The Dirac equation, the concept of quanta, and the description of interactions in quantum electrodynamics

Mario Bacelar Valente

Department of Philosophy, Logic and Philosophy of Science
University of Seville
mar.bacelar@gmail.com

Abstract

In this article the Dirac equation is used as a guideline to the historical emergence of the concept of quanta, associated with the quantum field. In Pascual Jordan's approach, electrons as quanta result from the quantization of a classical field described by the Dirac equation. With this quantization procedure – also used for the electromagnetic field – the concept of quanta becomes a central piece in quantum electrodynamics. This does not seem to avoid the apparent impossibility of using the concept of quanta when interacting fields are considered together as a closed system. In this article it is defended that the type of analysis that leads to so drastic conclusions is avoidable if we look beyond the mathematical structure of the theory and take into account the physical ideas inscribed in this mathematical structure. In this case we see that in quantum electrodynamics we are not considering a closed system of interacting fields, what we have is a description of the interactions between distinct fields. In this situation the concept of quanta is central, the Fock space being the natural mathematical structure that permits maintaining the concept of quanta when considering the interaction between the fields.

1. Introduction

Once upon a time, Richard P. Feynman wrote that "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 – Eg. 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" (quoted in Schweber, 1994, p. 407). He wrote this having in mind, in particular, the Dirac equation: $i\nabla\psi = m\psi$ (Feynman, 1961, p. 57).

In this article the Dirac equation will be used as a guideline to introduce the concept of quanta and to reveal its importance in the description of interactions in quantum electrodynamics. To this end the historical evolution and interpretation of the Dirac equation is considered. In section 2, I present the evolution of the Dirac equation from its first formulation as a relativistic wave equation for an electron, to a classical field equation from which an electron-positron quantum field is obtained. In this transition, the Dirac equation went from being a relativistic 'update' of the Schrödinger equation in the calculation of energy levels in atoms (basically of hydrogen) to becoming one of the cornerstones of the most successful quantum field theory: quantum electrodynamics. In section 3, I will try to clarify the relation between the different interpretations of the Dirac equation. In this way the results provided by Dirac's equation as a relativistic one-electron equation are reinterpreted from the perspective of the quantized Dirac field. Doing this, the importance of the concept of quanta in the description of bound states becomes clear. By contrast, bound states are usually only described at the level of the one-electron interpretation of the Dirac equation, which gives a distorted idea of the physical description of bound states that should be described from the perspective of quantum fields. In particular, an analysis of a two-body description of the hydrogen

atom reveals a distinctive feature of quantum electrodynamics: the interaction between fermions described as an exchange of photons.

In section 4, I will consider possible problems to the previous view. According to John Earman and Doreen Fraser, the Haag theorem seems to imply the mathematical inconsistency of the usual treatment of interactions in quantum electrodynamics using perturbative methods (Earman & Fraser, 2006), in which the concept of quanta is central. Fraser presents the idea that "once infinite renormalization counterterms are introduced … it is no longer possible to prove Haag's theorem" (Fraser, 2006, p. 2). I shall argue that simply calling attention to the fact that renormalization "renders the theory mathematically not well-defined" (Fraser, 2006, p. 171) does not provide any answer to the question: "why perturbation theory works as well as it does" (Earman & Fraser, 2006, p. 307). This is not an unimportant question, since we are facing the puzzling situation that the predictions of quantum electrodynamics are so accurate, while the theory from which they are derived appears to be mathematically inconsistent. In this section I will try to provide an answer to this question. In the process I will try to show that Earman & Fraser's conclusion that "Fock representations are generally inappropriate for interacting fields" (Earman & Fraser, 2006, p. 330) or Fraser's related conclusion that "an interacting system cannot be described in terms of quanta" (Fraser, 2008, p. 842), are not valid ones in quantum electrodynamics: we do not have an interacting system but two weakly interacting fields (systems), and their interaction is described in terms of quanta.

In a nutshell my argument will be as follows: it turns out that quantum electrodynamics, because the series expansion of the S-matrix is divergent, is unable to treat radiation and matter as one closed system, and can only give an approximate description (using a perturbative approach) of their interaction as distinct systems. If we try to close the gap and treat radiation and matter as one closed system, corresponding to an exact solution of the coupled non-linear Maxwell-Lorentz and Dirac equations, our perturbative approach fails. My view is that there is a one-to-one correspondence between having meaningful mathematics results and the physical assumptions used to set up the theory (in quantum electrodynamics I take part of the input physical concepts to be radiation and matter taken to be independent systems that are independently quantized due to the weak interaction between then). This is so, because the physical concepts are defined with their mathematical 'support', not beyond or independently of maths. In a situation where we have an infinity popping out we cannot have a well-defined use of the physical concepts involved. In the case of quantum electrodynamics this situation occurs when trying to give a full description of the (weak) interaction between matter and radiation, which would correspond to treat them as one closed system. I see the divergence of the S-matrix series expansion as a fingerprint of a tentative application of the mathematics of the theory beyond its physical content.

By focusing on the basic physical concepts used to set up the theory, and not solely on the (ill-defined) mathematical structure of the theory, we can make sense of the success of quantum electrodynamics (independently of the Haag theorem). This provides a different perspective than the one adopted by Earman & Fraser that not only contradicts several of their conclusions, but also, in my view, makes quantum electrodynamics more intelligible.

2. Different views on the Dirac equation

2.1. Dirac's equation as a one-electron equation

In the early days of 1928, Paul Dirac published a paper which presented what he considered to be a relativistic wave equation for one electron. The main guideline for obtaining Dirac's equation was, besides the requirement of being relativistic, to conform to the mathematical scheme of transformation theory. This implied an equation that is linear in the time derivative d/dt. Dirac also felt that "an appropriate formulation of quantum mechanics will only be possible when we succeed in treating space and time as equal to one another" (quoted in Kragh, 1990, p. 54). This means that space and time must appear in the equation on equal footing, as the coordinates of a Minkowski space-time.

Dirac ended up with a relativistic equation for a four-component wave function (Dirac, 1928, p. 615):

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

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 4 × 4 matrices that are an extension of the 2 × 2 Pauli matrices, and the 4 × 4 matrices ρ_1 and ρ_3 are also obtained from the Pauli matrices (Dirac, 1928, pp. 613-615). Spin could explain the existence of two of the four components of the wave function solution of the Dirac equation, but there were another two that had to be accounted for.

In the simplest case of a free electron, four independent solutions exist: two corresponding to electron states with positive energy and two states with negative energy. Dirac's first attitude was to reject these negative-energy solutions because they referred to a charge +e of the electron (Dirac, 1928, p. 618). This was possible in a first-order approximation. Looking at the exact solution we can see that the situation is far from that simple. The Dirac equation can be written as a set of two coupled differential equations for a pair of two-components wave functions ψ_A and ψ_B. 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 (ψ_B) are smaller that the upper two-components (ψ_A), roughly by a ratio of $\upsilon/2c$, where υ 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 $-\mathbf{p}$ 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, we have a four-component wave function (e. g. Sakurai, 1967, pp. 122-129), which, we can take to 'have', by resorting to the free electron case, positive and negative energy components, or as Dirac mentions, solutions referring to a charge –e and +e (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" (Dirac, 1958, p. 274)). 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" (Dirac, 1958, p. 274). 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-components wave equation (corresponding to a negative energy state), and as mentioned, the lower components are smaller than the upper components and can be disregarded. That is, we can only disregard two of the components of the four-component wave function in the non-relativistic limit.

Dirac's not very consistent idea of neglecting the negative-energy solutions was soon challenged when Oscar Klein showed that the simple case of a positive-energy wave incident on a potential barrier could give rise to a transmitted negative-energy wave, a result known as the Klein paradox (Mehra & Rechenberg, 2000, pp. 309-311). More importantly, to obtain the Klein-Nishina relativistic formula for the photon-electron scattering, the negative-energy solutions had to be considered. In Niels Bohr's words: "the striking confirmation which this formula has obtained became soon the main support for the essential correctness of Dirac's theory when it was apparently confronted with so many grave difficulties" (quoted in Kragh, 1990, p. 89). So, the existence of negative-energy solutions in the Dirac equation had to be properly addressed.

Dirac himself did this. He recognized the problem that "in the general case of an arbitrary varying electromagnetic field we can make no hard-and-fast separation of the solutions of the wave equation into those referring to positive and those to negative kinetic energy" (Dirac, 1930, p. 361), and proposed as a solution that "*all the states of negative energy are occupied except perhaps a few of small velocity*. […] We shall have an infinite number of electrons in negative-energy states, and indeed an infinite number per unit volume all over the world, but if their distribution is exactly uniform we should expect them to be unobservable" (Dirac, 1930, p. 362). In this first version of his negative-energy electron sea, the 'holes' in the ocean were identified as protons (Dirac, 1930, p. 363). A few months latter, a note by J. Robert Oppenheimer (1930) was published 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. The result was not very promising. Oppenheimer obtained a mean lifetime of a free electron in matter that was too low (Kragh, 1990, pp. 101-102).

What turned out to be more important was Hermann Weyl's demonstration that "according to Dirac's own theory of the electron the hole must necessarily have the same mass as an ordinary electron" (Kragh, 1990, p. 102). This paved the way to a second version of Dirac's sea in which "a hole, if there is 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" (Dirac, 1931, p. 61).

The reception of the 'hole' theory was not very warm. Some compared the negative-energy sea with the unobservable ether, others referred to the negative-energy electrons as donkey electrons because of their unusual 'dynamical' behaviour, and Wolfgang Pauli – not restrained by Dirac's views on the importance of the transformation theory – worked out with Victor F. Weisskopf a quantum field theory based on the Klein-Gordon equation, in which there was no need for a sea to take care of the concept of anti-particles (Kragh, 1990, pp. 111-114). In this work they followed recent approaches where the electrons and the anti-electrons (positrons) were described in the formalism of the theory on the same footing in a completely symmetrical way (e. g. Rugh, Zinkernagel & Cao, 1999, pp. 112-113).

This symmetrical treatment of the electrons and the positrons solved the newly created interpretation problem of Dirac's equation. Dirac's solution of the negative-energy difficulty makes it impossible to maintain a single electron interpretation of his equation, as can be seen in the derivation of the Klein-Nishina formula using the hole theory: In the scattering of a photon by a free electron, intermediate states with a negative-energy solution must be considered. For example, an intermediate state must be taken into account when 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" (Dirac, 1930, p.

363), and emits the outgoing photon. In the intermediate state we have three particles present (Sakurai, 1967, pp. 134-138), which means that the one-electron interpretation of Dirac's equation is not feasible.

2.2. The (classical) Dirac equation and the quantization of the Dirac field

When in 1927 Dirac developed a quantum treatment of the electromagnetic field he did this from two different approaches, which at the 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 (Schweber, 1994, pp. 9-11), 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 quantized particles. It seems that Dirac got to this approach by "playing about with Schrödinger equation" (quoted in Darrigol, 1986, p. 226). For Dirac this method, later called 'second quantization', turns out to be simply a different procedure to impose the Bose-Einstein statistics on the particles. He could instead have simply selected symmetrical wave functions as the physically admissible wave functions for this type of particle (Dirac, 1926, p. 672).

Dirac started with an assembly of N similar independent particles (subject to an interaction with an atomic system), whose wave function, solution of the Schrödinger equation, is $\psi = \Sigma_r a_r \psi_r$ (where ψ_r is the eigenfunction of a particle in the state r). Dirac took the expansion coefficients a_r to be 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 a particle in the state r), Dirac assumed that these variables were "q-numbers satisfying the usual quantum conditions instead of c-numbers" (Dirac, 1927, p. 251). This gives the false impression that an additional quantization scheme is being used, but what is being done is changing from a configuration space representation to an occupation number representation (Cao, 1997, pp. 166-167). The commutation relation $[b_r, b_s^*] = \delta_{rs}$ holding between b_r and b_r^* serves to impose the symmetrization of the wave functions. This implies that the particles obey Bose-Einstein statistics (Dirac, 1927, pp. 252-255; Schweber, 1994, p. 28).

As mentioned previously, in the final part of his work Dirac considered the quantization of a classical electromagnetic wave. Starting with the classical Hamiltonian describing an atom in interaction with radiation, the field was described by canonical variables, which in the quantization procedure where taken to be q-numbers satisfying the usual quantum commutation relations. With this procedure Dirac obtained a Hamiltonian for the quantized system, which was, according to his view, consistent with the Hamiltonian obtained using the particle view (Dirac, 1927, p. 263). Dirac considered that this work demonstrated the equivalence between a quantized electromagnetic wave and a system of bosons (light-quanta). Implicit in this conclusion is the identification of the quanta of energy with the particles (light-quanta). In order to get this result, a 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 (photons), in a state in which their momentum and energy are zero, from which the particle can jump from or into (Cao, 1997, pp. 163-164; Dirac, 1927, p. 261). The fact that Dirac considers that there is "a complete harmony between the wave and light-quantum description" (Dirac, 1927, p. 245), does not mean that he takes over this equivalence to the case of the electrons. In fact, as Dirac clearly states in this article, he makes a clear distinction between the case of electromagnetic radiation and matter. For Dirac there simply is no real de Broglie wave which, after quantization, permits the description of the electrons (Dirac, 1927, p. 247).

Strongly influenced by Louie de Broglie's views, Jordan made a very personal reading of Dirac's work. Jordan considered the method presented by Dirac as a legitimate procedure to obtain quantized waves, which could be applied to the case of de Broglie waves. That is, for Jordan, Dirac's method meant a first quantization of a classical wave. Jordan extended Dirac's approach to the case of matter (described by a 'classical' Schrödinger equation), and worked out a quantization procedure that could be applied in the case of Fermi-Dirac statistics (the statistics of the quantized particles, corresponding to antisymmetrical wave functions). Already in his first (incomplete) treatment of quantized matter waves, Jordan felt confident to state:

> The results we have reached hardly leave any doubt that … a quantum-mechanical wave theory of matter can be developed that represents electrons by quantized waves in the usual three-dimensional space. The natural formulation of the quantum theory of electrons will be attained by conceiving light and matter as interacting waves in three-dimensional space. The basic fact of electron theory, the existence of discrete electric particles, appears in this context as a characteristic quantum phenomenon; indeed it means exactly that matter waves occur only in discrete quantum states (quoted in Darrigol, 1986, p. 232).

The work of Jordan (and his collaborators) was extended by Werner Heisenberg and Pauli in the development of a relativistic quantum theory of fields in which the electromagnetic and matter fields were described by a classical Lagrangian and quantized by a new method (Miller, 1994, p. 31). The difficulties in the quantization of the Maxwell equations delayed for more than a year the completion of their work. When finally Heisenberg managed "to eliminate the difficulties by means of a formal trick" (quoted in Pais, 1986, p. 343), Dirac had already published his equation, and Heisenberg and Pauli adopted it in their quantum field treatment of the interaction of matter and the electromagnetic field.

After a subsequent refinement, mentioned above, the quantum field approach gave a different view on the negative-energy solutions without need for the hole theory. 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 (Schweber, 1961, pp. 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;[1] 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 (Schweber, 1961, p. 223; Miller, 1994, p. 56), and the expansion of the $\psi(\mathbf{x})$ operator is now given by

[1] 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, 1961, pp. 134-137; Gross, Runge & Heinonen, 1991, p. 21).

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

With this formulation there are no negative-energy states (identified as the 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 energy-momentum operator

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

and 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 the quanta identified as electrons and $n^+(p)$ is the number of quanta identified as positrons (Jauch & Rohrlich, 1976, p. 64). As Jordan proposed, the quantization of charge and subsequent emergence of a particle-like concept of an electron can be seen as a result of the quantization of the classical field.

In Heisenberg and Pauli's work, using the Lagrangian formalism, the system of interacting Dirac and Maxwell fields is described by a Lagrangian $L = L_0(A_\mu) + L_0(\psi) + L_I(\psi, A_\mu)$, where $L_0(\psi)$ and $L_0(A_\mu)$ are the Lagrangian for the free Dirac and Maxwell fields, and $L_I(\psi, A_\mu) = e\overline{\psi}\gamma^\mu\psi A_\mu$ is the interaction term. The variation of the Lagrangian with respect to the field operators results in a set of coupled equations. In this case, the Dirac equation in the presence of the electromagnetic field is given by $i(\nabla\!\!\!/ - eA\!\!\!/)\psi = m\psi$ (Feynman, 1961, p. 56). This apparently integrated treatment of the Dirac and Maxwell fields must not make us forget that the fields are quantized as free fields. From the quantization of each field it will turn out that the Lagrangian (or Hamiltonian) for each field and the interaction term become operators. It is important to keep in mind that quantum electrodynamics, as an interacting field theory, is developed as a theory that describes the interaction between distinct fields. Quantum electrodynamics was developed by considering two clearly separated fields from the start, which corresponds to having in the Lagrangian (the mathematical core of the theory) two separated pieces exclusive of each field. The relations between the two fields are determined by an interaction term that couples the Maxwell and Dirac equations.

3. Combining results from the different views on Dirac's equation

3.1. A quantized electron-positron field view on the solutions of the Dirac equation as a relativistic one-electron equation

When considering the exact solution of the one-electron Dirac equation in a central potential (the hydrogen atom) 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. As we

have seen, the quantized field perspective relates the (free) positive-energy solutions to the electrons and the (free) negative-energy solutions to the 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 model 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 in the presence of electromagnetic coupling $\psi = (\psi_A\ \psi_B)$, in the non-relativistic limit the lower two components ψ_B are smaller than the upper two ψ_A. When calculating matrix elements like $(\psi, \gamma_4\psi) = \psi_A^*\psi_A - \psi_B^*\psi_B$, neglecting terms of 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 (Mandl, 1957, pp. 214-215). 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, Leslie Foldy and Siegfried 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" (Foldy & Wouthuysen, 1949, p. 29). Foldy and Wouthuysen 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" (Foldy & Wouthuysen, 1949, p. 29).

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 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 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 Dirac one-electron equation in a way in which only positive-energy solutions are considered. Foldy and Wouthuysen applied their method to the case where a Dirac electron interacts with an external electromagnetic field. By making three canonical transformations and using only terms of order $(1/m)^2$ they obtained an 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 wave equation as in the Pauli theory.

3.2. A Dirac field approach to the hydrogen atom

In the previous method we are not taking into account directly the quantized Dirac field, and we are basically maintaining the inconsistent one-electron interpretation of

Dirac's equation. Also, Foldy and Wouthuysen definition of the positive-energy solutions was made by taking into account the classical limit, and this is not the only possible definition (Foldy & Wouthuysen, 1949, pp. 33-34). As we have seen the quantization of the (free) Dirac field leads to an association of the quanta to individual terms of a plane-wave expansion of the field corresponding to either positive or negative energy eigenvalues, which implies choosing another definition of positive-energy solutions.

This leads us to the necessity of taking a quantum field approach to the case of the hydrogen atom. One possibility is to use the so-called Furry, or bound interaction, representation within the external field approximation. This gives a method for calculating corrections to the energy levels of a bound electron determined by the Dirac equation as a relativistic one-electron equation (Berestetskii, Lifshitz & Pitaevskii, 1982, p. 487). But the starting point is the field operator defined by

$$\psi(x) = \sum_{r} u_r(\mathbf{x}) a_r e^{-iE_r t} + v_r(\mathbf{x}) b_r^* e^{iE_r t},$$

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 φ^μ a static external field (Jauch & Rohrlich, 1976, p. 313). 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 energy levels obtained by this method, and that in spite of identifying $u_a(\mathbf{x})$ as the electron's wave function, it contains both positive- and negative-energy components (Schweber, 1961, p. 566).

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.

The main working tool in quantum electrodynamics, the S-matrix, was designed for scattering problems where we have free particles in the beginning and free particles in the end of an interaction (scattering). This means that the S-matrix is not very appropriate to deal with the case of a bound particle, at least not in a direct way. Moreover, one of the most important characteristics of quantum field methods is that the interaction between fermions is represented by the exchange of photons: quanta of the electromagnetic field (e. g. Carson, 1996, pp. 127-129). If we make a model of the atom in which a classical Coulomb field gives the effect of the nucleus, this quanta view is lost (as in the external field method previously discussed). A way to overcome these difficulties is to address directly the two-body problem using the Bethe-Salpeter equation. In this method, the two-body problem is addressed by considering directly the two-particle propagator for an electron and a proton (which in the calculations is taken to be a 'big' positron with the same mass as the proton).[2] Considering a power series expansion of the two-particle propagator, the binding energy between a proton and an electron is basically calculated using what is known as the ladder approximation (Schweber, 1961, p. 713). Concerning this approximation, Hans Bethe and Edwin Salpeter remarked that "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

[2] There is an indirect method to calculate the energy levels of bound states from the S-matrix, which consists in determining the poles of the exact scattering amplitude. But in practice this approach leads to a summation of an infinite series of diagrams that corresponds to solving the Bethe-Salpeter equation (Berestetskii, Lifshitz & Pitaevskii, 1982, p. 554).

state an indefinite number of quanta may be exchanged *successively*. It is just such processes that the ladder-type graphs deal with" (Salpeter & Bethe, 1951, p. 1234). 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. We see that from a quantum field theory perspective, the description of the hydrogen atom (as a two-body problem) leads to a physical picture of the process going on in the atom quite different from the one obtained when using the Dirac equation as a one-electron equation. That is, we see, when going from a central potential approach to a two-body treatment of the hydrogen atom, the importance of the quanta concept in the description of interactions in quantum electrodynamics. The exchange of quanta is a basic characteristic of the description of physical processes in quantum electrodynamics.

4. From a theory of interacting fields to a theory describing the interaction between fields

The description of the interaction between the Dirac and Maxwell fields makes use of the Fock space of each field. In this section, I will restrict attention to scattering problems in which particles can be considered in a non-interacting state before and after the scattering process is over. The initial and final states of the scattered particles are states of the free-particle Fock space. The interaction process is described by the S-matrix perturbative approach in which only Fock states are considered. In the S-matrix approach we can use the interaction representation that enables an easy treatment of the interaction term as a perturbation to the free states of the fields. Let us consider, for example, the quantum electrodynamical treatment of the two-photon annihilation of an electron and a positron (Sakurai, 1967, pp. 204-208). The second-order S-matrix element is given by

$$
\begin{aligned}
S_{fi} = (-e)^2 \int d^4x_1 \int d^4x_2 \left\langle 2\gamma \left| A_\mu(x_1) A_\nu(x_2) \right| 0 \right\rangle (\gamma_\mu)_{\alpha\beta} (\gamma_\nu)_{\gamma\delta} \\
\times \Big[\left\langle 0 \left| \Psi^{(+)}_\beta(x_1) \overline{\Psi}^{(-)}_\gamma(x_2) \overline{\Psi}^{(+)}_\alpha(x_1) \overline{\Psi}^{(+)}_\delta(x_2) \right| e^- e^+ \right\rangle \theta(t_1 - t_2) \\
- \left\langle 0 \left| \Psi^{(+)}_\gamma(x_2) \overline{\Psi}^{(-)}_\beta(x_1) \overline{\Psi}^{(+)}_\alpha(x_1) \overline{\Psi}^{(+)}_\delta(x_2) \right| e^- e^+ \right\rangle \theta(t_2 - t_1) \Big].
\end{aligned}
$$

I will not go into the details of this expression but only address some relevant features for the case I am presenting. The expression is an example of the perturbative quantum electrodynamical treatment of the interaction between the Maxwell and Dirac fields. In it we have the description of the propagation of quanta between two space-time points x_1 and x_2, and components related to the non-interacting states of each field (Fock states). That is, we use only the Fock space for each field to calculate the interactions in quantum electrodynamics. Implicit in this expression is a subtle point about the perturbative procedure adopted in quantum electrodynamics: the idea that when far apart, particles do not interact and they can be described by states of the free-particle Fock space (disregarding their self-interaction). This idea can be given a more formal presentation by considering that the interaction between particles in a scattering process is adiabatically switched on in the remote past and adiabatically switched off in the remote future (Lippmann & Schwinger, 1950, p. 473; Bogoliubov & Shirkov, 1959, p. 197). It is the procedure of infinite past and infinite future separation (corresponding also to a spatial separation) of the particles – taken to be quanta of one of the fields (for example in the electron-electron scattering) or quanta of both fields (for example in the photon-electron scattering) – that permits a perturbative approach for the interactions as corrections to the free-particles states of each field by itself.

In the Fock space we have a direct connection of quanta to the normal modes of classical field configurations. It is from the Fourier splitting of a classical wave into positive and negative frequency normal modes, providing the basis for the configuration space, that the concept of quanta emerges (via a quantization procedure which results in associating a quanta to each normal mode). In the case of interacting fields it is not possible to make this Fourier expansion (Fraser 2006, p. 136). This implies that it is not possible to use the Fock representation for a free field to represent the interacting fields (Earman & Fraser, 2006, p. 322). Following this line of reasoning, this means that the quanta concept is unavailable when we consider fully interacting fields (Earman & Fraser, 2006, p. 330; Fraser, 2008, pp. 2-3). It could seem that the perturbative S-matrix approach could overcome this problem by providing a rigorous procedure to describe the interacting Dirac and Maxwell fields in the scattering problem by using only the Fock states representing free quanta from each field. According to the Haag theorem this is not possible.

From Haag's theorem (Haag, 1955) we know that we cannot have a unitary transformation that relates the field operators corresponding to the free Hamiltonian H and the interacting field Hamiltonian H_I. Considering that at t_0 the Heisenberg picture and the Dirac picture (interaction representation) coincide (Earman & Fraser, 2006, p. 320), it would seem that the state vector in the interaction representation, in the limit $t \to \pm\infty$, corresponds to free particles due to the fact that the interaction part of the Hamiltonian is negligible. But from Haag's theorem it seems that "at times $t = \pm\infty$, all the assumptions of the theorem hold for the Heisenberg representation, which represents an interaction, and for the interaction representation, which is a Fock representation for a free system" (Earman & Fraser, 2006, p. 322). In informal terms Haag's theorem implies that the state vectors in the interaction representation, that for $t \to \pm\infty$ are *supposed* to represent the free field, and the state vector in the Heisenberg representation for the interacting fields, are not in a common domain of both H and H_I (Schweber, 1961, p. 416).

From Haag's theorem we can conclude that if we have a free field at $t = -\infty$, the interaction representation describes also a free field at any time t_0. This means that we need to have a state of the fully interacting Hamiltonian from the start so that we can consistently give the interaction representation its usual interpretation as giving a different time dependency to the state vector and the operators (Schweber, 1961, p. 317).

Both the Heisenberg and Dirac pictures can hypothetically be used in either free or interacting systems, if we can separate the Hamiltonian in two parts. The change of representation does not change the physical situation whether it concerns free or interacting fields. There really is no "interaction picture's assumption that there is a time at which the representation for the interaction is unitarily equivalent to the Fock representation for a free system" (Fraser, 2006, p. 54).[3] This 'assumption' has nothing to do with the representation being used. The 'assumption' is that with an adiabatic

[3] Earman & Fraser's take it that in the infinite past and the infinite future "particles are assumed to be infinitely far apart and therefore not interacting" (Earman & Fraser, 2006, p. 321). In this case they consider that at $t=\pm\infty$ "the representation is taken to be a Fock representation" (p. 321). And this is taken to be a Dirac picture (interaction representation) assumption: "In the interaction picture ... at $t = \pm\infty$ the representation is the Fock representation for the free field" (p. 321). At this point I thus disagree with Earman & Fraser's presentation of the consequences of the Haag theorem. I simply do not think that the assumption that at infinite past we can take the state of a particular field to be a free state as described by a Fock space has to be seen as part of the assumptions of the Dirac picture. But this is a question of detail. The main aspect of their presentation is irrefutable: according to the Haag theorem the perturbative approach used in quantum electrodynamics to describe interactions is mathematically inconsistent.

switching on of the interaction, the state vector for the interacting fields can be constructed from the state vector of free fields (Schweber, 1961, p. 320):

$$\lim_{t\to-\infty} U(t_0,t)\varphi_a = \psi_a .$$

This is what is supposed to be achieved in the adiabatic switching on of the potential that 'connects' a free field Hamiltonian with the interacting field Hamiltonian (Jauch & Rohrlich, 1976, p. 134; Schweber, 1961, p. 322): $\lim_{t\to-\infty} H_\varepsilon(t) = \lim_{t\to-\infty}(H_0 + e^{-\varepsilon|t|}V(t)) = H_0$, where ε is a parameter that is taken to zero in the end of the computations. The point is that at infinite times before and after the adiabatic switching on/off of the interaction potential, the state vectors in either the Heisenberg or the Dirac picture are assumed to be describing free fields. The question here is not thus the representation being used but whether it is possible to connect the interacting state to a free field state. It seems clear from the consequences of Haag's theorem that the usual adiabatic switching on/off of the interaction will not do the trick.

It is well known that in the applications of quantum electrodynamics there are problems with divergent integrals. These problems are circumvented in practice with renormalization techniques in which (basically) all the divergent integrals appearing in the series expansion of the S-matrix are related to corrections to the mass and charge of the fermions. Since the value of the mass and the charge are not defined by the theory but result from measurements, the terms in the series expansion that are divergent (but formally should be smaller and smaller) are taken to be part of the observed mass and charge.

It might seem that by using renormalization techniques the consequences of Haag's theorem might be evaded because "once infinite renormalization counter terms are introduced, the interaction picture is not mathematically well-defined" (Fraser, 2006, p. 2). From this it might seem that "renormalization addresses this problem not by refining the assumptions, but by rendering the canonical framework mathematically ill-defined" (Fraser, 2006, p. 90). But it would be rather strange to say the least, that by considering an impoverished mathematical structure, suddenly, something physically equivalent to a unitary transformation connecting free and interacting field operators might emerge. In reality this argument does not provide any explanation for why the theory is effective, and this is a central question which has correctly been considered to be relevant: "there is, however, unfinished business in explaining why perturbation theory works as well as it does" (Earman & Fraser, 2006, p. 307).

If we followed Fraser's account we would be back where we started since, without taking into account the ill-defined mathematics of renormalization (and this is possible by considering only the lowest-order calculations), one can say that the adiabatic switching on/off of the interaction is also of doubtful mathematical rigor precisely due to the Haag theorem. This means that when imposing the adiabatic switching on/off we are developing an ill-defined approach. The primary question of the effectiveness of the theory would remain, independently of the probable impossibility of taking into account the Haag theorem in this ill-defined mathematical context (now due to the adiabatic switching on/off of the interaction).

It does not seem that the effectiveness of the perturbative approach, in spite of Haag's theorem, is a consequence of the need for a mathematical ill-defined mass and charge renormalization. At least there is no argument that shows how the efficiency of this scheme might result from developing it from a mathematical structure that is ill-defined due to renormalization. Moreover the renormalization technique is not the only

element that makes the theory mathematically ill-defined. It seems that we cannot arrive at any solid conclusions by analysing from a mathematical perspective, what appears to be an ill-defined mathematical structure.

My view is that, the explanation for the good results of the perturbation theory approach to the scattering (and bound state) problems in spite of the Haag theorem, will not be found in the mathematical structure of quantum electrodynamics. As will be seen below, the point about Haag's theorem is that the question of its applicability, is not even addressed in the way the physical theory is really used: *we have to consider the physical input assumptions of the theory together with its mathematical formulation and application, not the mathematical structure by itself.*

As mentioned in a previous section, the theory is developed from a canonical quantization of two independent classical fields. The description of the interaction between the fields is given, like in the classical counterpart, by an interaction term. Formally we can adopt whatever representation is mathematically possible. However, as I will show bellow, the use of the interaction representation is justified when we notice that, in the theory, we are considering different systems with an interaction that can be taken as a perturbation of their independent states. In this sense *the use of the interaction representation is clinched to the physical conceptual basis of the theory*. We simply use part of the Hamiltonian, which is possible since the theory was developed considering distinct parts, one corresponding to the free Dirac field, another to the free Maxwell field, and another to the interaction. This is the one pulled apart from the others in the interaction representation.

The description of scattering is developed from the theory considering an initial state corresponding to a limited number of free particles (quanta), and with an adiabatic switching on of the interaction between the fields, a fully interacting state ψ_A^+ is *apparently* obtained (ψ_A^+ would be an eigenstate of the interacting field Hamiltonian). The interacting state ψ_B^- that corresponds to a well-defined number of quanta in the final state is defined in an equivalent way. The scattering amplitude S_{AB} is given by (ψ_B^-, ψ_A^+) (Schweber, 1961, p. 323). The point is that we really do not work with these doubtful interacting states. What is going on is quite different. We are only considering a few terms of a perturbation expansion of the scattering matrix. We are always considering a weak interaction between distinct fields. For the description of their interaction it is not necessary to have a description of the fields as a closed interacting system. On the contrary, if we try to make a full description of the interaction considering all the terms of the power series expansion of the S-matrix, it can "at best only be an asymptotic expansion" (Schweber, 1961, p. 644).

According to Freeman Dyson, the series expansion of the S-matrix is divergent, and this has nothing to do with renormalization (Dyson, 1952b). That is, even if there were no divergent integrals appearing in the terms of the S-matrix, the series would still be divergent.[4] Dyson mentions that the "divergence will not prevent practical calculations being made with the series" (Dyson, 1952b, p. 632). However Dyson's view was that in a certain way the theory only provided a basis for developing the S-matrix series expansion (Cushing, 1986, p. 122). In Dyson's own words: "I had this rather positivistic view that all QED was the perturbative series. So if that failed you didn't really have a

[4] It is important to notice that Dyson's argument is a heuristic physical one, not a rigorous mathematical derivation. In my view this is not a deficiency of Dyson's argumentation, since we are considering a theory with an ill-defined mathematical structure. Also, even if a strict mathematical proof of the divergence of the S-matrix does not exist, further strong evidence in favour of Dyson's claim has been given in the last decades (e.g. Aramaki 1989, 91-92; West, 2000, 180-181; Jentschura, 2004, pp. 86-112; Caliceti et al, 2007, pp. 5-6).

theory" (quoted in Schweber, 1994, p. 565). This makes it even more imperative to justify the perturbative approach (which provides excellent agreement with experimental results), and to explain how to get rid of large-order terms of a divergent series that simply would make it impossible to use the theory.

Just from a mathematical perspective the use of only a few terms of a divergent series is difficult to defend, but by considering the input physical assumptions of the theory the justification of throwing away large-order terms will become clear. As we will see it is related to the weakness of the interaction between the Maxwell and Dirac fields (Mandl & Shaw, 1984, p. 95). *This by itself can explain why a few order terms in a perturbative approach can give so good results*. But it does not by itself justify throwing away terms that should be smaller and smaller but which will ultimately result in a divergent series.

As mentioned, the scattering amplitude S_{AB} can be written as (ψ_B^-, ψ_A^+), where we have $\lim_{t_0 \to -\infty} U(0,t_0)\varphi_A = \psi_A^+$ and $\lim_{t \to +\infty} U(0,t)\varphi_B = \psi_B^-$, where U is the time displacement unitary operator determined by solving the Schrödinger equation, φ_A and φ_B are the initial and final free states, and the in and out states ψ_A^+ and ψ_B^- are taken to be eigenstates of the full interacting Hamiltonian (Schweber, 1961, pp. 317-323). Formally these interacting states can be derived from the (complete) S-matrix. Since the S-matrix series expansion is divergent we can conclude that we cannot obtain these interacting states (Scharf, 1995, pp. 314-318), i.e. it is not even possible to find solutions of the coupled non-linear Maxwell-Lorentz and Dirac equations as (infinite) power series expansions in the coefficient *e* (the electric charge). To say in a different form, it is not possible to find a solution for a closed system of (electromagnetic) field and charged particles (Dirac field).

Let us recall that the theory was structured around the idea that the interaction between radiation and matter is weak. This we know from the fact that the coupling constant between the fields is small. We did not quantize the interacting system, but each field one by one; i.e. when we speak of weak interaction this has two related aspects: one is that the interaction term in the Lagrangian (or Hamiltonian) is small; the other is that from this we are justified to consider the quantization of radiation and matter separately as the quantization of free fields, and to take the interaction term as a perturbative one.[5] As Dyson mentioned, quantum electrodynamics "is in its nature a perturbation theory starting from the non-interacting fields as … unperturbed system[s]" (Dyson, 1952a, p. 79). Since we are starting from the non-interacting fields, we need to use the adiabatic switching on/off trick to 'connect' the two quantized systems and so to be able to describe their interaction. Now, as we have just seen, if we try to describe within the theory the full weak interaction between radiation and matter (corresponding to determine the eigenstates of the full interacting Hamiltonian) we get into trouble (i.e. we obtain divergent results).

We conclude then that, *in practice* (i.e. in the theory we really have and work with), the notion of weak interaction implies an intrinsic (essentially) approximate approach,

[5] It is important to remember that contrary to what formal presentations of the theory might led us to think (e.g. Dyson, 1952a, pp. 58-59; Källén, 1972, pp. 75-80), we do not start with a Lagrangian for an interacting system of radiation and matter and then due to practical problems in solving a set of coupled non-linear equations we feel forced to resort to perturbative calculations. On the contrary the theory was pretty much developed along two separated lines – one of then the quantization of the free electromagnetic field (a subject not addressed here) the other the development of a relativistic equation for the electron –, from the start taking into account, implicitly, that we were dealing with two clearly distinct weakly interacting physical entities: (quantized) charged particles and (quantized) electromagnetic field.

i.e. there is, *in practice*, an unbridgeable gap between the notion of weak interaction and the idea of a full (complete) description of the (weak) interaction (since a full description of the interaction would correspond to obtaining a complete expansion of the S-matrix, which is not possible). Thus the divergence of the S-matrix series expansion implies that we *are unable* to bridge the gap that exists between our starting physical assumption of two independent unperturbed systems and the (*ideal*) closed system of fully interacting radiation and matter; and there is a good reason for this.

As Earman & Fraser showed from a *formal* consideration of fully interacting fields (corresponding to an exact solution that we cannot obtain in the case of quantum electrodynamics), we cannot describe them in terms of the Fock representation for free fields. This means that, *formally*, for a closed system of interacting fields we cannot use all the physical input of quantum electrodynamics associated with the notion of weak interaction of radiation and matter (in particular our starting physical assumption of two non-interacting fields); i.e. the *formal* considerations imply an, *in principle*, *incompatibility* between the notion of *weak* interaction as implemented in the theory and the (formal) notion of a closed system of fully interacting fields. Also, as we have seen, considering the complete series expansion of the S-matrix would correspond to the description of a closed system of fully interacting fields. *Here the formal results are valid*, i.e. we would have two contradictory mathematical results. This is not the case since the series expansion of the S-matrix is divergent.

We see then that we disregard the large-order terms not simply for pragmatic reasons but for physical reasons. Including these terms would correspond to an improper use of the mathematical structure of the theory beyond its physical content (in a tentative description of a closed system of radiation and matter): in quantum electrodynamics we have the concepts of radiation and matter, and of a weak interaction between them, not of fully interacting fields.[6] At this point I make a *one-to-one correspondence between getting meaningful mathematical results and the way we set up the physical concepts in the theory*. In this case, (meaningful) approximate calculations of the *weak* interaction between different physical systems (radiation and matter).

I think that the situation we are facing here can be illuminated by recalling some of Bohr's views related to the Klein paradox and the problem with infinities in the theory (later addressed by renormalization). Regarding the Klein paradox, Bohr called the attention to the fact that it arises from not taking into account the elementary unit of electrical charge in the determination of the actual potential barrier, i.e. by not considering the physical concepts inscribed in the theory. According to Bohr, if calculations are made in which this fact and the limits in the determination of the electron's position are not taken into account, we would be facing the

> actual limit of applying the idea of potentials in connection with possible experimental arrangements. In fact, due to the existence of an elementary unit of electrical charge we cannot build up a potential barrier of any height and steepness desired without facing a definite atomic problem. (Quoted in Moyer, 1981, pp. 1057)

In a nutshell, to Bohr the paradox resulted from "an unlimited [mathematical] use of the concept of potentials in relativistic quantum mechanics" (quoted in Moyer, 1981, p. 1058; see also Darrigol, 1991, pp. 154-155).

[6] With this justification for disregarding the large-order terms of the S-matrix series expansion we are then in the situation of considering that the good empirical results of the theory follow simply – in the perturbative approach – from the weakness of the interaction between the two *separately* quantized fields.

An analogous situation occurs with the (renormalizable) infinities in the theory. According to Alexander Rueger presentation of Bohr's ideas,

> only for an [atomic] electron weakly interacting with the electromagnetic field could the radiation reaction, which would render the electron's orbit unstable, be ignored; as Bohr stressed repeatedly, strong interactions would make the idea of approximately stationary states of the electron in the atom impossible. (Rueger, 1992, pp. 317-318)

In these circumstances, Bohr recalls that "the whole attack on atomic problems … is an *essentially approximate procedure*, made possible only by the smallness of [the coupling constant]" (Bohr, 1932a, p. 378). That is, to Bohr

> the attempts to treat the radiation effects on rigorous lines by considering the atoms and the electromagnetic field as a closed quantum-mechanical system led to paradoxes arising from the appearance of an infinite energy of coupling between atoms and field (Bohr, 1932b, p. 66)

We have then, according to Bohr, that the physical conditions used to set up the theory imply an essentially approximate approach of an electron *weakly* interacting with the electromagnetic field. Like in the case of the Klein paradox if we extend the mathematical structure of the theory beyond its physical content we face a breakdown in the calculations. In my view, this is exactly the situation we are facing with the divergence of the S-matrix series expansion.

When trying to close the gap between two weakly interacting systems (described by an approximate approach) and fully interacting fields (corresponding to obtaining the exact solution for a closed quantized system of interacting fields), we face 'the actual limit of applying the idea' of non-interacting fields that are part of our physical input assumptions.

What we have then is the impossibility of defining within the theory a fully interacting state from the two fields that are defined and quantized as independent entities. Thus, in my view, we have a theory that is able, on an approximate level, to describe (using a few terms in perturbative calculations) the interaction between two separate fields, and not a theory describing as a whole a system of fully interacting fields. From a formal perspective the Haag theorem says that it is not possible to connect the separate fields with (fully) interacting fields when starting from the physical assumptions used to articulate the theory, i.e. from the notion of weak interaction as it is implemented in quantum electrodynamics.

In the theory the consequences of the Haag theorem are circumvented not because we are facing a "canonical framework mathematically ill-defined" (Fraser, 2006, p. 90) but because *we are not even trying to describe a system of (fully) interacting fields* (this eventual possibility is excluded *in practice* due to the divergence in the series expansion of the S-matrix and *in principle*, on formal grounds, by the above mentioned incompatibility between the notion of weak interaction as it is implemented in the theory and the formal results related to interacting fields); we are just tying to describe, by an 'essentially approximate procedure', the weak interaction between radiation and matter as distinct systems.[7] Thus, there is no conflict in quantum electrodynamics with the Haag theorem.

[7] However I agree with Fraser's view that we cannot apply Haag's theorem when working with a "canonical framework mathematically ill-defined" (Fraser, 2006, p. 90). What I do not agree with is Fraser's view that in some way renormalization addresses the problem of evading the consequences of Haag's theorem enabling the theory to be effective (also, as I mentioned, there are other factors which can be taken to render the theory mathematically ill-defined). That is because, in my view, in quantum electrodynamics, the problem of circumventing the Haag theorem is in reality the problem of explaining

Let us now go back to the problem of the role of the concept of quanta in the description of interactions within quantum electrodynamics. If, when calculating say the amplitude for electron-electron scattering, the complete S-matrix was (somehow) considered, then there would be an infinite number of terms corresponding to an infinite number of combinations of different quanta. One could say that in this case the quanta "type and number are not sharp" (Weingard, 1988, p. 46). The quanta description of interactions, as quanta exchange, would then appear to be a mathematical fiction due to the use of perturbation theory in the calculation of the scattering amplitude. However, when considering the scattering as really described in the theory, we can only use a few terms of the S-matrix expansion. There is simply no meaning to an exact S-matrix in the theory.[8]

As the divergence in the series expansion of the S-matrix shows, the Lagrangian of quantum electrodynamics does not provide us with the possibility of describing a system of interacting Dirac and Maxwell fields, but with the possibility to describe in an intrinsically approximate way the interaction between the two fields. The descriptions of interactions in the theory are based on the use of the Fock space for each field and the idea of (virtual) quanta exchange. There are no alternatives in quantum electrodynamics. From the start the theory was not developed to treat the question of fully interacting fields, but to treat the question of the interaction between distinct fields that are separately quantized. To consider that "Fock representations are generally inappropriate for interacting fields" (Earman & Fraser, 2006, p. 330), is, in the context of quantum electrodynamics, to turn upside down the theory as it was developed. The theory is built on top of the physical idea of independent entities whose interaction describes change in nature. When accepting this approach, and its intrinsic limitations, it is difficult to consider inappropriate, at least from an empirical point of view, the results of quantum electrodynamics; and so, contrary to Earman and Fraser's views, we can retain the concept of quanta in the description of interactions.

5. Conclusions

From a historical perspective, and by using the Dirac equation, I have tried to show how the concept of quanta emerges and what its role is in the description of interactions in quantum electrodynamics. To this end, I made a detour by not considering in detail the well-known case of the description of scattering processes in quantum electrodynamics, but the less analysed question of the quantum field theoretical description of a two-body system (like the hydrogen atom). This detour made it possible to connect the usual description of the hydrogen atom through the Dirac equation as a relativistic one-electron equation with the quantized Dirac field description (resulting from the quantization of a classical de Broglie wave described by the Dirac equation). Both in scattering and bound state problems, I think it becomes clear that the concept of quanta is fundamental in quantum electrodynamics since the physical processes are

---

how the theory gives good results while addressing the problem of the divergence of the S-matrix series expansion (necessary to justify the perturbative approach) and, at the same time, evading the consequences of Haag theorem. These are not separated problems. In reality, as we have seen, when addressing the first part of the problem, the consequences of Haag's theorem become irrelevant (independently of the fact that we are considering a mathematically ill-defined approach).

[8] There might appear to be ways of sidestepping this type of approach considering the Feynman path integral approach (Weingard, 1988, p. 54). But again, in the theory it is not possible to consider an infinite expansion of the transition amplitudes. In the mathematical expression for the transition amplitudes there are quanta propagators, and the interpretation of the propagators relating them to quanta cannot be overturned in a (finite expansion) calculation based on path integrals.

described as resulting from the exchange of virtual quanta between particles whose states are described using Fock spaces.

It turns out that this approach is not mathematically well-defined. In a nutshell, Haag's theorem implies that the perturbative approach used in quantum electrodynamics to treat the interaction between the quantized Maxwell and Dirac fields is not consistent. This brings up two big questions. How can the way the theory is implemented (giving a prominent role to the concept of quanta in the description of interactions) be justified? And even more importantly, "why perturbation theory works as well as it does" (Earman & Fraser, 2006, p. 307). Earman & Fraser do not provide any answer to this last question as regards quantum electrodynamics. And with respect to the first question, Fraser merely proposes to take shelter in the ill-defined mathematical structure of the theory to justify the inapplicability of the Haag theorem (Fraser, 2006, p. 90). However this argument has no explanatory power. It is an argument based solely on the mathematical structure of the theory (not taking into account its physical content) and it is not providing an explanation of the good results and soundness of the perturbative S-matrix approach.

According to Dyson the series expansion of the S-matrix, used in the description of scattering (and bound state) processes, is divergent. This indicates that the theory only provides a description of interactions using a few lower-order terms (which works well due to the small coupling constant between the fields). This means that the theory can only provide results if we are close to a free field situation. In other words, only when considering the interaction between two different fields as a small perturbation to their individual free states can the theory provide results in agreement with experiments.

We see that one has problems when trying to give a full description of the interaction, which corresponds to treating matter and radiation as one closed system. This would imply to go beyond the initial physical set up of the theory based on the idealization of totally non-interacting fields. In this way I think that quantum electrodynamics can be seen as providing only an approximate approach to the description of the interaction between two fields taken to be different physical systems. Only the lower-order terms of the series expansion can be kept. To take in the large-order terms would mean to disrupt the physical input assumptions provided by the implementation of the notion of weak interaction (i.e. the possibility of quantization of free fields and the description of their interaction perturbatively, by resort to the adiabatic switching on/off of the interaction). Due to the small coupling constant between the fields, the lower-order terms already provide good results.

From this perspective, how can Earman & Fraser's conundrum be solved while saving the use of quanta in the description of interactions? It is true that we are in an ill-defined mathematical context. However, we do not really need that to make the consequences of the theorem irrelevant in the theory. If we forget about Haag's theorem and set the machinery into motion, we face the situation that we cannot go from a free fields situation to a fully interacting fields situation (exactly as the Haag theorem says). This occurs because we are stretching the physical concepts too much and the calculations breakdown, i.e. we have a theory describing the weak interaction between different fields, not a theory describing fully interacting fields. In this way we are outside the scope of Haag's theorem.

Regarding the concept of quanta – which follows naturally from the quantization procedure – this is, as we have seen, a central concept in the quantum electrodynamical description of the weak interactions between the fields (as this description involves free-particle Fock spaces). However we are left with a tension regarding the concept of quanta. The point is that we start with the idealization of non-interacting fields, and, as

we have seen, we need the unphysical adiabatic switching on/off trick to set quantum electrodynamics as a perturbation theory into motion. The problem is that when addressing scattering problems, we take the particles (for example electrons) to be far apart before (and after) the scattering, and because of this not interacting (i.e. electrons as quanta of the Dirac field without self-interaction). In this way, we are describing the particles *observed* after the scattering process is over with the idealization of charged particles with a 'disconnected' charge, and this is rather unphysical (since implicit in the possibility of observation of an electron is the possibility of electromagnetic interaction with it).

References

Aramaki, S. (1989). Development of the renormalization theory in quantum electrodynamics (II). *Historia Scientiarum*, 37, 91-113.

Berestetskii, V. B., Lifshitz, E. M., & Pitaevskii, L. P. (1982). *Quantum electrodynamics*. Oxford: Butterworth-Heinemann.

Björken, J. D., & Drell, S. D. (1964). *Relativistic quantum mechanics*. New York: McGraw-Hill.

Bogoliubov, N. N., & Shirkov, D. V. (1959). *Introduction to the theory of quantized fields*. New York: Interscience Publishers.

Bohr, N. (1932a). Chemistry and the quantum theory of atomic constitution. In, J. Kalckar (Ed.), *Niels Bohr Collected works* (Vol. 6, pp. 371-408). Amsterdam: North-Holland.

Bohr, N. (1932b). Atomic stability and conservation laws. In Jan Faye & Henry J. Folse (Eds.), *The philosophical writings of Niels Bohr, Vol. 4* (1998, pp. 62-72). Woodbridge: Ox Bow Press.

Cao, T. Y. (1997). *Conceptual development of 20th century field theory*. Cambridge: Cambridge University Press.

Caliceti, E., Meyer-Hermann, M., Ribeca, P., Surzhykov, A., & Jentschura, U.D. (2007). From useful algorithms for slowly convergent series to physical predictions based on divergent perturbative expansions. *Physics Reports*, 446, 1-96.

Carson, C. (1996). The peculiar notion of exchange forces – II: from nuclear forces to QED, 1929-1950. *Studies In History and Philosophy of Modern Physics*, 27, 99-131.

Cushing, J. T. (1986). The importance of Heisenberg's matrix program for the theoretical high-energy physics of the 1950's. *Centaurus*, 29, 110-149.

Darrigol, O. (1986). The origin of quantized matter waves. *Historical Studies in the Physical and biological sciences*, 16/2, 198-253.

Darrigol, O. (1991). Cohérence et complétude de la mécanique quantique : l'exemple de «Bohr-Roselfeld». *Revue d'Histoire des Sciences*, 44, 137-179.

Dirac, P.A.M. (1926). On the theory of quantum mechanics. *Proceedings of the Royal Society of London A*, 112, 661-677.

Dirac, P.A.M. (1927). The quantum theory of the emission and absorption of radiation. *Proceedings of the Royal Society of London A*, 114, 243-265.

Dirac, P.A.M. (1928). The quantum theory of the electron. *Proceedings of the Royal Society of London A*, 117, 610-624.

Dirac, P.A.M. (1930). A theory of electrons and protons. *Proceedings of the Royal Society of London A*, 126, 360-365.

Dirac, P.A.M. (1931). Quantized singularities in the electromagnetic field. *Proceedings of the Royal Society of London A*, 133, 60-72.

Dirac, P. A. M. (1958). *The principles of quantum mechanics*. Oxford: Oxford University Press.

Dyson, F. J. (1952a). *Lecture notes on advanced quantum mechanics*. Cornell Laboratory of Nuclear Studies, Cornell University, Ithaca, N.Y.

Dyson, F. J. (1952b). Divergence of perturbation theory in quantum electrodynamics. *Physical Review*, 85, 631-632.

Earman, J., & Fraser, D. (2006). Haag's theorem and its implications for the foundations of quantum field theory. *Erkenntnis*, 64, 305-344.

Feynman, R. P. (1961). *Quantum electrodynamics*. Reading: Addison-Wesley.

Foldy, L. L., & Wouthuysen, S. A. (1949). On the Dirac theory of spin ½ particles and its non-relativistic limit. *Physical Review*, 78, 29-36.

Fraser, D. (2006). *Haag's theorem and the interpretation of quantum field theories with interaction*. Dissertation, University of Pittsburgh. Available at (http://etd.library.pitt.edu/ETD/available/etd-07042006-134120/).
Fraser, D. (2008). The fate of 'particles' in quantum field theories with interactions. *Studies in History and Philosophy of modern Physics*, 39, 841-859.
Gross, E. K. U., Runge, E., & Heinonen, O. (1991). *Many-particle theory*. Bristol: Adam Hilger.
Haag, R. (1955). On quantum field theories. *Det Kongelige Danske Videnskabernes Selskab, Matematisk-fysiske Meddelelser*, 29/12, 1-37.
Jauch, J. M., & Rohrlich, F. (1976). *The Theory of photons and electrons*. Berlin and New York: Springer-Verlag.
Jentschura, U. D. (2004). Quantum Electrodynamic Bound-State Calculations and Large-Order Perturbation Theory. *Arxiv preprint hep-ph/0306153*.
Källén, G. (1972). *Quantum electrodynamics*. Berlin and New York: Springer-Verlag.
Kragh, H. (1981). The genesis of Dirac's relativistic theory of electrons. *Archive for History of Exact Sciences*, 24, 31-67.
Kragh, H. (1990). *Dirac: a scientific biography*. Cambridge: Cambridge University Press.
Lippmann, B. A., & Schwinger, J. (1950). Variational principles for scattering processes I. *Physical Review*, 79, 469-480.
Mandl, F. (1957). *Quantum mechanics*. London: Butterworths.
Mandl, F., & Shaw, G. (1984). *Quantum field theory*. New York: Wiley.
Mehra, J., & Rechenberg, H. (2000). *The historical development of quantum theory*, Vol. 6. Berlin and New York: Springer-Verlag.
Miller, Arthur I. (1994). *Early quantum electrodynamics*. Cambridge: Cambridge University Press.
Oppenheimer, J. R. (1930). Two notes on the probability of radiative transition. *Physical Review*, 35, 939-947.
Moyer, D. F. (1981). Evaluations of Dirac's electron, 1928-1932. *American Journal of Physics*, 49, 1055-1062.
Pais, A. (1986). *Inward bound*. Oxford: Oxford University Press.
Rueger, A. (1992). Attitudes towards infinities: Responses to anomalies in quantum electrodynamics, 1927-1947. *Historical Studies in the Physical and Biological Sciences*, 22, 309-337.
Rugh, S. E., Zinkernagel, H., & Cao, T. Y. (1999). The Casimir effect and the interpretation of the vacuum. *Studies In History and Philosophy of Modern Physics*, 30, 111-139.
Salpeter, E. E., & Bethe, H. A. (1951). A relativistic equation for bound-state problems. *Physical Review*, 84, 1232-1242.
Sakurai, J. J. (1967). *Advanced quantum mechanics*. Reading: Addison-Wesley.
Scharf, G. (1995). *Finite quantum electrodynamics: the causal approach*. Berlin: Springer-Verlag.
Schweber, Silvan S. (1961). *An introduction to relativistic quantum field theory*. New York: Dover Publications.
Schweber, Silvan S. (1994). *QED and the men who made It: Dyson, Feynman, Schwinger, and Tomonaga*. Princeton: Princeton University Press.
Weingard, R. (1988). Virtual particles and the interpretation of quantum field theory. In. H. R. Brown, & R. Harré (Eds.), *Philosophical foundations of quantum field theory* (pp. 43-58). Oxford: Clarendon Press.
West, G. B. (2000). Perturbation theory, asymptotic series and the renormalisation group. *Physica A*, 279, 180-187.