Lorentz covariance of optical Dirac equation and spinorial photon field

Xiao-Lu You and Chun-Fang Li
Department of Physics, Shanghai University, 99 Shangda Road, 200444 Shanghai, People’s Republic of China
* Author to whom any correspondence should be addressed.
E-mail: cfl@shu.edu.cn

Keywords: optical Dirac equation, Lorentz covariance, relativistic condition, spinorial photon field

Abstract
In a recent paper (2014 New J. Phys. 16 093008) Barnett discussed the so-called optical Dirac equation and referred to the involved wave function as a spinor. But as he claimed explicitly, he did not really associate that wave function with a true spinor. Here we show that if the optical Dirac equation is interpreted as the dynamical equation for the photon in conventional quantum mechanics, the wave function, called the photon field, does transform under the Lorentz transformation as a spinor. For the optical Dirac equation to be Lorentz covariant, however, a constraint on the photon field is required, which can be cast into the form of Maxwell’s divergence equations. It is found that the spinorial photon field not only satisfies the principle of locality but also has the right dimensionality as is required by conventional quantum mechanics.

1. Introduction
The significance of formulating the quantum-mechanical equation for the photon cannot be overestimated. So even though Akhiezer and Berestetskii [1] claimed more than half a century ago that it is impossible to introduce a photon wave function in the position representation, such an effort has never been given up. But unfortunately it is still unclear [2] whether there is a photon wave function that, on one hand, satisfies the principle of locality [3] and, on the other hand, has the dimensionality that is required by quantum mechanics. For a comprehensive review of the photon wave function before 2005, readers may refer to the articles by Bialynicki-Birula [4] and by Keller [5].

Weinberg [6] once pointed out from the point of view of geometric object that a photon wave function cannot be a Lorentz vector nor an anti-symmetric Lorentz tensor. A well-familiar wave function that is neither a Lorentz vector nor a Lorentz tensor is the relativistic electron field. It is a Lorentz spinor. The Dirac equation for a free electron reads [7, 8]

$$i\hbar \frac{\partial \psi}{\partial t} = H_D \psi,$$

where $H_D = ic\gamma_0 \gamma \cdot p + \gamma_0 mc^2$ is the Hamiltonian, $c$ is the speed of light in vacuum,

$$\gamma_0 = \begin{pmatrix} I_2 & 0 \\ 0 & -I_2 \end{pmatrix}, \quad \gamma = \begin{pmatrix} 0 & -i\sigma \\ i\sigma & 0 \end{pmatrix},$$

$I_2$ is the two-by-two unit matrix, $\sigma_i (i = 1, 2, 3)$ are the Pauli matrices, and $p = -i\hbar \nabla$ is the momentum operator. Recently, Barnett [9] exploited the analogy between Maxwell’s equations in free space and the Dirac equation (1) to study the mechanical properties of light by discussing the so-called optical Dirac equation, which takes the form of

$$i\hbar \frac{\partial \Psi}{\partial t} = H \Psi,$$
where $H = i\gamma_0 \Gamma \cdot p$ is the Hamiltonian,

$$\Gamma_0 = \begin{pmatrix} I_3 & 0 \\ 0 & -I_3 \end{pmatrix}, \quad \Gamma = \begin{pmatrix} 0 & \Sigma \\ \Sigma & 0 \end{pmatrix},$$

$I_3$ is the three-by-three unit matrix, and $(\Sigma_k)_{ij} = -i\epsilon_{ijk}$ with $\epsilon_{ijk}$ the Levi-Civita pseudotensor. To emphasize the analogy with Dirac spinors for the electron, he also referred to the six-component wave function $\Psi$ in (2a) as a spinor. But he did not really associate that wave function with a true spinor. By the optical Dirac equation he actually meant that the upper and lower parts of the wave function are essentially the classical electric and magnetic fields, respectively [10, 11]. As is well known, the classical electromagnetic field transforms under the Lorentz transformation as an anti-symmetric tensor [12]. Here we will show that if the optical Dirac equation (2a) is interpreted as the dynamical equation for the photon in accordance with conventional quantum mechanics, the photon wave function $\Psi$ does transform under the Lorentz transformation as a spinor. The key point is that the optical Dirac equation (2a) is not relativistic by itself in that the $\Gamma$’s here have not the simple commutation relations of the $\gamma$’s in the Dirac equation (1). That is to say, it does not by itself guarantee the wave function to be a spinor under the Lorentz transformation. We will show, however, that there is a necessary and sufficient condition for (2a) to be relativistic. It is a time-independent constraint on the wave function,

$$(\Gamma \cdot p)^2 \Psi = p^2 \Psi. \tag{2b}$$

As a matter of fact, letting $\Psi = \begin{pmatrix} \zeta \\ \eta \end{pmatrix}$ in accordance with the concrete forms of the matrices $\Gamma_0$ and $\Gamma$ and making use of the identity

$$(\Sigma \cdot a)b = ia \times b, \tag{3}$$

the optical Dirac equation (2a) can be cast into

$$i\hbar \frac{\partial \zeta}{\partial t} = -cp \times \eta, \quad i\hbar \frac{\partial \eta}{\partial t} = cp \times \zeta,$$

or, equivalently,

$$\frac{\partial \zeta}{\partial t} = c\nabla \times \eta, \tag{4a}$$

$$\frac{\partial \eta}{\partial t} = -c\nabla \times \zeta. \tag{4b}$$

Meanwhile, as we will show in section 3, the constraint equation (2b) can be cast into

$$\nabla \cdot \zeta = 0, \tag{4c}$$

$$\nabla \cdot \eta = 0, \tag{4d}$$

when (2a) or (4a) and (4b) is taken into account. This shows that the set of equations (2a) and (2b) is equivalent to the set of equations (4a)–(4d). It is seen that the equation set (4a)–(4d) is identical with the free-space Maxwell’s equations in form,

$$\varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} = \nabla \times \mathbf{H}, \tag{5a}$$

$$\mu_0 \frac{\partial \mathbf{H}}{\partial t} = -\nabla \times \mathbf{E}, \tag{5b}$$

$$\nabla \cdot \mathbf{E} = 0, \tag{5c}$$

$$\nabla \cdot \mathbf{H} = 0, \tag{5d}$$

if the following correspondences are assumed,

$$\zeta \sim \sqrt{\varepsilon_0} \mathbf{E}, \quad \eta \sim \sqrt{\mu_0} \mathbf{H}. \tag{6}$$

The Lorentz covariance of equation set (4a)–(4d) or (2a) and (2b) can thus be inferred from the Lorentz covariance [12] of Maxwell’s equations (5a)–(5d).

A well-known fact is that the electromagnetic field in Maxwell’s equations (5a)–(5d) transforms under the Lorentz transformation as an anti-symmetric tensor. From the correspondence between (4a)–(4d) and
(5a)–(5d) one might reason that the photon wave function $\Psi$ also transforms as a tensor. Nevertheless, Cook [13] has shown that if (4a)–(4d) are interpreted as the quantum-mechanical equations for the photon, the photon wave function does not transform as a tensor under the Lorentz transformation. What we will show is that the wave function in (4a)–(4d) or (2a) and (2b) can transform under the Lorentz transformation as a spinor. The spinor property of the wave function lifts (2a) and (2b) to the genuine quantum-mechanical equations for the photon in the sense that the wave function not only satisfies the principle of locality but also has the dimensions of the square-root of a probability density. By the principle of locality it is meant here, according to Bialynicki-Birula and Bialynicka-Birula [2], that the value of the photon wave function in space at the point $x$ only on its value at the point $x$ which was mapped into the point $x'$ by that transformation. The purpose of this paper is to formulate these quantum-mechanical equations by deriving the constraint equation (2b) and, on this basis, showing that the involved wave function transforms under the Lorentz transformation as a spinor. For the sake of clarity, we will refer to the spinorial wave function as the photon field.

2. Derivation of relativistic condition

The matrices $\Gamma_0$ and $\Gamma$ in (2a) are all Hermitian and have the following algebraic properties,

\begin{align}
\Gamma_0^2 &= 1, \\
\Gamma_0 \Gamma + \Gamma \Gamma_0 &= 0, \\
\Gamma_i \Gamma_j \Gamma_k + \Gamma_k \Gamma_i \Gamma_j = \Gamma_0 \delta_{jk} + \Gamma_0 \delta_{ij}, & i,j,k = 1,2,3. 
\end{align}

With the aid of (7a), the optical Dirac equation (2a) can be rewritten in the form

\begin{equation}
\Gamma_{\mu} p_{\mu} \Psi = 0, \quad \mu = 0,1,2,3, 
\end{equation}

where $p_0 = -i\hbar \partial / \partial x_0$, $x_0 = i c t$, and the summation convention is assumed. It is known that the Dirac equation (1) is equivalent to the Klein–Gordon equation of massive particles. For (8) to describe the dynamics of the photon, that is to say, for (8) to be relativistic, it should be equivalent to the Klein–Gordon equation of massless particles, $(p_0^2 + p^2)\Psi = 0$. Multiplying (8) by $\Gamma_{\mu} p_{\mu}$ on the left and considering (7a) and (7b) into account, one has

\begin{equation}
[p_0^2 + (\Gamma \cdot p)^2] \Psi = 0. 
\end{equation}

Instead of having the same commutation relations as those of the $\gamma$’s in the Dirac equation (1), the $\Gamma$’s here obey (7a)–(7c). Therefore, one cannot have $(\Gamma \cdot p)^2 = p^2$. For (9) to be the Klein–Gordon equation, $\Psi$ has to satisfy (2b). That is to say, (2b) is the necessary condition for (2a) to be relativistic. The other way around, if (2b) is satisfied, from (9) one will readily obtain the Klein–Gordon equation, meaning that (2b) is also the sufficient condition for (2a) to be relativistic.

From these discussions it is concluded that (2b) is the necessary and sufficient condition for (2a) to be relativistic. For this reason, we will refer to (2b) as the relativistic condition to distinguish from the dynamical equation (2a). Although it is known [9–11] that the dynamical equation is equivalent to the pair of time-dependent equations (4a) and (4b), the equivalence between the whole set of equations (2a) and (2b) and the whole set of equations (4a)–(4d) does not seem to have been proven before. That is to say, the equivalence between the time-independent relativistic condition (2b) and the pair of time-independent equations (4c) and (4d) does not seem to have been proven before. The former is usually regarded as the consequence of the latter [11]. Let us show below that (2b) is equivalent to (4c) and (4d) when (2a) or (4a) and (4b) is taken into account.

3. Equivalence between (2b) and (4c) and (4d)

To this end, we make use of (3) to rewrite (2b) in terms of $\zeta$ and $\eta$ as

\begin{equation}
p(p \cdot \zeta) = 0, \quad p(p \cdot \eta) = 0, \tag{10}
\end{equation}

or, equivalently,

\begin{equation}
\nabla(\nabla \cdot \zeta) = 0, \quad \nabla(\nabla \cdot \eta) = 0. \tag{11}
\end{equation}

From (2b) it follows that $H^2 = c^2 p^2$ by virtue of (7a) and (7b). By this it is meant that a photon with a nonzero energy cannot have a vanishing momentum. So when the energy is nonzero, from (10) we are led...
to \( p \cdot \zeta = 0 \) and \( p \cdot \eta = 0 \), which are nothing but the equation pair (4c) and (4d). Whereas when the energy is zero, that is to say, when \( H \Psi = 0 \), we have \( \frac{\partial \Psi}{\partial t} = 0 \) in accordance with (2a). This amounts to
\[
\nabla \times \zeta = 0, \quad \nabla \times \eta = 0
\]
(12)
as can be seen from (4a) and (4b). With the help of these equations, we readily obtain from (11)
\[
\nabla^2 \zeta = 0, \quad \nabla^2 \eta = 0.
\]
(13)
Since, according to Stratton [12], every function which is regular at infinity and which satisfies Laplace’s equation at all points of space is necessarily zero, we must have \( \zeta = 0 \) and \( \eta = 0 \), which satisfy (4c) and (4d). The relativistic condition (2b) is thus cast into the equation pair (4c) and (4d) whether the energy is nonzero or not. The other way around to derive (2b) from (4c) and (4d) is trivial [11]. In a word, the relativistic condition (2b) is equivalent to the pair of time-independent equations (4c) and (4d) if the optical Dirac equation (2a) is taken into account. That is to say, the equation set (2a) and (2b) is equivalent to the equation set (4a)–(4d).

It should be pointed out that if the energy does not vanish, (2b) is actually contained in (2a) or (4a) and (4b) [10]. In fact, from (2a) it follows that the photon field of nonzero energy is time-dependent, \( \frac{\partial \Psi}{\partial t} \neq 0 \), which means that both \( \zeta \) and \( \eta \) are time-dependent,

\[
\frac{\partial \zeta}{\partial t} \neq 0, \quad \frac{\partial \eta}{\partial t} \neq 0,
\]
due to the duality between \( \zeta \) and \( \eta \) in (2a). This can also be easily seen from (4a) and (4b) by a reductio ad absurdum proof. On account of the independence between the space and time variables, \( \nabla \cdot \zeta \) and \( \nabla \cdot \eta \) are also time-dependent unless they are equal to zero. On the other hand, from (4a) and (4b) it follows that
\[
\nabla \cdot \frac{\partial \zeta}{\partial t} = \frac{\partial}{\partial t}(\nabla \cdot \zeta) = 0, \quad \nabla \cdot \frac{\partial \eta}{\partial t} = \frac{\partial}{\partial t}(\nabla \cdot \eta) = 0.
\]
We therefore must have \( \nabla \cdot \zeta = \nabla \cdot \eta = 0 \). This proves our assertion. However, no theoretical description of the photon would be complete without taking account of the zero-energy state. This may help to understand why (2a) is not relativistic if \( \Psi \) is not constrained by (2b).

4. Photon field as a Lorentz spinor

4.1. Lorentz transformation of photon field

Now we are in a position to show the spinor property of the photon field in the quantum-mechanical equations (2a) and (2b). Consider a pure Lorentz transformation,

\[
x'_\mu = a_{\mu \nu} x_\nu, \quad (14)
\]
where the transformation matrix is a unitary matrix. Letting the primed reference system move relative to the unprimed reference system at a velocity \( v = vn \) in an arbitrarily given direction denoted by the unit vector \( n \), the time-space variables transform as follows,

\[
x'_0 = \gamma(x_0 - i\beta n \cdot x), \quad x'_1 = \gamma n(i\beta x_0 + n \cdot x) - n \times (n \times x),
\]
(15)
where \( \beta = v/c \) and \( \gamma = (1 - \beta^2)^{-1/2} \). For later convenience, we here write down the transformation laws [12] for the electromagnetic field under the Lorentz transformation (15),

\[
E' = (E \cdot n)n + \gamma(n \times E) \times n - \gamma\beta \epsilon \mu_0 \mathcal{H} \times n, \quad \mathcal{H}' = (\mathcal{H} \cdot n)n + \gamma(n \times \mathcal{H}) \times n + \gamma\beta \epsilon \mathcal{E} \times n.
\]
(16)
If attention is paid only to the correspondence between the equation set (4a)–(4d) and Maxwell’s equations (5a)–(5d), one is likely led to expressing the photon field \( \Psi \) as an anti-symmetric matrix [13]. Here we will follow Barnett [9] to take a different perspective. We will consider (2a) as the optical analog of the Dirac equation (1) and will show that if the photon field transforms under the Lorentz transformation (15) as follows,

\[
\Psi'(x', t') = \Lambda \Psi(x, t), \quad (17)
\]
where
\[
\Lambda = \exp(-i \Gamma_0 \mathbf{n} \cdot \mathbf{n} \chi)
\]
(18)
and $\chi$ is determined only by $\beta$ or $\gamma$, then the equation set (2a) and (2b) or (4a)–(4d) will be form-invariant.

In fact, from the identity $(\boldsymbol{\Sigma} \cdot \boldsymbol{n})^2 = 1 - \boldsymbol{n} \cdot \boldsymbol{n}$, where $\boldsymbol{n}$ is a dyadic, it is evident that $(\boldsymbol{\Sigma} \cdot \boldsymbol{n})^3 = \boldsymbol{\Sigma} \cdot \boldsymbol{n}$ by virtue of (3). As a result, one has

$$(\mathbf{\Gamma} \cdot \boldsymbol{n})^3 = \mathbf{\Gamma} \cdot \boldsymbol{n}.$$  

With the aid of this equality as well as of (7a) and (7b), the transformation matrix (18) can be expanded as

$$\Lambda = 1 + (\mathbf{\Gamma} \cdot \boldsymbol{n})^2(cosh \chi - 1) - i\mathbf{\gamma}_0 \mathbf{\Gamma} \cdot \boldsymbol{n} \sinh \chi.$$  

Letting $\Psi' = \begin{pmatrix} \zeta' \\ \eta' \end{pmatrix}$, we rewrite (17) as

$$\begin{pmatrix} \zeta' \\ \eta' \end{pmatrix} = \begin{pmatrix} 1 + (\boldsymbol{\Sigma} \cdot \boldsymbol{n})^2(cosh \chi - 1) & -i\boldsymbol{\Sigma} \cdot \boldsymbol{n} \sinh \chi \\ i\boldsymbol{\Sigma} \cdot \boldsymbol{n} \sinh \chi & 1 + (\boldsymbol{\Sigma} \cdot \boldsymbol{n})^2(cosh \chi - 1) \end{pmatrix} \begin{pmatrix} \zeta \\ \eta \end{pmatrix}.$$  

By use of (3), a straightforward calculation gives

$$\zeta' = (\zeta \cdot \boldsymbol{n})\boldsymbol{n} + (\boldsymbol{n} \times \zeta) \times \boldsymbol{n} \cosh \chi - \eta \times \boldsymbol{n} \sinh \chi,$$

$$\eta' = (\eta \cdot \boldsymbol{n})\boldsymbol{n} + (\boldsymbol{n} \times \eta) \times \boldsymbol{n} \cosh \chi + \zeta \times \boldsymbol{n} \sinh \chi.$$  

It is seen that if $\cosh \chi = \gamma$ and therefore $\sinh \chi = \gamma \beta$, these transformations are the same as the transformations (16) for the electromagnetic field by virtue of the correspondences (6). The equation set (4a)–(4d) is thus form-invariant under the transformations (19a) and (19b) the same as Maxwell’s equations (5a)–(5d) under the transformation (16). It is emphasized that the transformation (17) of the photon field under the Lorentz transformation differs from that of the electromagnetic field-strength tensor [12]. Instead, it is analogous to the transformation of the electron field in the Dirac theory. As is shown in (18), the generator of the infinitesimal Lorentz transformation of the photon field is $\Gamma_0 \mathbf{\Gamma}$, analogous to the generator $\gamma_0 \gamma/2$ of the infinitesimal Lorentz transformation of the electron field [7, 8]. In a word, the photon field transforms under the Lorentz transformation as a spinor in much the same way as the electron field does. It is important to note that the local transformation (17) or (19a) and (19b) is just what the principle of locality requires for the photon field with respect to the Lorentz transformation (15).

Apart from such a difference between the photon and electromagnetic fields, another subtle difference between the photon-field equations (2a) and (2b) and Maxwell’s equations (5a)–(5d) in Lorentz transformation also deserves mentioning. As is known, in terms of the electromagnetic field-strength tensor,

$$F_{\mu\nu} = \begin{pmatrix} 0 & iE_1/c & iE_2/c & iE_3/c \\ -iE_1/c & 0 & \mu_0 H_3 & -\mu_0 H_2 \\ -iE_2/c & -\mu_0 H_3 & 0 & \mu_0 H_1 \\ -iE_3/c & \mu_0 H_2 & -\mu_0 H_1 & 0 \end{pmatrix},$$

Maxwell’s equations (5a)–(5d) can also be cast into two equations. The first and third equations (5a) and (5c) are contained in the divergence equation,

$$\frac{\partial F_{\mu\nu}}{\partial x_\nu} = 0,$$

and the second and fourth equations (5b) and (5d) in the curl equation,

$$\frac{\partial F_{\mu\nu}}{\partial x_\lambda} + \frac{\partial F_{\nu\lambda}}{\partial x_\mu} + \frac{\partial F_{\lambda\mu}}{\partial x_\nu} = 0.$$  

These two equations are Lorentz covariant individually. But as has been shown in section 2, the optical Dirac equation (2a) is not Lorentz covariant without the help of the relativistic condition (2b). To appreciate how (2b) makes (2a) Lorentz covariant, let us have a look at the form-invariance of the photon-field equations (4a)–(4d) under the Lorentz transformations (15) and (19a) and (19b).

### 4.2. Lorentz covariance of photon-field equations

Our objective is to show that under the Lorentz transformations (15) and (19a) and (19b), the photon field $\Psi$ in the unprimed reference frame satisfies (4a)–(4d) if the photon field $\Psi'$ in the primed reference frame satisfies

$$\frac{\partial \zeta'}{\partial t'} = c\nabla' \times \eta',$$

(20a)
\[
\frac{\partial \eta'}{\partial x'} = -c \nabla' \times \zeta', \quad (20b)
\]
\[
\nabla' \cdot \zeta' = 0, \quad (20c)
\]
\[
\nabla' \cdot \eta' = 0, \quad (20d)
\]

where \( \nabla' \) means the gradient operator with respect to \( x' \). As is known, \( \partial / \partial x_\mu \) is a Lorentz four-vector,

\[
\frac{\partial}{\partial x'_\mu} = a_{\mu\nu} \frac{\partial}{\partial x_\nu}.
\]

By making use of the \( a_{\mu\nu} \) defined by (15), a straightforward calculation gives

\[
\frac{\partial}{\partial x'} = \gamma \left( \frac{\partial}{\partial t} + \beta c n \cdot \nabla \right), \quad (21a)
\]
\[
c \nabla' = \gamma n \left( \beta \frac{\partial}{\partial t} + c n \cdot \nabla \right) - c n \times (n \times \nabla), \quad (21b)
\]
in agreement with (15). Substituting (19a) and (19b) and (21a) and (21b) into (20a), one gets after a lengthy but direct calculation

\[
\frac{\partial \zeta}{\partial t} - c \nabla \times \eta = -\left( \gamma - 1 \right) \left( \frac{\partial \zeta}{\partial t} - c \nabla \times \eta \right) \cdot n + \gamma \beta c \nabla \cdot \zeta \right] n. \quad (22)
\]

On the other hand, substituting (19a) into (20c) and noticing (21b), one finds

\[
c \nabla \cdot \zeta = -\beta \left( \frac{\partial \zeta}{\partial t} - c \nabla \times \eta \right) \cdot n. \quad (23)
\]

Substituting it into (22), one has

\[
\frac{\partial \zeta}{\partial t} - c \nabla \times \eta + \left[ \gamma(1 - \beta^2) - 1 \right] \left( \frac{\partial \zeta}{\partial t} - c \nabla \times \eta \right) \cdot n = 0.
\]

Upon dot-multiplying this equation by the unit vector \( n \), one obtains

\[
\gamma(1 - \beta^2) \left( \frac{\partial \zeta}{\partial t} - c \nabla \times \eta \right) \cdot n = 0.
\]

Since \( \gamma(1 - \beta^2) \neq 0 \), one is left only with

\[
\left( \frac{\partial \zeta}{\partial t} - c \nabla \times \eta \right) \cdot n = 0.
\]

Therefore, from (23) one is led to \( \nabla \cdot \zeta = 0 \), which is exactly (4c). Furthermore, from (22) it follows that

\[
\frac{\partial \zeta}{\partial t} - c \nabla \times \eta = 0.
\]

This is nothing but (4a). The form-invariance of (4a) and (4c) under the Lorentz transformations (15) and (19a) and (19b) is thus proven. The form-invariance of (4b) and (4d) can also be proven in a similar way. It is clearly seen from these discussions that neither (4a) nor (4c) is Lorentz covariant individually. But when combined together, they are both Lorentz covariant. The same is true of (4b) and (4d). This demonstrates why the photon field needs to be constrained by (2b) for the optical Dirac equation (2a) to be relativistic.

As observed before, the transformation laws (19a) and (19b) for the photon field are the same as the transformation laws (16) for the electromagnetic field though the photon field and the electromagnetic field do not transform under the Lorentz transformation as the same geometric object. This ought to imply the spatially nonlocal relation [13, 14] between the spinorial photon field and the corresponding tensorial electromagnetic field. It is known in quantum mechanics that the expectation value of the photon Hamiltonian in a quantum state is given by [15]

\[
\langle H \rangle = i\hbar \int \zeta^* \frac{\partial \zeta}{\partial t} \, d^3x + i\hbar \int \eta^* \frac{\partial \eta}{\partial t} \, d^3x \quad (24)
\]
if the photon field is normalized as \( \int \Psi^\dagger \Psi \, d^3x = 1 \). According to the correspondence principle [16], this should be equal to the total energy of the corresponding electromagnetic field, which is expressed in terms of the electric and magnetic fields as

\[
E = \frac{\varepsilon_0}{2} \int \mathbf{E}^2 \, d^3x + \frac{\mu_0}{2} \int \mathbf{H}^2 \, d^3x.
\]

Due to the occurrence of the operator \( i\hbar \partial / \partial t \) in the integrands on the right side of (24), the photon field cannot be related to the electromagnetic field in a spatially local fashion. More importantly, as can be seen from (24), the spinorial photon field has precisely the dimensions of the square-root of a probability density. In a word, the spinor satisfying (4a)–(4d) or (2a) and (2b) indeed describes the photon field in accordance with quantum mechanics. It not only satisfies the principle of locality but also has the right dimensionality that is required by quantum mechanics.

5. Further discussions

We have thus successfully formulated a set of quantum-mechanical equations for the photon. It consists of two equations. The optical Dirac equation (2a) determines the photon dynamics. Equation (2b) appears as a constraint that makes the dynamical equation form-invariant under the Lorentz transformation. Although they can be cast into the form of Maxwell’s equations, the photon field is a spinorial field. It depends nonlocally on the electromagnetic field. To be honest, it was Cook [13] who first proposed the photon-field equations (4a)–(4d). Probably guided by the correspondence between (4a)–(4d) and Maxwell’s equations (5a)–(5d), he expressed the photon field as a 4 \times 4 antisymmetric matrix to discuss their Lorentz covariance and demonstrated that the photon field does not transform under the Lorentz transformation as a tensor. Later on Inagaki [14] reformulated these equations in terms of conventional quantum mechanics. He expressed the photon field as the 6 \times 1 column matrix \( \Psi \) and converted the time-dependent equations (4a) and (4b) into the Hamiltonian form (2a). But unfortunately, he did not seem to realize that this equation is not relativistic by itself and, consequently, was not able to find the spinor property of the photon field.

As mentioned before, Maxwell’s equations (5a)–(5d) are the same as the photon-field equations (4a)–(4d) in form. Upon expressing the electromagnetic field as a 6 \times 1 column matrix \( \Psi_M = \left( \begin{array}{c} \sqrt{\varepsilon_0} \mathbf{E} \\ \sqrt{\mu_0} \mathbf{H} \end{array} \right) \), by virtue of the correspondences (6), one can convert (5a)–(5d) into the following two equations,

\[
\begin{align*}
\frac{i\hbar}{\partial t} \Psi_M &= H \Psi_M, \\
(\mathbf{\Gamma} \cdot \mathbf{p})^2 \Psi_M &= m^2 \Psi_M.
\end{align*}
\]

They are identical with the photon-field equations (2a) and (2b) in form. Mathematically, it can be expected that under the Lorentz transformation (15), they are form-invariant if the column matrix is assumed to transform as follows,

\[
\Psi_M'(x', t') = \Lambda \Psi_M(x, t),
\]

where \( \Lambda \) is given by (18). But this does not mean that the column matrix \( \Psi_M \) in (25a) and (25b) can be viewed physically as a spinor as was discussed in [17]. Here are the reasons.

As discussed at the end of last section, in order for the photon field \( \Psi \) in (2a) and (2b) to be a spinor, it must be nonlocally related to the electromagnetic field. Nevertheless, the column matrix \( \Psi_M \) is essentially the electromagnetic field itself though (25a) and (25b) look the same as (2a) and (2b). It therefore cannot be a spinor. As a matter of fact, the electromagnetic field transforms under the Lorentz transformation as a tensor. It is on that basis that Cook [13] made use of the nonlocal relation between the photon field and the electromagnetic field to show that the photon field cannot be a tensor. We are now clear that the photon field is a spinor. On this basis, it is straightforward to follow Cook’s method to show that the electromagnetic field \( \Psi_M \) cannot be a spinor though equations (25a) and (25b) are identical with equations (2a) and (2b) in form. By the way, the photon field discussed here is to be distinguished from the four-component spinorial wave function [18–20] that is also related to the electromagnetic field in a spatially local fashion [4].
6. Conclusions and remarks

In conclusion, we showed that the so-called optical Dirac equation (2a) by Barnett can indeed be interpreted as the quantum-mechanical dynamical equation for the photon so long as the relativistic condition (2b) is considered into account. These two equations can be cast into (4a)–(4d), which are exactly the same as Maxwell’s equations (5a)–(5d) in form if the correspondences (6) are assumed. More importantly, the photon field transforms under the Lorentz transformation as a spinor in much the same way as the electron field in the Dirac equation (1). In particular, the generator $\Gamma_0$ of its infinitesimal Lorentz transformation is analogous to the generator $\gamma_0\gamma/2$ of the infinitesimal Lorentz transformation of the electron field. The spinor property of the photon field makes (2a) and (2b) serve as the genuine quantum-mechanical equations for the photon. It not only meets the need of the principle of locality but also allows the photon field to have the dimensions of the square-root of a probability density as is required by quantum mechanics. Besides, we also discussed that although Maxwell’s equations (5a)–(5d) about the electromagnetic field can be cast into (25a) and (25b), which look the same as (2a) and (2b) about the photon field, $\Psi_M$ cannot be a spinor.

The photon field is related to the electromagnetic field in a spatially nonlocal fashion. Traditionally, the nonlocal relation of the photon field with the electromagnetic field was regarded as a serious drawback for the theory [3–5]. Now we see that it just reflects the difference between the photon and electromagnetic fields in geometric object. Considering that the optical Dirac equation (2a) as the dynamical equation for the photon is not relativistic by itself, the implications of such a nonlocal relation can also be elucidated from the perspective of the relativistic condition (2b), a constraint that is unusual in conventional quantum mechanics. For example, from the optical Dirac equation (2a) one can easily derive the following continuity equation,

$$\frac{\partial J_\mu}{\partial x_\mu} = 0,$$

where

$$J_\mu = i c \bar{\Psi} \Gamma_\mu \Psi = i c \Psi^\dagger \Gamma_0 \Gamma_\mu \Psi,$$

$\bar{\Psi} = \Psi^\dagger \Gamma_0$, and the superscript $\dagger$ denotes conjugate transpose. The relativistic condition (2b) ensures that this equation is form-invariant under the Lorentz transformation. In other words, the quantum constraint (2b) guarantees $J_\mu$ to be a Lorentz four-vector. According to conventional quantum mechanics, if the photon field were not constrained by (2b), the positive-definite entity $\Psi^\dagger \Psi$ that corresponds to the time component $J_0 = i c \Psi^\dagger \Psi$ could be interpreted as the probability density of the photon in spatial space. But without (2b), there would be no quantum mechanics for the photon. So the roles of the constraint equation (2b) in quantum mechanics, especially its effects on the physical interpretation of the spinorial photon field $\Psi$, deserve more attention.

Acknowledgments

This work was supported in part by Science and Technology Commission of Shanghai Municipality (18ZR1415500).

Data availability statement

No new data were created or analysed in this study.

ORCID iDs

Xiao-Lu You https://orcid.org/0000-0002-2964-3291
Chun-Fang Li https://orcid.org/0000-0001-6897-8097

References

[1] Akhiezer A I and Berestetskii V B 1965 Quantum Electrodynamics (New York: Interscience)
[2] Bialynicki-Birula I and Bialynicka-Birula Z 2017 Quantum-mechanical description of optical beams J. Opt. 19 125201
[3] Pauli W 1980 General Principles of Quantum Mechanics (Berlin: Springer)
[4] Bialynicki-Birula I 1996 Photon wave function Prog. Opt. 36 245–94
[5] Keller O 2005 On the theory of spatial localization of photons Phys. Rep. 411 1–232
[6] Weinberg S 1995 *The Quantum Theory of Fields* vol 1 (Cambridge: Cambridge University Press)
[7] Dirac P A M 1947 *The Principle of Quantum Mechanics* 3rd edn (Oxford: Oxford University Press)
[8] Rose M E 1961 *Relativistic Electron Theory* (New York: Wiley)
[9] Barnett S M 2014 Optical Dirac equation *New J. Phys.* **16** 093008
[10] Darwin C G 1932 Notes on the theory of radiation *Proc. R. Soc. A* **136** 36–52
[11] Pryce M H L 1948 The mass-centre in the restricted theory of relativity and its connexion with the quantum theory of elementary particles *Proc. R. Soc. A* **195** 62–81
[12] Stratton J A 1941 *Electromagnetic Theory* (New York: McGraw-Hill)
[13] Cook R J 1982 Lorentz covariance of photon dynamics *Phys. Rev. A* **26** 2754–60
[14] Inagaki T 1994 Quantum-mechanical approach to a free photon *Phys. Rev. A* **49** 2839–43
[15] Sakurai J J 1985 *Modern Quantum Mechanics* (New York: Benjamin-Cummings)
[16] Merzbacher E 1998 *Quantum Mechanics* 3rd edn (New York: Wiley)
[17] Wang Z-Y, Xiong C-D and Qiu Q 2009 Photon wave function and Zitterbewegung *Phys. Rev. A* **80** 032118
[18] Laporte O and Uhlenbeck G E 1931 Application of spinor analysis to the Maxwell and Dirac equations *Phys. Rev.* **37** 1380–97
[19] Oppenheimer J R 1931 Note on light quanta and the electromagnetic field *Phys. Rev.* **38** 725–46
[20] Moses H E 1959 Solution of Maxwell’s equations in terms of a spinor notation: the direct and inverse problem *Phys. Rev.* **113** 1670–9