Chapter 2

Physical contents of Dirac equation: preliminary discussion

As we have noted in the preceding chapter, the prime motivation for finding an alternative to the Klein–Gordon equation was the requirement that the probability defined in terms of a quantum mechanical wave function should be positive. So, let us now examine this problem for the Dirac equation; for convenience, we return to the natural units. Eq. (1.23) then reads

$$i\,\frac{\partial\psi}{\partial t} = -i\vec{\alpha}\cdot\vec{\nabla}\psi + \beta m\psi \tag{2.1}$$

(we will use the standard representation (1.32) in what follows). Let us recall that ψ is a four-component wave function that is conventionally written as a column

$$\psi(x) = \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \\ \psi_3(x) \\ \psi_4(x) \end{pmatrix}. \tag{2.2}$$

Upon Hermitian conjugation of Eq. (2.1) one has

$$-i\,\frac{\partial\psi^\dagger}{\partial t} = i\,\vec{\nabla}\psi^\dagger\vec{\alpha} + m\psi^\dagger\beta\,, \tag{2.3}$$

where $\psi^\dagger = (\psi_1^*, \psi_2^*, \psi_3^*, \psi_4^*)$, and we have utilized the hermiticity property (1.27) of $\vec{\alpha}$ and β. Multiplying Eq. (2.1) by $\psi^\dagger$ from the left and (2.3) by ψ from the right, and taking then the difference of the two equations, one gets immediately

$$\frac{\partial}{\partial t}(\psi^\dagger\psi) + \vec{\nabla}(\psi^\dagger\vec{\alpha}\psi) = 0\,, \tag{2.4}$$

which is the anticipated continuity equation. Thus we may identify the probability density and the probability current as

$$\rho_{\text{Dirac}} = \psi^\dagger\psi\,, \qquad \vec{j}_{\text{Dirac}} = \psi^\dagger\vec{\alpha}\psi\,. \tag{2.5}$$

The positivity of the ρ_{Dirac} is obvious, since

$$\psi^\dagger\psi = |\psi_1|^2 + |\psi_2|^2 + |\psi_3|^2 + |\psi_4|^2\,. \tag{2.6}$$

This is an expected result, due to the fact that the Dirac equation (2.1) is, in a sense, "square root of Klein–Gordon equation"; more precisely, it is an evolution equation of the 1st order in time, having the form

$$i\,\frac{\partial\psi}{\partial t} = H\,\psi\,, \tag{2.7}$$

where H is the Dirac Hamiltonian

$$H = -i\vec{\alpha} \cdot \vec{\nabla} + \beta m \,. \tag{2.8}$$

Thus, the time evolution is generated by an operator of energy, as it should be, in accordance with the general principles of quantum theory.

A next issue is the angular momentum. Let us start with orbital angular momentum, defined in the standard way as $\vec{L} = \vec{x} \times \vec{p}$, where $\vec{p}$ is the (linear) momentum $\vec{p} = -i\vec{\nabla}$. As we know, $\vec{L}$ commutes with the non-relativistic Hamiltonian in the Schrödinger equation (1.4). For the Dirac Hamiltonian (2.8) one gets, employing the canonical commutation relation $[x^j, p^k] = i\delta^{jk}$,

$$[H, \vec{L}] = -i(\vec{\alpha} \times \vec{p}) \,. \tag{2.9}$$

Let us remark that the vector product in (2.9) is defined formally as usual, i.e.

$$(\vec{\alpha} \times \vec{p})^j = \varepsilon^{jkl} \alpha^k p^l \,.$$

So, apparently, there is something missing, since any decent angular momentum should be an integral of motion for the free particle, i.e. the corresponding operator should commute with the Hamiltonian. In other words, the fact that $[H, \vec{L}] \neq 0$ is a hint that we are on the right track towards the electron spin. A good candidate for such an additional ingredient of the full angular momentum is guessed quite easily. Let us consider the 4×4 matrices

$$\vec{S} = \frac{1}{2}\vec{\Sigma} \,, \qquad \vec{\Sigma} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} , \tag{2.10}$$

and recall that the Pauli matrices have the commutation relations

$$[\sigma_j, \sigma_k] = 2i\varepsilon_{jkl}\sigma_l \,. \tag{2.11}$$

This means that the matrices $\vec{S}$ defined by (2.10) satisfy

$$[S_j, S_k] = i\varepsilon_{jkl} S_l \,, \tag{2.12}$$

which, of course, is a set of commutation relations for components of an angular momentum. Needless to say, the matrices $\vec{S}$ possess eigenvalues $\pm 1/2$ (because $(\sigma_j)^2 = 1$ for $j = 1, 2, 3$). Now we may evaluate the commutator $[H, \vec{S}]$. Clearly, $\vec{S}$ commutes with the diagonal matrix β (see (1.32)). Concerning the commutator involving $\vec{\alpha}$, one gets first

$$[\alpha^j, \Sigma^k] = \begin{pmatrix} 0 & 2i\varepsilon^{jkl}\sigma^l \\ 2i\varepsilon^{jkl}\sigma^l & 0 \end{pmatrix} ,$$

so that

$$[H, \Sigma^k] = 2i(\vec{\alpha} \times \vec{p})^k \,. \tag{2.13}$$

Summarizing the results of our simple algebraic exercise, we have

$$\begin{aligned} [H, \vec{L}] &= -i(\vec{\alpha} \times \vec{p}) \,, \\ [H, \vec{S}] &= i(\vec{\alpha} \times \vec{p}) \,, \end{aligned} \tag{2.14}$$

and thus

$$[H, \vec{J}] = 0 \,, \tag{2.15}$$

with

$$\vec{J} = \vec{L} + \vec{S}\,. \tag{2.16}$$

Thus, in such a straightforward manner we have recovered the electron spin as a part of the conserved total angular momentum (2.16).

Let us now recall the problem of negative energy solutions of the Klein–Gordon equation, mentioned in the preceding chapter (cf. (1.12)). One may wonder whether the Dirac equation suffers an analogous difficulty. For clarifying this point, we are going to consider the solution of Eq. (2.1) in the form of a plane wave involving the usual factor $\exp\left[-i(Et - \vec{p}\cdot\vec{x})\right]$. To make our discussion as simple as possible, we will restrict ourselves to the case of a particle at rest, i.e. set $\vec{p} = 0$. Eq. (2.1) is then reduced to

$$i\,\frac{\partial\psi}{\partial t} = \beta m\psi\,. \tag{2.17}$$

Taking into account the block diagonal structure of the matrix β ((1.32), it is useful to split the ψ as

$$\psi = \begin{pmatrix}\varphi\\ \chi\end{pmatrix}, \tag{2.18}$$

where φ and χ are two-component column vectors. Eq. (2.17) is then recast as

$$i\,\frac{\partial\varphi}{\partial t} = m\varphi\,, \tag{2.19}$$

$$i\,\frac{\partial\chi}{\partial t} = -m\chi\,. \tag{2.20}$$

Thus, two linearly independent solutions of Eq. (2.19) may be written e.g. as

$$\varphi_{(1)} = e^{-imt}\begin{pmatrix}1\\0\end{pmatrix}, \qquad \varphi_{(2)} = e^{-imt}\begin{pmatrix}0\\1\end{pmatrix}, \tag{2.21}$$

and similarly for (2.20),

$$\chi_{(1)} = e^{imt}\begin{pmatrix}1\\0\end{pmatrix}, \qquad \chi_{(2)} = e^{imt}\begin{pmatrix}0\\1\end{pmatrix}. \tag{2.22}$$

In this way, we obtain a set of four independent solutions of Eq. (2.1)

$$\psi_{(1)} = e^{-imt}\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \quad \psi_{(2)} = e^{-imt}\begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \quad \psi_{(3)} = e^{imt}\begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \quad \psi_{(4)} = e^{imt}\begin{pmatrix}0\\0\\0\\1\end{pmatrix}. \tag{2.23}$$

Obviously, $\psi_{(1)}$ and $\psi_{(2)}$ correspond to the positive rest energy $E = m$, while $\psi_{(3)}$ and $\psi_{(4)}$ carry negative energy $E = -m$ (they are also characterized by the two possible spin projections to the third axis, up and down $(\pm 1/2)$). It is interesting to notice that in the considered case, the existence of the negative energy solutions is a consequence of the specific structure of the matrix β. If β were 4×4 unit matrix, we would have only a solution with positive energy. But, alas, β can never be the unit matrix because of the required anticommutation relations (1.26). As we have already noted in the preceding chapter, the appearance of negative energy solutions is a generic feature of the equations of relativistic quantum mechanics. We will discuss the plane-wave solutions of Dirac equation in detail later on.

The last topic that we are going to discuss here is a derivation of the spin magnetic moment of the electron. Soon after the birth of relativistic quantum mechanics this was indeed