# THE DIRAC EQUATION AND ITS INTERPRETATIONS[1]

Mário Bacelar Valente[2]

Abstract

In this paper, it is presented a historical account of the formulation of the quantum relativistic wave equation of an electron – the Dirac equation, issues regarding its interpretation that arose from the very beginning, and the later formulation of this equation in relation to a quantized electron-positron field, which implies a new interpretation. The way in which solutions obtained under each interpretation of the equation relate to one another is also considered for the simple case of hydrogen-like problems.

## 1. Introduction

Feynman once made the following comments, having in mind, in particular, the Dirac equation:

We know so very much and then subsume it into so very few equations that we can say we know very little (except these equations – e.g. Dirac, Maxwell, Schrod.). Then we think we have the physical picture with which to interpret the equations. But these are so very few equations that I have found that many physical pictures can give the same equations.[3]

In this paper, adopting a historical account, we will try to elaborate some 'pictures', or, more exactly, address the interpretations of the Dirac equation and consider issues regarding these interpretations taking into account applications to simple problems. In this way, we will try to get a glimpse of the 'so very much' that is condensed in so very little.

The meaning of the Dirac equation is not as simple as we might think. Since its first formulation, its meaning has changed from a relativistic wave equation for an electron to a classical field equation from which an electron-positron quantum field is derived – the Dirac field; in the process it went from being a relativistic 'update' of the Schrödinger equation in the calculation of energy levels in atoms (basically of hydrogen) to become one of the cornerstones of the most successful quantum field theory: quantum electrodynamics.

In sections 2 we make some remarks regarding previous efforts, before Dirac, to arrive at a relativistic wave equation. Section 3 presents the Dirac equation in its first formulation as a one-electron relativistic wave equation. Section 4 addresses Dirac's effort to deal with issues regarding this interpretation of his equation that led him to develop the so-called hole theory. In section 5 we address the field theoretical interpretation of Dirac's equation. This is the second formulation of the Dirac equation.

---

[1] This is a streamlined and slightly corrected version of earlier works.

[2] University Pablo de Olavide, Seville, Spain; email: mar.bacelar@gmail.com.

[3] Quoted in Silvan Schweber, *QED and the men who made It: Dyson, Feynman, Schwinger, and Tomonaga* (Princeton: Princeton University Press, 1994), on 407.

In this interpretation, the Dirac equation is related to the quantized Dirac field – its physics is, so to speak, expressed in the quantum field. Finally, in section 6 we will consider, in very simple terms, solutions obtained in one interpretation from the perspective of the other. Our example will be hydrogen-like problems. What kind of solution obtained with the one-electron interpretation of the Dirac equation fits with the quantum field interpretation? And what kind of quantum field solution 'corresponds' to this one-electron solution? There must be some rationale since we are working with the same equation even if in different settings.

## 2. Before the Dirac equation: some historical remarks

Schrödinger's first attempt at a wave equation was the development of a relativistic wave equation for the hydrogen atom. As in the case of the non-relativistic wave equation, Schrödinger considered the problem of determining the eigenvibrations of the hydrogen atom. Schrödinger took the de Broglie relations generalizing them to the case of an electron in a central Coulomb potential and by inserting them in a second-order differential equation obtained the relativistic wave equation

$$\Delta\psi = -\frac{4\pi^2 m^2 c^2}{h^2}\left[\left(\frac{h\nu}{mc^2}+\frac{e^2}{mc^2 r}\right)-1\right]\psi = 0\,.$$

Using this wave equation, Schrödinger determined the energy levels for the hydrogen atom and arrived at a result that was not in agreement with Sommerfeld's result for the hydrogen spectrum obtained, within the so-called old quantum theory, through the quantization of the relativistic Bohr atom. Sommerfeld's result was

$$E_{n,k} = \frac{mc^2}{\sqrt{1+\dfrac{\alpha^2}{\left(n-k-\sqrt{k^2-\alpha^2}\right)^2}}} - mc^2,$$

where $\alpha$ is the fine structure constant, n the principal quantum number, and k the azimuthal quantum number. Schrödinger obtained an expression that did not depend on n – k and k but on n – k + 1/2 and k – 1/2.[4] This meant the failure of Schrödinger's relativistic wave mechanics. Schrödinger set aside his attempt at a relativistic wavefunction and developed a non-relativistic wave equation.[5] The relativistic wave equation, later known by the name of Klein-Gordon equation, was presented by several

[4] Helge Kragh, "The genesis of Dirac's relativistic theory of electrons," *Archive for History of Exact Sciences* 24 (1981): 31-67, on 33.

[5] Erwin Schrödinger, "Quantisation as a problem of proper values—I," *Collected papers on wave mechanics* (Providence: Chelsea Publishing Company, 1982), on 1-12. Original work published in *Annalen der Physik* 79 (1926).

physicists during 1926, but since it did not give the fine structure of the hydrogen spectrum it was not accepted as the correct relativistic equation for an electron.[6]

Another factor that would complicate matters in what regards the wave mechanics description of the electrons was the discovery of spin. In Bohr's theory, each spectral term of the hydrogen atom is labeled by three quantum numbers n, k, and m. Due to an external magnetic field, a spectral term labeled by n and k splits into 2k + 1 levels, the new sub-levels being distinguished by the quantum number m. This is called the Zeeman effect.[7] Before 1900 it was already known that the spectral lines, later described by n and k, were not unique and in reality, when an external magnetic field was applied, they consisted of closely spaced lines: they showed a multiplet structure. This was called the anomalous Zeeman effect.[8] In 1920 Sommerfeld introduced a new quantum number j that enabled the classification of the different energy levels within one multiplet term (characterized by n and k). Under this new classification m is still related to the specification of sublevels but now of a level specified by n, k and j. Also, m must satisfy the inequality $-j \leq m \leq j$. To explain the spectroscopic evidence available Landé (and also Sommerfeld) set forward a tentative model in which it was assumed that the core of the atom had an angular momentum. There would then be a magnetic moment $\mu_K$ related to the orbital angular momentum **K** of the electron in the outermost orbit (the valence electron), given by $\mu_K = -\mathbf{K}$ (in units of the Bohr magneton). Also, there would be a magnetic moment $\mu_R$ associated with the core angular momentum **R** (corresponding to a quantum number related not to the electron but to the core). The relation between $\mu_R$ and **R** is given by $\mu_R = -g_0\mathbf{R}$, where $g_0$ has to be determined by fitting the model to experimental results. Pachen and Back's study of the Zeeman effect in the case of a strong magnetic field enabled us to set the value of $g_0$ as 2. The interaction between the two magnetic moments $\mu_K$ and $\mu_R$ of the atom leads to a slight energy change in the atomic energy levels, which results in the multiplet structure of the spectral lines.[9]

In 1924 Pauli showed that the association of $\mu_R$ with the atom's core was inconsistent and considered it to be associated with the valence electron. In this way, the four quantum numbers are all related to the electrons. Upon reading Pauli's ideas Uhlenbeck and Goudsmit proposed in 1925 to reinterpret the core angular momentum **R** as an intrinsic angular momentum of the electron. Uhlenbeck and Goudsmit had the idea that each quantum number should correspond to a degree of freedom of the electron. This led them to the idea of an intrinsic rotation of the electron, the spin, as a fourth degree of freedom to which a quantum number would be associated. As in the case of Landé's model, Uhlenbeck and Goudsmit got a discrepancy, by a factor of 2, between the theoretical and the experimental results in the case of doublet levels of alkali atoms. In 1926 Thomas presented a relativistic calculation where the missing 1/2 factor, later called the Thomas factor, was obtained.[10] So, by 1926 there was no relativistic wave equation for the electron and there was the further complication of having to account also for the electron's spin.

---

[6] Kragh, "Equation with many fathers. The Klein-Gordon equation in 1926," *American Journal of Physics* 52 (1984): 1024-1033.
[7] Sin-Itiro Tomonaga, *The story of spin* (Chicago: The University of Chicago Press, 1997), on 1-2.
[8] Max Jammer, *The conceptual development of quantum mechanics* (New York: McGraw-Hill, 1966), on 122.
[9] Tomonaga, 11-20.
[10] Jammer, 149-152

In 1927 Pauli attempted to incorporate spin into wave mechanics by considering a Schrödinger wave function depending also on a degree of freedom related to spin.[11] The intrinsic (spin) angular momentum in any direction takes only the two values ± ħ/2. This made Pauli consider a two-component wave function, with one component corresponding to an electron's spin-up $\psi(\mathbf{x}, +1/2)$ and another to an electron's spin-down $\psi(\mathbf{x}, -1/2)$. This two-component wave function must be the solution of two coupled equations with the form

$$\mathrm{H}\left(\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial \mathrm{q}}, \mathrm{s}\right)\psi = \mathrm{E}\psi\,.$$

The question was how to define in the wave equation the spin operators **s**. Pauli defined the spin operators as $s_x = 1/2\ \sigma_x$, $s_y = 1/2\ \sigma_y$, $s_z = 1/2\ \sigma_z$, where $\sigma_x, \sigma_y, \sigma_z$ are the so-called Pauli matrices:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\ \sigma_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix},\ \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

In the Hamiltonian for his equation, Pauli considered two new terms, besides the term corresponding to an electron (without spin) in a central potential, which was already present in Schrödinger's equation. One of these terms resulted from the interaction of an external magnetic field and the valence electron, which, as we have seen, possesses besides an orbital angular momentum an intrinsic spin momentum. The other resulted from the interaction between the spin magnetic moment and the central potential due to the orbital motion of the electron: the spin-orbit coupling. In this approach, Pauli considered only first-order relativistic corrections and the way the spin operators (and $g_0$ and the Thomas factor) were put into the Hamiltonian was arbitrary. Most importantly Pauli was unable to extend this approach into a fully relativistic form.

## 3. The one-electron interpretation of the Dirac equation

Things changed by the end of 1927 when Dirac was able to formulate a relativistic wave equation. In his first attempts towards a relativistic theory, Dirac consider a Klein-Gordon type equation written in terms of a relativistic Hamiltonian:[12]

$$\{\mathrm{p}_x^2 + \mathrm{p}_y^2 + \mathrm{p}_z^2 - \mathrm{E}/\mathrm{c}^2 + \mathrm{m}^2\mathrm{c}^2\}\psi = 0\,,$$

$$\left\{\frac{\partial^2}{\partial \mathrm{x}^2} + \frac{\partial^2}{\partial \mathrm{y}^2} + \frac{\partial^2}{\partial \mathrm{z}^2} - \frac{1}{\mathrm{c}^2}\frac{\partial^2}{\partial \mathrm{t}^2} - \frac{\mathrm{m}^2\mathrm{c}^2}{\mathrm{h}^2}\right\}\psi = 0\,.$$

Upon reading Dirac's articles using this equation, Ehrenfest asked Dirac in a letter on the motive for using a particular form for the Hamiltonian:

---

[11] Kragh, "*The genesis of Dirac's relativistic theory of electrons*," on 45-46

[12] Paul Dirac, "On the theory of quantum mechanics," *Proceedings of the Royal Society of London A* 112 (1926): 661-677.

Why do you write the Hamilton equation in the form:

$$E^2/c^2 - p_1^2 - p_2^2 - p_3^2 = m^2c^2$$

and not:

$$mc^2\sqrt{1-(p_1^2+p_2^2+p_3^2)/m^2c^2} = E$$

Does it make a difference?[13]

By that time also Pauli was proposing to adopt instead of a second-order equation a first-order equation involving a square-root.[14]

Dirac felt that neither form of the relativistic Hamiltonian was appropriate for the development of a relativistic wave equation. Dirac considered that this equation should maintain the formal structure of the Schrödinger equation.[15] On one side Dirac knew he needed an equation that was linear in the time derivative so that he could maintain in the relativistic case the statistical interpretation of the wave function adopted in the non-relativistic case.[16] On the other side, this meant, due to relativistic considerations, that the equation should be linear also in the spatial derivatives. According to Dirac "an appropriate formulation of quantum mechanics will only be possible when we succeed in treating space and time as equal to one another".[17] This means that space and time must appear in the equation as the coordinates of a Minkowski space-time.

It seems that what resulted fundamental to Dirac's development of his relativistic equation was Dirac's realization of the identity:

$$|\vec{p}| = \sqrt{p_1^2 + p_2^2 + p_3^2} = \sigma_1 p_1 + \sigma_2 p_2 + \sigma_3 p_3,$$

where $\sigma_1$, $\sigma_2$, $\sigma_3$ are the Pauli matrices. According to Dirac:

I was playing around with the three components $\sigma_1$, $\sigma_2$, $\sigma_3$, which I had used to describe the spin of an electron, and I noticed that if you formed the expression $\sigma_1 p_1 + \sigma_2 p_2 + \sigma_3 p_3$ and squared it, $p_1$, $p_2$ and $p_3$ being the three components of momentum, you got $p_1^2 + p_2^2 + p_3^2$, the square of the momentum. This was a pretty mathematical result. I was quite excited over it. It seemed that it must be of some importance.[18]

This mathematical identity was the insight that made Dirac search for a relativistic counterpart involving a term corresponding to the electron's rest mass. The problem facing Dirac was that with the Pauli matrices it was not possible to write down an expression with four squares:

---

[13] Kragh, *Dirac: a scientific biography* (Cambridge: Cambridge University Press, 1990), on 53.
[14] Jagdish Mehra and Helmut Rechenberg, *The historical development of quantum theory, Vol. 6, Part 1* (Berlin: Springer-Verlag, 2000), on 293.
[15] Kragh, "*Dirac: a scientific biography*," on 54.
[16] Mehra and Rechenberg, 294 ; Kragh, "*Dirac: a scientific biography*," on 64.
[17] Quoted in Kragh, "*Dirac: a scientific biography*," on 54.
[18] Quoted in Mehra and Rechenberg, 295.

$$|\vec{p}| = \sqrt{p_1^2 + p_2^2 + p_3^2 + m^2c^2} = \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \alpha_4 mc\,.$$

Dirac considered that from the wave equation developed using the previous expression $\{p_0 - (m^2c^2 + p_1^2 + p_2^2 + p_3^2)^{1/2}\}\psi = 0$ one should be able to recover the equation $\{p_0^2 - m^2c^2 - p_1^2 - p_2^2 - p_3^2\}\psi = 0$, "which is of a relativistically invariant form".[19] That is, Dirac expected his relativistic equation to contain the Klein-Gordon equation as its square since this equation involves the relativistic Hamiltonian in its normal invariant form. This implied a set of relations for the unknown coefficients:

$$\alpha_\mu\alpha_\nu + \alpha_\nu\alpha_\mu = 0 \ (\mu \neq \nu);\ \mu, \nu = 1, 2, 3, 4,$$

$$\alpha_\mu^2 = 1.$$

There is no set of four 2 × 2 matrices that satisfies the previous conditions. According to his recollections, at some point, Dirac realized that:

There was no need to stick to quantities, which can be represented by matrices with just two rows and columns. Why not go to four rows and columns? Mathematically there was no objection to this at all. Replacing the σ-matrices by four-row and column matrices, one could easily take the square root of the sum of four squares, or even five squares if one wanted to.[20]

With this insight, Dirac arrived at his relativistic wave equation. Dirac choose a representation where the coefficients are given by:

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

$$\rho_1 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \qquad \rho_2 = \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -i \\ i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{pmatrix} \qquad \rho_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$

and $\alpha_1 = \rho_1\sigma_1$, $\alpha_2 = \rho_1\sigma_2$, $\alpha_3 = \rho_1\sigma_3$, $\alpha_4 = \sigma_3$.

In this way, Dirac had immediately the relativistic wave equation for a free electron:

$$[p_0 - \rho_1(\boldsymbol{\sigma}, \mathbf{p}) - \rho_3 mc]\psi = 0\,,$$

[19] Dirac, *The principles of quantum mechanics* (Oxford: Oxford University Press, 1958), on 255.
[20] Quoted in Mehra and Rechenberg, 295.

where $p_0 = i\hbar\,\partial/(c\partial t)$ and $\mathbf{p} = (p_1, p_2, p_3)$, where $p_r = -i\hbar\,\partial/(c\partial x_r)$ with r = 1, 2, 3; $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ is a vector formed with the above $4 \times 4$ matrices.

Dirac generalized his equation to the case of an electron in an external electromagnetic field. Dirac followed the rule of replacing $p_0$ by $p_0 + e/c.A_0$ and $\mathbf{p}$ by $\mathbf{p} + e/c.\mathbf{A}$ (where $A_0$ and $\mathbf{A}$ are the scalar and vector potentials). This gives us the equation

$$[p_0 + \frac{e}{c}A_0 - \rho_1(\boldsymbol{\sigma}, \mathbf{p} + \frac{e}{c}\mathbf{A}) - \rho_3 mc]\psi = 0,$$

which is the "fundamental wave equation of the relativistic theory of the electron".[21]

Dirac had developed his equation by considering the relativistic Hamiltonian of a free point particle; that is, Dirac did not take into account in his Hamiltonian (as, for example, Pauli did in the derivation of his equation) any term related to the spin of the electron. It was a surprise to Dirac that "the simplest possible case did involve the spin".[22]

Dirac set out to explore the relationship between his wave equation with external potentials and the Klein-Gordon equation (based on a classical relativistic Hamiltonian), which was according to Dirac "the wave equation to be expected from analogy with the classical theory":[23]

$$\left\{\left(p_0 + \frac{e}{c}A_0\right)^2 - \left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right)^2 - m^2c^2\right\}\psi = 0.$$

By squaring his equation, Dirac obtained a differential equation that included the operator of the Klein-Gordon equation and two additional terms:

$$\left\{\left(p_0 + \frac{e}{c}A_0\right)^2 - \left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right)^2 - m^2c^2 - \frac{\hbar e}{c}(\boldsymbol{\sigma}, \mathbf{H}) + i\rho_1\frac{\hbar e}{c}(\boldsymbol{\sigma}, \mathbf{E})\right\}\psi = 0.$$

Dirac concluded that:

> The electron will therefore behave as though it has a magnetic moment $eh/4\pi mc \cdot \boldsymbol{\sigma}$ and an electric moment $ieh/4\pi mc \cdot \rho_1\boldsymbol{\sigma}$. This magnetic moment is just that assumed by the spinning electron model. The electric moment, being a pure imaginary, we should not expect to appear in the model. It is doubtful whether the electric moment has any physical meaning, since the Hamiltonian […] that we started from is real, and the imaginary part only appeared when we multiplied it up in an artificial way to make it resemble the Hamiltonian of previous theories.[24]

Dirac showed how this internal magnetic moment resulted from the electron having a spin angular momentum. According to Dirac:

---

[21] Dirac, "*The principles of quantum mechanics*," on 257.
[22] Quoted in Kragh, "*The genesis of Dirac's relativistic theory of electrons*," on 55.
[23] Dirac, "*The principles of quantum mechanics*," on 264.
[24] Dirac, "The quantum theory of the electron," *Proceedings of the Royal Society of London A* 117 (1928): 610-624, on 619.

The spin angular momentum does not give rise to any potential energy and therefore does not appear in the result of the preceding calculation. The simplest way of showing the existence of the spin angular momentum is to take the case of the motion of a free electron or an electron in a central field of force and determine the angular momentum integrals [of the motion].[25]

By setting $\mathbf{A} = 0$ and $A_0$ a function of the radius r, the Hamiltonian in Dirac's equation takes the form $H = -eA_0(r) + c\rho_1(\boldsymbol{\sigma}, \mathbf{p}) + \rho_3 mc^2$. Considering, for example, the equation of motion of the $x_1$-component of the orbital angular momentum $m_1 = x_2p_3 - x_3p_2$, we have $i\hbar\dot{m}_1 = i\hbar c\rho_1\{\sigma_2p_3 - \sigma_3p_2\}$. We see that $m_1$ is not a constant of the motion. By considering also the equation $i\hbar\dot{\sigma}_1 = 2ic\rho_1\{\sigma_3p_2 - \sigma_2p_3\}$, Dirac found that $\dot{m}_1 + \hbar/2 \cdot \dot{\sigma}_1 = 0$. This means that the vector $\mathbf{m} + \hbar/2 \cdot \boldsymbol{\sigma}$ is a constant of the motion. According to Dirac:

We can interpret this result by saying that the electron has a spin angular momentum of $\hbar/2 \cdot \boldsymbol{\sigma}$, which, added to the orbital angular momentum $\mathbf{m}$, gives the total angular momentum $\mathbf{M}$, which is a constant of the motion.[26]

In this way, Dirac found that his relativistic wave equation described an electron with a spin angular momentum (and corresponding magnetic moment).

Dirac did not provide an exact solution of his equation for an electron in a central potential. He only made a first-order calculation. In this treatment of the hydrogen atom, Dirac was able to obtain Pauli's results on the energy levels, but without using any arbitrary assumptions: the spin angular momentum, gyromagnetic ratio ($g_0$) and Thomas factor all came out right. Just a few weeks after the publication of Dirac's paper on the relativistic wave equation, it was shown independently by Darwin and Gordon that an exact solution of Dirac's equation gave an expression for the discrete energy levels of the hydrogen spectrum which was identical to Sommerfeld's original formula derived in 1915, which was in good agreement with experimental results.[27]

## 4. Problems with the one-electron interpretation: the negative energy solutions

Since Dirac was using $4 \times 4$ matrices in his equation, the wave function had four components (recall that Schrödinger's original equation had one component, and Pauli's equation had two components due to the spin degree of freedom). Initially, Dirac thought that he could simply drop two of the components, since "half of the solutions must be rejected as referring to the charge +e on the electron".[28] This was possible in a first-order approximation. Looking at the exact solution we can see that the situation is far from that simple. Dirac's equation can be written as a set of two coupled differential equations for a pair of two-component wave functions $\psi_A$ and $\psi_B$, where Dirac's wave function is given by

$$\psi = \begin{pmatrix} \psi_A \\ \psi_B \end{pmatrix}.$$

[25] Dirac, "*The principles of quantum mechanics*," on 266.
[26] Dirac, "*The quantum theory of the electron*," on 620.
[27] Mehra and Rechenberg, 300.
[28] Dirac, "*The quantum theory of the electron*," on 618.

These two wave functions are essential in the solution of Dirac's equation and cannot be disregarded in the calculation of the energy levels. What happens is that the lower two-components ($\psi_B$) are smaller than the upper two-components ($\psi_A$), roughly by a ratio of $\upsilon/2c$, where $\upsilon$ is the 'velocity' of the orbiting electron in Bohr's theory. Now, when considering the solution of Dirac's equation for a free electron with momentum **p**, we see that there are two solutions corresponding to electron states with momentum **p** and energy $E_p$, and two solutions corresponding to states with momentum –**p** and energy $-E_p$, or, as Dirac mentioned, to an electron with charge +e (and positive energy $E_p$). In the case of the exact solution for an electron in an external electromagnetic field, as we have seen, we have a four-component wave function, which we can say, by resort to the free electron case, to 'have' positive and negative energy components, or, as Dirac mentions, solutions referring to a charge –e and +e. In this way, we must take, as Dirac did, the reference to positive and negative energy components as a "rough one, applying to the case when such a separation is approximately possible".[29] When taking the non-relativistic limit of Dirac's equation, the equation for the upper two-components takes the form of the Schrödinger-Pauli two-component wave equation (corresponding to a negative energy state), and as mentioned, the lower components are smaller than the upper components and can be disregarded. This is the procedure taken by Darwin to 'derive' Schrödinger's equation from Dirac's equation.[30] That is, we can only disregard two of the components of the four-component wave function in the non-relativistic limit. Now, Dirac's results are not non-relativistic. The whole point of Dirac's approach was to develop and apply a relativistic wave equation.

Dirac soon recognized that there was a fundamental difficulty with his equation. Already in his early work with the Klein-Gordon equation, Dirac had noticed the possibility of solutions corresponding to a charge +e (this is a general characteristic of any relativistic equation due to the relativistic formula for the Hamiltonian involving $E^2$). The problem is that when considering any small external electromagnetic field, "in general a perturbation will cause transitions from states with E positive to states with E negative".[31] Accordingly, Dirac considered the following:

> Such a transition would appear experimentally as the electron suddenly changing its charge from –e to +e, a phenomenon which has not been observed. The true relativity wave equation should thus be such that its solutions split up into two non-combining sets, referring respectively to the charge –e and the charge +e.[32]

As we have seen, also in the case of an electron in a central potential as described by Dirac's equation, that is not the case. Dirac knew that; he recognized at the beginning of his paper on the relativistic wave equation that he was unable to remove this difficulty and considered that his theory "is therefore still only an approximation".[33] However, Dirac expected "the probability of these transitions [to be] extremely small".[34] That was not the case. Soon afterward Heisenberg showed that the probability for transitions in

---

[29] Dirac, "*The principles of quantum mechanics*," on 274. According to Dirac, "it is not possible, of course, with an arbitrary electromagnetic field, to separate the solutions of [the relativistic wave equation] definitely into those referring to positive and those referring to negative values [of energy], as such a separation would imply that transitions from one kind to the other do not occur" (Ibid., 274).
[30] Mehra and Rechenberg, 302.
[31] Quoted in Mehra and Rechenberg, 306.
[32] Dirac, "*The quantum theory of the electron*," on 612.
[33] Ibid, 612.
[34] Quoted in Mehra and Rechenberg, 306.

which an electron in a state corresponding to a charge –e goes into a state corresponding to a charge +e (a negative energy state) was much larger than Dirac's estimation. Also, Heisenberg showed that the negative-energy states were necessary to obtain the correct dispersion formulae.[35] The problem with the negative-energy solutions was highlighted by Klein, when, at the end of 1928, he showed that the simple case of a positive-energy wave incident on a potential barrier could give rise to a transmitted negative-energy wave.[36]

By the end of 1929 Dirac had found a way to solve the '± difficulty' of his electron theory. In late March 1929, Heisenberg wrote to Dirac mentioning that Weyl thought he had a solution to the ± difficulty and asked Dirac if he knew and could give him any details about it. Weyl's idea consisted of suggesting that the two extra components might be ascribed not to the electron but to the proton (Dirac's equation would be in this way describing simultaneously electrons and protons). Dirac worked on Weyl's ideas and set forward a new interpretation of his equation that might provide a solution to the problem of the negative energy solutions.

Dirac's first written accounts of his new views were made in an exchange of letters with Bohr. In a letter to Dirac from November 24, Bohr had put forward the idea that there might not be a strict conservation of momentum and energy in some nuclear processes and that this might lead to a solution of the negative energy problem, resulting from the fact that, according to Klein, the potential that confines the electron in the nucleus induces transitions to negative energy states (by that time it was believed that there were electrons in the nucleus of the atoms). Dirac answered, on November 26, that "I should prefer to keep rigorous conservation of energy at all costs", and that "there is a simple way of avoiding the difficulty of electrons having negative kinetic energy".[37] Dirac then elaborates on his views recognizing that "the most stable states for the electron are those of negative energy with very high velocity".[38] What to do in this situation? Dirac proposes the following:

Let us now suppose there are so many electrons in the world that all these most stable states are occupied. The Pauli principle will then compel some electrons to remain in less stable states. For example, if all the states of –νe energy are occupied and also few of +νe energy, those electrons with +νe energy will be unable to make transitions to states of –νe energy and will therefore have to behave quite properly.[39]

In this account what would be a proton? Let us follow Dirac's reasoning:

It seems reasonable to assume that not all the states of negative energy are occupied, but that there are a few vacancies or 'holes'. Such a hole which can be described by a wave function like an X-ray orbit would appear experimentally as a thing with +νe energy, since to make the hole disappear (i.e. to fill it up,) one would have to put –νe energy into it. Further, one can easily see that such a hole would move in an electromagnetic field as though it had a +νe charge. These holes I believe to be the protons. When an electron of +νe energy drops into a hole and fills it up, we have an electron and proton disappearing simultaneously and emitting their energy in the form of radiation.[40]

[35] Ibid., 306-307.
[36] Ibid., 309-311.
[37] Quoted in Kragh, "*Dirac: a scientific biography*," on 90.
[38] Ibid., 91.
[39] Ibid., 91.
[40] Ibid., 91.

There was still the difficulty of the asymmetry in the masses of the proton and the electron. Dirac speculated about a physical mechanism that might explain the difference in mass.[41] In any case, he expressed confidence that this could be sorted out:

I have not yet been able to get a relativistic formulation of this. One can hope, however, that a proper theory of this will enable one to calculate the ratio of the masses of proton and electron.[42]

Bohr answered a few days later (on December 5), after discussing Dirac's letter with Klein. According to Bohr, the Klein paradox results from an unlimited mathematical application of the concept of potential (or more generally of field) in relativistic quantum mechanics. Bohr called attention to the necessity of taking into account the elementary unit of electrical charge in the determination of the actual potential barrier. Calculations made not taking into account this fact and the limits in the determination of the electron's position (and its associated uncertainty in momentum; and according to Bohr also in energy) would be beyond the 'actual limit of applicability' of the theory and any 'possible experimental arrangement'. Accordingly, Bohr considered that the elimination of Klein's paradox passed through an essential limitation of the mathematical use of the concept of field.[43] However, Bohr considered that within the domain of applicability of the theory, the concepts (being used in a restricted context) and the results obtained were consistent. Due to this, the problem of the transitions to negative-energy states would not occur in 'all the successful applications of the theory'. In this way, there would be no need for the hole theory.

Dirac did not agree with Bohr's views. In a letter sent to Bohr on December 9, Dirac maintained that the negative-energy states have a role in the theory. According to him:

There is one case where transitions of electrons from positive to negative energy levels does give rise to serious practical difficulties, as has been pointed out to me by Waller. This is the case of the scattering of radiation by an electron, free or bound. […] If one says the states of negative energy have no physical meaning, then one cannot see how the scattering can occur.

On my new theory the state of negative energy has a physical meaning, but the electron cannot jump down into it because it is already occupied. There is, however, a new kind of double transition now taking place […] This new kind of process just makes up for those excluded and restores the validity of the scattering formulas derived on the assumption of the possibility of intermediate states of negative energy.[44]

A paper containing Dirac's hole interpretation of his equation was published in early 1930. In this work Dirac first made clear his departure from Weyl's original idea of associating the negative energy solutions directly to the protons – the negative-energy states do not correspond to protons; it is holes in an infinite distribution of negative-energy electrons that would be the protons. In this paper, Dirac gave, in particular, an account of the scattering of a photon by a free electron according to his hole theory.

[41] Thomas Pashby, "Dirac's prediction of the positron: a case study for the current realism debate," *Perspectives on science* 20 (2012): 440-475, on 449.
[42] Quoted in Kragh, "*Dirac: a scientific biography*," on 91.
[43] In this respect see also Olivier Darrigol, "Cohérence et complétude de la mécanique quantique: l'exemple de « Bohr-Rosenfeld »," *Revue d'histoire des sciences* 44 (1991): 137-179, on 154-155.
[44] Quoted in Kragh, "*Dirac: a scientific biography*," on 93.

Dirac mentions that in a scattering process two types of intermediate states can occur. One case is as follows:

First an absorption of a photon with the electron simultaneously jumping to any state, and then an emission with the electron jumping into its final state, or else of first the emission and then the absorption.[45]

The other case is as follows:

One of the distribution of negative-energy electrons jumps up into the required final state for the electron with absorption (or emission) of a photon, and then the original positive-energy electron drops into the hole formed by the first transition with emission (or absorption) of a photon. Such processes result in a final state of the whole system indistinguishable from the final state with the more direct processes, in which the same electron makes two successive jumps.[46]

In this second case, an intermediate state must be taken into account in which a negative-energy electron absorbs the incident photon and makes a transition to a state of positive-energy, leaving a hole present (that is seen as a positron). Then the free electron 'drops into the hole and fills it up' and emits the outgoing photon. In the intermediate state, we have three particles present.[47] This means that the one-electron interpretation of Dirac's equation is not consistent; we are faced with a many-body theory. Dirac with his hole theory changed the character of his equation from a single-electron equation into an equation describing a many-body – in reality infinite – electron system.[48]

Even before the publication of his paper on the hole theory, several physicists, knowing the general lines of Dirac's ideas, had a skeptical reaction to it. In particular, Heisenberg made a rough calculation of the electron-proton interaction in the new theory. Heisenberg concluded that the electron and the proton had to have the same mass.[49] Dirac had already recognized this problem in one of his letters to Bohr but expected that a future detailed theory of the interaction between electrons and protons (holes) might solve this difficulty. A further objection was given a few months later, in a note by Oppenheimer in which the author calculated the transition probability for the annihilation of an electron and a proton that corresponds to the filling of a hole in the sea.[50] The result was not very promising. Oppenheimer obtained a mean lifetime of a free electron in matter that was too low, and totally inconsistent with the observed stability of matter.[51] Oppenheimer's proposition was to "return to the assumption of two independent elementary particles of opposite charge".[52] That is, to consider the electron and the proton as dissociated particles, each one being related to its 'Dirac sea' of negative-energy particles. In this way, there would not be a problem of a possible

[45] Dirac, "A theory of electrons and protons," *Proceedings of the Royal Society of London A* 126 (1930): 360-365, on 364.
[46] Ibid., on 365.
[47] Jun Sakurai, *Advanced quantum mechanics* (Reading: Addison-Wesley, 1967), on 134-138.
[48] Jaroslav Zamastil and Jakub Benda, *Quantum mechanics and electrodynamics* (Cham: Springer, 2017), on 302.
[49] Kragh, "*Dirac: a scientific biography*," on 94. 4
[50] Julius Oppenheimer, "Two notes on the probability of radiative transitions," *Physical Review* 35 (1930): 939-947.
[51] Kragh, "*Dirac: a scientific biography*," on 101-102.
[52] Quoted in Kragh, "*Dirac: a scientific biography*," on 102.

proton-electron annihilation. More importantly, in 1931 Weyl published a paper in which he proved by symmetry properties of Dirac's equation that the negative-energy electrons must have the same mass as the positive-energy electrons. However, by this time this was already clear to Dirac as a result of his correspondence with several physicists.[53] In this same year Dirac rethought his hole theory in the face of the objections being made and presented a new view on the problem:

It thus appears that we must abandon the identification of the holes with protons and must find some other interpretation for them. Following Oppenheimer, we can assume that in the world as we know it, all, and not merely nearly all, of the negative-energy states for electrons are occupied. A hole, if there were one, would be a new kind of particle, unknown to experimental physics, having the same mass and opposite charge to an electron. We may call such a particle an anti-electron. We should not expect to find any of them in nature, on account of their rapid rate of recombination with electrons, but if they could be produced experimentally in high vacuum they would be quite stable and amenable to observation.[54]

In 1932 a brief article by Anderson was published presenting experimental evidence for a new kind of positively charged particle with a mass much smaller than that of the proton. These experimental findings were presented without taking into account Dirac's theory. By that time Blackett and Occhialini had independent evidence for the positrons and previous to publication, they discussed their findings with Dirac. They published their results after Anderson's publication, mentioning Dirac's theory and presenting some ideas regarding the origin of the positive particle. According to Blackett and Occhialini "one can imagine that negative and positive electrons may be born in pairs during the disintegration of light nuclei".[55] However, they were cautious regarding their proposed mechanism:

When the behaviour of the positive electrons has been investigated in more detail, it will be possible to test these predictions of Dirac's theory. There appears to be no evidence as yet against its validity, and in its favour is the fact that it predicts a time of life for the positive electron that is long enough for it to be observed in the cloud chamber but short enough to explain why it had not been discovered by other methods.[56]

After a thorough revision of the experimental evidence for positrons Blackett and Occhialini published another article where their support to Dirac's theory was stronger:

These conclusions as to the existence and the properties of positive electrons have been derived from the experimental data by the use of simple physical principles. That Dirac's theory of the electron predicts the existence of particles with just these properties, gives strong reason to believe in the essential correctness of his theory.[57]

---

[53] Aaron Wright, "A beautiful sea: P. A. M. Dirac's epistemology and ontology of the vacuum," *Annals of Science* 73 (2016): 225-256, on 234-238.
[54] Dirac, "Quantized singularities in the electromagnetic field," *Proceedings of the Royal Society of London A* 133 (1931): 60-72, on 61.
[55] Quoted in Schweber, "*QED and the men who made It*," on 69.
[56] Ibid., 69.
[57] Ibid., 69.

Even with the experimental evidence for the positron, there was resistance to Dirac's hole theory. Concerning this discovery, Bohr considered that "even if all this turns out to be true, of one thing I am certain: that it has nothing to do with Dirac's theory of holes!".[58] And Pauli wrote to Dirac saying: "I do not believe in your perception of 'holes', even if the existence of the 'antielectron' is proved".[59] It turns out they were right and a different – field theoretical – approach was possible without any need for an infinite sea of negative-energy particles with some holes in it.

## 5. The field theoretical interpretation of Dirac's equation

When in 1927 Dirac developed a quantum treatment of the electromagnetic field, he did this from two different approaches, which at a quantum level gave the same mathematical result. In the final part of his paper, Dirac extended Jordan's initial work on the quantization of the electromagnetic field.[60] Dirac began by resolving the radiation field into its Fourier components. Then, having made the Fourier expansion of the field, the situation is as follows:

We can consider the energy and phase of each component to be dynamical variables describing the radiation field […] we can suppose each $E_r$ and $\theta_r$ to form a pair of canonically conjugate variables […] satisfying the standard quantum conditions $\theta_r E_r - E_r\theta_r = i\hbar$ […] this assumption immediately gives light-quantum properties to the radiation.[61]

But the main aspect of Dirac's work was not the quantization of a wave; on the contrary, his paper is mainly a treatment of an assembly of identical particles. It seems that Dirac got to this approach by 'playing about with Schrödinger equation'. According to his words:

I remember the origin of that work was just playing about with equations. I was intending to get a theory of radiation at the time. I was just playing about with the Schrödinger equation. I got the idea of applying the quantization to it and worked out what it gave and found out it just gave the Bose statistics.[62]

In this method, later called 'second quantization', Dirac started with an assembly of N similar independent particles subjected to a perturbation, whose wave function is $\psi = \Sigma_r a_r \psi_r$ (where $\psi_r$ are the eigenfunctions of the free particles) and considered the expansion coefficients $a_r$ as canonical conjugates. Working with the canonical variables $b_r = a_r\, e^{-iw_r t/\hbar}$ and $b_r^* = a_r^*\, e^{iw_r t/\hbar}$ (where $w_r$ is the energy of the unperturbed state), Dirac assumed that these variables were "q-numbers satisfying the usual quantum conditions

[58] Quoted in Kragh, "*Dirac: a scientific biography*," on 112.
[59] Ibid., 112.
[60] Max Born, Werner Heisenberg, and Pascual Jordon, "On quantum mechanics II," in B. L. van der Waerden (ed.), *Sources of quantum mechanics* (New York: Dover Publications, 1967), on 321-385. Original work published in *Zeitschrift für Physik* 35 (1926). Schweber, "*QED and the men who made It*," on 11. Darrigol, "The origin of quantized matter waves," *Historical Studies in the Physical and Biological Sciences* 16 (1986): 198-253, on 220-221.
[61] Dirac, "The quantum theory of the emission and absorption of radiation," *Proceedings of the Royal Society of London A* 114 (1927): 243-265, on 244.
[62] Quoted in Darrigol, "La genèse du concept de champ quantique," *Annales de Physique* 9 (1984): 433-501, on 46.

instead of c-numbers".[63] This gives a false impression that a quantization scheme is being used, but what is being done is changing from a configuration space representation to an occupation number representation,[64] where the commutation relation $[b_r, b_s^*] = d_{rs}$ holding between $b_r$ and $b_r^*$ serves to impose the symmetrization of the configuration space wave function, which means that the particles obey Bose-Einstein statistics.[65] What Dirac thinks he demonstrates in this work, is the equivalence between a quantized electromagnetic wave and a system of bosons (light-quanta), but for this he makes an identification of the quanta of energy with the particles (light-quanta), and in order to get this result the particle cannot cease to exist when it is apparently absorbed, or created when it is emitted. It is, therefore, necessary to have an infinite sea of light-quanta in a state in which their momentum and energy are zero, from which the particle can jump from or into.[66] The fact that Dirac considers that there is "a complete harmony between the wave and light-quantum description",[67] does not mean that he takes over this equivalence to the case of the electrons. As Dirac clearly states in this article, he makes a sharp distinction between the case of electromagnetic radiation and matter. For Dirac, there simply is no real de Broglie wave that, after quantization, permits the description of the electrons.[68]

Jordan's reading of Dirac's work was quite different. Going back to the cumbersome method of second quantization that for Dirac was "nothing but a convenient way to take Bose statistics into account",[69] Jordan interpreted the scheme of second quantization as the quantization of a classical wave described by a classical wave equation that could be the Maxwell-Lorentz equations for the case of the electromagnetic field or a Schrödinger equation for the case of electrons seen not as particles but as de Broglie waves.[70] This view of Jordan had the advantage to make it possible to treat the quantized waves in a three-dimensional space (or four-dimensional Minkowski space-time) instead of using a 3n (multidimensional) space in the case of an n-particle system.[71]

In a paper sent to publication in July 1927, Jordan made a conceptual turnaround on Dirac's approach, and applied his method to the quantization of quanta obeying Pauli's exclusion principle; that is to electrons. Contrary to the case of photons (light quanta), in which there is no limit to the number of particles that may occupy the same state, the electrons (according to Pauli's exclusion principle) cannot be in the same state; that is, the occupation number for each possible state can only be 0 or 1 (another way of making this statement is to say that the electrons are fermions; that is, they obey Fermi-Dirac statistics). Jordan's approach was to take the method used by Dirac with the light quanta and to apply it to the electron as a de Broglie wave that was a solution of a (classical) Schrödinger equation. In this way, Dirac's 'second quantization' was from Jordan's perspective a 'first quantization' of a classical wave. This was simply implemented by taking the coefficients of each normal mode to be matrices satisfying certain commutation relations.

---

[63] Dirac, "*The quantum theory of the emission and absorption of radiation*," on 251.
[64] Tian Cao, *Conceptual development of 20th century field theory* (Cambridge: Cambridge University Press, 1997), on 166-167.
[65] Dirac, "*The quantum theory of the emission and absorption of radiation*," on 252-255. Schweber, "*QED and the men who made It*," on 28.
[66] Cao, 163-164.
[67] Dirac, "*The quantum theory of the emission and absorption of radiation*," on 245.
[68] Ibid., 247.
[69] Darrigol, "*The origin of quantized matter waves*," on 229.
[70] Ibid., 229-230.
[71] Schweber, "*QED and the men who made It*," on 36.

In this work, Jordan did not take correctly into account the phase factors in the matrices "necessary to guarantee that the creation operators for different energy states anticommute with one another".[72] This point was settled by the end of the year in a collaboration with Wigner, where the correct anticommutation relations were given. In any case, already in his first incomplete work, Jordan was able, by his quantization of the wave, to obtain "an evolution identical to the one given by anti-symmetrical wave functions in configuration space".[73] That is, Jordan showed the equivalence, for the electrons, of adopting as a classical starting point, not the electrons as particles but electrons as classical de Broglie waves. With this procedure, there was no need for a multidimensional abstract configuration space since it was possible with the wave approach to maintain the description solely in terms of a quantized wave described in a three-dimensional space. In this way, already in this first incomplete work, Jordan was able to conclude that "a quantum-mechanical wave theory of matter can be developed that represents electrons by quantum waves in the usual three-dimensional space".[74] This conclusion made possible for Jordan to make the 'programmatic' assertion: "The natural formulation of the quantum theory of electrons will be attained by conceiving light and matter as interacting waves in three-dimensional space".[75]

Heisenberg and Pauli adopted Jordan's approach in the development of a quantum field approach to the description of the interaction of radiation and matter. In their case they took Dirac's equation as a classical wave equation. The 'de Broglie' solution of Dirac's equation is then quantized according to the procedure set forward by Jordan. The elementary excitations (quanta) of the field resulting from the quantization are the particles. In a sense this field theoretical approach as applied by Heisenberg and Pauli is incomplete. The point is that the energy of the field can be negative. This is so due to the presence of negative energy quanta. This means it was still necessary to make use of Dirac's hole theory to make sense of the field quantization. The negative energy states were taken to be full and a positron was identified with an empty negative-energy state. That is, a positron in this field theoretical view was taken to be the lack of an elementary excitation in an infinite sea of negative-energy elementary excitations.

In November 1933, Fock published a paper where he made a symmetrical treatment of free electrons and positrons without using negative-energy particles, following a procedure developed by Heisenberg in 1931.[76] In a letter to Pauli from July 1933, Heisenberg had presented that same approach and later applied it in a paper published in 1934, considering the case where an external field was present.[77] In this paper, he required that "the symmetry of nature in the positive and negative charges should from the very beginning be expressed in the basic equations of [the] theory".[78] Taking the Dirac equation and its adjoint equation as classical field equations derived from a classical Lagrangian, an arbitrary field can be expanded in terms of the complete set of free-particle solutions:[79]

---

[72] Ibid., 37.
[73] Darrigol, "*The origin of quantized matter waves*," on 231.
[74] Quoted in Darrigol, "*The origin of quantized matter waves*," on 232.
[75] Ibid., 232.
[76] Abraham Pais, *Inward bound* (Oxford: Oxford University Press, 1986), on 379.
[77] Darrigol, "La genèse du concept de champ quantique," on 479; Arthur Miller, *Early quantum electrodynamics* (Cambridge: Cambridge University Press, 1994), on 63.
[78] Heisenberg, "Remarks on the Dirac theory of the positron," in Miller, on 169. Original work published in *Zeitchrift für Physik* 90 (1934).
[79] Schweber, *An introduction to relativistic quantum field theory* (New York: Dover Publications, 1961), on 222-223.

$$\psi(\mathbf{x}) = \int \frac{d^3p}{(2\pi)^{3/2}} \left( \frac{m}{E_{\mathbf{p}}} \right)^{1/2} \left\{ \sum_{r=1}^{2} b_r(\mathbf{p}) w^r(\mathbf{p}) e^{i\mathbf{p}\cdot\mathbf{x}} + \sum_{r=3}^{4} b_r(-\mathbf{p}) w^r(\mathbf{p}) e^{-i\mathbf{p}\cdot\mathbf{x}} \right\}.$$

The quantization scheme consists in replacing the expansion coefficients by operators satisfying the anticommutation relations $[b_n, b_m]_+ = [b_n^*, b_m^*]_+ = 0$ and $[b_n, b_m^*]_+ = \delta_{nm}$. With this procedure $\psi(\mathbf{x})$ and the adjoint spinor field $\psi^*(\mathbf{x})$ become operators that act on state vectors of a Fock space;[80] and $b_r(\mathbf{p})$ and $b_r^*(\mathbf{p})$ are interpreted as the annihilation and creation operators of an electron in the state ($\mathbf{p}$, r). Redefining the operators for the negative-energy states as $b_{r+2}(-\mathbf{p}) = d_r^*(\mathbf{p})$ and $b_{r+2}^*(-\mathbf{p}) = d_r(\mathbf{p})$ with r = 1, 2, these operators can be interpreted as the creation and annihilation operators for a positive-energy positron,[81] and the expansion of the $\psi(\mathbf{x})$ operator is now given by

$$\psi(\mathbf{x}) = \int \frac{d^3p}{(2\pi)^{3/2}} \left( \frac{m}{E_{\mathbf{p}}} \right)^{1/2} \sum_{r=1}^{2} \left\{ b_r(\mathbf{p}) w^r(\mathbf{p}) e^{i\mathbf{p}\cdot\mathbf{x}} + d_r^*(\mathbf{p}) \upsilon^r(\mathbf{p}) e^{-i\mathbf{p}\cdot\mathbf{x}} \right\}.$$

With this formulation, there are no negative-energy states (which are identified as positive-energy positrons) and so no need for the infinite sea of negative-energy electrons. Also, in the field operators $\psi(\mathbf{x})$ and $\psi^*(\mathbf{x})$ we have simultaneously components related to electrons and positrons. Let us consider the total charge operator

$$Q = \sum_r \int d^3p [-a_r^*(\mathbf{p}) a_r(\mathbf{p}) + b_r^*(\mathbf{p}) b_r(\mathbf{p})] = \int d^3p [n^+(\mathbf{p}) - n^-(\mathbf{p})],$$

where $n^-(p)$ is the number of quanta identified as electrons and $n^+(p)$ is the number of quanta identified as positrons.[82] As we can see from this expression, as Jordan proposed, the quantization of charge and the subsequent emergence of a particle-like concept of an electron can be seen as a result of the quantization of the classical field.

## 6. Combining results from the different interpretations of the Dirac equation

The field theoretical reinterpretation of Dirac's equation does not have issues of the like of the one-electron wave equation interpretation. The original one-electron interpretation is not consistent – at best we must regard Dirac's equation as a relativist wave equation as the equation of a many-body electron system. At this point, we might ask what is the relation between the original formulation of the Dirac equation with its final formulation? In what regards the equation itself we can see the Dirac one-electron equation as a 'semi-classical' equation resulting from using the so-called external field

[80] Considering the vacuum state, which is the state with no quanta, an n-quanta Hilbert space can be defined by n applications of creation operators. The Fock space is the (infinite) product of the n-quanta Hilbert spaces: $H^{(0)} \oplus H^{(1)} \oplus H^{(2)} \ldots$ (Schweber, "*An introduction to relativistic quantum field theory*," on 134-137. Eberhard Gross, Erich Runge, and Olle Heinonen, *Many-particle theory* (Bristol: Adam Hilger, 1991), on 21).

[81] Schweber, "*An introduction to relativistic quantum field theory*," on 223. Miller, 56.

[82] Josef Jauch and Fritz Rohrlich, *The Theory of photons and electrons* (Berlin and New York: Springer-Verlag, 1976), on 64.

approximation,[83] where there appears to be a classical potential within the quantum formalism, but that really is due to a quantum field theoretical description of the interaction with a very heavy charged particle (described by a quantum field) when its recoil is neglected.[84] It is within the external field approximation that a Dirac field operator equation with an 'external field' appears, and from which the relativistic one-electron equation with a 'classical potential' can be seen to emerge from the full quantum electrodynamics.[85]

But what about the solutions? What to make of the solutions of Dirac's equation as a one-electron relativistic wave equation from the quantum field perspective? We will just consider the very simple case of hydrogen-like problems. When considering the exact solution of the one-electron Dirac equation in a central potential we have a four-component wave function. In simple terms, we can say that in the four-component wave function we have components that, from a quantized field perspective, are related to both electrons and positrons. If we want, taking into account this perspective, a simple model for the hydrogen atom with only one electron, while using the Dirac equation as a relativistic one-electron equation, we must develop a method that uses only two-component positive-energy wave functions to describe the electron. This approach is also important if we want to make a clear connection between the relativistic and non-relativistic equations; that is, between the Dirac and the Schrödinger equations.

Considering the four-component solutions of the Dirac equation $\psi = (\psi_A\ \psi_B)$ in the presence of electromagnetic coupling; in the non-relativistic limit, the lower two components $\psi_B$ are smaller than the upper two $\psi_A$. When calculating matrix elements like $(\psi, \gamma_4\psi) = \psi_A^*\psi_A - \psi_B^*\psi_B$, neglecting terms of the order $(v/c)^2$, we obtain an expression only in terms of the large components $\psi_A^*\psi_A$, reducing the matrix element to its non-relativistic form in terms of two-component wave functions.[86] In the non-relativistic limit, the large components can be seen as the solution of the Schrödinger-Pauli two-component wave equation. Concerning this approach to the problem of the non-relativistic limit of the Dirac equation, Foldy and Wouthuysen argued that "the above method of demonstrating the equivalence of the Dirac and Pauli theories encounters difficulties […] when one wishes to go beyond the lowest order approximation".[87] They proposed a new method (using a different representation than Dirac's original one) which would not only provide better results for higher-order approximations but also the definition of new operators for position and spin "which pass over into the position and spin operators in the Pauli theory in the non-relativistic limit".[88]

In the case of the Dirac equation for a free electron, it is possible to perform a canonical transformation on the Hamiltonian that enables the decoupling of the positive- and negative-energy solutions of Dirac's equation, each one becoming associated to a two-component wave equation. This means we get two independent equations for two-component wave functions, and that we can identify the equation with a positive-energy solution as the Schrödinger-Pauli equation. The case of an electron interacting with an external electromagnetic field is more involved. The trick is to consider the electromagnetic field as a perturbation and to make a sequence of transformations to

[83] Ibid., 303.
[84] Schweber, "*An introduction to relativistic quantum field theory*," on 535.
[85] Jauch and Rohrlich, 307 and 313.
[86] Franz Mandl, *Quantum mechanics* (London: Butterworths, 1957), on 214-215.
[87] Lawrance Foldy and Siegfried Wouthuysen, "On the Dirac theory of spin ½ particles and its non-relativistic limit," *Physical Review* 78 (1949): 29-36, on 29.
[88] Ibid., 29.

obtain the separation of negative- and positive-energy solutions. In the non-relativistic limit, like in the previous method, the Schrödinger-Pauli equation is obtained.

It is then possible to rework the relativistic one-electron Dirac equation in a way in which only positive-energy solutions are considered. Foldy and Wouthuysen applied their method to the case where an electron interacts with an external electromagnetic field. By making three canonical transformations and using only terms of the order $(1/m)^2$ they obtained a Hamiltonian (incorporating relativistic correction to this order) that enabled a clear separation of positive- and negative-energy solutions. With this method, the non-relativistic limit of Dirac's equation results in two uncoupled equations one with positive-energy solutions and the other with negative-energy solutions. With their three canonical transformations, Foldy and Wouthuysen were able to obtain the same results (i.e. the same wave equation) as in the Pauli theory – as Dirac did in his original work.

Going the other way around; what would be the counterpart of this result in terms of a quantum field approach? One possibility might be to use the so-called Furry or bound interaction representation within the external field approximation. This gives a method for calculating corrections (due to a quantized electromagnetic field) to the energy levels of a bound electron (due to a static external potential) determined by the Dirac equation as a relativistic one-electron equation.[89] But the starting point is the field operator defined by

$$\psi(x) = \sum_{\mathrm{r}} u_{\mathrm{r}}(\mathbf{x}) a_{\mathrm{r}} e^{-\mathrm{i}E_{\mathrm{r}}t} + v_{\mathrm{r}}(\mathbf{x}) b_{\mathrm{r}}^{*} e^{\mathrm{i}E_{\mathrm{r}}t},$$

where $u_r(\mathbf{x})$ and $v_r(\mathbf{x})$ are obtained by solving the Dirac equation for a positive-energy particle representing the electron, and a negative-energy particle representing de positron: $Hu_a(\mathbf{x}) = E_a u_a(\mathbf{x})$ and $Hv_b(\mathbf{x}) = -E_b v_b(\mathbf{x})$, where $H = i\gamma^0\boldsymbol{\gamma}.\nabla - e\gamma^0\gamma_\mu\varphi^\mu + i\gamma^0 m$, with $\varphi^\mu$ a static external field.[90] The first equation is exactly the one solved in the case of the one-electron interpretation of the Dirac equation. This means that $E_a$ gives the positive energy levels obtained by this method and that in spite of identifying $u_a(\mathbf{x})$ as the electron's positive-energy wave function, it contains what in the limit of a free-particle solution are positive- and negative-energy components.[91] Now, what is needed is a method in which the free particle positive-energy characteristic of the electron is maintained during the interaction with no mixing of positive- and negative-energy components.

Another possibility is to address this issue within a fully field theoretical approach. To do so we consider our electron (as the quanta of the Dirac field) in interaction with a proton – the nucleus of the 'hydrogen atom'– which we describe also in terms of a Dirac field. In the calculation, the proton is taken to be a 'big' positron that has the same mass as that of the proton. One way to calculate the energy levels of the bound state of the electron (i.e. the binding energy between the 'proton' and the electron) is by using the so-called ladder approximation.[92] Concerning this approximation, Bethe and Salpeter remarked the following:

---

[89] Vladimir Berestetskii, Evgeny Lifshitz, and Lev Pitaevskii, *Quantum electrodynamics* (Oxford: Butterworth-Heinemann, 1982), on 487.
[90] Jauch and Rohrlich, 313.
[91] Schweber, "*An introduction to relativistic quantum field theory*," on 566.
[92] Ibid., 713. Walter Greiner and Joachim Reinhardt, *Quantum electrodynamics* (Berlin: Springer, 2009), on 329-333.

Although the probability for the exchange of a quantum during a small time interval is fairly small, during the infinite time of existence of the bound state an indefinite number of quanta may be exchanged *successively*. It is just such processes that the ladder-type graphs deal with.[93]

Thus, in the quantum field theory approach, the binding of the electron in the atom is achieved by an exchange of photons with the proton. By having a perspectival approach to solutions in each interpretation we see that while there are huge differences, there are solutions in both cases that make sense when looking from the other interpretation. If we start with the 'identification' electron ≡ positive-energy quantum with a negative charge, we can find solutions to the hydrogen problem in the one-electron formulation of the Dirac equation in which the electron is described by a positive-energy wave function. Corresponding to this, in the quantum field formulation we can find solutions to the hydrogen problem where the electron is described in terms of just this type of quanta of the Dirac field and no quanta identified as positrons are at sight.

## 7. Conclusion

The interpretation of the Dirac equation in relation to the quantized Dirac field does not lead to consistency issues as the original interpretation of the Dirac equation as a one-electron equation. We can best regard the Dirac equation as having a 'subsidiary' role in relation to the Dirac field – as a classical field equation of a de Broglie wave, that contains the physics that will be encapsulated in the quantized Dirac field. The 'original' equation – the relativistic wave equation – results from the so-called external field approximation in quantum electrodynamics. The solutions obtained within this approximation, which we traditionally address as solutions of relativistic quantum mechanics problems, can be seen as approximate solutions of relativistic quantum field problems. In this way, there is a 'rationale' to the solutions obtained adopting the original formulation of the Dirac equation. Here, we have only looked into the very simple case of hydrogen-like problems. Both, in full quantum electrodynamics as when adopting the external field approximation, we can describe an electron bound to a nucleus without having to include quanta identified as positrons or negative-energy wave functions.

[93] Edwin Salpeter and Hans Bethe, "A relativistic equation for bound-state problems," *Physical Review*, 84 (1951): 1232-1242, on 1234.