[Dirac equation] Derivation of the Dirac equation and the non-relativistic limit

Paul Dirac, motivated by the desire for describing a spin-1/2 particle (e.g., an electron) in an electromagnetic field which could not be done in a natural manner by the Klein-Gordon equation, derived a relativistic wave equation in 1928, now known as the Dirac equation, which is celebrated as one of the most significant accomplishment in physics of the 20th century. It's sole dependence on first order in the time derivative, $\partial/\partial t$, avoids the negative probability-current, $|\psi|^2 < 0$ and the negative-energy solutions as those appeared in the Klein-Gordon equation which is of quadratic in $\partial/\partial t$.

First, to include electromagnetic interaction, we substitute the momentum, $\vec{\mathbf{p}}$, for $\vec{\mathbf{p}} - e\vec{\mathbf{A}}/c$. Then for a spin-1/2 particle the kinetic energy, $\text{K.E.} = (\vec{\sigma}\cdot\vec{\mathbf{p}})^2 \big/ 2m$, becomes

$$\begin{aligned}\frac{1}{2m}\vec{\sigma}\cdot\left(\vec{\mathbf{p}} - \frac{e\vec{\mathbf{A}}}{c}\right)\vec{\sigma}\cdot\left(\vec{\mathbf{p}} - \frac{e\vec{\mathbf{A}}}{c}\right) &= \frac{1}{2m}\left(\vec{\mathbf{p}} - \frac{e\vec{\mathbf{A}}}{c}\right)^2 + \frac{i}{2m}\vec{\sigma}\cdot\left[\left(\vec{\mathbf{p}} - \frac{e\vec{\mathbf{A}}}{c}\right)\times\left(\vec{\mathbf{p}} - \frac{e\vec{\mathbf{A}}}{c}\right)\right] \\ &= \frac{1}{2m}\left(\vec{\mathbf{p}} - \frac{e\vec{\mathbf{A}}}{c}\right)^2 - \frac{e\hbar}{2mc}\vec{\sigma}\cdot\vec{\mathbf{B}}.\end{aligned} \tag{1}$$

We have used the relation $\vec{\mathbf{p}}\times\vec{\mathbf{A}} = -i\hbar\left(\vec{\nabla}\times\vec{\mathbf{A}}\right) - \vec{\mathbf{A}}\times\vec{\mathbf{p}}$ and $\vec{\mathbf{B}} = \vec{\nabla}\times\vec{\mathbf{A}}$. Note that (1) calls for a *g-factor*, $g = 2$ for the electron so that $\mu_{spin} = g\mu_B$, as observed experimentally.[1] Next, to derive the relativistic wave equation, we start with

$$\frac{E^2}{c^2} - \vec{\mathbf{p}}^2 = (mc)^2 \rightarrow \left(\frac{E}{c} - \vec{\sigma}\cdot\vec{\mathbf{p}}\right)\left(\frac{E}{c} - \vec{\sigma}\cdot\vec{\mathbf{p}}\right) = (mc)^2, \tag{2}$$

and use the operator forms

$$E = i\hbar\frac{\partial}{\partial t}, \tag{3}$$

and

$$\vec{\mathbf{p}} = -i\hbar\vec{\nabla}, \tag{4}$$

to obtain

$$\left(\frac{i\hbar}{c}\frac{\partial}{\partial t} + \vec{\sigma}\cdot i\hbar\vec{\nabla}\right)\left(\frac{i\hbar}{c}\frac{\partial}{\partial t} - \vec{\sigma}\cdot i\hbar\vec{\nabla}\right)\phi = \left(mc^2\right)\phi. \tag{5}$$

Recollection of the 2 x 2 Pauli matrices

---

[1] More precisely measured value is g = 2.002319..., also predicted by the theory of quantum electrodynamics, $g = 2\left(1 + \frac{\alpha}{2\pi} + ...\right)$, where α is the fine structure constant.

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{6}$$

implies the wavefunction in (5) has two components. In order to convert (5) to equations having linear time derivative, we define

$$\phi^{(R)} \equiv \frac{1}{mc}\left(\frac{i\hbar}{c}\frac{\partial}{\partial t} - \vec{\sigma}\cdot i\hbar\vec{\nabla}\right)\phi, \text{ and } \phi^{(L)} \equiv \phi. \tag{7}$$

Now multiply $\left(\frac{i\hbar}{c}\frac{\partial}{\partial t} + \vec{\sigma}\cdot i\hbar\vec{\nabla}\right)$ on the left to both sides of (5), we obtain two first-order equations,

$$\begin{aligned} \left(i\hbar\vec{\sigma}\cdot\vec{\nabla} - \frac{i\hbar}{c}\frac{\partial}{\partial t}\right)\phi^{(L)} &= -mc\phi^{(R)}, \\ \left(-i\hbar\vec{\sigma}\cdot\vec{\nabla} - \frac{i\hbar}{c}\frac{\partial}{\partial t}\right)\phi^{(R)} &= -mc\phi^{(L)}. \end{aligned} \tag{8}$$

Taking the sum and difference of these two equations and defining a four-component wavefunction

$$\psi \equiv \begin{pmatrix} \Psi_A \\ \Psi_B \end{pmatrix} = \begin{pmatrix} \phi^{(R)} + \phi^{(L)} \\ \phi^{(R)} - \phi^{(L)} \end{pmatrix}, \tag{9}$$

we have the Dirac equation:

$$\begin{pmatrix} -\frac{i\hbar}{c}\frac{\partial}{\partial t} & -i\hbar\vec{\sigma}\cdot\vec{\nabla} \\ i\hbar\vec{\sigma}\cdot\vec{\nabla} & \frac{i\hbar}{c}\frac{\partial}{\partial t} \end{pmatrix}\begin{pmatrix} \Psi_A \\ \Psi_B \end{pmatrix} = -mc\begin{pmatrix} \Psi_A \\ \Psi_B \end{pmatrix}, \tag{10}$$

or more concisely

$$\left(\gamma_\mu \frac{\partial}{\partial x_\mu} + \frac{mc}{\hbar}\right)\Psi = 0. \tag{11}$$

Here,

$$x_\mu = (\mathbf{\vec{x}}, ct), \quad \gamma_\mu = (\vec{\gamma}, \gamma_4); \quad \mu = 1,2,3,4. \tag{12}$$

$$\gamma_k = \begin{pmatrix} 0 & -i\sigma_k \\ i\sigma_k & 0 \end{pmatrix}, \quad \gamma_4 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \tag{13}$$

and I is a 2 x 2 unit matrix.

Next, we derive the nonrelativistic approximation of the Dirac equation for an electron. Again, using the replacement

$$-i\hbar\frac{\partial}{\partial x_\mu}\rightarrow -i\hbar\frac{\partial}{\partial x_\mu}-\frac{e}{c}A_\mu, \tag{14}$$

where $A_\mu=\left(\vec{\mathbf{A}},i\varphi\right)$ is the vector and scalar potential of the electromagnetic field and is assumed to be time independent. Furthermore, let ψ and E be the eigenfunction and eigenvalue of $i\hbar\,\partial/\partial t$, respectively, (10) can be written as (4.19) given in the book

$$\left\{\begin{array}{l}\left(E-e\phi-mc^2\right)\Psi_A-c\vec{\sigma}\cdot\left(\vec{\mathbf{p}}-\dfrac{e\vec{\mathbf{A}}}{c}\right)\Psi_B=0\\ \left(E-e\phi+mc^2\right)\Psi_B-c\vec{\sigma}\cdot\left(\vec{\mathbf{p}}-\dfrac{e\vec{\mathbf{A}}}{c}\right)\Psi_A=0\end{array}\right\}. \tag{4.19}$$

Moving toward the nonrelativistic limit, we use the second equation in (4.19) to eliminate Ψ_B and assume $E\approx mc^2$ and $|e\varphi|\ll mc^2$. Through an expansion of the result in powers of $\left(v/c\right)^2$ and take only the zeroth-order term, we obtain the *Schrödinger-Pauli wave function* in nonrelativistic quantum mechanics:

$$\left[\frac{1}{2m}\left(\vec{\mathbf{p}}-\frac{e\vec{\mathbf{A}}}{c}\right)^2-\frac{e\hbar}{2mc}\vec{\sigma}\cdot\vec{\mathbf{B}}+e\varphi\right]\Psi_A=E^{NR}\Psi_A, \tag{15}$$

which is equivalent to (1). It can be shown that Ψ_B is smaller than Ψ_A by a factor of $\sim v/2c$, hence Ψ_A and Ψ_B are called the *large* and *small component*s, respectively.

Eq. (15) is powerful in explaining many important phenomena in electrostatics, including the plane-wave solution of a free electron, the spin-orbit force, and the Darwin term in hydrogen atom. Therefore, it is reasonable to ask if the Dirac equation can be equally successful in applying to the proton and neutron. The answer is No! This stems from the fact that, unlike the electron, the proton and neutron are not elemental particles, i.e., they possess internal structures which require theoretical treatment beyond the Dirac equation. Nevertheless, it is still instructive to go over the result for the electron and then change the corresponding quantities of the electron to those for the neutron and see to what extent physical interpretation of the different contributions can be attained. This is done in {§4.2:132}.