Nonrelativistic Quantum Mechanics

We begin by recalling that a prescription for obtaining the Schrödinger equation for a free particle of mass \( m \) is to substitute into the classical energy momentum relation

\[
E = \frac{p^2}{2m} \tag{3.1}
\]

the differential operators

\[
E \rightarrow i\hbar \frac{\partial}{\partial t}, \quad p \rightarrow -i\hbar \nabla. \tag{3.2}
\]

The resulting operator equation is understood to act on a (complex) wavefunction \( \psi(x, t) \). That is (with \( \hbar = 1 \)),

\[
-i\frac{\partial \psi}{\partial t} + \frac{1}{2m} \nabla^2 \psi = 0. \tag{3.3}
\]

where we interpret 

\[
\rho = |\psi|^2
\]

as the probability density (\( |\psi|^2 d^3x \) gives the probability of finding the particle in a volume element \( d^3x \)).

We are often concerned with moving particles, for example, the collision of one particle with another. We therefore need to be able to calculate the density flux of a beam of particles, \( j \). Now from the conservation of probability, the rate of decrease of the number of particles in a given volume is equal to the total flux of particles out of that volume, that is,

\[
-\frac{\partial}{\partial t} \int_V \rho \, dV = \int_S j \cdot \hat{n} \, dS = \int_V \nabla \cdot j \, dV
\]

where the last equality is Gauss's theorem and \( \hat{n} \) is a unit vector along the outward normal to the element \( dS \) of the surface \( S \) enclosing volume \( V \). The probability and the flux densities are therefore related by the "continuity" equation

\[
\frac{\partial \rho}{\partial t} + \nabla \cdot j = 0. \tag{3.4}
\]

To determine the flux, we first form \( \partial \rho / \partial t \) by subtracting the wave equation, (3.3), multiplied by \( -i \psi^* \) from the complex conjugate equation multiplied by \( -i \psi \). We then obtain

\[
\frac{\partial \rho}{\partial t} - \frac{i}{2m} (\psi^* \nabla^2 \psi - \psi \nabla^2 \psi^*) = 0. \tag{3.5}
\]
Comparing this with (3.4), we identify the probability flux density as
\[ j = -i \frac{\hbar}{2m} (\psi^* \nabla \psi - \psi \nabla \psi^*). \]  
(3.6)

For example, a solution of (3.3),
\[ \psi = N e^{i p \cdot x - i E t}, \]  
(3.7)

which describes a free particle of energy \( E \) and momentum \( p \), has
\[ \rho = |N|^2, \quad j = \frac{p}{m} |N|^2. \]  
(3.8)

Note that the space-like components of \( A^\mu \) and \( A_\mu \) are \( A \) and \( -A \), respectively. The exception is
\[ \partial^\mu = \left( \frac{\partial}{\partial t}, -\nabla \right) \quad \text{and} \quad \partial_\mu = \left( \frac{\partial}{\partial t}, \nabla \right). \]  
(3.13)

which can be shown to transform like \( x^\mu = (t, \mathbf{x}) \) and \( x_\mu = (t, -\mathbf{x}) \), respectively. Thus, the covariant form of (3.2) is
\[ \rho^\mu \rightarrow i \partial^\mu. \]  
(3.14)

From \( \partial_\mu \) and \( \partial^\mu \) we can form the invariant (D'Alembertian) operator
\[ \Box^2 = \partial_\mu \partial^\mu. \]  
(3.15)

### 3.3 The Klein–Gordon Equation

Wave equation (3.3) violates Lorentz covariance and is not suitable for a particle moving relativistically. It is tempting to repeat the steps of Section 3.2, but starting from the relativistic energy–momentum relation (3.10).

\[ E^2 = p^2 + m^2. \]

Making the operator substitutions (3.2), we obtain

\[ -\frac{\partial^2 \phi}{\partial t^2} + \nabla^2 \phi = m^2 \phi. \]  
(3.16)

which is known as the Klein–Gordon equation (but could, more correctly, have been called the relativistic Schrödinger equation). Multiplying the Klein Gordon equation by \(-i\phi^*\) and the complex conjugate equation by \( -i\phi \), and subtracting, gives the relativistic analogue of (3.5)

\[ \frac{\partial}{\partial t} \left[ i \left( \phi^* \frac{\partial \phi}{\partial t} - \phi \frac{\partial \phi^*}{\partial t} \right) \right] + \nabla \cdot \left[ -i (\phi^* \nabla \phi - \phi \nabla \phi^*) \right] = 0. \]  
(3.17)

By comparison with (3.4), we identify the probability and the flux densities with the terms in square brackets. For example, for a free particle of energy \( E \) and momentum \( p \), described by the Klein–Gordon solution
\[ \phi = N e^{i p \cdot x - i E t}, \]
we find from (3.17) that [see (3.8)],
\[ \rho = i (-2iE)|N|^2 = 2E|N|^2 \]
\[ j = -i(2ip)|N|^2 = 2p|N|^2. \]  
(3.18)
We see that the probability density is proportional to $E$, the relativistic energy of the particle. (We defer the explanation of this for a moment.)

It is advantageous to express these results in four-vector notation. Not only are they then more concise, but also the covariance becomes explicit. Using the D'Alembertian operator, (3.15), the Klein–Gordon equation becomes

$$(\Box^2 + m^2)\phi = 0.$$  \hfill (3.19)

Moreover, the probability and the flux densities form a four-vector

$$j^\mu = (\rho, j) = i(\phi^* \partial^\mu \phi - \phi \partial^\mu \phi^*)$$  \hfill (3.20)

which satisfies the (covariant) continuity relation

$$\partial_\mu j^\mu = 0.$$  \hfill (3.21)

Taking the free particle solution

$$\phi = N e^{-i p \cdot x},$$  \hfill (3.22)

we have [see (3.18)]

$$j^\mu = 2 p^\mu |N|^2.$$  \hfill (3.23)

We noted that the probability density $\rho$ is the time-like component of a four-vector; $\rho$ is proportional to $E$. This result may be anticipated since under a Lorentz boost of velocity $v$, a volume element suffers a Lorentz contraction $d^3x \to d^3x v \sqrt{1 - v^2}$; and so, to keep $\rho d^3x$ invariant, we require $\rho$ to transform as the time-like component of a four-vector $\rho \to \rho \sqrt{1 - v^2}$.

So far, so good; but what are the energy eigenvalues of the Klein–Gordon equation? Substitution of (3.22) into (3.19) gives

$$E = \pm (p^2 + m^2)^{1/2}.$$  \hfill (3.24)

Thus, in addition to the acceptable $E > 0$ solutions, we have negative energy solutions. This looks at first like a total disaster, because transitions can occur to lower and lower (more negative) energies. A second problem is that the $E < 0$ solutions are associated with a negative probability density from (3.18). To summarize, the difficulties are

$$E < 0 \text{ solutions with } \rho < 0.$$

It is clear that this problem cannot be simply ignored. We cannot simply discard the negative energy solutions as we have to work with a complete set of states, and this set inevitably includes the unwanted states.
The Dirac Equation

We constructed the Feynman rules for particles (and antiparticles) described by wavefunctions $\phi$ that satisfy the Klein–Gordon equation. These wavefunctions do not have the required two-component structure to accommodate, for instance, the spins of the electron and positron. We are looking for a relativistic equation with solutions that have two-component structure for both particle and antiparticle. For some time, it was thought that the Klein–Gordon equation was the only relativistic generalization of the Schrödinger equation until Dirac discovered an alternative one. His goal was to write an equation which, unlike the Klein–Gordon equation, was linear in $\partial / \partial t$. In order to be covariant, it must then also be linear in $\nabla$ and has therefore the general form

$$H \psi = (\alpha \cdot \mathbf{P} + \beta m) \psi.$$  \hspace{1cm} (5.1)

The four coefficients $\beta$ and $\alpha_i (i = 1, 2, 3)$ are determined by the requirement that a free particle must satisfy the relativistic energy–momentum relation (3.10),

$$H^2 \psi = (\mathbf{P}^2 + m^2) \psi.$$ \hspace{1cm} (5.2)

Equations (5.1) and (5.2) represent the Dirac equation. We will show that its solutions have sufficiently rich structure to describe spin-$\frac{1}{2}$ particles and antiparticles.

The historical impact of Dirac’s suggestion is most profound and goes far beyond providing us with a relativistic equation to describe fermions, which is our current interest. Its study led to developments ranging from quantum field theory to semiconductors and beyond. Note, however, that Dirac’s original motivation for linearizing the Klein–Gordon equation in $\partial / \partial t$ was not to explain spin but to remove negative probability densities. The appearance of $\partial / \partial t$ in the probability (3.17) is indeed at the root of this “problem.” However, for us this feature of the Klein–Gordon equation is no longer a problem. It is a bonus that allows the correct treatment of antiparticles, at least when they have no spin!
Let us forget history and study how (5.1) describes leptons (or quarks) with spin. From (5.1), we have
\[
H^2 \psi = (\alpha_i P_i + \beta m)(\alpha_j P_j + \beta m) \psi
\]
\[
= (\alpha_i^2 P_i^2 + (\alpha_i \alpha_j + \alpha_j \alpha_i)P_iP_j + (\alpha_i \beta + \beta \alpha_i)P_i m + \beta^2 m^2) \psi,
\]
where we sum over repeated indices, with the condition \(i > j\) on the second term. Comparing with (5.2), we see that

- \(\alpha_1, \alpha_2, \alpha_3, \beta\) all anticommute with each other,
- \(\alpha_1^2 = \alpha_2^2 = \alpha_3^2 = \beta^2 = 1\). \hspace{1cm} (5.3)

Since the coefficients \(\alpha\) and \(\beta\) do not commute, they cannot simply be numbers, and we are led to consider matrices operating on a wavefunction \(\psi\), which is a multicomponent column vector.

**EXERCISE 5.1** Prove that the \(\alpha\) and \(\beta\) are hermitian, traceless matrices of even dimensionality, with eigenvalues \(\pm 1\).

The lowest dimensionality matrices satisfying all these requirements are \(4 \times 4\). The choice of the four matrices \((\alpha, \beta)\) is not unique. The Dirac–Pauli representation is most frequently used:
\[
\alpha = \begin{pmatrix} 0 & \sigma_1 \\ \sigma_2 & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}
\]
where \(I\) denotes the unit \(2 \times 2\) matrix (which is frequently written as \(1\)) and where \(\sigma\) are the Pauli matrices:
\[
\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
\] \hspace{1cm} (5.5)

Another possible representation, the Weyl representation, is
\[
\alpha = \begin{pmatrix} -\sigma & 0 \\ 0 & \sigma \end{pmatrix}, \quad \beta = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}.
\] \hspace{1cm} (5.6)

Most of the results are independent of the choice of representation. Certainly, all the physics depends only on the properties listed in (5.3). In fact, not until we exhibit explicit solutions of the Dirac equation in Section 5.3 will we use a particular representation. Unless stated otherwise, we shall always choose the Dirac–Pauli representation, (5.4).

A four-component column vector \(\psi\) which satisfies the Dirac equation (5.1) is called a Dirac spinor. We might have anticipated two independent solutions (particles and antiparticles), but instead we have four!

Maybe the surprise should not have been total. We know at least one other example where a field with more components appears when linearizing the
equation. The covariant Maxwell equations $\Box^2 A_\mu = 0$ are second-order but can be written in a linear form $\partial_\mu F^{\mu\nu} = 0$ by introducing the field strength $F_{\mu\nu}$, which has more components than $A_\mu$.

Covariant Form of the Dirac Equation. Dirac $\gamma$-Matrices

On multiplying Dirac's equation, (5.1), by $\beta$ from the left, we obtain

$$i\beta \frac{\partial \psi}{\partial t} = -i\beta \alpha \cdot \nabla \psi + m\psi,$$

which may be rewritten

$$\left(i\gamma^\mu \partial_\mu - m\right)\psi = 0,$$  \hspace{1cm} (5.7)

where we have introduced four Dirac $\gamma$-matrices

$$\gamma^\mu \equiv (\beta, \beta\alpha).$$  \hspace{1cm} (5.8)

Equation (5.7) is called the covariant form of the Dirac equation. We must wait until Section 5.2 to understand the sense in which the four $4 \times 4$ matrices $\gamma^0$, $\gamma^1$, $\gamma^2$, $\gamma^3$ are to be regarded as a four-vector. The Dirac equation is really four differential equations which couple the four components of a single column vector $\psi$:

$$\sum_{k=1}^{4} \left[ \sum_{\mu} i(\gamma^\mu)_{jk} \partial_\mu - m\delta_{jk} \right] \psi_k = 0.$$

Using (5.3) and (5.8), it is straightforward to show that the Dirac $\gamma$-matrices satisfy the anticommutation relations

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu}.$$  \hspace{1cm} (5.9)

Moreover, since $\gamma^0 = \beta$, we have

$$\gamma^{0\dagger} = \gamma^0, \hspace{1cm} (\gamma^0)^2 = I$$  \hspace{1cm} (5.10)

and

$$\gamma^{k\dagger} = (\beta\alpha^k)^\dagger = \alpha^k\beta = -\gamma^k \hspace{1cm} (\gamma^k)^2 = \beta\alpha^k\beta\alpha^k = -I$$

\hspace{1cm} $k = 1, 2, 3.$  \hspace{1cm} (5.11)

Note that the hermitian conjugation results can be summarized by

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