THE DIRAC EQUATION (A REVIEW)

We will try to find the relativistic wave equation for a particle. First, we introduce four dimensional notation for a vector by writing $x^\mu = (x^0, x^1, x^2, x^3) = (ct, x, y, z)$. The metric tensor of special relativity is $\eta_{\nu\mu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$. This allows us to introduce $x_\nu = \eta_{\nu\mu} x^\mu$ where $x_\nu = (x_0, x_1, x_2, x_3) = (x^0, -x^1, -x^2, -x^3)$. The summation convention has been used; when the same index appears as a superscript and a subscript, then that index is summed over. Thus, $x_\nu = \eta_{\nu 0} x^0 + \eta_{\nu 1} x^1 + \eta_{\nu 2} x^2 + \eta_{\nu 3} x^3$. The scalar product $x \cdot x$ is defined by $x \cdot x = x_\nu x^\nu = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = (x^0)^2 - \mathbf{x} \cdot \mathbf{x}$.

For a free relativistic particle, $E^2 = p^2 c^2 + m^2 c^4$. Using the correspondence that $E = i\hbar \partial/\partial t = i\hbar c \partial/\partial x^0$ and $\mathbf{p} = -i\hbar \nabla$ leads to

$$-\hbar^2 \frac{\partial^2 \phi}{\partial t^2} = -\hbar^2 c^2 \frac{\partial^2 \phi}{\partial (x^0)^2} = -\hbar^2 c^2 \nabla^2 \phi + m^2 c^4 \phi \,. \tag{1}$$

Substituting the plane wave solution $\phi \propto \exp(-ip \cdot x/\hbar)$ into Eq. (1), yields

Date: April 1, 2012.

$$E^2\phi = (p^2c^2 + m^2c^4)\phi, \tag{2}$$

and we recover $E^2 = p^2c^2+m^2c^4$, or $E = \pm\sqrt{p^2c^2 + m^2c^4}$ as is required. Eq. (1), an equation second order in all derivatives, was originally rejected because of the negative energy solution. Later, it was applied to mesons.

Another attempt to find a relativistic wave equation was made by Dirac. Observe that the Schrödinger equation is linear in the time derivative, but of second order in the spatial derivatives. Try to find a relativistic wave equation which is linear in all derivatives. For a free particle, try

$$ic\hbar\partial\phi/\partial x^0 = -ic\hbar(\alpha^1\partial\phi/\partial x^1+\alpha^2\partial\phi/\partial x^2+\alpha^3\partial\phi/\partial x^3)+\beta mc^2\phi = H\phi. \tag{3}$$

where $\alpha^1, \alpha^2, \alpha^3$, and β are to be determined, and H is the hamiltonian. Notice that if ϕ is replaced by the usual plane wave, then the coefficients of ϕ after differentiation have the dimensions of energy if $\alpha^i (i = 1, 2, 3)$ and β are dimensionless. Apply the operator below to Eq. (3)

$$ic\hbar\partial/\partial x^0 - ic\hbar(\alpha^1\partial/\partial x^1 + \alpha^2\partial/\partial x^2 + \alpha^3\partial/\partial x^3) + \beta mc^2, \tag{4}$$

and find

$$-\hbar^2 c^2 \frac{\partial^2\phi}{\partial(x^0)^2} = -\hbar^2 c^2 \sum_{i,j=1}^{3} \frac{(\alpha^j\alpha^i + \alpha^i\alpha^j)}{2} \frac{\partial}{\partial x^j}\frac{\partial\phi}{\partial x^i} - i\hbar m c^3 \sum_{i=1}^{3} (\alpha^i\beta + \beta\alpha^i)\frac{\partial\phi}{\partial x^i} + \beta^2 m^2 c^4 \phi. \quad (5)$$

Eq. (5) shall be required to reduce to Eq. (1). Notice that the first term on the right hand side of Eq. (5) will contain

$$(\alpha^1)^2 \frac{\partial^2\phi}{\partial(x^1)^2} + (\alpha^2)^2 \frac{\partial^2\phi}{\partial(x^2)^2} + (\alpha^3)^2 \frac{\partial^2\phi}{\partial(x^3)^2} \quad (6)$$

plus cross terms such as

$$(\alpha^1\alpha^2 + \alpha^2\alpha^1)\frac{\partial^2\phi}{\partial x^1 \partial x^2}. \quad (7)$$

In order that Eq. (5) reduces to Eq. (1), $(\alpha^1)^2 = (\alpha^2)^2 = (\alpha^3)^2 = 1$, $\beta^2 = 1$, $(\alpha^1\alpha^2 + \alpha^2\alpha^1) = 0$, and the other cross terms vanish including $\alpha^i\beta + \beta\alpha^i$. The cross terms can not vanish if the α^i and β are numbers. Matrices can anti-commute, so we assume they are matrices. This means that the 1 in the equations $(\alpha^i)^2 = \beta^2 = 1$ is actually the unit matrix, and $\partial\phi/\partial x^0$ in Eq. (3) is multiplied by the unit matrix. In addition, take ϕ to be a column matrix. This allows spin to appear in the theory naturally.

Suppose now that λ^i_j is an eigenvalue of the N by N matrix α^i where N is to be determined. Here j indicates which eigenvalue. Let χ^i_j be the eigenvector associated with the eigenvalue λ^i_j. χ^i_j will be a 1 by N column matrix. The eigenvalue equation takes the form $\alpha^i\chi^i_j = \lambda^i_j\chi^j_j$. Apply α^i to the above equation, and find

$$(\alpha^i)^2\chi^i_j = \chi^i_j = \alpha^i\lambda^i_j\chi^i_j = (\lambda^i_j)^2\chi^i_j. \tag{8}$$

Thus $(\lambda^i_j)^2 = 1$, and $\lambda^i_j = \pm 1$. It will be shown next that N is even. Start with $\alpha^i\beta = -\beta\alpha^i$, then $\beta\alpha^i\beta = -\beta^2\alpha^i = -\alpha^i$. Take the trace of both sides, and find $\mathrm{Tr}\,\beta\alpha^i\beta$ = -$\mathrm{Tr}\,\alpha^i$. Use the trace property that $\mathrm{Tr}\,\beta\alpha^i\beta$ =$\mathrm{Tr}\alpha^i\beta\beta$, and $\beta^2 = 1$ to show that $\mathrm{Tr}\,\alpha^i = -\mathrm{Tr}\,\alpha^i$, or $\mathrm{Tr}\,\alpha^i = 0$. The Trace of a matrix is just the sum of its eigenvalues. Since the trace is zero, N must be even. $N = 2$ is ruled out because there are not four 2 by 2 anti-commuting matrices. We proceed with $N = 4$. Thus, the α^i and β are four by four matrices, and ϕ is a one by four column matrix.

It is convenient to multiply Eq. (3) by the matrix β. Introduce the notation $\gamma^0 = \beta$, and $\gamma^i = \beta\alpha^i$. Then Eq. (3) can be written

$$i\hbar\big(\gamma^0\frac{\partial}{\partial x^0} + \gamma^1\frac{\partial}{\partial x^1} + \gamma^2\frac{\partial}{\partial x^2} + \gamma^3\frac{\partial}{\partial x^3}\big)\phi - mc\phi = 0 \tag{9}$$

where $mc\phi$ is multiplied by the unit matrix. We assume the plane wave solution $\phi \propto u \exp(-ip \cdot x/\hbar)$ where u is the one by four column matrix $\begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix}$ with u_i to be determined. Notice that $(\gamma^i)^2 = \beta\,\alpha^i \beta\,\alpha^i = -\alpha^i\,\beta^2\,\alpha^i = -(\alpha^i)^2 = -1$ and $(\gamma^0)^2 = 1$. Similarly, it can be shown that $\gamma^i\gamma^j + \gamma^j\gamma^i = 0$. These results can be written $\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}$ where $g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$. The product of a four by four matrix with a four by four matrix is a four by four matrix, so each entry in $g^{\mu\nu}$ is a four by four diagonal matrix. A possible representation for the matrices is $\gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}$ and $\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ where the σ^i are the Pauli matrices. Recall that $\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, and $\sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. As an illustration, we now calculate $\gamma^1\gamma^3 + \gamma^3\gamma^1$ in the chosen representation.

$$\gamma^1\gamma^3 = \begin{pmatrix} 0 & \sigma^1 \\ -\sigma^1 & 0 \end{pmatrix} \begin{pmatrix} 0 & \sigma^3 \\ -\sigma^3 & 0 \end{pmatrix} = \begin{pmatrix} -\sigma^1\sigma^3 & 0 \\ 0 & -\sigma^1\sigma^3 \end{pmatrix}. \tag{10}$$

Similarly, $\gamma^3\gamma^1 = \begin{pmatrix} -\sigma^3\sigma^1 & 0 \\ 0 & -\sigma^3\sigma^1 \end{pmatrix}$. Next, note that

$$\sigma^1\sigma^3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{11}$$

Similarly, $\sigma^3\sigma^1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Finally,

$$\gamma^1\gamma^3 + \gamma^3\gamma^1 = \begin{pmatrix} -\sigma^1\sigma^3 - \sigma^3\sigma^1 & 0 \\ 0 & -\sigma^1\sigma^3 - \sigma^3\sigma^1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \quad (12)$$

The other identities can be similarly verified.

For $\mathbf{p} = 0$, the Dirac equation takes the form

$$\left(i\hbar\gamma^0 \frac{\partial}{\partial x^0} - mc\right)\phi = 0 \quad (13)$$

where $\phi \propto u \exp(-ip^0x^0/\hbar)$ and u is a 1 by 4 matrix described above. Eq. (13) can be written

$$i\hbar\frac{\partial \exp(-ip^0x^0/\hbar)}{\partial x^0} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} - mc\exp(-ip^0x^0/\hbar) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = 0. \quad (14)$$

Carrying out the matrix multiplication leads to

$$(p^0 - mc)u_1 = 0,\ (p^0 - mc)u_2 = 0,\ (-p^0 - mc)u_3 = 0,\ (-p^0 - mc)u_4 = 0. \quad (15)$$

One possible set of solutions is found by setting three of the u_i to zero, and setting the remaining one equal to one. Thus, the four solutions to Eq. (13) are

$$\phi^1 = \exp(-imcx^0/\hbar)\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \phi^2 = \exp(-imcx^0/\hbar)\begin{pmatrix}0\\1\\0\\0\end{pmatrix}$$

$$\phi^3 = \exp(+imcx^0/\hbar)\begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \phi^4 = \exp(+imcx^0/\hbar)\begin{pmatrix}0\\0\\0\\1\end{pmatrix}. \quad (16)$$

Since $p^0 = mc$ for the first two solutions, we shall associate a positive energy spin up electron with ϕ^1, and we shall associate a positive energy spin down electron with ϕ^2. The two negative energy solutions will be discussed in more detail when positrons are treated.

For $\mathbf{p} \neq 0$, assume that a solution to Eq. (9) will take the form $\phi \propto u \exp(-ip \cdot x/\hbar)$. Here $u = \begin{pmatrix}u_1\\u_2\\u_3\\u_4\end{pmatrix}$ and the four vector p is given by $p^\mu = (p^0 = E/c, p^1, p^2, p^3)$. Substitution into Eq. (9) gives the following result:

$$(p^0 - mc)u_1 - (p^1 - ip^2)u_4 - p^3 u_3 = 0 \quad (17)$$

$$(p^0 - mc)u_2 - (p^1 + ip^2)u_3 + p^3 u_4 = 0 \quad (18)$$

$$-(p^0 + mc)u_3 + (p^1 - ip^2)u_2 + p^3 u_1 = 0 \quad (19)$$

$$-(p^0 + mc)u_4 + (p^1 + ip^2)u_1 - p^3 u_2 = 0. \quad (20)$$

For the positive energy spin up solution, set $u_2 = 0$ and find $u_3 = p^3 u_1/(p^0 + mc)$ and $u_4 = (p^1 + ip^2)u_1/(p^0 + mc)$. So

$$\phi^1 \propto \exp\left(-ip \cdot x/\hbar\right) \begin{pmatrix} u_1 \\ 0 \\ p^3 c u_1/E + mc^2) \\ (p^1 + ip^2) c u_1/(E + mc^2) \end{pmatrix} \tag{21}$$

In order to agree with reference (6) in the Mott Rutherford paper, u_1 is set equal to $\sqrt{(E + mc^2)/(2mc^2)}$. Write $\phi^1(x) \propto w^1(p) \exp\left(-ip \cdot x/\hbar\right)$ where

$$w^1(p) = \sqrt{(E + mc^2)/(2mc^2)} \begin{pmatrix} 1 \\ 0 \\ p^3 c/(E + mc^2) \\ (p^1 + ip^2) c/(E + mc^2) \end{pmatrix} . \tag{22}$$

Set $\bar{w} = w^\dagger \gamma^0$ and notice that

$$\bar{w}^1 w^1 = (w^1)^\dagger \gamma^0 w^1 = (E+mc^2)/(2mc^2) \begin{pmatrix} 1 & 0 & p^3 c/(E + mc^2) & (p^1 - ip^2) c/(E + mc^2) \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ p^3 c/(E + mc^2) \\ (p^1 + ip^2) c/(E + mc^2) \end{pmatrix} = 1 . \tag{23}$$

A similar calculation yields $(w^1)^\dagger w^1 = E/mc^2$. Introduce the four vector $k = p/\hbar$. Then $\phi^1(x) \propto w^1(p) \exp\left(-ik \cdot x\right)$. The normalization that $\int (\phi^1_2)^\dagger \phi^1_1 \, d^3x = \delta^3(\mathbf{k_2} - \mathbf{k_1})$ requires that $\phi^1(x) = \sqrt{mc^2/(2\pi)^3 E}\, w^1(p) \exp\left(-ik \cdot x\right)$. Alternatively, if box normalization is used, $\int (\phi^1_2)^\dagger \phi^1_1 \, d^3x = \delta_{\mathbf{k_2},\mathbf{k_1}}$ and $\phi^1(x) = \sqrt{mc^2/(EV)}\, w^1(p) \exp\left(-ik \cdot x\right)$.

For the positive energy spin down solution, set $u_1 = 0$ and find

$$\phi^2 \propto \exp\left(-ik\cdot x\right)\begin{pmatrix} 0 \\ u_2 \\ (p^1+ip^2)cu_2/(E+mc^2) \\ -p^3cu_2/(E+mc^2) \end{pmatrix} \tag{24}$$

Set u_2 equal to $\sqrt{(E+mc^2)/(2mc^2)}$. Write $\phi^2(x) \propto w^2(p)\exp\left(-ik\cdot x\right)$ where

$$w^2(p) = \sqrt{(E+mc^2)/(2mc^2)}\begin{pmatrix} 0 \\ 1 \\ (p^1-ip^2)c/(E+mc^2) \\ -p^3c/(E+mc^2) \end{pmatrix}. \tag{25}$$

Notice that $\bar{w}^2w^2 = 1$, and $\bar{w}^1w^2 = \bar{w}^2w^1 = 0$. Also, $(w^2)^\dagger w^2 = E/mc^2$, and $(w^2)^\dagger w^1 = (w^1)^\dagger w^2 = 0$. Box normalization will require $\phi^2(x) = \sqrt{mc^2/(EV)}\, w^2(p)\exp\left(-ik\cdot x\right)$. Introduce the notation $u(p, s_z = +\hbar/2) = w^1(p)$ and $u(p, s_z = -\hbar/2) = w^2(p)$. For an arbitrary spin s, write $\phi(x) = \sqrt{mc^2/(EV)}\, u(p,s)\exp\left(-ik\cdot x\right)$ where the column matrix $u(p,s)$ is a linear combination of $w^1(p)$ and $w^2(p)$.

This paper justifies the wave functions used in the paper on Mott-Rutherford scattering.

In special relativity, the three-dimensional vector potential and the potential, $A^0 = \Phi$, are combined to form a four-vector $A = (A^0, A^1, A^2, A^3)$. The replacement $p \to -i\hbar\gamma^\mu\partial/\partial x^\mu - e\gamma^\mu A_\mu$ yields the Dirac equation

$$i\hbar\gamma^\mu\frac{\partial\psi}{\partial x^\mu} - e\gamma^\mu A_\mu\psi - mc\psi = 0. \tag{26}$$

for a particle in an electro-magnetic field. This is the equation that we will use to solve scattering problems in QED. The next two review papers will show how this equation is used to calculate the S-matrix.