Dirac's equation and the nature of quantum field theory

2012 Phys. Scr. 2012 014010

(http://iopscience.iop.org/1402-4896/2012/T151/014010)

IOP PUBLISHING PHYSICA SCRIPTA

Phys. Scr. **T151** (2012) 014010 (9pp) doi:10.1088/0031-8949/2012/T151/014010

Dirac's equation and the nature of quantum field theory

Arkady Plotnitsky

Theory and Cultural Studies Program, Purdue University, W Lafayette, IN 47907, USA

E-mail: plotnits@purdue.edu

Received 2 May 2012
Accepted for publication 27 June 2012
Published 30 November 2012
Online at stacks.iop.org/PhysScr/T151/014010

Abstract

This paper re-examines the key aspects of Dirac's derivation of his relativistic equation for the electron in order advance our understanding of the nature of quantum field theory. Dirac's derivation, the paper argues, follows the key principles behind Heisenberg's discovery of quantum mechanics, which, the paper also argues, transformed the nature of both theoretical and experimental physics vis-à-vis classical physics and relativity. However, the limit theory (a crucial consideration for both Dirac and Heisenberg) in the case of Dirac's theory was quantum mechanics, specifically, Schrödinger's equation, while in the case of quantum mechanics, in Heisenberg's version, the limit theory was classical mechanics. Dirac had to find a new equation, Dirac's equation, along with a new type of quantum variables, while Heisenberg, to find new theory, was able to use the equations of classical physics, applied to different, quantum-mechanical variables. In this respect, Dirac's task was more similar to that of Schrödinger in his work on his version of quantum mechanics. Dirac's equation reflects a more complex character of quantum electrodynamics or quantum field theory in general and of the corresponding (high-energy) experimental quantum physics vis-à-vis that of quantum mechanics and the (low-energy) experimental quantum physics. The final section examines this greater complexity and its implications for fundamental physics.

PACS numbers: 01.65.+g, 01.70.+w, 03.65.−w, 03.65.Ud, 03.65.TA, 03.70.+k

1. The mathematical correspondence principle, quantum variables and probability

Heisenberg's thinking leading him to his discovery of quantum mechanics was defined by three key elements, clearly apparent in his paper announcing his discovery (Heisenberg 1925). The same elements, I argue here, also defined Dirac's work on quantum electrodynamics, most especially his discovery of his relativistic equation for the electron, my primary concern in this paper. Dirac's derivation of the equation was significantly influenced by Heisenberg's paper. As is well known, Dirac read the paper very carefully in 1925, and it inspired his 1925–6 work on quantum mechanics and quantum electrodynamics, from which, especially his transformation theory (independently discovered by Jordan), his work on his equation grew. It is true that Dirac's work on his equation was significantly indebted to other developments in quantum mechanics, most especially Schrödinger's equation, the transformation theory, and Pauli's spin theory. Nevertheless, Heisenberg's thinking in his 1925 paper, and specifically the three key points in question, exerted a profound influence on Dirac's work on his equation. These points are as follows:

(1) *The mathematical correspondence principle.* Stemming from Bohr's correspondence principle, this principle states that one should maintain the consistency, 'correspondence,' between quantum mechanics and classical physics. More specifically, in the region, such as for large quantum numbers for electrons in atoms (when electrons are far away from the nucleus), where one could use classical physics in dealing with quantum processes, the predictions of classical and quantum theory should coincide. The correspondence principle, used more heuristically and ad hoc by Bohr and others before quantum mechanics, was also given a more rigorous, mathematical form by Heisenberg. The mathematical correspondence principle (here extended to quantum electrodynamics and other forms of quantum field theory) requires recovering the equations and variables of the old theory, classical mechanics in the case of quantum

mechanics and quantum mechanics in the case of quantum field theory, in the limit region where the old theory could be used. Dirac was aware of this version of the correspondence principle from his first paper on quantum mechanics on. As he said there: '*The correspondence between the quantum and classical theories lies not so much in the limiting agreement when* $h \Rightarrow 0$ *as in the fact that the mathematical operations on the two theories obey in many cases the same* [*formal*] *laws*' (Dirac 1925, p 315). Rather than becoming obsolete after quantum mechanics, as has been sometimes argued, the principle, in this mathematical form, has continued to play a major role in the development of quantum theory. It still does, for example, in string and brane theories, where the corresponding limit theory is quantum field theory.

(2) *The introduction of the new type variables.* Arguably most centrally, both discoveries, that of Heisenberg and that of Dirac, were characterized by the introduction of the new types of variables.

(*QM*) In the case of quantum mechanics, these were matrix variables with *complex* coefficients, essentially operators in Hilbert spaces over *complex* numbers (apparently, unavoidable in quantum mechanics) versus classical physical variables, which are differential functions of *real* variables. Heisenberg formally retained the equations of classical physics.

(*QED*) In the case of quantum electrodynamics, these were Dirac's spinors and multi-component wave functions, which, jointly, entail more complex operator variables, and a more complex structure of the corresponding Hilbert spaces, again, over complex numbers (apparently equally unavoidable in quantum field theory). In contrast to Heisenberg, Dirac also introduced a new equation, Dirac's equation, formally different from Schrödinger's equation, which, in accordance with the mathematical correspondence principle, is a far non-relativistic limit of Dirac's theory, via Pauli's theory, the immediate non-relativistic limit of Dirac's theory.

(3) *A probabilistically predictive character of the theory.* This change in *mathematical* variables was accompanied by a fundamental change in physics: the variables and equations of quantum mechanics and quantum electrodynamics no longer described, even by way of idealization, the properties and behavior of quantum objects themselves, in the way classical physics or relativity do for classical objects. Instead, the formalism only predicts the outcomes of events, in general probabilistically (even in the case of individual events), and of statistical correlations between some of these events, thus establishing the new type of relationships between mathematics and physics.

This last feature is, arguably, the most controversial feature of quantum theory, from quantum mechanics to quantum field theory, a feature famously unacceptable to Einstein. Nevertheless, the probabilistic character of quantum theory is in accordance with the observable experimental data. For, identically prepared quantum experiments (in terms of the condition of the apparatuses involved), in general lead to different recordings of their outcomes, which makes predicting these outcomes, unavoidably probabilistic, although, again, certain multiplicities of quantum events also exhibit statistical correlations (not found in classical physics). In some respects, quantum phenomena are more remarkable for these correlations than for the irreducible randomness of individual quantum events. Perhaps the greatest of many enigmas of quantum physics is how random individual events combine into (statistically) ordered multiplicities under certain conditions, such as, most famously, those of the EPR (Einstein–Podolsky–Rosen) type experiments, considered in Bell's theorem.

Heisenberg's revolutionary thinking established a new way of doing theoretical physics, and, as a consequence, it redefined experimental physics as well. The practice of experimental physics no longer consists, as in classical experiments, in tracking the independent behavior of the systems considered, but in *unavoidably* creating configurations of experimental technology that reflect the fact that what happens is *unavoidably* defined by what kinds of experiments we perform, how we affect quantum objects, rather than only by their independent behavior. My emphasis on 'unavoidably' reflects the fact that, while the behavior of classical physical objects is sometimes affected by experimental technology, in general we can observe classical physical objects, such as, planets moving around the sun, without appreciably affecting their behavior. This does not appear to be possible in quantum experiments. That identically prepared quantum experiments lead to different outcomes, thus making our predictions unavoidably probabilistic, appears to be correlative to the irreducible role of measuring instruments in quantum experiments (e.g., Bohr 1987, vol 1, p 93). The practice of theoretical physics no longer consists, as in classical physics or relativity, in offering an idealized mathematical description of quantum objects and their behavior. Instead it consists in developing mathematical machinery that is able to predict, in general (again, in accordance with what obtains in experiments) probabilistically, the outcomes of quantum events and of correlations between some of these events.

The situation takes a more radical form in quantum field theory and the experimental physics in the corresponding (high) energy regimes. Although I shall, in this paper, primarily discuss Dirac's theory, my argument could be extended to all forms of quantum field theory. While retaining, at least in the present view, Heisenberg's non-realist and non-causal epistemology of quantum mechanics, quantum field theory is characterized by, correlatively:

1. more complex configurations of phenomena observed and hence measuring apparatuses involved, and thus more complex configurations of effects of the interactions between quantum objects and measuring instruments;
2. a more complex nature of the mathematical formalism of theory, in part reflected in the necessity of renormalization (although renormalization will not be considered here);

3. a more complex character of the quantum-field-theoretical predictions and, hence, of the relationships between the mathematical formalism and the measuring instruments involved.

Before I proceed to a discussion of how this situation played itself out in and in part emerged from Dirac's relativistic theory of the electron, I would like to make a few brief qualifications concerning my assumptions in this paper. I assume, first, that quantum theory–quantum mechanics and quantum field theory (culminating in the standard model of particle physics)—is our best theory of the ultimate constitution of nature, considered apart from gravity. We do not thus far have workable quantum theories that incorporate gravity and, hence, establish proper connections between quantum theory and general relativity. All of the theories that we have thus far, such as string and brane theories, that aim to achieve this goal remain hypothetical. I see quantum theory as a probabilistically predictive and not descriptive and, also, indeed as a consequence, not causal, theory of quantum objects and their behavior. There are alternative interpretations of quantum mechanics and quantum field theory, which are thought by their proponents to be more acceptable, especially on the account of realism and causality. There are also alternative *theories* of quantum phenomena that are deemed preferable on the same grounds, such as Bohmian mechanics, which contains the Dirac-type equation for the electron. These alternatives and the seemingly interminable debates concerning the epistemology of quantum theory, will not, however, be considered here. I will only be concerned with the standard versions of quantum mechanics and quantum field theory, and of the non-realist and non-causal interpretations of these theories, given that both Heisenberg and Dirac adopted this type of interpretation. Finally, although Heisenberg's first paper on quantum mechanics was essential to Dirac's thinking and thus considering this paper would be helpful to my argument, I shall, given my space limit, bypass it, and move directly to Dirac's derivation of his relativistic equation for the electron.

2. The symmetry of space and time, and the calculus of spinors

The three key elements were, in Dirac's view, especially required for a *relativistic quantum* equation for a free electron, that is, an equation dealing with the higher levels of energy of the electron at which the effect of special relativity theory cannot be neglected in the way they can be in quantum mechanics. The first, *relativistic*, element is that time and space must enter symmetrically, which is required by relativity but which is not the case in Schrödinger's equation, which contains the first derivative of time and the second derivatives of coordinates. The second, *quantum-theoretical*, element is the first order derivative in time, an element required by quantum-theoretical considerations, captured by the quantum-mechanics formalism, specifically by Schrödinger's equation, and related to several other key features of the formalism. Among these features are the non-commutativity of certain quantum variables, linear superposition, and the conservation of the probability current (which entails positive definite probability density) and, correlatively, the probabilistic character of the predictions enabled by the formalism. The third key element was that, by the mathematical correspondence principle (as applied in quantum electrodynamics) the non-relativistic limit of a relativistic equation for the electron needed to be Schrödinger's equation. It follows, then, that to be both relativistic and quantum, a relativistic equation for the electron must be the first order differential equation in both space and time, since quantum theory requires the first order derivative in time, and relativity requires that space and time must enter symmetrically, and indeed that space and time must be interchangeable.

Although this seems simple enough, at least in retrospect, it appears that at the time only Dirac thought of the situation in this way. This thinking was greatly helped by the transformation theory (his 'darling,' as he called it), especially as concerned linearity in $\partial/\partial t$, and the positive definite probability density, both central to the transformation theory (Dirac 1962a). His famous conversation on the subject with Bohr at the time is revealing: 'Bohr: What are you working on? Dirac: I am trying to get a relativistic theory of the electron. Bohr: But Klein already solved that problem' (Dirac 1962b) Dirac, naturally, disagreed, and, for the reasons just explained, it is clear why he did, and why Bohr should have known better. The Klein–Gordon equation is relativistic and symmetrical in space and time, but it is not first-order in either, since both enter via the second derivative $\frac{\partial^2}{\partial t^2}$. One can derive the continuity equation from it, but the probability density is not positive definite. By the same token, the Klein–Gordon equation does not give us the correct equation, Schrödinger's equation

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\Psi = \hat{H}\Psi$$

in the non-relativistic limit. Schrödinger, who appears to be the first to have written down the Klein—Gordon equation in the process of his discovery of wave quantum mechanics, abandoned it in view of the incorrect predictions it gave in the non-relativistic limit. The roles of the first order derivative in time and of the probability-density considerations were not apparent to Schrödinger (who resisted the probabilistic view of his equation later on); they came into play in with Born's probabilistic interpretation of the wave function. Dirac's equation, on the other hand, does convert into Schrödinger's equation in the non-relativistic limit, which was, again, a crucial part of Dirac's thinking, following Heisenberg's approach in quantum mechanics. Technically, as its immediate non-relativistic limit, Dirac's theory converts into Pauli's spin-matrix theory, while Schrödinger's equation (which does not contain spin) was the limit of Pauli's theory, if one neglects spin. To summarize the conditions Dirac's equation had to fulfill:

(1) relativistic requirements, in particular the symmetry of space and time;
(2) to be the first-order linear differential equation in time, which means the first-order in both space *and* time by (1);
(3) the probability density must be positive definite and the probability current must be conserved;
(4) Schrödinger's equation should be the non-relativistic limit of the theory.

The conditions (2), (3) and (4) are correlative. None of these conditions is satisfied by the Klein–Gordon equation. This is not to say that the task of deriving the equation itself becomes simple once these requirements are in place; quite the contrary, this derivation involved highly original and non-trivial moves. Most of these moves were, I argue, parallel to those of Heisenberg in his paper introducing quantum mechanics. However, Dirac's mathematical task was more difficult because conditions (1), (2), (3) and (4), require both new variables, as in Heisenberg's scheme, and, in contrast to Heisenberg's scheme (which used the equations of classical mechanics), a new equation, in this respect similarly to Schrödinger's approach. In a way, Dirac's derivation of his equation combined Heisenberg's and Schrödinger's approaches, in the spirit of Dirac's transformation theory. As in Heisenberg, Dirac's new variables proved to be matrix-type variables, but of a more complex character, involving the so-called spinors and the multi-component wave function. The latter was a crucial concept, already discovered by Pauli in his non-relativistic theory of spin. As Heisenberg's matrices, Dirac's spinors had never been used in physics previously, although they were introduced in mathematics by Clifford about 50 years earlier (following the work of Hermann Grassmann on the so-called exterior algebras). But, just as Heisenberg in the case of matrices, Dirac was unaware of their existence and reinvented them in deriving his equation.

In spite of the elegance of its famous compact form, $i\gamma \cdot \delta\psi = m\psi$, reproduced on the plate in the Westminster Abbey commemorating Dirac, Dirac's equation encodes an extremely complex multi-component Hilbert-space machinery. It may also be noted that, unlike that of the Klein–Gordon equation, the Lorentz invariance of Dirac's equation is non-trivial, and it was surprising at the time, given the nature of the equation. Mathematically, the problem confronting Dirac may be seen as in terms of taking a square root of the Klein–Cordon equation (the solutions of which are, again, complex quantities, a mathematically crucial fact here), which also implies that every solution of Dirac's equation is a solution of the Klein–Gordon equation while the opposite is not true. Dirac used this fact in his derivation. I shall follow Dirac's paper because it reflects the key aspects of quantum-theoretical thinking that I want to address. First, however, I shall give a general summary. The equation, as introduced by Dirac, is

$$\left(\beta mc^2 + \sum_{k=1}^{3} \alpha_k p_k c\right) \psi(x,t) = i\hbar \frac{\partial \psi(x,t)}{\partial t}.$$

The new mathematical elements here, which never previously occurred in physics (quantum mechanics included), are the 4×4 matrices α_k and β and the four component wave function ψ. The Dirac matrices are all Hermitian,

$$\alpha_i^2 = \beta^2 = I_4$$

(I_4 is the identity matrix), and mutually anticommute:

$$\begin{gathered}\alpha_i \beta + \beta \alpha_i = 0,\\ \alpha_i \alpha_j + \alpha_j \alpha_i = 0.\end{gathered}$$

The above single symbolic equation unfolds into four coupled linear first-order partial differential equations for the four quantities that make up the wave function. The matrices form a basis of the corresponding Clifford algebra. Indeed, following Dirac's work and subsequent developments of quantum theory, one can think of Clifford algebras as *quantizations* of Grassmann's exterior algebras, in the same way that the Weyl algebra is a quantization of the symmetric algebra. $\boldsymbol{p}$ is the momentum operator in Schrödinger's sense, but in a more complicated Hilbert space than in the standard quantum mechanics. The wave function $\psi(t, \boldsymbol{x})$ takes value in a Hilbert space $X = C^4$ (Dirac's spinors are elements of X). For each t, $\psi(t, \boldsymbol{x})$ is an element of $H = L^2(R^3; X) = L^2(R^3) \otimes X = L^2(R^3) \otimes C^4$. I shall comment on the significance of this mathematical architecture below, merely noting now that it allows one to properly predict the probabilities of quantum-electrodynamical (high-energy) events, which have a greater complexity than quantum-mechanical (low-energy) events.

Dirac begins his paper by commenting on previous relativistic treatments of the electron, specifically the Klein–Gordon equation and its insufficiencies. He says in particular:

> [The Gordon–Klein approach] appears to be satisfactory as far as emission and absorption of radiation are concerned, but is not so general as the interpretation of the non-relativi[stic] quantum mechanics, which has been developed [specifically in Dirac's and Jordan's transformation theory] sufficiently to enable one to answer the question: What is the probability of any dynamical variables at any specified time having a value laying between any specified limits, when the system is represented by a given wave function ψ_n? The Gordon–Klein interpretation can answer such questions if they refer to the position of the electron ... but not if they refer to its momentum, or angular momentum, or any other dynamic variable. We would expect the interpretation of the relativi[stic] theory to be just as general as that the non-relativi[stic] theory. (Dirac 1928, pp 611–612)

The term 'interpretation' means here a mathematical representation of the physical situation (quantum-mechanical or quantum-electrodynamical), rather than, as is more common, a physical or philosophical interpretation of a given quantum formalism cum the phenomena it relates to. Dirac's statement does not mean that a physical description of quantum processes in space and time is provided, as against only predictions, in general, probabilistic, of the outcomes of quantum experiments. Dirac, as is clear from this passage, thought the capacity of a given theory to enable such predictions sufficient, and sufficiently general if such predictions are possible for any dynamic variable. Dirac then argues for the first order derivative in time, missing in the Klein–Gordon equation, as a proper starting point for the relativis[tic] theory of the electron. He says: 'The general interpretation of non-relativi[stic] quantum mechanics is based of the transformation theory, and is made possible by the wave equation being of the form

$$(H - W)\psi = 0, \tag{1}$$

i.e. being linear in W or $\frac{\partial}{\partial t}$, so that the wave function at any time determines the wave function at any later time. The wave function of the relativi[stic] theory must also be linear in W if the general interpretation is to be possible' (Dirac 1928, p 612). Before Dirac proceed to his derivation of his equation, he comments on 'the second difficulty' of the Klein–Gordon's equation:

> [The equation] refers equally well to an electron with charge e as to one with charge $-e$. If one considers for definitiveness the limiting case of large quantum numbers one would find that some of the solutions of the wave equation are wave packets moving in the way a particle of $-e$ would move on the classical theory, while other are wave packets moving in the way a particle with charge e would move classically. For this second class of solutions W has a negative value. One gets over the difficulty on the classical theory by arbitrarily excluding those solutions that have a negative W. One cannot do this on the quantum theory, since in general a perturbation will cause transitions from state with W positive to states with W negative. Such a transition would appear experimentally as the electron suddenly changes its charge from $-e$ to e, a phenomenon which has not been observed. The true relativi[stic] wave equation should thus be such than its solution split into two non-combining sets, referring respectively to the charge $-e$ and the charge e. ... In the present paper, we shall only be concerned with the removal of the first of these difficulties. The resulting theory is therefore still only an approximation, but it appears to be good enough to account for all the duplexity phenomena without arbitrary assumptions (p 612).

Dirac's theory, thus, inherits this second problem of the Klein–Gordon theory, because, as I said, mathematically every solution of Dirac's equation is a solution of the Klein–Gordon equation, of which Dirac's equation is essentially a squire root (the opposite is, again, not true). The difficulty that does not appear in the low-energy quantum regimes, and, one might add, it disappears at the low-energy limit of Dirac's theory, since his equation converts into Schrödinger's equations. This problem had ultimately proven to be a good thing. Dirac did not know this at the moment, but, as he will eventually have learned, the theory is much better because of this difficulty. That 'in general a perturbation will cause transitions from state with W positive to states with W negative,' and that 'such a transition would appear experimentally as the electron suddenly changes its charge from $-e$ to e' is what actually happens, and it will have been experimentally established in a year or so. Antimatter was staring right into Dirac's eyes, but it took a few years to realize that it is antimatter and that this type of transitions (eventually understood in terms of the creation and annihilation of particles, and virtual particle formation) defines high-energy regimes in quantum physics.

Dirac is now ready to present his derivation of his equation, guided by the two key ideas in question: the invariance under a Lorentz transformation and the equivalence of whatever the new one finds to Schrödinger's equation $(H-W)\ \psi = 0$ (equation (1) above) in the limit of large quantum numbers (p 613). In the case of the absence of the external field, which Dirac considers first and to which I shall restrict myself here, since it is sufficient for my main argument, equation the Klein–Gordon equation 'reduces to

$$(-p_0^2 + \mathbf{p}^2 + m^2c^2)\psi = 0, \tag{3}$$

if one puts

$$p_0 = \frac{W}{c} = \mathrm{i}h\frac{\partial}{c\partial t} \quad (\text{p } 613)'.$$

Next Dirac uses the symmetry between time, p_0, and space, p_1, p_2 and p_3, required by relativity, which implies that because the Hamiltonian one needs is linear in p_0, 'it must also be linear in p_1, p_2, and p_3.' He then says:

> [the necessary] wave equation is therefore in the form
>
> $$(p_0 + \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \beta)\psi = 0, \tag{4}$$
>
> where for the present all that is known about the dynamical variables or operators α_1, α_2, α_3, and β is that they are independent of p_0, p_1, p_2 and p_3, i.e. that they commute with t, x_1, x_2 and x_3. Since we are considering the case of a particle moving in empty space, so that all points in space are equivalent, we should expect the Hamiltonian not to involve t, x_1, x_2 and x_3. This means that α_1, α_2, α_3 and β are independent of t, x_1, x_2 and x_3, i.e. that they commute with of p_0, p_1, p_2 and p_3. We are therefore obliged to have other dynamical variables besides the co-ordinates and momenta of the electron, in order that α_1, α_2, α_3 and β may be functions of them. The wave function ψ must then involve more variables that merely x_1, x_2, x_3 and t.
>
> Equation (4) leads to
>
> $$\begin{aligned}0 &= (-p_0 + \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \beta)(p_0 \alpha_1 p_1 + \alpha_2 p_2 \\ &\quad + \alpha_3 p_3 + \beta)\psi = \Big[-p_0^2 + \sum \alpha_1^2 p_1^2 + \Sigma(\alpha_1\alpha_2 \\ &\quad + \alpha_2\alpha_1) p_1 p_2 + \beta^2 + \Sigma\,(\alpha_1\beta + \beta\alpha_1\Big]\ \psi,\end{aligned} \tag{5}$$
>
> where Σ refers to cyclic permutation of the suffixes 1, 2, 3. (p 613)

Let us pose here to observe and admire Dirac's way of thinking, manifest in this passage and throughout his derivation of his equation. This is 'very Dirac'—mathematically elegant and physically profound, for example, in the remarkable and far reaching conclusion that 'the wave function ψ must then involve more variables than merely x_1, x_2, x_3, t.' Taking advantage of non-commutativity in (5), one of Dirac's fortes, is worth a special notice. Equation (4), a square root of (3), the Klein–Gordon equation (a form of 'consistency' Dirac uses throughout the paper), is already Dirac's equation in abstract algebraic terms. One will now need to find α_n and β, to find the actual form of the equation. Physics as much follows mathematics as mathematics physics. While, as I argue, inspired by Heisenberg's thinking, Dirac's more formal approach is different from Heisenberg's calculations, equally profound physically but more straightforward and cumbersome, in his paper; and unlike in Dirac, in Heisenberg's mathematics nearly always follows physics. This is not say that mathematics is less important for Heisenberg, but only that one does not find in Heisenberg the same kind of, to use Dirac's word, *play* with of abstract structures that one finds in Dirac (Dirac 1962a, b). Dirac's earlier work on q-numbers quantum-mechanics formalism manifested the same use of the power of mathematical formalization. Dirac proceeds as follows:

> [Equation (5)] agrees with (3) [the Klein–Gordon equation in the absence of the external field
>
> $$(-p_0^2+\mathbf{p}^2+m^2c^2)\psi=0] \text{ if}$$
> $$\alpha_r=1, \alpha_r\alpha_s+\alpha_s\alpha_r=0\ (r\neq s)\ r,s\ =\ 1,2,3$$
> $$\beta^2=m^2c^2, \quad \alpha_r\beta+\beta\alpha_r=0.$$
>
> If we put $\beta=\alpha_4 mc$, these conditions become
>
> $$\alpha_\mu^2=1, \alpha_\mu\alpha_\nu+\alpha_\nu\alpha_\mu=0(\mu\neq\nu), \quad \mu,\nu=1,2,3,4. \qquad (6) \quad (\text{p } 613)$$

Thus, as I said, there is also a partial mathematical correspondence with the Klein–Gordon equation (that between a function of a complex variables and its square root), which allows Dirac to derive certain necessary algebraic conditions upon α_μ and β. Dirac will now state that 'we can suppose α_μ's to be expressed in some matrix scheme, the matrix elements of α_μ being, say, α_μ $(\zeta'\ \zeta'')$' (p 613). This supposition is not surprising given both the formal mathematical considerations (such anti-commuting relations between them) and the preceding history of matrix mechanics, including Dirac's own previous work. We know or may safely assume from Dirac's account of his work on his equations that matrix manipulation, 'playing with equations,' as he called it, was one of his starting points (Dirac 1962a, b). In addition, Pauli's theory, which is about to enter Dirac's argument, provided a handy example of a matrix scheme.[1] It is clear that matrix algebra of some sort is a good candidate for α_μ. Still, Dirac's ways of thinking is worth a further reflection. Dirac gets extraordinary mileage from considering the formal properties of the variables involved, even before considering what these variables actually are and as a way to gauging what they should be, which is his next step. He used the same approach—begin with the necessary formal properties and then find the actually variables—earlier in developing his q-number formalism. This is, as I said, both analogous to Heisenberg's approach and yet more formally oriented. That Heisenberg had to find new variables formally satisfying the classical equations may be seen as a partially formal task. However, the primary guidance for finding these variables was provided by certain experimentally established physical conditions (Bohr's energy rules and the Rydberg–Ritz frequency rules). In Dirac, while physical conditions of relativity and quantum mechanics do play a role, the primary driving force is the formal properties of variables and, especially, equations, in particular, the linearity and non-commutativity of the formalism. These, along with the role of complex numbers, are the defining mathematical properties of quantum theory. This approach was, again, used by Dirac in most of his work on quantum theory, especially his q-number formalism of quantum mechanics and the transformation theory.

Another remarkable consequence of the necessity of the particular matrix variables required by Dirac, appears next. (We still do not know what these variables actually are!) For if we indeed 'suppose α_μ's to be expressed in some matrix scheme, the matrix elements of α_μ being, say, $\alpha_\mu(\zeta'\ \zeta'')$,' then 'the wave function ψ must be a function of ζ as well as x_1, x_2, x_3 and t. The result of α_μ multiplied into ψ will be a function $(\alpha_\mu,\ \psi)$ of x_1, x_2, x_3, t, ζ defined by

$$(\alpha_\mu, \psi)\,(x,t,\zeta)\ =\ \Sigma_{\zeta'}\alpha_\mu(\zeta\zeta')\,\psi\,(x,t,\zeta')\text{'}(\text{p } 614).$$

Dirac is now ready 'for finding four matrices α_μ to satisfy the conditions (6),' those de facto forming the Clifford algebra, and for finding the actual form of variables that satisfy formal equation (4) or (5). Dirac considers first the three Pauli spin matrices, which satisfy the conditions (6), but not the equations (4) or (5), which needs four matrices. He says:

> We make use of the matrices
>
> $$\sigma_1=\begin{bmatrix}0&1\\1&0\end{bmatrix},\quad \sigma_2=\begin{bmatrix}0&-\mathrm{i}\\\mathrm{i}&0\end{bmatrix},\quad \sigma_3=\begin{bmatrix}1&0\\0&1\end{bmatrix},$$
>
> which Pauli introduced to describe the three components of the spin angular momentum. These matrices have just the properties
>
> $$\sigma_r^2=1\quad \sigma_r\sigma_s+\sigma_s\sigma_r=0\quad (r\neq s), \qquad (7)$$
>
> that we require for our α's. We cannot, however, just take the σ's to be the thereof our α's, because then it would not be possible to find the fourth. We must extend the σ's in a diagonal matter to bring in tow more rows and columns, so that we can introduced three more matrices ρ_1, ρ_2, ρ_3 of the same form as σ_1, σ_2, σ_3, but referring to different rows and columns, thus:
>
> $$\sigma_1=\begin{bmatrix}0&1&0&0\\1&0&0&0\\0&0&0&1\\0&0&1&0\end{bmatrix},\quad \sigma_2=\begin{bmatrix}0&-\mathrm{i}&0&0\\i&0&0&0\\0&0&0&-\mathrm{i}\\0&0&i&0\end{bmatrix},$$
> $$\sigma_3=\begin{bmatrix}1&0&0&0\\0&-1&0&0\\0&0&1&0\\0&0&0&-1\end{bmatrix}.$$
> $$\rho_1=\begin{bmatrix}0&0&1&0\\0&0&0&1\\1&0&0&0\\0&1&0&0\end{bmatrix},\quad \rho_2=\begin{bmatrix}0&0&-\mathrm{i}&0\\0&0&0&-\mathrm{i}\\i&0&0&0\\0&i&0&0\end{bmatrix},$$
> $$\rho_3=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{bmatrix}.$$
>
> The ρ's are obtained from σ's by interchanging the second and the third row, and the second and the third

[1] On Dirac and Pauli's theory, see (Kragh 1990 pp 55–56, 60).

columns. We now have, in addition to equations (7)

$$\rho_r^2 = 1 \qquad \rho_r\rho_s + \rho_s\rho_r = 0\,(r \neq s),$$
$$\text{and also} \quad \rho_r\sigma_t = \sigma_t\rho_r. \qquad (7') \quad (\text{p } 615)$$

These matrices are Dirac's great mathematical invention, parallel to Heisenberg's invention of his new matrix variables for quantum mechanics. Dirac's matrices form the basis of the corresponding Clifford algebra and define the mathematical architecture, mentioned above, where the multi-component relativistic wave function for the electron must defined. The entities they transformed are different from either vectors or tensors and are called spinors, introduced, as mathematical objects, by Cartan in 1913. Pauli, again, introduced the two-component non-relativistic wave function, which was necessary to incorporate spin, but he did so phenomenologically, rather than from the first principles, as was done by Dirac.

The rest of the derivation of Dirac's equation is a nearly routine exercise, with a few elegant but easy matrix manipulations. Dirac still needs to prove the relativistic invariance and the conservation of the probability current, and to consider the case of the external field, none of which is automatic, but is the standard textbook material at this point. The most fundamental and profound aspects of Dirac's thinking are contained in the parts of his paper just discussed, and I shall in the remainder of this paper considers the implications of Dirac's mathematical architecture for physics. I close this section by summing up my discussion thus far. Dirac uses the key elements of Heisenberg's approach to quantum mechanics:

(1) the mathematical correspondence principle, with Pauli's theory (the immediate limit) and Schrödinger's equation (the far limit) as the non-relativistic limits of his theory;
(2) the introduction of a new type of matrix variables, and spinors on which these matrices act; while Dirac is, again, helped here by Pauli's theory, he displays a highly original and non-trivial way of mathematical thinking;
(3) the assumption of the probabilistically predictive and not descriptive character of the theory he aims to construct.

The main difference from Heisenberg's discovery of quantum mechanics is:

(4) the appearance of a new equation, and of a *new type* of equation.

A few qualifications may be in order concerning Pauli's theory and Schrödinger's equation as the quantum-mechanical limits of Dirac's theory. Given the preceding discussion, it is easy to surmise that Pauli's theory and, if one further neglects spin, Schrödinger's equation, as itself the limit of Pauli's theory, would appear as the non-relativistic limits (immediate and far) of Dirac's theory. Pauli's matrices are contained in Dirac's matrices, and one can, accordingly, split Dirac's matrices (or the corresponding spinors) into small and large components. One does need a few calculations to rigorously establish this correspondence, but they are more or less straightforward. What is important is that, while Pauli's argument was phenomenological, Dirac had a theoretical argument from the first principles, which contained spin (also for positrons) and which also suggested that spin might have been the consequence of bringing together quantum mechanics to relativity. Dirac's theory also explained or at least more rigorously justified, the irreducible role of complex numbers and the necessity of a complex wave function in quantum theory from the geometry of the relativistic space-time via his spinor algebra. This role appeared phenomenologically, almost mysteriously, in quantum mechanics. It is important, however, that splitting the Dirac spinors and matrices into large and small components only applies at a low-energy approximation. Rigorously, one needs all of them. It is the whole composition of Dirac's scheme that reflects new physical phenomena found in the high-energy relativistic regime.

3. The architecture of mathematics and the architecture of matter in quantum field theory

Dirac's equation encodes a complex mathematical architecture. The Hilbert space associated with a given quantum systems in Dirac's theory is a tensor product of the infinite-dimensional Hilbert space (encoding the mathematics of continuous variables) and a finite-dimensional Hilbert space over complex numbers, which, in contrast to the two-dimensional (2D) Hilbert space of in Pauli's theory, C^2, is 4D in Dirac's theory, C^4. (Spin is contained by the theory automatically.) Dirac's wave function $\psi\,(t, \boldsymbol{x})$ takes value in a Hilbert space $X = C^4$ (Dirac's spinors are elements of X). For each t, $\psi\,(t, \boldsymbol{x})$ is an element of

$$H = L^2(R^3; X) = L^2(R^3) \otimes X = L^2(R^3) \otimes C^4.$$

Other forms of quantum field theory give this type of architecture an even greater complexity. It was, again, Dirac's move to 4D vectors and matrices that was especially crucial. Dirac's own recollection present this as merely an intuitive (why not?) guess: 'I suddenly 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?' (Dirac 1977, p 142). A flash of brilliant mathematical intuition it might have been, and it is difficult and perhaps ultimately impossible to account for such guesses. Some factors, however, may be plausibly conjectured to provide a meaningful insight into how some developments of mathematics and physics might have (one cannot be entirely certain here) shaped Dirac's thinking here. One can easily see several formal reasons for this move that have to do with the structure of the Clifford algebra and related abstract mathematical considerations, which would make four-by-four matrices a necessary move mathematically. Dirac appears not to have been aware of the Clifford algebras at the time, or, as, say, Weyl was, of the group-representational considerations (Lie groups and algebras). Dirac could have also easily seen that 3×3 matrices would not do; and he might have also been guided by the symmetry of his matrices as extending Pauli's matrices. 4×4 becomes quite natural, then. Dirac, however, was also familiar with some developments in abstract algebra that might have helped his guess. Thus, we know that he was familiar, even if relatively superficially, with non-commutative geometrical structures, which helped him, unlike say, initially Pauli, to accept the non-commutativity found in Heisenberg's theory and to realize its essential role in quantum mechanics.[2]

It is difficult to overestimate the significance of this mathematical architecture, which amounted to a very radical view of matter, physically especially manifest in the existence of antimatter. This architecture mathematically responds, and in fact led a discovery of, the following physical situation (again, keeping in mind that, just as that of quantum mechanics, Dirac's formalism only provides probabilities for the outcomes of quantum events, experimentally registered in the corresponding measuring technology).

Suppose that one arranges for an emission of an electron, at a given high energy, from a source and then performs a measurement at a certain distance from that source. Placing a photographic plate at this point would do. The probability of the outcome would be properly predicted by quantum electrodynamics. But what will be the outcome? The answer is not what our classical or even our quantum-mechanical intuition would expect, and this unexpected answer was a revolutionary discovery of quantum electrodynamics. Let us consider, first, what happens if we deal with a classical and then a quantum object in the same type of arrangement.

We can take as a model of the classical situation a small ball that hits a metal plate, which can be considered as either a position or a momentum measurement, or indeed a simultaneous measurement of both, at time t. In classical mechanics we can deal directly with the objects involved, rather than with their effects upon measuring instruments. The place of the collision could, at least in an idealized representation of the situation, be predicted exactly by classical mechanics, and we can repeat the experiment with the same outcome on an identical or even the same object. Most importantly, regardless of where we place the plate, we always find the same object, at least in a well-defined experimental situation, which is shielded from significant outside interferences, such as, for example, those that can deflect or even destroy the ball earlier.

By contrast, if we deal with an electron as a quantum object in the quantum-mechanical regime we cannot predict the place of collision exactly and, correlatively, exactly repeat the experiment on the same electron. Also correlatively, we cannot simultaneously predict, or measure, the position and the momentum of an electron, which makes the situation correlative to the uncertainty relations. Indeed, there is a non-zero probability that we will not observe such a collision at all, or that if we do, that a different electron (coming from somewhere else) is involved. It is also not possible to distinguish two observed traces as belonging to two difference objects of the same type. Unlike in the classical case, in dealing with quantum objects, there is no way to improve the conditions or the precision of the experiment to avoid this situation. Quantum mechanics, however, gives us correct probabilities for such events. Mathematically, this is accomplished by defining the corresponding Hilbert space, $H = L^2(R^3) \otimes C$ with the position and other operators as observables, and writing down Schrödinger's equation for the state vector $|\psi\rangle$ (in the case of pure state), and using Born's or similar rules to obtain the probabilities of possible outcomes.

Once the process occurs at a high energy and is governed by quantum electrodynamics, the situation is still different, even radically different. One might find, in the corresponding region, not only an electron, as in classical physics, or an electron or nothing, as in the quantum-mechanical regime, but also other particles: a positron, a photon, an electron–positron pair. Just as does quantum mechanics, quantum-electrodynamics, beginning with Dirac's equation, rigorously predicts which among such events can occur, and with what probability, and, in the present view it can only predict such probabilities, or statistical correlations between certain quantum events. In order to do so, however, the corresponding Hilbert-space machinery becomes much more complex, essentially making the wave function ψ a four-component Hilbert-space vector, as opposed to a one-component Hilbert-space vector, as in quantum mechanics. This Hilbert space is, as noted, $H = L^2(R^3; X) = L^2(R^3) \otimes X = L^2(R^3) \otimes C^4$ and the operators are defined accordingly. This structure naturally allows for a more complex structure of predictions (which are still probabilistic) corresponding to the situation just explained, usually considered in terms of virtual particle formation and Feynman's diagrams.

Once we move to still higher energies or different domains governed by quantum field theory the panoply of possible outcomes becomes much greater. The Hilbert spaces involved would be given a yet more complex structure, in relation to the appropriate Lie groups and their representations, defining (when these representations are irreducible) different elementary particles. In the case of Dirac's equation we only have electron, positron, and photon. It follows that in quantum field theory an investigation of a particular type of quantum object irreducibly involves not only other particles of the same type but also other types of particles. This qualification is important because the identity of particles within each type is strictly maintained in quantum field theory, as it is in quantum mechanics. In either theory one cannot distinguish different particles of the same type, such as electrons. One can never be certain that one encounters the same electron in the experiment just described even in the quantum-mechanical situation, although the probability that it would in fact be a different electron is low in the quantum-mechanical regime in comparison to that in the regime of quantum electrodynamics. In quantum field theory, it is as if instead of identifiable moving objects and motions of the type studied in classical physics, we encounter a continuous emergence and disappearance, creation and annihilation, of particles from point to point, theoretically governed by the concept of virtual particle formation. The operators used to predict the probability of such events, are the creation and annihilation operators. This view, thus, clearly takes us beyond quantum mechanics. For, while the latter questions the applicability of classical concepts, such as objects (particles or waves) and motion, at the quantum level, it still preserves the identity of quantum objects.

The introduction of this new mathematical formalism, involving more complex Hilbert spaces and operator algebras, was a momentous event in the history of quantum physics,

[2] On this points see (Mehra and Rechenberg 2001, vol 4 pp 131–147) and (Plotnitsky 2009 pp 117–188).

comparable to that of Heisenberg's introduction of his matrix variables. To cite Bohr's assessment of Dirac's theory: 'Dirac's ingenious quantum theory of the electron offered a most striking illustration of the power and fertility of the general quantum-mechanical way of description. In the phenomena of creation and annihilation of electron pairs we have in fact to do with new fundamental features of atomicity, which are intimately connected with the non-classical aspects of quantum statistics expressed in the exclusion principle, and which have demanded a still more far-reaching [than in quantum mechanics] renunciation of explanation in terms of a pictorial representation [of the type found in classical physics]' (Bohr 1987, vol 2, p 63). Heisenberg was even more emphatic. He saw Dirac's theory as an even more radical revolution than quantum mechanics was. In the early 1970s, Heisenberg, who made major contributions to quantum field theory in the meantime, spoke of Dirac's discovery of antimatter as 'perhaps the biggest change of all the big changes in physics of our century . . . because it changed our whole picture of matter. . . . It was one of the most spectacular consequence of Dirac's discovery that the old concept of the elementary particle [based on their stable identity] collapsed completely' (Heisenberg 1989, pp 31–33).

Quantum field theory made remarkable progress since its introduction or since Heisenberg's remark, a progress resulting, for example, in the electroweak unification and the quark model of nuclear forces, developments that commenced around the time of these remarks. Many predictions of the theory, from quarks to electroweak bosons and the concept of confinement and asymptotic freedom, to name just a few, were spectacular, and, since its introduction, the field has garnered arguably the greatest number of Nobel Prizes in physics. It was also quantum field theory that led to string and then brane theories, the current stratosphere of theoretical physics. However, the essential mathematical and experimental architecture of the theory, as considered here, have remained in place. Quantum mechanics and then higher-level quantum theories continue classical physics insofar as it is, just as classical physics, from Galileo on, and then relativity have been, the experimental-mathematical science of nature. On the other hand, quantum theory, at least in the interpretations of the type discussed here, breaks with both classical physics and relativity by establishing radically new relationships between mathematics and physics, or mathematics and nature. The mathematics of quantum theory is able to predict correctly the experimental data in question without offering a description of the physical processes responsible for these data. This is of course remarkable, and as a number of physicists, beginning with Heisenberg (Heisenberg 1930, p 11), have noted, we have been extraordinarily lucky that our mathematics works, that nature responds to our mathematical physics, in the absence of such a description, thus allowing us to bring mathematics and physics together in a new way, vis-à-vis classical physics or relativity. Many, beginning, again, with Einstein, have found this epistemological situation deeply unsatisfactory and even disturbing; and the debate beginning the debates concerning the epistemology of quantum theory have never subsided or lost any of their intensity. It is not my aim to enter these debates here. It is conceivable that the future development of fundamental physics will bring about a more classical (realist and causal) alternatives, as Einstein hoped, although a more radical departure from classical epistemology than that enacted by quantum mechanics and quantum field theory (in the interpretation discussed here) is not inconceivable either. Instead, I would like to close by noting that the 'miracles' of quantum theory are far from being a matter of luck alone. They are, just as all of the best physics, from Galileo on, the products of extraordinary creative thinking in theoretical physics, such as that of Heisenberg and Dirac, or in experimental physics, from Geissler's vacuum tube and Rühmnikorff's coil at the birth of particle physics in the mid-nineteenth century to the Linear Hadron Collider, designed to test the limits of quantum field theory and perhaps give rise to new theoretical physics. Will it be some form of quantum field theory, or of string or brane theory, or of loop quantum gravity, to name some currently prominent candidates, or will it be something else, perhaps something yet unheard of? It is difficult to predict. One could be reasonably certain, however, that discovering this physics and making it work will require some equally extraordinary creative thinking, and probably quite a bit of luck as well.

References

Bohr N 1987 *The Philosophical Writings of Niels Bohr* (Woodbridge, CT: Ox Bow Press)

Dirac P A M 1925 The fundamental equations of quantum mechanics *Sources of Quantum Mechanics* ed B L van der Warden (New York: Dover; 1968) pp 307–20

Dirac P A M 1928 The quantum theory of the electron *Proc. R. Soc.* A **177** 610

Dirac P A M 1962a *Hungarian Academy of Science Report* KFKI-1997-62

Dirac P A M 1962b Interview with T. Kuhn, April 1, 1962 *Niels Bohr Archive* (Copenhagen, and College Park, MD: American Institute of Physics) online at http://www.aip.org/history/ohilist

Dirac P A M 1977 Recollections of an exciting era *History of the Twentieth-Century Physics* ed C Weiner (New York: Academic) pp 109–46

Heisenberg W 1925 Quantum-theoretical re-interpretation of kinematical and mechanical relations *Sources of Quantum Mechanics* ed B L Van der Waerden (New York: Dover; 1968) pp 261–77

Heisenberg W 1930 *The Physical Principles of the Quantum Theory* translated by Eckhart K and Hoyt F C (New York: Dover; repreinted in 1949)

Heisenberg W 1989 *Encounters with Einstein, and Other Essays on People, Places, and Particles* (Princeton, NJ: Princeton University Press)

Kragh H 1990 *Dirac: A Scientific Biography* (Cambridge: Cambridge University Press)

Mehra J and Rechenberg H 2001 *The Historical Development of Quantum Theory* (Berlin: Springer)

Plotnitsky A 2009 *Epistemology and Probability: Bohr, Heisenberg, Schrödinger and the Nature of Quantum-Theoretical Thinking* (New York: Springer)