Second-order wave equation for spin- $\frac{1}{2}$ fields

Nicola Cufaro Petroni

Istituto Nazionale di Fisica Nucleare e Dipartimento di Fisica dell'Università, 70126 Bari, Italy

Philippe Gueret

Institut de Mathématiques Pures et Appliées, Université P. et M. Curie, 4, Pl. Jussieu, 75230 Paris, France

Jean Pierre Vigier and Anastasios Kyprianidis

Institut H. Poincaré, Laboratoire de Physique Théorique, 11, rue P. et M. Curie, 75231 Paris, France (Received 26 November 1984)

Some features of the second-order Feynman—Gell-Mann wave equation are discussed in order to show that (1) a statistical interpretation in terms of particles and antiparticles is possible, (2) the Dirac condition must be considered too restrictive, (3) no unwanted results arise from the use of this equation.

I. INTRODUCTION

In a series of recent papers¹ the authors proposed a theory of spin- $\frac{1}{2}$ particles based on a second-order wave equation: the so-called four-component Feynman—Gell-Mann equation² that, in the presence of an external electromagnetic field A_{μ} , takes the following form ($\hbar = c = 1$):

$$\begin{split} &[(i\partial - eA)(i\partial - eA) - m^2]\psi \\ &= &[(i\partial_{\mu} - eA_{\mu})(i\partial^{\mu} - eA^{\mu}) - \frac{1}{2}eF_{\mu\nu}\sigma^{\mu\nu} - m^2]\psi = 0 \;. \end{split}$$

Beyond a series of well-known remarks² about the use of (1.1), the main reasons that induced the authors to adopt it instead of the Dirac equation

$$(i\partial - eA - m)\psi = 0 \tag{1.2}$$

are the following.

- (a) If we are looking for a causal interpretation of the quantum equations, a classical analogy can be found only starting from a second-order differential equation.
- (b) If we want to interpret the appearance of a quantum potential^{1,3} in the classical equations as the global effect of a stochastic process induced on the particle by a subquantum medium, namely, the so-called Dirac aether,⁴ we must use only second-order differential equations.⁵
- (c) Both of the preceding steps are essential in a coherent causal physical interpretation⁶ of the nonlocal quantum effects suggested by the experiments⁷ on the Einstein-Podolsky-Rosen (EPR) paradox.⁸

The present paper is devoted to a deeper examination of the problems connected with the use of (1.1) and in particular to an analysis of the features of the conserved density. First of all we briefly discuss in Sec. II the relations between the sets of solutions of (1.1) and (1.2). In Sec. III we show that the nonpositive conserved densities obtained from (1.1) are differences of positive densities connected with the presence of particles and antiparticles and that a statistical interpretation is possible. In Sec. IV we prove

that the Dirac equation (1.2), considered as a positivity condition for the densities, is a sufficient but unnecessary condition. Finally, in Sec. V, we analyze the problem if the presence in (1.1) of more solutions than in (1.2) can give way to some unphysical prediction.

II. PARTICLES AND ANTIPARTICLES

Let us start with a discussion of the relations between the four-spinors ψ solutions of (1.1) and those solutions of (1.2). From this standpoint we can define the following sets:

$$\mathcal{F} = \{ \psi \mid (I - \mathcal{D}^2)\psi = 0 \} ,$$

$$\mathcal{D}_{\pm} = \{ \psi \mid (I \pm \mathcal{D})\psi = 0 \}$$
(2.1)

with

$$D_{\mu} = \frac{1}{m} (i\partial_{\mu} - eA_{\mu}) . \tag{2.2}$$

Of course \mathcal{F} is the set of the solutions of (1.1) and \mathcal{Q}_{-} is the set of the solutions of (1.2). Moreover, to show what \mathcal{Q}_{+} is, the following propositions can be proved:

$$P_1: \mathscr{D}_+ \cap \mathscr{D}_- = \{\psi = 0\} ,$$

$$P_2: \mathscr{D}_+ \cup \mathscr{D}_- \subset \mathscr{F}$$
,

$$P_3: \forall \psi_+ \in \mathscr{Q}_+ \exists \mid \psi_- \in \mathscr{Q}_- \ni '\psi_+ = \gamma_5 \psi_-$$

To prove P_1 we must only remark that if $\psi \in \mathcal{D}_+ \cap \mathcal{D}_-$ we will have at once

$$(I + D)\psi = 0, (I - D)\psi = 0$$
 (2.3)

that immediately give $\psi=0$. Furthermore, it can be shown that $\mathscr{D}_{\pm}\subseteq\mathscr{F}$: in fact, if $\psi\in\mathscr{D}_{\pm}$, we have also that

$$(I - \cancel{D}^2)\psi = (I \mp \cancel{D})(I \pm \cancel{D})\psi = 0, \qquad (2.4)$$

namely that $\psi \in \mathcal{F}$. As a consequence, $\mathcal{Q}_+ \cup \mathcal{Q}_- \subseteq \mathcal{F}$.

However, if $\psi_{+} \in \mathscr{D}_{+}$ and $\psi_{-} \in \mathscr{D}_{-}$, each linear combination $\psi = a\psi_{+} + b\psi_{-}$ is an element of \mathscr{F} , but, because of P_{1} , it cannot belong to \mathscr{D}_{+} or \mathscr{D}_{-} . Hence $\mathscr{D}_{+} \cup \mathscr{D}_{-} \neq \mathscr{F}$ that completely proves P_{2} .

To prove P_3 we remark that, if $\psi_+ \in \mathcal{Q}_+$, $\psi_- = \gamma_5 \psi_+$ belongs to \mathcal{Q}_- since

$$(I - D)\psi_{-} = \gamma_{5}(I + D)\psi_{+} = 0.$$
 (2.5)

Moreover $\gamma_5\psi_-=\psi_+$, so that $\psi_-=\gamma_5\psi_+$ is exactly the spinor that we look for. Furthermore, if it would be possible to find two spinors ψ_- and ψ'_- , such that $\gamma_5\psi_-=\gamma_5\psi'_-=\psi_+$, we should also admit that $\gamma_5(\psi_--\psi'_-)=0$, or equivalently that $\psi_-=\psi'_-$.

By recalling what γ_5 represents for the symmetry operations on spinors, we can conclude that \mathcal{D}_+ contains the antiparticle wave functions moving backward in space-time and with the sign of energy inverted with respect to the particle wave functions [solutions of the Dirac equation (1.2)] belonging to \mathcal{D}_- . In other words, we would say that \mathcal{F} contains both particle and antiparticle solutions and we will show now that it also contains all their superpositions. In fact it can be proved that

$$P_4$$
: $\forall \psi \in \mathcal{F} \exists \mid \psi_+ \in \mathcal{D}_+ \land \psi_- \in \mathcal{D}_- \ni \psi = \psi_+ + \psi_-$,

since it is very easy to see that, if $\psi \in \mathcal{F}$, the spinors $\psi_{\pm} = \frac{1}{2}(I \mp \mathcal{D})\psi$ belonging to \mathcal{D}_{\pm} are such that $\psi_{+} + \psi_{-} = \psi$.

III. CONSERVED DENSITY

The mixing of particles and antiparticles in the general solutions of (1.1) is particularly evident in the form of the conserved density. It is well known¹ that Eq. (1.1) can be deduced from the scalar Lagrangian density

$$\mathcal{L} = \overline{(i\partial - e\mathbf{A})}\overline{\psi}(i\partial - e\mathbf{A})\psi - m\overline{\psi}\psi \tag{3.1}$$

so that the conserved current density is

$$J_{\mu} = \frac{1}{m} \operatorname{Re}[\overline{\psi} \gamma_{\mu} (i\partial - eA) \psi] ; \qquad (3.2)$$

that, of course, does not coincide with the Dirac current $\overline{\psi}\gamma_{\mu}\psi$ unless the spinor ψ is a solution of (1.2). As a consequence the conserved density J_0 will not be positive definite. This feature, which is common to all the relativistic second-order quantum equations, forbids a direct statistical interpretation of (1.1) with the conserved density playing the role of a probability density.

The way out proposed by Dirac⁹ was, in some sense, the restriction of the physically acceptable states to the solutions of (1.2), so that the conserved density becomes $\psi^{\dagger}\psi \geq 0$. We will examine in the next section whether this Dirac condition can be considered general enough to contain all the physically meaningful solutions leading to a positive density. Here we will limit ourselves to analyze another interpretation of the appearance of nonpositive densities: J_0 is not a probability density, but an auxiliary function which obeys many relations we would expect from such a probability. In fact it behaves like an average charge density where, for mixtures of particles and antiparticles, "charge" must be understood in the widest

sense, i.e., as a certain property which distinguishes between particles and antiparticles which are identical to each other in all other respects (electric charge, baryon number, etc.). In this sense we can calculate averages of physical observables exactly as in the ordinary probability calculus, but using a "probability measure" which is not positive.

From this standpoint the connection between the non-positivity of J_0 and the mixing of particles and antiparticles is better understood if we consider that in the preceding section it was determined that each $\psi \in \mathcal{F}$ is a superposition of ψ_- and ψ_+ , namely of a particle and an antiparticle state. As a consequence, we get from P_4 that

$$J_{\mu} = \operatorname{Re}[(\overline{\psi}_{+} + \overline{\psi}_{-})\gamma_{\mu}\mathcal{D}(\psi_{+} + \psi_{-})]$$

$$= \operatorname{Re}[(\overline{\psi}_{+} + \overline{\psi}_{-})\gamma_{\mu}(-\psi_{+} + \psi_{-})]$$

$$= \overline{\psi}_{-}\gamma_{\mu}\psi_{-} - \overline{\psi}_{+}\gamma_{\mu}\psi_{+}, \qquad (3.3)$$

i.e., the conserved current J_{μ} is the difference of two Dirac-type conserved currents for particles and antiparticles. Moreover

$$J_0 = \psi_-^{\dagger} \psi_- - \psi_+^{\dagger} \psi_+ , \qquad (3.4)$$

which means that the nonpositive conserved density J_0 is always a difference of two positive Dirac densities for particles and antiparticles. The result (3.4) is perfectly coherent with the point of view that considers J_0 as a nonpositive measure, because a classical result of the measure theory¹¹ states that each real measure is the difference of two positive measures.

IV. POSITIVITY CONDITIONS FOR THE DENSITY

In this section we will show that the Dirac condition (1.2), which selects in \mathcal{F} solutions with $J_0 \geq 0$, is too restrictive in the sense that it is a sufficient but unnecessary condition. This result will be proved in a particular case by showing that, for the free fields, the more general positivity condition admits states which are not solutions of (1.2).

In fact, if $A_{\mu} = 0$, (1.1) becomes

$$(\Box + m^2)\psi = 0 , \qquad (4.1)$$

so that the general form of a plane-wave solution of (4.1) is

$$\psi_p(x) = Ne^{ip \cdot x}u , \qquad (4.2)$$

where N is a constant and u is an arbitrary four-spinor constant in space-time. Of course, in order to have a solution of (4.1), p_{μ} must satisfy the relation $p_{\mu}p^{\mu}=m^2$ so that there will be solutions with both positive and negative energies:

$$E = \pm m \left[1 + (\mathbf{p}/m)^2 \right]^{1/2}. \tag{4.3}$$

The main difference with the plane-wave solutions of the free Dirac equation is that no a priori restrictions on u such as⁹

$$(p\pm m)u=0 (4.4)$$

are requested, so that there is no a priori connection be-

tween the sign of $\overline{u}u$ and the sign of the energy.

However we will show that there is a condition, less restrictive than (4.4), connecting the sign of $\overline{u}u$ and the sign of the energy, that gives a positive conserved density.

First of all let us use spinors u normalized in the sense that (if $\overline{u}u$ is not identically zero)

$$\overline{u}u = \overline{\psi}\psi/N = \pm 1 . \tag{4.5}$$

We can now separate the solutions of (4.1) with positive and negative $\overline{u}u$ by introducing a label $\epsilon = \pm 1$ so that

$$\overline{u}_{\varepsilon}u_{\varepsilon} = \varepsilon$$
 (4.6)

By writing down our four-spinors u in terms of twocomponent elements, we get as a more general form

$$u_{\epsilon} = \begin{bmatrix} H_{\epsilon}(\alpha)\xi \\ H_{-\epsilon}(\alpha)\eta \end{bmatrix} \tag{4.7}$$

with

$$\xi^{\dagger} \xi = \eta^{\dagger} \eta = 1 ,$$

$$H_{\epsilon}(\alpha) = \frac{1}{2} (e^{\alpha} + \epsilon e^{-\alpha}), \quad \alpha \in [0, \infty] ,$$
(4.8)

where the two-component quantities ξ , η and the parameter α are fixed only by supplementary informations like polarization, etc. Now the free Lagrangian density is

$$\mathcal{L}_{0} = i \overline{\partial \psi} i \partial \psi - m^{2} \overline{\psi} \psi \tag{4.9}$$

so that the free conserved current for plane waves be-

$$j_{\mu} = \frac{1}{m} \operatorname{Re}(\overline{\psi}\gamma_{\mu}i\partial\psi) = \frac{1}{m} \operatorname{Re}(\overline{\psi}\gamma_{\mu}p\psi) = \frac{1}{m}p_{\mu}\overline{\psi}\psi$$
$$= \frac{1}{m}N^{2}p_{\mu}\overline{u}_{\epsilon}u_{\epsilon} = \frac{1}{m}\epsilon N^{2}p_{\mu} \qquad (4.10)$$

and the conserved density is

$$j_0 = \frac{1}{m} \epsilon N^2 E$$
 (4.11)

Of course the more general positivity condition for j_0 is

$$\epsilon = \operatorname{sgn}(E)$$
 (4.12)

and hence the states

$$\psi(x) = Ne^{ip \cdot x} u_{\epsilon = \operatorname{sgn}(E)} \tag{4.13}$$

always lead to a positive j_0 .

We can show now, in a simple particular case, that there are spinors of the form (4.13) which are not solutions of the free Dirac equation. In fact, if we consider the solutions (4.13) at rest, we get from (4.7) the following spinors:

with ξ and η arbitrary two-component quantities and $\alpha \in [0, \infty]$. On the other hand, the free solutions at rest of the Dirac equation are restricted by (4.4) so that we have

$$\psi_{+}^{D} = N_{+}e^{-imt} \begin{bmatrix} \xi \\ 0 \end{bmatrix} \quad (E = m \ge 0) ,$$

$$\psi_{-}^{D} = N_{-}e^{+imt} \begin{bmatrix} 0 \\ \eta \end{bmatrix} \quad (E = -m \le 0) .$$

$$(4.15)$$

In other words, only the value $\alpha = 0$ is possible for Dirac solutions at rest and hence the condition imposed by (4.4) must be considered sufficient but unnecessary in order to have positive conserved densities.

V. MOTT CROSS SECTION

In the preceding sections it was established that it cannot be forbidden in principle to use solutions of (1.1) which do not obey to the Dirac equation (1.2). However one could suspect that all the physically meaningful states be solutions of (1.2). In fact this standpoint could be supported by the trivial remark that Dirac's equation works. Therefore we will discuss at the end of this paper if the use of (1.1) can lead to some unwanted physical prediction. Of course an ultimate answer to this question is not yet possible, but we must remark that, in all known cases, the use of (1.1) leads to very reasonable results, from the domain of atomic physics¹² to that of the quantized theory of fields.2,13

Here we will limit ourselves to present, as an example, the outlines of the calculation of the Coulomb cross section to order e^2 (Mott cross section) starting from (1.1). If we consider an external electromagnetic field of the

$$A_0 = -\frac{Ze}{4\pi} \frac{1}{r}, \quad A_k = 0 \tag{5.1}$$

at the first order in e^2 , Eq. (1.1) becomes

$$(\Box + m^2)\psi = i\frac{Ze^2}{4\pi} \left[\frac{2}{r} \frac{\partial}{\partial t} - \frac{1}{r^3} \mathbf{r} \cdot \boldsymbol{\alpha} \right] \psi , \qquad (5.2)$$

where, if σ_k are the Pauli matrices,

$$\alpha = \begin{bmatrix} 0 & \sigma \\ \sigma & 0 \end{bmatrix} . \tag{5.3}$$

$$\psi_i = N_i e^{-ip_i \cdot x} u_i, \quad \psi_f = N_f e^{-ip_f \cdot x} u_f$$
 (5.4)

$$S_{fi} \propto \int d^4x \, \overline{\psi}_f \left[\frac{2}{r^3} \frac{\partial}{\partial t} - \frac{1}{r^3} \mathbf{r} \cdot \boldsymbol{\alpha} \right] \psi_i \propto \int d^3x \, \overline{u}_f \left[-\frac{2i}{r} E_i - \frac{1}{r^3} \mathbf{r} \cdot \boldsymbol{\alpha} \right] u_i e^{-i(\mathbf{p}_f - \mathbf{p}_i) \cdot \mathbf{r}} \delta(E_f - E_i)$$

$$\propto \frac{1}{a^2} \left[u_f (2E_i - \mathbf{q} \cdot \boldsymbol{\alpha}) u_i \right] \delta(E_f - E_i)$$
(5.5)

with $q = p_f - p_i$. We have now

$$\overline{u}_{f}(2E_{i} - \mathbf{q} \cdot \boldsymbol{\alpha})u_{i}\delta(E_{f} - E_{i}) = \overline{u}_{f}[\gamma_{0}(E_{f}\gamma_{0} - \mathbf{p}_{f} \cdot \boldsymbol{\gamma}) + (E_{i}\gamma_{0} - \mathbf{p}_{i} \cdot \boldsymbol{\gamma})\gamma_{0}]u_{i}\delta(E_{f} - E_{i})$$

$$= \overline{u}_{f}(\gamma_{0}\boldsymbol{p}_{f} + \boldsymbol{p}_{i}\gamma_{0})u_{i}\delta(E_{f} - E_{i}), \qquad (5.6)$$

and hence

$$S_{fi} \propto \frac{1}{q^2} \overline{u}_f (\gamma_0 p_f + p_i \gamma_0) u_i \delta(E_f - E_i) . \qquad (5.7)$$

In calculating the cross section we must use $|S_{fi}|^2$ and sum over the unobserved states. However, differently from the calculation based on the Dirac equation, there is no connection between the constant spinors u_f and u_i and the momenta p_f and p_i (see Sec. IV) and hence we must sum over the complete orthonormal sets of initial and final spinors $\{u_{i,\alpha}\}$ and $\{u_{f,\beta}\}$. Of course we will have

$$\sum_{\alpha} u_{i,\alpha} \overline{u}_{i,\alpha} = \sum_{\beta} u_{f,\beta} \overline{u}_{f,\beta} = I$$
 (5.8)

so that our cross section will be

$$\frac{d\sigma}{d\Omega} \propto \sum_{\alpha,\beta} |\overline{u}_f(\gamma_0 p_f + p_i \gamma_0) u_i|^2$$

$$= \text{Tr}[(\gamma_0 p_f + p_i \gamma_0)(p_f \gamma_0 + \gamma_0 p_i)]$$

$$= 2 \text{Tr}(m^2 + \gamma_0 p_i \gamma_0 p_f), \qquad (5.9)$$

which, as is well known, leads exactly to the Mott formula. In other words, starting from (1.1) we are obliged to use (5.8) instead of the sum over polarization states, but this modification of the standard calculation is exactly compensated by the appearance of the term $\mathbf{r} \cdot \boldsymbol{\alpha} / r$ in (5.2).

VI. CONCLUSIONS

We remark, in conclusion, that, up to now, there are no unwanted results emerging from the broadening of the set of the solutions of the Dirac equation (1.2) to that of the four-component Feynman—Gell-Mann equation (1.1). Or better still, \mathcal{F} also contains spinors that are not solutions of (1.2), but that constitute the base for the V-A theory of the weak interactions. Moreover, a statistical interpretation of (1.1), different from the interpretation that identifies the conserved density directly with a probability density, is possible and the appearance of negative values for the density can be interpreted on the ground of the mixing of particles and antiparticles. Anyway, the limitations imposed to the solutions of (1.1) by the condition

(1.2) in order to get positive densities must be considered much too severe, so that, also in this respect, we should widen our horizon beyond the boundaries of \mathcal{D}_{-} .

Of course a crucial test, to judge whether or not we can completely adopt (1.1) instead of (1.2) for fermions, is the analysis of the bound states of well-known quantum systems such as H atoms.

First of all we will point out that (1.1) can be considered as a generalization of a Klein-Gordon equation on a four-spinor ψ taking into account "the interaction of the electromagnetic field with an electric and a magnetic dipole moment, collectively called 'Dirac moment of the electron'... If the electromagnetic field is sufficiently weak, the effect of this term on the energy eigenvalue is small and can be calculated by approximation methods which involves first order perturbation theory and an expansion in inverse power of c, the velocity of light."12 Hence we can claim that the eigenvalues of (1.1) for the bound states of an H atom will not be different, apart from the perturbation spin terms, from the eigenvalues of a Klein-Gordon equation with external Coulomb interaction. On the other hand it is well known 12 that this Klein-Gordon spectrum of an H atom coincides with the corresponding Dirac spectrum, exactly apart from the spin terms. From these preliminary considerations can be deduced that the eigenvalues of (1.1) for an external Coulomb interaction will reproduce the Dirac case without differences.

However it will be very interesting to solve exactly (1.1) for an H atom because it is evident that this second-order equation allows different selections of the complete set of commuting observables used in classifying the eigenstates. For instance it is possible for (1.1), but impossible for (1.2)to work out solutions that are also eigenstates of γ_5 ; in other words, it is possible to solve our bound-state problem also in the framework of the two-component formalism of the so-called Kramers equation that, as is well known, "gives the same results in standard problems as the usual Dirac theory, but often the calculations are simpler and the physics more transparent."13 The importance of this analysis lies, beyond its own interest, in the fact that, in this case, we would look for solutions of (1.1) exactly in the same domain in which Feynman and Gell-Mann worked their theory of the Fermi interaction:² a complete exposition of this problem will constitute the argument of a forthcoming paper.

- ¹N. Cufaro Petroni, Ph. Gueret, and J. P. Vigier, Nuovo Cimento B 81, 243 (1984); Phys. Rev. D 30, 495 (1984).
- ²R. P. Feynman and M. Gell-Mann, Phys. Rev. 109, 193 (1958).
- ³D. Bohm, Phys. Rev. **85**, 166 (1952); **85**, 180 (1952); **89**, 458 (1953); D. Bohm and J. P. Vigier, *ibid.* **96**, 208 (1954).
- 4P. A. M. Dirac, Nature 168, 906 (1951); 169, 702 (1952); K. P. Sinha, C. Sivaram, and E. C. G. Sudarshan, Found. Phys. 6, 65 (1976); 6, 717 (1976); J. P. Vigier, Astron. Nachr. 303, 55 (1982); N. Cufaro Petroni and J. P. Vigier, Found. Phys. 13, 253 (1983).
- ⁵W. Lehr and J. Park, J. Math. Phys. 18, 1235 (1977); F. Guerra and P. Ruggiero, Lett. Nuovo Cimento 23, 529 (1978); N. Cufaro Petroni and J. P. Vigier, Phys. Lett. 73A, 289 (1979); 81A, 12 (1981); J. Phys. A 17, 599 (1984).
- ⁶N. Cufaro Petroni, Ph. Droz-Vincent, and J. P. Vigier, Lett. Nuovo Cimento 31, 415 (1981); N. Cufaro Petroni and J. P. Vigier, Phys. Lett. 88A, 272 (1982); 93A, 383 (1983).
- ⁷A. Aspect, G. Grangier, and G. Roger, Phys. Rev. Lett. 47, 460

- (1981); **49**, 91 (1982); A. Aspect, J. Dalibard, and G. Roger, *ibid*. **49**, 1804 (1982).
- ⁸A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).
- ⁹J. D. Bjorken and S. D. Drell, *Relativistic Quantum Mechanics* (McGraw-Hill, New York, 1964).
- 10E. P. Wigner, Phys. Rev. 40, 749 (1932); J. P. Vigier and Ya.
 P. Terletskij, Zh. Eksp. Teor. Fiz. 40, 508 (1961) [Sov. Phys. JETP 13, 356 (1961)].
- ¹¹N. Bourbaki, Elements de Mathémathiques; XIII—Intégration (Hermann, Paris, 1965), Sect. 4.
- ¹²H. A. Bethe and E. E. Salpeter, Quantum Mechanics of Oneand Two-Electron Atoms (Academic, New York, 1957); H. A. Bethe, Intermediate Quantum Mechanics (Benjamin, New York, 1964).
- ¹³L. Brown, Phys. Rev. 111, 957 (1958); Lectures in Theoretical Physics (Interscience, New York, 1962), Vol. IV.