PHYSICAL REVIEW A VOLUME 30, NUMBER 4 OCTOBER 1984

Quantization of electrodynamics in nonlinear dielectric media

Mark Hillery
Institute for Modern Optics, Department of Physics and Astronomy, University of New Mexico, Albuquerque, New Mexico 87131

Leonard D. Mlodinow*†
California Institute of Technology, Pasadena, California 91125
and Max-Planck-Institut für Quantenoptik, D-8046 Garching bei München, West Germany
(Received 19 August 1983; revised manuscript received 13 April 1984)

We point out some problems with the usual quantum-mechanical theory of electrodynamics in nonlinear dielectric media which is used in nonlinear optics. In order to understand these problems, the Hamiltonian formulation of the theory is examined. It is found that many of the difficulties in the usual theory are a result of the fact that the canonical momentum for the interacting theory is not the same as that for the free electromagnetic field theory.

I. INTRODUCTION

Electrodynamics in a dielectric medium is described by the macroscopic Maxwell equations

$$\vec{\nabla}\cdot\vec{B}=0\,, \tag{1.1a}$$

$$\vec{\nabla}\times\vec{E}=-\dot{\vec{B}}\,, \tag{1.1b}$$

$$\vec{\nabla}\cdot\vec{D}=\rho_{\rm ex}\,, \tag{1.1c}$$

$$\vec{\nabla}\times\vec{B}=\dot{\vec{D}}+\vec{j}_{\rm ex}\,. \tag{1.1d}$$

Here $\vec{D}=\vec{E}+\vec{P}$ is the displacement field, $\rho_{\rm ex}$ and $\vec{j}_{\rm ex}$ represent sources that are not considered part of the dielectric medium, $\vec{P}$ is the polarization of the medium, and Heaviside-Lorentz units have been used. The polarization is a function of the electric field which may be written as a power series, i.e.,

$$\vec{P}=\chi^{(1)}{:}\vec{E}+\chi^{(2)}{:}\vec{E}\vec{E}+\chi^{(3)}{:}\vec{E}\vec{E}\vec{E}\,. \tag{1.2}$$

The quantities $\chi^{(j)}$ are the $(j+1)$th-rank susceptibility tensors.[1] Throughout this paper we will assume that the medium is lossless, nondispersive, and uniform. Under these conditions the quantities $\chi^{(j)}$ are symmetric tensors. These equations form the basis of the theory of nonlinear optical effects in matter.

For most applications in nonlinear optics the electromagnetic field can be treated classically. When one is interested in the photon statistics of the field, however, a quantum-mechanical description is necessary. It is then necessary to find a quantum theory capable of describing the effects inherent in Eqs. (1.1) and (1.2) above.

A quantum formulation of the linear macroscopic theory was given as early as 1948 in a series of papers by Jauch and Watson.[2] They used their theory to analyze Čerenkov radiation and also discussed the problems associated with dispersive media. The standard macroscopic quantum theory of electrodynamics in a nonlinear medium which is used in nonlinear optics is due to Shen[3] and was elaborated upon by Tucker and Walls.[4]

One can approach the problem of deriving a macroscopic quantum theory in two ways. One way would be to begin from a quantum microscopic theory and then to eliminate the matter degrees of freedom in order to obtain a radiation theory which depends only upon the bulk properties of the material. This method was explored in the linear case by Hopfield.[5] The other approach is to quantize the macroscopic classical theory. This is the approach of the standard theory. The Hamiltonian is taken to be

$$H=H_{\rm EM}+H_I\,,$$

where

$$\begin{aligned}H_{\rm EM}&=\tfrac{1}{2}\int d^3x\,(\vec{E}^2+\vec{B}^2)\,,\\ H_I&=-\int d^3x\,\vec{E}\cdot\vec{P}\,.\end{aligned} \tag{1.3}$$

The electric and magnetic fields are expressed in terms of the vector potential $\vec{A}$:

$$\vec{E}=-\dot{\vec{A}},\quad \vec{B}=\vec{\nabla}\times\vec{A}\,, \tag{1.4}$$

where $\vec{A}$ is assumed to satisfy

$$\vec{\nabla}\cdot\vec{A}=0\,. \tag{1.5}$$

The theory is quantized by replacing the classical vector potential by the following operator, familiar from ordinary QED:

$$\begin{aligned}\vec{A}(\vec{x},t)=\frac{1}{\sqrt{V}}\sum_{\vec{k},\alpha}\frac{1}{\sqrt{2\omega_{\vec{k}}}}&[a_{\vec{k},\alpha}(t)\hat{\epsilon}^{(\alpha)}_{\vec{k}}e^{i\vec{k}\cdot\vec{x}}\\ &+a^{\dagger}_{\vec{k},\alpha}(t)\hat{\epsilon}^{(\alpha)}_{\vec{k}}e^{-i\vec{k}\cdot\vec{x}}]\,,\\ \vec{k}\cdot\hat{\epsilon}^{(\alpha)}_{\vec{k}}=0,\quad [a_{\vec{k},\alpha}(t),a^{\dagger}_{\vec{k}',\alpha'}(t)]&=\delta_{\vec{k}\vec{k}'}\delta_{\alpha\alpha'}\,.\end{aligned} \tag{1.6}$$

There are a number of problems with this theory, but the greatest is that it is not consistent with Eqs. (1.1). For example, Eqs. (1.4) and (1.5) imply that $\vec{\nabla}\cdot\vec{E}=0$, not

$\vec{\nabla}\cdot\vec{D}=0$ as should be the case. Equation (1.1b) is also not a result of this theory. If we compute $\dot{\vec{B}}$ we find

$$\dot{B}_j=-i[B_j,H]=-\epsilon_{jkl}\frac{\partial E_l}{\partial x_k}+2\int d^3x'\,\epsilon_{jkl}\frac{\partial}{\partial x_k}\delta^{\rm tr}_{lm}(\vec{x}-\vec{x}\,')\chi^{(1)}_{mn}E_n(\vec{x}\,',t)$$
$$+3\int d^3x'\,\epsilon_{jkl}\frac{\partial}{\partial x_k}\delta^{\rm tr}_{lm}(\vec{x}-\vec{x}\,')\chi^{(2)}_{mnp}E_n(\vec{x}\,',t)E_p(\vec{x}\,',t)+\cdots$$
$$=[-\vec{\nabla}\times\vec{E}+2\vec{\nabla}\times(\underline{\chi}^{(1)}{:}\vec{E})+3\vec{\nabla}\times(\underline{\chi}^{(2)}{:}\vec{E}\vec{E})+\cdots]_j\ , \tag{1.7}$$

where $\delta^{\rm tr}_{jk}(\vec{x})$ is the transverse delta function,[6] ϵ_{jkl} is the completely antisymmetric tensor of rank 3, and repeated indexes are always summed over. Equation (1.7) does not agree with Eq. (1.1b).

II. CANONICAL APPROACH

In order to find out what the problem is we return to the classical theory and examine its Hamiltonian formulation. We first find a Lagrangian which gives the proper equations of motion and then calculate the Hamiltonian. One can then quantize the theory in the canonical way. An appropriate Lagrangian density is

$$L(A,\dot{A})=\tfrac{1}{2}(\vec{E}^2-\vec{B}^2)+\tfrac{1}{2}\chi^{(1)}_{ij}E_iE_j+\tfrac{1}{3}\chi^{(2)}_{ijk}E_iE_jE_k+\tfrac{1}{4}\chi^{(3)}_{ijkl}E_iE_jE_kE_l+\cdots\ , \tag{2.1}$$

where A is the four-vector $A=(A_0,\vec{A})$ and

$$\vec{E}=-\dot{\vec{A}}-\vec{\nabla}A_0,\quad \vec{B}=\vec{\nabla}\times\vec{A}\ . \tag{2.2}$$

Equations (2.2) imply Eqs. (1.1a) and (1.1b) whereas Lagrange's equations

$$\frac{\partial}{\partial x^\mu}\frac{\partial L}{\partial(\partial_\mu A_j)}-\frac{\partial L}{\partial A_j}=0 \tag{2.3}$$

imply Eq. (1.1d) (without external sources), and Lagrange's equation

$$\frac{\partial}{\partial x^\mu}\frac{\partial L}{\partial(\partial_\mu A_0)}-\frac{\partial L}{\partial A_0}=0 \tag{2.4}$$

implies Gauss's law Eq. (1.1c) (again without external sources).

The momentum canonical to A is $\Pi=(\Pi_0,\vec{\Pi})$ where

$$\Pi_0=\frac{\partial L}{\partial(\partial_0 A_0)}=0,\quad \Pi_i=\frac{\partial L}{\partial(\partial_0 A_i)}=-D_i\ . \tag{2.5}$$

The vanishing of Π_0 indicates that this system is constrained, and we must proceed with caution. The correct Hamiltonian and revised Poisson brackets (Dirac brackets) may be obtained by following the prescription given by Dirac for constrained Hamiltonian systems.[7–9] This is done in Appendix A.

One may obtain the same result using the heuristic arguments often applied to ordinary electrodynamics.[6] The vanishing of Π_0 indicates that A_0 is not an independent field. In macroscopic electrodynamics the divergence of $\vec{A}$ in the temporal gauge is not time independent, so that one cannot impose both the temporal-gauge condition

$$A_0=0 \tag{2.6}$$

and the Coulomb-gauge condition

$$\vec{\nabla}\cdot\vec{A}=0 \tag{2.7}$$

(as was done in the standard theory). Thus, if we maintain the Coulomb gauge, we must solve for A_0 in terms of the independent fields:

$$A_0=-\nabla^{-2}(\vec{\nabla}\cdot\vec{E})\ . \tag{2.8}$$

This is also true in ordinary electrodynamics, when charges are present. In that case Eq. (2.8) reads

$$A_0(\vec{x},t)=\frac{1}{4\pi}\int d^3x'\frac{\rho(\vec{x}\,',t)}{|\vec{x}-\vec{x}\,'|}\ , \tag{2.9}$$

where ρ is the charge density. In the present case, however, the situation is more complicated: instead of ρ, one has $\vec{\nabla}\cdot\vec{E}$, where for the purposes of the Hamiltonian formulation, E must be expressed in terms of the canonical momentum $-\vec{D}$. Toward that end we define the tensors $\underline{\beta}^{(i)}$ by

$$E_i=\beta^{(1)}_{ij}D_j+\beta^{(2)}_{ijk}D_jD_k+\cdots\ . \tag{2.10}$$

Using Eq. (1.2) and the definition of $\vec{D}$ we may solve for the $\underline{\beta}^{(i)}$ tensors in terms of the susceptibility tensors $\underline{\chi}^{(i)}$. For example,

$$\begin{aligned}&\underline{\beta}^{(1)}=[\underline{1}+\underline{\chi}^{(1)}]^{-1}\ ,\\&\beta^{(2)}_{imn}=-\beta^{(1)}_{ji}\beta^{(1)}_{km}\beta^{(1)}_{ln}\chi^{(2)}_{jkl}\ ,\\&\beta^{(3)}_{ijkl}=-\beta^{(1)}_{ai}\beta^{(1)}_{bj}\beta^{(1)}_{ck}\beta^{(1)}_{dl}\chi^{(3)}_{abcd}+2\beta^{(2)}_{ijs}\beta^{(2)}_{ckl}\chi^{(1)}_{cs}\ .\end{aligned} \tag{2.11}$$

One can then substitute Eq. (2.10) into Eq. (2.8) in order to solve for A_0 in terms of the canonical momenta.

Since A_0 is not an independent dynamical variable we lose Gauss's law as an equation of motion. It follows, however, from the other equations of motion [Eq. (1.1d) in the absence of external current] that the divergence of $\vec{D}$ is time independent; hence Gauss's law may be imposed as a constraint on the system, having the status of an initial condition.

We can now derive the Hamiltonian. We have that

$$\begin{aligned}H(\vec{A},\vec{\Pi})&=\int d^3x(\Pi_l\dot{A}_l-L)\\&=H_{\rm EM}+\int d^3x\left(\tfrac{1}{2}\chi^{(1)}_{ij}E_iE_j+\tfrac{2}{3}\chi^{(2)}_{ijk}E_iE_jE_k+\tfrac{3}{4}\chi^{(3)}_{ijkl}E_iE_jE_kE_l+\ldots\right)\\&\quad+\int d^3x\,\vec{D}\cdot\vec{\nabla}A_0\ .\end{aligned} \tag{2.12}$$

An integration by parts combined with the constraint $\vec{\nabla}\cdot\vec{D}=0$ allows us to eliminate the last term leaving

$$H(\vec{A},\Pi)=H_{\rm EM}+\int d^3x\left(\tfrac{1}{2}\chi_{ij}^{(1)}E_iE_j+\tfrac{2}{3}\chi_{ijk}^{(2)}E_iE_jE_k+\tfrac{3}{4}\chi_{ijkl}^{(3)}E_iE_jE_kE_l+\cdots\right), \tag{2.13}$$

where $\vec{E}$ and $\vec{B}$ are to be considered as functions of $\vec{A}$ and $\vec{\Pi}$ [see Eqs. (2.2) and (2.10)]. The Hamiltonian can also be expressed in terms of the displacement field as

$$H(\vec{A},\Pi)=\tfrac{1}{2}\int d^3x\,\vec{B}^2+\int d^3x\left(\tfrac{1}{2}\beta_{ij}^{(1)}D_iD_j+\tfrac{1}{3}\beta_{ijk}^{(2)}D_iD_jD_k+\tfrac{1}{4}\beta_{ijkl}^{(3)}D_iD_jD_kD_l+\cdots\right). \tag{2.14}$$

The expression in Eq. (2.13) should be compared with the expression in Eq. (1.3). It should also be noted that for the case of a linear dielectric only Eq. (2.13) reduces to

$$H_{\rm lin}(\vec{A},\vec{\Pi})=\tfrac{1}{2}\int d^3x(\vec{E}\cdot\vec{D}+\vec{B}^2) \tag{2.15}$$

which is the expected result.

Before proceeding it should be noted that there is an alternative procedure for deriving a Hamiltonian theory which is simpler than the one which has just been presented.[10] It makes use of a different type of vector potential. We discuss it in Appendix B. We have chosen to present the theory in terms of the standard vector potential in this section because it more closely parallels what is done in deriving a Hamiltonian formulation of the free electromagnetic field.

III. QUANTIZATION

The theory can now be quantized in the usual way. As is the case with the free electromagnetic field, the canonical commutation relations must be altered due to the vanishing of the divergences of A and Π. The commutation relations

$$[A_i(\vec{x},t),\Pi_j(\vec{x}\,',t)]=i\delta_{ij}^{\rm tr}(\vec{x}-\vec{x}\,') \tag{3.1}$$

are consistent with these constraints. A rigorous treatment, following Dirac's prescription confirms the validity of these results. The commutation relations used in the standard theory, Eq. (1.6), correspond to Eq. (3.1), but with $\vec{\Pi}=-\vec{E}$, rather than $\vec{\Pi}=-\vec{D}$. The fact that $-\vec{D}$ rather than $-\vec{E}$ is the proper canonical momentum was first noted in a somewhat different context by Born and Infeld.[11,12] It has also come up more recently in the discussion of Hamiltonians describing the interaction of the electromagnetic field with atoms.[13,14]

As in the case of free QED one can perform a mode expansion of the fields and define creation and annihilation operators. In particular one has for $a_{\vec{k},\alpha}(t)$

$$a_{\vec{k},\alpha}(t)=\frac{1}{\sqrt{V}}\int d^3x\,e^{-i\vec{k}\cdot\vec{x}}\hat{\epsilon}_{\vec{k}}^{(\alpha)}\cdot\left[\left(\frac{\omega_{\vec{k}}}{2}\right)^{1/2}\vec{A}(\vec{x},t)-\frac{i}{\sqrt{2\omega_{\vec{k}}}}\vec{D}(\vec{x},t)\right] \tag{3.2}$$

with a similar expression for $a_{\vec{k},\alpha}^{\dagger}(t)$. These operators obey the commutation relations given in Eq. (1.6). Note that because $a_{\vec{k},\alpha}(t)$ depends on $\vec{D}$ (in the free-field case $\vec{D}$ is replaced by $-\dot{\vec{A}}$) it contains both field and matter degrees of freedom. It represents, therefore, a collective matter-field mode. This must be kept in mind when interpreting the theory. In principle any difficulties in interpretation can be dealt with by confining the medium to part of the quantization volume and placing the field source and the detector outside of the medium. This corresponds to the actual experimental set up. The field is generated in free space by the source, propagates through the medium, and is measured by the detector in free space. Thus, the state preparation and measurement take place in regions which are free from the complicating effects of the nonlinear medium. This approach, however, requires the consideration of propagation effects.

We now come to the question of operator ordering. How should the operators in the Hamiltonian be ordered? To consider this question let us look at the part of the Hamiltonian, Eq. (2.14), which is quadratic in the field operators

$$H_2=\tfrac{1}{2}\int d^3x(\vec{D}^2+\vec{B}^2)+\tfrac{1}{2}\int d^3x\,\beta_{ij}^{(1)\prime}D_iD_j\,, \tag{3.3}$$

where we have defined $\beta_{ij}^{(1)}=\delta_{ij}+\beta_{ij}^{(1)\prime}$. The operators $a_{\vec{k},\alpha}$ defined in Eq. (3.2) diagonalize the first term of Eq. (3.3), i.e., if one solves Eq. (3.2) and its conjugate for $\vec{A}(\vec{x},t)$ and $\vec{D}(\vec{x},t)$ and substitutes them into the first term of Eq. (3.3) one finds (after subtracting an infinite constant)

$$\tfrac{1}{2}\int d^3x(\vec{D}^2+\vec{B}^2)=\sum_{\vec{k},\alpha}\omega_{\vec{k}}a_{\vec{k},\alpha}^{\dagger}a_{\vec{k},\alpha}\,. \tag{3.4}$$

It is also possible to find annihilation operators $b_{\vec{k},\alpha}$ (where $b_{\vec{k},\alpha}$ is a linear combination of $a_{\vec{k},\alpha'}$, $a_{\vec{k},\alpha'}^{\dagger}$, $a_{-\vec{k},\alpha'}$, and $a_{-\vec{k},\alpha'}^{\dagger}$) which will diagonalize H_2. The nonlinear part of the Hamiltonian, Eq. (2.14) can be expressed in terms of the $a_{\vec{k},\alpha}$'s or the $b_{\vec{k},\alpha}$'s. We can then normal order with respect to the $a_{\vec{k},\alpha}$'s or with respect to the $b_{\vec{k},\alpha}$'s. The results are not the same. If we follow the usual procedure in quantum optics and restrict the Hamiltonian to only a few modes the effect of these different orderings will show up. The nonlinear terms would, for example, produce different frequency shifts in the Hamiltonian normally ordered with respect to $a_{\vec{k},\alpha}$ than in that normally ordered with respect to $b_{\vec{k},\alpha}$. Whether a full renormalization scheme is feasible and would correct this is not at all clear.

In standard QED physical considerations dictate the ordering. There, normal ordering with respect to the field operators simply subtracts an infinite constant from the Hamiltonian. The normal ordering of the matter (electron and positron) operators comes from demanding that the expectation value of the current operator in the physical vacuum vanish. One can show by invoking the charge conjugation invariance of the physical vacuum that if the current operator is normally ordered then this requirement is satisfied.[6]

It is not obvious how to find a workable ordering criterion for the theory we are considering or, in fact, whether one exists. The resolution of this problem would seem to require the consideration of a microscopic theory, i.e., a model for the medium.

IV. CONCLUSION

We have discussed problems in the usual quantum theory for the electromagnetic field in a nonlinear dielectric medium, and derived an alternative according to the canonical approach. We encountered problems in the quantum theory related to operator ordering.

Further difficulties are encountered if one wants to generalize this theory by including dispersion. Strictly speaking, the expansion, Eq. (1.2), of the polarization holds only in the static case. It is a good approximation if the electric field varies much more slowly than the time it takes for the atoms in the dielectric to adjust themselves. For more rapidly varying fields, the polarization at a time t depends, not only on the value of the field at that time, but on its value at all previous times. For example, for a linear dispersive system, the polarization is[15]

$$\vec{P}(t)=\int_0^{\infty} d\tau \chi^{(1)}(\tau):\vec{E}(t-\tau) . \tag{4.1}$$

This dependence on previous times makes the Hamiltonian formulation of this theory, and, hence, its quantization, problematic.

These problems lead us to the conclusion that the best way to proceed is to consider a microscopic theory, i.e., fields plus matter, and to see if an effective macroscopic theory can be developed. The discussion in Sec. III suggests that this effective theory should have as its basic objects collective matter-field modes. In a future publication we will show how such a theory can be derived for a medium consisting of two-level atoms.

ACKNOWLEDGMENTS

We would like to thank W. Becker, P. Breithenlohner, A. Caticha, N. Caticha, C. Caves, J. Cresser, J. K. McIver, and E. Seiler for interesting discussions. We would also like to thank I. Białynicki-Birula and S. Stenholm for useful comments. This research was supported by U.S. Kirkland Air Force Base under Contract No. F29601-82-K-0017 and by the National Science Foundation under Grant No. MCS-81-19979. One of us (L.D.M.) would like to acknowledge the Alexander von Humboldt Foundation for financial support.

APPENDIX A

A rigorous derivation of the results in Sec. II can be obtained by employing the Dirac quantization procedure for constrained Hamiltonian systems.[7–9] The calculation is similar to that for ordinary QED.[8] The vanishing of Π_0 is a primary constraint on the system. Since Π_0 must vanish for all times, its commutation with the Hamiltonian must also vanish, which leads to the secondary constraint

$$\vec{\nabla}\cdot\vec{\Pi}=0 \tag{A1}$$

which is just Gauss's law. This constraint commutes with the Hamiltonian, so there are no further secondary constraints. Let us write these as

$$\zeta_1=\Pi_0=0,\quad \xi_1=\vec{\nabla}\cdot\vec{\Pi}=0 . \tag{A2}$$

It is easy to see that these constraints have a vanishing Poisson bracket with each other, hence they are first-class constraints, and the equations of motion will involve two arbitrary functions of time. This is just the gauge freedom. We fix the gauge by introducing one "gauge constraint" for each first-class constraint. We shall use the radiation gauge [Eqs. (2.7) and (2.8)], which we write as

$$\zeta_2=A_0+\nabla^{-2}(\vec{\nabla}\cdot\vec{E})=0,\quad \xi_2=\vec{\nabla}\cdot\vec{A}=0 \tag{A3}$$

(remembering that $\vec{E}$ is to be considered a function of $\vec{\Pi}$). None of the constraints in Eqs. (A2) and (A3) has zero Poisson bracket with all the others, so, rather than two first-class constraints, we now have four second-class constraints. The matrix of their Poisson brackets is nonsingular, and may be used to calculate the Dirac bracket.

To calculate the Dirac bracket, we first define the operator $\mathscr{E}$ to have the kernel

$$e_{ij}(\vec{x},\vec{y})=\{\xi_i(\vec{x}),\xi_j(\vec{y})\} . \tag{A4}$$

The symbol $\{\,,\,\}$ denotes the Poisson bracket. Then we define a modified Poisson bracket by

$$\begin{aligned}\{A(\vec{x}),B(\vec{y})\}' &\equiv\{A(\vec{x}),B(\vec{y})\} - \sum_{i,j}\int d^3u\int d^3v\{A(\vec{x}),\xi_i(\vec{u})\}\\ &\quad\times e_{ij}^{-1}(\vec{u},\vec{v})\{\xi_j(\vec{v}),B(\vec{y})\} ,\end{aligned} \tag{A5}$$

where $e_{ij}^{-1}(\vec{x},\vec{y})$ is the kernel for $\mathscr{E}^{-1}$. Similarly, we define the operator $\mathscr{F}$ to have the kernel

$$f_{ij}(\vec{x},\vec{y})=\{\zeta_i(\vec{x}),\zeta_j(\vec{y})\} . \tag{A6}$$

The Dirac bracket is then defined by[8]

$$\begin{aligned}\{A(\vec{x}),B(\vec{y})\}_{\rm Dirac} &\equiv\{A(\vec{x}),B(\vec{y})\}' - \sum_{i,j}\int d^3u\int d^3v\{A(\vec{x}),\zeta_i(\vec{u})\}'\\ &\quad\times f_{ij}^{-1}(\vec{u},\vec{v})\{\zeta_i(\vec{v}),B(\vec{y})\}' .\end{aligned} \tag{A7}$$

Note that of the constraints, only ζ_2 differs from that of ordinary QED, and that the operators $\mathscr{E}$ and $\mathscr{F}$ are identical to their QED analogs. It is thus easily verified that the Dirac brackets between components of $\vec{A}$ and $\vec{\Pi}$ are unchanged from those of ordinary QED, except for the difference in the definition of $\vec{\Pi}$. Since all the constraints are second class, we may consider them to hold in the strong sense and use the canonical Hamiltonian given in Eq. (2.14). To quantize, we equate the Dirac bracket with i times the commutator, which results in Eq. (3.1).

APPENDIX B

One may simplify quantization of the macroscopic Maxwell theory by redefining the vector potential. We replace Eq. (2.2) by

$$\vec{D}=\vec{\nabla}\times\vec{\Lambda},\quad \vec{B}=\dot{\vec{\Lambda}}+\vec{\nabla}\Lambda_0 . \tag{B1}$$

Equations (B1) imply Eqs. (1.1c) and (1.1d) without external sources, whereas Eqs. (1.1a) and (1.1b) determine the Lagrangian.

If we compare Eqs. (1.1c) and (1.1d) to Eqs. (1.1a) and (1.1b), the choice of working with A or $\Lambda=(\Lambda_0,\vec{\Lambda})$ appears to make no essential difference. The use of Λ, however, greatly simplifies the quantization of the theory. In both cases the Lagrangian is quadratic in $\vec{B}$, but not in $\vec{D}$ (or $\vec{E}$). When the vector potential Λ is used, the Lagrangian is, therefore, quadratic in time derivatives of the field.

In order to distinguish between A and Λ, we call the latter the "dual potential." Technically speaking, we should call it the dual potential only in the absence of a dielectric ($\vec{D}=\vec{E}$). In that case, Maxwell's equations are invariant under the transformation

$$\vec{E}\to\vec{B},\quad \vec{B}\to-\vec{E}\ , \tag{B2}$$

which is the same as the duality transformation of the electromagnetic field tensor[16] $F_{\lambda\rho}=\partial A_\rho/\partial x^\lambda-\partial A_\lambda/\partial x^\rho$:

$$F^*_{\mu\nu}=\epsilon^{\lambda\rho\mu\nu}F_{\lambda\rho}\ . \tag{B3}$$

Here $\epsilon^{\lambda\rho\mu\nu}$ is the totally antisymmetric tenor of rank 4. This transformation exchanges A and Λ. In the presence of a dielectric, (B2) no longer represents a symmetry of Maxwell's equation, but we may still use Λ, provided we define it in terms of $\vec{D}$ and $\vec{B}$, as in Eq. (B1).

Apart from the restriction to dispersionless media, there is one other drawback of this formulation of the theory. That arises from the fact that inclusion of an external charge or current in Eqs. (1.1c) and (1.1d) invalidates the definition Eq. (B1). In the case of a stationary current ($\vec{\nabla}\cdot\vec{j}=0=\partial\rho/\partial t$) this can be remedied, but only at the expense of a more complicated theory.

The usual expression for the polarization density, Eq. (1.2) is no longer convenient. Instead, we write

$$\vec{P}=\underline{\eta}^{(1)}{:}\vec{D}+\underline{\eta}^{(2)}{:}\vec{D}\vec{D}+\underline{\eta}^{(3)}{:}\vec{D}\vec{D}\vec{D}+\cdots\ . \tag{B4}$$

One can calculate the η's from the X's, and vice versa. They are simply related to the β's we used in Sec. II [Eq. (2.11)]:

$$\underline{\eta}^{(1)}=\underline{1}-\underline{\beta}^{(1)},\quad \underline{\eta}^{(j)}=-\underline{\beta}^{(j)},\quad j=2,3,\ldots\ . \tag{B5}$$

The Lagrangian density that reproduces Eqs. (1.1a) and (1.1b) is

$$\mathscr{L}=\mathscr{L}_{\text{free}}+\tfrac{1}{2}\vec{D}\cdot\underline{\eta}^{(1)}{:}\vec{D}+\tfrac{1}{3}\vec{D}\cdot\underline{\eta}^{(2)}{:}\vec{D}\vec{D}+\tfrac{1}{4}\vec{D}\cdot\underline{\eta}^{(3)}{:}\vec{D}\vec{D}\vec{D}+\cdots\ ,\qquad \mathscr{L}_{\text{free}}=\tfrac{1}{2}(\vec{B}^2-\vec{D}^2)\ . \tag{B6}$$

Variation with respect to Λ_0 gives

$$\vec{\nabla}\cdot\vec{B}=0 \tag{B7}$$

and variation with respect to $\vec{\Lambda}$ gives

$$\vec{\nabla}\times\vec{E}=-\dot{\vec{B}}\ . \tag{B8}$$

The canonical momenta are

$$\Pi_0=\frac{\partial\mathscr{L}}{\partial(\partial_0\Lambda_0)}=0,\quad \Pi_j=\frac{\partial\mathscr{L}}{\partial(\partial_0\Lambda_j)}=-B_j\ . \tag{B9}$$

These are the same as in the free case, because no time derivative of Λ appears in the interaction Lagrangian. This is what makes quantization of the theory easier. The canonical Hamiltonian density is now

$$\mathscr{H}=\mathscr{H}_{\text{free}}-\tfrac{1}{2}\vec{D}\cdot\underline{\eta}^{(1)}{:}\vec{D}-\tfrac{1}{3}\vec{D}\cdot\underline{\eta}^{(2)}{:}\vec{D}\vec{D}-\tfrac{1}{4}\vec{D}\cdot\underline{\eta}^{(3)}{:}\vec{D}\vec{D}\vec{D}-\vec{B}\cdot\vec{\nabla}\Lambda_0\ ,\qquad \mathscr{H}_{\text{free}}=\tfrac{1}{2}(\vec{B}^2+\vec{D}^2)\ . \tag{B10}$$

Note that the vanishing of Π_0 leads to the secondary constraint $\vec{\nabla}\cdot\vec{\Pi}=0$, which is just Eq. (B7). Before quantizing the theory, we fix the gauge. The physical fields are unchanged under the transformation

$$\vec{\Lambda}\to\vec{\Lambda}+\vec{\nabla}\Theta,\quad \Lambda_0\to\Lambda_0-\dot{\Theta}\ , \tag{B11}$$

where Θ is an arbitrary function of space-time. We may thus still impose the Coulomb gauge. To eliminate Λ_0, choose Θ to be a solution of

$$\dot{\Theta}=\Lambda_0\ ; \tag{B12}$$

this determines Θ up to an arbitrary function of $\vec{x}$, which we denote $\theta(x)$. Since $\vec{\nabla}\cdot\vec{\Lambda}$ is time independent,

$$\frac{\partial}{\partial t}(\vec{\nabla}\cdot\vec{\Lambda})=\vec{\nabla}\cdot\dot{\vec{\Lambda}}=\vec{\nabla}\cdot\vec{B}=0\ , \tag{B13}$$

we may choose θ so that $\vec{\nabla}\cdot\vec{\Lambda}=0$. This theory can therefore be quantized in exactly the same manner as free QED. The commutation relations are

$$[\dot{\Lambda}_i(\vec{x},t),\Lambda_j(\vec{x}\,',t)]=i\delta^{\text{tr}}_{ij}(\vec{x}-\vec{x}\,')\ . \tag{B14}$$

It is easily confirmed that Eq. (B14) and the Heisenberg equations lead to (B8) and (1.1d) (with $\vec{j}_{\text{ex}}=\vec{0}$) as the quantum-mechanical equations of motion.

*On leave from the California Institute of Technology, Pasadena, CA 91125.

†Present address: Max Planck Institut für Physik und Astro physik, Föhringer Ring 6, 8 München 40, West Germany.

[1]N. Bloembergen, *Nonlinear Optics* (Benjamin, Reading, Mass., 1965).
[2]J. M. Jauch and K. M. Watson, Phys. Rev. **74**, 950 (1948); **74**, 1485 (1948); **75**, 1249 (1948).
[3]Y. R. Shen, Phys. Rev. **155**, 921 (1967).
[4]J. Tucker and D. F. Walls, Phys. Rev. **178**, 2036 (1969).
[5]J. J. Hopfield, Phys. Rev. **178**, 1555 (1958).
[6]J. D. Bjorken and S. D. Drell, *Relativistic Quantum Fields* (McGraw-Hill, New York, 1965).
[7]P. A. M. Dirac, *Lectures on Quantum Mechanics* (Belfer Graduate School of Science, New York, 1964).
[8]K. Sundermeyer, *Constrained Dynamics* (Springer, Berlin, 1982).
[9]E. C. G. Sudarshan and N. Mukunda, *Classical Dynamics. A Modern Perspective* (Wiley, New York, 1974).
[10]C. Caves (private communication).
[11]M. Born and L. Infeld, Proc. R. Soc. London, **144**, 425 (1934); **147**, 522 (1934); **150**, 141 (1935).
[12]I. Białynicki-Birula and Z. Białynicki-Birula, *Quantum Electrodynamics* (Pergamon, New York, 1975).
[13]J. Savolainen and S. Stenholm, Am. J. Phys. **40**, 667 (1972).
[14]H. P. Healy, *Non-Relativistic Quantum Electrodynamics* (Academic, New York, 1982).
[15]L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Pergamon, New York, 1960).
[16]L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields* (Pergamon, New York, 1975).