

# Theory of optical long-baseline interferometry on polarized sources

G. Perrin<sup>\*,\*\*\*</sup> 

LIRA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Université Paris Cité, 5 place Jules Janssen, 92195 Meudon, France

Received 24 December 2025 / Accepted 19 February 2026

## ABSTRACT

The effects of the polarization characteristics of beam trains in optical long-baseline interferometers are well known and have led to difficulties in measuring the spatial coherence of astronomical sources in the past. This has been overcome by designing symmetrical optical trains. With the advent of interferometers using large telescopes, observations of faint sources with high degrees of polarization have become even more possible. As in the radio domain, where radiation processes usually lead to high polarization rates, a description of coherence for polarized or unpolarized sources observed with non-polarization neutral interferometers is necessary. A theory of optical long-baseline interferometry fully taking into account the polarization characteristics of beam trains and those of the sources is presented in this paper, building on concepts developed for radio aperture synthesis. The concept of generalized Mueller matrix is introduced for the case of multi-aperture interferometers leading to a simple matrix relationship between the observed Stokes visibilities, as they are disturbed by the instrument polarization characteristics, and the object Stokes visibilities. This relationship is applied to the case of single-mode interferometers. The formalism also shows that classical complex visibilities (squared moduli, phases and closure phases) need to be debiased from polarization crosstalk, even when the source is not polarized as in this case ghost polarized visibilities are created.

**Key words.** instrumentation: high angular resolution – instrumentation: interferometers – methods: analytical – techniques: polarimetric

## 1. Introduction

Long-baseline interferometers measure spatial coherence between distant sampling points. Although the basis of theory for long-baseline interferometry is the same for the radio and optical domains, there are very important differences in their practical implementation and in the processing of data to get images. Because radio waves can be directly measured at the focus of antennas, there is no need to propagate beams down to a common recombination point as is the case in the optical domain, making optical interferometers more complex systems than radio interferometers. A consequence of this important difference is that the polarizing properties of optical interferometers can be very complex, with a few tens of optical components that reflect or transmit the beams. This had huge consequences in the past for the ability to detect fringes, as was experienced by Labeyrie before his first historical detection of long-baseline fringes at optical wavelengths (Labeyrie 1975). This also had consequences for how optical interferometers are designed to avoid the effects of differential birefringence between the beams that can kill fringe contrast, of differential rotation between polarization planes, or of differences in coatings, for example, which can have the same effect (Traub 1988; Rousselet-Perraut et al. 1996): interferometers need to be symmetric, which still increases the number of optical reflections. These effects also have consequences for the process to measure complex

visibilities on polarized sources especially as it is now possible to observe even more sources with high degrees of polarization thanks to the access to much fainter sources where radiation can be produced by processes such as synchrotron radiation (see for example the results of GRAVITY on flares at the Galactic Center in GRAVITY Collaboration (2018)). Some authors have addressed these issues in the past, including Elias (2001, 2004) and Buscher et al. (2009), with Jones and Mueller matrices to propagate polarization states and Stokes parameters across the interferometer, or more recently GRAVITY Collaboration (2024), who studied the polarization characteristics of the Very Large Telescope Interferometer (VLTI) and their effects. Shuai et al. (2025) recently carried out a similar work as in the present paper and applied it to the MIRC-X and MYSTIC instruments of the CHARA interferometer. These works use different notations or definitions, even sometimes for similar quantities. The goal of this paper is to establish a new comprehensive formalism for optical long-baseline interferometry generalized to polarizing interferometers and polarized sources, partly based on what has been done for decades in the radio domain (see for example the monograph on radio interferometry and aperture synthesis by Thompson et al. (2017) for a complete description).

The paper is organized as follows: the basic formalism of coherence in an interferometer for polarized waves is given in Section 2, the propagation of coherence taking into account the polarizing properties of an interferometer is addressed in Section 3, and the consequences for the calibration of complex visibilities for single-mode interferometers is described in Section 4. Examples of instrumental polarization effects are given in Section 5 as are examples of coherence for some

\* Formerly LESIA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Université Paris Cité, 5 place Jules Janssen, 92195 Meudon, France.

\*\* Corresponding author: [guy.perrin@obspm.fr](mailto:guy.perrin@obspm.fr)

specific sources. Some conclusions are given in the last section.

## 2. Formalism of the polarized waves and coherence in an interferometer

In this paper, monochromatic polarized waves propagate in vacuum and are described as complex vectors called Jones vectors of the form

$$\vec{E}(\vec{r}, \lambda, t) = e^{-i(\vec{k} \cdot \vec{r} - \omega t)} \begin{bmatrix} E_x \\ E_y \end{bmatrix}, \quad (1)$$

where  $x$  and  $y$  are the two orthogonal axes along which polarizations are split or analyzed in the beam-combining instrument.

The wave vector,  $\vec{k}$ , is pointing in the direction of propagation and its norm is equal to  $2\pi/\lambda$ . The pulsation,  $\omega$ , relates to the wavelength,  $\lambda$ , through  $\omega = 2\pi c/\lambda$ . Monochromatic waves are not strictly speaking physical and should be integrated across a finite bandpass to define quasi-monochromatic waves in the sense of [Goodman \(2000\)](#) with the consequence of systematically having an additional integral in all the equations to follow. The choice is made to consider monochromatic waves only and ignore temporal coherence effects. In case these effects are important to consider for a specific application, all waves and correlations can be integrated across a given bandwidth a posteriori. Dependency on wavelength is therefore implicit in what follows. Since only spatial coherence is addressed here, correlations between waves are systematically measured at the same times and therefore with no differential time dependency. As a consequence, the  $e^{-i\omega t}$  term is also dropped from now on.

The waves are emitted by astronomical sources with some spatial extension. The above Jones vector is therefore the integral of a spatial distribution of Jones vectors dependent on astronomical coordinates,  $\alpha$  and  $\delta$ , over the spatial extension of the source:

$$\vec{E}(\vec{r}) = \iint_{source} \vec{E}(\alpha, \delta) d\alpha d\delta, \quad (2)$$

where the cursive style is used for the spatial wave distribution. In what follows, the limitation to the source spatial extension is always implicit. In practice it is also limited by the field of view of the instrument or by the fundamental mode in the case of a single-mode interferometer, as is also the case in radio. The waves are now expanded in unpolarized and polarized components:

$$\vec{E}(\vec{r}) = e^{-i\vec{k} \cdot \vec{r}} \begin{bmatrix} E_{np,x} + E_{p,x} \\ E_{np,y} + E_{p,y} \end{bmatrix}. \quad (3)$$

This is done for all beams in an interferometer identified by a superscript,  $n$ , in  $\vec{E}^n(\vec{r})$  for beam  $n$ . Long baseline interferometry is about measuring the spatial coherence of the source at the points where the waves are collected by telescopes. This coherence is mixed with polarization coherence and both can be characterized by the coherence matrix of [Born & Wolf \(1999\)](#):

$$C = \begin{bmatrix} \langle E_x E_x^* \rangle & \langle E_x E_y^* \rangle \\ \langle E_y E_x^* \rangle & \langle E_y E_y^* \rangle \end{bmatrix}, \quad (4)$$

where the superscript  $\star$  is the complex conjugation and the terms of the matrix are correlations between the wave components.

Note that this matrix can also be written as the average product of the wave vector by its conjugate transpose noted with the dagger:

$$C = \langle \vec{E} \cdot \vec{E}^\dagger \rangle. \quad (5)$$

In the optical, the terms of the matrix can be measured by splitting polarizations with a Wollaston prism ( $\langle E_x E_x^* \rangle$  and  $\langle E_y E_y^* \rangle$ ), by rotating the polarizations by  $45^\circ$  with a half-wave plate before splitting ( $\langle E_x E_y^* + E_y E_x^* \rangle$ ), and by applying a quarter wave-plate rotated by  $45^\circ$  before splitting ( $\langle E_x E_y^* - E_y E_x^* \rangle$ ). The total flux from the source is obtained by taking the trace of the coherence matrix  $I = \text{tr}(C) = \langle |E_x|^2 \rangle + \langle |E_y|^2 \rangle$ . The polarized intensity is  $PI = \langle |E_{p,x}|^2 \rangle + \langle |E_{p,y}|^2 \rangle$ , with  $P$ , the degree of polarization.

Since the polarized and unpolarized waves are not coherent between each other and since the  $x$  and  $y$  components of the unpolarized waves are incoherent, the matrix can also be written:

$$C = \begin{bmatrix} \langle E_{np,x} E_{np,x}^* \rangle & 0 \\ 0 & \langle E_{np,y} E_{np,y}^* \rangle \end{bmatrix} + \begin{bmatrix} \langle E_{p,x} E_{p,x}^* \rangle & \langle E_{p,x} E_{p,y}^* \rangle \\ \langle E_{p,y} E_{p,x}^* \rangle & \langle E_{p,y} E_{p,y}^* \rangle \end{bmatrix}. \quad (6)$$

For the unpolarized part of the wave, the first matrix reads as

$$C_{np} = (1 - P) \frac{I}{2} \times I_2, \quad (7)$$

with  $I_2$  the  $2 \times 2$  identity matrix. The second matrix can be written with the Stokes parameters  $Q$ ,  $U$ , and  $V$ :

$$C_p = \frac{1}{2} \begin{bmatrix} PI + Q & U + iV \\ U - iV & PI - Q \end{bmatrix}. \quad (8)$$

The definition used for the Stokes parameters is the one of [Hamaker & Bregman \(1996\)](#). It is aligned with the IAU convention of 1974 when  $x$  is the declination axis increasing toward the north and  $y$  the right ascension axis increasing toward the east, and takes into account the definition of the harmonic exponential in Eq. (1) for the sign and expression of the  $V$  Stokes parameter. [Shuai et al. \(2025\)](#) have chosen different conventions; hence, the differences in signs in the coherence matrix. Note that the trace of this part of the coherence matrix is equal to the polarized intensity:  $\text{tr}(C_p) = PI$ . The sum of the two matrices equates to half the Stokes brightness matrix of [Hamaker \(2000\)](#):

$$C = \frac{1}{2} \begin{bmatrix} I + Q & U + iV \\ U - iV & I - Q \end{bmatrix}. \quad (9)$$

Stokes parameter spatial distributions can also be defined assuming the emitter is spatially incoherent, the products of double integrals in Eq. (4) being replaced by a single double integral of the products of wave components as was done in [Goodman \(2000\)](#):

$$\begin{cases} I = \iint I(\alpha, \delta) d\alpha d\delta = \iint \langle \mathcal{E}_x \mathcal{E}_x^* + \mathcal{E}_y \mathcal{E}_y^* \rangle d\alpha d\delta \\ Q = \iint Q(\alpha, \delta) d\alpha d\delta = \iint \langle \mathcal{E}_{p,x} \mathcal{E}_{p,x}^* - \mathcal{E}_{p,y} \mathcal{E}_{p,y}^* \rangle d\alpha d\delta \\ U = \iint U(\alpha, \delta) d\alpha d\delta = \iint \langle \mathcal{E}_{p,x} \mathcal{E}_{p,y}^* + \mathcal{E}_{p,y} \mathcal{E}_{p,x}^* \rangle d\alpha d\delta \\ V = \iint V(\alpha, \delta) d\alpha d\delta = \iint -i \langle \mathcal{E}_{p,x} \mathcal{E}_{p,y}^* - \mathcal{E}_{p,y} \mathcal{E}_{p,x}^* \rangle d\alpha d\delta \end{cases}. \quad (10)$$

These relations can be extended to the case of an interferometer with correlations in Eq. (4) replaced by cross-correlations between beams  $n$  and  $m$  as was done by Hamaker & Bregman (1996):

$$C^{nm} = \begin{bmatrix} \langle E_x^n E_x^{m*} \rangle & \langle E_x^n E_y^{m*} \rangle \\ \langle E_y^n E_x^{m*} \rangle & \langle E_y^n E_y^{m*} \rangle \end{bmatrix}, \quad (11)$$

yielding

$$C^{nm} = \frac{I}{2} V_I^{nm} I_2 + \frac{I}{2} \begin{bmatrix} V_Q^{nm} & V_U^{nm} + iV_V^{nm} \\ V_U^{nm} - iV_V^{nm} & -V_Q^{nm} \end{bmatrix}, \quad (12)$$

where  $V_I^{nm}$  is the classical visibility defined in fundamental monographs dealing with coherence (Goodman 2000; Born & Wolf 1999) and given by the Zernike-van Cittert theorem. In this paper, coherence is associated with correlations of waves and is a coherent flux, while visibilities are systematically associated with normalized correlations, and therefore have no dimensions and are unitless. This is applied in the same spirit to the generalization of the visibility for the Stokes parameters  $V_Q$ ,  $V_U$ , and  $V_V$ . Note that, in the radio domain, the normalization of these quantities is different in other papers or between papers, in which the coherence matrix can be called a visibility matrix; for example, with a factor of 2 in addition (this is the case in Smirnov (2011) and Hamaker (2000), for example). Also, the Stokes visibilities were first introduced by Morris et al. (1964), in which for example the complex visibility function for the brightness distribution of  $Q$  was normalized by the integral of  $Q$ , which is to say by  $Q$ ; similar expressions were given in Conway & Kronberg (1969) but were not normalized. There is therefore nothing new introduced at this stage in this paper except a different and consistent way to define the physical quantities used in this work. There are physical reasons for that. Because of the effect of atmospheric turbulence in the optical, the flux of photons injected in single-mode fibers is fluctuating with time and is also affected by static aberrations. The coherent flux is therefore fluctuating, while the normalized coherent flux is stable in single-mode interferometry. This is less the case when going to longer wavelengths, starting in the mid-infrared where coherence fluctuations are less severe and where it becomes possible to work with coherent fluxes as in radio. In what follows, coherent fluxes (correlations) are therefore systematically normalized with respect to total intensity to get visibilities, which is a difference compared to how the data of radio interferometers are calibrated. Hence the definitions used in this paper:

$$\begin{cases} V_I(u, v) = \iint I(\alpha, \delta) e^{-2i\pi(\alpha u + \delta v)} d\alpha d\delta / I \\ V_Q(u, v) = \iint Q(\alpha, \delta) e^{-2i\pi(\alpha u + \delta v)} d\alpha d\delta / I \\ V_U(u, v) = \iint U(\alpha, \delta) e^{-2i\pi(\alpha u + \delta v)} d\alpha d\delta / I \\ V_V(u, v) = \iint V(\alpha, \delta) e^{-2i\pi(\alpha u + \delta v)} d\alpha d\delta / I \end{cases} \quad (13)$$

$(u, v)$  are the coordinates of the spatial frequencies given by the projection of  $\vec{r}_n - \vec{r}_m$  on the plane of the sky in the direction of the object according to the Zernike-van Cittert theorem (a quadratic phasor needs to be introduced in the radio domain in these expressions as the mode of the antennas gives access to a much wider field of view). As for the traditional Stokes parameters, the Stokes visibilities being second-order moments are only sensitive to the differential phases between the waves.

The coherence matrix above mixes both spatial and polarization coherences. It can be obtained without assuming that the source is point-like. It is clear from Eq. (12) that the classical visibility,  $V_I^{nm}$ , is biased by the polarization characteristics of the

source if these are not taken into account to properly measure it. The Stokes visibilities of the object can be fully determined by inversion of Eq. (12), which can be rewritten as

$$\begin{bmatrix} \langle E_x^n E_x^{m*} \rangle \\ \langle E_x^n E_y^{m*} \rangle \\ \langle E_y^n E_x^{m*} \rangle \\ \langle E_y^n E_y^{m*} \rangle \end{bmatrix} = I \times \frac{1}{2} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & i \\ 0 & 0 & 1 & -i \\ 1 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} V_I^{nm} \\ V_Q^{nm} \\ V_U^{nm} \\ V_V^{nm} \end{bmatrix}. \quad (14)$$

The  $4 \times 4$  matrix times  $1/2$  is called  $T_{I_2}$  and is used in Section 3. In what follows, the column on the left is called the coherence vector or the cross-coherence vector. When  $n = m$ , it is called the self-coherence vector. The determinant of the  $4 \times 4$  matrix is equal to  $i/4$  and is nonzero. Note that the imaginary determinant encodes the  $\pi/2$  phase difference between linear and circular analyzer bases and has no other physical meaning. Note also that for unpolarized sources  $Q$ ,  $U$ , and  $V$  are zero and  $V_I^{nm}$  is directly measured with the first and/or last correlation or the sum of the two, the problem then being scalar and not vectorial anymore. In the general case, the matrix can be inverted and the Stokes visibilities can be obtained from the components of the coherence matrix through the relation

$$\begin{bmatrix} V_I^{nm} \\ V_Q^{nm} \\ V_U^{nm} \\ V_V^{nm} \end{bmatrix} = \frac{1}{I} \times \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & -i & i & 0 \end{bmatrix} \begin{bmatrix} \langle E_x^n E_x^{m*} \rangle \\ \langle E_x^n E_y^{m*} \rangle \\ \langle E_y^n E_x^{m*} \rangle \\ \langle E_y^n E_y^{m*} \rangle \end{bmatrix}. \quad (15)$$

The other particular case is when the source is only linearly polarized as circularly polarized light is often much fainter than linearly polarized light. In this case in which  $V$  can be neglected with respect to  $Q$  and  $U$ ,  $V$  is set equal to zero and only three equations need to be inverted, which can be written as

$$\begin{bmatrix} \langle E_x^n E_x^{m*} \rangle \\ \langle E_x^n E_y^{m*} \rangle + \langle E_y^n E_x^{m*} \rangle \\ \langle E_y^n E_y^{m*} \rangle \end{bmatrix} = I \times \frac{1}{2} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 2 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} V_I^{nm} \\ V_Q^{nm} \\ V_U^{nm} \end{bmatrix}, \quad (16)$$

the determinant of the  $3 \times 3$  matrix now being 4 and the middle coherence term being given by the difference of the Wollaston polarizer outputs after the action of a  $22.5^\circ$  rotated half-wave plate. The visibilities are then given by

$$\begin{bmatrix} V_I^{nm} \\ V_Q^{nm} \\ V_U^{nm} \end{bmatrix} = \frac{1}{I} \times \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \langle E_x^n E_x^{m*} \rangle \\ \langle E_x^n E_y^{m*} \rangle + \langle E_y^n E_x^{m*} \rangle \\ \langle E_y^n E_y^{m*} \rangle \end{bmatrix}. \quad (17)$$

### 3. Propagation of coherence in an optical interferometer

The propagation of waves in an optical system can be described with the formalism of  $2 \times 2$  complex Jones matrices. If the input wave collected by a telescope is denoted as  $\vec{E}$  and the optical system to transport the collected beam down to the beam combination point is described by the matrix  $J$ , then the exit beam reads as

$$\vec{E} \sim J \vec{E}. \quad (18)$$

The  $\sim$  symbol is used for propagated quantities in this paper. Using Eq. (5) for the definition of the coherence matrix from the wave vectors, one gets that the coherence matrix,  $C^{nm}$ , is

transported as the following propagated coherence matrix across the optical system:

$$\tilde{C}^{nm} = J^n C^{nm} J^{m\dagger}. \quad (19)$$

The above equation is known as the radio interferometer measurement equation (RIME), in for example Smirnov (2011), and was first introduced by Hamaker et al. (1996). The equivalent of Eq. (12) after propagation of the beams is written as

$$\tilde{C}^{nm} = \frac{I}{2} V_I^{nm} J^n \cdot J^{m\dagger} + \frac{I}{2} J^n \cdot \begin{bmatrix} V_Q^{nm} & V_U^{nm} + iV_V^{nm} \\ V_U^{nm} - iV_V^{nm} & -V_Q^{nm} \end{bmatrix} \cdot J^{m\dagger}. \quad (20)$$

Since in the radio domain the effect of instrumentation on the polarization can be more easily described by a few complex factors, the Jones propagation matrices are quickly expanded as the product of a few simple matrices at this stage.

As in Eq. (12), the equation is written splitting the classical part with the source visibility,  $V_I^{nm}$ , from the part with the Stokes visibilities, which can be understood as a bias if they are neglected. This equation therefore has a major consequence: even in the case of an unpolarized source, the propagated coherence matrix may have nonzero antidiagonal terms, which should only exist for polarized sources. It therefore exhibits some polarization crosstalk, which creates biases and ghost polarized visibilities that need to be subtracted in the calibration process.

This relation can be rearranged in a system of equations:

$$\begin{bmatrix} \langle \tilde{E}_x^n \tilde{E}_x^{m*} \rangle \\ \langle \tilde{E}_x^n \tilde{E}_y^{m*} \rangle \\ \langle \tilde{E}_y^n \tilde{E}_x^{m*} \rangle \\ \langle \tilde{E}_y^n \tilde{E}_y^{m*} \rangle \end{bmatrix} = I \times T^{nm} \begin{bmatrix} V_I^{nm} \\ V_Q^{nm} \\ V_U^{nm} \\ V_V^{nm} \end{bmatrix}. \quad (21)$$

In what follows, the  $4 \times 1$  coherence and visibility column vectors of the above equation are denoted as  $\tilde{C}^{nm}$  and  $\mathcal{V}^{nm}$  and the following equation is written:

$$\tilde{C}^{nm} = I \times T^{nm} \cdot \mathcal{V}^{nm}. \quad (22)$$

A similar expression can be derived with a  $3 \times 3$  matrix in the case in which circular polarization can be neglected. The transfer matrix,  $T^{nm}$ , expression is the following (splitting the 4x4 matrix into two lines):

$$T^{nm} = \frac{1}{2} \begin{bmatrix} [J_{11}^n J_{11}^{m*} + J_{12}^n J_{12}^{m*}] & [J_{11}^n J_{11}^{m*} - J_{12}^n J_{12}^{m*}] \\ [J_{11}^n J_{21}^{m*} + J_{12}^n J_{22}^{m*}] & [J_{11}^n J_{21}^{m*} - J_{12}^n J_{22}^{m*}] \\ [J_{21}^n J_{11}^{m*} + J_{22}^n J_{12}^{m*}] & [J_{21}^n J_{11}^{m*} - J_{22}^n J_{12}^{m*}] \\ [J_{21}^n J_{21}^{m*} + J_{22}^n J_{22}^{m*}] & [J_{21}^n J_{21}^{m*} - J_{22}^n J_{22}^{m*}] \end{bmatrix} \quad (23)$$

$$\begin{bmatrix} [J_{11}^n J_{12}^{m*} + J_{12}^n J_{11}^{m*}] & i[J_{11}^n J_{12}^{m*} - J_{12}^n J_{11}^{m*}] \\ [J_{11}^n J_{22}^{m*} + J_{12}^n J_{21}^{m*}] & i[J_{11}^n J_{22}^{m*} - J_{12}^n J_{21}^{m*}] \\ [J_{21}^n J_{12}^{m*} + J_{22}^n J_{11}^{m*}] & i[J_{21}^n J_{12}^{m*} - J_{22}^n J_{11}^{m*}] \\ [J_{21}^n J_{22}^{m*} + J_{22}^n J_{21}^{m*}] & i[J_{21}^n J_{22}^{m*} - J_{22}^n J_{21}^{m*}] \end{bmatrix}.$$

The columns of  $T^{nm}$  can be computed very conveniently with the Pauli matrices,  $\sigma_0, \sigma_1, \sigma_2, \sigma_3$  (with  $\sigma_0 = I_2$ ). As matter of fact, the first column is obtained by reorganizing the terms of  $\text{Col}_1 = J^n \sigma_0 J^{m\dagger}$  into a single column with on the first line  $\text{Col}_1[1,1]$ , on the second line  $\text{Col}_1[1,2]$ , on the third line  $\text{Col}_1[2,1]$ , and on the last line  $\text{Col}_1[2,2]$ . The same is then done for  $\text{Col}_2 = J^n \sigma_3 J^{m\dagger}$ ,  $\text{Col}_3 = J^n \sigma_1 J^{m\dagger}$ , and  $\text{Col}_4 = J^n \sigma_2 J^{m\dagger}$ .

The  $T_{I_2}$  matrix introduced in Eq. (14) of Section 2 is equal to the transfer matrix,  $T^{nm}$ , in the special case in which  $J_n = J_m = I_2$ . And one gets the 3x3 matrix of Eq. (16) by suppressing the

fourth column and adding the second and third lines. For  $n = m$ ,  $\mathcal{S} = I \mathcal{V}^{mm}$  is the Stokes parameter vector (it is independent of  $n$ ) and  $T_{I_2}^{-1} T^{mm}$  is the Mueller matrix describing the propagation of the Stokes parameters for beam  $n$ . By analogy, I defined the generalized Mueller matrix,  $M^{nm} = T_{I_2}^{-1} T^{nm}$ , to describe the propagation of the Stokes visibilities in the optical system according to the equation

$$\tilde{\mathcal{V}}^{nm} = M^{nm} \cdot \mathcal{V}^{nm}. \quad (24)$$

## 4. Calibration of visibilities of polarized beams

### 4.1. Visibility moduli

#### 4.1.1. Case of polarized sources

In this section, it is assumed that the interferometer is single-mode, i.e., that the beams are filtered with single-mode fibers before beam combination ensuring that phase fluctuations in the pupil plane are traded against intensity fluctuations which can be calibrated, as demonstrated by the FLUOR instrument (Coudé du Foresto et al. 1997; Perrin et al. 1998). In practice, the Jones matrices,  $J^n$ , are measured to within a complex factor, which can be called a gain and denoted as  $g_n$ . Eq. (24) is then modified as

$$\tilde{\mathcal{V}}^{nm} = g_n g_m^* M^{nm} \cdot \mathcal{V}^{nm}, \quad (25)$$

or, equivalently, since  $T^{nm} = T_{I_2} \cdot M^{nm}$ , Eq. (22) can be modified as

$$\tilde{C}^{nm} = I g_n g_m^* T^{nm} \cdot \mathcal{V}^{nm}. \quad (26)$$

The same applies for the self-coherence measured for each beam, which here in the case of beam  $n$  is

$$\tilde{\mathcal{P}}^n = |g_n|^2 T^{nn} \cdot \mathcal{S}. \quad (27)$$

$\mathcal{S} = I \mathcal{V}^{nn} = I \mathcal{V}^{nn}$  is the Stokes parameter vector. Some extra gains may need to be introduced depending on the type of beam combination and the calibration protocol, but these are calibrated on an internal source and are not mentioned here to keep formulas as simple as possible. The self-coherence for beam  $n$  is denoted as  $\mathcal{P}^n$ , signifying photometry for beam  $n$  for a reason that will become obvious in the next section. The equation is modified by multiplying the two members by the inverse of matrix  $T^{nn}$  and by taking the scalar product with  $\mathcal{S}$ :

$$\mathcal{S}^\dagger \cdot T^{nn-1} \cdot \tilde{\mathcal{P}}^n = |g_n|^2 \mathcal{S}^\dagger \cdot \mathcal{S}. \quad (28)$$

$\mathcal{S}^\dagger \cdot \mathcal{S}$  takes the simple expression

$$\mathcal{S}^\dagger \cdot \mathcal{S} = I^2 + Q^2 + U^2 + V^2 = (1 + P^2) I^2; \quad (29)$$

hence, the value of the square of the modulus of the gain,

$$|g_n|^2 = \frac{1}{(1 + P^2) I^2} \mathcal{S}^\dagger \cdot T^{nn-1} \cdot \tilde{\mathcal{P}}^n. \quad (30)$$

The same is obtained for beam  $m$ . Eq. (25) can be inverted to

$$\mathcal{V}^{nm} = \frac{1}{g_n g_m^*} M^{nm-1} \cdot \tilde{\mathcal{V}}^{nm}. \quad (31)$$

And one gets for the moduli of the Stokes visibilities (the modulus applied to the vector is the modulus of its components):

$$|\mathcal{V}^{nm}| = \frac{I^2 (1 + P^2) |M^{nm-1} \cdot \tilde{\mathcal{V}}^{nm}|}{\sqrt{\mathcal{S}^\dagger \cdot T^{nn-1} \cdot \tilde{\mathcal{P}}^n} \sqrt{\mathcal{S}^\dagger \cdot T^{mm-1} \cdot \tilde{\mathcal{P}}^m}}. \quad (32)$$

Note that the measurement of the Stokes parameters usually leads to  $\mathcal{S}$  normalized by the measured total intensity. In the same vein, either one measures the coherent Stokes flux,  $I\mathcal{V}^{nm}$ , and the self-coherent flux,  $\tilde{\mathcal{P}}^m$ , or the Stokes visibilities,  $\tilde{\mathcal{V}}^{nm}$ , and the normalized self-coherent flux,  $\tilde{\mathcal{P}}^m/I$ . In either case, the modulus of the Stokes visibility is independent of the absolute value of the total intensity,  $I$ , in the above formula.

#### 4.1.2. Case of unpolarized sources

In absence of polarization crosstalk, i.e., when  $\langle \tilde{E}_x^n \tilde{E}_y^{m*} \rangle = \langle \tilde{E}_y^n \tilde{E}_x^{m*} \rangle = 0$ , the above equations take a simpler form in the case in which the degree of polarization is  $P = 0$  or, equivalently, when the Stokes parameters  $Q = U = V = 0$ . This leads to scalar values for the visibilities, coherence and self-coherence as in this case the visibility is only  $V_I^{nm}$  and the first and last element of the self-coherence vector are the only nonzero terms and can be summed to obtain the total intensity detected by the instrument for a given beam, and hence the term of photometry in this case. The same applies as well for the cross-coherence and one sums the first and last terms of the coherence vector to get the total coherent intensity  $I^{nm} = \tilde{I}_x^{nm} + \tilde{I}_y^{nm}$ . As a particular case of the above equations, one gets for the coherent intensity

$$\tilde{I}^{nm} = \frac{I}{2} V_I^{nm} \left( \sum_{i,j} J_{ij}^n J_{ij}^{m*} \right) \times g_n g_m^*, \quad (33)$$

and for the photometries

$$\tilde{\mathcal{P}}^{n,m} = \frac{I}{2} \left( \sum_{i,j} |J_{ij}^{n,m}|^2 \right) \times |g_{n,m}|^2. \quad (34)$$

So that the equivalent of Eq. (32) is written as

$$|V_I^{nm}| = \frac{|\tilde{I}^{nm}|}{\sqrt{\tilde{\mathcal{P}}^n \tilde{\mathcal{P}}^m}} \times \frac{\sqrt{\sum_{i,j} |J_{ij}^n|^2 \sum_{i,j} |J_{ij}^m|^2}}{|\sum_{i,j} J_{ij}^n J_{ij}^{m*}|}. \quad (35)$$

The ratio after the times sign is the reciprocal of  $\sqrt{\kappa_n \kappa_m}$  of Coudé du Foresto et al. (1997).

In case of polarization crosstalk, as in Buscher et al. (2009), these expressions need to be debiased using the 4D Stokes visibility formalism presented in Section 4.1.1 as ghost polarized visibilities are produced by the optical system.

#### 4.2. Visibility phases and closure phases

Eq. (25) can be modified multiplying both sides by  $M^{nm-1}$ , yielding

$$M^{nm-1} \cdot \tilde{\mathcal{V}}^{nm} = g_n g_m^* \mathcal{V}^{nm}. \quad (36)$$

This equation shows how to get access to visibilities from which the effects of polarization crosstalks have been removed. The remaining phase bias due to the gain product,  $g_n g_m^*$ , can be calibrated out by subtracting the phase observed on a reference source in the case of a dual-field instrument such as GRAVITY.

For single-field instruments, the above relation allows one to obtain an unbiased closure phase quantity. Taking the Hadamard product of the above column vector for baselines  $nm$ ,  $ml$ , and  $ln$ , one gets

$$\begin{aligned} (M^{nm-1} \cdot \tilde{\mathcal{V}}^{nm}) \odot (M^{ml-1} \cdot \tilde{\mathcal{V}}^{ml}) \odot (M^{ln-1} \cdot \tilde{\mathcal{V}}^{ln}) \\ = |g_n g_m g_l|^2 \mathcal{V}^{nm} \odot \mathcal{V}^{ml} \odot \mathcal{V}^{ln}, \end{aligned} \quad (37)$$

and therefore the closure phase relation for all Stokes visibilities:

$$\begin{aligned} \text{Arg} \left( (M^{nm-1} \cdot \tilde{\mathcal{V}}^{nm}) \odot (M^{ml-1} \cdot \tilde{\mathcal{V}}^{ml}) \odot (M^{ln-1} \cdot \tilde{\mathcal{V}}^{ln}) \right) \\ = \text{Arg} \left( \mathcal{V}^{nm} \odot \mathcal{V}^{ml} \odot \mathcal{V}^{ln} \right). \end{aligned} \quad (38)$$

## 5. Examples of polarized waves and of interferometer polarizing properties

In this section, some specific cases are given to illustrate the above theory and also make a link with some previous publications.

### 5.1. Effect of beam rotation and birefringence

The process of detection is a major difference between radio interferometry and optical interferometry as in the latter case only intensity can be detected (except for wavelengths beyond 10  $\mu\text{m}$  where heterodyne techniques can be employed, as was the case with the Infrared Spatial Interferometer Hale et al. (2000)) and therefore beams must be transported down to the beam combination point. A consequence is the complexity of beam trains, which must include, on top of that, delay lines to compensate for the external variable optical path difference. Modern optical interferometers are almost perfectly symmetric but still have some imperfections, causing small contrast losses. In Perrin (2024), I showed that, at least in the case of the VLTI, the Jones matrices of beam trains can be approximated with a good accuracy by what I called quasi-unitary matrices of the form

$$U^q(\theta, \psi, \varphi_1, \varphi_2, \tau_1, \tau_2) = R_\theta R_\psi \Delta(\tau_1 e^{i\varphi_1}, \tau_2 e^{i\varphi_2}) R_{-\psi}, \quad (39)$$

where  $R_\theta$  is a global rotation of angle  $\theta$ , and  $\Delta$  is a diagonal matrix describing the polarizing properties of the system (diattenuation  $\tau_2 - \tau_1$  and birefringence  $\varphi_2 - \varphi_1$ ) described for its neutral axis rotated by the angle  $\psi$  with respect to the reference axes. This case was studied by Rousselet-Perraut et al. (1997) for the GI2T interferometer considering identical neutral axes orientations in all beams but with different phases and coefficients and different global rotations using the following quasi-unitary matrix for beam  $i$ :  $J_i = U^q(\theta_i, 0, \Phi_{s,i}, \Phi_{p,i}, R_{s,i}, R_{p,i})$ . In a simpler version, they considered the effects of differential beam rotation only ( $J_i = U^q(\theta_i, 0, 0, 0, 1, 1)$ ) and of differential birefringence only ( $J_i = U^q(0, 0, \phi_{s,i}, \phi_{p,i}, 1, 1)$ ), first assuming no attenuation and then mixing all effects (Rousselet-Perraut et al. 1996). Shuai et al. (2025) have modeled the polarization properties of CHARA by decomposing the beam train in groups of optical elements, each described by unitary or quasi-unitary Jones matrices with beam rotation, diattenuation, and retardance, and applied their model to observations of polarized sources. In Perrin (2025), I also considered quasi-unitary matrixes to discuss biases produced by retardance and diattenuation in the measurements of long-baseline interferometers for point-like sources.

### 5.2. Case of symmetric interferometers without diattenuation

For such interferometers, all the Jones matrices are exactly proportional to unitary matrices and are of the type:  $J_i = U^q(\theta, 0, \varphi_1, \varphi_2, |g|, |g|) = |g| U(\theta, 0, \varphi_1, \varphi_2)$ , independently of  $i$  and where  $U(\theta, 0, \varphi_1, \varphi_2)$  is a unitary matrix (all unitary matrices can be written as the product of rotation matrices with a diagonal

unitary matrix, see Perrin (2024)). Eq. (20) becomes in this case

$$\tilde{C}^{nm} = |g|^2 \frac{I}{2} V_I^{nm} I_2 + |g|^2 \frac{I}{2} U \cdot \begin{bmatrix} V_Q^{nm} & V_U^{nm} + iV_V^{nm} \\ V_U^{nm} - iV_V^{nm} & -V_Q^{nm} \end{bmatrix} \cdot U^\dagger. \quad (40)$$

The effect on  $V_I^{nm}$  of the polarization properties of the interferometer cancels but it does not cancel for the polarized part of the visibilities. This shows that even when an interferometer is perfectly symmetric and without diattenuation, the optical beam properties need to be taken into account to calibrate the visibilities, unless the degree of polarization of the source is null.

### 5.3. Case of a polarized point source

Here, the polarized emitter is a point source, in addition to an unpolarized spatial intensity distribution. The Stokes visibilities are constant and written as  $Q/I$ ,  $U/I$  and  $V/I$ . Hence the expression of Eq. (20) in this case is

$$\tilde{C}^{nm} = \frac{I}{2} V_I^{nm} J^n \cdot J^{m\dagger} + \frac{1}{2} J^n \cdot \begin{bmatrix} Q & U + iV \\ U - iV & -Q \end{bmatrix} \cdot J^{m\dagger}. \quad (41)$$

As in traditional interferometry, the visibility of the unpolarized source is offset by the visibility of the polarized point source but also with an additional term that is proportional to the Stokes parameters of the point source, the second term in the above equation.

### 5.4. Case of sources with uniform polarization properties

It is now supposed that the polarized emitter is an extended source but with uniform polarization properties across the object. In this case, the Stokes visibilities (except for  $I$ ) are proportional to the Fourier transform of the support of the object,  $A(\alpha, \delta)$ , whose integral is the area covered by the object,  $\mathcal{A}$ , and are written as

$$\begin{cases} V_Q(u, v) = QFT[A](u, v)/(\mathcal{A}I) \\ V_U(u, v) = UFT[A](u, v)/(\mathcal{A}I) \\ V_V(u, v) = VFT[A](u, v)/(\mathcal{A}I) \end{cases} \quad (42)$$

Eq. (20) then becomes

$$\tilde{C}^{nm} = \frac{I}{2} V_I^{nm} J^n \cdot J^{m\dagger} + \frac{1}{2} \frac{FT[A]}{\mathcal{A}} J^n \cdot \begin{bmatrix} Q & U + iV \\ U - iV & -Q \end{bmatrix} \cdot J^{m\dagger}. \quad (43)$$

The more extended the object, the narrower the visibility function,  $\frac{FT[A]}{\mathcal{A}}$ . This function becomes negligible past a spatial frequency that is proportional to the reciprocal of size of the object. Beyond this value, the cross-coherence is almost not biased by the Stokes visibilities but the visibilities are biased by the source polarization after the photometric calibration if this term is neglected.

## 6. Conclusions

Based on previous results on aperture synthesis in radio astronomy, a full and consistent formalism for long baseline inter-

ferometry in the optical domain, taking into account the polarization properties of the interferometer and the polarimetric characteristics of the source, is established. Spatial coherence is shown to be linked to the Stokes visibilities through a simple matrix relationship for each baseline of the interferometer. The Stokes visibilities are a generalization of the classical visibility function but they are applied to the Stokes parameters, and therefore they describe the spatial distribution of the polarimetric characteristics of the astronomical source. A generalized Mueller matrix is introduced to describe the propagation of the Stokes visibilities across the instrument as an analogy with the Mueller matrix that describes the propagation of the Stokes parameters in the single-beam case. The matrix describes the polarization properties of the interferometer for a given baseline using the Jones formalism. This simple formula shows that both the polarimetry characteristics of the source and those of the instrument need to be taken into account to calibrate the visibilities. It is in particular shown that complex visibilities and derived quantities such as the squared moduli, differential phases, and closure phases need to be debiased from polarization crosstalk, even when the source is not polarized as in this case ghost polarized visibilities are produced by the instrument and need to be removed in the calibration process using the generalized Mueller matrix. The matrix formalism therefore provides a method of deriving the object Stokes visibilities from the observed visibilities, and therefore yields constraints on the spatial distribution of the Stokes parameters. The principles of the calibration in the case of a single-mode interferometer are presented in the paper.

*Acknowledgements.* The author is thankful to Steve Ridgway, Karine Perraut and Oleg Smirnov for their careful reading and their suggestions to improve the article. The author also thanks the anonymous referee for the very useful comments that led to the last version of the paper.

## References

- Born, M., & Wolf, E. 1999, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light* (Cambridge University Press), 7th edn.
- Buscher, D., Baron, F., & Haniff, C. 2009, *PASP*, **121**, 45
- Conway, R. G., & Kronberg, P. P. 1969, *MNRAS*, **142**, 11
- Coudé du Foresto, V., Ridgway, S., & Mariotti, J. M. 1997, *A&AS*, **121**, 379
- Elias, Nicholas M., I. 2001, *ApJ*, **549**, 647
- Elias, Nicholas M., I. 2004, *ApJ*, **611**, 1175
- Goodman, J. W. 2000, *Statistical Optics* (Wiley)
- GRAVITY Collaboration (Abuter, R., et al.) 2018, *A&A*, **618**, L10
- GRAVITY Collaboration (Widmann, F., et al.) 2024, *A&A*, **681**, A115
- Hale, D. D. S., Bester, M., Danchi, W. C., et al. 2000, *ApJ*, **537**, 998
- Hamaker, J. P. 2000, *A&AS*, **143**, 515
- Hamaker, J. P., & Bregman, J. D. 1996, *A&AS*, **117**, 161
- Hamaker, J. P., Bregman, J. D., & Sault, R. J. 1996, *A&AS*, **117**, 137
- Labeyrie, A. 1975, *ApJ*, **196**, L71
- Morris, D., Radhakrishnan, V., & Seielstad, G. A. 1964, *ApJ*, **139**, 551
- Perrin, G. 2024, *J. Opt. Soc. Am. A*, **41**, 708
- Perrin, G. 2025, *A&A*, **695**, A157
- Perrin, G., Coudé du Foresto, V., Ridgway, S. T., et al. 1998, *A&A*, **331**, 619
- Rousselot-Perraut, K., Mourard, D., & Vakili, F. 1997, *Opt. Eng.*, **36**, 980
- Rousselot-Perraut, K., Vakili, F., & Mourard, D. 1996, *Opt. Eng.*, **35**, 2943
- Shuai, L., Monnier, J. D., Setterholm, B. R., et al. 2025, *AJ*, **170**, 344
- Smirnov, O. M. 2011, *A&A*, **527**, A106
- Thompson, A. R., Moran, J. M., & Swenson, G. W. 2017, *Interferometry and Synthesis in Radio Astronomy*, 3rd edn., Astronomy and astrophysics library (Cham, Switzerland: Springer International Publishing)
- Traub, W. A. 1988, in *European Southern Observatory Conference and Workshop Proceedings*, 29, ed. F. Merkle, 1029