Pure Appl. Opt. **7** (1998) 941–951. Printed in the UK PII: S0963-9659(98)91420-9

Beam coherence–polarization matrix

F Gori†, M Santarsiero†, S Vicalvi‡, R Borghi§ and G Guattari§

† Dipartimento di Fisica, Università degli Studi Roma Tre, and Istituto Nazionale per la Fisica della Materia, Via della Vasca Navale 84, I-00146 Rome, Italy

‡ Dipartimento di Fisica, Università La Sapienza, P le Aldo Moro 2, I-00185 Rome, Italy

§ Dipartimento di Ingegneria Elettronica, Università degli Studi Roma Tre, and Istituto Nazionale per la Fisica della Materia, Via della Vasca Navale 84, I-00146 Rome, Italy

Received 5 February 1998, in final form 25 May 1998

Abstract. We present an approach for describing the properties of a quasi-monochromatic, beam-like field that is both partially polarized and partially coherent from the spatial standpoint. It is based on the use of a single 2×2 matrix, called the beam coherence–polarization matrix, whose elements have the form of mutual intensities. This approach, which can be viewed as an approximate form of Wolf's general tensorial theory of coherence, appears to be very simple, yet it is able to cover significant aspects of the beam behaviour that would not be accounted for by a scalar theory or by a local polarization matrix approach. A peculiar interference law applying to mutual intensities is derived. We show through simple examples how this approach leads to distinguish fields that would appear identical in a scalar treatment or in a local polarization matrix description. Hints for extensions are given.

1. Introduction

The study of the polarization properties of light is of the utmost importance in optics, both conceptually and for applications [1–3]. It is a fascinating subject that often leads to results of great elegance and significance. As a matter of fact, some of the most intriguing topics in classical and quantum optics, such as Pancharatnam's phase [4–6] and the optical version of the EPR paradox [1, 7], to name only two, have to do with polarization. Therefore, it is no surprise that this subject, with its extraordinary richness, keeps on attracting the attention of scientists [8–10]. We are interested here in partially polarized light, which includes of course completely polarized and completely unpolarized light as particular cases. In the simplest situation we deal with a quasi-monochromatic plane wave. Polarization properties can then be described by a 2×2 matrix known as the coherence matrix or the polarization matrix [1, 11]. Here, we shall adopt the latter name. Essentially, the polarization matrix is a set of four numbers (two of them being real and the other two complex conjugate) that give the self- and mixed correlations, at equal times, of the Cartesian components of the electric field of the wave. Equivalently, we can use Stokes parameters [1]. This type of description can also be used whenever the light field, without being a plane wave, is in the form of a well collimated, uniform beam [1]. In this case one assumes that the field is approximately the same at any point in a beam section. This implies that the polarization is constant and that spatial coherence is complete across a transverse plane. Indeed, it is amazing how many different phenomena can be handled by this simple model. At present, however, beams that depart from the previous hypotheses are becoming of interest [12–14]. One then wonders which approach is to be used. An answer was furnished

0963-9659/98/050941+11$19.50

long ago by Wolf [15] through his general tensorial theory of the electromagnetic field. Generally speaking, this is a rather demanding approach since it involves the use of three different 3×3 matrices† whose elements are functions of six (scalar) space variables. However, a drastic simplification can be made if the beam structure of the field is taken into account. In this case, the electromagnetic field is essentially transverse with respect to the propagation axis and this allows us to neglect the longitudinal vector components. In addition, we may base our description on one matrix only, much in the same spirit as in the ordinary use of the polarization matrix. We then pass from three 3×3 matrices to a single 2×2 matrix. For reasons to be discussed later, we shall refer to it as the beam coherence–polarization matrix (or BCP matrix, for the sake of brevity). A brief sketch of it has been presented in [16]. This type of approach has already been used by James, who showed how the polarization of a Gaussian Schell-model beam can change upon propagation [17]. Furthermore, partially polarized, partially coherent beams have been studied recently by Martínez-Herrero *et al* by using an approach based on the Wigner distribution function [18].

Here we discuss the basis of an approach based on the BCP matrix. The elements of such a matrix are similar to mutual intensities [1]. Thanks to symmetry properties only three of the matrix elements are required, each of them obeying well known wave equations. After a brief discussion of the main properties of the BCP matrix we shall examine its use in the analysis of experiments in which anisotropic optical elements (either spatially uniform or not) are inserted in the beam path. This will lead us to establish an interference law of a new type in that it applies to mutual intensity functions rather than to optical intensities only. As an example of the use of the present approach we shall discuss three beams that would be indistinguishable in the scalar treatment, while exhibiting measurable differences in the BCP matrix description. Throughout this paper we will draw from the theory of the ordinary polarization matrix as well as from the general tensorial approach by Wolf. The reader can consult chapter 6 of [1] for a complete exposition of the present state of the vectorial theory of coherence. Explicit references to such work will be made at specific points. A few hints for extensions will be given in the final section.

2. The beam coherence–polarization matrix

Many partially coherent light fields discussed in the scalar theory of coherence propagate in the form of beams, i.e. they remain essentially concentrated around a mean axis without a large angular spread. This applies to light fields generated by lasers oscillating on several transverse modes [19–21] as well as to partially coherent fields obtained through synthesis procedures [22, 23]. Conditions to be satisfied for beam-like propagation are discussed in [1]. Essentially, both the coherence length across the source and the linear dimensions of the source itself must be large with respect to the wavelength‡. Typical angular widths of a beam are of the order of a few milliradians or less. With similar angular apertures the components of the electric and magnetic vectors along the beam axis have amplitudes two or three orders of magnitude smaller than the amplitudes of the vectors themselves. We can then neglect, to a good approximation, the longitudinal vector components. We shall refer to this as the quasi-transversality approximation.

† Actually four matrices appear in the general theory, but two of them are equivalent.

‡ It may be useful to recall that filtered sunlight arriving at the Earth's surface has a coherence (transverse) length of the order of 100 wavelengths [1].

There is another important point to be made. In the usual matrix theory of well collimated, uniform, quasi-monochromatic beams only one vectorial quantity† is used to describe the polarization properties of the field. Under the assumption of stationarity for the light field, the essential property of such a vector is that the time average of its square must be representative of the response signal of the detector used in the experiments. Such a single-vector description of the wave field has proved to be fully satisfactory. Physically, this can be explained by noting that, since the field is constituted by a set of plane waves with slightly different directions, the Poynting vector can be assumed to be proportional to the square of the vectorial quantity used to describe the field, as for a single plane wave. The same holds true for the cases of our interest. Accordingly, we shall use an identical approach.

Let us then introduce the covariance matrix of the electric field vector [1]. Since we want to consider a typical cross section of the beam, we adopt the following system of coordinates. The z-axis is along the effective direction of propagation of the beam. Then, at any plane $z =$ constant, we denote by $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ the position vectors, lying in that plane, of two typical points. In the quasi-transversality approximation only the x- and y-components of the electric vector are considered and the corresponding covariance matrix across the plane $z =$ constant, at time delay τ, is

$$\hat{\Gamma}^{(E)}(\boldsymbol{r}_1, \boldsymbol{r}_2, z; \tau) = \begin{pmatrix} \Gamma_{xx}^{(E)}(\boldsymbol{r}_1, \boldsymbol{r}_2, z; \tau) & \Gamma_{xy}^{(E)}(\boldsymbol{r}_1, \boldsymbol{r}_2, z; \tau) \\ \Gamma_{yx}^{(E)}(\boldsymbol{r}_1, \boldsymbol{r}_2, z; \tau) & \Gamma_{yy}^{(E)}(\boldsymbol{r}_1, \boldsymbol{r}_2, z; \tau) \end{pmatrix} \tag{1}$$

where

$$\Gamma_{\alpha\beta}^{(E)}(\boldsymbol{r}_1, \boldsymbol{r}_2, z; \tau) = \left\langle E_\alpha^*(\boldsymbol{r}_1, z; t)\, E_\beta(\boldsymbol{r}_2, z; t + \tau)\right\rangle. \tag{2}$$

Here and in the following, α and β denote the subscripts x or y. The time variable is denoted by t. E_x and E_y are the components of the complex electric field vector [1] in the x- and y-directions, respectively, and the asterisk denotes a complex conjugate. The field is assumed to be stationary and ergodic. Angular brackets indicate a time average. Here and below, we denote a matrix by a symbol surmounted by a caret.

We assume the field to be quasi-monochromatic with mean frequency $\bar{\nu}$ and effective bandwidth $\Delta\nu$. Furthermore, we suppose that the time delays τ introduced by the experimental apparatus are small compared to the coherence time $1/\Delta\nu$. In this case it can be easily shown that the following approximation holds:

$$\Gamma_{\alpha\beta}^{(E)}(\boldsymbol{r}_1, \boldsymbol{r}_2, z; \tau) = J_{\alpha\beta}(\boldsymbol{r}_1, \boldsymbol{r}_2, z) \exp(-2\pi \mathrm{i}\bar{\nu}\tau) \tag{3}$$

where

$$J_{\alpha\beta}(\boldsymbol{r}_1, \boldsymbol{r}_2, z) = \left\langle E_\alpha^*(\boldsymbol{r}_1, z; t)\, E_\beta(\boldsymbol{r}_2, z; t)\right\rangle \tag{4}$$

are the elements of the covariance matrix at equal times. Consequently, the polarization and spatial coherence properties are specified by the matrix

$$\hat{J}(\boldsymbol{r}_1, \boldsymbol{r}_2, z) = \begin{pmatrix} J_{xx}(\boldsymbol{r}_1, \boldsymbol{r}_2, z) & J_{xy}(\boldsymbol{r}_1, \boldsymbol{r}_2, z) \\ J_{yx}(\boldsymbol{r}_1, \boldsymbol{r}_2, z) & J_{yy}(\boldsymbol{r}_1, \boldsymbol{r}_2, z) \end{pmatrix}. \tag{5}$$

† Typically the electric vector is used but, depending on the detection process, other vectorial quantities, e.g. the vector potential [1], could be more suitable.

Equation (5) can be written in a more compact form by introducing the row vector $\boldsymbol{E}(\boldsymbol{r}, z; t) = \left[E_x(\boldsymbol{r}, z; t) \quad E_y(\boldsymbol{r}, z; t)\right]$, thus giving [1]

$$\hat{J}(\boldsymbol{r}_1, \boldsymbol{r}_2, z) = \langle \boldsymbol{E}^{\dagger}(\boldsymbol{r}_1, z; t)\, \boldsymbol{E}(\boldsymbol{r}_2, z; t)\rangle \tag{6}$$

where the dagger denotes the Hermitian conjugate.

Matrix (5) shares with the ordinary polarization matrix [1] the fact that it refers to light beams, but differs from it since its elements are functions of two position vectors instead of being constant parameters of the beam. This is because the present matrix formulation does not rely on the hypothesis of complete spatial coherence across the beam section. On the other hand, the general Wolf coherence matrices for the electromagnetic field also have elements depending on two position vectors, but differ from (5) since they have nine elements instead of four. This is because the general matrix formulation does not require the hypothesis that the field is beam-like. Therefore, we can distinguish matrix (5) from other coherence matrices by adopting the name beam coherence–polarization matrix (BCP matrix).

Of course the BCP matrix includes the ordinary polarization matrix as the particular case in which $J_{\alpha\beta}$ is independent of the space coordinates for all points within the beam. In turn, the BCP matrix can be thought of as a particular case of the general coherence matrix of the electric vector when the z-component of the field can be neglected and the quasi-monochromaticity hypothesis applies.

As shown by equation (4), the elements of $\hat{J}$ have the same structure as the mutual intensity used in the scalar theory of coherence [1, 11]. More precisely, J_{xx} (J_{yy}) is the mutual intensity that would characterize the beam if the y-component (x-component) of the field were eliminated (for example, by a linear polarizer). On the other hand, J_{xy} and J_{yx} account for the correlations that exist at two distinct points between the x- and the y-components. Only one of these functions is needed since it follows from equation (4) that

$$J_{yx}(\boldsymbol{r}_1, \boldsymbol{r}_2, z) = J^{*}_{xy}(\boldsymbol{r}_2, \boldsymbol{r}_1, z). \tag{7}$$

The optical intensity at a point $(\boldsymbol{r}, z)$ is made up of two contributions corresponding to the orthogonal components of the electric vector. Disregarding a proportionality factor, as in scalar theory, we shall compute the optical intensity $I(\boldsymbol{r}, z)$ as

$$I(\boldsymbol{r}, z) = J_{xx}(\boldsymbol{r}, \boldsymbol{r}, z) + J_{yy}(\boldsymbol{r}, \boldsymbol{r}, z) = \mathrm{Tr}\left\{\hat{J}(\boldsymbol{r}, \boldsymbol{r}, z)\right\} \tag{8}$$

where Tr denotes the trace.

Using the Schwarz inequality it can be easily proved that

$$\left|J_{\alpha\beta}(\boldsymbol{r}_1, \boldsymbol{r}_2, z)\right|^2 \leqslant J_{\alpha\alpha}(\boldsymbol{r}_1, \boldsymbol{r}_1, z)\, J_{\beta\beta}(\boldsymbol{r}_2, \boldsymbol{r}_2, z). \tag{9}$$

This allows us to introduce a normalized version of the elements of the BCP matrix through the definitions

$$j_{\alpha\beta}(\boldsymbol{r}_1, \boldsymbol{r}_2, z) = \frac{J_{\alpha\beta}(\boldsymbol{r}_1, \boldsymbol{r}_2, z)}{\sqrt{J_{\alpha\alpha}(\boldsymbol{r}_1, \boldsymbol{r}_1, z)\, J_{\beta\beta}(\boldsymbol{r}_2, \boldsymbol{r}_2, z)}}. \tag{10}$$

Equation (9) implies that $\left|j_{\alpha\beta}\right| \leqslant 1$. The diagonal elements j_{xx} and j_{yy} are the complex degrees of coherence for the x- and y-components of the field, respectively. As for j_{xy}, it represents the degree of cross-coherence between the x-component at $(\boldsymbol{r}_1, z)$ and the y-component at $(\boldsymbol{r}_2, z)$ (and a similar meaning can be given to j_{yx}). We shall see later on that j_{xy} and j_{yx} are measurable through suitable interference experiments.

It is worthwhile to remark that $J_{xy}(\boldsymbol{r}_1, \boldsymbol{r}_2, z)$ is not locally connected to $J_{xx}(\boldsymbol{r}_1, \boldsymbol{r}_2, z)$ and $J_{yy}(\boldsymbol{r}_1, \boldsymbol{r}_2, z)$ except when $\boldsymbol{r}_1 = \boldsymbol{r}_2$ (see equation (9)). This means, for example, that

$J_{xy}(\boldsymbol{r}_1, \boldsymbol{r}_2, z)$ can be different from zero even if (at the same points) J_{xx} and J_{yy} vanish. There is, however, an important global connection that can be derived from the general semi-definiteness conditions obeyed by the coherence matrices [1]. Denoting by $f_1(\boldsymbol{r})$ and $f_2(\boldsymbol{r})$ two arbitrary well behaved functions, the following inequality holds:

$$\iint \big[f_1^*(\boldsymbol{r}_1)\, f_1(\boldsymbol{r}_2)\, J_{xx}(\boldsymbol{r}_1, \boldsymbol{r}_2, z) + f_2^*(\boldsymbol{r}_1)\, f_2(\boldsymbol{r}_2)\, J_{yy}(\boldsymbol{r}_1, \boldsymbol{r}_2, z) + f_1^*(\boldsymbol{r}_1)\, f_2(\boldsymbol{r}_2)\, J_{xy}(\boldsymbol{r}_1, \boldsymbol{r}_2, z) + f_2^*(\boldsymbol{r}_1)\, f_1(\boldsymbol{r}_2)\, J_{yx}(\boldsymbol{r}_1, \boldsymbol{r}_2, z) \big]\, \mathrm{d}\boldsymbol{r}_1\, \mathrm{d}\boldsymbol{r}_2 \geqslant 0 \tag{11}$$

where the integrals are extended to the beam section. Using equation (7) it is seen at once that equation (11) can be written as

$$\iint \big[f_1^*(\boldsymbol{r}_1) f_1(\boldsymbol{r}_2)\, J_{xx}(\boldsymbol{r}_1, \boldsymbol{r}_2, z) + f_2^*(\boldsymbol{r}_1)\, f_2(\boldsymbol{r}_2)\, J_{yy}(\boldsymbol{r}_1, \boldsymbol{r}_2, z) + 2\,\mathrm{Re}\left\{ f_1^*(\boldsymbol{r}_1) f_2(\boldsymbol{r}_2)\, J_{xy}(\boldsymbol{r}_1, \boldsymbol{r}_2, z) \right\} \big]\, \mathrm{d}\boldsymbol{r}_1\, \mathrm{d}\boldsymbol{r}_2 \geqslant 0 \tag{12}$$

where Re{} stands for the real part. Letting $f_2 = 0$ ($f_1 = 0$) we see that J_{xx} (J_{yy}) is a semi-definite positive kernel. Furthermore, equation (12) shows that the functions $J_{\alpha\beta}$ cannot be chosen independently from one another.

The local polarization properties of the beam at a typical point of a cross section are specified by the BCP matrix with equal arguments. We shall refer to it as the local polarization matrix. Further, the degree of polarization is given by [1, 11]

$$P(\boldsymbol{r}, z) = \sqrt{1 - \frac{4 \det\left\{ \hat{J}(\boldsymbol{r}, \boldsymbol{r}, z) \right\}}{\left(\mathrm{Tr}\left\{ \hat{J}(\boldsymbol{r}, \boldsymbol{r}, z) \right\} \right)^2}} \tag{13}$$

where det stands for determinant. It should be noted that P will, in general, be different from one point to another.

3. Interference and propagation of partially coherent, partially polarized beams

Let us discuss how the BCP matrix changes when some anisotropic optical element is inserted in the beam path. We suppose that such an element is a flat transparency, i.e. that its thickness can be neglected. Then it can be represented by a Jones matrix [1, 24] of the form

$$\hat{T}(\boldsymbol{r}) = \begin{pmatrix} a(\boldsymbol{r}) & b(\boldsymbol{r}) \\ c(\boldsymbol{r}) & d(\boldsymbol{r}) \end{pmatrix} \tag{14}$$

which differs from the usual form only because the matrix elements can be functions of the position across the element plane. Simple examples are offered by a linear polarizer covering only a portion of the beam section or by a lens made of a birefringent material. When the anisotropic element is inserted in the beam path at a certain plane $z =$ constant, the BCP matrix of the field emerging from it, say $\hat{J}'$, can be evaluated through the relation

$$\hat{J}'(\boldsymbol{r}_1, \boldsymbol{r}_2, z) = \hat{T}^{\dagger}(\boldsymbol{r}_1)\, \hat{J}(\boldsymbol{r}_1, \boldsymbol{r}_2, z)\, \hat{T}(\boldsymbol{r}_2). \tag{15}$$

Equation (15) can be derived by the same procedure used for the ordinary polarization matrix and indeed is quite similar to a relation holding for the latter [1], except that the matrices multiplying $\hat{J}$ on the right and on the left have different arguments.

A relation of the form (15) also holds for determining the form taken by $\hat{J}$ when the axes x and y are rotated, provided that $\hat{T}$ represents the rotation matrix.

Let us use equations (8) and (15) for evaluating the intensity emerging from a uniform linear polarizer, such as a sheet of polarizing film. If the transmission axis of the polarizer forms an angle α (anti-clockwise) with the x-axis, the matrix has the form

$$\hat{T} = \begin{pmatrix} C^2 & CS \\ CS & S^2 \end{pmatrix} \tag{16}$$

where $C = \cos\alpha$ and $S = \sin\alpha$ and the dependence on $\boldsymbol{r}$ has been dropped because the polarizer is uniform. On inserting from equation (16) into (15) and evaluating the trace of $\hat{J}'$ we find that the intensity, say $I_\alpha(\boldsymbol{r}, z)$, after the polarizer is

$$I_\alpha(\boldsymbol{r}, z) = J_{xx}(\boldsymbol{r}, \boldsymbol{r}, z)\, C^2 + J_{yy}(\boldsymbol{r}, \boldsymbol{r}, z)\, S^2 + 2\,\mathrm{Re}\left\{J_{xy}(\boldsymbol{r}, \boldsymbol{r}, z)\right\} CS. \tag{17}$$

This relation is also similar to the one that is found using the ordinary polarization matrix [1]. Again, the difference is that here all the involved quantities depend on the position.

Let us now give a brief discussion on the role played by the BCP matrix in the analysis of the classical Young interference experiment. Since in our procedure we have simply to rephrase a type of analysis carried out in a full way in several references (see e.g. [1, 11, 25]) we can omit most details of the calculation. Let the Young mask be located at $z = z_m$ with pinholes at $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$. Further, let $z = z_p$ be the observation plane. The observed optical intensity is the sum of two contributions due to the (orthogonal) x- and y-components of the field.

We then easily find

$$\begin{aligned} I(\boldsymbol{r}, z_p) = {} & |K_1|^2 \left[J_{xx}(\boldsymbol{r}_1, \boldsymbol{r}_1, z_m) + J_{yy}(\boldsymbol{r}_1, \boldsymbol{r}_1, z_m)\right] \\ & + |K_2|^2 \left[J_{xx}(\boldsymbol{r}_2, \boldsymbol{r}_2, z_m) + J_{yy}(\boldsymbol{r}_2, \boldsymbol{r}_2, z_m)\right] \\ & + 2\,\mathrm{Re}\left\{K_1^* K_2 \left[J_{xx}(\boldsymbol{r}_1, \boldsymbol{r}_2, z_m) + J_{yy}(\boldsymbol{r}_1, \boldsymbol{r}_2, z_m)\right] \exp(-\mathrm{i}2\pi\bar{\nu}\tau)\right\}. \end{aligned} \tag{18}$$

The quantities K_1 and K_2 are purely imaginary numbers that depend on the size of the pinholes and on the geometry [1, 11]. The time delay τ is given by

$$\tau = \frac{1}{c}\left\{\sqrt{(\boldsymbol{r} - \boldsymbol{r}_1)^2 + \left(z_p - z_m\right)^2} - \sqrt{(\boldsymbol{r} - \boldsymbol{r}_2)^2 + \left(z_p - z_m\right)^2}\right\} \tag{19}$$

where c is the speed of light. Equation (18) can be written in the more compact form

$$I(\boldsymbol{r}, z_p) = |K_1|^2\, I(\boldsymbol{r}_1, z_m) + |K_2|^2\, I(\boldsymbol{r}_2, z_m) + 2\,\mathrm{Re}\left\{K_1^* K_2 J_{\mathrm{eq}}(\boldsymbol{r}_1, \boldsymbol{r}_2, z_m) \exp(-\mathrm{i}2\pi\bar{\nu}\tau)\right\} \tag{20}$$

where we used equation (8) and we have defined an equivalent mutual intensity at the pinholes as

$$J_{\mathrm{eq}}(\boldsymbol{r}_1, \boldsymbol{r}_2, z_m) = J_{xx}(\boldsymbol{r}_1, \boldsymbol{r}_2, z_m) + J_{yy}(\boldsymbol{r}_1, \boldsymbol{r}_2, z_m). \tag{21}$$

It is seen that equation (20) is exactly of the same form that one obtains in the scalar theory provided that we take as the mutual intensity the sum of the corresponding functions for the x- and y-components of the field vector. This is physically very sound. After all, if no anisotropic element is used, we should see the superposition of the interference patterns produced by the x- and y-components of the field at the pinholes. It may be worthwhile to note that rule (21) for the equivalent mutual intensity is the same as what we would obtain in scalar theory for the superposition of two uncorrelated fields. Here, however, the origin is different because $J_{xy}(\boldsymbol{r}_1, \boldsymbol{r}_2, z_m)$ will, in general, be different from zero. Rule (21) simply stems from the geometrical orthogonality of the two components of the field vector. Just as in the case of the intensity distribution across the beam (see equation (17)), the role of anti-diagonal elements of the BCP matrix appears when some optical anisotropy occurs.

As a simple example, suppose the Young mask is covered with a uniform linear polarizer set at an angle φ with respect to the x-axis. On evaluating the intensity distribution at the observation plane we find

$$I_\varphi(\boldsymbol{r}, z_p) = |K_1|^2\, I_\varphi(\boldsymbol{r}_1, z_m) + |K_2|^2\, I_\varphi(\boldsymbol{r}_2, z_m) \\ +2\,\mathrm{Re}\left\{K_1^* K_2\, J_{\varphi,\mathrm{eq}}(\boldsymbol{r}_1, \boldsymbol{r}_2, z_m)\exp(-\mathrm{i}2\pi\bar{\nu}\tau)\right\} \tag{22}$$

where I_φ is computed through equation (17) and the equivalent mutual intensity is given by

$$J_{\varphi,\mathrm{eq}}(\boldsymbol{r}_1, \boldsymbol{r}_2, z_m) = J_{xx}(\boldsymbol{r}_1, \boldsymbol{r}_2, z_m)\, C^2 + J_{yy}(\boldsymbol{r}_1, \boldsymbol{r}_2, z_m)\, S^2 + 2\,\mathrm{Re}\left\{J_{xy}(\boldsymbol{r}_1, \boldsymbol{r}_2, z_m)\right\} CS. \tag{23}$$

This result deserves some comment. We have found that the Young interference pattern seen after the polarizer can be cast in the usual form if we introduce a suitable equivalent mutual intensity. The latter is furnished by equation (23), which represents a new type of interference law. Indeed what is going to interfere in equation (23) for $\boldsymbol{r}_1 \neq \boldsymbol{r}_2$ is not a set of optical intensities but of mutual intensities. In other words, the correlation between the fields emerging at $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ from the polarizer is determined by an interplay among the self-correlations of the x- and the y-components and the mixed correlation between them. For example, it may well happen that while J_{xx}, J_{yy} and J_{xy} are all different from zero, the equivalent mutual intensity vanishes. Paradoxically, the overall effect of interference, in such cases, is to destroy correlation between the fields at the pinholes. Needless to say, equation (23) includes equation (17) as a particular case for $\boldsymbol{r}_1 = \boldsymbol{r}_2$.

We finally note that, setting $\varphi = 0$ $(\pi/2)$, $J_{\varphi,\mathrm{eq}}$ becomes equal to J_{xx} (J_{yy}). The diagonal elements of the BCP matrix can then be obtained from intensity measurements (see equation (22)).

The similarity between any element of the BCP matrix and the ordinary mutual intensity allows us to exhaust the propagation problem with a few words. Starting from the knowledge of the BCP matrix at a certain plane $z = z_1$, the evaluation at any successive plane $z = z_2$ $(z_2 > z_1)$ can be made by using, for each matrix element, the formulae holding for the propagation of the mutual intensity. We simply note that, while in certain cases it will be possible to use exact formulae, in most cases the very hypothesis of beam propagation will suggest the use of paraxial propagation formulae. Of course the propagation problem could not be treated using the local polarization matrix only.

4. Generalized Young interference experiments

It is of some interest to extend the previous formalism to treat Young interference when the two pinholes are covered by different anisotropic elements. As we shall see, this gives the key for measuring the off-diagonal element of the BCP matrix. Let $\hat{T}_j$ be the matrix of the anisotropic element placed onto the jth pinhole $(j = 1, 2)$. Further, let $\boldsymbol{E}'(\boldsymbol{r}_j, z_m; t)$ be the electric field at point $\boldsymbol{r}_j$ on the transverse plane immediately after the pinhole j $(j = 1, 2)$. Following Mandel and Wolf [1], we can write

$$\begin{aligned} \boldsymbol{E}'(\boldsymbol{r}_1, z_m; t) &= \boldsymbol{E}(\boldsymbol{r}_1, z_m; t)\, \hat{T}_1 \\ \boldsymbol{E}'(\boldsymbol{r}_2, z_m; t) &= \boldsymbol{E}(\boldsymbol{r}_2, z_m; t)\, \hat{T}_2 \end{aligned} \tag{24}$$

where $\boldsymbol{E}(\boldsymbol{r}_j, z_m; t)$ is the electric field impinging on the Young interferometer at $\boldsymbol{r}_j$. Following analogous lines as for the previous section, the BCP matrix at the point $(\boldsymbol{r}, z_p)$ is given by (see equation (6))

$$\hat{J}(\boldsymbol{r}, \boldsymbol{r}, z_p) = |K_1|^2\, \langle \boldsymbol{E}'^{\dagger}(\boldsymbol{r}_1, z_m; t)\, \boldsymbol{E}'(\boldsymbol{r}_1, z_m; t)\rangle + |K_2|^2\, \langle \boldsymbol{E}'^{\dagger}(\boldsymbol{r}_2, z_m; t)\, \boldsymbol{E}'(\boldsymbol{r}_2, z_m; t)\rangle \\ +2\,\mathrm{Re}\left\{K_1^* K_2 \langle \boldsymbol{E}'^{\dagger}(\boldsymbol{r}_1, z_m; t)\, \boldsymbol{E}'(\boldsymbol{r}_2, z_m; t)\rangle \exp(-\mathrm{i}2\pi\bar{\nu}\tau)\right\}. \tag{25}$$

Taking equations (24) and (6) into account, we obtain

$$\hat{J}(\boldsymbol{r}, \boldsymbol{r}, z_p) = |K_1|^2\, \hat{T}_1^{\dagger}\, \hat{J}(\boldsymbol{r}_1, \boldsymbol{r}_1, z_m)\, \hat{T}_1 + |K_2|^2\, \hat{T}_2^{\dagger}\, \hat{J}(\boldsymbol{r}_2, \boldsymbol{r}_2, z_m)\, \hat{T}_2 \\ +2\,\mathrm{Re}\left\{K_1^* K_2\, \hat{T}_1^{\dagger}\, \hat{J}(\boldsymbol{r}_1, \boldsymbol{r}_2, z_m)\, \hat{T}_2 \exp(-\mathrm{i}2\pi\bar{\nu}\tau)\right\} \tag{26}$$

which represents a generalized interference law. Obviously, equation (26) leads to equation (22) in the particular case when both $\hat{T}_1$ and $\hat{T}_2$ coincide with the matrix of a linear polarizer set at one and the same angle φ.

A suitable choice of $\hat{T}_1$ and $\hat{T}_2$ can lead to selecting the off-diagonal elements of the BCP matrix. As a simple example let us consider the case in which $\hat{T}_1$ and $\hat{T}_2$ are given by

$$\hat{T}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad \hat{T}_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{27}$$

Such matrices describe an x-oriented linear polarizer ($\hat{T}_1$) and a y-oriented linear polarizer followed by a $-\pi/2$ rotator ($\hat{T}_2$). On inserting equation (27) into (26) and taking the trace of both sides, we obtain

$$I(\boldsymbol{r}, z_p) = |K_1|^2\, J_{xx}(\boldsymbol{r}_1, \boldsymbol{r}_1, z_m) + |K_2|^2\, J_{yy}(\boldsymbol{r}_2, \boldsymbol{r}_2, z_m) \\ +2\,\mathrm{Re}\left\{K_1^* K_2\; J_{xy}(\boldsymbol{r}_1, \boldsymbol{r}_2, z_m) \exp(-\mathrm{i}2\pi\bar{\nu}\tau)\right\}. \tag{28}$$

This shows how a modified Young interferometer may be used to measure J_{xy}.

5. Examples

In this section we will show through simple examples how physically different beams can appear identical in a scalar description while their differences are revealed by the BCP matrix. To this end, let us consider three field distributions specified, at the plane $z = 0$, by the following sets of Cartesian components:

$$\begin{cases} E_x^{(1)}(\boldsymbol{r}, 0; t) = a(t)\dfrac{r}{v\sqrt{2}} \exp\left(-\dfrac{r^2}{v^2}\right) \exp(\mathrm{i}\vartheta) \\ E_y^{(1)}(\boldsymbol{r}, 0; t) = b(t)\dfrac{r}{v\sqrt{2}} \exp\left(-\dfrac{r^2}{v^2}\right) \exp(-\mathrm{i}\vartheta) \end{cases} \tag{29}$$

$$\begin{cases} E_x^{(2)}(\boldsymbol{r}, 0; t) = \left[a(t)\exp(\mathrm{i}\vartheta) + b(t)\exp(-\mathrm{i}\vartheta)\right] \dfrac{r}{2v} \exp\left(-\dfrac{r^2}{v^2}\right) \\ E_y^{(2)}(\boldsymbol{r}, 0; t) = \mathrm{i}\left[a(t)\exp(\mathrm{i}\vartheta) - b(t)\exp(-\mathrm{i}\vartheta)\right] \dfrac{r}{2v} \exp\left(-\dfrac{r^2}{v^2}\right) \end{cases} \tag{30}$$

$$\begin{cases} E_x^{(3)}(\boldsymbol{r}, 0; t) = a(t)\dfrac{r}{v} \exp\left(-\dfrac{r^2}{v^2}\right) \cos\vartheta \\ E_y^{(3)}(\boldsymbol{r}, 0; t) = b(t)\dfrac{r}{v} \exp\left(-\dfrac{r^2}{v^2}\right) \sin\vartheta \end{cases} \tag{31}$$

where r and ϑ are the polar coordinates associated with $\boldsymbol{r}$ and v is a positive constant. The functions $a(t)$ and $b(t)$ are uncorrelated quasi-monochromatic signals such that

$$\langle |a|^2 \rangle = \langle |b|^2 \rangle = I_0 \qquad \langle a^* b \rangle = 0 \tag{32}$$

with positive I_0. We shall refer to the field corresponding to equation (29) (equations (30) and (31)) as field 1 (2, 3). Each of the functions appearing in equation (29) will be recognized as a particular Laguerre–Gauss mode with spot size v, taken at its waist [26].

Both modes exhibit a vortex, but the first has charge 1, whereas the second has charge -1 [27]. Field 1 is thus the incoherent superposition of these two modes polarized along the x- and the y-axis, respectively. Similarly, field 2 is obtained through (incoherent) superposition of the same two modes except that they are now circularly polarized with opposite helicity. Finally, field 3 is the incoherent superposition of linearly polarized Laguerre–Gauss modes with angular amplitude modulation.

Using equations (29)–(31) and taking equation (32) into account, the elements of the PCB matrix are easily evaluated for fields 1–3 and turn out to be

$$\begin{cases} J_{xx}^{(1)}(\boldsymbol{r}_1, \boldsymbol{r}_2, 0) = I_0 \dfrac{r_1 r_2}{2v^2} \exp\left[-\dfrac{r_1^2 + r_2^2}{v^2} + \mathrm{i}\,(\vartheta_2 - \vartheta_1)\right] \\ J_{yy}^{(1)}(\boldsymbol{r}_1, \boldsymbol{r}_2, 0) = \left[J_{xx}^{(1)}(\boldsymbol{r}_1, \boldsymbol{r}_2, 0)\right]^* \\ J_{xy}^{(1)}(\boldsymbol{r}_1, \boldsymbol{r}_2, 0) = 0 \end{cases} \tag{33}$$

$$\begin{cases} J_{xx}^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2, 0) = I_0 \dfrac{r_1 r_2}{2v^2} \exp\left(-\dfrac{r_1^2 + r_2^2}{v^2}\right) \cos(\vartheta_2 - \vartheta_1) \\ J_{yy}^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2, 0) = J_{xx}^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2, 0) \\ J_{xy}^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2, 0) = -I_0 \dfrac{r_1 r_2}{2v^2} \exp\left(-\dfrac{r_1^2 + r_2^2}{v^2}\right) \sin(\vartheta_2 - \vartheta_1) \end{cases} \tag{34}$$

$$\begin{cases} J_{xx}^{(3)}(\boldsymbol{r}_1, \boldsymbol{r}_2, 0) = I_0 \dfrac{r_1 r_2}{v^2} \exp\left(-\dfrac{r_1^2 + r_2^2}{v^2}\right) \cos\vartheta_2 \cos\vartheta_1 \\ J_{yy}^{(3)}(\boldsymbol{r}_1, \boldsymbol{r}_2, 0) = I_0 \dfrac{r_1 r_2}{v^2} \exp\left(-\dfrac{r_1^2 + r_2^2}{v^2}\right) \sin\vartheta_2 \sin\vartheta_1 \\ J_{xy}^{(3)}(\boldsymbol{r}_1, \boldsymbol{r}_2, 0) = 0. \end{cases} \tag{35}$$

On inserting equations (33)–(35) into equations (21) and (8) we can evaluate the equivalent mutual intensities and the optical intensity distributions associated with fields 1–3. It is easily seen that the results are the same for all of them. More explicitly, we obtain

$$J_{\mathrm{eq}}^{(j)}(\boldsymbol{r}_1, \boldsymbol{r}_2, 0) = I_0 \frac{r_1 r_2}{v^2} \exp\left(-\frac{r_1^2 + r_2^2}{v^2}\right) \cos(\vartheta_2 - \vartheta_1) \qquad (j = 1, 2, 3) \tag{36}$$

$$I^{(j)}(\boldsymbol{r}, 0) = I_0 \frac{r^2}{v^2} \exp\left(-\frac{2r^2}{v^2}\right) \qquad (j = 1, 2, 3). \tag{37}$$

Using the well known propagation laws of the Laguerre–Gauss modes, the mutual intensities and the optical intensities associated with fields 1–3 could be easily computed at any plane $z =$ constant. Without doing this, we simply remark that the field distributions specified by equations (29)–(31) give rise to shape-invariant beams in the sense of coherence theory [28]. This essentially means that, at any plane $z =$ constant, the mutual intensity remains of the same functional form as across the plane $z = 0$, except for a transverse scale factor and a quadratic phase factor depending on $r_2^2 - r_1^2$. As a consequence, the beams produced by fields 1–3 behave as identical in the realm of scalar coherence theory. In other words, they cannot be distinguished from one another if only intensity and/or scalar coherence measurements are performed (without inserting anisotropic elements on the beam path). In contrast, they will appear different if the elements of their BCP matrices are measured. One might guess that, without introducing the BCP matrix, differences among the three beams would be revealed by the use of the local polarization matrix. This is partly true. As a matter of fact, fields 1, 2 could be distinguished from field 3 because the former are not polarized, while the latter has a variable degree of polarization (as can be easily seen through equations (13) and (33)–(35)). On the other hand, fields 1 and 2 remain indistinguishable unless the BCP matrix description is adopted.

6. Conclusions

For a general optical field, a complete description of the correlations among the electromagnetic vectors requires the use of a tensorial theory of coherence [1]. However, for quasi-monochromatic, beam-like fields, a much simpler, albeit approximate, approach can be based on the BCP matrix. As we saw in this paper, it can capture several significant features of beams that are both partially polarized and partially coherent from the spatial standpoint. In particular, it leads us to discover a peculiar interference law that applies to mutual intensities and not only to optical intensities. Furthermore, it affords a simple tool for describing Young's interferometers modified through insertion of anisotropic elements. Propagation problems can be easily tackled since each BCP matrix element behaves as a mutual intensity. In this connection, it should be remarked that, differently from the examples dealt with in this paper, in which we deliberately chose shape-invariant fields, the propagation process can lead to drastic changes of the BCP matrix elements.

For the sake of brevity and simplicity, we limited ourselves to quasi-monochromatic fields, further assuming all delays of interest to be small compared to the coherence time. The removal of such limitations as well as the introduction of suitable cross-spectral densities should not offer any difficulties.

The scalar theory of coherence applied to beam-like fields has led to the discovery of a host of interesting models and phenomena. Let us simply quote the Schell–Gauss beams [1, 29] and the occurrence of the Wolf effect [1, 30]; similarly, phenomena describable by the ordinary polarization matrix fill volumes [2, 3]. The BCP matrix, which is the adaption of the tensorial Wolf theory [1] to beam-like fields, establishes a link between these two topics. Therefore, it should prove to be useful in several applications.

Nonetheless, the application of this type of formalism to radiation fields not describable by the beam model requires some caution. The need for a complete vectorial treatment for certain cases, such as black-body radiation [31], has been demonstrated (see for example [1, 32]).

References

[1] Mandel L and Wolf E 1995 *Optical Coherence and Quantum Optics* (Cambridge: Cambridge University Press)
[2] Shurcliff W A 1962 *Polarization Light* (Cambridge, MA: Harvard University Press)
[3] Azzam R M A and Bashara N M 1986 *Ellipsometry and Polarized Light* (Amsterdam: North-Holland)
[4] Pancharatnam S 1956 *Proc. Ind. Acad. Sci.* A **44** 247
[5] Berry M V 1984 *Proc. R. Soc.* A **392** 45
[6] Schmitzer H, Klein S and Dultz W 1993 *Phys. Rev. Lett.* **71** 1530
[7] Aspect A, Dalibord J and Roger G 1982 *Phys. Rev. Lett.* **49** 1804
[8] Simon P and Mukunda N 1990 *Phys. Lett.* A **140** 165
[9] Barakat R 1996 *Opt. Commun.* **123** 443
[10] Tudor T 1997 *J. Opt. Soc. Am.* A **14** 2013
[11] Born M and Wolf E 1980 *Principles of Optics* (Oxford: Pergamon)
[12] Erdogan T and Hall D G 1990 *J. Appl. Phys.* **68** 1435
[13] Tidwell S C, Ford D H and Kimura W D 1992 *Opt. Eng.* **31** 1527
[14] Greene P L and Hall D G 1996 *J. Opt. Soc. Am.* A **13** 962

[15] Wolf E 1954 *Nuovo Cimento* **12** 884
[16] Gori F 1998 *Opt. Lett.* **23** 241
[17] James D F V 1994 *J. Opt. Soc. Am.* A **11** 1641
[18] Martínez-Herrero R, Mejías P M and Movilla J M 1997 *Opt. Lett.* **22** 206
[19] Gori F 1980 *Opt. Commun.* **34** 301
[20] Starikov A and Wolf E 1982 *J. Opt. Soc. Am.* **72** 923
[21] Gase R 1991 *J. Mod. Opt.* **38** 1107
[22] Simon R and Mukunda N 1993 *J. Opt. Soc. Am.* A **10** 95
[23] Friberg A T, Tervonen E and Turunen J 1994 *J. Opt. Soc. Am.* A **11** 1818
[24] Jones R C 1941 *J. Opt. Soc. Am.* **31** 488
[25] James D F V and Wolf E 1998 *Opt. Commun.* **145** 1
[26] Siegman A E 1986 *Lasers* (Mill Valley, CA: University Science Books)
[27] Indebetouw G 1993 *J. Mod. Opt.* **40** 73
[28] Gori F 1984 *Coherence and Quantum Optics* ed L Mandel and E Wolf (New York: Plenum) p 363
[29] Foley J T and Zubairy M S 1978 *Opt. Commun.* **26** 297
Saleh B E A 1979 *Opt. Commun.* **30** 135
Farina J D, Narducci L M and Collett E 1980 *Opt. Commun.* **32** 203
Carter W H and Wolf E 1981 *Opt. Acta* **28** 227
Carter W H and Wolf E 1981 *Opt. Acta* **28** 245
Indebetouw G 1989 *J. Mod. Opt.* **36** 251
Sundar K, Mukunda N and Simon R 1995 *J. Opt. Soc. Am.* A **12** 560
[30] Wolf E and James D F V 1996 *Rep. Prog. Phys.* **59** 771
Wolf E 1986 *Phys. Rev. Lett.* **56** 1370
Dačić Z and Wolf E 1988 *J. Opt. Soc. Am.* A **5** 1118
Kandpal H C, Vaishya J S and Joshi K C 1989 *Opt. Commun.* **73** 169
Gori F, Marcopoli G-L and Santarsiero M 1990 *Opt. Commun.* **81** 123
Palma C and Cincotti G 1997 *J. Opt. Soc. Am.* A **14** 1885
[31] Jaiswal A K, Agrawal G P and Metha C L 1973 *Nuovo Cimento* **15b** 295
[32] James D F V 1994 *Opt. Commun.* **109** 209