INTRODUCTION

Quantization of the electromagnetic field is traditionally introduced at the level of second quantization: the classical field variables are replaced by field operators. I believe that the reasons why a first-quantized theory of photons has never been fully developed are mainly historical. Had Dirac discovered his relativistic wave equation [1] prior to his quantization of the electromagnetic field [2], he would have noticed and most probably further explored a great similarity between the wave equation for the electron (or even better for the neutrino) and the Maxwell equations. As it happened, this similarity was noticed later (for the first time apparently by Majorana [3]) and played no role in the development of the quantum theory of electromagnetism because the quantized electromagnetic field has been introduced from the very beginning and accounted for all quantum properties electromagnetic radiation. Subsequently quantum electrodynamics has become so successful in explaining with utmost accuracy all experiments within its range of applicability that there was no need to search for an alternative formulation that would employ the concept of the photon wave function. Considering our trust in quantum electrodynamics and our familiarity with its formal apparatus one may even ask if there is any justification at all for, what it essentially amounts to, a reconstruction of the notion of the photon function from QED, only to face a not so familiar object whose properties are yet to be uncovered.

I can see three reasons why the introduction of a photon wave function is worth the effort. First of all, one achieves a unified description of all particles (massive and massless) at both levels (first and second) of quantization. This is particularly attractive from the pedagogical point of view since one can use the same mathematical tools to study the quantum properties of massive and massless particles. Second, various aspects of the photon dynamics (for example, the angular momentum of the photon, evanescent waves, or the propagation of photons in noninertial frames of reference) can be described much more easily within the framework of photon wave mechanics without bringing in the formalism of field quantization. Third, armed with the concept of the photon wave function and all the tools of the standard wave mechanics, one can come up with some new methods of describing photons. For example, one can study quantum mechanical eigenvalue problems for the photons, one can introduce an analog of the Wigner function for the photon, and one can find a counterpart of the hydrodynamic formulation of wave mechanics for photons. Much of what I will say here is not new; it can be obtained by "reverse engineering" from textbooks on quantum electrodynamics. There will be, however, a significant change in emphasis leading to some new conclusions.

Photon wave functions appear in disguise in the standard formulation of quantum electrodynamics as the coefficients (called mode functions) in an expansion of the electromagnetic field operators into photon creation and annihilation operators. It takes some effort to recognize genuine wave functions
in these coefficient functions since the mode functions are always assumed to describe monochromatic radiation [4–7]. Thus, contrary to the spirit of wave mechanics, arbitrary superpositions of the mode functions are forbidden since they, in general, lead to nonstationary wave packets. This restriction may, however, be removed (see for example [8]); the only essential property of the coefficient functions being that they form a complete (and preferably also an orthonormal) set.

In this paper, I shall argue that not only one may reconstruct photon wave functions from standard quantum electrodynamics but that one can set up a consistent wave mechanics of photons that could be used to describe various quantum effects, independently of the formalism of second quantization. In other words, in constructing a full quantum theory of photons one may also proceed, as in quantum theory of massive particles, through two stages. At the first stage one introduces wave functions and a wave equation obeyed by these wave functions. At the second stage one upgrades the wave functions to the level of field operators in order to deal more effectively with states involving many particles and to allow for processes in which the number of particles is not conserved.

The very concept of the photon wave function is not new, but strangely enough it has never been systematically explored. Many textbooks on quantum mechanics start the introduction to quantum theory with a discussion of photon polarization measurements (cf. for example [9–12]) but a complete photon wave function never makes its appearance, as if the notion of a wave function was restricted only to the description of the polarization states in a simple two-dimensional Hilbert space and could not describe the photon propagation in full. In some textbooks (cf. for example [13–15]) the possibility of introducing a space-dependent wave function for the photon is even explicitly rejected.

I shall introduce a wave function for the photon by reviving and extending the mode of description of the electromagnetic field based on the complex form of the Maxwell equations known already to Riemann [16,17]. The complex vector that appears in this description will be shown to have the properties that one would associate with a one-photon wave function, including of course a probabilistic interpretation. Additional properties of the photon wave function are described in my recent article [18] while its relation to full quantum electrodynamics is discussed in our book [8].

The approach adopted here is to be contrasted with that of Landau and Peierls [19] and Cook [20]. The Landau-Peierls and Cook wave functions are highly nonlocal objects. The nonlocality in space is introduced by dividing the Fourier transform of the electromagnetic field by \( \sqrt{k} \) (cf. Eq. (43)). It has been already noted by Pauli [21], that these nonlocal wave functions have a serious drawback since they do not transform under Lorentz transformations as well-defined local geometric objects; their values taken at a point in one coordinate system depend on all their values in another coordinate system.

PHOTON WAVE FUNCTION AND ITS TIME EVOLUTION

My discussion of the photon wave function begins with writing the Maxwell equation for a homogeneous medium in a compact form

\[
i \partial_t \mathbf{F}(\mathbf{r}, t) = c \nabla \times \mathbf{F}(\mathbf{r}, t),
\]

\[
\nabla \cdot \mathbf{F}(\mathbf{r}, t) = 0,
\]

by combining, as Riemann have done, two real vectors \( \mathbf{D} \) and \( \mathbf{B} \) into one complex-valued vector \( \mathbf{F} \),

\[
\mathbf{F}(\mathbf{r}, t) = \frac{1}{2} \left( \frac{\mathbf{D}(\mathbf{r}, t)}{\sqrt{\epsilon_0}} + i \frac{\mathbf{B}(\mathbf{r}, t)}{\sqrt{\mu_0}} \right).
\]

The square roots of \( \epsilon \) and \( \mu \) are needed to match the dimensions of the two terms and an additional factor of \( 1/2 \) is introduced for future convenience.

The evolution equation (1) for \( \mathbf{F} \) and the transversality condition (2) can also be written in a matrix form

\[
i \partial_t \mathbf{F} = c \left( s \cdot \frac{1}{i} \nabla \right) \mathbf{F},
\]

\[
(s \cdot \nabla) s_j \mathbf{F} = \nabla_j \mathbf{F},
\]

where \( s_i \)’s are the spin matrices for a spin-1 particle

\[
(s_i)_{kl} = -i \varepsilon_{ikl},
\]
$s_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}$,  $s_y = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}$,  $s_z = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.  (7)

Upon the multiplication of both sides of Eq. (4) by $\hbar$, one can convert it to the Schrödinger form

$$i\hbar \partial_t \phi = H\phi \equiv c(s \cdot p)\phi,$$  (8)

where $p = (h/i)\nabla$. The scalar product $cs \cdot p$ plays the role of the Hamiltonian in wave mechanics of the photon. Note, that $cs$ may also be interpreted as a velocity operator since $(i/h)[r, H] = cs$.

The form (8) of the wave equation is quite universal. Replacing the matrices $s$ by the Pauli matrices one obtains the Weyl equation for the wave function of a neutrino [22]. With the appropriate choice of spin matrices, the wave equations for higher-spin massless particles can also be cast into this form (cf. for example [23]). Upon the identification of the components of the vector $F$ with components of a symmetric relativistic spinor $\phi_{AB}$

$$\phi_{00} = -F_x + iF_y, \quad \phi_{01} = F_z, \quad \phi_{11} = F_x - iF_y,$$  (9)

Maxwell equations become a member of a universal set of relativistic wave equations describing the propagation of massless particles of any spin. All these equations have the form [23]

$$\sigma^{\mu A} \nabla_\mu \phi_{AB;1} \cdots B_{n-1} = 0,$$  (10)

where four matrices $\sigma^{\mu A}$ represent a relativistic extension of Pauli matrices ($\sigma^0$ is the unit matrix and the remaining ones are ordinary Pauli matrices). This universality provides a strong argument in favor of interpreting the solutions of Eq. (4) as photon wave functions. After all one has no qualms about accepting the solutions of the Weyl equation as neutrino wave functions.

Eqs. (4) and (5) are completely equivalent to the Maxwell equations but when treated as the equations for the photon wave function they suffer from one serious shortcoming: they possess only half of the positive-energy solutions needed to describe all polarization states of a photon. The same is true in the general case; Eqs. (10) describe only left-handed particles.

Positive energy solutions are identified as those solutions whose wave function has the form $\exp(-i\omega t)F(r)$ with positive values of $\omega$. The substitution of this form into (4) yields

$$c(s \cdot \frac{\hbar}{i} \nabla)\phi = \hbar \omega \phi.$$

(11)

This equation says that for positive energy solutions the helicity, i.e. the projection of the spin of the photon on the direction of its momentum, is always positive. The sign of helicity can be traced back to the choice of either $F$ or its complex conjugate as the photon wave function. Thus, the choice of a sign of the imaginary part in (3) is equivalent to choosing one helicity (right-handed or left-handed) over the other. One needs, however, wave functions of the photons of both helicities to describe both circular polarizations: right and left. This requires a doubling of the wave-function components: the two helicities will be described by upper and lower components of one wave function. In empty space both polarization states propagate independently; the two parts of the photon wave function satisfy two separate evolution equations with opposite signs of the Hamiltonian. By introducing a six-component wave function $F$ made of $F_{\pm}$ as its upper/lower components,

$$F = \begin{pmatrix} F_+ \\ F_- \end{pmatrix},$$  (12)

the two evolution equations can be combined together

$$i\hbar \partial_t F = c\rho_3(s \cdot \frac{\hbar}{i} \nabla)F,$$  (13)

where $\rho_3$ is a member of a set of Pauli-type matrices $\rho_\mu$. The matrix $\rho_3$ gives plus/minus sign when acting on upper/lower components while $\rho_1$ and $\rho_2$ (to be used later) interchange these components.

The doubling of the components of the electron wave function in the relativistic case, as compared to the nonrelativistic case, is justified by the existence of antiparticles. Antiparticles are described by the negative energy solutions of the Dirac equation. Photons, however, do not have antiparticles so that the negative energy part of the wave function should not carry any additional information. The doubling of the components of the wave function (12) requires, therefore, an auxiliary condition that restores the original number of degrees of freedom. This is achieved by demanding that the complex conjugation of the wave function has the same effect as an interchange of upper and lower components,
\[ F = \rho_1 F^*. \] (14)

As an independent quantity one may take just the positive energy part \( F^+ \) of the wave field \( F \) (called the analytic signal in classical theory). This part represents the true photon wave function and that part is to be identified with the matrix elements of the field operators between the vacuum state and one-photon states that make their appearance in quantum electrodynamics.

In what follows I shall use the symbol \( \Psi \) to denote the positive energy part of the function \( F \). Owing to the condition (14), the negative frequency part is obtained by complex conjugation and by an interchange of the upper and lower components of the positive frequency part. In the language of particle physics, complex conjugation combined with multiplication by \( \rho_1 \) would be called charge conjugation and Eq. (14) is the condition of invariance under charge conjugation.

Maxwell equations written in the form (13) exhibit a close analogy with the Dirac equation written in the chiral representation of the Dirac matrices (cf. for example [24,8]). In this representation the four-component wave function \( \psi \) is made of two relativistic spinors \( \phi \) and \( \chi \) — the analogs of \( F^+ \) and \( F^- \) — describing two helicities,

\[ \psi = \begin{pmatrix} \phi \\ \chi \end{pmatrix}. \] (15)

The Dirac equation for in this representation takes the form

\[ i\hbar \partial_t \psi = c \rho_3 \left( \sigma \cdot \frac{\hbar}{i} \nabla \right) \psi + mc^2 \rho_1 \psi. \] (16)

For the Dirac particle the mass term always induces a mixing of the two states of polarization while for the photons, as I shall show later, the mixing is induced only by an external influence, by the medium.

**CONSERVED QUANTITIES AND CORRESPONDING QUANTUM OPERATORS**

I shall try now to make the probabilistic interpretation of the photon wave function \( \Psi \) plausible by comparing the expressions for basic observables in classical and in quantum theories. The fundamental conserved quantities characterizing classical electromagnetic field — the energy \( E \), the momentum \( P \), and the angular momentum \( M \) — are given as the following space integrals

\[ E = \int d^3r \mathcal{H}(r), \quad P = \int d^3r \mathcal{P}(r), \quad M = \int d^3r \mathbf{r} \times \mathcal{P}(r), \] (17)

that involve the local densities of energy \( \mathcal{H}(r) \) and momentum \( \mathcal{P}(r) \)

\[ \mathcal{H} = \frac{\mathbf{D} \cdot \mathbf{D}}{2\epsilon} + \frac{\mathbf{B} \cdot \mathbf{B}}{2\mu} = \mathcal{F}^\dag \cdot \mathcal{F}, \quad \mathcal{P} = \mathbf{D} \times \mathbf{B} = \mathcal{F}^\dag \rho_3 \mathcal{F}/c. \] (18)

The conservation laws for \( E, P, \) and \( M \) follow from the continuity equations satisfied by \( \mathcal{H}(r) \) and \( \mathcal{P}(r) \),

\[ \partial_t \mathcal{H} = -\nabla \cdot S, \quad \partial_t \mathcal{P}_i = -\nabla_k T_{ik}, \] (19)

where \( S \) is the energy flux and \( T_{ij} \) is the Maxwell stress tensor,

\[ S = c \mathcal{F}^\dag \rho_3 \mathcal{F}, \quad T_{ij} = \mathcal{F}^\dag (s_i s_j + s_j s_i) \mathcal{F} - \delta_{ij} \mathcal{F}^\dag \cdot \mathcal{F}, \] (20)

provided the wave function is localized in space so that the integration by parts does not introduce any boundary terms.

Now, I shall analyze the energy, momentum, and the angular momentum from a quantum mechanical point of view applying the same methods as in wave mechanics of massive particles, by exhibiting the corresponding operators, their eigenvalues and eigenfunctions. The quantum operators representing these observables are taken to be (cf. for example [9]) the generators (up to a factor of \( \hbar/\imath \)) of the appropriate infinitesimal transformations of the wave function. The time translation gives the Hamiltonian or the energy, space translations give the momentum, and rotations give the angular momentum.

The momentum operator \( \hat{p} \) and its eigenvalue problem have exactly the same form as in quantum mechanics of massive particles

\[ \hat{p} = \frac{\hbar}{\imath} \nabla, \quad \frac{\hbar}{\imath} \partial_t \Psi = \hbar k_i \Psi, \] (21)
where $k_i$'s are the components of a wave vector.

The time evolution Eq. (13) enables one to identify the quantum mechanical Hamiltonian operator $\hat{H}$ as,

$$\hat{H} = c\rho^3 (\hat{s} \cdot \hat{p}).$$  \hspace{1cm} (22)

The eigenvalue problem for this operator has the form

$$c\rho^3 (s \cdot \nabla) \Psi = \hbar \omega \Psi.$$  \hspace{1cm} (23)

The total angular momentum operator for spinning particles consists of two parts: the orbital angular momentum and the spin angular momentum,

$$\hat{J} = \hat{r} \times \hat{p} + \hbar \hat{s}.$$  \hspace{1cm} (24)

It is worth noting that it is the total angular momentum $\hat{J}$, defined as the sum of the orbital and the spin part, that commutes with the photon Hamiltonian (22) and, therefore, is a constant of motion. The eigenvalue problem for the angular momentum contains, as usual, the eigenvalue problem for the $z$-component of the total angular momentum

$$\hat{J}_z \Psi = \hbar M \Psi,$$  \hspace{1cm} (25)

and the eigenvalue problem for the square of the total angular momentum

$$\hat{J}^2 \Psi = \hbar^2 J(J + 1) \Psi.$$  \hspace{1cm} (26)

The solutions of Eqs. (25) and (26) are well known vector spherical harmonics (cf. for example [25,26]).

At first look, the classical expressions (17) for the energy, momentum, and angular momentum are different from their quantum mechanical counterparts. In the next section I shall show that these differences can be reconciled by adopting an appropriate metric for the evaluation of expectation values.

**EXPECTATION VALUES AND TRANSITION PROBABILITIES**

A correspondence between quantum and classical descriptions of the radiation field requires that the expectation values of the quantum-mechanical operators $\hat{H}$, $\hat{p}$, or $\hat{M}$ be equal to the classical values of the field energy, momentum, and angular momentum. I shall show that this can be achieved by an appropriate choice of the scalar product for the wave functions $\Psi$ used in the construction of expectation values. To this end, I invoke the transversality condition (5) for $\Psi$ to obtain the following relations

$$\hat{H} \rho^3 s/c \Psi = \hat{p} \Psi, \quad \hat{H} \rho^3 \hat{r} \times s/c \Psi = (\hat{r} \times \hat{p} + s) \Psi.$$  \hspace{1cm} (27)

With the help of these relations (after a division by $\hat{H}$), one can write the classical quantities (17) in the form of expectation values of the corresponding quantum-mechanical operators

$$E = \langle E \rangle, \quad P = \langle P \rangle, \quad M = \langle M \rangle,$$  \hspace{1cm} (28)

where

$$\langle E \rangle = \int d^3 r \Psi^\dagger \frac{\hat{H}}{H} \Psi, \quad \langle P \rangle = \int d^3 r \Psi^\dagger \frac{\hat{p}}{H} \Psi, \quad \langle M \rangle = \int d^3 r \Psi^\dagger \left( \hat{r} \times \hat{p} + s \right) \Psi.$$  \hspace{1cm} (29)

All these expressions are bilinear in $\Psi$, like all expectation values in wave mechanics of massive particles, and the only difference is a systematic appearance of the Hamiltonian in the denominator. Therefore, it is natural to assume that the division by the Hamiltonian should be included in the definition of the expectation values or in the definition of the scalar product $\langle \Psi_1 | \Psi_2 \rangle_H$ of two wave functions,

$$\langle \Psi_1 | \Psi_2 \rangle_H = \int d^3 r \Psi_1^\dagger \frac{1}{H} \Psi_2.$$  \hspace{1cm} (30)

I have added the subscript $H$ in the symbol of the scalar product as a reminder of the necessary division by $H$. With this definition of the scalar product, one obtains

$$\langle E \rangle = \langle \Psi | \hat{H} \Psi \rangle_H, \quad \langle P \rangle = \langle \Psi | \hat{p} \Psi \rangle_H, \quad \langle M \rangle = \langle \Psi | (\hat{r} \times \hat{p} + s) \Psi \rangle_H.$$  \hspace{1cm} (31)
The uniqueness of this definition of the scalar product can be proven in the framework of geometric quantization along similar lines [27] as has been done for the scalar field [28,29].

The scalar product of the wave function with itself — the expectation value of the unit operator — gives the squared norm of the wave function  \( \langle \Psi | \Psi \rangle = \langle \Psi | \hat{\mathcal{H}} | \Psi \rangle \). Note that the norm of the wave function is a dimensionless quantity, a pure number. Therefore, by multiplying a photon wave function by an appropriate numerical factor one may always normalize the wave function to 1. Such a normalization is suitable for one-photon states. In general, for unnormalized wave functions one may interpret the norm squared as the total number of photons  \( N \),

\[
N = \int d^3r \, \Psi^\dagger \frac{1}{\hat{\mathcal{H}}} \Psi. \tag{32}
\]

According to general rules of quantum theory the scalar product  \( \langle \Psi_1 | \Psi_2 \rangle \) of two photon wave functions (both normalized to 1) is a transition amplitude. Its squared absolute value is the probability to find the photon in the state  \( \Psi_2 \), when the photon is known to be in the state  \( \Psi_1 \).

Effective evaluation of expectation values and transition probabilities requires a diagonalization of the Hamiltonian. Only then one can give a meaning to the operation of a division by  \( \hat{\mathcal{H}} \). I shall perform now this task in the simplest case of a photon propagating in empty space — the free field case.

In this case, the Hamiltonian eigenvalue problem for the photon wave function can be easily solved by Fourier transformation, since it reduces to a set of algebraic equations for the Fourier components  \( \tilde{\Psi} \) of the wave function,

\[
\Psi(r) = \int \frac{d^3k}{(2\pi)^3} e^{ik \cdot r} \tilde{\Psi}(k), \tag{33}
\]

\[c_{\rho \lambda}(s, k) \tilde{\Psi} = \omega \tilde{\Psi}. \tag{34}\]

The solution of these equations can be expressed in terms of two vectors  \( f_+ \) and  \( f_- \) representing the upper and lower components of  \( \tilde{\Psi} \). In vector notation, the equations for these components resulting from the eigenvalue problem (34) are

\[i c k \times f_\pm = \pm \omega f_\pm. \tag{35}\]

Note that the solutions of these equations automatically satisfy the condition  \( k \cdot f_\pm = 0 \) that expresses the transversality condition (2) in terms of the Fourier transform. For positive values of  \( \omega \) there is just one solution for each sign. The corresponding complex normalized vectors will be denoted by  \( e_\pm \),

\[e_+^\dagger \cdot e_+ = 1, \quad e_-^\dagger \cdot e_- = 1, \quad e_+^\dagger \cdot e_- = 0. \tag{36}\]

Finally, the Fourier transform of the photon wave function, satisfying the transversality condition, can be written in terms of two complex functions  \( f_\pm(k) \)

\[\tilde{\Psi}(k) = \begin{pmatrix} e_+ f_+(k) \\ e_- f_-(k) \end{pmatrix}. \tag{37}\]

The functions  \( f_\pm(k) \) are clearly the two components of the photon wave function in momentum space.

The division by  \( \hat{\mathcal{H}} \) can easily be performed in the Fourier representation and the expression for the average photon number  \( N \) reads

\[N = \int \frac{d^3k}{(2\pi)^3} \left| f_+(k) \right|^2 + \left| f_-(k) \right|^2. \tag{38}\]

This expression can be transformed back to the coordinate representation and yields the following nonlocal formula for the photon number  \( N \) (or the norm of the wave function) and the scalar product of two wave functions

\[N = \frac{1}{2\pi^2} \int \int d^3r_1 d^3r_2 \frac{\Psi^\dagger(r_1) \cdot \Psi(r_2)}{|r_1 - r_2|^2}, \tag{39}\]

\[\langle \Psi_1 | \Psi_2 \rangle = \frac{1}{2\pi^2} \int \int d^3r_1 d^3r_2 \frac{\tilde{\Psi}_1^\dagger(r_1) \cdot \tilde{\Psi}_2(r_2)}{|r_1 - r_2|^2}. \tag{40}\]

One must learn to live with these nonlocal expressions, because according to an old no-go theorem by Bargmann and Wigner [30], no local expression for the scalar product exists. Apart from nonlocality, the norm (39) has appealing properties: it is fully invariant not only under the Lorentz transformations but also under all conformal transformations [31]. The scalar product (40) has been recently used by Kaiser [32] as a basis for wavelet electrodynamics.
LOCAL VERSUS NONLOCAL WAVE FUNCTIONS

The scalar product for photon wave functions (40) can be written in a seemingly local form

$$\langle \Psi_1 | \Psi_2 \rangle_H = \int d^3r \Phi_1^\dagger(r)\Phi_2(r),$$

(41)

where the new functions $\Phi$ are related to the original wave functions $\Psi$ through the formula

$$\Phi(r) = \frac{1}{\sqrt{H}}\Psi(r).$$

(42)

In the free-field case the symbolic operation of dividing by square root of the Hamiltonian can be converted into an integral

$$\Phi(\vec{r}) = \frac{1}{2(2\pi)^{3/2}} \int \frac{d^3r'}{|\vec{r} - \vec{r}'|^{5/2}} \Psi(\vec{r}').$$

(43)

The new function $\Phi$ is exactly the wave function introduced in [19,20]. The relation (43) shows clearly that $\Phi$ is a highly nonlocal object. In the case of a photon moving in a medium, the formula (42) becomes purely symbolic and its evaluation requires the knowledge of a complete set of wave functions in a medium. Thus, there is no real advantage in using the nonlocal wave functions $\Phi$. Even though the scalar product looks simpler, the determination of the wave function becomes complicated and its physical interpretation is not clear at all. Essentially, it can be used only for calculations of scalar products. No simple interpretation can be given to $|\Phi(r)|^2$ evaluated at a point. It cannot be called a probability density since it does not even satisfy a continuity equation. On the other hand one may associate a direct probabilistic interpretation with the modulus squared $|\Psi(r)|^2$ of the local wave function. After a division by the expectation value of the energy it becomes a genuine probability density $\rho_E(r)$,

$$\rho_E(r) = \frac{|\Psi(r)|^2}{\langle E \rangle},$$

(44)

normalized to 1 and satisfying, on account of (19), a continuity equation. One is, therefore, justified in assuming that it describes in probabilistic terms the energy distribution in space associated with a photon [18]. All the differences between the two wave functions basically disappear for stationary, monochromatic states: both function become proportional to each other [18].

It is worthwhile to point out here that for massive particles the modulus squared of the wave function gives the generic probability to find a particle only in the simplest case of nonrelativistic wave mechanics. In the relativistic case there are various probabilities associated with the Dirac wave function. The probability to find mass is different from the probability to find charge, and that is different from the probability to find energy of a particle.

WAVE MECHANICS OF PHOTONS IN A MEDIUM

For the electromagnetic field in a medium characterized by constant values of $\epsilon$ and $\mu$, the Maxwell equations give rise to the same simple evolution equation for the wave function as in free space, except that $\epsilon_0$ and $\mu_0$ in (3) are replaced by their values in a medium. The “big” wave function $\Psi$ introduced to account for the two states of polarization satisfies a linear wave equation in an arbitrary linear medium. In free space, in an isotropic homogeneous medium, or in a gravitational field the two states of polarization propagate independently resulting in two independent wave equations. In an inhomogeneous medium, however, the two helicities mix and the corresponding wave equations become coupled. I shall write down the wave equation in an inhomogeneous medium in the form (upper and lower parts of $\Psi$ are treated as three-dimensional vectors to make use of the nabla operation)

$$i(\partial_t + \partial_t L)\Psi = c\rho_3(\nabla + \nabla L) \times \Psi,$$

(45)

where the speed of light $c$ is constructed from the local values of permittivity and permeability,

$$c = 1/\sqrt{\epsilon \mu},$$

(46)

and the matrix $L$ has the form,

$$L = \frac{1}{2} \left( \ln \sqrt{\epsilon \mu} + \rho_1 \ln \sqrt{\frac{\epsilon}{\mu}} \right).$$

(47)
One can check that the operator on the right hand side of Eq. (45) is Hermitian so that it can be viewed as a quantum mechanical Hamiltonian in the presence of an inhomogeneous medium. The transversality condition (2) in an inhomogeneous medium is also modified and it takes on the form,

\[(\nabla + \nabla L) \cdot \mathcal{F} = 0. \tag{48}\]

Note that the speed of light \(1/\sqrt{\epsilon \mu}\) may vary (as it does in the gravitational field) without causing the mixing of polarizations. It is only the varying “conductance of the medium” \(\sqrt{\epsilon/\mu}\) (the sole justification for the use of this name is the right dimensionality of 1/Ohm) that mixes the two states of polarization.

The wave equation (45) and the transversality condition (48) are invariant, as was the case in empty space, under the charge symmetry operation. Therefore, we can impose now the same restriction (14) as in empty space on the photon wave function that will guarantee that also photons propagating in a medium will not have distinct antiparticles.

The photon wave equations in an inhomogeneous medium do not look very appealing but that is due to a phenomenological character of macroscopic electrodynamics. The propagation of a photon in a medium is a succession of absorptions and subsequent emissions of the photon by the charges that form the medium. The number of photons of a given helicity is, in general, not conserved in these processes and that accounts for all the complications. The photon wave equations in an inhomogeneous medium describes, in actual fact, the propagation of some collective excitations of the whole system and not just the propagation of pure photons.

### BOUND STATES OF A PHOTON IN AN OPTICAL FIBER

I shall show now that the notion of the photon wave function enables one to describe the propagation of a photon in an optical fiber in close analogy with Schrödinger wave mechanics of massive particles moving in a constant magnetic field. Let me consider an infinite optical fiber of diameter \(a\) characterized by a dielectric permittivity \(\epsilon\). The eigenvalue problem for the photon Hamiltonian operator can be treated as that for a particle moving in a potential well. Owing to the symmetry of the problem, one might include in the set of commuting operators, in addition to the Hamiltonian, also the projections of the momentum operator and the total angular momentum on the direction of fiber axis. In cylindrical coordinates the eigenvalue equations for the \(z\)-component of momentum and angular momentum have the form

\[-i\partial_z \Psi = k_z \Psi, \tag{49}\]

\[-i\partial_\phi \Psi = M \Psi. \tag{50}\]

Assuming that the photon wave function belongs to the eigenvalues \(\hbar k_z\) and \(\hbar M\) of these operators, one can separate out the dependence on \(z\) and \(\phi\)

\[\Psi = \exp(ik_z z) \exp(iM\phi)f(\rho). \tag{51}\]

In order to write down the eigenvalue equation for the Hamiltonian (in a homogeneous medium) in terms of the components of \(f\) one needs only the formula for the curl in cylindrical coordinates (cf. for example [33]). The result can be written in the form

\[-\frac{M}{\rho} f_z + k_z f_\phi = \frac{\omega}{c} (i f_\rho), \tag{52}\]

\[-\partial_\rho f_z + k_z (i f_\phi) = \frac{\omega}{c} f_\rho, \tag{53}\]

\[\frac{1}{\rho} \partial_\rho \rho f_\phi - \frac{M}{\rho} (i f_\rho) = \frac{\omega}{c} f_z. \tag{54}\]

Every solution of these equations automatically satisfies the divergence condition (2). Eqs. (52–54) lead to a Bessel equation for \(f_z\)

\[\left[\partial_z^2 + \frac{1}{\rho} \partial_\rho - \frac{m^2}{\rho^2} + k_z^2\right] f_z = 0, \tag{55}\]

where \(k_z^2 = \omega^2/c^2 - k_\perp^2\). Photon wave functions obey Eq. (55) inside the fiber with one value of \(k_\perp\) and in the surrounding free space with different values of \(k_\perp\) since the value of \(c\) is different in these two regions. Depending on whether \(k_\perp\) is real or imaginary a general solution of this equation is a linear combinations of Bessel functions of the first kind \(J_M(\rho)\) and the second kind \(Y_M(\rho)\) or a linear
combinations of modified Bessel functions $I_M(\rho)$ and $K_M(\rho)$. In full analogy with the problem of a potential well in quantum mechanics, one can search for bound states in the transverse direction by matching a regular oscillatory solution inside (that means the $J_M(\rho)$ function) with an exponentially damped solution outside the fiber (that means the $K_M(\rho)$ function). The matching conditions, well known from classical electromagnetic theory, are the continuity conditions for the $E_z$ and $H_z$ field components at the surface of the fiber, when $\rho = a$. Bound states are possible because the speed of light is greater in the vacuum than inside the fiber and therefore it may happen that $k_\perp$ is real inside and imaginary outside the fiber. Since there are two matching conditions and only one ratio of the amplitudes inside and outside the fiber, both conditions can be satisfied only for a set of discrete eigenvalues of the photon energy $\hbar \omega$. It is worth noting that in order to have an imaginary $k_\perp$ one must have a nonvanishing $k_z$. Thus, a photon may well be bound in the plane perpendicular to the fiber, but it is always moving freely along the fiber in full analogy with a charged particle moving in a homogeneous magnetic field.

CONCLUSIONS

I have shown here how to develop a formalism of “first quantization” to describe photons. This formalism is so similar to that of standard wave mechanics that various problems encountered in quantum optics can be handled by analogy with quantum mechanics. No new physical results were obtained here, because the photon wave function is already implicitly present, but very seldom explicitly recognized, in the standard formulation of quantum electrodynamics. The identification of the photon wave function enables one to bring the quantum theory of photons into line with theories of other elementary particles. A distinctive feature of the photon wave functions is that there is no local expression for the probability density and that the evaluation of the scalar product requires, in general, a diagonalization of the Hamiltonian. Despite an apparent nonlocality associated with the scalar product, the photon wave function described here is a perfectly local object. It transforms as a local field not only under Lorentz but also under conformal transformations and it describes the propagation of photons in a given gravitational field [18] in a local manner.

Finally, I would like to make the following disclaimer. It has not been my intention to question the fact that the quantized theory of electromagnetic field describes fully and correctly wave-mechanical properties of photons. I just wanted to inform the quantum optics community that there exists also a formalism describing a wide class of phenomena involving quantum properties of photons that does not require field quantization. I do believe that the photon wave function is a very useful concept. It is true that it has its limitations, but the same holds for the Schrödinger wave function of, say, an electron; a comprehensive description requires the quantization of the electron field.

ACKNOWLEDGMENTS

I would like to thank Wolfgang Schleich for hospitality and the Alexander von Humboldt Foundation for making my stay at the University of Ulm possible.