 Relation between relativistic quantum mechanics and classical electromagnetic field theory

Johannes F. Geurdes
C van der Lijnstraat 164, 2593 NN Den Haag, The Netherlands
(Received 25 October 1994)

The objective of this report is twofold. In the first place, it aims to demonstrate that a four-dimensional local U(1) gauge invariant relativistic quantum mechanical Dirac-type equation is derivable from the equations for the classical electromagnetic field. In the second place, the transformational consequences of this local U(1) invariance are used to obtain solutions of different Maxwell equations.

PACS number(s): 41.20.Bt, 03.65.—w

I. INTRODUCTION

In the early history of quantum mechanics, some physicists tried to interpret the formalism of quantum mechanics in terms of classical physics. Two examples are the hydrodynamical interpretations of Madelung and Korn [1,2]. Despite the fact that those classical interpretations fell short on the conceptual side [3], they revealed interesting similarities in mathematical structure between quantum mechanics and classical physics. The Madelung-type transformation was applied in, for instance, Bohm's hidden variable interpretation of nonrelativistic quantum mechanics [4] and, more recently, in Vigier's study on neutron interferometry [5] and in Kyprianidis's study of the Sutherland paradox [6]. The Madelung transformation was also applied in Purcell's study of Lie-Backlund transformations of quantum fluids [7] and in Perrie's study of the mapping of classical wave systems to nonlinear Schrödinger equations in which nonlinear Schrödinger equations are obtained as resultant Euler-Lagrange equations [8]. In the present paper, as in a previous one [9], the connections between quantum and classical physics are studied further. It will be found that an equation similar to a four-dimensional relativistic quantum mechanical Dirac equation can be derived from classical field theory. In addition, the obtained formal relation can be used to derive solutions of different Maxwell equations. The structure of this report is as follows. In Sec. II, the problem is defined. In Sec. III, the necessary mathematics is briefly outlined. Then, in Sec. IV a Dirac-type equation is derived, and in Sec. V, the U(1) invariance of this equation is employed to obtain solutions of generally nonsolenoidal Maxwell equations. In Sec. VI, the obtained result is briefly discussed.

II. PRELIMINARIES

For mathematical convenience, the dielectric constant and the magnetic permeability are chosen to be unity. Moreover, $\kappa = c = 1$ and the factor $4\pi$ is included in the charge density and the electrical current vector. With this system of units, the Maxwell equations can be written as

$$\nabla \times \vec{F} = i \left( \frac{\partial \vec{E}}{\partial t} + \vec{j} \right), \quad \nabla \cdot \vec{F} = q,$$

with $F^k = E^k + iB^k$ ($k = 1, 2, 3$) and $(E^1, E^2, E^3)$ the electrical vector, $(B^1, B^2, B^3)$ the magnetic vector, and $(j^1, j^2, j^3)$ the electrical current vector. $q$ is the charge density. Note that it is assumed that the fields depend on the array $(x_1, x_2, x_3, t)$, with $x_k$ the spatial coordinates $(k = 1, 2, 3)$ and $t$ the time. Later on, use will be made of $x = (x_1, x_2, x_3, x_4)$, with $x_4 = it$. Those coordinates are useful for relativistic quantum mechanics. The dependence will be made explicit in special cases only. Suppose we start with a solenoidal system of equations with a charge density that is independent of time. This means we may write

$$q(\vec{x}) = \nabla \cdot \vec{Q}(\vec{x}),$$

$$\vec{j} = \nabla \times \vec{J}.$$

Moreover, suppose that there is a vector $(C^1, C^2, C^3)$ such that

$$\vec{F} = \vec{Q}(\vec{x}) + \nabla \times \vec{C}.$$

Because the “div on curl” operation vanishes, the divergence condition on the $\vec{F}$ vector, given in (1) applies. It can also be concluded from (2) that the system is solenoidal. Substitution of (2) and (3) in (1) then gives

$$\nabla \times \left[ \vec{F} - i \frac{\partial \vec{C}}{\partial t} - i \cdot \vec{J} \right] = 0.$$

Because the “curl on grad” operation produces the zero vector, we are allowed to introduce a function $\phi$ such that

$$\vec{F} - i \frac{\partial \vec{C}}{\partial t} - i \cdot \vec{J} = \nabla \phi.$$

Moreover, it is assumed that

$$\nabla \cdot \vec{C} = i \frac{\partial \phi}{\partial t}.$$

Hence, taking the divergence of (5) and using (6) and the divergence condition on the $F$ vector, it follows that

$$\nabla^2 \phi = q(\vec{x}) - i \nabla \cdot \vec{J} = -\omega.$$

Suppose the $F$ vector, the current vector, and the charge density can be decomposed as

$$\vec{F} = \sum_{\mu=1}^{4} \vec{F}^\mu, \quad \vec{j} = \sum_{\mu=1}^{4} j^\mu, \quad q = \sum_{\mu=1}^{4} q^\mu.$$
and that, for each component, an equation similar to (1) applies. This entails the functions $\delta^\mu$, such that
\[
\Box^2 \phi^\mu = -\omega^\mu, \quad \mu = 1, 2, 3, 4.
\] (9)
Or, in the form of four-dimensional vectors,
\[
\Box^2 \tilde{\phi} = -\vec{\omega},
\] (10)
where the tilde indicates a four-dimensional vector. This equation will be the basis for further analysis.

III. DEFINITIONS AND THEOREMS

In this section the necessary mathematics will be outlined. First, use will be made of two $4 \times 4$ matrices $J_1$ and $J_2$. The entries of those matrices are defined by
\[
(J_1)_{\mu, \lambda} = \delta^\mu_{\lambda, 1} + \delta^\mu_{\lambda, 2},
\] (11)
\[
(J_2)_{\mu, \lambda} = \delta^\mu_{\lambda, 3} + \delta^\mu_{\lambda, 4}
\]
with $\delta_{\mu, \lambda}$ Kronecker's $\delta$.

In the second place, a matrix $S$ is defined by
\[
S = I - \gamma^1 \gamma^2 - \gamma^3 \gamma^4,
\] (12)
with $\gamma^\mu$ the Dirac matrices ($\mu = 1, 2, 3, 4$) and $I$ the $4 \times 4$ unity matrix. For $k = 1, 2, 3$, the Dirac matrices are equal to
\[
\gamma^k = \begin{pmatrix} O & \sigma^k \\ \sigma^k & O \end{pmatrix},
\] (13)
with $O$ the $2 \times 2$ zero matrix and $\sigma^k$ ($k = 1, 2, 3$) the Pauli matrices
\[
\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
\] (14)
The matrix $\gamma^4$ is defined by $J_1 - J_2$.

From (13) it then follows that
\[
\gamma^1 S = S \gamma^1, \quad \gamma^2 S = S \gamma^2, \quad \gamma^3 S = S \gamma^3, \quad \gamma^4 S = S \gamma^4.
\] (15)
Or, more generally, when $f(1) = 2$, $f(2) = 3$, $f(3) = 1$, $f(4) = 4$,
\[
\gamma^f \mu S = S \gamma^\mu, \quad \mu = 1, 2, 3, 4.
\] (16)

Furthermore, $S$ is the sum of two other matrices $S_1$ and $S_2$ with
\[
S_a = \begin{pmatrix} s_a & O \\ O & s_a \end{pmatrix}, \quad a = 1, 2,
\] (17)
and the $2 \times 2$ matrices $s_1$ and $s_2$ are defined by
\[
s_1 = \begin{pmatrix} 1 - i & 0 \\ 0 & 1 + i \end{pmatrix}, \quad s_2 = \begin{pmatrix} 0 & -1 - i \\ 1 - i & 0 \end{pmatrix}.
\] (18)
Moreover, a matrix $T$ is defined by $T = I - S$. From this definition and using $\gamma^\mu \gamma^\lambda + \gamma^\lambda \gamma^\mu = 2I \delta^\mu_{\lambda, 4}$, it follows that $T^2 = -3I$.

Third, a matrix $\mathcal{R}(X)$ is associated with a matrix $X$ such that
\[
Y = \mathcal{R}(X) \equiv Y_{\mu, \lambda} = \sqrt{X_{\mu, \lambda}}.
\] (19)

Using this notation, we may present the following theorems for two arbitrary four-vectors ($f^1, f^2, f^3, f^4$) and ($g^1, g^2, g^3, g^4$). The first theorem is
\[
\gamma^\mu (f \otimes g) = [\mathcal{R}(\gamma^\mu) f \otimes [\mathcal{R}(\gamma^\mu) g]],
\] (20)
with the $\mu$th component of the direct product defined by
\[
(f \otimes g)^\mu = f^\mu g^\mu,
\] (21)
for $\mu = 1, 2, 3, 4$. The proof of the above theorem is based on the fact that the Dirac matrices for each column or row vector contain only one entry unequal to zero. Two other theorems can be proved similarly. The first is
\[
\gamma^\mu S_a (f \otimes g) = [\mathcal{R}(\gamma^\mu S_a) f \otimes [\mathcal{R}(\gamma^\mu S_a) g]],
\] (22)
with $a = 1, 2$ and the $S_a$ as defined previously. The second is
\[
\gamma^\mu J_a (f \otimes g) = [\mathcal{R}(\gamma^\mu J_a) f \otimes [\mathcal{R}(\gamma^\mu J_a) g]],
\] (23)
with $a = 1, 2$. In addition, we have
\[
J_a (f \otimes g) = (J_a f \otimes g = f \otimes (J_a g)) = (J_a f) \otimes (J_a g).
\] (24)
In the following a theorem will be presented that is crucial to the derivation. Suppose there are three different types of four-vectors,

\[
\text{In (25), } a, b, c = 1, 2 \text{ and } \mu = 1, 2, 3, 4. \text{ Those vectors are supposed to be related by}
\]
\[
\mathcal{R}(\gamma^\mu J_a) \tilde{\psi}_a = \mathcal{R}(\gamma^f \mu S_1) \tilde{\psi}_a + \mathcal{R}(\gamma^\mu) (\xi_{b, a} + \xi_{b, a} \tilde{\chi}_{b, a})
\]
with $b = 1 \rightarrow b = 2, \quad b = 2 \rightarrow b = 1$.

Observing (23) we may write
\[
\gamma^\mu J_a (\tilde{\psi}_1 \otimes \tilde{\psi}_2) = [\mathcal{R}(\gamma^\mu J_a) \tilde{\psi}_1] \otimes [\mathcal{R}(\gamma^\mu J_a) \tilde{\psi}_2].
\] (27)

Substitution of (26) in (27) and making use of (22) and the summation convention for $c, d = 1, 2$, with $\delta^\mu_{c, d}$ Kronecker's $\delta$, then gives
\[
\gamma^\mu J_a (\tilde{\psi}_1 \otimes \tilde{\psi}_2) = \gamma^f \mu S_1 (\tilde{\psi}_1 \otimes \tilde{\psi}_2) + (1 - \delta^\mu_{c, d}) \gamma^f \mu (\xi_{c, d} \tilde{\chi}_{d, a})
\]
\[
+ \gamma^f \mu (\xi_{c, d} \tilde{\chi}_{d, a} + \xi_{b, a} \tilde{\chi}_{b, a})
\]
\[
+ (1 - \delta^\mu_{c, d}) [\mathcal{R}(\gamma^\mu) S_1] \tilde{\psi}_c.
\] (28)
After rewriting the terms of (28) we arrive at
\[ T^\gamma_{\mu}(\bar{\Psi}_1 \otimes \bar{\Psi}_2) \]
\[ = (1 - \delta^{c,d}) \delta^{a,b} \gamma^\mu (\bar{\Psi}_{1,a} - \frac{1}{2} \bar{\Psi}_{a,c}) \otimes \bar{\Psi}_{d,b} + (1 - \delta^{c,d}) \delta^{a,b} \gamma^\mu (\bar{\Psi}_{1,a} - \frac{1}{2} \bar{\Psi}_{a,c}) \otimes \bar{\Psi}_{d,b} \],
(29)
with the summation convention for \(a,b,c,d = 1,2\) and
\[ \bar{\Psi}_{c,a} = J_a \bar{\Psi}_c \].
(30)

This mathematical outline ends by noting that the application of the "\( R \)" operation to the Pauli matrices gives the following relations:
\[ \mathcal{R}(i\sigma^k) = \mathcal{R}(i\sigma^k)^\dagger = \mathcal{R}(i\sigma^k)^\dagger \mathcal{R}(i\sigma^k) = I_2 \]
while
\[ \mathcal{R}(-I_2) = \mathcal{R}(-I_2)^\dagger = \mathcal{R}(-I_2)^\dagger \mathcal{R}(-I_2) = I_2 \].
(31)

Moreover, we remark that in this paper the summation convention for greek letters also applies unless special emphasis is wanted.

IV. DERIVATION OF A U(1) INVARIANT DIRAC-TYPE EQUATION

In order to derive a Dirac equation, we first identify the vector on the left-hand side of (7), containing the functions \( \phi^\mu \), with the first term on the right-hand side of Eq. (29) (summation convention for \(a,b,c,d = 1,2\),
\[ \phi = (1 - \delta^{a,b}) \delta^{c,d} (\bar{\Psi}_{a,c} - \frac{1}{2} \bar{\Psi}_{a,c}) \otimes \bar{\Psi}_{d,b} \].
(33)

The second term on the right-hand side of (29) is written, for short, as
\[ \bar{U}^\mu = (1 - \delta^{a,b}) \delta^{c,d} (\bar{\Psi}_{a,c} - \frac{1}{2} \bar{\Psi}_{a,c}) \otimes \bar{\Psi}_{d,b} \].
(34)

Hence, (29) may be rewritten as
\[ T^\gamma_{\mu}(\bar{\Psi}_1 \otimes \bar{\Psi}_2) = \gamma^\mu \phi + \gamma^\mu \bar{U}^\mu \].
(35)

In addition, a gauge covariant derivative is defined by
\[ \mathcal{D}_\mu = \partial_\mu - \bar{G}_\mu \]
and \( G_\mu(x) \) a gauge function. With the definition of the \( U^\mu \) vector and the gauge covariant derivative, the right-hand side of (7) can be identified as
\[ \bar{\omega} = \gamma^\nu \mathcal{D}_\nu \sum_{\mu=1}^4 \gamma^\mu \mathcal{D}_\mu \bar{U}^\mu + \gamma^\mu \gamma^\nu (\bar{G}_\mu \bar{G}_\nu - \partial_\mu \bar{G}_\nu - \bar{G}_\mu \partial_\nu \bar{G}_\nu) \bar{\phi} \].
(36)

Operating gauge covariant differentiation on the proper terms we obtain
\[ T^\gamma_{\mu}(\bar{\Psi}_1 \otimes \bar{\Psi}_2) = \gamma^\mu \bar{G}_\mu + \sum_{\mu=1}^4 \gamma^\mu \mathcal{D}_\mu \bar{U}^\mu \].
Employing the operator \( \gamma^\nu \mathcal{D}_\nu \) on both sides of the previous equation and making use of (37), the following equation can be derived:
\[ \gamma^\nu \mathcal{D}_\nu \gamma^\mu \mathcal{D}_\mu (\bar{\Psi}_1 \otimes \bar{\Psi}_2) = 0 \].
(39)

Note that this equation for a certain, yet unspecified subset of cases, contains the following U(1) gauge invariant equation:
\[ \gamma^\mu \bar{G}_\mu = 0 \] with \( \bar{\Psi} = (\bar{\Psi}_1 \otimes \bar{\Psi}_2) \),
(40)
which can be identified as a massless free-particle Dirac equation with gauge covariant derivatives. Hence, from a set of classical field equations, an equation similar in form to an equation central to relativistic quantum mechanics can be derived. The U(1) invariance of Eqs. (39) and (40) can be illustrated by noting that if
\[ \bar{\Psi} \to e^{iR} \bar{\Psi}, \quad \bar{G}_\mu \to \bar{G}_\mu + i \partial_\mu R \],
(41)
both Lagrangian functions
\[ L_1(x) = \bar{\Psi}^\dagger \gamma^\nu (\gamma^\mu \mathcal{D}_\mu - i \partial_\mu) \bar{\Psi} \]
\[ L_2(x) = \bar{\Psi}^\dagger \gamma^\nu (\gamma^\mu \mathcal{D}_\mu - i \gamma^\nu \mathcal{D}_\nu) \bar{\Psi} \]
remain invariant [10].

V. APPLICATION OF U(1) INVARIANCE

With the transformation in (41), new Maxwell equations can be derived from the original solenoidal system. Because the \( \Psi \) vector transforms like in (41), we may take
\[ \bar{\Psi}_a \to e^{iR/2} \bar{\Psi}_a, \quad a = 1,2 \].
(42)

Moreover, let us assume that \( (a,b) = (1,2) \)
\[ \bar{\Psi}_{a,b} \to e^{iR/2} \bar{\Psi}_{a,b} \].
(43)

In this case, a new vector \( \phi^\nu \) can be obtained from (33), using the previous two equations,
\[ \phi^\nu = e^{iR} \bar{\phi} \].
(44)

This means that for each \( \mu \)th component of the new vector \( \phi^\nu \) the gradient can be written, using (5), as
\[ \mathcal{D} \phi^\mu = i \phi^\mu \mathcal{D} R + e^{iR} \left[ \mathcal{D} \phi \mu - i \mathcal{D} C \mu \right] + i \mathcal{D} \mathcal{J}^\mu \],
(45)

with \( R = R(x) \). This equation can be rewritten as
\[ \mathcal{D} \phi^\mu = e^{iR} \mathcal{D} \phi^\mu + i e^{iR} \mathcal{D} \mathcal{J}^\mu \]
\[ - e^{iR} \mathcal{D} \mathcal{C} \mu \mathcal{D} R - i \mathcal{D} \mathcal{J}^\mu e^{iR} \].
(46)

If the \( F \), the \( Q \), and the \( j \) vector in the primed system are written as
\[ F^\mu = e^{iR} \mathcal{D} \mathcal{J}^\mu + i \mathcal{D} \mathcal{Q}^\mu \]
with \( \mathcal{Q}^\mu = e^{iR} \mathcal{D} \mathcal{J}^\mu \)
\[ Q^\mu = e^{iR} \mathcal{D} \mathcal{J}^\mu + i \mathcal{D} \mathcal{Q}^\mu \]
\[ j^\mu = \frac{\partial \mathcal{Q}^\mu}{\partial t} + \mathcal{D} \mathcal{J}^\mu \]
\[ j^\mu = \frac{\partial \mathcal{Q}^\mu}{\partial t} + \mathcal{D} \mathcal{J}^\mu \]
\[ j^\mu = \frac{\partial \mathcal{Q}^\mu}{\partial t} + \mathcal{D} \mathcal{J}^\mu \]
then, taking the curl of (47) and noting that \( (\mathcal{C}^k)^\mu = e^{iR} \mathcal{C}^k \)
\( (k = 1,2,3) \), the following irregular set of Maxwell equations can be obtained:
\[ \vec{\nabla} \times \vec{E}' + \frac{\partial \vec{B}'}{\partial t} = -\text{Im} \vec{j}', \quad \vec{\nabla} \cdot \vec{B}' = \text{Im} q' , \tag{49} \]
\[ \vec{\nabla} \times \vec{B}' - \frac{\partial \vec{E}'}{\partial t} = \text{Re} \vec{j}', \quad \vec{\nabla} \cdot \vec{E}' = \text{Re} q' . \]

This irregular system, which resembles the system of Maxwell equations in Dirac's monopole theory [11], can be transformed into a regular system as follows. Suppose we define a double primed system as
\[ \vec{B}'' = \vec{B}' - \text{Im} \vec{Q}' , \tag{50} \]
and
\[ \vec{E}'' = \vec{E}' + \vec{X} , \tag{51} \]
with \((X^1, X^2, X^3)\),
\[ \vec{X} = \text{Im} \left[ e^{iR} \left( \vec{J} - i\vec{C} \frac{\partial \vec{R}}{\partial t} \right) \right] . \tag{52} \]

Then it follows that in the double primed system the irregular Maxwell equations (49) can be rewritten as
\[ \vec{\nabla} \times \vec{E}'' + \frac{\partial \vec{B}''}{\partial t} = 0 , \quad \vec{\nabla} \cdot \vec{B}'' = 0 , \tag{53} \]
\[ \vec{\nabla} \times \vec{B}'' - \frac{\partial \vec{E}''}{\partial t} = \vec{j}'', \quad \vec{\nabla} \cdot \vec{E}'' = q'' . \]

In this case we have
\[ q'' = \text{Re} q' + \vec{\nabla} \cdot \left( \text{Im} \left[ e^{iR} \left( \vec{J} - i\vec{C} \frac{\partial \vec{R}}{\partial t} \right) \right] \right) \tag{54} \]
for the charge density and
\[ \vec{j}'' = \text{Re} \vec{j}' - \vec{\nabla} \times (\text{Im} \vec{Q}') \]
\[ - \frac{\partial}{\partial t} \left[ \text{Im} \left[ e^{iR} \left( \vec{j} - i\vec{C} \frac{\partial \vec{R}}{\partial t} \right) \right] \right] \tag{55} \]
for the current density vector. Note that Eqs. (33) and (34) determine the \(\psi, \xi, \) and \(\lambda\) vectors. Because of the dimensions of the system, we actually have \((1 \times 4) + (4 \times 4) = 20\) equations. Moreover, there are, according to (26), only four "free" \(\lambda\) vectors. Because there are \(2 \times 2\) \(\varepsilon_{b, c}\) vectors and \(2 \psi_b\) vectors, we have in totality \((4 \times 4) + (4 \times 4) + (2 \times 4) = 40\) variables available. Hence, the proposed derivation and transformation can be performed in principle.

VI. CONCLUSION

In this Brief Report two things have been demonstrated. In the first place, an equation similar to a relativistic quantum mechanical Dirac equation can be obtained from classical field theory. Hence, classical field theory and relativistic quantum mechanics are more deeply connected than is generally assumed. A similar finding was obtained by Moses [12]. However, Moses used a six-dimensional theory, whereas in the present Brief Report a four-dimensional theory is used. In the second place, the local U(1) gauge invariance of the Dirac-type equation associated with the classical field equations is used to obtain solutions of generally nonsolenoidal Maxwell equations from the original solenoidal system. This adds to our tactics for deriving analytical solutions of Maxwell electromagnetic field equations.