Relativistic Quantum Theories

Frederick J. Ernst

February 21, 2005

Abstract

These special lectures, which might be subtitled "Much Ado about Nothing", were delivered in 1973, while I was a teacher of physics at the Illinois Institute of Technology in Chicago.

1 The Klein-Gordon Equation

The Schrödinger wave equation which describes a particle of mass m in a potential $V(\vec{x})$,

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V(\vec{x})\psi = -\frac{\hbar}{i}\frac{\partial\psi}{\partial t}, \tag{1}$$

cannot be valid in all Minkowski frames of reference, for under a Lorentz transformation this equation does not retain its original form. Right from the earliest days of wave mechanics an attempt was made to replace the Schrödinger wave equation with a relativistically covariant equation. We shall see that such attempts are frought with difficulties of interpretation, and that these difficulties led to the development of quantum field theory, in particular, quantum electrodynamics.

Temporarily let us consider a free particle ($V = 0$). A conspicuous problem with the Schrödinger equation

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

is the appearance of a first time derivative and second space derivatives. In a manifestly Lorentz covariant theory one expects time and space coordinates to appear in roughly equivalent ways.

One fruitful idea involved the following correspondence:

$$\frac{p^2}{2m} = E \quad \Rightarrow \quad -\frac{\hbar^2}{2m}\nabla^2\psi = -\frac{\hbar}{i}\frac{\partial\psi}{\partial t}, \tag{3}$$

$$p^2c^2 + m^2c^4 = E^2 \quad \Rightarrow \quad -\hbar^2c^2\nabla^2\psi + m^2c^4\psi = -\hbar^2\frac{\partial^2\psi}{\partial t^2}. \tag{4}$$

The latter wave equation may be written

$$\nabla^2\psi - \frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} - \left(\frac{mc}{\hbar}\right)^2\psi = 0, \tag{5}$$

which is very similar to the wave equation of electromagnetic theory,

$$\nabla^2\psi - \frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} = 0. \tag{6}$$

The extra term which distinguishes the Klein-Gordon equation makes that equation no less Lorentz covariant than the electromagnetic wave equation.

Suppose we attempt to formulate a quantum theory in which the free particle would be described by the Klein-Gordon equation. How should the wave function ψ be interpreted in such a theory? Afterall, one cannot simply assume $\psi^*\psi$ gives the probability density as in Schrödinger theory, for $\int_v \psi^*\psi dv$ may not be a conserved quantity in the new theory. Let us see, however, what quantity *is* conserved.

Multiplying the Klein-Gordon equation by ψ^*, we get

$$\psi^*\nabla^2\psi - \frac{1}{c^2}\psi^*\frac{\partial^2\psi}{\partial t^2} - \left(\frac{mc}{\hbar}\right)^2\psi^*\psi = 0. \tag{7}$$

In addition, the complex conjugate equation is

$$\psi\nabla^2\psi^* - \frac{1}{c^2}\psi\frac{\partial^2\psi^*}{\partial t^2} - \left(\frac{mc}{\hbar}\right)^2\psi\psi^* = 0. \tag{8}$$

Subtracting one equation from the other, we get

$$\psi^*\nabla^2\psi - \psi\nabla^2\psi^* - \frac{1}{c^2}\left\{\psi^*\frac{\partial^2\psi}{\partial t^2} - \psi\frac{\partial^2\psi^*}{\partial t^2}\right\} = 0, \tag{9}$$

or

$$\vec{\nabla}\cdot\left[\psi^*\vec{\nabla}\psi - \psi\vec{\nabla}\psi^*\right] - \frac{1}{c^2}\frac{\partial}{\partial t}\left[\psi^*\frac{\partial\psi}{\partial t} - \psi\frac{\partial\psi^*}{\partial t}\right] = 0. \tag{10}$$

Multiplying by $\hbar/(2mi)$, we get

$$\vec{\nabla} \cdot \left\{ \frac{\hbar}{2mi} \left(\psi^* \vec{\nabla} \psi - \psi \vec{\nabla} \psi^* \right) \right\} + \frac{\partial}{\partial t} \left\{ -\frac{\hbar}{2mc^2 i} \left(\psi^* \frac{\partial \psi}{\partial t} - \psi \frac{\partial \psi^*}{\partial t} \right) \right\} = 0, \quad (11)$$

which should be compared to the corresponding equation of nonrelativistic quantum mechanics,

$$\vec{\nabla} \cdot \left\{ \frac{\hbar}{2mi} \left(\psi^* \vec{\nabla} \psi - \psi \vec{\nabla} \psi^* \right) \right\} + \frac{\partial}{\partial t} \left(\psi^* \psi \right) = 0. \quad (12)$$

The probability flux vector

$$\vec{j} = \frac{\hbar}{2mi} \left(\psi^* \vec{\nabla} \psi - \psi \vec{\nabla} \psi^* \right) \quad (13)$$

is unchanged, but the probability density must be

$$\rho = -\frac{\hbar}{2mc^2 i} \left(\psi^* \frac{\partial \psi}{\partial t} - \psi \frac{\partial \psi^*}{\partial t} \right) \quad (14)$$

in the relativistic theory instead of

$$\rho = \psi^* \psi. \quad (15)$$

In the case of an energy eigenstate, where $\psi \sim \exp(-iEt/\hbar)$, the probability density for the Klein-Gordon equation reduces to

$$\rho = \frac{E}{mc^2} \psi^* \psi, \quad (16)$$

so, if $E \approx mc^2$, it looks like the theory reduces to the nonrelativistic result. Unfortunately, this is a drastic over-simplification of a complicated question.

In the nonrelativistic theory it suffices to specify ψ at one time. The Schrödinger equation can then be used to generate ψ at a later time. This is because the Schrödinger equation is of the first degree in time derivatives. Since the Klein-Gordon equation is of the second degree in time derivatives, one must specify not only ψ but also $\partial \psi / \partial t$ at the initial time. This is, however, the source of profound difficulties, since it means that ρ may be negative as well as positive. This is rather inconsistent with the point of view that ρ represents probability density, an intrinsically positive quantity.

Restriction to Positive Energy States

When one separates the Klein-Gordon equation in Cartesian coordinates, one obtains plane wave solutions

$$\psi(\vec{x},t) = e^{i(\vec{k}\cdot\vec{x}-\omega t)}, \tag{17}$$

where

$$-|\vec{k}|^2 + \frac{1}{c^2}\omega^2 = \left(\frac{mc}{\hbar}\right)^2. \tag{18}$$

For any given $\vec{k}$, there are *two* possible values of ω; namely,

$$\omega = \pm\sqrt{\left(\frac{mc^2}{\hbar}\right)^2 + c^2|\vec{k}|^2}. \tag{19}$$

The lower sign corresponds to a negative energy particle. Note that we are referring here to total energy *including the rest energy.*

The most general solution of the Klein-Gordon equation may be represented as a linear superposition of plane waves, including negative energy plane waves as well as positive energy plane waves. Let us, however, restrict admissible states to those which may be represented as linear superpositions of *positive* energy plane waves alone.

The admissible states constitute a *linear vector space*; i.e., if ψ_1 and ψ_2 are admissible, then so are $\lambda\psi_1$ and $\psi_1+\psi_2$, where λ is any complex number.

How should the scalar product of two admissible Klein-Gordon wave functions be defined? We are guided by the requirement that the normalization of wave functions be preserved in time. We shall adopt the definition

$$(\phi,\psi)_t = i\int_t d^3x \left\{\phi^*(\vec{x})\frac{\partial\psi(\vec{x})}{\partial t} - \frac{\partial\phi^*(\vec{x})}{\partial t}\psi(\vec{x})\right\}, \tag{20}$$

which possesses all the properties usually required of a scalar product; namely,

1. $(\phi,\psi) = (\psi,\phi)^*$,
2. $(\phi,\lambda\psi) = \lambda(\phi,\psi)$,
3. $(\phi,\psi_1+\psi_2) = (\phi,\psi_1) + (\phi,\psi_2)$,
4. $(\phi,\phi) \geq 0$.

The last property, $(\phi, \phi) \geq 0$, is the only one which is not obviously satisfied by the proposed scalar product. In order to establish this property, it is convenient to express the scalar product in terms of the momentum representation wave functions.

The Momentum Representation

When we pass from Schrödinger representation to momentum representation, we should like to preserve the Lorentz invariant character of the wave function. If one were to write simply

$$\phi(\vec{x}, t) = \frac{1}{(2\pi)^{3/2}} \int d^3k e^{i(\vec{k}\cdot\vec{x} - \omega(\vec{k})t)} \tilde{\phi}(\vec{k}), \tag{21}$$

then the function $\tilde{\phi}$ so obtained would in no sense be Lorentz invariant. The relation between $\tilde{\phi}$ in one inertial frame and $\tilde{\phi}$ in some other inertial frame would be unnecessarily complicated.

To avoid this complication we first note that the volume element d^3k is not a Lorentz invariant quantity, but $d^3k/\omega(\vec{k})$ is so invariant. [Exercise: Prove this allegation.]

We shall accordingly write

$$\phi(\vec{x}, t) = \frac{1}{(2\pi)^{3/2}} \int \frac{d^3k}{\sqrt{2}\omega(\vec{k})} e^{i(\vec{k}\cdot\vec{x} - \omega(\vec{k})t)} \tilde{\phi}(\vec{k}), \tag{22}$$

where the factor $1/\sqrt{2}$ has been introduced for later convenience. When $\tilde{\phi}(\vec{k})$ is so defined, it transforms in a reasonable way under Lorentz transformations. [Exercise: In what sense might $\tilde{\phi}(\vec{k})$ be called invariant?]

The scalar product $(\phi, \psi)_t$ may now be evaluated in the momentum representation. We note that

$$\begin{aligned} i\frac{\partial \psi}{\partial t} &= \frac{1}{(2\pi)^{3/2}} \int \frac{d^3k}{\sqrt{2}\omega(\vec{k})} \omega(\vec{k}) e^{i(\vec{k}\cdot\vec{x} - \omega(\vec{k})t)} \tilde{\psi}(\vec{k}) \\ &= \frac{1}{(2\pi)^{3/2}} \int \frac{d^3k}{\sqrt{2}} e^{i(\vec{k}\cdot\vec{x} - \omega(\vec{k})t)} \tilde{\psi}(\vec{k}). \end{aligned} \tag{23}$$

Thus,

$$\begin{aligned} i\phi^* \frac{\partial \psi}{\partial t} &= \frac{1}{2(2\pi)^3} \int d^3k \int d^3k' e^{i(\vec{k} - \vec{k}')\cdot\vec{x}} \\ &\qquad e^{-i(\omega(\vec{k}) - \omega(\vec{k}'))t} \, \frac{\tilde{\phi}(\vec{k}')^* \tilde{\psi}(\vec{k})}{\omega(\vec{k}')}, \end{aligned} \tag{24}$$

and, taking the complex conjugate (with $\phi \leftrightarrow \psi$),

$$
\begin{aligned}
-i\frac{\partial\phi^*}{\partial t}\psi &= \frac{1}{2(2\pi)^3}\int d^3k \int d^3k' e^{-i(\vec{k}-\vec{k}')\cdot\vec{x}} \\
&\quad e^{i(\omega(\vec{k})-\omega(\vec{k}'))t}\,\frac{\tilde{\psi}(\vec{k}')\tilde{\phi}(\vec{k})^*}{\omega(\vec{k}')}. \qquad (25)
\end{aligned}
$$

Adding, and integrating over d^3x, we get

$$
\begin{aligned}
(\phi,\psi)_t &= \frac{1}{2}\int d^3k \int d^3k' \delta(\vec{k}-\vec{k}')e^{-i(\omega(\vec{k})-\omega(\vec{k}'))t}\frac{\tilde{\phi}(\vec{k}')^*\tilde{\psi}(\vec{k})}{\omega(\vec{k}')} \\
&\quad +\frac{1}{2}\int d^3k \int d^3k' \delta(\vec{k}-\vec{k}')e^{i(\omega(\vec{k})-\omega(\vec{k}'))t}\frac{\tilde{\psi}(\vec{k}')\tilde{\phi}(\vec{k})^*}{\omega(\vec{k}')} \\
&= \int \frac{d^3k}{\omega(\vec{k})}\tilde{\phi}(\vec{k})^*\tilde{\psi}(\vec{k}). \qquad (26)
\end{aligned}
$$

Notice the natural appearance of the Lorentz invariant $d^3k/\omega(\vec{k})$. Furthermore, it is evident that the scalar product is conserved in time providing $\phi(\vec{x},t)$ and $\psi(\vec{x},t)$ satisfy the Klein-Gordon equation. In this case one may drop the subscript t on (ϕ,ψ).

Specializing to the case $\psi = \phi$, we have

$$
(\phi,\phi) = \int \frac{d^3k}{\omega(\vec{k})}|\tilde{\phi}(\vec{k})|^2 \geq 0, \qquad (27)
$$

since our Hilbert space has been restricted to positive energy states, for which $\omega(\vec{k}) \geq 0$.

The Position Operator

The restriction to positive energy states has implications you might not anticipate. For example, you may think a position eigenfunction would simply be a delta function, as in nonrelativistic theory. But there is no way you can form a delta function by taking a linear combination of positive energy plane waves. The negative energy plane waves must be employed as well for that. This strongly suggests that $\vec{x}$ cannot be the position operator after all.

It is easier to work in terms of momentum representation, where the scalar product is simpler. Consider the operator $i\partial/\partial\vec{k}$ and the scalar product

$$(\phi, (i\partial/\partial\vec{k})\psi) = \int \frac{d^3k}{\omega(\vec{k})}\tilde{\phi}(\vec{k})^* i\frac{\partial}{\partial\vec{k}}\tilde{\psi}(\vec{k}). \tag{28}$$

Neglecting certain surface terms, integration by parts yields

$$\begin{aligned}
(\phi, (i\partial/\partial\vec{k})\psi) &= \int \frac{d^3k}{\omega(\vec{k})}\left[i\frac{\partial}{\partial\vec{k}}\tilde{\phi}(\vec{k})\right]^* \tilde{\psi}(\vec{k}) \\
&\quad - i\int d^3k\tilde{\phi}(\vec{k})^*\tilde{\psi}(\vec{k})\frac{\partial}{\partial\vec{k}}\left(\frac{1}{\omega(\vec{k})}\right) \\
&= \Big((i\partial/\partial\vec{k})\phi, \psi\Big) + i\int d^3k\tilde{\phi}(\vec{k})^*\tilde{\psi}(\vec{k})\left[\frac{c^2\vec{k}}{(\omega(\vec{k}))^3}\right].
\end{aligned} \tag{29}$$

Because of the last term $i\partial/\partial\vec{k}$ is not a self-adjoint operator. However, it is apparent that

$$\vec{X}_{\rm op} := i\frac{\partial}{\partial\vec{k}} - \frac{ic^2\vec{k}}{2(\omega(\vec{k}))^2} \tag{30}$$

is a self-adjoint operator, which could be identified as *the position operator*.

For completeness, let us write $\vec{X}_{\rm op}$ in the Schrödinger representation in order to see whether or not it corresponds to multiplication by $\vec{x}$.

$$\begin{aligned}
&\frac{1}{(2\pi)^{3/2}}\int \frac{d^3k}{\sqrt{2}\omega(\vec{k})}e^{i(\vec{k}\cdot\vec{x}-\omega(\vec{k})t)}\vec{X}_{\rm op}\tilde{\phi}(\vec{k}) = \\
&\qquad \frac{1}{(2\pi)^{3/2}}\int \frac{d^3k}{\sqrt{2}\omega(\vec{k})}e^{i(\vec{k}\cdot\vec{x}-\omega(\vec{k})t)}\left[i\frac{\partial}{\partial\vec{k}} - \frac{ic^2\vec{k}}{2(\omega(\vec{k}))^2}\right]\tilde{\phi}(\vec{k}) \\
&= \frac{1}{(2\pi)^{3/2}}\int \frac{d^3k}{\sqrt{2}}\left\{-i\frac{\partial}{\partial\vec{k}}\left[\frac{e^{i(\vec{k}\cdot\vec{x}-\omega(\vec{k})t)}}{\omega(\vec{k})}\right]\right. \\
&\qquad \left. - \frac{ic^2\vec{k}}{2(\omega(\vec{k}))^2}\frac{e^{i(\vec{k}\cdot\vec{x}-\omega(\vec{k})t)}}{\omega(\vec{k})}\right\}\tilde{\phi}(\vec{k}).
\end{aligned} \tag{31}$$

This can be expressed in turn as

$$\frac{1}{(2\pi)^{3/2}} \int \frac{d^3k}{\sqrt{2}\omega(\vec{k})} e^{i(\vec{k}\cdot\vec{x}-\omega(\vec{k})t)} \vec{X}_{\rm op}\tilde{\phi}(\vec{k}) =$$
$$\vec{x}\left[\frac{1}{(2\pi)^{3/2}} \int \frac{d^3k}{\sqrt{2}\omega(\vec{k})} e^{i(\vec{k}\cdot\vec{x}-\omega(\vec{k})t)} \tilde{\phi}(\vec{k})\right] \tag{32}$$
$$+ \frac{1}{(2\pi)^{3/2}} \int \frac{d^3k}{\sqrt{2}} \left\{ -i\frac{\partial}{\partial \vec{k}} \left[\frac{e^{-i\omega(\vec{k})t}}{\omega(\vec{k})}\right] - \frac{ic^2\vec{k}}{2(\omega(\vec{k}))^2} \frac{e^{-i\omega(\vec{k})t}}{\omega(\vec{k})} \right\} e^{i\vec{k}\cdot\vec{x}} \tilde{\phi}(\vec{k}).$$

The first term is just $\vec{x}\phi(\vec{x},t)$, but one can easily check that the second term does not vanish, so $\vec{X}_{\rm op}$ does *not* correspond simply to multiplication by $\vec{x}$, and therefore the true position eigenfunctions are *not* simply delta functions. [Exercise: For what momentum representation wave function $\psi_{\vec{x}}(\vec{k})$ is the eigenvalue equation

$$\vec{X}_{\rm op}\psi_{\vec{x}}(\vec{k}) = \vec{x}\psi_{\vec{x}}(\vec{k}) \tag{33}$$

satisfied? What does this position eigenfunction look like in the Schrödinger representation?] In this theory the position eigenfunctions are nonvanishing over a distance of the order of the Compton wave length $\hbar/(mc)$.

Klein-Gordon Particle Subject to an Interaction

If the Klein-Gordon particle is charged, it will interact electromagnetically with other charged particles. Such a possibility is readily taken into account by modifying the Klein-Gordon equation in a Lorentz covariant manner. This time we proceed from the expression

$$E = \sqrt{(mc^2)^2 + c^2|\vec{p} - (q/c)\vec{A}|^2} + q\varphi, \tag{34}$$

where $\vec{A}$ is the electromagnetic vector potential and φ is the scalar potential. You will recognize this as the classical Hamiltonian of a relativistic particle in an electromagnetic field. To avoid the square root we write this in the form

$$c^2|\vec{p} - (q/c)\vec{A}|^2 - (E - q\phi)^2 + (mc^2)^2 = 0. \tag{35}$$

The substitutions $\vec{p} = (\hbar/i)\vec{\nabla}$ and $E = -(\hbar/i)\partial/\partial t$ this time yield the modified Klein-Gordon equation

$$|-i\hbar c\vec{\nabla} - q\vec{A}|^2\psi - \left(i\hbar\frac{\partial}{\partial t} - q\varphi\right)^2 \psi + (mc^2)^2\psi 0, \tag{36}$$

or

$$\left\{\left|\vec{\nabla} - \frac{iq}{\hbar c}\vec{A}\right|^2 - \frac{1}{c^2}\left(\frac{\partial}{\partial t} + \frac{iq}{\hbar}\varphi\right)^2 - \left(\frac{mc}{\hbar}\right)^2\right\}\psi = 0. \tag{37}$$

Because $\vec{A}$ and $-\varphi$ constitute four components of a 4-vector A_μ, this modified Klein-Gordon equation is still Lorentz invariant.

Following the technique we developed earlier, the conserved probability density may be identified as

$$\rho = \frac{i\hbar}{2mc^2}\left\{\psi^*\frac{\partial\psi}{\partial t} - \left(\frac{\partial\psi}{\partial t}\right)^*\psi\right\} - \frac{q}{mc^2}\varphi\psi^*\psi. \tag{38}$$

In particular, if one considers for ψ an energy eigenfunction $\psi \sim \exp(-iEt/\hbar)$, this complicated expression reduces to

$$\rho = \frac{E - q\varphi}{mc^2}\psi^*\psi. \tag{39}$$

Here emerges still another spectre to haunt us, for suppose φ were the potential of the Coulomb interaction. At sufficiently small separation distances $q\varphi = qq'/r$ becomes large (in the case of like charges) and hence $\rho < 0$ then.

Application to Hydrogen Atom

Suppose $-e\varphi = -e^2/r$ (where $q = -e$). Then

$$\nabla^2\psi - \frac{1}{c^2}\left(\frac{\partial}{\partial t} - \frac{ie^2}{\hbar r}\right)^2\psi - \left(\frac{mc}{\hbar}\right)^2\psi = 0, \tag{40}$$

or

$$\nabla^2\psi - \frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} + \frac{2ie^2}{\hbar c^2 r}\frac{\partial\psi}{\partial t} + \frac{e^4}{\hbar^2c^2r^2}\psi - \left(\frac{mc}{\hbar}\right)^2\psi = 0. \tag{41}$$

For an energy eigenfunction, $\psi \sim \exp(-iEt/\hbar)$, so the equation reduces to

$$\nabla^2\psi + \frac{E^2}{\hbar^2c^2}\psi + \frac{2e^2E}{\hbar^2c^2r}\psi + \frac{e^4}{\hbar^2c^2r^2}\psi - \left(\frac{mc}{\hbar}\right)^2\psi = 0. \tag{42}$$

This equation may be separated in spherical polar coordinates, where

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \frac{L^2}{\hbar^2r^2} \tag{43}$$

with the angular momentum operator L defined by

$$L^2 = -\hbar^2 \left\{ \frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \left(\sin\theta \frac{\partial}{\partial\theta} \right) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\varphi^2} \right\}. \tag{44}$$

Obviously, the angular dependence of the separated solution must be a spherical harmonic $Y_\ell^m(\theta, \varphi)$. The radial function R then satisfies

$$\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) - \frac{\ell(\ell+1)}{r^2} R + \frac{E^2}{\hbar^2 c^2} R + \frac{2e^2 E}{\hbar^2 c^2 r} R + \frac{e^4}{\hbar^2 c^2 r^2} R - \left(\frac{mc}{\hbar} \right)^2 R = 0. \tag{45}$$

Terms of different character are clearly distinguished by writing this in the following form:

$$\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) - \frac{\ell(\ell+1) - \alpha^2}{r^2} R + \frac{2}{a} \frac{E}{mc^2} \frac{1}{r} R + \left(\frac{mc}{\hbar} \right)^2 \left[\frac{E^2 - (mc^2)^2}{(mc^2)^2} \right] R = 0, \tag{46}$$

where $\alpha = e^2/(\hbar c)$ is the *fine structure constant* (approximately 1/137) and $a = \hbar^2/(me^2)$ is the so-called *Bohr radius*. This clearly shows the dimensional consistency of the equation. It also facilitates a reduction to the nonrelativistic case.

Actually, the structure of this radial equation is very similar to that of the corresponding Schrödinger equation. It is not surprising, therefore, that it can be solved by using similar techniques. [Exercise: Complete the derivation of the energy levels according to the Klein-Gordon theory.] Unfortunately, the energy levels turn out all wrong for the hydrogen atom.

The Klein-Gordon equation fell into disrepute until it was revitalized as a field equation for π mesons. We shall turn our attention now to the Dirac relativistic electron theory, which dates back to 1928.

2 The Dirac Equation

In Dirac's theory an attempt is made to avoid the appearance of second time derivatives in the basic wave equation. This is accomplished by insisting that the Hamiltonian H be *linear* in the momentum. That is,

$$H\psi = -\frac{\hbar}{i} \frac{\partial}{\partial t} \psi, \tag{47}$$

where

$$H := c\vec{\alpha}\cdot\vec{p} + mc^2\beta. \tag{48}$$

The factors c and mc^2 are introduced in order to make $\vec{\alpha}$ and β dimensionless. When one inserts $\vec{p} = (\hbar/i)\vec{\nabla}$, one gets the Dirac wave equation

$$-i\hbar c\vec{\alpha}\cdot\vec{\nabla}\psi + mc^2\beta\psi = +i\hbar\frac{\partial\psi}{\partial t}. \tag{49}$$

Here the spatial derivatives and the time-derivative enter in a similar way, so there is some hope that the equation might be made manifestly covariant under Lorentz transformations.

The coefficients $\vec{\alpha}$ and β are determined by requiring

$$H^2 = (c\vec{\alpha}\cdot\vec{p} + mc^2\beta)^2 = c^2|\vec{p}|^2 + (mc^2)^2. \tag{50}$$

Comparing the two sides of this equation, we find

$$\begin{aligned} \alpha^i\alpha^j + \alpha j\alpha^i &= 2\delta^{ij}, && (51)\\ \alpha^i\beta + \beta\alpha^i &= 0 && (52)\\ \beta^2 &= 1. && (53) \end{aligned}$$

If $\vec{\alpha}$ and β are ordinary numbers (i.e., commutative), these conditions cannot be satisfied. Therefore, one is obliged to introduce objects obeying a noncommutative algebra. The simplest example of objects which satisfy the necessary rules are certain 4×4 matrices. Suppose

$$\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix} \text{ and } \beta = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \tag{54}$$

where $\vec{\sigma}$ represents the 2×2 Pauli spin matrices and I represents a 2×2 unit matrix. Then, we have

$$\beta^2 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \tag{55}$$

and

$$\begin{aligned} \vec{\alpha}\beta + \beta\vec{\alpha} &= \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix}\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} + \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}\begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -\vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix} + \begin{pmatrix} 0 & \vec{\sigma} \\ -\vec{\sigma} & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \qquad (56) \end{aligned}$$

while

$$\begin{aligned}
\alpha^i\alpha^j + \alpha^j\alpha^i &= \begin{pmatrix} 0 & \sigma^i \\ \sigma^i & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma^j \\ \sigma^j & 0 \end{pmatrix} + \begin{pmatrix} 0 & \sigma^j \\ \sigma^j & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma^i \\ \sigma^i & 0 \end{pmatrix} \\
&= \begin{pmatrix} \sigma^i\sigma^j & 0 \\ 0 & \sigma^i\sigma^j \end{pmatrix} + \begin{pmatrix} \sigma^j\sigma^i & 0 \\ 0 & \sigma^j\sigma^i \end{pmatrix} \\
&= \begin{pmatrix} \sigma^i\sigma^j + \sigma^j\sigma^i & 0 \\ 0 & \sigma^i\sigma^j + \sigma^j\sigma^i \end{pmatrix} \\
&= \begin{pmatrix} 2\delta^{ij} & 0 \\ 0 & 2\delta^{ij} \end{pmatrix}. \qquad (57)
\end{aligned}$$

Hence, these matrices provide an example of objects having the desired commutation properties. We shall accordingly think henceforth of $\vec{\alpha}$ and β as certain 4×4 matrices.

Naturally, we do not want to end up with 16 equations for ψ. We shall accordingly consider ψ itself to be a 4 component object upon which $\vec{\alpha}$ and β may act. This means that the equation

$$H\psi = -\frac{\hbar}{i}\frac{\partial}{\partial t}\psi \qquad (58)$$

will, in Dirac's theory, be a set of four coupled equations for four unknowns, the components of ψ.

Although this state of affairs may sound peculiar, consider the similar situation posed by Maxwell's equations, which we know are Lorentz covariant. In Gaussian units the inhomogeneous Maxwell equations read

$$\vec{\nabla} \cdot \vec{D} = 4\pi\rho, \quad \vec{\nabla} \times \vec{H} = \frac{4\pi}{c}\vec{j} + \frac{1}{c}\frac{\partial \vec{D}}{\partial t}, \qquad (59)$$

while the homogeneous Maxwell equations read

$$\vec{\nabla} \times \vec{E} = -\frac{1}{c}\frac{\partial \vec{B}}{\partial t}, \quad \vec{\nabla} \cdot \vec{B} = 0. \qquad (60)$$

In the case of a vacuum, $\rho = 0$, $\vec{j} = 0$, $\vec{D} = \vec{E}$ and $\vec{B} = \vec{H}$. Hence, in this case, it is possible to reexpress the Maxwell equations in the form

$$\vec{\nabla} \cdot \vec{\psi} = 0 \text{ and } \vec{\nabla} \times \vec{\psi} = +\frac{i}{c}\frac{\partial \vec{\psi}}{\partial t}, \qquad (61)$$

where

$$\vec{\psi} := \vec{E} + i\vec{H}. \tag{62}$$

While the first equation merely is a constraint upon the initial value of $\vec{\psi}$, the second set of 3 equations determines the temporal evolution of the three components of $\vec{\psi}$. In spite of the fact that we have expressed this equation in vector notation rather than matrix notation, the similarity to the Dirac equation should be fairly apparent. In the light of this the structure of the Dirac equation looks quite reasonable.

Dirac γ-Matrices

The Dirac equation

$$(c\vec{\alpha} \cdot \vec{p} + mc^2\beta)\psi = -\frac{\hbar}{i}\frac{\partial \psi}{\partial t} \tag{63}$$

can be made to look manifestly Lorentz covariant by introducing four matrices

$$\gamma^0 = -i\beta, \quad \gamma^i = -i\beta\alpha^i \quad (i = 1, 2, 3). \tag{64}$$

Multiplying the Dirac equation by β and using $\beta^2 = I$, we get

$$(ic\vec{\gamma} \cdot \vec{p} + mc^2 I)\psi = -\frac{\hbar}{i}\gamma^0\frac{\partial \psi}{\partial t}, \tag{65}$$

so, since $\vec{p} = (\hbar/i)\vec{\nabla}$, we have

$$(+\hbar c\vec{\gamma} \cdot \vec{\nabla} + mc^2 I)\psi = -\hbar\gamma^0\frac{\partial \psi}{\partial t}, \tag{66}$$

or

$$\left(\vec{\gamma} \cdot \vec{\nabla} + \gamma^0 \frac{1}{c}\frac{\partial}{\partial t}\right)\psi + \left(\frac{mc}{\hbar}\right)\psi = 0. \tag{67}$$

One can then introduce the 4-vector gradient

$$\partial_\mu = \left(\frac{1}{c}\frac{\partial}{\partial t}, \vec{\nabla}\right) \tag{68}$$

and write the Dirac equation in the elegant form

$$\gamma^\mu \partial_\mu \psi + \left(\frac{mc}{\hbar}\right)\psi = 0, \tag{69}$$

where a sum over $\mu = 0, 1, 2, 3$ is tacitly implied.

The commutation properties of the γ-matrices may be inferred from those of $\vec{\alpha}$ and β. Thus, we have

$$(\gamma^0)^2 = (-i\beta)^2 = -\beta^2 = -1, \tag{70}$$

$$\begin{aligned} \gamma^0\vec{\gamma} + \vec{\gamma}\gamma^0 &= (-i\beta)(-i\beta\vec{\alpha}) + (-i\beta\vec{\alpha})(-i\beta) \\ &= -\beta^2\vec{\alpha} - \beta\vec{\alpha}\beta \\ &= -\beta(\beta\vec{\alpha} + \vec{\alpha}\beta) = 0, \end{aligned} \tag{71}$$

and

$$\begin{aligned} \gamma^i\gamma^j + \gamma^j\gamma^i &= (-i\beta\alpha^i)(-i\beta\alpha^j) + (-i\beta\alpha^j)(-i\beta\alpha^i) \\ &= -\beta\alpha^i\beta\alpha^j - \beta\alpha^j\beta\alpha^i \\ &= \beta^2(\alpha^i\alpha^j + \alpha^j\alpha^i) = 2\delta^{ij}. \end{aligned} \tag{72}$$

These conditions are readily summarized by the single relation

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}, \tag{73}$$

where $g_{\mu\nu}$ is the metric tensor and $g^{\mu\nu}$ is its inverse.

Sometimes it is convenient to employ an explicit representation of the γ-matrices such as

$$\gamma^0 = \begin{pmatrix} -iI & 0 \\ 0 & iI \end{pmatrix}, \quad \vec{\gamma} = \begin{pmatrix} 0 & -i\vec{\sigma} \\ i\vec{\sigma} & 0 \end{pmatrix}. \tag{74}$$

At other times it is sufficient to know the commutation relations obeyed by the γ-matrices.

Conserved Probability Density

The Dirac equation may be written in the form

$$c\vec{\alpha} \cdot \vec{\nabla}\psi + \frac{\partial}{\partial t}\psi = -i\frac{mc^2}{\hbar}\beta\psi. \tag{75}$$

Multiplying by the row vector $\psi^\dagger$ on the left, we obtain

$$c\psi^\dagger\vec{\alpha} \cdot \vec{\nabla}\psi + \psi^\dagger\frac{\partial}{\partial t}\psi = -i\frac{mc^2}{\hbar}\psi^\dagger\beta\psi. \tag{76}$$

The complex conjugate of this equation is

$$c\vec{\nabla}\psi^\dagger \cdot \vec{\alpha}\psi + \frac{\partial}{\partial t}(\psi^\dagger)\psi = +i\frac{mc^2}{\hbar}\psi^\dagger\beta\psi. \tag{77}$$

Adding the two equations together, we get a continuity equation,

$$\vec{\nabla} \cdot (c\psi^\dagger\vec{\alpha}\psi) + \frac{\partial}{\partial t}(\psi^\dagger\psi) = 0, \tag{78}$$

which suggests identifying

$$\rho = \psi^\dagger\psi \tag{79}$$

as the conserved probability density and

$$\vec{j} = c\psi^\dagger\vec{\alpha}\psi \tag{80}$$

as the probability flux density. It is possible to introduce a 4-vector probability density $j^\mu = (c\rho, \vec{j})$, in terms of which the continuity equation assumes the neat form

$$\partial_\mu j^\mu = 0. \tag{81}$$

Introducing the Dirac γ-matrices, we may also write

$$j^\mu = ic\bar{\psi}\gamma^\mu\psi, \tag{82}$$

where $\bar{\psi} := \psi^\dagger\beta$.

Positive Energy Plane Wave Solutions of Dirac Equation

We shall now seek plane wave solutions of the Dirac equation,

$$\gamma^\mu\partial_\mu\psi + \left(\frac{mc}{\hbar}\right)\psi = 0. \tag{83}$$

Suppose $\psi = u(\vec{k})\exp(ik_\mu x^\mu)$, where

$$k_0 = -\sqrt{(mc/\hbar)^2 + |\vec{k}|^2}. \tag{84}$$

The column vector (spinor) $u(\vec{k})$ must satisfy the algebraic condition

$$\left(i\gamma^\mu k_\mu + \frac{mc}{\hbar}I\right)u(\vec{k}) = 0, \tag{85}$$

which may be recognized as four homogeneous equations in the four unknown components of $u(\vec{k})$. Because the equations are homogeneous, there is a nontrivial solution $u(\vec{k})$ only if

$$\det\left(i\gamma^\mu k_\mu + \frac{mc}{\hbar} I\right) = 0. \tag{86}$$

[Exercise: Verify that this determinantal condition is satisfied.] Because this determinantal condition is satisfied, $u(\vec{k})$ may be determined up to a multiplicative normalization factor by solving the algebraic equations.

The 4×4 matrix $i\gamma^\mu k_\mu + (mc/\hbar)I$ may be expressed in terms of Pauli spin matrices, since

$$\gamma^0 = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \text{ and } \vec{\gamma} = \begin{pmatrix} 0 & -i\vec{\sigma} \\ i\vec{\sigma} & 0 \end{pmatrix}. \tag{87}$$

One finds that

$$i\gamma^\mu k_\mu + \frac{mc}{\hbar} I = \begin{pmatrix} \left(\frac{mc}{\hbar} - \frac{\omega}{c}\right) I & \vec{\sigma}\cdot\vec{k} \\ -\vec{\sigma}\cdot\vec{k} & \left(\frac{mc}{\hbar} + \frac{\omega}{c}\right) I \end{pmatrix}. \tag{88}$$

If we express $u(\vec{k})$ in the form

$$u(\vec{k}) = \begin{pmatrix} A \\ B \end{pmatrix}, \tag{89}$$

then the algebraic equations governing A and B are the following:

$$\left(\frac{mc}{\hbar} - \frac{\omega}{c}\right) A + \vec{\sigma}\cdot\vec{k}B = 0 \tag{90}$$

$$-\vec{\sigma}\cdot\vec{k}A + \left(\frac{mc}{\hbar} + \frac{\omega}{c}\right) B = 0. \tag{91}$$

These equations both imply the same thing; namely,

$$B = \frac{\vec{\sigma}\cdot\vec{k}}{\frac{mc}{\hbar} + \frac{\omega}{c}} A, \tag{92}$$

where the two-component spinor A is completely arbitrary. Obviously, it can be expressed as a linear combination of

$$X_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } X_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{93}$$

Therefore, we have two basic positive energy Dirac spinors,

$$u_\pm(\vec{k}) = N \begin{pmatrix} X_\pm \\ \frac{\vec{\sigma}\cdot\vec{k}}{\frac{mc}{\hbar}+\frac{\omega}{c}} X_\pm \end{pmatrix}. \tag{94}$$

Here N is a normalization constant, which is customarily chosen in such a way that

$$\bar{u}_r(\vec{k})u_s(\vec{k}) = \delta_{rs}. \tag{95}$$

[Exercise: Show that

$$N = \sqrt{\frac{\left(\frac{mc}{\hbar}+\frac{\omega}{c}\right)}{2\left(\frac{mc}{\hbar}\right)}}$$

up to a phase factor.] In the nonrelativistic limit $|\vec{k}| << \omega/c$, so

$$u_\pm(\vec{k}) \to \begin{pmatrix} X_\pm \\ 0 \end{pmatrix}.$$

For this reason one frequently refers to the first two components of $u(\vec{k})$ as the "large components", the last two as the "small components".

The fact that there are *two* linearly independent positive energy states is associated with the fact that in the Dirac theory the electron *automatically* has spin $1/2\,\hbar$. We shall see later that this is not the only outstanding achievement of Dirac's theory. It also accounts for the gyromagnetic ratio of $2[q/(2mc)]$ for spin angular momentum, a result which remained a mystery in the nonrelativistic theory.

Electromagnetic Interactions

As in the case of the Klein-Gordon equation, we may introduce the electromagnetic interaction of the Dirac particle by replacing $\vec{\nabla}$ by $\vec{\nabla} - iq/(\hbar c)\vec{A}$ and $\partial/\partial t$ by $\partial/\partial t + i(q/\hbar)\varphi$. If we introduce the 4-vector potential

$$A^\mu = \left(\varphi, \vec{A}\right), \tag{96}$$

then the substitution simply amounts to

$$\partial_\mu \to \partial_\mu - \frac{iq}{\hbar c}A_\mu. \tag{97}$$

The modified Dirac equation is accordingly

$$\gamma^\mu \left(\partial_\mu - \frac{iq}{\hbar c} A_\mu \right) \psi + \frac{mc}{\hbar} \psi = 0. \tag{98}$$

Without immersing ourselves too deeply in the difficult problems associated with such applications as the Dirac theory of the hydrogen atom, we can demonstrate the $g = 2$ result by deriving from this Dirac equation an equation similar to the Klein-Gordon equation. To do this one simply applies the operator

$$\gamma^\mu \left(\partial_\mu - \frac{iq}{\hbar c} A_\mu \right) - \frac{mc}{\hbar} I$$

to the Dirac equation and makes use of the commutation properties of the γ-matrices; viz.,

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu}. \tag{99}$$

It is convenient to introduce a symbol $\sigma^{\mu\nu}$ such that

$$\gamma^\mu \gamma^\nu - \gamma^\nu \gamma^\mu = 2i\sigma^{\mu\nu}. \tag{100}$$

Then, one has simply

$$\gamma^\mu \gamma^\nu = g^{\mu\nu} + i\sigma^{\mu\nu}. \tag{101}$$

Proceeding with the calculation, we have

$$\begin{aligned}
&\left[\gamma^\mu \left(\partial_\mu - \frac{iq}{\hbar c} A_\mu \right) - \frac{mc}{\hbar} \right] \left[\gamma^\nu \left(\partial_\nu - \frac{iq}{\hbar c} A_\nu \right) + \frac{mc}{\hbar} \right] \psi \\
&= \gamma^\mu \gamma^\nu \left(\partial_\mu - \frac{iq}{\hbar c} A_\mu \right) \left(\partial_\nu - \frac{iq}{\hbar c} A_\nu \right) \psi - \left(\frac{mc}{\hbar} \right)^2 \psi \\
&= \left(\partial^\mu - \frac{iq}{\hbar c} A^\mu \right) \left(\partial_\mu - \frac{iq}{\hbar c} A_\mu \right) \psi - \left(\frac{mc}{\hbar} \right)^2 \psi \\
&\quad + i\sigma^{\mu\nu} \left(\partial_\mu - \frac{iq}{\hbar c} A_\mu \right) \left(\partial_\nu - \frac{iq}{\hbar c} A_\nu \right) \psi \\
&= \left(\partial^\mu - \frac{iq}{\hbar c} A^\mu \right) \left(\partial_\mu - \frac{iq}{\hbar c} A_\mu \right) \psi - \left(\frac{mc}{\hbar} \right)^2 \psi \\
&\quad + \frac{q}{\hbar c} \sigma^{\mu\nu} \left[\partial_\mu (A_\nu \psi) + A_\mu (\partial_\nu \psi) \right] \\
&= \left(\partial^\mu - \frac{iq}{\hbar c} A^\mu \right) \left(\partial_\mu - \frac{iq}{\hbar c} A_\mu \right) \psi - \left(\frac{mc}{\hbar} \right)^2 \psi \\
&\quad + \frac{q}{\hbar c} \sigma^{\mu\nu} (\partial_\mu A_\nu) \psi
\end{aligned}$$

$$
\begin{aligned}
&= \left(\partial^\mu - \frac{iq}{\hbar c}A^\mu\right)\left(\partial_\mu - \frac{iq}{\hbar c}A_\mu\right)\psi - \left(\frac{mc}{\hbar}\right)^2\psi \\
&\quad + \frac{q}{2\hbar c}\sigma^{\mu\nu}(\partial_\mu A_\nu - \partial_\nu A_\mu)\psi \\
&= \left(\partial^\mu - \frac{iq}{\hbar c}A^\mu\right)\left(\partial_\mu - \frac{iq}{\hbar c}A_\mu\right)\psi - \left(\frac{mc}{\hbar}\right)^2\psi \\
&\quad + \frac{q}{2\hbar c}\sigma^{\mu\nu}F_{\mu\nu}\psi = 0, \qquad (102)
\end{aligned}
$$

where liberal use has been made of the fact that $\sigma^{\mu\nu} = -\sigma^{\nu\mu}$.

The Klein-Gordon equation in the presence of electromagnetic interaction is simply

$$
\left(\partial^\mu - \frac{iq}{\hbar c}A^\mu\right)\left(\partial_\mu - \frac{iq}{\hbar c}A_\mu\right)\psi - \left(\frac{mc}{\hbar}\right)^2\psi = 0. \qquad (103)
$$

While the Klein-Gordon-like equation which we have derived does not tell the whole story of the Dirac equation, it does suggest the occurrence of additional electromagnetic effects, due to the extra term

$$
+\frac{q}{2\hbar c}\sigma^{\mu\nu}F_{\mu\nu}\psi.
$$

Here $F_{\mu\nu}$ is the electromagnetic field tensor, the components of which are $\vec{E}$ and $\vec{H}$, or more accurately $\vec{B}$. The magnetic components involve the spacelike indices exclusively, so let us consider $\sigma^{\mu\nu}$ when both μ and ν are spacelike indices (i and j). In this case we may write

$$
\sigma^{ij} = \begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix}, \qquad (104)
$$

where i, j, k is a cyclic permutation of $1, 2, 3$. We also have $F_{ij} = B_k$ where i, j, k is a cyclic permutation of $1, 2, 3$. Thus,

$$
\sigma^{ij}F_{ij} = 2\begin{pmatrix} \vec{\sigma}\cdot\vec{B} & 0 \\ 0 & \vec{\sigma}\cdot\vec{B} \end{pmatrix}, \qquad (105)
$$

and the extra term in the Klein-Gordon-like equation which arises due to the external magnetic field is equal to

$$
+\frac{q}{\hbar c}\begin{pmatrix} \vec{\sigma}\cdot\vec{B} & 0 \\ 0 & \vec{\sigma}\cdot\vec{B} \end{pmatrix}\psi.
$$

Since we are concerned primarily with the spin interaction with an external magnetic field, let us suppress the other electromagnetic interactions which arise in the Klein-Gordon equation too. We then have simply

$$\nabla^2\psi - \frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} - \left(\frac{mc}{\hbar}\right)^2\psi + \frac{q}{\hbar c}\begin{pmatrix} \vec{\sigma}\cdot\vec{B} & 0 \\ 0 & \vec{\sigma}\cdot\vec{B} \end{pmatrix}\psi = 0. \tag{106}$$

Notice that the "large components" of ψ satisfy the same equation as the "small components" of ψ; namely,

$$\nabla^2\psi - \frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} - \left(\frac{mc}{\hbar}\right)^2\psi + \frac{q}{\hbar c}\vec{\sigma}\cdot\vec{B}\psi = 0 \tag{107}$$

Let us restrict attention to an energy eigenstate, in which case

$$\psi \sim e^{-iEt/\hbar} \tag{108}$$

and the equation becomes

$$\nabla^2\psi + \left(\frac{E}{\hbar c}\right)^2\psi - \left(\frac{mc}{\hbar}\right)^2\psi + \frac{q}{\hbar c}\vec{\sigma}\cdot\vec{B}\psi = 0. \tag{109}$$

In the nonrelativistic limit

$$E = mc^2 + E_N R, \tag{110}$$

where $E_N R$ is small compared to mc^2. Thus, the equation reduces to

$$\nabla^2\psi + \frac{2m}{\hbar^2}E_N R\psi + \frac{q}{\hbar c}\vec{\sigma}\cdot\vec{B}\psi = 0, \tag{111}$$

or

$$-\frac{\hbar^2}{2m}\nabla^2\psi - \frac{\hbar q}{2mc}\vec{\sigma}\cdot\vec{B}\psi = E_N R\psi. \tag{112}$$

This last equation, however, is just the Schrödinger equation for a particle of spin 1/2 and magnetic moment

$$\vec{\mu} = \frac{q}{mc}\vec{S}. \tag{113}$$

The gyromagnetic ratio is *automatically* equal to q/mc, or twice the classical value. Thus, not only does the Dirac equation imply spin 1/2, but it also implies the $g = 2$ value which was so mysterious in the nonrelativistic Schrödinger theory.

Further confidence in the Dirac equation can be gained by solving the hydrogen atom problem, which it is very successful. However, this requires skills at computation which I feel are inappropriate for an undergraduate quantum mechanics course. Of the same character are considerations of the physically meaningful operators representing position, etc.

Negative Energy Solutions

In spite of many spectacular successes, the Dirac theory of the electron is haunted by the spectre of negative energy states. Plane wave solutions of this type have the form

$$v(\vec{k})e^{ik_\mu x^\mu},$$

where the first two components of $v(\vec{k})$ are small, and the second pair are large. Following the same procedure as we employed earlier for the positive energy states, we may construct a pair of linearly independent negative energy spinors,

$$v_\pm(\vec{k}) = \mp N \begin{pmatrix} \frac{\vec{\sigma}\cdot\vec{k}}{\frac{mc}{\hbar}+\frac{\omega}{c}} X_\mp \\ X_\mp \end{pmatrix}, \tag{114}$$

normalized so that

$$\bar{v}_r(\vec{k})v_s(\vec{k}) = -\delta_{rs} \tag{115}$$

by choosing

$$N = \sqrt{\frac{\left(\frac{mc}{\hbar}+\frac{\omega}{c}\right)}{2\left(\frac{mc}{\hbar}\right)}}. \tag{116}$$

If one considers a Dirac equation electron in an atom interacting with a sinusoidal electromagnetic wave, one finds that there is a distinct possibility of the electron undergoing a transition from a positive to a negative energy state, even as an electron can undergo transitions from a higher energy level to a lower one. Dirac took the bold step of imagining that in some sense all negative energy states were normally filled, so that the Pauli principle prevented the unobserved transitions to negative energy states. He did not pretend to understand why the consequent infinite negative charge density did not have physically observable effects. However, he did speculate that a sufficiently energetic photon might raise an electron out of the infinite sea to become a normal positive energy electron. An energy greater than $2mc^2$ would be necessary to accomplish this. The hole left in the infinite sea of negative energy electrons would have the appearance of a positively charged particle of the same mass as an electron. In 1932 such a positively charged particle was observed experimentally for the first time, presenting another striking success of Dirac's electron theory. However, the discovery of electron-positron pair creation also made clear the need to develop a theory involving a variable number of particles. Thus we are led to the birth of quantum field theory, especially quantum electrodynamics.

References

1. Bjorken and Drell, *Relativistic Quantum Mechanics* (McGraw-Hill)
2. Messiah, *Quantum Mechanics, Vol. 2* (Wiley)
3. Schweber, *Relativistic Quantum Field Theory* (Row, Peterson)

(Note: Every book employs its own unique conventions. Beware of lifting formulas out of unfamiliar books; rather lift the ideas therein.)