

# On the Wave Function of the Photon

Iwo Białynicki-Birula

*Centrum Fizyki Teoretycznej, Polska Akademia Nauk  
Lotników 32/46, 02-668 Warsaw, Poland*

It is shown that the wave function of the photon is a very useful concept with interesting ramifications. It plays a major role in bridging the gap between classical theory of electromagnetism and quantum electrodynamics. The photon wave function can also be used to construct a phase-space distribution for the photon (Wigner function) that describes the transport of radiation.

Published in Acta Physica Polonica **86**, 97-116 (1994).

## I. INTRODUCTION

It is believed that certain matrix elements of the electromagnetic field operators in quantum electrodynamics, in close analogy with nonrelativistic quantum theory of massive particles, may be treated as photon wave functions. In this paper, I would like to push this interpretation even further by arguing that one can set up a consistent wave mechanics of photons that could be often used as a convenient tool in the description of electromagnetic fields, *independently* of the formalism of second quantization. In other words, in constructing quantum theories of photons we may proceed, as in quantum theory of massive particles, through two stages. At the first stage we introduce wave functions and a wave equation obeyed by these wave functions. At the second stage we upgrade the wave functions to the level of field operators in order to deal more effectively with the states involving many particles and to allow for processes in which the number of particles is not conserved. An important additional consequence of having a genuine wave function for the photon is a possibility to define an analog of the Wigner function for the photon with its semiclassical interpretation as a (quasi) distribution function in the phase space.

The very concept of the photon wave function is not new, but strangely enough it has never been systematically explored. Some textbooks on quantum mechanics start the introduction to quantum theory with a discussion of photon polarization measurements (cf., for example [1–4]). Dirac, in particular, writes a lot about the role of the wave function in the description of quantum interference phenomena for photons (in this context he uses the now famous phrase: “Each photon interferes only with itself”), but in his exposition the photon wave function never takes on a specific mathematical form.

It is true that in the description of polarization simple prototype two-component wave functions are often used to describe various polarization states of the photon and with their help the preparation and the measurement of polarization is thoroughly explained. However, after such a heuristic introduction to quantum theory, the authors go on to the study of massive particles and if they ever return to a quantum theory of photons it is always within the formalism of second quantization with creation and annihilation operators. In some textbooks (cf., for example [5–7]) one may even find statements that completely negate the possibility of introducing a wave function for the photon.

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. This form was known already at the beginning of the century [8,9] and was later rediscovered by Majorana [10] who explored the analogy between the Dirac equation and the Maxwell equations. 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 also an acceptable probabilistic interpretation. Photons are much different from massive particles. They are also different from neutrinos since the photon number does not obey a conservation law. These differences are a source of complications, especially when photons propagate in a medium. However, even these complications have a certain value: they teach us something new about the nature of photons.

The approach adopted in this paper is distinctly different from the line of investigation started by Landau and Peierls [11] and continued more recently by Cook [13]. The Landau-Peierls and Cook wave functions are nonlocal objects. The nonlocality is introduced by operating on the electromagnetic fields with the integral operator  $(-\Delta)^{-1/4}$ ,

$$((-\Delta)^{-1/4}f)(\mathbf{r}) = \pi \int \frac{d^3r'}{(2\pi|\mathbf{r}-\mathbf{r}'|)^{5/2}} f(\mathbf{r}'). \quad (1)$$

This integral operator changes the dimension from  $L^{-2}$  to  $L^{-3/2}$ , so that the modulus squared of the Landau-Peierls wave function may be interpreted as a probability density to find a photon. However, as has been already noted by Pauli [12], these nonlocal wave functions have serious drawbacks. First, they do not transform under Lorentz transformations as tensors or any other well-defined local geometric objects. Second, a nonlocal wave function taken at a point in one coordinate system depends on the values of this wave function everywhere in another coordinate system. Third, the probability density defined with the use of a nonlocal wave function does not correspond to the probability of interaction of the electromagnetic field with localized charges. Vanishing of the wave function at a definite point has in Pauli's words [12] "*no direct physical significance*" because the electromagnetic field does act on charges at the points where the probability to find a photon vanishes.

## II. AN ANALOGY BETWEEN THE MAXWELL EQUATIONS AND THE DIRAC EQUATION

I shall introduce the photon wave function and its evolution equation with the help of a heuristic argument used by Dirac [1,14] in his derivation of the relativistic theory of the electron. There are only two slight modifications that need to be made in Dirac's original argument. First, my point of departure will be the d'Alembert wave equation,

$$\left(\frac{1}{c^2}\partial_t^2 - \Delta\right)\phi = 0, \quad (2)$$

instead of the Klein-Gordon equation that describes the propagation of massive particles. Second, I shall not stop after finding the simplest relativistic wave equation, namely the one describing a massless spin- $\frac{1}{2}$  particle, but I shall continue a step further to obtain a relativistic equation for a massless spin-1 particle.

In order to implement this plan we need a set of three matrices, say  $\alpha_x, \alpha_y$ , and  $\alpha_z$  which satisfy the Dirac anticommutation relations

$$\alpha_i\alpha_j + \alpha_j\alpha_i = 2\delta_{ij} \mathbf{I}, \quad (3)$$

where  $\mathbf{I}$  is the unit matrix. For each set of such matrices we can take a square root of the Laplacian using the following formula that is a direct consequence of the relations (3)

$$\Delta \mathbf{1} = (\boldsymbol{\alpha} \cdot \nabla)^2. \quad (4)$$

The simplest choice, of course, is to take as the matrices  $\alpha_i$  the set of well known  $2 \times 2$  Pauli matrices. Under this choice we obtain the Weyl wave equation [15] for spin- $\frac{1}{2}$  particles describing the propagation of neutrinos

$$i\hbar\partial_t\phi = \frac{\hbar}{i}\boldsymbol{\sigma} \cdot \nabla\phi. \quad (5)$$

In order to obtain the wave equation for photons one might try to replace the Pauli matrices by the  $3 \times 3$  Hermitian matrices  $S_i$  representing infinitesimal rotations for spin-1 particles. These matrices, however, satisfy somewhat different anticommutation relations as compared to (3), with extra terms on the right hand side. One may write down these relations explicitly choosing a specific representation of spin matrices  $S_i$ . For our purpose it is convenient to use the representation in which the three components of spin-1 wave function transform under rotations as Cartesian components of a vector. In this representation the matrix elements of spin matrices can be expressed in terms of the antisymmetric Levi-Civita symbol  $\varepsilon_{ijk}$ ,

$$(S_i)_{kl} = -i\varepsilon_{ikl}, \quad (6)$$

$$S_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad S_y = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}, \quad S_z = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (7)$$

Matrix elements of the anticommutation relations for  $S_i$  are

$$[(S_i)(S_j) + (S_j)(S_i)]_{ab} = 2\delta_{ij}\delta_{ab} - \delta_{ai}\delta_{bj} - \delta_{aj}\delta_{bi}. \quad (8)$$

At first it might seem that extra terms on the right hand side of this equation completely rule out the choice of spin matrices  $S_i$  as candidates for the task of taking the square root of the Laplacian. However, we should remember that photons, being massless, are quite different from massive particles with spin 1. They have only two linearly

independent states of polarization (for example, they are left-handed or right-handed) instead of three. That means that photon wave functions must obey one auxiliary condition which eliminates the unwanted state of polarization. This special property will enable us after all to use spin-1 matrices to take the square root of the Laplacian. To see how this comes about let us multiply both sides of Eq.(8) by two nabla operators  $\nabla_i$  and  $\nabla_j$  and perform a summation over the indices  $i$  and  $j$

$$[(\mathbf{S}\cdot\nabla)(\mathbf{S}\cdot\nabla)]_{ab} = \Delta\delta_{ab} - \nabla_a\nabla_b. \quad (9)$$

This relation is nothing else but a matrix version of a well known identity from vector analysis

$$-\nabla \times (\nabla \times \mathbf{F}) = \Delta\mathbf{F} - \nabla(\nabla\cdot\mathbf{F}). \quad (10)$$

Thus, for all vectors  $\mathbf{F}$  whose divergence vanishes,

$$\nabla\cdot\mathbf{F} = 0, \quad (11)$$

we can imitate Dirac and take square root of the Laplacian in the form

$$\sqrt{\Delta}\mathbf{F} = \pm(\mathbf{S}\cdot\nabla)\mathbf{F}. \quad (12)$$

This leads to a Dirac-like equation for the photon wave function  $\mathbf{F}$  that looks like the Weyl equation (5) for neutrinos except that the Pauli matrices that are appropriate for spin-1/2 are replaced by spin-1 matrices,

$$i\hbar\partial_t\mathbf{F} = c\left(\mathbf{S}\cdot\frac{\hbar}{i}\nabla\right)\mathbf{F}. \quad (13)$$

Thus, the Hamiltonian operator for the photon emerges in this approach as the scalar product of the momentum vector and the spin vector multiplied by speed of light. The auxiliary condition (11) can also be expressed in terms of spin and momentum operators in the form

$$(\mathbf{S}\cdot\nabla)S_j\mathbf{F} = \nabla_j\mathbf{F}. \quad (14)$$

For definiteness, I have chosen the sign in Eq.(13) to correspond to positive helicity — as in the equation for the antineutrino. Since the matrices  $S_i$  are purely imaginary, reversal of the sign is equivalent to complex conjugation of the wave function. I shall discuss this problem in detail in the next section.

In order to find the relation between the photon wave equation and the Maxwell equations I shall first get rid of the Planck constant by dividing both sides of Eq.(13) by  $\hbar$ . Next, I shall rewrite this equation in vector notation using the identity

$$(\mathbf{S}\cdot\mathbf{a})\mathbf{F} = i\mathbf{a} \times \mathbf{F}, \quad (15)$$

to exhibit a familiar pattern involving the curl operation

$$i\partial_t\mathbf{F}(\mathbf{r}, t) = c\nabla \times \mathbf{F}(\mathbf{r}, t). \quad (16)$$

After separating the real and the imaginary part,

$$\partial_t\Re\mathbf{F} = c\nabla \times \Im\mathbf{F}, \quad (17a)$$

$$\partial_t\Im\mathbf{F} = -c\nabla \times \Re\mathbf{F}, \quad (17b)$$

the above equations become the Maxwell equations, provided we identify the real and the imaginary part of  $\mathbf{F}$  with the vectors of electric displacement  $\mathbf{D}$  and magnetic induction  $\mathbf{B}$ ,

$$\mathbf{F}(\mathbf{r}, t) = \frac{1}{\sqrt{2}} \left( \frac{\mathbf{D}(\mathbf{r}, t)}{\sqrt{\epsilon_0}} + i\frac{\mathbf{B}(\mathbf{r}, t)}{\sqrt{\mu_0}} \right). \quad (18)$$

The square roots of  $\epsilon$  and  $\mu$  are needed to match the dimensions of the two terms and an additional factor of  $\sqrt{2}$  is introduced to make the modulus of  $\mathbf{F}$  equal simply to the energy density (cf. formulas (28)-(31)). My sign convention for  $\mathbf{F}$  is motivated by the existing connection with the spinor calculus (Cf. Sec. XI). The choice of the opposite sign, as adopted, for example, by Majorana results in an interchange of  $\mathbf{F}$  with its complex conjugate leading to a reversal of helicity. However, as I show in the next section, if the polarization states of the photon are to be properly accounted for, both these choices must be included on equal footing.

I believe that the main lesson that we can learn from this derivation is that, despite of their purely classical origin, Maxwell equations bear a very strong resemblance to quantum wave equations that describe time evolution of a wave function in relativistic theory. The same line of argument that led Dirac to the discovery of his relativistic equation for the electron can be used to derive Maxwell equations. The analogy with the Weyl equation for the neutrino is particularly striking; both equations share the same property of not containing Planck's constant. This analogy can be used to introduce the complete photon wave function.

### III. THE WAVE FUNCTION OF THE PHOTON

Stationary solutions of the photon wave equation (13) with positive frequency satisfy the following eigenvalue condition

$$c\left(\mathbf{S}\cdot\frac{\hbar}{i}\nabla\right)\mathbf{F}=\hbar\omega\mathbf{F}. \quad (19)$$

This equation says that the projection of the spin on the direction of momentum (helicity) is positive. Thus, the choice of sign in the wave equation is equivalent to choosing one helicity (right-handed or left-handed) over the other. For neutrinos such a choice is based on the experimental observation that all neutrinos are left-handed, while all antineutrinos are right-handed. The two choices are kept separate by the law of lepton conservation that forbids the superposition of neutrino and antineutrino wave functions. For photons such a distinction does not exist. Photons do not have antiparticles, there is no photon conservation law. As a consequence, the wave functions of left-handed and right-handed photons can be freely superposed. Moreover, the electromagnetic interactions are invariant under the parity transformation that interchanges the states with opposite helicities. Therefore, a change of polarization should be associated with the parity transformation and not with a particle-antiparticle transition. These facts require an important modification of the proposed approach. Of course, if we are only interested in translating the Maxwell equations into a complex form, we can choose either  $\mathbf{F}$  or  $\mathbf{F}^*$  as a primary object. In both cases we obtain a one-to-one correspondence between the real field vectors  $\mathbf{D}$  and  $\mathbf{B}$  and their complex combination. However, in order to treat such an object as a bona fide photon wave function we must be able to superpose different polarization states without changing the sign of the energy (frequency). This can only be done when both polarizations are described by different components of *the same wave function*. Since photons may have both signs of helicity, corresponding to left-handed and right-handed polarizations, we need wave functions for the photons of both helicities. In order to account for all photon states, two independent parts of the wave function, say  $\mathbf{F}_+$  and  $\mathbf{F}_-$ , have to be introduced. For the same reason we introduce two-component wave functions to describe two polarization states of spin 1/2 massive particles. The two parts of the photon wave function in empty space (I shall show in Sec. XII that the presence of a medium changes the situation) satisfy two separate evolution equations with opposite signs of the Hamiltonian,

$$i\hbar\partial_t\mathbf{F}_+=c\left(\mathbf{S}\cdot\frac{\hbar}{i}\nabla\right)\mathbf{F}_+, \quad (20a)$$

$$i\hbar\partial_t\mathbf{F}_-=-c\left(\mathbf{S}\cdot\frac{\hbar}{i}\nabla\right)\mathbf{F}_-. \quad (20b)$$

By introducing a six-component wave function  $\mathcal{F}$  made of  $\mathbf{F}_\pm$  as its upper and lower components,

$$\mathcal{F}=\begin{pmatrix} \mathbf{F}_+ \\ \mathbf{F}_- \end{pmatrix}, \quad (21)$$

the two evolution equations can be cast into a compact form

$$i\hbar\partial_t\mathcal{F}=c\rho_3\left(\mathbf{S}\cdot\frac{\hbar}{i}\nabla\right)\mathcal{F}, \quad (22)$$

where  $\rho_3$  is a member of a set of Pauli-like matrices  $\rho_i$ . The matrix  $\rho_3$  produces plus/minus sign when acting on upper/lower components while  $\rho_1$  interchanges these components.

With the understanding that the upper and the lower components of  $\mathcal{F}$  form vectors, one may also use the vector notation for the evolution equation (22) and the subsidiary condition

$$i\hbar\partial_t\mathcal{F}=c\rho_3\hbar\nabla\times\mathcal{F}, \quad (23)$$

$$\nabla\cdot\mathcal{F}=0. \quad (24)$$

The doubling of the components of the wave function requires an auxiliary condition that would restore the original number of degrees of freedom. This is achieved by demanding that complex conjugation of the wave function has the same effect as an interchange of upper and lower components,

$$\mathcal{F}=\rho_1\mathcal{F}^*. \quad (25)$$

This condition may also be interpreted as a mathematical form of the statement that photons do not have antiparticles so that the negative frequency part of the wave function does not carry any additional information. Therefore, as

an independent quantity we may consider only the positive energy part  $\mathcal{F}^{(+)}$  of the wave field  $\mathcal{F}$  (called an analytic signal in classical theory). This part represents the true photon wave function and that is the part to be identified with the matrix elements of the field operators between the vacuum and one-photon states that appear in quantum electrodynamics. In what follows I shall often use the symbol  $\mathcal{F}$  to denote the positive energy part of the photon wave field. Owing to the condition (25), the negative frequency part is obtained just 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. (25) is the condition of invariance under charge conjugation.

The analogy with the relativistic electron theory that I have been pursuing in the previous section becomes the closest when the photon wave equation is written in the form (22) for then it directly compares with the Dirac equation written in the chiral representation of the Dirac matrices. In this representation the bispinor is made of two relativistic spinors (cf. for example [17,16]),

$$\psi = \begin{pmatrix} \phi_A \\ \chi^{A'} \end{pmatrix}. \quad (26)$$

and the Dirac equation for the bispinor  $\psi$  has the form

$$i\hbar\partial_t\psi = c\rho_3\left(\boldsymbol{\sigma}\cdot\frac{\hbar}{i}\nabla\right)\psi + mc^2\rho_1\psi. \quad (27)$$

An auxiliary condition of the type (25) is not imposed on the electron wave function, because the electrons do have distinct antiparticles. Such a condition would be required for neutral fermions that do not have distinct antiparticles (the so called massive Majorana neutrinos). For the electrons, even in free space, the mass term induces mixing of two states of polarization while for the photons, as I shall show in Sec. XII, mixing is induced only by an external influence, by the medium.

#### IV. CONSERVATION LAWS

The fundamental quantities characterizing the photon: the energy density  $\mathcal{H}$ , the momentum density  $\mathcal{P}$ , the energy flux  $\mathcal{S}$  and the Maxwell stress tensor  $\mathcal{T}_{ij}$  can be expressed as the following combinations of the components of the photon wave function,

$$\mathcal{H} = \mathcal{F}^\dagger \cdot \mathcal{F}, \quad (28)$$

$$\mathcal{P} = \mathcal{F}^\dagger \rho_3 \mathbf{S} \mathcal{F} / c, \quad (29)$$

$$\mathcal{S} = c \mathcal{F}^\dagger \rho_3 \mathbf{S} \mathcal{F}, \quad (30)$$

$$\mathcal{T}_{ij} = \mathcal{F}^\dagger (S_i S_j + S_j S_i) \mathcal{F} - \delta_{ij} \mathcal{F}^\dagger \cdot \mathcal{F}. \quad (31)$$

All these expressions are bilinear in the wave function  $\mathbf{F}$ , exactly as in wave mechanics of massive particles. The evolution equation for  $\mathcal{F}$  leads to the following continuity equations involving the field quantities (28)-(31) that are just well known local versions of energy and momentum conservation

$$\partial_t \mathcal{H} = -\nabla \cdot \mathcal{S}, \quad (32)$$

$$\partial_t \mathcal{P}_i = -\nabla_k \mathcal{T}_{ik}. \quad (33)$$

It follows from the continuity equations (32) and (33) that the following integrals representing the total energy  $E$ , momentum  $\mathbf{P}$ , and angular momentum  $\mathbf{M}$  of the field are time independent

$$E = \int d^3r \mathcal{F}^\dagger \cdot \mathcal{F}, \quad (34)$$

$$\mathbf{P} = \int d^3r \mathcal{F}^\dagger \rho_3 \mathbf{S} \mathcal{F} / c, \quad (35)$$

$$\mathbf{M} = \int d^3r \mathbf{r} \times (\mathcal{F}^\dagger \rho_3 \mathbf{S} \mathcal{F}) / c. \quad (36)$$

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

In addition to these well known conserved quantities there is an additional conserved quantity that can also be expressed as a bilinear form built from the photon wave function. In contradistinction to expressions (34)-(36), this one is not obtained by integrating a local density, but has a form of a double integral,

$$\Lambda = \frac{-i}{\hbar c} \int d^3 r_1 \int d^3 r_2 \frac{\mathcal{F}^\dagger(\mathbf{r}_1) \rho_3(\mathbf{S} \cdot \nabla) \mathcal{F}(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}. \quad (37)$$

This constant of motion can be shown to represent the difference between the average number of right-handed and left-handed photons. I shall say more about conserved quantities in Sec. X.

## V. PROBABILISTIC INTERPRETATION OF THE WAVE FUNCTION

Probabilistic interpretation of the photon wave function is possible without destroying locality if one takes the energy density as a relative measure of the probability density to find the photon. Since photons are not charged, this seems to be a sensible choice. After all, one detects photons by absorbing their energy: the photon is where its energy is localized. The normalized value  $\rho(\mathbf{r}, t)$  of the probability density is

$$\rho(\mathbf{r}, t) = \frac{\mathcal{H}(\mathbf{r}, t)}{\int d^3 r \mathcal{H}}, \quad (38)$$

$$\int d^3 r \rho(\mathbf{r}, t) = 1, \quad (39)$$

and energy conservation automatically gives the conservation of total probability. The probability current is given by the Poynting vector (normalized again by the total energy)

$$\mathbf{j}(\mathbf{r}, t) = \frac{\mathcal{S}(\mathbf{r}, t)}{\int d^3 r \mathcal{H}}. \quad (40)$$

satisfying together with the probability density the continuity equation

$$\partial_t \rho + \nabla \cdot \mathbf{j} = 0. \quad (41)$$

The nonlocal photon wave functions proposed by Landau-Peierls and by Cook do not give a local conservation of probability.

For stationary states of the photon the proper normalization condition for the photon wave function will be

$$\int d^3 r \mathcal{H} = \hbar \omega, \quad (42)$$

expressing the fact that the wave function describes only a single photon. This condition must be handled with some care, since stationary states in infinite space are described by non-normalizable wave functions. Therefore, Eq.(42) strictly speaking can hold only in a finite volume. In the formulas representing expectation values (cf. for example Eqs.(48) and (49)), we may, however, in most cases take the limit of infinite volume.

## VI. EIGENVALUE PROBLEMS FOR THE PHOTON WAVE FUNCTION

Wave mechanics of photons is amenable to the same analysis as is customarily applied in wave mechanics of massive particles. In particular, we may study eigenvalues and eigenfunctions of various interesting observables. The most important examples of such observables are momentum, angular momentum and energy.

In the discussion of the Hamiltonian, we have already encountered the quantum mechanical momentum operator  $(\hbar/i)\nabla$ . The eigenvalue problem for the photon momentum operator has the standard form

$$\hat{p}_i \mathcal{F} = \hbar k_i \mathcal{F}, \quad (43)$$

where  $k_i$  are components of the wave vector.

The eigenvalue problem for the photon angular momentum also has the same form as in quantum mechanics of massive particles. The total angular momentum operator consists of two parts: the orbital angular momentum and

the spin angular momentum. It is worth noting that it is the total angular momentum  $\hat{\mathbf{J}}$  defined as the sum of the orbital and the spin part,

$$\hat{\mathbf{J}} = \mathbf{r} \times \frac{\hbar}{i} \nabla + \hbar \mathbf{S}, \quad (44)$$

that commutes with the Hamiltonian and, therefore, is a constant of motion. The eigenvalue problem for angular momentum contains, as usual, the eigenvalue problem for the  $z$ -component of the total angular momentum

$$\hat{J}_z \mathcal{F} = \hbar M \mathcal{F}, \quad (45)$$

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

$$\hat{\mathbf{J}}^2 \mathcal{F} = \hbar^2 J(J+1) \mathcal{F}. \quad (46)$$

The solutions of Eqs. (45) and (46) are well known vector spherical harmonics (cf. for example [18]).

An interesting problem is that of a relationship between the wave mechanical expectation values  $\langle \mathbf{P} \rangle$  and  $\langle \mathbf{M} \rangle$  of the momentum and angular momentum operator and the total momentum and angular momentum of the electromagnetic field defined as the integrals (35) and (36). At first the corresponding expressions look completely different, but for stationary photon states they turn out to produce identical results. In order to see this one may use the following relation obtained from Eqs.(16) and (19) that are valid for wave functions of stationary states

$$\mathcal{F} = \rho_3 \frac{c}{\omega} \nabla \times \mathcal{F}. \quad (47)$$

Upon substituting the r.h.s. of this equality into the formula (35) for the field momentum, after an integration by parts and with the use of the transversality condition (11), I transform (35) into the quantum mechanical expression for an average value

$$\langle \mathbf{P} \rangle = \frac{1}{\hbar \omega} \int d^3 r \mathcal{F}^\dagger \frac{\hbar}{i} \nabla \mathcal{F} = \frac{\int d^3 r \mathcal{F}^\dagger \frac{\hbar}{i} \nabla \mathcal{F}}{\int d^3 r \mathcal{F}^\dagger \cdot \mathcal{F}}, \quad (48)$$

where in the last term I have used the normalization condition (42) for stationary states. In a similar manner, the classical formula (36) is transformed into an expression for the average value of the angular momentum

$$\langle \mathbf{M} \rangle = \frac{\int d^3 r \mathcal{F}^\dagger [\mathbf{r} \times \frac{\hbar}{i} \nabla + \hbar \mathbf{S}] \mathcal{F}}{\int d^3 r \mathcal{F}^\dagger \cdot \mathcal{F}}. \quad (49)$$

These results show that in order to convert for stationary states the standard classical expressions for the total momentum and angular momentum of the electromagnetic field into the expressions used in wave mechanics we only need to use the appropriate normalization of the wave field. For nonstationary states the two methods for calculating momentum and angular momentum can not be compared because we lack the normalization prescription for the classical fields: we can not tell how many photons are in a state. It is worth noting that the final formulas (48) and (49) have exactly the same form as in Schrödinger's wave mechanics; the matrix  $\rho_3$  has disappeared.

The eigenvalue problem for the photon energy operator can be set up along similar lines. We can search for stationary solutions of photon wave equations closely following the path travelled in nonrelativistic wave mechanics of massive particles. This procedure usually involves selecting a set of operators commuting with the Hamiltonian and then solving the appropriate set of eigenvalue equations. I shall illustrate this approach in the next section with an example of photon propagation along an infinite cylindrical optical fiber.

## VII. PHOTON PROPAGATION ALONG AN OPTICAL FIBER AS A QUANTUM MECHANICAL BOUND STATE PROBLEM

Let us consider an infinite optical fiber of diameter  $a$  characterized by a dielectric permittivity  $\epsilon$ . 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

$$-i\partial_z \mathcal{F} = k_z \mathcal{F}, \quad (50a)$$

$$-i\partial_\varphi \mathcal{F} = M \mathcal{F}. \quad (50b)$$

Assuming that the photon wave function belongs to the eigenvalues  $\hbar k_z$  and  $\hbar M$  of these operators we can separate out the dependence on  $z$  and  $\varphi$

$$\mathcal{F} = \exp(ik_z z) \exp(iM\varphi) f(\rho). \quad (51)$$

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 [19]). The result can be written in the form

$$-\frac{M}{\rho} f_z + k_z f_\varphi = \frac{\omega}{c} (i f_\rho), \quad (52a)$$

$$-\partial_\rho f_z + k_z (i f_\rho) = \frac{\omega}{c} f_\varphi, \quad (52b)$$

$$\frac{1}{\rho} \partial_\rho \rho f_\varphi - \frac{M}{\rho} (i f_\rho) = \frac{\omega}{c} f_z. \quad (52c)$$

Every solution of these equations automatically satisfies the divergence condition (11). Eqs.(52) lead to a Bessel equation for  $f_z$

$$\left[ \partial_\rho^2 + \frac{1}{\rho} \partial_\rho - \frac{m^2}{\rho^2} + k_\perp^2 \right] f_z = 0, \quad (53)$$

where  $k_\perp^2 = \omega^2/c^2 - k_z^2$ . Photon wave functions obey Eq.(53) inside the fiber with one value of  $k_\perp$  and in the surrounding free space with different values of  $k_\perp$ . 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, we 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  $k_\perp$  we 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.

## VIII. PHOTON WAVE FUNCTION IN NON-CARTESIAN COORDINATE SYSTEMS AND IN CURVED SPACE

It has been observed some time ago [20,21] that the equations determining the propagation of the electromagnetic field in arbitrary coordinate systems including also the case of curved spacetime may be written in the form of Maxwell equations, with all the information about the spacetime geometry contained in the constitutive relations connecting the  $\mathbf{E}$ ,  $\mathbf{B}$  and  $\mathbf{D}$ ,  $\mathbf{H}$  fields. The description of the electromagnetic field in terms of the wave function adds to this old observation one interesting new element. In contradistinction to the case of an inhomogeneous medium (cf. Sec. XII, in the gravitational field the two photon helicities do not mix. This follows from the fact that the constitutive relations induced by an arbitrary metric tensor  $g_{\mu\nu}$  can be written as a single equation connecting two complex vectors: the vector  $\mathbf{F}$  defined by the formula (18) and a new vector  $\mathbf{G}$  defined as

$$\mathbf{G}(\mathbf{r}, t) = \frac{1}{\sqrt{2}} \left( \frac{\mathbf{E}(\mathbf{r}, t)}{\sqrt{\mu_0}} + i \frac{\mathbf{H}(\mathbf{r}, t)}{\sqrt{\epsilon_0}} \right). \quad (54)$$

The constitutive relations for two complex vectors  $\mathbf{F}$  and  $\mathbf{G}$  have in curved spacetime the following form

$$F^i = -\frac{1}{g_{00}} (\sqrt{-g} g^{ij} + i g_{0k} \epsilon^{ikj}) G_j, \quad (55)$$

$$G_i = -\frac{1}{g^{00}} (g_{ij} / \sqrt{-g} - i g^{0k} \epsilon_{ikj}) F^j. \quad (56)$$

The complex form of the Maxwell equations expressed in terms of vectors  $\mathbf{G}$  and  $\mathbf{F}$  in curved spacetime the same as in flat space

$$i\partial_t \mathbf{F} = \nabla \times \mathbf{G}, \quad (57)$$

$$\nabla \cdot \mathbf{F} = 0, \quad (58)$$

In these equations the differentiations with respect to time and the nabla operation involve only *ordinary* (not covariant) derivatives as in flat space. The only difference is, of course, in the form of the constituent relations (56) which contain all the information about the gravitational field. Since the relation between  $\mathbf{G}$  and  $\mathbf{F}$  is linear, we may write again two separate wave equations for the two helicity states as before. Combining these two equations, we obtain the following wave equation and the subsidiary condition for the full photon wave function  $\mathcal{F}$

$$i\partial_t \mathcal{F} = \rho_3 \nabla \times \mathcal{G}, \quad (59)$$

$$\nabla \cdot \mathcal{F} = 0, \quad (60)$$

where

$$\mathcal{G}_i = -\frac{1}{g_{00}} (g_{ij}/\sqrt{-g} - i\rho_3 g^{0k} \varepsilon_{ikj}) \mathcal{F}^j. \quad (61)$$

These equations contain only the matrix  $\rho_3$  that does not mix the helicity states.

## IX. PHASE-SPACE DISTRIBUTION FUNCTION FOR THE PHOTON AND ITS EVOLUTION EQUATIONS

Having introduced the photon wave function, I can follow the procedure known from nonrelativistic wave mechanics and define a distribution function in the phase-space: an analog of the Wigner function [22] in wave mechanics. This is achieved by introducing an appropriate Fourier transform of the product of the wave function and its complex conjugate. Fourier transforms of the electromagnetic fields similar to the Wigner function in quantum mechanics have been introduced before in optics, first by Walther [23] in the two-dimensional context of radiative transfer theory and then by Wolf [24] and by Sudarshan [25–27] in the three-dimensional case. In all these works phase-space distribution functions were defined for stationary states of the electromagnetic field only and they were treated as functions of frequency. By contrast, in my more general approach the distribution functions will be generically time dependent since they will be constructed from arbitrary wave functions. This will allow me to give a meaning to the notion of an instantaneous state of the electromagnetic field and to study its the time evolution.

The only formal difference between the standard definition [22] of the Wigner function in quantum mechanics and the present case is the presence of vector indices. Thus, the photon distribution function (PDF) in phase space will not be a single scalar function but rather a  $6 \times 6$  matrix defined as follows

$$W_{ab}(\mathbf{r}, \mathbf{k}, t) = \int d^3s e^{-i\mathbf{k}\mathbf{s}} \mathcal{F}_a(\mathbf{r} + \mathbf{s}/2, t) \mathcal{F}_b^*(\mathbf{r} - \mathbf{s}/2, t). \quad (62)$$

Similar multi-component distribution functions arise also for a Dirac particle and I shall borrow from a recent publication [28] some of the techniques to deal with such functions.

It follows directly from the definition (62) that the components of  $W_{ab}$  form a Hermitian matrix. Every such  $6 \times 6$  matrix can be written in the following block form

$$W_{ab} = \begin{pmatrix} W_{ij}^0 + W_{ij}^3 & W_{ij}^1 - iW_{ij}^2 \\ W_{ij}^1 + iW_{ij}^2 & W_{ij}^0 - W_{ij}^3 \end{pmatrix}, \quad (63)$$

where all  $3 \times 3$  matrices  $W_{ij}^\alpha$  are hermitian. This decomposition can also be expressed in terms of the  $\rho$  matrices

$$W_{ab} = \rho_0 W_{ij}^0 + \rho_1 W_{ij}^1 + \rho_2 W_{ij}^2 + \rho_3 W_{ij}^3. \quad (64)$$

The vector indices  $i$  and  $j$  refer to components within the upper and lower parts of the wave function and the matrices  $\rho$  act on these parts as a whole. The most general PDF, as seen from this analysis, is quite complicated. In general, when the medium induces transitions between the two polarization states, all components of this function are needed. However, when the photons propagate in free space only a small subset of these components is sufficient to account for

all important dynamic properties of photon beams. As I show below the full phase-space dynamics can be described in a self-contained manner by just one scalar function and one vector function. To this end let me introduce the following reduced PDF

$$W_{ij}(\mathbf{r}, \mathbf{k}, t) = \int d^3s e^{-i\mathbf{k}\mathbf{s}} F_i(\mathbf{r} + \mathbf{s}/2, t) F_j^*(\mathbf{r} - \mathbf{s}/2, t), \quad (65)$$

where  $F_i$  are the components of the original complex vector (18). This vector contains full information about the classical electromagnetic field and the PDF (65) will suffice to account for the dynamics. The matrix  $W_{ij}^\alpha$  can be decomposed into a real symmetric tensor and a real vector according to the formula

$$W_{ij}^\alpha = w_{ij}^\alpha + \frac{c}{2i} \epsilon_{ijk} u_k^\alpha. \quad (66)$$

The tensor corresponds to the symmetric part and the vector to the antisymmetric part of  $W_{ij}^\alpha$ . I have separated out a factor of  $c$  in the second term since the vector  $\mathbf{u}$  will be shown to be related to the momentum density, while the trace of  $w_{ij}$  is related to the energy density.

The equations satisfied by the components of the PDF in free space can be obtained from Maxwell equations (16) and (11) for the vector  $\mathbf{F}$ ,

$$\partial_t W_{ij} = c \left( \mathbf{k} + \frac{i}{2} \nabla \right)_m \epsilon_{mil} W_{lj} - c \left( \mathbf{k} - \frac{i}{2} \nabla \right)_m W_{il} \epsilon_{mlj}, \quad (67)$$

$$\left( \mathbf{k} + \frac{i}{2} \nabla \right)_i W_{ij} = 0 = \left( \mathbf{k} - \frac{i}{2} \nabla \right)_j W_{ij}. \quad (68)$$

This leads to the following set of coupled evolution equations for the real components  $w_{ij}$  and  $u_i$

$$\partial_t w_{ij} = -c \epsilon_{ilk} k_l w_{kj} - c \epsilon_{jlk} k_l w_{ki} - \frac{1}{2} c^2 (\nabla_i u_j + \nabla_j u_i - \delta_{ij} \nabla_k u_k), \quad (69a)$$

$$\partial_t u_i = -c \epsilon_{ijk} k_j u_k - \frac{1}{2} (\nabla_j w_{ij} - \nabla_i w_{jj}), \quad (69b)$$

and to the subsidiary conditions

$$c \epsilon_{ijk} k_j u_k = \nabla_j w_{ij}, \quad (70a)$$

$$c \epsilon_{ijk} \nabla_j u_k = 4k_j w_{ij}, \quad (70b)$$

The  $k$ -dependent terms in the evolution equations describe a uniform rotation of the vector and of the tensor around the wave vector  $k$  so that these terms can be eliminated by “going to a rotating coordinate system”.

With the help of the subsidiary conditions (70) we can eliminate the remaining components and obtain from the evolution equations (69) equations for  $w = \sum w_{ii}$  and  $\mathbf{u}$

$$\partial_t w = -c^2 \nabla_i u, \quad (71a)$$

$$\partial_t u_i = -2c \epsilon_{ijk} k_j u_k - \nabla_i w, \quad (71b)$$

These evolution equations do form a simple, self-contained set. However, as it is always the case with the phase-space distribution functions in wave mechanics, not all solutions of these equations are admissible. Only those distribution functions are allowed that can be represented in the form (65) at the initial time (with the vector  $\mathbf{F}$  satisfying the subsidiary condition (70a)).

## X. INVARIANCE GROUP AND ITS GENERATORS

The Maxwell equations in free space are invariant under the full 15-parameter group of conformal transformations that includes the Poincaré group as its 10 parameter subgroup. This invariance is shared by the equations (69) and (70) for the PDF which were obtained from the Maxwell equations and it leads, by Noether’s theorem, to the existence of 15 conserved quantities — the generators of the invariance group (see, for example, [16]). All these constants of motion can be expressed in the uniform manner as phase-space integrals involving only the four components  $w(\mathbf{r}, \mathbf{k}, t)$  and  $\mathbf{u}(\mathbf{r}, \mathbf{k}, t)$  of the PDF  $W_{ij}$ :

$$E = \int d\Gamma w, \quad (72a)$$

$$\mathbf{P} = \int d\Gamma \mathbf{u}, \quad (72b)$$

$$\mathbf{M} = \int d\Gamma \mathbf{r} \times \mathbf{u}, \quad (72c)$$

$$\mathbf{N} = \int d\Gamma \mathbf{r} w - c^2 t \mathbf{P}, \quad (72d)$$

$$D = \int d\Gamma \mathbf{r} \cdot \mathbf{u} - tE, \quad (72e)$$

$$K_0 = \int d\Gamma (\mathbf{r}^2 + c^2 t^2) w - 2c^2 t \int d\Gamma \mathbf{r} \cdot \mathbf{u}, \quad (72f)$$

$$\mathbf{K} = c \int d\Gamma \mathbf{r}^2 \mathbf{u} - c^3 t^2 \mathbf{P} - 2c \int d\Gamma \mathbf{r}(\mathbf{r} \cdot \mathbf{u}) + 2ct \int d\Gamma \mathbf{r} w, \quad (72g)$$

where  $d\Gamma = d^3r d^3k/(2\pi)^3$ . The constants of motion  $E, \mathbf{P}, \mathbf{M}$ , and  $\mathbf{N}$  have a clear physical interpretation of the energy, momentum, angular momentum, and the moment of energy (the center of mass vector multiplied by the total energy  $E$ ). However, there is no direct interpretation of  $D, K_0$ , and  $\mathbf{K}$  except as certain moments of the momentum distribution and the energy distribution, or as generators of the conformal transformations. In order to prove the time independence of all quantities (72) we can use the evolution equations and the subsidiary conditions satisfied by the components of the distribution function.

## XI. PHOTON WAVE FUNCTION AND THE SPINOR CALCULUS

For those who are accustomed to using the spinorial description of relativistic particles, the photon wave function is a familiar object. It is the self-dual (or anti self-dual) part of the electromagnetic field tensor (cf. ref. [29]). To be more precise, the components of the vector  $\mathbf{F}$  are related to the components of a second rank symmetric spinor  $\phi_{AB}$

$$\phi_{00} = -F_x + iF_y, \quad (73a)$$

$$\phi_{01} = F_z, \quad (73b)$$

$$\phi_{11} = F_x + iF_y, \quad (73c)$$

and the components of the complex conjugate vector  $\mathbf{F}^*$  are related to the components of a second rank symmetric primed spinor  $\phi^{A'B'}$ ,

$$\phi^{0'0'} = -F_x^* - iF_y^*, \quad (74a)$$

$$\phi^{0'1'} = -F_z^*, \quad (74b)$$

$$\phi^{1'1'} = F_x^* - iF_y^*. \quad (74c)$$

In the spinorial formalism, our observation that in curved space both helicities propagate without mixing is a simple consequence of the fact that both spinors  $\phi_{AB}$  and  $\phi^{A'B'}$  satisfy separate equations of the form [29]

$$\sigma^{\mu C' A} \nabla_\mu \phi_{AB} = 0, \quad (75a)$$

$$\sigma^\mu_{C A'} \nabla_\mu \phi^{A' B'} = 0, \quad (75b)$$

where the soldering form  $\sigma^{\mu A' B}$  is related to the metric tensor through the formulas

$$\sigma^{\mu A' B} \sigma^\nu_{A' C} + \sigma^{\nu A' B} \sigma^\mu_{A' C} = 2g^{\mu\nu} \delta_C^B. \quad (76)$$

The covariant spinorial derivative  $\nabla_\mu$  acts on spinors with the help of the spinorial affine connections  $\Gamma_{\mu A}^B$  and  $\Gamma_{\mu}^{A' B'}$  (cf., for example, [31])

$$(\nabla_\mu \psi)_A = \partial_\mu \psi_A - \Gamma_{\mu A}^B \psi_B, \quad (77a)$$

$$(\nabla_\mu \phi)^{A'} = \partial_\mu \phi^{A'} + \Gamma_{\mu}^{A' B'} \phi^{B'}, \quad (77b)$$

$$(\nabla_\mu \phi)_{AB} = \partial_\mu \phi_{AB} - \Gamma_{\mu A}^C \phi_{CB} - \Gamma_{\mu B}^C \phi_{AC}, \quad (77c)$$

that can be expressed in terms of the soldering form

$$\Gamma_{\mu A}{}^B = \frac{1}{4}(\partial_\lambda g_{\rho\mu} + \frac{1}{2}\sigma_{\lambda C'D}\partial_\mu\sigma_\rho{}^{C'D})S^{\lambda\rho}{}_A{}^B, \quad (78)$$

where the generators of Lorentz transformations  $S^{\mu\nu}{}_A{}^B$  are defined as

$$S^{\mu\nu}{}_A{}^B = \frac{1}{2}(\sigma^\mu{}_{C'A}\sigma^\nu{}^{C'B} - \sigma^\nu{}_{C'A}\sigma^\mu{}^{C'B}). \quad (79)$$

The existence of a direct relationship between the three components of the wave function  $\mathbf{F}$  and the three independent components of the symmetric spinor  $\phi_{AB}$  signifies that the photon wave function in free space (also in a curved space), unlike the wave function of Landau and Peierls, is a truly relativistic geometric object. The components of  $\mathbf{F}$  at a point  $x$  in one coordinate system are linear combinations of the components in another coordinate system taken at the corresponding point  $'x$  for under a Lorentz transformation the second rank spinor changes according to the formula

$$'\phi_{AB}('x) = S_A{}^C S_B{}^D \phi_{CD}(x). \quad (80)$$

where  $S_A{}^C$  is a  $2 \times 2$  matrix of the fundamental (spinor) representation of the Lorentz group.

Thus, from the point of view of the representation theory of the Lorentz and Poincaré groups [30], our mathematical scheme looks painfully simple: the photon wave functions for a given helicity is just a three component field that transforms under irreducible representations (1, 0) or (0, 1) of the proper Lorentz group (without reflections). In order to accommodate reflections we must combine both representations and introduce the object  $\Phi$ ,

$$\Phi = \begin{pmatrix} \phi_{AB} \\ \phi^{A'B'} \end{pmatrix}, \quad (81)$$

that is a counterpart of the Dirac bispinor (26). This is exactly our complete photon wave function  $\mathcal{F}$  describing both polarizations (up to a transformation from the vectorial to the spinorial components given by Eqs.(73)and (74)). I have used the same symbol to denote upper and lower components since, unlike the case of the Dirac bispinor (26), they are connected by the complex conjugation (25).

The Maxwell equations expressed in spinor notation and the Weyl equation provide just the simplest examples taken out of a hierarchy of wave equations for massless fields described by symmetric spinors  $\phi_{B_1 B_2 \dots B_n}$  or  $\phi^{B'_1 B'_2 \dots B'_n}$ . All these equations have the form [29]

$$\sigma^{\mu C'A} \nabla_\mu \phi_{AB_1 B_2 \dots B_{n-1}} = 0, \quad (82a)$$

$$\sigma^\mu{}_{C'A'} \nabla_\mu \phi^{A'B'_1 B'_2 \dots B'_{n-1}} = 0. \quad (82b)$$

## XII. WAVE MECHANICS OF PHOTONS IN AN INHOMOGENEOUS MEDIUM

Having made the identification of the real and the imaginary parts of the photon wave function with the electromagnetic field vectors we may turn the argument around and use the Maxwell equations to obtain the wave equation for the photon wave function in a general linear 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 vector  $\mathbf{F}$  as in free space, except that  $\epsilon_0$  and  $\mu_0$  in (18) are replaced by their values in a medium. I have taken advantage of this fact in the description of the photon propagation in a fiber. However, when the values of  $\epsilon$  and  $\mu$  are space- and/or time-dependent the vector  $\mathbf{F}$  ceases to satisfy an equation by itself but the evolution equations involve also its complex conjugate. This, at first sight, gives nonlinear wave equations and threatens to ruin our whole scheme, but the doubling of components of the photon wave function introduced in Sec. III solves the problem. The “big” wave function  $\mathcal{F}$  introduced to account for the two states of polarization satisfies a linear wave equation in an arbitrary linear medium. In free space, in a homogeneous medium, or in 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  $\mathcal{F}$  are treated as three-dimensional vectors to make use of the nabla operation)

$$i(\partial_t + \partial_t L)\mathcal{F} = c\rho_3(\nabla + \nabla L) \times \mathcal{F}, \quad (83)$$

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

$$c = 1/\sqrt{\epsilon\mu}, \quad (84)$$

and the matrix  $L$  has the form,

$$L = \frac{1}{2} \left( \ln \sqrt{\epsilon\mu} + \rho_1 \ln \sqrt{\frac{\epsilon}{\mu}} \right). \quad (85)$$

One can check that the operator on the right hand side of Eq.(83) is Hermitian so that it can be viewed as a quantum mechanical Hamiltonian in the presence of an inhomogeneous medium. The transversality condition (11) in an inhomogeneous medium is also modified and it takes on the form,

$$(\nabla + \nabla L) \cdot \mathcal{F} = 0. \quad (86)$$

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 (83) and the transversality condition (86) are invariant, as was the case in empty space, under the charge symmetry operation

$$\mathcal{F} \rightarrow \rho_1 \mathcal{F}^*. \quad (87)$$

Therefore, we can impose now the same restriction (25) 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 is describing in actual fact a propagation of some collective excitations of the whole system and not just of pure photons.

### XIII. CONCLUSIONS

I do not want to claim that the notion of the photon wave function automatically leads to some new physics. It might, however, play a useful pedagogical role, providing a ‘‘missing link’’ between classical Maxwell theory and quantum electrodynamics. With the help of this concept one may unify the treatment of massive and massless particles in quantum theory. In particular, some solutions of the Maxwell equations receive a fresh interpretation as stationary photon states closely resembling analogous states of massive particles. One may also be able to give a direct physical meaning to some essentially complex solutions of the Maxwell equations. Such solutions arise, for example, when one imposes boundary conditions describing absorption of photons at the boundary.

As has been pointed out by Eberly [32], there should also exist a connection between the photon wave mechanics and semiclassical radiation theory. Namely, by introducing material interactions microscopically, via the complex polarization vector  $\vec{P} + i\vec{M}$  given in terms of the Schrödinger wave function, instead of macroscopically via  $\epsilon$  and  $\mu$ , one may obtain the coupled Maxwell-Schrödinger wave equations. The difference with more conventional semiclassical or neoclassical radiation theory would arise from the probabilistic character of the radiation field.

There is finally the global aspect of quantum theory that is handled much more easily at the level of wave mechanics than at the level of local field operators. The requirement that all wave functions must be regular and single valued plays an important role in quantum theory of particles with mass. I believe that the notion of photon wave function might be of importance whenever the global properties of photon wave mechanics are relevant. Such problems arise, for example, when the propagation of probability waves takes place in multiply connected regions. In particular, one may look for an analog of the Aharonov-Bohm effect for photons.

### XIV. ACKNOWLEDGMENTS

I would like to acknowledge the hospitality extended to me by Max-Planck-Institut für Quantenoptik in Garching where this work has been completed. I would like to thank Marlan O.Scully for many discussions and also for convincing me that a paper on the photon wave function should be written. This research has been supported by the KBN Grant 2-2418-92-03.

- 
- [1] P. A. M. Dirac, *The Principles of Quantum Theory*, 4 ed., Clarendon Press, Oxford 1958.
- [2] G. Baym, *Lectures on Quantum Mechanics*, Benjamin, Reading 1969.
- [3] H. J. Lipkin, *Quantum Mechanics*, North-Holland, Amsterdam 1973.
- [4] C. Cohen-Tannoudji, B. Diu, and F. Laloë, *Quantum Mechanics*, Vol.I, Wiley, New York 1977.
- [5] H. A. Kramers, *Quantentheorie des Elektrons und der Strahlung* in Hand- und Jahrbuch der Chemischen Physik, Eucken-Wolf, Leipzig, 1938 (English translation *Quantum Mechanics*, North-Holland, Amsterdam 1957).
- [6] D. Bohm, *Quantum Theory*, Constable, London 1954.
- [7] E. A. Power, *Introductory Quantum Electrodynamics*, Longmans, London 1964.
- [8] L. Silberstein, Ann. d. Phys.**22**,579 (1907); **24**, 783 (1907). In the second paper Silberstein writes that the complex form of Maxwell equations has been known before. He refers there to the second volume of the lecture notes on the differential equations of mathematical physics by B. Riemann that were edited and published by H. Weber in 1901.
- [9] H. Bateman, *The Mathematical Analysis of Electrical and Optical Wave Motion on the Basis of Maxwell's Equations*, Cambridge, 1915 (reprinted by Dover, New York 1955).
- [10] E. Majorana (unpublished notes), quoted after R. Mignani, E. Recami, and M. Baldo, Lett.Nuovo Cimento **11**, 568 (1974).
- [11] L. D. Landau and R. Peierls, Z. Phys.**62**, 188 (1930).
- [12] W. Pauli, *Prinzipien der Quantentheorie*, Handbuch der Physik, Vol.24, Springer, Berlin, 1933, (English translation: *General Principles of Quantum Mechanics*, Springer, Berlin, 1980).
- [13] R. J. Cook, Phys. Rev. **A25**, 2164 (1982); **26**, 2754 (1982).
- [14] P. A. M. Dirac, The quantum theory of the electron, Proc. Roy. Soc.(London)**A117**, 610 (1928);**A118**, 351 (1928).
- [15] H. Weyl, Z. Phys.**56**, 330 (1929).
- [16] I. Bialynicki-Birula and Z. Bialynicka-Birula, *Quantum Electrodynamics*, Pergamon, Oxford 1976.
- [17] C. Itzykson and J.-B. Zuber, *Quantum Field Theory*, McGraw-Hill, New York 1980, p.50.
- [18] L. C. Biedenharn and J. D. Louck, *Angular Momentum in Quantum Physics*, Addison-Wesley, Reading 1981.
- [19] J. D. Jackson, *Classical Electrodynamics*, Wiley, New York 1975.
- [20] G. B. Skrotskii, Soviet Physics Doklady **2**, 226 (1957).
- [21] J. Plebanski, Phys. Rev.**118**, 1396 (1959) (for a pedagogical review see W. Schleich and M. O. Scully, General Relativity and Modern Optics in *New Trends in Atomic Physics*, Les Houches, Session XXXVIII, Eds. G. Grynberg and R. Stora, Elsevier, Amsterdam 1984, p. 995).
- [22] E. P. Wigner, Phys. Rev.**40**, 749 (1932).
- [23] A. Walther, J. Opt. Soc. Am.**68**, 1256 (1968).
- [24] E. Wolf, Phys. Rev. **D13**, 869 (1976).
- [25] E. C. G. Sudarshan, Phys. Lett. **A73**, 269 (1979).
- [26] E. C. G. Sudarshan, Physica **A96**, 315 (1981).
- [27] E. C. G. Sudarshan, Phys. Rev. **A23**, 2803 (1981).
- [28] I. Bialynicki-Birula, P. Górnicki, and J. Rafelski, Phys. Rev. **D44**, 1825 (1991).
- [29] R. Penrose and W. Rindler, *Spinors and Space-Time*, Cambridge University Press, Cambridge 1984, Vol.I, Ch. 5.
- [30] R. F. Streater and A. S. Wightman, *PCT, Spin and Statistics, and All That*, Benjamin, Reading 1978.
- [31] W. L. Bade and H. Jehle, Rev. Mod. Phys. **25**, 714 (1953).
- [32] J. H. Eberly, private communication.