PHY489 Lecture 14

The Dirac Equation

First a reminder (hopefully) from non-relativistic quantum mechanics:

$$E = \frac{\left|\vec{p}\right|^2}{2m} + V$$

To express this in the form of a wave equation we make the operator substitutions

$$\vec{p} \to -i\hbar\vec{\nabla} \qquad E \to i\hbar\frac{\partial}{\partial t}$$

and allow these to act on a wavefunction ψ

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V\psi = i\hbar\frac{\partial}{\partial t}\psi \quad \text{(Schrödinger equation)}$$

One can do the same thing for the relativistic case, starting with the usual relationship between the relativistic energy and momentum:

$$E^2 = \left|\vec{p}\right|^2 c^2 + m^2c^4 \quad \text{or} \quad p_\mu p^\mu - m^2c^2 = 0 \quad \text{(in relativistic notation)}$$

Starting with this, we make the operator substitution $p_\mu \to i\hbar\partial_\mu$ with $\partial_\mu \equiv \dfrac{\partial}{\partial x^\mu}$

and allow the resulting expression to act on a wavefunction ψ :

$$-\hbar^2\partial^\mu\partial_\mu\psi - m^2c^2\psi = 0 \quad \Rightarrow \boxed{-\frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} + \nabla^2\psi = \left(\frac{mc}{\hbar}\right)^2\psi}$$ (Klein-Gordon equation)

$$x^\mu = \left(ct, \quad x, \quad y, \quad z \right) \qquad x_\mu = g_{\mu\nu}x^\nu = \left(ct, \quad -x, \quad -y, \quad -z \right)$$

$$\frac{\partial}{\partial x^\mu} = \left(\frac{1}{c}\frac{\partial}{\partial t}, \quad \frac{\partial}{\partial x}, \quad \frac{\partial}{\partial y}, \quad \frac{\partial}{\partial z} \right) \quad = \quad \left(\frac{1}{c}\frac{\partial}{\partial t}, \quad \vec{\nabla} \right)$$

$$\frac{\partial}{\partial x_\mu} = \left(\frac{1}{c}\frac{\partial}{\partial t}, \quad -\frac{\partial}{\partial x}, \quad -\frac{\partial}{\partial y}, \quad -\frac{\partial}{\partial z} \right) \quad = \quad \left(\frac{1}{c}\frac{\partial}{\partial t}, \quad -\vec{\nabla} \right)$$

$$\partial_\mu\partial^\mu = \partial^\mu\partial_\mu = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2 \equiv \Box^2$$

This equation describes free particles without spin (*e.g.* spin 0). We didn't need it for the ABC model because we had the Feynman rules.

For historical reasons (see pg 227… or other texts) Dirac was looking for a relativistic wave equation that was linear in the time derivative (like the Schrödinger equation). His strategy was to try to factorize the expression for the energy/momentum relation that we began with: $p_\mu p^\mu - m^2c^2 = 0$

To see how this might work, consider first the simple case of a particle at rest:

$$p^\mu = \left(\; p^0, \quad \vec{0} \; \right)$$

$$p_\mu p^\mu - m^2c^2 = 0 \;\; \Rightarrow \;\; \left(p^0\right)^2 - m^2c^2 = 0 \;\; \Rightarrow \;\; \boxed{\left(p^0 + m\right)\left(p^0 - m\right) = 0}$$

This equation is easily satisfied by $p^0 = \pm mc$

For $\vec{p} \neq 0$ the factorization is less straightforward. Can try something like

$$\left(p^{\mu}p_{\mu} - m^2c^2\right) = \underbrace{\left(\beta^{\kappa}p_{\kappa} + mc\right)\left(\gamma^{\lambda}p_{\lambda} - mc\right)}_{\beta^{\kappa}\gamma^{\lambda}p_{\kappa}p_{\lambda} - m^2c^2 - mc\left(\beta^{\kappa}p_{\kappa} - \gamma^{\lambda}p_{\lambda}\right)}$$

Since there are no terms linear in p (or m) we require that $\gamma^k = \beta^k$, so we get

$$p^{\mu}p_{\mu} = \gamma^{\kappa}\gamma^{\lambda}p_{\kappa}p_{\lambda}$$ or, more explicitly:

$$\begin{aligned}\left(p_0\right)^2 - \left(p_1\right)^2 - \left(p_2\right)^2 - \left(p_3\right)^2 = &\left(\gamma^0\right)^2\left(p_0\right)^2 + \left(\gamma^1\right)^2\left(p_1\right)^2 + \left(\gamma^2\right)^2\left(p_2\right)^2 + \left(\gamma^3\right)^2\left(p_3\right)^2 \\ &+\left(\gamma^0\gamma^1 + \gamma^1\gamma^0\right)\left(p_0p_1\right) \\ &+\left(\gamma^0\gamma^2 + \gamma^2\gamma^0\right)\left(p_0p_2\right) \\ &+\left(\gamma^0\gamma^3 + \gamma^3\gamma^0\right)\left(p_0p_3\right) \\ &+\ldots\ldots\ldots\ldots\end{aligned}$$

The Gamma Matrices

Could chose $\gamma^0=1$, $\gamma^1=\gamma^2=\gamma^3=i$ but that leaves cross-terms that do NOT vanish if γ^i is simply a number. What we need is for the γ^i factors to anti-commute, *e.g.*

$$\left.\begin{aligned}&\left(\gamma^0\right)^2=1 \qquad \left(\gamma^1\right)^2=\left(\gamma^2\right)^2=\left(\gamma^3\right)^2=-1\\[2ex]&\gamma^\mu\gamma^\nu+\gamma^\nu\gamma^\mu=0 \quad \text{for} \quad \mu\neq\nu\end{aligned}\right\} \qquad \left\{\gamma^\mu,\gamma^\nu\right\}=2g^{\mu\nu}$$

Can't do this with numbers since they commute ($AB=BA$ always) but we can do it with matrices (which do not, in general, commute).

The smallest matrices that will work are 4x4. There are various conventions: we will follow that of the text (which follows the text by Bjorken & Drell).

$$\left.\gamma^0=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} \qquad \gamma^i=\begin{pmatrix}0 & \sigma_i\\ -\sigma_i & 0\end{pmatrix}\right\}$$

Here, by convention, each element is a 2x2 matrix, and the σ_i's are the Pauli spin matrices.

NB: these four objects are NOT the elements of a four-vector

Now we have $p^\mu p_\mu - m^2c^2 = \left(\gamma^\kappa p_\kappa + mc\right)\left(\gamma^\lambda p_\lambda - mc\right) = 0$ as desired.

Can choose either factor to define the Dirac equation.

By convention $\gamma^\mu p_\mu - mc = 0$.

Making the usual operator substitution $p_\mu \to i\hbar\partial_\mu$ and allowing the resulting expression to operate on the (now necessarily four component) wavefunction ψ yields the Dirac equation:

$$i\hbar\gamma^\mu\partial_\mu\psi - mc\psi = 0 \qquad \psi = \begin{pmatrix} \psi_0 \\ \psi_1 \\ \psi_2 \\ \psi_3 \end{pmatrix}$$

is called a Dirac spinor (more on this later)

It is important to note that while the Dirac spinor ψ is a four component object, is it NOT a four-vector. It's transformation properties are important however, and we will discuss these later on.

So we have what we wanted: a relativistic wave equation linear in $\dfrac{\partial}{\partial t}$:

$$\left(i\hbar\gamma^{\mu}\partial_{\mu} - mc\right)\psi = 0$$

Where ψ is a four component object called a Dirac spinor.

Transformation Properties of Dirac Spinor ψ

We will state the transformation properties without proof. Under a Lorentz transformation (along x) we have:

$$\psi \to \psi' = S\psi \qquad S = \begin{pmatrix} a_+ & a_-\sigma_1 \\ a_-\sigma_1 & a_+ \end{pmatrix}$$

Note that this is a 4x4 matrix. The σ_i are the 2x2 Pauli spin matrices (here we use σ_1 for the case of a boost along x).

with $a_\pm = \pm\sqrt{(\gamma \pm 1)/2}$ and $\gamma = (1-\beta^2)^{-1/2}$

Note that the wavefunction ψ is not an observable. What we want to construct (from this) are quantities that do have well defined transformation properties. For example, try to form a Lorentz scalar:

$$\psi^\dagger \psi = \begin{pmatrix} \psi_1^* & \psi_2^* & \psi_3^* & \psi_4^* \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} = |\psi_1|^2 + |\psi_2|^2 + |\psi_3|^2 + |\psi_4|^2$$

This is more or less the same thing we tried when we tried to define an invariant based on the four-vector x^μ.

Actually, didn't do this this year, but you can try it and see that it doesn't work

Transformation Properties of $\psi^\dagger\psi$

How does $\psi^\dagger\psi$ transform under a Lorentz transformation (again along x)?

$$\left(\psi^\dagger\psi\right) \to \left(\psi^\dagger\psi\right)' = \left(\psi'\right)^\dagger \psi' = \psi^\dagger S^\dagger S \psi \quad \neq \quad \psi^\dagger\psi \quad \text{since } S^\dagger S \neq 1$$

$$S^\dagger S = \begin{pmatrix} a_+ & 0 & 0 & a_- \\ 0 & a_+ & a_- & 0 \\ 0 & a_- & a_+ & 0 \\ a_- & 0 & 0 & a_+ \end{pmatrix} \begin{pmatrix} a_+ & 0 & 0 & a_- \\ 0 & a_+ & a_- & 0 \\ 0 & a_- & a_+ & 0 \\ a_- & 0 & 0 & a_+ \end{pmatrix} = \begin{pmatrix} a_+^2 + a_-^2 & 0 & 0 & 2a_+a_- \\ 0 & a_+^2 + a_-^2 & 2a_+a_- & 0 \\ 0 & 2a_+a_- & a_+^2 + a_-^2 & 0 \\ 2a_+a_- & 0 & 0 & a_+^2 + a_-^2 \end{pmatrix}$$

$$= \begin{pmatrix} a_+^2 + a_-^2 & 2a_-a_+\sigma_1 \\ 2a_-a_+\sigma_1 & a_+^2 + a_-^2 \end{pmatrix} = \gamma \begin{pmatrix} 1 & -\beta\sigma_1 \\ -\beta\sigma_1 & 1 \end{pmatrix}$$

[*e.g.* in 2x2 form]

since

$$\begin{cases} a_+^2 + a_-^2 = \frac{1}{2}\left(\gamma + 1\right) + \frac{1}{2}\left(\gamma - 1\right) = \gamma \\ 2a_+a_- = -\sqrt{\gamma - 1}\sqrt{\gamma + 1} = -\sqrt{\gamma^2 - 1} = -\sqrt{\frac{1}{1-\beta^2} - 1} = -\sqrt{\frac{1 - 1 + \beta^2}{1-\beta^2}} = -\sqrt{\beta^2\gamma^2} = -\beta\gamma \end{cases}$$

Transformation Properties $\psi^\dagger\psi$

So $\psi^\dagger\psi$ does not form an invariant (under Lorentz transformations).

[Just as the sum of the squares of the components of a four vector did not, when we were looking at four-vector invariants. In that case, needed a relative negative sign on the spatial components, for which we introduced the covariant four vector to define the invariant $a^\mu a_\mu$.]

Need to do something similar here. Note that the elements of $S^\dagger S$ are either

$$a_+^2 + a_-^2 \quad \text{or} \quad 2a_+a_- \;(\text{from } a_+a_- + a_-a_+)$$

If the sign were reversed in each case, we would get

$$a_+^2 - a_-^2 = \frac{1}{2}(\gamma+1) - \frac{1}{2}(\gamma-1) = 1 \qquad\qquad a_+a_- - a_-a_+ = 0$$

So we can get what we need if we introduce a matrix that effects the required change of sign: (γ^0 will do the trick).

The Adjoint Spinor

Define the adjoint spinor as $\bar{\psi} \equiv \psi^{\dagger}\gamma^{0}$

e.g. $$\begin{pmatrix} \psi_1^* & \psi_2^* & \psi_3^* & \psi_4^* \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} \psi_1^* & \psi_2^* & -\psi_3^* & -\psi_4^* \end{pmatrix}$$

The quantity $\bar{\psi}\psi = \psi^{\dagger}\gamma^{0}\psi$ is a Lorentz invariant scalar quantity, since

$$\left(\bar{\psi}\psi\right) \to \left(\bar{\psi}\psi\right)' = \left(\psi^{\dagger}\right)' \gamma^{0}\psi' = \psi^{\dagger}S^{\dagger}\gamma^{0}S\psi \quad = \quad \bar{\psi}\psi$$

$$\left[S^{\dagger}\gamma^{0}S = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \gamma^{0} \right]$$

So we know that this quantity is a Lorentz scalar. We can also ask whether is it a regular scalar or a pseudoscalar. That is, how does it transform under parity?

To determine how the quantity $\bar{\psi}\psi$ behaves under parity, first need to know how ψ transforms:

Quote without proof: under parity we have $\psi \to \psi' = \gamma^0\psi$

$$\left(\bar{\psi}\psi\right) \to \left(\bar{\psi}\psi\right)' = \left(\psi^{\dagger}\right)' \gamma^0\psi' = \psi^{\dagger}\left(\gamma^0\right)^{\dagger}\gamma^0\gamma^0\psi \;=\; \psi^{\dagger}\gamma^0\psi = \bar{\psi}\psi$$

since $\gamma^{0\dagger} = \gamma^0$ $\quad \left(\gamma^0\right)^2 = 1$ [e.g. the 4x4 identity matrix]

Can also make a pseudoscalar combination: $\bar{\psi}\gamma^5\psi$ with $\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3$

$$\gamma^5 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$ in the usual 2x2 form

Parity Transformation of $\bar{\psi}\gamma^5\psi$

Under parity, we have

$$\left(\bar{\psi}\gamma^5\psi\right) \to \left(\bar{\psi}\gamma^5\psi\right)' = \left(\psi^\dagger\right)' \gamma^0\gamma^5\psi' = \psi^\dagger\gamma^0\gamma^0\gamma^5\gamma^0\psi \;\;=\;\; \psi^\dagger\gamma^5\gamma^0\psi$$

$$= \; -\psi^\dagger\gamma^0\gamma^5\psi = -\bar{\psi}\gamma^5\psi \quad \text{since} \quad \gamma^0\gamma^5 = -\gamma^5\gamma^0$$

γ^0 anti-commutes with γ^1, γ^2, γ^3 and commutes with itself, so

$$\gamma^5\gamma^0 = i\gamma^0\gamma^1\gamma^2\gamma^3\gamma^0 = \left(-1\right)^3 \gamma^0 i\gamma^0\gamma^1\gamma^2\gamma^3 = -\gamma^0\gamma^5$$

γ^5 anti-commutes with all other γ matrices

$$\left\{\gamma^\mu, \gamma^5\right\} = 0 \qquad \left(\gamma^5\right)^2 = 1 \qquad \left(\gamma^5\right)^\dagger = \gamma^5$$

Can also show that $\bar{\psi}\gamma^5\psi$ is Lorentz invariant (Problem 7.12)

Lorentz Invariance of $\bar{\psi}\gamma^5\psi$

Under Lorentz transformation, we have

$$\left(\bar{\psi}\gamma^5\psi\right) \to \left(\bar{\psi}\gamma^5\psi\right)' = \left(\psi^\dagger\right)'\gamma^0\gamma^5\psi' = \psi^\dagger S^\dagger\gamma^0\gamma^5 S\psi$$

Can show that S, γ^5 commute:

$$S\gamma^5 = \begin{pmatrix} a_+ & a_-\sigma_1 \\ a_-\sigma_1 & a_+ \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} a_-\sigma_1 & a_+ \\ a_+ & a_-\sigma_1 \end{pmatrix}$$

$$\gamma^5 S = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_+ & a_-\sigma_1 \\ a_-\sigma_1 & a_+ \end{pmatrix} = \begin{pmatrix} a_-\sigma_1 & a_+ \\ a_+ & a_-\sigma_1 \end{pmatrix}$$

So $\quad \psi^\dagger S^\dagger\gamma^0\gamma^5 S\psi = \psi^\dagger S^\dagger\gamma^0 S\gamma^5\psi = \psi^\dagger\gamma^0\gamma^5\psi = \bar{\psi}\gamma^5\psi$

Reminder: the form of S used is for a Lorentz transformation along x.

Other Lorentz Invariant Quantities

There are 16 products of the form $\psi_i^* \psi_j$. Linear combinations of these can be used to form quantities having distinct transformation properties:

	Quantity	*Components*	*Parity*
Scalar	$\bar{\psi}\psi$	1	$+$
Pseudoscalar	$\bar{\psi}\gamma^5\psi$	1	$-$
Vector	$\bar{\psi}\gamma^\mu\psi$	4	*spatial* $-$
Pseudovector	$\bar{\psi}\gamma^5\gamma^\mu\psi$	4	*spatial* $+$
Tensor	$\bar{\psi}\sigma^{\mu\nu}\psi$	6	

$$\sigma^{\mu\nu} \equiv \frac{i}{2}\left[\gamma^\mu, \gamma^\nu\right] = \frac{i}{2}\left(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu\right)$$

[traceless, antisymmetric, so 6 components]