8

Quantum theory and spinors

In this chapter we study the application of geometric algebra to both non-relativistic and relativistic quantum mechanics. We concentrate on the quantum theory of spin-1/2 particles, whose dynamics is described by the Pauli and Dirac equations. For interactions where spin and relativity are not important the dynamics reduces to that of the Schrödinger equation. There are many good textbooks describing this topic and we will make no attempt to cover it here. We assume, furthermore, that most readers have a basic understanding of quantum mechanics, and are familiar with the concepts of states and operators.

Both the Pauli and Dirac matrices arise naturally as representations of the geometric algebras of space and spacetime. It is no surprise, then, that much of quantum theory finds a natural expression within geometric algebra. To achieve this, however, one must reconsider the standard interpretation of the quantum spin operators. Like much discussion of the interpretation of quantum theory, certain issues raised here are controversial. There is no question about the validity of our algebraic approach, however, and little doubt about its advantages. Whether the algebraic simplifications obtained here are indicative of a deeper structure embedded in quantum mechanics is an open question.

In this chapter we only consider the quantum theory of single particles in background fields. Multiparticle systems are considered in the following chapter. Amongst the results discussed in this section are the angular separation of the Dirac equation, and a method of calculating cross sections that avoids the need for spin sums. Both of these results are used in chapter 14 for studying the behaviour of fermions in gravitational backgrounds.

8.1 Non-relativistic quantum spin

The Stern–Gerlach experiment was the first to demonstrate the quantum nature of the magnetic moment. In this experiment a beam of particles passes through

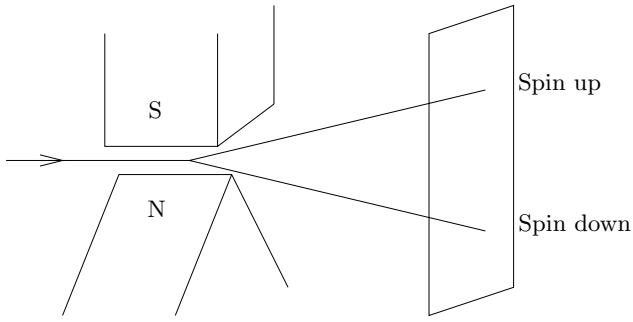

Figure 8.1 *The Stern–Gerlach experiment.* A particle beam is sent through a highly non-uniform $\boldsymbol{B}$ field. What emerges is a set of discrete, evenly-spaced beams.

a non-uniform magnetic field $\boldsymbol{B}$. Classically, one would expect the force on each particle to be governed by the equation

$$\boldsymbol{f} = \boldsymbol{\mu}\cdot\boldsymbol{\nabla}\boldsymbol{B}, \tag{8.1}$$

where $\boldsymbol{\mu}$ is the magnetic moment. This would give rise to a continuous distribution after passing through the field. Instead, what is observed is a number of evenly-spaced discrete bands (figure 8.1). The magnetic moment is quantised in the same manner as angular momentum.

When silver atoms are used to make up the beam there is a further surprise: only two beams emerge on the far side. Silver atoms contain a single electron in their outermost shell, so it looks as if electrons have an intrinsic angular momentum which can take only two values. This is known as its *spin*, though no classical picture should be inferred from this name. The double-valued nature of the spin suggests that the electron's wavefunction should contain two terms, representing a superposition of the possible spin states,

$$|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle, \tag{8.2}$$

where α and β are complex numbers. Such a state can be represented in matrix form as the *spinor*

$$|\psi\rangle = \begin{pmatrix}\alpha\\ \beta\end{pmatrix}. \tag{8.3}$$

If we align the z axis with the spin-up direction, then the operator returning the spin along the z axis must be

$$\hat{s}_3 = \lambda\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}, \tag{8.4}$$

where λ is to be determined. The spin is added to the orbital angular momentum

to give a conserved total angular momentum operator $\hat{\boldsymbol{j}} = \hat{\boldsymbol{l}} + \hat{\boldsymbol{s}}$. For this to make sense the spin operators should have the same commutation relations as the angular momentum operators $\hat{l}_i$,

$$\hat{l}_i = -i\hbar\epsilon_{ijk}x_j\partial_k, \quad [\hat{l}_i, \hat{l}_j] = i\hbar\epsilon_{ijk}\hat{l}_k. \tag{8.5}$$

This is sufficient to specify the remaining operators, up to an arbitrary phase (see exercise 8.1). The result is that the spin operators are given by

$$\hat{s}_k = \tfrac{1}{2}\hbar\hat{\sigma}_k, \tag{8.6}$$

where the $\hat{\sigma}_k$ are the familiar *Pauli matrices*

$$\hat{\sigma}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{\sigma}_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{8.7}$$

The 'hat' notation is used to record the fact that these are viewed explicitly as matrix operators, rather than as elements of a geometric algebra. The Pauli matrices satisfy the commutation relations,

$$[\hat{\sigma}_i, \hat{\sigma}_j] = 2i\epsilon_{ijk}\hat{\sigma}_k. \tag{8.8}$$

They also have the property that two different matrices *anticommute*,

$$\hat{\sigma}_1\hat{\sigma}_2 + \hat{\sigma}_2\hat{\sigma}_1 = 0, \quad \text{etc.} \tag{8.9}$$

and all of the matrices square to the identity matrix,

$$\hat{\sigma}_1^2 = \hat{\sigma}_2^2 = \hat{\sigma}_3^2 = \mathsf{I}. \tag{8.10}$$

These are precisely the relations obeyed by a set of orthonormal vectors in space. We denote such a set by $\{\boldsymbol{\sigma}_k\}$. The crucial distinction is that the Pauli matrices are operators in quantum isospace, whereas the $\{\boldsymbol{\sigma}_k\}$ are vectors in real space.

The $\hat{\sigma}_k$ operators act on two-component complex spinors as described in equation (8.3). Spinors belong to two-dimensional complex vector space, so have four real degrees of freedom. A natural question to ask is whether an equivalent representation can be found in terms of real multivectors, such that the matrix action is replaced by multiplication by the $\{\boldsymbol{\sigma}_k\}$ vectors. To find a natural way to do this we consider the observables of a spinor. These are the eigenvalues of Hermitian operators and, for two-state systems, the relevant operators are the Pauli matrices. We therefore form the three observables

$$s_k = \tfrac{1}{2}\hbar n_k = \langle\psi|\hat{s}_k|\psi\rangle. \tag{8.11}$$

The n_k are the components of a single vector in the quantum theory of spin. Focusing attention on the components of this vector, we have

$$\begin{aligned} n_1 &= \langle\psi|\hat{\sigma}_1|\psi\rangle = \alpha\beta^* + \alpha^*\beta, \\ n_2 &= \langle\psi|\hat{\sigma}_2|\psi\rangle = i(\alpha\beta^* - \alpha^*\beta), \\ n_3 &= \langle\psi|\hat{\sigma}_3|\psi\rangle = \alpha\alpha^* - \beta\beta^*. \end{aligned} \tag{8.12}$$

The magnitude of the vector with components n_k is

$$\begin{aligned}|\boldsymbol{n}|^2 &= (\alpha\beta^* + \alpha^*\beta)^2 - (\alpha\beta^* - \alpha^*\beta)^2 + (\alpha\alpha^* - \beta\beta^*)^2 \\ &= (|\alpha|^2 + |\beta|^2)^2 = \langle\psi|\psi\rangle^2.\end{aligned} \tag{8.13}$$

So, provided the state is normalised to 1, the vector $\boldsymbol{n}$ must have unit length. We can therefore introduce polar coordinates and write

$$\begin{aligned}n_1 &= \sin(\theta)\cos(\phi), \\ n_2 &= \sin(\theta)\sin(\phi), \\ n_3 &= \cos(\theta).\end{aligned} \tag{8.14}$$

Comparing equation (8.14) with equation (8.12) we see that we must have

$$\alpha = \cos(\theta/2)\mathrm{e}^{i\gamma}, \qquad \beta = \sin(\theta/2)\mathrm{e}^{i\delta} \tag{8.15}$$

where $\delta - \gamma = \phi$. It follows that the spinor can be written in terms of the polar coordinates of the vector observable as

$$|\psi\rangle = \begin{pmatrix}\cos(\theta/2)\mathrm{e}^{-i\phi/2} \\ \sin(\theta/2)\mathrm{e}^{i\phi/2}\end{pmatrix}\mathrm{e}^{i(\gamma+\delta)/2}. \tag{8.16}$$

The overall phase factor can be ignored, and what remains is a description in terms of half-angles. This suggests a strong analogy with *rotors*. To investigate this analogy, we use the idea that polar coordinates can be viewed as part of an instruction to rotate the 3 axis onto the chosen vector. To expose this we write the vector $\boldsymbol{n}$ as

$$\boldsymbol{n} = \sin(\theta)\big(\cos(\phi)\boldsymbol{\sigma}_1 + \sin(\phi)\boldsymbol{\sigma}_2\big) + \cos(\theta)\boldsymbol{\sigma}_3. \tag{8.17}$$

This can be written

$$\boldsymbol{n} = R\boldsymbol{\sigma}_3 R^\dagger, \tag{8.18}$$

where

$$R = \mathrm{e}^{-\phi I\boldsymbol{\sigma}_3/2}\mathrm{e}^{-\theta I\boldsymbol{\sigma}_2/2}. \tag{8.19}$$

This suggests that there should be a natural map between the normalised spinor of equation (8.16) and the rotor R. Both belong to linear spaces of real dimension four and both are normalised. Expanding out the rotor R the following one-to-one map is found:

$$|\psi\rangle = \begin{pmatrix}a^0 + ia^3 \\ -a^2 + ia^1\end{pmatrix} \leftrightarrow \psi = a^0 + a^k I\boldsymbol{\sigma}_k. \tag{8.20}$$

This map will enable us to perform all operations involving spinors without leaving the geometric algebra of space. Throughout this chapter we use the $\leftrightarrow$ symbol to denote a one-to-one map between conventional quantum mechanics and the multivector equivalent. We will continue to refer to the multivector ψ

as a spinor. On this scheme the spin-up and spin-down basis states $|\uparrow\rangle$ and $|\downarrow\rangle$ become

$$|\uparrow\rangle \leftrightarrow 1 \qquad |\downarrow\rangle \leftrightarrow -I\boldsymbol{\sigma}_2. \tag{8.21}$$

One can immediately see for these that the vectors of observables have components $(0, 0, \pm 1)$, as required.

8.1.1 Pauli operators

Now that a suitable one-to-one map has been found, we need to find a representation for Pauli operators acting on the multivector version of a spinor. It turns out that the action of the quantum $\hat{\sigma}_k$ operators on a state $|\psi\rangle$ is equivalent to the following operation on ψ:

$$\hat{\sigma}_k|\psi\rangle \leftrightarrow \boldsymbol{\sigma}_k\psi\boldsymbol{\sigma}_3 \quad (k = 1, 2, 3). \tag{8.22}$$

The $\boldsymbol{\sigma}_3$ on the right-hand side ensures that the multivector remains in the even subalgebra. The choice of vector does not break rotational covariance, in the same way that choosing the $\hat{\sigma}_3$ matrix to be diagonal does not alter the rotational covariance of the Pauli theory. One can explicitly verify that the translation procedure of equation (8.20) and equation (8.22) is consistent by routine computation; for example

$$\hat{\sigma}_1|\psi\rangle = \begin{pmatrix} -a^2 + ia^1 \\ a^0 + ia^3 \end{pmatrix} \leftrightarrow -a^2 + a^1 I\boldsymbol{\sigma}_3 - a^0 I\boldsymbol{\sigma}_2 + a^3 I\boldsymbol{\sigma}_1 = \boldsymbol{\sigma}_1\psi\boldsymbol{\sigma}_3. \tag{8.23}$$

The remaining cases, for $\hat{\sigma}_2$ and $\hat{\sigma}_3$ can be checked equally easily.

Now that we have a translation for the action of the Pauli matrices, we can find the equivalent of multiplying by the unit imaginary i. To find this we note that

$$\hat{\sigma}_1\hat{\sigma}_2\hat{\sigma}_3 = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}, \tag{8.24}$$

so multiplication of both components of $|\psi\rangle$ by i can be achieved by multiplying by the product of the three matrix operators. We therefore arrive at the translation

$$i|\psi\rangle \leftrightarrow \boldsymbol{\sigma}_1\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3\psi(\boldsymbol{\sigma}_3)^3 = \psi I\boldsymbol{\sigma}_3. \tag{8.25}$$

So, on this scheme, the unit imaginary of quantum theory is replaced by right multiplication by the *bivector* $I\boldsymbol{\sigma}_3$. This is certainly suggestive, though it should be borne in mind that this conclusion is a feature of our chosen representation. The appearance of the bivector $I\boldsymbol{\sigma}_3$ is to be expected, since the vector of observables $\boldsymbol{s} = s_k\boldsymbol{\sigma}_k$ was formed by rotating the $\boldsymbol{\sigma}_3$ vector. This vector is unchanged by rotations in the $I\boldsymbol{\sigma}_3$ plane, which provides a geometric picture of phase invariance.

8.1.2 Observables in the Pauli theory

We next need to establish the quantum inner product for our multivector form of a spinor. We first note that the Hermitian adjoint operation has $\hat{\sigma}_k^\dagger = \hat{\sigma}_k$, and reverses the order of all products. This is precisely the same as the reversion operation for multivectors in three dimensions, so the dagger symbol can be used consistently for both operations. The quantum inner product is

$$\langle\psi|\phi\rangle = (\psi_1^*, \psi_2^*)\begin{pmatrix}\phi_1\\ \phi_2\end{pmatrix} = \psi_1^*\phi_1 + \psi_2^*\phi_2, \tag{8.26}$$

where we ignore spatial integrals. For a wide range of problems the spatial and spin components of the wave function can be separated. If this is not the case then the quantum inner product should also contain an integral over all space. The result of the real part of the inner product is reproduced by

$$\mathrm{Re}\langle\psi|\phi\rangle \;\leftrightarrow\; \langle\psi^\dagger\phi\rangle, \tag{8.27}$$

so that, for example,

$$\langle\psi|\psi\rangle \;\leftrightarrow\; \langle\psi^\dagger\psi\rangle = \langle(a^0 - a^j I\boldsymbol{\sigma}_j)(a^0 + a^k I\boldsymbol{\sigma}_k)\rangle = \sum_{\alpha=0}^{3} a^\alpha a^\alpha. \tag{8.28}$$

Since

$$\langle\psi|\phi\rangle = \mathrm{Re}\langle\psi|\phi\rangle - i\mathrm{Re}\langle\psi|i\phi\rangle, \tag{8.29}$$

the full inner product can be written

$$\langle\psi|\phi\rangle \;\leftrightarrow\; \langle\psi^\dagger\phi\rangle - \langle\psi^\dagger\phi I\boldsymbol{\sigma}_3\rangle I\boldsymbol{\sigma}_3. \tag{8.30}$$

The right-hand side projects out the 1 and $I\boldsymbol{\sigma}_3$ components from the geometric product $\psi^\dagger\phi$. The result of this projection on a multivector A is written $\langle A\rangle_q$. For even-grade multivectors in three dimensions this projection has the simple form

$$\langle A\rangle_q = \tfrac{1}{2}(A + \boldsymbol{\sigma}_3 A\boldsymbol{\sigma}_3). \tag{8.31}$$

If the result of an inner product is used to multiply a second multivector, one has to remember to keep the terms in $I\boldsymbol{\sigma}_3$ to the right of the multivector. This might appear a slightly clumsy procedure at first, but it is easy to establish conventions so that manipulations are just as efficient as in the standard treatment. Furthermore, the fact that all manipulations are now performed within the geometric algebra framework offers a number of new ways to simplify the analysis of a range of problems.

8.1.3 The spin vector

As a check on the consistency of our scheme, we return to the expectation value of the spin in the k-direction, $\langle\psi|\hat{s}_k|\psi\rangle$. For this we require

$$\langle\psi|\hat{\sigma}_k|\psi\rangle \leftrightarrow \langle\psi^\dagger\boldsymbol{\sigma}_k\psi\boldsymbol{\sigma}_3\rangle - \langle\psi^\dagger\boldsymbol{\sigma}_k\psi I\rangle I\boldsymbol{\sigma}_3. \tag{8.32}$$

Since $\psi^\dagger I\boldsymbol{\sigma}_k\psi$ reverses to give minus itself it has zero scalar part, so the final term on the right-hand side vanishes. This is to be expected, as the $\hat{\sigma}_k$ are Hermitian operators. For the remaining term we note that in three dimensions $\psi\boldsymbol{\sigma}_3\psi^\dagger$ is both odd-grade and reverses to itself, so is a pure vector. We therefore define the spin vector

$$\boldsymbol{s} = \tfrac{1}{2}\hbar\psi\boldsymbol{\sigma}_3\psi^\dagger. \tag{8.33}$$

The quantum expectation now reduces to

$$\langle\psi|\hat{s}_k|\psi\rangle = \tfrac{1}{2}\hbar\langle\boldsymbol{\sigma}_k\psi\boldsymbol{\sigma}_3\psi^\dagger\rangle = \boldsymbol{\sigma}_k\cdot\boldsymbol{s}. \tag{8.34}$$

This new expression has a rather different interpretation to that usually encountered in quantum theory. Rather than forming the expectation value of a quantum operator, we are simply projecting out the kth component of the vector $\boldsymbol{s}$. Working with the vector $\boldsymbol{s}$ may appear to raise questions about whether we are free to talk about all three components of the spin vector. This is in fact consistent with the results of spin measurements, if we view the spin measurement apparatus as acting more as a spin polariser. This is discussed in Doran *et al.* (1996b).

The rotor description introduced at the start of this section is recovered by first defining the scalar

$$\rho = \psi\psi^\dagger. \tag{8.35}$$

The spinor ψ then decomposes into

$$\psi = \rho^{1/2}R, \tag{8.36}$$

where $R = \rho^{-1/2}\psi$. The multivector R satisfies $RR^\dagger = 1$, so is a rotor. In this approach, Pauli spinors are nothing but unnormalised rotors. The spin vector $\boldsymbol{s}$ can now be written as

$$\boldsymbol{s} = \tfrac{1}{2}\hbar\rho R\boldsymbol{\sigma}_3R^\dagger, \tag{8.37}$$

which recovers the form of equation (8.18).

The double-sided construction of the expectation value of equation (8.32) contains an instruction to rotate the fixed $\boldsymbol{\sigma}_3$ axis into the spin direction and dilate it. It might appear here that we are singling out some preferred direction in space. But in fact all we are doing is utilising an idea from rigid-body dynamics, as discussed in section 3.4.3. The $\boldsymbol{\sigma}_3$ on the right of ψ represents a vector in a 'reference' frame. All physical vectors, like $\boldsymbol{s}$, are obtained by rotating this frame

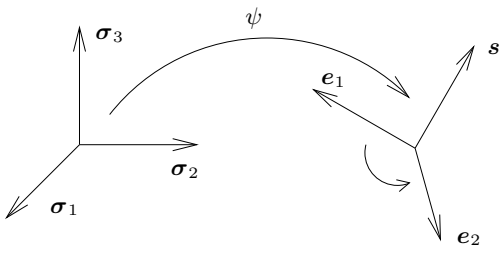

Figure 8.2 *The spin vector.* The normalised spinor ψ transforms the initial reference frame onto the frame $\{\boldsymbol{e}_k\}$. The vector $\boldsymbol{e}_3$ is the spin vector. A phase transformation of ψ generates a rotation in the $\boldsymbol{e}_1\boldsymbol{e}_2$ plane. Such a transformation is unobservable, so the $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ vectors are also unobservable.

onto the physical values (see figure 8.2). There is nothing special about $\boldsymbol{\sigma}_3$ — one can choose any (constant) reference frame and use the appropriate rotation onto $\boldsymbol{s}$, in the same way that there is nothing special about the orientation of the reference configuration of a rigid body. In rigid-body mechanics this freedom is usually employed to align the reference configuration with the initial state of the body. In quantum theory the convention is to work with the z axis as the reference vector.

8.1.4 Rotating spinors

Suppose that the vector $\boldsymbol{s}$ is to be rotated to a new vector $R_0\boldsymbol{s}R_0^\dagger$. To achieve this the spinor ψ must transform according to

$$\psi \mapsto R_0\psi. \tag{8.38}$$

Now suppose that for R_0 we use the rotor R_θ,

$$R_\theta = \exp(-\hat{B}\theta/2), \tag{8.39}$$

where $\hat{B}^2 = -1$ is a constant bivector. The resulting spinor is

$$\psi' = R_\theta\psi = \mathrm{e}^{-\hat{B}\theta/2}\psi. \tag{8.40}$$

We now start to increase θ from 0 through to 2π, so that $\theta = 2\pi$ corresponds to a 2π rotation, bringing all observables back to their original values. But under this we see that ψ transforms to

$$\psi' = \mathrm{e}^{-\hat{B}\pi}\psi = \big(\cos(\pi) - \hat{B}\sin(\pi)\big)\psi = -\psi. \tag{8.41}$$

The spinor changes sign! If a spin vector is rotated through 2π, the wavefunction does not come back to itself, but instead transforms to minus its original value.

This change of sign of a state vector under 2π rotations is the distinguishing property of spin-1/2 fermions in quantum theory. Once one sees the rotor derivation of this result, however, it is rather less mysterious. Indeed, there are classical phenomena involving systems of linked rotations that show precisely the same property. One example is the 4π symmetry observed when rotating an arm holding a tray. For a more detailed discussion if this point, see chapter 41 of *Gravitation* by Misner, Thorne & Wheeler (1973). A linear space which is acted on in a single-sided manner by rotors forms a carrier space for a spin representation of the rotation group. Elements of such a space are generally called spinors, which is why that name is adopted for our representation in terms of even multivectors.

8.1.5 Quantum particles in a magnetic field

Particles with non-zero spin also have a magnetic moment which is proportional to the spin. This is expressed as the operator relation

$$\hat{\mu}_k = \gamma \hat{s}_k, \tag{8.42}$$

where $\hat{\mu}_k$ is the magnetic moment operator, γ is the gyromagnetic ratio and $\hat{s}_k$ is the spin operator. The gyromagnetic ratio is usually written in the form

$$\gamma = g\frac{q}{2m}, \tag{8.43}$$

where m is the particle mass, q is the charge and g is the reduced gyromagnetic ratio. The reduced gyromagnetic ratios are determined experimentally to be

electron	$g_e = 2$	(actually $2(1 + \alpha/2\pi + \cdots)$),
proton	$g_p = 5.587$,	
neutron	$g_n = -3.826$	(using proton charge).

The value for the neutron is negative because its spin and magnetic moment are antiparallel. All of the above are spin-1/2 particles for which we have $\hat{s}_k = (\hbar/2)\hat{\sigma}_k$.

Now suppose that the particle is placed in a magnetic field, and that all of the spatial dynamics has been separated out. We introduce the Hamiltonian operator

$$\hat{H} = -\tfrac{1}{2}\gamma\hbar B_k\hat{\sigma}_k = -\hat{\mu}_k B_k. \tag{8.44}$$

The spin state at time t is then written as

$$|\psi(t)\rangle = \alpha(t)|\uparrow\rangle + \beta(t)|\downarrow\rangle, \tag{8.45}$$

with α and β general complex coefficients. The dynamical equation for these coefficients is given by the time-dependent Schrödinger equation

$$\hat{H}|\psi\rangle = i\hbar\frac{d|\psi\rangle}{dt}. \tag{8.46}$$

This equation can be hard to analyse, conventionally, because it involves a pair of coupled differential equations for α and β. Instead, let us see what the Schrödinger equation looks like in the geometric algebra formulation. We first write the equation in the form

$$\frac{d|\psi\rangle}{dt} = \tfrac{1}{2}\gamma i B_k\hat{\sigma}_k|\psi\rangle. \tag{8.47}$$

Now replacing $|\psi\rangle$ by the multivector ψ we see that the left-hand side is simply $\dot{\psi}$, where the dot denotes the time derivative. The right-hand side involves multiplication of the spinor $|\psi\rangle$ by $i\hat{\sigma}_k$, which we replace by

$$i\hat{\sigma}_k|\psi\rangle \leftrightarrow \boldsymbol{\sigma}_k\psi\boldsymbol{\sigma}_3(I\boldsymbol{\sigma}_3) = I\boldsymbol{\sigma}_k\psi. \tag{8.48}$$

The Schrödinger equation (8.46) is therefore simply

$$\dot{\psi} = \tfrac{1}{2}\gamma B_k I\boldsymbol{\sigma}_k\psi = \tfrac{1}{2}\gamma I\boldsymbol{B}\psi, \tag{8.49}$$

where $\boldsymbol{B} = B_k\boldsymbol{\sigma}_k$. If we now decompose ψ into $\rho^{1/2}R$ we see that

$$\dot{\psi}\psi^\dagger = \tfrac{1}{2}\dot{\rho} + \rho\dot{R}R^\dagger = \tfrac{1}{2}\rho\gamma I\boldsymbol{B}. \tag{8.50}$$

The right-hand side is a bivector, so ρ must be constant. This is to be expected, as the evolution should be unitary. The dynamics now reduces to

$$\dot{R} = \tfrac{1}{2}\gamma I\boldsymbol{B}R, \tag{8.51}$$

so the quantum theory of a spin-1/2 particle in a magnetic field reduces to a simple rotor equation. This is very natural, if one thinks about the behaviour of particles in magnetic fields, and is an important justification for our approach.

Recovering a rotor equation explains the difficulty of the traditional analysis based on a pair of coupled equations for the components of $|\psi\rangle$. This approach fails to capture the fact that there is a rotor underlying the dynamics, and so carries along a redundant degree of freedom in the normalisation. In addition, the separation of a rotor into a pair of components is far from natural. For example, suppose that $\boldsymbol{B}$ is a constant field. The rotor equation integrates immediately to give

$$\psi(t) = \mathrm{e}^{\gamma I\boldsymbol{B}t/2}\psi_0. \tag{8.52}$$

The spin vector $\boldsymbol{s}$ therefore just precesses in the $I\boldsymbol{B}$ plane at a rate $\omega_0 = \gamma|\boldsymbol{B}|$. Even this simple result is rather more difficult to establish when working with the components of $|\psi\rangle$.

8.1.6 NMR and magnetic resonance imaging

A more interesting example of a particle in a magnetic field is provided by nuclear magnetic resonance, or NMR. Suppose that the $\boldsymbol{B}$ field includes an oscillatory field $(B_1\cos(\omega t), B_1\sin(\omega t), 0)$ together with a constant field along the z axis. This oscillatory field induces transitions (spin-flips) between the up and down states, which differ in energy because of the constant component of the field. This is a very interesting system of great practical importance. It is the basis of magnetic resonance imaging and Rabi molecular beam spectroscopy.

To study this system we first write the $\boldsymbol{B}$ field as

$$B_1\big(\cos(\omega t)\boldsymbol{\sigma}_1 + \sin(\omega t)\boldsymbol{\sigma}_2\big) + B_0\boldsymbol{\sigma}_3 = S(B_1\boldsymbol{\sigma}_1 + B_0\boldsymbol{\sigma}_3)S^\dagger, \tag{8.53}$$

where

$$S = \mathrm{e}^{-\omega t I\boldsymbol{\sigma}_3/2}. \tag{8.54}$$

We now define

$$\boldsymbol{B}_c = B_1\boldsymbol{\sigma}_1 + B_0\boldsymbol{\sigma}_3 \tag{8.55}$$

so that we can write $\boldsymbol{B} = S\boldsymbol{B}_cS^\dagger$. The rotor equation now simplifies to

$$S^\dagger\dot{\psi} = \tfrac{1}{2}\gamma I\boldsymbol{B}_cS^\dagger\psi, \tag{8.56}$$

where we have pre-multiplied by $S^\dagger$, and we continue to use ψ for the normalised rotor. Now noting that

$$\dot{S}^\dagger = \tfrac{1}{2}\omega I\boldsymbol{\sigma}_3S^\dagger \tag{8.57}$$

we see that

$$\frac{d}{dt}(S^\dagger\psi) = \tfrac{1}{2}(\gamma I\boldsymbol{B}_c + \omega I\boldsymbol{\sigma}_3)S^\dagger\psi. \tag{8.58}$$

It is now $S^\dagger\psi$ that satisfies a rotor equation with a constant field. The solution is straightforward:

$$S^\dagger\psi(t) = \exp\Big(\tfrac{1}{2}\gamma t\, I\boldsymbol{B}_c + \tfrac{1}{2}\omega t\, I\boldsymbol{\sigma}_3\Big)\psi_0, \tag{8.59}$$

and we arrive at

$$\psi(t) = \exp\Big(-\tfrac{1}{2}\omega t\, I\boldsymbol{\sigma}_3\Big)\exp\Big(\tfrac{1}{2}(\omega_0+\omega)t\, I\boldsymbol{\sigma}_3 + \tfrac{1}{2}\omega_1 t\, I\boldsymbol{\sigma}_1\Big)\psi_0, \tag{8.60}$$

where $\omega_1 = \gamma B_1$. There are three separate frequencies in this solution, which contains a wealth of interesting physics.

To complete our analysis we must relate our solution to the results of experiments. Suppose that at time $t = 0$ we switch on the oscillating field. The particle is initially in a spin-up state, so $\psi_0 = 1$, which also ensures that the state is normalised. The probability that at time t the particle is in the spin-down state is

$$P_\downarrow = |\langle\downarrow|\psi(t)\rangle|^2. \tag{8.61}$$

We therefore need to form the inner product

$$\langle\downarrow|\psi(t)\rangle \leftrightarrow \langle I\boldsymbol{\sigma}_2\psi\rangle_q = \langle I\boldsymbol{\sigma}_2\psi\rangle - I\boldsymbol{\sigma}_3\langle I\boldsymbol{\sigma}_1\psi\rangle. \tag{8.62}$$

To find this inner product we write

$$\psi(t) = \mathrm{e}^{-\omega t I\boldsymbol{\sigma}_3/2}\big(\cos(\alpha t/2) + I\hat{\boldsymbol{B}}\sin(\alpha t/2)\big), \tag{8.63}$$

where

$$\hat{\boldsymbol{B}} = \frac{(\omega_0+\omega)\boldsymbol{\sigma}_3 + \omega_1\boldsymbol{\sigma}_1}{\alpha} \quad \text{and} \quad \alpha = \sqrt{(\omega+\omega_0)^2+\omega_1^2}. \tag{8.64}$$

The only term giving a contribution in the $I\boldsymbol{\sigma}_1$ and $I\boldsymbol{\sigma}_2$ planes is that in $\omega_1 I\boldsymbol{\sigma}_1/\alpha$. We therefore have

$$\langle I\boldsymbol{\sigma}_2\psi\rangle_q = \frac{\omega_1\sin(\alpha t/2)}{\alpha}\,\mathrm{e}^{-\omega t I\boldsymbol{\sigma}_3/2} I\boldsymbol{\sigma}_3 \tag{8.65}$$

and the probability is immediately

$$P_\downarrow = \left(\frac{\omega_1\sin(\alpha t/2)}{\alpha}\right)^2. \tag{8.66}$$

The maximum value is at $\alpha t = \pi$, and the probability at this time is maximised by choosing α as small as possible. This is achieved by setting $\omega = -\omega_0 = -\gamma B_0$. This is the *spin resonance condition* which is the basis of NMR spectroscopy.

8.2 Relativistic quantum states

The relativistic quantum dynamics of a spin-1/2 particle is described by the Dirac theory. The Dirac matrix operators are

$$\hat{\gamma}_0 = \begin{pmatrix} \mathsf{I} & 0 \\ 0 & -\mathsf{I} \end{pmatrix}, \quad \hat{\gamma}_k = \begin{pmatrix} 0 & -\hat{\sigma}_k \\ \hat{\sigma}_k & 0 \end{pmatrix}, \quad \hat{\gamma}_5 = \begin{pmatrix} 0 & \mathsf{I} \\ \mathsf{I} & 0 \end{pmatrix}, \tag{8.67}$$

where $\hat{\gamma}_5 = -i\hat{\gamma}_0\hat{\gamma}_1\hat{\gamma}_2\hat{\gamma}_3$ and I is the 2×2 identity matrix. These matrices act on Dirac spinors, which have four complex components (eight real degrees of freedom). We follow an analogous procedure to the Pauli case and map these spinors onto elements of the eight-dimensional even subalgebra of the spacetime algebra. Dirac spinors can be visualised as decomposing into 'upper' and 'lower' components,

$$|\psi\rangle = \begin{pmatrix} |\phi\rangle \\ |\eta\rangle \end{pmatrix}, \tag{8.68}$$

where $|\phi\rangle$ and $|\eta\rangle$ are a pair of two-component spinors. We already know how to represent these as multivectors ϕ and η, which lie in the space of scalars +

relative bivectors. Our map from the Dirac spinor onto an element of the full eight-dimensional subalgebra is simply

$$|\psi\rangle = \begin{pmatrix} |\phi\rangle \\ |\eta\rangle \end{pmatrix} \leftrightarrow \psi = \phi + \eta\boldsymbol{\sigma}_3. \tag{8.69}$$

The action of the Dirac matrix operators now becomes,

$$\begin{aligned} \hat{\gamma}_\mu|\psi\rangle &\leftrightarrow \gamma_\mu\psi\gamma_0 \quad (\mu = 0, \ldots, 3), \\ i|\psi\rangle &\leftrightarrow \psi I\boldsymbol{\sigma}_3, \\ \hat{\gamma}_5|\psi\rangle &\leftrightarrow \psi\boldsymbol{\sigma}_3. \end{aligned} \tag{8.70}$$

Again, verifying the details of this map is a matter of routine computation. One feature is that we now have two 'reference' vectors that can appear on the right-hand side of ψ: γ_0 and γ_3. That is, the relative vector $\boldsymbol{\sigma}_3$ used in the Pauli theory has been decomposed into a spacelike and a timelike direction. As in the Pauli theory, these reference vectors multiplying ψ from the right do not break Lorentz covariance, as all observables are formed by rotating these reference vectors onto the frame of observables. Since $I\boldsymbol{\sigma}_3$ and γ_0 commute, our use of right-multiplication by $I\boldsymbol{\sigma}_3$ for the complex structure remains consistent.

The goal of our approach is to perform all calculations without ever having to introduce an explicit matrix representation. The explicit map of equation (8.69) is for column spinors written in the Dirac–Pauli representation, but it is a simple matter to establish similar maps for other representations. All one needs to do is find the unitary matrix which transforms the second representation into the Dirac–Pauli one, and then apply the map of equation (8.69). All of the matrix operators are then guaranteed to have the equivalence defined in equation (8.70). Certain other operations, such as complex conjugation, depend on the particular representation. But rather than think of these as the same operation in different representations, it is simpler to view them as different operations which can be applied to the multivector ψ.

In order to discuss the observables of the Dirac theory, we must first distinguish between the Hermitian and Dirac adjoints. The Hermitian adjoint is written as usual as $\langle\psi|$. The Dirac adjoint is written as $\langle\bar{\psi}|$ and is defined by

$$\langle\bar{\psi}| = (\langle\psi_u|, -\langle\psi_l|), \tag{8.71}$$

where the subscripts u and l refer to the upper and lower components. It is the Dirac adjoint which gives Lorentz-covariant observables. The Dirac inner product decomposes into

$$\langle\bar{\psi}|\phi\rangle = \langle\psi_u|\phi_u\rangle - \langle\psi_l|\phi_l\rangle. \tag{8.72}$$

This has the equivalent form

$$\langle\psi_u^\dagger\phi_u\rangle_q - \langle\psi_l^\dagger\phi_l\rangle_q = \langle(\psi_u^\dagger - \boldsymbol{\sigma}_3\psi_l^\dagger)(\phi_u + \phi_l\boldsymbol{\sigma}_3)\rangle_q = \langle\tilde{\psi}\phi\rangle_q. \tag{8.73}$$

So the Dirac adjoint is replaced by the manifestly covariant operation of spacetime reversion in the spacetime algebra formulation. The Hermitian adjoint now becomes

$$\langle\psi| \leftrightarrow \psi^\dagger = \gamma_0\tilde{\psi}\gamma_0, \tag{8.74}$$

which defines the meaning of the dagger symbol in the full spacetime algebra. Clearly, this operation requires singling out a preferred timelike vector, so is not covariant. In the relative space defined by γ_0, the Hermitian adjoint reduces to the non-relativistic reverse operation, so our notation is consistent with the use of the dagger for the reverse in three-dimensional space.

We can now look at the main observables formed from a Dirac spinor. The first is the current

$$J_\mu = \langle\bar{\psi}|\hat{\gamma}_\mu|\psi\rangle \leftrightarrow \langle\tilde{\psi}\gamma_\mu\psi\gamma_0\rangle - \langle\tilde{\psi}\gamma_\mu\psi I\gamma_3\rangle I\boldsymbol{\sigma}_3. \tag{8.75}$$

The final term contains $\langle\gamma_\mu\psi I\gamma_3\tilde{\psi}\rangle$. This vanishes because $\psi I\gamma_3\tilde{\psi}$ is odd-grade and reverses to minus itself, so is a pure trivector. Similarly, $\psi\gamma_0\tilde{\psi}$ is a pure vector, and we are left with

$$J_\mu = \langle\bar{\psi}|\hat{\gamma}_\mu|\psi\rangle \leftrightarrow \gamma_\mu\cdot(\psi\gamma_0\tilde{\psi}). \tag{8.76}$$

As with the Pauli theory, the operation of taking the expectation value of a matrix operator is replaced by that of picking out a component of a vector. We can therefore reconstitute the full vector J and write

$$J = \psi\gamma_0\tilde{\psi} \tag{8.77}$$

for the first of our observables.

To gain some further insight into the form of J, and its formation from ψ, we introduce the scalar + pseudoscalar quantity $\psi\tilde{\psi}$ as

$$\psi\tilde{\psi} = \rho\mathrm{e}^{I\beta}. \tag{8.78}$$

Factoring this out from ψ, we define the spacetime rotor R:

$$R = \psi\rho^{-1/2}\mathrm{e}^{-I\beta/2}, \qquad R\tilde{R} = 1. \tag{8.79}$$

(If $\rho = 0$ a slightly different procedure can be used.) We have now decomposed the spinor ψ into

$$\psi = \rho^{1/2}\mathrm{e}^{I\beta/2}R, \tag{8.80}$$

which separates out a density ρ and the rotor R. The remaining factor of β is curious. It turns out that plane-wave particle states have $\beta = 0$, whereas antiparticle states have $\beta = \pi$. The picture for bound state wavefunctions is more complicated, however, and β appears to act as a remnant of multiparticle

Bilinear covariant	Standard form	STA equivalent	Frame-free form
Scalar	$\langle\bar{\psi}\vert\psi\rangle$	$\langle\psi\tilde{\psi}\rangle$	$\rho\cos(\beta)$
Vector	$\langle\bar{\psi}\vert\hat{\gamma}_\mu\vert\psi\rangle$	$\gamma_\mu\cdot(\psi\gamma_0\tilde{\psi})$	$\psi\gamma_0\tilde{\psi}=J$
Bivector	$\langle\bar{\psi}\vert i\hat{\gamma}_{\mu\nu}\vert\psi\rangle$	$(\gamma_\mu\wedge\gamma_\nu)\cdot(\psi I\boldsymbol{\sigma}_3\tilde{\psi})$	$\psi I\boldsymbol{\sigma}_3\tilde{\psi}=S$
Pseudovector	$\langle\bar{\psi}\vert\hat{\gamma}_\mu\hat{\gamma}_5\vert\psi\rangle$	$\gamma_\mu\cdot(\psi\gamma_3\tilde{\psi})$	$\psi\gamma_3\tilde{\psi}=s$
Pseudoscalar	$\langle\bar{\psi}\vert i\hat{\gamma}_5\vert\psi\rangle$	$\langle\psi\tilde{\psi}I\rangle$	$-\rho\sin(\beta)$

Table 8.1 *Observables in the Dirac theory.* The standard expressions for the bilinear covariants are shown, together with their spacetime algebra (STA) equivalents.

effects from the full quantum field theory. With this decomposition of ψ, the current becomes

$$J=\psi\gamma_0\tilde{\psi}=\rho e^{I\beta/2}R\gamma_0\tilde{R}e^{I\beta/2}=\rho R\gamma_0\tilde{R}. \tag{8.81}$$

So the rotor is now an instruction to rotate γ_0 onto the direction of the current. This is precisely the picture we adopted in section 5.5 for studying the dynamics of a relativistic point particle.

A similar picture emerges for the spin. In relativistic mechanics angular momentum is a bivector quantity. Accordingly, the spin observables form a rank-2 antisymmetric tensor, with components given by

$$\langle\bar{\psi}\vert i\tfrac{1}{2}(\hat{\gamma}_\mu\hat{\gamma}_\nu-\hat{\gamma}_\nu\hat{\gamma}_\mu)\vert\psi\rangle\ \leftrightarrow\ \langle\tilde{\psi}\gamma_\mu\wedge\gamma_\nu\psi I\boldsymbol{\sigma}_3\rangle_q=\langle\gamma_\mu\wedge\gamma_\nu\psi I\boldsymbol{\sigma}_3\tilde{\psi}\rangle, \tag{8.82}$$

where again there is no imaginary component. This time we are picking out the components of the spin bivector S, given by

$$S=\psi I\boldsymbol{\sigma}_3\tilde{\psi}. \tag{8.83}$$

This is the natural spacetime generalisation of the Pauli result of equation (8.18). (Factors of $\hbar/2$ can always be inserted when required.) There are five such observables in all, which are summarised in Table 8.1. Of particular interest is the spin vector $s=\rho R\gamma_3\tilde{R}$. This justifies the classical model of spin introduced in section 5.5.6, where it was shown that the rotor form of the Lorentz force law naturally gives rise to a reduced gyromagnetic ratio of $g=2$.

8.3 The Dirac equation

While much of the preceding discussion is both suggestive about the role of spinors in quantum theory, and algebraically very useful, one has to remember that quantum mechanics deals with *wave equations*. We therefore need to

construct a relativistic wave equation for our Dirac spinor ψ, where ψ is an element of the eight-dimensional even subalgebra of the spacetime algebra. The relativistic wave equation for a spin-1/2 particle is the *Dirac equation*. This is a first-order wave equation, which is both Lorentz-invariant and has a future-pointing conserved current.

Like Pauli spinors, ψ is also subject to a single-sided rotor transformation law, $\psi \mapsto R\psi$, where R is a Lorentz rotor. To write down a covariant equation, we can therefore only place other covariant objects on the left of ψ. The available objects are any scalar or pseudoscalar, the vector derivative ∇ and any gauge fields describing interactions. On the right of ψ we can place combinations of γ_0, γ_3 and $I\boldsymbol{\sigma}_3$. The first equation we could write down is simply

$$\nabla\psi = 0. \tag{8.84}$$

This is the spacetime generalisation of the Cauchy–Riemann equations, as described in section 6.3. Remarkably, this equation does describe the behaviour of fermions — it is the wave equation for a (massless) *neutrino*. Any solution to this decomposes into two separate solutions by writing

$$\psi = \psi\tfrac{1}{2}(1 + \boldsymbol{\sigma}_3) + \psi\tfrac{1}{2}(1 - \boldsymbol{\sigma}_3) = \psi_+ + \psi_-. \tag{8.85}$$

The separate solutions ψ_+ and ψ_- are the right-handed and left-handed helicity eigenstates. For neutrinos, nature only appears to make use of the left-handed solutions. A more complete treatment of this subject involves the *electroweak* theory. (In fact, recent experiments point towards neutrinos carrying a small mass, whose origin can be explained by an interaction with the Higgs field.)

The formal operator identification of $i\partial_\mu$ with p_μ tells us that any wavefunction for a free massive particle should satisfy the Klein–Gordon equation $\nabla^2\psi = -m^2\psi$. We therefore need to add to the right-hand side of equation (8.84) a term that is linear in the particle mass m and that generates $-m^2\psi$ on squaring the operator. The natural covariant vector to form on the left of ψ is the momentum $\gamma^\mu p_\mu$. In terms of this operator we are led to an equation of the form

$$p\psi = m\psi a_0, \tag{8.86}$$

where a_0 is some multivector to be determined. It is immediately clear that a_0 must have odd grade, and must square to $+1$. The obvious candidate is γ_0, so that ψ contains a rotor to transform γ_0 to the velocity p/m. We therefore arrive at the equation

$$\nabla\psi I\boldsymbol{\sigma}_3 = m\psi\gamma_0. \tag{8.87}$$

This is the *Dirac equation* in its spacetime algebra form. This is easily seen to be equivalent to the matrix form of the equation

$$\hat{\gamma}^u\mu(\partial_\mu - ieA_\mu)|\psi\rangle = m|\psi\rangle, \tag{8.88}$$

where the electromagnetic vector potential has been included. The full Dirac equation is now

$$\nabla\psi I\boldsymbol{\sigma}_3 - eA\psi = m\psi\gamma_0. \tag{8.89}$$

A remarkable feature of this formulation is that the equation and all of its observables have been captured in the *real* algebra of spacetime, with no need for a unit imaginary. This suggests that interpretations of quantum mechanics that place great significance in the need for complex numbers are wide off the mark.

8.3.1 Symmetries and currents

The subject of the symmetries of the Dirac equation, and their conjugate currents, is discussed more fully in chapter 12. Here we highlight the main results. There are three important discrete symmetry operations: charge conjugation, parity and time reversal, denoted C, P and T respectively. Following the conventions of Bjorken & Drell (1964) we find that

$$\begin{aligned}
\hat{P}|\psi\rangle &\leftrightarrow \gamma_0\psi(\bar{x})\gamma_0, \\
\hat{C}|\psi\rangle &\leftrightarrow \psi\boldsymbol{\sigma}_1, \\
\hat{T}|\psi\rangle &\leftrightarrow I\gamma_0\psi(-\bar{x})\gamma_1,
\end{aligned} \tag{8.90}$$

where $\bar{x} = \gamma_0 x \gamma_0$ is (minus) the reflection of x in the timelike γ_0 axis. The combined CPT symmetry corresponds to

$$\psi \mapsto -I\psi(-x) \tag{8.91}$$

so that CPT symmetry does not require singling out a preferred timelike vector.

Amongst the continuous symmetries of the Dirac equation, the most significant is local electromagnetic gauge invariance. The equation is unchanged in physical content if we make the simultaneous replacements

$$\psi \mapsto \psi \mathrm{e}^{\alpha I\boldsymbol{\sigma}_3}, \qquad eA \mapsto eA - \nabla\alpha. \tag{8.92}$$

The conserved current conjugate to this symmetry is the Dirac current $J = \psi\gamma_0\tilde{\psi}$. This satisfies

$$\begin{aligned}
\nabla\cdot J &= \langle\nabla\psi\gamma_0\tilde{\psi}\rangle + \langle\psi\gamma_0\dot{\tilde{\psi}}\dot{\nabla}\rangle \\
&= -2\langle(eA\psi\gamma_0 + m\psi)I\boldsymbol{\sigma}_3\tilde{\psi}\rangle \\
&= 0
\end{aligned} \tag{8.93}$$

and so is conserved even in the presence of a background field. This is important. It means that single fermions cannot be created or destroyed. This feature was initially viewed as a great strength of the Dirac equation, though ultimately it is its biggest weakness. Fermion pairs, such as an electron and a positron, can be created and destroyed — a process which cannot be explained by the Dirac

equation alone. These are many-body problems and are described by *quantum field theory*.

The timelike component of J in the γ_0 frame, say, is

$$J_0 = \gamma_0 \cdot J = \langle \gamma_0 \tilde{\psi} \gamma_0 \psi \rangle = \langle \psi^\dagger \psi \rangle > 0, \tag{8.94}$$

which is *positive definite*. This is interpreted as a probability density, and localised wave functions are usually normalised such that

$$\int d^3x \, J_0 = 1. \tag{8.95}$$

Arriving at a relativistic theory with a consistent probabilistic interpretation was Dirac's original goal.

8.3.2 Plane-wave states

A positive energy plane-wave state is defined by

$$\psi = \psi_0 \mathrm{e}^{-I\boldsymbol{\sigma}_3 p \cdot x}, \tag{8.96}$$

where ψ_0 is a constant spinor. The Dirac equation (8.87) tells us that ψ_0 satisfies

$$p\psi_0 = m\psi_0\gamma_0, \tag{8.97}$$

and post-multiplying by $\tilde{\psi}_0$ we see that

$$p\psi_0\tilde{\psi}_0 = mJ. \tag{8.98}$$

Recalling that we have $\psi\tilde{\psi} = \rho \mathrm{e}^{i\beta}$, and noting that both p and J are vectors, we see that we must have $\exp(i\beta) = \pm 1$. For positive energy states the timelike component of p is positive, as is the timelike component of J, so we take the positive solution $\beta = 0$. It follows that ψ_0 is then simply a rotor with a normalisation constant. The proper boost L taking $m\gamma_0$ onto the momentum has

$$p = mL\gamma_0\tilde{L} = mL^2\gamma_0, \tag{8.99}$$

and from section 5.4.4 the solution is

$$L = \frac{m + p\gamma_0}{[2m(m + p\cdot\gamma_0)]^{1/2}} = \frac{E + m + \boldsymbol{p}}{[2m(E+m)]^{1/2}}, \tag{8.100}$$

where $p\gamma_0 = E + \boldsymbol{p}$. The full spinor ψ_0 is LU, where U is a spatial rotor in the γ_0 frame, so is a Pauli spinor.

Negative-energy solutions have a phase factor of $\exp(+I\boldsymbol{\sigma}_3 p \cdot x)$, with $E = \gamma_0 \cdot p > 0$. For these we have $-p\psi\tilde{\psi} = mJ$ so it is clear that we now need $\beta = \pi$. Positive and negative energy plane wave states can therefore be summarised by

$$\begin{aligned} \text{positive energy:} \quad & \psi^{(+)}(x) = L(p)U_r \mathrm{e}^{-I\boldsymbol{\sigma}_3 p\cdot x}, \\ \text{negative energy:} \quad & \psi^{(-)}(x) = L(p)U_r I \mathrm{e}^{I\boldsymbol{\sigma}_3 p\cdot x}, \end{aligned} \tag{8.101}$$

with $L(p)$ given by equation (8.100). The subscript r on the spatial rotors labels the spin state, with $U_0 = 1$, $U_1 = -I\boldsymbol{\sigma}_2$. These plane wave solutions are the fundamental components of *scattering theory.*

8.3.3 Hamiltonian form and the Pauli equation

The problem of how to best formulate operator techniques within spacetime algebra is little more than a question of finding a good notation. We could of course borrow the traditional Dirac 'bra-ket' notation, but we have already seen that the bilinear covariants are better handled without it. It is easier instead to just juxtapose the operator and the wavefunction on which it acts. But we saw in section 8.2 that the operators often act double-sidedly on the spinor ψ. This is not a problem, as the only permitted right-sided operations are multiplication by γ_0 or $I\boldsymbol{\sigma}_3$, and these operations commute. Our notation can therefore safely suppress these right-sided multiplications and gather all operations on the left. The overhat notation is useful to achieve this and we define

$$\hat{\gamma}_\mu \psi = \gamma_\mu \psi \gamma_0. \tag{8.102}$$

It should be borne in mind that all operations are now defined in the spacetime algebra, so the $\hat{\gamma}_\mu$ are not to be read as matrix operators, as they were in section 8.2. Of course, the action of the operators in either system is identical.

It is also useful to have a symbol for the operation of right-sided multiplication by $I\boldsymbol{\sigma}_3$. The symbol j carries the correct connotations of an operator that commutes with all others and squares to -1, and we define

$$j\psi = \psi I\boldsymbol{\sigma}_3. \tag{8.103}$$

The Dirac equation can now be written in the 'operator' form

$$j\hat{\nabla}\psi - e\hat{A}\psi = m\psi, \tag{8.104}$$

where

$$\hat{\nabla}\psi = \nabla\psi\gamma_0 \quad \text{and} \quad \hat{A}\psi = A\psi\gamma_0. \tag{8.105}$$

Writing the Dirac equation in the form (8.104) does not add anything new, but does confirm that we have an efficient notation for handling operators. One might ask why we have preferred the j symbol over the more obvious i. One reason is historical. In much of the spacetime algebra literature it has been common practice to denote the spacetime pseudoscalar with a small i. We now feel that this is a misleading notation, but it is commonplace. In addition, there are occasions when we may wish to formally complexify the spacetime algebra, as was the case for electromagnetic scattering, covered in section 7.5. To avoid confusion with either of these cases we have chosen to denote right-multiplication of ψ by $I\boldsymbol{\sigma}_3$ as $j\psi$ in both this and the following chapter.

To express the Dirac equation in Hamiltonian form we simply multiply from the left by γ_0. The resulting equation, with the dimensional constants temporarily put back in, is

$$j\hbar\partial_t\psi = c\hat{\boldsymbol{p}}\psi + eV\psi - ce\boldsymbol{A}\psi + mc^2\bar{\psi}, \tag{8.106}$$

where

$$\begin{aligned} \hat{\boldsymbol{p}}\psi &= -j\hbar\boldsymbol{\nabla}\psi, \\ \bar{\psi} &= \gamma_0\psi\gamma_0, \\ \gamma_0 A &= V - c\boldsymbol{A}. \end{aligned} \tag{8.107}$$

Choosing a Hamiltonian is a non-covariant operation, since it picks out a preferred timelike direction. The Hamiltonian relative to the γ_0 direction is the operator on the right-hand side of equation (8.106).

As an application of the Hamiltonian formulation, consider the non-relativistic reduction of the Dirac equation. This can be achieved formally via the *Foldy–Wouthuysen* transformation. For details we refer the reader to Itzykson & Zuber (1980). While the theoretical motivation for this transformation is clear, it can be hard to compute in all but the simplest cases. A simpler approach, dating back to Feynman, is to separate out the fast-oscillating component of the waves and then split into separate equations for the Pauli-even and Pauli-odd components of ψ. We write (with $\hbar = 1$ and the factors of c kept in)

$$\psi = (\phi + \eta)\mathrm{e}^{-I\boldsymbol{\sigma}_3 mc^2 t}, \tag{8.108}$$

where $\bar{\phi} = \phi$ (Pauli-even) and $\bar{\eta} = -\eta$ (Pauli-odd). The Dirac equation (8.106) now splits into the two equations

$$\begin{aligned} \mathcal{E}\phi - c\mathcal{O}\eta &= 0, \\ (\mathcal{E} + 2mc^2)\eta - c\mathcal{O}\phi &= 0, \end{aligned} \tag{8.109}$$

where

$$\begin{aligned} \mathcal{E}\phi &= (j\partial_t - eV)\phi, \\ \mathcal{O}\phi &= (\hat{\boldsymbol{p}} - e\boldsymbol{A})\phi. \end{aligned} \tag{8.110}$$

The formal solution to the second of equations (8.109) is

$$\eta = \frac{1}{2mc}\left(1 + \frac{\mathcal{E}}{2mc^2}\right)^{-1}\mathcal{O}\phi, \tag{8.111}$$

where the inverse on the right-hand side denotes a power series. Provided the expectation value of $\mathcal{E}$ is smaller than $2mc^2$ (which it is in the non-relativistic limit) the series should converge. The remaining equation for ϕ is

$$\mathcal{E}\phi - \frac{\mathcal{O}}{2m}\left(1 - \frac{\mathcal{E}}{2mc^2} + \cdots\right)\mathcal{O}\phi = 0, \tag{8.112}$$

which can be expanded out to the desired order of magnitude. There is little point in going beyond the first relativistic correction, so we approximate equation (8.112) by

$$\mathcal{E}\phi + \frac{\mathcal{O}\mathcal{E}\mathcal{O}}{4m^2c^2}\phi = \frac{\mathcal{O}^2}{2m}\phi. \tag{8.113}$$

We seek an equation of the form $\mathcal{E}\phi = \mathcal{H}\phi$, where $\mathcal{H}$ is the non-relativistic Hamiltonian. We therefore need to replace the $\mathcal{O}\mathcal{E}\mathcal{O}$ term in equation (8.113) by a term that does not involve $\mathcal{E}$. To achieve this we write

$$2\mathcal{O}\mathcal{E}\mathcal{O} = [\mathcal{O}, [\mathcal{E}, \mathcal{O}]] + \mathcal{E}\mathcal{O}^2 + \mathcal{O}^2\mathcal{E} \tag{8.114}$$

so that equation (8.113) becomes

$$\mathcal{E}\phi = \frac{\mathcal{O}^2}{2m}\phi - \frac{\mathcal{E}\mathcal{O}^2 + \mathcal{O}^2\mathcal{E}}{8m^2c^2}\phi - \frac{1}{8m^2c^2}[\mathcal{O}, [\mathcal{E}, \mathcal{O}]]\phi. \tag{8.115}$$

We can now make the approximation

$$\mathcal{E}\phi \approx \frac{\mathcal{O}^2}{2m}\phi, \tag{8.116}$$

so that equation (8.113) can be approximated by

$$\mathcal{E}\phi = \frac{\mathcal{O}^2}{2m}\phi - \frac{1}{8m^2c^2}[\mathcal{O}, [\mathcal{E}, \mathcal{O}]]\phi - \frac{\mathcal{O}^4}{8m^3c^2}\phi, \tag{8.117}$$

which is valid to order c^{-2}.

To evaluate the commutators we first need

$$[\mathcal{E}, \mathcal{O}] = -je(\partial_t \boldsymbol{A} + \boldsymbol{\nabla} V) = je\boldsymbol{E}. \tag{8.118}$$

There are no time derivatives left in this commutator, so we do achieve a sensible non-relativistic Hamiltonian. The full commutator required in equation (8.117) is

$$\begin{aligned}[\mathcal{O}, [\mathcal{E}, \mathcal{O}]] &= [-j\boldsymbol{\nabla} - e\boldsymbol{A}, je\boldsymbol{E}] \\ &= (e\boldsymbol{\nabla}\boldsymbol{E}) - 2e\boldsymbol{E}\wedge\boldsymbol{\nabla} - 2je^2\boldsymbol{A}\wedge\boldsymbol{E}.\end{aligned} \tag{8.119}$$

The various operators (8.110) and (8.119) can now be substituted into equation (8.117) to yield the Pauli equation

$$\begin{aligned}\frac{\partial\phi}{\partial t}I\boldsymbol{\sigma}_3 = \frac{1}{2m}(\hat{\boldsymbol{p}} - e\boldsymbol{A})^2\phi + eV\phi - \frac{\hat{\boldsymbol{p}}^4}{8m^3c^2}\phi \\ - \frac{1}{8m^2c^2}\big(e(\boldsymbol{\nabla}\boldsymbol{E} - 2\boldsymbol{E}\wedge\boldsymbol{\nabla})\phi - 2e^2\boldsymbol{A}\wedge\boldsymbol{E}\phi I\boldsymbol{\sigma}_3\big),\end{aligned} \tag{8.120}$$

which is written entirely in the geometric algebra of three-dimensional space. In the standard approach, the geometric product in the $\boldsymbol{\nabla}\boldsymbol{E}$ term of equation (8.120) is split into a 'spin-orbit' term $\boldsymbol{\nabla}\wedge\boldsymbol{E}$ and the 'Darwin' term $\boldsymbol{\nabla}\cdot\boldsymbol{E}$.

The spacetime algebra approach reveals that these terms arise from a single source.

A similar approximation scheme can be adopted for the observables of the Dirac theory. For example the current, $\psi\gamma_0\tilde{\psi}$, has a three-vector part:

$$\boldsymbol{J} = (\psi\gamma_0\tilde{\psi})\wedge\gamma_0 = \phi\eta^\dagger + \eta\phi^\dagger. \tag{8.121}$$

This is approximated to leading order by

$$\boldsymbol{J} \approx -\frac{1}{m}(\langle\boldsymbol{\nabla}\phi I\boldsymbol{\sigma}_3\phi^\dagger\rangle_1 - \boldsymbol{A}\phi\phi^\dagger), \tag{8.122}$$

where the $\langle\,\rangle_1$ projects onto the grade-1 components of the Pauli algebra. Not all applications of the Pauli theory correctly identify (8.122) as the conserved current in the Pauli theory — an inconsistency first pointed out by Hestenes & Gurtler (1971).

8.4 Central potentials

Suppose now that we restrict our discussion to problems described by a central potential $V = V(r)$, $\boldsymbol{A} = 0$, where $r = |\boldsymbol{x}|$. The full Hamiltonian, denoted $\mathcal{H}$, reduces to

$$j\hbar\partial_t\psi = \mathcal{H}\psi = -j\boldsymbol{\nabla}\psi + eV(r)\psi + m\bar{\psi}. \tag{8.123}$$

Quantum states are classified in terms of eigenstates of operators that commute with the Hamiltonian $\mathcal{H}$, because the accompanying quantum numbers are conserved in time. Of particular importance are the angular-momentum operators $\hat{L}_i$, defined by

$$\hat{L}_i = -i\epsilon_{ijk}x_j\partial_k. \tag{8.124}$$

These are the components of the bivector operator $i\boldsymbol{x}\wedge\boldsymbol{\nabla}$. We therefore define the operators

$$L_B = jB\cdot(\boldsymbol{x}\wedge\boldsymbol{\nabla}), \tag{8.125}$$

where B is a relative bivector. Throughout this section interior and exterior products refer to the (Pauli) algebra of space. Writing $B = I\boldsymbol{\sigma}_i$ recovers the component form. The L_B operators satisfy the commutation relations

$$[L_{B_1}, L_{B_2}] = -jL_{B_1\times B_2}, \tag{8.126}$$

where $B_1\times B_2$ denotes the commutator product. The angular-momentum commutation relations directly encode the bivector commutation relations, which are those of the Lie algebra of the rotation group (see chapter 11). One naturally expects this group to arise as it represents a symmetry of the potential.

If we now form the commutator of L_B with the Hamiltonian $\mathcal{H}$ we obtain a

result that is, initially, disconcerting. The scalar operator L_B commutes with the bar operator $\psi \mapsto \bar{\psi}$, but for the momentum term we find that

$$[B\cdot(\boldsymbol{x}\wedge\boldsymbol{\nabla}),\boldsymbol{\nabla}] = -\dot{\boldsymbol{\nabla}}B\cdot(\dot{\boldsymbol{x}}\wedge\boldsymbol{\nabla}) = B\times\boldsymbol{\nabla}. \tag{8.127}$$

The commutator does not vanish, so orbital angular momentum does not yield a conserved quantum number in relativistic physics. But, since $B\times\boldsymbol{\nabla} = \frac{1}{2}(B\boldsymbol{\nabla} - \boldsymbol{\nabla}B)$, we can write equation (8.127) as

$$[B\cdot(\boldsymbol{x}\wedge\boldsymbol{\nabla}) - \tfrac{1}{2}B, \mathcal{H}] = 0. \tag{8.128}$$

We therefore recover a conserved angular momentum operator by defining

$$J_B = L_B - \tfrac{1}{2}jB. \tag{8.129}$$

In conventional notation this is

$$\hat{J}_i = \hat{L}_i + \tfrac{1}{2}\hat{\Sigma}_i, \tag{8.130}$$

where $\hat{\Sigma}_i = (i/2)\epsilon_{ijk}\hat{\gamma}_j\hat{\gamma}_k$. The extra term of $B/2$ accounts for the spin-1/2 nature of Dirac particles. If we look for eigenstates of the J_3 operator, we see that the spin contribution to this is

$$-\tfrac{1}{2}jI\boldsymbol{\sigma}_3\psi = \tfrac{1}{2}\boldsymbol{\sigma}_3\psi\boldsymbol{\sigma}_3. \tag{8.131}$$

In the non-relativistic Pauli theory the eigenstates of this operator are simply 1 and $-I\boldsymbol{\sigma}_2$, with eigenvalues $\pm 1/2$. In the relativistic theory the separate spin and orbital operators are not conserved, and it is only the combined J_B operators that commute with the Hamiltonian.

The geometric algebra derivation employed here highlights some interesting features. Stripping away all of the extraneous terms, the result rests solely on the commutation properties of the $B\cdot(\boldsymbol{x}\wedge\boldsymbol{\nabla})$ and $\boldsymbol{\nabla}$ operators. The factor of $1/2$ would therefore be present in any dimension, and so has no special relation to the three-dimensional rotation group. Furthermore, in writing $J_B = L_B - \frac{1}{2}jB$ we are forming an explicit sum of a scalar and a bivector. The standard notation of equation (8.130) encourages us to view these as the sum of two vector operators!

8.4.1 Spherical monogenics

The spherical monogenics play a key role in the solution of the Dirac equation for problems with radial symmetry. These are Pauli spinors (even elements of the Pauli algebra) that satisfy the eigenvalue equation

$$-\boldsymbol{x}\wedge\boldsymbol{\nabla}\psi = l\psi. \tag{8.132}$$

These functions arise naturally as solutions of the three-dimensional generalisation of the Cauchy–Riemann equations

$$\boldsymbol{\nabla}\Psi = 0. \tag{8.133}$$

Solutions of this equation are known in the Clifford analysis literature as monogenics. Looking for solutions which separate into $\Psi = r^l\psi(\theta,\phi)$ yields equation (8.132), where (r,θ,ϕ) is a standard set of polar coordinates. The solutions of equation (8.132) are called spherical monogenics, or *spin-weighted spherical harmonics* (with weight 1/2).

To analyse the properties of equation (8.132) we first note that

$$[J_B, \boldsymbol{x}\wedge\boldsymbol{\nabla}] = 0, \tag{8.134}$$

which is proved in the same manner as equation (8.128). It follows that ψ can simultaneously be an eigenstate of the $\boldsymbol{x}\wedge\boldsymbol{\nabla}$ operator and one of the J_B operators. To simplify the notation we now define

$$J_k\psi = J_{I\sigma_k}\psi = \big((I\sigma_k)\cdot(\boldsymbol{x}\wedge\boldsymbol{\nabla}) - \tfrac{1}{2}I\sigma_k\big)\psi I\boldsymbol{\sigma}_3. \tag{8.135}$$

We choose ψ to be an eigenstate of J_3. We label this state as $\psi(l,\mu)$, so

$$-\boldsymbol{x}\wedge\boldsymbol{\nabla}\psi(l,\mu) = l\psi(l,\mu), \qquad J_3\psi(l,\mu) = \mu\psi(l,\mu). \tag{8.136}$$

The J_i operators satisfy

$$\begin{aligned} J_iJ_i\psi(l,\mu) &= 3/4\psi - 2\boldsymbol{x}\wedge\boldsymbol{\nabla}\psi + \boldsymbol{x}\wedge\boldsymbol{\nabla}(\boldsymbol{x}\wedge\boldsymbol{\nabla}\psi) \\ &= (l+1/2)(l+3/2)\psi(l,\mu), \end{aligned} \tag{8.137}$$

so the $\psi(l,\mu)$ are also eigenstates of J_iJ_i.

We next introduce the ladder operators J_+ and J_-, defined by

$$\begin{aligned} J_+ &= J_1 + jJ_2, \\ J_- &= J_1 - jJ_2. \end{aligned} \tag{8.138}$$

It is a simple matter to prove the following results:

$$\begin{aligned} [J_+, J_-] &= 2J_3, \qquad & J_iJ_i &= J_-J_+ + J_3 + {J_3}^2, \\ [J_\pm, J_3] &= \mp J_\pm, \qquad & J_iJ_i &= J_+J_- - J_3 + {J_3}^2. \end{aligned} \tag{8.139}$$

The raising operator J_+ increases the eigenvalue of J_3 by an integer. But, for fixed l, μ must ultimately attain some maximum value. Denoting this value as μ_+, we must reach a state for which

$$J_+\psi(l,\mu_+) = 0. \tag{8.140}$$

Acting on this state with J_iJ_i and using one of the results in equation (8.139) we find that

$$(l+1/2)(l+3/2) = \mu_+(\mu_+ + 1). \tag{8.141}$$

Since l is positive and μ_+ represents an upper bound, it follows that

$$\mu_+ = l + 1/2. \tag{8.142}$$

There must similarly be a lowest eigenvalue of J_3 and a corresponding state with

$$J_-\psi(l,\mu_-) = 0. \tag{8.143}$$

In this case we find that

$$(l+1/2)(l+3/2) = \mu_-(\mu_- - 1), \tag{8.144}$$

hence $\mu_- = -(l+1/2)$. The spectrum of eigenvalues of J_3 therefore ranges from $(l+1/2)$ to $-(l+1/2)$, a total of $2(l+1)$ states. Since the J_3 eigenvalues are always of the form (integer $+1/2$), it is simpler to label the spherical monogenics with a pair of integers. We therefore write the spherical monogenics as ψ_l^m, where

$$-\boldsymbol{x}\wedge\boldsymbol{\nabla}\psi_l^m = l\psi_l^m \qquad l \geq 0 \tag{8.145}$$

and

$$J_3\psi_l^m = (m+\tfrac{1}{2})\psi_l^m \qquad -1-l \leq m \leq l. \tag{8.146}$$

To find an explicit form for the ψ_l^m we first construct the highest m case. This satisfies

$$J_+\psi_l^l = 0 \tag{8.147}$$

and it is not hard to see that this equation is solved by

$$\psi_l^l \propto \sin^l(\theta)\,\mathrm{e}^{-l\phi I\boldsymbol{\sigma}_3}. \tag{8.148}$$

This is the angular part of the monogenic function $(x+yI\boldsymbol{\sigma}_3)^l$. Introducing a convenient factor, we write

$$\psi_l^l = (2l+1)P_l^l(\cos(\theta))\,\mathrm{e}^{l\phi I\boldsymbol{\sigma}_3}. \tag{8.149}$$

Our convention for the associated Legendre polynomials follows Gradshteyn & Ryzhik (1994), so we have

$$P_l^m(x) = \frac{(-1)^m}{2^l l!}(1-x^2)^{m/2}\frac{d^{l+m}}{dx^{l+m}}(x^2-1)^l. \tag{8.150}$$

(Some useful recursion relations for the associated Legendre polynomials are discussed in the exercises.) The lowering operator J_- has the following effect on ψ:

$$J_-\psi = \big(-\partial_\theta\psi + \cot(\theta)\,\partial_\phi\psi I\boldsymbol{\sigma}_3\big)\mathrm{e}^{-\phi I\boldsymbol{\sigma}_3} - I\boldsymbol{\sigma}_2\,\tfrac{1}{2}(\psi + \boldsymbol{\sigma}_3\psi\boldsymbol{\sigma}_3). \tag{8.151}$$

The final term just projects out the $\{1, I\boldsymbol{\sigma}_3\}$ terms and multiplies them by $-I\boldsymbol{\sigma}_2$. This is the analog of the lowering matrix in the standard formalism. The derivatives acting on ψ_l^l form

$$\big(-\partial_\theta\psi_l^l + \cot(\theta)\,\partial_\phi\psi_l^l I\boldsymbol{\sigma}_3\big)\mathrm{e}^{-\phi I\boldsymbol{\sigma}_3} = (2l+1)2lP_l^{l-1}(\cos(\theta))\mathrm{e}^{(l-1)\phi I\boldsymbol{\sigma}_3}, \tag{8.152}$$

and, if we use the result that

$$\boldsymbol{\sigma}_\phi = \boldsymbol{\sigma}_2 \mathrm{e}^{\phi I \boldsymbol{\sigma}_3}, \tag{8.153}$$

we find that

$$\psi_l^{l-1} \propto \big(2lP_l^{l-1}(\cos(\theta)) - P_l^l(\cos(\theta)) I\boldsymbol{\sigma}_\phi\big)\mathrm{e}^{(l-1)\phi I\boldsymbol{\sigma}_3}. \tag{8.154}$$

Proceeding in this manner, we are led to the following formula for the spherical monogenics:

$$\psi_l^m = \big((l+m+1)P_l^m(\cos(\theta)) - P_l^{m+1}(\cos(\theta)) I\boldsymbol{\sigma}_\phi\big)\mathrm{e}^{m\phi I\boldsymbol{\sigma}_3}, \tag{8.155}$$

in which l is a positive integer or zero, m ranges from $-(l+1)$ to l and the P_l^m are taken to be zero if $|m| > l$. The positive- and negative-m states are related by

$$P_l^{-m}(x) = (-1)^m \frac{(l-m)!}{(l+m)!} P_l^m(x), \tag{8.156}$$

from which it can be shown that

$$\psi_l^m(-I\boldsymbol{\sigma}_2) = (-1)^m \frac{(l+m+1)!}{(l-m)!} \psi_l^{-(m+1)}. \tag{8.157}$$

The spherical monogenics presented here are unnormalised. Normalisation factors are not hard to compute, and we find that

$$\int_0^\pi d\theta \int_0^{2\pi} d\phi \, \sin(\theta)\, \psi_l^m \psi_l^{m\,\dagger} = 4\pi \frac{(l+m+1)!}{(l-m)!}. \tag{8.158}$$

If $\boldsymbol{\sigma}_r$ denotes the unit radial vector, $\boldsymbol{\sigma}_r = \boldsymbol{x}/r$ we find that

$$\boldsymbol{x} \wedge \boldsymbol{\nabla} \boldsymbol{\sigma}_r = 2\boldsymbol{\sigma}_r. \tag{8.159}$$

It follows that

$$-\boldsymbol{x} \wedge \boldsymbol{\nabla}(\boldsymbol{\sigma}_r \psi \boldsymbol{\sigma}_3) = -(l+2)\boldsymbol{\sigma}_r \psi \boldsymbol{\sigma}_3, \tag{8.160}$$

which provides an equation for the negative-l eigenstates. The possible eigenvalues and degeneracies are summarised in Table 8.2. One curious feature of this table is that we appear to be missing a line for the eigenvalue $l = -1$. In fact solutions for this case do exist, but they contain singularities which render them unnormalisable. For example, the functions

$$\frac{I\sigma_\phi}{\sin(\theta)}, \quad \text{and} \quad \frac{\mathrm{e}^{-I\boldsymbol{\sigma}_3\phi}}{\sin(\theta)} \tag{8.161}$$

have $l = -1$ and J_3 eigenvalues $+1/2$ and $-1/2$ respectively. Both solutions are singular along the z axis, however, which limits their physical relevance.

l	Eigenvalues of J_3	Degeneracy
⋮	⋮	⋮
2	$5/2, \ldots, -5/2$	6
1	$3/2, \ldots, -3/2$	4
0	$1/2, \ldots, -1/2$	2
(-1)	?	?
-2	$1/2, \ldots, -1/2$	2
⋮	⋮	⋮

Table 8.2 *Eigenvalues and degeneracies for the ψ_l^m monogenics.*

8.4.2 The radial equations

We can use the angular monogenics to construct eigenfunctions of the Dirac Hamiltonian of equation (8.123). Since the J_B operators commute with $\mathcal{H}$, ψ can be placed in an eigenstate of J_3. The operator J_iJ_i must also commute with $\mathcal{H}$, so $(l+1/2)(l+3/2)$ is a good quantum number. The operator $\boldsymbol{x}\wedge\boldsymbol{\nabla}$ does not commute with $\mathcal{H}$, however, so both the ψ_l^m and $\boldsymbol{\sigma}_r\psi_l^m\boldsymbol{\sigma}_3$ monogenics are needed in the solution. While $\boldsymbol{x}\wedge\boldsymbol{\nabla}$ does not commute with $\mathcal{H}$, the operator

$$\hat{K} = \hat{\gamma}_0(1 - \boldsymbol{x}\wedge\boldsymbol{\nabla}) \tag{8.162}$$

does, as follows from

$$[\hat{\gamma}_0(1 - \boldsymbol{x}\wedge\boldsymbol{\nabla}), \boldsymbol{\nabla}] = 2\hat{\gamma}_0\boldsymbol{\nabla} - \hat{\gamma}_0\dot{\boldsymbol{\nabla}}\dot{\boldsymbol{x}}\wedge\boldsymbol{\nabla} = 0. \tag{8.163}$$

We should therefore work with eigenstates of the $\hat{K}$ operator. This implies that $\psi(\boldsymbol{x})$ can be written for positive l as either

$$\psi(\boldsymbol{x}, l+1) = \psi_l^m u(r) + \boldsymbol{\sigma}_r\psi_l^m v(r) I\boldsymbol{\sigma}_3 \tag{8.164}$$

or

$$\psi(\boldsymbol{x}, -(l+1)) = \boldsymbol{\sigma}_r\psi_l^m\boldsymbol{\sigma}_3 u(r) + \psi_l^m I v(r). \tag{8.165}$$

In both cases the second label in $\psi(\boldsymbol{x}, l+1)$ specifies the eigenvalue of $\hat{K}$. It is useful to denote this by κ, so we have

$$\hat{K}\psi = \kappa\psi, \qquad \kappa = \ldots, -2, -1, 1, 2, \ldots \tag{8.166}$$

and κ is a non-zero positive or negative integer.

In equations (8.164) and (8.165) the radial functions $u(r)$ and $v(r)$ are 'complex' combinations of 1 and $I\boldsymbol{\sigma}_3$. In the case of the Hamiltonian of (8.123), with $V(r)$ real, it turns out that the real and imaginary equations decouple, and it is

sufficient to treat $u(r)$ and $v(r)$ as real, scalar quantities. On substituting our trial functions into the Hamiltonian, we find that the radial equations reduce to

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} (\kappa-1)/r & -(E-eV(r)+m) \\ E-eV(r)-m & (-\kappa-1)/r \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}. \tag{8.167}$$

The same equation holds for all values of κ. This successfully separates the Dirac equation in any radially-symmetric potential. As one might expect, we arrive at a pair of coupled first-order equations, as opposed to the single second-order equation familiar from Schrödinger theory.

8.4.3 The hydrogen atom

The radial equations describing the relativistic quantum theory of the hydrogen atom are obtained simply by setting $eV = -Z\alpha/r$, where $\alpha = e^2/4\pi$ is the fine structure constant and Z is the atomic charge. The solution of the radial equations is described in most textbooks on relativistic quantum mechanics. The conclusion is that the radial dependence is governed by a pair of *hypergeometric functions,* which generalise the Laguerre polynomials of the non-relativistic theory. Rather than reproduce the analysis here, we instead present a more direct method of solving the equations, first given by Eddington (1936) in his unconventional *Relativity Theory of Protons and Electrons.*

We start with the equation

$$-j\boldsymbol{\nabla}\psi - \frac{Z\alpha}{r}\psi + m\hat{\gamma}_0\psi = E\psi. \tag{8.168}$$

We assume that ψ is in an eigenstate of $\hat{K}$, so we can write

$$\boldsymbol{x}\wedge\boldsymbol{\nabla}\psi = \psi - \kappa\hat{\gamma}_0\psi. \tag{8.169}$$

We now pre-multiply the Dirac equation by $j\boldsymbol{x}$ and rearrange to find

$$r\partial_r\psi + \psi - \kappa\hat{\gamma}_0\psi = j\boldsymbol{x}\left(E + \frac{Z\alpha}{r}\right)\psi - jm\boldsymbol{x}\hat{\gamma}_0\psi. \tag{8.170}$$

On introducing the reduced function $\Psi = r\psi$ the equation simplifies to

$$\partial_r\Psi = j\boldsymbol{\sigma}_r(E - m\hat{\gamma}_0)\Psi + \frac{1}{r}(jZ\alpha\boldsymbol{\sigma}_r + \kappa\hat{\gamma}_0)\Psi. \tag{8.171}$$

We accordingly define the two operators

$$\hat{F} = -j\boldsymbol{\sigma}_r(E - m\hat{\gamma}_0), \qquad \hat{G} = -(jZ\alpha\boldsymbol{\sigma}_r + \kappa\hat{\gamma}_0), \tag{8.172}$$

so that the Dirac equation reduces to

$$\partial_r\Psi + \left(\hat{F} + \frac{\hat{G}}{r}\right)\Psi = 0. \tag{8.173}$$

The $\hat{F}$ and $\hat{G}$ operators satisfy

$$\begin{aligned}\hat{F}^2 &= m^2 - E^2 = f^2,\\ \hat{G}^2 &= \kappa^2 - (Z\alpha)^2 = \nu^2,\end{aligned} \tag{8.174}$$

which define f and ν. The operators also satisfy the anticommutation relation

$$\hat{F}\hat{G} + \hat{G}\hat{F} = -2Z\alpha E. \tag{8.175}$$

The next step is to transform to the dimensionless variable $x = fr$ and remove the large-x behaviour by setting

$$\Psi = \Phi \mathrm{e}^{-x}. \tag{8.176}$$

The function Φ now satisfies

$$\partial_x \Phi + \frac{\hat{G}}{x}\Phi + \Big(\frac{\hat{F}}{f} - 1\Big)\Phi = 0. \tag{8.177}$$

We are now in a position to consider a power series solution, so we set

$$\Phi = x^s \sum_{n=0} C_n x^n, \tag{8.178}$$

where the C_n are all multivectors. (In Eddington's original notation these are his 'e-numbers'.) The recursion relation is first-order and is given simply by

$$(n + s + \hat{G})C_n = -\left(\frac{\hat{F}}{f} - 1\right) C_{n-1}. \tag{8.179}$$

Setting $n = 0$ we see that

$$\big(s + \hat{G}\big)C_0 = 0. \tag{8.180}$$

Acting on this equation with the operator $\big(s - \hat{G}\big)$ we see that we must have $s^2 = \hat{G}^2 = \nu^2$. We set $s = \nu$ in order that the wavefunction is well behaved at the origin.

With the small and large x behaviour now separated out, all that remains is the power series. One can show that, in order for ψ to fall to zero at large distances, the series must terminate. We therefore set $C_{n+1} = 0$, and it follows that

$$\left(\frac{\hat{F}}{f} - 1\right) C_n = 0, \quad \text{or} \quad \hat{F}C_n = fC_n. \tag{8.181}$$

But we also have

$$\left(\frac{\hat{F}}{f} + 1\right)(n + \nu + \hat{G})C_n = -\left(\frac{\hat{F}}{f} + 1\right)\left(\frac{\hat{F}}{f} - 1\right) C_{n-1} = 0, \tag{8.182}$$

so

$$\left(2(n+\nu)+\hat{G}+\frac{\hat{F}}{f}\hat{G}\right)C_n=0. \tag{8.183}$$

If we write this as

$$\left(2(n+\nu)+\frac{1}{f}(\hat{G}\hat{F}+\hat{F}\hat{G})\right)C_n=0, \tag{8.184}$$

we find that we must have

$$n+\nu-\frac{Z\alpha E}{f}=0. \tag{8.185}$$

This is precisely our energy quantisation condition. The equation is equivalent to

$$\frac{E}{(m^2-E^2)^{1/2}}=\frac{n+\nu}{Z\alpha}, \tag{8.186}$$

which rearranges to the standard formula

$$E^2=m^2\left(1-\frac{(Z\alpha)^2}{n^2+2n\nu+\kappa^2}\right), \tag{8.187}$$

where n is a non-negative integer.

The non-relativistic formula for the energy levels is recovered by first recalling that $\alpha\approx 1/137$ is small. We can therefore approximate to

$$\nu\approx|\kappa|=l+1, \tag{8.188}$$

where $l\geq 0$ and

$$E\approx m\left(1-\frac{(Z\alpha)^2}{2}\frac{1}{n^2+2n(l+1)+(l+1)^2}\right). \tag{8.189}$$

Subtracting off the rest mass energy we are left with the non-relativistic expression

$$E_{NR}=-m\frac{(Z\alpha)^2}{2}\frac{1}{(n+l+1)^2}=-\frac{mZ^2e^4}{32\pi^2\epsilon_0^2\hbar^2}\frac{1}{n'^2}, \tag{8.190}$$

where $n'=n+l+1$ and the dimensional constants have been reinserted. We have recovered the familiar Bohr formula for the energy levels. This derivation shows that the relativistic quantum number n differs from the Bohr quantum number n'.

Expanding to next order we find that

$$E_{NR}=-m\frac{(Z\alpha)^2}{2n'^2}-m\frac{(Z\alpha)^4}{2n'^4}\left(\frac{n'}{l+1}-\frac{3}{4}\right). \tag{8.191}$$

The first relativistic correction shows that the binding energy is increased slightly from the non-relativistic value, and also introduces some dependence on the

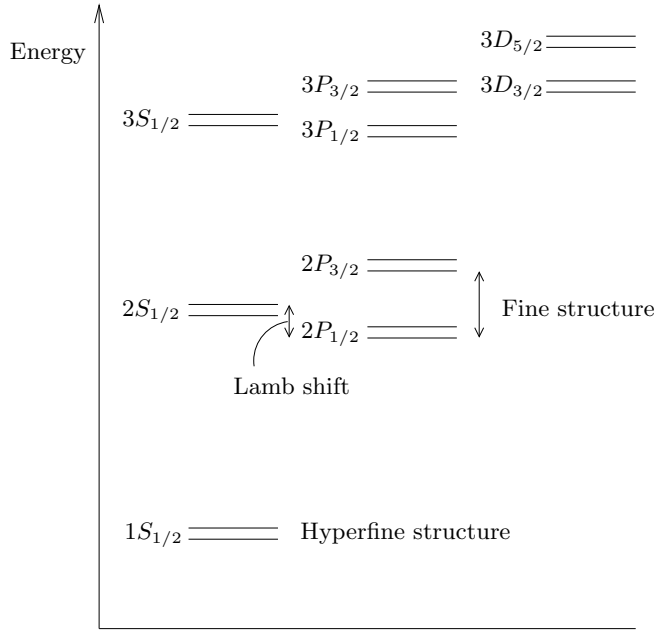

Figure 8.3 *Hydrogen atom energy levels.* The diagram illustrates how various degeneracies are broken by relativistic and spin effects. The Dirac equation accounts for the fine structure. The hyperfine structure is due to interaction with the magnetic moment of the nucleus. The Lamb shift is explained by quantum field theory. It lifts the degeneracy between the $S_{1/2}$ and $P_{1/2}$ states.

angular quantum number l. This lifts some degeneracies present in the non-relativistic solution. The various corrections contributing to the energy levels are shown in figure 8.3. A more complete analysis also requires replacing the electron mass m by the reduced mass of the two-body system. This introduces corrections of the same order of the relativistic corrections, but only affects the overall scale.

8.5 Scattering theory

Many of the experimental tests of Dirac theory, and quantum electrodynamics in general, are based on the results of scattering. Here we see how our new formulation can help to simplify these calculations through its handling of spin.

To aid this analysis it is useful to introduce the energy projection operators

$$\Lambda_{\pm}\psi = \frac{1}{2m}(m\psi \pm p\psi\gamma_0), \tag{8.192}$$

which project onto particle and antiparticle states.

A key role in relativistic quantum theory is played by Feynman propagators, which provide a means of imposing causal boundary conditions. We start by replacing the Dirac equation with the integral equation

$$\psi(x) = \psi_i(x) + e\int d^4x'\, S_F(x-x')A(x')\psi(x')\gamma_0, \tag{8.193}$$

where $\psi_i(x)$ is the asymptotic in-state and solves the free-particle equation, and $S_F(x-x')$ is the propagator. Substituting (8.193) into the Dirac equation, we find that $S_F(x-x')$ must satisfy

$$j\nabla_x S_F(x-x')\psi(x')\gamma_0 - mS_F(x-x')\psi(x') = \delta^4(x-x')\psi(x'). \tag{8.194}$$

The solution to this equation is

$$S_F(x-x')\psi(x') = \int \frac{d^4p}{(2\pi)^4}\,\frac{p\psi(x')\gamma_0 + m\psi(x')}{p^2 - m^2 + j\epsilon}\mathrm{e}^{-jp\cdot(x-x')}. \tag{8.195}$$

The factor of $j\epsilon$ is a mnemonic device to tell us how to negotiate the poles in the complex energy integral, which is performed first. The factor ensures positive-frequency waves propagate into the future ($t > t'$) and negative-frequency waves propagate into the past ($t' > t$). The result of performing the energy integration is summarised in the expression

$$S_F(x) = -2mj\int \frac{d^3\boldsymbol{p}}{(2\pi)^3}\frac{1}{2E}\Big(\theta(t)\Lambda_+\mathrm{e}^{-jp\cdot x} + \theta(-t)\Lambda_-\mathrm{e}^{jp\cdot x}\Big), \tag{8.196}$$

where $E = +\sqrt{\boldsymbol{p}^2 + m^2}$.

There are other choices of relativistic propagator, which may be appropriate in other settings. For classical electromagnetism, for example, it is necessary to work with retarded propagators. If one constructs a closed spacetime surface integral, with boundary conditions consistent with the field equations, then the choice of propagator is irrelevant, since they all differ by a spacetime monogenic function. In most applications, however, we do not work like this. Instead we work with initial data, which we seek to propagate to a later time in such a way that the final result is consistent with imposing causal boundary conditions. In this case one has to use the Feynman propagator for quantum fields.

8.5.1 Electron scattering

In scattering calculations we write the wavefunction as the sum of an incoming plane wave and a scattered beam,

$$\psi(x) = \psi_i(x) + \psi_{diff}(x). \tag{8.197}$$

At asymptotically large times ψ_{diff} is given by

$$\psi_{diff}(x) = -2mje \int d^4x' \int \frac{d^3\boldsymbol{p}}{(2\pi)^3}\frac{1}{2E}\Lambda_+\big(A(x')\psi(x')\gamma_0\big)\mathrm{e}^{-jp\cdot(x-x')}. \tag{8.198}$$

This can be written as a sum over final states

$$\psi_{diff}(x) = \int \frac{d^3\boldsymbol{p}_f}{(2\pi)^3}\frac{1}{2E_f}\psi_f \mathrm{e}^{-jp_f\cdot x}, \tag{8.199}$$

where the final states are plane waves with

$$\psi_f = -je\int d^4x'\big(p_f A(x')\psi(x') + mA(x')\psi(x')\gamma_0\big)\mathrm{e}^{jp_f\cdot x'}. \tag{8.200}$$

The number of scattered particles is given by (recalling that $J = \psi\gamma_0\tilde{\psi}$)

$$\int d^3\boldsymbol{x}\,\gamma_0\cdot J_{diff} = \int \frac{d^3\boldsymbol{p}_f}{(2\pi)^3}\frac{1}{2E_f}\left(\frac{\gamma_0\cdot J_f}{2E_f}\right) = \int \frac{d^3\boldsymbol{p}_f}{(2\pi)^3}\frac{1}{2E_f}N_f, \tag{8.201}$$

where N_f is the number density per Lorentz-invariant phase space interval:

$$N_f = \frac{\gamma_0\cdot J_f}{2E_f} = \frac{\gamma_0\cdot(\psi_f\gamma_0\tilde{\psi}_f)}{2E_f} = \frac{\rho_f}{2m}. \tag{8.202}$$

The integral equation (8.193) is the basis for a perturbative approach to solving the Dirac equation in an external field. We seek the full propagator S_A which satisfies

$$\big(j\nabla_2 - eA(x_2)\big)S_A(x_2,x_1)\gamma_0 - mS_A(x_2,x_1) = \delta^4(x_2-x_1). \tag{8.203}$$

The iterative solution to this is provided by

$$\begin{aligned} S_A(x_f,x_i) &= S_F(x_f-x_i) + \int d^4x_1\, S_F(x_f-x_1)e\hat{A}(x_1)S_F(x_1-x_i) \\ &+ \iint d^4x_1\, d^4x_2\, S_F(x_f-x_1)e\hat{A}(x_1)S_F(x_1-x_2)e\hat{A}(x_2)S_F(x_2-x_i) + \cdots, \end{aligned} \tag{8.204}$$

which is the basis for a diagrammatic representation of a scattering calculation. In the Born approximation we work to first order and truncate the series for S_A

after the first interaction term. Assuming incident plane waves of momentum p_i, so that $\psi_i(x) = \psi_i \exp(-jp_i \cdot x)$, we find that the final states become

$$\begin{aligned}\psi_f &= -je \int d^4x' \left(p_f A(x') + A(x')p_i\right)\psi_i e^{jq\cdot x'} \\ &= -je\big(p_f A(q) + A(q)p_i\big)\psi_i,\end{aligned} \tag{8.205}$$

where $q = p_f - p_i$ is the change in momentum, and $A(q)$ is the Fourier transform of the electromagnetic potential. The form of the result here is quite typical, and in general we can write

$$\psi_f = S_{fi}\psi_i, \tag{8.206}$$

where S_{fi} is the *scattering operator*. This is a multivector that takes initial states into final states. Since both ψ_i and ψ_f are plane-wave particle states, we must have

$$S_{fi}\tilde{S}_{fi} = \rho_{fi}, \tag{8.207}$$

where ρ_{fi} is a scalar quantity (which determines the cross section). We can therefore decompose S_{fi} as

$$S_{fi} = \rho_{fi}^{1/2} R_{fi}, \tag{8.208}$$

where R_{fi} is a rotor. This rotor takes the initial momentum to the final momentum,

$$R_{fi}p_i\tilde{R}_{fi} = p_f. \tag{8.209}$$

8.5.2 Spin effects in scattering

The multivector S_{fi} depends on the initial and final momenta and, in some cases, the initial spin. The final spin is determined from the initial spin by the rotation encoded in S_{fi}. If s_i and s_f denote the initial and final (unit) spin vectors, we have

$$s_f = R_{fi}s_i\tilde{R}_{fi}. \tag{8.210}$$

Sometimes it is of greater interest to separate out the boost terms in R_{fi} to isolate a pure rotation in the γ_0 frame. This tells us directly what happens to the spin vector in the electron's rest frame. With L_i and L_f the appropriate pure boosts, we define the rest spin scattering operator

$$U_{fi} = \tilde{L}_f R_{fi} L_i. \tag{8.211}$$

This satisfies

$$U_{fi}\gamma_0\tilde{U}_{fi} = \frac{1}{m}\tilde{L}_f R_{fi}p_i\tilde{R}_{fi}L_f = \gamma_0, \tag{8.212}$$

so is a pure rotation in the γ_0 frame.

The fact that $p_f S_{fi} = S_{fi} p_i$ ensures that S_{fi} is always of the form

$$S_{fi} = -j(p_f M + M p_i), \tag{8.213}$$

where M is an odd-grade multivector. In the Born approximation of equation (8.205), for example, we have $M = eA(q)$. In general, M can contain both real and imaginary terms, so we must write

$$S_{fi}\psi_i = -j\big(p_f(M_r + jM_j) + (M_r + jM_j)p_i\big)\psi_i, \tag{8.214}$$

where M_j and M_r are independent of j. We can now use

$$j\psi_i = \psi_i I\boldsymbol{\sigma}_3 = \hat{S}_i\psi_i, \tag{8.215}$$

where $\hat{S}_i$ is the initial unit spin bivector. Since $\hat{S}_i$ and p_i commute, S_{fi} can still be written in the form of equation (8.213), with

$$M = M_r + M_j\hat{S}_i. \tag{8.216}$$

So M remains a real multivector, which now depends on the initial spin. This scheme is helpful if we are interested in any spin-dependent features of the scattering process.

8.5.3 Positron scattering and pair annihilation

Adapting the preceding results to positron scattering is straightforward. In this case a negative-energy plane wave arrives from the future and scatters into the past, so we set

$$\psi_i(x) = \psi_2 e^{jp_i\cdot x}, \quad \psi_f(x) = \psi_f e^{jp_f\cdot x}. \tag{8.217}$$

In this case repeating the analysis gives

$$S_{fi}\psi_i = -j(-p_f M\psi_i + M\psi_i\gamma_0), \tag{8.218}$$

which we can write as

$$S_{fi} = j(p_f M + M p_i). \tag{8.219}$$

This amounts to simply swapping the sign of S_{fi}. In the Born approximation, q is replaced by $-q$ in the Fourier transform of $A(x)$, which will alter the factor M if $A(x)$ is complex.

The other case to consider is when the incoming electron is scattered into the past, corresponding to pair annihilation. In this case we have

$$S_{fi} = -j(-p_2 M + M p_1), \tag{8.220}$$

where p_1 and p_2 are the incoming momenta of the electron and positron respectively. We decompose S_{fi} as

$$S_{fi} = \rho_{fi}^{1/2} I R_{fi}, \tag{8.221}$$

since S_{fi} must now contain a factor of I to map electrons into positrons. This form for S_{fi} implies that

$$S_{fi}\tilde{S}_{fi} = -\rho_{fi}. \tag{8.222}$$

The minus sign reflects the fact that the transformation between initial and final momenta is not proper orthochronous.

8.5.4 Cross sections

We must now relate our results to the cross sections measured in experiments. The scattering rate into the final states, per unit volume, per unit time, is given by

$$W_{fi} = \frac{1}{VT}N_f = \frac{1}{VT}\frac{\gamma_0 \cdot J_f}{2E_f} = \frac{\rho_f}{2mVT}, \tag{8.223}$$

where V and T denote the total volume and time respectively. The density ρ_f is given by

$$\rho_f = |S_{fi}\tilde{S}_{fi}|\rho_i = \rho_{fi}\rho_i. \tag{8.224}$$

Here $S_{fi}\tilde{S}_{fi} = \pm\rho_{fi}$, where the plus sign corresponds to electron to electron and positron to positron scattering, and the minus sign to electron–positron annihilation.

The differential cross section is defined as

$$d\sigma = \frac{W_{fi}}{\text{target density} \times \text{incident flux}}. \tag{8.225}$$

When S_{fi} is of the form

$$S_{fi} = -j(2\pi)^4\delta^4(P_f - P_i)T_{fi}, \tag{8.226}$$

where the δ-function ensures conservation of total momentum, we have

$$|S_{fi}|^2 = VT(2\pi)^4\delta^4(P_f - P_i)|T_{fi}|^2. \tag{8.227}$$

Working in the J_i frame the target density is just ρ_i so, writing the incident flux as χ, we have

$$d\sigma = \frac{1}{2m\chi}(2\pi)^4\delta^4(P_f - P_i)|T_{fi}|^2. \tag{8.228}$$

Alternatively we may be interested in an elastic scattering with just energy conservation ($E_f = E_i$) and

$$S_{fi} = -j2\pi\delta(E_f - E_i)T_{fi}. \tag{8.229}$$

In this case

$$|S_{fi}|^2 = 2\pi T\delta(E_f - E_i)|T_{fi}|^2. \tag{8.230}$$

A target density of $1/V$ and an incident flux of $|\boldsymbol{J}_i| = \rho_i|\boldsymbol{p}_i|/m$ then gives

$$d\sigma = \frac{\pi}{|\boldsymbol{p}_i|}\delta(E_f - E_i)|T_{fi}|^2. \tag{8.231}$$

The total cross section is obtained by integrating over the available phase space. For the case of a single particle scattering elastically we find that

$$\sigma = \int \frac{d^3\boldsymbol{p}_f}{(2\pi)^3}\frac{1}{2E_f}\frac{\pi}{|\boldsymbol{p}_i|}\delta(E_f - E_i)|T_{fi}|^2 = \int d\Omega\,\frac{|T_{fi}|^2}{16\pi^2}. \tag{8.232}$$

This is usually expressed in terms of the differential cross section per solid angle:

$$\frac{d\sigma}{d\Omega_f} = \frac{|T_{fi}|^2}{16\pi^2}. \tag{8.233}$$

8.5.5 Coulomb scattering

As an application of our formalism consider Coulomb scattering from a nucleus, with the external field defined by

$$A(x) = \frac{-Ze}{4\pi|\boldsymbol{x}|}\gamma_0. \tag{8.234}$$

Working with the first Born approximation, M is given by $M = eA(q)$, where $A(q)$ is the Fourier transform of $A(x)$ given by

$$A(q) = -\frac{2\pi Ze}{\boldsymbol{q}^2}\delta(E_f - E_i)\gamma_0 \tag{8.235}$$

and $q\cdot\gamma_0 = E_f - E_i$. Writing

$$S_{fi} = -j2\pi\delta(E_f - E_i)T_{fi} \tag{8.236}$$

and using energy conservation we find that

$$T_{fi} = -\frac{Ze^2}{\boldsymbol{q}^2}(2E + \boldsymbol{q}). \tag{8.237}$$

The cross section is therefore given by the Mott scattering formula:

$$\frac{d\sigma}{d\Omega_f} = \frac{Z^2\alpha^2}{\boldsymbol{q}^4}(4E^2 - \boldsymbol{q}^2) = \frac{Z^2\alpha^2}{4\boldsymbol{p}^2\beta^2\sin^4(\theta/2)}\left(1 - \beta^2\sin^2(\theta/2)\right), \tag{8.238}$$

where

$$\boldsymbol{q}^2 = (\boldsymbol{p}_f - \boldsymbol{p}_i)^2 = 2\boldsymbol{p}^2\big(1 - \cos(\theta)\big) \quad \text{and} \quad \beta = |\boldsymbol{p}|/E. \tag{8.239}$$

The angle θ measures the deviation between the incoming and scattered beams. In the low velocity limit the Mott result reduces to the Rutherford formula. The result is independent of the sign of the nuclear charge and, to this order, is obtained for both electron and positron scattering.

A significant feature of this derivation is that no spin sums are required. Instead, all the spin dependence is contained in the directional information in T_{fi}. As well as being computationally more efficient, this method for organising cross section calculations offers deeper insights into the structure of the theory. For Coulomb scattering the spin information is contained in the rotor

$$R_{fi} = \frac{p_f\gamma_0 + \gamma_0 p_i}{4E^2 - \boldsymbol{q}^2} \propto L_f^2 + \tilde{L}_i^2, \tag{8.240}$$

where L_f and L_i are the pure boosts from γ_0 to p_f and p_i respectively. The behaviour of the rest spin is governed by the unnormalised rotor

$$U_{fi} = \tilde{L}_f(L_f^2 + \tilde{L}_i^2)L_i = L_f L_i + \tilde{L}_f\tilde{L}_i, = 2\big((E+m)^2 + \boldsymbol{p}_f\boldsymbol{p}_i\big). \tag{8.241}$$

It follows that the rest-spin vector precesses in the $\boldsymbol{p}_f \wedge \boldsymbol{p}_i$ plane through an angle δ, where

$$\tan(\delta/2) = \frac{\sin(\theta)}{(E+m)/(E-m) + \cos(\theta)}. \tag{8.242}$$

This method of calculating the spin precession for Coulomb scattering was first described by Hestenes (1982a).

8.5.6 Compton scattering

Compton scattering is the process in which an electron scatters off a photon. To lowest order there are two Feynman diagrams to consider, shown in figure 8.4. The preceding analysis follows through with little modification, and gives rise two terms of the form

$$M_1 = e^2 \iint d^4x_1\, d^4x_2\, \frac{d^4p}{(2\pi)^4}\, A_1(x_1)\frac{pA_2(x_2) + A_2(x_2)p_i}{p^2 - m^2 + j\epsilon} \times \mathrm{e}^{jx_1\cdot(p_f - p)}\mathrm{e}^{jx_2\cdot(p - p_i)}, \tag{8.243}$$

where

$$A(x) = \epsilon \mathrm{e}^{\mp jk\cdot x} \tag{8.244}$$

is the (complex) vector potential. The vector ϵ denotes the polarisation state, so $k\cdot\epsilon = 0$ and $\epsilon^2 = -1$. In relativistic quantum theory there appears to be no alternative but to work with a fully complex vector potential.

Performing the integrations and summing the two contributions we arrive at

$$M = e^2(2\pi)^4\delta^4(P)\left(\epsilon_f\frac{(p_i + k_i)\epsilon_i + \epsilon_i p_i}{2k_i\cdot p_i} - \epsilon_i\frac{(p_i - k_f)\epsilon_f + \epsilon_f p_i}{2p_i\cdot k_f}\right), \tag{8.245}$$

where $P = p_f + k_f - p_i - k_i$, so that the δ-function enforces momentum conservation. Gauge invariance means that we can set $p_i\cdot\epsilon_i = p_i\cdot\epsilon_f = 0$, in which case

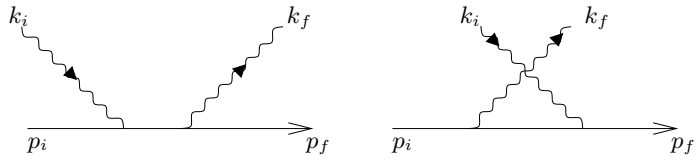

Figure 8.4 *Compton scattering.* Two diagrams contribute to the amplitude, to lowest order.

M simplifies to

$$M = e^2(2\pi)^4\delta^4(p_f + k_f - p_i - k_i)\left(\frac{\epsilon_f k_i \epsilon_i}{2k_i\cdot p_i} + \frac{\epsilon_i k_f \epsilon_f}{2p_i\cdot k_f}\right). \tag{8.246}$$

We now set

$$S_{fi} = -j(2\pi)^4\delta^4(p_f + k_f - p_i - k_i)T_{fi}, \tag{8.247}$$

so that the cross section is given by equation (8.228). After a little work, and making use of momentum conservation, we find that

$$|T_{fi}|^2 = e^4\left(4(\epsilon_i\cdot\epsilon_f)^2 - 2 + \frac{p_i\cdot k_f}{p_i\cdot k_i} + \frac{p_i\cdot k_i}{p_i\cdot k_f}\right). \tag{8.248}$$

This is all that is required to calculate the cross section in any desired frame. Again, this derivation applies regardless of the initial electron spin.

The same scheme can be applied to a wide range of relativistic scattering problems. In all cases the spacetime algebra formulation provides a simpler and clearer method for handling the spin, as it does not force us to work with a preferred basis set. In section 14.4.1 the same formalism is applied to scattering from a black hole. At some point, however, it is necessary to face questions of second quantisation and the construction of a relativistic multiparticle quantum theory. This is discussed in the following chapter.

8.6 Notes

A significant amount of new notation was introduced in this chapter, relating to how spinors are handled in spacetime algebra. Much of this is important in later chapters, and the most useful results of this approach are summarised in table 8.3.

Quantum mechanics has probably been the most widely researched application of geometric algebra to date. Many authors have carried out investigations into whether the spacetime algebra formulation of Dirac theory offers any deeper insights into the nature of quantum theory. Among the most interesting of these are Hestenes' work on zitterbewegung (1990), and his comments on the nature

Pauli spinors	$\|\psi\rangle = \begin{pmatrix} a^0 + ia^3 \\ -a^2 + ia^1 \end{pmatrix} \leftrightarrow \psi = a^0 + a^k I\boldsymbol{\sigma}_k$
Pauli operators	$\hat{\sigma}_k\|\psi\rangle \leftrightarrow \boldsymbol{\sigma}_k\psi\boldsymbol{\sigma}_3$ $i\|\psi\rangle \leftrightarrow \psi I\boldsymbol{\sigma}_3 = j\psi$ $\langle\psi\|\psi'\rangle \leftrightarrow \langle\psi^\dagger\psi'\rangle_q = \frac{1}{2}(\psi^\dagger\psi' + \boldsymbol{\sigma}_3\psi^\dagger\psi'\boldsymbol{\sigma}_3)$
Pauli observables	$\rho = \psi\psi^\dagger$ $\boldsymbol{s} = \frac{1}{2}\psi\boldsymbol{\sigma}_3\psi^\dagger$
Dirac spinors	$\begin{pmatrix} \|\phi\rangle \\ \|\eta\rangle \end{pmatrix} \leftrightarrow \psi = \phi + \eta\boldsymbol{\sigma}_3$
Dirac operators	$\hat{\gamma}_\mu\|\psi\rangle \leftrightarrow \gamma_\mu\psi\gamma_0$ $j\|\psi\rangle \leftrightarrow \psi i\boldsymbol{\sigma}_3$ $\hat{\gamma}_5\|\psi\rangle \leftrightarrow \psi\boldsymbol{\sigma}_3$ $\langle\tilde{\psi}\|\psi'\rangle \leftrightarrow \langle\tilde{\psi}\psi'\rangle_q$
Dirac equation	$\nabla\psi I\boldsymbol{\sigma}_3 - eA\psi = m\psi\gamma_0$
Dirac observables	$\rho e^{i\beta} = \psi\tilde{\psi} \qquad J = \psi\gamma_0\tilde{\psi}$ $S = \psi i\boldsymbol{\sigma}_3\tilde{\psi} \quad s = \psi\gamma_3\tilde{\psi}$
Plane-wave states	$\psi^{(+)}(x) = L(p)\Phi e^{-i\boldsymbol{\sigma}_3 p\cdot x}$ $\psi^{(-)}(x) = L(p)\Phi\boldsymbol{\sigma}_3 e^{i\boldsymbol{\sigma}_3 p\cdot x}$ $L(p) = (p\gamma_0 + m)/\sqrt{2m(E+m)}$

Table 8.3 *Quantum states and operators.* This table summarises the main features of the spacetime algebra representation of Pauli and Dirac spinors and operators.

of the electroweak group (1982b). Many authors have advocated spacetime algebra as a better computational tool for Dirac theory than the explicit matrix formulation (augmented with various spin sum rules). A summary of these ideas is contained in the paper 'Electron scattering without spin sums' by Lewis *et al.* (2001). Elsewhere, a similar approach has been applied to modelling a spin measurement (Challinor *et al.* 1996) and to the results of tunnelling experiments (Gull *et al.* 1993b). Much of this work is summarised in the review 'Spacetime algebra and electron physics' by Doran *et al.* (1996b).

There is no shortage of good textbooks describing standard formulations of Dirac theory and quantum electrodynamics. We particularly made use of the classic texts by Itzykson & Zuber (1980), and Bjorken & Drell (1964). For a detailed exposition of the solution of the Dirac equation in various backgrounds one can do little better than Greiner's *Relativistic Quantum Mechanics* (1990).

Also recommended is Grandy's *Relativistic Quantum Mechanics of Leptons and Fields* (1991) which, unusually, does not shy away from the more problematic areas of the conceptual foundations of quantum field theory.

8.7 Exercises

8.1 The spin matrix operators $\hat{s}_k$ are defined as a set of 2×2 Hermitian matrices satisfying the commutation relations $[\hat{s}_i, \hat{s}_j] = i\hbar\epsilon_{ijk}\hat{s}_k$. Given that $\hat{s}_3$ is defined by

$$\hat{s}_3 = \lambda \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

show that the remaining matrices are unique, up to an overall choice of phase. Find λ and show that we can choose the phase such that $\hat{s}_k = \hbar/2\,\hat{\sigma}_k$.

8.2 Verify that the equivalence between Pauli spinors and even multivectors defined in equation (8.20) is consistent with the operator equivalences

$$\hat{\sigma}_k|\psi\rangle \;\leftrightarrow\; \boldsymbol{\sigma}_k\psi\boldsymbol{\sigma}_3 \quad (k = 1, 2, 3).$$

8.3 Suppose that two spin-1/2 states are represented by the even multivectors ϕ and ψ, and the accompanying spin vectors are

$$\boldsymbol{s}_1 = \phi\boldsymbol{\sigma}_3\tilde{\phi} \quad \text{and} \quad \boldsymbol{s}_2 = \psi\boldsymbol{\sigma}_3\tilde{\psi}.$$

Prove that the quantum mechanical formula for the probability of measuring state ϕ in state ψ satisfies

$$P = \frac{|\langle\phi|\psi\rangle|^2}{\langle\phi|\phi\rangle\langle\psi|\psi\rangle} = \tfrac{1}{2}\big(1 + \cos(\theta)\big)$$

where θ is the angle between $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$.

8.4 Verify that the Pauli inner product is invariant under both spatial rotations and gauge transformations (i.e. rotations in the $I\boldsymbol{\sigma}_3$ plane applied to the right of the spinor ψ). Repeat the analysis for Dirac spinors.

8.5 Prove that the angular momentum operators $L_B = jB\cdot(\boldsymbol{x}\wedge\boldsymbol{\nabla})$ satisfy

$$[L_{B_1}, L_{B_2}] = -jL_{B_1\times B_2}.$$

8.6 Prove that, in any dimension,

$$[B\cdot(x\wedge\nabla) - \tfrac{1}{2}B, \nabla] = 0,$$

where B is a bivector.

8.7 The Majorana representation is defined in terms of a set of real matrices. Prove that the complex conjugation operation in this representation has the spacetime algebra equivalent

$$|\psi\rangle^*_{Maj} \leftrightarrow \psi\boldsymbol{\sigma}_2.$$

Confirm that this anticommutes with the operation of multiplying by the imaginary.

8.8 Prove that the associated Legendre polynomials satisfy the following recursion relations:

$$(1-x^2)\frac{dP_l^m(x)}{dx} + mxP_l^m(x) = -(1-x^2)^{1/2}P_l^{m+1}(x),$$

$$(1-x^2)\frac{dP_l^m(x)}{dx} - mxP_l^m(x) = (1-x^2)^{1/2}(l+m)(l-m+1)P_l^{m-1}(x).$$

8.9 Prove that the spherical monogenics satisfy

$$\int d\Omega \, \langle \psi_l^{m\dagger}\psi_{l'}^{m'}\rangle_q = \delta^{mm'}\delta_{ll'}4\pi\frac{(l+m+1)!}{(l-m)!}.$$

8.10 From the result of equation (8.248), show that the cross section for scattering of a photon of a free electron (initially at rest) is determined by the Klein–Nishina formula

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4m^2}\left(\frac{\omega_f}{\omega_i}\right)^2\left(\frac{\omega_f}{\omega_i} + \frac{\omega_i}{\omega_f} + 4(\epsilon_f\cdot\epsilon_i)^2 - 2\right).$$