Paraxial-wave optics and relativistic front description. II. The vector theory

N. Mukunda*
Institute of Theoretical Physics, S-41296, Göteborg, Sweden

R. Simon†
Department of Physics, Indian Institute of Science, Bangalore 560012, Karnataka, India

E. C. G. Sudarshan‡
Institute of Theoretical Physics, S-41296, Göteborg, Sweden
(Received 20 December 1982)

With the extension of the work of the preceding paper, the relativistic front form for Maxwell's equations for electromagnetism is developed and shown to be particularly suited to the description of paraxial waves. The generators of the Poincaré group in a form applicable directly to the electric and magnetic field vectors are derived. It is shown that the effect of a thin lens on a paraxial electromagnetic wave is given by a six-dimensional transformation matrix, constructed out of certain special generators of the Poincaré group. The method of construction guarantees that the free propagation of such waves as well as their transmission through ideal optical systems can be described in terms of the metaplectic group, exactly as found for scalar waves by Bacry and Cadilhac. An alternative formulation in terms of a vector potential is also constructed. It is chosen in a gauge suggested by the front form and by the requirement that the lens transformation matrix act locally in space. Pencils of light with accompanying polarization are defined for statistical states in terms of the two-point correlation function of the vector potential. Their propagation and transmission through lenses are briefly considered in the paraxial limit. This paper extends Fourier optics and completes it by formulating it for the Maxwell field. We stress that the derivations depend explicitly on the "henochromatic" idealization as well as the identification of the ideal lens with a quadratic phase shift and are heuristic to this extent.

I. INTRODUCTION

In the preceding paper[1] (hereafter referred to as I) we have set up a general formalism, based on the front form of relativistic dynamics,[2] for the treatment of paraxial-wave propagation problems in optics. The treatment was restricted to the case of scalar waves for simplicity, and we analyzed in detail the group-theoretical basis underlying the front form and the paraxial limit. It was shown that both the free propagation of such waves and their passage through ideal optical systems was very similar in mathematical structure to the quantum mechanics of free nonrelativistic particles in two dimensions encountering harmonic impulses. Moreover, the significance of the metaplectic group for this class of problems, first realized in the work of Bacry and Cadilhac,[3] was traced back to the structure of the Poincaré group: It arises from the fact that in the Lie algebra of the Poincaré group there is a subalgebra isomorphic to (a central extension of) the two-dimensional Galilei algebra, and this is exposed by the front form.[4] We also analyzed the behavior of generalized light rays[5] in this context and presented their extremely simple free propagation behavior as well as their passage through ideal lenses.

The present paper extends this work to the complete electromagnetic field described by Maxwell's equations, so that a satisfactory account of polarization in paraxial wave problems can be given. Since in I we have explained in considerable detail how the front form description of wave propagation is related to the more familiar one employing separate space and time coordinates, we shall emphasize in this paper just those features that are particularly associated with the Maxwell system. We shall extend the group theoretical discussion given in I to determine the effect of ideal optical systems on Maxwellian waves, and we shall find that this procedure gives physically correct results. Thus the relevance of the metaplectic group carries over to vector waves.

We use the same metric conventions and terminology for the Poincaré group as in I. Section II puts the Maxwell equations into the front form,[6] pointing out that in contrast to the instant form there are now equal numbers of constraint conditions and dynamical equations. The generators of the Poincaré group, in a form suitable for application to the six independent components of the Maxwell field tensor, are worked out. Paraxial solutions to Maxwell's equations are described in Sec. III. It is a very useful feature of the front form that it shows one how to rearrange the field components in particular combinations and in a specific sequence, which makes the description of paraxial waves most natural. The action of a lens on such waves is determined by the principle that in the lens transformation function of scalar wave theory,[7] the role of the transverse position coordinates must now be played by the conjugates to transverse momentum provided by the Galilean subalgebra of the Poincaré algebra. For vector waves described in terms of the six field strengths $\vec{E},\vec{B}$, each thin lens is then represented by a corresponding transformation matrix, rather than by a function, of dimension six. For the simplest case of an axially symmetric lens this matrix is explicitly computed, and it is shown that its action on an incident wave gives physically expected results. In Sec. IV we give a treatment based on the vector potential, leading to simpler matrix algebra. We describe a new gauge associated with the front form and particularly suited to paraxial waves, and calculate the lens transformation matrix, now three dimensional, again for the axially symmetric case. It is verified that the treatments of Secs. III and IV give mutually consistent results. In Sec. V we show how to define generalized rays of light with polarization properties, in terms of the two-point correlation tensors of the electromagnetic field. Their free propagation and passage through thin lenses is briefly described, always working within the paraxial approximation. Finally we summarize our work in Sec. VI, and especially provide a simple reason for the correctness of our principle based on the underlying group theory.

II. MAXWELL'S EQUATIONS IN THE FRONT FORM

The free Maxwell equations in conventional three-dimensional notation appropriate to the instant form are

$$\partial_0\vec{E}-\vec{\nabla}\times\vec{B}=0\ , \tag{2.1a}$$

$$\partial_0\vec{B}+\vec{\nabla}\times\vec{E}=0\ , \tag{2.1b}$$

$$\vec{\nabla}\cdot\vec{E}=0\ , \tag{2.1c}$$

$$\vec{\nabla}\cdot\vec{B}=0\ . \tag{2.1d}$$

In (2.1a) and (2.1b) we have genuine equations of motion for $\vec{E}$ and $\vec{B}$; (2.1c) and (2.1d) are two constraint equations to be obeyed at each time. Of course, the former ensure that the latter are maintained in time. If a conserved external current j^μ is included as a source, the equation of motion (2.1a) and the constraint (2.1c) acquire $-\vec{j}$ and j^0, respectively, on their right-hand sides, while (2.1b) and (2.1d) are unchanged. As parts of a relativistic covariant antisymmetric tensor field $F_{\mu\nu}$, we identify $\vec{E}$ and $\vec{B}$ by

$$E_j=F_{j0}\ ,\ \ B_j=\tfrac{1}{2}\epsilon_{jkl}F_{kl}\ . \tag{2.2}$$

To express these equations in the front form, we need a set of conventions for defining new components of vectors and tensors. Just as the space-time coordinates x^0,x^3 are replaced by the combinations $\tau=\frac{1}{2}(x^0+x^3)$ and $\sigma=x^0-x^3$, for any contravariant vector X^μ we define

$$X^\tau=\tfrac{1}{2}(X^0+X^3)\ ,\ \ X^\sigma=X^0-X^3\ . \tag{2.3}$$

Similarly for a covariant vector Y_μ we define

$$Y_\sigma=\tfrac{1}{2}(Y_0-Y_3)\ ,\ \ Y_\tau=Y_0+Y_3\ , \tag{2.4}$$

so that the invariant takes the form (indices $a,b,\ldots$, take values 1,2)

$$X^\mu Y_\mu\equiv X_jY_j+X^0Y_0=X_aY_a+X^\sigma Y_\sigma+X^\tau Y_\tau\ . \tag{2.5}$$

The accompanying form of the metric and the rules for raising and lowering indices σ,τ are

$$g_{ab}=\delta_{ab}\ ,\ \ g_{\sigma\tau}=g_{\tau\sigma}=-1\ , \tag{2.6a}$$

$$X^\sigma=-X_\tau\ ,\ \ X^\tau=-X_\sigma\ ,\ \ X^a=X_a\ . \tag{2.6b}$$

Partial derivatives with respect to σ and τ are

$$\partial_\sigma\equiv\frac{\partial}{\partial\sigma}=\tfrac{1}{2}(\partial_0-\partial_3)\ ,\ \ \partial_\tau\equiv\frac{\partial}{\partial\tau}=\partial_0+\partial_3\ . \tag{2.7}$$

Consistent with Eq. (2.6a) we have $\partial^\sigma=-\partial_\tau,\partial^\tau=-\partial_\sigma$. We now introduce combinations U_a,V_a of the transverse components E_a,B_a in the following way:

$$U_a=F_{\sigma a}=-\tfrac{1}{2}(E_a+\epsilon_{ab}B_b)\ , \tag{2.8a}$$

$$V_a=F_{\tau a}=-(E_a-\epsilon_{ab}B_b)\ . \tag{2.8b}$$

The remaining two components E_3,B_3 are carried along unchanged, noting only that

$$E_3=F_{\tau\sigma}\ ,\ \ B_3=\tfrac{1}{2}\epsilon_{ab}F_{ab}\ . \tag{2.9}$$

From these definitions we can see that in the front form (B_3,U_a) form a natural "magnetic" triplet of field components, while (E_3,V_a) form an "electric" triplet. Note that the components U_a are defined in terms of $F_{\mu\nu}$ in the same way in which the Galilean generators G_a are defined starting with the homogeneous Lorentz generators $M_{\mu\nu}$, since $G_a=M_{\sigma a}$. We will see in the next section that the sets of field components $U_a;E_3$ and $B_3;V_a$ are very well suited to describe paraxial waves.

With these definitions it is straightforward to rewrite all the Maxwell equations in the front variables. We now find that there are four equations of motion specifying the τ derivatives of certain field components, and four constraint equations:

$$\partial_\tau E_3=\partial_a V_a\ , \tag{2.10a}$$

$$\partial_\tau B_3=\epsilon_{ab}\partial_a V_b\ , \tag{2.10b}$$

$$\partial_\tau U_a=-\tfrac{1}{2}(\partial_a E_3+\epsilon_{ab}\partial_b B_3)\ , \tag{2.10c}$$

$$\partial_\sigma E_3=-\partial_a U_a\ , \tag{2.10d}$$

$$\partial_\sigma B_3=\epsilon_{ab}\partial_a U_b\ , \tag{2.10e}$$

$$\partial_\sigma V_a=\tfrac{1}{2}(\partial_a E_3-\epsilon_{ab}\partial_b B_3)\ . \tag{2.10f}$$

As for the maintenance of the constraints in τ, we find that (2.10f) together with the equations of motion (2.10a)–(2.10c) ensures this for (2.10d) and (2.10e); while (2.10f) itself is maintained because of the wave equation for V_a.

Let us now compute the generators of the Poincaré group suitable for action on the field tensor $F_{\mu\nu}$. Under the infinitesimal Poincaré transformation

$$x^\mu\to x'^\mu=x^\mu+\omega^{\mu\nu}x_\nu+a^\mu\ , \tag{2.11}$$

the geometrical transformation rule for $F_{\mu\nu}$ is

$$F'_{\mu\nu}(x')=F_{\mu\nu}(x)+\omega_\mu{}^\lambda F_{\lambda\nu}(x)-\omega_\nu{}^\lambda F_{\lambda\mu}(x)\ . \tag{2.12}$$

This means that the change in functional form in $F_{\mu\nu}$ is of amount

$$\begin{aligned}\delta F_{\mu\nu}(x)=&-(\omega^{\alpha\beta}x_\beta+a^\alpha)\partial_\alpha F_{\mu\nu}(x)+\omega_\mu{}^\lambda F_{\lambda\nu}(x)\\&-\omega_\nu{}^\lambda F_{\lambda\mu}(x)\ .\end{aligned} \tag{2.13}$$

In I we set up a generator G to accompany the transformation (2.11) (Ref. 8):

$$\begin{aligned}G&=\tfrac{1}{2}\omega^{\mu\nu}M_{\mu\nu}-a^\mu P_\mu\\&=\omega^{0j}K_j+\omega_{23}J_1+\omega_{31}J_2+\omega_{12}J_3-a^\mu P_\mu\ .\end{aligned} \tag{2.14}$$

The basic generators $M_{\mu\nu}, P_\mu$ must now be determined in such a way that on applying iG to $F_{\mu\nu}$ we get just the change $\delta F_{\mu\nu}$ of Eq. (2.13). To be specific, let us arrange the components $\vec{E},\vec{B}$ of $F_{\mu\nu}$ into a six-component column vector $\binom{E_j}{B_j}$. Then each of the generators $M_{\mu\nu}, P_\mu$ is simultaneously a differential operator on space-time variables and a six-dimensional matrix. Wherever the unit matrix in six dimensions is to appear we do not explicitly indicate it but let it be understood. Then the various generators are

$$
\begin{aligned}
P_\mu &= -i\partial_\mu ,\\
\vec{J} &= -i\vec{x}\times\vec{\nabla} + \begin{bmatrix} \vec{S} & 0 \\ 0 & \vec{S} \end{bmatrix} ,\\
\vec{K} &= i(x^0\vec{\nabla} + \vec{x}\partial_0) + \begin{bmatrix} 0 & \vec{S} \\ -\vec{S} & 0 \end{bmatrix} .
\end{aligned}
\tag{2.15}
$$

Here $\vec{S}$ denotes the triplet of Hermitian three-dimensional spin-1 matrices taken in the Cartesian form, i.e.,

$$(S_j)_{kl} = -i\epsilon_{jkl} . \tag{2.16}$$

It is a straightforward matter to rewrite these generators $M_{\mu\nu}, P_\mu$ suitable for application to a column vector in which one lists the components of $F_{\mu\nu}$ in the sequence, say, U_a, E_3, B_3, V_a, but we shall leave them in the above form.

For determining the effect of a thin lens on an incident paraxial wave we recall from I that the particular generators $G_a = M_{\sigma a}$ are needed. They can be easily found from Eq. (2.15) and are

$$
\begin{aligned}
G_a &= Mx_a - \tau P_a + G_a^{(\mathrm{spin})} ,\\
G_1^{(\mathrm{spin})} &= \frac{1}{2}\begin{bmatrix} -S_2 & S_1 \\ -S_1 & -S_2 \end{bmatrix} ,\\
G_2^{(\mathrm{spin})} &= \frac{1}{2}\begin{bmatrix} S_1 & S_2 \\ -S_2 & S_1 \end{bmatrix} ,\\
M &= \tfrac{1}{2}(P^0 + P^3) = i\partial_\sigma .
\end{aligned}
\tag{2.17}
$$

We shall use these generators and explore their properties in the next section.

III. PARAXIAL SOLUTIONS AND THE LENS TRANSFORMATION

We saw in I for the scalar case that since we are dealing with waves traveling precisely with the speed of light, there is a class of solutions of the wave equation that cannot be described in the format of an initial value problem with respect to the front label τ. These are antiaxial waves depending on τ alone but not on σ and x_a. For the Maxwell system (2.10), these solutions are

$$U_a = E_3 = B_3 = 0 , \quad V_a(\sigma; x_\perp; \tau) = f_a(\tau) , \tag{3.1}$$

with the two functions $f_a(\tau)$ arbitrary. In wave-number space such solutions involve only wave vectors for which $k^0 = -k^3$, $k_a = 0$, and on these the mass operator M of the two-dimensional Galilei group vanishes identically. Out of the seven generators of the Poincaré group associated with transformations leaving the front invariant, G_a, P_a, and M annihilate each solution of the type (3.1), while J_3 and $K_3 - \tau(P^0 - P^3)$ rotate and scale it, respectively. As in I, we will here exclude such solutions from consideration.

We turn now to setting up a convenient description of paraxial solutions of Eqs. (2.10), corresponding to waves propagating roughly along the positive x^3 axis. We are interested only in analytic signal solutions, and the terms quasihenochromatic, henochromatic, paraxial, etc., will be used in the same sense as in I. An exactly axial wave has vanishing E_3, B_3, V_a while each U_a is an arbitrary function of σ; so this is a τ-independent solution. One therefore expects that for the paraxial case the important components of the field are U_a, while E_3, B_3, and V_a will be smaller in comparison. A quasihenochromatic paraxial wave will involve a mean value $\mathscr{M}_0$ and a spread $\Delta\mathscr{M}$ in values of M, and a range of values $|k_\perp| \lesssim \Delta k$ of the transverse wave vector, which will obey as in I the conditions

$$\Delta\mathscr{M} \ll \mathscr{M}_0 , \quad \Delta k \ll \mathscr{M}_0 . \tag{3.2}$$

In the region of space-time where σ and τ obey[3]

$$
\begin{aligned}
|\sigma| &\ll \frac{2\pi}{\mathscr{M}_0}\left[\frac{\mathscr{M}_0}{\Delta\mathscr{M}}\right] ,\\
|\tau| &\ll \frac{4\pi}{\mathscr{M}_0}\left[\frac{\mathscr{M}_0}{\Delta\mathscr{M}}\right]\left[\frac{\mathscr{M}_0}{\Delta k}\right]^2 ,
\end{aligned}
\tag{3.3}
$$

all components of the field will be henochromatic to good approximation. For U_a in this region we assume the forms

$$U_a(\sigma; x_\perp; \tau) \approx e^{-i\mathscr{M}_0\sigma} \int d^2k_\perp f_a(k_\perp) \exp[i(k_\perp\cdot x_\perp - k_\perp^2\tau/2\mathscr{M}_0)] , \tag{3.4}$$

where the two functions $f_a(k_\perp)$ are nonzero only for $|k_\perp| \lesssim \Delta k$. From the constraint equations (2.10d)–(2.10f) we are able to determine all the other field components in the region (3.3) directly in terms of U_a:

$$
\begin{aligned}
E_3(\sigma; x_\perp; \tau) &\approx \frac{e^{-i\mathscr{M}_0\sigma}}{\mathscr{M}_0} \int d^2k_\perp k_a f_a(k_\perp) \exp\left[i\left[k_\perp\cdot x_\perp - \frac{k_\perp^2\tau}{2\mathscr{M}_0}\right]\right] ,\\
B_3(\sigma; x_\perp; \tau) &\approx -\frac{e^{-i\mathscr{M}_0\sigma}}{\mathscr{M}_0} \int d^2k_\perp \epsilon_{ab} k_a f_b(k_\perp) \exp\left[i\left[k_\perp\cdot x_\perp - \frac{k_\perp^2\tau}{2\mathscr{M}_0}\right]\right] ,\\
V_a(\sigma; x_\perp; \tau) &\approx -\frac{e^{-i\mathscr{M}_0\sigma}}{2\mathscr{M}_0^2} \int d^2k_\perp [(k_1^2 - k_2^2)\rho_3 + 2k_1k_2\rho_1]_{ab} f_b(k_\perp) \exp\left[i\left[k_\perp\cdot x_\perp - \frac{k_\perp^2\tau}{2\mathscr{M}_0}\right]\right] .
\end{aligned}
\tag{3.5}
$$

To simplify the expressions for V_a we have used a pair of Pauli matrices ρ_1 and ρ_3. One can now check that the remaining Maxwell equations, the true equations of motion (2.10a)–(2.10c) in the front form, are all obeyed by the expressions (3.4) and (3.5), with no conditions arising on $f_a(k_\perp)$.

The paraxial nature of the wave is determined by the ratio $\Delta k/\mathcal{M}_0$, which is assumed to be a very small number. Viewing this as a controlling parameter we see that in the henochromatic case the field components E_3,B_3 are 1 order of magnitude smaller than U_a, while V_a are 2 orders of magnitude smaller:

$$E_3,B_3 \sim \frac{\Delta k}{\mathcal{M}_0} U_a \ , \quad V_a \sim \left[\frac{\Delta k}{\mathcal{M}_0}\right]^2 U_a \ . \tag{3.6}$$

In addition each individual field component depends most strongly on σ, relatively weakly by the factor $\Delta k/\mathcal{M}_0$ on $x_\perp$, and even more weakly by the factor $(\Delta k/\mathcal{M}_0)^2$ on τ. From (3.6) we can say that in the paraxial region the equality

$$E_a \approx \epsilon_{ab} B_b \approx -U_a \tag{3.7}$$

is good up to and including first-order terms in $\Delta k/\mathcal{M}_0$. The evolution of each component of the field in τ is given of course by a Schrödinger-type equation:

$$i\frac{\partial}{\partial\tau}(U_b,E_3,B_3,V_b) = \frac{P_a P_a}{2\mathcal{M}_0}(U_b,E_3,B_3,V_b) \ . \tag{3.8}$$

We shall now calculate, to lowest nontrivial order, the effect of a thin circular lens of focal length f on a quasihenochromatic paraxial electromagnetic wave incident on it from the left. The lens will be assumed to be placed centrally on and normally to the axis. This calculation will be based on the idea that the relevance of the metaplectic group for such problems, disclosed by the work of Bacry and Cadilhac for scalar waves,[3] must be maintained in a natural way for vector waves. The way to achieve this has been recognized already in I. After writing the generators of the Poincaré group for the Maxwell case in the form suited to the front variables, one isolates the (2 + 1)-dimensional Galilean subalgebra. This subalgebra supplies us with canonical conjugates G_a/M to the transverse "momenta" P_a. We now use these in place of the transverse position coordinates x_a in the lens transformation law for scalar waves described in I. We need to assume as in I that the wave is such that the parameters $\mathcal{M}_0, \Delta\mathcal{M}, \Delta k$ obey

$$\frac{\Delta k}{\mathcal{M}_0} \lesssim 2\left[\frac{\Delta\mathcal{M}}{\mathcal{M}_0}\right]^{1/2} . \tag{3.9}$$

This will permit us to say, when Eqs. (3.3) hold, that a lens located in ordinary space at, say, $x^3=0$, can be thought of to good approximation as being "located" at $\tau=0$ in the front language.[9]

For a scalar paraxial wave incident on the lens, the effect of the lens is to introduce a phase factor[7]

$$e^{i\varphi(x_\perp)} \ , \quad \varphi(x_\perp) \approx \mathcal{M}_0(n\Delta_0 - x_\perp^2/2f) \ . \tag{3.10}$$

Here, n and Δ_0 are the refractive index and lens thickness, respectively, and it is understood as part of the paraxial approximation that this expression for φ must be used only for $|x_\perp| \ll f$. From Eq. (2.17) on setting $\tau=0$ we get the analog of x_a for the Maxwell case:

$$\begin{aligned} x_a \to \frac{G_a(0)}{M} &= x_a + \frac{1}{M} G_a^{(\text{spin})} \\ &\approx x_a + \frac{1}{\mathcal{M}_0} G_a^{(\text{spin})} \ . \end{aligned} \tag{3.11}$$

We are therefore led to suggest that when a paraxial Maxwell wave encounters the lens, the effect will be described by a lens phase transformation matrix

$$\Omega(x_\perp) = \exp\left[-\frac{i\mathcal{M}_0}{2f}\left[x_\perp + \frac{1}{\mathcal{M}_0} G_\perp^{(\text{spin})}\right]^2\right] , \tag{3.12}$$

where a constant inessential phase has been omitted. We now evaluate this matrix: The algebraic properties of $G_a^{(\text{spin})}$ allow us to do so exactly.

Let us introduce a set of auxiliary Pauli matrices ρ_1,ρ_2,ρ_3 such that

$$\rho_1^2 = \rho_2^2 = \rho_3^2 = 1 \ , \quad \rho_1\rho_2 = -\rho_2\rho_1 \ ,$$

etc. We may then write the six-dimensional matrices $G_a^{(\text{spin})}$ as Kronecker products of ρ's and S's in this way:

$$\begin{aligned} G_1^{(\text{spin})} &= \tfrac{1}{2}(i\rho_2 S_1 - S_2) \ , \\ G_2^{(\text{spin})} &= \tfrac{1}{2}(S_1 + i\rho_2 S_2) \ . \end{aligned} \tag{3.13}$$

From here we see that

$$G_2^{(\text{spin})} = -i\rho_2 G_1^{(\text{spin})} \ . \tag{3.14}$$

It is obvious that $G_1^{(\text{spin})}$ and $G_2^{(\text{spin})}$ commute: This is consistent with their being parts of the commuting Galilean boost generators. Moreover, since S_j are the generators of the spin-1 representation of the rotation group, S_1 and S_2 obey the relations[10]

$$\begin{aligned} &S_1^3 = S_1 \ , \quad S_2^3 = S_2 \ , \\ &S_1^2 S_2 + S_1 S_2 S_1 + S_2 S_1^2 = S_2 \ , \\ &S_1 S_2^2 + S_2 S_1 S_2 + S_2^2 S_1 = S_1 \ . \end{aligned} \tag{3.15}$$

Using (3.14) and (3.15) we have the results

$$\begin{aligned} &G_a^{(\text{spin})} G_a^{(\text{spin})} = 0 \ , \\ &G_a^{(\text{spin})} G_b^{(\text{spin})} G_c^{(\text{spin})} = 0 \ . \end{aligned} \tag{3.16}$$

As a consequence, we get a closed-form expression for $\Omega(x_\perp)$; reinstating the constant phase it is

$$\Omega(x_\perp)=e^{i\varphi(x_\perp)}\exp\left[-\frac{i}{f}x_a G_a^{(\mathrm{spin})}\right]=e^{i\varphi(x_\perp)}\left[1-\frac{i}{f}x_a G_a^{(\mathrm{spin})}-\frac{1}{2f^2}(x_a G_a^{(\mathrm{spin})})^2\right]$$

$$=e^{i\varphi(x_\perp)}\begin{bmatrix} 1+\frac{(y^2-x^2)}{8f^2} & \frac{-xy}{4f^2} & \frac{-x}{2f} & \frac{xy}{4f^2} & \frac{(y^2-x^2)}{8f^2} & \frac{y}{2f} \\ \frac{-xy}{4f^2} & 1+\frac{(x^2-y^2)}{8f^2} & \frac{-y}{2f} & \frac{(y^2-x^2)}{8f^2} & \frac{-xy}{4f^2} & \frac{-x}{2f} \\ \frac{x}{2f} & \frac{y}{2f} & 1 & \frac{-y}{2f} & \frac{x}{2f} & 0 \\ \frac{-xy}{4f^2} & \frac{(x^2-y^2)}{8f^2} & \frac{-y}{2f} & 1+\frac{(y^2-x^2)}{8f^2} & \frac{-xy}{4f^2} & \frac{-x}{2f} \\ \frac{(x^2-y^2)}{8f^2} & \frac{xy}{4f^2} & \frac{x}{2f} & \frac{-xy}{4f^2} & 1+\frac{(x^2-y^2)}{8f^2} & \frac{-y}{2f} \\ \frac{y}{2f} & \frac{-x}{2f} & 0 & \frac{x}{2f} & \frac{y}{2f} & 1 \end{bmatrix} . \qquad (3.17)$$

This matrix applied to the column vector made up of the functions E_j,B_j for the incident wave is expected to yield the column vector of the wave after passage through the lens.

Since we are only interested in exhibiting how the principle expressed by Eq. (3.11) works, we shall retain only the lowest-order term x_a/f of the paraxial aproximation and neglect the quadratic terms in x and y in $\Omega(x)$, apart from the piece $e^{i\varphi}$. We also remember from Eq. (3.7) that we can set $E_1=B_2$, $E_2=-B_1$ to leading order, and treat E_3 and B_3 as being small quantities of first order relative to E_a,B_a. Then after dropping a factor $e^{-i\mathcal{M}_0\sigma}$ common to all the field components, we find that the outgoing E',B' are related to the incident E,B by

$$\begin{aligned} E_a'(x_\perp;0)&\approx e^{i\varphi(x_\perp)}E_a(x_\perp;0)\ ,\\ B_a'(x_\perp;0)&\approx e^{i\varphi(x_\perp)}B_a(x_\perp;0)\ ,\\ E_3'(x_\perp;0)&\approx e^{i\varphi(x_\perp)}[E_3(x_\perp;0)+(x_a/f)E_a(x_\perp;0)]\ ,\\ B_3'(x_\perp;0)&\approx e^{i\varphi(x_\perp)}[B_3(x_\perp;0)+(x_a/f)B_a(x_\perp;0)]\ . \end{aligned} \qquad (3.18)$$

These formulas show that for a paraxial incident wave, the passage through a lens contributes a small additional axial component which agrees with what one obtains from a more direct calculation based on the condition that the Maxwell equations (2.1c) and (2.1d) must be maintained both before and after the action of the lens.[11] The natural geometrical interpretation of this result will become clear in Sec. V.

To assure ourselves of the correctness of Eqs. (3.18), it is interesting to consider the following simple situation. Let the incident wave be a strictly axial plane wave, so that to the immediate left of the lens we have $E_1=B_2$, $E_2=-B_1$, $E_3=B_3=0$. Then on passage through the lens the wave picks up small nonzero axial components E_3',B_3'. These show up vividly in the Poynting vector of the outgoing wave, the components of which are found to be

$$\mathrm{Re}\{\vec{E}'(x_\perp;0)\times[\vec{B}'(x_\perp;0)]^*\} \approx \tfrac{1}{2}(|\vec{E}|^2+|\vec{B}|^2)(-x_\perp/f,1)\ . \qquad (3.19)$$

For a converging lens with $f>0$ we see that at each $x_\perp$ to the immediate right of the lens (assuming $|x_\perp|\ll f$) the outgoing Poynting vector points exactly to the focus $(0,0,f)$, which is just what is expected. [If the terms x_aE_a,x_aB_a in E_3',B_3' in Eqs. (3.18) had been absent, it is clear that the outgoing Poynting vector, like the incoming one, would have been parallel to the system axis at each $x_\perp$.] At the level of the field components we see for example that the electric vector $\vec{E}'(x_\perp;0)$ is orthogonal to $(-x_\perp,f)$ which is the vector leading from $(x_\perp,0)$ to the focal point $(0,0,f)$, and similarly for $\vec{B}'(x_\perp;0)$. Thus for an incident axial plane wave the lens transformation (3.18) yields an outgoing wave which locally can be described as a set of vector plane waves, all directed to the focal point. This justifies the extension to vector waves of the well-known Debye integral representation for focused fields.[12]

In principle similar calculations can be carried out for other situations of interest. The identification of the action of various kinds of lenses (in the quadratic phase approximation) with corresponding elements of the group SL(2,R) when one has axial symmetry [and more generally the group Sp(4,R)], and even of free propagation according to Eq. (3.8) with an element of this group, goes through *with no changes at all* compared to the scalar case.

IV. RADIATION GAUGE IN THE FRONT FORM

The description of paraxial solutions of the Maxwell equations and their passage through thin lenses given in the preceding section is physically transparent since the behavior of all six field components $F_{\mu\nu}$ was specified. However, it is somewhat unwieldy in that one has to contend with six-dimensional matrices, so one may try to give a more economical treatment using a vector potential. For the class of problems one is interested in here, namely, paraxial beams passing through optical systems, one must choose the gauge judiciously, so that the particular generators G_a in the Poincaré algebra act in a simple way. We show in this section how this is to be done.

To begin with, we recall the Poincaré transformation properties of the radiation gauge vector potential suited to the instant form. For free fields, this potential is defined by the conditions

$$A_0(x)=0, \quad \vec{\nabla}\cdot\vec{A}(x)=0 \tag{4.1}$$

so that the usual relations $F_{\mu\nu}=\partial_\mu A_\nu-\partial_\nu A_\mu$ are in this case

$$\vec{E}=-\partial_0\vec{A}, \quad B=\vec{\nabla}\times\vec{A}\ . \tag{4.2}$$

It is well known that this vector potential transforms simply under space-time translations and spatial rotations, but has a nonlocal behavior under pure Lorentz transformations. In fact, the various generators of the Poincaré group act in this gauge in the following ways[13]:

$$\begin{aligned}
P_\mu\vec{A}(x)&=-i\partial_\mu\vec{A}(x)\ ,\\
(J_j\vec{A})_k(x)&=-i(\vec{x}\times\vec{\nabla})_jA_k(x)+(S_j)_{kl}A_l(x)\\
&=-i[(\vec{x}\times\vec{\nabla})_j\delta_{kl}+\epsilon_{jkl}]A_l(x)\ ,\\
(K_j\vec{A})_k(x)&=i(x^0\partial_j+x_j\partial_0)A_k(x)\\
&\quad+\frac{i}{4\pi}\partial_k\int d^3x'\frac{1}{|\vec{x}-\vec{x}'|}\partial_0A_j(\vec{x}',x^0)\ .
\end{aligned} \tag{4.3}$$

It follows that the nonlocality of the action of K_j will be present in the combinations G_a as well, so this gauge is unsuitable for the present situation. We need a gauge which is natural from the front point of view.

We could try setting one of the two components A_τ or A_σ equal to zero. From the Lagrangian point of view it is natural to make the choice $A_\tau=0$,[6] but it turns out that this still involves a nonlocal action of G_a. (This will become clear later.) However, this problem does not arise for the choice $A_\sigma=0$.[14] Therefore, for free fields we shall *define* the vector potential in the *front form radiation gauge* by the conditions

$$\begin{aligned}
&A_\sigma(x)=0,\quad A_0(x)=A_3(x)=\tfrac{1}{2}A_\tau(x)\ ,\\
&\partial^\mu A_\mu(x)=0\rightarrow\partial_aA_a(x)=\partial_\sigma A_\tau(x)\ .
\end{aligned} \tag{4.4}$$

Then the various components of $F_{\mu\nu}$ are given by

$$\begin{aligned}
U_a&=\partial_\sigma A_a\ ,\\
E_3&=-\partial_\sigma A_\tau=-\partial_aA_a\ ,\\
B_3&=\epsilon_{ab}\partial_aA_b\ ,\\
V_a&=\partial_\tau A_a-\partial_aA_\tau\ .
\end{aligned} \tag{4.5}$$

One easily checks that the Maxwell equations in the front form, (2.10), behave in the following way when these expressions are put in: The equation of motion (2.10b) and the constraints (2.10d), (2.10e) are identically obeyed; the equation of motion (2.10a) leads to the wave equation for A_τ; the equations of motion (2.10c) as well as the constraints (2.10f) lead to the wave equations for A_a.

The forms of the generators of the Poincaré group, suitable for application to this vector potential, can be obtained in a straightforward way. For those infinitesimal transformations which preserve the conditions (4.4) when we naively tranform A_μ as though it were a four-vector field, there is no difficulty at all. In other cases, we find that conditions (4.4) can be restored by a suitable gauge transformation after we have first transformed A_μ as a four-vector. In this way we get for the generators acting on the column vector

$$\begin{bmatrix}A_1\\A_2\\A_\tau\end{bmatrix}$$

the following expressions:

$$P_a=-i\partial_a,\quad M=\tfrac{1}{2}(P^0+P^3)=i\partial_\sigma,\quad H=P^0-P^3=i\partial_\tau, \tag{4.6a}$$

$$J_3=-i\epsilon_{ab}x_a\partial_b+S_3\ , \tag{4.6b}$$

$$G_a=Mx_a-\tau P_a+\widetilde{G}_a^{(\mathrm{spin})},\quad \widetilde{G}_1^{(\mathrm{spin})}=\begin{bmatrix}0&0&0\\0&0&0\\i&0&0\end{bmatrix},\quad \widetilde{G}_2^{(\mathrm{spin})}=\begin{bmatrix}0&0&0\\0&0&0\\0&i&0\end{bmatrix}, \tag{4.6c}$$

$$K_3=\tau H-\sigma M+\begin{bmatrix}0&0&0\\0&0&0\\0&0&i\end{bmatrix}, \tag{4.6d}$$

$$\begin{aligned}
(K_1+J_2)\begin{bmatrix}A_1(\sigma;x_\perp;\tau)\\A_2(\sigma;x_\perp;\tau)\\A_\tau(\sigma;x_\perp;\tau)\end{bmatrix}&=\left[Hx_1-\sigma P_1+\begin{bmatrix}0&0&i\\0&0&0\\0&0&0\end{bmatrix}\right]\begin{bmatrix}A_1\\A_2\\A_\tau\end{bmatrix}-\frac{i}{\pi}\begin{bmatrix}\partial_1\\\partial_2\\\partial_\tau\end{bmatrix}\int d^2x'_\perp\ln|x_\perp-x'_\perp|\,\partial_\tau A_1(\sigma;x'_\perp;\tau)\ ,\\
(K_2-J_1)\begin{bmatrix}A_1(\sigma;x_\perp;\tau)\\A_2(\sigma;x_\perp;\tau)\\A_\tau(\sigma;x_\perp;\tau)\end{bmatrix}&=\left[Hx_2-\sigma P_2+\begin{bmatrix}0&0&0\\0&0&i\\0&0&0\end{bmatrix}\right]\begin{bmatrix}A_1\\A_2\\A_\tau\end{bmatrix}-\frac{i}{\pi}\begin{bmatrix}\partial_1\\\partial_2\\\partial_\tau\end{bmatrix}\int d^2x'_\perp\ln|x_\perp-x'_\perp|\,\partial_\tau A_2(\sigma;x'_\perp;\tau)\ .
\end{aligned} \tag{4.6e}$$

In comparison with the situation described by Eq. (4.3) and pertaining to the instant form, we notice two interesting points: (i) out of the ten generators only two have a nonlocal action in the front form radiation gauge; (ii) this nonlocality is with respect to two of the "spatial" variables x_a in the front alone, and not with respect to σ. From (4.6c) and (4.6e) it is also clear that if we had chosen the gauge $A_\tau=0$ rather than $A_\sigma=0$, which amounts to interchanging the roles of σ and τ, G_a would have had a nonlocal action but K_1+J_2 and K_2-J_1 would become simple.

We want now to set up a vector potential that will yield the quasihenochromatic paraxial solution [(3.4) and (3.5)] of the Maxwell equations, and are interested only in the region (3.3) where it is effectively henochromatic. This potential can be easily guessed from Eqs. (4.5) and it is

$$\begin{aligned} A_a(\sigma;x_\perp;\tau) &\approx \frac{i}{\mathcal{M}_0}U_a(\sigma;x_\perp;\tau)\ , \\ A_\tau(\sigma;x_\perp;\tau) &\approx \frac{-1}{\mathcal{M}_0^2}\partial_a U_a(\sigma;x_\perp;\tau)\ . \end{aligned} \tag{4.7}$$

One can check that the gauge condition (4.4) as well as all of Eqs. (4.5) are obeyed. We see that A_τ is smaller than A_a by 1 order of magnitude in the small quantity $\Delta k/\mathcal{M}_0$:

$$A_\tau \sim \frac{\Delta k}{\mathcal{M}_0}A_a\ , \tag{4.8}$$

which must be compared to the relationships (3.6) among the field components. From a physical point of view, because of Eq. (3.7) we can say that up to and including first-order terms in $\Delta k/\mathcal{M}_0$ the vector potential coincides with the electric field:

$$A_a \approx \frac{-i}{\mathcal{M}_0}E_a,\ \ A_\tau \approx \frac{-i}{\mathcal{M}_0}E_3\ . \tag{4.9}$$

Of course, it obeys the free propagation equation (3.8) just like the field components.

In this formalism the lens transformation matrix is three dimensional and its computation is algebraically even simpler than in the formalism of the preceding section working with field components. This is because the matrix terms in G_a of Eq. (4.6c) obey

$$\begin{aligned} (\widetilde{G}_1^{(\mathrm{spin})})^2 &= \widetilde{G}_1^{(\mathrm{spin})}\widetilde{G}_2^{(\mathrm{spin})} = \widetilde{G}_2^{(\mathrm{spin})}\widetilde{G}_1^{(\mathrm{spin})} \\ &= (\widetilde{G}_2^{(\mathrm{spin})})^2 = 0\ . \end{aligned} \tag{4.10}$$

We therefore expect that in the paraxial approximation to leading order the effect of a thin circular lens of focal length f on the vector potential of an incident henochromatic paraxial wave is given by the matrix

$$\begin{aligned} \widetilde{\Omega}(x_\perp) &= \exp\left[i\mathcal{M}_0 n\Delta_0 - \frac{i\mathcal{M}_0}{2f}\left[x_\perp + \frac{1}{\mathcal{M}_0}(\widetilde{G}_\perp^{(\mathrm{spin})})^2 \right] \right] \\ &= \exp\{ i[\varphi(x_\perp) - (x_a/f)\widetilde{G}_a^{(\mathrm{spin})}] \} \\ &= e^{i\varphi(x_\perp)} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ x/f & y/f & 1 \end{bmatrix}\ . \end{aligned} \tag{4.11}$$

Therefore, the outgoing vector potential is related to the incident one by

$$A_a' = e^{i\varphi}A_a,\ \ A_\tau' = e^{i\varphi}(A_\tau + x_a A_a/f)\ . \tag{4.12}$$

It is easy to verify that the gauge condition (4.4) is maintained by this transformation, and when we take account of (4.9) we see that we have recovered Eqs. (3.18) for the components of the electric field. To leading order one can also confirm that the changes in B_a, B_3 contained in (3.18) are properly reproduced; making use of Eqs. (3.7) and (4.9), we have for B_3:

$$\begin{aligned} B_3' &= \epsilon_{ab}\partial_a A_b' \simeq \epsilon_{ab}\partial_a(e^{i\varphi}A_b) \\ &= e^{i\varphi}\epsilon_{ab}[\partial_a A_b - i(\mathcal{M}_0/f)x_a A_b] \\ &\simeq e^{i\varphi}[B_3 + (x_a/f)B_a]\ . \end{aligned} \tag{4.13}$$

Thus, to leading order, the six-dimensional lens transformation matrix of Eq. (3.17) and the simpler three-dimensional one of Eq. (4.11) give mutually consistent results.

V. PENCILS OF ELECTROMAGNETIC RAYS IN PARAXIAL OPTICS

The second-order coherence properties of a general statistical state of the Maxwell field are described by suitable correlation tensors[15] that generalize the two-point correlation function of the scalar treatment. These tensors could be defined in terms of the components of the electric and magnetic field vectors or, more simply, in terms of the vector potential. We shall work with the latter and shall briefly indicate how the concept of generalized pencils of rays[5] is set up for vector waves in the paraxial limit.

Let the subscripts $\alpha,\beta,\ldots,$ run over the values $1,2,\tau$. Consider an ensemble of quasihenochromatic paraxial waves with characteristic parameters $\mathcal{M}_0$, $\Delta\mathcal{M}$, Δk. In the space-time region (3.3) where it is effectively henochromatic, the representative vector potential in the front form radiation gauge can be written as

$$A_\alpha(\sigma;x_\perp;\tau) \approx e^{-i\mathcal{M}_0\sigma}\mathcal{A}_\alpha(x_\perp;\tau)\ , \tag{5.1}$$

and we remember that $\mathcal{A}_\tau$ is smaller than $\mathcal{A}_a$ by a factor $\Delta k/\mathcal{M}_0$. For any two points on the same front we define the correlation tensor Γ as

$$\begin{aligned} \Gamma_{\alpha\beta}(\sigma_1,x_{1_\perp};\sigma_2,x_{2_\perp};\tau) &= \langle [A_\alpha(\sigma_1;x_{1_\perp};\tau)]^* A_\beta(\sigma_2;x_{2_\perp};\tau)\rangle \\ &\approx e^{i\mathcal{M}_0(\sigma_1-\sigma_2)}\Gamma_{\alpha\beta}^{(0)}(x_{1_\perp};x_{2_\perp};\tau)\ , \\ \Gamma_{\alpha\beta}^{(0)}(x_{1_\perp};x_{2_\perp};\tau) &= \langle [\mathcal{A}_\alpha(x_{1_\perp};\tau)]^*\mathcal{A}_\beta(x_{2_\perp};\tau)\rangle\ . \end{aligned} \tag{5.2}$$

The angular brackets denote an average over the ensemble. The leading elements of this matrix are $\Gamma_{ab}^{(0)}$; the elements $\Gamma_{a\tau}^{(0)}$ and $\Gamma_{\tau b}^{(0)}$ are smaller by a factor $\Delta k/\mathcal{M}_0$; while the last element $\Gamma_{\tau\tau}^{(0)}$, being smaller by $(\Delta k/\mathcal{M}_0)^2$, will be neglected. The Wolf matrix of generalized rays of light is now defined as a Wigner-Moyal tranform of $\Gamma^{(0)}$:

$$W_{\alpha\beta}(x_\perp;p_\perp;\tau)=(2\pi)^{-2}\int d^2\xi_\perp e^{ip_\perp\cdot\xi_\perp}\Gamma^{(0)}_{\alpha\beta}(x_\perp+\tfrac{1}{2}\xi_\perp;x_\perp-\tfrac{1}{2}\xi_\perp;\tau)\ . \tag{5.3}$$

This is a Hermitian, but not pointwise positive definite, matrix. Since the paraxial condition is assumed to hold over the entire ensemble, it is easy to see that $W_{\alpha\beta}(x_\perp;p_\perp;\tau)$ is nonzero only for $p_\perp$ in the range $|p_\perp|\lesssim\Delta k\ll\mathscr{M}_0$. Thus we can interpret $W_{ab}(x_\perp;p_\perp;\tau)$ as representing in matrix form the intensity of transversely polarized generalized rays of light at the point $x_\perp$ in the front τ, traveling in the transverse direction $p_\perp$.[16] As in the scalar case this intensity is guaranteed to be real but may not be positive. The elements $W_{a\tau}(x_\perp;p_\perp;\tau)$ describe the correlation, at the point $x_\perp$, between the transverse and the longitudinal rays having a common direction $p_\perp$. $W_{\tau\tau}(x_\perp;p_\perp;\tau)$ must be interpreted as the intensity of generalized longitudinal rays, but since it is of order $(\Delta k/\mathscr{M}_0)^2$ we neglect it. In interpreting the matrix W in this way, we of course make use of Eqs. (4.9).

The gauge condition (4.4) on the vector potential leads to a condition on the matrix W which can be expressed as a determination of $W_{a\tau}$ in terms of W_{ab}:

$$W_{a\tau}(x_\perp;p_\perp;\tau)\approx\frac{i}{\mathscr{M}_0}\left[\frac{1}{2}\frac{\partial}{\partial x_b}+ip_b\right]W_{ab}(x_\perp;p_\perp;\tau)\ . \tag{5.4}$$

This equation lends itself to the following interpretation. For a uniform beam for which we can neglect the variation with respect to $x_\perp$, except at the very edges, Eq. (5.4) states that the polarization is perpendicular to the paraxial ray direction given by $p^\tau\approx\mathscr{M}_0,p_\perp$. Near the edges of the beam the polarization is no longer strictly transverse, a result already known from previous work.[5] The free propagation law for W is as simple as in the scalar case, because Eq. (3.8) applies to the vector potential also:

$$W_{\alpha\beta}(x_\perp;p_\perp;\tau)\approx W_{\alpha\beta}(x_\perp-\tau p_\perp/\mathscr{M}_0;p_\perp;0)\ . \tag{5.5}$$

This means that in free space generalized rays travel in straight lines with unchanging polarization properties, in the paraxial approximation.

Finally we see how the matrix W is changed by a thin circular lens placed centrally and normally on the axis at $x^3=0$. The effect on the "reduced" vector potential $\mathscr{A}_\alpha$ is given by Eqs. (4.11) and (4.12) and is

$$\mathscr{A}'_\alpha(x_\perp;0)\approx e^{i\varphi(x_\perp)}T_{\alpha\beta}(x_\perp)\mathscr{A}_\beta(x_\perp;0)\ ,$$

$$T_{\alpha\beta}(x_\perp)=\delta_{\alpha\beta}+\delta_{\alpha\tau}\delta_{\beta b}x_b/f\ . \tag{5.6}$$

It follows that the change produced in the correlation tensor $\Gamma^{(0)}$ is, in matrix form,

$$\Gamma^{(0)\prime}(x_{1_\perp};x_{2_\perp};0)=e^{i[\varphi(x_{2_\perp})-\varphi(x_{1_\perp})]}[T(x_{1_\perp})]^*\times\Gamma^{(0)}(x_{1_\perp};x_{2_\perp};0)[T(x_{2_\perp})]^{\sim} \tag{5.7}$$

where the asterisk and the tilde denote complex conjugation and the transpose, respectively. From here we can easily calculate the change in W and find

$$W'_{ab}(x_\perp;p_\perp;0)=W_{ab}\left[x_\perp;p_\perp+\frac{\mathscr{M}_0}{f}x_\perp;0\right]\ , \tag{5.8a}$$

$$W'_{a\tau}(x_\perp;p_\perp;0)=W_{a\tau}\left[x_\perp;p_\perp+\frac{\mathscr{M}_0}{f}x_\perp;0\right]+\frac{1}{f}\left[x_b+\frac{i}{2}\frac{\partial}{\partial p_b}\right]W_{ab}\left[x_\perp;p_\perp+\frac{\mathscr{M}_0}{f}x_\perp;0\right]\ . \tag{5.8b}$$

We see, as in the scalar case described in I, that the generalized rays are bent by the lens in a simple geometrical way. The lens action given above preserves the connection (5.4) found from the gauge condition on the vector potential. Since Eq. (5.4) guarantees that the polarization is perpendicular to the paraxial direction, it follows that the lens action given by Eqs. (5.8) is such as to preserve this condition.

VI. CONCLUDING REMARKS

In the two papers of this series we have developed a general formalism which, we believe, is ideally suited for analysis of all paraxial-wave optical problems. The use of relativistic ideas has led us in a natural way to a classification of the space-time coordinates as well as of field components of the Maxwell field according to their importance in terms of the basic parameters $\Delta k/\mathscr{M}_0$ of a paraxial beam.

The relevance of the metaplectic group—the group of linear transformations on canonical variables preserving their commutation relations—for paraxial scalar wave propagation through optical system was discerned by Bacry and Cadilhac.[3] We have used the following principle in establishing this connection for electromagnetic waves: In place of the transverse coordinate x_a which occurs in the transformation function representing a given optical system, we must substitute an operator $G_a(0)/M$, where $G_a(0)$ taken at $\tau=0$ and M are particular generators of the Poincaré group:

$$x_a\to\frac{G_a(0)}{M}=\begin{cases}x_a+\dfrac{1}{M}G_a^{(\mathrm{spin})}\\[2mm] x_a+\dfrac{1}{M}\tilde{G}_a^{(\mathrm{spin})}\ ,\end{cases} \tag{6.1}$$

for fields $\vec{E},\vec{B}$ and for vector potential A, respectively. Then each ideal optical system gets represented by a transformation matrix. In both cases above, $G_a(0)/M$ and the transverse momenta P_a form two canonically conjugate pairs but they are reducible in the operator sense. This rule has led to correct and consistent results, and it is natural to search for a simple explanation of this fact. This is actually not hard to find. The essential physical point is that—whether we speak in terms of the field strengths $\vec{E},\vec{B}$ or the vector potential A—each transverse component propagates through every ideal optical system as though it were a scalar wave, while the axial components propagate *almost as if* they were also a scalar wave.[17] The departure from such behavior for these latter components is *just enough* to ensure that the constraint equations (2.1c) and (2.1d) on $\vec{E},\vec{B}$ or the gauge condition (4.4) on A are maintained both before and after the encounter with the lens. Now for paraxial situations, if we make use of Eqs. (4.9) and neglect terms of second order in $\Delta k/\mathscr{M}_0$, we find

$$\vec{\nabla}\cdot\vec{E}=\partial_a E_a+(\tfrac{1}{2}\partial_\tau-\partial_\sigma)E_3\approx\partial_a E_a-\partial_\sigma E_3$$
$$\approx i\mathscr{M}_0(\partial_a A_a-\partial_\sigma A_\tau)\ , \qquad (6.2)$$

while $\vec{\nabla}\cdot\vec{B}$ vanishes identically when $\vec{B}$ is expressed in terms of A. So in the front formalism both the gauge condition (4.4) on A and the Maxwell equation $\vec{\nabla}\cdot\vec{E}=0$ are constraint conditions since they do not involve derivatives of any field quantities with respect to τ. Now we pointed out in I that for the scalar wave equation a solution remains a solution if we apply any function $F(M_{\mu\nu},P_\mu)$ of the Poincaré generators to it, provided F has no explict dependence on space-time variables. An analogous statement is true for vector waves. Any function of the six-dimensional matrix generators (2.15) can be applied to a column vector made up of an $\vec{E}$ and a $\vec{B}$ obeying Maxwell's equations, and the result will be another solution; or we can apply a function of the three-dimensoinal matrix generators (4.6) to a vector potential (A_a,A_τ) obeying the gauge condition (4.4) and the wave equation, and the result will also obey both. This is a consequence of the linearity and the Poincaré invariance of the respective equations. While the function F of the generators must not carry any explicit τ dependence, the generators themselves, or at least some of them like $G_a(\tau)$, may carry explicit dependences on τ. But if one seeks only to maintain the constraint relation $\vec{\nabla}\cdot\vec{E}=0$, then one can relax the conditions mentioned above and apply a function of $G_a(0)/M$, rather than $G_a(\tau)/M$, to fields obeying this constraint, and the constraint will be maintained.

The formal similarity of paraxial problems to a nonrelativistic quantum-mechanical "particle" in two dimensions persists in going from the scalar theory to Maxwell's equations.[18] If we use the approach based on the vector potential, we can say that the particle can have "helicity" ± 1 or 0. But while in the scalar theory every thin lens imparted a harmonic impulse which is just a phase change and so a unitary transformation, with the vector potential the transformation matrix, for example in Eq. (4.11), is not unitary. Nevertheless is would be worthwhile computing the transformation matrices for various configurations of interest. It would also be interesting to examine the detailed behavior of generalized pencils of light endowed with polarization, in situations where the optical system has a nontrivial effect on the state of polarization. We hope to come back to these questions and to other applications of our formalism elsewhere.

ACKNOWLEDGMENTS

Two of us (N.M. and E.C.G.S.) thank Professor Jan S. Nilsson for providing gracious hospitality and excellent working conditions during the completion of this work. This work was supported by U. S. Department of Energy No. DE-AS05-76ER03992.

---

*Permanent address: Centre for Theoretical Studies and Department of Physics, Indian Institute of Science, Bangalore 560012, Karnataka, India.

†On leave of absence from the Department of Physics, The American College, Madurai, Tamil Nadu, India.

‡Permanent address: Department of Physics, Center for Particle Theory, University of Texas at Austin, Austin, Texas 78712.

[1]E. C. G. Sudarshan, R. Simon, and N. Mukunda, Phys. Rev. A 28, 2931 (1983), hereafter referred to as I.

[2]P. A. M. Dirac, Rev. Mod. Phys. 21, 392 (1949).

[3]H. Bacry and M. Cadilhac, Phys. Rev. A 23, 2533 (1981); M. Nazarathy and J. Shamir, J. Opt. Soc. Am. 72, 356 (1982).

[4]L. Susskind, Phys. Rev. 165, 1535 (1968); J. B. Kogut and D. E. Soper, Phys. Rev. D 1, 2901 (1970); L. C. Biedenharn and H. van Dam, *ibid.* 9, 471 (1974).

[5]E. C. G. Sudarshan, Phys. Lett. 73A, 269 (1979); Physica 96A, 315 (1979); Phys. Rev. A 23, 2802 (1981).

[6]For a treatment within the canonical formalism, see A. J. Hanson, T. Regge, and C. Teitelboim, *Constrained Hamiltonian Systems* (Accademia Nazionale dei Lincei, Rome, 1976).

[7]J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, New York, 1968); A. Papoulis, *Systems and Tranforms with Applications in Optics* (McGraw-Hill, New York, 1968).

[8]We remind the reader that this generator has only a geometric meaning in the present context and not a dynamical one.

[9]The condition (3.9) ensures that in the region of space-time where a quasihenochromatic paraxial wave can be regarded as being henochromatic, it is to good approximation monochromatic in the ordinary sense as well.

[10]All these relations arise from the general one $(S_1\cos\theta+S_2\sin\theta)^3=S_1\cos\theta+S_2\sin\theta$.

[11]Note that both the longitudinal and the transverse components are in phase. In this context, compare M. Lax, W. H. Louisell, and W. B. Knight, Phys. Rev. A 11, 1365 (1975).

[12]M. Born and E. Wolf, *Principles of Optics* (Pergamon, Oxford, 1964), Sec. 8.8.1.

[13]See, for instance, J. D. Bjorken and S. D. Drell, *Relativsitic Quantum Fields* (McGraw-Hill, New York, 1965), Chap. 14.

[14]This is the choice made in J. B. Kogut and D. E. Soper, Ref. 4 above.

[15]M. Born and E. Wolf, *Principles of Optics* (Pergamon, Oxford, 1964), Chap. X; J. R. Klauder and E. C. G. Sudarshan, *Fun-*

damentals of Quantum Optics (Benjamin, New York, 1968).

[16]To leading order, the complete wave vector is $p^{\tau} \approx \mathscr{M}_0$, $p_a, p^{\sigma} = -p_{\tau} \approx 0$.

[17]The longitudinal components are E_3, B_3, or A_{τ} depending on the formulation one is using.

[18]The wave function in quantum mechanics must be normalized. The Wolf function, on the other hand, is not normalized, but its integral over $x_{\perp}$ and $p_{\perp}$ is the total intensity carried by the beam and can assume any non-negative value.