Relativistic Quantum Mechanics

Special relativity and quantum mechanics fit together nicely. We have already handled one relativistic particle, the photon. In this section, we learn how to treat other particles when their velocities are no longer negligible compared to the speed of light.

Relativity Review Basically, we want to formulate quantum mechanics in a way that relativistic invariance is manifest. In classical physics, this is accomplished by the use of 4-vectors. To establish notation, we briefly review classical relativity in the language of 4-vectors.

The coordinate 3-vector and time are the basic components of the 4-vector x^μ, specified by

$$x^\mu = (x^0, \vec{x}), \quad \vec{x} = (x^1, x^2, x^3), \quad x^0 = ct. \tag{1}$$

This is saying that we will take the familiar 3-vector $\vec{x}$ and regard its three components as part of the 4-vector. What we often call x is now x^1, y is x^2, and z is x^3. Together with $x^0 \equiv ct$, these make up the basic space-time 4-vector.

In special relativity, we need to distinguish upper and lower indices. So along with x^μ we also have x_μ, specified by

$$x_\mu = (x^0, -\vec{x})$$

The scalar product of a 4-vector with itself involves both x^μ and x_μ. We define

$$x \cdot x = x_\mu x^\mu = (x^0)^2 - \vec{x} \cdot \vec{x}, \tag{2}$$

where the convention is that repeated upper and lower indices are summed, so $x_\mu x^\mu$ really means

$$\sum_{\mu=0}^{3} x_\mu x^\mu$$

More formally, the relation between x_μ and x^μ can be written using the metric tensor $g_{\mu\nu}$,

$$x_\mu = g_{\mu\nu} x^\nu,$$

or equivalently,

$$x^\mu = g^{\mu\nu} x_\nu,$$

where again there is a sum on the repeated index ν. By making use of Eq.(2) we find that the components of $g_{\mu\nu}$ are given by

$$g_{\mu\nu} : \begin{cases} g_{00} = 1 \\ g_{kk} = -1 \\ g_{\mu\nu} = 0, \quad \mu \neq \nu \end{cases},$$

and further that $g_{\mu\nu} = g^{\mu\nu}$. It is easy to remember the rule for raising and lower indices; for a 0 or time index, raised and lowered versions are the same, whereas for a spacial index (1,2,3), raised and lowered versions differ by a minus sign, e.g. $x_1 = -x^1$.

3-vectors and 4-vectors The discussion above was for the space-time 4-vector x^μ. How do we find other 4-vectors? In ordinary three dimensional physics, we have a number of 3-vectors. They are usually denoted by an over-arrow, e.g. $\vec{V}$. A given $\vec{V}$ has three components. Index placement is of no concern when dealing with ordinary 3-vectors, so the components of $\vec{V}$ could be written in several equivalent ways, e.g. $\vec{V} = (V_1, V_2, V_3)$ or $\vec{V} = (V^1, V^2, V^3)$ are equally valid, so here $V^1 = V_1$, etc and for writing convenience, usually use the lower index form, $\vec{V} = (V_1, V_2, V_3)$.

The story changes when we go to 4-vectors and incorporate velocity transformations as well as rotations. Here, since 4-vector scalar product involves some minus signs, it is necessary to distinguish upper and lower components, so if $\vec{V}$ is part of a 4-vector, in most cases we write $V^\mu = (V^0, \vec{V}) = (V^0, V^1, V^2, V^3)$. **Now it does matter where the indices are placed,** so $V_0 = V^0$, but $V_k = -V^k$, $k = 1, 2, 3$. The 4-vector x^μ is the prime example of this. Below is a list of other cases where the generalization from 3-vector to 4-vector works just like it does for x^μ.

Energy-momentum

$$p^\mu = (p^0, \vec{p}), \quad p^0 = \frac{E}{c} \tag{3}$$

Scalar-Vector Potential

$$A^\mu = (A^0, \vec{A}), \quad A^0 = \phi$$

Charge-Current Density

$$J^\mu = (J^0, \vec{J}), \quad J^0 = \rho c$$

Examining this list might suggest that all 3-vectors are merely the spacial components of 4-vectors. This is certainly **not true** for some important 3-vectors. The electric and magnetic fields as well as any form of angular momentum, orbital, spin, or total are **not** spacial components of 4-vectors, but in fact become parts of second rank tensors in special relativity. The electromagnetic case is discussed in the next section.

There is one last important case where the 3-vector does generalize to a 4-vector, but the index placement is different than the list above. That is the case of the gradient operator, $\vec{\nabla}$. We may write $\vec{\nabla}$ in various ways, e.g.

$$\vec{\nabla} = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})$$

is fine in ordinary 3-dimensional physics. But for generalization to special relativity, we replace x, y, z by x^1, x^2, x^3. In this form, $\vec{\nabla}$ reads

$$\vec{\nabla} = (\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3})$$

The index placement here suggests that $\vec{\nabla}$ is part of a 4-vector generalization of three dimensional gradient, but with **lower indices**. We can see that this is correct as follows. Let us write

$$\vec{\nabla} = (\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}) \equiv (\nabla_1, \nabla_2, \nabla_3)$$

Now take the scalar product of momentum and coordinate 4-vectors, $p \cdot x$, and apply say ∇_1 to it. We have

$$\nabla_1(p \cdot x) = \frac{\partial}{\partial x^1}(p^0 x^0 - \sum_k p^k x^k) = \frac{\partial}{\partial x^1}(-p^1 x^1) = -p^1 = +p_1$$

or for any three dimensional component,

$$\nabla_k(p \cdot x) = p_k$$

If we now generalize and define a 4-vector gradient, we define

$$\nabla_\mu = (\nabla_0, \nabla_1, \nabla_2, \nabla_3)$$

with the property that

$$\nabla_\mu(p \cdot x) = p_\mu.$$

To see what ∇_0 is, we use $\nabla_0(p \cdot x) = p_0$ which requires that

$$\nabla_0 = \frac{\partial}{\partial x^0}$$

To summarize, $\vec{\nabla}$ does indeed become part of a 4 dimensional gradient, which we may write as follows,

$$\nabla_\mu = (\nabla_0, \nabla_1, \nabla_2, \nabla_3) = (\nabla_0, \vec{\nabla}).$$

Implicit is that

$$\nabla_\mu = \frac{\partial}{\partial x^\mu}.$$

The only difference comparing ∇_μ to x^μ, p^μ, etc. is the index placement. **The notation ∂_μ is often used for ∇_μ. We will use ∂_μ in the rest of these notes.**

Classical Electromagnetism Maxwell's equations are the foundation of special relativity, but their usual 3-vector form does not make this obvious. In that form, Ampere's and Gauss' Laws are

$$\vec{\nabla} \times \vec{B} = \frac{1}{c}\frac{\partial \vec{E}}{\partial t} + \frac{4\pi}{c}\vec{J},$$

and

$$\vec{\nabla} \cdot \vec{E} = 4\pi\rho.$$

We can unify these by introducing a second rank tensor, defined by

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

To see how $F_{\mu\nu}$ determines the electric and magnetic fields, we first write out F_{12}. This gives

$$F_{12} = \partial_1 A_2 - \partial_2 A_1 = -(\partial_1 A^2 - \partial_2 A^1) = -(\vec{\nabla} \times \vec{B})_3 = -B_3.$$

The other components of the magnetic field follow in a similar way, so we have

$$F_{12} = -B_3, \ \ F_{23} = -B_1, \ \ F_{31} = -B_2$$

The electric field is in the space-time components. Writing out F_{01}, we have

$$F_{01} = \partial_0 A_1 - \partial_1 A_0 = -\frac{1}{c}\frac{\partial A^1}{\partial t} - \frac{\partial \phi}{\partial x^1}, \ \ = E_1,$$

so

$$F_{01} = E_1, \ \ F_{02} = E_2, \ \ F_{03} = E_3.$$

The tensor $F_{\mu\nu}$ is antisymmetric, $F_{\mu\nu} = -F_{\nu\mu}$, and so has 6 components, which comprise the two 3-vectors, $\vec{E}$, and $\vec{B}$.

With $F_{\mu\nu}$ in hand, it is easy to check that the following covariant equation contains both Ampere's and Gauss' laws,

$$\partial^\mu F_{\mu\nu} = \frac{4\pi}{c} J_\nu. \tag{4}$$

As an example, consider Eq.(4) for $\nu = 1$. We have

$$\partial^0 F_{01} + \partial^2 F_{21} + \partial^3 F_{31} = \frac{4\pi}{c} J_1.$$

Using results from above, this is the same as

$$\frac{1}{c}\frac{\partial E_1}{\partial t} - \nabla_2 B_3 + \nabla_3 B_2 = -\frac{4\pi}{c} J^1,$$

which is the 1-component of Ampere's law.

Relativistic Plane Waves Treating plane waves is a simple way to deduce relativistic wave equations. Suppose we have a one photon state, $|\vec{k}, \lambda >$. From the previous notes on photons, the matrix element of the vector potential is (using box normalization)

$$< 0|\vec{A}|\vec{k}\lambda >= \sqrt{\frac{4\pi\hbar c^2}{2\omega_k V}} \exp(i\vec{k}\cdot\vec{x} - i\omega_{\vec{k}} t)\vec{\epsilon}_{\vec{k}}.$$

The photoelectric effect and the Compton effect establish the relation between $\omega, \vec{k}$ and energy-momentum. (Prior to the experimental results, de Broglie also used relativistic reasoning to introduce his famous wavelength.) We have

$$\vec{p} = \hbar\vec{k}, \ \ E(\vec{p}) = \hbar\omega_{\vec{k}} = |\vec{p}|c.$$

Rewriting the photon plane wave space-time dependence in terms of energy-momentum gives

$$\exp(i\vec{k}\cdot\vec{x} - i\omega_{\vec{k}} t) = \exp(\frac{i}{\hbar}\vec{p}\cdot\vec{x} - \frac{i}{\hbar}E(\vec{p})t) = \exp(-i\frac{p\cdot x}{\hbar}). \tag{5}$$

The middle term in Eq.(5) is very familiar in non-relativistic quantum mechanics, where we use the formula $E(\vec{p}) = \vec{p}\cdot\vec{p}/2m$. The example of the photon implies that this general form is also valid for a relativistic particle; we need only replace the non-relativistic formula for $E(\vec{p})$ with its relativistic counterpart.

For a massive particle, we have

$$p\cdot p = (\frac{E(\vec{p})}{c})^2 - \vec{p}\cdot\vec{p} = (m_0 c)^2,$$

which gives the Einstein formula,

$$(E(\vec{p}))^2 = (m_0 c^2)^2 + \vec{p}c\cdot\vec{p}c. \tag{6}$$

So as far as its space-time dependence is concerned, a relativistic particle of definite 4-momentum will be described by

$$\exp(-\frac{i}{\hbar}p\cdot x), \text{ where } p^0 = \frac{E(\vec{p})}{c} = \frac{1}{c}\sqrt{((m_0c^2)^2 + \vec{p}\cdot\vec{p}c^2}.$$

We already know the factors that accompany the plane wave factor for photons. The corresponding factors for massive particles, including spin 1/2, will be deduced later in this section. We will also learn the meaning of taking the minus sign in the energy equation,

$$E(\vec{p}) = -\sqrt{((m_0c^2)^2 + \vec{p}\cdot\vec{p}c^2}.$$

Operators Once we have a plane wave, in non-relativistic quantum mechanics, we can bring down the 3-momentum by acting with the momentum *operator*,

$$\vec{P} = \frac{\hbar\vec{\nabla}}{i}. \tag{7}$$

In order to use this to define a 4-vector operator, we need to do a little index manipulation. Using the rule for that for the momentum 3-vector, its components have upper indices (see Eq.(3), we can write Eq.(7) as

$$P^k = \frac{\hbar\partial_k}{i}, \text{ or } P_k = i\hbar\partial_k,$$

where we lowered the index on P^k and used the fact that $\nabla_k = \partial_k$. Having gotten the indices in the same position on both sides of the equation, the generalization to a 4-momentum operator is straightforward,

$$P_\mu \equiv i\hbar\partial_\mu,$$

with components

$$P_0 = i\hbar\partial_0, \;\; P_k = i\hbar\partial_k = i\hbar\nabla_k.$$

Acting with P_μ on a plane wave gives

$$P_\mu \exp(-i\frac{p\cdot x}{\hbar}) = i\hbar\partial_\mu \exp(-i\frac{p\cdot x}{\hbar}) = \partial_\mu(p\cdot x)\exp(-i\frac{p\cdot x}{\hbar}) = p_\mu \exp(-i\frac{p\cdot x}{\hbar}),$$

which is generalizes the corresponding result in non-relativistic quantum mechanics. Applying the 4-momentum operator again, we have

$$P\cdot P \exp(-i\frac{p\cdot x}{\hbar}) = p\cdot p \exp(-i\frac{p\cdot x}{\hbar}) = (m_0 c)^2 \exp(-i\frac{p\cdot x}{\hbar}).$$

(It is important to note that for this equation to be valid, only 3 of the 4 components of the 4-vector p are independent.) Expressing the 4-vector operator P in terms of differential operators gives

$$-\hbar^2 \partial\cdot\partial \exp(-i\frac{p\cdot x}{\hbar}) = (m_0 c)^2 \exp(-i\frac{p\cdot x}{\hbar}). \tag{8}$$

At this point, we can go beyond plane waves by superposition. If we superpose plane waves satisfying Eq.(8) to build up a function $\phi(x)$, then ϕ will satisfy the same differential equation,

$$\partial\cdot\partial\phi(x) + (\frac{m_0 c}{\hbar})^2\phi(x) = 0, \tag{9}$$

where $\partial\cdot\partial = \partial_0^2 - \nabla^2$. Eq.(9) is a relativistically invariant equation known as the Klein-Gordon equation. For a spinless particle of rest mass m_0, , the Klein-Gordon equation can be used as a relativistic generalization of the free Schrödinger equation.

Current 4-vector and normalized plane waves for Klein-Gordon equation Although there are no stable charged scalar particles in nature, investigating simple probability questions and plane wave normalization for such a particle can clarify the same questions for Dirac particles, which of course have spin 1/2.

Suppose we have a single relativistic spinless particle of mass m_0, carrying a conserved quantum number like electric charge. The function $\phi(x^0, \vec{x})$ is then complex, just like the ordinary Schrödinger wave function, $\Psi(t, \vec{x})$. We know that a plane wave of our scalar particle has the space-time dependence

$$\exp(-\frac{i}{\hbar}p\cdot x) = \exp(-\frac{iE_p t}{\hbar} + \frac{i\vec{p}\cdot\vec{x}}{\hbar}),$$

where $E_p = \sqrt{(m_0c^2)^2 + (pc)^2}$. Further, it is obvious we may superpose plane waves to build up more general solutions of the Klein-Gordon equation. How do we normalize such solutions, and do the analog of single particle quantum mechanics on them? In non-relativistic quantum mechanics, we define the scalar product for a solution of the Schrödinger equation by

$$< \Psi|\Psi >= \int d^3\vec{x}(\Psi^*(t,\vec{x})\Psi(t,\vec{x})). \tag{10}$$

To understand what to write for our scalar particle it is useful to review the concept of so-called "probability current." Non-relativistic particles are conserved and as such carry a conserved quantum number, and there is a conserved current which goes along with any such conserved quantity. In non-relativistic quantum mechanics, we have

$$\rho(t,\vec{x}) \equiv \Psi^*(t,\vec{x})\Psi(t,\vec{x}), \quad \vec{J}(t,\vec{x}) \equiv \frac{\hbar}{2mi}\Psi^*(t,\vec{x}) \overleftrightarrow{\nabla} \Psi(t,\vec{x}), \tag{11}$$

where

$$\Psi^*(t,\vec{x}) \overleftrightarrow{\nabla} \Psi(t,\vec{x}) \equiv \Psi^*(t,\vec{x})\vec{\nabla}\Psi(t,\vec{x}) - (\vec{\nabla}\Psi^*(t,\vec{x}))\Psi(t,\vec{x})$$

The Schrödinger equation guarantees that

$$\frac{\partial \rho}{\partial t} + \vec{\nabla}\cdot\vec{J} = 0, \tag{12}$$

and the scalar product of Eq.(10) is the same as

$$<\Psi|\Psi>= \int d^3\vec{x}\rho(t,\vec{x}). \tag{13}$$

To find the correct definition of scalar product for our charged scalar field ϕ, we follow a similar route. Particle number is not a good concept in relativistic quantum mechanics, but for our current we may simply use a multiple of the electromagnetic current. Guided by the non-relativistic formula, we define probability current

$$J_\mu = i\hbar\phi^* \overleftrightarrow{\partial_\mu} \phi, \tag{14}$$

where

$$\phi^* \overleftrightarrow{\partial_\mu} \phi = \phi^*\partial_\mu\phi - (\partial_\mu\phi^*)\phi.$$

(The relation of J_μ as defined in Eq.(14) to the actual electromagnetic current will become clear later.) It is easy to check that if ϕ satisfies the Klein-Gordon equation, Eq.(9), that J_μ is conserved, i.e.

$$\partial^\mu J_\mu = \partial_\mu J^\mu = \partial_0 J_0 + \vec{\nabla}\cdot\vec{J} = 0. \tag{15}$$

(Note that again the lower index nature of ∇_μ is playing a role here; there is a **plus** sign before $\vec{\nabla}\cdot\vec{J}$ in Eq.(15).) The individual components of J^μ are

$$J^0 = J_0 = i\hbar\phi^* \overleftrightarrow{\partial_0} \phi, \quad \vec{J} = \frac{\hbar}{i}\phi^* \overleftrightarrow{\nabla} \phi. \tag{16}$$

What we want in the scalar product is ρ. Noting that

$$\partial_0 J^0 = \partial_t \frac{J^0}{c},$$

we have that $\rho = J^0/c$. We now can define the quantum mechanical scalar product as

$$<\phi|\phi>= \int d^3\vec{x}\rho = \frac{1}{c^2}\int d^3\vec{x}\phi^* i\hbar \overleftrightarrow{\partial_t} \phi \tag{17}$$

An important special case of Eq.(17) is using it to correctly normalize a plane wave. Going back to the non-relativistic case, we would write

$$\Psi_{\vec{p}}(t,\vec{x}) = \frac{1}{(2\pi\hbar)^{3/2}}\exp(-\frac{iE_pt}{\hbar}+\frac{i\vec{p}\cdot\vec{x}}{\hbar}), \quad (18)$$

where here of course $E_p = p^2/2m$. The standard normalization is that

$$\int d^3\vec{x}\Psi^*_{p'}\Psi_p = \delta^3(\vec{p}-\vec{p'}). \quad (19)$$

For our relativistic scalar particle, let us write

$$\phi_p = \frac{N_p}{(2\pi\hbar)^{3/2}}\exp(-\frac{iE_pt}{\hbar}+\frac{i\vec{p}\cdot\vec{x}}{\hbar}), \quad (20)$$

(here $E_p = \sqrt{(m_0c^2)^2+(pc)^2}$.) We will determine the normalization factor N_p by demanding that

$$<\phi_{p'}|\phi_p>= i\frac{\hbar}{c^2}\int d^3\vec{x}(\phi^*_{p'}\overset{\leftrightarrow}{\partial}_t\phi_p) = \delta^3(\vec{p}-\vec{p'})$$

Using the formula Eq.(20), we have

$$<\phi_{p'}|\phi_p>= \frac{1}{(2\pi\hbar)^{3/2}}(N_{p'}N_p)(\frac{E_{p'}+E_p}{c^2})\int d^3\vec{x}\exp(\frac{i}{\hbar}(\vec{p}-\vec{p'})\cdot\vec{x})) = N_p^2\frac{2E_p}{c^2}\delta^3(\vec{p}-\vec{p'}). \quad (21)$$

Demanding that the result be $\delta^3\vec{p}-\vec{p'})$, we find

$$N_p = \sqrt{\frac{c^2}{2E_p}},$$

so a correctly normalized plane wave for the Klein-Gordon equation is

$$\phi_p(x_0,\vec{x}) = \frac{1}{(2\pi\hbar)^{3/2}}\sqrt{\frac{c^2}{2E_p}}\exp(-\frac{iE_pt}{\hbar}+\frac{i\vec{p}\cdot\vec{x}}{\hbar}), \quad (22)$$

for $E_p = \sqrt{(m_0c^2)^2+(pc)^2}$.) The result in Eq.(22) is more general than the present discussion would imply. Imagine a real process in which a scalar particle of the type we are discussing is incident. The particle may be destroyed, or other particles may be produced. Nevertheless, before the interaction occurs, the particle is in a plane wave state, and Eq.(22) would be the correct factor for such a particle in the initial state. To change normalization to a delta function in wave vector $\vec{k}$ instead of momentum $\vec{p}$, we merely replace $(2\pi\hbar)^{3/2}$ by $(2\pi)^{3/2}$, and to go to box normalization replace $(2\pi)^{3/2}$ by $1/\sqrt{V}$.

Dirac Equation For spin 1/2, even in the non-relativistic case, the wave function has two components. In the relativistic case, the corresponding object Ψ will be a four component spinor. For a free particle with mass m_0, each component will be required to satisfy the Klein-Gordon equation. Dirac arrived at his equation by demanding in addition that an equation linear in partial derivatives also be satisfied. His reasoning was that the time dependent Schrödinger equation is linear in the time derivative, so it was natural to look for a relativistic generalization also linear in ∂_t, or ∂_0. What he did is often described as "taking the square root" of the Klein-Gordon equation. Before going to that case, it is easy to illustrate a similar move with the ordinary three dimensional laplacian. Consider a two component spinor satisfying the equation

$$-\nabla^2\Psi = \kappa^2\Psi, \tag{23}$$

for some real number κ. Suppose we want to take the "square root" of this equation. Form the combination

$$\vec{\nabla}\cdot\vec{\sigma},$$

where $\vec{\sigma}$ is a 3-vector made up of the three Pauli matrices. Now multiply these together, giving

$$\vec{\nabla}\cdot\vec{\sigma}\vec{\nabla}\cdot\vec{\sigma} = \nabla_k\sigma^k\nabla_j\sigma^j = \frac{1}{2}\nabla_k\nabla_j(\sigma^k\sigma^j + \sigma^j\sigma^k).$$

The Pauli matrices satisfy

$$(\sigma^k\sigma^j + \sigma^j\sigma^k) = 2\delta^{kj}I,$$

so we have

$$\vec{\nabla}\cdot\vec{\sigma}\vec{\nabla}\cdot\vec{\sigma} = \nabla^2 I,$$

where I is the 2×2 identity matrix. We could now demand that the two component spinor Ψ satisfy not just Eq.(23), but also the linear equation

$$i\vec{\nabla}\cdot\vec{\sigma}\Psi = \kappa\Psi.$$

This example has no obvious physical significance, but it does illustrate the method Dirac used to find his equation.

The Klein-Gordon equation takes us back and forth between $P\cdot P$ and $(m_0c)^2$, represented symbolically as follows,

$$(i\hbar\partial_\mu)(i\hbar\partial^\mu) \leftrightarrow (m_0c)^2.$$

To take the "square root," we introduce a set of matrices γ^μ and require that

$$i\hbar\partial_\mu\gamma^\mu \leftrightarrow m_0c,$$

or as a concrete equation,

$$i\hbar\partial_\mu\gamma^\mu\Psi(x) = i\hbar\partial\cdot\gamma\Psi = P\cdot\gamma\Psi = m_0c\Psi(x).$$

Applying $P \cdot \gamma$ again, we have

$$P \cdot \gamma P \cdot \gamma \Psi = m_0 c P \cdot \gamma \Psi(x) = (m_0 c)^2 \Psi(x). \tag{24}$$

Since at this point, we are describing a free particle with mass m_0, every component of Ψ must satisfy the Klein-Gordon equation,

$$P \cdot P \Psi = (m_0 c)^2 \Psi(x). \tag{25}$$

Comparing Eqs.(24) and (25), we must have

$$P \cdot \gamma P \cdot \gamma = P \cdot P,$$

or in terms of differential operators after canceling factors of $i\hbar$, we have

$$\partial_\mu \gamma^\mu \partial_\nu \gamma^\nu = \partial_\mu \partial_\nu \gamma^\mu \gamma^\nu = \partial \cdot \partial = \partial_\mu \partial_\nu g^{\mu\nu}.$$

Now $\partial_\mu \partial_\nu = \partial_\nu \partial_\mu$. Using this and relabeling gives

$$\partial_\mu \partial_\nu \gamma^\mu \gamma^\nu = \frac{1}{2} \partial_\mu \partial_\nu (\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu) = \partial_\mu \partial_\nu g^{\mu\nu}.$$

Comparing the two terms in this equation, we finally have

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2 g^{\mu\nu} I. \tag{26}$$

where I is the identity matrix. At this point, it is not obvious how many components Ψ has. It turns out that to find a set of four matrices satisfying Eq.(26) requires that Ψ have at least four components, or that the γ^μ be 4×4 matrices. There are many representations of the γ^μ, just as we might use other representations than the one introduced by Pauli for 2×2 matrices. The representation of the γ^μ which is most useful for seeing how the Dirac equation reduces to the Schrödinger equation in the non-relativistic limit is the following:

$$\gamma^0 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \tag{27}$$

$$\gamma^k = \begin{pmatrix} 0 & \sigma^k \\ -\sigma^k & 0 \end{pmatrix},$$

where the γ^μ are represented as blocks of 2×2 matrices, and the σ^k are Pauli matrices.

Let us now explore the solution of the Dirac equation,

$$i\hbar \partial_\mu \gamma^\mu \Psi(x) = m_0 c \Psi(x) \tag{28}$$

for a free particle. For a free particle with positive energy in a plane wave state, we may write $\Psi(x)$ as

$$\Psi(x) = u(\vec{p}) \exp(-\frac{i}{\hbar} p \cdot x).$$

It is important to keep in mind that p^0 is not a free variable (as is $\vec{p}$) but is constrained by $p^0 c = E(\vec{p})$, where $E(\vec{p})$ is given by the Einstein formula, Eq.(6). This is why u is denoted as $u(\vec{p})$.

Since the γ^μ are represented in 2×2 blocks, it is natural to represent Ψ, or equivalently, $u(\vec{p})$ in terms of 2-component spinors. The standard notation for this is

$$u(\vec{p}) = N \begin{pmatrix} \phi(\vec{p}) \\ \chi(\vec{p}) \end{pmatrix},$$

where $\phi(\vec{p})$ and $\chi(\vec{p})$ are 2-component spinors or column vectors, exactly like those used to describe spin 1/2 in non-relativistic quantum mechanics, and N is a normalization factor. (N will be determined below. Explicit spin labels will also be added further on.) Now $i\hbar\partial_\mu$ brings down p_μ, so Eq.(28) becomes

$$p \cdot \gamma u(\vec{p}) = m_0 c\, u(\vec{p}),$$

or

$$\left(\gamma^0 E(\vec{p}) - \vec{\gamma} \cdot \vec{p}\right) u(\vec{p}) = m_0 c\, u(\vec{p}).$$

Using the representation of Eq.(27) for the $\gamma-$matrices, we have

$$\begin{pmatrix} E(\vec{p}) & -\vec{\sigma} \cdot \vec{p}c \\ & \\ \vec{\sigma} \cdot \vec{p}c & -E(\vec{p}) \end{pmatrix} \begin{pmatrix} \phi \\ \\ \chi \end{pmatrix} = m_0 c^2 \begin{pmatrix} \phi \\ \\ \chi \end{pmatrix}.$$

Writing this out as two equations for the spinors ϕ and χ, we have

$$E(\vec{p})\phi - \vec{\sigma} \cdot \vec{p}c\, \chi = m_0 c^2 \phi,$$

and

$$-E(\vec{p})\chi + \vec{\sigma} \cdot \vec{p}c\, \phi = m_0 c^2 \chi.$$

These two equations contain the same information, and show that ϕ and χ are not independent. Solving for χ, we have

$$\chi = \frac{\vec{\sigma} \cdot \vec{p}c}{m_0 c^2 + E(\vec{p})} \phi. \tag{29}$$

This equation shows that in the non-relativistic limit,

$$\frac{|\vec{p}|}{m_0 c} << 1,$$

χ becomes $O(v/c)$ relative to ϕ, where $v/c = |\vec{p}|/m_0 c$. This means that in the extreme non-relativistic limit, we have only a two-component spinor as expected for a spin 1/2 particle.

The Dirac Bar Operation

Current Let us consider first a non-relativistic particle with spin 1/2, described by a 2-component wave function $\Psi(\vec{x}, t)$, satisfying the free Schrödinger equation

$$i\hbar\partial_t\Psi = -\frac{\hbar^2\nabla^2}{2m}\Psi$$

Defining

$$\rho = \Psi^\dagger\Psi, \quad \vec{J} = \frac{\hbar}{2mi}\left(\Psi^\dagger\vec{\nabla}\Psi - \vec{\nabla}\Psi^\dagger\Psi\right),$$

the Schrödinger equation guarantees that

$$\frac{\partial\rho}{\partial t} + \vec{\nabla}\cdot\vec{J} = 0.$$

The generalization to the Dirac equation involves a 4-current J^μ, satisfying

$$\partial_\mu J^\mu = 0.$$

By analogy with the non-relativistic case, we expect J^μ to be constructed out of bilinears in the 4-component Dirac spinor Ψ, and conservation of J^μ to be guaranteed by the Dirac equation. The correct expression for J^μ is

$$J^\mu \equiv \bar{\Psi}\gamma^\mu\Psi, \tag{30}$$

where $\bar{\Psi}$ is not quite the adjoint of Ψ, but is defined by

$$\bar{\Psi} = \Psi^\dagger\Gamma, \tag{31}$$

where Γ is a 4×4 matrix to be specified shortly. When we calculate $\partial_\mu J^\mu$, we clearly need the Dirac equation for $\bar{\Psi}$ as well as for Ψ. To see why we need Γ at all, let us start by taking the adjoint in the 4×4 matrix sense, of the Dirac equation. We have

$$-i\hbar\partial_\mu\Psi^\dagger(\gamma^\mu)^\dagger = \Psi^\dagger m_0 c \tag{32}$$

The γ^μ are not self-adjoint matrices. From Eq.(27) it is easy to see that

$$\gamma^0 = (\gamma^0)^\dagger,$$

$$\gamma^k = -(\gamma^k)^\dagger, \quad k = 1, 2, 3$$

We define the matrix Γ by requiring that

$$\Gamma^{-1}(\gamma^\mu)^\dagger\Gamma = \gamma^\mu. \tag{33}$$

From the representation of Eq.(27) and making use of the anti-commutation relations Eq.(26), we see that we can take

$$\Gamma = \Gamma^{-1} = \gamma^0.$$

Now from Eq.(31), we can write $\Psi^\dagger = \bar{\Psi}\Gamma^{-1}$. Using this and Eq.(33) in Eq.(32), we have as the equation satisfied by $\bar{\Psi}$,

$$-i\hbar\partial_\mu \bar{\Psi}\gamma^\mu = \bar{\Psi} m_0 c. \tag{34}$$

Returning to the Dirac current, we have finally have the current conservation equation,

$$i\hbar\partial_\mu J^\mu = i\hbar\partial_\mu \bar{\Psi}\gamma^\mu\Psi = \bar{\Psi}(m_0c - m_0c)\Psi = 0.$$

Probability and Scalar Product The current we have just defined allows definition of a probability density for a Dirac spinor. As in the non-relativistic case, the 0th component of the current can be thought of as a probability density. We have

$$J^0 = \bar{\Psi}\gamma^0\Psi,$$

but since $\Gamma = \gamma^0$, we have

$$J^0 = \Psi^\dagger\Psi.$$

This resembles the result in non-relativistic quantum mechanics for spin 1/2, but here we use the full 4-component Dirac spinor. The quantity $\Psi^\dagger\Psi$ is clearly non-negative and allows a probability interpretation. If we have two Dirac spinor states, Ψ_1 and Ψ_2, we may define a scalar product of these two states by

$$< \Psi_2|\Psi_1 > \equiv \int d\vec{x}\Psi_2^\dagger(\vec{x},t)\Psi_1(\vec{x},t).$$

Note that the scalar product or probability amplitude is defined just as in the non-relativistic case at a given *time.*

Spin Indices and Normalization of Dirac Spinors Let us first deal with spin indices. Since ϕ is a 2-component spin 1/2 spinor, it would seem natural to give it an index s, referring to "up" or "down" along some arbitrarily chosen z-axis. While this can be done, it is much more useful to choose the axis of spin quantization to be along the 3-momentum of the particle. This is a "helicity" basis, completely analogous to circular polarization states for a photon. We will use λ to denote the helicity basis index, where $\lambda = \pm 1/2$ are the allowed values. The 2-spinor ϕ_λ then satisfies by definition,

$$\hat{p}\cdot\vec{\sigma}\phi_\lambda = 2\lambda\phi_\lambda,$$

and we will choose $\phi_\lambda^\dagger\phi_\lambda = 1$. Eq.(29) determining χ now becomes very simple. We have

$$\chi_\lambda = \left(\frac{|\vec{p}|c}{m_0c^2 + E(\vec{p})}\right) 2\lambda\phi_\lambda$$

Using the helicity basis, both ϕ and χ have definite spin components along $\vec{p}$.

Finally, let us normalize the plane wave spinor, which we now denote as $u_\lambda(\vec{p})$. Various choices are possible, but the most common choice of norm is

$$\bar{u}_\lambda(\vec{p})u_{\lambda'}(\vec{p}) = \delta_{\lambda\lambda'}.$$

Writing this out, we have

$$|N|^2 \begin{pmatrix} \phi_\lambda^\dagger & \chi_\lambda^\dagger \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \begin{pmatrix} \phi_\lambda \\ \chi_\lambda \end{pmatrix} = 1.$$

Using Eq.(29) this gives

$$|N|^2 \phi_\lambda^\dagger \phi_\lambda \left(1 - \frac{\vec{p}c \cdot \vec{p}c}{(m_0c^2 + E(\vec{p}))^2}\right). = 1$$

Since ϕ_λ is normalized to unity, this leaves

$$|N|^2 \left(\frac{2m_0c^2}{m_0c^2 + E(\vec{p})}\right) = 1,$$

where we used the Einstein formula Eq.(6). Solving for N, we have

$$u_\lambda(\vec{p}) = \sqrt{\frac{m_0c^2 + E(\vec{p})}{2m_0c^2}} \begin{pmatrix} \phi_\lambda \\ \chi_\lambda \end{pmatrix}.$$

For a free particle, we will normalize states as usual to δ functions in 3-momentum. So if we have two solutions $\Psi_{\vec{p}\lambda}(\vec{x}, t)$ and $\Psi_{\vec{p}'\lambda'}(\vec{x}, t)$, we require

$$< \Psi_{\vec{p}'\lambda'} | \Psi_{\vec{p}\lambda} >= \int d\vec{x} \Psi^\dagger_{\vec{p}'\lambda'}(\vec{x}, t) \Psi_{\vec{p}\lambda}(\vec{x}, t) = \delta^3(\vec{p} - \vec{p}')\delta_{\lambda\lambda'} \quad (35)$$

The spinor $\Psi_{\vec{p}\lambda}(x)$ must contain a factor of $u_\lambda(\vec{p})$, but since we have normalized the $u_\lambda(\vec{p})$ in a way which takes no account of the overall normalization of our plane wave states, we must allow for another normalization constant, $\tilde{N}$. We may write

$$\Psi_{\vec{p}\lambda}(x) = \tilde{N} u_\lambda(\vec{p}) \exp(-\frac{i}{\hbar} p \cdot x).$$

Using this form for both $\Psi_{\vec{p}\lambda}(\vec{x}, t)$ and $\Psi_{\vec{p}'\lambda'}(\vec{x}, t)$, in Eq.(35) we arrive at

$$(2\pi\hbar)^3 \delta^3(\vec{p} - \vec{p}') |\tilde{N}|^2 u^\dagger_{\lambda'}(\vec{p}) u_\lambda(\vec{p}) = \delta^3(\vec{p} - \vec{p}')\delta_{\lambda\lambda'} \quad (36)$$

Now, using formulas established above it is easy to show that

$$u^\dagger_{\lambda'}(\vec{p}) u_\lambda(\vec{p}) = \left(\frac{E(\vec{p})}{m_0c^2}\right) \delta_{\lambda\lambda'}$$

Using this in Eq.(36) gives

$$\tilde{N} = \sqrt{\frac{m_0 c^2}{E(\vec{p})}}(2\pi\hbar)^{-3/2}.$$

We finally have

$$\Psi_{\vec{p}\lambda}(x) = (\frac{1}{2\pi\hbar})^{3/2}\sqrt{\frac{m_0 c^2}{E(\vec{p})}}u_\lambda(\vec{p})\exp(-\frac{i}{\hbar}p\cdot x). \tag{37}$$

The factor $(2\pi\hbar)^{-3/2}$ is so that our plane wave states are normalized to a δ function in 3-momentum, i.e.

$$< \Psi_{\vec{p}'\lambda'}|\Psi_{\vec{p}\lambda} >= \int d\vec{x}\Psi^\dagger_{\vec{p}'\lambda'}(\vec{x},t)\Psi_{\vec{p}\lambda}(\vec{x},t) = \delta^3(\vec{p}-\vec{p}')\delta_{\lambda'\lambda}$$

In our previous work, we have consistently normalized using the wave vector $\vec{k}$ rather than the 3-momentum, $\vec{p}$. If it is desired to normalize to a δ function in $\vec{k}$ instead of $\vec{p}$, we simply remove the factor of $\hbar^{-3/2}$ in Eq.(37). Finally for normalization in a box with periodic boundary conditions, we replace $(2\pi)^{-3/2}$ by $V^{-1/2}$, so the factor for an initial particle in a scattering process is

$$\Psi_{\vec{p}\lambda}(x) = (\frac{1}{\sqrt{V}})\sqrt{\frac{m_0 c^2}{E(\vec{p})}}u_\lambda(\vec{p})\exp(-\frac{i}{\hbar}p\cdot x), \tag{38}$$

and the corresponding factor for a final particle is

$$\Psi^\dagger_{\vec{p}\lambda}(x) = (\frac{1}{\sqrt{V}})\sqrt{\frac{m_0 c^2}{E(\vec{p})}}u^\dagger_\lambda(\vec{p})\exp(\frac{i}{\hbar}p\cdot x), \tag{39}$$

Coupling to External Electromagnetic Fields In non relativistic quantum mechanics, we couple to an external vector potential by the replacement

$$\frac{\hbar}{i}\vec{\nabla} \to \frac{\hbar}{i}\vec{\nabla} - \frac{q}{c}\vec{A},$$

where q is the particle's charge, or in components,

$$\frac{\hbar}{i}\partial_k \to \frac{\hbar}{i}\partial_k - \frac{q}{c}A^k.$$

(Recall that $\vec{\nabla}$ has components ∂_k, while $\vec{A}$ has components A^k.) Rearranging the indices prior to generalization, our replacement is

$$i\hbar\partial_k \to i\hbar\partial_k - \frac{q}{c}A_k.$$

The relativistic generalization is clearly

$$i\hbar\partial_\mu \to i\hbar\partial_\mu - \frac{q}{c}A_\mu.$$

The Dirac equation for a particle in an external 4-potential is then

$$\left(i\hbar\partial_\mu - \frac{q}{c}A_\mu\right)\gamma^\mu\Psi = m_0 c\Psi$$

We will explore the consequences of this in the next section.

Hamiltonian for Dirac Equation Let us suppose there is an external vector potential present, as well as an external scalar potential. The latter does not occur in atomic physics, but does occur in nuclear and high energy physics. **NOTE** For the most part in this section, we will work in natural units with $\hbar = c = 1$, and we will denote m_0 by m.

The Dirac equation now reads

$$\left(i\partial_\mu - qA_\mu\right)\gamma^\mu\Psi = (m + V_S)\Psi \tag{40}$$

Following Dirac's original treatment, we want to transform this into the form

$$i\partial_t\Psi = H\Psi.$$

To do so, we write out the various terms in Eq.(40), (note that with $c = 1$, $\partial_0 = \partial_t$)

$$\left(i\partial_t\gamma^0 + (i\partial_k - qA_k)\gamma^k)\right)\Psi = (m + V_S)\Psi.$$

Now multiply this equation on the left by γ^0, and move all terms except the term in ∂_t to the right hand side. This gives

$$i\partial_t\Psi = -(i\partial_k - qA_k)\gamma^0\gamma^k\Psi + qA_0\Psi + (m + V_S)\gamma^0\Psi$$

The right hand side is the Hamiltonian acting Ψ. We may re-arrange this as follows. Define

$$\alpha^k \equiv \gamma^0\gamma^k = \begin{pmatrix} 0 & \sigma^k \\ \sigma^k & 0 \end{pmatrix}.$$

Returning to ordinary 3-vector notation, we have

$$i\partial_t\Psi = \left\{ (\frac{\vec{\nabla}}{i} - q\vec{A}) \cdot \vec{\alpha} + qA_0 + (m + V_S)\gamma^0 \right\}\Psi. \tag{41}$$

The quantity in { } is the Hamiltonian for the Dirac equation. We may break H up as usual,

$$H = H_0 + V,$$

where

$$H_0 = (\frac{\vec{\nabla}}{i}) \cdot \vec{\alpha} + m\gamma^0,$$

and

$$V = -q\vec{A} \cdot \vec{\alpha} + qA_0 + V_S\gamma^0$$

Scattering may be treated using V and the $U(\infty, -\infty)$ operator. To first order in V, the matrix element for scattering for a Dirac particle with initial momentum and helicity $\vec{p}, \lambda$ to a final $\vec{p'}, \lambda'$, would be

$$< f|U(\infty, -\infty)|i >= -i \int_{-\infty}^{\infty} dt < f|e^{iH_0 t} V e^{-iH_0 t}|i >,$$

The exponentials in H_0 just give initial and final particle energies, so

$$< f|e^{iH_0 t} V e^{-iH_0 t}|i >= \int d\vec{x} \sqrt{\frac{m}{E'V}} \left(u^\dagger_{\lambda'}(\vec{p'})\, V(\vec{x}) u_\lambda(\vec{p}) \right) \sqrt{\frac{m}{EV}}\, e^{-i(p-p')\cdot x},$$

where we took the case of a static V, although that is not necessary. We again emphasize that $p^0 = E(\vec{p})$, and likewise for $(p')^0$. Aside from the additional complexity from Dirac spinors rather than 2-component spinors, this formula has an overall similarity to the case of first order scattering of a spin 1/2 particle in non relativistic quantum mechanics. Returning to the matrix element of $U(\infty, -\infty)$, the combination of space and time integrations is an integral d^4x over all space-time.

Magnetic Moment One of the great successes of the Dirac equation is that it gives a prediction for the magnetic moment of the electron. The value the Dirac equation gives is then corrected by quantum electrodynamic effects, which generate a series of corrections in powers of the fine structure constant. This section treats the prediction of the Dirac equation itself.

Returning to Eq.(41), assuming an energy eigenstate and breaking Ψ into 2-component spinors, we have

$$E\begin{pmatrix}\phi\\ \chi\end{pmatrix} = \begin{pmatrix}0 & (c\vec{P}-q\vec{A})\cdot\sigma\\ (c\vec{P}-q\vec{A})\cdot\sigma & 0\end{pmatrix}\begin{pmatrix}\phi\\ \chi\end{pmatrix} \tag{42}$$
$$+\begin{pmatrix}qA_0+V_S+mc^2 & 0\\ 0 & qA_0-V_S-mc^2\end{pmatrix}\begin{pmatrix}\phi\\ \chi\end{pmatrix}$$

In this equation, the momentum operator $\vec{P}$ is the familiar expression from non relativistic quantum mechanics,

$$\vec{P} = \frac{\hbar}{i}\vec{\nabla},$$

and factors of c and $\hbar$ have been restored. Writing out the equations for ϕ and χ from Eq.(42) gives

$$(E - mc^2 - V_S - qA_0)\phi = \left((c\vec{P} - q\vec{A})\cdot\sigma\right)\chi,$$

and

$$(E + mc^2 + V_S - qA_0)\chi = ((c\vec{P} - q\vec{A})\cdot\sigma)\phi.$$

For extracting the value of the magnetic moment a non relativistic approximation is adequate. We set $E = mc^2 + \epsilon$, where ϵ is taken to be very small compared to mc^2. Keeping only the biggest terms in the coefficient of χ in the χ equation, we have

$$\chi = \frac{1}{2mc^2}((c\vec{P} - q\vec{A})\cdot\sigma)\phi.$$

Substituting this expression in the ϕ equation gives

$$\epsilon\phi = \left(\frac{1}{2m}(\vec{P} - \frac{q}{c}\vec{A})\cdot\sigma(\vec{P} - \frac{q}{c}\vec{A})\cdot\sigma\right)\phi + (V_S + qA_0)\phi \tag{43}$$

Our interest is in the first term, the one involving $\vec{A}$, but as a side comment, we note that in this non relativistic approximation, there is no way to distinguish qA_0 from V_S, or in words, there is no way to distinguish the 0th component of a 4-vector from a Lorentz scalar. This is as it should be of course.

Turning to the term involving $\vec{A}$, we note the following identity for Pauli matrices,

$$\sigma^j\sigma^k = \delta^{jk} + i\epsilon^{jkl}\sigma^l,$$

so if we have an expression $\vec{C}\cdot\vec{\sigma}\vec{D}\cdot\vec{\sigma}$, we obtain

$$\vec{C}\cdot\vec{\sigma}\vec{D}\cdot\vec{\sigma} = \vec{C}\cdot\vec{D} + i\epsilon^{jkl}C^jD^k\sigma^l$$

Using this identity in Eq.(43), the term analogous to $\vec{C}\cdot\vec{D}$, just gives

$$\frac{1}{2m}(\vec{P} - \frac{q}{c}\vec{A})\cdot(\vec{P} - \frac{q}{c}\vec{A}),$$

which is the familiar term in the non relativistic Hamiltonian for coupling a charge to an external magnetic field. The magnetic moment coupling is contained in the term analogous to $i\epsilon^{jkl}C^j D^k \sigma^l$. Writing out the σ^3 term from Eq.(43) gives

$$\frac{i}{2m}\left((P^1-\frac{q}{c}A^1)(P^2-\frac{q}{c}A^2)-(P^2-\frac{q}{c}A^2)(P^1-\frac{q}{c}A^1)\right)\sigma^3$$

$$=-\frac{q\hbar}{2mc}(\partial_1 A^2-\partial_2 A^1)\sigma^3=-\frac{q\hbar}{2mc}B^3\sigma^3$$

There are analogous terms involving other components of $\vec{\sigma}$ and corresponding components of $\vec{B}$. For an electron, $q=-e$, so we have for the magnetic moment of the electron

$$\vec{\mu}_{elect}=-\frac{e\hbar}{2mc}\vec{\sigma}.$$

The factor multiplying $\vec{\sigma}$ is the Bohr magneton, so if we write

$$\vec{\mu}_{elect}=-(\frac{e\hbar}{2mc})g_{elect}\,\vec{S},\quad \text{where } \vec{S}=\frac{1}{2}\vec{\sigma},$$

we have $g_{elect}=2$. The conclusion is that the Dirac equation gives the g factor of the electron, which otherwise has to be fed in from experiment.