84 - 41

DOE-ER40200-052

July, 1984

Paraxial Maxwell beams: Transformation by general linear optical systems*

by

N. Mukunda+
Institute of Theoretical Physics, S-412 96 Göteborg, Sweden

R. Simon*
Centre for Theoretical Studies, Indian Institute of Science
Bangalore 560 012, India

E.C.G. Sudarshan++
Institute of Theoretical Physics, S-412 96 Göteborg, Sweden

*Supported in part by the U.S. Department of Energy under Contract DE-AS05 - 76 ERO 3992.

+Permanent address: Centre for Theoretical Studies and Dept. of Physics, Indian Institute of Science, Bangalore 560 012, India.

*Permanent address: Dept. of Physics, The American College, Madurai 625 002, India.

++Permanent address: Dept. of Physics, The University of Texas, Austin, Texas, 78712, USA.

ABSTRACT

Extending the work of earlier papers on the relativistic front description of paraxial optics, we generalize the Jones calculus of axial plane waves to describe the action of the most general linear optical system on paraxial Maxwell fields. Several examples are worked out and in each case it is shown that the formalism leads to physically correct results. The importance of retaining the small components of the field vectors along the axis of the system for a consistent description is emphasized.

1. INTRODUCTION

Several formalisms are available for the description of the polarization states of light fields and their transformation by linear optical systems. The Jones calculus[1] and the Poincaré sphere method[2] give descriptions of fully polarized light while the coherency matrix method[3] and the Chandrasekhar-Mueller-Stokes calculus[4] are capable of handling partially polarized fields, of which unpolarized and polarized fields are opposite extremes. In the conventional formulation the transformation of the coherency matrix by an optical system is through the Jones matrix of the system, but this can only handle systems which do not have any statistical element. The Chandrasekhar-Mueller-Stokes method on the other hand is capable of dealing with systems having statistical features, such as depolarizing systems. Further, it deals only with real measurable quantities. The necessary and sufficient condition on a Mueller matrix for it to be derivable from a Jones matrix is also known.[5]

All these standard methods are based on the assumption that the radiation fields of concern are (superpositions of) strictly axial plane waves propagating along the axis of the system. Thus it is assumed that there is no component of the field vector along the axis of the system. In many realistic situations however the actual fields are paraxial beams rather than purely axial. Even the simplest optical system like a thin lens maps an axial field into a (converging or diverging) paraxial field. Hence a consistent description of polarization states and their transformation by optical systems should deal with paraxial fields.

We have recently set up[6,7] a general formalism, based on the front form of relativistic dynamics[8], for the treatment of paraxial wave propagation problems in optics. The Maxwell field can be represented by a six-component column vector made up of the components of the electric and magnetic field vectors $\underline{E}$, $\underline{B}$. By making judicious use of the Poincaré generators for the Maxwell field in the front form, we developed a method by which linear optical systems can be represented as 6 x 6 matrix operators acting on the column vector representing paraxial fields. The particular systems dealt with were all those definable within the framework of scalar paraxial optics, and our analysis based on the Poincaré group gave an unambiguous prescription to determine their representation and action on the Maxwell field. A slightly simpler description uses the vector potential, chosen in a special gauge suited to the front form. Since such a potential has only three independent components, paraxial fields and optical systems get represented by three-component column vectors and 3 x 3 matrices respectively. These two methods are of course mutually consistent. The way they lead to an unambiguous generalization of conventional scalar Fourier optics to the complete Maxwell fields has been elaborated elsewhere.[9]

In the present paper we use this formalism to describe the action of general linear optical systems on paraxial Maxwell beams. This gives a generalization of the Jones calculus based on the assumption of axial plane waves, to the physically more correct and consistent paraxial fields. Section 2 reviews the description of paraxial solutions of the complete Maxwell equations, which we have derived earlier. Such solutions, specified either by the $\underline{E}$ field or the $\underline{B}$ field, are expressed in a compact form by separating the independent field components from the dependent ones. The form of the most general linear system that can act on such fields is developed. This clearly goes beyond the class of systems describable within scalar optics. Each possible Jones matrix is shown to be uniquely and consistently extended to an operator designed to act on a paraxial $\underline{E}$ field, and an accompanying operator acting on the $\underline{B}$ field. In Section 3 some illustrative examples are worked out: the rotator, polarizers, retarders and thin lenses. In each case the way the formalism handles all components of the field in the proper way is clearly brought out. Section 4 contains some concluding remarks.

2. PARAXIAL MAXWELL FIELDS AND ACTION OF LINEAR OPTICAL SYSTEMS

Let us consider a paraxial quasimonochromatic electromagnetic wave propagating along the positive z-axis of a cartesian coordinate system. We denote the mean wave number and the transverse spread by $k = \omega/c$ and Δk respectively, $\Delta k << k$. For such a wave we have shown from front form analysis that upto and including terms of first order in the small quantity $\frac{\Delta k}{k}$ we have the approximate equalities:

$$E_a \approx \epsilon_{ab} B_b \,, \qquad \text{(a)}$$

$$E_3 \approx \frac{i}{k} \partial_a E_a \,, \qquad \text{(b)}$$

$$B_3 \approx \frac{i}{k} \partial_a B_a \qquad \text{(c)} \qquad (2.1)$$

Here $\underline{E}$ and $\underline{B}$ are the positive frequency (''analytic signal'') parts of the real fields, subscripts a,b,... run over the transverse values 1,2; and $\epsilon_{ab} = -\epsilon_{ba}$ with $\epsilon_{12} = 1$. In the radiation gauge appropriate to the front form, we have the condition $A_0 = A_3$; and for the above wave we have the approximate equalities

$$A_a \approx \frac{-i}{k} E_a \,,$$

$$A_0 \approx \frac{-i}{2k} E_3 \qquad (2.2)$$

valid to the same degree of accuracy as (2.1). Therefore a paraxial Maxwell wave can be adequately represented by the 3-component electric column vector

$$\underline{E}(x) = \begin{pmatrix} E_1(x) \\ E_2(x) \\ E_3(x) \end{pmatrix} \approx \begin{pmatrix} E_1(x) \\ E_2(x) \\ \frac{i}{k}\,\partial_a\, E_a(x) \end{pmatrix} , \tag{2.3}$$

or equally well by the magnetic column vector

$$\underline{B}(x) = \begin{pmatrix} B_1(x) \\ B_2(x) \\ B_3(x) \end{pmatrix} \approx \begin{pmatrix} -E_2(x) \\ E_1(x) \\ \frac{-i}{k}\,\epsilon_{ab}\,\partial_a\, E_b(x) \end{pmatrix} . \tag{2.4}$$

Thus, as long as we work only to the accuracy $\frac{\Delta k}{k}$, such a field is completely specified by the two independent analytic signals $E_a(x)$ with narrow angular spectra peaked about the positive z-axis.

The longitudinal components E_3, B_3 are smaller than the transverse components E_a, B_a by a factor $\frac{\Delta k}{k}$. For the purpose of the present paper we note that (2.1b) and (2.1c) are immediate consequences of the Maxwell equations

$$\underline{\nabla} \cdot \underline{E} = \underline{\nabla} \cdot \underline{B} = 0 \quad , \tag{2.5}$$

when one works only to the accuracy $\frac{\Delta k}{k}$. In a similar way, (2.1a) is an easy consequence of the Maxwell equation

$$\frac{1}{c} \frac{\partial}{\partial t} \underline{B} + \underline{\nabla} \wedge \underline{E} = 0. \tag{2.6}$$

The analysis based on the relativistic front form leads us to the following way of expressing the structure of the column vector $\underline{E}$ (x) in (2.3). Out of the leading components of the electric field we set up a special ''transverse'' column vector as

$$\underline{E}_T (x) = \begin{pmatrix} E_1(x) \\ E_2(x) \\ 0 \end{pmatrix} . \tag{2.7}$$

We then introduce the transverse ''momentum'' operators

$$P_a = \frac{-i}{k} \frac{\partial}{\partial x_a} \tag{2.8}$$

and two 3 x 3 matrices

$$G_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} , \qquad G_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & i & 0 \end{pmatrix} . \tag{2.9}$$

The usefulness of these matrices will be soon evident; they are parts of certain special combinations of the generators of the

Poincaré group. We combine them with the transverse coordinates x_a to define also

$$Q_a = x_a + \frac{i}{k} G_a \ . \qquad (2.10)$$

The matrices G_a are nilpotent and in fact satisfy

$$G_a G_b = 0 \ . \qquad (2.11)$$

As a result, among Q_a and P_a we have the algebraic relations

$$[Q_a, Q_b] = [P_a, P_b] = 0,$$

$$[Q_a, P_b] = \frac{i}{k} \delta_{ab} \ . \qquad (2.12)$$

The way in which E_3 is determined by E_a in (2.3) shows that the complete column vector $\underline{E}$ can be reconstructed from $\underline{E}_T$ by applying an operator built out of G_a and P_a:

$$\underline{E} = e^{i G_a P_a} \underline{E}_T \ . \qquad (2.13)$$

Thus any paraxial Maxwell beam is completely specified by the column vector $\underline{E}_T$ with independent first and second components and vanishing third component. (Except for a strictly axial wave, $\underline{E}_T$ itself is not an allowed electric field vector). Hence the most general linear optical system preserving the paraxial property can be represented by a linear transformation on $\underline{E}_T$ maintaining the form just noted:

$$\underline{E}_T \rightarrow \underline{E}_T' = \begin{pmatrix} E_1'(x_\perp) \\ E_2'(x_\perp) \\ 0 \end{pmatrix} = \Omega_T \begin{pmatrix} E_1(x_\perp) \\ E_2(x_\perp) \\ 0 \end{pmatrix} = \Omega_T \underline{E}_T . \tag{2.14}$$

Here, $x_\perp$ is the transverse coordinate in the input and output planes, and we have suppressed the dependence on z and t. Clearly, Ω_T is a 3 x 3 matrix, each of whose elements can be a function of $x_\perp$ and $P_\perp$; as a consequence in general the effects of the elements of Ω_T on E_1 and E_2 are to be given by suitable integral transformations. The demand that the third entry in $\underline{E}_T'$ be zero requires that the first and second elements in the third row of Ω_T vanish:

$$(\Omega_T)_{31} = (\Omega_T)_{32} = 0 . \tag{2.15}$$

By the same token, i.e. because of the form of $\underline{E}_T$, the elements in the third column of Ω_T are irrelevant. For the present we take them to be zero, and comment on this choice later. The general Ω_T is then

$$\Omega_T = \left(\begin{array}{cc|c} \multicolumn{2}{c|}{\multirow{2}{*}{J}} & 0 \\ & & 0 \\ \hline 0 & 0 & 0 \end{array}\right) , \tag{2.16}$$

where J is the familiar 2 x 2 Jones matrix of the optical system. Each of the four elements of J could depend on both $x_\perp$ and $P_\perp$.

The change in the complete column vector $\underline{E}$ when $\underline{E}_T$ experiences the change (2.14) is given by an operator arising out of Ω_T and the universal operator in (2.13) connecting $\underline{E}$ and $\underline{E}_T$:

$$\underline{E}' = \Omega \underline{E}$$

$$\Omega = e^{i\, G_a\, P_a}\, \Omega_T\, e^{-i\, G_b\, P_b} \quad . \tag{2.17}$$

We note that for a thin lens of focal length f,

$$J = e^{-\frac{ik}{2f}\, x_\perp^2} \cdot \mathbb{1} \quad ; \tag{2.18}$$

while for free propagation through a distance d,

$$J = e^{-\frac{i\,k\,d}{2}\, P_\perp^2} \cdot \mathbb{1} \quad . \tag{2.19}$$

In both (2.18) and (2.19), J is a multiple of the 2 x 2 unit matrix. These are therefore systems definable within scalar Fourier optics. For the most general system of this kind, J is a single linear operator $t(x_\perp, P_\perp)$ times the 2 x 2 unit matrix. Our earlier analysis has given an unambiguous rule to pass from scalar to vector Fourier optics for such systems: it is to replace $x_\perp$ by $Q_\perp$ of (2.10) within $t(x_\perp, P_\perp)$. The effect on $\underline{E}$ is then given by a 3 x 3 matrix operator $t(Q_\perp, P_\perp)$. (In fact, because the choice of $(\Omega_T)_{33}$ is free, we can take Ω to be $t(Q_\perp, P_\perp)$.) We now proceed to deal with the most general J, thus encompassing optical systems not definable within the scalar theory.

Along with the 2 x 2 unit matrix, the three Pauli matrices σ_1, σ_2, σ_3 give a basis of 2 x 2 matrices in terms of which any J can be expanded:

$$J = j_0(x_\perp, P_\perp) \cdot \mathbb{1}_{2 \times 2} + \sum_{r=1}^{3} j_r(x_\perp, P_\perp)\, \sigma_r \quad . \tag{2.20}$$

The coefficients j_0, j_r can be recovered by

$$j_0 = \frac{1}{2} \operatorname{Tr} J, \quad j_r = \frac{1}{2} \operatorname{Tr} J\, \sigma_r \quad . \tag{2.21}$$

We now use the freedom available in the choice of $(\Omega_T)_{33}$ and (instead of (2.16)) take

$$(\Omega_T)_{33} = j_0(x_\perp, P_\perp) \quad . \tag{2.22}$$

This makes the scalar optics limit, when $j_r = 0$, appear as simple as possible. The σ_r also are enlarged to 3 x 3 matrices $\bar{\sigma}_r$ by

$$\bar{\sigma}_r = \left(\begin{array}{cc|c} \multicolumn{2}{c|}{\multirow{2}{*}{σ_r}} & 0 \\ & & 0 \\ \hline 0 & 0 & 0 \end{array}\right) \quad . \tag{2.23}$$

Then the Ω_T corresponding to J of (2.20) can be taken as

$$\Omega_T = j_0(x_\perp, P_\perp) \cdot \mathbb{1}_{3 \times 3} + \sum_{r=1}^{3} j_r(x_\perp, P_\perp)\, \bar{\sigma}_r \quad . \tag{2.29}$$

It is now clear that we need to compute, just once, the three matrix operators

$$S_r = e^{i\, G_a\, P_a}\, \bar{\sigma}_r\, e^{-i\, G_b\, P_b} \quad ; \tag{2.25}$$

and then the operator Ω is immediately obtained from the 2 x 2 Jones matrix as

$$\Omega = j_0 \ (Q_\perp, P_\perp) + \sum_{r=1}^{3} j_r(Q_\perp, P_\perp) \ S_r \quad . \tag{2.26}$$

Here the 3 x 3 unit matrix accompanying j_0 is not explicitly indicated, since in general $j_0(Q_\perp, P_\perp)$ is itself a 3 x 3 matrix operator, as is each $j_r(Q_\perp, P_\perp)$. By virtue of the property (2.11) as well as

$$\bar{\sigma}_r \, e^{-i \, G_a \, P_a} = \bar{\sigma}_r \quad , \tag{2.27}$$

S_r are easily obtained:

$$S_1 = \left(\begin{array}{cc|c} \multicolumn{2}{c|}{\multirow{2}{*}{σ_1}} & 0 \\ & & 0 \\ \hline -P_2 & -P_1 & 0 \end{array}\right), \quad S_2 = \left(\begin{array}{cc|c} \multicolumn{2}{c|}{\multirow{2}{*}{σ_2}} & 0 \\ & & 0 \\ \hline -iP_2 & iP_1 & 0 \end{array}\right), \quad S_3 = \left(\begin{array}{cc|c} \multicolumn{2}{c|}{\multirow{2}{*}{σ_3}} & 0 \\ & & 0 \\ \hline -P_1 & P_2 & 0 \end{array}\right) \quad \ldots \tag{2.28}$$

By the method of construction it is guaranteed that for arbitrary input j_0 and j_r, as long as they do not violate the paraxial nature of the wave, the operator Ω of (2.26) can be applied to any incoming allowed electric field vector and the result will be another allowed electric field vector.

In passing we may note that even though $j_r(Q_\perp, P_\perp)$ and S_r in

(2.26) are 3 x 3 matrix operators, they commute with one another since they are similarity transforms of the sets $j_r(x_\perp, P_\perp)$ and $\bar{\sigma}_r$ which commute with one another. However the different j_r do not in general commute with one another, whether the arguments be $x_\perp$, $P_\perp$ or $Q_\perp$, $P_\perp$; and neither do the S_r.

To conclude this Section, we develop the transformation rule for the magnetic vector field $\underline{B}$ that accompanies (2.17). This is useful when one wishes to trace the changes undergone by the Poynting vector. For the paraxial case, we have a relation exactly similar to (2.13) for $\underline{B}$:

$$\underline{B} = e^{i\,G_a\,P_a}\,\underline{B}_T\,,$$

$$\underline{B}_T = \begin{pmatrix} B_1 \\ B_2 \\ 0 \end{pmatrix} = \begin{pmatrix} -E_2 \\ E_1 \\ 0 \end{pmatrix}\,. \qquad (2.29)$$

It is now easily checked that when the transverse electric components are transformed by the Jones matrix J expanded in the form (2.20) with coefficients j_0, j_r the transverse magnetic components are transformed by a matrix J_B which differs from J only in that j_1 and j_3 are reversed in sign while j_0 and j_2 are unchanged. In view of (2.29), we can thus write the action of an optical system on $\underline{E}$ and $\underline{B}$ compactly as:

$$\underline{E}' = \Omega\underline{E}\,, \qquad \underline{B}' = \Omega_B\,\underline{B}\,,$$

$$\Omega = j_0(Q_\perp, P_\perp) + \sum_{r=1}^{3} j_r(Q_\perp, P_\perp)\,S_r\,,$$

$$\Omega_B = j_0(Q_\perp, P_\perp) + \sum_{r=1}^{3} \epsilon_r\, j_r(Q_\perp, P_\perp)\, S_r$$

$$\epsilon_1 = \epsilon_3 = -1, \qquad \epsilon_2 = 1 \quad . \tag{2.30}$$

Obviously, for systems describable within scalar theory, we have $\Omega = \Omega_B$. This equality persists of course also for all systems for which $j_1 = j_3 = 0$.

3. EXAMPLES

To illustrate the general formalism of the previous Sections, we describe here the cases of the rotator, the polarizer, the retarder and the (symmetric) thin lens. In the first three examples, the Jones matrices are purely numerical with no dependence on either $x_\perp$ or $P_\perp$; in the last example, there is only an $x_\perp$ -dependence.

Rotator

Consider an optical system which rotates the plane of polarization by an amount θ. The corresponding Jones matrix is

$$J = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

$$j_0 = \cos\theta, \quad j_1 = j_3 = 0, \quad j_2 = i \sin\theta. \tag{3.1}$$

Hence 3 x 3 matrix operators Ω and Ω_B are equal:

$$\Omega = \Omega_B = \cos\theta + i \sin\theta \cdot S_2$$

$$= \left(\begin{array}{cc|c} \multicolumn{2}{c|}{\multirow{2}{*}{J}} & 0 \\ & & 0 \\ \hline \sin\theta\, P_2 & -\sin\theta\, P_1 & \cos\theta \end{array}\right). \tag{3.2}$$

The elements $(\Omega)_{31}$ and $(\Omega)_{32}$ are operators with respect to $x_\perp$ dependences.

Now assume an incoming paraxial plane wave with propagation vector $\underline{k}$ in the $x_1 - x_3$ plane at an angle α to the x_3 -axis, where $|\alpha|<<1$. Let the electric vector be polarized normal to the $x_1 - x_3$ plane. Since $\underline{k} = k(\alpha,0,1)$, the incident electric and magnetic fields are

$$\underline{E} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} u \, , \qquad \underline{B} = \begin{pmatrix} -1 \\ 0 \\ \alpha \end{pmatrix} u \, ,$$

$$u(x,t) = u_0 \, e^{ik(x_3 + \alpha x_1) - i\omega t} \, . \tag{3.3}$$

Thus, $\underline{E}$, $\underline{B}$ and $\underline{k}$ form an orthogonal right handed triplet. The amplitude $u(\underline{x},t)$ obeys

$$P_1 \, u = \alpha u, \quad P_2 u = 0 \quad . \tag{3.4}$$

By applying Ω of (3.2) to the fields $\underline{E}$, $\underline{B}$, we find that the outgoing fields after action by the rotator are

$$\underline{E}' = \Omega \, \underline{E} = \begin{pmatrix} \sin\theta \\ \cos\theta \\ -\alpha \sin\theta \end{pmatrix} u \, , \quad \underline{B}' = \Omega \, \underline{B} = \begin{pmatrix} -\cos\theta \\ \sin\theta \\ \alpha \cos\theta \end{pmatrix} u \, . \tag{3.5}$$

The output is a plane wave with the same propagation vector $\underline{k}$ as the input, and $\underline{E}'$, $\underline{B}'$, $\underline{k}$ also form a right hand orthogonal triplet (to first order in α). We also note that $\underline{E}'$ and $\underline{B}'$ are properly related to $\underline{E}$ and $\underline{B}$ through a rotation of amount θ about $\underline{k}$. In the particular case $\theta = \pi/2$, we find that $\underline{B}$ loses its longitudinal component while $\underline{E}$ picks up such a component:

$$\theta = \pi/2: \qquad \underline{E}' = -\underline{B} \ , \qquad \underline{B}' = \underline{E} \ . \tag{3.6}$$

These results illustrate that the formalism handles the longitudinal components properly.

Polarizers and Retarders

A polarizer attenuates two mutually perpendicular components of the transverse $\underline{E}$ field by different amounts, while a retarder introduces a phase difference between them. In either case we can choose the x_1, x_2 axes to be along these eigendirections, so both systems correspond to diagonal numerical Jones matrices:

$$J = \begin{pmatrix} j_0 + j_3 & 0 \\ 0 & j_0 - j_3 \end{pmatrix} . \tag{3.7}$$

Both j_0 and j_3 are numerical parameters. For a retarder, we have

$$j_0 = \text{Cos}\ \delta/2, \qquad j_3 = i\ \text{Sin}\ \delta/2 \ , \tag{3.8}$$

with real δ. For a (partial) polarizer, both j_0 and j_3 are real and obey

$$0 \leq j_0 \pm j_3 \leq 1 \quad . \tag{3.9}$$

The x_1-polarizer corresponds to $j_0 = j_3 = 1/2$.

The matrices Ω and Ω_B are now different:

$$\Omega = j_0 + j_3 S_3 = \left(\begin{array}{cc|c} \multicolumn{2}{c|}{\multirow{2}{*}{J}} & 0 \\ & & 0 \\ \hline -j_3 P_1 & j_3 P_2 & j_0 \end{array}\right) ,$$

$$\Omega_B = j_0 - j_3 S_3 = \left(\begin{array}{cc|c} j_0 - j_3 & 0 & 0 \\ 0 & j_0 + j_3 & 0 \\ \hline j_3 P_1 & -j_3 P_2 & j_0 \end{array}\right) \tag{3.10}$$

As an example let us take first the case of a retarder and let the input be again a paraxial plane wave with propagation vector $\underline{k} = k(\alpha,\beta,1)$, $|\alpha|$, $|\beta| << 1$. Let it be plane polarized with the $\underline{E}$ vector making an angle of $\pi/4$ radians with the x_1-axis. We have

$$\underline{E} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ -\alpha-\beta \end{pmatrix} v \quad ,$$

$$v(\underline{x},t) = v_0 \, e^{ik(\alpha x_1 + \beta x_2 + x_3) - i\omega t} \quad . \tag{3.11}$$

The amplitude v now obeys

$$P_1 v = \alpha v, \qquad P_2 v = \beta v \quad . \tag{3.12}$$

We note that all the components of $\underline{E}$ are in phase. After action by the retarder, eq. (3.8), we have the output electric field

$$\underline{E}' = \Omega \underline{E} = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{i\delta/2} \\ e^{-i\delta/2} \\ -\alpha e^{i\delta/2} - \beta e^{-i\delta/2} \end{pmatrix} v \quad . \tag{3.13}$$

Thus Ω has not only retarded the phase of E_2 relative to E_1, but has also altered the longitudinal component E_3 by just the right amount so that $\underline{E}'$ remains orthogonal to the (unchanged) propagation vector $\underline{k}$.

To illustrate the action of the polarizer, eq. (3.9), let us take an input paraxial plane wave $\underline{k} = k(\alpha,0,1)$, with the electric vector at an angle of $\pi/4$ radians with the x_1-axis:

$$\underline{E} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ -\alpha \end{pmatrix} u \quad . \tag{3.14}$$

The amplitude u is given in (3.3), and it obeys (3.4). For an x_2-polarizer, $j_0 = -j_3 = 1/2$, we find:

$$\underline{E}' = \Omega\,\underline{E} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} u \tag{3.15}$$

which correctly describes a plane wave with unaltered wave vector $\underline{k}$. The important point to note is that Ω acting on $\underline{E}$ not only annihilates the component E_1 but also properly reduces E_3 to zero, so that $\underline{E}'$ remains orthogonal to $\underline{k}$.

Thin Lens

As a final illustration, we consider the action of a thin lens whose 2 x 2 Jones matrix is given in (2.18):

$$j_0 = e^{\frac{-ik}{2f} x_\perp^2} , \qquad j_r = 0 \quad . \tag{3.16}$$

Here we have a dependence on $x_\perp$. It follows from (2.30) that

$$\Omega = \Omega_B = j_0(Q_\perp) = e^{\frac{-ik}{2f} (x_\perp + \frac{i}{k} G_\perp)^2}$$

$$= e^{\frac{-ik}{2f} x_\perp^2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{x_1}{f} & \frac{x_2}{f} & 1 \end{pmatrix} \quad . \tag{3.17}$$

As an interesting application, let us compute the effect of the lens on the Poynting vector $\underline{P} = \mathrm{Re}\ \underline{E} \wedge \underline{B}^*$. In a paraxial field composed of plane waves whose propagation directions all make small angles with the x_3-axis, the Poynting vector has a major longitudinal component and small transverse components:

$$|P_a(x_\perp)| << P_3(x_\perp) \quad . \tag{3.18}$$

The change in $\underline{E}$ and $\underline{B}$ are:

$$\begin{pmatrix} E'_a(x_\perp) \\ E'_3(x_\perp) \\ B'_a(x_\perp) \\ B'_3(x_\perp) \end{pmatrix} \approx e^{\frac{-ik}{2f} x_\perp^2} \begin{pmatrix} E_a(x_\perp) \\ E_3(x_\perp) + \frac{x_a}{f} E_a(x_\perp) \\ B_a(x_\perp) \\ B_3(x_\perp) + \frac{x_a}{f} B_a(x_\perp) \end{pmatrix} \quad . \tag{3.19}$$

From this the output Poynting vector is related to the input by

$$P'_3(x_\perp) \approx P_3(x_\perp) \quad ,$$

$$P'_a(x_\perp) \approx P_a(x_\perp) - \frac{x_a}{f} P_3(x_\perp) \quad . \tag{3.20}$$

One identifies this to be formally the same as the ray transfer equation for a thin lens in geometrical optics. Thus for a lens located at $x_3 = 0$ and an input field generated by a point source located at the paraxial point $(a_\perp, -u)$, $|a_\perp| << u$, $u > f$, the Poynting vector at points $x_\perp$ over the plane immediately before the lens is proportional to

$$\underline{P}(x_\perp) = (\frac{1}{u}(x_\perp - a_\perp), 1). \tag{3.21}$$

Using (3.20), the Poynting vector over the plane immediately after the lens is proportional to

$$\underline{P}'(x_\perp) = (\frac{1}{u}(x_\perp - a_\perp) - \frac{x_\perp}{f}, 1)$$

$$= (\frac{1}{v}(b_\perp - x_\perp), 1) \tag{3.22}$$

where v and $b_\perp$ are given by

$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f}, \qquad b_\perp = -\frac{v}{u} a_\perp . \tag{3.23}$$

The output Poynting vectors (3.22) for various $x_\perp$ are all directed towards the point $(b_\perp, v)$. Thus the thin lens and magnification formulae (3.23) have been derived from Poynting vector (energy flow) considerations, for which obviously the longitudinal components E_3 and B_3 had to be handled consistently.[10]

4. CONCLUDING REMARKS

Our analysis in the preceding Sections generalizes the Jones calculus designed to deal with the transverse electric field components of an axial plane wave to one which describes in a consistent way all the components of the field vectors of a paraxial wave. In this process we have also developed a procedure for constructing the Jones matrix acting on the magnetic vector given the one to act on the electric vector. We have applied our formalism to several simple examples and in each case we have shown that it leads to physically expected results, and that the small longitudinal components are essential for a consistent description.

The present analysis can be easily extended to generalize the Chandrasekhar-Mueller-Stokes calculus for the fields and optical systems. It should be noted that earlier attempts[11] at such a generalization have dealt with only the fields and not the optical systems.

Generalized light rays endowed with polarization properties[12] have proved to give an exact geometrical picture of the polarization properties of radiation fields and we have recently analysed the behaviour of these rays under action by optical systems which are describable within scalar Fourier optics[7,9]. The present formalism can be applied to analyse the behaviour of these polarized rays under the action of the most general linear optical system. Work in this direction is in progress and will be reported elsewhere.

ACKNOWLEDGEMENTS

One of the authors (R.S.) is grateful to the Research and Development Committee of the American College for financial support and to the Centre for Theoretical Studies, Indian Institute of Science for a visitorship during which a part of this work was carried out. The other two authors are grateful to Prof. Jan Nilsson for his gracious hospitality at the Institute of Theoretical Physics, Göteborg, where this work was completed.

REFERENCES

1. R.C. Jones, J.Opt.Soc.Am. 31, 488, 493, 500 (1941); 32, 486 (1942).

2. S. Pancharatnam, Proc. Indian Acad. Sci. A 44, 398 (1956).

3. E. Wolf, Nuovo Cim. 12, 884 (1954); 13, 1165 (1959); G.B. Parrent, Jr. and P. Roman, Nuovo Cim. 15, 370 (1960); N. Wiener, Acta Math. 55, 118 (1930).

4. G.G. Stokes, Trans. Cambridge Phil. Soc. 9, 399 (1852); H. Mueller, J. Opt. Soc. Am. 38, 661 (1948); S. Chandrasekhar, Radiative Transfer (Dover, New York, 1950) Ch.1.

5. R. Simon, Opt. Commun. 42, 293 (1982); R. Barakat, Opt. Commun. 38, 159 (1981).

6. E.C.G. Sudarshan, R. Simon and N. Mukunda, Phys. Rev. A 28, 2921 (1983).

7. N. Mukunda, R. Simon and E.C.G. Sudarshan, Phys. Rev. A 28, 2933 (1983).

8. P.A.M. Dirac, Rev. Mod. Phys. 21, 392 (1949).

9. N. Mukunda, R. Simon and E.C.G. Sudarshan, J. Opt. Soc. Am. (submitted).

10. See in this connection M. Lax, W.H. Louisell and W.H. McKnight, Phys. Rev. A 11, 455 (1975).

11. P. Roman, Nuovo Cim. 13, 974 (1959); G. Ramachandran, M.V.N. Murthy and K.S. Mallesh, Pramana 15, 356 (1980).

12. E.C.G. Sudarshan, Physica 96 A, 315 (1979); Phys. Rev. A 23, 2802 (1981).