

Physics 221B
Spring 2020
Notes 49
Solutions of the Dirac Equation and Their Properties[†]

1. Introduction

In Notes 45 we introduced the Dirac equation in much the same manner as Dirac himself did, with the motivation of curing the problems of the Klein-Gordon equation. We saw that the Dirac equation, unlike the Klein-Gordon equation, admits a conserved 4-current with a nonnegative definite time component, so that it can be interpreted as a probability 4-current. On the other hand, the Klein-Gordon equation has one problem that is shared with the Dirac equation, the existence of negative energy solutions, for which we have no interpretation as yet. We might suppose that the negative energy solutions could simply be declared to be nonphysical, but this turns out to be impossible, as we will see. In addition, the velocity operator in the Dirac theory is strange, being a purely spin operator when we might have expected something purely spatial on the nonrelativistic analogy. The same velocity operator appears in the spatial part of the probability current. It is however encouraging that the Dirac equation automatically produces the correct interaction of the spin with a magnetic field, with the (almost exactly) correct g -factor for an electron. To summarize, as a candidate relativistic wave equation for the electron, the Dirac equation presents some remarkable successes, some strange features, and some others with no obvious physical interpretation.

The purpose of these notes is to explore some solutions of the Dirac equation and their properties, in order to accumulate some results that will be useful for later work, to reveal some further successes of the Dirac equation, and to explore some of the puzzling aspects of the theory in order to understand them better and to lay the foundation for their ultimate resolution.

We begin by exploring the properties of free particle solutions of the Dirac equation, including the negative energy solutions. The free particle solutions form a basis of orthonormal eigenfunctions that will be useful for later work, especially when we turn to the second quantization of the Dirac field. We do not manage at this time to obtain a physical interpretation of the negative energy free particle solutions, but we learn some things about them, such as their energy spectrum and the fact that, in a certain sense, they have negative momentum and spin as well as negative energy. We then turn to the Gordon decomposition of the current, which gives some understanding as to why the velocity appears as a spin operator in the Dirac theory. Then we present a discussion of the Zitterbewegung, a kind of jittering motion that the Dirac electron undergoes on a rapid time scale,

[†] Links to the other sets of notes can be found at:
<http://bohr.physics.berkeley.edu/classes/221/1920/221.html>.

even in the case of a free particle. The Zitterbewegung is related to the Darwin term in atomic physics (see Notes 24). Finally, we discuss central force problems in general and then the exact solution of the hydrogen atom for the Dirac equation. This gives us not only all the fine structure corrections we saw in Notes 24 but also negative energy solutions. For these, however, as in the case of the free particle, we have no interpretation as yet.

The net effect of this set of notes will be to give us experience with some exact solutions of the Dirac equation (the free particle and the hydrogen atom), some understanding of some of the strange features (the velocity operator), some new strange features (the Zitterbewegung), and yet still no physical interpretation of the negative energy solutions (that will come later).

In the following treatment of free particle solutions it will be important to distinguish the momentum operator (3-vector or 4-vector) from the corresponding vectors of c -numbers. Therefore we will put an “op” subscript on the momentum when an operator is intended, such as $\mathbf{p}_{\text{op}} = -i\hbar\nabla$ or $p_{\text{op}}^\mu = i\hbar\partial^\mu$, and omit the “op” subscript when a c -number is intended. For other operators the distinction will not be necessary so we will omit the “op” subscripts. We will use hats on unit vectors, as usual.

2. Free Particle Solutions of the Pauli Equation

The Pauli Hamiltonian for a free particle of spin $\frac{1}{2}$ is $H = \mathbf{p}_{\text{op}}^2/2m$ and the wave function is a 2-component spinor. A complete set of commuting observables that also commute with the Hamiltonian is $(\mathbf{p}_{\text{op}}, S_z)$. The eigenfunctions of H that are also eigenfunctions of these observables are

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{i(\mathbf{p}\cdot\mathbf{x}-Et)/\hbar}, \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{i(\mathbf{p}\cdot\mathbf{x}-Et)/\hbar}, \quad (1)$$

where \mathbf{p} is the momentum eigenvalue, where the first solution is an eigenfunction of S_z with eigenvalue $+\hbar/2$ and the second with eigenvalue $-\hbar/2$. We have included the time dependence in these solutions, where $E = \mathbf{p}^2/2m$.

3. Dirac Free Particle Solutions and Spin

The Dirac Hamiltonian for a free particle is

$$H = c\mathbf{p}_{\text{op}} \cdot \boldsymbol{\alpha} + \beta mc^2. \quad (2)$$

In comparison to the Pauli theory, we might expect to find eigenfunctions of H that are also eigenfunctions of \mathbf{p}_{op} and $(\hbar/2)\Sigma_3$, the latter being the Dirac generalization of the Pauli operator $S_z = (\hbar/2)\sigma_3$. This does not work, however, for although \mathbf{p}_{op} and $\boldsymbol{\Sigma}$ commute with each other (one is purely spatial and the other purely spin), and although \mathbf{p}_{op} commutes with H , $\boldsymbol{\Sigma}$ does not commute with H . This is because $\boldsymbol{\Sigma}$ generates spin rotations, which rotate the vector $\boldsymbol{\alpha}$ but not the spatial vector \mathbf{p}_{op} , so the dot product $\mathbf{p}_{\text{op}} \cdot \boldsymbol{\alpha}$ is not invariant. On the other hand, orbital angular momentum \mathbf{L} generates spatial rotations that rotate \mathbf{p}_{op} but not $\boldsymbol{\alpha}$, so again $\mathbf{p}_{\text{op}} \cdot \boldsymbol{\alpha}$ is not

invariant. Obviously if we want to find an angular momentum operator that does commute with H , it is $\mathbf{J} = \mathbf{L} + (\hbar/2)\boldsymbol{\Sigma}$, the total angular momentum (orbital plus spin) of the Dirac particle. See Sec. 47.17.

Although orbital and spin angular momentum are defined for the Dirac particle, they are not separately conserved, even for a free particle. This is a reflection of the more intimate coupling of orbital and spin degrees of freedom in the relativistic theory, as compared to nonrelativistic quantum mechanics. Recall that for the photon, orbital and spin angular momentum are separately not even defined, although \mathbf{J} is.

On account of these considerations, we can expect spin to play a more complicated role in the theory of free particle solutions of the Dirac equation than it does in the Pauli theory.

4. Free Particle Solutions of the Dirac Equation at Rest

There are four linearly independent solutions of the time-dependent Dirac equation for a free particle at rest. These are given in Eq. (45.31), reproduced here:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} e^{-imc^2 t/\hbar}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} e^{-imc^2 t/\hbar}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} e^{+imc^2 t/\hbar}, \quad \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} e^{+imc^2 t/\hbar}. \quad (3)$$

The first two of these have energy $+mc^2$, and the last two have energy $-mc^2$, as we find by applying either H or $i\hbar \partial/\partial t$. We have as yet no physical interpretation for the solutions of negative energy, but they will turn out to be important so we must keep them and investigate their properties.

The solutions at rest, Eqs. (3), are eigenstates of Σ_3 , that is, the spin is polarized in the $\pm \hat{\mathbf{z}}$ direction in the rest frame. In following work we will want to be more general, and to allow the spin to be polarized in an arbitrary direction in the rest frame. Let $\hat{\mathbf{s}}$ be a unit vector in the rest frame of the particle. If (θ, ϕ) are the angles of $\hat{\mathbf{s}}$ in spherical coordinates, then we define the 2-component (Pauli) spinor $\chi(\hat{\mathbf{s}})$ as the spinor “pointing in” the direction $\hat{\mathbf{s}}$, as explained in Sec. 12.14. According to Eq. (12.61), this spinor is

$$\chi(\hat{\mathbf{s}}) = \begin{pmatrix} e^{-i\phi/2} \cos(\theta/2) \\ e^{i\phi/2} \sin(\theta/2) \end{pmatrix}, \quad (4)$$

to within an overall phase which will not be important to us. In this notation the eigenspinors of $\hat{\mathbf{s}} \cdot \boldsymbol{\sigma}$ are $\chi(\pm \hat{\mathbf{s}})$,

$$(\hat{\mathbf{s}} \cdot \boldsymbol{\sigma})\chi(\hat{\mathbf{s}}) = \chi(\hat{\mathbf{s}}), \quad (\hat{\mathbf{s}} \cdot \boldsymbol{\sigma})\chi(-\hat{\mathbf{s}}) = -\chi(-\hat{\mathbf{s}}), \quad (5)$$

the second of which follows by making the replacement $\hat{\mathbf{s}} \rightarrow -\hat{\mathbf{s}}$ in the first. For a fixed value of $\hat{\mathbf{s}}$, the spinors $\chi(\pm \hat{\mathbf{s}})$ form an orthonormal basis in the space of 2-component spinors.

Now we define two types of 4-component spinors, here broken into their upper and lower 2-component spinors,

$$u_0(\hat{\mathbf{s}}) = \begin{pmatrix} \chi(\hat{\mathbf{s}}) \\ 0 \end{pmatrix}, \quad v_0(\hat{\mathbf{s}}) = \begin{pmatrix} 0 \\ \chi(-\hat{\mathbf{s}}) \end{pmatrix}, \quad (6)$$

where the 0 subscript on u_0 and v_0 means, “at rest.” The odd thing about these definitions is that we have used $\chi(-\hat{\mathbf{s}})$ in the definition of v_0 , that is, with a minus sign on the direction of spin. The reason for this choice will be explained momentarily. These spinors are eigenspinors of the Dirac matrix $\hat{\mathbf{s}} \cdot \boldsymbol{\Sigma}$, where

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix} \quad (7)$$

(this is Eq. (47.61)). That is,

$$(\hat{\mathbf{s}} \cdot \boldsymbol{\Sigma})u_0(\hat{\mathbf{s}}) = u_0(\hat{\mathbf{s}}), \quad (\hat{\mathbf{s}} \cdot \boldsymbol{\Sigma})v_0(\hat{\mathbf{s}}) = -v_0(\hat{\mathbf{s}}). \quad (8)$$

Then the most general free-particle solutions of the Dirac equation when the particle is at rest can be written

$$u_0(\hat{\mathbf{s}}) e^{-imc^2 t/\hbar}, \quad v_0(\hat{\mathbf{s}}) e^{+imc^2 t/\hbar}, \quad (9)$$

of which the first has energy $+mc^2$ and the second $-mc^2$. These are the positive- and negative-energy solutions, respectively. If $\hat{\mathbf{s}}$ is chosen to run over two oppositely oriented unit vectors in the rest frame, then the solutions (9) constitute a complete set of free particle solutions at rest, with definite spins in the rest frame.

5. Boosting the Solutions at Rest

To create free particle solutions in some state of motion we apply a boost to the free particle at rest. We first consider this problem from a classical standpoint. Let \mathbf{p} be the desired 3-momentum of our particle, and $\Lambda(\mathbf{p})$ the boost that will take the particle at rest into one with momentum \mathbf{p} . The 4-momentum of the particle at rest is

$$p_0^\mu = \begin{pmatrix} mc \\ \mathbf{0} \end{pmatrix}, \quad (10)$$

where again the 0-subscript means “at rest” and the 4-vector is given in $(1+3)$ -notation. The 4-momentum of the particle after the boost is

$$p^\mu = \begin{pmatrix} E/c \\ \mathbf{p} \end{pmatrix}, \quad (11)$$

where

$$E = E(\mathbf{p}) = \sqrt{m^2 c^4 + c^2 \mathbf{p}^2}. \quad (12)$$

This is of course the standard definition of the energy of a free particle in relativity theory, but in the following it will be important to remember that the symbol E stands for the positive square root in Eq. (12), even when we are talking about negative energy solutions. Also, in the following both the energy E and the 4-momentum p^μ will be understood to be functions of the 3-momentum \mathbf{p} , which by itself characterizes the state of motion of the particle after the boost.

We can put the boost $\Lambda(\mathbf{p})$ into axis-rapidity form, as in Sec. 46.5. The boost must be in the direction of the momentum \mathbf{p} so the boost axis is

$$\hat{\mathbf{b}} = \frac{\mathbf{p}}{|\mathbf{p}|}, \quad (13)$$

where $\hat{\mathbf{b}}$ is a unit vector, and its rapidity λ is given implicitly as a function of \mathbf{p} by

$$\cosh \lambda = \frac{E}{mc^2}, \quad \sinh \lambda = \frac{|\mathbf{p}|}{mc}, \quad \tanh \lambda = \frac{c|\mathbf{p}|}{E} = \frac{|\mathbf{v}|}{c}, \quad (14)$$

where \mathbf{v} is the velocity of the particle after the boost. With this axis and rapidity, we have $\Lambda(\mathbf{p}) = \Lambda(\hat{\mathbf{b}}, \lambda)$. The action of a boost in axis-rapidity notation on an arbitrary 4-vector is given explicitly by Eq. (46.34). Applying that equation, we can check that

$$p^\mu = \Lambda(\mathbf{p})^\mu{}_\nu p_0^\nu. \quad (15)$$

We shall sometimes write this as $p = \Lambda p_0$, using simply p and p_0 for the 4-momenta with components p^μ and p_0^μ .

We turn now to the quantum side of the boost, that is, boosting the free particle solutions at rest, Eqs. (9), of the Dirac equation. In this process it will be convenient to use covariant notation insofar as possible. As for the exponents in Eq. (9), notice that

$$mc^2 t = (mc)(ct) = p_{0\mu} x^\mu = p_0 \cdot x, \quad (16)$$

in the notation of Eq. (46.104) for the Minkowski scalar product of two 4-vectors. Therefore we can write the positive and negative energy solutions for a free particle at rest as

$$u_0(\hat{\mathbf{s}}) e^{-i(p_0 \cdot x)/\hbar}, \quad v_0(\hat{\mathbf{s}}) e^{+i(p_0 \cdot x)/\hbar}. \quad (17)$$

Under the Lorentz transformation $\Lambda(\mathbf{p})$, a Dirac 4-component wave function $\psi(x)$ is mapped into $D(\mathbf{p})\psi(\Lambda(\mathbf{p})^{-1}x)$, as explained in Sec. 47.4, where we define

$$D(\mathbf{p}) = D(\Lambda(\mathbf{p})). \quad (18)$$

Under this mapping, the exponents in Eq. (17) transform according to

$$p_0 \cdot x \mapsto p_0 \cdot (\Lambda^{-1}x) = (\Lambda p_0) \cdot x = p \cdot x = p_\mu x^\mu, \quad (19)$$

where we use the identity (46.105). This calculation is essentially the same as that presented in Sec. 47.5. As for the matrix $D(\mathbf{p})$, the D -matrix in axis-rapidity notation is given by Eq. (47.75). Here we express the hyperbolic functions of half-angles in terms of the energy and momentum of the particle after the boost, as follows:

$$\begin{aligned} \cosh \frac{\lambda}{2} &= \sqrt{\frac{\cosh \lambda + 1}{2}} = \sqrt{\frac{E + mc^2}{2mc^2}}, \\ \sinh \frac{\lambda}{2} &= \sqrt{\frac{\cosh \lambda - 1}{2}} = \sqrt{\frac{E - mc^2}{2mc^2}}, \\ \tanh \frac{\lambda}{2} &= \sqrt{\frac{E - mc^2}{E + mc^2}} = \sqrt{\frac{E^2 - m^2 c^4}{(E + mc^2)^2}} = \frac{c|\mathbf{p}|}{E + mc^2}. \end{aligned} \quad (20)$$

Substituting these into Eq. (47.75), we obtain the explicit form of the D -matrix,

$$D(\mathbf{p}) = \sqrt{\frac{E + mc^2}{2mc^2}} \begin{pmatrix} 1 & \frac{c\mathbf{p} \cdot \boldsymbol{\sigma}}{E + mc^2} \\ \frac{c\mathbf{p} \cdot \boldsymbol{\sigma}}{E + mc^2} & 1 \end{pmatrix}. \quad (21)$$

Under the Lorentz transformation $D(\mathbf{p})$ acts on the spinors $u_0(\hat{\mathbf{s}})$, $v_0(\hat{\mathbf{s}})$, giving us boosted spinors that we will write as

$$u(\mathbf{p}, \hat{\mathbf{s}}) = D(\mathbf{p})u_0(\hat{\mathbf{s}}), \quad v(\mathbf{p}, \hat{\mathbf{s}}) = D(\mathbf{p})v_0(\hat{\mathbf{s}}), \quad (22)$$

without the 0-subscript. The boosted positive and negative energy solutions are now

$$u(\mathbf{p}, \hat{\mathbf{s}}) e^{-i(p \cdot x)/\hbar}, \quad v(\mathbf{p}, \hat{\mathbf{s}}) e^{+i(p \cdot x)/\hbar}. \quad (23)$$

If we apply H or $i\hbar\partial/\partial t$ to these to obtain their energies, we find that the positive energy solutions have energy E , given in terms of the momentum \mathbf{p} which parameterizes the boost by Eq. (12), while the negative energy solutions have energy $-E$. We see that when we boost a negative energy solution at rest, we obtain a solution whose energy is even more negative. Thus, the energy spectrum of the free particle Dirac Hamiltonian consists of a positive energy part ranging from $+mc^2$ to ∞ , and a negative energy part ranging from $-mc^2$ to $-\infty$, as illustrated in Fig. 1. There is an energy gap between $-mc^2$ and $+mc^2$.

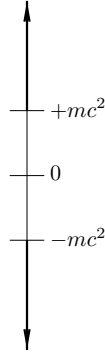


Fig. 1. The energy spectrum of the free particle Dirac Hamiltonian consists of all positive energies $\geq +mc^2$ and all negative energies $\leq -mc^2$. There is an energy gap between $-mc^2$ and $+mc^2$.

The free particle solutions (23) are not only eigenfunctions of energy, they are also eigenfunctions of momentum. If we apply the operator $\mathbf{p}_{\text{op}} = -i\hbar\nabla$ to these solutions, we find that positive energy solutions are eigenfunctions of momentum with eigenvalue \mathbf{p} , while the negative energy solutions are eigenfunctions of momentum with eigenvalue $-\mathbf{p}$. We see that when we boost a negative energy solution at rest by a Lorentz transformation parameterized by \mathbf{p} , we obtain a solution with momentum $-\mathbf{p}$ (in the opposite direction of the boost). In a certain sense, the negative energy solutions have not only negative energy but also negative momentum.

We see that the negative energy solutions $v(\mathbf{p}, \hat{\mathbf{s}})e^{i(p \cdot x)/\hbar}$ have a momentum eigenvalue that is the opposite of the momentum label \mathbf{p} . The same is true for the spin label, on account of the minus sign introduced into our definition of $v_0(\hat{\mathbf{s}})$ in Eq. (6). The purpose of that minus sign was consistency in the labeling of the negative energy solutions; the way we have done it, both the spin and momentum eigenvalues of those solutions are the opposite of the spin and momentum labels. This convention will be convenient when we come to hole theory, in Notes 50.

6. Covariant Notation for Free Particle Solutions

Instead of labeling the u - and v -spinors by the 3-vectors \mathbf{p} and $\hat{\mathbf{s}}$, it is convenient to use 4-vectors, as a part of a notation that is as covariant as possible. As for the momentum, we will simply label the spinors by p , which stands for the 4-momentum p^μ in Eq. (15). Since the 4-momentum p is a function of the 3-momentum \mathbf{p} , and since the 3-momentum \mathbf{p} is the spatial part of the 4-momentum p , the new momentum label carries the same information as the old.

As for the spin, let us go back to the spinors before the boost, and promote the spin 3-vector $\hat{\mathbf{s}}$ into a 4-vector s_0 by defining

$$s_0^\mu = \begin{pmatrix} 0 \\ \hat{\mathbf{s}} \end{pmatrix}, \quad (24)$$

that is, s_0^μ is a purely spatial unit vector in the rest frame of the particle. Again, the 0-subscript means “at rest.” Notice that in the rest frame, the spin s_0 is purely space-like, while the momentum p_0 is purely time-like, so we have the dot products,

$$p_0 \cdot p_0 = m^2 c^2, \quad p_0 \cdot s_0 = 0, \quad s_0 \cdot s_0 = -1, \quad (25)$$

the last being equivalent to $\hat{\mathbf{s}} \cdot \hat{\mathbf{s}} = 1$.

Now when we boost the particle we boost the 4-vector s_0 along with it, writing

$$s = \Lambda(\mathbf{p})s_0, \quad \text{or} \quad s^\mu = \Lambda(\mathbf{p})^\mu{}_\nu s_0^\nu. \quad (26)$$

After the boost the 4-vector s is a unit space-like vector in the rest frame of the particle, that is, it lies in the 3-dimensional hyperplane that is Minkowski orthogonal to the 4-momentum p . Because the Minkowski scalar product is preserved under Lorentz transformations, the boosted 4-vectors p and s satisfy the same dot product relations as the 4-vectors p_0 and s_0 at rest, that is,

$$p \cdot p = m^2 c^2, \quad p \cdot s = 0, \quad s \cdot s = -1. \quad (27)$$

We will henceforth label the boosted spinors (22) by the 4-vectors p and s instead of the 3-vectors \mathbf{p} and $\hat{\mathbf{s}}$, that is, as $u(p, s)$ and $v(p, s)$. The notation is slightly illogical, since u and v are not functions of two independent 4-vectors p and s , rather p is a function of the 3-momentum \mathbf{p} and p and s must satisfy Eqs. (27). Nevertheless, the notation is useful and convenient. Likewise, when we need to refer to the spinors at rest, we will henceforth write them as $u(p_0, s_0)$, $v(p_0, s_0)$ rather than $u_0(\hat{\mathbf{s}})$, $v_0(\hat{\mathbf{s}})$, as above.

Altogether, in this covariant notation, a complete set of solutions of the free particle Dirac equation is obtained from

$$u(p, s)e^{-i(p \cdot x)/\hbar}, \quad v(p, s)e^{+i(p \cdot x)/\hbar}, \quad (28)$$

by letting \mathbf{p} run over all of 3-momentum space and allowing s to run over two oppositely oriented unit vectors orthogonal to p , that is, in the rest frame of the particle.

Notice that if we apply the energy-momentum 4-vector of operators $p_{\text{op}}^\mu = i\hbar \partial^\mu$ to these solutions we obtain

$$\begin{aligned} p_{\text{op}}^\mu u(p, s)e^{-i(p \cdot x)/\hbar} &= +p^\mu u(p, s)e^{-i(p \cdot x)/\hbar}, \\ p_{\text{op}}^\mu v(p, s)e^{+i(p \cdot x)/\hbar} &= -p^\mu v(p, s)e^{+i(p \cdot x)/\hbar}. \end{aligned} \quad (29)$$

For the negative energy solutions, the energy-momentum eigenvalues are opposite the energy-momentum label p of the solutions.

7. Geometrical Interpretation of the Spinors $u(p, s)$ and $v(p, s)$

A geometrical interpretation may be given of the manner in which the spinors $u(p, s)$ and $v(p, s)$ are parameterized. This takes place in 4-dimensional momentum space with coordinates p^μ , that is, E and \mathbf{p} . This space is sketched in Fig. 2, in which the z -component is suppressed to make a drawing possible.

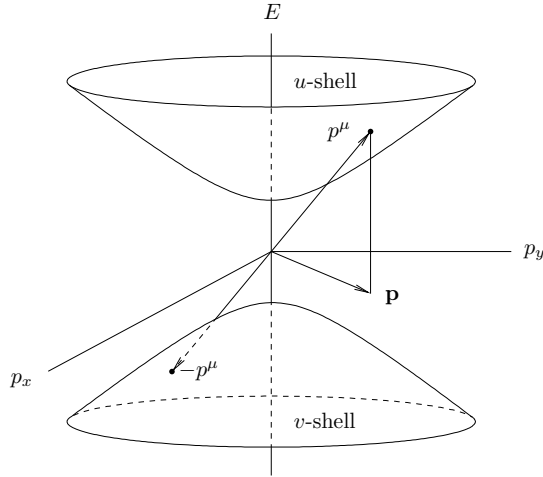


Fig. 2. The forward and backward mass shells, the surfaces upon which $E^2 = m^2c^4 + c^2\mathbf{p}^2$, are illustrated, in 4-momentum space. Given a 3-momentum \mathbf{p} (illustrated in only two dimensions in the figure), we can erect the vertical in the E direction, which intersects the forward mass shell at the 4-momentum p^μ . Directly opposite in 4-momentum space is the 4-momentum $-p^\mu$, on the backward mass shell. The u -spinors are associated with the point p^μ and the v -spinors with the point $-p^\mu$, so we call the forward mass shell the “ u -shell” and the backward one the “ v -shell.”

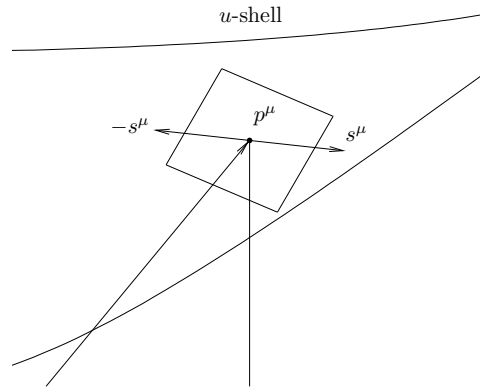


Fig. 3. An enlarged view of the point with coordinates p^μ on the forward (u) mass shell. The “plane” (really a 3-dimensional space) tangent to the mass shell at this point contains vectors that are Minkowski orthogonal to p^μ , that is, they are purely spatial in the rest frame of the particle. Two opposite unit vectors s^μ and $-s^\mu$ are chosen in this plane which specify the polarization of the spin.

Figure 3 contains a magnified picture of the point on the forward or u - mass shell with coordinates p^μ . The plane tangent to the mass shell at the point is illustrated. In the diagram it is a plane, but really it is a 3-dimensional space. Vectors in this plane are Minkowski orthogonal to p^μ , so they represent 4-vectors that are purely spatial in the rest frame of the particle. Two unit vectors in this plane are illustrated, s^μ and $-s^\mu$, which represent two polarizations for the spin, and correspond to the spinors $u(p, \pm s)$. A similar diagram applies on the backward or v -shell. By convention, the spinor $v(p, s)$ has 4-momentum $-p$ and is polarized in the direction $-s$.

8. The Spinors $u(p, s)$ and $v(p, s)$ as Covariant Eigenspinors

The spinors $u(p, s)$ and $v(p, s)$ can be characterized in covariant notation as eigenspinors of certain Dirac matrices. First, the spinors $u(p_0, s_0)$ and $v(p_0, s_0)$ at rest are eigenspinors of γ^0 , as we see from Eqs. (47.14) and (6),

$$\gamma^0 u(p_0, s_0) = u(p_0, s_0), \quad \gamma^0 v(p_0, s_0) = -v(p_0, s_0). \quad (30)$$

The eigenspaces of γ^0 are 2-dimensional. But

$$\gamma^0 = \frac{p_0 \cdot \gamma}{mc} = \frac{\not{p}_0}{mc}, \quad (31)$$

so we have

$$\not{p}_0 u(p_0, s_0) = mc u(p_0, s_0), \quad (32a)$$

$$\not{p}_0 v(p_0, s_0) = -mc v(p_0, s_0). \quad (32b)$$

Let us now boost these equations. Applying $D(\mathbf{p})$ to Eq. (32a), we find

$$\begin{aligned} D(\mathbf{p}) \not{p}_0 u(p_0, s_0) &= D(\mathbf{p})(p_0 \cdot \gamma) D(\mathbf{p})^{-1} D(\mathbf{p}) u(p_0, s_0) = (p_0 \cdot (\Lambda^{-1} \gamma)) u(p, s) \\ &= ((\Lambda p_0) \cdot \gamma) u(p, s) = (p \cdot \gamma) u(p, s) = \not{p} u(p, s) = mc D(\mathbf{p}) u_0 = mc u(p, s), \end{aligned} \quad (33)$$

where we use Eq. (47.42) in the second equality. Applying $D(\mathbf{p})$ in a similar manner to Eq. (32b) gives us altogether

$$(\not{p} - mc) u(p, s) = 0, \quad (\not{p} + mc) v(p, s) = 0. \quad (34)$$

The spinors $u(p, s)$ and $v(p, s)$ are eigenspinors of \not{p} with eigenvalues $+mc$ and $-mc$, respectively. Equations (32) are special cases of Eqs. (34) (for the particle at rest). Equations (34) gives us covariant way of characterizing the p -dependence of the spinors $u(p, s)$ and $v(p, s)$.

Please note that Eqs. (34) are purely spinor equations with no space-time dependence, and that \not{p} involves the c -number 4-vector p^μ , not the operator vector p_{op}^μ . The free particle Dirac equation in covariant form involves the operator 4-vector (see Eq. (47.8) or (47.13)),

$$(\not{p}_{\text{op}} - mc) \psi(x) = 0. \quad (35)$$

When we apply this to a positive energy solution, we get

$$(\not{p}_{\text{op}} - mc) u(p, s) e^{-i(p \cdot x)/\hbar} = (\not{p} - mc) u(p, s) e^{-i(p \cdot x)/\hbar} = 0, \quad (36)$$

since p_{op}^μ just brings out p^μ from the exponent. When we apply it to a negative energy solution, we get

$$(\not{p}_{\text{op}} - mc)v(p, s)e^{+i(p \cdot x)/\hbar} = (-\not{p} - mc)v(p, s)e^{+i(p \cdot x)/\hbar} = 0, \quad (37)$$

since p_{op}^μ brings out $-p^\mu$ from the exponent. We have used Eqs. (34) in the final step in both these results.

Equations (34) can be used to construct projection operators Π_\pm , which when applied to an arbitrary spinor project out the positive and negative energy components. These are

$$\Pi_\pm = \frac{mc \pm \not{p}}{2mc}, \quad (38)$$

and they satisfy

$$(\Pi_+)^2 = \Pi_+, \quad (\Pi_-)^2 = \Pi_-, \quad \Pi_+\Pi_- = \Pi_-\Pi_+ = 0, \quad \Pi_+ + \Pi_- = 1, \quad (39)$$

which says that they are complementary projectors. They also satisfy

$$\Pi_+u(p, s) = u(p, s), \quad \Pi_+v(p, s) = 0, \quad \Pi_-u(p, s) = 0, \quad \Pi_-v(p, s) = v(p, s), \quad (40)$$

which shows how they project out the positive and negative energy components of an arbitrary linear combination of positive and negative energy spinors.

One can also give a covariant way of characterizing the s -dependence of $u(p, s)$ and $v(p, s)$. It turns out that these spinors are eigenspinors of $\not{s}\gamma_5$, and that projection operators can be constructed that project out the positive or negative components of spin along the spin vector s . See Prob. 1.

9. Orthonormality Relations for the Spinors $u(p, s)$ and $v(p, s)$

The spinors $u(p, s)$ and $v(p, s)$ satisfy two types of orthonormality relations, one the adjoint and the other the Hermitian. The adjoint relations are a bit easier to derive, so we present these first. The relations are

$$\bar{u}(p, s)u(p, s') = \delta_{ss'}, \quad (41a)$$

$$\bar{u}(p, s)v(p, s') = \bar{v}(p, s)u(p, s') = 0, \quad (41b)$$

$$\bar{v}(p, s)v(p, s') = -\delta_{ss'}, \quad (41c)$$

where s' takes on only two values, $\pm s$, so that s' is a polarization the same as s or in the opposite direction in the rest frame. There is also the completeness relation,

$$1 = \sum_{\pm s} [u(p, s)\bar{u}(p, s) - v(p, s)\bar{v}(p, s)], \quad (42)$$

where 1 represents the 4×4 identity matrix, where the sum is taken over a polarization vector and its opposite, and the outer product of a column spinor times a row spinor is implied.

Equation (41a) may be proved as follows:

$$\begin{aligned}
\bar{u}(p, s)u(p, s') &= u^\dagger(p, s)\gamma^0 u(p, s') = u^\dagger(p_0, s_0)D(\mathbf{p})^\dagger \gamma^0 D(\mathbf{p})u(p_0, s'_0) \\
&= u^\dagger(p_0, s_0)\gamma^0 \gamma^0 D(\mathbf{p})^\dagger \gamma^0 D(\mathbf{p})u(p_0, s'_0) \\
&= u^\dagger(p_0, s_0)\gamma^0 D(\mathbf{p})^{-1} D(\mathbf{p})u(p_0, s'_0) = \bar{u}(p_0, s_0)u(p_0, s'_0),
\end{aligned} \tag{43}$$

where we use Eq. (47.78). Thus $\bar{u}(p, s)u(p, s')$ is unchanged when referred to the rest frame. But

$$\bar{u}(p_0, s_0)u(p_0, s'_0) = u^\dagger(p_0, s_0)\gamma^0 u(p_0, s'_0) = u^\dagger(p_0, s_0)u(p_0, s'_0) = (\chi^\dagger(\hat{\mathbf{s}}), 0) \begin{pmatrix} \chi(\hat{\mathbf{s}}') \\ 0 \end{pmatrix} = \delta_{ss'}, \tag{44}$$

where we use Eq. (30). Equations (41b) and (41c) may be proved similarly, as may the completeness relation (42).

The Hermitian orthonormality relations involve a modified 4-vector \tilde{p} , defined by $\tilde{p} = \Lambda(-\mathbf{p})p_0$, that is,

$$\tilde{p}^\mu = \Lambda(-\mathbf{p})^\mu{}_\nu p_0^\nu = \begin{pmatrix} E/c \\ -\mathbf{p} \end{pmatrix}. \tag{45}$$

These are

$$u^\dagger(p, s)u(p, s') = \frac{E}{mc^2} \delta_{ss'}, \tag{46a}$$

$$u^\dagger(p, s)v(\tilde{p}, s') = v^\dagger(p, s)u(\tilde{p}, s') = 0, \tag{46b}$$

$$v^\dagger(p, s)v(p, s') = \frac{E}{mc^2} \delta_{ss'}, \tag{46c}$$

where again, for given s , the vector s' takes on only the two values, $\pm s$. The proofs of these relations will be left as an exercise.

10. The Gordon Decomposition of the Current

The next topic gives us some insight into why the velocity operator $\mathbf{v} = c\boldsymbol{\alpha}$ is a purely spin operator in the Dirac theory, in contrast to the nonrelativistic theory. The velocity operator appears in the Heisenberg equations of motion, in the probability current, and other places. See the discussion in Sec. 45.8.

We will work with the covariant version of the Dirac equation for a particle interacting with an external electromagnetic field, Eq. (47.11). We write this equation in the form,

$$i\hbar \gamma^\nu \frac{\partial \psi}{\partial x^\nu} - \frac{q}{c} A_\nu \gamma^\nu \psi = mc \psi. \tag{47}$$

We multiply this on the left by $\bar{\psi}\gamma^\mu$, which causes the current to appear on the right hand side,

$$i\hbar \left(\bar{\psi}\gamma^\mu \gamma^\nu \frac{\partial \psi}{\partial x^\nu} \right) - \frac{q}{c} A_\nu (\bar{\psi}\gamma^\mu \gamma^\nu \psi) = mc (\bar{\psi}\gamma^\mu \psi) = mJ^\mu. \tag{48}$$

The right hand side of this equation is real but the left hand side is not obviously so. To bring out the reality, we take the complex conjugate of Eq. (48), which, after the use of Eq. (47.79), can be written

$$-i\hbar\left(\frac{\partial\bar{\psi}}{\partial x^\nu}\gamma^\nu\gamma^\mu\psi\right) - \frac{q}{c}A_\nu(\bar{\psi}\gamma^\nu\gamma^\mu\psi) = mJ^\mu. \quad (49)$$

We add these two together. The terms involving A_ν give

$$-\frac{q}{c}A_\nu[\bar{\psi}(\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu)\psi] = -\frac{q}{c}A_\nu[\bar{\psi}(2g^{\mu\nu})\psi] = -2\bar{\psi}\left(\frac{q}{c}A^\mu\right)\psi. \quad (50)$$

As for the terms involving the derivatives of ψ , we use

$$\gamma^\mu\gamma^\nu = \frac{1}{2}\{\gamma^\mu, \gamma^\nu\} + \frac{1}{2}[\gamma^\mu, \gamma^\nu] = g^{\mu\nu} - i\sigma^{\mu\nu} \quad (51)$$

(see Eqs. (47.18) and (47.55)), to obtain

$$i\hbar\left[\bar{\psi}(g^{\mu\nu} - i\sigma^{\mu\nu})\frac{\partial\psi}{\partial x^\nu}\right] - i\hbar\left[\frac{\partial\bar{\psi}}{\partial x^\nu}(g^{\mu\nu} + i\sigma^{\mu\nu})\psi\right] = 2\text{Re}(\bar{\psi}i\hbar\partial^\mu\psi) + \hbar\frac{\partial}{\partial x^\nu}(\bar{\psi}\sigma^{\mu\nu}\psi). \quad (52)$$

Altogether, we obtain

$$J^\mu = \frac{1}{m}\text{Re}\left[\bar{\psi}\left(i\hbar\partial^\mu - \frac{q}{c}A^\mu\right)\psi\right] + \frac{\hbar}{2m}\frac{\partial}{\partial x^\nu}(\bar{\psi}\sigma^{\mu\nu}\psi), \quad (53)$$

after dividing by $2m$.

Equation (53) is the Gordon decomposition of the Dirac current. The first term is obviously a generalization of the Klein-Gordon current (see Eq. (44.27)), with the addition of the interaction with an electromagnetic field. The Klein-Gordon current, in turn, is a relativistic generalization of the probability current for a spinless, nonrelativistic particle, given in Eq. (5.57). This term alone is called the *convection* current, since it involves only the purely spatial operator $p_{\text{op}}^\mu = i\hbar\partial^\mu$, which describes the state of motion of the particle. The second term in Eq. (53) is called the *magnetization* current, since it is the relativistic generalization of the magnetization current that was introduced in Sec. 18.6.

In summary, the Gordon decomposition of the current shows us that the velocity operator $\dot{\mathbf{x}} = c\boldsymbol{\alpha}$ in the Dirac theory, a purely spin operator, corresponds to the entire current of the nonrelativistic theory, including the magnetization current. We may point out that the Dirac version of the classical $\mathbf{v}\times\mathbf{B}$ force that a charged particle experiences in a magnetic field is $c\boldsymbol{\alpha}\times\mathbf{B}$, as shown by Eq. (45.41). That is, the magnetization current contributes to the $\mathbf{v}\times\mathbf{B}$ force, which is what we would expect physically.

11. The Zitterbewegung

Other strange features of the Dirac velocity operator $\mathbf{v} = \dot{\mathbf{x}} = c\boldsymbol{\alpha}$ that were discussed in Sec. 45.8 included the fact that the velocity is not constant, even for a free particle, and the fact that the different components of the velocity do not commute with one other. We return now to the

Heisenberg equations of motion, the subject of Sec. 45.8, working for simplicity with the case of a free particle.

Specializing the results of Sec. 45.8 to the case of a free particle, we obtain the Heisenberg equations of motion for \mathbf{x} and $\mathbf{p} = \boldsymbol{\pi}$, that is, $\dot{\mathbf{x}} = c\boldsymbol{\alpha}$, $\dot{\mathbf{p}} = 0$ (we henceforth omit the hats on the momentum operator). As for $\boldsymbol{\alpha}$, we work out the Heisenberg equations of motion as follows:

$$i\hbar \dot{\alpha}_i = [\alpha_i, H] = \alpha_i H - H \alpha_i = 2\alpha_i H - \alpha_i H - H \alpha_i = 2\alpha_i H - \{\alpha_i, H\}. \quad (54)$$

We write it this way because the anticommutator of $\boldsymbol{\alpha}$ with H is easier to work out than the commutator. Using Eqs. (45.8) we get the anticommutator,

$$\{\alpha_i, H\} = c\{\alpha_i, \boldsymbol{\alpha} \cdot \mathbf{p}\} = c\{\alpha_i, \alpha_j\} p_j = 2c \delta_{ij} p_j = 2cp_i. \quad (55)$$

Thus we have

$$\dot{\boldsymbol{\alpha}} = -\frac{2i}{\hbar}(\boldsymbol{\alpha}H - c\mathbf{p}). \quad (56)$$

What makes this equation easy to solve is the fact that both H and \mathbf{p} are constants of motion. One way to solve an operator equation like Eq. (56) is to interpret all the quantities as c -numbers, then to solve the c -number equation as an ordinary differential equation, then to use that solution to guess the solution to the operator equation, which can be checked by back-substitution. In this case, the quantities H and \mathbf{p} in the c -number equation are taken as constants. Of course operators do not commute in general while c -numbers do, so we can expect that the c -number equation can be wrong by an ordering of operators, which we must take into account when making the guess. In this case the procedure is straightforward; Eq. (56) interpreted as a c -number equation can be easily solved by Lagrange's method of variation of parameters, and the solution makes it easy to guess the solution to the operator equation. In this way we find the solution to the Heisenberg equation of motion (56),

$$\boldsymbol{\alpha}(t) = \frac{c\mathbf{p}}{H} + \left[\boldsymbol{\alpha}(0) - \frac{c\mathbf{p}}{H} \right] e^{-2iHt/\hbar}. \quad (57)$$

This can be checked by back substitution.

In this solution $c\mathbf{p}/H$ means the operator $c\mathbf{p}$ multiplied by the operator $1/H$. It does not matter which order we do the multiplication, because \mathbf{p} and H commute (otherwise the notation $c\mathbf{p}/H$ would be ambiguous). Also, we note that $1/H$ is defined because the spectrum of H (see Fig. 1) has an energy gap from $-mc^2$ to $+mc^2$ and does not include $E = 0$.

To interpret this equation we multiply through by c to get the velocity operator,

$$\mathbf{v}(t) = \frac{c^2\mathbf{p}}{H} + \left[\mathbf{v}(0) - \frac{c^2\mathbf{p}}{H} \right] e^{-2iHt/\hbar}. \quad (58)$$

We integrate this with respect to time, obtaining

$$\mathbf{x}(t) = \mathbf{x}(0) + \frac{c^2\mathbf{p}t}{H} + \left[\mathbf{v}(0) - \frac{c^2\mathbf{p}}{H} \right] \left[\frac{e^{-2iHt/\hbar} - 1}{(-2iH/\hbar)} \right], \quad (59)$$

the solution for the position operator $\mathbf{x}(t)$ in the Heisenberg picture.

The first term of Eq. (58), interpreted classically, is exactly the classical expression for the velocity as a function of the momentum in special relativity. See Eq. (45.39). It is, moreover, constant in time, as we expect for the velocity of a free particle in relativistic mechanics. We know, however, that the velocity does not commute with the Hamiltonian in the Dirac theory, so the velocity is not really a constant. The time dependence is contained in the second term in Eq. (58), which contains the factor $e^{-2iHt/\hbar}$.

To understand this time dependence better, we take the expectation value of Eq. (58) with respect to a state ψ that contains mainly positive energy solutions of energy close to $+mc^2$, that is, an essentially nonrelativistic state. This is all in the Heisenberg picture, so ψ is the initial wave function. Then the Hamiltonian applied to this state brings out a quantity of order mc^2 , and the operator $e^{-2iHt/\hbar}$ behaves like the phase factor $e^{-2imc^2t/\hbar}$. This factor oscillates on a time scale of order $\tau = \hbar/mc^2 \approx 1.3 \times 10^{-21}$ sec, which is of order $\alpha^2 \approx 1/(137)^2$ times smaller than the (already short) time scale in atomic units. See Table 17.1.

For $t \ll \tau$ we can replace the exponential $e^{-2iHt/\hbar}$ by simply 1, and Eqs. (58) and (59) become

$$\mathbf{v}(t) = \mathbf{v}(0), \quad t \ll \tau, \quad (60)$$

$$\mathbf{x}(t) = \mathbf{x}(0) + \mathbf{v}(0)t, \quad t \ll \tau. \quad (61)$$

On this time scale the expectation value of \mathbf{x} moves in a straight line. Also, since the eigenvalues of each component of α are ± 1 , the electron moves at a velocity comparable to the speed of light, so the distance covered in time τ , before the phase factor $e^{-2imc^2t/\hbar}$ has a chance to change sign, is of order

$$c\tau = \frac{\hbar}{mc}, \quad (62)$$

that is, the Compton wavelength.

If the particle is essentially nonrelativistic then the the expectation value of first term in Eq. (58) is $\ll c$, so it is small in comparison to the second term. But on time scales $\gg \tau$, the second term averages to zero, while the integral of the first term continues to accumulate. Thus on longer time scales, we have

$$\mathbf{v}(t) = \frac{c^2 \mathbf{p}}{E}, \quad t \gg \tau, \quad (63)$$

$$\mathbf{x}(t) = \mathbf{x}(0) + \frac{c^2 \mathbf{p} t}{E}, \quad t \gg \tau, \quad (64)$$

and the expectation value of the particle position follows a straight line at constant velocity, as we would expect from classical relativistic mechanics.

The rapid motion of the particle on a time scale of τ and a length scale of the Compton wavelength was first pointed out by Schrödinger in 1930, who called it the *Zitterbewegung*, German for “jittering motion.” An interesting feature of the *Zitterbewegung* is that it is not present in a wave packet consisting solely of positive energy solutions. See Prob. 5. The *Zitterbewegung*, when present, arises from the interference between the positive and negative energy solutions.

For this reason we may question if the negative energy solutions are real. Perhaps we should just declare them to be nonphysical, then we would not have to interpret them and the Zitterbewegung would go away.

Unfortunately, this does not work, for quite a few reasons. First of all, we can solve the Dirac equation for other problems besides the free particle. For example, the Dirac equation for the hydrogen atom is exactly solvable, and it gives positive energy solutions in good agreement with experiment, including the fine structure. These positive energy solutions can be expanded as linear combinations of free particle solutions, but the positive energy free particle solutions by themselves do not form a complete set. Thus, the expansion of (say) the ground state wave function of the Dirac hydrogen atom as a linear combination of free particle solutions will necessarily contain some negative energy components. In addition, the solution of the Dirac hydrogen atom also possesses negative energy solutions, and we have (at this point) no more physical interpretation for them than we do for the negative energy solutions for the free particle. But if we are going to declare them nonphysical, we now have a problem: Are we going to declare the negative energy free particle solutions nonphysical, or the negative energy hydrogen atom solutions? These two classes of negative energy solutions do not span the same subspace of the Dirac Hilbert space.

On the other hand, if we accept the negative energy free particle solutions, then we must have Zitterbewegung in any state that is a linear combination of both positive and negative free particle solutions. But we have just seen that the Dirac hydrogen atom wave functions are precisely such linear combinations. So does the Zitterbewegung exist in hydrogen? The answer is yes, it lies behind the Darwin term, as explained in Sec. 24.2. We will make a more careful analysis of the relation between the Darwin and other fine structure terms and the Dirac equation in a later set of notes, but for now we can see that there is no option to naively discard the negative energy solutions.

Ultimately we will see that the negative energy solutions of the free particle Dirac equation are related to positrons, the antiparticle of the electron, and that the Zitterbewegung is related to the fact that when an electron interacts with a potential, such as the potential of the nucleus in the hydrogen atom, there arises some finite amplitude for the formation of an electron-positron pair. That is, for brief periods of time, at least, there exist three particles instead of just one: two electrons and one positron. But in the early history of the Dirac equation, no one suspected the existence of antimatter, so these interpretations lay in the future.

12. Central Force Motion and Complete Sets of Commuting Observables

In the case of a central force potential the time-independent Dirac equation can be reduced to a pair of one-dimensional, first-order, coupled radial wave equations, a generalization of the single second-order radial wave equation that arises in the nonrelativistic theory (see Eqs. (16.7) and (16.11) for the latter). In some cases the Dirac radial wave equations can be solved in terms of standard special functions of mathematical physics. The main example that we are interested in is the Coulomb potential, but part of the analysis applies to any central force problem so we treat

that first. The Coulomb potential (for hydrogen-like atoms) is covered in Secs. 16–18. For the rest of these notes we use natural units ($\hbar = c = 1$), except at the end where we restore ordinary units when presenting results.

We seek the energy eigenvalues and eigenfunctions of the Dirac Hamiltonian in a central force potential, that is, we wish to solve

$$H\psi = E\psi, \quad (65)$$

where

$$H = \boldsymbol{\alpha} \cdot \mathbf{p} + V(r) + m\beta. \quad (66)$$

We identify the potential $V(r)$ with $q\Phi(r)$ in problems with electrostatic fields.

We appeal to the fine structure model of atoms (see Notes 24) for hints as to how to proceed, since this model incorporates relativistic effects to order $(v/c)^2$, relative to nonrelativistic energies. We recall that a complete set of commuting observables in the fine structure model is (H, L^2, J^2, J_z) . Of these, H , J^2 and J_z are commuting observables also in the relativistic (Dirac) problem, since $\mathbf{J} = \mathbf{L} + (1/2)\boldsymbol{\Sigma}$ commutes with the Hamiltonian (66). However, $[L^2, H] \neq 0$, as we find by a direct calculation of the commutator, so L^2 cannot be used as a member of the complete set in the relativistic theory and must be replaced by some other operator. In a moment we will see what the choices for this (fourth) operator are.

Accepting that J^2 and J_z will be members of our complete set, we begin by looking for the most general eigenfunction of these with quantum numbers (jm_j) . We work in the Dirac-Pauli representation of the Dirac matrices, in which the matrices $\boldsymbol{\Sigma}$ are block-diagonal,

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix}. \quad (67)$$

We break the 4-component Dirac spinor ψ into two 2-component Pauli spinors,

$$\psi = \begin{pmatrix} \phi \\ \chi \end{pmatrix}, \quad (68)$$

so that the action of the Dirac angular momentum $\mathbf{J} = \mathbf{L} + (1/2)\boldsymbol{\Sigma}$ on ψ is equivalent to the action of two Pauli angular momenta $\mathbf{J} = \mathbf{L} + (1/2)\boldsymbol{\sigma}$ acting separately on ϕ and χ . Thus the eigenfunctions of J^2 and J_z in the Dirac theory are 4-component spinors ψ in which both ϕ and χ are eigenfunctions of the Pauli operators J^2 and J_z , with the same quantum numbers (jm_j) .

Eigenfunctions of (J^2, J_z) in the Pauli theory are conveniently expressed in terms of *spinor spherical harmonics*, which are 2-component spinors defined over the unit sphere. We define them by

$$\mathcal{Y}_{jm_j}^\ell(\theta, \phi) = \sum_{m_\ell m_s} Y_{\ell m_\ell}(\theta, \phi) \chi_{m_s} \langle \ell s m_\ell m_s | jm_j \rangle, \quad (69)$$

where χ_{m_s} for $m_s = \pm\frac{1}{2}$ are the spin-up and spin-down spinors (the unit vectors in spin space), and where in the Clebsch-Gordan coefficient s has the value $\frac{1}{2}$. The spinor spherical harmonics are eigenfunctions of L^2 as well as J^2 and J_z , with quantum number ℓ , which, for a given value

of j can take on only the values $\ell = j \pm 1/2$. Of course, j is a half-integer, $j = \frac{1}{2}, \frac{3}{2}, \dots$. We are seeing eigenfunctions of L^2 here because they arise when we construct eigenfunctions of J^2 and J_z by the usual method of coupling $\ell \otimes \frac{1}{2}$, but since $[L^2, H] \neq 0$ we do not expect to find energy eigenfunctions of the Dirac Hamiltonian that are also eigenfunctions of L^2 . In any case, the spin spherical harmonics are eigenfunctions of L^2 only as Pauli spinors, not as Dirac spinors.

The Clebsch-Gordan coefficients in Eq. (69) can be worked out by the method of Prob. 18.3. These allow us to write out the explicit forms of the spin spherical harmonics,

$$\begin{aligned}\mathcal{Y}_{jm_j}^{j-1/2} &= \frac{1}{\sqrt{2j}} \begin{pmatrix} \sqrt{j+m_j} Y_{j-1/2, m_j-1/2} \\ \sqrt{j-m_j} Y_{j-1/2, m_j+1/2} \end{pmatrix}, \\ \mathcal{Y}_{jm_j}^{j+1/2} &= \frac{1}{\sqrt{2j+2}} \begin{pmatrix} -\sqrt{j-m_j+1} Y_{j+1/2, m_j-1/2} \\ \sqrt{j+m_j+1} Y_{j+1/2, m_j+1/2} \end{pmatrix},\end{aligned}\tag{70}$$

where each \mathcal{Y} and each Y is understood as a function of (θ, ϕ) . We note that the spinor spherical harmonics are orthonormal,

$$\int d\Omega (\mathcal{Y}_{jm_j}^\ell)^\dagger \mathcal{Y}_{j'm_j'}^{\ell'} = \delta_{\ell\ell'} \delta_{jj'} \delta_{m_j m_j'}.\tag{71}$$

We can now write down the most general 2-component eigenfunction of (J^2, J_z) with quantum numbers (jm_j) . It must be a linear combination of $\mathcal{Y}_{jm_j}^\ell$ for the two values $\ell = j \pm \frac{1}{2}$, where each of the coefficients is allowed to be an arbitrary function of r ,

$$A(r)\mathcal{Y}_{jm_j}^{j-1/2}(\theta, \phi) + B(r)\mathcal{Y}_{jm_j}^{j+1/2}(\theta, \phi).\tag{72}$$

From this it follows that the most general 4-component spinor that is an eigenfunction of the Dirac operators (J^2, J_z) with eigenvalues (jm_j) is of the form,

$$\psi_{jm_j}(\mathbf{x}) = \begin{pmatrix} A(r)\mathcal{Y}_{jm_j}^{j-1/2} + B(r)\mathcal{Y}_{jm_j}^{j+1/2} \\ C(r)\mathcal{Y}_{jm_j}^{j-1/2} + D(r)\mathcal{Y}_{jm_j}^{j+1/2} \end{pmatrix},\tag{73}$$

where $A(r)$, $B(r)$, $C(r)$ and $D(r)$ are four arbitrary radial wave functions and where the upper and lower 2-component spinors have the same (jm_j) values. The radial wave functions will be determined by demanding that ψ be an eigenfunction of the remaining members of our complete set of commuting observables.

13. The Fourth Observable and the Operator K

Now we can see some choices for the fourth observable in the complete set of the Dirac theory, the one that will replace L^2 in the fine structure model. The simplest choice is parity, which in the Dirac theory is discussed in Sec. 47.14. We write the Dirac parity operator as

$$\pi = \pi_s \gamma^0,\tag{74}$$

where π_s takes care of the spatial part of parity,

$$(\pi_s \psi)(\mathbf{x}) = \psi(-\mathbf{x}), \quad (75)$$

and where γ^0 takes care of the spin part. The parity operator has an arbitrary \pm sign (see Sec. 47.14), which we have chosen in Eq. (75) so as to make the intrinsic parity of the electron $+1$. In the case of 2-component (Pauli) spinors, the parity operator consists of π_s alone, which takes $\phi(\mathbf{x})$ into $\phi(-\mathbf{x})$. See Notes 20.

When we apply π_s to the Dirac spinor (73) it brings out $(-1)^\ell$ in each term, that is, $(-1)^{j\pm 1/2}$. Thus the A and B terms acquire opposite signs, as do the C and D terms. But when we multiply the spinor by γ^0 , the lower 2-component spinor acquires an additional phase of -1 . Therefore if we require that the Dirac spinor (73) be an eigenstate of parity, then either $B = C = 0$, in which case the parity is $(-1)^{j-1/2}$, or else $A = D = 0$, in which case the parity is $(-1)^{j+1/2}$. We will write the simultaneous eigenfunctions of (π, J^2, J_z) as

$$\psi_{\kappa j m_j}(\mathbf{x}) = \begin{pmatrix} F(r) \mathcal{Y}_{j m_j}^{j-1/2} \\ iG(r) \mathcal{Y}_{j m_j}^{j+1/2} \end{pmatrix}, \quad \psi_{\kappa j m_j}(\mathbf{x}) = \begin{pmatrix} F(r) \mathcal{Y}_{j m_j}^{j+1/2} \\ iG(r) \mathcal{Y}_{j m_j}^{j-1/2} \end{pmatrix}, \quad (76)$$

where the first and second spinors have parity $(-1)^{j-1/2}$ and $(-1)^{j+1/2}$, respectively, where the arbitrary radial wave functions are now written as F and G and where for later convenience we have split off a factor of i in the lower component. We see that indeed these spinors are not eigenfunctions of L^2 , which has a different value in the upper and lower components.

We have introduced a new quantum number κ in Eq. (76), which is related to the eigenvalue of a new operator K . This operator serves as well as parity for creating a complete set, but it is not as obvious a choice. It is defined by

$$K = \beta(\mathbf{L} \cdot \boldsymbol{\Sigma} + 1). \quad (77)$$

This is a scalar under spatial rotations so it commutes with \mathbf{J} . It also commutes with parity π , that is, it is a true scalar, not a pseudoscalar.

The fact that it also commutes with H follows from a straightforward but not very illuminating calculation. First note the commutator $[\beta, \boldsymbol{\Sigma}] = 0$ and the identity,

$$\alpha_i \Sigma_j = \Sigma_i \alpha_j = \delta_{ij} \gamma_5 + i\epsilon_{ijk} \alpha_k, \quad (78)$$

which follows from the definitions (45.11) and (47.104) and which implies

$$\{\alpha_i, \Sigma_j\} = 2\delta_{ij} \gamma_5, \quad [\alpha_i, \Sigma_j] = 2i\epsilon_{ijk} \alpha_k. \quad (79)$$

To compute $[K, H]$ we need the commutator,

$$\begin{aligned} [\beta(\mathbf{L} \cdot \boldsymbol{\Sigma}), \boldsymbol{\alpha} \cdot \mathbf{p}] &= \beta(\mathbf{L} \cdot \boldsymbol{\Sigma})(\boldsymbol{\alpha} \cdot \mathbf{p}) - (\boldsymbol{\alpha} \cdot \mathbf{p})\beta(\mathbf{L} \cdot \boldsymbol{\Sigma}) = \beta[(\mathbf{L} \cdot \boldsymbol{\Sigma})(\boldsymbol{\alpha} \cdot \mathbf{p}) + (\boldsymbol{\alpha} \cdot \mathbf{p})(\mathbf{L} \cdot \boldsymbol{\Sigma})] \\ &= \beta\{\mathbf{L} \cdot \boldsymbol{\Sigma}, \boldsymbol{\alpha} \cdot \mathbf{p}\}, \end{aligned} \quad (80)$$

which is expressed in terms of an anticommutator as shown. As for the latter, we migrate factors in the second term,

$$(\boldsymbol{\alpha} \cdot \mathbf{p})(\mathbf{L} \cdot \boldsymbol{\Sigma}) = \alpha_i p_i L_j \Sigma_j = \alpha_i L_j p_i \Sigma_j + i\epsilon_{ijk} \alpha_i p_k \Sigma_j, \quad (81)$$

using Eq. (19.20). We continue to migrate factors in the first term in this expression,

$$\alpha_i L_j p_i \Sigma_j = \alpha_i L_j \Sigma_j p_i = L_j \alpha_i \Sigma_j p_i = -L_j \Sigma_j \alpha_i p_i + 2\delta_{ij} \gamma_5 L_j p_i = -(\mathbf{L} \cdot \boldsymbol{\Sigma})(\boldsymbol{\alpha} \cdot \mathbf{p}), \quad (82)$$

since $\mathbf{L} \cdot \mathbf{p} = 0$. This cancels the first term of the anticommutator in Eq. (80). As for the second term in Eq. (81), it is

$$i\epsilon_{ijk} \alpha_i p_k \Sigma_j = i\epsilon_{ijk} \alpha_i \Sigma_j p_k = i\epsilon_{ijk} (\delta_{ij} \gamma_5 + i\epsilon_{ij\ell} \alpha_\ell) p_k = -2\delta_{k\ell} \alpha_\ell p_k, \quad (83)$$

which gives finally for the commutator (80),

$$[\beta(\mathbf{L} \cdot \boldsymbol{\Sigma}), \boldsymbol{\alpha} \cdot \mathbf{p}] = -2\beta(\boldsymbol{\alpha} \cdot \mathbf{p}). \quad (84)$$

The other commutators we need are $[K, \beta] = 0$ and

$$[\beta, \boldsymbol{\alpha} \cdot \mathbf{p}] = 2\beta(\boldsymbol{\alpha} \cdot \mathbf{p}), \quad (85)$$

so that on adding Eqs. (84) and (85) and assembling the pieces we obtain $[K, H] = 0$.

In view of these commutators we will henceforth take (H, K, J^2, J_z) as our complete set of commuting operators, although parity π could be used in place of K .

To obtain the spectrum of K we first note that for the Dirac angular momentum $\mathbf{J} = \mathbf{L} + (1/2)\boldsymbol{\Sigma}$ we have

$$J^2 = L^2 + \mathbf{L} \cdot \boldsymbol{\Sigma} + \frac{3}{4}. \quad (86)$$

Next we look at the square of K ,

$$K^2 = \beta(\mathbf{L} \cdot \boldsymbol{\Sigma} + 1)\beta(\mathbf{L} \cdot \boldsymbol{\Sigma} + 1) = (\mathbf{L} \cdot \boldsymbol{\Sigma} + 1)^2 = (\mathbf{L} \cdot \boldsymbol{\Sigma})^2 + 2\mathbf{L} \cdot \boldsymbol{\Sigma} + 1. \quad (87)$$

We evaluate the first term of this result with the help of Eq. (1.148),

$$(\mathbf{L} \cdot \boldsymbol{\Sigma})^2 = L^2 + i\boldsymbol{\Sigma} \cdot (\mathbf{L} \times \mathbf{L}), \quad (88)$$

which is a little tricky, since $\mathbf{L} \times \mathbf{L} = i\mathbf{L}$ is not zero, rather it is equivalent to the commutation relations $[L_i, L_j] = i\epsilon_{ijk} L_k$. Thus we obtain

$$K^2 = L^2 + \mathbf{L} \cdot \boldsymbol{\Sigma} + 1 = J^2 + \frac{1}{4}, \quad (89)$$

and when we apply K^2 to an eigenstate of J^2 with eigenvalue j , it brings out $j(j+1)+1/4 = (j+1/2)^2$. Therefore the eigenvalues of K are $\pm(j+1/2)$.

To apply K to the wave functions (76) we need the action of the Pauli operator $\mathbf{L} \cdot \boldsymbol{\sigma}$ on the spinor spherical harmonics. Since the Pauli angular momentum is given by $\mathbf{J} = \mathbf{L} + (1/2)\boldsymbol{\sigma}$, we have

$$J^2 = L^2 + \mathbf{L} \cdot \boldsymbol{\sigma} + \frac{3}{4}, \quad (90)$$

which looks like Eq. (86) but is a Pauli, not a Dirac, operator. This implies

$$(\mathbf{L} \cdot \boldsymbol{\sigma}) \mathcal{Y}_{jm_j}^\ell = \left[j(j+1) - \ell(\ell+1) - \frac{3}{4} \right] \mathcal{Y}_{jm_j}^\ell, \quad (91)$$

which in turn implies, for the two values $\ell = j \pm 1/2$,

$$(\mathbf{L} \cdot \boldsymbol{\sigma} + 1) \mathcal{Y}_{jm_j}^\ell = \begin{cases} +(j + \frac{1}{2}) \mathcal{Y}_{jm_j}^\ell, & \ell = j - \frac{1}{2}, \\ -(j + \frac{1}{2}) \mathcal{Y}_{jm_j}^\ell, & \ell = j + \frac{1}{2}, \end{cases} \quad (92)$$

These show that if we apply the Pauli operator $\mathbf{L} \cdot \boldsymbol{\sigma} + 1$ to the upper and lower components of the Dirac spinors (76), we bring out a factor of $+(j + \frac{1}{2})$ in one component and $-(j + \frac{1}{2})$ in the other. But then when we multiply by β , the entire Dirac spinor acquires an overall factor, which is the eigenvalue $(j + \frac{1}{2})$ of K for the first spinor in Eq. (76) and $-(j + \frac{1}{2})$ for the second. Thus the eigenvalue of K distinguishes the two spinors in Eq. (76). It is customary in the literature to denote the negative of the eigenvalue of K by κ , which has the value $\kappa = -(j + \frac{1}{2})$ for the first spinor in Eq. (76) and $\kappa = +(j + \frac{1}{2})$ in the second. The extra minus sign is confusing, but we fear greater confusion if we change the convention. Table 1 summarizes the relations among κ , the parity, and the angular momentum quantum number ℓ attached to the upper and lower components of the Dirac spinors (76).

κ	$\ell(\text{upper})$	$\ell(\text{lower})$	π
$-(j + \frac{1}{2})$	$j - \frac{1}{2}$	$j + \frac{1}{2}$	$(-1)^{j-1/2}$
$+(j + \frac{1}{2})$	$j + \frac{1}{2}$	$j - \frac{1}{2}$	$(-1)^{j+1/2}$

Table 1. Values of ℓ quantum number for upper and lower spinors and overall parity, as a function of k , the eigenvalue of K .

Recall that in the nonrelativistic limit, the upper 2-component spinor of the Dirac theory becomes the Pauli spinor of the nonrelativistic theory. See Sec. 45.9. In the case of the Dirac spinors (76), we see that the ℓ quantum number of the nonrelativistic limit is the one labeled $\ell(\text{upper})$ in Table 1, and that the radial wave function $F(r)$ of the relativistic theory becomes $R(r)$ of the nonrelativistic theory. The radial wave function $R(r)$ is discussed in Sec. 16.2.

It is common in the literature to label the eigenfunctions of (K, J^2, J_z) by $\psi_{\ell jm_j}$ rather than $\psi_{\kappa jm_j}$ as we have done here, where ℓ refers to $\ell(\text{upper})$ in Table 1. This has the advantage of labeling the solutions of the Dirac equation by the same quantum numbers that apply in the nonrelativistic limit, but of course it does not imply that the Dirac wave functions are eigenfunctions of L^2 .

14. The Action of $\boldsymbol{\sigma} \cdot \mathbf{p}$

Equation (76) now gives us the simultaneous eigenfunctions of (K, J^2, J_z) . When we demand that these also be eigenfunctions of H we obtain equations that will determine the radial wave

functions $F(r)$ and $G(r)$. We substitute the spinors (76) into the Dirac equation, obtaining

$$\left[\begin{pmatrix} 0 & \boldsymbol{\sigma} \cdot \mathbf{p} \\ \boldsymbol{\sigma} \cdot \mathbf{p} & 0 \end{pmatrix} + m \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + V(r) - E \right] \begin{pmatrix} F(r) \mathcal{Y}_{jm_j}^{j\pm 1/2} \\ iG(r) \mathcal{Y}_{jm_j}^{j\pm 1/2} \end{pmatrix} = 0, \quad (93)$$

where the upper sign refers to the first spinor (76) and the lower sign to the second. The corresponding values of κ are shown in Table 1.

This makes it evident that we will need to evaluate expressions of the form $(\boldsymbol{\sigma} \cdot \mathbf{p})A(r)\mathcal{Y}_{jm_j}^{j\pm 1/2}$, where $A(r)$ is some radial function. It is easier to evaluate $(\boldsymbol{\sigma} \cdot \mathbf{x})A(r)\mathcal{Y}_{jm_j}^{j\pm 1/2}$ (with \mathbf{x} instead of \mathbf{p}), because it is only a multiplicative operator on the \mathbf{x} -space wave functions and not a differential operator. Moreover, once we have the latter we can obtain the former by using the identity,

$$(\boldsymbol{\sigma} \cdot \mathbf{p})(\boldsymbol{\sigma} \cdot \mathbf{x}) = \mathbf{p} \cdot \mathbf{x} - i(\boldsymbol{\sigma} \cdot \mathbf{L}) = -i \left[r \frac{\partial}{\partial r} + 3 + \boldsymbol{\sigma} \cdot \mathbf{L} \right], \quad (94)$$

where we use Eq. (1.148) and the fact that $x_i(\partial/\partial x_i) = r(\partial/\partial r)$.

Therefore we turn our attention to evaluating

$$(\boldsymbol{\sigma} \cdot \mathbf{x})A(r)\mathcal{Y}_{jm_j}^{j\pm 1/2}(\theta, \phi) = rA(r) \begin{pmatrix} \cos \theta & \sin \theta e^{-i\phi} \\ \sin \theta e^{i\phi} & -\cos \theta \end{pmatrix} \mathcal{Y}_{jm_j}^{j\pm 1/2}(\theta, \phi). \quad (95)$$

We see that under the action of $(\boldsymbol{\sigma} \cdot \mathbf{x})$ the radial part of the wave function $A(r)$ just gets multiplied by r . As for the angular part, it will be desirable to expand it as a linear combination of spin spherical harmonics. To do this we freeze the radial coordinate r and work with the Hilbert space of spinors over a sphere, in which the spin spherical harmonics $\mathcal{Y}_{jm_j}^\ell$ form a complete orthonormal basis. We denote these in ket language by $|\ell jm_j\rangle$, where $j = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots$, $m_j = -j, \dots, +j$, and $\ell = j \pm \frac{1}{2}$. The radial coordinate r can be regarded as just a constant for this part of the calculation.

In this notation we wish to evaluate

$$(\boldsymbol{\sigma} \cdot \mathbf{x})|\ell jm_j\rangle = \sum_{\ell' j' m'_j} |\ell' j' m'_j\rangle \langle \ell' j' m'_j | \boldsymbol{\sigma} \cdot \mathbf{x} | \ell jm_j \rangle. \quad (96)$$

This sum simplifies. Since $\boldsymbol{\sigma} \cdot \mathbf{x}$ is a scalar under rotations generated by \mathbf{J} , it commutes with \mathbf{J} and hence with both J^2 and J_z , so its matrix elements in the basis $|\ell jm_j\rangle$ are diagonal in both j and m_j . Moreover it anticommutes with parity π_s (it is a pseudoscalar, not a scalar), so it must map a state with quantum number $\ell = j \pm \frac{1}{2}$ into one with quantum number $\ell = j \mp \frac{1}{2}$. This means that ℓ' in the sum (96) can only take on the value $\ell' = \tilde{\ell}$, where if $\ell = j \pm \frac{1}{2}$ then $\tilde{\ell} = j \mp \frac{1}{2}$. Thus the sum (96) reduces to a single term, which in the two cases $\ell = j \pm \frac{1}{2}$ can be written

$$\begin{aligned} (\boldsymbol{\sigma} \cdot \mathbf{x})|j - \tfrac{1}{2}, jm_j\rangle &= \alpha r |j + \tfrac{1}{2}, jm_j\rangle, \\ (\boldsymbol{\sigma} \cdot \mathbf{x})|j + \tfrac{1}{2}, jm_j\rangle &= \beta r |j - \tfrac{1}{2}, jm_j\rangle, \end{aligned} \quad (97)$$

where α and β are given by

$$\alpha r = \langle j + \tfrac{1}{2}, jm_j | (\boldsymbol{\sigma} \cdot \mathbf{x}) | j - \tfrac{1}{2}, jm_j \rangle, \quad \beta r = \langle j - \tfrac{1}{2}, jm_j | (\boldsymbol{\sigma} \cdot \mathbf{x}) | j + \tfrac{1}{2}, jm_j \rangle. \quad (98)$$

But since $\boldsymbol{\sigma} \cdot \mathbf{x}$ is Hermitian, these imply $\beta = \alpha^*$. Finally, since $(\boldsymbol{\sigma} \cdot \mathbf{x})^2 = r^2$, we see that α is a phase factor, $|\alpha|^2 = 1$.

There is some effort in evaluating the final phase α . A straightforward method is simply to apply the matrix in Eq. (95) to the spinors (70), noting that the components of the matrix can be written in terms of the $Y_{\ell m}$'s for $\ell = 1$, as in Eq. (19.52). The product of $Y_{\ell m}$'s can then be expressed in terms of single $Y_{\ell m}$'s by means of the 3- $Y_{\ell m}$ formula, Eq. (18.59), and finally the result can be expanded in terms of spinor spherical harmonics. After some algebra along these lines, we find $\alpha = -1$. We summarize the results by writing

$$(\boldsymbol{\sigma} \cdot \mathbf{x})A(r)\mathcal{Y}_{jm_j}^\ell = -rA(r)\mathcal{Y}_{jm_j}^{\tilde{\ell}}, \quad (99)$$

where again if $\ell = j \pm \frac{1}{2}$, then $\tilde{\ell} = j \mp \frac{1}{2}$.

Now we divide both sides of Eq. (99) by r and apply $(\boldsymbol{\sigma} \cdot \mathbf{p})$, using Eqs. (94) and (91). We find

$$-i\left[r\frac{\partial}{\partial r} + 3 + j(j+1) - \ell(\ell+1) - \frac{3}{4}\right]\left[\frac{A(r)}{r}\right]\mathcal{Y}_{jm_j}^\ell = -(\boldsymbol{\sigma} \cdot \mathbf{p})A(r)\mathcal{Y}_{jm_j}^{\tilde{\ell}}. \quad (100)$$

Writing this out for the two cases $\ell = j \pm \frac{1}{2}$, we obtain

$$\begin{aligned} (\boldsymbol{\sigma} \cdot \mathbf{p})A(r)\mathcal{Y}_{jm_j}^{j-1/2} &= i\left[\frac{dA}{dr} - \frac{j-1/2}{r}A(r)\right]\mathcal{Y}_{jm_j}^{j+1/2}, \\ (\boldsymbol{\sigma} \cdot \mathbf{p})A(r)\mathcal{Y}_{jm_j}^{j+1/2} &= i\left[\frac{dA}{dr} + \frac{j+3/2}{r}A(r)\right]\mathcal{Y}_{jm_j}^{j-1/2}, \end{aligned} \quad (101)$$

which is the desired action of $\boldsymbol{\sigma} \cdot \mathbf{p}$ on 2-component eigenfunctions of (L^2, J^2, J_z) .

15. The Radial Wave Equations

Now we may substitute Eq. (101) into the Dirac equation (93). While we are at it we will express the radial wave functions $F(r)$, $G(r)$ in terms of alternative radial wave functions $f(r)$ and $g(r)$, defined by

$$F(r) = \frac{f(r)}{r}, \quad G(r) = \frac{g(r)}{r}, \quad (102)$$

which are just like the transformation (16.10) used in the nonrelativistic theory to create what we called version two of the radial wave equation.

The result is two coupled, 2-component equations equations for the upper and lower components of the Dirac spinor, from which the common spinor spherical harmonic can be canceled, leaving two coupled radial wave equations for $f(r)$ and $g(r)$. These are

$$\begin{aligned} \frac{df}{dr} \mp \frac{j+1/2}{r}f(r) &= [m + E - V(r)]g(r), \\ \frac{dg}{dr} \pm \frac{j+1/2}{r}g(r) &= [m - E + V(r)]f(r), \end{aligned} \quad (103)$$

where as in Eq. (93) the upper sign corresponds to the first spinor (76) and the lower sign to the second. These equations can also be written in terms of the quantum number κ ,

$$\frac{df}{dr} + \frac{\kappa}{r}f(r) = [m + E - V(r)]g(r), \quad (104a)$$

$$\frac{dg}{dr} - \frac{\kappa}{r}g(r) = [m - E + V(r)]f(r), \quad (104b)$$

where $\kappa = \mp(j + 1/2)$, as indicated in Table 1. These are the radial wave equations in the Dirac theory that replace version two of the radial wave equation, Eq. (16.11), in the Schrödinger theory.

The nonrelativistic limit of these equations is obtained by the method of Sec. 45.9. That is, we approximate $E \approx m \gg V$ so that the right hand side of Eq. (104a) becomes $2mg(r)$, which allows that equation to be solved for g in terms of f . This is then substituted into Eq. (104b) to eliminate g and obtain an equation purely for f . This turns out to be version two of the radial Schrödinger equation, Eq. (16.11).

16. The Hydrogen Atom

We now specialize to the case $V(r) = -Ze^2/r$ for a hydrogen-like atom. We will solve the Dirac equation (65) for the bound states in this potential. Note that in natural units $e^2 = \alpha \approx 1/137$, so that $Ze^2 = Z\alpha$ is a dimensionless parameter of the problem. This parameter ranges from about $1/137$ for hydrogen to about $2/3$ for uranium and beyond, remaining, however, always less than unity for all known nuclei. As we will see, the ground state of the Dirac equation is a bound state only for $Z\alpha < 1$, that is, $Z < \approx 137$, and becomes unbound for higher Z . This lends some interest to the question of whether nuclei with $Z > 137$ exist or could be made, even briefly. Note, however, that the potential $Z\alpha/r$ is an idealization and that in the real world it must be modified at small radii because nuclei are not point charges. The corrections required for this become more important as Z increases. See Sec. 17.8 and Prob. 22.1.

Since the probability density in the Dirac theory is $\psi^\dagger\psi$, the normalization integral for normalized bound states is

$$\int d^3\mathbf{x} \psi^\dagger(\mathbf{x})\psi(\mathbf{x}) = 1. \quad (105)$$

If we specialize this to Dirac wave functions of the form (76) and carry out the integral in spherical coordinates, we have

$$\int d^3\mathbf{x} \psi^\dagger\psi = \int_0^\infty r^2 dr (|F(r)|^2 + |G(r)|^2) = \int_0^\infty dr (|f(r)|^2 + |g(r)|^2) = 1, \quad (106)$$

where we use Eq. (71). We see that a Dirac central force eigenfunction is normalizable if and only if both $f(r)$ and $g(r)$ are normalizable as one-dimensional wave functions on the interval $[0, \infty)$.

We first look at the Dirac equation when r is large, so that the potential energy and the terms that involve κ/r can be neglected. Then Eqs. (104) become

$$\frac{df}{dr} = (m + E)g, \quad \frac{dg}{dr} = (m - E)f, \quad (107)$$

or,

$$\frac{d^2 f}{dr^2} = (m^2 - E^2)f, \quad \frac{d^2 g}{dr^2} = (m^2 - E^2)g. \quad (108)$$

If $|E| > m$ these have oscillatory solutions that are not normalizable and cannot represent bound states. Likewise if $|E| = m$ the solutions are of the form $c_1 + c_2 r$, which are not normalizable. We see that the Dirac hydrogen-like atom has a continuous spectrum for $E \geq m$, which corresponds to the positive energy, continuous spectrum of the nonrelativistic hydrogen atom. But the Dirac equation also possesses a continuous spectrum for $E \leq -m$, that is, it possesses negative energy solutions, just like the free particle.

Finally, if $|E| < m$ then the solutions go as

$$f(r), g(r) \sim \exp(\pm \sqrt{m^2 - E^2} r), \quad r \rightarrow \infty, \quad (109)$$

which are normalizable if we choose the minus sign. We see that bound states are possible only in the range $|E| < m$, that is, $-m < E < m$, and that the eigenfunctions must behave asymptotically as $\exp(-\sqrt{m^2 - E^2} r)$.

We define a new radial variable,

$$\rho = 2\sqrt{m^2 - E^2} r, \quad (110)$$

so that the asymptotic behavior of the bound state wave functions will be $e^{-\rho/2}$. Then Eqs. (104) become

$$\begin{aligned} \frac{df}{d\rho} + \frac{\kappa}{\rho} f &= \left(\frac{1}{2\gamma} + \frac{Z\alpha}{\rho} \right) f, \\ \frac{dg}{d\rho} - \frac{\kappa}{\rho} g &= \left(\frac{\gamma}{2} - \frac{Z\alpha}{\rho} \right), \end{aligned} \quad (111)$$

where

$$\gamma = \sqrt{\frac{m - E}{m + E}}. \quad (112)$$

Next we define new radial wave functions $a(\rho)$ and $b(\rho)$, in which the asymptotic behavior has been stripped off,

$$f(r) = e^{-\rho/2} a(\rho), \quad g(r) = e^{-\rho/2} b(\rho), \quad (113)$$

which cause Eqs. (111) to become

$$\frac{da}{d\rho} - \frac{a}{2} + \frac{\kappa}{\rho} a = \left(\frac{1}{2\gamma} + \frac{Z\alpha}{\rho} \right) b, \quad (114a)$$

$$\frac{db}{d\rho} - \frac{b}{2} - \frac{\kappa}{\rho} b = \left(\frac{\gamma}{2} - \frac{Z\alpha}{\rho} \right) a. \quad (114b)$$

We examine the behavior of the solutions near $\rho = 0$. We assume that the dominant behavior of $a(\rho)$ and $b(\rho)$ when ρ is small is $a_0 \rho^s$ and $b_0 \rho^t$, respectively, for some powers s and t , where $a_0 \neq 0$ and $b_0 \neq 0$. Then the left hand side of Eq. (114a) is dominated by $(s + \kappa)a_0 \rho^{s-1}$ and the right hand side by $b_0 Z\alpha \rho^{t-1}$. Therefore if $t < s$ the right hand side dominates, and we must have $b_0 Z\alpha = 0$ or

$b_0 = 0$. But this contradicts our assumption about the leading behavior, so we cannot have $t < s$ and we must have $s \geq t$. A similar argument applied to Eq. (114b) shows that we must have $t \geq s$. Taken together, these imply $t = s$. We see that both $a(\rho)$ and $b(\rho)$ begin with the same power s .

It then follows that to leading order both sides of both Eqs. (114) go as ρ^{s-1} , and that the coefficients must satisfy

$$\begin{aligned}(s + \kappa)a_0 &= Z\alpha b_0, \\ (s - \kappa)b_0 &= -Z\alpha a_0\end{aligned}\tag{115}$$

or

$$\begin{pmatrix} s + \kappa & -Z\alpha \\ Z\alpha & s - \kappa \end{pmatrix} \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} = 0.\tag{116}$$

Thus, the determinant $s^2 - \kappa^2 + (Z\alpha)^2$ of the matrix must vanish, or,

$$s = \pm \sqrt{\kappa^2 - (Z\alpha)^2} = \pm \sqrt{(j + 1/2)^2 - (Z\alpha)^2}.\tag{117}$$

Because the smallest value of $j + 1/2$ is 1, we see that the exponent s will be real for all j as long as $Z\alpha < 1$.

The case $Z\alpha > 1$ does not occur in real atoms, and even if it did, the potential would not be pure Coulomb because nuclei are not point charges. Nevertheless there is some interest in solving the problem anyway. We see that $Z\alpha > 1$ leads to purely imaginary exponents s for small values of j , which, as it turns out, produce wave functions that are oscillatory and not normalizable. Thus, the usual ground state of hydrogen would disappear as a bound state for such high values of Z , and be replaced by unbound states. This strange behavior for high Z is related to what is called the “supercritical vacuum,” and is properly understood by methods of quantum field theory. In a sense what happens is that for high Z the electric field becomes strong enough to pull electron-positron pairs out of the vacuum, and thus a multiparticle situation is created.

Similar effects occur in a square well potential, if it is deep enough. You would think that the deeper the well gets, the more tightly bound the electron would be, but in the Dirac equation that is not so, since for deep enough wells we start to get unbound solutions. This effect is related to what is called the “Klein paradox,” another one of the puzzling aspects of the Dirac equation, when treated as the wave equation for a single, relativistic particle. In these notes we assume $Z\alpha < 1$, which means that s is real for all j .

Now it turns out that we can reject the minus sign in Eq. (117) on grounds of the regularity of the wave function. If we require that the amount of probability in a region near $\rho = 0$ be finite, then the integral

$$\int_0^\epsilon \rho^{2s} d\rho\tag{118}$$

must be finite for some $\epsilon > 0$. But this implies $2s > -1$, or $s > -1/2$. If we demand in addition that the expectation value of the potential energy term $Z\alpha/\rho$ be finite, then we obtain $2s - 1 > -1$, or $s > 0$. These regularity conditions are discussed in greater detail by W. Greiner, *Relativistic Quantum Mechanics*. We henceforth take only the positive root for s in Eq. (117).

If $Z\alpha$ is small then s is approximately $j + 1/2$, an integer, but with a small negative correction. With the correction s is not an integer, so the wave function behaves as a fractional power of r near the origin. In the case $j = 1/2$, s is a fraction less than 1, which means that the wave functions F and G (see Eq. (102)) go like a small negative power of r near the origin. That is, they diverge at $r = 0$, unlike the nonrelativistic solution, which is regular there. In spite of the divergence, the net probability in a neighborhood of the origin is finite, as is the expectation value of the potential, as noted above.

17. Recursion Relations and Energy Levels

We now write $a(\rho)$ and $b(\rho)$ as power series,

$$a(\rho) = \rho^s \sum_{k=0}^{\infty} a_k \rho^k, \quad b(\rho) = \rho^s \sum_{k=0}^{\infty} b_k \rho^k, \quad (119)$$

and substitute these into Eqs. (114). By equating coefficients of the different powers of ρ we obtain recursion relations for the coefficients a_n and b_n . These are conveniently written in matrix-vector form,

$$A_k x_k = B x_{k-1}, \quad (120)$$

where

$$A_k = \begin{pmatrix} k + s + \kappa & -Z\alpha \\ Z\alpha & k + s - \kappa \end{pmatrix}, \quad B = \frac{1}{2} \begin{pmatrix} 1 & 1/\gamma \\ \gamma & 1 \end{pmatrix}, \quad (121)$$

and

$$x_k = \begin{pmatrix} a_k \\ b_k \end{pmatrix}. \quad (122)$$

It is understood that $x_k = 0$ for $k < 0$. In the following we will use letters w , x , etc., at the end of the alphabet for 2-component vectors, and those near the beginning, a , b , etc., for scalars, for example, a_k and b_k are the components of x_k .

Setting $k = 0$ in Eq. (120) we obtain $A_0 x_0 = 0$, which is Eq. (116). Since $x_0 \neq 0$ this implies $\det A_0 = 0$, which as shown above determines s . As for x_0 , it is any nonzero vector in the one-dimensional kernel of A_0 . For $k > 0$ the matrices A_k are nonsingular, since

$$\det A_k = k(2s + k), \quad (123)$$

so we can write the recursion relation as

$$x_k = A_k^{-1} B x_{k-1}, \quad k > 0. \quad (124)$$

The matrix B is also singular. We pay special attention to its image, a one-dimensional space spanned by

$$w = \begin{pmatrix} 1 \\ \gamma \end{pmatrix}. \quad (125)$$

Vector w is an eigenvector of B with eigenvalue 1,

$$Bw = w. \quad (126)$$

The kernel of A_0 is conveniently handled with the adjoint (or adjugate) of A_0 . If M is any 2×2 matrix, then the adjoint \tilde{M} is defined by

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \tilde{M} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, \quad (127)$$

so that $M\tilde{M} = (\det M)I$. Thus $A_0\tilde{A}_0 = 0$ and \tilde{A}_0 maps any vector into a vector in the kernel of A_0 . It is convenient to choose

$$x_0 = \tilde{A}_0 w = \begin{pmatrix} s - \kappa & Z\alpha \\ -Z\alpha & s + \kappa \end{pmatrix} \begin{pmatrix} 1 \\ \gamma \end{pmatrix} = \begin{pmatrix} s - \kappa + \gamma Z\alpha \\ -Z\alpha + \gamma(s + \kappa) \end{pmatrix}. \quad (128)$$

In the following we will need Bx_0 , which must be a multiple of w . Using Eqs. (121) and (128), we find

$$Bx_0 = -\nu w, \quad (129)$$

where

$$\nu = \frac{Z\alpha}{2} \left(\frac{1}{\gamma} - \gamma \right) - s. \quad (130)$$

Also, for $k > 0$ we write

$$A_k^{-1} = \frac{\tilde{A}_k}{\det A_k} = \frac{1}{k(2s+k)} \begin{pmatrix} k+s-\kappa & Z\alpha \\ -Z\alpha & k+s+\kappa \end{pmatrix} = \frac{1}{k(2s+k)} (\tilde{A}_0 + kI). \quad (131)$$

Now to apply the recursion relation (124) for $k = 1$ we must evaluate

$$(\tilde{A}_0 + I)Bx_0 = -\nu(\tilde{A}_0 + I)w = -\nu(x_0 + w), \quad (132)$$

so that

$$x_1 = A_1^{-1}Bx_0 = \frac{-\nu}{1 \cdot (2s+1)}(x_0 + w). \quad (133)$$

Iterating the recursion relation in a similar manner, we obtain

$$x_k = \frac{(-1)^k \nu(\nu-1) \dots (\nu-k+1)}{k!(2s+1)(2s+2) \dots (2s+k)} (x_0 + kw). \quad (134)$$

If ν is not an integer ≥ 0 then as k gets large the coefficient in front of x_0 approaches $1/k!$, so that the series for both $a(\rho)$ and $b(\rho)$ behave like e^ρ . This overwhelms the factor of $e^{-\rho/2}$ that we have stripped off, making both $f(\rho)$ and $g(\rho)$ diverges as $e^{\rho/2}$ as $\rho \rightarrow \infty$. These solutions are not normalizable and must be rejected. We conclude that ν must be a nonnegative integer, $\nu = N$, with $N \geq 0$. Then $x_k = 0$ for $k > N$, and both $a(\rho)$ and $b(\rho)$ are polynomials in ρ of degree N . These are dominated at large ρ by $e^{-\rho/2}$, so that the wave functions are normalizable and acceptable as bound states.

Now since

$$\frac{1}{\gamma} - \gamma = \sqrt{\frac{m+E}{m-E}} - \sqrt{\frac{m-E}{m+E}} = \frac{2E}{\sqrt{m^2 - E^2}}, \quad (135)$$

Eq. (131) becomes

$$N + s = \frac{Z\alpha E}{\sqrt{m^2 - E^2}}, \quad (136)$$

which we solve for E to obtain,

$$E = \frac{mc^2}{\sqrt{1 + \frac{(Z\alpha)^2}{(N+s)^2}}}, \quad (137)$$

where we have restored ordinary units. These are the energy eigenvalues of the bound states of the Dirac hydrogen atom; they are functions of the two quantum numbers, j and N . The bound state energies are all positive and lie in the interval $0 < E < mc^2$.

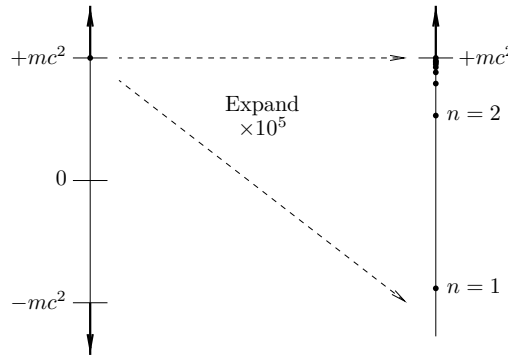


Fig. 4. The energy spectrum of the Dirac hydrogen atom, here illustrated for $Z = 1$. The usual hydrogen ground state $n = 1$ lies 13.6 eV below the rest mass-energy, mc^2 , and the usual unbound states of the Schrödinger theory lie in a continuum above $E = mc^2$. However, there also exists another continuum of unbound (negative energy) states with $E \leq -mc^2$.

To connect these quantum numbers with those of the nonrelativistic hydrogen atom, we expand in powers of $Z\alpha$. For this we only need s to lowest order,

$$s = \sqrt{\kappa^2 - (Z\alpha)^2} \approx |\kappa| = j + 1/2, \quad (138)$$

so that

$$E = mc^2 \left[1 - \frac{(Z\alpha)^2}{2(N + j + \frac{1}{2})^2} \right] = mc^2 - \frac{Ze^2}{2an^2}, \quad (139)$$

where $a = \hbar^2/mZe^2$ (see Table 17.2) is the effective Bohr radius and where we set

$$n = N + j + \frac{1}{2} \quad (140)$$

which is identified with the usual principal quantum number of the nonrelativistic problem. To second order in $Z\alpha$, the Dirac energy levels are the rest-mass-energy mc^2 of the electron, shifted downward by the binding energies of the Schrödinger theory in the electrostatic model.

Expressed in terms of n instead of N , Eq. (137) becomes the Dirac hydrogen energy levels as cited in Eq. (24.51). As discussed in Notes 24, the energy levels (137), when expanded to fourth order in $Z\alpha$, yield all the fine structure corrections obtained by perturbation theory in those notes. This means that the Dirac equation automatically incorporates the spin-orbit corrections, including the factor of $1/2$ due to Thomas precession, as well as the not-so-obvious Darwin term. All this complexity in the Schrödinger-Pauli theory is contained in the simple equation (47.13). The entire spectrum of the Dirac hydrogen atom is illustrated in Fig. 4, including the two continua of unbound states with $E \geq mc^2$ and $E \leq -mc^2$.

18. Energy Eigenfunctions

It is now straightforward to work out the energy eigenfunctions. We combine Eqs. (113), (119), (122) and (134) with $\nu = N$ to obtain

$$\begin{aligned} \begin{pmatrix} f(\rho) \\ g(\rho) \end{pmatrix} &= e^{-\rho/2} \begin{pmatrix} a(\rho) \\ b(\rho) \end{pmatrix} = \rho^s e^{-\rho/2} \sum_{k=0}^N \rho^k x_k \\ &= \frac{\Gamma(2s+1) N!}{\Gamma(N+2s+1)} \rho^s e^{-\rho/2} \left(x_0 + w \rho \frac{d}{d\rho} \right) L_N^{2s}(\rho), \end{aligned} \quad (141)$$

where we use Abramowitz and Stegun, *Handbook of Mathematical Functions*, Eq. (22.3.9), for the definition of the associated Laguerre polynomials. In the present context that definition can be written,

$$L_N^{2s}(\rho) = \sum_{k=0}^N \frac{(-1)^k \Gamma(N+2s+1)}{(N-k)! \Gamma(k+2s+1)} \frac{\rho^k}{k!}. \quad (142)$$

Also, the 2-vectors x_0 and w are given by Eqs. (125) and (128). In these formulas it is necessary to use gamma functions instead of factorials, because the upper index of the Laguerre polynomials ($2s$) is not an integer.

To explore the nonrelativistic limit it is convenient to use the identities,

$$\rho \frac{d}{d\rho} L_N^{2s}(\rho) = (N+2s) L_N^{2s-1}(\rho) - 2s L_N^{2s}(\rho) = -\rho L_{N-1}^{2s+1}(\rho), \quad (143)$$

of which the first is easily derived from the definition (142) and the second is Abramowitz and Stegun, Eq. (22.8.6) (and which is also easily derived from the definition). Then we have two expressions for the radial wave functions,

$$\begin{pmatrix} f(\rho) \\ g(\rho) \end{pmatrix} = \rho^s e^{-\rho/2} [(x_0 - 2sw) L_N^{2s}(\rho) + w(N+2s) L_N^{2s-1}(\rho)] \quad (144a)$$

$$= \rho^s e^{-\rho/2} [x_0 L_N^{2s}(\rho) - \rho w L_{N-1}^{2s+1}(\rho)], \quad (144b)$$

where we drop the constant prefactors seen on the right in Eq. (141) since at this point we are not interested in the normalization. The two forms of the solution (144) are useful for taking the

nonrelativistic limit, in which we expand everything to lowest order in $Z\alpha$. The form (144a) is useful when $\ell = j - \frac{1}{2}$, in which case the second term dominates and produces $\rho^{\ell+1}e^{-\rho/2}L_{n-\ell-1}^{2\ell+1}(\rho)$ to leading order in $Z\alpha$, which is the nonrelativistic wave function (see Eq. (17.16)); and the form (144b) is useful when $\ell = j + \frac{1}{2}$ in which case the second term again dominates and produces the same nonrelativistic wave function.

The normalization of the Dirac wave functions may also be obtained in closed form; it may be found in the references.

Problems

1. Show that $u(p, \pm s)$ and $v(p, \pm s)$ are eigenspinors of $\not{\gamma}_5$ with eigenvalue ± 1 . Thus we can define spin projectors,

$$\Sigma_{\pm} = \frac{1 \pm \not{\gamma}_5}{2}, \quad (145)$$

which project out the positive and negative spin components of an arbitrary 4-spinor of a free particle. As far as the negative energy spinors $v(p, s)$ are concerned, notice that these project out the components whose labels are $+s$ and $-s$, while the labels are the negatives of the spin eigenvalues.

2. The orthonormality and completeness relations (41), (42) and (46) will be important in future applications.

(a) Prove Eqs. (41b), (41c) and (42).

(b) Prove Eqs. (46). It may help to note that the D matrix for a pure boost is Hermitian, and that $D(\mathbf{p})^{-1} = D(-\mathbf{p})$.

3. Show that

$$\begin{aligned} \Pi_+ &= \sum_{\pm s} u(p, s) \bar{u}(p, s), \\ \Pi_- &= - \sum_{\pm s} v(p, s) \bar{v}(p, s), \end{aligned} \quad (146)$$

that is, the positive and negative energy projectors are the first and second terms in Eq. (42), respectively. As in that equation, the summation is taken over two spin vectors that point in opposite directions in the rest frame.

4. In Sec. 45.9 we carried out a nonrelativistic approximation on the Dirac equation. The essence of the approximation was Eq. (45.46), which expresses the lower two components χ of the Dirac spinor in terms of the upper two components ϕ . Notice that this equation has a relative error of order $(v/c)^2$, when the solution of the Dirac equation has an energy which is $+mc^2$ plus a small (nonrelativistic) correction of order $(v/c)^2$.

Apply this approximation to the Dirac current $J^\mu = c\bar{\psi}\gamma^\mu\psi$ to get the probability density and probability current in the Pauli theory, to order $(v/c)^2$. See if the answer includes a magnetization current. See Sec. 18.6.

5. An arbitrary free particle solution of the Dirac equation is a linear combination of the positive and negative energy solutions, which we can write as

$$\psi(\mathbf{x}, t) = \int d^3\mathbf{p} \sum_{\pm s} \left[A(p, s) u(p, s) e^{-i(p \cdot x)/\hbar} + B(p, s) v(p, s) e^{+i(p \cdot x)/\hbar} \right], \quad (147)$$

where $A(p, s)$ and $B(p, s)$ are expansion coefficients. We write these as if they depend on two 4-vectors p and s , but the notation is like the dependence of $u(p, s)$ and $v(p, s)$, that is, the 4-vector p is really a function of the 3-momentum \mathbf{p} (the variable of integration) and s is taken over only two opposite unit vectors in the rest frame. This wave function is in the Schrödinger picture, since the explicit time-dependence is given.

Compute the expectation value of the velocity operator $\mathbf{v} = c\boldsymbol{\alpha}$ to get $\langle \mathbf{v} \rangle(t)$, and show that it is independent of t and contains no Zitterbewegung if the wave packet consists purely of positive energy solutions (that is, if $B(p, s) = 0$).

6. Consider a Dirac electron ($q = -e$) in a uniform magnetic field, $\mathbf{B} = B_0 \hat{\mathbf{z}}$. Choose the gauge,

$$\mathbf{A} = B_0 x \hat{\mathbf{y}}, \quad (148)$$

which is translationally invariant in the y -direction. This means that p_y will be a constant of the motion (although you must distinguish between the kinetic and canonical momentum).

Here is some background on the nonrelativistic problem, which will save you some time. Ignore the spin and the motion in the z -direction, for simplicity. Then the Schrödinger equation is

$$\frac{1}{2m} \left[\hat{p}_x^2 + \left(\hat{p}_y + m\omega x \right)^2 \right] \psi(x, y) = E \psi(x, y). \quad (149)$$

Here we put hats on the momentum operators to distinguish them from the corresponding eigenvalues (where relevant), which are c -numbers. For example, $\hat{p}_y = -i\hbar \partial/\partial y$. Also, we define

$$\omega = \frac{eB_0}{mc}. \quad (150)$$

Then the wave equation (149) is separable, and has the solution,

$$\psi(x, y) = e^{ip_y y/\hbar} u_n(\xi), \quad (151)$$

where p_y is the eigenvalue of \hat{p}_y , where

$$\xi = x + \frac{p_y}{m\omega}, \quad (152)$$

and where u_n is the usual, normalized Hermite function for the one-dimensional harmonic oscillator with frequency ω ,

$$u_n(\xi) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \frac{1}{\sqrt{n!2^n}} H_n\left(\sqrt{\frac{m\omega}{\hbar}}\xi\right) \exp\left(-\frac{m\omega\xi^2}{2\hbar}\right). \quad (153)$$

Here H_n is the usual Hermite polynomial, defined by Eq. (8.44). The energy eigenvalue for the eigenfunction (151) is

$$E = (n + \frac{1}{2})\hbar\omega, \quad (154)$$

where $n = 0, 1, 2, \dots$ is the Landau level. The energy is independent of the quantum number p_y . The wave function is like a ridge in the x - y plane, centered on $x = -p_y/m\omega$ (i.e., on $\xi = 0$), and running in the y -direction.

(a) Solve the Dirac equation for the relativistic electron in the same magnetic field. This time you must include the z -motion and the spin. Express the energy in terms of the quantum numbers (n, p_y, p_z, m_s) . Write out explicitly a complete set of positive energy solutions as 4-component spinors. You need not normalize these solutions, and you may ignore the negative energy solutions.

(b) Consider the motion of a Dirac electron in the field,

$$\mathbf{B} = B_1 \hat{\mathbf{z}}, \quad \mathbf{E} = E_1 \hat{\mathbf{x}}, \quad (155)$$

where $0 < E_1 < B_1$. The solution of the Dirac equation for this problem can be obtained from the solution to part (a) by using the transformation law (47.26). Find matrices Λ and $D(\Lambda)$ which will cause $\psi'(x)$ to be the solution in the field (155) if $\psi(x)$ is the solution in the purely magnetic field of part (a). You will also need to find a relation between (E_1, B_1) and B_0 . You need not write out the solution $\psi'(x)$ explicitly, but do find the energy eigenvalues in terms of the same quantum numbers (n, p_y, p_z, m_s) as in part (a).