Physics 613 Lecture 3 Jan. 30, 2014

Having discussed the various particles of the standard model and the symmetries which classify them, it is time to turn to the question of how they behave, that is, what equations govern their motion.

We are dealing with quantum mechanics here, so of course the states in which the system can exist live in a large Hilbert space on which physical quantities act as linear operators. But unlike non-relativistic quantum mechanics, these states are not just configurations of a fixed number of particles, because in relativity particles can be created and destroyed. But even though the states can represent amorphous configurations, they can be decomposed into irreducible representations of any symmetry of the physics.

Because we assume physics is invariant under spatial and time translations, the total 4-momentum of the state is conserved, and because the generators $\hat{P}^\mu$ of these translations commute and can be simultaneously diagonalized as operators in the space of all states of the system, we can decompose the states into states with definite values p^μ of $\hat{P}^\mu$. That is, these states are eigenvectors of the operators $\hat{P}^\mu$. But when we consider the bigger Poincaré symmetry group, including rotations and Lorentz boosts, the momentum operator does not commute with rotations or boosts, so the individual momentum states must be combined into bigger representations. But $\sum_\mu \hat{P}_\mu \hat{P}^\mu$ does commute with all these symmetries, so an irreducible representation will have a definite (single) value for $p^2 = m^2$.

[Note: We are using relativistic (4-D) notation, which requires vectors to be either contravariant ($x^\mu = (x^0, x^1, \cdots, x^3) = (t, x, y, z)$) or covariant ($p_\mu = (p_0, \cdots p_3) = (E, -\vec{p})$. We recall that the (Minkowski) square of a vector or the dot product of two vectors is defined as the Lorentz scalar

$$\mathbf{W} \cdot \mathbf{V} = W^0 V^0 - \vec{W} \cdot \vec{V} = \sum_{\mu\nu} g_{\mu\nu} W^\mu V^\nu,$$

where the metric tensor $g_{\mu\nu}$ is defined by $g_{00} = 1$, $g_{11} = g_{22} = g_{33} = -1$, with $g_{\mu\nu} = 0$ for $\mu \neq \nu$. Minkowsky sums of the form $\sum_\mu W_\mu V^\mu$ occur so often that we adopt the Einstein convention, that whenever we see a term with an index occurring once covariantly and once contravariantly (once down and once up) we understand that index to be summed (from 0 to 3) over without needing a summation sign. Einstein said of this "I have made a great discovery in mathematics".]

$\hat{P}^2$ is a Casimir operator of the Poincaré group. That is, it is made from the generators and commutes with all of them (and hence with every element of the group). This is why we can label representations with a definite value of p^2. Besides the momentum generators $\hat{P}^\mu$, we have the generators of the Lorentz group, the group which transforms vectors V^μ as observed by an inertial observer $\mathcal{O}$ to the vector V'^μ describing the same physics by another inertial observer $\mathcal{O}'$. This is given by a matrix $\Lambda^\mu{}_\nu$ with

$$V'^\mu = \Lambda^\mu{}_\nu V^\nu.$$

The infinitesimal generators for these finite transformations are linear combinations of a basis set of six matrices, as you will show for homework. They can be labelled with a pair $(\alpha\beta)$ of indices which each go from 0 to 3, but with $\hat{L}^{(\alpha\beta)} = -\hat{L}^{(\beta\alpha)}$. Note each of these can be considered a abstract generator of a symmetry of the whole physics, but when acting on a vector, each $L^{(\alpha\beta)}$ is a 4×4 matrix, really $L^{(\alpha\beta)\mu}{}_\nu$.

In the first homework assignment, you will show that the space of infinitesimal Lorentz transformations is in fact six dimensional, and you will verify

$$L^{(\alpha\beta)\mu}{}_\nu = ig^{\alpha\mu}\delta^\beta_\nu - ig^{\beta\mu}\delta^\alpha_\nu,$$

and find the Lie algebra of these generators. In the next homework, you will also see how $\hat{L}^{(0j)}$ generates boosts in the j direction and $\epsilon_{jk\ell}\hat{L}^{(jk)}$ generates rotations about the ℓ axis.

What other properties should be fixed for what constitutes a particle? That is, what commutes with the Poincaré group? From the generators we can form the Pauli-Lubanski vector

$$\hat{W}_\mu = \frac{1}{2}\,\epsilon_{\mu\nu\rho\sigma}\hat{P}^\nu \hat{L}^{(\rho\sigma)}.$$

One can show that the square, $\hat{W}^\mu\hat{W}_\mu$ commutes with all the generators of the Poincaré group, so it is also a Casimir operator and irreducible representations will have a definite value. Here $\epsilon^{\mu\nu\rho\sigma}$ is the totally antisymmetric Levi-Civita symbol[1], with $\epsilon^{0123} = 1$. We can understand the W^2 invariant for states of positive mass by going to the rest frame, so $p^\mu = (m, 0, 0, 0)$

[1] If you are not familiar with these antisymmetric symbols and how they are used, both in three dimensions, in four and in general, see the notes on "Using ϵ's and determinants" on the "Supplementary Notes" page from the web site.

and $W_j = \frac{1}{2} m \epsilon_{jk\ell} L^{(k\ell)}$, which is just m times the angular momentum in the j direction, or the spin.

The rotational invariance of the physical laws tells us that single particle states must behave under rotations as a representation of the rotation group, labelled by that spin, half of a natural number[2]. As we have seen, the three historic particles e, p and n, and many of the hadronic resonances, have spin 1/2, though the photon has spin 1 and the pion spin 0. We will need to develop equations describing particles with each of these spins.

In non-relativistic quantum mechanics, the behavior of particles is determined by the Schrödinger equation, which can be understood as a deBroglie interpretation of the classical hamiltonian

$$E = H = \frac{p^2}{2m} + V(\vec{r}) \tag{1}$$

where $E \to i\hbar \frac{\partial}{\partial t}$ and $\vec{p} \to -i\hbar\vec{\nabla}$, and this differential operator is supposed to act on the wave function. Thus

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

which is the Schrödinger equation.

Of course this relation is inappropriate for relativistic physics. In the first place, the very idea of a potential $V(\vec{r}_1, \vec{r}_2)$, which gives the force on one particle at time t in terms of the position of another particle at the same time, violates the relativistic concept that information cannot travel faster than light. The only exception is a delta function potential. If we ignore interactions, we can take the relativistic connection between energy and momentum

$$E^2 = c^2 \vec{p}^{\,2} + m^2 c^4, \tag{3}$$

which gives the differential operator equation

$$\hbar^2 \frac{\partial^2}{\partial t^2} - c^2 \hbar^2 \vec{\nabla}^{\,2} + m^2 c^4 \equiv 0$$

when acting on some function, sort of the wave function but now we will call it the particle's field.

[2] A natural number is 0 or a positive integer.

As we will be dealing exclusively with relativistic physics it is absurd to use different units for time (x^0) and for space ($\vec{x}$), so we will set $c = 1$ and measure time in meters. Also, as we will be strictly quantum-mechanical, we might as well set $\hbar = 1$ and measure energy in inverse meters or vice versa. In fact, we will generally use eV for units (or MeV or GeV, etc.)

Finally, we use the differential operator notation

$$\partial_\mu := \frac{\partial}{\partial x^\mu}.$$

Note the relative positions of the μ's, which gives us

$$\partial_\mu x^\nu = \delta^\nu_\mu.$$

Okay, now that we have simplified our notation, the differential equation which emerges from the correct kinematical connection of energy and momentum (3) is the Klein-Gordon equation:

$$\left(\partial_\mu \partial^\mu + m^2\right) \Phi = 0.$$

Klein and Gordon proposed this equation in 1926.

Notice that the Klein-Gordon equation acts on a single scalar field, and has solutions $\Phi \propto e^{i\vec{k}\cdot\vec{x} - i\omega t} = e^{-ik_\mu x^\mu}$, with definite values $\hat{P}^\mu = k^\mu$, and furthermore that $\hat{P}^\mu \hat{P}_\mu = m^2$ has a definite value. Under a Poincaré transformation Φ will be transformed to have a different, Lorentz transformed, k'^μ and will also pick up a phase from translation, but still have the same mass. In general, whenever we have a symmetry to the underlying physics, states of a system which obey the equations of motion must be transformed by the symmetry into other states which evolve according to the same equations of motion. As a consequence, the possible states of a system form representations of the symmetry group. Now we have rotational symmetry as part of Poincaré invariance. We are used to this idea from atomic physics, where rotational symmetry tells us that, in the absence of external fields, the states belong to multiplets of fixed angular momentum L^2. Such a multiplet can have several states $|\ell, m\rangle$ of different $L_z = \hbar m$, which are transformed into each other by rotations (other than along z). Because L^2 is invariant under the rotations, the states can be decomposed into separate *irreducible* representations with $L^2 = \hbar^2 \ell(\ell + 1)$ for integer ℓ. If we consider spin as well, it is J^2 which describes the multiplets, and the j in $J^2 = \hbar^2 j(j + 1)$ may be integer or half-integer. For our states of relativistic physics, this idea pertains to the spin. The Φ of our Klein-Gordon equation is spin 0, but there are other possibilities.

0.1 The Dirac Equation

For a scalar (spin 0) particle there is just one state with each value of $\vec{p}$, and the Klein-Gordon equation gives all the constraint needed. But for particles with higher spin, there is a multiplet of states for each $\vec{p}$. Each component will satisfy the Klein-Gordon equation, but the full equations of motion provide constraints between the components, which the Klein-Gordon equation does not do. This should be familiar from the Dirac equation for spin 1/2. You will recall that Dirac, trying to find an equation first order in $\partial/\partial t$ so as to avoid negative energy solutions, postulated

$$i\frac{\partial\psi(\vec{x},t)}{\partial t} = \left(-i\vec{\alpha}\cdot\vec{\nabla} + \beta m\right)\psi(\vec{x},t) \tag{4}$$

and found that for the solutions to be covariant under Lorentz transformations, we must have the three components of $\vec{\alpha}$ anticommute with each other and with β, and the square of each one to be 1. That is,

$$\{\alpha_j, \alpha_k\} := \alpha_j\alpha_k + \alpha_k\alpha_j = 2\delta_{jk}, \qquad \{\alpha_j, \beta\} = 0, \qquad \beta^2 = 1. \tag{5}$$

This, of course, means that α_j and β can't be ordinary numbers, but can be matrices. The minimum dimension for these matrices is 4, and one conventional choice for these 4×4 matrices is

$$\alpha_j = \begin{pmatrix} \mathbf{0} & \sigma_j \\ \sigma_j & \mathbf{0} \end{pmatrix}, \qquad \beta = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix}, \qquad \text{Dirac rep.} \tag{6}$$

where each of the entries is a 2×2 matrix, and the σ_j are the usual Pauli spin matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{7}$$

It is important to keep in mind that the particular solution (6) is not unique. The function ψ is a four-dimensional representation of the Lorentz group, and the values of each component depend on the basis vectors we use to expand objects in this four-dimensional space. Had we used a different basis, in which $\psi_a \to \sum_b U_{ab}\psi_b$, with U a unitary matrix, the equation would be satisfied with transformed $\vec{\alpha}' = U\vec{\alpha}U^{-1}$, $\beta' = U\vec{\beta}U^{-1}$. The new α' and β' would automatically satisfy (5). The representation (6) is particularly useful if we are dealing with particles of low energy, while for ultrarelativistic particles it is often more useful to use the Weyl representation

$$\alpha_j = \begin{pmatrix} -\sigma_j & \mathbf{0} \\ \mathbf{0} & \sigma_j \end{pmatrix}, \qquad \beta = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{pmatrix}, \qquad \text{Weyl rep.} \tag{8}$$

This is a course in relativistic physics, so we ought to convert our equation, written in three dimensional language distinguishing t from $\vec{x}$, into four dimensional language. If we multiply (4) on the left with β, we get

$$\left(i\gamma^\mu \frac{\partial}{\partial x^\mu} - m \right) \phi(x^\nu) = 0, \tag{9}$$

where $\gamma^0 = \beta$, and $\gamma^j = \beta\alpha_j$ for $j = 1, 2, 3$. In the Dirac representation,

$$\gamma^0 = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix}, \qquad \gamma^j = \begin{pmatrix} \mathbf{0} & \sigma_j \\ -\sigma_j & \mathbf{0} \end{pmatrix}, \qquad \text{Dirac rep.}$$

From (5) and observing that
$\{\beta\alpha^j, \beta\alpha^k\} = \beta\{\alpha^j, \beta\}\alpha^k + \beta\{\alpha^k, \beta\}\alpha^j - \beta^2\{\alpha^j, \alpha^k\} = -2\delta_{jk}$, we have

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}. \tag{10}$$

Also note that the $\vec{\alpha}$ and β are hermitian, but γ^j is antihermitian due to the reversal of the β and α, while of course $\gamma^0 = \beta$ is still hermitean.

If ψ is representing a state of definite momentum, $\psi(x^\mu) \propto e^{-ip_\nu x^\nu} = e^{-ip\cdot x}$, then $\frac{\partial}{\partial x^\mu} \to -ip_\mu$, and the Dirac equation becomes $(\gamma^\mu p_\mu - m)\tilde{\psi}(p^\nu) = 0$. If we distinguish the upper two components as ϕ and the lower two as χ, so

$$\psi = \begin{pmatrix} \phi \\ \chi \end{pmatrix}, \qquad 0 = (\gamma^\mu p_\mu - m)\,\psi = \begin{pmatrix} (E-m)\mathbf{1} & -\vec{\sigma}\cdot\vec{p} \\ \vec{\sigma}\cdot\vec{p} & -(E+m)\mathbf{1} \end{pmatrix} \begin{pmatrix} \phi \\ \chi \end{pmatrix},$$

or $(E-m)\,\phi = \vec{\sigma}\cdot\vec{p}\,\chi$, $(E+m)\,\chi = \vec{\sigma}\cdot\vec{p}\,\phi$. Plugging one into the other, we find, of course,

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

because[3] for any 3-vector $\vec{V}$,

$$(\vec{V}\cdot\vec{\sigma}\,)^2 = \sum_{jk} V_j V_k \sigma_j \sigma_k = \frac{1}{2}\sum_{jk} V_j V_k \{\sigma_j, \sigma_k\} = \sum_{jk} V_j V_k \delta_{jk} = V^2.$$

[3] More generally, as $\sigma_j\sigma_k = \delta_{jk} + i\epsilon_{jk\ell}\sigma_\ell$, we have $(\vec{V}\cdot\vec{\sigma})(\vec{W}\cdot\vec{\sigma}) = V_j W_k(\delta_{jk} + i\epsilon_{jk\ell}\sigma_\ell) = \vec{V}\cdot\vec{W} + i(\vec{V}\times\vec{W})\cdot\vec{\sigma}$.

But then $(E^2-\vec{p}^{\,2}-m^2)\,\phi = (p_\mu p^\mu - m^2)\,\phi = 0$, and ϕ obeys the Klein-Gordon equation. So does χ.

In non-relativistic physics we are used to considering the probability density as $\rho(\vec{x}) = \psi^\dagger(\vec{x})\psi(\vec{x})$. If we ask how that changes with time,

$$\frac{\partial\rho}{\partial t}(\vec{x}) = \frac{\partial\psi^\dagger}{\partial t}\psi + \psi^\dagger\frac{\partial\psi}{\partial t}.$$

From (4) $\frac{\partial\psi}{\partial t} = -\vec{\alpha}\cdot\vec{\nabla}\psi - i\beta m\psi$, and the hermitian conjugate gives $\frac{\partial\psi^\dagger}{\partial t} = -\vec{\nabla}\psi^\dagger\cdot\vec{\alpha}^\dagger + i\psi^\dagger\beta^\dagger m$. Both $\vec{\alpha}$ and β are hermitian, so we can drop the daggers on them, and we have

$$\frac{\partial\rho}{\partial t}(\vec{x}) = \left(-\vec{\nabla}\psi^\dagger\cdot\vec{\alpha} + im\psi^\dagger\beta\right)\psi + \psi^\dagger\left(-\vec{\alpha}\cdot\vec{\nabla}\psi - im\beta\psi\right) = -\vec{\nabla}\cdot\left(\psi^\dagger\vec{\alpha}\psi\right).$$

Thus if we define

$$j^\mu = (\psi^\dagger\psi, \psi^\dagger\vec{\alpha}\psi) \text{ we see } \frac{\partial j^\mu}{\partial x^\mu} = 0$$

or j^μ is a conserved current.

Lorentz symmetry

We see that the Dirac equation has, for each 3-momentum $\vec{p}$, positive energy solutions with an arbitrary two component ϕ, so two particle states, which are of course the two spin states. There are also two negative energy states. For $\vec{p} = 0$ these are particularly simple, as $\chi = 0$ for the positive energy states and $\phi = 0$ for the negative energy ones.

To find the more general form and better understand the spins, we ask how Lorentz transformations act on the wave functions. First lets review what it means for a field to be a scalar, vector, or other representation of rotations. Considering rotations is a passive sense, as a change of coordinates, we have that a rotation R changes the coordinates $\vec{x} \to \vec{x}\,'$ with $x'_j = R_{jk}x_k$. When we say that a field ψ_j transforms according to a representation $M_{jk}(R)$, $\psi'_j(\vec{x}\,') = M_{jk}(R)\psi_k(\vec{x})$. Notice the functions ψ and ψ' are evaluated for different arguments. For a scalar field $M = 1$, the new field at the point $\vec{x}\,'$ is the same as the old field at the point $\vec{x}$, unchanged at the same physical point, but the functions ψ' and ψ are different. Considered actively, by which I mean that the rotation is not a change of coordinate systems but a real rotation of the physics, we see that the rotation does change the field.

The same considerations apply to Lorentz transformations on fields which are functions on space-time, x^μ. A Lorentz transformation Λ acts as a linear transformation on the coordinates, $x^\mu \to \Lambda^\mu{}_\nu x^\nu$, and a field $\psi_a(x^\mu)$ which transforms according to an irreducible representation $D_{ab}(\Lambda)$ is transformed into ψ', with

$$\psi'_a(x'^\mu) = D_{ab}(\Lambda)\psi_b(x^\mu).$$

For the Dirac equation, if Lorentz invariance of the physics is to hold, the Lorentz transform of a field $\psi(x^\mu)$ which obeys the Dirac equation must also obey the Dirac equation, with the same m. Thus

$$\begin{aligned}
\gamma^\alpha \frac{\partial \psi'}{\partial x'^\alpha}\bigg|_{x'^\mu} &= -im\, \psi'|_{x'^\mu} \\
\gamma^\alpha \frac{\partial x^\beta}{\partial x'^\alpha}\bigg|_{x^\nu} \frac{\partial}{\partial x^\beta}\psi'(x'(x))\bigg|_{x^\nu} &= -im\, \psi'|_{x'^\nu} \\
\gamma^\alpha \frac{\partial x^\beta}{\partial x'^\alpha}\bigg|_{x^\nu} D(\Lambda)\, \frac{\partial}{\partial x^\beta}\psi\bigg|_{x^\nu} &= -im\, D(\Lambda)\psi|_{x^\nu}
\end{aligned}$$

Multiplying on the left by $D^{-1}(\Lambda)$ and using the Dirac equation on the right,

$$\begin{aligned}
D^{-1}(\Lambda)\gamma^\alpha \frac{\partial x^\beta}{\partial x'^\alpha} D(\Lambda)\partial_\beta \psi &= \gamma^\beta \partial_\beta \psi \\
\text{so} \qquad D^{-1}(\Lambda)\gamma^\alpha \frac{\partial x^\beta}{\partial x'^\alpha} D(\Lambda) &= \gamma^\beta .
\end{aligned}$$

Multiply $\dfrac{\partial x'^\rho}{\partial x^\beta} = \Lambda^\rho{}_\beta$ to get

$$D^{-1}(\Lambda)\gamma^\rho D(\Lambda) = \Lambda^\rho{}_\beta \gamma^\beta. \tag{11}$$

This tells us what it means for γ^ρ to act like a contravariant vector, even though its matrix elements are fixed numbers. It tells us as well what the representation $D(\Lambda)$ must be. Consider an infinitesimal Lorentz transformation $\Lambda^\mu{}_\nu = \delta^\mu_\nu - i\epsilon L^{(\alpha\beta)\mu}{}_\nu$ and $D_{ab} = \delta_{ab} - i\epsilon S_{ab}(L^{(\alpha\beta)})$, and we see

$$\left[\delta_{ab} + i\epsilon S_{ab}(L^{(\alpha\beta)})\right] \gamma^\rho_{bc} \left[\delta_{cd} - i\epsilon S_{cd}(L^{(\alpha\beta)})\right] = \left[\delta^\rho_\sigma - i\epsilon L^{(\alpha\beta)\rho}{}_\sigma\right] \gamma^\sigma_{ad}$$

which gives, to first order in ϵ,

$$\left[S(L^{(\alpha\beta)}), \gamma^\rho\right] = -L^{(\alpha\beta)\rho}{}_\sigma \gamma^\sigma = -ig^{\alpha\rho}\gamma^\beta + ig^{\beta\rho}\gamma^\alpha.$$

Then what is S? If ρ is neither α nor β, it commutes, and if it is one of these, we get the other. So try

$$S(L^{(\alpha\beta)}) = \frac{i}{4}\left[\gamma^\alpha, \gamma^\beta\right],$$

and indeed

$$\begin{aligned}\left[S(L^{(\alpha\beta)}), \gamma^\rho\right] &= \frac{i}{4}\left[\left[\gamma^\alpha, \gamma^\beta\right], \gamma^\rho\right] \\ &= \frac{i}{2}\left(\gamma^\alpha\left\{\gamma^\beta, \gamma^\rho\right\} - \gamma^\beta\left\{\gamma^\alpha, \gamma^\rho\right\}\right) = i\gamma^\alpha g^{\beta\rho} - i\gamma^\beta g^{\alpha\rho}\end{aligned}$$

in agreement with what is needed.

Now, in particular, consider a rotation about the z axis, which takes $t \to t' = t$, $z \to z' = z$, and $\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$, with $\Lambda = e^{-i\theta J_z}$ with the generator $J_z = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ when acting on a contravariant vector. In our four dimensional language $J_z = L^{(12)}$, and so it acts as $\frac{1}{2}\Sigma_z = \frac{i}{4}[\gamma^1, \gamma^2] = \frac{1}{2}\begin{pmatrix} \sigma_z & 0 \\ 0 & \sigma_z \end{pmatrix}$ where we define

$$\Sigma_j = \begin{pmatrix} \sigma_j & 0 \\ 0 & \sigma_j \end{pmatrix}$$

We see that spin is just what we expect.