

9.6: Dirac's Theory

The real breakthrough toward the quantum relativistic theory of electrons (and other spin- 1/2 fermions) was achieved in 1928 by P. A. M. Dirac. For that time, the structure of his theory was highly nontrivial. Namely, while formally preserving, in the coordinate representation, the same Schrödingerpicture equation of quantum dynamics as in the non-relativistic quantum mechanics, ⁵¹

$$i\hbar \frac{\partial \Psi}{\partial t} = \hat{H} \Psi, \quad (9.6.1)$$

it postulates that the wavefunction Ψ it describes is not a scalar complex function of time and coordinates, but a four-component column-vector (sometimes called the bispinor) of such functions, its Hermitian-conjugate bispinor Ψ^\dagger being a 4-component row-vector of their complex conjugates:

$$\Psi = \begin{pmatrix} \Psi_1(\mathbf{r}, t) \\ \Psi_2(\mathbf{r}, t) \\ \Psi_3(\mathbf{r}, t) \\ \Psi_4(\mathbf{r}, t) \end{pmatrix}, \quad \Psi^\dagger = (\Psi_1^*(\mathbf{r}, t), \quad \Psi_2^*(\mathbf{r}, t), \quad \Psi_3^*(\mathbf{r}, t), \quad \Psi_4^*(\mathbf{r}, t)) \quad (9.6.2)$$

and that the Hamiltonian participating in Eq. (95) is a 4×4 matrix defined in the Hilbert space of bispinors Ψ . For a free particle, the postulated Hamiltonian looks amazingly simple: ⁵²

$$\hat{H} = c\hat{\alpha} \cdot \hat{\mathbf{p}} + \hat{\beta}mc^2. \quad (9.6.3)$$

where $\hat{\mathbf{p}} = -i\hbar\nabla$ is the same 3D vector operator of momentum as in the non-relativistic case, while the operators $\hat{\alpha}$ and $\hat{\beta}$ may be represented in the following shorthand 2×2 form:

$$\hat{\alpha} \equiv \begin{pmatrix} \hat{0} & \hat{\boldsymbol{\sigma}} \\ \hat{\boldsymbol{\sigma}} & \hat{0} \end{pmatrix}, \quad \hat{\beta} \equiv \begin{pmatrix} \hat{I} & \hat{0} \\ \hat{0} & -\hat{I} \end{pmatrix} \quad (9.6.4)$$

The operator $\hat{\alpha}$, composed of the Pauli vector operators $\hat{\boldsymbol{\sigma}}$, is also a vector in the usual 3D space, with each of its 3 Cartesian components being a 4×4 matrix. The particular form of the 2×2 matrices corresponding to the operators $\hat{\boldsymbol{\sigma}}$ and \hat{I} in Eq. (98a) depends on the basis selected for the spin state representation; for example, in the standard z -basis, in which the Cartesian components of $\hat{\boldsymbol{\sigma}}$ are represented by the Pauli matrices (4.105), the 4×4 matrix form of Eq. (98a) is

$$\alpha_x = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad \alpha_y = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}, \quad \alpha_z = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad (9.6.5)$$

$$\beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (9.98 \text{ b})$$

It is straightforward to use Eqs. (98) to verify that the matrices $\alpha_x, \alpha_y, \alpha_z$ and β satisfy the following relations:

$$\alpha_x^2 = \alpha_y^2 = \alpha_z^2 = \beta^2 = \mathbf{I},$$

$$\alpha_x \alpha_y + \alpha_y \alpha_x = \alpha_y \alpha_z + \alpha_z \alpha_y = \alpha_z \alpha_x + \alpha_x \alpha_z = \alpha_x \beta + \beta \alpha_x = \alpha_y \beta + \beta \alpha_y = \alpha_z \beta + \beta \alpha_z = 0,$$

i.e. anticommute.

Using these commutation relations, and acting essentially as in Sec. 1.4, it is straightforward to show that any solution to the Dirac equation obeys the probability conservation law, i.e. the continuity equation (1.52), with the probability density:

$$w = \Psi^\dagger \Psi \quad (9.6.6)$$

and the probability current,

$$\mathbf{j} = \Psi^\dagger c \hat{\alpha} \Psi \quad (9.6.7)$$

looking almost as in the non-relativistic wave mechanics - cf. Eqs. (1.22) and (1.47). Note, however, the Hermitian conjugation used in these formulas instead of the complex conjugation, to form the scalars w , j_x , j_y , and j_z from the 4-component state vectors (96).

This close similarity is extended to the fundamental, plane-wave solutions of the Dirac equations in free space. Indeed, plugging such solution, in the form

$$\Psi = \mathbf{u} e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} \equiv \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} \quad (9.6.8)$$

into Eqs. (95) and (97), we see that they are indeed satisfied, provided that a system of four coupled, linear algebraic equations for four complex c -number amplitudes $u_{1,2,3,4}$ is satisfied. The condition of its consistency yields the same dispersion relation (87), i.e. the same two-branch diagram shown in Fig. 6, as follows from the Klein-Gordon equation. The difference is that plugging each value of ω , given by Eq. (87), back into the system of the linear equations for four amplitudes u , we get two solutions for their vector $\mathbf{u} \equiv (u_1, u_2, u_3, u_4)$ for each of the two energy branches - see Fig. 6 again. In the standard z basis of spin operators, they may be represented as follows:

$$\text{for } E = E_+ > 0: \quad \mathbf{u}_{+\uparrow} = c_{+\uparrow} \begin{pmatrix} 1 \\ 0 \\ \frac{cp_z}{E_+ + mc^2} \\ \frac{cp_+}{E_+ + mc^2} \end{pmatrix}, \quad \mathbf{u}_{+\downarrow} = c_{+\downarrow} \begin{pmatrix} 0 \\ 1 \\ \frac{cp_-}{E_+ + mc^2} \\ \frac{-cp_z}{E_+ + mc^2} \end{pmatrix}, \quad (9.6.9)$$

$$\text{for } E = E_- < 0: \quad \mathbf{u}_{-\uparrow} = c_{-\uparrow} \begin{pmatrix} \frac{cp_z}{E_- - mc^2} \\ \frac{cp_+}{E_- - mc^2} \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{u}_{-\downarrow} = c_{-\downarrow} \begin{pmatrix} \frac{cp_-}{E_- - mc^2} \\ \frac{-cp_z}{E_- - mc^2} \\ 0 \\ 1 \end{pmatrix}, \quad (9.6.10)$$

where $p_{\pm} \equiv p_x \pm ip_y$, and c_{\pm} are normalization coefficients.

The simplest interpretation of these solutions is that Eq. (103), with the vectors \mathbf{u}_+ given by Eq. (104a), represents a spin-1/2 particle (say, an electron), while with the vectors \mathbf{u}_- given by Eq. (104b), it represents an antiparticle (a positron), and the two solutions for each particle, indexed with opposite arrows, correspond to two possible directions of the spin-1/2, $\sigma_z = \pm 1$, i.e. $S_z = \pm \hbar/2$. This interpretation is indeed solid in the non-relativistic limit, when two last components of the vector (104a), and two first components of the vector (104b) are negligibly small:

$$\mathbf{u}_{+\uparrow} \rightarrow \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{u}_{+\downarrow} \rightarrow \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{u}_{-\uparrow} \rightarrow \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{u}_{-\downarrow} \rightarrow \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \text{for } \frac{p_{x,y,z}}{mc} \rightarrow 0 \quad (9.6.11)$$

However, at arbitrary energies, the physical picture is more complex. To show this, let us use the Dirac equation to calculate the Heisenberg-picture law of time evolution of the operator of some Cartesian component of the orbital angular momentum $\mathbf{L} \equiv \mathbf{r} \times \mathbf{p}$, for example of $L_x = yp_z - zp_y$, taking into account that the Dirac operators (98a) commute with those of \mathbf{r} and \mathbf{p} , and also the Heisenberg commutation relations (2.14):

$$i\hbar \frac{\partial \hat{L}_x}{\partial t} = [\hat{L}_x, \hat{H}] = c\hat{\boldsymbol{\alpha}} \cdot [(\hat{y}\hat{p}_z - \hat{z}\hat{p}_y), \hat{\mathbf{p}}] = -i\hbar c (\hat{\alpha}_z \hat{p}_y - \hat{\alpha}_y \hat{p}_z) \quad (9.6.12)$$

with similar relations for two other Cartesian components. Since the right-hand side of these equations is different from zero, the orbital momentum is generally not conserved - even for a free particle! Let us, however, consider the following vector operator,

$$\hat{\mathbf{S}} \equiv \frac{\hbar}{2} \begin{pmatrix} \hat{\sigma} & \hat{0} \\ \hat{0} & \hat{\sigma} \end{pmatrix}. \quad (9.6.13)$$

According to Eqs. (4.105), its Cartesian components, in the z -basis, are represented by 4×4 matrices

$$S_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad S_y = \frac{\hbar}{2} \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix}, \quad S_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (9.6.14)$$

Let us calculate the Heisenberg-picture law of time evolution of these components, for example

$$i\hbar \frac{\partial \hat{S}_x}{\partial t} = [\hat{S}_x, \hat{H}] = c [\hat{S}_x, (\hat{\alpha}_x \hat{p}_x + \hat{\alpha}_y \hat{p}_y + \hat{\alpha}_z \hat{p}_z)] \quad (9.6.15)$$

A direct calculation of the commutators of the matrices (98) and (107) yields

$$[\hat{S}_x, \hat{\alpha}_x] = 0, \quad [\hat{S}_x, \hat{\alpha}_y] = i\hbar \hat{\alpha}_z, \quad [\hat{S}_x, \hat{\alpha}_z] = -i\hbar \hat{\alpha}_y \quad (9.6.16)$$

so that we finally get

$$i\hbar \frac{\partial \hat{S}_x}{\partial t} = i\hbar c (\hat{\alpha}_z \hat{p}_y - \hat{\alpha}_y \hat{p}_z), \quad (9.6.17)$$

with similar expressions for the other two components of the operator. Comparing this result with Eq. (106), we see that any Cartesian component of the operator defined similarly to Eq. (5.170),

$$\hat{\mathbf{J}} \equiv \hat{\mathbf{L}} + \hat{\mathbf{S}} \quad (9.6.18)$$

is an integral of motion,⁵³ so that this operator may be interpreted as the one representing the total angular momentum of the particle. Hence, the operator (107) may be interpreted as the spin operator of a spin-1/2 particle (e.g., electron). As it follows from the last of Eq. (107b), in the non-relativistic limit the columns (105) represent the eigenkets of the z -component of that operator, with eigenstates $S_z = \pm\hbar/2$, with the sign corresponding to on the arrow index. So, the Dirac theory provides a justification for spin-1/2 - or, somewhat more humbly, replaces the Pauli Hamiltonian postulate (4.163) with that of a simpler (and hence more plausible), Lorentz-invariant Hamiltonian (97).

Note, however, that this simple interpretation, fully separating a particle from its antiparticle, is not valid for the exact solutions (103)-(104), so that generally the eigenstates of the Dirac Hamiltonian are certain linear (coherent) superpositions of the components describing the particle and its antiparticle - each with both directions of spin. This fact leads to several interesting effects, including the so-called Klien paradox at the reflection of a relativistic electron from a potential barrier.⁵⁴

⁵¹ After the "naturally-relativistic" form of the Klein-Gordon equation (84), this apparent return to the nonrelativistic Schrödinger equation may look very counter-intuitive. However, it becomes a bit less surprising taking into account the fact (whose proof is left for the reader's exercise) that Eq. (84) may be also recast into the form (95) for a two-component column-vector Ψ (sometimes called spinor), with a Hamiltonian which may be represented by a 2×2 matrix - and hence expressed via the Pauli matrices (4.105) and the identity matrix I .

⁵² Moreover, if the time derivative participating in Eq. (95), and the three coordinate derivatives participating (via the momentum operator) in Eq. (97), are merged into one 4-vector operator $\partial/\partial x_k \equiv \{\nabla, \partial/\partial(ct)\}$, the Dirac equation (95) may be rewritten in an even simpler, manifestly Lorentz-invariant 4-vector form (with the implied summation over the repeated index $k = 1, \dots, 4$ - see, e.g., EM Sec. 9.4):

$$\left(\hat{\gamma}_k \frac{\partial}{\partial x_k} + \mu \right) \Psi = 0, \quad \text{where } \hat{\gamma} \equiv \{\hat{\gamma}_1, \hat{\gamma}_2, \hat{\gamma}_3\} = \begin{pmatrix} 0 & -i\hat{\sigma} \\ i\hat{\sigma} & 0 \end{pmatrix}, \quad \hat{\gamma}_4 = \hat{\beta} \quad (9.6.19)$$

where $\mu \equiv mc/\hbar$ - just as in Eq. (84). Note also that, very counter-intuitively, the Dirac Hamiltonian (97) is linear in the momentum, while the non-relativistic Hamiltonian of a particle, as well as the relativistic Schrödinger equation, are quadratic in p . In my humble opinion, the Dirac theory (including the concept of antiparticles it has inspired) may compete for the title of the most revolutionary theoretical idea in physics of all times, despite such strong contenders as Newton's laws, Maxwell's equations, Gibbs' statistical distribution, Bohr's theory of the hydrogen atom, and Einstein's general relativity.

⁵³ It is straightforward to show that this result remains valid for a particle in any central field $U(r)$.

⁵⁴ See, e.g., A. Calogeracos and N. Dombey, Contemp. Phys. 40, 313 (1999).

This page titled [9.6: Dirac's Theory](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Konstantin K. Likharev](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.