JOURNAL OF MODERN OPTICS, 1995, VOL. 42, NO. 9, 1875–1899

Polarization effects in heterodyne interferometry

J. M. DE FREITAS AND M. A. PLAYER
Engineering Physics Group, Department of Engineering,
University of Aberdeen, Aberdeen AB9 2UE, Scotland

(*Received 18 November 1994*)

Abstract. A detailed analysis of the polarization effects which lead to nonlinearity in the non-ideal optical heterodyne interferometer is presented. Extensive use is made of the coherency matrix representation by setting up a 'cross-coherency matrix' representation. A generalized treatment of periodic phase errors (nonlinearity) is then presented. Individual contributions to the nonlinearity have been characterized as either 'independent' or 'dependent' phase errors. In the single-pass plane-mirror heterodyne system, to which the approach is applied, phase errors for rotational misalignment of the nominally orthogonal linearly polarized input states, beam splitter leakage, non-orthogonality, ellipticity and the effect of misaligned polarizer–mixer are explicitly considered. The latter effect is found to produce nonlinearity only when in combination with any one of the first three and is therefore a dependent phase error. The nonlinearity arising from ellipticity is identical with that from rotational misalignment except that it has an offset. Rotational misalignment and ellipticity produce nonlinearity at the second harmonic and are second order for practical set-ups. It is also found that combinations of positive (anticlockwise) and negative (clockwise) angular misalignments of the azimuth of the states, non-orthogonality and misorientations of the polarizer–mixer, all relative to the polarizing beam splitter axes, lead to different peak-to-peak nonlinearities in the given system.

1. Introduction

In the polarization coded laser heterodyne interferometer, the input beam consists of two angular frequencies ω and $\omega + \Delta\omega$, each of which is ideally associated exclusively with the transverse electric (TE) or transverse magnetic (TM) polarization state. Ideally, the polarizing beam splitter (PBS) separates the two frequencies (polarization states) so that the TE and the TM modes travel different optical paths in pure form. The difference between the two paths is the optical path-length change and is carried by the heterodyne signal as a phase lead or lag δ, with respect to a reference signal. In an ideal interferometer (the relative) signal I received at the measurement photodetector is

$$I = 1 + \cos(\Delta\omega t + \delta), \tag{1}$$

where $\Delta\omega$ is the beat frequency which results from the mixing of the TE and TM modes via a polarizer–mixer placed in front of the photodetector. The reference signal I_R is obtained by sampling the input beam with an ordinary beam splitter and then 'beating' the two modes in the reference detector, to obtain $I_R = 1 + \cos(\Delta\omega t)$ (figure 1).

Most commonly, the heterodyne beam is derived from the spectral splitting that occurs when a laser medium is placed in a magnetic field (Zeeman splitting); with

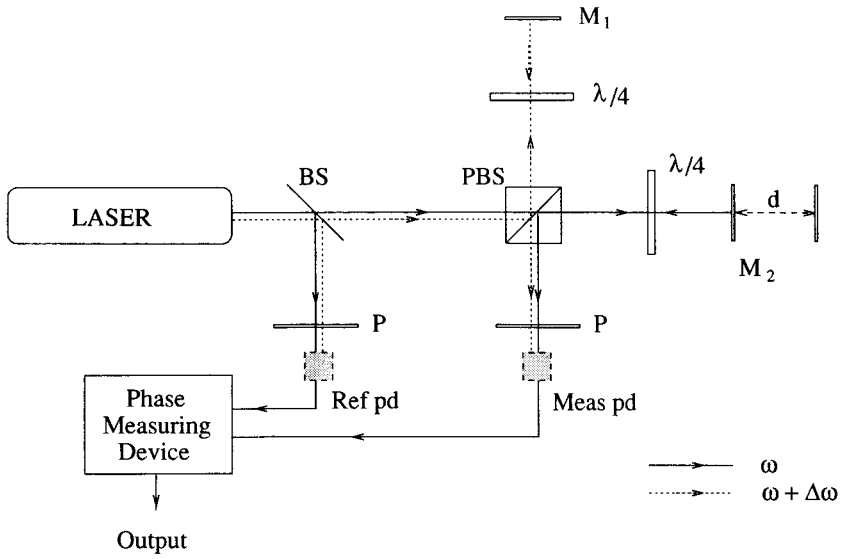

Figure 1. Layout of a polarization coded heterodyne interferometer: P, polarizer–mixer; PBS, polarizing beam splitter; M, mirror; pd, photodetector. Phase delay $\delta = 4\pi d/\lambda$.

suitably positioned half- and quarter-wave plates, nominally orthogonal linearly polarized states are obtained. The considerations of this paper also apply to heterodyne systems using acousto-optic modulators, provided that beam separation is achieved through polarization coding. If the polarization modes associated with the two optical frequencies are not ideal (i.e. they may be elliptical, non-orthogonal, rotationally misaligned, or a combination of these), then strong and weak components of both frequencies will propagate in both arms, leading to frequency mixing of the wrong components. This effect ultimately results in a small periodic error (nonlinearity) which limits high precision measurements of optical path lengths [1–9].

It was first suggested by Quenelle [1] that, in practice, PBSs are imperfect and allow through to each arm a small amount of the polarization state intended for the other arm, and that elliptical states may cause nonlinearity. Sutton [2] carried out experimental investigations and confirmed the predictions of Quenelle, finding about 5 nm peak-to-peak variations in nonlinearity. A periodicity of one cycle per 2π change in optical path length was immediately evident; however, Sutton also observed a weaker second-harmonic component of the nonlinearity but came to no firm conclusion as to its origin. Bobroff [3] also made experimental investigations which confirmed the existence of the periodic errors of Quenelle [1]. In a related paper, Rosenbluth and Bobroff [4] carried out a phasor analysis of the optical sources of nonlinearity in heterodyne interferometers. Although they pointed out that a purely scalar treatment was inadequate for a full quantitative treatment of nonlinearity, they concluded that 'frequency mixing is not a major cause of nonlinearity unless the mixing is asymmetric between the two arms of the interferometer' [4]. Moreover, they obtained the surprising result that many commonly cited causes of frequency mixing were not in themselves sufficient to produce first-order non-linearity. One example cited by these workers was the effect of rotational misalignment of the input polarization states on nonlinearity; for a

symmetric interferometer no nonlinearity is predicted. More recently, using a coherency matrix analysis [8], the present authors have shown that rotational misalignment in the absence of any 'asymmetry' can produce appreciable nonlinearity and does so at the second harmonic (i.e. two cycles per 2π change in optical phase), although in principle it is a second-order effect for practical interferometer layouts. Experimental results pertaining to the effects of rotational misalignment of the input states were provided by Picoto and Sacconi [7]. Other investigations were carried out by Xie and Wu [6, 10] who provided a matrix analysis and identified the source of non-orthogonality and ellipticity of the laser output as the birefringence and dichroism loss ratio of the laser cavity.

A significant approach to the understanding of phase-dependent polarization effects was introduced by Berry [11] and later shown to be related to the Pancharatnam phase [12–14]. Geometrical (topological) phase effects occur when the state of polarization evolves along a closed path on the Poincaré sphere. However, in classical interferometry treatment of polarization effects through the use of Jones and coherency matrices will also yield the same results and is therefore adequate for a detailed study of polarization effects in heterodyne interferometry.

The aim of this paper is to provide a detailed model of polarization effects in the heterodyne interferometer, expanding the coherency matrix treatment given initially in [8]. We specifically show that some effects, although second order in nature, can give rise to substantial nonlinearity when combined with other first-order or second-order effects. The model assumes that the differential heterodyne optical path error (due to separation of the heterodyne spectral lines) is very much smaller than the optical path length. We consider only those effects related to the polarization phenomena in a non-ideal interferometer. In a sense these effects are geometrical, but discussion will relate only to the geometry of the polarized states, as opposed to the interferometer geometrical layout effects such as Abbé errors. No effort is made to cast this work into the geometrical (Berry's or Pancharatnam's) phase interpretation, as has been done for the heterodyne fibre optic fed interferometer by Bergamin *et al.* [9].

2. Theoretical background

This section develops the basis of the Jones calculus used in the analysis. This is then converted into the coherency matrix representation, and the appropriate phase extracted. A key feature of the coherency matrix representation, is that it can represent partially polarized light, with the ability to assess the noise properties of the optical system (for example [15]), although this is beyond the scope of the present paper. The Jones representation limits consideration to non-depolarizing systems.

2.1. *Approximations and the Jones matrices for an interferometer*

In a general two-beam interferometer, the Jones matrix for the 'reflection' (TE or reference) and 'transmission' (TM or measurement) arms of the interferometer may be written, following Fymat [15, 16], as

$$\mathbf{D}^{(r)} = [d_{ij}^{(r)}] \ (i,j = 1,2), \tag{2}$$

$$\mathbf{D}^{(t)} = [d_{ij}^{(t)}] \ (i,j = 1,2), \tag{3}$$

where $d_{ij}^{(r)}$ $(i,j = 1,2)$ and $d_{ij}^{(t)}$ are reflection and transmission coefficients respectively for the interferometer as a whole. Thus, for example, $\mathbf{D}^{(r)}$ may represent reflection

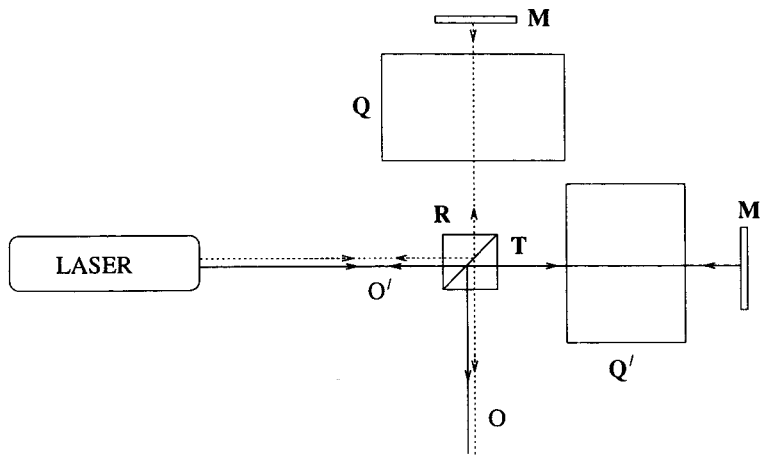

Figure 2. Jones matrix description of the general interferometer. Output may be at **O** or **O'** depending on **Q** and **Q'**.

R in the cube PBS, followed by transmission **Q** through optical element(s) in the reference arm, reflection **M** in the mirror and finally a retrace of the beam path, where **R**, **Q** and **M** are the Jones matrices for appropriate optical components. Then, $\mathbf{D}^{(r)} = \mathbf{RQMQR}$; similarly one may write the Jones representation for transmission as $\mathbf{D}^{(t)} = \mathbf{TQ'MQ'T}$, where **T** and **Q'** are the Jones matrices for transmission through the beam splitter and the optical elements in the transmission arm respectively (figure 2).

Let δ_t and δ_r be the optical phase changes that occur in the measurement (TM) arm (subscript t) and reference arm (TE) (subscript r) respectively. Using superscripts α and β on the δ values to denote the phase change associated with the incident field $\mathbf{E}^{(\alpha,\mathrm{i})}$ at angular frequency ω, and $\mathbf{E}^{(\beta,\mathrm{i})}$ at frequency $\omega + \Delta\omega$, then the Jones matrix **D** for the whole interferometer can be written for each field as

$$\begin{aligned} \mathbf{D}^{(\alpha)} &= [\mathbf{D}^{(r)} + \mathbf{D}^{(t)} \exp(-\mathrm{i}\delta^{(\alpha)})] \exp(-\mathrm{i}\delta_r^{(\alpha)}), \\ \mathbf{D}^{(\beta)} &= [\mathbf{D}^{(r)} + \mathbf{D}^{(t)} \exp(-\mathrm{i}\delta^{(\beta)})] \exp(-\mathrm{i}\delta_r^{(\beta)}), \end{aligned} \tag{4}$$

where $\delta^{(u)} = \delta_t^{(u)} - \delta_r^{(u)}$ (with $u = \alpha, \beta$). Although the optical path lengths traversed by the fields in each arm may be identical, they nonetheless yield different optical phases because of their slightly differing wavelengths $\lambda^{(u)}$ (with $u = \alpha, \beta$). In general, one has

$$\delta_v^{(u)} = \frac{4\pi}{\lambda^{(u)}} \int_0^{L_v} n_v(\lambda^{(u)}, l)\, \mathrm{d}l \quad (u = \alpha, \beta; v = t, r), \tag{5}$$

where $n_v(\lambda^{(u)}, l)$ is the refractive index along the beam path taken by wave field $\mathbf{E}^{(u,\mathrm{i})}$ ($u = \alpha, \beta$) and L_v is the path length along the beam path.

The difference between the magnitudes of $\delta^{(\alpha)}$ and $\delta^{(\beta)}$ (which can be considered as a measure of the heterodyne 'beat' error) is usually small, and to a first-order approximation is (appendix A)

$$\delta^{(\alpha)} - \delta^{(\beta)} \approx \frac{\Delta\lambda}{\lambda}(\delta_{\mathrm{g},t} + \delta_0), \tag{6}$$

where $\Delta\lambda$ is the difference between the two input wavelengths, δ_0 is an offset phase

corresponding to the interferometer path difference and $\delta_{g,t}$ is a group optical phase change corresponding to dispersive material in (say) the measurement (transmission) arm and is given by,

$$\delta_{g,t} = \frac{4\pi}{\lambda} \int_{L_t}^{L_t + d} n_{g,t}(\lambda, l)\, \mathrm{d}l, \tag{7}$$

where $n_{g,t}$ is the group refractive index of the dispersive material, of length d. Thus, by ensuring that $\Delta\lambda/\lambda$ is sufficiently small, which is commonly the case, an approximation can be made for the Jones matrices such that both $\mathbf{D}^{(\alpha)}$ and $\mathbf{D}^{(\beta)}$ can be replaced by a single Jones representation

$$\mathbf{D} = [\mathbf{D}^{(r)} + \mathbf{D}^{(t)} \exp(-\mathrm{i}\delta)] \exp(-\mathrm{i}\delta_r), \tag{8}$$

Note also that, in the process of finding the coherency matrix (see section below), $\mathbf{D}$ will be multiplied by its Hermitian conjugate and, as such, the free multiplier $\exp(-\mathrm{i}\delta_r)$ will always be cancelled and can be neglected leading to the expression in [8]:

$$\mathbf{D} = \mathbf{D}^{(r)} + \mathbf{D}^{(t)} \exp(-\mathrm{i}\delta). \tag{9}$$

The wave-field $\mathbf{E}$ returning to the measurement photodetector is then

$$\mathbf{E} = \mathbf{PD}(\mathbf{E}^{(\alpha,\mathrm{i})} + \mathbf{E}^{(\beta,\mathrm{i})}), \tag{10}$$

where $\mathbf{P}$ is the Jones matrix for the polarizer–mixer, which is intended to mix the polarized components in optimum proportions. The input wave fields are

$$\mathbf{E}^{(\alpha,\mathrm{i})} = \begin{bmatrix} E_x^{(\alpha,\mathrm{i})} \\ E_y^{(\alpha,\mathrm{i})} \end{bmatrix} \exp[-\mathrm{i}(\omega + \Delta\omega)t], \tag{11}$$

$$\mathbf{E}^{(\beta,\mathrm{i})} = \begin{bmatrix} E_x^{(\beta,\mathrm{i})} \\ E_y^{(\beta,\mathrm{i})} \end{bmatrix} \exp(-\mathrm{i}\omega t), \tag{12}$$

where $E_x^{(u,\mathrm{i})}$ and $E_y^{(u,\mathrm{i})}$ $(u = \alpha, \beta)$ are incident orthogonal components (superscript i) of the amplitude of the electromagnetic fields, and t denotes time.

2.2. *Coherency matrix representation*

The coherency matrix $\mathbf{J}$ describing interference and coherence concepts in terms of the wave-field vector $\mathbf{E}$ (ω, t), is given by [17]

$$\mathbf{J} = \langle \mathbf{E} \times \mathbf{E}^{\dagger} \rangle = \begin{bmatrix} \langle E_x E_x^{*} \rangle & \langle E_x E_y^{*} \rangle \\ \langle E_y E_x^{*} \rangle & \langle E_y E_y^{*} \rangle \end{bmatrix}, \tag{13}$$

where the matrix elements are denoted by $J_{ij} = \langle E_i E_j^{*} \rangle$ $(i, j = x, y)$, $\times$ stands for the Kronecker product, the dagger denotes Hermitian conjugation, the angular brackets represent infinite time averaging, and the asterisk stands for complex conjugation. It is readily seen that the average intensity of the beam is given by $I = J_{xx} + J_{yy} = \mathrm{Tr}(\mathbf{J})$.

The coherency matrix describing the interaction of two wave fields $\mathbf{E}^{(\alpha)}(\omega, t)$ and $\mathbf{E}^{(\beta)}$ $(\omega + \Delta\omega, t)$, with different frequencies will be denoted by $\mathbf{J}^{(uv)}$ $(u, v = \alpha, \beta)$. Note that there are four possible matrix representations describing, firstly, interactions between $\mathbf{E}^{(\alpha)}$ and $\mathbf{E}^{(\beta)}$ with themselves (i.e. 'self-mixing'), which are

Table 1. Phasors associated with **J**.

Coherency matrix	$M_{\Delta\omega}(t)$
$\mathbf{J}^{(\alpha\beta,\,i)}$	$\exp(-i\Delta\omega t)$
$\mathbf{J}^{(\alpha\alpha,\,i)}$	1
$\mathbf{J}^{(\beta\beta,\,i)}$	1
$\mathbf{J}^{(\beta\alpha,\,i)}$	$\exp(i\Delta\omega t)$

termed *homodyne* interference, and 'cross-mixing' between $\mathbf{E}^{(\alpha)}$ and $\mathbf{E}^{(\beta)}$, referred to as *heterodyne* mixing. This can be written in a straightforward notation as

$$\mathbf{J}^{(uv)} = \langle \mathbf{E}^{(u)} \times \mathbf{E}^{(v)\dagger} \rangle \quad (u, v = \alpha, \beta)$$

$$= \begin{bmatrix} \langle E_x^{(u)} E_x^{(v)*} \rangle & \langle E_x^{(u)} E_y^{(v)*} \rangle \\ \langle E_y^{(u)} E_x^{(v)*} \rangle & \langle E_y^{(u)} E_y^{(v)*} \rangle \end{bmatrix} M_{\Delta\omega}^{uv}(t), \tag{14}$$

where $M_{\Delta\omega}^{(uv)}(t)$ is the characteristic function or phasor associated with the matrix; neglecting noise effects, $M_{\Delta\omega}^{(uv)} = \exp(\pm i\overline{\Delta\omega}t)$ when $u \neq v$ and is unitary when $u = v$, where $\overline{\Delta\omega}$ is the average heterodyne beat frequency (see appendix B and table 1). In the rest of the paper we shall use $\Delta\omega$ instead of $\overline{\Delta\omega}$.

The matrix elements of $\mathbf{J}^{(uv)}$ will be denoted by $J_{ij}^{(uv)} = \langle E_i^{(u)} E_j^{(uv)*} \rangle$ $(i, j = x, y)$ and $(u, v = \alpha, \beta)$. Although for the homodyne case $J_{xy}^{(uv)} = J_{yx}^{(uv)*}$, this is generally not the case for $J_{xy}^{(uv)}$, $u \neq v$. For example, $J_{xy}^{(\alpha\beta)} = \langle E_x^{(\alpha)} E_y^{(\beta)*} \rangle$ is clearly not equal to $J_{yx}^{(\alpha\beta)*} = \langle E_y^{(\alpha)*} E_x^{(\beta)} \rangle$, but it should be obvious that $J_{ij}^{(uv)} = J_{ji}^{(vu)*}$. The following results may therefore be stated.

(*a*) *Given partially polarized waves* $\mathbf{E}^{(u)}$ and $\mathbf{E}^{(v)}$ *with orthogonal components* $E_x^{(u)*}, E_y^{(u)*}$ *and* $E_x^{(v)*}, E_y^{(v)*}$ *respectively, and defining the cross-coherency element* $J_{ij}^{(uv)} = \langle E_i^{(u)} E_j^{(v*)} \rangle$ $(i, j = x, y)$ then

$$J_{ij}^{(uv)} = J_{ji}^{(vu)*}. \tag{15}$$

Provided that the cross-coherency matrix is defined as in equation (14), we then have from (*a*) the following.

(*b*) *For cross-coherency matrices,* $\mathbf{J}^{(uv)} = \mathbf{J}^{(vu)\dagger}$.

The interpretation of $J_{ij}^{(uv)}$ is also straightforward; the matrix element represents the coherence between the i coordinate component of wave field u, and the j coordinate component of wave field v.

Thus the coherency matrix **J** of the output wave field of the heterodyne system is, using (10) in equation (13),

$$\mathbf{J} = \mathbf{P}\mathbf{D}\mathbf{J}^{(\alpha\beta,\,i))}\mathbf{D}^{\dagger}\mathbf{P}^{\dagger} + \mathbf{P}\mathbf{D}\mathbf{J}^{(\alpha\alpha,\,i)}\mathbf{D}^{\dagger}\mathbf{P}^{\dagger} + \mathbf{P}\mathbf{D}\mathbf{J}^{(\beta\alpha,\,i))}\mathbf{D}^{\dagger}\mathbf{P}^{\dagger} + \mathbf{P}\mathbf{D}\mathbf{J}^{(\beta\beta,\,i)}\mathbf{D}^{\dagger}\mathbf{P}^{\dagger}. \tag{16}$$

The first and third terms of the above equation describe optical mixing and other influences of the interferometer optics on the partial coherence properties of the two initial (heterodyne) polarization states, $\mathbf{E}^{(\alpha,\,i)}$ and $\mathbf{E}^{(\beta,\,i)}$. The second and fourth terms describe the effects of the interferometer optics on the partial self-coherency properties of similar initial (homodyne) states. The description of reflectivity and transmissivity of the interferometer is contained in the Jones matrix **D** (equation (9)). A similar treatment of the cross-coherency vector has been carried out in appendix

D, so that the corresponding analysis in terms of the Mueller representation and Stokes vectors is easily accessible.

2.2.1. *Intensity*

The intensity I observed at the photodetector is

$$I = \mathrm{Tr}(\mathbf{PDJ}^{(\alpha\beta,\,\mathrm{i})}\mathbf{D}^{\dagger}\mathbf{P}^{\dagger}) + \mathrm{Tr}(\mathbf{PDJ}^{(\alpha\alpha,\,\mathrm{i})}\mathbf{D}^{\dagger}\mathbf{P}^{\dagger}) + \mathrm{Tr}(\mathbf{PDJ}^{(\beta\alpha,\,\mathrm{i})}\mathbf{D}^{\dagger}\mathbf{P}^{\dagger}) + \mathrm{Tr}(\mathbf{PDJ}^{(\beta\beta,\,\mathrm{i})}\mathbf{D}^{\dagger}\mathbf{P}^{\dagger}). \tag{17}$$

It can be shown (appendix C) that, for the matrix $\mathbf{J}^{(uv)}$ $(u, v = \alpha, \beta)$,

$$\mathrm{Tr}(\mathbf{PDJ}^{(uv)}\mathbf{D}^{\dagger}\mathbf{P}^{\dagger}) = [\gamma_0^{(uv)} + \gamma_1^{(uv)}\exp(-\mathrm{i}\delta) + \gamma_2^{(uv)}\exp(\mathrm{i}\delta)]\, M_{\Delta\omega}^{(uv)}(t), \tag{18}$$

where $M_{\Delta\omega}^{(uv)}$ is the characteristic function or phasor associated with the coherence states (uv) (table 1). Note that only the cross-coherency matrices (i.e. those with (α, β) and (β, α)) have a time-dependent characteristic function.

Now using the result (b) above in equation (18), one finds that

$$\gamma_0^{(uv)} = \gamma_0^{(vu)*}, \tag{19}$$

$$\gamma_1^{(uv)} = \gamma_2^{(vu)*}, \tag{20}$$

$$M_{\Delta\omega}^{(uv)} = M_{\Delta\omega}^{(vu)*}, \tag{21}$$

which allows equation (18) to be written in the generalized form for the heterodyne interferometer as

$$I = K_0 + 2\mathscr{R}\{K_\delta\exp(-\mathrm{i}\delta) + K_{\Delta\omega}\exp(-\mathrm{i}\Delta\omega t) + K_{\Delta\omega+\delta}\exp[-\mathrm{i}(\Delta\omega t + \delta)] + K_{\Delta\omega-\delta}\exp[-\mathrm{i}(\Delta\omega t - \delta)]\}, \tag{22}$$

where $\mathscr{R}\{\}$ means the real part of $\{\}$, $K_0 = \gamma_0^{(\alpha\alpha)} + \gamma_0^{(\beta\beta)}$, $K_\delta = \gamma_1^{(\alpha\alpha)} + \gamma_1^{(\beta\beta)}$, $K_{\Delta\omega} = \gamma_0^{(\alpha\beta)}$, $K_{\Delta\omega+\delta} = \gamma_1^{(\alpha\beta)}$, $K_{\Delta\omega-\delta} = \gamma_2^{(\alpha\beta)}$, and the K values in general are complex (see also appendix C). The term in $\exp(-\mathrm{i}\delta)$ results from homodyne interference between components at ω traversing the TE and TM arms, plus similar effects involving components at $\omega + \Delta\omega$. The term in $\exp(-\mathrm{i}\Delta\omega t)$ (usually considered to be the main 'nuisance term') results from interference of $\omega + \Delta\omega$ with ω components in each arm separately. The term in $\exp[-\mathrm{i}(\Delta\omega t + \delta)]$, the main photodetector signal term, arises from interference of the main $\omega + \Delta\omega$ component traversing the TM arm with the ω component traversing the TE arm. The final term in $\exp[-\mathrm{i}(\Delta\omega t - \delta)]$ results from ω components in the TM arm (from, in general, both leakage and misalignment effects) interfering with $\omega + \Delta\omega$ components in the TE arm. The parameter δ in the phasor terms is the phase change introduced by the Jones matrix $\mathbf{D}$.

2.2.2. *Phase errors*

In this section, general formulae are given for the measured phase ϕ and the resulting phase error δ_ϵ for cases when the K in equation (22) are real and complex respectively. Note that the K are real for linear states only, and complex for elliptically polarized light.

(a) *Real K; linear states.* It is usual in heterodyne interferometry to pass the measurement photodetector signal through a high-pass filter to remove low-frequency components lower than $\Delta\omega$. In so doing we arrive at the a.c. measurement signal $I_{\mathrm{A.C.}}$. Since the K in equation (22) are now real then

$$I_{\text{A.C.}} = K_{\Delta\omega}\cos(\Delta\omega t) + K_{\Delta\omega+\delta}\cos(\Delta\omega t + \delta) + K_{\Delta\omega-\delta}\cos(\Delta\omega t - \delta). \quad (23)$$

We reiterate here that the main photodetector signal is represented by the term in $\cos(\Delta\omega t + \delta)$, while the first and last terms introduce phase errors into the measurement system. $I_{\text{A.C.}}$ is therefore suitably rewritten as

$$I_{\text{A.C.}} = \mathscr{A}\cos(\Delta\omega t + \phi) \quad (24)$$

where $\mathscr{A}$ is the measurement signal amplitude and ϕ is the measured phase and is given by

$$\phi = \tan^{-1}\left(\frac{K_-\tan\delta}{K_{\Delta\omega}\sec\delta + K_+}\right), \quad (25)$$

where $K_+ = K_{\Delta\omega+\delta} + K_{\Delta\omega-\delta}$ and $K_- = K_{\Delta\omega+\delta} - K_{\Delta\omega-\delta}$. The square of the amplitude $\mathscr{A}$ is given by the sum of the squares of the numerator and denominator within the large parentheses in equation (25). The definition of the phase error δ_ϵ is somewhat arbitrary; however, retaining the framework of [8], we have

$$\delta_\epsilon = \phi - \delta. \quad (26)$$

It is a straightforward matter to show that the resulting phase error (nonlinearity) is·

$$\delta_\epsilon = \tan^{-1}\left(\frac{(K_- - K_+ - K_{\Delta\omega}\sec\delta)\tan\delta}{K_{\Delta\omega}\sec\delta + K_+ + K_-\tan^2\delta}\right). \quad (27)$$

(*b*) *Complex K; elliptical states.* In the most general case (i.e. where the K are complex), one may write $K_{\Delta\omega} = k_{\Delta\omega}\exp(-i\delta_{\Delta\omega})$, $K_{\Delta\omega+\delta} = k_{\Delta\omega+}\exp(-i\delta_{\Delta\omega+})$, $K_{\Delta\omega-\delta} = k_{\Delta\omega-}\exp(-i\delta_{\Delta\omega-})$. The a.c. part of equation (22) becomes

$$I_{\text{A.C.}} = k_{\Delta\omega}\cos(\Delta\omega t + \delta_{\Delta\omega}) + k_{\Delta\omega+}\cos(\Delta\omega t + \delta + \delta_{\Delta\omega+}) + k_{\Delta\omega-}\cos(\Delta\omega t - \delta + \delta_{\Delta\omega-}). \quad (28)$$

Such a system is easily reduced to give the generalized phase error as

$$\delta_\epsilon = \tan^{-1}\left(\frac{k_{\Delta\omega}\sin(\delta_{\Delta\omega} - \delta) + k_{\Delta\omega+}\sin\delta_{\Delta\omega+} + k_{\Delta\omega-}\sin(\delta_{\Delta\omega-} - 2\delta)}{k_{\Delta\omega}\cos(\delta_{\Delta\omega} - \delta) + k_{\Delta\omega+}\cos\delta_{\Delta\omega+} + k_{\Delta\omega-}\cos(\delta_{\Delta\omega-} - 2\delta)}\right). \quad (29)$$

It is clear from equation (29) that $k_{\Delta\omega}$ introduces nonlinearity at the fundamental, while $k_{\Delta\omega-}$ introduces nonlinearity at the second harmonic. For a proper interferometer set-up these two parameters will be small in comparison with $k_{\Delta\omega+}$, the main signal amplitude. The phase $\delta_{\Delta\omega+}$ of $K_{\Delta\omega+\delta}$ represents just a constant offset between ϕ and δ.

The opportunity is taken here to introduce the concept of *independent* and *dependent* phase errors. Returning to equation (25), it is clear that $\phi = \delta$ and $\delta_\epsilon = 0$ when $K_{\Delta\omega} = K_{\Delta\omega-\delta} = 0$, that is the ideal case. However, if for a specific non-ideal effect, for example rotational misalignment in the plane-mirror heterodyne interferometer, $K_{\Delta\omega} \neq K_{\Delta\omega-\delta}$ and both are not simultaneously zero, then the effect is independent and produces *independent* phase errors. This is in contrast with *dependent* phase errors where, by itself, the non-ideal effect does not independently produce phase errors but only in conjunction

with an independent effect. Moreover, two dependent phase effects cannot produce nonlinearity. These two types of phase error will be discussed in the next section in relation to the plane mirror system.

3. The plane-mirror interferometer

The simplest interferometer layout possible is one in which $\mathbf{Q} = \mathbf{Q}' = \mathbf{1}$, the unit matrix (figure 2). In such a case, the output returns in the same direction as the input. Moreover, the matrix $\mathbf{D}$ is diagonal. In more complex interferometer configurations involving quarter-wave plates and retroreflectors (cube corners), $\mathbf{D}$ is not diagonal since these optical components induce rotations in one or both of the linearly polarized states [18, 19], or by improper alignment of the wave plates. Although this particular study is limited to the plane-mirror configuration, it is believed that it is a fairly good representation of most single-pass interferometer configurations. First, a detailed description is given for the output of the interferometer observed at the measurement photodetector, followed by the treatment of special cases affecting nonlinearity in the interferometer.

3.1. *Parametrization*

3.1.1. *Generalized Jones input states.*

The axes of the PBS are taken as reference axes for the nominally polarized heterodyne input $\mathbf{E}^{(\alpha,\mathrm{i})}$ and $\mathbf{E}^{(\beta,\mathrm{i})}$. These may be written in the generalised format of Shurcliff [20] as

$$\mathbf{E}^{(\alpha,\mathrm{i})} = \begin{pmatrix} \cos R \exp(-\tfrac{1}{2}\mathrm{i}\mu_1) \\ \sin R \exp(\tfrac{1}{2}\mathrm{i}\mu_1) \end{pmatrix} \exp[-\mathrm{i}(\omega + \Delta\omega)t], \tag{30}$$

$$\mathbf{E}^{(\beta,\mathrm{i})} = \begin{pmatrix} -\sin(R + \Delta R) \exp(-\tfrac{1}{2}\mathrm{i}\mu_2) \\ \cos(R + \Delta R) \exp(\tfrac{1}{2}\mathrm{i}\mu_2) \end{pmatrix} \exp(-\mathrm{i}\omega t) \tag{31}$$

where R is the angle between the diagonal of the rectangle circumscribing the ellipse and the x axis (figure 3). R is related to the azimuth α of the major axis of the elliptical

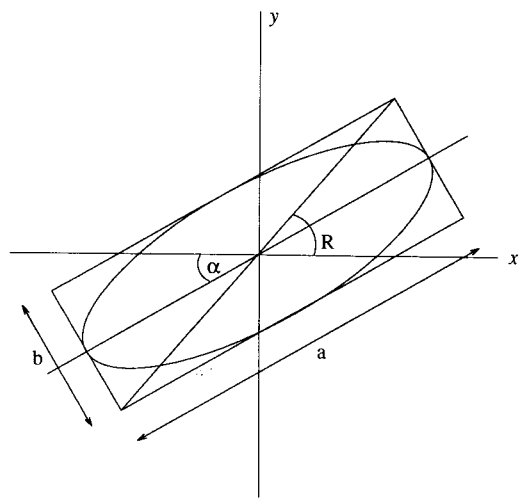

Figure 3. Description of the ellipse and definition of parameters. x and y are the axes of the beam splitter.

state through

$$\tan R = \left(\frac{1 - \cos(2\beta)\cos(2\alpha)}{1 + \cos(2\beta)\cos(2\alpha)}\right)^{1/2}, \tag{32}$$

where $\tan\beta = b/a$ is the ellipticity of the state. For a given ellipticity, if the azimuth α changes to $\alpha + \Delta\alpha$, then R will change to $R + \Delta R$ where

$$\tan(R + \Delta R) = \left(\frac{1 - \cos(2\beta)\cos[2(\alpha + \Delta\alpha)]}{1 + \cos(2\beta)\cos[2(\alpha + \Delta\alpha)]}\right)^{1/2}. \tag{33}$$

Note that a variation in ellipticity β will also yield a change in R, its magnitude depending on azimuth α. The parameters μ_i $(i = 1, 2)$ are related to the ellipticity through $\mu = \tan^{-1}[\tan(2\beta)/\sin(2\alpha)]$.

Orthogonality of the elliptical states occurs when $\mathbf{E}^{(\alpha,i)\dagger}\mathbf{E}^{(\beta,i)} = 0$, which leads to

$$\tan R \cot(R + \Delta R)\exp[-\mathrm{i}(\mu_2 - \mu_1)] = 1. \tag{34}$$

Here two conditions must be satisfied simultaneously for orthogonality, namely $\mu_1 = \mu_2$ and $\Delta R = 0$. These two conditions are satisfied by heterodyning via an ideal Zeeman process. Clearly $\Delta R = 0$ when both $\Delta\beta$ and $\Delta\alpha$ are simultaneously zero. In particular, equations (30) and (31) become linear states when $\mu_1 = \mu_2 = 0$. In such a case, $R = \alpha$. Moreover, when β (ellipticity) is small, $R \approx \alpha$. However, it must be emphasized that R is not necessarily zero if α is zero; this can be seen from equation (32) since $R = \beta$ when $\alpha = 0$. The significance of this is that an effect akin to rotational misalignment can occur, based on the presence of ellipticity alone, even though the major axes of the two elliptical states are orthogonal and aligned with the axes of the PBS.

3.1.2. *Interferometer matrix* **D**

Leakage through the beam splitter is denoted by terms $\epsilon_x^{1/2}$ and $\epsilon_y^{1/2}$ for the x- and y-axis components of both input wave fields $\mathbf{E}^{(\alpha,i)}$ and $\mathbf{E}^{(\beta,i)}$. The ideal and non-ideal Jones matrices for optical components related to the interferometer are given in table 2. The non-ideal polarizer–mixer is described by a parameter η_P; this is related to the 'efficiency' of the device which is defined as $1 - |\eta_P|$. Under ideal circumstances, $\eta_P = 0$ and the polarizer–mixer is at azimuth $\theta = 45°$, where it causes the largest components of both polarization states (and frequencies) returning to the input, to be optimally mixed and averaged on the surface of the measurement photodetector. The non-ideal Jones matrix **D** for the plane-mirror interferometer (as developed in section 2) is therefore simply

$$\mathbf{D} = \begin{bmatrix} \exp(-\mathrm{i}\delta) + \epsilon_x & 0 \\ 0 & -[1 + \epsilon_y \exp(-\mathrm{i}\delta)] \end{bmatrix}, \tag{35}$$

where δ is the optical phase change acquired following a change in optical path in the measurement arm. This reduces to the ideal Jones representation for **D** when $\epsilon_x = \epsilon_y = 0$.

3.2. *Nonlinearities in the plane-mirror system*

This section identifies contributions from various non-ideal optical polarization effects which lead to phase errors (nonlinearities) in the plane-mirror heterodyne interferometer. First, a general description of the measured phase error is given, which will be followed by the shorter individual source effects.

Table 2. Ideal and non-ideal Jones representations of some optical components.

Optical component	Jones representation	
	Ideal	Non-ideal
PBS	$\mathbf{T}_{\mathrm{PBS}}=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}$	$\mathbf{T}'_{\mathrm{PBS}}=\begin{bmatrix}1 & 0\\ 0 & \epsilon_y^{1/2}\end{bmatrix}$
	$\mathbf{R}_{\mathrm{PBS}}=\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}$	$\mathbf{R}'_{\mathrm{PBS}}=\begin{bmatrix}\epsilon_x^{1/2} & 0\\ 0 & 1\end{bmatrix}$
Mirror	$\mathbf{M}=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$	$\mathbf{M}=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$
Polarizer–mixer	$\mathbf{P}_\theta=\mathbf{R}_\theta\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\mathbf{R}_{-\theta}$	$\mathbf{P}'_\theta=\mathbf{R}_\theta\begin{bmatrix}1 & 0\\ 0 & \eta_{\mathrm{P}}\end{bmatrix}\mathbf{R}_{-\theta}$
	$\mathbf{R}_\theta=\begin{bmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{bmatrix}$	
Plane-mirror interferometer	$\mathbf{D}^{(r)}=\mathbf{R}_{\mathrm{PBS}}\mathbf{M}\mathbf{R}_{\mathrm{PBS}}$ $\mathbf{D}^{(t)}=\mathbf{T}_{\mathrm{PBS}}\mathbf{M}\mathbf{T}_{\mathrm{PBS}}$	$\mathbf{D}^{(r)}=\mathbf{R}'_{\mathrm{PBS}}\mathbf{M}\mathbf{R}'_{\mathrm{PBS}}$ $\mathbf{D}^{(t)}=\mathbf{T}'_{\mathrm{PBS}}\mathbf{M}\mathbf{T}'_{\mathrm{PBS}}$
$\mathbf{D}=\mathbf{D}^{(r)}+\mathbf{D}^{(t)}\exp(-\mathrm{i}\delta)$	$\mathbf{D}=\begin{bmatrix}\exp(-\mathrm{i}\delta) & 0\\ 0 & -1\end{bmatrix}$	$\mathbf{D}=\begin{bmatrix}\exp(-\mathrm{i}\delta)+\epsilon_x & 0\\ 0 & -[1+\epsilon_y\exp(-\mathrm{i}\delta)]\end{bmatrix}$

Substitution of equations (30), (31) and (35) into equation (17) and neglecting squares and cross-products of ϵ_x and ϵ_y leads to specific expressions for $\delta_{\Delta\omega}$, $\delta_{\Delta\omega+}$ and $\delta_{\Delta\omega-}$ of equations (28) and (29):

$$\delta_{\Delta\omega}=\tan^{-1}\left(\frac{(A-D)\sin\delta'+(B-C)\sin\delta''}{(A+D)\cos\delta'+(B+C)\cos\delta''}\right), \tag{36}$$

with

$$A=-2\cos^2\theta\cos R\sin(R+\Delta R),$$
$$B=-(\epsilon_x+\epsilon_y)\sin(2\theta)\cos R\cos(R+\Delta R),$$
$$C=(\epsilon_x+\epsilon_y)\sin(2\theta)\sin R\sin(R+\Delta R),$$
$$D=2\sin^2\theta\sin R\cos(R+\Delta R);$$

$$\delta_{\Delta\omega+}=\tan^{-1}\left(\frac{A'\sin\delta'+B'\sin\delta''-C'\sin\delta'}{A'\cos\delta'+B'\cos\delta''+C'\cos\delta'}\right), \tag{37}$$

with

$$A'=-2\epsilon_x\cos^2\theta\cos R\sin(R+\Delta R),$$
$$B'=-\sin(2\theta)\cos R\cos(R+\Delta R),$$
$$C'=2\epsilon_y\sin^2\theta\sin R\cos(R+\Delta R);$$

$$\delta_{\Delta\omega-}=\tan^{-1}\left(\frac{A''\sin\delta'-B''\sin\delta''-C''\sin\delta'}{A''\cos\delta'+B''\cos\delta''+C''\cos\delta'}\right). \tag{38}$$

with

$$A''=-2\epsilon_x\cos^2\theta\cos R\sin(R+\Delta R),$$
$$B''=-\sin(2\theta)\sin R\sin(R+\Delta R),$$
$$C''=2\epsilon_y\sin^2\theta\sin R\cos(R+\Delta R);$$

where θ is the azimuth of the polarizer–mixer (see table 2), ϵ_x, ϵ_y, R and ΔR are as

before, $\delta' = \frac{1}{2}(\mu_1 - \mu_2)$ and $\delta'' = \frac{1}{2}(\mu_1 + \mu_2)$. The corresponding amplitudes $k_{\Delta\omega}$, $k_{\Delta\omega+}$ and $k_{\Delta\omega-}$ can be found by evaluating the square root of the sum of the squares of the numerator and denominator within large parentheses for each of equations (36), (37) and (38). Of course, when the inputs are not elliptically polarized (i.e. $\delta' = \delta'' = 0$), the K in equation (22) are real and are given by

$$K_{\Delta\omega} = A + B + C + D, \tag{39}$$

$$K_{\Delta\omega+\delta} = A' + B' + C', \tag{40}$$

$$K_{\Delta\omega-\delta} = A'' + B'' + C''. \tag{41}$$

3.2.1. *Effects involving linearly polarized states*

The effects involving linearly polarized states are relatively straightforward to obtain. For linear states, $R = \alpha$ and $\beta = 0$; therefore $\delta' = \delta'' = 0$. Moreover, the extent of non-orthogonality, for example, is simply described in terms of a single parameter $\Delta\alpha$ in contrast with elliptical states. By allowing a specific effect to influence the ideal interferometer, we are able to observe the extent and nature of that non-ideal effect.

(*a*) *Rotational misalignment.* In [8], it was shown that, when the input polarization states were rotated about the input beam relative to the beam splitter, the phase error was (using equation (27))

$$\delta_\epsilon = \tan^{-1}\left(\frac{[\sec(2\alpha) - 1]\tan\delta}{1 + \sec(2\alpha)\tan^2\delta}\right) \tag{42}$$

(where α is used in place of R in [8]). It was also pointed out in [8] that the second-harmonic nonlinearity, which equation (42) represents, is the result of the non-zero amplitude introduced by $K_{\Delta\omega-\delta}$, for non-zero rotational misalignment angle α. Note that, for linearly polarized light, only rotational misalignment can independently produce second-harmonic nonlinearity, although beam splitter leakage and non-orthogonality also contribute to second-harmonic nonlinearity.

(*b*) *Misalignment of polarizer–mixer.* For an ideal interferometer with a polarizer–mixer at arbitrary azimuth θ, one finds that only $K_{\Delta\omega+\delta}$ is non-zero; this shows that misalignment of the polarizer–mixer alone does not independently produce non-linearity. However, in cooperation with an independent effect, such as non-zero α, a significant amount of nonlinearity may be introduced, in this case, first harmonics in addition to second harmonics. In fact, $K_{\Delta\omega} = -\sin(2\alpha)\cos(2\theta)$, $K_{\Delta\omega+\delta} = -\sin(2\theta)\cos^2\alpha$, and $K_{\Delta\omega-\delta} = \sin(2\theta)\sin^2\alpha$, and the phase error is (equation (27))

$$\delta_\epsilon = \tan^{-1}\left(\frac{\{[\sec(2\alpha) - 1] - \tan(2\alpha)\cot(2\theta)\sec\delta\}\tan\delta}{[1 + \sec(2\alpha)\tan^2\delta] + \tan(2\alpha)\cot(2\theta)\sec\delta}\right). \tag{43}$$

Clearly, equation (43) reduces to equation (42) when $\theta = \pi/4$. A plot of equation (43) is shown in figure 4 where $\theta = \pi/4 + \alpha$, as can occur in some laser metrology systems, when the laser source, polarizer–mixer and photodetectors are found in a single unit.

(*c*) *Effect of non-orthogonality.* Non-orthogonality will independently produce nonlinearity. Moreover, we find that $K_{\Delta\omega} = -\sin\Delta\alpha$, $K_{\Delta\omega+\delta} = -\cos\Delta\alpha$, and $K_{\Delta\omega-\delta} = 0$, so that we have from slight rearrangement of equation (27):

$$\delta_\epsilon = -\tan^{-1}\left(\frac{\tan\delta}{1+\cot\Delta\alpha\sec\delta}\right). \tag{44}$$

Non-orthogonality is therefore an independent effect, dominantly producing fundamental nonlinearity. The extremum of equation (44) is given by

$$\delta_{\epsilon,\max} = \tan^{-1}\left(\frac{\tan\Delta\alpha}{(1-\tan^2\Delta\alpha)^{1/2}}\right), \tag{45}$$

which is approximately equal to $\Delta\alpha$ for small $\Delta\alpha$. This turns out to be particularly significant for the plane-mirror interferometer as $\Delta\alpha$ up to 3° has

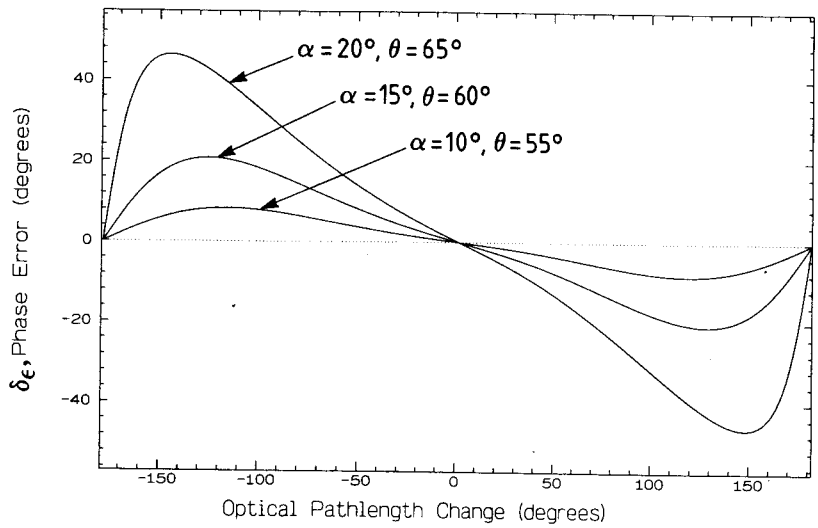

Figure 4. Phase errors due to a combination of input orthogonal states rotated through angle α (relative to the beam splitter) and misaligned polarizer–mixer with azimuth θ.

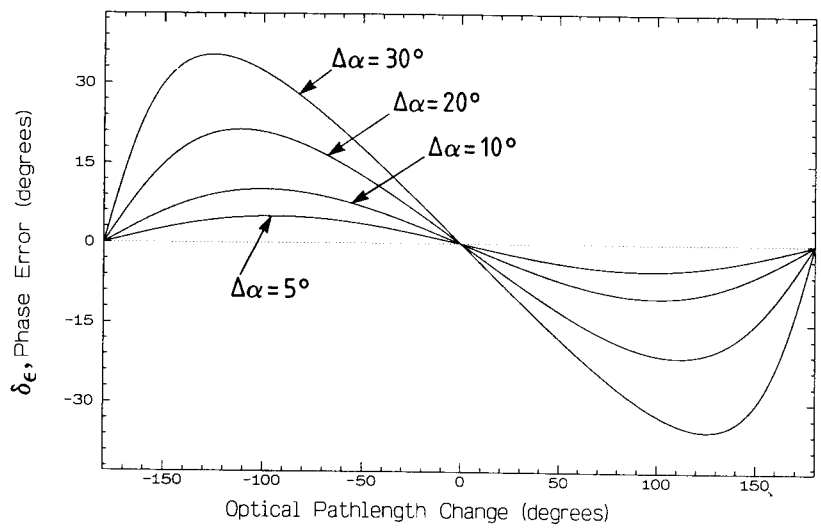

Figure 5. Phase errors due to non-orthogonality alone.

been observed in [4]. A worst-case peak-to-peak error of 6° is estimated. A plot of equation (45) is shown in figure 5.

(*d*) *Effect of beam splitter leakage.* In this case $K_{\Delta\omega} = (\epsilon_x + \epsilon_y)$, $K_{\Delta\omega+\delta} = -1$ and $K_{\Delta\omega-\delta} = 0$, leading to phase nonlinearity (equation (27)):

$$\delta_\epsilon = -\tan^{-1}\left(\frac{(\epsilon_x + \epsilon_y)\tan\delta}{(\epsilon_x + \epsilon_y) + \sec\delta}\right). \tag{46}$$

The extremum of equation (46) is

$$\delta_{\epsilon,\max} = \tan^{-1}\left(\frac{\epsilon_x + \epsilon_y}{[1 - \tan^2(\epsilon_x + \epsilon_y)]^{1/2}}\right), \tag{47}$$

which is approximately equal to $\epsilon_x + \epsilon_y$ for small ϵ_x and ϵ_y. This also turns out to be significant for two reasons: firstly the error is first order in the leakage terms, and secondly the error is identical with $\delta_{\epsilon,\text{leak}}$ calculated from a factorization of **D**, the interferometer Jones matrix, into a retarder and variable partial linear polarizer (appendix E). Polarizing beam splitter leakage parameters are often unique to each device, but some manufacturers quote for intensity leakage of the TM (p or transmission) state into the TE (s or reflection) state as 1% and 2% respectively. A worst-case estimate of the peak-to-peak nonlinearity is therefore 3·4°. As is readily seen from equation (46), the effects of beam splitter leakage are independent.

3.2.2. *Effects involving elliptically polarized light*

The major difference from linear states is that, since $\beta \neq 0$, the parameter R is related to both α and β through equation (32). Because these two parameters are bound up in R, individual expressions for nonlinearity will involve R, as well as the phases μ_1 and μ_2.

(*a*) *Ellipticity and rotational misalignment.* The phase error resulting from ellipticity alone (i.e. assuming accurately aligned orthogonal states) in terms of R is given by

$$\delta_\epsilon = \tan^{-1}\left(\frac{\sec(2R)\tan(\delta+\delta'') - \tan\delta + \tan(2R)\sec(\delta+\delta'')\sin\delta'}{1 + \tan\delta\,[\sec(2R)\tan(\delta+\delta'') + \tan(2R)\sec(\delta+\delta'')\sin\delta']}\right). \tag{48}$$

It was pointed out earlier that, although the major axes of the elliptical states are orthogonal and aligned (i.e. $\Delta\alpha = 0$ and $\alpha = 0$), we still have $R = \beta$, the ellipticity. Moreover, $\delta' = 0$ and $\delta'' = \mu$ from the orthogonality requirements. Under these conditions, equation (48) becomes

$$\delta_\epsilon = \tan^{-1}\left(\frac{\sec(2\beta)\tan(\delta+\mu) - \tan\delta}{1 + \tan\delta\sec(2\beta)\tan(\delta+\mu)}\right). \tag{49}$$

This result is identical with rotational misalignment of linearly polarized states at angle $\alpha = \beta$ and shows that ellipticity gives an independent effect at the second harmonic. Note, however, that the effect from ellipticity has an offset of μ.

Attempts to find analytical expressions for other effects in analogy with linear states lead to cumbersome expressions and are best instead studied directly through equations (36)–(38).

(*b*) *Second-harmonic nonlinearity*. For elliptical states, second-harmonic nonlinearity occurs when $K_{\Delta\omega-} \neq 0$. The only parameter which independently supports second-harmonic nonlinearity is R. However, equation (38) shows that, if $R = 0$ but ΔR, ϵ_x and ϵ_y are non-zero, then together they contribute to second-harmonic nonlinearity, although none of them can do so independently.

The collapse of elliptical states to linear states when $\beta = 0$ leads to $R = \alpha$, corresponding to rotational misalignment. On the other hand, if $\alpha = 0$, then $R = \beta$. This point was alluded to earlier and could be expressed as

$$R = \begin{cases} 0, & \alpha = 0 \text{ and } \beta = 0, \\ \text{constant}, & \alpha \neq 0 \text{ or } \beta \neq 0. \end{cases} \tag{50}$$

Thus we find that rotational misalignment and ellipticity can each independently support second-harmonic nonlinearity.

3.3. *Non-linearity in a practical interferometer layout*

We now consider the nonlinearity in a practical plane-mirror set-up, based on a Zeeman split-frequency heterodyne source. An ellipticity ratio of semiminor to semimajor axis of 0·01 ($\beta = 0{\cdot}58°$) was used, with $\alpha = 2°$, $\Delta\alpha = 3°$, $\epsilon_x = 0{\cdot}02$ and $\epsilon_y = 0{\cdot}01$. We determine $\mu_1 = \tan^{-1}[\tan(2\beta)/\sin(2\alpha)]$, and similarly for μ_2 from β and $\alpha + \Delta\alpha$. R and $R + \Delta R$ are then found from equations (32) and (33) respectively. In laser metrology systems where the detector, polarizer and source are physically linked into a single unit, then the polarizer misalignment is coupled to the rotational misalignment of the input polarization states; thus $\theta = \pi/4 + \alpha$. The values for δ' and δ'' are then determined and substituted along with R, $R + \Delta R$, θ, ϵ_x and ϵ_y into equations (36)–(38) to give the amplitude and phase for the complex K values. The phase-amplitude results for the K are then used in equation (29) to yield the phase error δ_ϵ.

An interesting question arises with regard to the importance of the sense of misalignment and non-orthogonality. In previous analyses in the literature, this was never explicitly considered. However, there are four possible input configurations to the beam splitter (figure 6); these occur for various combinations of anticlockwise (positive) or clockwise (negative) rotational misalignments and non-orthogonality of the input states (which have the same input ellipticity β). Substitution of the above values for the practical set-up shows that there are four distinct nonlinearity effects (figure 7). This result is not surprising since the combination of odd (asymmetric) functions, particularly terms involving $\sin R$ and $\sin \Delta R$, can significantly alter the behaviour of δ_ϵ.

4. Extension to more general interferometers

Although equation (22) for intensity and equation (29) for phase error are general, the plane-mirror interferometer treated in detail in section 3 is obviously a special case. A complete generalization requires fully general Jones matrices $\mathbf{D}^{(u)}$ ($u = r, t$) to be written for both paths of the interferometer, and the question arises as to whether explicit relationships can be established between phase errors and the parameters of those matrices.

One approach relies on the following parametrization of the Jones matrices.

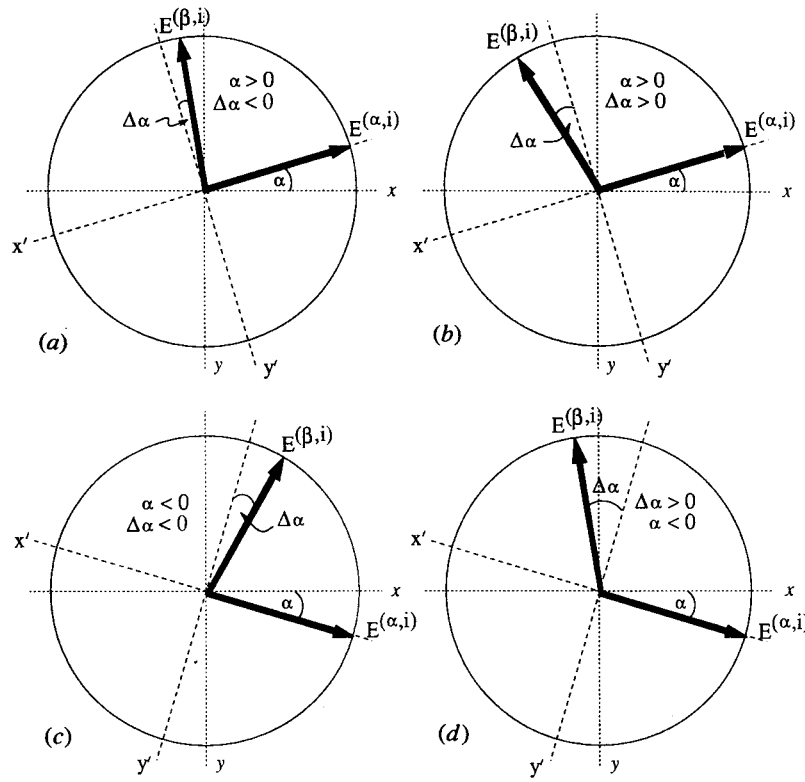

Figure 6. Input configuration of polarization states. x and y are beam splitter axes. The arrows are vectors along the major axes of the general elliptically polarized heterodyne states. Angle α is measured relative to the x axis while $\Delta\alpha$ is measured relative to y', the y axis rotated through α.

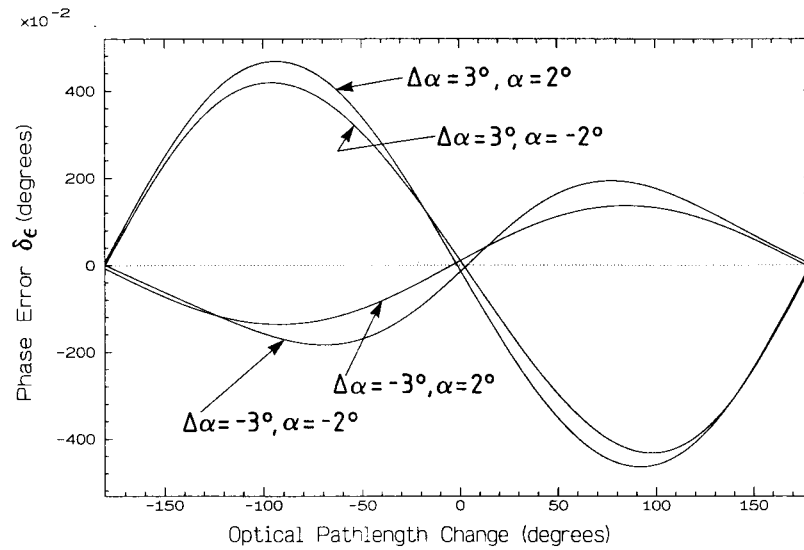

Figure 7. Phase errors due to combinations of non-orthogonality, rotational misalignment, ellipticity, PBS leakage and polarizer–mixer misorientation.

It is straightforward to verify that any Jones matrix can be expressed as

$$\mathbf{D}^{(u)} = a\mathbf{WH}, \tag{51}$$

where a is a complex amplitude factor, $\mathbf{W}$ a member of $SU(2)$ and $\mathbf{H}$ is Hermitian; alternatively, using a unitary transformation to diagonalize $\mathbf{H}$,

$$\mathbf{D}^{(u)} = a\mathbf{UPV}, \tag{52}$$

where $\mathbf{P}$ is a real diagonal matrix (corresponding to a partial linear polarizer), $\mathbf{V}$ is a restricted member of $SU(2)$ with only two parameters, and $\mathbf{U}$ is a general member of $SU(2)$. Similarly, the matrix $\mathbf{D}$ for the full interferometer can be written in the form of equation (51) or (52). This parametrization corresponds to the well known result for the Lorentz group, which is of course homomorphic to the group of 2×2 complex matrices, that a general Lorentz transformation reduces to a single special transformation (corresponding to $\mathbf{P}$) and spatial rotations (corresponding to $\mathbf{U}$ and $\mathbf{V}$).

In some cases this parametrization does allow a useful re-expression of $\mathbf{D}$. For instance, in the plane-mirror interferometer the expression for $\mathbf{D}$ in equation (34) can be readily rewritten as the product of a partial linear polarizer ($\mathbf{P}$) and a linear retarder ($\mathbf{V}$), where both elements depend on the interferometer path difference δ (appendix E). Also, for a system in which both r and t paths, and the overall interferometer are all lossless, then $\mathbf{D}^{(r)}$, $\mathbf{D}^{(t)}$ and $\mathbf{D}$ all reduce to members of $SU(2)$, allowing the parameters of $\mathbf{D}$ to be obtained straightforwardly. Such a description is appropriate, for instance, to transport through a highly birefringent fibre as discussed by Bergamin *et al.* [9]. More generally, however, while the parametrization procedure based on equations (51) and (52) does provide algorithms to obtain the parameters of $\mathbf{D}$, explicit relationships between phase errors and their sources are difficult to establish. In general, then, it appears that the K parameters in equation (29) must be obtained by consideration of a particular interferometer configuration on an individual basis.

The form of equation (51) can, however, be applied to a treatment of beam splitters, which somewhat extends the generality of the discussion of section 3. Again, it is not easy to find relationships between the parameters of a completely general beam splitter. However, consider a lossless beam splitter in which inputs (expressed as Jones vectors $\mathbf{i}_2$ and $\mathbf{i}_3$) at ports 2 and 3, say, produce outputs $\mathbf{d}_1$ and $\mathbf{d}_4$ at ports 1 and 4; further suppose path 2 to 1 (representing, say, reflection) is equivalent to path 3 to 4, and path 3 to 1 (representing, say, transmission) is equivalent to path 2 to 4. Such a device, modelling a dielectric beam splitter with negligible spurious reflection losses, can be represented by

$$\begin{bmatrix}\mathbf{d}_1\\ \mathbf{d}_4\end{bmatrix} = \begin{bmatrix}\mathbf{R} & \mathbf{T}\\ \mathbf{T} & \mathbf{R}\end{bmatrix}\begin{bmatrix}\mathbf{i}_2\\ \mathbf{i}_3\end{bmatrix}, \tag{53}$$

in which $\mathbf{T}$ and $\mathbf{R}$ are Jones matrices for transmission and reflection paths, and the condition that the beam splitter is lossless then requires

$$\begin{aligned}\mathbf{R}^\dagger\mathbf{R} + \mathbf{T}^\dagger\mathbf{T} &= \mathbf{1},\\ \mathbf{R}^\dagger\mathbf{T} + \mathbf{T}^\dagger\mathbf{R} &= \mathbf{0},\end{aligned} \tag{54}$$

where $\mathbf{1}$ is the 2×2 unit matrix. Writing $\mathbf{R}$ and $\mathbf{T}$ in the form of equation (52), it is straightforward to show that equation (54) then implies $\mathbf{R}$ and $\mathbf{T}$ must share the

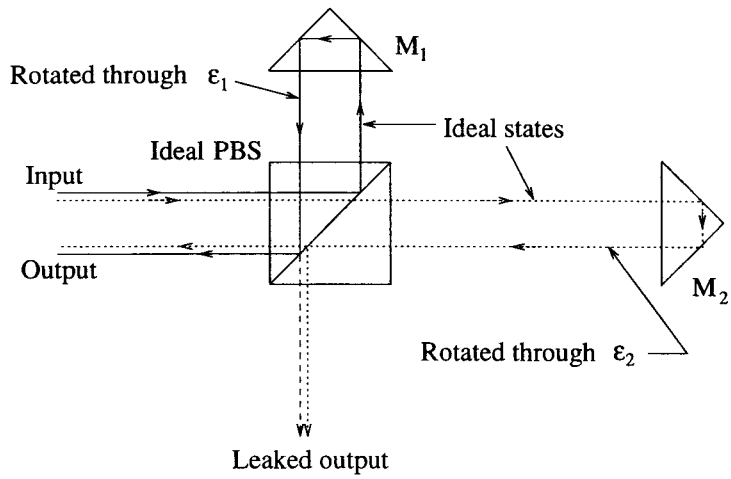

Figure 8. Leaked components appearing at an output port of an ideal PBS, following rotation of the returning polarized states by the cube corner reflectors. The output intensity $I = \cos^2 \epsilon_1 + \cos^2 \epsilon_2 \cos(\Delta\omega t + \delta)$. The leaked intensity $I_{\text{leak}} = \sin^2 \epsilon_1 + \sin^2 \epsilon_2 \cos(\Delta\omega t + \delta)$.

same $SU(2)$ matrices **U** and **V** and that, to within an overall phase factor, they can be expressed as

$$\begin{aligned} \mathbf{R} &= \mathbf{U}\mathbf{P}_R\mathbf{V} \\ \mathbf{T} &= \mathrm{i}\mathbf{U}\mathbf{P}_T\mathbf{V}. \end{aligned} \tag{55}$$

where $\mathbf{P}_R$ and $\mathbf{P}_T$ are real and diagonal. Hence, to within the $SU(2)$ elements **U** and **V** which are common to both reflection and transmission paths (and which can be readily combined with other such elements appearing in the interferometer system), a beam splitter within these fairly realistic restrictions can be represented by real diagonal $\mathbf{R} = \mathbf{P}_R$ and $\mathbf{T} = \mathbf{P}_T$, as used in section 3.

The most substantive omission from the treatment of section 3 therefore is the case where $SU(2)$ elements are included in both arms, which will have a major effect only if the beam splitter is substantially imperfect. As an illustration, if **Q** (figure 2) induces rotational misalignment of the state returning to the beam splitter from **M**, then only the visibility is affected (figure 8).

5. Conclusion

We have undertaken a detailed analysis of the polarization effects which lead to nonlinearity in the heterodyne interferometer. Our approach is based on Jones matrix and coherency matrix analyses, leading to a general expression for output intensity (equation (22)) and corresponding expressions for phase error (equation (27) or (29)). Although these results are straightforward, they show explicitly the origins of the fundamental and second-harmonic nonlinearity in terms of the Jones matrices of system components. This is sufficient to provide a complete analysis of the plane-mirror interferometer. Although analytical results for more general interferometers are elusive, it does appear that the separation of effects shown in equation (22) provides a good basis for calculating results in individual cases.

The choice of a coherency matrix approach, rather than direct use of Jones

calculus, was made to provide a more generalized framework. Although this has not been exploited in the presented paper, it should for instance allow calculation of noise resulting from partial coherence of laser modes.

Acknowledgments

This work has been supported by a grant from Stanford University, USA, as part of the Gravity Probe-B programme.

Appendix A. 'Beat' and dispersion effects

Equation (5) may be rewritten to include explicitly a path length d in a material of refractive index n_t in, say, the measurement arm, as

$$\delta_t^{(u)} = \frac{4\pi}{\lambda^{(u)}}\left(\int_0^{L_t} n_t(\lambda^{(u)}, l)\,\mathrm{d}l + \int_{L_t}^{L_t+d} n_t(\lambda^{(u)}, l)\,\mathrm{d}l\right) \quad (u = \alpha, \beta). \tag{A 1}$$

The optical phase changes acquired in the interferometer are $\delta^{(\alpha)} = \delta_t^{(\alpha)} - \delta_r^{(\alpha)}$ and $\delta^{(\beta)} = \delta_t^{(\beta)} - \delta_r^{(\beta)}$. Thus the difference between optical phase changes for wave fields labelled α and β is

$$\delta^{(\alpha)} - \delta^{(\beta)} = \frac{4\pi}{\lambda^{(\alpha)}}\int_{L_t}^{L_t+d} n_t(\lambda^{(\alpha)}, l)\,\mathrm{d}l - \frac{4\pi}{\lambda^{(\beta)}}\int_{L_t}^{L_t+d} n_t(\lambda^{(\beta)}, l)\,\mathrm{d}l + \delta_0^{(\alpha)} - \delta_0^{(\beta)} \tag{A 2}$$

where, for $u = \alpha, \beta$,

$$\delta_0^{(u)} = \frac{4\pi}{\lambda^{(u)}}\int_0^{L_t} n_t(\lambda^{(u)}, l)\,\mathrm{d}l - \frac{4\pi}{\lambda^{(u)}}\int_0^{L_r} n_r(\lambda^{(u)}, l)\,\mathrm{d}l. \tag{A 3}$$

Note that, if the interferometer is balanced, for example $L_t = L_r$, and $n_t = n_r$, then $\delta_0^{(\alpha)} = \delta_0^{(\beta)} = 0$. Now, making the substitution $\lambda^{(\alpha)} = \lambda$ and $\lambda^{(\beta)} = \lambda + \Delta\lambda$ in equation (A 2),

$$\delta^{(\alpha)} - \delta^{(\beta)} = -4\pi\int_{L_t}^{L_t+d}\left(\frac{n_t(\lambda+\Delta\lambda, l)}{\lambda+\Delta\lambda} - \frac{n_t(\lambda, l)}{\lambda}\right)\mathrm{d}l + \delta_0^{(\alpha)} - \delta_0^{(\beta)}. \tag{A 4}$$

Writing

$$\frac{n_t(\lambda+\Delta\lambda, l)}{\lambda+\Delta\lambda} - \frac{n_t(\lambda, l)}{\lambda} = \frac{\lambda n_t(\lambda+\Delta\lambda, l) - (\lambda+\Delta\lambda)n_t(\lambda, l)}{\lambda(\lambda+\Delta\lambda)\Delta\lambda}\Delta\lambda \tag{A 5}$$

$$\approx \frac{\partial}{\partial\lambda}\left(\frac{n_t(\lambda, l)}{\lambda}\right)\Delta\lambda \tag{A 6}$$

and substituting into equation (A 4) gives

$$\delta^{(\alpha)} - \delta^{(\beta)} = -4\pi\,\Delta\lambda\int_{L_t}^{L_t+d}\frac{\partial}{\partial\lambda}\left(\frac{n_t(\lambda+l)}{\lambda}\right)\mathrm{d}l + \delta_0^{(\alpha)} - \delta_0^{(\beta)}. \tag{A 7}$$

Using the first-order dispersion approximation

$$\frac{\partial n_t(\lambda, l)}{\partial\lambda} \approx \frac{n_t(\lambda, l) - n_{g,t}(\lambda, l)}{\lambda} \tag{A 8}$$

in the derivative expansion in equation (A 7) yields

$$\delta^{(\alpha)} - \delta^{(\beta)} \approx \frac{\Delta\lambda}{\lambda}\delta_{g,t} + \delta_0^{(\alpha)} - \delta_0^{(\beta)}. \tag{A 9}$$

To get an idea of the order of magnitude of $\delta_0^{(\alpha)} - \delta_0^{(\beta)}$, let $n_t = n_r = \bar{n}$, the average refractive index for both wave fields, so that

$$\delta_0^{(\alpha)} = \frac{4\pi}{\lambda} \bar{n}(L_t - L_r) \tag{A 10}$$

$$\delta_0^{(\beta)} = \frac{4\pi}{\lambda + \Delta\lambda} \bar{n}(L_t - L_r). \tag{A 11}$$

Using $\lambda \gg \Delta\lambda$,

$$\delta_0^{(\alpha)} - \delta_0^{(\beta)} \approx \frac{4\pi}{\lambda} \bar{n}(L_t - L_r) \frac{\Delta\lambda}{\lambda} \tag{A 12}$$

$$\approx \delta_0 \frac{\Delta\lambda}{\lambda}. \tag{A 13}$$

Thus $\delta_0^{(\alpha)} - \delta_0^{(\beta)}$ is closely equal to $\delta^{(\alpha)} - \delta^{(\beta)}$ and both are small compared with δ_0. Finally, we rewrite equation (A 9) as

$$\delta^{(\alpha)} - \delta^{(\beta)} \approx \frac{\Delta\lambda}{\lambda} (\delta_{g,t} + \delta_0). \tag{A 14}$$

Note that this treatment indicates only the scale of such effects. In general, the cross-coherency matrices will contain phasors involving for example $\delta_r^{(\alpha)} - \delta_r^{(\beta)}$ and $\delta_t^{(\alpha)} - \delta_r^{(\beta)}$ which depend on total path, as well as path difference.

Appendix B. Derivation of phasor terms for cross-coherency matrices

In this appendix we give a simplistic derivation of the phasor terms associated with the cross-coherency matrices $\mathbf{J}^{(uv)}(u, v = \alpha, \beta)$. Starting with equations (11), (12) and (14), we find for example that, for $u = \alpha$ and $v = \beta$,

$$J_{xx}^{(\alpha\beta)} = \langle E_x^{(\alpha)} E_x^{(\beta)*} \exp(-i\Delta\omega t) \rangle. \tag{B 1}$$

Treating $\Delta\omega$ as a random time-dependent function with mean $\overline{\Delta\omega}$, then we may write

$$\Delta\omega(t) = \overline{\Delta\omega}\, t + \varphi_n(t) \tag{B 2}$$

where $\varphi_n(t)$ is a phase noise term. We therefore have, rewriting equation (B 1),

$$J_{xx}^{(\alpha\beta)} = \langle E_x^{(\alpha)} E_x^{(\beta)*} \exp[-i\varphi_n(t)] \rangle \exp(-i\overline{\Delta\omega}\, t), \tag{B 3}$$

whereupon, assuming independence and that $\varphi_n(t)$ follows a zero mean Gaussian distribution, then

$$\begin{aligned} J_{xx}^{(\alpha\beta)} &= \langle E_x^{(\alpha)} E_x^{(\beta)*} \rangle \langle \exp[-i\varphi_n(t)] \rangle \exp(-i\overline{\Delta\omega}\, t) \\ &= \langle E_x^{(\alpha)} E_x^{(\beta)*} \rangle \exp(-i\overline{\varphi_n} - \sigma_\varphi^2) \exp(-i\overline{\Delta\omega}\, t) \qquad \text{(B 4)} \\ &= \langle E_x^{(\alpha)} E_x^{(\beta)*} \rangle \exp(-\sigma_\varphi^2) \exp(-i\overline{\Delta\omega}\, t), \qquad \text{(B 5)} \end{aligned}$$

since $\overline{\varphi_n} = 0$ by hypothesis, and where σ_φ^2 is the variance of the noise process. A similar analysis can be carried out for $(uv) = (\beta\alpha)$, $(\alpha\alpha)$ and $(\beta\beta)$. The phasors associated with particular coherency matrices are given in table 1.

Appendix C. Evaluation of Tr (PDJD†P†)

Let $\mathbf{P} = [p_{ij}]$ and $\mathbf{D} = [d_{ij}]$ $(i,j = 1, 2)$, then we can write

$$\mathrm{Tr}\,(\mathbf{PDJD}^\dagger\mathbf{P}^\dagger) = (a_0 J_{xx} + a_1 J_{xy} + a_2 J_{yx} + a_3 J_{yy}) M_{\Delta\omega}(t) \tag{C 1}$$

where $M_{\Delta\omega}(t)$ is the phasor associated with $\mathbf{J}$ (see table 1), and the a are combinations of the elements of $\mathbf{P}$ and $\mathbf{D}$. These are obtained from straightforward multiplication $\mathbf{PDJD}^\dagger\mathbf{P}^\dagger$. Note that $a_1 = a_2^*$ and that a_0 and a_3 are real. If $\mathbf{D}$ is diagonal, then $d_{12} = d_{21} = 0$ as is the case for $\mathbf{Q} = \mathbf{1}$ in the plane-mirror interferometer, and

$$\begin{aligned} a_0 &= (p_{11}^2 + p_{21}^2)|d_{11}|^2, \\ a_1 &= (p_{11}p_{12} + p_{21}p_{22}) d_{11} d_{22}^*, \\ a_2 &= (p_{11}p_{12} + p_{21}p_{22}) d_{22} d_{11}^*, \\ a_3 &= (p_{12}^2 + p_{22}^2)|d_{22}|^2. \end{aligned} \tag{C 2}$$

The matrices $\mathbf{D}^{(r)}$ and $\mathbf{D}^{(t)}$ are given by

$$\mathbf{D}^{(r)} = \begin{bmatrix} r_\perp & 0 \\ 0 & r_\parallel \end{bmatrix}, \tag{C 3}$$

$$\mathbf{D}^{(t)} = \begin{bmatrix} t_\perp & 0 \\ 0 & t_\parallel \end{bmatrix}. \tag{C 4}$$

Then using $\mathbf{D} = \mathbf{D}^{(r)} + \mathbf{D}^{(t)} \exp(-\mathrm{i}\delta)$, we have

$$\mathbf{D} = \begin{bmatrix} r_\perp + t_\perp \exp(-\mathrm{i}\delta) & 0 \\ 0 & r_\parallel + t_\parallel \exp(-\mathrm{i}\delta) \end{bmatrix}. \tag{C 5}$$

For the polarizer–mixer (with azimuth θ),

$$\begin{aligned} \mathbf{P} &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \eta_\mathrm{P} \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \\ &= \begin{bmatrix} 1 - (1-\eta_\mathrm{P})\sin^2\theta & (1-\eta_\mathrm{P})\sin\theta\cos\theta \\ (1-\eta_\mathrm{P})\sin\theta\cos\theta & 1 - (1-\eta_\mathrm{P})\cos^2\theta \end{bmatrix} \end{aligned} \tag{C 6}$$

which yields,

$$\begin{aligned} a_0 &= [(\cos^2\theta + \eta_\mathrm{P}\sin^2\theta)^2 + (1-\eta_\mathrm{P})^2 \sin^2\theta\cos^2\theta][|r_\perp + t_\perp \exp(-\mathrm{i}\delta)|^2] \\ &= (\cos^2\theta + \eta_\mathrm{P}^2 \sin^2\theta)[|r_\perp|^2 + |t_\perp|^2 + r_\perp t_\perp^* \exp(\mathrm{i}\delta) + r_\perp^* t_\perp \exp(-\mathrm{i}\delta)] \\ &= f_0 + f_\delta \exp(-\mathrm{i}\delta) + f_\delta^* \exp(\mathrm{i}\delta), \end{aligned} \tag{C 7}$$

$$\begin{aligned} a_1 &= [(\cos^2\theta + \eta_\mathrm{P}\sin^2\theta)(1-\eta_\mathrm{P})\sin\theta\cos\theta + (\eta_\mathrm{P}\cos^2\theta + \sin^2\theta) \\ &\quad \times (1-\eta_\mathrm{P})\sin\theta\cos\theta](r_\perp + t_\perp \exp(-\mathrm{i}\delta))[r_\parallel^* + t_\parallel^* \exp(\mathrm{i}\delta)] \\ &= [(1-\eta_\mathrm{P}^2)\sin\theta\cos\theta][r_\perp r_\parallel^* + t_\perp t_\parallel^* + r_\perp t_\parallel^* \exp(\mathrm{i}\delta) + t_\perp r_\parallel^* \exp(-\mathrm{i}\delta)] \\ &= g_0 + g_1 \exp(-\mathrm{i}\delta) + g_2^* \exp(\mathrm{i}\delta) \end{aligned} \tag{C 8}$$

$$a_2 = a_1^* = g_0^* + g_1^* \exp(\mathrm{i}\delta) + g_2^* \exp(-\mathrm{i}\delta) \tag{C 9}$$

$$\begin{aligned} a_3 &= [(\eta_\mathrm{P}\cos^2\theta + \sin^2\theta)^2 + (1-\eta_\mathrm{P})^2 \sin^2\theta\cos^2\theta]|(r_\parallel + t_\parallel \exp(-\mathrm{i}\delta)|^2 \\ &= (\eta_\mathrm{P}^2\cos^2\theta + \sin^2\theta)[|r_\parallel|^2 + |t_\parallel|^2 + r_\parallel t_\parallel^* \exp(\mathrm{i}\delta) + r_\parallel^* t_\parallel \exp(-\mathrm{i}\delta)] \\ &= h_0 + h_\delta \exp(-\mathrm{i}\delta) + h_\delta^* \exp(\mathrm{i}\delta). \end{aligned} \tag{C 10}$$

Substituting for a_0, a_1, a_2 and a_3 in equation (C 1) yields

$$\begin{aligned}\mathrm{Tr}\,(\mathbf{PDJD}^{\dagger}\mathbf{P}^{\dagger}) = [&(f_0 J_{xx} + g_0 J_{xy} + g_0^{*} J_{yx} + h_0 J_{yy}) \\ &+ (f_\delta J_{xx} + g_1 J_{xy} + g_2^{*} J_{yx} + h_\delta J_{yy}) \exp(-\mathrm{i}\delta) \\ &+ (f_\delta^{*} J_{xx} + g_2 J_{xy} + g_1^{*} J_{yx} + h_\delta^{*} J_{yy}) \exp(\mathrm{i}\delta)] M_{\Delta\omega}(t) \\ = [&\gamma_0 + \gamma_1 \exp(-\mathrm{i}\delta) + \gamma_2 \exp(\mathrm{i}\delta)] M_{\Delta\omega}(t). \end{aligned} \tag{C 11}$$

Appendix D. The Mueller representation and the Stokes vectors

In this appendix a slightly different approach is used, where the coherency *vector* is manipulated in contrast with the coherency *matrix* in section 2.2. The advantage of this is that there is a straightforward method of converting coherency vectors to the Mueller and Stokes representations.

The cross-coherency vector $\mathcal{J}^{(uv)}$ is defined as

$$\mathcal{J}^{(uv)} = \langle \mathbf{E}^{(u)} \times \mathbf{E}^{(v)*} \rangle = \begin{bmatrix} J_{xx}^{(uv)} \\ J_{xy}^{(uv)} \\ J_{yx}^{(uv)} \\ J_{yy}^{(uv)} \end{bmatrix} M_{\Delta\omega}^{(uv)}. \tag{D 1}$$

Note here that since the discussion of section 2.2 demands that $J_{ij}^{(uv)} = J_{ji}^{(vu)*}$ $(i, j = x, y)$, then $\mathcal{J}^{(uv)} \neq \mathcal{J}^{(vu)*}$ when $u \neq v$. In order to maintain consistency in the use of cross-coherency vectors, a new transformation **B** must be introduced, so that (following on from (*b*) in section 2.2).

(*c*) *For cross-coherency vectors* $\mathcal{J}^{(uv)} = \mathbf{B}\mathcal{J}^{(vu)*}$, *where* **B** *is a unitary transformation and is given by*

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{D 2}$$

B performs the operation equivalent to matrix transposition by allowing $J_{xy}^{(uv)}$ to swap places with $J_{yx}^{(uv)}$ in equation (D 1) so that, when result (*a*) (section 2.2) is applied to the elements of $\mathcal{J}^{(uv)}$, the result will be consistent with the definition of the coherency vector $\mathcal{J}^{(vu)}$ and its complex conjugate.

It is well known [21] that the coherency vector is easily transformed to the Stokes vector $\mathbf{S}^{(uv,\,\mathrm{i})}$ through the transformation

$$\mathbf{S}^{(uv,\,\mathrm{i})} = \mathbf{A}\mathcal{J}^{(uv,\,\mathrm{i})}, \tag{D 3}$$

where

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & \mathrm{i} & -\mathrm{i} & 0 \end{bmatrix}. \tag{D 4}$$

Moreover, for an optical system possessing a Jones matrix **Q**, the corresponding Mueller matrix **M** is given by [21]

$$\mathbf{M} = \mathbf{A}(\mathbf{Q} \times \mathbf{Q}^*)\mathbf{A}^{-1}. \tag{D 5}$$

Furthermore one writes the output Stokes vector $\mathbf{S}^{(uv)}$, in terms of the input Stokes vector $\mathbf{S}^{(uv,i)}$, as

$$\mathbf{S}^{(uv)} = \mathbf{M}\mathbf{S}^{(uv,i)}. \tag{D 6}$$

We also have $\mathbf{S}^{(uv)} = \mathbf{AB}\mathscr{J}^{(vu)^*}$ but, since $\mathbf{AB} = \mathbf{A}^*$, one gets the following result

(*d*) *For 'mixed' Stokes vectors*

$$\mathbf{S}^{(uv)} = \mathbf{S}^{(vu)^*}.$$

If N different frequencies are input to the optical system, then the input coherency vector is written as

$$\mathscr{J} = \sum_{u,v=1}^{N} \langle \mathbf{E}^{(u)} \times \mathbf{E}^{(v)^*} \rangle = \sum_{u,v=1}^{N} \mathscr{J}^{(uv,i)}, \tag{D 7}$$

and therefore the total output Stokes vector denoted by **S** is

$$\mathbf{S} = \mathbf{MA}\mathscr{J}. \tag{D 8}$$

However, substitution of equations (D 3) and (D 7) into equation (D 8) yields the result

$$\mathbf{S} = \mathbf{M} \sum_{u,v=1}^{N} \mathbf{S}^{(uv,i)}. \tag{D 9}$$

Thus the output Stokes vector for a system of N frequencies is the Mueller transform of the sum of the Stokes vectors of the pairwise permutations of the input wavefields.

Appendix E. Factorization of the Jones matrix for an interferometer

The ideal interferometer matrix **D** can be written as

$$\mathbf{D} = \begin{bmatrix} \exp[-\tfrac{1}{2}\mathrm{i}(\delta - \pi)] & 0 \\ 0 & \exp[\tfrac{1}{2}\mathrm{i}(\delta - \pi)] \end{bmatrix}, \tag{E 1}$$

which is identical with a linear retarder of retardation $\delta - \pi$ with azimuth 0°. Thus an ideal plane-mirror heterodyne interferometer could be described as the combination of a variable linear retarder followed by a polarizer–mixer in front of the measurement photodetector. Secondly, the Jones matrix for non-ideal **D** could be written as

$$\mathbf{D} = \begin{bmatrix} (1 + 2\epsilon_x \cos\delta + \epsilon_x^2)\exp(-\mathrm{i}\Theta_x) & 0 \\ 0 & -(1 + 2\epsilon_y \cos\delta + \epsilon_y^2)\exp(-\mathrm{i}\Theta_y) \end{bmatrix} \tag{E 2}$$

and is easily factorized, neglecting overall phasor terms, to give

$$\mathbf{D} = \begin{bmatrix} (1 + 2\epsilon_x \cos\delta + \epsilon_x^2) & 0 \\ 0 & (1 + 2\epsilon_y \cos\delta + \epsilon_y^2) \end{bmatrix} \times \begin{bmatrix} \exp[-\tfrac{1}{2}\mathrm{i}(\Theta_x - \Theta_y - \pi)] & 0 \\ 0 & \exp[\tfrac{1}{2}\mathrm{i}(\Theta_x - \Theta_y - \pi)] \end{bmatrix}, \tag{E 3}$$

where

$$\Theta_x = \tan^{-1}\left(\frac{\tan\delta}{1+\epsilon_x \sec\delta}\right),$$

$$\Theta_y = \tan^{-1}\left(\frac{\tan\delta}{1+(1/\epsilon_y)\sec\delta}\right).$$

This factorization suggests that the introduction of leakage into the polarizing beam splitter can be described by the combination of a variable-intensity non-polarizing beam splitter followed by a retarder of retardation $\Theta_x - \Theta_y - \pi$. Note that, as ϵ_x and ϵ_y approach zero, $\Theta_x \to \delta$ and $\Theta_y \to 0$. Thus the parameter $\Theta_x - \Theta_y$ replaces δ in equation (E 1) and can be re-expressed as $\Theta_x - \Theta_y = \delta + \delta_{\epsilon,\text{leak}}$, where $\delta_{\epsilon,\text{leak}}$ is a retardation error resulting from leakage in the beam splitter, which depends on δ. Bruning [22] also pointed out a similar distortion which occurs when extraneous light fields enter the scanning Twyman–Green interferometer. The importance of this new retardation is that the ideal variable retarder acquires a phase error which depends on δ and contributes to the nonlinearity of the interferometer. Moreover, the variable-intensity non-polarizing beam splitter also introduces phase errors indirectly since it is a function of δ, the main measurement parameter, as well. The final phase error results from the mixing of the leaked components with the main components in each arm separately.

Neglecting squares and cross-products of ϵ_x and ϵ_y, the non-ideal Mueller matrix **M** for the interferometer (corresponding to non-ideal Jones matrix **D**) is written as

$$\mathbf{M} = \begin{bmatrix} 1+(\epsilon_x+\epsilon_y)\cos\delta & (\epsilon_x-\epsilon_y)\cos\delta & 0 & 0 \\ (\epsilon_x-\epsilon_y)\cos\delta & 1+(\epsilon_x+\epsilon_y)\cos\delta & 0 & 0 \\ 0 & 0 & -[(\epsilon_x+\epsilon_y)+\cos\delta] & -[(\epsilon_x+\epsilon_y)+\sin\delta] \\ 0 & 0 & (\epsilon_x+\epsilon_y)+\sin\delta & -[(\epsilon_x+\epsilon_y)+\cos\delta] \end{bmatrix}. \tag{E 4}$$

(This must be multiplied by the Mueller matrix for the polarizer–mixer to describe output at the photodetector.) Like the non-ideal Jones matrix **D**, the non-ideal matrix **M** describes a distorted linear variable retarder, as can be seen when **M** reduces to the ideal linear retarder for $\epsilon_x = \epsilon_y = 0$. These results can be viewed as examples of the general parametrization given in section 4.

References

[1] Quenelle, R. C., 1983, *Hewlett–Packard J.*, **31,** 10.
[2] Sutton, C. M., 1987, *J. phys. E,* **20,** 1290.
[3] Bobroff, N., 1987, *Appl. Optics,* **26,** 2676.
[4] Rosenbluth, A. E., and Bobroff, N., 1990, *Precision Engng.*, **12,** 7.
[5] Augustyn, W., and Davis, P., 1990, *J. vac. Sci. Technol.* B, **8,** 2032.
[6] Xie, Y., and Wu, Y.-Z., 1992, *Appl. Optics,* **31,** 881.
[7] Picoto, G. B., and Sacconi, A., 1991, *Quantum Electron. Plasma Phys.*, **29,** 361.
[8] De Freitas, J. M., and Player, M. A., 1993, *Meas. Sci. Technol.*, **4,** 1173.
[9] Bergamin, A., Cavagnero, G., and Mana, G., 1992, *J. mod. Optics,* **39,** 2053.
[10] Xie, Y., and Wu, Y.-Z., 1989, *Appl. Optics,* **28,** 2043.
[11] Berry, M. V., 1984, *Proc. R. Soc.* A, **392,** 45.

[12] BERRY, M. V., 1987, *J. mod. Optics*, **34**, 1401.
[13] KWIAT, P. G., and CHIAO, R. Y., 1991, *Phys. Rev. Lett.*, **66**, 588.
[14] TOMITA, A., and CHIAO, R. Y., 1986, *Phys. Rev. Lett.*, **57**, 937.
[15] FYMAT, A. L., 1972, *Appl. Optics*, **11**, 161.
[16] FYMAT, A. L., 1971, *Appl. Optics*, **10**, 2499.
[17] BORN, M., and WOLF, E., 1964, *Principles of Optics* (New York: Macmillan), sections 1.5 and 10.8.
[18] FYMAT, A. L., 1971, *Appl. Optics*, **10**, 2711.
[19] PLAYER, M. A., 1988, *J. mod. Optics*, **35**, 1813.
[20] SHURCLIFF, W. A., 1962, *Polarised Light: Production and Use* (Cambridge, Massachusetts: Harvard University Press).
[21] AZZAM, R. M. A., and BASHARA, N. M., 1987, *Ellipsometry and Polarised Light* (Amsterdam: North-Holland).
[22] BRUNING, J. H., 1978, *Optical Shop Testing*, edited by D. Malacara (New York: Wiley), chapter 13.