Understanding radio polarimetry.
III. Interpreting the IAU/IEEE definitions of the Stokes parameters

J.P. Hamaker and J.D. Bregman

Netherlands Foundation for Research in Astronomy, Postbus 2, 7990 AA Dwingeloo, The Netherlands

Abstract. — In two companion papers (Paper I, Hamaker et al. 1996; Paper II, Sault et al. 1996), a new theory of radio-interferometric polarimetry and its application to the calibration of interferometer arrays are presented. To complete our study of radio polarimetry, we examine here the definition of the Stokes parameters adopted by Commission 40 of the IAU (1974) and the way this definition works out in the mathematical equations. Using the formalism of Paper I, we give a simplified derivation of the frequently-cited 'black-box' formula originally derived by Morris et al. (1964). We show that their original version is in error in the sign of Stokes V, the correct sign being that given by Weiler (1973) and Thompson et al. (1986).

1. Introduction

In a companion paper (Hamaker et al. 1996, Paper I) we have presented a theory that describes the operation of a polarimetric radio interferometer in terms of the properties of its constituent elements and in doing so unifies the heretofore disjoint realms of radio and optical polarimetry. In a second paper (Sault et al., Paper II) we apply this theory along with theorems borrowed from optical polarimetry to the problem of calibrating an interferometer array such as an aperture-synthesis telescope.

In practical applications, the theory must be supplemented by precise definitions of the coordinate frames and the Stokes parameters that are used. This problem was first addressed by the Institute of Radio Engineers in 1942; the most recent version of their definition was published in 1969 (IEEE 1969). For radio-astronomical applications, the IAU (1974) endorses the IEEE standard, supplementing it with definitions of the Cartesian coordinate frame shown in Fig. 1 and of the sign of the Stokes parameter V.

Most published work on actual polarimetric interferometer observations infers the source's Stokes-parameter brightness distributions from a formula derived by Morris et al. (1964). Weiler (1973) rederives their result, agreeing except for the sign of Stokes V. Thompson et al. (1987)

include his version in their textbook, even though they suggest in their wording that they agree with Morris et al. Clearly the situation needs to be clarified; starting from a complete interpretation of the definitions, we are in a good position to do so. We shall show Weiler's version indeed to be the correct one.

2. The Stokes parameters in a single point in the field

The definition of the Stokes parameters most frequently found in the literature is in terms of the auto- and cross-correlations of the x and y components of the oscillating electrical field vectors in a Cartesian frame whose z axis is along the direction of propagation. Following the notation of Paper I, we represent the components of the electric field by their time-varying complex amplitudes $e_x(t), e_y(t)$. The Stokes parameters are then customarily defined by (e.g. Born & Wolf; Thompson et al. 1986):

$$
\begin{aligned}
I &= \; <|e_x|^2 + |e_y|^2> \\
Q &= \; <|e_x|^2 - |e_y|^2> \\
U &= \; 2<|e_x||e_y|\cos\delta> \\
V &= \; 2<|e_x||e_y|\sin\delta>
\end{aligned}
\tag{1}
$$

where $<>$ denotes a time average and δ is the phase difference between the two oscillating components, i.e. the argument of $e_x e_y^*$ or its conjugate.
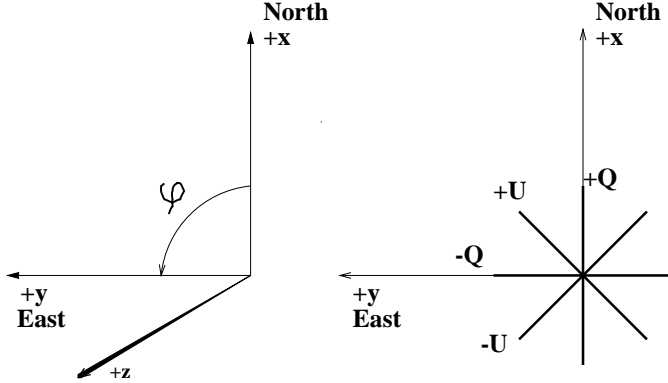
Fig. 1. The IAU coordinate system. Left: The right-handed xyz coordinate system seen from the position of the observer; radiation from the source propagates in the positive z direction. Right: The orientations of the electric amplitude vectors for Stokes Q and U on the plane of the sky

From these equations, it is immediately clear that I, Q and U are unambiguously defined once the xy coordinate frame is chosen. The IAU (1974) defines a conventional right-handed coordinate system (Fig. 1) with the x and y axes in the plane of the sky toward the North and East, respectively, and the positive z axis along the line of sight toward the observer; position angle is measured counterclockwise from the North through the East. It immediately follows from Eq. (1) that

- Positive Q is along the x axis.
- Positive U is along the bisectrix of the positive x and y axes, in position angle $\pi/4$.

These results are graphically summarized in Fig. 1.

For the definition of V in Eq. (1) to be unambiguous, we must somehow define the sign of δ. There are several ways of doing this; most references either use an implicit definition or leave the matter unresolved. As we shall show, the same correct definition may result in different formulae; it is the failure to clearly distinguish between the physics of the problem and its mathematical representation that makes many existing treatments of circular polarization confusing.

The correct approach is shown by the IAU (1974) definition, which stipulates:

- For right-handed circular polarization, the position angle of the electric vector at any point *increases* with time; this implies that the y component of the field *lags* the x component. Also, the electric vectors along the line of sight at any instant in time form a *left*-handed screw.
- Stokes V is positive for right-handed circular polarization.

How these definitions translate into mathematical formulae depends on how one defines the harmonic exponential with which the amplitudes e_x, e_y are multiplied to obtain the instantaneous field values (Simmons & Guttman 1970). Assuming ω to be positive, one has two choices:

$$\exp \pm i(\omega t - kz) \tag{2}$$

Both choices are valid, provided

- One makes sure to define the imaginary parts of all quantities appearing in the calculations with signs appropriate to the sign selected in Eq. (2).
- One attaches the correct corresponding physical interpretation to the results of the calculations.

In particular, the following mathematical representations must follow the choice of the sign in Eq. (2):

- For a right-handed circular wave, the y component lagging the x component is represented by

$$e_y = \mp i e_x$$

In other words, a unity-amplitude right-circular signal vector is given in xy coordinates by

$$e^+ = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \mp i \end{pmatrix} \tag{3}$$

- We want a definition of Stokes V equivalent to Eq. (1) but avoiding the complications of dealing with the phase difference δ. To this end we recast the equation in an equivalent form and adjust its sign to yield V correctly, i.e. positive for the right-circular wave of Eq. (3):

$$V = \mp i < (e_x e_y^* - e_y e_x^*) > \tag{4}$$

The adjustment (i.e. taking $\mp$ rather than $\pm$) is equivalent to defining the sign of δ in Eq. (1).

- As discussed in Born & Wolf (1964) and Thompson et al. (1986), a receptor is characterized by its *orientation* ϕ and *ellipticity* ϵ. The sign of ϕ is defined by the IAU convention. The references quoted formulate the definition of ϵ in terms of equations using the upper-sign convention of Eq. (2). We complete this definition by postulating that the ellipticity thus defined is a *physical* parameter whose sign is independent of the mathematical representation used.

Thompson et al. (1986) show that, for the IAU definitions listed above, a receptor with *positive* ϵ is more sensitive to a *left*-circular wave traveling in the positive z direction; for $\epsilon = \pi/4$, the response to a right-circular wave is zero. Thus, the receptor response for orientation $\phi = 0$ is defined by a *row* vector d_A:

$$v_A = d_A \, e_A \;\; ; \;\; d_A = (\cos \epsilon, \mp i \sin \epsilon) \tag{5}$$

and we see that indeed $v_A = 0$ for $\epsilon = \pi/4$ and a right-circular wave as defined by Eq. (3).

In the expressions given, the upper signs are the ones used in Papers I and II and by other authors quoted.

3. The circular-polarization coordinate frame

Another point to be settled is the order of the coordinate axes in the circular coordinate frame. This choice can be made independently of the preceding considerations. The order used in Paper I is (right, left) and this determines the form of the linear-to-circular transformation. In the notation of Paper I, this transformation is given by

$$\boldsymbol{C}_\mathrm{A} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & \pm i \\ 1 & \mp i \end{pmatrix}$$

Indeed, for a right-circular signal, Eq. (3), we have

$$\boldsymbol{e}_\mathrm{A}^\odot = \boldsymbol{C}_\mathrm{A} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \mp i \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

To remind the reader of the particular order chosen, we follow Paper I in referring to these coordinates as *circular-rl*.

4. Stokes visibilities in an interferometer

An interferometer is modelled by the positions of its two antenna elements A and B. The generalization of the Stokes-parameter definitions for a single point in the radiation field to those for an interferometer is then straightforward: All we have to do is to refer the two terms in each correlation product in Eqs. (1) and (4) to the two antennas:

$$
\begin{aligned}
I &= \; <e_{\mathrm{A}x}e_{\mathrm{B}x}^* + e_{\mathrm{A}y}e_{\mathrm{B}y}^*> \\
Q &= \; <e_{\mathrm{A}x}e_{\mathrm{B}x}^* - e_{\mathrm{A}y}e_{\mathrm{B}y}^*> \\
U &= \; <e_{\mathrm{A}x}e_{\mathrm{B}y}^* + e_{\mathrm{A}y}e_{\mathrm{B}x}^*> \\
V &= \mp i(\; <e_{\mathrm{A}x}e_{\mathrm{B}y}^* - e_{\mathrm{A}y}e_{\mathrm{B}x}^*>)
\end{aligned}
$$

(6)

and note that, for a source at infinite distance, these parameters depend only on the relative position $\boldsymbol{r}$ of the antennas, the *baseline vector*. Equation (6) is the definition used in Paper I, where we refer to these generalized Stokes parameters as Stokes *visibilities*. Appendix C of that paper discusses their Fourier-transform relationship to the Stokes-parameter brightness distributions of the source observed.

The Stokes visibilities are complex quantities and as such are subject to the sign ambiguity discussed in Sect. 2. According to a well-known theorem from Fourier theory, conjugating the visibilities results in an inversion of the corresponding sky image, i.e. a replacement of $B(\boldsymbol{l})$ with $B(-\boldsymbol{l})$ for all positions $\boldsymbol{l}$ in the image. This inversion may be compensated for by selecting the appropriate sign in the Fourier kernel or, equivalently, in the definition of the baseline vector between telescopes A and B. That selection can be made independently of our choice of the sign of ωt, and may take into account additional conjugations that may occur in the hardware (e.g. phase reversals in frequency conversions). In practice it is often made empirically.

5. The black-box formula of Morris et al.

All radio polarimetry work has until now been based on a formula published first by Morris et al. (1964). This formula gives the response to an input Stokes visibility of a simple interferometer consisting of two receptors, with given orientations and ellipticities, and a voltage multiplier. We call it the 'black-box formula' because it treats the entire interferometer as a black box without illuminating its inner workings.

Rederiving the formula, Weiler (1973) finds an unexplained difference with Morris et al. in the sign of Stokes V. Thompson et al. (1986) obtain the same result, yet their text suggests that they agree with Morris et al. To settle the controversy, we give an independent derivation of the formula on the basis of the definitions discussed above and show it in more detail than the other authors did. We conclude that Weiler's form is the correct one. Our derivation is also of interest in its use of circular-rl coordinates to simplify the calculations.

We recall from Paper I that for each signal vector and transformation matrix in the two-dimensional *signal domain* there is a corresponding four-dimensional entity in the *coherency domain*, which is related to the former by an *outer multiplication*. The notational convention for distinguishing the two- and four-dimensional entities is to use antenna subscripts A and B for the former and no subscripts for the latter. Superscripts are used to distinguish between representations in different coordinate systems: + for geometric-xy, $\odot$ for circular-rl and S for Stokes coordinates.

Following the authors cited, we consider two antennas, each with one receptor. Let the orientation and ellipticity for the A receptor be ϕ_A and ϵ_A, respectively. Including these two parameters and setting $\phi_\mathrm{A} = 0$ to begin with, we reproduce Eq. (5) in the form:

$$\boldsymbol{d}_\mathrm{A}(\epsilon_\mathrm{A}, 0) = (\cos\epsilon_\mathrm{A}, \mp i\sin\epsilon_\mathrm{A})$$

For considering receptors at arbitrary position angles ϕ, it is convenient to transform our problem to circular-rl coordinates where the rotation matrix is a diagonal one. Recall from Paper I that

$$\boldsymbol{e}_\mathrm{A}^\odot = \boldsymbol{C}_\mathrm{A}\boldsymbol{e}_\mathrm{A}^+ \tag{7}$$

For the response vector $\boldsymbol{d}$ the representations in the linear and circular-rl coordinate systems are related by

$$\boldsymbol{d}_\mathrm{A}^\odot(\epsilon_\mathrm{A}, 0)\,\boldsymbol{e}_\mathrm{A}^\odot = \boldsymbol{d}_\mathrm{A}^+(\epsilon_\mathrm{A}, 0)\,\boldsymbol{e}_\mathrm{A}^+$$

Substituting

$$\alpha_\mathrm{A} = \epsilon_\mathrm{A} + \pi/4\;, \tag{8}$$

and Eq. (7) we get

$$\boldsymbol{d}_\mathrm{A}^\odot\,(\alpha_\mathrm{A}, 0) = \boldsymbol{d}_\mathrm{A}^+\,(\epsilon_\mathrm{A}, 0)\,\boldsymbol{C}_\mathrm{A}^{-1} = (\cos\alpha_\mathrm{A},\;\sin\alpha_\mathrm{A})$$

The simplest way to represent a rotation of the receptor over an angle ϕ_A is by noting that it is equivalent to rotating the incident signal over an angle $-\phi_A$. For the rotation matrix we have

$$\boldsymbol{R}_A^\odot(-\phi_A) = \boldsymbol{C}_A \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \boldsymbol{C}_A^{-1} = \begin{pmatrix} e^{\mp i\phi_A} & 0 \\ 0 & e^{\pm i\phi_A} \end{pmatrix}$$

We can easily see that this form is correct: Rotation of the receptor in the direction in which the electric vector of the right-circular wave rotates will result in a phase retardation of the received signal for such a wave.

The response vector for arbitrary ϕ_A now becomes:

$$\begin{aligned} \boldsymbol{d}_A^\odot(\alpha_A, \phi_A) &= \boldsymbol{d}_A^\odot(\alpha_A, 0)\,\boldsymbol{R}_A^\odot(-\phi_A) \\ &= (e^{\mp i\phi_A}\cos\alpha_A,\ e^{\pm\phi_A}\sin\alpha_A) \end{aligned}$$

The interferometer response vector $\boldsymbol{d}^\odot$ is the direct product of the responses $\boldsymbol{d}_A^\odot$ and $\boldsymbol{d}_B^\odot$:

$$\begin{aligned} \boldsymbol{d}^\odot(\ & \alpha_A, \alpha_B, \phi_A, \phi_B) = \boldsymbol{d}_A^\odot(\alpha_A, \phi_A) \otimes \boldsymbol{d}_B^{\odot*}(\alpha_B, \phi_B) = \\ (\ & e^{\mp i(\phi_A-\phi_B)}\cos\alpha_A\cos\alpha_B,\ e^{\mp i(\phi_A+\phi_B)}\cos\alpha_A\sin\alpha_B, \\ & e^{\pm i(\phi_A+\phi_B)}\sin\alpha_A\cos\alpha_B,\ e^{\pm i(\phi_A-\phi_B)}\sin\alpha_A\sin\alpha_B) \end{aligned}$$

$$(9)$$

The response v in terms of the Stokes representation $\boldsymbol{e}^S$ of the coherency vector is then

$$v = \boldsymbol{d}^\odot \boldsymbol{S}^\odot \boldsymbol{e}^S \tag{10}$$

where $\boldsymbol{S}^\odot$ is the matrix that converts from Stokes to circular-rl coordinates. From Paper I we take the result

$$\boldsymbol{S}^\odot = \boldsymbol{S}\boldsymbol{C}^S = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & \pm i & 0 \\ 0 & 1 & \mp i & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix} \tag{11}$$

Substituting Eqs. (9) and (11) in Eq. (10) we obtain

$2v =$
$I\ [e^{\mp i(\phi_A-\phi_B)}\cos\alpha_A\cos\alpha_B + e^{\pm i(\phi_A-\phi_B)}\sin\alpha_A\sin\alpha_B]$
$+Q\ [e^{\mp i(\phi_A+\phi_B)}\cos\alpha_A\sin\alpha_B +\ e^{\pm i(\phi_A+\phi_B)}\sin\alpha_A\cos\alpha_B]$
$+U\ i[e^{\mp i(\phi_A+\phi_B)}\cos\alpha_A\sin\alpha_B - e^{\pm i(\phi_A+\phi_B)}\sin\alpha_A\cos\alpha_B]$
$+V\ [e^{\mp i(\phi_A-\phi_B)}\cos\alpha_A\cos\alpha_B - e^{\pm i(\phi_A-\phi_B)}\sin\alpha_A\sin\alpha_B]$

Expanding exponentials, regrouping the terms, and substituting the identies

$$\begin{aligned} \alpha_A - \alpha_B &= \epsilon_A - \epsilon_B \\ \cos(\alpha_A+\alpha_B) &= -\sin(\epsilon_A+\epsilon_B) \\ \sin(\alpha_A+\alpha_B) &=\ \cos(\epsilon_A+\epsilon_B) \end{aligned}$$

that follow from Eq. (8), we finally obtain

$2v =$

$\qquad I[\cos(\phi_A-\phi_B)\quad \cos(\epsilon_A-\epsilon_B)$
$\qquad\qquad \pm i\sin(\phi_A-\phi_B)\sin(\epsilon_A+\epsilon_B)]$
$\quad +Q[\cos(\phi_A+\phi_B)\quad \cos(\epsilon_A+\epsilon_B)$
$\qquad\qquad \pm i\sin(\phi_A+\phi_B)\sin(\epsilon_A-\epsilon_B)]$
$\quad +U[\sin(\phi_A+\phi_B)\quad \cos(\epsilon_A+\epsilon_B)$
$\qquad\qquad \mp i\cos(\phi_A+\phi_B)\sin(\epsilon_A-\epsilon_B)]$
$\quad +V[-\cos(\phi_A-\phi_B)\quad \sin(\epsilon_A+\epsilon_B)$
$\qquad\qquad \mp i\sin(\phi_A-\phi_B)\cos(\epsilon_A-\epsilon_B)]$

$$(12)$$

For the upper signs, this result is identical with that of Weiler (1973). We may verify the correctness of the sign of V by considering two simple examples.

In the first one, we assume two right-circular receptors ($\epsilon = -\pi/4$) in position angle 0:

$$d_A = d_B = \frac{1}{\sqrt{2}}(1,\ \pm i)$$

The responses to right-circular radiation are $v_A = v_B = 1$, so the cross-corration $v = 1$; the same result follows from Eq. (12).

In the second example, we rotate receptor A over $\pi/2$ to obtain

$$v_A = \frac{1}{\sqrt{2}}(1, \pm i)\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ \mp i \end{pmatrix} = \mp i$$

The cross-correlation with v_B gives $v = \mp i$; again, this agrees with Eq. (12).

We emphasize once more that the sign ambiguities do not represent ambiguities in the physics, but only reflect the fact that it admits of two mutually equivalent mathematical representation.

6. Conclusion

In this paper we have presented a unified treatment of both the physical definitions pertaining to the description of polarized radiation and their precipitation in mathematical form. We have shown that, even though the physical definitions of the IAU (1974) when combined with the standard formulae for I, Q and U define the Stokes parameters uniquely, they leave room for two similar but distinct mathematical representations. By clearly distinguishing between the physical and mathematical aspects of the definitions, we hope to have eliminated the confusion that has long surrounded the literature on radio polarimetry.

In particular, we have settled a long-standing question concerning the 'black-box' formula of Morris et al. (1964), two versions of which had lived in a peaceful yet unsatisfactory coexistence for over twenty years.

References

Born M., Wolf E., 1964, Principles of Optics. Pergamon Press

Hamaker J.P., Bregman J.D., Sault R.J., 1996, A&AS 117, 137

IAU, 1974, Transactions of the IAU Vol. 15B (1973) 166

IEEE, 1969, Standard Definitions of Terms for Radio Wave Propagation, IEEE Trans. AP-17, 270-275

Morris D., Radhakrishnan V., Seielstad G.A., 1964, ApJ 139, 551-559. The derivation of the black-box formula is reproduced in Weiler (1973) and Thompson et al. (1986)

Sault R.J., Hamaker J.P., Bregman J.D., 1996, A&AS 117, 149

Simmons J.W., Guttman M.J., 1970, States, waves and photons: A modern introduction to light. Addison-Wesley

Thompson A.R., Moran J.M., Swenson G.W. Jr, 1986, Interferometry and Synthesis in Radio Astronomy. John Wiley & Sons, New York

Weiler K.W., 1973, A&A 26, 404-407