). Note too that for $s$-waves and $j = \frac{1}{2}$, the square root in (5.117) becomes imaginary unless $Z < 1/\alpha$.

In the strong-field limit, $Z > 1/\alpha \approx 137$ and the breakdown of the bound-state Dirac solutions is similar to the Klein–Gordon case with $Z > 1/2\alpha$. Once again, the breakdown has a classical analog for $(Z\alpha)^2 > L^2$ with multi-particle quantum states modifying the single particle orbit at short distances. These effects are not fully understood, nor have they been detected experimentally.

For $Z\alpha \ll 1$, one can expand (5.116) and (5.117) in powers of $(Z\alpha)$ to find the fine-structure corrections to the Bohr formula for a single-electron atom ($n = 1, 2, \ldots$),

$$E_{\text{NR}} = E - m = -\frac{m(Z\alpha)^2}{2n^2}\left[1 + \frac{(Z\alpha)^2}{n}\left(\frac{1}{j+\frac{1}{2}} - \frac{3}{4n}\right)\right], \tag{5.118}$$

now degenerate for a given value of $j$. We shall return to a detailed study of the hydrogen energy levels shortly. Finally, an analysis of the Dirac wave functions (proportional to associated Laguerre functions), based upon this two-component approach, depends upon (5.103). [See Auvil and Brown (1977).] Alternatively, the entire (and more complicated) four-component analysis of $H\psi = E\psi$ can be carried out in a straightforward manner [see e.g. Bethe and Salpeter (1957)]. For our purposes, the above analysis of the energy levels alone will suffice for our later work.

Before leaving this topic, it is worth mentioning that the two-component form of (5.93) also can be used to obtain the constants of the motion for spin-$\frac{1}{2}$ problems other than the one-electron Coulomb atom. Consider, for example, an electron of momentum $\mathbf{p}$ moving through a region of constant magnetic field, $\mathbf{B} = B\hat{\mathbf{e}}_3$ and $\mathbf{A} = \frac{1}{2}\mathbf{B} \times \mathbf{r} = \frac{1}{2}B(-y\hat{\mathbf{e}}_1 + x\hat{\mathbf{e}}_2)$ with $A_0 = 0$. Applying the projection operator (5.104) in the Dirac–Pauli representation, the four-component equation (5.93) once more decouples into two identical two component equations:

$$[E^2 - (\mathbf{p} - e\mathbf{A})^2 - m^2 + eB\sigma_3]\varphi = 0. \tag{5.119}$$

Since $\nabla \cdot \mathbf{A} = 0$ and $\mathbf{A} \cdot \mathbf{p} = \frac{1}{2}\mathbf{B} \cdot \mathbf{L}$, we may write (5.119) in the form

$$[p_x^2 + p_y^2 + (eB/2)^2(x^2 + y^2)]\varphi = [E^2 - p_z^2 - m^2 + eB(L_3 + \sigma_3)]\varphi. \tag{5.120}$$

Separating the $z$ variable from the $x$ and $y$ variables in (5.120), we replace the operator $p_z^2$ by the eigenvalue $p_3^2$, and the magnetic-moment operator (proportional to $L_3 + \sigma_3$) by the Zeeman-splitting eigenvalues $l_3 \pm 1$ on the right-hand side of (5.120). The left-hand side of (5.120) has the structure of a nonrelativistic, two-dimensional harmonic-oscillator hamiltonian with a potential

$$V = \tfrac{1}{2}k(x^2 + y^2) = \left(\frac{1}{2m}\right)\left|\frac{eB}{2}\right|^2 (x^2 + y^2), \tag{5.121}$$

corresponding to an angular frequency of

$$\omega = \sqrt{k/m} = |eB|/2m. \tag{5.122}$$

Since the eigenvalues of such an harmonic oscillator are known to be of the form

$$E_{\text{h.o.}} = (n_x + \tfrac{1}{2})\omega + (n_y + \tfrac{1}{2})\omega = n\omega, \tag{5.123}$$

where $n = n_x + n_y + 1 = 1, 2, 3, \ldots$, we may identify the right-hand side of (5.120) as $2mE_{\text{h.o.}}$. This leads to the constant of motion

$$E^2 = p_3^2 + m^2 + n|eB| - (l_3 \pm 1)eB, \tag{5.124}$$

with $E^2 \geq p_3^2 + m^2$ implying $n \geq |l_3| + 1$. Thus we see that the energy and angular-momentum eigenvalues of a spin-$\frac{1}{2}$ particle in a magnetic field can be found from (5.93) in a manner similar to those for the Klein–Gordon or Dirac atom. The two-component and four-component bispinor wave functions for this problem are proportional to Hermite polynomials, as might be expected from the structure of (5.120). (See Problem 5.10.)

**Nonrelativistic Reduction.** While the exact energy eigenvalues for a relativistic spin-$\frac{1}{2}$ particle in an external field can always be obtained as in the last section, it is nonetheless of interest to expand the Dirac equation (5.92) in powers of $E_{\text{NR}}/m$, where $E_{\text{NR}} \equiv E - m$. Starting with the Dirac hamiltonian ($V = eA_0$, $\boldsymbol{\pi} = \mathbf{p} - e\mathbf{A}$)

$$H = \boldsymbol{\alpha}\cdot\boldsymbol{\pi} + \beta m + V, \tag{5.125}$$

we may express $H\psi = E\psi$ in the Dirac–Pauli representation as

$$\begin{pmatrix} E - m - V & -\boldsymbol{\sigma}\cdot\boldsymbol{\pi} \\ -\boldsymbol{\sigma}\cdot\boldsymbol{\pi} & E + m - V \end{pmatrix}\begin{pmatrix} \varphi \\ \chi \end{pmatrix} = 0. \tag{5.126}$$

This results in two coupled two-component spinor equations

$$(E - m - V)\varphi = \boldsymbol{\sigma}\cdot\boldsymbol{\pi}\chi, \qquad (E + m - V)\chi = \boldsymbol{\sigma}\cdot\boldsymbol{\pi}\varphi. \tag{5.127}$$

For weak potentials $V \ll E \approx m$, $\varphi \propto V^{-1}\chi$ is large while $\chi \propto m^{-1}\varphi$ is small. Identifying then $\varphi$ as the relativistic extension of the nonrelativistic Schrödinger wave function, we eliminate $\chi$ in (5.127) to find ($E_{\text{NR}} = E - m$)

$$H_{\text{NR}}\varphi = E_{\text{NR}}\varphi, \tag{5.128a}$$

$$H_{\text{NR}} = \frac{1}{2m}\boldsymbol{\sigma}\cdot\boldsymbol{\pi}\left(1 + \frac{E_{\text{NR}} - V}{2m}\right)^{-1}\boldsymbol{\sigma}\cdot\boldsymbol{\pi} + V. \tag{5.128b}$$

Assuming $(E_{\text{NR}} - V)/2m \ll 1$, we expand

$$\left(1 + \frac{E_{\text{NR}} - V}{2m}\right)^{-1} \approx 1 - \frac{E_{\text{NR}} - V}{2m}. \tag{5.129}$$

The leading term in (5.129) then converts (5.128b) to

$$H_{\mathrm{NR}}^{(1)} = \frac{(\boldsymbol{\sigma}\cdot\boldsymbol{\pi})^2}{2m} + V = \frac{\boldsymbol{\pi}^2}{2m} - \frac{e}{2m}\boldsymbol{\sigma}\cdot\mathbf{B} + V, \tag{5.130}$$

where we have used (5.13) to obtain the Pauli magnetic-moment term with $g = 2$, a result now expected.

To next order in the expansion (5.129), we set $\mathbf{A} = 0$ to isolate the correction terms due to a weak electrostatic potential $V$:

$$H_{\mathrm{NR}}^{(2)} = \frac{\mathbf{p}^2}{2m} + V + H', \tag{5.131a}$$

$$\begin{aligned} 4m^2 H' &= -\boldsymbol{\sigma}\cdot\mathbf{p}(E_{\mathrm{NR}} - V)\boldsymbol{\sigma}\cdot\mathbf{p} \\ &= -(E_{\mathrm{NR}} - V)\mathbf{p}^2 + i\boldsymbol{\sigma}\cdot(\mathbf{p}V)\times\mathbf{p} + (\mathbf{p}V)\cdot\mathbf{p}, \end{aligned} \tag{5.131b}$$

where $\sigma_i\sigma_j = \delta_{ij} + i\varepsilon_{ijk}\sigma_k$ and $p_i V p_i = V\mathbf{p}^2 + (p_i V)p_i$ have been used to obtain (5.131b). The first term in (5.131b) represents a relativistic momentum correction, since to leading order $E_{\mathrm{NR}} - V \approx \mathbf{p}^2/2m$ means that

$$H'_{\mathrm{rel}} = -(E_{\mathrm{NR}} - V)\frac{\mathbf{p}^2}{4m^2} \approx -\frac{p^4}{8m^3}. \tag{5.132}$$

This potential then corresponds to the usual kinetic-energy correction

$$T = m(\gamma - 1) \approx \frac{\mathbf{p}^2}{2m} - \frac{p^4}{8m^3} + \cdots. \tag{5.133}$$

The second term in (5.131b) can be written for a spherically symmetric potential $V = V(r)$ as

$$i\boldsymbol{\sigma}\cdot(\mathbf{p}V)\times\mathbf{p} = \left(\frac{dV}{dr}\right)\boldsymbol{\sigma}\cdot(\hat{\mathbf{r}}\times\mathbf{p}) = \left(\frac{1}{r}\frac{dV}{dr}\right)\boldsymbol{\sigma}\cdot\mathbf{L}. \tag{5.134}$$

Consequently this term generates the familiar spin–orbit interaction in (5.131),

$$H'_{\mathrm{s.o.}} = \frac{1}{4m^2}\left(\frac{1}{r}\frac{dV}{dr}\right)\boldsymbol{\sigma}\cdot\mathbf{L}. \tag{5.135}$$

Note that this relativistic derivation of $H'_{\mathrm{s.o.}}$ includes the troublesome factor of $\frac{1}{2}$ referred to as the Thomas precession, a factor which must be introduced in a subtle manner in the usual nonrelativistic derivation of (5.135) when transforming from the rest frame of the electron to the laboratory i.e. rest frame of the nucleus (Problem 5.9).

Finally, we interpret the third term in (5.131b), but first note that $(\mathbf{p}V)\cdot\mathbf{p} = -(\nabla V)\cdot\nabla$ is not an hermitian operator. To convert it to hermitian form, we sandwich this term between a spatial integral over wave functions $\varphi^*$ and $\varphi$ in a symmetric fashion and then integrate by parts.

Discarding the resulting surface term, we are led to the real integrand

$$(\mathbf{p}V) \cdot \mathbf{p} \to -\tfrac{1}{2}[\varphi^*(\nabla V) \cdot \nabla\varphi + \varphi(\nabla V) \cdot \nabla\varphi^*] \to \tfrac{1}{2}\varphi^*(\nabla^2 V)\varphi. \quad (5.136)$$

Alternatively, one can understand (5.136) as due to a "renormalization" of the large-component wave function $\varphi$, no longer the exact nonrelativistic Schrödinger solution to $O(p^2/m^2)$, but related to it by $\psi_{\mathrm{NR}} = (1 + \mathbf{p}^2/8m^2)\varphi$ [see e.g. Sakurai (1967)]. In either case one is led to the hermitian *Darwin* interaction

$$H'_{\mathrm{Dar.}} = \frac{1}{8m^2}(\nabla^2 V). \quad (5.137)$$

While this hamiltonian does not have an obvious analog in the usual nonrelativistic formulation, it can be understood as a consequence of the relativistic effect of *Zitterbewegung*. To see this, account for the electron jittery motion to order of the Compton wavelength, $\delta r \lesssim m^{-1}$, by the expansion

$$H_{Zitt.} \equiv V(\mathbf{r} + \delta\mathbf{r}) - V(\mathbf{r}) = \tfrac{1}{2}\,\delta r_i\,\delta r_j \frac{\partial^2 V}{\partial r_i\,\partial r_j} + \cdots \approx \tfrac{1}{6}(\delta r)^2\nabla^2 V, \quad (5.138)$$

where symmetry dictates that the first-order term in (5.138) vanishes and the second-order derivatives are $\frac{1}{3}\nabla^2$. Comparing (5.138) with (5.137), we see that $(m\,\delta r \sim 1)$

$$H'_{\mathrm{Dar.}} \approx \tfrac{3}{4}(m\,\delta r)^{-2} H_{Zitt.} \sim H_{Zitt.}. \quad (5.139)$$

**Fine-Structure Energy Levels of Hydrogen.** To cap the discussion on bound states, we return to the Coulomb energy levels (5.118) for a Dirac atom, now specializing to the hydrogen atom with $Z = 1$:

$$E_{\mathrm{NR}} = -\frac{m\alpha^2}{2n^2}\left[1 + \frac{\alpha^2}{n}\left(\frac{1}{j + \frac{1}{2}} - \frac{3}{4n}\right)\right]. \quad (5.140)$$

To understand how (5.140) arises in the context of Schrödinger theory, we write $H_{\mathrm{NR}} = H_0 + H'$, where $H_0$ is the lowest-order Schrödinger hamiltonian with $V = -\alpha/r$, which generates the Bohr energy levels for $n = 1, 2, \ldots$ (structure):

$$E_n^0 = -\frac{m\alpha^2}{2n^2} = -\frac{\alpha}{2a_0 n^2}. \quad (5.141)$$

The small splittings of (5.141) to $O(\alpha^2)$ are the fine-structure corrections caused by $H'$ as given by (5.131b), viz.,

$$H' = H'_{\mathrm{rel}} + H'_{\mathrm{s.o.}} + H'_{\mathrm{Dar.}}, \quad (5.142a)$$

generating the first-order perturbation-theory shifts

$$\Delta E^{\mathrm{f.s.}}_{njl} = \langle nl|H'|nl\rangle = \langle nl|H'_{\mathrm{rel.}}|nl\rangle + \langle nl|H'_{\mathrm{s.o.}}|nl\rangle + \langle nl|H'_{\mathrm{Dar.}}|nl\rangle, \quad (5.142b)$$

where $|nl\rangle$ refers to the usual unperturbed Laguerre eigenfunction solutions $\psi^0_{nl}(r)$.

To calculate the relativistic momentum-correction shift, note that

$$\frac{1}{2m}\mathbf{p}^2\psi^0_{nl} = \left(E^0_n + \frac{\alpha}{r}\right)\psi^0_{nl} \tag{5.143a}$$

implies the square

$$\frac{1}{4m^2}\langle nl|p^4|nl\rangle = \left\langle nl\left|\left(E^0_n + \frac{\alpha}{r}\right)^2\right|nl\right\rangle, \tag{5.143b}$$

where one may easily verify that

$$\langle nl|r^{-1}|nl\rangle = \frac{m\alpha}{n^2}, \qquad \langle nl|r^{-2}|nl\rangle = \frac{m^2\alpha^2}{n^3(l+\frac{1}{2})}. \tag{5.144}$$

Then combining (5.143) and (5.144) with (5.132) we find

$$\langle nl|H'_{\text{rel.}}|nl\rangle = -\frac{1}{8m^3}\langle nl|p^4|nl\rangle = -\frac{m\alpha^4}{2n^3}\left[\frac{1}{l+\frac{1}{2}} - \frac{3}{4n}\right], \tag{5.145}$$

which breaks the $l$-degeneracy and is in fact the entire fine-structure shift for the Klein–Gordon atom (4.35).

The spin–orbit shift is found from (5.111a), (5.135) with $dV/dr = \alpha/r^2$ and

$$\langle nl|r^{-3}|nl\rangle = \frac{m^3\alpha^3}{n^3 l(l+\frac{1}{2})(l+1)}, \tag{5.146}$$

giving

$$\begin{aligned}\langle nl|H'_{\text{s.o.}}|nl\rangle &= \frac{\alpha}{4m^2}\left\langle nl\left|\frac{\boldsymbol{\sigma}\cdot\mathbf{L}}{r^3}\right|nl\right\rangle \\ &= \pm\frac{m\alpha^4}{4n^3}\frac{1}{(j+\frac{1}{2})(l+\frac{1}{2})}(1-\delta_{l,0})\end{aligned} \tag{5.147}$$

for $j = l \pm \frac{1}{2}$, but zero for $l = 0$.

Lastly, $\nabla^2 V = 4\pi\alpha\delta^3(\mathbf{r})$, and at the origin

$$|\psi^0_{nl}(0)|^2 = \frac{m^3\alpha^3}{\pi n^3}\delta_{l,0}, \tag{5.148}$$

which converts (5.137) to the Darwin shift

$$\langle nl|H'_{\text{Dar.}}|nl\rangle = \frac{\pi\alpha}{2m^2}\langle nl|\delta^3(\mathbf{r})|nl\rangle = \frac{m\alpha^4}{2n^3}\delta_{l,0}. \tag{5.149}$$

While this Darwin shift is nonvanishing only for $s$-states, the important point is that it is *precisely* what the spin-orbit shift would be for $l = 0$ if $1 - \delta_{l,0}$ were replaced by 1 in (5.147). In a sense then, the effect of *Zitterbewegung* on the Dirac atom corresponds to a "continuation in $l$" of the spin–orbit energy shift down to $l = 0$.

[redacted]

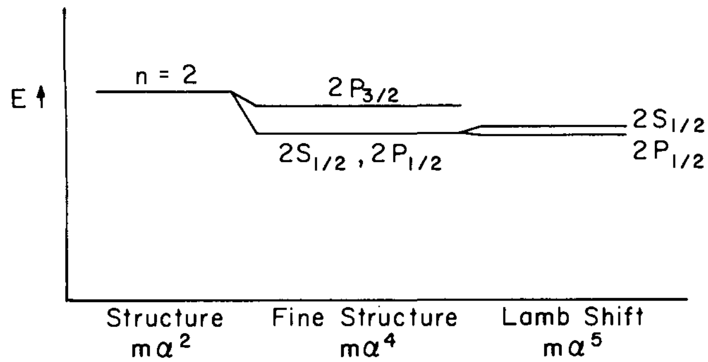



**Figure 5.1** The $n = 2$ energy-level splittings in hydrogen.

$2S_{\frac{1}{2}}$ and $2P_{\frac{1}{2}}$ energy levels are not degenerate. Rather, the $2S_{\frac{1}{2}}$ line is shifted upward relative to the $2P_{\frac{1}{2}}$ state by some 1058 Mc/sec, as pictured in Figure 5.1. While the theory of this "Lamb shift" is now well understood, a detailed explanation must be postponed until Chapter 15. Qualitatively speaking, this shift is caused by the "cloud" of photons and virtual electron-positron pairs which surround the bound electron, modifying its form factors as given in (5.97) and the resulting Coulomb interaction with the proton nucleus. Suffice it to say now that theory and experiment agree exactly to five significant figures, a truly remarkable result rarely equaled in modern science.

## 5.E Wave Equations for Other Fermi Particles

Thus far we have explored the consequences of the Dirac equation for massive, spin-$\frac{1}{2}$ fermions (e.g., electrons, protons). We now consider wave equations for massless spin-$\frac{1}{2}$ fermions (neutrinos) and also massive spin-$\frac{3}{2}$ fermions (e.g., the $\Delta$, 33 "resonance").

**Massless Spin-$\frac{1}{2}$ Particles.** Recall from the discussion in Section 3.D that the two helicity states for a massive spin-$\frac{1}{2}$ particle, $\lambda = \pm\frac{1}{2}$, become reduced to just one helicity state as $m \to 0$. The explicit structure of the two-component spinor boost operator in (3.94) shows that the $\lambda = \frac{1}{2}$ right-handed massless state (with spin parallel to the momentum) survives for the $(0, \frac{1}{2})$ representation, while $\lambda = -\frac{1}{2}$ survives for the $(\frac{1}{2}, 0)$ irreducible representation of the homogeneous Lorentz group. Experiment alone (see Chapter 13) has determined that the massless neutrino is left handed and the antineutrino is right handed. Consequently the two-component $(\frac{1}{2}, 0)$ free-particle wave function $\phi_{L\nu}$ for a neutrino satisfies a wave equation inferred from the boost in (3.94):

$$(i\partial_t + \boldsymbol{\sigma} \cdot \mathbf{p})\varphi_{L\nu}(x) = 0, \tag{5.155a}$$

whereas the $(0, \frac{1}{2})$ right-handed antineutrino wave equation is

$$(i\partial_t - \boldsymbol{\sigma} \cdot \mathbf{p})\varphi_{R\bar{\nu}}(x) = 0. \tag{5.155b}$$

These equations also follow from the $m = 0$ limit of (5.17).

While these wave equations were first obtained by Weyl (1929), they were rejected for 28 years because of noninvariance under spatial reflection ($\varphi_L \to \varphi_R$, $\varphi_R \to \varphi_L$). We shall return to this subject in Chapter 6. To link (5.155) with helicity eigenstates, note that the plane-wave solution of (5.155a) for *positive*-energy states, proportional to $e^{-ip \cdot x}$ with $E = |\mathbf{p}|$, satisfies

$$\boldsymbol{\sigma} \cdot \hat{\mathbf{p}}\varphi_{L\nu}(\mathbf{p}) = -\varphi_{L\nu}(\mathbf{p}) \tag{5.156a}$$

in accordance with a left-handed neutrino. Likewise, the plane-wave solution of (5.155b) for *positive*-energy states, proportional to $(e^{-ip \cdot x})^*$ with $E = |\mathbf{p}|$, obeys

$$\boldsymbol{\sigma} \cdot \hat{\mathbf{p}}\varphi_{R\bar{\nu}}(\mathbf{p}) = \varphi_{R\bar{\nu}}(\mathbf{p}), \tag{5.156b}$$

corresponding to a right-handed antineutrino. On the other hand, the plane-wave solution of (5.156b) for *negative*-energy neutrino states, also proportional to $e^{ip \cdot x}$ with $E = |\mathbf{p}|$, satisfies

$$\boldsymbol{\sigma} \cdot \hat{\mathbf{p}}\varphi_{\nu -}(-\mathbf{p}) = \varphi_{\nu -}(-\mathbf{p}). \tag{5.156c}$$

Comparing (5.156b) and (5.156c), we see that

$$\varphi_{R\bar{\nu}}(\mathbf{p}) \propto \varphi_{\nu -}(-\mathbf{p}). \tag{5.157}$$

This is our first concrete example of the Feynman interpretation, identifying a negative-energy particle state with a positive-energy antiparticle state. More examples will be given in Chapter 6.

It is possible, and extremely useful, to couch these two-component neutrino equations in four-component Dirac language. One way to proceed is to note that (5.155) corresponds to (5.17) with $m = 0$ along with the identification $\varphi_L \to \varphi_{L\nu}$, $\varphi_R \to \varphi_{R\bar{\nu}}$. We employ the (extreme relativistic) Weyl representation for the $\gamma$-matrices, (5.36), and apply it to the projection operator $P_{\pm} = \frac{1}{2}(1 \pm i\gamma_5)$ of (5.73):

$$P_+ = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \qquad P_- = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{5.158}$$

Then we define the Weyl-representation bispinors

$$u_{L\nu}(\mathbf{p}) = P_- u_{L\nu}(\mathbf{p}) = \begin{pmatrix} \varphi_{L\nu}(\mathbf{p}) \\ 0 \end{pmatrix}, \tag{5.159a}$$

$$v_{R\bar{\nu}}(\mathbf{p}) = P_+ v_{R\bar{\nu}}(\mathbf{p}) = \begin{pmatrix} 0 \\ e^{-i\phi}\varphi_{\nu -}(-\mathbf{p}) \end{pmatrix} \tag{5.159b}$$

with the positive-energy bispinor $u_{L\nu}(\mathbf{p})$ describing a left-handed neutrino and the negative-energy bispinor $v_{R\bar{\nu}}(\mathbf{p})$ describing a right-handed antineutrino, and both having momentum $\mathbf{p}$. As for the four-component dynamical equations that these bispinors satisfy, since $P_+$ projects out the lower negative-energy components in the Weyl representation and $u_{L\nu}$ has no such component, it is obvious that

$$P_+ u_{L\nu} = \tfrac{1}{2}(1 + i\gamma_5)u_{L\nu} = 0, \tag{5.160a}$$

and similarly,

$$P_- v_{R\bar{\nu}} = \tfrac{1}{2}(1 - i\gamma_5)v_{R\bar{\nu}} = 0. \tag{5.160b}$$

We may also proceed by starting with the free-particle Dirac equations (5.58) for $m = 0$,

$$\not{p}u(\mathbf{p}) = \not{p}v(\mathbf{p}) = 0, \tag{5.161}$$

then multiplying (5.161) on the left by $i\gamma_5\gamma_0$ and using $i\gamma_5\gamma_0\boldsymbol{\gamma} = \boldsymbol{\sigma}$, $E = |\mathbf{p}|$ we obtain

$$\boldsymbol{\sigma}\cdot\hat{\mathbf{p}}u(\mathbf{p}) = i\gamma_5 u(\mathbf{p}), \qquad \boldsymbol{\sigma}\cdot\hat{\mathbf{p}}v(\mathbf{p}) = -i\gamma_5 v(\mathbf{p}). \tag{5.162}$$

Next observe that the bispinor helicity eigenstates (5.65) become in this case

$$\boldsymbol{\sigma}\cdot\hat{\mathbf{p}}u_{L\nu}(\mathbf{p}) = -u_{L\nu}(\mathbf{p}), \qquad -\boldsymbol{\sigma}\cdot\hat{\mathbf{p}}v_{R\bar{\nu}}(\mathbf{p}) = v_{R\bar{\nu}}(\mathbf{p}). \tag{5.163}$$

Combining (5.162) with (5.163), we are once more led to (5.160).

Another useful bispinor expression for massless fermions is the probability current density $j_\mu = \bar{\psi}\gamma_\mu\psi$. Applying (5.159), we may write the momentum-space current for Weyl neutrinos obeying (5.155) as (dropping the factor of $\frac{1}{2}$ in $P_-$ by convention)

$$\langle \mathbf{p}'|j_\mu^{L\nu}|\mathbf{p}\rangle = \bar{u}_{L\nu}(\mathbf{p}')\gamma_\mu(1 - i\gamma_5)u_{L\nu}(\mathbf{p}). \tag{5.164}$$

Since $\gamma_\mu(1 - i\gamma_5) = (1 + i\gamma_5)\gamma_\mu$, the latter $\gamma$-matrix combinations used in (5.164) can also be expressed as $\bar{u}\bar{P}_- = \bar{u}P_+$. For an antineutrino current, it turns out that the analog of (5.164) is

$$\langle \mathbf{p}'|j_\mu^{R\bar{\nu}}|\mathbf{p}\rangle = -\bar{v}_{R\bar{\nu}}(\mathbf{p})\gamma_\mu(1 + i\gamma_5)v_{R\bar{\nu}}(\mathbf{p}'). \tag{5.165}$$

The projection operator $P_+$ in (5.165) again follows from (5.159), but the reason for reversing the momentum in the spinors in (5.165) must await the discussion of charge conjugation in Chapter 6.

Finally, note that even for $q^2 = (p' - p)^2 \neq 0$, we assume that (5.164) and (5.165) do not develop anomalous magnetic-moment contributions or form factors, because neutrinos interact only weakly with matter (Chapter 13). Such weak interaction experiments detect two types of neutrinos, associated with electrons and muons, respectively.

**Spin-$\frac{3}{2}$ Particles.** It is also possible, and again useful, to construct a Dirac bispinor wave function for a spin-$\frac{3}{2}$ free particle. In Section 2.E we built up a spin-$\frac{3}{2}$ two-component spinor by the Clebsch–Gordan combination (2.62) of a spin-1 polarization vector and a spin-$\frac{1}{2}$ spinor, with the spin-$\frac{1}{2}$ combination ($1 \times \frac{1}{2} = \frac{3}{2} + \frac{1}{2}$) removed by the condition (2.63), $\sigma_i\varphi_i = 0$. These rest-frame statements can be expressed in bispinor language by using the Dirac–Pauli representation and then boosting up to a general momentum frame with $p^2 = m^2$ (recall $-\delta_{ij} \to g_{\mu\nu} - p_\mu p_\nu/m^2$). This leads to the helicity sum

$$u_\mu^{(\lambda)}(p) = \sum_{\lambda'\lambda''} \langle 1\tfrac{1}{2}, \lambda'\lambda''|\tfrac{3}{2}\lambda\rangle \varepsilon_\mu^{(\lambda')}(\mathbf{p})u^{(\lambda'')}(\mathbf{p}), \tag{5.166}$$

where such spin-$\frac{3}{2}$ covariant bispinors satisfy a free-particle Dirac equation (Rarita and Schwinger 1941, Auvil and Brehm 1966)

$$(\not{p} - m)u_\mu(\mathbf{p}) = 0, \tag{5.167}$$

with (2.63) replaced by (see Problem 5.11)

$$\gamma^\mu u_\mu(\mathbf{p}) = 0. \tag{5.168}$$

Given (5.167) and (5.168), $u_\mu$ automatically obeys the weaker conditions

$$(p^2 - m^2)u_\mu(\mathbf{p}) = 0, \qquad p^\mu u_\mu(\mathbf{p}) = 0. \tag{5.169}$$

A polarization or spin sum can then be formed from these spin-$\frac{3}{2}$ bispinors in analogy with the spin-1 polarization sum (4.46) and the Dirac projection operator (5.69). In the rest frame this projection operator is

$$\mathscr{P}_{ij}^{(\frac{3}{2})} = \sum_{\lambda}^{4} \varphi_i^{(\lambda)}\varphi_j^{(\lambda)*} = \delta_{ij} - \tfrac{1}{3}\sigma_i\sigma_j, \tag{5.170}$$

because (2.63) requires $\sigma_i \mathscr{P}_{ij} = \mathscr{P}_{ij}\sigma_j = 0$, a condition obviously satisfied by (5.170), since $\sigma_i\sigma_i = 3$. The normalization of (5.170) is chosen so that $\mathscr{P}_{ij}\mathscr{P}_{jk} = \mathscr{P}_{ik}$. Then the boosted version of (5.170) is (see Problem 5.11)

$$\mathscr{P}_{\mu\nu}^{(\frac{3}{2})} = \sum_{\lambda}^{4} u_\mu^{(\lambda)}(\mathbf{p})\bar{u}_\nu^{(\lambda)}(\mathbf{p}) \tag{5.171a}$$

$$= -\left[\left(g_{\mu\nu} - \frac{p_\mu p_\nu}{m^2}\right)(\not{p} + m) + \frac{1}{3}\left(\gamma_\mu + \frac{p_\mu}{m}\right)(\not{p} - m)\left(\gamma_\nu + \frac{p_\nu}{m}\right)\right] \tag{5.171b}$$

$$= -\left[g_{\mu\nu} - \tfrac{1}{3}\gamma_\mu\gamma_\nu + \frac{1}{3m}(\gamma_\mu p_\nu - p_\mu\gamma_\nu) - \frac{2p_\mu p_\nu}{3m^2}\right](\not{p} + m). \tag{5.171c}$$

As long as the particles are on mass shell, this scheme can be extended to arbitrarily high spin [see e.g. Fronsdal (1958), Scadron (1968)].

There are, however, alternative formulations of spin-$\frac{3}{2}$ wave functions. Moreover, for $p^2 \neq m^2$ or for wave equations involving interactions, ambiguities arise and such theories are no longer unique. In the context of lagrangian field theory, interactions involving massive particles with spin-$\frac{3}{2}$ and higher (and sometimes even spin 1) are "nonrenormalizable" (see Chapter 15). Luckily, nature has been kind enough to see to it that most of the fundamental particles detected so far have low spin.

General references on the Dirac equation are: Dirac (1958), Hamilton (1959), Schweber (1961), Bjorken and Drell (1964), Muirhead (1965), Sakurai (1967), Pilkuhn (1967), Bethe and Jackiw (1968), Schiff (1968), Baym (1969), Berestetskii et al. (1971), Jauch and Rohrlich (1976).

# CHAPTER VI

# THE KLEIN-GORDON EQUATION

## VI.1 DERIVATION AND COVARIANCE

The requirements which special relativity imposes upon quantum mechanics are both fascinating and far-reaching† We begin our discussion of these effects by considering the wave equation obeyed by particles of zero spin, examples of which are provided by $\pi$, $K$, $\eta$ mesons, etc.

### Relativistic Schrödinger Equation

We review first the heuristic "derivation" of the Schrödinger equation which results from writing the standard non-relativistic relation between energy and momentum

$$E = \frac{\vec{p}^2}{2m} \tag{1.1}$$

and making the correspondence between the energy-momentum four-vector $p^\mu = (E, \vec{p})$ and the four-dimensional gradient operator $\partial^\mu$

$$i\partial^\mu = i\left(\frac{\partial}{\partial t}, -\vec{\nabla}\right) \sim p^\mu = (E, \vec{p}) \quad . \tag{1.2}$$

We have then

$$i\frac{\partial}{\partial t}\psi = -\frac{1}{2m}\vec{\nabla}^2\psi \tag{1.3}$$

which is the Schrödinger equation for a free particle. [Recall that if we make a Lorentz transformation to a frame $S'$ moving with velocity $v\hat{k}$ with respect to frame $S$, then

$$t' = \frac{t - vz}{\sqrt{1-\vec{v}^2}} \quad , \quad z' = \frac{z - vt}{\sqrt{1-\vec{v}^2}} \quad , \quad x' = x \quad , \quad y' = y \quad . \tag{1.4}$$

According to the chain rule

$$\begin{aligned}
\partial'^0 &= \frac{\partial}{\partial t'} = \frac{\partial t}{\partial t'}\frac{\partial}{\partial t} + \frac{\partial z}{\partial t'}\frac{\partial}{\partial z} = \frac{\partial t}{\partial t'}\partial^0 - \frac{\partial z}{\partial t'}\partial^3 \\
\partial'^3 &= -\frac{\partial}{\partial z'} = -\frac{\partial t}{\partial z'}\frac{\partial}{\partial t} - \frac{\partial z}{\partial z'}\frac{\partial}{\partial z} = -\frac{\partial t}{\partial z'}\partial^0 + \frac{\partial z}{\partial t'}\partial^3 \\
\partial'^1 &= \partial^1 \\
\partial'^2 &= \partial^2 \quad .
\end{aligned} \tag{1.5}$$

† Much of our discussion in this section is based on corresponding material found in Gordon Baym's *Lectures on Quantum Mechanics* [Ba69].

From the inverse Lorentz transformation

$$t = \frac{t' + vz'}{\sqrt{1 - \vec{v}^2}} \quad , \quad z = \frac{z' + vt'}{\sqrt{1 - \vec{v}^2}} \quad , \quad x = x' \quad , \quad y = y' \tag{1.6}$$

we find

$$\frac{\partial t}{\partial t'} = \frac{\partial z}{\partial z'} = \frac{1}{\sqrt{1 - \vec{v}^2}} \quad , \quad \frac{\partial t}{\partial z'} = \frac{\partial z}{\partial t'} = \frac{v}{\sqrt{1 - \vec{v}^2}} \quad . \tag{1.7}$$

Hence

$$\partial'^0 = \frac{\partial^0 - v\partial^3}{\sqrt{1 - \vec{v}^2}} \quad , \quad \partial'^3 = \frac{\partial^3 - v\partial^0}{\sqrt{1 - \vec{v}^2}} \quad , \quad \partial'^1 = \partial^1 \quad , \quad \partial'^2 = \partial^2 \tag{1.8}$$

so that $\partial^\mu$ is indeed a four-vector. Note that $\partial_\mu = \left(\frac{\partial}{\partial t}, \vec{\nabla}\right)$ does *not* have this property. The minus sign in Eq. 1.2 is essential.]

Now consider how Eq. 1.3 might be modified by the strictures of special relativity, wherein the relation between energy and momentum is

$$E = \sqrt{m^2 + \vec{p}^2} \quad . \tag{1.9}$$

As a first guess, we might try the wave equation

$$i\frac{\partial}{\partial t}\psi = \sqrt{m^2 - \vec{\nabla}^2}\,\psi \quad . \tag{1.10}$$

Seeing the square root with an operator inside is a bit peculiar, but this operation is well-defined if we write $\psi$ as a Fourier transform

$$\psi(\vec{x}, t) = \int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}\cdot\vec{x}} \phi(\vec{p}, t) \quad . \tag{1.11}$$

Then

$$\begin{aligned} i\frac{\partial}{\partial t}\psi(\vec{x}, t) &= \int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}\cdot\vec{x}} \sqrt{m^2 + \vec{p}^2}\,\phi(\vec{p}, t) \\ &= \int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}\cdot\vec{x}} \sqrt{m^2 + \vec{p}^2} \int d^3x'\, e^{-i\vec{p}\cdot\vec{x}'} \psi(\vec{x}', t) \quad . \end{aligned} \tag{1.12}$$

Interchanging orders of integration, this becomes

$$i\frac{\partial}{\partial t}\psi(\vec{x}, t) = \int d^3x'\, K(\vec{x}, \vec{x}')\psi(\vec{x}', t) \tag{1.13}$$

with

$$K(x, \vec{x}') = \int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}\cdot(\vec{x}-\vec{x}')} \sqrt{m^2 + \vec{p}^2} \quad . \tag{1.14}$$

For large $|\vec{x}-\vec{x}'|$ most values of $p$ except for those with $p \lesssim \frac{1}{|x-x'|}$ will lead to rapid oscillation of the exponential and consequently a very small value for the integral. In fact, the integral will be sizable only for $|x-x'| \lesssim \frac{1}{m}$, but this leads to a severe problem. Eq. 1.13 may be used via Taylor's expansion to relate $\psi(x, t+\delta t)$ to values of $\psi\left(x' \sim x \pm \frac{1}{m}, t\right)$

$$\begin{aligned}\psi(\vec{x}, t+\delta t) &= \psi(\vec{x}, t) + \delta t \frac{\partial \psi(\vec{x}, t)}{\partial t} \\ &= \psi(\vec{x}, t) - i\delta t \int d^3x' \, K(\vec{x}, \vec{x}')\psi(\vec{x}'t) \quad .\end{aligned} \tag{1.15}$$

This means that values of $\psi\left(\vec{x} \pm \frac{1}{m}, t\right)$ are affecting $\psi\left(\vec{x}, t+\delta t\right)$ even though these two spacetime points are outside the forward light cone, *i.e.*, since $\delta t$ can be made very small,

$$\left(\delta t^2 - \frac{1}{m^2}\right) < 0 \tag{1.16}$$

which violates causality. We must then abandon Eq. 1.10 as a possible relativistic wave equation.

## Klein-Gordon Equation

Next try squaring the energy, momentum relation, yielding

$$E^2 = \vec{p}^{\,2} + m^2 \quad . \tag{1.17}$$

The quantum mechanical analog becomes

$$-\frac{\partial^2}{\partial t^2}\phi(\vec{x}, t) = \left(-\vec{\nabla}^2 + m^2\right)\phi(\vec{x}, t) \tag{1.18}$$

or

$$\left(\Box + m^2\right)\phi(\vec{x}, t) = 0 \tag{1.19}$$

where

$$\Box = \frac{\partial^2}{\partial t^2} - \vec{\nabla}^2 \tag{1.20}$$

is the D'Alembertian. We see that Eq. 1.19 is simply the wave equation with the addition of a term involving $m^2$. This differential equation is properly relativistic and is called the Klein–Gordon equation.

That it is relativistically covariant can be seen by transforming from the original frame $S$ to a new frame $S'$. Since $\partial_\mu$ is a four-vector—

$$\Box = \partial_\mu \partial^\mu = \partial'_\mu \partial^{\mu'} = \Box' \tag{1.21}$$

—then according to an observer in $S'$, the equation reads

$$\left(\Box' + m^2\right)\phi(\vec{x}', t') = 0 \tag{1.22}$$

which has the same form as in the original frame $S$. This covariance of the equation is required by special relativity, since otherwise the form of the equation could be used in order to determine how fast one is moving.

We first look for plane wave solutions of the Klein–Gordon equation, having the form

$$\phi(\vec{x},t) = \exp\left(i\vec{p}\cdot\vec{x} - iEt\right) = \exp(-ip_\mu x^\mu) \quad . \tag{1.23}$$

Since

$$\begin{aligned} 0 &= \left(\frac{\partial^2}{\partial t^2} - \vec{\nabla}^2 + m^2\right)\exp\left(i\vec{p}\cdot\vec{x} - iEt\right) \\ &= \left(-E^2 + \vec{p}^2 + m^2\right)\exp\left(i\vec{p}\cdot\vec{x} - iEt\right) \end{aligned} \tag{1.24}$$

we see that $\phi(\vec{x},t)$ does indeed satisfy the Klein–Gordon equation provided

$$E^2 = \vec{p}^2 + m^2 \tag{1.25}$$

which was our starting point. However, the wavefunction $\phi(\vec{x},t)$ depends not upon $E^2$, but rather upon $E$, which has the values

$$E = \pm\sqrt{\vec{p}^2 + m^2} \quad . \tag{1.26}$$

The energy can be positive or ***negative***, which was somewhat disconcerting to the researchers first studying this equation. An additional problem was the inability to construct a conserved probability density. For the Schrödinger equation one has a probability density

$$\rho = \psi^*\psi \tag{1.27}$$

and probability current density

$$\vec{j} = \frac{i}{2m}\left(\psi^*\vec{\nabla}\psi - \left(\vec{\nabla}\psi^*\right)\psi\right) \tag{1.28}$$

with the property that

$$\frac{\partial}{\partial t}\rho + \vec{\nabla}\cdot\vec{j} = 0 \quad . \tag{1.29}$$

This insures that

$$\frac{d}{dt}\int_V \rho d^3r = -\int_S \vec{j}\cdot d\vec{S} \tag{1.30}$$

and says that any probability which flows out of the volume $V$ must pass through the surface. Thus probability is *locally* conserved, which is also required in a relativistic theory. (Simultaneous appearance and disappearance of probability at two spacelike separated points in one frame would not be simultaneous in another and would thus lead to trouble.)

It is easy to construct a conserved current density for the relativistic case via

$$j^\mu = (\rho,\vec{j}) = \frac{i}{2m}\left(\phi^*\partial^\mu\phi - (\partial^\mu\phi^*\phi)\right) \quad . \tag{1.31}$$

We observe that

$$\begin{aligned}\partial_\mu j^\mu = \frac{\partial\rho}{\partial t} + \vec{\nabla}\cdot\vec{j} &= \frac{i}{2m}\left((\partial^\mu\phi^*)\,\partial_\mu\phi + \phi^*\Box\phi - (\Box\phi^*)\,\phi - (\partial_\mu\phi^*)\,\partial^\mu\phi\right)\\ &= \frac{i}{2m}\left(-m^2\phi^*\phi + m^2\phi^*\phi\right) = 0\end{aligned} \tag{1.32}$$

and since, $\partial_\mu j^\mu$ is a scalar, this local conservation holds in all frames. The problem is that if we calculate $j^\mu$ for the plane wave solution $\phi(\vec{x},t)$ we find

$$j^\mu = \left(\frac{E}{m},\frac{\vec{p}}{m}\right) = \frac{p^\mu}{m} \quad , \tag{1.33}$$

whereby

$$\rho = \frac{E}{m} \tag{1.34}$$

can be either positive or negative, depending upon the sign of the energy. Such a negative "probability" density is obviously unsatisfactory.

Another peculiar feature of the relativistic wave equation is that it is *second* order in time, as opposed to the Schrödinger equation which is first order. This difference is an important one. For the Schrödinger equation, this means that, given the state vector $|\psi(0)>$ at time $t = 0$, one can determine the state at all future times, via

$$|\psi(t)\rangle = e^{-iHt}|\psi(0)\rangle \quad . \tag{1.35}$$

On the other hand, for the Klein–Gordon equation one requires *two* initial conditions—both the wavefunction $\phi(\vec{x},0)$ *and* its time derivative $\dot{\phi}(\vec{x},0)$.

These problems are resolved by the realization that a properly relativistic wave equation involves of necessity *both* particle *and* antiparticle degrees of freedom. If we identify the positive energy solution $\phi_{E>0}(\vec{x},t)$ with the particle, then the corresponding antiparticle solution is constructed from the negative energy solution via

$$\phi_{\text{antiparticle}}(\vec{x},t) = \phi^*_{E<0}(\vec{x},t) \quad . \tag{1.36}$$

Then, for example, in the case of a plane wave

$$\phi_{\text{antiparticle}}(\vec{x},t) = \left(e^{i\vec{p}\cdot\vec{x}+i|E|t}\right)^* = e^{-i\vec{p}\cdot\vec{x}-i|E|t} \tag{1.37}$$

which corresponds to positive energy time development.

This identification is made secure by writing the Klein–Gordon equation in the presence of a vector potential $A^\mu$. Via the aforementioned minimal substitution

$$i\nabla^\mu \to i\nabla^\mu - eA^\mu \tag{1.38}$$

where $e$ is the particle charge, the Klein–Gordon equation becomes, for a particle (positive energy) solution

$$\left((\nabla_\mu + ieA_\mu)(\nabla^\mu + ieA^\mu) + m^2\right)\phi_{E>0}(\vec{x},t) = 0 \quad . \tag{1.39}$$

Taking the complex conjugate equation for a negative energy solution, we find

$$\begin{aligned}&\left((\nabla_\mu + ieA_\mu)^* (\nabla^\mu + ieA^\mu)^* + m^2\right) \phi^*_{E<0}(x,t)\\ &= \left((\nabla_\mu - ieA_\mu)(\nabla^\mu - ieA^\mu) + m^2\right) \phi_{\text{antiparticle}}(\vec{x},t) = 0\end{aligned} \tag{1.40}$$

which is the Klein–Gordon equation for a particle of opposite charge. The existence of both particle *and* antiparticle degrees of freedom in the wavefunction is the reason that one needs *two* boundary conditions at time $t = 0$ in order to predict the future behavior.

The antiparticle degrees of freedom also resolve the problem of the negative "probability" density. Multiplying by the particle charge $e$, we identify

$$j^\mu = \frac{ie}{2m}\left(\phi^* \partial^\mu \phi - (\partial^\mu \phi^*)\,\phi\right) \tag{1.41}$$

as the electromagnetic current density. Then $\rho$ is the *charge* density which is positive (negative) for positive (negative) energy, *i.e.*, particle (antiparticle) solutions, as expected.

Returning to the plane wave solutions, we note that they are clearly Lorentz invariant since $p_\mu x^\mu$ is a Lorentz scalar. Thus an observer in the particle rest frame ($E = m$, $\vec{p} = 0$) writes the solution as $e^{-imt}$, while an observer in a frame moving with velocity $\vec{v} = -v\hat{k}$ with respect to the rest frame sees

$$E' = \frac{m}{\sqrt{1-\vec{v}^2}} \quad , \quad \vec{p}' = \frac{mv\hat{k}}{\sqrt{1-\vec{v}^2}} \tag{1.42}$$

and writes the wave function as

$$e^{i\vec{p}'\cdot\vec{x}' - iE't'} \tag{1.43}$$

which is identical since $p_\mu x^\mu$ has the same value in both frames.

The form of the current density is as expected since if we consider a region of space $d^3x$ in the rest frame containing charge

$$dq = \rho(\vec{x},t)d^3x \tag{1.44}$$

then in the primed frame the corresponding volume is

$$d^3x' = \sqrt{1-\vec{v}^2}\, d^3x \tag{1.45}$$

since the dimension of the volume in the direction of the boost is reduced by Lorentz contraction. The same amount of charge must be contained in this contracted region so

$$dq = \rho'(\vec{x}',t')d^3x' \ . \tag{1.46}$$

We find then

$$\frac{\rho'(\vec{x}',t')}{\rho(x,t)} = \frac{1}{\sqrt{1-\vec{v}^2}} = \frac{E'}{m} \ . \tag{1.47}$$

Also, we expect the spatial current density to be given by

$$\begin{aligned}\vec{j}'(\vec{x}',t') &= \rho(\vec{x}',t')\vec{v} = \rho(\vec{x}',t')\frac{\vec{p}'}{E'} \\ &= \frac{\vec{p}'}{m}\rho(\vec{x},t) \quad .\end{aligned} \tag{1.48}$$

These anticipated behaviors are obviously satisfied by the form

$$j^\mu = e\frac{p^\mu}{m} \tag{1.49}$$

arising from the plane wave solution.

### Two-Component Form

In order to deal with a more general class of solutions it is useful to write the Klein–Gordon equation in two-component form

$$\begin{aligned}\chi_1 &= \frac{1}{2}\left[\phi + \frac{i}{m}\left(\partial^0 + ieA^0\right)\phi\right] \\ \chi_2 &= \frac{1}{2}\left[\phi - \frac{i}{m}\left(\partial^0 + ieA^0\right)\phi\right] \quad .\end{aligned} \tag{1.50}$$

In terms of this notation the charge density becomes

$$\begin{aligned}\rho &= \frac{ie}{2m}\left[\phi^*\left(\partial^0 + ieA^0\right)\phi - \left(\left(\partial^0 - ieA^0\right)\phi^*\right)\phi\right] \\ &= \frac{ie}{2m}\left[-im\left(\chi_1 - \chi_2\right)\left(\chi_1 + \chi_2\right)^* - im\left(\chi_1 - \chi_2\right)^*\left(\chi_1 + \chi_2\right)\right] \\ &= e\left(|\chi_1|^2 - |\chi_2|^2\right) \quad .\end{aligned} \tag{1.51}$$

The components $\chi_1$, $\chi_2$ obey the coupled equations

$$\begin{aligned}\left(i\partial^0 - eA^0\right)\chi_1 &= \left[\frac{1}{2m}\left(-i\vec{\nabla} - e\vec{A}\right)^2 + \frac{m}{2}\right]\left(\chi_1 + \chi_2\right) + \frac{m}{2}\left(\chi_1 - \chi_2\right) \\ &= \frac{1}{2m}\left(-i\vec{\nabla} - e\vec{A}\right)^2\left(\chi_1 + \chi_2\right) + m\chi_1\end{aligned} \tag{1.52}$$

$$\begin{aligned}\left(i\partial^0 - eA^0\right)\chi_2 &= -\left[\frac{1}{2m}\left(-i\vec{\nabla} - e\vec{A}\right)^2 + \frac{m}{2}\right]\left(\chi_1 + \chi_2\right) + \frac{m}{2}\left(\chi_1 - \chi_2\right) \\ &= -\frac{1}{2m}\left(-i\vec{\nabla} - e\vec{A}\right)^2\left(\chi_1 + \chi_2\right) - m\chi_2\end{aligned} \tag{1.53}$$

These results are displayed most conveniently by defining a two-component "spinor"

$$\chi = \begin{pmatrix}\chi_1 \\ \chi_2\end{pmatrix} \tag{1.54}$$

in terms of which the Klein–Gordon equation hass the form

$$\left(i\partial^0 - eA^0\right)\chi = \left[\frac{1}{2m}\left(-i\vec{\nabla} - e\vec{A}\right)^2(\tau_3 + i\tau_2) + m\tau_3\right]\chi \tag{1.55}$$

where

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad , \quad \tau_3 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{1.56}$$

are the Pauli matrices. The charge density may be written as

$$\rho = e\chi^\dagger \tau_3 \chi \tag{1.57}$$

and the normalization condition becomes

$$\langle\chi|\chi\rangle = \int d^3x\, \chi^\dagger \tau_3 \chi = \pm 1, \tag{1.58}$$

the sign being determined by whether we start with particles (+) or antiparticles (−).

The Klein–Gordon Hamiltonian

$$H = \frac{1}{2m}\left(-i\vec{\nabla} - e\vec{A}\right)^2(\tau_3 + i\tau_2) + m\tau_3 + eA^0 \tag{1.59}$$

does not appear to be Hermitian since

$$(\tau_3 + i\tau_2)^\dagger = \tau_3 - i\tau_2 \neq \tau_3 + i\tau_2 \quad . \tag{1.60}$$

However, since the norm is defined using $\tau_3$—*cf.* Eq. 1.58—we have

$$\langle\chi'|H|\chi\rangle = \int d^3x\, \chi^{\dagger\prime}(x)\tau_3 H\chi(x) \tag{1.61}$$

and

$$\begin{aligned} \langle\chi'|H|\chi\rangle^* &= \left(\int d^3x \chi'^\dagger(x)\tau_3 H\chi(x)\right)^* \\ &= \int d^3x\, \chi^\dagger(x) H^\dagger \tau_3 \chi'(x) = \int d^3x \chi^\dagger(x)\tau_3\left(\tau_3 H^\dagger \tau_3\right)\chi'(x) \\ &= \left\langle\chi|\tau_3 H^\dagger \tau_3|\chi'\right\rangle \quad . \end{aligned} \tag{1.62}$$

The Hamiltonian is thus Hermitian provided

$$\tau_3 H^\dagger \tau_3 = H \quad , \tag{1.63}$$

and since

$$\begin{aligned} \tau_3(\tau_3 + i\tau_2)^\dagger \tau_3 &= \tau_3(\tau_3 - i\tau_2)\tau_3 \\ &= \tau_3 + i\tau_2 \end{aligned} \tag{1.64}$$

this condition is satisfied.

Now let's take a look at the free particle solutions ($A^0 = 0$, $\vec{A} = 0$) in terms of this notation. A positive energy solution with momentum $\vec{p}$ and normalized to unit density is given by

$$\phi(x) = \sqrt{\frac{m}{E}} e^{i\vec{p}\cdot\vec{x} - iEt} \equiv \sqrt{\frac{m}{E}} e^{-ip\cdot x} \quad . \tag{1.65}$$

This can be written in two component form as

$$\phi(x) = \chi^{(+)}(\vec{p})\, e^{-ip\cdot x} \tag{1.66}$$

where

$$\chi^{(+)}(\vec{p}) = \frac{1}{2\sqrt{mE}} \begin{pmatrix} m+E \\ m-E \end{pmatrix} \quad . \tag{1.67}$$

A corresponding negative energy solution

$$\phi(x) = \sqrt{\frac{m}{|E|}} e^{-i\vec{p}\cdot\vec{x} + i|E|t} \tag{1.68}$$

has the two component form

$$\chi^{(-)}(\vec{p}) = \frac{1}{2\sqrt{m|E|}} \begin{pmatrix} m-|E| \\ m+|E| \end{pmatrix} \quad . \tag{1.69}$$

Of course, $\chi^{(+)}(\vec{p})$ is orthogonal to $\chi^{(-)}(\vec{p})$

$$\left\langle \chi^{(+)}(\vec{p}) | \chi^{(-)}(\vec{p}) \right\rangle = \chi^{(+)\dagger}(\vec{p}) \tau_3 \chi^{(-)}(\vec{p}) = 0 \tag{1.70}$$

and, by completeness, any wavepacket can be expanded in terms of a linear combination of positive and negative energy solutions. That is, we can write

$$\begin{aligned} \phi(\vec{x},t) &= \int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}\cdot\vec{x}} \left( a_{\vec{p}}^{(+)}(t)\chi^{(+)}(\vec{p}) + a_{\vec{p}}^{(-)}(t)\chi^{(-)}(-\vec{p}) \right) \\ &= \int \frac{d^3p}{(2\pi)^3} \left( a_{\vec{p}}^{(+)}(t)\chi^{(+)}(\vec{p}) e^{i\vec{p}\cdot\vec{x}} + a_{\vec{p}}^{(-)}(t)\chi^{(-)}(\vec{p}) e^{-i\vec{p}\cdot\vec{x}} \right) \quad . \end{aligned} \tag{1.71}$$

Clearly then $a_{\vec{p}}^{(+)}(t)$ is the amplitude to be in the positive charge, positive energy state $\chi^{(+)}(\vec{p})$ while $a_{\vec{p}}^{(-)}(t)$ is the amplitude to be in negative charge, negative energy state $\chi^{(-)}(\vec{p})$. These amplitudes are given by

$$\begin{aligned} a_{\vec{p}}^{(+)}(t) &= \int d^3x\, e^{-i\vec{p}\cdot\vec{x}} \chi^{(+)\dagger}(\vec{p}) \tau_3 \phi(\vec{x},t) \\ a_{\vec{p}}^{(-)}(t) &= -\int d^3x\, e^{i\vec{p}\cdot\vec{x}} \chi^{(-)\dagger}(\vec{p}) \tau_3 \phi(\vec{x},t) \quad . \end{aligned} \tag{1.72}$$

If the wavefunction is normalized to $\pm 1$ we have

$$\begin{aligned}\pm 1 = \langle\phi|\phi\rangle &= \int d^3x\, \phi^\dagger(\vec{x},t)\tau_3\phi(\vec{x},t) \\ &= \int \frac{d^3p}{(2\pi)^3}\left(\left|a_{\vec{p}}^{(+)}(t)\right|^2 - \left|a_{\vec{p}}^{(-)}(t)\right|^2\right)\end{aligned} \tag{1.73}$$

while the energy and momentum expectation values are given by

$$\begin{aligned}E = \langle\phi|H_0|\phi\rangle &= \int d^3x\, \phi^\dagger(\vec{x},t)\tau_3 H_0\phi(\vec{x},t) \\ &= \int \frac{d^3p}{(2\pi)^3}\sqrt{m^2+\vec{p}^{\,2}}\left(\left|a_{\vec{p}}^{(+)}(t)\right|^2 + \left|a_{\vec{p}}^{(-)}(t)\right|^2\right)\end{aligned} \tag{1.74}$$

$$\begin{aligned}\vec{P} = \left\langle\phi\right| - i\vec{\nabla}\left|\phi\right\rangle &= \int d^3x\, \phi^\dagger(\vec{x},t)\tau_3 - i\vec{\nabla}_x\phi(\vec{x},t) \\ &= \int \frac{d^3p}{(2\pi)^3}\vec{p}\left(\left|a_{\vec{p}}^{(+)}(t)\right|^2 + \left|a_{\vec{p}}^{(-)}(t)\right|^2\right) \quad .\end{aligned} \tag{1.75}$$

## PROBLEM VI.1.1

### The Free Klein–Gordon Particle in a Magnetic Field

Consider a free charged Klein-Gordon particle of mass m and charge e immersed in a uniform magnetic field B in the z direction. Using the gauge

$$\vec{A} = \frac{1}{2}(\vec{B}\times\vec{r})$$

show that motion is quantized with energy

$$E_n = \sqrt{m^2+p_z^2+eB(2n+1)} \quad n = 0,1,2,\ldots$$

## PROBLEM VI.1.2

### Pair Production by a Time Varying Electric Field

A rapidly varying electric field can lead to the creation of particle-antiparticle pairs. Calculate to lowest order in $\alpha$ the probability per unit volume per unit time of producing such pairs in the presence of an external electric field

$$\vec{E}(t) = \hat{e}_x a\cos\omega t$$

and show that

$$\text{Prob.} = VT\frac{\alpha a^2}{6}\left(1-\frac{4m^2}{\omega^2}\right)^{\frac{3}{2}}\theta(\omega-2m)$$

Suggestion: Use as an interaction potential the usual form

$$H_{\text{int}} = e\int d^3x j_\mu A^\mu$$

where

$$j_\mu = \frac{i}{2m}(\phi^* \partial_\mu \phi - \partial \phi^* \phi)$$
$$\vec{A}(t) = -\hat{e}_x \frac{a}{\omega} \sin\omega t$$

Utilize normalized plane wave solutions of the Klein-Gordon equation

$$\phi(x) = \sqrt{\frac{m}{E}} \exp(i\vec{p} \cdot \vec{x} - iEt) \quad \text{with} \quad E = \sqrt{\vec{p}^2 + m^2}$$

and simple first order perturbation theory

$$\text{Amp} = -i \int_{-\frac{T}{2}}^{\frac{T}{2}} < f|H_{\text{int}}(t)|0 > dt$$

## VI.2 KLEIN'S PARADOX AND ZITTERBEWEGUNG

So far everything seems quite reasonable. However, there exist at least a few unexpected results within this formalism.

### Zitterbewegung

Consider the construction of a wavepacket containing *only* positive energy components.

$$\phi(\vec{x}, t) = \int \frac{d^3p}{(2\pi)^3} e^{i\vec{p} \cdot \vec{x}} a_p^{(+)}(t) \chi^{(+)}(\vec{p}) \tag{2.1}$$

with

$$a_{\vec{p}}^{(+)}(t) = e^{-iE_p t} f\left((\vec{p} - \vec{p}_0)^2\right) \quad . \tag{2.2}$$

We choose some function $f$ which is peaked about the origin, say

$$f\left((\vec{p} - \vec{p}_0)^2\right) \sim N \exp - \frac{(\vec{p} - \vec{p}_0)^2}{\Delta^2} \quad . \tag{2.3}$$

and consider the expectation value of the position operator $\vec{x}$

$$\begin{aligned} \vec{X}(t) = \langle \phi | \vec{x} | \phi \rangle &= \int d^3x \, \phi^\dagger(\vec{x}, t) \tau_3 \vec{x} \phi(\vec{x}, t) \\ &= \int d^3x \int \frac{d^3p}{(2\pi)^3} \int \frac{d^3p'}{(2\pi)^3} \vec{x} \, e^{i(\vec{p}' - \vec{p}) \cdot \vec{x}} a_{\vec{p}}^{(+)*}(t) a_{\vec{p}'}^{(+)}(t) \\ &\quad \times \chi^{(+)\dagger}(\vec{p}) \tau_3 \chi^{(+)}(\vec{p}') \quad . \end{aligned} \tag{2.4}$$

If we write $\vec{x} \, e^{i\vec{p}' \cdot \vec{x}} = -i \vec{\nabla}_{\vec{p}'} e^{i\vec{p}' \cdot \vec{x}}$ and integrate by parts

$$\begin{aligned} \vec{X}(t) &= i \int \frac{d^3p}{(2\pi)^3} a_{\vec{p}}^{(+)*}(t) \vec{\nabla}_{\vec{p}} a_{\vec{p}}^{(+)}(t) \\ &\quad + i \int \frac{d^3p}{(2\pi)^3} \left| a_p^{(+)}(t) \right|^2 \chi^{(+)\dagger}(\vec{p}) \tau_3 \nabla_{\vec{p}} \chi^{(+)}(\vec{p}) \quad . \end{aligned} \tag{2.5}$$

However,

$$\vec{\nabla}_{\vec{p}}\chi^{(+)}(\vec{p}) = \frac{1}{2\sqrt{mE_p}}\frac{\vec{p}}{E_p}\begin{pmatrix}1\\-1\end{pmatrix} - \frac{1}{4\sqrt{mE_p^3}}\frac{\vec{p}}{E_p}\begin{pmatrix}m+E_p\\m-E_p\end{pmatrix}$$
$$= -\frac{\vec{p}}{2E_p^2}\chi^{(-)}(\vec{p}) \tag{2.6}$$

so that the term $\chi^{(+)\dagger}(\vec{p})\tau_3\vec{\nabla}_{\vec{p}}\chi^{(+)}(\vec{p})$ vanishes by orthogonality. Then

$$\begin{aligned}\vec{X}(t) &= i\int\frac{d^3p}{(2\pi)^3}a_{\vec{p}}^{(+)*}(t)\vec{\nabla}_{\vec{p}}a_{\vec{p}}^{(+)}(t)\\ &= \int\frac{d^3p}{(2\pi)^3}\left(\frac{\vec{p}}{E_p}tf^2\left((\vec{p}-\vec{p}_0)^2\right)+i2(\vec{p}-\vec{p}_0)f\left((\vec{p}-\vec{p}_0)^2\right)f'\left((\vec{p}-\vec{p}_0)^2\right)\right)\\ &\cong \frac{\vec{p}_0}{E_{p_0}}t\ .\end{aligned} \tag{2.7}$$

Thus the wave packet begins from the origin at $t = 0$ and moves with uniform velocity $\vec{v}_0 = \frac{\vec{p}_0}{E_{p_0}}$, as expected. Now consider the width of the packet. For simplicity pick $\vec{p} = 0$. Then

$$\left\langle(\vec{x}-\bar{\vec{x}})^2\right\rangle = \left\langle\phi|\vec{x}^2|\phi\right\rangle = -\int\frac{d^3p}{(2\pi)^3}\left[a_{\vec{p}}^{(+)*}(t)\vec{\nabla}_{\vec{p}}^2a_{\vec{p}}^{(+)}(t)+\frac{\vec{p}^2}{4E_p^4}\left|a_{\vec{p}}^{(+)}(t)\right|^2\right] \tag{2.8}$$

which has two components. By dimensional analysis the first piece gives the expected result

$$-\int\frac{d^3p}{(2\pi)^3}a_{\vec{p}}^{(+)*}(t)\vec{\nabla}_{\vec{p}}^2a_{\vec{p}}^{(+)}(t)\sim\frac{1}{\Delta^2} \tag{2.9}$$

where $\Delta$ is determined by the functional form of $f(\vec{p}^2)$. The second piece is more interesting, yielding

$$-\int\frac{d^3p}{(2\pi)^3}\frac{\vec{p}^2}{4E_p^4}\left|a_{\vec{p}}^{(+)}(t)\right|^2\sim\frac{1}{m^2}K\left(\frac{m}{\Delta}\right)\ . \tag{2.10}$$

where K is a function depending on the specific form of $a_{\vec{p}}^{(+)}(t)$. We see that even if $\Delta$ is very large so that the first component of the width is very small, there is still a contribution $\delta x \sim \frac{1}{m}$ of order the Compton wavelength of the particle. This represents a minimum width of the wavepacket and cannot be made smaller so long as only positive energy components are present.

We may explore this point further by attempting to construct a wavepacket localized at the origin

$$\phi(\vec{x}, t=0) = \delta^3(x)\begin{pmatrix}\rho\\\sigma\end{pmatrix} \tag{2.11}$$

where $\rho, \sigma$ are arbitrary numbers. We find then

$$\begin{aligned} a_{\vec{p}}^{(+)}(t) &= e^{-iE_p t}\frac{1}{2\sqrt{mE_p}}\left[(E_p+m)\rho+(E_p-m)\sigma\right] \\ a_{\vec{p}}^{(-)*}(t) &= -e^{iE_p t}\frac{1}{2\sqrt{mE_p}}\left[(E_p-m)\rho+(E_p+m)\sigma\right] \quad , \end{aligned} \tag{2.12}$$

so that $a_{\vec{p}}^{(-)}(t) \neq 0$. That is, if one wishes to localize a wavepacket within a distance smaller than a Compton wavelength negative energy components are *required*.

Such negative energy components arise naturally in another way as well. Suppose we start with a positive-energy-only wavepacket and apply the position operator $\vec{x}$. Then, as shown earlier

$$\begin{aligned} \vec{x}\phi(\vec{x},t) &= \vec{x}\int \frac{d^3p}{(2\pi)^3}e^{i\vec{p}\cdot\vec{x}}a_{\vec{p}}^{(+)}(t)\chi^{(+)}(\vec{p}) \\ &= i\int \frac{d^3p}{(2\pi)^3}e^{i\vec{p}\cdot\vec{x}}\left[\chi^{(+)}(\vec{p})\vec{\nabla}_{\vec{p}}a_{\vec{p}}^{(+)}(t) - \chi^{(-)}(\vec{p})\frac{\vec{p}}{2E_p^2}a_{\vec{p}}^{(+)}(t)\right] \quad . \end{aligned} \tag{2.13}$$

so that the position operator introduces a negative energy piece into the wavefunction even if there was none present originally. Equivalently, multiplication by the potential energy $eA^0(\vec{x})$ introduces such negative energy states, so that whenever a wavepacket interacts with a potential we should not be surprised to find negative energy states appearing.

This mixture of positive and negative energy components in the wavepacket has an interesting consequence if we evaluate the expectation value of the position operator

$$\begin{aligned} \langle\phi(t)|\vec{x}|\phi(t)\rangle &= \int \frac{d^3p}{(2\pi)^3}\,\frac{\vec{p}}{E_p}t\left(\left|a_{\vec{p}}^{(+)}(t)\right|^2+\left|a_{\vec{p}}^{(-)}(t)\right|^2\right) \\ &\quad - \mathrm{Re}\int \frac{d^3p}{(2\pi)^3}\,\frac{\vec{p}}{E_p^2}a_{\vec{p}}^{(+)*}(t)a_{\vec{p}}^{(-)}(t) \quad . \end{aligned} \tag{2.14}$$

The first component represents just the expected uniform velocity motion of the packet. However, the second piece is more interesting. Since

$$a_{\vec{p}}^{(+)*}(t)a_{\vec{p}}^{(-)}(t) \sim e^{2i|E_p|t} \tag{2.15}$$

this term represents a rapid—$\omega \geq 2m$—wiggling of the position of the particle about its central location due to the interference of positive and negative energy components. This rapid movement—called zitterbewegung or jitter motion—is the price one pays for localization with $\delta x \lesssim \frac{1}{m}$, or for interaction with a potential. In the latter case, since positive and negative energies correspond to positive and negative charges, the particle and antiparticle components of the wavepacket travel in opposite directions. Thus the interference damps out after a time $\Delta t \sim \frac{1}{m}$ once interaction with the potential has ceased. An exception is when the potential is very strong—$V - E > m$. This problem is called Klein's paradox for reasons which will become apparent.

## Klein's Paradox

Imagine a Klein–Gordon particle of mass $m$, charge e, and energy $E = \sqrt{p^2 + m^2}$ incident from the left upon a potential step

$$V(x) = V_0\theta(x) = eA^0(x) \tag{2.16}$$

located at the origin, as shown in Figure VI.1. Since the potential

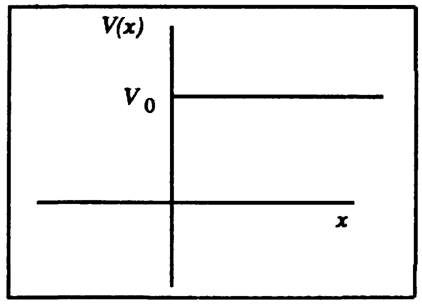



*Fig. VI.1: Step potential used for the discussion of the Klein paradox.*

is constant, the solutions can be represented in terms of plane waves. We look for stationary state solutions

$$\phi(x) = \begin{cases} a\, e^{ikx} + b\, e^{-ikx} & x < 0 \qquad k = \sqrt{E^2 - m^2} \\ c\, e^{iqx} & x > 0 \qquad q = \sqrt{(E - V_0)^2 - m^2} \ . \end{cases} \tag{2.17}$$

These are seen to be solutions of the Klein–Gordon equation in the regions $x > 0$, $x < 0$, respectively. Now, as usual, match the wavefunction and its first spatial derivative at the origin, yielding

$$\begin{aligned} \phi(0^+) &= c = a + b = \phi(0^-) \\ \phi'(0^+) &= iqc = ik(a - b) = \phi'(0^-) \end{aligned} \tag{2.18}$$

whose solution is

$$\frac{c}{a} = \frac{2}{1 + \frac{q}{k}} \ , \qquad \frac{b}{a} = \frac{1 - \frac{q}{k}}{1 + \frac{q}{k}} \ . \tag{2.19}$$

The transmission and reflection coefficients are calculated as

$$\begin{aligned} T &= \frac{q}{k}\left|\frac{c}{a}\right|^2 = 4\frac{q}{k}\,\frac{1}{\left|1 + \frac{q}{k}\right|^2} \\ R &= \left|\frac{b}{a}\right|^2 = \left|\frac{1 - \frac{q}{k}}{1 + \frac{q}{k}}\right|^2 \ . \end{aligned} \tag{2.20}$$

If the kinetic energy $E - m$ is above the height of the barrier—$E - m > V_0$—then $k, q$ are both real and positive, yielding

$$T = 4\frac{qk}{(k + q)^2} \ , \qquad R = \left(\frac{k - q}{k + q}\right)^2 \ , \qquad R + T = 1 \ . \tag{2.21}$$

Thus the incident beam is partly reflected and partly transmitted, as expected from our experience in the corresponding non-relativistic problem. On the other hand, if the incident kinetic energy is less than the height of the barrier, but $|V_0 - E| < m$ we see that $q$ is imaginary. Then

$$R = 1 \ , \qquad T = 0 \ , \qquad R + T = 1 \tag{2.22}$$

which again agrees with the non-relativistic analog.

Suppose, however, that there exists a *very* strong potential—$V_0 > E + m$. In this case $q$ becomes real again but *negative*. Then

$$R = \left(\frac{k + |q|}{k - |q|}\right)^2 > 1 \ , \qquad T = -\frac{4k|q|}{(k - |q|)^2} < 0 \tag{2.23}$$

but

$$R + T = 1 \ . \tag{2.24}$$

Probability is still conserved, but only at the cost of a *negative* transmission coefficient and a reflection coefficient which exceeds unity. This is the paradoxical result which confronted Klein and others.

In light of our present knowledge there exists no paradox. In the case that

$$V_0 - E > m \tag{2.25}$$

the potential is sufficiently strong to create particle-antiparticle pairs. The antiparticles are *attracted* by the potential and create a negatively charged current moving to the right. This is the origin of the negative transmission coefficient. The particles, on the other hand, are reflected from the barrier and combine with the incident particle beam (which is completely reflected) leading to a positively charged current, moving to the left and with magnitude greater than that of the incident beam. Thus $R > 1$, as found.

Another way of thinking of this is in terms of what happens to the energy spectrum of the Klein–Gordon equation when a potential $V > 0$ is turned on adiabatically [Sa 67]. When $V = 0$ this spectrum ranges from $m < E < \infty$ and $-\infty < E < -m$. Now consider a positive energy solution as shown in Figure VI.2. As $V$ is increased from zero, this energy level first finds itself in the forbidden region where solutions are strongly damped. However, when $E < V - m$ we are again in a region of oscillatory solutions. From Figure VI.2 it is clear that even though $E > 0$ this is essentially an antiparticle solution. The tunneling from region I to region III should be considered a transition from a particle state (when $V = 0$) to an antiparticle state (when $V > E + m$) as described above.

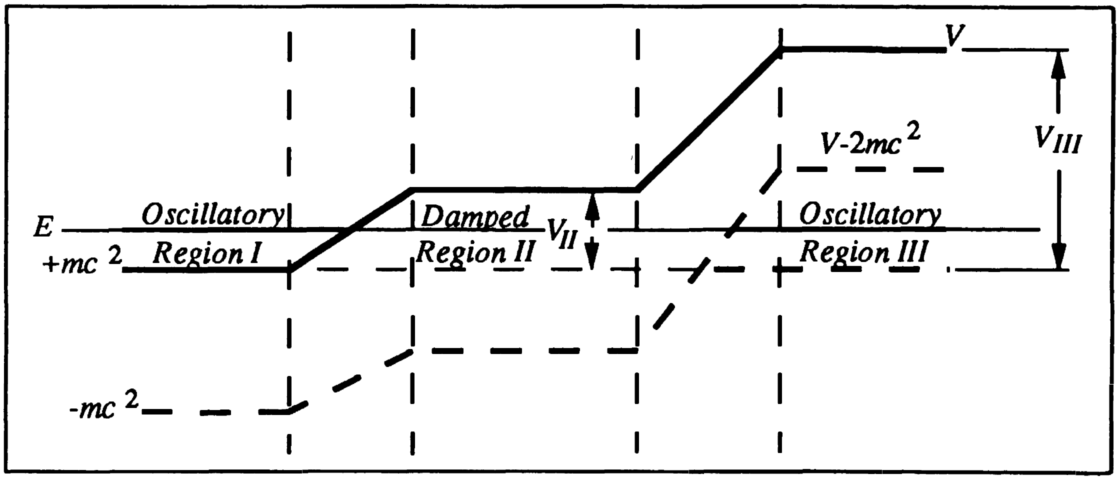



*Fig. VI.2 Klein-Gordon energy levels in the presence of a potential.*

There is no mystery then. Any problems which arise come from attempting to apply a simple single-particle wavefunction picture to what is obviously a many-body situation. The correct way in which to handle all the subtlety of this problem is via the formalism of quantum field theory. Nevertheless, the elementary wavefunction paradigm allows an accurate sketch of the physics involved.

## VI.3 THE COULOMB SOLUTION: MESONIC ATOMS

The Coulomb bound state problem can also straightforwardly analyzed. Before looking at the exact solution, however, it is useful to make the connection with the Schrödinger formalism by deriving an effective Hamiltonian for the situation that the Klein–Gordon particle is non-relativistic.

### Effective Schrödinger Equation

Consider the two component formalism—$\chi_1, \chi_2$—and look for a stationary state solution with

$$i\partial^0 \chi_j \approx (m+W)\chi_j \qquad j = 1,2 \tag{3.1}$$

where $W << m$. Then

$$\begin{aligned}(m+W-e\phi)\chi_1 &= \left(m + \frac{1}{2m}(\vec{p}-e\vec{A})^2\right)\chi_1 + \frac{1}{2m}(\vec{p}-e\vec{A})^2\chi_2 \\ (m+W-e\phi)\chi_2 &= -\left(m + \frac{1}{2m}(\vec{p}-e\vec{A})^2\right)\chi_2 - \frac{1}{2m}(\vec{p}-e\vec{A})^2\chi_1 \quad .\end{aligned} \tag{3.2}$$

Solving the second equation for $\chi_2$ we find

$$\begin{aligned}\chi_2 &\approx -\frac{1}{2m+W-e\phi}\,\frac{1}{2m}(\vec{p}-e\vec{A})^2\chi_1 \\ &\approx -\frac{1}{4m^2}\left(1-\frac{W-e\phi}{2m}\right)(\vec{p}-e\vec{A})^2\chi_1 \quad .\end{aligned} \tag{3.3}$$

Substitution into the top equation yields

$$(W - e\phi)\chi_1 \cong \frac{1}{2m}(\vec{p} - e\vec{A})^2\chi_1 - \frac{1}{8m^3}(\vec{p} - e\vec{A})^2 \left(1 - \frac{W - e\phi}{2m}\right)(\vec{p} - e\vec{A})^2\chi_1 \quad (3.4)$$

or

$$W\left(1 - \frac{1}{16m^4}(\vec{p} - e\vec{A})^4\right)\chi_1$$
$$= \left(\frac{(\vec{p} - e\vec{A})^2}{2m} + e\phi - \frac{(\vec{p} - e\vec{A})^4}{8m^2} - (\vec{p} - e\vec{A})^2\frac{e\phi}{16m^4}(\vec{p} - e\vec{A})^2\right)\chi_1 \ . \quad (3.5)$$

However, $\chi_1$ is not normalized to unity. Rather

$$1 = \int d^3x \left(\chi_1^\dagger\chi_1 - \chi_2^\dagger\chi_2\right) \approx \int d^3x\, \chi_1^\dagger \left(1 - \frac{(\vec{p} - e\vec{A})^4}{16m^4}\right)\chi_1$$
$$\approx \int d^3x\, \chi_1'^\dagger\chi_1' \quad (3.6)$$

where we have defined

$$\chi_1' = \left(1 - \frac{(\vec{p} - e\vec{A})^4}{32m^4}\right)\chi_1 \ . \quad (3.7)$$

Thus multiply Eq. 3.5 by the factor $\left(1 + \frac{(\vec{p}-e\vec{A})^4}{32m^4}\right)$, yielding

$$W\chi_1' = \left(\frac{(\vec{p} - e\vec{A})^2}{2m} - \frac{(\vec{p} - e\vec{A})^4}{8m^3} - \frac{1}{16m^4}(\vec{p} - e\vec{A})^2 e\phi(\vec{p} - e\vec{A})^2\right.$$
$$\left. + \left(1 + \frac{(\vec{p} - e\vec{A})^4}{32m^4}\right)e\phi\left(1 + \frac{(\vec{p} - e\vec{A})^4}{32m^4}\right)\right)\chi_1' + \ldots$$
$$= \left(\frac{(\vec{p} - e\vec{A})^2}{2m} - \frac{(\vec{p} - e\vec{A})^4}{8m^3} + e\phi + \frac{1}{32m^4}\left[(\vec{p} - e\vec{A})^2, \left[(\vec{p} - e\vec{A})^2, e\phi\right]\right]\right)\chi_1' \ . \quad (3.8)$$

The effective Schrödinger Hamiltonian is then

$$H = \frac{1}{2m}(-i\vec{\nabla} - e\vec{A})^2 - \frac{1}{8m^3}(-i\vec{\nabla} - e\vec{A})^4 + e\phi$$
$$+ \frac{1}{32m^4}\left[(-i\vec{\nabla} - e\vec{A})^2, \left[(-i\vec{\nabla} - eA)^2, e\phi\right]\right] \ . \quad (3.9)$$

and we can identify its various components as follows:

a) The terms

$$\frac{(\vec{p} - e\vec{A})^2}{2m} + e\phi \quad (3.10)$$

represent the usual non-relativistic energy.

b) Relativistically the kinetic energy is

$$T = \sqrt{m^2 + (\vec{p} - e\vec{A})^2} - m \tag{3.11}$$

whose non-relativistic approximation is

$$T = \frac{(\vec{p} - e\vec{A})^2}{2m} - \frac{(\vec{p} - e\vec{A})^4}{8m^3} + \dots \ . \tag{3.12}$$

Thus the piece $-\frac{1}{8m^3}(\vec{p} - e\vec{A})^4$ represents a relativistic $\mathcal{O}\left(\frac{v^2}{c^2}\right)$ correction to the usual kinetic energy.

c) The origin of the double commutator term is more subtle and is associated with the zitterbewegung motion discussed earlier. By completeness we can expand a bound state solution in terms of plane wave solutions. This expansion will involve, in general, a combination of positive and negative energy solutions and will thus lead to zitterbewegung motion of magnitude $(\delta x)^2 \sim \frac{1}{m^2}$ about the usual trajectory. This leads, in general, to a shift in the energy of magnitude

$$\begin{aligned} V(x + \delta x) - V(x) &\approx \nabla_i V(x) \langle \delta x_i \rangle + \frac{1}{2!} \nabla_i \nabla_j V(x) \langle \delta x_i \delta x_j \rangle + \dots \\ &\approx \frac{1}{6m^2} \vec{\nabla}^2 V(x) + \dots \end{aligned} \tag{3.13}$$

In the case of spin-1/2 we shall see that a term of precisely this form—the Darwin term—appears in the effective Hamiltonian. For the case of spinless particles the zitterbewegung term arises, however, in $\mathcal{O}\left(\frac{v^4}{c^4}\right)$.

## Mesonic Atoms

A spin-zero "atom" actually exists in nature when a $\pi^-$ or $K^-$ meson is captured by a nucleus. This object is called a "pionic" or "kaonic" atom. We can calculate the energy levels of such a system approximately by treating the $\vec{p}^4$ term as a perturbation. In lowest order we have

$$H_0 = \frac{\vec{p}^2}{2m} - \frac{Z\alpha}{r} \tag{3.14}$$

which is simply hydrogen-like atomic Hamiltonian, yielding eigenvalues

$$E_n = -\frac{Z^2\alpha^2}{2n^2} m \tag{3.15}$$

and eigenfunctions $\psi_n(\vec{r})$ with

$$H_0 \psi_n(\vec{r}) = \left( \frac{\vec{p}^2}{2m} - \frac{Z\alpha}{r} \right) \psi_n(\vec{r}) = E_n \psi_n(\vec{r}). \tag{3.16}$$

Then

$$\begin{aligned}\frac{\vec{p}^4}{8m^3}\psi_n(r) &= \frac{1}{2m}\left(E_n + \frac{Z\alpha}{r}\right)^2 \psi_n(r) \\ &= \frac{1}{2m}\left(m^2\frac{Z^4\alpha^4}{4n^4} - m\frac{Z^3\alpha^3}{n^2 r} + \frac{Z^2\alpha^2}{r^2}\right)\psi_n(r) \ .\end{aligned} \tag{3.17}$$

Since

$$\left\langle \frac{1}{r}\right\rangle = m\frac{Z\alpha}{n^2} \ , \quad \left\langle \frac{1}{r^2}\right\rangle = m^2\frac{Z^2\alpha^2}{n^3\left(\ell+\frac{1}{2}\right)} \tag{3.18}$$

we find

$$\begin{aligned}\left\langle \frac{\vec{p}^4}{8m^3}\right\rangle &= \frac{1}{2m}\left(m^2\frac{Z^4\alpha^4}{4n^4} - m^2\frac{Z^4\alpha^4}{n^4} + m^2\frac{Z^4\alpha^4}{n^3\left(\ell+\frac{1}{2}\right)}\right) \\ &= \frac{mZ^4\alpha^4}{2n^3}\left(\frac{1}{\ell+\frac{1}{2}} - \frac{3}{4n}\right) \ .\end{aligned} \tag{3.19}$$

Also, we note that

$$\left[\vec{p}^2, \phi(\vec{r})\right] = -Ze\delta^3(r) \ . \tag{3.20}$$

Then

$$\left\langle \psi_n\right| \left[\vec{p}^2, \left[\vec{p}^2, \phi\right]\right] \left|\psi_n\right\rangle = Ze\int d^3r\, \delta^3(r)\left(\left(\vec{\nabla}^2\psi_n^*(r)\right)\psi_n(r) - \psi_n^*(r)\vec{\nabla}^2\psi_n(r)\right) = 0 \tag{3.21}$$

so the Darwin term does not contribute.

The energy levels become

$$E_{n\ell} = -m\frac{Z^2\alpha^2}{2n^2}\left[1 + \frac{Z^2\alpha^2}{n^2}\left(\frac{n}{\ell+\frac{1}{2}} - \frac{3}{4}\right) + \dots\right] \ . \tag{3.22}$$

We observe that the $\vec{p}^4$ term acts as a fine structure term, removing the $\ell$ degeneracy of the hydrogen atom and lowering the energy for states of smaller angular momentum.

We can also solve this system exactly. Using

$$e\phi(r) = -\frac{Z\alpha}{r} \tag{3.23}$$

the Klein–Gordon equation becomes

$$\left(\left(E + \frac{Z\alpha}{r}\right)^2 + \vec{\nabla}^2 - m^2\right)\psi(\vec{r}) = 0 \ . \tag{3.24}$$

If we look for solutions having a definite angular momentum $\ell$

$$\psi(\vec{r}) = \psi_\ell(r) Y_\ell^m(\hat{r}) \ . \tag{3.25}$$

we require

$$\begin{aligned}\left(\left(E+\frac{Z\alpha}{r}\right)^2+\frac{1}{r}\frac{\partial^2}{\partial r^2}r-\frac{\ell(\ell+1)}{r^2}-m^2\right)\psi_\ell(r)&\\ =\left(E^2-m^2+2\frac{Z\alpha E}{r}+\frac{1}{r}\frac{\partial^2}{\partial r^2}r-\frac{\ell(\ell+1)-Z^2\alpha^2}{r^2}\right)\psi_\ell(r)=0\ .&\end{aligned} \tag{3.26}$$

Then making the identification

$$\begin{aligned}\ell(\ell+1)-Z^2\alpha^2&=\ell'(\ell'+1)\\ E^2-m^2=k^2&=2m'E'\\ E&=m'\end{aligned} \tag{3.27}$$

Eq. 3.26 becomes

$$\left(\frac{1}{r}\frac{\partial^2}{\partial r^2}r-\frac{\ell'(\ell'+1)}{r^2}+\frac{2m'Z\alpha}{r}+k^2\right)\psi(r)=0 \tag{3.28}$$

which is identical to the differential equation which arises in the usual Schrödinger equation solution of the hydrogen atom, except that in the present case $\ell'$ is *not* an integer. It is necessary then to analytically continue the hydrogen solutions in order to apply them here. We have

$$E'_s=-\frac{Z^2\alpha^2}{2s^2}m' \tag{3.29}$$

and

$$E=\frac{m}{\left(1+\frac{Z^2\alpha^2}{s^2}\right)^{1/2}}\ . \tag{3.30}$$

However, $s$ is not an integer but is defined rather in terms of $\ell'$ †—

$$s=n+\ell'-\ell\ , \tag{3.31}$$

with $\ell'$ determined by

$$\ell'(\ell'+1)+\frac{1}{4}=\left(\ell'+\frac{1}{2}\right)^2=\ell(\ell+1)-Z^2\alpha^2+\frac{1}{4}$$

*i.e.*

,

$$\ell'=-\frac{1}{2}\pm\sqrt{\left(\ell+\frac{1}{2}\right)^2-Z^2\alpha^2}\ . \tag{3.32}$$

† The positive sign in front of $\ell'-\ell$ is determined by the hydrogen atom condition that $n-\ell$ is a non-negative integer.

Although mathematically either sign is allowed, only the positive sign allows a normalizable solution as $\alpha \to 0$. Then

$$s = n - \frac{1}{2} - \ell + \sqrt{\left(\ell + \frac{1}{2}\right)^2 - Z^2\alpha^2} \tag{3.33}$$

so that

$$E = m\left[1 + \frac{Z^2\alpha^2}{\left(n - \ell - \frac{1}{2} + \sqrt{(\ell + \frac{1}{2})^2 - Z^2\alpha^2}\right)^2}\right]^{-1/2} \tag{3.34}$$

Noting that

$$s \approx n - \frac{1}{2}\frac{Z^2\alpha^2}{\ell + \frac{1}{2}} \tag{3.35}$$

we find

$$\begin{aligned} E &\approx m\left(1 - \frac{1}{2}\frac{Z^2\alpha^2}{s^2} + \frac{3}{8}\frac{Z^4\alpha^4}{s^4} + \ldots\right) \\ &\approx m\left(1 - \frac{1}{2}\frac{Z^2\alpha^2}{n^2}\left(1 + \frac{Z^2\alpha^2}{n^2}\left(\frac{n}{\ell + \frac{1}{2}} - \frac{3}{4}\right)\right)\ldots\right) \quad . \end{aligned} \tag{3.36}$$

which is in complete agreement with Eq. 3.22 obtained perturbatively.

In comparing this prediction to experimental data on pionic atoms, various corrections are required:

i) the reduced mass

$$\mu = \frac{mM}{m + M} \tag{3.37}$$

must be utilized in place of the pion mass $m$;

ii) account must be taken of the fact that the central nucleus is not a point charge but has a radius $R \sim 1.2 \times A^{\frac{1}{3}}$fm;

iii) correction must be made for the so-called vacuum polarization, wherein a virtual photon, responsible for the Coulomb potential between pion and nucleus, transforms temporarily into an electron–positron pair;

iv) finally, since the radius of the lowest Bohr orbit

$$r_\pi = \frac{1}{Z\alpha m_\pi} \tag{3.38}$$

is $m_\pi/m_e \sim 300$ times smaller than the corresponding electron Bohr radius, the pion wavefunction has significant overlap with the central nucleus, requiring a correction factor for the strong pion-nuclear interaction.

When these modifications are made, agreement is excellent over a wide range of nuclei.

## PROBLEM VI.3.1

### Relativistic Zeeman Effect

Suppose a pionic atom is placed in a uniform magnetic field, described by the vector potential

$$\vec{A} = \frac{1}{2}(\vec{B} \times \vec{r}) \ .$$

i) Neglecting the quadratic term (justify this) show that this problem can be exactly solved to yield the energy levels

$$E = E_{B=0}(1 - 2\omega_L \frac{m}{m_\pi})^{\frac{1}{2}}$$

where $\omega_L = eB/2m_\pi$ is the Larmor frequency and m is the eigenvalue of L along the direction of the magnetic field.

ii) Evaluate the nonrelativistic limit of the Klein-Gordon equation in this case and show that the effective Hamiltonian is

$$H = H_{B=0} - \frac{e}{2m_\pi}\vec{L} \cdot \vec{B}(1 - \frac{\vec{p}^2}{2m_\pi^2} + ...)$$

Thus the usual Bohr magneton $e/2m_\pi$ is reduced by relativistic effects to

$$\frac{e}{2m_\pi}(1 - \frac{\vec{p}^2}{2m_\pi^2})$$

iii) Calculate the energy shift induced by the magnetic field using perturbation theory and show that this result agrees with the exact answer to first order in $\vec{B}$.

## PROBLEM VI.3.2

### High Z Mesonic Atoms

We have seen that the energy levels of a mesonic atom are given by

$$E = m\left[1 + \frac{Z^2\alpha^2}{(n - \ell - \frac{1}{2} + \sqrt{(\ell + \frac{1}{2})^2 - Z^2\alpha^2})^2}\right]^{-\frac{1}{2}} \ .$$

i) Show that the ground state energy for any mesonic atom heavier that Z=69 is complex. Explain what this complex energy means.

ii) Mesonic atoms have been well studied at places like Los Alamos and it has been found that the ground states of atoms as heavy as lead (Z=82) or Uranium (Z=92) are quite stable. How do you reconcile this fact with the result obtained above? Be as quantitative as you can.

# CHAPTER VII

# THE DIRAC EQUATION

## VII.1 DERIVATION AND COVARIANCE

Historically the Klein–Gordon equation was written down before the Dirac equation. However, it was abandoned for a period of time due to the problems with negative energy states and the inability to construct a positive definite probability density. Dirac then developed his formalism and demonstrated the connection between negative energy states and antiparticles, at which point the Klein–Gordon equation was resurrected. It is in this spirit then that we interrupt our discussion of the Klein–Gordon equation to present the Dirac equation.

### Intuitive Derivation

Consider the Schrödinger equation which describes a spin-1/2 particle in the presence of an electromagnetic field. The wavefunction $\psi(\vec{x},t)$ is a two-component object

$$\psi(\vec{x},t) = \begin{pmatrix} \psi_1(\vec{x},t) \\ \psi_2(\vec{x},t) \end{pmatrix} \tag{1.1}$$

describing a particle with spin

$$\vec{S} = \left\langle \psi \left| \frac{1}{2}\vec{\sigma} \right| \psi \right\rangle \quad . \tag{1.2}$$

In addition to the pieces of the Hamiltonian

$$H \sim \frac{\left(\vec{p} - e\vec{A}\right)^2}{2m} + e\phi \tag{1.3}$$

expected from the spinless case, we must also append a term which accounts for the interaction of the magnetic moment with a magnetic field. Recalling that for an electron the moment is given by

$$\vec{\mu} = \frac{g_e e}{2m}\vec{S} \tag{1.4}$$

with gyromagnetic ratio $g_e = 2$ Bohr magnetons, this leads to an additional term in the Hamiltonian

$$\begin{aligned} H' &= -\vec{\mu} \cdot \vec{B} = -\frac{e}{m}\vec{S} \cdot \vec{B} \\ &= -\frac{e}{2m}\vec{\sigma} \cdot \vec{B} \end{aligned} \tag{1.5}$$

so that the Schrödinger equation becomes

$$\left(\frac{1}{2m}\left(\vec{p} - e\vec{A}\right)^2 + e\phi - \frac{e}{2m}\vec{\sigma} \cdot \vec{\nabla} \times \vec{A}\right) \psi(\vec{x},t) = i\frac{\partial}{\partial t}\psi(\vec{x},t) \quad . \tag{1.6}$$

Using the identity

$$\sigma_i\sigma_j = \delta_{ij} + i\epsilon_{ijk}\sigma_k \tag{1.7}$$

we note that

$$\begin{aligned}\vec{\sigma}\cdot\left(\vec{p}-e\vec{A}\right)\vec{\sigma}\cdot\left(\vec{p}-e\vec{A}\right) &= \left(\vec{p}-e\vec{A}\right)\cdot\left(\vec{p}-e\vec{A}\right)+i\vec{\sigma}\cdot\left(\vec{p}-e\vec{A}\right)\times\left(\vec{p}-e\vec{A}\right)\\ &= \left(\vec{p}-e\vec{A}\right)^2 - e\vec{\sigma}\cdot\vec{\nabla}\times\vec{A}\end{aligned} \tag{1.8}$$

where $\vec{p}$ denotes the operator $-i\vec{\nabla}$. Thus we may write the Schrödinger equation in the suggestive form

$$\left(\frac{1}{2m}\vec{\sigma}\cdot\left(\vec{p}-e\vec{A}\right)\vec{\sigma}\cdot\left(\vec{p}-e\vec{A}\right)+e\phi\right)\psi = i\frac{\partial\psi}{\partial t} \quad . \tag{1.9}$$

Defining the four-vector

$$\pi_\mu = i\nabla_\mu - eA_\mu \tag{1.10}$$

we can write the previously discussed wave equations as

| | | |
|---|---|---|
| Non-Relativistic | Spin 0 | $\frac{1}{2m}\vec{\pi}\cdot\vec{\pi}\psi = \pi_0\psi$ |
| Relativistic | Spin 0 | $\left(\pi_0^2 - \vec{\pi}\cdot\vec{\pi} - m^2\right)\psi = 0$ |
| Non-Relativistic | Spin $\frac{1}{2}$ | $\frac{1}{2m}\vec{\sigma}\cdot\vec{\pi}\,\vec{\sigma}\cdot\vec{\pi}\psi = \pi_0\psi$ |

and from this tabulation we might well guess that the relativistic version of the spin-1/2 equation would take the form

$$\left(\pi_0^2 - \vec{\sigma}\cdot\vec{\pi}\,\vec{\sigma}\cdot\vec{\pi} - m^2\right)\psi = 0 \quad . \tag{1.11}$$

In fact this is almost but not quite right. Rather the correct form of the Dirac equation is given by [Fe 62]

$$\left(\left(\pi_0 - \vec{\sigma}\cdot\vec{\pi}\right)\left(\pi_0 + \vec{\sigma}\cdot\vec{\pi}\right) - m^2\right)\psi = 0 \tag{1.12}$$

where here $\psi$ is a two-component spinor. [Note that if $\pi_0, \vec{\pi}$ were simply numbers and not operators we would have

$$\begin{aligned}\left(\pi_0 - \vec{\sigma}\cdot\vec{\pi}\right)\left(\pi_0 + \vec{\sigma}\cdot\vec{\pi}\right) &= \pi_0^2 - \vec{\sigma}\cdot\vec{\pi}\,\vec{\sigma}\cdot\vec{\pi}\\ &= \pi_0^2 - \vec{\pi}^2 \quad .\end{aligned} \tag{1.13}$$

However, this is not in general the case.]

Eq. 1.12, while correct, is *not* the conventional form of the Dirac equation. Instead Dirac chose to write his equation in Hamiltonian form—*i.e.*, first order in time. As we saw in the case of the Klein–Gordon equation, this requires a doubling of the number of components. We must deal with a *four* component spinor —

two components for particles, two for antiparticles. We define then a pair of *two* component objects $\rho, \chi$ such that

$$\begin{aligned}(\pi_0 + \vec{\sigma}\cdot\vec{\pi})\,\rho &= m\chi \\ (\pi_0 - \vec{\sigma}\cdot\vec{\pi})\,\chi &= m\rho\end{aligned} \tag{1.14}$$

Eq. 1.14 is an equivalent version of Eq. 1.12, but is still not the standard form of the Dirac equation. Instead the conventional version is found via use of the linear combinations

$$\psi_a = \chi + \rho \ , \qquad \psi_b = \chi - \rho \tag{1.15}$$

which satisfy

$$\begin{aligned}\pi_0\psi_a - \vec{\sigma}\cdot\vec{\pi}\psi_b &= m\psi_a \\ \vec{\sigma}\cdot\vec{\pi}\psi_a - \pi_0\psi_b &= m\psi_b \ .\end{aligned} \tag{1.16}$$

Eq. 1.16 can be represented most succinctly by employing a *four* component object

$$\psi \equiv \begin{pmatrix}\psi_a \\ \psi_b\end{pmatrix} \tag{1.17}$$

and the $4\times 4$ matrices

$$\gamma^0 = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & -1\end{array}\right) \tag{1.18a}$$

$$\vec{\gamma} = \left(\begin{array}{c|c} 0 & \vec{\sigma} \\ \hline -\vec{\sigma} & 0\end{array}\right) \tag{1.18b}$$

in terms of which we can write

$$\left(\gamma^0\pi_0 - \vec{\gamma}\cdot\vec{\pi}\right)\psi = m\psi \ . \tag{1.19}$$

[Check:

$$\begin{aligned}\left(\gamma^0\pi_0 - \vec{\gamma}\cdot\vec{\pi}\right)\psi &= \begin{pmatrix}\pi_0 & -\vec{\sigma}\cdot\vec{\pi} \\ \vec{\sigma}\cdot\vec{\pi} & -\pi_0\end{pmatrix}\begin{pmatrix}\psi_a \\ \psi_b\end{pmatrix} \\ &= m\begin{pmatrix}\psi_a \\ \psi_b\end{pmatrix} = m\psi\end{aligned} \tag{1.20}$$

which agrees with the coupled equations in Eq. 1.16.]
This is the conventional form of the Dirac equation and is usually written as

$$\left(\gamma^\mu\pi_\mu - m\right)\psi = 0 \ . \tag{1.21}$$

Often in the literature one finds this result expressed via a shorthand notation due to Feynman wherein an arbitrary four vector $A^\mu$ contracted with the Dirac matrices $\gamma^\mu$ is denoted by using a slash through the four-vector

$$A_\mu \gamma^\mu \equiv \not{A} \quad . \tag{1.22}$$

Then the Dirac equation assumes the simple form

$$(\not{\pi} - m)\,\psi = (i\,\not{\nabla} - e\,\not{A} - m)\,\psi = 0 \quad . \tag{1.23}$$

It is useful at this point to identify certain properties of the $\gamma^\mu$ matrices which we shall later exploit. We note

$$\gamma^{0\dagger} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \gamma^0 \quad , \qquad \vec{\gamma}^\dagger = \begin{pmatrix} 0 & -\vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix} = -\vec{\gamma} \quad . \tag{1.24}$$

Also

$$(\gamma^0)^2 = \begin{pmatrix} 1^2 & 0 \\ 0 & (-1)^2 \end{pmatrix} = 1 \quad , \qquad (\gamma^i)^2 = \begin{pmatrix} -\sigma_i^2 & 0 \\ 0 & -\sigma_i^2 \end{pmatrix} = -1 \quad . \tag{1.25}$$

Since any two different $\gamma$'s anticommute

$$\begin{aligned} \gamma^0\gamma^i + \gamma^i\gamma^0 &= \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix} + \begin{pmatrix} 0 & -\sigma_i \\ -\sigma_i & 0 \end{pmatrix} = 0 \\ \gamma^i\gamma^j + \gamma^j\gamma^i &= \begin{pmatrix} -\sigma_i\sigma_j - \sigma_j\sigma_i & 0 \\ 0 & -\sigma_i\sigma_j - \sigma_j\sigma_i \end{pmatrix} = 0 \ \text{ if } i \neq j \end{aligned} \tag{1.26}$$

we can represent Eqs. 1.25, 1.26 in terms of the relation

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu} \tag{1.27}$$

where

$$\eta^{\mu\nu} = \eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{1.28}$$

is the metric tensor.

### Hamiltonian Form

It is important to note that Dirac's original presentation of the relativistic equation was somewhat different than given above and was written in Hamiltonian form

$$i\frac{\partial \psi}{\partial t} = H_D \psi \quad . \tag{1.29}$$

We can reproduce this version by noting that since $(\gamma^0)^2 = 1$

$$\gamma^0 \not{\pi} = \gamma^0 \left(\gamma^0 \pi_0 - \vec{\gamma} \cdot \vec{\pi}\right) = \pi_0 - \gamma^0 \vec{\gamma} \cdot \vec{\pi} \quad . \tag{1.30}$$

Thus Eq. 1.23 becomes

$$\left(\pi_0 - \gamma^0 \vec{\gamma} \cdot \vec{\pi} - \gamma^0 m\right) \psi = 0 \quad . \tag{1.31}$$

Dirac's notation was to define

$$\beta \equiv \gamma^0 \qquad \vec{\alpha} \equiv \gamma^0 \vec{\gamma} \tag{1.32}$$

so that

$$\begin{aligned} i\frac{\partial}{\partial t}\psi &= \left(\vec{\alpha} \cdot \left(\vec{p} - e\vec{A}\right) + e\phi + \beta m\right) \psi \\ &\equiv H_D \psi \quad . \end{aligned} \tag{1.33}$$

Here

$$\beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad \vec{\alpha} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & \vec{\sigma} \\ -\vec{\sigma} & 0 \end{pmatrix} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix} \tag{1.34}$$

so that $\beta^\dagger = \beta$ and $\vec{\alpha}^\dagger = \vec{\alpha}$— Dirac's Hamiltonian $H_D$ is explicitly Hermitian.

## Covariance

The crucial issue, of course, is the covariance of the equation — does it have an identical form in all Lorentz frames? That is, if in one frame the Dirac equation is written

$$\left(\gamma_\mu \left(i\nabla^\mu - eA^\mu(x)\right) - m\right) \psi(x) = 0 \tag{1.35}$$

does it in some other frame read

$$\left(\gamma_\mu \left(i\nabla^{\mu'} - eA^{\mu'}(x')\right) - m\right) \psi'(x') = 0 \tag{1.36}$$

where

$$x'^\mu = a^\mu{}_\nu x^\nu \tag{1.37}$$

is the point into which $x$ transforms and, since $A^\mu$, $\nabla^\mu$ are also four vectors

$$A'^\mu = a^\mu{}_\nu A^\nu \quad , \qquad \nabla'^\mu = a^\mu{}_\nu \nabla^\nu \quad . \tag{1.38}$$

Although the Dirac matrices $\gamma^\mu$ are written with Greek indices, they are *not* four vectors. Rather, they have the same value in *every* frame. On the other hand, the Dirac spinor $\psi$ does change under a Lorentz transformation

$$\psi(x) \longrightarrow \psi'(x') = S(a)\psi(x) \tag{1.39}$$

where $S$ represents an as yet underdetermined matrix function of $a^\mu{}_\nu$. One should not be surprised that the spinor undergoes such a change. Indeed even in the case of a non-relativistic two-component spinor, there exists such an effect. For example, under a rotation by angle $\delta\phi$ about an axis specified by $\hat{n}$, we have

$$\begin{aligned}\psi(x) = \begin{pmatrix}\psi_1(x)\\ \psi_2(x)\end{pmatrix} \longrightarrow \psi'(x') &= \exp\left(i\frac{\delta\phi}{2}\vec{\sigma}\cdot\hat{n}\right)\begin{pmatrix}\psi_1(x)\\ \psi_2(x)\end{pmatrix}\\ &= \exp\left[i\delta\phi\hat{n}\cdot\left(\vec{L}+\frac{\vec{\sigma}}{2}\right)\right]\begin{pmatrix}\psi_1(x')\\ \psi_2(x')\end{pmatrix}\end{aligned} \quad . \tag{1.40}$$

where $\vec{L} = \vec{r}\times\vec{p}$ is the orbital angular momentum operator.

Similarly in the relativistic case we seek an operator $S(a)$ such that $\psi'(x') = S(a)\psi(x)$ and

$$(\gamma_\mu i\nabla'^\mu - m)\,S(a)\psi(x) = (\gamma_\mu i a^\mu{}_\nu\nabla^\nu - m)\,S(a)\psi(x) = 0 \tag{1.41}$$

If we multiply on the left by $S^{-1}(a)$, yielding

$$\left(S^{-1}(a)\gamma_\mu S(a)a^\mu{}_\nu i\nabla^\nu - m\right)\psi(x) = 0 \quad . \tag{1.42}$$

then in order to reproduce the original form of the Dirac equation, it is required that

$$S^{-1}(a)\gamma_\mu S(a)a^\mu{}_\nu = \gamma_\nu \quad . \tag{1.43}$$

Noting that the covariant component of a four-vector must transform as

$$x_{\nu'} = x_\lambda (a^{-1})^\lambda{}_\nu \tag{1.44}$$

in order that

$$x'_\nu x'^\nu = x_\lambda (a^{-1})^\lambda{}_\nu \times a^\nu{}_\epsilon x^\epsilon = x_\lambda \delta^\lambda{}_\epsilon x^\epsilon = x_\lambda x^\lambda \quad , \tag{1.45}$$

we see that Eq. 1.43 can be written in the alternate form

$$\gamma_\nu (a^{-1})^\nu{}_\lambda = S^{-1}(a)\gamma_\mu S(a)\, a^\mu{}_\nu (a^{-1})^\nu{}_\lambda = S^{-1}(a)\gamma_\lambda S(a) \tag{1.46}$$

or in terms of its contravariant version

$$S^{-1}(a)\gamma^\lambda S(a) = a^\lambda{}_\nu \gamma^\nu \quad . \tag{1.47}$$

Rather than present a detailed derivation, we shall merely quote the appropriate forms for $S(a)$:

i) Rotations:

Consider a rotation by angle $\phi$ about the $z$-axis, with

$$\begin{aligned} x'^1 &= x^1 \cos\phi + x^2 \sin\phi \\ x'^2 &= -x^1 \sin\phi + x^2 \cos\phi \\ x'^3 &= x^3 \\ x'^0 &= x^0 \quad . \end{aligned} \tag{1.48}$$

Then

$$\begin{aligned} S &= \exp\left(-\frac{\phi}{2}\gamma^1\gamma^2\right) = \cos\frac{\phi}{2} - \gamma^1\gamma^2 \sin\frac{\phi}{2} \\ S^{-1} &= \exp\left(\frac{\phi}{2}\gamma^1\gamma^2\right) = \cos\frac{\phi}{2} + \gamma^1\gamma^2 \sin\frac{\phi}{2} \end{aligned} \tag{1.49}$$

Check:

$$\begin{aligned} S^{-1}\gamma^0 S &= \gamma^0 S^{-1} S = \gamma^0 \qquad \text{since } [\gamma^1\gamma^2, \gamma^0] = 0 \\ S^{-1}\gamma^3 S &= \gamma^3 S^{-1} S = \gamma^3 \qquad \text{since } [\gamma^1\gamma^2, \gamma^3] = 0 \quad . \end{aligned} \tag{1.50}$$

On the other hand, since $\{\gamma^1\gamma^2, \gamma^i\} = 0 \quad i = 1, 2$

$$\begin{aligned} S^{-1}\gamma^1 S &= \gamma^1 (S)^2 = \gamma^1 \left(\cos^2\frac{\phi}{2} - \sin^2\frac{\phi}{2} - 2\sin\frac{\phi}{2}\cos\frac{\phi}{2}\gamma^1\gamma^2\right) \\ &= \cos\phi\gamma^1 + \sin\phi\gamma^2 \end{aligned} \tag{1.51a}$$

$$\begin{aligned} S^{-1}\gamma^2 S &= \gamma^2 (S)^2 = \gamma^2 \left(\cos^2\frac{\phi}{2} - \sin^2\frac{\phi}{2} - 2\sin\frac{\phi}{2}\cos\frac{\phi}{2}\gamma^1\gamma^2\right) \\ &= \cos\phi\gamma^2 - \sin\phi\gamma^1 \quad . \end{aligned} \tag{1.51b}$$

Note also that

$$(\gamma^1\gamma^2)^\dagger = \gamma^{2\dagger}\gamma^{1\dagger} = -\gamma^1\gamma^2 \tag{1.52}$$

so that

$$S^\dagger = S^{-1} \quad . \tag{1.53}$$

Since under rotations $d^3x' = d^3x$ and

$$\psi'^\dagger \psi' = \psi^\dagger S^{-1} S \psi = \psi^\dagger \psi \tag{1.54}$$

we see that the normalization $\int d^3x\, \psi^\dagger \psi$ is preserved.

ii) Lorentz Boost:

Consider a Lorentz transformation with velocity $v$ along the $z$-axis, with

$$\begin{aligned} x'^0 &= \cosh\theta x^0 - \sinh\theta x^3 \\ x'^3 &= \cosh\theta x^3 - \sinh\theta x^0 \\ x'^2 &= x^2 \\ x'^1 &= x^1 \end{aligned} \tag{1.55}$$

where we have defined

$$\cosh\theta = \frac{1}{\sqrt{1-v^2}} \quad \text{and} \quad \sinh\theta = \frac{v}{\sqrt{1-v^2}} \quad . \tag{1.56}$$

[Note

$$\cosh^2\theta - \sinh^2\theta = \frac{1}{1-v^2} - \frac{v^2}{1-v^2} = 1 \quad .] \tag{1.57}$$

Then

$$\begin{aligned} S &= \exp\frac{\theta}{2}\gamma^3\gamma^0 = \cosh\frac{\theta}{2} + \gamma^3\gamma^0 \sinh\frac{\theta}{2} \\ S^{-1} &= \exp -\frac{\theta}{2}\gamma^3\gamma^0 = \cosh\frac{\theta}{2} - \gamma^3\gamma^0\sinh\frac{\theta}{2} \quad . \end{aligned} \tag{1.58}$$

Check: Since $[\gamma^3\gamma^0, \gamma^1] = 0 \quad i = 1, 2$

$$\begin{aligned} S^{-1}\gamma^1 S &= \gamma^1 S^{-1} S = \gamma^1 \\ S^{-1}\gamma^2 S &= \gamma^2 S^{-1} S = \gamma^2 \quad . \end{aligned} \tag{1.59}$$

Since $\{\gamma^3\gamma^0, \gamma^i\} = 0 \quad i = 0, 3$

$$\begin{aligned} S^{-1}\gamma^0 S = \gamma^0 (S)^2 &= \gamma^0\left(\cosh^2\frac{\theta}{2} + \sinh^2\frac{\theta}{2} + 2\cosh\frac{\theta}{2}\sinh\frac{\theta}{2}\gamma^3\gamma^0\right) \\ &= \gamma^0\cosh\theta - \gamma^3\sinh\theta \end{aligned} \tag{1.60a}$$

$$\begin{aligned} S^{-1}\gamma^3 S = \gamma^3 (S)^2 &= \gamma^3\left(\cosh^2\frac{\theta}{2} + \sinh^2\frac{\theta}{2} + 2\cosh\frac{\theta}{2}\sinh\frac{\theta}{2}\gamma^3\gamma^0\right) \\ &= \gamma^3\cosh\theta - \gamma^0\sinh\theta \end{aligned} \tag{1.60b}$$

Note also that

$$(\gamma^3\gamma^0)^\dagger = \gamma^{0\dagger}\gamma^{3\dagger} = -\gamma^0\gamma^3 = \gamma^3\gamma^0 \tag{1.61}$$

so that

$$S^\dagger = S \quad . \tag{1.62}$$

Then

$$\psi'^\dagger\psi' = \psi^\dagger S^\dagger S\psi = \psi^\dagger (S)^2\psi \neq \psi^\dagger\psi \quad . \tag{1.63}$$

so that the normalization is *not* preserved. However, we should *expect* a change to occur since because of Lorentz contraction $d^3x' \neq d^3x$. Rather we should have

$$\psi'^\dagger\psi' d^3x' = \psi^\dagger\psi\, d^3x \tag{1.64}$$

which does not require unitarity of $S$.

We have shown then via i) and ii) that for arbitrary rotations and boosts it is possible to construct an operator $S$ such that the Dirac equation *is* covariant. For later use we note that covariance can also be verified under a rather different kind of transformation:

iii) Spatial Inversion: (*i.e.*, Parity Transformation) with

$$\begin{aligned} x'^0 &= x^0 \\ \vec{x}' &= -\vec{x} \end{aligned} \quad . \tag{1.65}$$

Then

$$S = S^{-1} = S^\dagger = \gamma^0 \quad . \tag{1.66}$$

Check:

$$\begin{aligned} S^{-1}\gamma^0 S &= \gamma^0 S^{-1} S = \gamma^0 \\ S^{-1}\vec{\gamma} S &= -\vec{\gamma} S^{-1} S = -\vec{\gamma} \quad . \end{aligned} \tag{1.67}$$

Of course, $\psi'^\dagger \psi' = \psi^\dagger S^\dagger S \psi = \psi^\dagger \psi$ so that the normalization is preserved.

**Conserved Current Density**

There exists one additional requirement on $S$ — that we be able to construct a properly conserved probability current density—$j^\mu = (\rho, \vec{j})$ —satisfying

$$\nabla^\mu j_\mu = \frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{j} = 0 \quad . \tag{1.68}$$

in all frames. For the probability density we expect

$$\rho = \psi^\dagger \psi \tag{1.69}$$

which is properly non-negative. However, does there exist a corresponding $\vec{j}$? In order to answer this question we note that

$$\begin{aligned} \psi^\dagger \left( i\frac{\partial \psi}{\partial t} + \left( i\vec{\alpha} \cdot \vec{\nabla} - \beta m \right) \psi \right) &= 0 \\ \left( -i\frac{\partial \psi^\dagger}{\partial t} + \psi^\dagger \left( -i\vec{\alpha} \cdot \overleftarrow{\nabla} - \beta m \right) \right) \psi &= 0 \quad . \end{aligned} \tag{1.70}$$

Subtracting, we find

$$\begin{aligned} i\left( \psi^\dagger \frac{\partial \psi}{\partial t} + \frac{\partial \psi^\dagger}{\partial t} \psi \right) + i\psi^\dagger \left( \vec{\alpha} \cdot \vec{\nabla} + \vec{\alpha} \cdot \overleftarrow{\nabla} \right) \psi &= 0 \\ = i\left( \frac{\partial}{\partial t} \psi^\dagger \psi + \vec{\nabla} \cdot \psi^\dagger \vec{\alpha} \psi \right) \quad , & \end{aligned} \tag{1.71}$$

whereby we identify

$$\vec{j} = \psi^\dagger \vec{\alpha} \psi \quad . \tag{1.72}$$

In order that the conservation equation

$$\nabla_\mu j^\mu = 0 \tag{1.73}$$

be Lorentz invariant and thus valid in an arbitrary frame, it is necessary that the four-current density

$$j^\mu = (\psi^\dagger\psi,\ \psi^\dagger\vec{\alpha}\psi) \tag{1.74}$$

transform as a four-vector. It is conventional to express the current density not in terms of $\psi^\dagger$ but rather in terms of

$$\bar{\psi} \equiv \psi^\dagger\gamma^0 \quad . \tag{1.75}$$

Then

$$j^\mu = \bar{\psi}\gamma^\mu\psi \quad . \tag{1.76}$$

which in a new frame becomes

$$\begin{aligned} j'^\mu &= \psi^\dagger S^\dagger\gamma^0\gamma^\mu S\psi = \psi^\dagger(\gamma^0)^2 S^\dagger\gamma^0\gamma^\mu S\psi \\ &= \bar{\psi}\gamma^0 S^\dagger\gamma^0\gamma^\mu S\psi \quad . \end{aligned} \tag{1.77}$$

Since we already know that

$$S^{-1}\gamma^\mu S = \gamma'^\mu \tag{1.78}$$

the current density $\bar{\psi}\gamma^\mu\psi$ will be a four-vector provided

$$\gamma^0 S^\dagger\gamma^0 = S^{-1} \quad . \tag{1.79}$$

This requirement is easily verified:

i) Rotations: $[\gamma^0, \gamma^i\gamma^j] = 0 \quad i,j = 1,2,3$
Then

$$\gamma^0 S^\dagger\gamma^0 = S^\dagger = S^{-1} \quad . \tag{1.80}$$

ii) Boosts: $\{\gamma^0, \gamma^0\gamma^i\} = 0 \quad i = 1,2,3$
Then

$$\gamma^0 S^\dagger\gamma^0 = \gamma^0 S\gamma^0 = S^{-1} \quad . \tag{1.81}$$

iii) Spatial Inversion: Then

$$\gamma^0 S^\dagger\gamma^0 = S^{-1} \quad . \tag{1.82}$$

We have in general

$$j^\mu = \bar{\psi}\gamma^\mu\psi \longrightarrow \bar{\psi}\gamma'^\mu\psi = a^\mu{}_\nu j^\nu = j'^\mu, \tag{1.83}$$

so that, as required,

$$0 = \nabla^\mu j_\mu \longrightarrow \nabla^{\mu\prime} j'_\mu = 0 \tag{1.84}$$

*i.e.*, current conservation obtains in all frames. Eq. 1.84 guarentees that the normalization is preserved in time. Thus

$$\frac{d}{dt}\int d^3x\,\rho = -\int_{\mathrm{Vol}} d^3x\,\nabla\cdot\vec{j} = -\int_{\mathrm{Surf}} \vec{j}\cdot d\vec{S} \tag{1.85}$$

where we have used Gauss' theorem. For a localized wavefunction the current density $\vec{j}$ vanishes on a sufficiently large surface and we find

$$\frac{d}{dt}\int d^3x\,\rho = 0 \qquad \text{q.e.d.} \tag{1.86}$$

Since

$$\rho = \psi^\dagger\psi \tag{1.87}$$

is non-negative we shall for the moment be able to interpret $\psi$ in terms of a single particle wavefunction with $\rho$ as the probability density. Later on we shall find that at least some of the same problems which plagued the Klein–Gordon equation occur in the Dirac case. The ultimate solution to these difficulties is quantization of the Dirac field. For now, however, we proceed with the single particle wavefunction interpretation.

## VII.2 BILINEAR FORMS

We have seen that although the $\gamma^\mu$ are constant matrices which have the same value in *all* frames, the bilinear quantity

$$j^\mu = \bar{\psi}\gamma^\mu\psi \tag{2.1}$$

transforms as a four-vector—under a Lorentz transformation

$$x^\mu \to x'^\mu = a^\mu{}_\nu x^\nu \tag{2.2}$$

we find

$$j^\mu \longrightarrow j'^\mu = a^\mu{}_\nu j^\nu \tag{2.3}$$

### Completeness

Since $\bar{\psi}$, $\psi$ are simply four-component row, column vectors, a bilinear

$$\bar{\psi}O\psi \tag{2.4}$$

can be always decomposed into a combination of $4 \times 4 = 16$ linearly independent matrices. Defining

$$\gamma_5 = -i\gamma^0\gamma^1\gamma^2\gamma^3 = -i\begin{pmatrix} 0 & \sigma_1 \\ \sigma_1 & 0 \end{pmatrix}\begin{pmatrix} -i\sigma_1 & 0 \\ 0 & -i\sigma_1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \tag{2.5}$$

and the antisymmetric tensor

$$\sigma^{\mu\nu} = \frac{i}{2}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu) \tag{2.6}$$

with

$$\begin{gathered} \sigma^{00} = \sigma^{ii} = 0 \\ \sigma^{0i} = i\begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix} \qquad \sigma^{ij} = \epsilon_{ijk}\begin{pmatrix} \sigma_k & 0 \\ 0 & \sigma_k \end{pmatrix} \end{gathered} \tag{2.7}$$

we may choose these sixteen matrices to be

$$1 \qquad \gamma^\mu \qquad \sigma^{\mu\nu} \qquad \gamma^\mu\gamma_5 \qquad \gamma_5 \tag{2.8}$$

There exist four $\gamma^\mu$'s, four $\gamma^\mu\gamma_5$'s, one unit operator and one $\gamma_5$ matrix. The $4 \times 4$ matrices $\sigma^{\mu\nu}$ are antisymmetric in $\mu, \nu$. Since an antisymmetric $4 \times 4$ matrix must have the structure

$$A^{\mu\nu} \sim \begin{pmatrix} 0 & c_1 & c_2 & c_3 \\ -c_1 & 0 & c_4 & c_5 \\ -c_2 & -c_4 & 0 & c_6 \\ -c_3 & -c_5 & -c_6 & 0 \end{pmatrix} \tag{2.9}$$

we see that there exist only six *independent* elements. Thus the total number of linearly independent matrices is found to be

$$\begin{array}{ccccccccccc} 1 & & \gamma^\mu & & \sigma^{\mu\nu} & & \gamma^\mu\gamma_5 & & \gamma_5 & & \\ 1 & + & 4 & + & 6 & + & 4 & + & 1 & = & 16 \end{array}$$

as required.

## Transformation Properties

All Dirac matrices are simply constants and have the same value in all Lorentz frames. However, when contracted with $\bar{\psi}, \psi$ different bilinears have their own distinct transformation properties and we study each in turn:

$$1: \qquad \bar{\psi}\psi \longrightarrow (\psi^\dagger S^\dagger)\,\gamma^0 S\psi = \psi^\dagger\gamma^0(\gamma^0 S^\dagger\gamma^0)S\psi = \bar{\psi}S^{-1}S\psi = \bar{\psi}\psi \ . \tag{2.10}$$

Thus $\bar{\psi}\psi$ is a Lorentz scalar, transforming into itself under boosts, rotations and spatial inversions.

Before looking at $\gamma_5$ we note that since the matrix is a product of the four Dirac matrices, it must correspondingly anticommute with any of these—

$$\{\gamma_5, \gamma^\mu\} = 0 \tag{2.11}$$

—so that

$$\gamma_5: \qquad \bar{\psi}\gamma_5\psi \longrightarrow \psi^\dagger S^\dagger\gamma^0\gamma_5 S\psi = \psi^\dagger\gamma^0\left(\gamma^0 S^\dagger\gamma^0\right)\gamma_5 S\psi = \bar{\psi}S^{-1}\gamma_5 S\psi \ . \tag{2.12}$$

Since $\gamma_5$ anticommutes with any single $\gamma^\mu$, it must commute with a product

$$[\gamma_5, \gamma^\mu\gamma^\nu] = 0 \ . \tag{2.13}$$

Thus for rotations and/or boosts

$$\bar{\psi}\gamma_5\psi \underset{\substack{\text{Rot}\\ \text{Boost}}}{\longrightarrow} \bar{\psi}\gamma_5 S^{-1}S\psi = \bar{\psi}\gamma_5\psi \tag{2.14}$$

but under spatial inversion

$$\bar{\psi}\gamma_5\psi \underset{P}{\longrightarrow} -\bar{\psi}\gamma_5 S^{-1}S\psi = -\bar{\psi}\gamma_5\psi \quad . \tag{2.15}$$

The bilinear $\bar{\psi}\gamma_5\psi$ then transforms as a pseudoscalar.

Classical physics examples of scalar and pseudoscalar quantities are charge (scalar) and "magnetic charge" (pseudoscalar)—if it exists! In order to see this, note that an ordinary four-vector such as $x^\mu = (t, \vec{x})$ has the behavior under a spatial inversion

$$x^\mu = (t, \vec{x}) \underset{P}{\longrightarrow} (t, -\vec{x}) = x_\mu \tag{2.16}$$

Since $A^\mu$ itself is a four-vector then

$$A^\mu = (\phi, \vec{A}) \underset{P}{\longrightarrow} (\phi, -\vec{A}) = A_\mu \quad . \tag{2.17}$$

This means, however, that

$$\begin{aligned}\vec{E} &= -\vec{\nabla}\phi - \frac{\partial \vec{A}}{\partial t} \underset{P}{\longrightarrow} -(-\vec{\nabla})\phi - \frac{\partial}{\partial t}(-\vec{A}) = -\vec{E} \\ \vec{B} &= \vec{\nabla} \times \vec{A} \underset{P}{\longrightarrow} (-\vec{\nabla}) \times (-\vec{A}) = \vec{B}\end{aligned} \tag{2.18}$$

so that $\vec{E}$ and $\vec{B}$ behave oppositely under spatial inversion. Consider the electric field from a point charge. Since

$$\vec{E} = \frac{q}{4\pi r^2}\hat{r} \underset{P}{\longrightarrow} \frac{q_P}{4\pi r^2}(-\hat{r}) = -\vec{E} \tag{2.19}$$

we see that $q_P$ (the electric charge in the inverted frame) must be the same as $q$—$q$ is a scalar. However, if a magnetic monopole were to exist, with magnetic charge $g$

$$\vec{B} = \frac{g}{r^2}\hat{r} \underset{P}{\longrightarrow} \frac{g_P}{r^2}(-\hat{r}) = +\vec{B} \tag{2.20}$$

which requires that

$$g = -g_P \tag{2.21}$$

*i.e.*, magnetic charge is a pseudoscalar quantity.

Now consider the transformation properties of $\bar{\psi}\gamma^\mu\psi, \bar{\psi}\gamma^\mu\gamma_5\psi$.

$$\begin{aligned}\gamma^\mu: \qquad \bar{\psi}\gamma^\mu\psi \underset{\substack{\text{rot}\\ \text{boost}}}{\longrightarrow} \psi^\dagger S^\dagger\gamma^0\gamma^\mu S\psi &= \psi^\dagger\gamma^0\left(\gamma^0 S^\dagger\gamma^0\right)\gamma^\mu S\psi \\ &= \bar{\psi}S^{-1}\gamma^\mu S\psi = a^\mu{}_\nu\bar{\psi}\gamma^\nu\psi = \bar{\psi}\gamma'^\mu\psi\end{aligned} \tag{2.22}$$

so that $\bar{\psi}\gamma^\mu\psi$ transforms as a Lorentz four-vector. Also, under a spatial inversion we note that

$$\bar{\psi}\gamma^\mu\psi \underset{P}{\longrightarrow} \bar{\psi}\gamma^0\gamma^\mu\gamma^0\psi = \bar{\psi}\gamma_\mu\psi \quad . \tag{2.23}$$

Thus $\bar{\psi}\gamma^\mu\psi$ transforms under parity in the same manner as $x^\mu = (t, \vec{x})$ and is a four-vector or polar vector.

If, however, we consider $\bar{\psi}\gamma^\mu\gamma_5\psi$

$$\begin{aligned} \gamma^\mu\gamma_5: \qquad \bar{\psi}\gamma^\mu\gamma_5\psi \underset{\substack{\text{rot}\\ \text{boost}}}{\longrightarrow} \psi^\dagger S^\dagger\gamma^0\gamma^\mu\gamma_5 S\psi &= \psi^\dagger\gamma^0\left(\gamma^0 S^\dagger\gamma^0\right)\gamma^\mu\gamma_5 S\psi \\ &= \bar{\psi}S^{-1}\gamma^\mu S\gamma_5\psi = \bar{\psi}\gamma'^\mu\gamma_5\psi \end{aligned} \tag{2.24}$$

but

$$\bar{\psi}\gamma^\mu\gamma_5\psi \underset{P}{\longrightarrow} \bar{\psi}\gamma^0\gamma^\mu\gamma_5\gamma^0\psi = -\bar{\psi}\gamma_\mu\gamma_5\psi \quad . \tag{2.25}$$

Thus under ordinary Lorentz transformations (rotations/boosts) this quantity transforms like a four-vector. However, under spatial inversion an extra minus sign arises. Such a quantity is termed a "pseudo-vector" or axial vector.

There exist many examples of polar vectors in classical physics, such as

$$\begin{aligned} \vec{r} &\underset{P}{\longrightarrow} -\vec{r} \\ \vec{v} = \frac{d\vec{r}}{dt} &\underset{P}{\longrightarrow} -\frac{d\vec{r}}{dt} = -\vec{v} \\ \vec{a} = \frac{d^2\vec{r}}{dt^2} &\underset{P}{\longrightarrow} -\frac{d^2\vec{r}}{dt^2} = -\vec{a} \quad . \end{aligned} \tag{2.26}$$

Perhaps the most familiar example of an axial vector is angular momentum

$$\vec{L} = \vec{r} \times m\vec{v} \underset{P}{\longrightarrow} (-\vec{r}) \times (-m\vec{v}) = +\vec{L} \quad . \tag{2.27}$$

Similarly one requires that the spin be an axial vector

$$S \underset{P}{\longrightarrow} +\vec{S} \tag{2.28}$$

in order that the total angular momentum

$$\vec{J} = \vec{L} + \vec{S} \tag{2.29}$$

have the property of transforming into itself under spatial inversion.

Finally, for the bilinear $\bar{\psi}\sigma^{\mu\nu}\psi$

$$\begin{aligned} \sigma_{\mu\nu}: \qquad \bar{\psi}\sigma^{\mu\nu}\psi = \bar{\psi}\frac{i}{2}\left(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu\right)\psi &\underset{\substack{\text{rot}\\ \text{boost}}}{\longrightarrow} \psi^\dagger S^\dagger\gamma^0\frac{i}{2}\left(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu\right)S\psi \\ &= \psi^\dagger\gamma^0\gamma^0 S^\dagger\gamma^0\frac{i}{2}\left(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu\right)S\psi = \bar{\psi}S^{-1}\frac{i}{2}\left(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu\right)S\psi \\ &= \bar{\psi}\frac{i}{2}\left(S^{-1}\gamma^\mu S\, S^{-1}\gamma^\nu S - S^{-1}\gamma^\nu S\, S^{-1}\gamma^\mu S\right)\psi \\ &= \bar{\psi}\frac{i}{2}a^\mu{}_\lambda a^\nu{}_\sigma\left(\gamma^\lambda\gamma^\sigma - \gamma^\sigma\gamma^\lambda\right)\psi = \bar{\psi}\sigma'^{\mu\nu}\psi \quad . \end{aligned} \tag{2.30}$$

Under spatial inversion

$$\bar{\psi}\sigma^{\mu\nu}\psi \underset{P}{\longrightarrow} \bar{\psi}\gamma^0\sigma^{\mu\nu}\gamma^0\psi = \bar{\psi}\sigma_{\mu\nu}\psi \quad . \tag{2.31}$$

so that $\bar{\psi}\sigma^{\mu\nu}\psi$ transforms as a an antisymmetric second rank Lorentz tensor, an example of which in the sector of classical physics is the electromagnetic field tensor $F^{\mu\nu}$.

We have already seen the utility of one of these bilinear forms—the four-vector $\bar{\psi}\gamma^\mu\psi$ has been identified as the conserved probability current density. Other uses will arise in subsequent discussions.

## VII.3 NONRELATIVISTIC REDUCTION

It is helpful to construct the effective Schrödinger equation which applies in the non-relativistic limit of the Dirac equation, as the resulting form must contain various familiar structures.

### Effective Schrödinger Equation

We seek a (positive energy) stationary state solution

$$\psi(\vec{x},t) = \psi(\vec{x})\exp(-iEt) \tag{3.1}$$

with

$$E = m + W \quad \text{and} \quad W << m \quad . \tag{3.2}$$

Then for the two-component coupled equations involving $\psi_a, \psi_b$ we find

$$\begin{aligned} (m + W - e\phi)\,\psi_a - \vec{\sigma}\cdot\vec{\pi}\psi_b &= m\psi_a \\ \vec{\sigma}\cdot\vec{\pi}\psi_a - (m + W - e\phi)\,\psi_b &= m\psi_b \quad . \end{aligned} \tag{3.3}$$

We may solve the second of these equations for $\psi_b$ in terms of $\psi_a$

$$\psi_b = \frac{1}{2m + W - e\phi}\vec{\sigma}\cdot\vec{\pi}\psi_a \quad . \tag{3.4}$$

If $W, e\phi$ are both much smaller than $m$, then since $\pi \sim mv$ we have

$$\psi_b \sim v\psi_a \ll \psi_a \quad \text{for} \quad \frac{v}{c} \ll 1 \quad . \tag{3.5}$$

For this reason $\psi_a$ ($\psi_b$) is often called the large (small) component of the Dirac equation. Substitution of Eq. 3.4 into Eq. 3.3a gives a relation for $\psi_a$ alone

$$\begin{aligned} &\left(\vec{\sigma}\cdot\vec{\pi}\frac{1}{2m + W - e\phi}\vec{\sigma}\cdot\vec{\pi} + e\phi\right)\psi_a = W\psi_a \\ &\approx \left(\frac{1}{2m}\vec{\sigma}\cdot\vec{\pi}\,\vec{\sigma}\cdot\vec{\pi} + \frac{1}{(2m)^2}\vec{\sigma}\cdot\vec{\pi}\,(e\phi - W)\,\vec{\sigma}\cdot\vec{\pi} + e\phi\right)\psi_a \quad . \end{aligned} \tag{3.6}$$

This is not quite the effective Schrödinger equation since $\psi_a$ is not normalized to unity. Instead we have

$$\begin{aligned} 1 &= \int d^3x\, \psi^\dagger(\vec{x})\psi(\vec{x}) = \int d^3x \left(\psi_a^\dagger\psi_a + \psi_b^\dagger\psi_b\right) \\ &\approx \int d^3x\, \psi_a^\dagger \left(1 + \frac{\vec{\sigma}\cdot\vec{\pi}}{2m}\frac{\vec{\sigma}\cdot\vec{\pi}}{2m}\right)\psi_a \approx \int d^3x\, \chi^\dagger\chi \end{aligned} \tag{3.7}$$

where we have defined

$$\chi \equiv \left(1 + \frac{(\vec{\sigma}\cdot\vec{\pi})^2}{8m^2}\right)\psi_a \quad . \tag{3.8}$$

Then if we add

$$\frac{1}{8m^2}\left[(\vec{\sigma}\cdot\vec{\pi})^2(W - e\phi) + (W - e\phi)(\vec{\sigma}\cdot\vec{\pi})^2\right]\psi_a \tag{3.9}$$

to each side of the Eq. 3.6 we find

$$\begin{aligned} &\left(\frac{1}{2m}(\vec{\sigma}\cdot\vec{\pi})^2 + \frac{1}{8m^2}\left((\vec{\sigma}\cdot\vec{\pi})^2(W - e\phi)\right.\right. \\ &\left.\left. -2\vec{\sigma}\cdot\vec{\pi}(W - e\phi)\vec{\sigma}\cdot\vec{\pi} + (W - e\phi)(\vec{\sigma}\cdot\vec{\pi})^2\right)\right)\psi_a \\ &\cong \left(1 + \frac{(\vec{\sigma}\cdot\vec{\pi})^2}{8m^2}\right)(W - e\phi)\left(1 + \frac{(\vec{\sigma}\cdot\vec{\pi})}{8m^2}\right)\psi_a \end{aligned} \tag{3.10}$$

or in terms of the wavefunction $\chi$

$$\begin{aligned} &\left(1 + \frac{(\vec{\sigma}\cdot\vec{\pi})^2}{8m^2}\right)(W - e\phi)\chi \cong \frac{1}{2m}(\vec{\sigma}\cdot\vec{\pi})^2\left(1 - \frac{(\vec{\sigma}\cdot\vec{\pi})^2}{8m^2}\right)\chi \\ &+ \frac{1}{8m^2}\left((\vec{\sigma}\cdot\vec{\pi})^2(W - e\phi) - 2\vec{\sigma}\cdot\vec{\pi}(W - e\phi)\vec{\sigma}\cdot\vec{\pi} + (W - e\phi)(\vec{\sigma}\cdot\vec{\pi})^2\right)\chi \quad . \end{aligned} \tag{3.11}$$

We can rewrite Eq. 3.11 by use of the identity

$$\begin{aligned} \hat{A}^2\hat{B} - 2\hat{A}\hat{B}\hat{A} + \hat{B}\hat{A}^2 &= \hat{A}(\hat{A}\hat{B} - \hat{B}\hat{A}) - (\hat{A}\hat{B} - \hat{B}\hat{A})\hat{A} \\ &= \left[\hat{A}, [\hat{A}, \hat{B}]\right] \end{aligned} \tag{3.12}$$

as

$$\begin{aligned} (W - e\phi)\chi &\cong \frac{1}{2m}(\vec{\sigma}\cdot\vec{\pi})^2\left(1 - \frac{(\vec{\sigma}\cdot\vec{\pi})^2}{4m^2}\right)\chi \\ &\quad + \frac{1}{8m^2}\left[(\vec{\sigma}\cdot\vec{\pi}), [(\vec{\sigma}\cdot\vec{\pi}), W - e\phi]\right]\chi \quad . \end{aligned} \tag{3.13}$$

Since, assuming $\vec{A}(x)$ to be independent of $t$,

$$\begin{aligned} [\vec{\sigma}\cdot\vec{\pi}, [\vec{\sigma}\cdot\vec{\pi}, W - e\phi]] &= \left[\vec{\sigma}\cdot\vec{\pi}, ie\vec{\sigma}\cdot\vec{\nabla}\phi\right] = -ie\left[\vec{\sigma}\cdot\vec{\pi}, \vec{\sigma}\cdot\vec{E}\right] \\ &= e\left(-\vec{\nabla}\cdot\vec{E} + 2\vec{\sigma}\cdot\vec{\pi}\times\vec{E}\right) \end{aligned} \tag{3.14}$$

we have, finally

$$\begin{aligned} W\chi \cong &\left(\frac{1}{2m}(\vec{p} - e\vec{A})^2 - \frac{1}{8m^3}(\vec{p} - e\vec{A})^4\right. \\ &\left. - \frac{e}{2(m + W)}\vec{\sigma}\cdot\vec{B} + e\phi - \frac{e}{8m^2}\left(\vec{\nabla}\cdot\vec{E} - 2\vec{\sigma}\cdot\vec{\pi}\times\vec{E}\right)\right)\chi \end{aligned} \tag{3.15}$$

which is the effective Schrödinger equation we seek.

### Interpretation

Interpretation of the various terms can now be undertaken. Thus in the spinless case, we evaluated the relativistic Hamiltonian to be used in the presence of an electromagnetic field described by the vector potential $A_\mu = (\phi, \vec{A})$, yielding

$$\begin{aligned} H &= \sqrt{m^2 + (\vec{p} - e\vec{A})^2} + e\phi \\ &\approx m + \frac{1}{2m}(\vec{p} - e\vec{A})^2 - \frac{1}{8m^3}(p - e\vec{A})^4 + e\phi \dots \quad . \end{aligned} \tag{3.16}$$

in the non-relativistic limit. Likewise in Eq. 3.15 we recognize

$$\frac{1}{2m}(\vec{p} - e\vec{A})^2 + e\phi \tag{3.17}$$

as the usual Schrödinger Hamiltonian and

$$-\frac{1}{8m^3}\left(\vec{p} - e\vec{A}\right)^4 \tag{3.18}$$

as a relativistic correction to the kinetic energy. Similarly we identify the term

$$-\frac{e}{2(m+W)}\vec{\sigma} \cdot \vec{B} \approx -\frac{e}{2m}\vec{\sigma} \cdot \vec{B} \tag{3.19}$$

as the energy of interaction of the magnetic moment of a spin-1/2 particle (with gyromagnetic ratio g=2) with an external magnetic field.† As discussed above, this arises automatically if we use

$$\frac{1}{2m}\vec{\sigma} \cdot \pi \ \vec{\sigma} \cdot \vec{\pi} = \frac{1}{2m}(\vec{p} - e\vec{A})^2 - \frac{e}{2m}\vec{\sigma} \cdot \vec{B} \tag{3.20}$$

for the kinetic energy term in the Hamiltonian rather than the simple spinless form

$$\frac{1}{2m}\vec{\pi} \cdot \vec{\pi} = \frac{1}{2m}\left(\vec{p} - e\vec{A}\right)^2 \quad . \tag{3.21}$$

We can also identify the remaining terms in a straight-forward fashion. For example, the operator

$$-\frac{e}{4m^2}\vec{\sigma} \cdot \vec{E} \times \vec{p} \tag{3.22}$$

is simply the usual spin-orbit term since for $\phi = \phi(r)$

$$-\frac{e}{4m^2}\vec{\sigma} \cdot \vec{E} \times \vec{p} = \frac{e}{4m^2}\,\frac{1}{r}\,\frac{d\phi}{dr}\vec{\sigma} \cdot \vec{r} \times \vec{p} = \frac{e}{4m^2}\,\frac{1}{r}\,\frac{d\phi}{dr}\vec{\sigma} \cdot \vec{L} \quad . \tag{3.23}$$

The "classical" derivation of this form involves looking at the problem of the "atom" from the perspective of the electron rest frame, in which case the central nucleus is seen to be orbiting at the distance of a Bohr radius. In this frame, however, there exists a *magnetic* field due to the Lorentz transformation of the electromagnetic fields involved

$$\vec{B} = -\vec{v} \times \vec{E} \quad . \tag{3.24}$$

† Note that the relativistic Bohr magneton is $\frac{e}{2E}$.

The energy associated with the interaction of the electron magnetic moment with this induced field is

$$H = -\vec{\mu} \cdot \vec{B} = \frac{e}{2m} \vec{\sigma} \cdot \vec{v} \times \vec{E} = \frac{e}{2m^2} \frac{1}{r} \frac{d\phi}{dr} \vec{\sigma} \cdot \vec{L} \ , \tag{3.25}$$

which differs from Eq. 3.23 by a factor of two due to the "Thomas precession."

The point is [Ja 80] that in an inertial frame the energy

$$U = -\vec{\mu} \cdot \vec{B} \tag{3.26}$$

associated with the interaction of a magnetic moment $\vec{\mu}$ with a magnetic field $\vec{B}$ corresponds to an equation of motion

$$\frac{d\vec{S}}{dt} = \vec{S} \times \frac{e\vec{B}}{m} \tag{3.27}$$

where the have used the relation

$$\vec{\mu} = \frac{e}{m} \vec{S} \tag{3.28}$$

between the spin and magnetic moment. However, in a non-inertial frame of reference one has

$$\frac{d\vec{S}}{dt} = \left.\frac{d\vec{S}}{dt}\right|_{\text{non-rotating}} - \vec{\omega} \times \vec{S} \tag{3.29}$$

where $\vec{\omega}$ is the angular velocity of the rotating frame, and the corresponding interaction energy is

$$U' = -\vec{S} \cdot \vec{\omega} \ . \tag{3.30}$$

In order to find the angular velocity $\vec{\omega}$ corresponding to our case, consider the trajectory of an accelerating electron as shown in Figure VII.1.

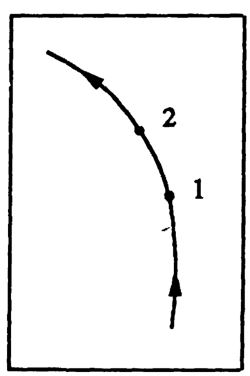



*Fig. VII.1: Trajectory of an accelerating electron.*

Suppose that at time $t$ the electron is at position #1 with velocity $\vec{v}$ while at time $t + \delta t$ the electron is located at position #2 with velocity $\vec{v} + \delta\vec{v}$. Let the laboratory frame be denoted by $W$, the rest frames of the electron at positions #1 and #2 by $W_1$ and $W_2$, respectively. Then one can reach $W_1$ from $W$ by a simple Lorentz transformation with velocity $\vec{v}$

$$W \xrightarrow[\vec{v}]{} W_1 \tag{3.31}$$

while one can reach $W_2$ from $W$ by a Lorentz transformation with velocity $\vec{v} + \delta\vec{v}$

$$W \xrightarrow[\vec{v}+\delta\vec{v}]{} W_2 \quad . \tag{3.32}$$

However, the transformation between the frames $W_1$ and $W_2$ is in general a combination of a boost plus a rotation. To first order in $\delta\vec{v}$ we find

$$\begin{aligned} t_2 &= t_1 - \vec{x}_1 \cdot \frac{1}{\sqrt{1-v^2}}\left(\delta\vec{v} + \left(\frac{1}{\sqrt{1-v^2}} - 1\right)\hat{v}\hat{v}\cdot\delta\vec{v}\right) \\ \vec{x}_2 &= \vec{x}_1 - t_1\frac{1}{\sqrt{1-v^2}}\left(\delta\vec{v} + \left(\frac{1}{\sqrt{1-v^2}} - 1\right)\hat{v}\hat{v}\cdot\delta\vec{v}\right) \\ &\quad + \left(\frac{1}{\sqrt{1-v^2}} - 1\right)\vec{x}_1 \times (\vec{v}\times\delta\vec{v})\,\frac{1}{v^2} \end{aligned} \tag{3.33}$$

which corresponds to a Lorentz transformation with velocity

$$\Delta\vec{v} = \frac{1}{\sqrt{1-v^2}}\left(\delta\vec{v} + \left(\frac{1}{\sqrt{1-v^2}} - 1\right)\hat{v}\hat{v}\cdot\delta\vec{v}\right) \tag{3.34}$$

accompanied by a rotation through angle

$$\delta\vec{\theta} = \left(\frac{1}{\sqrt{1-v^2}} - 1\right)\vec{v}\times\delta\vec{v}\frac{1}{v^2} \quad . \tag{3.35}$$

The corresponding angular velocity is

$$\vec{\omega} = \frac{\delta\vec{\theta}}{\delta t} = \left(\frac{1}{\sqrt{1-v^2}} - 1\right)\vec{v}\times\vec{a}\frac{1}{v^2} \approx \frac{1}{2}\vec{v}\times\vec{a} \quad , \tag{3.36}$$

leading to an additional interaction energy

$$\begin{aligned} U' &= -\vec{\omega}\cdot\vec{S} = -\frac{1}{2}\vec{v}\times\vec{a}\cdot\vec{S} \\ &= -\frac{1}{2}\vec{v}\times -\frac{e}{r}\frac{d\phi}{dr}\vec{r}\frac{1}{m}\cdot\vec{S} \\ &= -\frac{e}{4m^2}\frac{1}{r}\frac{d\phi}{dr}\vec{L}\cdot\vec{\sigma}. \end{aligned} \tag{3.37}$$

Adding Eqs. 3.25 and 3.37 we find

$$U_{\text{tot}} = \frac{e}{4m^2}\frac{1}{r}\frac{d\phi}{dr}\vec{L}\cdot\vec{\sigma} \tag{3.38}$$

as found in reduction of the Dirac equation.

The origin of the final piece of the effective interaction, the Darwin term

$$-\frac{e}{8m^2}\vec{\nabla}\cdot\vec{E}, \tag{3.39}$$

has already been noted in our discussion of the Klein–Gordon equation, where we showed that the zitterbewegung motion associated with the interference between positive and negative energy components leads to a shift in the potential energy in the amount

$$\begin{aligned}\Delta U &\simeq \frac{1}{2!}\delta r_i\,\delta r_j\frac{\partial^2}{\partial r_i\partial r_j}e\phi(\vec{r}) \sim \frac{e}{2\cdot 3m^2}\vec{\nabla}^2\phi\\ &= -\frac{e}{6m^2}\vec{\nabla}\cdot\vec{E}\ .\end{aligned} \tag{3.40}$$

Except for the factor of 1/6 rather than 1/8 this is clearly the additional term under discussion. Because of this identification the Darwin or zitterbewegung term has no classical analogy.

### Hydrogen Atom Energy Levels: Perturbative Approach

Now examine the effects of these perturbations on the energy levels of the hydrogen atom. For the zitterbewegung term we find that only $S$-waves are affected

$$\begin{aligned}(\Delta E)^{n\ell}_{\text{zitt}} &= -\frac{e}{8m^2}\left\langle\vec{\nabla}\cdot\vec{E}\right\rangle_{n\ell} = \frac{e^2}{8m^2}\left\langle\delta^3(r)\right\rangle_{n\ell}\\ &= \frac{e^2}{8m^2}\int d^3r\,\delta^3(r)\,\psi^*_{n\ell}(\vec{r})\psi_{n\ell}(\vec{r})\\ &= \frac{\pi\alpha}{2m^2}\delta_{\ell 0}\left|\psi_{n0}(0)\right|^2 = \frac{\pi\alpha}{2m^2}\delta_{\ell 0}\left(\frac{1}{\pi n^3a_0^3}\right)\\ &= m\frac{\alpha^4}{2n^3}\delta_{\ell 0}\ .\end{aligned} \tag{3.41}$$

while for the spin-orbit term $S$-waves are not altered in energy, but other angular momentum states are shifted. Recalling that for a given value of orbital angular momentum $\ell$, the total angular momentum

$$\vec{J} = \vec{L} + \vec{S} \tag{3.42}$$

can have the value $\ell + 1/2$ or $\ell - 1/2$, we find

$$\begin{aligned}(\Delta E)^{n\ell}_{so} &= \frac{\alpha}{4m^2}\left\langle r^{-3}\right\rangle_{n\ell}2\vec{S}\cdot\vec{L}\\ &= \frac{\alpha}{4m^2}\left\langle r^{-3}\right\rangle_{n\ell}\left(J^2 - L^2 - S^2\right) = \frac{\alpha}{4m^2}\left\langle r^{-3}\right\rangle_{n\ell}\begin{cases}\ell & j=\ell+\frac{1}{2}\\ -(\ell+1) & j=\ell-\frac{1}{2}\end{cases}\\ &= \frac{\alpha}{4m^2}\left(\frac{1}{n^3a_0^3\,\ell(\ell+1)(\ell+\frac{1}{2})}\right)\begin{cases}\ell & j=\ell+\frac{1}{2}\\ -(\ell+1) & j=\ell-\frac{1}{2}\end{cases}\\ &= m\frac{\alpha^4}{2n^3(2\ell+1)}\begin{cases}(\ell+1)^{-1} & j=\ell+\frac{1}{2}\\ -\ell^{-1} & j=\ell-\frac{1}{2}\end{cases}\ .\end{aligned} \tag{3.43}$$

Finally, we note that for the relativistic kinetic energy term

$$
\begin{aligned}
(\Delta E)^{n\ell}_{p^4} &= -\frac{1}{2m}\left\langle\left(\frac{p^2}{2m}\right)^2\right\rangle_{n\ell} \\
&= -\frac{1}{2m}\left\langle(E_n - V(r))^2\right\rangle_{n\ell} \\
&= -\frac{1}{2m}\left\langle E_n^2 - 2E_n V(r) + V^2(r)\right\rangle_{n\ell} \quad .
\end{aligned}
\tag{3.44}
$$

According to the virial theorem

$$
\langle V(r)\rangle_{n\ell} = 2E_n \tag{3.45}
$$

and by direct calculation

$$
\begin{aligned}
\left\langle V^2(r)\right\rangle_{n\ell} = \alpha^2\left\langle r^{-2}\right\rangle_{n\ell} &= \alpha^2\frac{1}{n^3 a_0^2(\ell+1/2)} \\
&= m^2\frac{\alpha^4}{n^3(\ell+1/2)} \quad .
\end{aligned}
\tag{3.46}
$$

Thus

$$
(\Delta E)^{n\ell}_{p^4} = -m\frac{\alpha^4}{2n^3}\left(\frac{1}{\ell+1/2} - \frac{3}{4n}\right) \quad . \tag{3.47}
$$

Our final result then, writing $E = m + W$ is

$$
W = -m\frac{\alpha^2}{2n^2}\left(1 + \frac{\alpha^2}{n}\left(\frac{1}{j+1/2} - \frac{3}{4n}\right) + \ldots\right) \quad . \tag{3.48}
$$

We observe that the energy depends only upon $j$ — it is independent of $\ell$. For example, the $2S_{1/2}$ and $2P_{1/2}$ states are degenerate, as used earlier in our discussions of the Lamb shift. Also, we see that the shift has a similar form to that found in the case of the Klein–Gordon atom, but with

$$
\begin{array}{ccc}
K-G & & \text{Dirac} \\
 & \text{replaced by} & \\
\ell + \frac{1}{2} & & j + \frac{1}{2}
\end{array} \quad . \tag{3.49}
$$

This may seem like a small difference, but it is easily measurable and agreement with experiment for the hydrogen atom *requires* the electron to be a spin-1/2 particle.

### PROBLEM VII.3.1

### The Runge-Lenz Vector and the Hydrogen Atom

One of the remarkable features of motion in a Coulomb (or gravitational) field is that there is no precession of the classical orbits. By use of the so-called Runge-Lenz vector it is possible both to understand this result and to derive the energy levels of a hydrogen atom without use of any differential equations.

i) Verify this result by showing classically that

$$\frac{d\vec{R}}{dt} = 0 \qquad \text{and} \qquad \vec{L}\cdot\vec{R} = 0$$

where

$$\vec{R} = \frac{1}{m}\vec{p}\times\vec{L} - \frac{\alpha}{r}\vec{r}$$

is the Runge–Lenz vector. Show that for gravitational motion $\vec{R}$ points along perihelion so that no precession takes place in the planetary orbits for an exact 1/r potential.

ii) If we define quantum mechanically

$$\vec{R} = \frac{1}{2m}(\vec{p}\times\vec{L} - \vec{L}\times\vec{p}) - \frac{\alpha}{r}\vec{r}$$

show that

$$[H,\vec{L}] = [H,\vec{R}] = 0 \qquad \vec{R}\cdot\vec{L} = \vec{L}\cdot\vec{R} = 0$$

iii) Verify that

$$R^2 = \alpha^2 + \frac{2}{m}H(L^2+1)$$

so that the Hamiltonian can be written in terms of two constants of the motion.

iv) Define $\vec{K} = \sqrt{\frac{-m}{2H}}\vec{R}$ and

$$\vec{M} = \frac{1}{2}(\vec{L}+\vec{K}) \qquad \vec{N} = \frac{1}{2}(\vec{L}-\vec{K}).$$

Show that

$$[M_i, M_j] = i\epsilon_{ijk}M_k$$
$$[N_i, N_j] = i\epsilon_{ijk}N_k$$
$$[M_i, N_j] = 0$$

Thus $\vec{M}$ and $\vec{N}$ obey commutation relations for angular momenta and commute with each other and also with the Hamiltonian

$$H = -\frac{m\alpha^2}{2(2M^2+2N^2+1)} \quad .$$

We can then find simultaneous eigenstates of $H, M^2, N^2, M_z, N_z$ with

$$H|E,m,n,m_z,n_z> = E|E,m,n,m_z,n_z>$$

$$M^2|E,m,n,m_z,n_z> = m(m+1)|E,m,n,m_z,n_z>$$

$$N^2|E,m,n,m_z,n_z> = n(n+1)|E,m,n,m_z,n_z>$$

$$M_z|e,m,n,m_z,n_z>= m_z|e,m,n,m_z,n_z>$$

$$N_z|E,m,n,m_z,n_z>= n_z|E,m,n,m_z,n_z> .$$

v) Show that $\vec{R}\cdot\vec{L}=0$ implies $\vec{K}\cdot\vec{L}=0$ and hence that $M^2=N^2$.

vi) Show that this in turn implies

$$E=-\frac{m\alpha^2}{2k^2}$$

where k=1,2,3,... with degeneracy factor $2n^2$, as required.

**PROBLEM VII.3.2**

**The Anomalous Magnetic Moment**

The Dirac equation describing the interaction of a proton or neutron with an applied external electromagnetic field has an additional term

$$\left(i\not\nabla - Q_i \not A + \frac{\kappa_i|e|}{4m}\sigma_{\mu\nu}F^{\mu\nu} - m\right)\psi(x)=0$$

involving the so-called anomalous magnetic moment. (For the proton, of course, $Q_i=|e|$ and for the neutron $Q_i=0$.)

i) Verify that the choice

$$\kappa_p=1.79\quad,\qquad \kappa_n=-1.91$$

corresponds to the observed magnetic moments of these particles, and

ii) show that the additional interaction disturbs neither the Lorentz covariance of the equation nor the hermiticity of the Hamiltonian.

**PROBLEM VII.3.3**

**The Aharonov–Casher Effect**

In Section III.3 we discussed the Aharonov-Bohm effect which shows that in quantum mechanics the behavior of particles can be altered by the presence of a non-zero vector potential even though the magnetic field vanishes in all regions of space accessible to these particles. More recently Aharonov and Casher [AhC 84] pointed out another interesting quantum mechanical process whereby the behavior of *magnetic* dipoles is altered by the presence of an *electric* field. In this problem we explore this effect.

i) The Dirac equation which describes an electron in the presence of an external vector potential $A_\mu$ is

$$(i\not\nabla - e\not A - m)\psi(x)=0$$

Writing the energy as $E = m + W$ and making a nonrelativistic reduction as done in the text, show that the effective Schrodinger equation which results is

$$\frac{\vec{\sigma}\cdot(\vec{p}-e\vec{A})\vec{\sigma}\cdot(\vec{p}-e\vec{A})}{2m}\psi(x) = W\psi(x)$$

ii) For the standard version of the A-B effect one uses an infinite solenoid positioned along the z-axis so that in cylindrical coordinates

$$\vec{A}(x) = \frac{BR^2}{2r}\hat{e}_\phi \qquad r > R$$

where R is the solenoidal radius. Show that for this geometry the effective Hamiltonian becomes

$$H_{\text{eff}} = \frac{1}{2m}\sum_{i=1}^{2}(p_i - eA_i)^2$$

iii) As shown in problem VII.3.2, the Dirac equation describing the interaction of a neutral spin 1/2 particle with an external electromagnetic field is given by

$$(i\nabla\!\!\!/ + \frac{\kappa|e|}{4m}\sigma_{\mu\nu}F^{\mu\nu} - m)\psi(x) = 0$$

where $\kappa$ is the magnetic moment. Now consider a beam of neutrons polarized along the z-axis interacting with a line charge with charge per unit length $\lambda$ aligned along the z-direction. Perform a nonrelativistic reduction of the above relativisitic equation and demonstrate that the effective Schrodinger equation is

$$\frac{\vec{\sigma}\cdot(\vec{p}-i\kappa'\vec{E})\vec{\sigma}\cdot(\vec{p}+i\kappa'\vec{E})}{2m}\psi(x) = W\psi(x)$$

where $\kappa' = \kappa|e|/2m$ and

$$\vec{E} = \frac{\lambda}{2\pi r}\hat{r}$$

is the electric field generated by the line charge.

iv) Show that this Hamiltonian is equivalent to the form

$$H_{\text{eff}} = \frac{(\vec{p}-\vec{E}\times\vec{\mu})^2}{2m} - \frac{\kappa'^2E^2}{2m}$$

for the geometry at hand, where $\vec{\mu} = \kappa'\vec{\sigma}$. Replacing $H_{\text{eff}} \to H_{\text{eff}}' \equiv \chi_i^\dagger H_{\text{eff}}\chi_i$ verify that

$$H_{\text{eff}}' = \frac{1}{2m}\sum_{i=1}^{2}[p_i - (\vec{E}\times\kappa'\hat{e}_z)_i]^2 \quad ,$$

which is completely equivalent to the A-B Hamiltonian provided we make the replacement $\kappa'E \to eA$.

We observe then that the Hamiltonian is the same and hence there must exist an effect on the magnetic dipoles for the A-C situation in complete analogy to that on electric charges for the A-B geometry. Recently this prediction was confirmed experimentally [Ci 89].

## VII.4 COULOMB SOLUTION

Although we have derived the $2S_{1/2} - 2P_{1/2}$ degeneracy within the context of first order perturbation theory, the result is more general and is valid to *all orders* in the fine structure constant $\alpha$, as we shall demonstrate.

### Hydrogen Atom Energy Levels: Relativistic Approach

We begin with the Dirac equation in its usual representation

$$(\gamma_\mu \pi^\mu - m)\,\psi(x) = 0 \quad \text{where} \quad \pi_\mu = i\nabla_\mu - eA_\mu \tag{4.1}$$

and introduce the projection operators

$$P_1, P_2 = \frac{1}{2}(1-\gamma_5)\,,\frac{1}{2}(1+\gamma_5) \quad . \tag{4.2}$$

[Note that since $\gamma_5^2 = 1$ we have $P_1^2 = P_2^2 = 1$, $P_1P_2 = P_2P_1 = 0$.] If we define

$$\psi_1 \equiv P_1\psi \quad , \qquad \psi_2 \equiv P_2\psi \tag{4.3}$$

then since

$$P_2\gamma_\mu = \gamma_\mu P_1 \tag{4.4}$$

we find

$$\begin{aligned} P_2\gamma_\mu\pi^\mu\psi &= \gamma_\mu\pi^\mu P_1\psi = \gamma_\mu\pi^\mu\psi_1 \\ &= mP_2\psi = m\psi_2 \end{aligned} \tag{4.5}$$

or

$$\psi_2 = \frac{1}{m}\gamma_\mu\pi^\mu\psi_1 \quad . \tag{4.6}$$

Also, since $P_1 + P_2 = 1$ we find

$$\psi = \psi_1 + \psi_2 = \left(1 + \frac{1}{m}\gamma_\mu\pi^\mu\right)\psi_1 \tag{4.7}$$

so that knowledge of $\psi_1$ is equivalent to knowledge of $\psi$ itself. From Eq. 4.7 we see that $\psi_1$ obeys the equation

$$m\,(\gamma_\mu\pi^\mu - m)\,\psi = (\gamma_\mu\pi^\mu - m)\,(\gamma_\mu\pi^\mu + m)\,\psi_1 = 0 \quad . \tag{4.8}$$

Also since

$$P_1 = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \tag{4.9}$$

a stationary state solution $\psi_1$ having definite orbital angular momentum $\ell$ must be of the form

$$\psi_1 = \begin{pmatrix} \chi_\ell(\vec{r}) \\ \chi_\ell(\vec{r}) \end{pmatrix} e^{-iEt} \tag{4.10}$$

where the $\chi_\ell$'s represent *two*-component spinors which satisfy the equation

$$\left(\frac{1}{r}\frac{\partial^2}{\partial r^2}r + 2\alpha\frac{E}{r} - \frac{\ell(\ell+1) - \alpha^2 - i\alpha\vec{\sigma}\cdot\hat{r}}{r^2} + E^2 - m^2\right)\chi_\ell(\vec{r}) = 0 \ . \tag{4.11}$$

Except for the term involving the Pauli spin matrix $\vec{\sigma}$, this differential equation is identical to that studied in the case of the Klein–Gordon atom

$$\left(\frac{1}{r}\frac{\partial^2}{\partial r^2}r - \frac{\ell'(\ell'+1)}{r^2} + \frac{2m'\alpha}{r} + k^2\right)\psi(r) = 0 \tag{4.12}$$

provided we make the substitution

$$\begin{aligned} \ell(\ell+1) - \alpha^2 &= \ell'(\ell'+1) \\ E^2 - m^2 &= k^2 = 2m'E' \\ E &= m' \end{aligned} \tag{4.13}$$

Eq. 4.11 then becomes

$$\left(\frac{1}{r}\frac{\partial^2}{\partial r^2}r - \frac{\ell'(\ell'+1)}{r^2} + \frac{2m'\alpha}{r} + k^2 + i\alpha\frac{1}{r^2}\vec{\sigma}\cdot\hat{r}\right)\chi_\ell(\vec{r}) = 0 \tag{4.14}$$

and can be diagonalized in terms of functions $\phi^{\pm}_{jm}$ which are eigenstates of $J^2$, $L^2$, $S^2$, $J_z$:

$$\phi^{\pm}_{jm} = \sum_{p,q} C^{p\,q;m}_{\ell\frac{1}{2};j=\ell\pm\frac{1}{2}} Y^p_\ell(\theta,\phi)\chi^q_{\frac{1}{2}} \ . \tag{4.15}$$

Explicitly, we find

$$\begin{aligned} \phi^+_{j,m} &= \begin{pmatrix} ((j+m)/2j)^{1/2} & Y^{m-1/2}_{j-1/2}(\theta,\phi) \\ ((j-m)/2j)^{1/2} & Y^{m+1/2}_{j-1/2}(\theta,\phi) \end{pmatrix} \\ \phi^-_{j,m} &= \begin{pmatrix} ((j+1-m)/2(j+1))^{1/2} & Y^{m-1/2}_{j+1/2}(\theta,\phi) \\ ((j+1+m)/2(j+1))^{1/2} & Y^{m+1/2}_{j+1/2}(\theta\,\phi) \end{pmatrix} \end{aligned} \tag{4.16}$$

Observe that $\phi^{\pm}_{j,m}$ have opposite parities—if $\Pi$ is the spatial inversion operator

$$\Pi\phi^{\pm}_{j,m} = (-1)^{j\mp 1/2}\phi^{\pm}_{j,m} \tag{4.17}$$

—and that they are connected via the operator $\vec{\sigma}\cdot\hat{r}$

$$\vec{\sigma}\cdot\hat{r}\phi^{\pm}_{j,m} = \phi^{\mp}_{j,m} \ . \tag{4.18}$$

This result is clear since

i) $\vec{\sigma}\cdot\hat{r}$ is odd under parity;

ii) $(\vec{\sigma}\cdot\hat{r})^2 = 1$;

iii) $\left[\vec{J}, \vec{\sigma}\cdot\hat{r}\right] = 0$;

so that eigenstates of $\vec{J}$ are also eigenstates of $\vec{\sigma}\cdot\hat{r}$. [For the latter result note that for $J_z$:

$$\begin{aligned} L_z = -i\frac{\partial}{\partial\phi}: \quad & [L_z, \sigma_x \sin\theta\cos\phi + \sigma_y \sin\theta\sin\phi + \sigma_z\cos\theta] \\ & = -i\left(-\sigma_x\sin\theta\sin\phi + \sigma_y\sin\theta\cos\phi\right) \\ S_z = \frac{\sigma_z}{2}: \quad & [S_z, \sigma_x \sin\theta\cos\phi + \sigma_y \sin\theta\sin\phi + \sigma_z\cos\theta] \\ & = i\sigma_y\sin\theta\cos\phi - i\sigma_x\sin\theta\sin\phi \quad . \end{aligned} \tag{4.19}$$

Thus (and other components may be proved similarly)

$$[J_z, \vec{\sigma}\cdot\hat{r}] = [L_z, \vec{\sigma}\cdot\hat{r}] + [S_z, \vec{\sigma}\cdot\hat{r}] = 0 \quad . \tag{4.20}$$

Now choose the linear combinations

$$F^{\pm}_{j,m} \equiv \phi^{\pm}_{j,m} \pm \frac{i}{\alpha}\left(j + \frac{1}{2} - s\right)\phi^{\mp}_{j,m} \tag{4.21}$$

where

$$s = \left(\left(j+\frac{1}{2}\right)^2 - \alpha^2\right)^{1/2} \quad . \tag{4.22}$$

Then we have

$$\begin{aligned} \left(L^2 - \alpha^2 - i\alpha\vec{\sigma}\cdot\hat{r}\right) F^{\pm}_{j,m} = & \left(L^2 - \alpha^2 \pm \left(j + \frac{1}{2} - s\right)\right)\phi^{\pm}_{j,m} \\ & + \left(\pm\frac{i}{\alpha}\left(j+\frac{1}{2} - s\right)\left(L^2-\alpha^2\right) - i\alpha\right)\phi^{\mp}_{j,m} \quad . \end{aligned} \tag{4.23}$$

But

$$\left(L^2 - \alpha^2 \pm \left(j + \frac{1}{2} - s\right)\right)\phi^{\pm}_{j,m} = u_{\pm}\left(u_{\pm}+1\right)\phi^{\pm}_{j,m} \tag{4.24}$$

with

$$u_{\pm} = s - \frac{1}{2} \mp \frac{1}{2} \quad . \tag{4.25}$$

[Check:

$$\begin{aligned} \left(L^2 - \alpha^2 \pm \left(j+\frac{1}{2}\right)\right)\phi^{\pm}_{j,m} & = \begin{cases} \left(\left(j-\frac{1}{2}\right)\left(j+\frac{1}{2}\right) - \alpha^2 + j + \frac{1}{2}\right)\phi^{+}_{j,m} \\ \left(\left(j+\frac{1}{2}\right)\left(j+\frac{3}{2}\right) - \alpha^2 - j - \frac{1}{2}\right)\phi^{+}_{j,m} \end{cases} \\ & = \left(\left(j+\frac{1}{2}\right)^2 - \alpha^2\right)\phi^{\pm}_{j,m} = s^2\phi^{\pm}_{j,m} \end{aligned} \tag{4.26}$$

$$s^2 - s = s(s-1) = u_+(u_+ + 1)$$
$$s^2 + s = s(s+1) = u_-(u_- + 1) \qquad \text{q.e.d.]}$$

Also

$$\begin{aligned}
&\left(\pm\frac{i}{\alpha}\left(j+\frac{1}{2}-s\right)\left(L^2-\alpha^2\right)-i\alpha\right)\phi^{\mp}_{j,m} \\
&= \pm\frac{i}{\alpha}\left(j+\frac{1}{2}-s\right)\left[L^2-\alpha^2 \mp \frac{\alpha^2}{j+\frac{1}{2}-s}\right]\phi^{\pm}_{j,m} \\
&= \pm\frac{i}{\alpha}\left(j+\frac{1}{2}-s\right)\left[L^2-\alpha^2 \mp \left(j+\frac{1}{2}+s\right)\right]\phi^{\mp}_{j,m} \\
&= \pm\frac{i}{\alpha}\left(j+\frac{1}{2}-s\right)u_\pm\left(u_\pm+1\right)\phi^{\mp}_{j,m} \ .
\end{aligned} \tag{4.27}$$

Thus

$$\left(L^2-\alpha^2-i\alpha\vec{\sigma}\cdot\hat{r}\right)F^{\pm}_{j,m} = u_\pm\left(u_\pm+1\right)F^{\pm}_{j,m} \tag{4.28}$$

and Eq. 4.14 becomes

$$\left(\frac{1}{r}\frac{\partial^2}{\partial r^2}r+\frac{2m'\alpha}{r}+k^2-\frac{u_\pm(u_\pm+1)}{r^2}\right)F^{\pm}_{j,m}=0 \tag{4.29}$$

which is completely identical to the Klein–Gordon case provided we make the substitution

$$\ell'(\ell'+1) \longrightarrow u_\pm(u_\pm+1) \ . \tag{4.30}$$

The associated energy levels are then given by (*cf.* Eq. VI.3.34)

$$\begin{aligned}
E_{n,j} &= m\left(1+\frac{\alpha^2}{\left(n+s-\left(j+\frac{1}{2}\right)\right)^2}\right)^{-1/2} \qquad n=1,2,\ldots \\
&\approx m\left(1-\frac{1}{2}\frac{\alpha^2}{n^2}\left(1+\frac{\alpha^2}{n^2}\left(\frac{n}{j+\frac{1}{2}}-\frac{3}{2}\right)+\ldots\right)\right) \ .
\end{aligned} \tag{4.31}$$

which is identical to that found in the Klein-Gordon case, but with

$$\begin{matrix} K-G & & \text{Dirac} \\ & \text{replaced by} & \\ \ell+\frac{1}{2} & & j+\frac{1}{2} \end{matrix} \tag{4.32}$$

in agreement with experiment for the hydrogen atom. Also, in the Dirac case there exists a twofold degeneracy in all levels with the same value of $j$ and $n$, so that *e.g.*, as claimed earlier, the $2P_{1/2}$ and $2S_{1/2}$ levels are degenerate to all orders in $\alpha$.

## VII.5 PLANE WAVE SOLUTIONS

It is particularly useful to examine plane wave solutions of the Dirac equation, corresponding to a freely moving particle, since it is with such wavefunctions that one can develop an intuitive feel for the physics.

### Derivation

We begin by writing

$$\begin{aligned}\psi(x) &= u(p)\,e^{-iEt+i\vec{p}\cdot\vec{x}} \\ &= u(p)\,e^{-ip\cdot x}\end{aligned} \tag{5.1}$$

where $u(p)$ is a four-component spinor. Since

$$i\nabla^\mu\,e^{-ip\cdot x} = p^\mu\,e^{-ip\cdot x} \tag{5.2}$$

the free particle Dirac equation

$$(i\gamma_\mu\nabla^\mu - m)\,\psi(x) = 0 \tag{5.3}$$

becomes an algebraic relation

$$(\gamma_\mu p^\mu - m)u(p) = (\not{p} - m)u(p) = 0 \tag{5.4}$$

which is equivalent to four linear homogeneous equations. In order for a solution to exist, we must require $\det(\not{p} - m)$ to vanish, and one could proceed to solve the system formally in this fashion. However, we shall utilize an alternative, more intuitive, approach, writing

$$u(p) = \begin{pmatrix} u_a \\ u_b \end{pmatrix} \tag{5.5}$$

where $u_a$, $u_b$ are both two-component spinors. Then the validity of the Dirac equation requires

$$\left(\gamma^0 E - \vec{\gamma}\cdot\vec{p}\right)u(p) = \begin{pmatrix} E & -\vec{\sigma}\cdot\vec{p} \\ \vec{\sigma}\cdot\vec{p} & -E \end{pmatrix}\begin{pmatrix} u_a \\ u_b \end{pmatrix} = m\begin{pmatrix} u_a \\ u_b \end{pmatrix} \tag{5.6}$$

which yields the pair of coupled equations

$$\begin{aligned} Eu_a - \vec{\sigma}\cdot\vec{p}u_b &= mu_a \\ -Eu_b + \vec{\sigma}\cdot\vec{p}u_a &= mu_b \quad . \end{aligned} \tag{5.7}$$

There exist two relations between the spinors $u_a, u_b$ which must *both* be satisfied

$$u_a = \frac{\vec{\sigma}\cdot\vec{p}}{E-m}u_b \quad , \qquad u_b = \frac{\vec{\sigma}\cdot\vec{p}}{E+m}u_a \quad . \tag{5.8}$$

We thus require

$$u_b = \frac{\vec{\sigma}\cdot\vec{p}}{E+m}u_a = \frac{\vec{\sigma}\cdot\vec{p}\,\vec{\sigma}\cdot\vec{p}}{E^2-m^2}u_b \quad . \tag{5.9}$$

Using the identity

$$\vec{\sigma}\cdot\vec{p}\,\vec{\sigma}\cdot\vec{p} = \vec{p}^2 \tag{5.10}$$

we see that Eq. 5.9 is satisfied provided that

$$E^2 - m^2 = \vec{p}^2 \tag{5.11}$$

which is the desired relation between the relativistic energy $E$ and momentum $\vec{p}$. Of course, for a given value of the momentum $\vec{p}$ there exist *two* solutions for the energy

$$E = \pm\sqrt{\vec{p}^2 + m^2} \tag{5.12}$$

as found in the case of the Klein–Gordon equation.

Consider first the positive energy solutions. We define

$$u_a = N\chi \tag{5.13}$$

where $\chi$ is a two-component spinor and $N$ is a normalization constant. The lower component $u_b$ is then

$$u_b = N\frac{\vec{\sigma}\cdot\vec{p}}{E+m}\chi \quad . \tag{5.14}$$

and our solution takes the form

$$u(p) = N\begin{pmatrix} \chi \\ \frac{\vec{\sigma}\cdot\vec{p}}{E+m}\chi \end{pmatrix} \quad . \tag{5.15}$$

There exist two linearly independent spinors $u(p)$ corresponding to the two linearly independent values for $\chi$ — call these $\chi_1$ and $\chi_2$ (*e.g.* we could take $\chi_1 = \binom{1}{0}$ and $\chi_2 = \binom{0}{1}$) with $\chi_1^\dagger\chi_2 = \chi_2^\dagger\chi_1 = 0$ and $\chi_1^\dagger\chi_1 = \chi_2^\dagger\chi_2 = 1$.

In order to normalize the Dirac spinor, one's first thought might be to place a normalization condition upon $u^\dagger(p)u(p)$. However, this is a non-relativistic way of thinking and would not be Lorentz invariant. Defining

$$\bar{u} \equiv u^\dagger\gamma^0 \tag{5.16}$$

we have already shown that $\bar{u}u$ is a Lorentz scalar quantity. Thus we may set

$$\bar{u}(p)u(p) = 1 \quad . \tag{5.17}$$

as our normalization condition and this will hold in all frames. Since

$$u(p) = N\begin{pmatrix} \chi \\ \frac{\vec{\sigma}\cdot\vec{p}}{E+m}\chi \end{pmatrix} \quad , \qquad \bar{u}(p) = N^*\left(\chi^\dagger, -\chi^\dagger\frac{\vec{\sigma}\cdot\vec{p}}{E+m}\right) \tag{5.18}$$

Eq. 5.17 becomes

$$\begin{aligned} 1 &= |N|^2\left(1 - \frac{\vec{\sigma}\cdot\vec{p}\,\vec{\sigma}\cdot\vec{p}}{(E+m)^2}\right) = |N|^2\left(1 - \frac{\vec{p}^2}{(E+m)^2}\right) \\ &= |N|^2\left(1 - \frac{E^2-m^2}{(E+m)^2}\right) = |N|^2\left(1 - \frac{E-m}{E+m}\right) = |N|^2\frac{2m}{E+m} \quad . \end{aligned} \tag{5.19}$$

Thus we choose

$$N = \sqrt{\frac{E+m}{2m}} \tag{5.20}$$

so

$$u(p) = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} \chi \\ \frac{\vec{\sigma}\cdot\vec{p}}{E+m}\chi \end{pmatrix} \quad . \tag{5.21}$$

We have then *two* linearly independent positive energy solutions $u_1(p), u_2(p)$ with

$$\bar{u}_1(p)u_1(p) = \bar{u}_2(p)u_2(p) = 1 \tag{5.22a}$$

and

$$\bar{u}_1(\vec{p})u_2(\vec{p}) = \bar{u}_2(\vec{p})u_1(\vec{p}) = 0 \quad . \tag{5.22b}$$

Sometimes one picks the corresponding Pauli spinors $\chi_1$, $\chi_2$ by the requirement that

$$\vec{\sigma}\cdot\hat{p}\chi_1 = \chi_1 \quad , \qquad \vec{\sigma}\cdot\hat{p}\chi_2 = -\chi_2 \quad . \tag{5.23}$$

Then the Dirac spinor constructed using $\chi_1$ is said to be in a positive "helicity" state, while that constructed from $\chi_2$ is said to have negative "helicity." Positive, negative helicity corresponds to the spin being parallel, antiparallel to the direction of momentum.

Now consider the two remaining linearly independent negative energy solutions. We shall construct these spinors for the case that $\vec{p}$ is the negative of what it was above. That is, we define

$$p'^{\mu} = (-E, -\vec{p}) = -p^{\mu} \quad , \tag{5.24}$$

and assume a solution of the form

$$\psi'(x) = v(p)\, e^{-ip'\cdot x} = v(p)\, e^{ip\cdot x} \quad . \tag{5.25}$$

As before, the Dirac equation becomes an algebraic equation

$$(\gamma_\mu p^\mu + m)v(p) = 0 \quad . \tag{5.26}$$

Writing

$$v = \begin{pmatrix} v_a \\ v_b \end{pmatrix} \quad \text{we have} \quad \begin{array}{c} -Ev_a + \vec{\sigma}\cdot\vec{p}v_b = mv_a \\ -\vec{\sigma}\cdot\vec{p}v_a + Ev_b = mv_b \end{array} \tag{5.27}$$

which yields

$$v_a = \frac{\vec{\sigma}\cdot\vec{p}}{E+m}v_b \quad , \qquad v_b = \frac{\vec{\sigma}\cdot\vec{p}}{E-m}v_a \quad . \tag{5.28}$$

Picking

$$v_b = N\chi' \quad . \tag{5.29}$$

we find

$$v_a = N\frac{\vec{\sigma}\cdot\vec{p}}{E+m}\chi' \tag{5.30}$$

and

$$v(p) = N \begin{pmatrix} \frac{\vec{\sigma}\cdot\vec{p}}{E+m}\chi' \\ \chi' \end{pmatrix} \quad . \tag{5.31}$$

Thus

$$\begin{aligned} \bar{v}(p)v(p) &= |N|^2 \left(\chi'^{\dagger}\frac{\vec{\sigma}\cdot\vec{p}}{E+m}, -\chi'^{\dagger}\right) \begin{pmatrix} \frac{\vec{\sigma}\cdot\vec{p}}{E+m}\chi' \\ \chi' \end{pmatrix} \\ &= -|N|^2\chi'^{\dagger}\left(-\frac{\vec{\sigma}\cdot\vec{p}\,\vec{\sigma}\cdot\vec{p}}{(E+m)^2}+1\right)\chi' = -|N|^2\frac{2m}{E+m} \end{aligned} \tag{5.32}$$

so that if we normalize to

$$\bar{v}(p)v(p) = -1 \tag{5.33}$$

we have, as before

$$N = \sqrt{\frac{E+m}{2m}} \quad . \tag{5.34}$$

Note that the positive and negative energy solutions are orthogonal in that

$$\bar{u}(p)v(p) = |N|^2 \left(\chi^{\dagger}, -\chi^{\dagger}\frac{\vec{\sigma}\cdot\vec{p}}{E+m}\right) \begin{pmatrix} \frac{\vec{\sigma}\cdot\vec{p}}{E+m}\chi' \\ \chi' \end{pmatrix} = 0 \quad . \tag{5.35}$$

We can summarize these results concisely as

$$\begin{aligned} \not{p}\, u(p) &= mu(p) \qquad & \not{p}\, v(p) &= -mv(p) \\ \bar{u}(p)u(p) &= 1 \qquad & \bar{v}(p)v(p) &= -1 \end{aligned} \tag{5.36}$$

with the orthogonality condition

$$\bar{v}(p)u(p) = \bar{u}(p)v(p) = 0 \quad . \tag{5.37}$$

One can also write the Dirac equation in its conjugate form . Thus

$$\gamma_\mu p^\mu u(p) = mu(p) \qquad \text{implies} \qquad u^\dagger(p)\gamma_\mu^\dagger p^\mu = mu^\dagger(p) \quad . \tag{5.38}$$

Since

$$\gamma_\mu^\dagger = \gamma^\mu \tag{5.39}$$

and

$$\gamma_\mu^\dagger\gamma^0 = \gamma^\mu\gamma^0 = \gamma^0\gamma_\mu \tag{5.40}$$

we can write Eq. 5.38 as

$$\bar{u}(p)\,\not{p} = m\bar{u}(p) \qquad \text{and also} \qquad \bar{v}(p)\,\not{p} = -mv(p) \quad . \tag{5.41}$$

## Boosts

We can understand the form of these plane-wave spinors in an alternative fashion by using the boost operator derived earlier. Considering a positive energy electron at rest, we have the spinor

$$u(\vec{p}=0)=\begin{pmatrix}\chi\\0\end{pmatrix}\quad . \tag{5.42}$$

If we view this state from a frame moving with velocity

$$\vec{v}=-v\hat{k} \tag{5.43}$$

so that

$$t'=\frac{t+vz}{\sqrt{1-v^2}}\quad,\quad z'=\frac{z+vt}{\sqrt{1-v^2}}\quad,\quad y'=y\quad,\quad x'=x \tag{5.44}$$

the corresponding Lorentz transformation matrix is

$$S(a)=\exp-\frac{\theta}{2}\gamma^3\gamma^0\quad\text{with}\quad\cosh\theta=\frac{E}{m}=\frac{1}{\sqrt{1-v^2}}\quad,\quad\sinh\theta=\frac{p}{m}=\frac{v}{\sqrt{1-v^2}} \tag{5.45}$$

and the Dirac spinor as viewed in this frame becomes

$$u(p)=S(a)u(\vec{p}=0)=\left(\cosh\frac{\theta}{2}-\gamma^3\gamma^0\sinh\frac{\theta}{2}\right)\begin{pmatrix}\chi\\0\end{pmatrix}=\begin{pmatrix}\cosh\frac{\theta}{2}\chi\\ \sinh\frac{\theta}{2}\sigma_3\chi\end{pmatrix}\quad . \tag{5.46}$$

Since

$$\begin{aligned}\cosh^2\frac{\theta}{2}&=\frac{1}{2}\left(1+\cosh\theta\right)=\frac{E+m}{2m}\\ \sinh^2\frac{\theta}{2}&=\frac{1}{2}\left(\cosh\theta-1\right)=\frac{E-m}{2m}\end{aligned} \tag{5.47}$$

we have

$$u(p)=\sqrt{\frac{E+m}{2m}}\begin{pmatrix}\chi\\ \sqrt{\frac{E-m}{E+m}}\vec{\sigma}\cdot\hat{p}\chi\end{pmatrix}=\sqrt{\frac{E+m}{2m}}\begin{pmatrix}\chi\\ \frac{\vec{\sigma}\cdot\vec{p}}{E+m}\chi\end{pmatrix} \tag{5.48}$$

in agreement with Eq. 5.21. Similarly for a negative energy spinor, we find

$$v(\vec{p})=S(a)v(\vec{p}=0)=S(a)\begin{pmatrix}0\\ \chi\end{pmatrix}=\sqrt{\frac{E+m}{2m}}\begin{pmatrix}\frac{\vec{\sigma}\cdot\vec{p}}{E+m}\chi\\ \chi\end{pmatrix} \tag{5.49}$$

which agrees with Eq. 5.31.

## Rotations

It is also useful to examine the free Dirac spinors from the point of view of rotations in order to get a feel for the physics involved. This is most easily displayed in the case of an *infinitesimal* rotation by angle $\delta\phi$ about say the $z$-axis

$$\vec{x}' = \vec{x} + \delta\vec{x} \qquad \delta\vec{x} = (-x_2\delta\phi, x_1\delta\phi, 0) = -\vec{x} \times \delta\vec{\phi} \tag{5.50}$$

For a spinless particle, *e.g.* the Klein–Gordon equation, it is well-known that the angular momentum operator $\vec{L}$ is the "generator" of rotations in that after rotation we have†

$$\phi'(x) = \exp\left(i\delta\vec{\phi} \cdot \vec{L}\right) \phi(x) \approx \left(1 + i\delta\vec{\phi} \cdot \vec{L}\right) \phi(x) \ . \tag{5.51}$$

On the other hand for the Dirac case, we expect rotations to be generated by the *total* angular momentum $\vec{J} = \vec{L} + \frac{1}{2}\Sigma$

$$\begin{aligned}\psi'(x) &= \exp\left(i\delta\vec{\phi} \cdot \vec{J}\right) \psi(x) \\ &\cong \left(1 + i\delta\vec{\phi} \cdot \left(\vec{L} + \frac{1}{2}\vec{\Sigma}\right)\right) \psi(x) \ ,\end{aligned} \tag{5.52}$$

where

$$\frac{1}{2}\vec{\Sigma} = \begin{pmatrix} \frac{1}{2}\vec{\sigma} & 0 \\ 0 & \frac{1}{2}\vec{\sigma} \end{pmatrix} \tag{5.53}$$

is the spin operator. We can verify this conjecture since we know the form of the Dirac solution under rotations

$$\psi'(x') = S(\delta\vec{\phi})\psi(x) \tag{5.54}$$

where

$$S(\delta\vec{\phi}) = \exp\left(-\frac{\delta\phi}{2}\gamma^1\gamma^2\right) \cong 1 - \gamma^1\gamma^2\frac{\delta\phi}{2} \ . \tag{5.55}$$

Note that

$$\gamma^1\gamma^2 = \begin{pmatrix} 0 & \sigma_1 \\ -\sigma_1 & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0 \end{pmatrix} = \begin{pmatrix} -i\sigma_3 & 0 \\ 0 & -i\sigma_3 \end{pmatrix} \ . \tag{5.56}$$

Then

$$\gamma^i\gamma^j \equiv -i\epsilon_{ijk}\Sigma_k \tag{5.57}$$

† Usually Eq. 5.51 is written in the equivalent form

$$\phi'(x') = \phi(x)$$

where x is the point which rotates into x'.

and

$$\psi'(x') = \left(1 - x_2\delta\phi\frac{\partial}{\partial x_1} + x_1\delta\phi\frac{\partial}{\partial x_2} + i\frac{1}{2}\Sigma_3\delta\phi\right)\psi(x')$$
$$= \left(1 + i\delta\phi\left(x_1p_2 - x_2p_1 + \frac{1}{2}\Sigma_3\right)\right)\psi(x') \tag{5.58}$$
$$= \left(1 + i\delta\vec{\phi}\cdot\left(\vec{L} + \frac{1}{2}\vec{\Sigma}\right)\right)\psi(x')$$

as expected.

We also verify that although

$$\begin{aligned}\left[H, \vec{L}\right] &= \left[\vec{\alpha}\cdot\left(\vec{p} - e\vec{A}\right) + e\phi + \beta m, \vec{r}\times\vec{p}\right] \\ &= -i\vec{\alpha}\times\vec{p} + e\left[\phi, \vec{L}\right] - e[\vec{\alpha}\cdot\vec{A}, \vec{L}] \neq 0\end{aligned} \tag{5.59}$$

and

$$\begin{aligned}\left[H, \frac{1}{2}\vec{\Sigma}\right] &= \left[\vec{\alpha}\cdot\left(\vec{p} - e\vec{A}\right) + e\phi + \beta m, \frac{1}{2}\vec{\Sigma}\right] \\ &= i\vec{\alpha}\times\left(\vec{p} - e\vec{A}\right) \neq 0\end{aligned} \tag{5.60}$$

if we consider a particle moving under the influence of a spherically symmetric potential $\phi(r)$ so that

$$\vec{A} = 0 \ , \qquad \left[\vec{L}, \phi(r)\right] = 0 \tag{5.61}$$

we find

$$\left[H, \vec{J}\right] = \left[H, \vec{L} + \frac{1}{2}\vec{\Sigma}\right] = 0 \ . \tag{5.62}$$

—the total angular momentum is a constant of the motion, as expected.

## Helicity

It is particularly interesting to examine the time development of the helicity $\vec{\Sigma}\cdot\left(\vec{p} - e\vec{A}\right)$, where $\vec{p} - e\vec{A}$ is the so-called "mechanical" momentum $m\dot{\vec{x}}$ (as opposed to the canonical momentum $\vec{p}$). Since

$$\begin{aligned}\left[H, \vec{p} - e\vec{A}\right] &= \left[\vec{\alpha}\cdot\left(\vec{p} - e\vec{A}\right) + e\phi + \beta m, \vec{p} - e\vec{A}\right] \\ &= ie\vec{\nabla}\phi - ie\vec{\alpha}\times(\vec{\nabla}\times\vec{A})\end{aligned} \tag{5.63}$$

we have

$$\begin{aligned}\left[H, \vec{\Sigma}\cdot(\vec{p} - e\vec{A})\right] &= \left[H, \vec{\Sigma}\right]\cdot(\vec{p} - e\vec{A}) + \vec{\Sigma}\cdot\left[H, \vec{p} - e\vec{A}\right] \\ &= ie\vec{\Sigma}\cdot\vec{\nabla}\phi \ .\end{aligned} \tag{5.64}$$

For a free particle — $\phi = 0$, $\vec{A} = 0$ — we find

$$\left[H, \vec{\Sigma}\cdot\vec{p}\right] = 0 \tag{5.65}$$

so that helicity is a constant of the motion. However, this result is more general. Imagine an electron moving in a region where $\vec{E} = 0$ (take $\vec{\nabla}\phi = \partial\vec{A}/\partial t = 0$) but $\vec{B} \neq 0$. It follows then from Eq. 5.64 that

$$\left[H, \vec{\Sigma} \cdot (\vec{p} - e\vec{A})\right] = 0 \tag{5.66}$$

which means that the helicity will be also be unchanged in this circumstance. In particular imagine a longitudinally polarized electron moving in a uniform magnetic field $\vec{B}$. Then the trajectory of the electron will be a circle as shown in Figure VII.2, and the spin vector will exactly track the mechanical momentum.

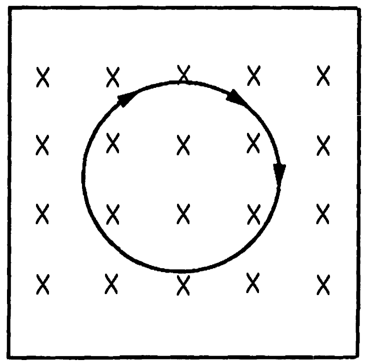


*Fig. VII.2: The electron trajectory in the presence of a uniform magnetic field is a circle of radius $mv/eB$.*

The reason for this is clear. According to classical physics the rotation frequency of the electron — the Larmor frequency $\omega_L$ — is found to be

$$\omega_L = \frac{v}{r} = \frac{eB}{m} \qquad (i.e.\ ,\ \frac{mv^2}{r} = evB)\ . \tag{5.67}$$

On the other hand, the spin precession frequency $\omega_S$ is determined from the torque equation

$$\frac{d\vec{S}}{dt} = \vec{S} \times \frac{g_e e\vec{B}}{2m} \tag{5.68}$$

which yields

$$\omega_S = \frac{g_e eB}{2m}\ . \tag{5.69}$$

We see that the Larmor and spin precession frequencies are identical (for $g_e$=2), and this is the origin of the constancy of the helicity.

### Current Density

Finally, consider the current density

$$j^\mu = \bar{\psi}\gamma^\mu\psi\ . \tag{5.70}$$

Between identical plane wave states—$u(p)\,e^{-ip\cdot x}$—$j^\mu$ assumes the form

$$\begin{aligned} j^\mu &= \bar{u}(p)\gamma^\mu u(p) = \frac{1}{2}\bar{u}(p)\left(\frac{1}{m}\,\not{p}\,\gamma^\mu + \gamma^\mu\,\not{p}\,\frac{1}{m}\right)u(p) \\ &= (\text{using the identity } \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu}) \quad \bar{u}(p)\frac{p^\mu}{m}u(p) \\ &= \frac{p^\mu}{m} \quad . \end{aligned} \tag{5.71}$$

This proportionality to $p^\mu/m$ is easy to understand. Since the probability density is $j^0(x)$ the probability of finding this particle in a volume $d^3x$ in the rest frame is

$$j^0(x)d^3x = \frac{m}{m}d^3x = d^3x \quad . \tag{5.72}$$

As viewed from a frame moving with velocity

$$\vec{v} = -v\hat{k} \tag{5.73}$$

the probability density becomes

$$j'^0(x') = \frac{E}{m} = \gamma \qquad \text{where} \quad \gamma = \frac{1}{\sqrt{1-v^2}} \quad . \tag{5.74}$$

However, because of the Lorentz contraction, the region of space being examined is correspondingly smaller

$$d^3x' = \frac{1}{\gamma}d^3x \quad . \tag{5.75}$$

so that the probability of being found within this volume is unchanged, as required.

$$j'^0(x')d^3x' = \frac{E}{m}d^3x' = \gamma \times \frac{1}{\gamma}d^3x = d^3x = j^0(x)d^3x \quad . \tag{5.76}$$

The corresponding three-vector component of the current density is also found as expected via

$$\vec{j}'(x') = j'^0(x')\vec{v} = \frac{E}{m}\cdot\frac{\vec{p}}{E} = \frac{\vec{p}}{m} \quad . \tag{5.77}$$

Considering the so-called transition current density, taken between different solutions of the Dirac equation $\psi_i, \psi_f$, we have

$$j^\mu(x) = \bar{\psi}_f(x)\gamma^\mu\psi_i(x) = \frac{1}{2m}\left(\bar{\psi}_f(x)\gamma^\mu i\gamma^\nu\partial_\nu\psi_i(x) - (i\partial_\nu\bar{\psi}_f(x))\gamma^\nu\gamma^\mu\psi_i(x)\right) \quad . \tag{5.78}$$

Now write

$$i\gamma^\mu\gamma^\nu = \frac{i}{2}\{\gamma^\mu,\gamma^\nu\} + \frac{i}{2}[\gamma^\mu,\gamma^\nu] = i\eta^{\mu\nu} + \sigma^{\mu\nu} \quad . \tag{5.79}$$

so that

$$\begin{aligned} j^\mu(x) &= \frac{i}{2m}\left(\bar{\psi}_f(x)\partial^\mu\psi_i(x) - \partial^\mu\bar{\psi}_f(x)\psi_i(x)\right) \\ &+ \frac{1}{2m}\left(\bar{\psi}_f(x)\sigma^{\mu\nu}\partial_\nu\psi_i(x) + \partial_\nu\bar{\psi}_f(x)\sigma^{\mu\nu}\psi_i(x)\right) \equiv j^{(1)\mu} + j^{(2)\mu} \quad . \end{aligned} \tag{5.80}$$

We observe that there exists a more or less natural separation of the current density into two components:

i) For the first piece — $j^{(1)\mu}$ — since for positive energy solutions the lower component of the Dirac spinor is $\mathcal{O}(v/c)$ compared to the upper component, we have

$$\begin{aligned}\bar{\psi}\psi &\equiv \left(\psi^\dagger_{\text{upper}}\psi^\dagger_{\text{lower}}\right)\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}\psi_{\text{upper}}\\ \psi_{\text{lower}}\end{pmatrix}\\ &= \psi^\dagger_{\text{upper}}\psi_{\text{upper}}\left(1+\mathcal{O}\left(\frac{v^2}{c^2}\right)\right) \approx \psi^\dagger\psi \quad .\end{aligned} \tag{5.81}$$

We can write this contribution to the transition density as

$$j^{(1)\mu} \simeq \frac{i}{2m}\left(\psi^\dagger\partial_\mu\psi - \partial_\mu\psi^\dagger\psi\right) \tag{5.82}$$

which is identical to the usual Schrödinger form of the current density.

ii) For the second piece — $j^{(2)\mu}$ — we note that its contribution to a Lagrange density would be of the form

$$\mathcal{L}_{\text{int}} = -ej^\mu A_\mu \sim -\frac{e}{2m}A_\mu\partial_\nu(\bar{\psi}_f\sigma^{\mu\nu}\psi_i) \quad . \tag{5.83}$$

Writing Eq. 5.83 as

$$\mathcal{L}_{\text{int}} = -\frac{e}{2m}\left(\partial_\nu\left(A_\mu\bar{\psi}_f\sigma^{\mu\nu}\psi_i\right) - \partial_\nu A_\mu\bar{\psi}_f\sigma^{\mu\nu}\psi_i\right) \quad . \tag{5.84}$$

we see that the first piece may be discarded, as it is a total derivative and contributes only a constant to the Lagrangian. For the second term we may use the antisymmetry of $\sigma^{\mu\nu}$ to write

$$\begin{aligned}\mathcal{L}_{\text{int}} &= \frac{e}{2m}\,\frac{1}{2}\left(\partial_\nu A_\mu - \partial_\mu A_\nu\right)\bar{\psi}_f\sigma^{\mu\nu}\psi_i\\ &= -\frac{e}{4m}F_{\mu\nu}\bar{\psi}_f\sigma^{\mu\nu}\psi_i \quad .\end{aligned} \tag{5.85}$$

Looking at the non-relativistic limit, we have

$$\begin{aligned}\bar{\psi}_f\sigma^{ij}\psi_i &\sim \epsilon_{ijk}\psi^\dagger_{\text{upper}}\sigma_k\psi_{\text{upper}}\\ \bar{\psi}_f\sigma^{0i}\psi_i &\sim \mathcal{O}\left(\frac{v}{c}\right)\psi^\dagger_{\text{upper}}\sigma_i\psi_{\text{upper}} \quad .\end{aligned} \tag{5.86}$$

Hence, keeping only the piece involving $\sigma_{ij}$ and noting that

$$F_{ij} = -\epsilon_{ijk}B_k \tag{5.87}$$

we find

$$\begin{aligned}\mathcal{L}_{\text{int}} &= \frac{e}{4m}\epsilon_{ijk}B_k\epsilon_{ij\ell}\psi^\dagger_{\text{upper}}\sigma_\ell\psi_{\text{upper}}\\ &= \frac{e}{2m}\vec{B}\cdot\psi^\dagger_{\text{upper}}\vec{\sigma}\psi_{\text{upper}} \quad .\end{aligned} \tag{5.88}$$

The corresponding Hamiltonian density is

$$\mathcal{H}_{\text{int}} = -\mathcal{L}_{\text{int}} = -\frac{e}{2m}\vec{B} \cdot \psi^\dagger_{\text{upper}}\vec{\sigma}\psi_{\text{upper}} \tag{5.89}$$

which is the usual energy of interaction of the electron magnetic moment

$$\vec{\mu} = \frac{e}{2m}\psi^\dagger\vec{\sigma}\psi \tag{5.90}$$

with an external magnetic field. For these reasons $j_\mu^{(1)}$ is called the "convection" current density while $j_\mu^{(2)}$ is referred to as the "magnetization" current density.

Experimental values of the gyromagnetic ratio for real spin-1/2 particles are found to be

$$\begin{aligned} \text{electron} \quad & g_{\text{exp}} = 2\left(1 + \frac{\alpha}{2\pi} + \ldots\right) \\ \text{proton} \quad & g_{\text{exp}} = 2\,(1 + 1.79) \\ \text{neutron} \quad & g_{\text{exp}} = 2\,(0 - 1.91) \quad . \end{aligned} \tag{5.91}$$

Does this mean that these are *not* Dirac particles? Not really. In the case of the proton and neutron, a microscopic view of these systems reveals that they are far from being simple pointlike spin-1/2 structures. Rather they are composed of three pointlike particles called quarks and the bound state wavefunction, which extends over distances of the order of $10^{-13}$ cm, clearly does not represent a pointlike structure. At a second level particles like the proton/neutron can fragment virtually into a nucleon-meson system with which the photon can interact, as shown in Figure VII.3.

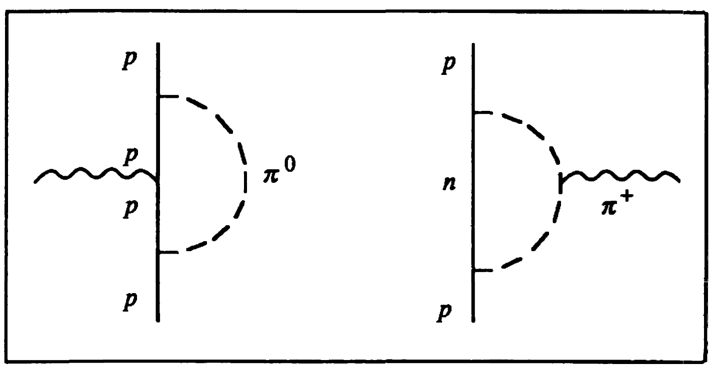



*Fig. VII.3: Mesonic corrections to the nucleon-photon interaction.*

Again we should expect substantial deviations from the result g=2 expected for a pointlike particle. In the case of the electron, there exists no quark substructure to deal with (as far as we know the electron really is a point particle). However, the electron can fragment into an $e - \gamma$ system as shown in Figure VII.4

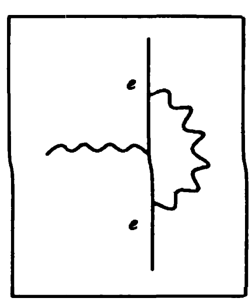



*Fig. VII.4: Radiative corrections to the electron-photon interaction.*

which yields a modification of the $g$-factor, but only at $\mathcal{O}(\alpha)$, as will be shown in the next chapter.

**PROBLEM VII.5.1**

**Electric Dipole Moment of the Electron [BjD 64]**

Suppose that the electron had a static electric dipole moment $d$ analogous to its magnetic moment.

i) Show that this could be accomodated by modifying the Dirac equation to become

$$(i \nabla\!\!\!/ - e A\!\!\!/ - i\frac{ed}{4m}\sigma_{\mu\nu}\gamma_5 F^{\mu\nu} - m)\psi(x) = 0 \ .$$

ii) Demonstrate that this equation is covariant but *not* invariant under a parity transformation.

iii) Show that this interaction would lead to a mixing between the $2S_{\frac{1}{2}}$ and $2P_{\frac{1}{2}}$ levels of the hydrogen atom and from the observed agreement between calculated and measured values of the Lamb shift at the level of a 0.05 MHz obtain an upper bound on the electric dipole moment of the electron.

Note: The relevant matrix element vanishes if the nonrelativistic wavefunctions are used—you will need the appropriate relativistic solutions.

**PROBLEM VII.5.2**

**Chirality and the Dirac Equation**

The operators

$$P_L = \frac{1}{2}(1+\gamma_5) \qquad P_R = \frac{1}{2}(1-\gamma_5)$$

are projection operators which are said to identify states of definite chirality (handedness).

i) Show that $P_L, P_R$ are legitimate projection operators in that

$$P_L^2 = P_L \qquad P_R^2 = P_R \qquad P_L P_R = P_R P_L = 0 \ .$$

ii) Demonstrate that in the limit of high energy—$E/m >> 1$—or equivalently in the massless limit that the Dirac spinors for positive helicity (right-handed) and negative helicity (left-handed) states of momentum $\vec{p}$ are given by

$$u_\pm(\vec{p}) = \sqrt{\frac{1}{2}}\begin{pmatrix} \chi_{\pm\hat{p}} \\ \pm\chi_{\pm\hat{p}} \end{pmatrix}$$

where $\chi_{\pm\hat{p}}$ are spinors such that

$$\vec{\sigma}\cdot\hat{p}\chi_{\pm\hat{p}} = \pm\chi_{\pm\hat{p}} \ .$$

Note: In order to conveniently deal with massless particles, it is important to use the normalization $u(p)^\dagger u(p) = 1$. The appropriate Dirac spinors can then be found by multiplying the usual forms by the factor $\sqrt{\frac{m}{E}}$. Demonstrate this.

iii) Show that

$$P_L u_-(p) = u_-(p) \qquad P_R u_+(p) = u_+(p) \qquad P_L u_+(p) = P_R u_-(p) = 0$$

so that the chirality operator is equivalent to the helicity operator in this limit.

**PROBLEM VII.5.3**

**Electron in a Magnetic Field**

Consider an electron immersed in a uniform magnetic field

$$\vec{B} = B_0 \hat{k} \quad .$$

i) Obtain the most general four-component positive energy eigenfunctions and demonstrate that the energy eigenvalues are given by

$$E = \sqrt{m^2 + p_3^2 + 2neB_0} \qquad n = 0, 1, 2, \ldots$$

ii) Compare your answer with what is expected nonrelativistically.

## VIII.6 NEGATIVE ENERGY SOLUTIONS AND ANTIPARTICLES

It is important at this point to address the question of the meaning of the negative energy solutions. We know that quantum mechanics is based upon the prescription

$$p^\mu \longrightarrow i\partial^\mu \quad \text{not} \quad p^\mu \longrightarrow -i\partial^\mu \quad . \tag{6.1}$$

But physical observables are real numbers and so cannot depend on this choice of $+i$ vs. $-i$. Does this freedom correspond to any freedom in the physical world? One answer is yes — it represents the particle/antiparticle duality seen in nature. We have already observed this duality in the case of the Klein–Gordon equation, wherein a particle solution corresponds to the positive energy Klein–Gordon wavefunction

$$\phi_{\text{part}}(x) = \phi_{E>0}(x) \tag{6.2}$$

while the antiparticle solution corresponds to the complex conjugate of the negative energy solution

$$\phi_{\text{antipart}}(x) = \phi^*_{E<0}(x) \quad . \tag{6.3}$$

### Antiparticles and the Dirac Equation

Similar results obtain for the Dirac equation. If we begin with a positive energy solution

$$\psi(x) = e^{-iEt}\psi(\vec{x}) \qquad E > 0 \tag{6.4}$$

which satisfies the equation

$$(i \not\nabla - e \not A - m)\, \psi(x) = 0 \tag{6.5}$$

then we may identify this as the wavefunction of a particle with charge $e$.

On the other hand, if we consider the negative energy solution

$$\psi(x) = e^{-iEt} \psi(\vec{x}) \qquad E = -W < 0 \tag{6.6}$$

and take the complex conjugate

$$\psi^*(x) = e^{-iWt} \psi^*(\vec{x}) \tag{6.7}$$

we see that this wavefunction obeys the differential equation

$$\left(-i\nabla^\mu \gamma_\mu^* - eA^\mu \gamma_\mu^* - m\right) \psi^*(x) = 0 \ . \tag{6.8}$$

In order to reproduce the Dirac equation, we need a transformation under which

$$\gamma_\mu^* \to -\gamma_\mu \ . \tag{6.9}$$

That is, we seek the "charge conjugation" operator $C\gamma^0$ which satisfies

$$\left(C\gamma^0\right) \gamma_\mu^* \left(C\gamma^0\right)^{-1} = -\gamma_\mu \ . \tag{6.10}$$

Since

$$\gamma_0^* = \gamma^0 \ , \quad \gamma_1^* = \gamma_1 \ , \quad \gamma_2^* = -\gamma_2 \ , \quad \gamma_3^* = \gamma_3 \tag{6.11}$$

we observe that the choice

$$C\gamma^0 = i\gamma_2 \ , \quad \left(C\gamma^0\right)^{-1} = i\gamma_2 \tag{6.12}$$

will suffice. Under this operation the Dirac equation becomes

$$C\gamma^0 \left(-i\nabla^\mu \gamma_\mu^* - eA^\mu \gamma_\mu^* - m\right) \psi^*(x) = \left(i\nabla^\mu \gamma_\mu + eA^\mu \gamma_\mu - m\right) C\gamma^0 \psi^*(x) = 0 \ . \tag{6.13}$$

If we define the antiparticle solution to be

$$\psi_{\text{antipart}}(x) = C\gamma^0 \psi^*(x) \tag{6.14}$$

we see that $\psi_{\text{antipart}}(x)$ obeys the Dirac equation for a particle of charge $-e$ and carries the correct (positive energy) time evolution.

[Note: Since

$$\psi^\dagger = \tilde{\psi}^* \tag{6.15}$$

where $\tilde{}$ indicates the transpose we have

$$\tilde{\bar{\psi}} = \tilde{\gamma}^0 \tilde{\psi}^\dagger = \gamma^0 \psi^* \ . \tag{6.16}$$

Thus the antiparticle solution is sometimes written as

$$\psi_{\text{antiparticle}}(x) = C\bar{\bar{\psi}} \quad .] \tag{6.17}$$

An interesting feature has to do with the intrinsic parities of these particle/antiparticle solutions. For a particle (positive energy) solution $\psi(\vec{x}, t)$ we have (*cf.* sect. VII.2)

$$\begin{aligned}\Pi\psi(\vec{x}, t) = \gamma^0\psi(-\vec{x}, t) &= \gamma^0 \begin{pmatrix} \psi_a(-\vec{x}, t) \\ \psi_b(-\vec{x}, t) \end{pmatrix} \\ &= \begin{pmatrix} \psi_a(-\vec{x}, t) \\ -\psi_b(-\vec{x}, t) \end{pmatrix} \quad ,\end{aligned} \tag{6.18}$$

where $\Pi$ is the spatial inversion or parity operator. Thus upper and lower components of the wavefunction then have differing behaviors under a parity transform. However, this is to be expected, since we have

$$\psi_b = \frac{\left(\vec{\sigma}\cdot(-i\vec{\nabla} - e\vec{A}\right)}{E+m}\psi_a \quad . \tag{6.19}$$

and the extra minus sign arises from the behavior of $\vec{\nabla}$, $\vec{A}$ under parity

$$\vec{\nabla}, \vec{A} \underset{P}{\longrightarrow} -\vec{\nabla}, -\vec{A} \quad . \tag{6.20}$$

However, if we look at the analogous negative energy or antiparticle solution, we find

$$P\psi_{\text{antipart}}(\vec{x}, t) = \gamma^0 C\gamma^0\psi^*(-\vec{x}, t) = -C\gamma^0 \begin{pmatrix} \psi_a^*(-\vec{x}, t) \\ -\psi_b^*(-\vec{x}, t) \end{pmatrix} \quad . \tag{6.21}$$

Thus the intrinsic parity of Dirac particles and antiparticles are opposite! (This is to be contrasted to the Klein–Gordon case wherein one finds identical parities for particle/antiparticle solutions.) This feature is verified experimentally in study of positronium decay or in the observation that ground ($S$-wave) states of quark-antiquark bound systems — *i.e.*, $\pi$, $K$, $\eta$ mesons — are determined to be *pseudo*scalar rather than scalar quantities.

## Dirac Sea

When Dirac found these negative energy solutions, he did not in the beginning understand their significance. (Recall that the Dirac equation was written down in 1928, but Anderson did not find the positron until 1932.) At first Dirac was worried that, since nature always prefers to lower the energy, positive energy electrons would radiate photons (of energy $\gtrsim 2m$) and fall into negative energy states. However, once in a negative energy state the electron could reduce its energy even further by radiating additional photons, thus lowering its energy indefinitely. Since this does not happen — experimentally the electron is stable— Dirac was faced with a dilemma. He solved this crisis by postulating that *all* negative energy states were already filled. Then, according to the Pauli exclusion principle, positive energy states are unable to make transitions to negative energy levels. Nevertheless, it is possible for an energetic ($\omega \gtrsim 2m$) photon to cause a negative energy electron to make a transition to a positive energy state leaving a "hole" in the set of of negative energy states.

The vacuum (*i.e.*, lowest energy state) in this picture consists of a "Dirac sea" filled negative energy states, and the net charge of a given state must be defined with respect to the vacuum. Thus a hole state behaves as if it had charge

$$Q_{\text{hole}} = (Q_{\text{vacuum}} - (e)) - Q_{\text{vacuum}} = -e \tag{6.22}$$

*i.e.*, the *negative* of the electron charge! (Of course, $Q_{\text{vacuum}}$ is infinite but we have seen such infinite renormalizations before.) Similarly if the momentum of this (negative energy) state is $\vec{p}$, the hole, upon renormalization, will behave as if it has momentum

$$\vec{P}_{\text{hole}} = \left(\vec{P}_{\text{vacuum}} - \vec{p}\right) - \vec{P}_{\text{vacuum}} = -\vec{p} \ . \tag{6.23}$$

(In this case we expect $\vec{P}_{\text{vacuum}} = 0$ since for each negative energy state with momentum $\vec{p}$ there is another with momentum $-\vec{p}$.) Finally, for the energy and spin we have

$$\begin{aligned} E_{\text{hole}} &= (E_{\text{vacuum}} - (-E)) - E_{\text{vacuum}} = +E \\ \frac{1}{2}\vec{\Sigma}_{\text{hole}} &= \left(\frac{1}{2}\vec{\Sigma}_{\text{vacuum}} - \frac{1}{2}\vec{\Sigma}\right) - \frac{1}{2}\vec{\Sigma}_{\text{vacuum}} = -\frac{1}{2}\vec{\Sigma} \ . \end{aligned} \tag{6.24}$$

so that the hole state behaves as a positive energy, positive charge state of momentum $-\vec{p}$ and spin $-\frac{1}{2}\vec{\Sigma}$. We recognize this state as a positron, whose existence was predicted by Dirac *prior* to its discovery. (Actually, Dirac first identified this antiparticle solution with the proton, but soon realized that its mass must be identical to that of the electron.)

We see then that the process by which an energetic photon ejects a negative energy electron from the Dirac sea, knocking it into a positive energy level and leaving behind a negative energy hole is to be interpreted as the process of pair creation

$$\gamma \longrightarrow e^+ e^- \ . \tag{6.25}$$

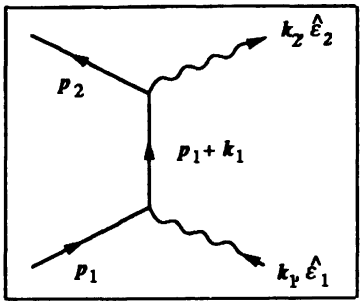
Since only a very energetic ($\omega \gtrsim 2m$) photon can bring about such a transition, one might be tempted to think that antiparticle states should not play an important role in low energy quantum mechanics. However, this is *not* correct. Consider the scattering of photons by a free positive energy electron. (This is the relativistic analog of the Thomson scattering process discussed previously.) Writing the Dirac equation in Hamiltonian form

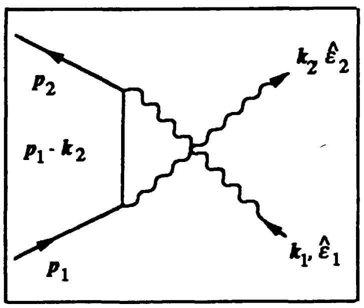
$$\begin{aligned} i\frac{\partial}{\partial t}\psi &= \left(\vec{\alpha} \cdot \vec{p} + \beta m + e\gamma^0 \not{A}\right) \psi \\ &\equiv (H_0 + V)\,\psi \end{aligned} \tag{6.26}$$

we recognize the interaction potential as

$$V = e\gamma^0 \not{A} \quad . \tag{6.27}$$

and can apply canonical time-dependent perturbation theory. Two obvious diagrams which arise in second order are shown below




$$\left\langle \vec{p}_2; \vec{k}_2, \hat{\epsilon}_2 \left| \hat{V} \frac{1}{E_{\vec{p}_1} + \omega_1 - \hat{H}_0} \hat{V} \right| \vec{p}_1; \vec{k}_1, \hat{\epsilon}_1 \right\rangle = \frac{e^2}{\sqrt{2\omega_1 2\omega_2}}$$

$$\times \frac{\left\langle \vec{p}_2 \left| \vec{\alpha} \cdot \hat{\epsilon}_2^* e^{-i\vec{k}_2 \cdot \vec{x}} \right| \vec{p}_1 + \vec{k}_1 \right\rangle \left\langle \vec{p}_1 + \vec{k}_1 \left| \vec{\alpha} \cdot \hat{\epsilon}_1 e^{i\vec{k}_1 \cdot \vec{x}} \right| p_1 \right\rangle}{E_{\vec{p}_1} + \omega_1 - E_{\vec{p}_1 + \vec{k}_1}}$$

$$\times \frac{\left\langle \vec{p}_2 \left| \vec{\alpha} \cdot \hat{\epsilon}_1 e^{i\vec{k}_1 \cdot \vec{x}} \right| \vec{p}_1 - \vec{k}_2 \right\rangle \left\langle \vec{p}_1 - \vec{k}_2 \left| \vec{\alpha} \cdot \hat{\epsilon}_2^* e^{-i\vec{k}_2 \cdot \vec{x}} \right| \vec{p}_1 \right\rangle}{E_{\vec{p}_1} - \omega_2 - E_{\vec{p}_1 - \vec{k}_2}} \tag{6.28}$$

which correspond to similar diagrams discussed in the analogous non-relativistic case. However, as Eq. 6.27 contains no term in $A^2$, there exists no analog of the seagull diagrams for the relativistic situation. Also since in the non-relativistic limit

$$u(p) \longrightarrow \begin{pmatrix} \chi \\ 0 \end{pmatrix} \tag{6.29}$$

while

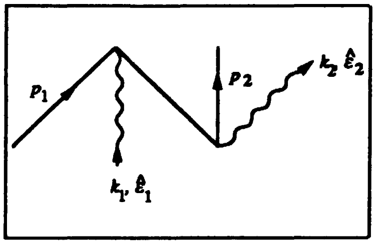
$$\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix} \tag{6.30}$$

couples upper and lower components so that

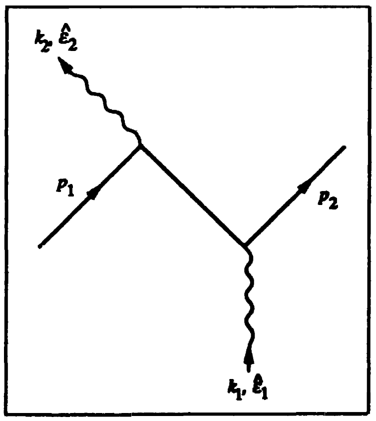
$$u^\dagger(p')\vec{\alpha}u(p) \approx 0, \tag{6.31}$$

there exists essentially *no* contribution to the scattering from the above pole diagrams.

The paradox is resolved if we include the contribution from the diagrams involving negative energy solutions as shown below




$$\times\ \frac{\left\langle \vec{p}_2 \left| \vec{\alpha}\cdot\hat{\epsilon}_2^* e^{-i\vec{k}_2\cdot\vec{x}} \right| \overline{\vec{p}_2+\vec{k}_2} \right\rangle \left\langle \overline{\vec{p}_2+\vec{k}_2} \left| \vec{\alpha}\cdot\hat{\epsilon}_1 e^{i\vec{k}_1\cdot\vec{x}} \right| \vec{p}_1 \right\rangle}{-E_{\vec{p}_2}-\omega_2-\left|E_{\vec{p}_2+\vec{k}_2}\right|}$$




$$\times\ \frac{\left\langle \vec{p}_2 \left| \vec{\alpha}\cdot\hat{\epsilon}_1 e^{i\vec{k}_1\cdot\vec{x}} \right| \overline{\vec{p}_2-\vec{k}_1} \right\rangle \left\langle \overline{\vec{p}_2-\vec{k}_1} \left| \vec{\alpha}\cdot\hat{\epsilon}_2^* e^{-i\vec{k}_2\cdot\vec{x}} \right| \vec{p}_1 \right\rangle}{-E_{\vec{p}_2}+\omega_1-\left|E_{\vec{p}_2-\vec{k}_1}\right|} \tag{6.32}$$

Then if $\omega \ll m$ and working in the non-relativistic limit where

$$v^\dagger(p')\vec{\alpha}u(p) \approx \chi'^\dagger\vec{\sigma}\chi \tag{6.33}$$

the sum of these negative energy diagrams yields

$$\begin{aligned}
\text{Amp} &\approx \frac{e^2}{\sqrt{2\omega_1 2\omega_2}}\frac{1}{2m}\sum_{i=1}^{2}\left(\chi_2^\dagger\vec{\sigma}\cdot\hat{\epsilon}_2^*\chi_i\chi_i^\dagger\vec{\sigma}\cdot\hat{\epsilon}_1\chi_1+\chi_2^\dagger\vec{\sigma}\cdot\hat{\epsilon}_1\chi_i\chi_i^\dagger\vec{\sigma}\cdot\hat{\epsilon}_2^*\chi_1\right)\\
&= \frac{e^2}{\sqrt{2\omega_1 2\omega_2}}\frac{1}{2m}\chi_2^\dagger\left(\vec{\sigma}\cdot\hat{\epsilon}_2^*\vec{\sigma}\cdot\hat{\epsilon}_1+\vec{\sigma}\cdot\hat{\epsilon}_1\vec{\sigma}\cdot\hat{\epsilon}_2^*\right)\chi_1\\
&= \frac{e^2}{\sqrt{2\omega_1 2\omega_2}}\frac{1}{2m}2\hat{\epsilon}_2^*\cdot\hat{\epsilon}_1\chi_2^\dagger\chi_1
\end{aligned} \tag{6.34}$$

which is identical to the seagull contribution found in the analogous non-relativistic case. We observe then that the inclusion of these negative energy states is absolutely crucial. Without these pieces the Dirac picture of Compton scattering would not reduce to the simple Thomson process.

## Zitterbewegung

We have already discussed the presence of zitterbewegung associated with the origin of the $\vec{\nabla}\cdot\vec{E}$ term in the effective non-relativistic Hamiltonian. We can study this phenomenon in more detail by examining the velocity operator $\vec{\alpha}$. That the Dirac matrix $\vec{\alpha}$ is related to the relativistic velocity is clear from the relation

$$\psi^\dagger(x)\vec{\alpha}\psi(x) = |N|^2\bar{u}(\vec{p})\vec{\gamma}u(\vec{p}) = |N|^2\frac{\vec{p}}{m} \tag{6.35}$$

valid for the plane wave solution

$$\psi(x) = N u(p) e^{-ip\cdot x} \quad . \tag{6.36}$$

The constant $N$ is determined by the normalization condition

$$1 = \int d^3x\, \rho(\vec{x},t) = |N|^2 \frac{E}{m} \int d^3x = |N|^2 \frac{E}{m} \tag{6.37}$$

using unit volume $\left(\int d^3x = 1\right)$. Thus

$$|N|^2 = \frac{m}{E} \tag{6.38}$$

and

$$\psi^\dagger(x)\vec{\alpha}\psi(x) = \frac{\vec{p}}{E} = \vec{v} \quad . \tag{6.39}$$

[Note: This is also suggested by the relation

$$\begin{aligned} \left[H, \vec{p} - e\vec{A}\right] &= ie\vec{\nabla}\phi - ie\vec{\alpha}\times\vec{B} \\ &= -ie\left(\vec{E} + \vec{\alpha}\times\vec{B}\right) \quad \text{if} \quad \frac{\partial \vec{A}}{\partial t} = 0 \end{aligned} \tag{6.40}$$

since

$$\frac{d}{dt}\mathcal{O} = i[H,\mathcal{O}] \quad .] \tag{6.41}$$

For a free particle, using the Heisenberg representation, we have

$$\frac{d}{dt}\vec{x} = i\,[H,\vec{x}] = \vec{\alpha} \tag{6.42}$$

and

$$\begin{aligned} \frac{d}{dt}\vec{\alpha} &= i\,[H,\vec{\alpha}] = i(-2m\vec{\alpha}\beta + 2i\vec{\Sigma}\times\vec{p}) \\ &= i\,(-2\vec{\alpha}H + 2\vec{p}) \end{aligned} \quad . \tag{6.43}$$

We see that

$$\vec{\alpha}(t) = \vec{p}H^{-1} + \left(\vec{\alpha}(0) - \vec{p}H^{-1}\right)e^{-2iHt} \quad . \tag{6.44}$$

[Check:

$$\begin{aligned} \frac{d\vec{\alpha}}{dt} &= \left(\vec{\alpha}(0) - \vec{p}H^{-1}\right)\times(-2iH)\,e^{-2iHt} \\ &= -2i\left(\vec{\alpha}(t) - \vec{p}H^{-1}\right)H \\ &= i\,(-2\vec{\alpha}H + 2\vec{p}) \qquad ] \end{aligned} \tag{6.45}$$

and thereby

$$\vec{x}(t) = \vec{x}(0) + \vec{p}H^{-1}t + \frac{i}{2}\left(\vec{\alpha}(0) - \vec{p}H^{-1}\right)H^{-1}\left(e^{-2iHt} - 1\right) \quad . \tag{6.46}$$