Relativistic Quantum Mechanics

Lecture 36

Physics 342
Quantum Mechanics I

Wednesday, April 28th, 2010

We know that the Schrödinger equation logically "replaces" Newton's second law (if we insist on a strict classical to quantum correspondence), in the sense that it is an equation that governs massive particle time-evolution for quantum mechanics. It suffers, as does Newton's law, from bad behavior in the (special) relativistic limit. It is possible to generate the classical (meaning non-quantum here) kinematics appropriate to special relativity, but the theory is incomplete without the promise of relativistically correct potentials. The same is true on the quantum side – we will now look at how to make a wave equation that is relativistically viable, but without a quantum theory of fields, we lack a complete description.

36.1 Candidates

What do we require, at bare minimum, of a quantum mechanical "theory"? Well, if we imagine that the target is a field $\Psi(\mathbf{r}, t)$ that contains all of the information associated with a system, given some initial starting point (after a measurement, say), then we know that the equations governing $\Psi(\mathbf{r}, t)$ must be at most first order in time. If, in addition, we want the wavefunction to have a probabilistic interpretation, with $\Psi^*(\mathbf{r}, t)\,\Psi(\mathbf{r}, t)$ a probability density, then the spatial derivatives must form a Hermitian operator.

As an example, and to see how these constraints work, take a generic "free particle" equation:

$$\frac{\partial \Psi}{\partial t} = \alpha\,\frac{\partial^2 \Psi}{\partial x^2} \tag{36.1}$$

in one dimension. We have no additional information regarding potentials or additional physical inputs, so we know the solution by separation of variables:

$$\Psi(x,t) = \left(A\, e^{\kappa\, x} + B\, e^{-\kappa\, x}\right)\, e^{\alpha\, \kappa^2\, t}. \tag{36.2}$$

Suppose we take α and κ to be real, then if $\alpha > 0$, this solution grows as time goes on. If $\alpha < 0$, the solution decays, and the wave function dies, the particle ceases to be anywhere. Neither of these configurations is particularly useful, as far as probabilities go, so we make $\alpha \in \mathbb{C}$, and purely imaginary. This is to say that whatever the wave equation form, it should have a Hermitian operator on its right-hand side, and a single derivative on the left:

$$\frac{\partial \Psi}{\partial t} = H\, \Psi \tag{36.3}$$

with H Hermitian (in the above, this means α purely imaginary). What we will end up additionally requiring is that (for reasons of relativity), H have no higher than first derivatives in it (this makes sense, if we are treating space and time as equivalent objects, and time only has one derivative, there should be only single spatial derivatives).

36.2 Free Particle Relativistic Kinematics

The free particle action, as we have seen before, is

$$S = -m\, c^2 \int \sqrt{1 - \frac{v^2}{c^2}}\, dt, \tag{36.4}$$

with Lagrangian given by the integrand, and Hamiltonian (energy) then given by

$$H = \frac{m\, c^2}{\sqrt{1 - \frac{v^2}{c^2}}} \qquad \mathbf{p} = \frac{m\, \mathbf{v}}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{36.5}$$

and this can be conveniently written in terms of p alone

$$H = \sqrt{p^2\, c^2 + m^2\, c^4}. \tag{36.6}$$

Suppose we just made the replacement, $\mathbf{p} \rightarrow -i\, \hbar\, \nabla$ – then we have an operator:

$$H = c\, \sqrt{-\hbar\, \nabla^2 + m^2\, c^2}, \tag{36.7}$$

which is sub-optimal, unless you have particularly good ideas about taking the square root of a differential operator. It can be done, but the extension to electricity and magnetism introduces additional issues.

Suppose, to avoid these issues, we agree to deal with the square of the above (setting $\hbar = c = 1$ for ease of use, now)

$$H^2 = -\nabla^2 + m^2, \tag{36.8}$$

and, based on the Schrödinger inspired identification, $H\,\Psi = i\,\hbar\,\frac{\partial \Psi}{\partial t} = E\,\Psi$, we start with a free-particle equation, for relativistic quantum mechanics, that looks like:

$$-\frac{\partial^2 \Psi(\mathbf{r},t)}{\partial t^2} + \nabla^2\,\Psi(\mathbf{r},t) - m^2\,\Psi(\mathbf{r},t) = 0. \tag{36.9}$$

This treats time and space in an equivalent manner, but does not satisfy our constraint that the wavefunction alone carries all information about the system – we also need its time-derivative (the above Klein-Gordon equation is second order in time).

36.2.1 Aside: Manifest Relativistic Covariance of K-G

We can see from its form that the Klein-Gordon equation (36.9) is a Lorentz scalar. Introduce the Minkowski metric in Cartesian coordinates:

$$g_{\mu\nu} \doteq \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{36.10}$$

and the covariant form of the gradient operator in $D = 3 + 1$:

$$\partial_\mu \doteq \begin{pmatrix} \partial_0 \\ \partial_x \\ \partial_y \\ \partial_z \end{pmatrix}. \tag{36.11}$$

Then the scalar Laplacian is $\partial_\mu\,g^{\mu\nu}\,\partial_\nu = -\frac{\partial^2}{\partial t^2} + \nabla^2 \equiv \Box^2$, so the Klein-Gordon equation can be written explicitly as a scalar:

$$\Box^2\,\Psi(\mathbf{r},t) - m^2\,\Psi(\mathbf{r},t) = 0 \tag{36.12}$$

and is therefore valid in all frames related by Lorentz transformations.

36.2.2 Splitting the Klein-Gordon Equation

We can make the above into a pair of first-order equations in time by introducing an auxiliary field $\frac{\partial \Psi}{\partial t}$, and then supposing that the wavefunction has multiple components. We do this symmetrically by making two complex fields:

$$\Lambda = \Psi + \frac{i}{m}\frac{\partial \Psi}{\partial t} \qquad \Phi = \Psi - \frac{i}{m}\frac{\partial \Psi}{\partial t} \tag{36.13}$$

or

$$\Psi = \frac{1}{2}\left(\Lambda + \Phi\right) \qquad \frac{\partial \Psi}{\partial t} = -\frac{1}{2}\, i\, m\left(\Lambda - \Phi\right) \tag{36.14}$$

so that the Klein-Gordon equation reads:

$$\begin{aligned} 0 &= \frac{1}{2}\, i\, m\, \dot{\Lambda} + \frac{1}{2}\nabla^2 \Lambda - \frac{1}{2}\, m^2\, \Lambda \\ 0 &= -\frac{1}{2}\, i\, m\, \dot{\Phi} + \frac{1}{2}\nabla^2 \Phi - \frac{1}{2}\, m^2\, \Phi, \end{aligned} \tag{36.15}$$

it may not look like much, but the lesson is clear: When we impose relativistic restrictions, in order to recover a theory with a clearly defined wavefunction, we may require (i.e. require) multiple components in the wavefunction itself.

36.3 The Free Dirac Equation

We now go the other way around – we will construct the Dirac equation by starting from the assumption that it is first-order in time, and has a Hermitian operator setting its spatial dependence. We will allow for the possibility that the wavefunction contains multiple components. In fact, we know it must since we are dealing, in the end, with an electron, and this requires spin information (an attached "internal" vector representing up and down).

So start with Ψ having d components, and satisfying:

$$i\,\frac{\partial \Psi}{\partial t} = H\,\Psi. \tag{36.16}$$

Since H must be first order in the spatial derivatives, we have the following generic option:

$$H = \mathbf{a}\cdot\mathbf{p} + \alpha\, m. \tag{36.17}$$

This expression has at most first derivatives (through $\mathbf{p} \sim \nabla$), and we allow the vector $\mathbf{a}$ and the scalar α to be Hermitian operators that act on the spin of the electron (our Pauli spin matrices, for example).

In order to constrain $\mathbf{p}$ and α, we require that the above "reduce" to the Klein-Gordon equation – that is, it must represent (spin aside) the correct relativistic energy-momentum relationship. Writing the Hamiltonian equation suggestively as (using the Einstein summation convention for the dot products):

$$(E - a_i\, p_i - \alpha\, m)\, \Psi = 0 \tag{36.18}$$

if we multiply on the left by $(E + a_j\, p_j + \alpha\, m)$[1] , then (omitting Ψ)

$$E^2 - a_j\, p_j\, a_i\, p_i - m\, (a_j\, p_j\, \alpha + \alpha\, a_i\, p_i) - \alpha^2\, m^2 = 0. \tag{36.20}$$

We have been careful about the ordering since these are meant to be operators (in the end). The set $(\alpha, \mathbf{a})$ and $\mathbf{p}$ do not talk to each other since the former act on the spin space, so we are free to re-order these as we like:

$$E^2 - a_j\, a_i\, p_j\, p_i - m\, (a_j\, \alpha + \alpha\, a_j)\; p_j - \alpha^2\, m^2 = 0. \tag{36.21}$$

Finally, to compare with the Hamiltonian $E^2 = p^2 + m^2$, we need to factor out the overall p^2 contribution. This can be done via

$$E^2 - (a_j\, a_j)(p_k\, p_k) - \alpha^2\, m^2 - \sum_{k<\ell} (a_k\, a_\ell + a_\ell\, a_k)\; p_k\, p_\ell - m\, (a_j\, \alpha + \alpha\, a_j)\; p_j = 0. \tag{36.22}$$

The first three terms represent the correct relativistic expression, provided $a_j\, a_j = \mathbb{I}$ and $\alpha^2 = \mathbb{I}$. In order to get rid of the remaining terms, we must have

$$\begin{aligned} a_k\, a_\ell + a_\ell\, a_k &= 0 \qquad k \neq \ell \\ a_j\, \alpha + \alpha\, a_j &= 0. \end{aligned} \tag{36.23}$$

If we can do that, then we will have a relativistic theory.

[1] Think of the factorization:

$$\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} = \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial x}\right)\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial x}\right) \tag{36.19}$$

which we would write as $-(E + p)\,(E - p)$.

36.4 Adding E&M

Let's go back to the relativistic form of the Lagrangian and attempt to couple E&M – in our current units:

$$L = -m\sqrt{1-v^2} \tag{36.24}$$

we know this is a Lorentz scalar, by construction, so the question is – what scalar quantity can we add that will encorporate E&M?

There are two natural four-vectors that we can combine – one is v^μ, the four-velocity, and in addition, we have the four-potential from E&M:

$$A^\mu \doteq \begin{pmatrix} \phi \\ A_x \\ A_y \\ A_z \end{pmatrix}. \tag{36.25}$$

Then the only thing we can make is $q\, v^\mu\, g_{\mu\nu}\, A^\nu$. We haven't specified a parametrization for $v^\mu = \dot{x}^\mu$. Thinking back to the relativistic action, what we are proposing is the addition of a term:

$$q\int \dot{x}^\mu(\lambda)\, g_{\mu\nu}\, A^\nu\, d\lambda, \tag{36.26}$$

but this term is itself re-parametrization invariant – take $\frac{dx^\mu}{d\lambda} = \frac{dx^\mu}{dt}\frac{dt}{d\lambda}$, then (36.26) will read

$$q\int \frac{dx^\mu}{dt}\, g_{\mu\nu}\, A^\nu\, dt \tag{36.27}$$

and can be combined with the free-particle action (both are parametrized using the coordinate t). The integrand that we should add to the Lagrangian is now simple:

$$q\,\frac{dx^\mu}{dt}\, g_{\mu\nu}\, A^\nu = -q\,\phi + q\,\mathbf{v}\cdot\mathbf{A} \tag{36.28}$$

This is just the usual term corresponding to the Lorentz force, no surprise, since we know that the Lorentz force is already relativistically covariant. Our final, electromagnetically coupled Lagrangian reads:

$$L = -m\sqrt{1-v^2} - q\,\phi + q\,\mathbf{v}\cdot\mathbf{A}. \tag{36.29}$$

Now when we form the Hamiltonian, we have to take:

$$H = \frac{\partial L}{\partial \mathbf{v}}\cdot\mathbf{v} - L \tag{36.30}$$

with

$$\frac{\partial L}{\partial \mathbf{v}} = \mathbf{p} - q\mathbf{A}. \tag{36.31}$$

From classical mechanics, we know what will happen (even in this relativistic setting) – we take $\mathbf{p} \longrightarrow \mathbf{p} - q\,\mathbf{A}$ and add $q\,\phi$ to the free particle Hamiltonian. In our quantum mechanical Hamiltonian, we will have

$$E = \mathbf{a} \cdot (\mathbf{p} - q\,\mathbf{A}) + \alpha\, m + q\,\phi. \tag{36.32}$$

Finally, the Dirac equation, with these replacements, reads (now taking $q \to -q$ to account for the negative charge of the electron):

$$[(E + q\,\phi) - \mathbf{a} \cdot (\mathbf{p} + q\,\mathbf{A}) - \alpha\, m]\,\Psi = 0, \tag{36.33}$$

or, inputting the usual $\mathbf{p} = -i\,\nabla$, $E = i\,\frac{\partial}{\partial t}$, and tabulating the side-constraints:

$$\boxed{\begin{aligned} 0 &= \left[\left(i\,\frac{\partial}{\partial t} + q\,\phi\right) - \mathbf{a} \cdot (-i\,\nabla + q\,\mathbf{A}) - \alpha\, m\right]\,\Psi(\mathbf{r}, t) \\ \mathbb{I} &= a_j \cdot a_j = \alpha\,\alpha \\ 0 &= a_k\, a_\ell + a_\ell\, a_k \qquad k \neq \ell \\ 0 &= a_j\,\alpha + \alpha\, a_j \end{aligned}} \tag{36.34}$$