

# Chapter 4

## Geometrical Relationships, Polarimetry, and the Interferometer Measurement Equation

In this chapter, we start to examine some of the practical aspects of interferometry. These include baselines, antenna mounts and beam shapes, and the response to polarized radiation, all of which involve geometric considerations and coordinate systems. The discussion is concentrated on Earth-based arrays with tracking antennas, which illustrate the principles involved, although the same principles apply to other systems such as those that include one or more antennas in Earth orbit.

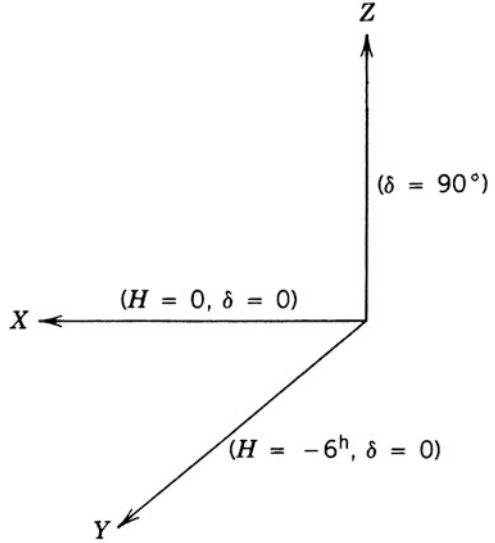
### 4.1 Antenna Spacing Coordinates and $(u, v)$ Loci

Various coordinate systems are used to specify the relative positions of the antennas in an array, and of these, one of the more convenient for terrestrial arrays is shown in Fig. 4.1. A right-handed Cartesian coordinate system is used, where  $X$  and  $Y$  are measured in a plane parallel to the Earth's equator,  $X$  in the meridian plane<sup>1</sup> (defined as the plane through the poles of the Earth and the reference point in the array),  $Y$  toward the east, and  $Z$  toward the north pole. In terms of hour angle  $H$  and declination  $\delta$ , coordinates  $(X, Y, Z)$  are measured toward  $(H = 0, \delta = 0)$ ,  $(H = -6^{\text{h}}, \delta = 0)$ , and  $(\delta = 90^{\circ})$ , respectively. If  $(X_{\lambda}, Y_{\lambda}, Z_{\lambda})$  are the components

---

<sup>1</sup>In VLBI observations, it is customary to set the  $X$  axis in the Greenwich meridian, in which case  $H$  is measured with respect to that meridian rather than a local one.

**Fig. 4.1** The  $(X, Y, Z)$  coordinate system for specification of relative positions of antennas. Directions of the axes specified are in terms of hour angle  $H$  and declination  $\delta$ .



of  $\mathbf{D}_\lambda$  in the  $(X, Y, Z)$  system, the components  $(u, v, w)$  are given by

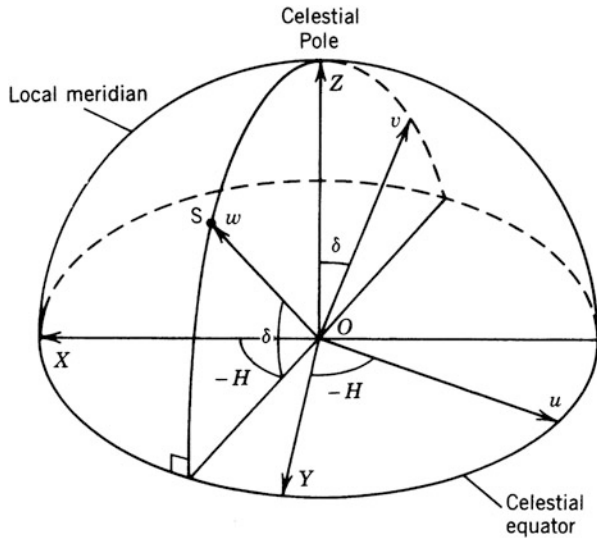
$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} \sin H & \cos H & 0 \\ -\sin \delta \cos H & \sin \delta \sin H & \cos \delta \\ \cos \delta \cos H & -\cos \delta \sin H & \sin \delta \end{bmatrix} \begin{bmatrix} X_\lambda \\ Y_\lambda \\ Z_\lambda \end{bmatrix}. \quad (4.1)$$

Here  $(H, \delta)$  are usually the hour angle and declination of the phase reference position. The elements of the transformation matrix given above are the direction cosines of the  $(u, v, w)$  axes with respect to the  $(X, Y, Z)$  axes and can easily be derived from the relationships in Fig. 4.2. Another method of specifying the baseline vector is in terms of its length,  $D$ , and the hour angle and declination,  $(h, d)$ , of the intersection of the baseline direction with the Northern Celestial Hemisphere. The coordinates in the  $(X, Y, Z)$  system are then given by

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = D \begin{bmatrix} \cos d \cos h \\ -\cos d \sin h \\ \sin d \end{bmatrix}. \quad (4.2)$$

The coordinates in the  $(u, v, w)$  system are, from Eqs. 4.1 and 4.2,

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = D_\lambda \begin{bmatrix} \cos d \sin(H - h) \\ \sin d \cos \delta - \cos d \sin \delta \cos(H - h) \\ \sin d \sin \delta + \cos d \cos \delta \cos(H - h) \end{bmatrix}. \quad (4.3)$$



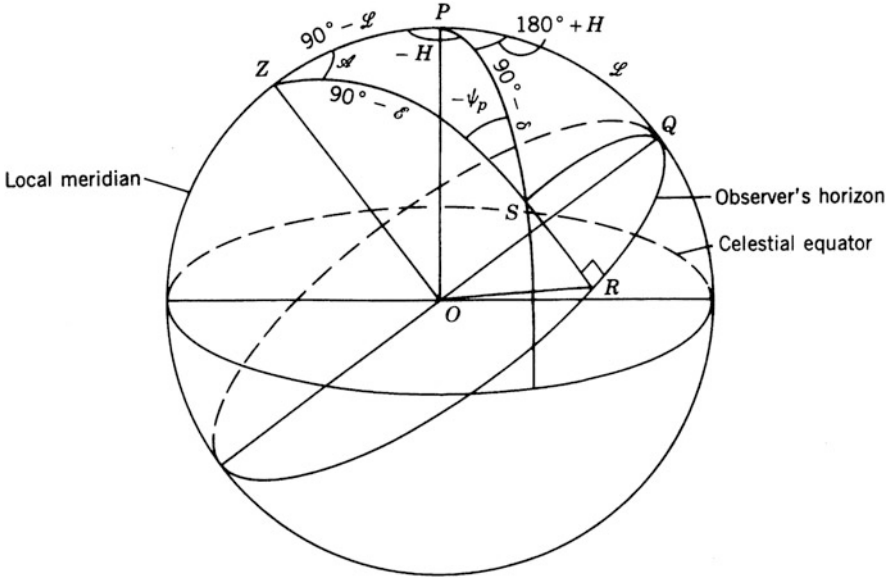
**Fig. 4.2** Relationships between the  $(X, Y, Z)$  and  $(u, v, w)$  coordinate systems. The  $(u, v, w)$  system is defined for observation in the direction of the point  $S$ , which has hour angle and declination  $H$  and  $\delta$ . As shown,  $S$  is in the eastern half of the hemisphere and  $H$  is therefore negative. The direction cosines in the transformation matrix in Eq. (4.1) follow from the relationships in this diagram. The relationship in Eq. (4.2) can also be derived if we let  $S$  represent the direction of the baseline and put the baseline coordinates  $(h, d)$  for  $(H, \delta)$ .

The  $(D, h, d)$  system was used more widely in the earlier literature, particularly for instruments involving only two antennas; see, for example, Rowson (1963).

When the  $(X, Y, Z)$  components of a new baseline are first established, the usual practice is to determine the elevation  $\mathcal{E}$ , azimuth  $\mathcal{A}$ , and length of the baseline by field surveying techniques. Figure 4.3 shows the relationship between  $(\mathcal{E}, \mathcal{A})$  and other coordinate systems; see also Appendix 4.1. For latitude  $\mathcal{L}$ , using Eqs. (4.2) and (A4.2), we obtain

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = D \begin{bmatrix} \cos \mathcal{L} \sin \mathcal{E} - \sin \mathcal{L} \cos \mathcal{E} \cos \mathcal{A} \\ \cos \mathcal{E} \sin \mathcal{A} \\ \sin \mathcal{L} \sin \mathcal{E} + \cos \mathcal{L} \cos \mathcal{E} \cos \mathcal{A} \end{bmatrix}. \quad (4.4)$$

Examination of Eq. (4.1) or (4.3) shows that the locus of the projected antenna spacing components  $u$  and  $v$  defines an ellipse with hour angle as the variable. Let



**Fig. 4.3** Relationship between the celestial coordinates  $(H, \delta)$  and the elevation and azimuth  $(\mathcal{E}, \mathcal{A})$  of a point  $S$  as seen by an observer at latitude  $\mathcal{L}$ .  $P$  is the celestial pole and  $Z$  the observer's zenith. The parallactic angle  $\psi_p$  is the position angle of the observer's vertical on the sky measured from north toward east. The lengths of the arcs measured in terms of angles subtended at the center of the sphere  $O$  are as follows:

$$\begin{aligned} ZP &= 90^\circ - \mathcal{L} & PQ &= \mathcal{L} & SR &= \mathcal{E} & RQ &= \mathcal{A} \\ SZ &= 90^\circ - \mathcal{E} & SP &= 90^\circ - \delta & SQ &= \cos^{-1}(\cos \mathcal{E} \cos \mathcal{A}). \end{aligned}$$

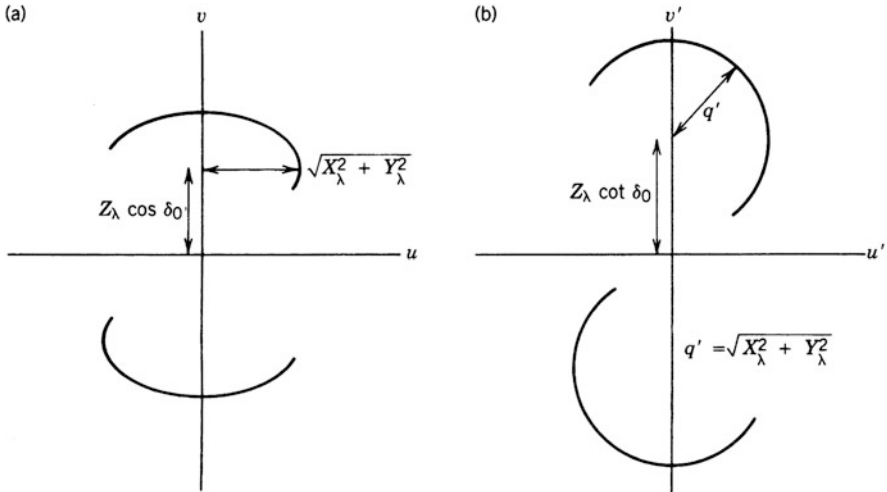
The required relationships can be obtained by application of the sine and cosine rules for spherical triangles to  $ZPS$  and  $PQS$  and are given in Appendix 4.1. Note that with  $S$  in the eastern half of the observer's sky, as shown,  $H$  and  $\psi_p$  are negative.

$(H_0, \delta_0)$  be the phase reference position. Then from Eq. (4.1), we have

$$u^2 + \left( \frac{v - Z_\lambda \cos \delta_0}{\sin \delta_0} \right)^2 = X_\lambda^2 + Y_\lambda^2. \quad (4.5)$$

In the  $(u, v)$  plane, Eq. (4.5) defines an ellipse<sup>2</sup> with the semimajor axis equal to  $\sqrt{X_\lambda^2 + Y_\lambda^2}$ , and the semiminor axis equal to  $\sin \delta_0 \sqrt{X_\lambda^2 + Y_\lambda^2}$ , as in Fig. 4.4a. The ellipse is centered on the  $v$  axis at  $(u, v) = (0, Z_\lambda \cos \delta_0)$ . The arc of the ellipse that is traced out during any observation depends on the azimuth, elevation, and latitude of the baseline; the declination of the source; and the range of hour angle covered,

<sup>2</sup>The first mention of elliptical loci appears to have been by Rowson (1963).



**Fig. 4.4** (a) Spacing vector locus in the  $(u, v)$  plane from Eq. (4.5). (b) Spacing vector locus in the  $(u', v')$  plane from Eq. (4.8). The lower arc in each diagram represents the locus of conjugate values of visibility. Unless the source is circumpolar, the cutoff at the horizon limits the lengths of the arcs.

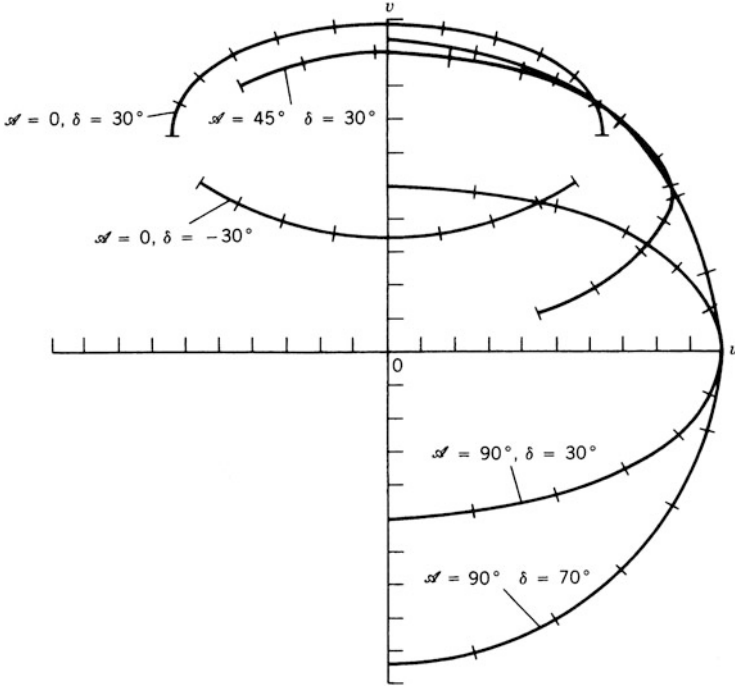
as illustrated in Fig. 4.5. Since  $\mathcal{V}(-u, -v) = \mathcal{V}^*(u, v)$ , any observation supplies simultaneous measurements on two arcs, which are part of the same ellipse only if  $Z_\lambda = 0$ .

## 4.2 ( $u', v'$ ) Plane

The  $(u', v')$  plane, which was introduced in Sect. 3.1.2 with regard to east–west baselines, is also useful in discussing certain aspects of the behavior of arrays in general. This plane is normal to the direction of the pole and can be envisaged as the equatorial plane of the Earth. For non-east–west baselines, we can also consider the projection of the spacing vectors onto the  $(u', v')$  plane. All such projected vectors sweep out circular loci as the Earth rotates. The spacing components in the  $(u', v')$  plane are derived from those in the  $(u, v)$  plane by the transformation  $u' = u, v' = v \operatorname{cosec} \delta_0$ . In terms of the components of the baseline  $(X_\lambda, Y_\lambda, Z_\lambda)$  for two antennas, we obtain from Eq. (4.1)

$$u' = X_\lambda \sin H_0 + Y_\lambda \cos H_0 \tag{4.6}$$

$$v' = -X_\lambda \cos H_0 + Y_\lambda \sin H_0 + Z_\lambda \cot \delta_0 . \tag{4.7}$$



**Fig. 4.5** Examples of  $(u, v)$  loci to show the variation with baseline azimuth  $\mathcal{A}$  and observing declination  $\delta$  (the baseline elevation  $\mathcal{E}$  is zero). The baseline length in all cases is equal to the length of the axes measured from the origin. The tracking range is  $-4$  to  $+4$  h for  $\delta = -30^\circ$ , and  $-6$  to  $+6$  h in all other cases. Marks along the loci indicate 1-h intervals in tracking. Note the change in ellipticity for east–west baselines ( $\mathcal{A} = 90^\circ$ ) with  $\delta = 30^\circ$  and with  $\delta = 70^\circ$ . The loci are calculated for latitude  $40^\circ$ .

The loci are circles centered on  $(0, Z_\lambda \cot \delta_0)$ , with radii  $q'$  given by

$$q'^2 = u'^2 + (v' - Z_\lambda \cot \delta_0)^2 = X_\lambda^2 + Y_\lambda^2, \quad (4.8)$$

as shown in Fig. 4.4b. The projected spacing vectors that generate the loci rotate with constant angular velocity  $\omega_e$ , the rotation velocity of the Earth, which is easier to visualize than the elliptic motion in the  $(u, v)$  plane. In particular, problems involving the effect of time, such as the averaging of visibility data, are conveniently dealt with in the  $(u', v')$  plane. Examples of its use will be found in Sects. 4.4, 6.4.2, and 16.3.2. In Fourier transformation, the conjugate variables of  $(u', v')$  are  $(l', m')$ , where  $l' = l$  and  $m' = m \sin \delta_0$ , that is, the image plane is compressed by a factor  $\sin \delta_0$  in the  $m$  direction.

### 4.3 Fringe Frequency

The component  $w$  of the baseline represents the path difference to the two antennas for a plane wave incident from the phase reference position. The corresponding time delay is  $w/v_0$ , where  $v_0$  is the center frequency of the observing band. The relative phase of the signals at the two antennas changes by  $2\pi$  radians when  $w$  changes by unity. Thus, the frequency of the oscillations at the output of the correlator that combines the signals is

$$\frac{dw}{dt} = \frac{dw}{dH} \frac{dH}{dt} = -\omega_e [X_\lambda \cos \delta \sin H + Y_\lambda \cos \delta \cos H] = -\omega_e u \cos \delta, \quad (4.9)$$

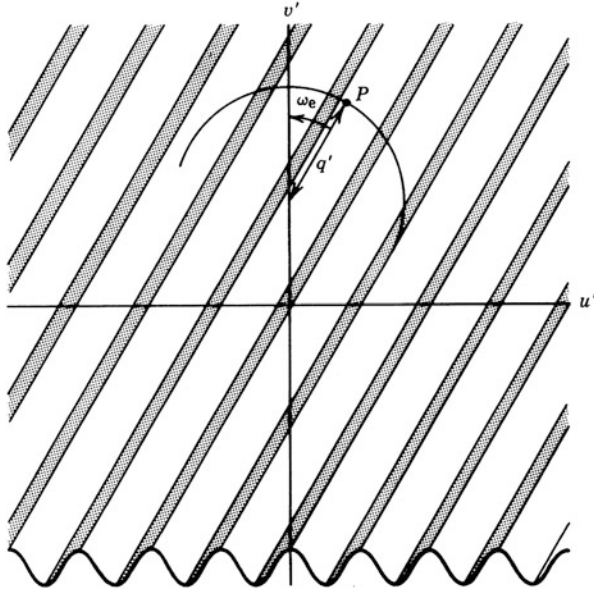
where  $\omega_e = dH/dt = 7.29115 \times 10^{-15} \text{ rad s}^{-1} = \omega_e$  is the rotation velocity of the Earth with respect to the fixed stars: for greater accuracy, see Seidelmann (1992). The sign of  $dw/dt$  indicates whether the phase is increasing or decreasing with time. The result shown above applies to the case in which the signals suffer no time-varying instrumental phase changes between the antennas and the correlator inputs. In an array in which the antennas track a source, time delays to compensate for the space path differences  $w$  are applied to maintain correlation of the signals. If an exact compensating delay were introduced in the radio frequency section of the receivers, the relative phases of the signals at the correlator input would remain constant, and the correlator output would show no fringes. However, except in some low-frequency systems like LOFAR (de Vos et al. 2009), the compensating delays are usually introduced at an intermediate frequency, of which the band center  $\nu_d$  is much less than the observing frequency  $\nu_0$ . The adjustment of the compensating delay introduces a rate of phase change  $2\pi \nu_d(dw/dt)/\nu_0 = -\omega_e u (\cos \delta) \nu_d/\nu_0$ . The resulting fringe frequency at the correlator output is

$$\nu_f = \frac{dw}{dt} \left( 1 \mp \frac{\nu_d}{\nu_0} \right) = -\omega_e u \cos \delta \left( 1 \mp \frac{\nu_d}{\nu_0} \right), \quad (4.10)$$

where the negative sign refers to upper-sideband reception and the positive sign to lower-sideband reception; these distinctions and the double-sideband case are explained in Sect. 6.1.8. From Eq. (4.3), the right side of Eq. (4.10) is equal to  $-\omega_e D \cos d \cos \delta \sin(H - h)(\nu_0 \mp \nu_d)/c$ . Note that  $(\nu_0 \mp \nu_d)$  is usually determined by one or more local oscillator frequencies.

### 4.4 Visibility Frequencies

As explained in Sect. 3.1, the phase of the complex visibility is measured with respect to that of a hypothetical point source at the phase reference position. The fringe-frequency variations do not appear in the visibility function, but slower



**Fig. 4.6** The  $(u', v')$  plane showing sinusoidal corrugations that represent the visibility of a point source. For simplicity, only the real part of the visibility is included. The most rapid variation in the visibility is encountered at the point  $P$ , where the direction of the spacing locus is normal to the ridges in the visibility.  $\omega_e$  is the rotation velocity of the Earth.

variations occur that depend on the position of the radiating sources within the field. We now examine the maximum temporal frequency of the visibility variations. Consider a point source represented by the delta function  $\delta(l_1, m_1)$ . The visibility function is the Fourier transform of  $\delta(l_1, m_1)$ , which is

$$e^{-j2\pi(ul_1 + vm_1)} = \cos 2\pi(ul_1 + vm_1) - j \sin 2\pi(ul_1 + vm_1) . \tag{4.11}$$

This expression represents two sets of sinusoidal corrugations, one real and one imaginary. The corrugations represented by the real part of Eq. (4.11) are shown in  $(u', v')$  coordinates in Fig. 4.6, where the arguments of the trigonometric functions in Eq. (4.11) become  $2\pi(u'l_1 + v'm_1 \sin \delta_0)$ . The frequency of the corrugations in terms of cycles per unit distance in the  $(u', v')$  plane is  $l_1$  in the  $u'$  direction,  $m_1 \sin \delta_0$  in the  $v'$  direction, and

$$r'_1 = \sqrt{l_1^2 + m_1^2 \sin^2 \delta_0} \tag{4.12}$$

in the direction of most rapid variations. Expression (4.12) is maximized at the pole and then becomes equal to  $r_1$ , which is the angular distance of the source from the  $(l, m)$  origin. For any antenna pair, the spatial frequency locus in the  $(u', v')$  plane is a circle of radius  $q'$  generated by a vector rotating with angular velocity  $\omega_e$ , where



$q'$  is as defined in Eq. (4.8). From Fig. 4.6, it is clear that the temporal variation of the measured visibility is greatest at the point  $P$  and is equal to  $\omega_e r'_1 q'$ . This is a useful result, since if  $r_1$  represents a position at the edge of the field to be imaged, it indicates that to follow the most rapid variations, the visibility must be sampled at time intervals sufficiently small compared with  $(\omega_e r'_1 q')^{-1}$ . Also, we may wish to alternate between two frequencies or polarizations during an observation, and these changes must be made on a similarly short timescale. Note that this requirement is also covered by the sampling theorem in Sect. 5.2.1.

## 4.5 Calibration of the Baseline

The position parameters ( $X, Y, Z$ ) for each antenna relative to a common reference point can usually be established to a few centimeters or millimeters by a conventional engineering survey. Except at long wavelengths, the accuracy required is greater than this. We must be able to compute the phase at any hour angle for a point source at the phase reference position to an accuracy of, say,  $1^\circ$  and subtract it from the observed phase. This reference phase is represented by the factor  $e^{j2\pi w}$  in Eq. (3.7), and it is therefore necessary to calculate  $w$  to  $1/360$  of the observing wavelength. The baseline parameters can be obtained to the required accuracy from observations of calibration sources for which the positions are accurately known. The phase of such a calibrator observed at the phase reference position ( $H_0, \delta_0$ ) should ideally be zero. However, if practical uncertainties are taken into account, the measured phase is, from Eq. (4.1),

$$2\pi \Delta w + \phi_{\text{in}} = 2\pi(\cos \delta_0 \cos H_0 \Delta X_\lambda - \cos \delta_0 \sin H_0 \Delta Y_\lambda + \sin \delta_0 \Delta Z_\lambda) + \phi_{\text{in}} , \quad (4.13)$$

where the prefix  $\Delta$  indicates the uncertainty in the associated quantity, and  $\phi_{\text{in}}$  is an instrumental phase term for the two antennas involved. If a calibrator is observed over a wide range of hour angle,  $\Delta X_\lambda$  and  $\Delta Y_\lambda$  can be obtained from the even and odd components, respectively, of the phase variation with  $H_0$ . To measure  $\Delta Z_\lambda$ , calibrators at more than one declination must be included. A possible procedure is to observe several calibrators at different declinations, repeating a cycle of observations for several hours. For the  $k$ th observation, we can write, from Eq. (4.13),

$$a_k \Delta X_\lambda + b_k \Delta Y_\lambda + c_k \Delta Z_\lambda + \phi_{\text{in}} = \phi_k , \quad (4.14)$$

where  $a_k, b_k$ , and  $c_k$  are known source parameters, and  $\phi_k$  is the measured phase. The calibrator source position need not be accurately known since the phase measurements can be used to estimate both the source positions and the baselines. Techniques for this analysis are discussed in Sect. 12.2. In practice, the instrumental phase  $\phi_{\text{in}}$  will vary slowly with time: instrumental stability is discussed in Chap. 7.

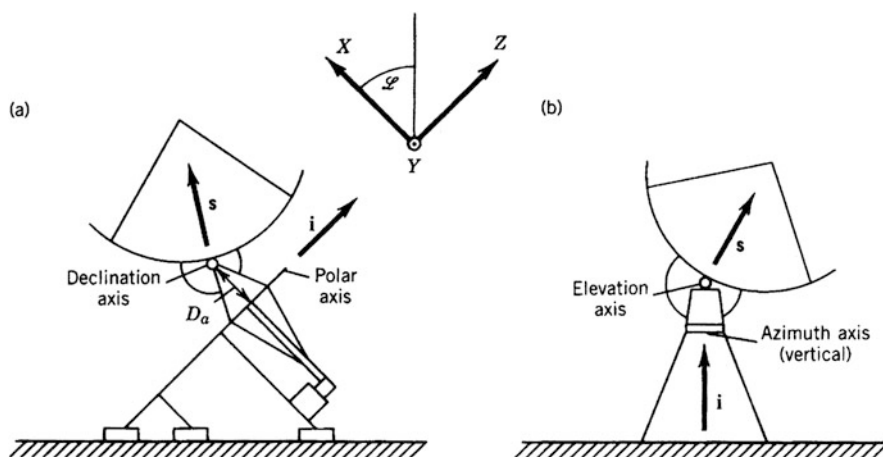
Also, there will be atmospheric phase variations, which are discussed in Chap. 13. These effects set the final limit on the attainable accuracy in observing both calibrators and sources under investigation.

Measurement of baseline parameters to an accuracy of order 1 part in  $10^7$  (e.g., 3 mm in 30 km) implies timing accuracy of order  $10^{-7}\omega_e^{-1} \simeq 1$  ms. Timekeeping is discussed in Sects. 9.5.8 and 12.3.3.

## 4.6 Antennas

### 4.6.1 Antenna Mounts

In discussing the dependence of the measured phase on the baseline components, we have ignored any effects introduced by the antennas, which is tantamount to assuming that the antennas are identical and their effects on the signals cancel out. This, however, is only approximately true. In most synthesis arrays, the antennas must have collecting areas of tens or hundreds of square meters for reasons of sensitivity. Except for dipole arrays at meter wavelengths, the antennas required are large structures that must be capable of accurately tracking a radio source across the sky. Tracking antennas are almost always constructed either on equatorial mounts (also called polar mounts) or on altazimuth mounts, as illustrated in Fig. 4.7. In an equatorial mount, the polar axis is parallel to the Earth's axis of rotation, and tracking a source requires only that the antenna be turned about the polar axis at the



**Fig. 4.7** Schematic diagrams of antennas on (a) equatorial (polar) and (b) altazimuth mounts. In the positions shown, the declination and elevation axes are normal to the plane of the page. In the equatorial mount, there is a distance  $D_a$  between the two rotational axes, but in the altazimuth mount, the axes often intersect, as shown.

sidereal rate. Equatorial mounts are mechanically more difficult to construct than altazimuth ones and are found mainly on antennas built prior to the introduction of computers for control and coordinate conversion.

In most tracking arrays used in radio astronomy, the antennas are circularly symmetrical reflectors. A desirable feature is that the axis of symmetry of the reflecting surface intersect both the rotation axes of the mount. If this is not the case, pointing motions will cause the antenna to have a component of motion along the direction of the beam. It is then necessary to take account of phase changes associated with small pointing corrections, which may differ from one antenna to another. In most antenna mounts, however, whether of equatorial or altazimuth type, the reflector axis intersects the rotation axes with sufficient precision that phase errors of this type are negligible.

It is convenient but not essential that the two rotation axes of the mount intersect. The intersection point then provides an appropriate reference point for defining the baseline between antennas, since whatever direction in which the antenna points, its aperture plane is always the same distance from that point as measured along the axis of the beam. In most large equatorially mounted antennas, the polar and declination axes do not intersect. In many cases, there is an offset of several meters between the polar and declination axes. Wade (1970) considered the implication of this offset for high-accuracy phase measurements and showed that it is necessary to take account of variations in the offset distance and in the accuracy of alignment of the polar axis. These results can be obtained as follows. Let  $\mathbf{i}$  and  $\mathbf{s}$  be unit vectors in the direction of the polar axis and the direction of the source under observation, respectively, and let  $\mathbf{D}_a$  be the spacing vector between the two axes measured perpendicular to  $\mathbf{i}$  (see Fig. 4.7a). The quantity that we need to compute is the projection of  $\mathbf{D}_a$  in the direction of observation,  $\mathbf{D}_a \cdot \mathbf{s}$ . Since  $\mathbf{D}_a$  is perpendicular to  $\mathbf{i}$ , the cosine of the angle between  $\mathbf{D}_a$  and  $\mathbf{s}$  is  $\sqrt{1 - (\mathbf{i} \cdot \mathbf{s})^2}$ . Thus,

$$\mathbf{D}_a \cdot \mathbf{s} = D_a \sqrt{1 - (\mathbf{i} \cdot \mathbf{s})^2}, \quad (4.15)$$

where  $D_a$  is the magnitude of  $\mathbf{D}_a$ . In the  $(X, Y, Z)$  coordinate system in which the baseline components are measured,  $\mathbf{i}$  has direction cosines  $(i_X, i_Y, i_Z)$ , and  $\mathbf{s}$  has direction cosines given by the transformation matrix on the right side of Eq. (4.2), but with  $h$  and  $d$  replaced by  $H$  and  $\delta$ , which refer to the direction of observation. If the polar axis is correctly aligned to within about 1 arcmin,  $i_X$  and  $i_Y$  are of order  $10^{-3}$  and  $i_Z \simeq 1$ . Thus, we can use the direction cosines to evaluate Eq. (4.15), and ignoring second-order terms in  $i_X$  and  $i_Y$ , we obtain

$$\mathbf{D}_a \cdot \mathbf{s} = D_a(\cos \delta - i_X \sin \delta \cos H + i_Y \sin \delta \sin H). \quad (4.16)$$

If the magnitude of  $\mathbf{D}_a$  is expressed in wavelengths, the difference in the values of  $\mathbf{D}_a \cdot \mathbf{s}$  for the two antennas must be added to the  $w$  component of the baseline given by Eq. (4.1) when calculating the reference phase at the field center. To do this, it is first necessary to determine the unknown constants in Eq. (4.16), which can be done

by adding a term of the form  $2\pi(\alpha \cos \delta_0 + \beta \sin \delta_0 \cos H_0 + \gamma \sin \delta_0 \sin H_0)$  to the right side of Eq. (4.13) and extending the solution to include  $\alpha$ ,  $\beta$ , and  $\gamma$ . The result then represents the differences in the corresponding mechanical dimensions of the two antennas. Note that the terms in  $i_X$  and  $i_Y$  in Eq. (4.16) are important only when  $D_a$  is large. If  $D_a$  is no more than one wavelength, it should be possible to ignore them.

The preceding analysis can be extended to the case of an altazimuth mount by letting  $\mathbf{i}$  represent the direction of the azimuth axis, as in Fig. 4.7b. Then  $i_X = \cos(\mathcal{L} + \varepsilon)$ ,  $i_Y = \sin \varepsilon'$ , and  $i_Z = \sin(\mathcal{L} + \varepsilon)$ , where  $\mathcal{L}$  is the latitude and  $\varepsilon$  and  $\varepsilon'$  are, respectively, the tilt errors in the  $XZ$  plane and in the plane containing the  $Y$  axis and the local vertical. The errors again should be quantities of order  $10^{-3}$ . In many altazimuth mounts, the axes are designed to intersect, and  $D_a$  represents only a structural tolerance. Thus, we assume that  $D_a$  is small enough to allow terms in  $i_Y D_a$  and  $\varepsilon D_a$  to be ignored, and evaluation of Eq. (4.15) gives

$$\mathbf{D}_a \cdot \mathbf{s} = D_a [1 - (\sin \mathcal{L} \sin \delta + \cos \mathcal{L} \cos \delta \cos H)^2] = D_a \cos \mathcal{E}, \quad (4.17)$$

where  $\mathcal{E}$  is the elevation of direction  $\mathbf{s}$ : see Eq. (A4.1) of Appendix 4.1. Correction terms of this form can be added to the expressions for the baseline calibration and for  $w$ .

### 4.6.2 Beamwidth and Beam-Shape Effects

The interpretation of data taken with arrays containing antennas with nonidentical beamwidths is not always a straightforward matter. Each antenna pair responds to an effective intensity distribution that is the product of the actual intensity of the sky and the geometric mean of the normalized beam profiles. If different pairs of antennas respond to different effective distributions, then, in principle, the Fourier transform relationship between  $I(l, m)$  and  $\mathcal{V}(u, v)$  cannot be applied to the ensemble of observations. Mixed arrays are sometimes used in VLBI when it is necessary to make use of antennas that have different designs. However, in VLBI studies, the source structure under investigation is very small compared with the widths of the antenna beams, so the differences in the beams can usually be ignored. If cases arise in which different beams are used and the source is not small compared with beamwidths, it is possible to restrict the measurements to the field defined by the narrowest beam by convolution of the visibility data with an appropriate function in the  $(u, v)$  plane.

A problem similar to that of unmatched beams occurs if the antennas have altazimuth mounts and the beam contours are not circularly symmetrical about the nominal beam axis. As a point in the sky is tracked using an altazimuth mount, the beam rotates with respect to the sky about this nominal axis. This rotation does not occur for equatorial mounts. The angle between the vertical at the antenna and the

direction of north at the point being observed (defined by the great circle through the point and the North Pole) is the parallactic angle  $\psi_p$  in Fig. 4.3. Application of the sine rule to the spherical triangle  $ZPS$  gives

$$\frac{-\sin \psi_p}{\cos \mathcal{L}} = \frac{-\sin H}{\cos \mathcal{E}} = \frac{\sin \mathcal{A}}{\cos \delta}, \quad (4.18)$$

which can be combined with Eq. (A4.1) or (A4.2) to express  $\psi_p$  as a function of  $(\mathcal{A}, \mathcal{E})$  or  $(H, \delta)$ . If the beam has elongated contours and width comparable to the source under observation, rotation of the beam causes the effective intensity distribution to vary with hour angle. This is particularly serious in the case of observations to reveal the structure of the most distant Universe, for which foreground sources need to be accurately removed. For the Australia Pathfinder Array (DeBoer et al. 2009), the 12-m-diameter antennas have altazimuth mounts, with a third axis that allows the reflector, feed supports, and feeds to be rotated about the reflector axis so the beam pattern and the angle of polarization remain fixed relative to the sky.

## 4.7 Polarimetry

Polarization measurements are very important in radio astronomy. Most synchrotron radiation shows a small degree of polarization that indicates the distribution of the magnetic fields within the source. As noted in Chap. 1, this polarization is generally linear (plane) and can vary in magnitude and position angle over the source. As frequency is increased, the percentage polarization often increases because the depolarizing action of Faraday rotation is reduced. Polarization of radio emission also results from the Zeeman effect in atoms and molecules, cyclotron radiation and plasma oscillations in the solar atmosphere, and Brewster angle effects at planetary surfaces. The measure of polarization that is almost universally used in astronomy is the set of four parameters introduced by Sir George Stokes in 1852. It is assumed here that readers have some familiarity with the concept of Stokes parameters or can refer to one of numerous texts that describe them [e.g., Born and Wolf (1999); Kraus and Carver (1973); Wilson et al. (2013)].

Stokes parameters are related to the amplitudes of the components of the electric field,  $E_x$  and  $E_y$ , resolved in two perpendicular directions normal to the direction of propagation. Thus, if  $E_x$  and  $E_y$  are represented by  $\mathcal{E}_x(t) \cos[2\pi\nu t + \delta_x(t)]$  and  $\mathcal{E}_y(t) \cos[2\pi\nu t + \delta_y(t)]$ , respectively, Stokes parameters are defined as follows:

$$\begin{aligned} I &= \langle \mathcal{E}_x^2(t) \rangle + \langle \mathcal{E}_y^2(t) \rangle \\ Q &= \langle \mathcal{E}_x^2(t) \rangle - \langle \mathcal{E}_y^2(t) \rangle \\ U &= 2 \langle \mathcal{E}_x(t) \mathcal{E}_y(t) \cos [\delta_x(t) - \delta_y(t)] \rangle \\ V &= 2 \langle \mathcal{E}_x(t) \mathcal{E}_y(t) \sin [\delta_x(t) - \delta_y(t)] \rangle, \end{aligned} \quad (4.19)$$

where the angular brackets denote the expectation or time average. This averaging is necessary because in radio astronomy, we are dealing with fields that vary with time in a random manner. Of the four parameters,  $I$  is a measure of the total intensity of the wave,  $Q$  and  $U$  represent the linearly polarized component, and  $V$  represents the circularly polarized component. Stokes parameters can be converted to a measure of polarization with a more direct physical interpretation as follows:

$$m_\ell = \frac{\sqrt{Q^2 + U^2}}{I} \quad (4.20)$$

$$m_c = \frac{V}{I} \quad (4.21)$$

$$m_t = \frac{\sqrt{Q^2 + U^2 + V^2}}{I} \quad (4.22)$$

$$\theta = \frac{1}{2} \tan^{-1} \left( \frac{U}{Q} \right), \quad 0 \leq \theta \leq \pi, \quad (4.23)$$

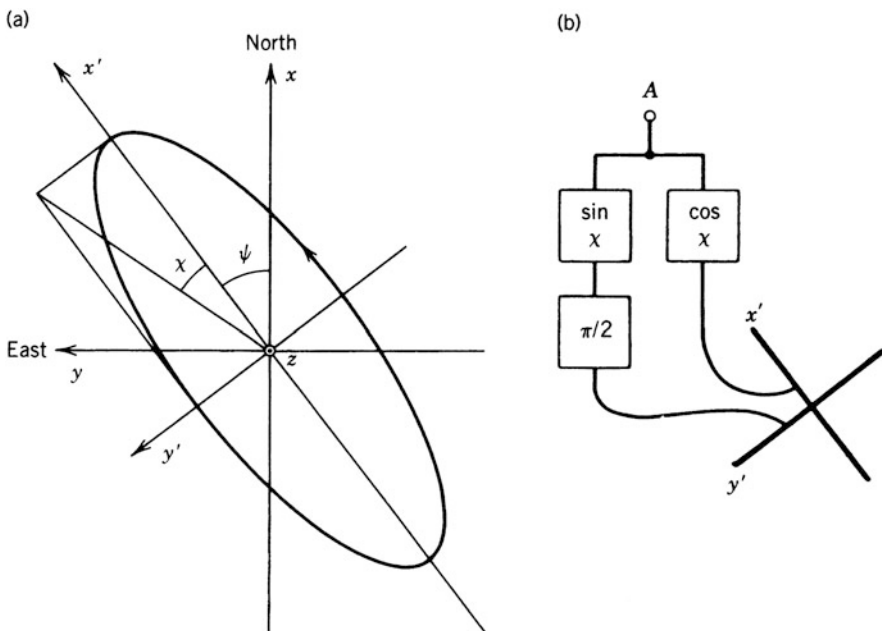
where  $m_\ell$ ,  $m_c$ , and  $m_t$  are the degrees of linear, circular, and total polarization, respectively, and  $\theta$  is the position angle of the plane of linear polarization. For monochromatic signals,  $m_t = 1$  and the polarization can be fully specified by just three parameters. For random signals such as those of cosmic origin,  $m_t \leq 1$ , and all four parameters are required. The Stokes parameters all have the dimensions of flux density or intensity, and they propagate in the same manner as the electromagnetic field. Thus, they can be determined by measurement or calculation at any point along a wave path, and their relative magnitudes define the state of polarization at that point. Stokes parameters combine additively for independent waves. When they are used to specify the total radiation from any point on a source,  $I$ , which measures the total intensity, is always positive, but  $Q$ ,  $U$ , and  $V$  can take both positive and negative values depending on the position angle or sense of rotation of the polarization. The corresponding visibility values measured with an interferometer are complex quantities, as will be discussed later.

In considering the response of interferometers and arrays, up to this point we have ignored the question of polarization. This simplification can be justified by the assumption that we have been dealing with completely unpolarized radiation for which only the parameter  $I$  is nonzero. In that case, the response of an interferometer with identically polarized antennas is proportional to the total flux density of the radiation. As will be shown below, in the more general case, the response is proportional to a linear combination of two or more Stokes parameters, where the combination is determined by the polarizations of the antennas. By observing with different states of polarization of the antennas, it is possible to separate the responses to the four parameters and determine the corresponding components of the visibility. The variation of each parameter over the source can thus be imaged individually, and the polarization of the radiation emitted at any point can be

determined. There are alternative methods of describing the polarization state of a wave, of which the coherency matrix is perhaps the most important (Ko 1967a,b). However, the classical treatment in terms of Stokes parameters remains widely used by astronomers, and we therefore follow it here.

### 4.7.1 Antenna Polarization Ellipse

The polarization of an antenna in either transmission or reception can be described in general by stating that the electric vector of a transmitted signal traces out an elliptical locus in the wavefront plane. Most antennas are designed so that the ellipse approximates a line or circle, corresponding to linear or circular polarization, in the central part of the main beam. However, precisely linear or circular responses are hardly achievable in practice. As shown in Fig. 4.8, the essential characteristics of the polarization ellipse are given by the position angle  $\psi$  of the major axis, and by



**Fig. 4.8** (a) Description of the general state of polarization of an antenna in terms of the characteristics of the ellipse generated by the electric vector in the transmission of a sinusoidal signal. The position angle  $\psi$  of the major axis is measured with respect to the  $x$  axis, which points toward the direction of north on the sky. A wave approaching from the sky is traveling toward the reader, in the direction of the positive  $z$  axis. For such a wave, the arrow on the ellipse indicates the direction of right-handed polarization. (b) Model antenna that radiates the electric field represented by the ellipse in (a) when a signal is applied to the terminal  $A$ .  $\cos \chi$  and  $\sin \chi$  indicate the amplitudes of the voltage responses of the units shown, and  $\pi/2$  indicates a phase lag.

the axial ratio, which it is convenient to express as the tangent of an angle  $\chi$ , where  $-\pi/4 \leq \chi \leq \pi/4$ .

An antenna of arbitrary polarization can be modeled in terms of two idealized dipoles as shown in Fig. 4.8b. Consider *transmitting* with this antenna by applying a signal waveform to the terminal  $A$ . The signals to the dipoles pass through networks with voltage responses proportional to  $\cos \chi$  and  $\sin \chi$ , and the signal to the  $y'$  dipole also passes through a network that introduces a  $\pi/2$  phase lag. Thus, the antenna produces field components of amplitude  $\mathcal{E}_{x'}$  and  $\mathcal{E}_{y'}$  in phase quadrature along the directions of the major and minor axes of the ellipse. If the antenna input is a radio frequency sine wave  $V_0 \cos 2\pi \nu t$ , then the field components are

$$\begin{aligned} \mathcal{E}_{x'} \cos(2\pi \nu t) &\propto V_0 \cos \chi \cos(2\pi \nu t) \\ \mathcal{E}_{y'} \sin(2\pi \nu t) &\propto V_0 \sin \chi \sin(2\pi \nu t) . \end{aligned} \quad (4.24)$$

In these equations, the  $y'$  component lags the  $x'$  component by  $\pi/2$ . If  $\chi = \pi/4$ , the radiated electric vector traces a circular locus with the sense of rotation from the  $x'$  axis to the  $y'$  axis (i.e., *counterclockwise* in Fig. 4.8a). This is consistent with the quarter-cycle delay in the signal to the  $y'$  dipole. Then a wave propagating in the positive  $z'$  direction of a right-handed coordinate system (i.e., toward the reader in Fig. 4.8a) is right circularly polarized in the IEEE (1977) definition. (This definition is now widely adopted, but in some of the older literature, such a wave would be defined as left circularly polarized.) The International Astronomical Union (IAU 1974) has adopted the IEEE definition and states that the position angle of the electric vector on the sky should be measured from north through east with reference to the system of right ascension and declination. The IAU also states that “the polarization of incoming radiation, for which the position angle,  $\theta$ , of the electric vector, measured at a fixed point in space, increases with time, is described as right-handed and positive.” Note that Stokes parameters in Eq. (4.19) specify only the field in the  $(x, y)$  plane, and to determine whether a circularly polarized wave is left- or right-handed, the direction of propagation must be given. From Eq. (4.19) and the definitions of  $E_x$  and  $E_y$  that precede them, a wave traveling in the positive  $z$  direction in right-handed coordinates is right circularly polarized for positive  $V$ .

In reception, an electric vector that rotates in a *clockwise* direction in Fig. 4.8 produces a voltage in the  $y'$  dipole that leads the voltage in the  $x'$  dipole by  $\pi/2$  in phase, and the two signals therefore combine in phase at  $A$ . For counterclockwise rotation, the signals at  $A$  are in antiphase and cancel one another. Thus, the antenna in Fig. 4.8 receives right-handed waves incident from the positive  $z$  direction (that is, traveling toward negative  $z$ ), and it transmits right-handed polarization in the direction toward positive  $z$ . To receive a right-handed wave propagating down from the sky (in the positive  $z$  direction), the polarity of one of the dipoles must be reversed, which requires that  $\chi = -\pi/4$ .



To determine the interferometer response, we begin by considering the output of the antenna modeled in Fig. 4.8b. We define the field components in complex form:

$$\begin{aligned} E_x(t) &= \mathcal{E}_x(t) e^{j[2\pi\nu t + \delta_x(t)]} , \\ E_y(t) &= \mathcal{E}_y(t) e^{j[2\pi\nu t + \delta_y(t)]} . \end{aligned} \quad (4.25)$$

The signal voltage received at  $A$  in Fig. 4.8b, expressed in complex form, is

$$V' = E_{x'} \cos \chi - jE_{y'} \sin \chi , \quad (4.26)$$

where the factor  $-j$  represents the  $\pi/2$  phase lag applied to the  $y'$  signal, for the fields represented by Eq. (4.25). Now we need to specify the polarization of the incident wave in terms of Stokes parameters. In accordance with IAU (1974), the axes used are in the directions of north and east on the sky, which are represented by  $x$  and  $y$  in Fig. 4.8a. In terms of the field in the  $x$  and  $y$  directions, the components of the field in the  $x'$  and  $y'$  directions are

$$\begin{aligned} E_{x'}(t) &= [\mathcal{E}_x(t) e^{j\delta_x(t)} \cos \psi + \mathcal{E}_y(t) e^{j\delta_y(t)} \sin \psi] e^{j2\pi\nu t} \\ E_{y'}(t) &= [-\mathcal{E}_x(t) e^{j\delta_x(t)} \sin \psi + \mathcal{E}_y(t) e^{j\delta_y(t)} \cos \psi] e^{j2\pi\nu t} . \end{aligned} \quad (4.27)$$

Derivation of the response at the output of the correlator for antennas  $m$  and  $n$  of an array involves straightforward manipulation of some rather lengthy expressions that are not reproduced here. The steps are as follows:

1. Substitute  $E_{x'}$  and  $E_{y'}$  from Eq. (4.27) into Eq. (4.26) to obtain the output of each antenna.
2. Indicate values of  $\psi$ ,  $\chi$ , and  $V'$  for the two antennas by subscripts  $m$  and  $n$  and calculate the correlator output,  $R_{mn} = G_{mn} \langle V'_m V'_n{}^* \rangle$ , where  $G_{mn}$  is an instrumental gain factor.
3. Substitute Stokes parameters for  $\mathcal{E}_x$ ,  $\mathcal{E}_y$ ,  $\delta_x$ ,  $\delta_y$  using Eq. (4.19) as follows:

$$\begin{aligned} \langle (\mathcal{E}_x e^{j\delta_x})(\mathcal{E}_x e^{j\delta_x})^* \rangle &= \langle \mathcal{E}_x^2 \rangle = \frac{1}{2}(I + Q) \\ \langle (\mathcal{E}_y e^{j\delta_y})(\mathcal{E}_y e^{j\delta_y})^* \rangle &= \langle \mathcal{E}_y^2 \rangle = \frac{1}{2}(I - Q) \\ \langle (\mathcal{E}_x e^{j\delta_x})(\mathcal{E}_y e^{j\delta_y})^* \rangle &= \langle \mathcal{E}_x \mathcal{E}_y e^{j(\delta_x - \delta_y)} \rangle = \frac{1}{2}(U + jV) \\ \langle (\mathcal{E}_x e^{j\delta_x})^*(\mathcal{E}_y e^{j\delta_y}) \rangle &= \langle \mathcal{E}_x \mathcal{E}_y e^{-j(\delta_x - \delta_y)} \rangle = \frac{1}{2}(U - jV) . \end{aligned} \quad (4.28)$$

The result is

$$\begin{aligned}
 R_{mn} = \frac{1}{2} G_{mn} \{ & I_v [\cos(\psi_m - \psi_n) \cos(\chi_m - \chi_n) + j \sin(\psi_m - \psi_n) \sin(\chi_m + \chi_n)] \\
 & + Q_v [\cos(\psi_m + \psi_n) \cos(\chi_m + \chi_n) + j \sin(\psi_m + \psi_n) \sin(\chi_m - \chi_n)] \\
 & + U_v [\sin(\psi_m + \psi_n) \cos(\chi_m + \chi_n) - j \cos(\psi_m + \psi_n) \sin(\chi_m - \chi_n)] \\
 & - V_v [\cos(\psi_m - \psi_n) \sin(\chi_m + \chi_n) + j \sin(\psi_m - \psi_n) \cos(\chi_m - \chi_n)] \} . \quad (4.29)
 \end{aligned}$$

In this equation, a subscript  $v$  has been added to Stokes parameter symbols to indicate that they represent the complex visibility for the distribution of the corresponding parameter over the source, not simply the intensity or brightness of the radiation. Equation (4.29) is a useful general formula that applies to all cases. It was originally derived by Morris et al. (1964) and later by Weiler (1973). In the derivation by Morris et al., the sign of  $V_v$  is opposite to that given by Weiler and in Eq. (4.29). This difference results from the convention for the sense of rotation for circular polarization. In the convention we have followed in Fig. 4.8, two identical antennas both adjusted to receive right circularly polarized radiation would have parameters  $\psi_m = \psi_n$  and  $\chi_m = \chi_n = -\pi/4$ . In Eq. (4.29), these values correspond to a positive sign for  $V_v$ . Thus, in Eq. (4.29), positive  $V_v$  represents right circular polarization incident from the sky, which is in agreement with the IAU definition in 1973 (IAU 1974). The derivation by Morris et al. predates the IAU definition and follows the commonly used convention at that time, in which the sign for  $V$  was the reverse of that in the IAU definition. Note that in what follows, the factor 1/2 in Eq. (4.29) is omitted and considered to be subsumed within the overall gain factor. Equation (4.29) was the main basis for polarization measurements in radio interferometry for at least three decades until an alternative formulation was developed by Hamaker et al. (1996). This later formulation is introduced in Sect. 4.8.

### 4.7.2 Stokes Visibilities

As noted above, the symbols  $I_v$ ,  $Q_v$ ,  $U_v$ , and  $V_v$  in Eq. (4.29) refer to the corresponding visibility values as measured by the spaced antennas. We shall therefore refer to these quantities as *Stokes visibilities*, following the nomenclature of Hamaker et al. (1996). Stokes visibilities are the quantities required in imaging polarized emission, and they can be derived from the correlator output values by using Eq. (4.29). This equation is considerably simplified when the nominal polarization characteristics of practical antennas are inserted. First, consider the case in which both antennas are identically polarized. Then  $\chi_m = \chi_n$ ,  $\psi_m = \psi_n$ , and Eq. (4.29) becomes

$$R_{mn} = G_{mn} [I_v + Q_v \cos 2\psi_m \cos 2\chi_m + U_v \sin 2\psi_m \cos 2\chi_m - V_v \sin 2\chi_m] . \quad (4.30)$$

In considering linearly polarized antennas, it is convenient to use subscripts  $x$  and  $y$  to indicate two orthogonal planes of polarization. For example,  $R_{xy}$  represents the correlator output for antenna  $m$  with polarization  $x$  and antenna  $n$  with polarization  $y$ . For linearly polarized antennas,  $\chi_m = \chi_n = 0$ . Consider two antennas, each with separate outputs for linear polarizations  $x$  and  $y$ . Then for parallel polarizations, omitting gain constants, we obtain from Eq. (4.30)

$$R_{xx} = I_v + Q_v \cos 2\psi_m + U_v \sin 2\psi_m . \quad (4.31)$$

Here,  $\psi_m$  is the position angle of the antenna polarization measured from celestial north in the direction of east. The  $y$  polarization angle is equal to the  $x$  polarization angle plus  $\pi/2$ . For  $\psi_m$  equal to  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ , and  $135^\circ$ , the output  $R_{xx}$  is proportional to  $(I_v + Q_v)$ ,  $(I_v + U_v)$ ,  $(I_v - Q_v)$ , and  $(I_v - U_v)$ , respectively. By using antennas with these polarization angles,  $I_v$ ,  $Q_v$ , and  $U_v$ , but not  $V_v$ , can be measured. In many cases, circular polarization is negligibly small, and the inability to measure  $V_v$  is not a serious problem. However,  $Q_v$  and  $U_v$  are often only a few percent of  $I_v$ , and in attempting to measure them with identical feeds, one faces the usual problems of measuring a small difference in two much larger quantities. The same is true if one attempts to measure  $V_v$  using identical circular feeds for which  $\chi = \pm\pi/4$  and the response is proportional to  $(I_v \mp V_v)$ . These problems are reduced by using oppositely polarized feeds to measure  $Q_v$ ,  $U_v$ , or  $V_v$ . For an example of measurement of  $V_v$ , see Weiler and Raimond (1976).

With oppositely polarized feeds, we insert in Eq. (4.29)  $\psi_n = \psi_m + \pi/2$ , and  $\chi_m = -\chi_n$ . For linear polarization, the  $\chi$  terms are zero and the planes of polarization orthogonal. The antennas are then described as cross-polarized, as typified by crossed dipoles. Omitting constant gain factors and using the  $x$  and  $y$  subscripts defined above, we obtain for the correlator output

$$\begin{aligned} R_{xy} &= -Q_v \sin 2\psi_m + U_v \cos 2\psi_m + jV_v \\ R_{yx} &= -Q_v \sin 2\psi_m + U_v \cos 2\psi_m - jV_v , \end{aligned} \quad (4.32)$$

where  $\psi_m$  refers to the angle of the plane of polarization in the direction ( $x$  or  $y$ ) indicated by the first subscript of the  $R$  term in the same equation. Then for  $\psi_m$  equal to  $0^\circ$  and  $45^\circ$ , the  $R_{xy}$  response is proportional to  $(U_v + jV_v)$  and  $(-Q_v + jV_v)$ . If  $V_v$  is assumed to be zero, this suffices to measure the polarized component. If both antennas provide outputs for cross-polarized signals, the outputs of which go to two separate receiving channels at each antenna, four correlators can be used for each antenna pair. These provide responses for both crossed and parallel pairs, as listed in Table 4.1. Thus, if the planes of polarization can be periodically rotated through  $45^\circ$  as indicated by position angles I and II in Table 4.1, for example, by rotating antenna feeds, then  $Q_v$ ,  $U_v$ , and  $V_v$  can be measured without taking differences between responses involving  $I_v$ . The use of rotating feeds has, however, proved to be of limited practicality. Rotating the feed relative to the main reflector is likely to have a small but significant effect on the beam shape and polarization properties. This is

**Table 4.1** Stokes visibilities vs. position angles

Position angles		Stokes visibilities	
$m$	$n$	measured	
$0^\circ$	$0^\circ$	$I_v + Q_v$	Position angle I
$0^\circ$	$90^\circ$	$U_v + jV_v$	"
$90^\circ$	$0^\circ$	$U_v - jV_v$	"
$90^\circ$	$90^\circ$	$I_v - Q_v$	"
$45^\circ$	$45^\circ$	$I_v + U_v$	Position angle II
$45^\circ$	$135^\circ$	$-Q_v + jV_v$	"
$135^\circ$	$45^\circ$	$-Q_v - jV_v$	"
$135^\circ$	$135^\circ$	$I_v - U_v$	"

because the rotation will cause deviations from circular symmetry in the radiation pattern of the feeds to interact differently with the shadowing effects of the focal support structure and any departures from circular symmetry in the main reflector. Furthermore, in radio astronomy systems designed for the greatest sensitivity, the feed together with the low-noise amplifiers and a cryogenically refrigerated Dewar are often built as one monolithic unit that cannot easily be rotated. However, for antennas on altazimuth mounts, the variation of the parallactic angle with hour angle causes the antenna response pattern to rotate on the sky as a source is tracked in hour angle. Conway and Kronberg (1969) pointed out this advantage of altazimuth mounts, which enables instrumental effects to be distinguished from the true polarization of the source if observations continue for a period of several hours.

An example of a different arrangement of linearly polarized feeds, which has been used at the Westerbork Synthesis Radio Telescope, is described by Weiler (1973). The antennas are equatorially mounted and the parallactic angle of the polarization remains fixed as a source is tracked. The outputs of the antennas that are movable on rail track are correlated with those from the antennas in fixed locations. Table 4.2 shows the measurements when the position angles of the planes of polarization for the movable antennas are  $45^\circ$  and  $135^\circ$  and those of the fixed antennas  $0^\circ$  and  $90^\circ$ . Although the responses are reduced by a factor of  $\sqrt{2}$  relative to those in Table 4.1, there is no loss in sensitivity since each Stokes visibility appears at all four correlator outputs. Note that since only signals from antennas with different polarization configurations are cross-correlated, this scheme does not make use of all possible polarization products.

Opposite circularly polarized feeds offer certain advantages for measurements of linear polarization. In determining the responses, an arbitrary position angle  $\psi_m$  for antenna  $m$  is included to represent the effect of rotation caused, for example, by an altazimuth antenna mount. If the antennas provide simultaneous outputs for opposite

**Table 4.2** Stokes visibilities vs. position angles

Position angles		Stokes visibilities measured
$m$	$n$	
$0^\circ$	$45^\circ$	$(I_v + Q_v + U_v + jV_v)/\sqrt{2}$
$0^\circ$	$135^\circ$	$(-I_v - Q_v + U_v + jV_v)/\sqrt{2}$
$90^\circ$	$45^\circ$	$(I_v - Q_v + U_v - jV_v)/\sqrt{2}$
$90^\circ$	$135^\circ$	$(I_v - Q_v - U_v + jV_v)/\sqrt{2}$

senses of rotation (denoted by  $r$  and  $\ell$ ) and four correlation products are generated for each antenna pair, the outputs are proportional to the quantities in Table 4.3.

Here, we have made  $\psi_\ell = \psi_r + \pi/2$ , and  $\chi = -\pi/4$  for right circular polarization and  $\chi = \pi/4$  for left circular. The feeds need not be rotated during an observation, and the responses to  $Q_v$  and  $U_v$  are separated from those to  $I_v$ . The expressions in Table 4.3 can be simplified by choosing values of  $\psi_r$  such as  $\pi/2$ ,  $\pi/4$ , or 0. For example, if  $\psi_r = 0$ , the sum of the  $r\ell$  and  $\ell r$  responses is a measure of Stokes visibility  $U_v$ . Again, the effects of the rotation of the position angle with altazimuth mounts must be taken into account. Conway and Kronberg (1969) appear to have been the first to use an interferometer with circularly polarized antennas to measure linear polarization in weakly polarized sources. Circularly polarized antennas have since been commonly used in radio astronomy.

### 4.7.3 Instrumental Polarization

The responses with the various combinations of linearly and circularly polarized antennas discussed above are derived on the assumption that the polarization is exactly linear or circular and that the position angles of the linear feeds are exactly determined. This is not the case in practice, and the polarization ellipse can never be maintained as a perfect circle or straight line. The nonideal characteristics of the antennas cause an unpolarized source to appear polarized and are therefore referred to as *instrumental polarization*. The effect of these deviations from ideal behavior

**Table 4.3** Stokes visibilities vs. sense of rotation

Sense of rotation		Stokes visibilities measured
$m$	$n$	
$r$	$r$	$I_v + V_v$
$r$	$\ell$	$(-jQ_v + U_v)e^{-j2\psi_m}$
$\ell$	$r$	$(-jQ_v - U_v)e^{j2\psi_m}$
$\ell$	$\ell$	$I_v - V_v$

can be calculated from Eq. (4.29) if the deviations are known. In the expressions in Tables 4.1–4.3, the responses given are only the major terms, and if the instrumental terms are included, all four Stokes visibilities are, in general, involved. For example, consider the case of crossed linear feeds with nominal position angles  $0^\circ$  and  $90^\circ$ . Let the actual values of  $\psi$  and  $\chi$  be such that  $(\psi_x + \psi_y) = \pi/2 + \Delta\psi^+$ ,  $(\psi_x - \psi_y) = -\pi/2 + \Delta\psi^-$ ,  $\chi_x + \chi_y = \Delta\chi^+$ , and  $\chi_x - \chi_y = \Delta\chi^-$ . Then from Eq. (4.29),

$$R_{xy} \simeq I_v(\Delta\psi^- - j\Delta\chi^+) - Q_v(\Delta\psi^+ - j\Delta\chi^-) + U_v + jV_v . \quad (4.33)$$

Generally, antennas can be adjusted so that the  $\Delta$  terms are no more than  $\sim 1^\circ$ , and here we have assumed that they are small enough that their cosines can be approximated by unity, their sines by the angles, and products of two sines by zero. Instrumental polarization is often different for the antennas even if they are structurally similar, and corrections must be made to the visibility data before they are combined into an image.

Although we have derived expressions for deviations of the antenna polarizations from the ideal in terms of the ellipticity and orientation of the polarization ellipse in Eq. (4.29), it is not necessary to know these parameters for the antennas so long as it is possible to remove the instrumental effects from the measurements, so that they do not appear in the final image. In calibrating the antenna responses, an approach that is widely preferred is to specify the instrumental polarization in terms of the response of the antenna to a wave of polarization that is orthogonal or opposite-handed with respect to the nominal antenna response. Thus, for linearly polarized antennas, following the analysis of Sault et al. (1991), we can write

$$v'_x = v_x + D_x v_y \quad \text{and} \quad v'_y = v_y + D_y v_x , \quad (4.34)$$

where subscripts  $x$  and  $y$  indicate two orthogonal planes of polarization, the  $v'$  terms indicate the signal received, the  $v$  terms indicate the signal that would be received with an ideally polarized antenna, and the  $D$  terms indicate the response of the real antenna to the polarization orthogonal to the nominal polarization. The  $D$  terms are often described as the *leakage* of the orthogonal polarization into the antenna (Bignell 1982) and represent the instrumental polarization. For each polarization state, the leakage is specified by one complex number, that is, the same number of terms as the two real numbers required to specify the ellipticity and orientation of the polarization ellipse. In Appendix 4.2, expressions for  $D_x$  and  $D_y$  are derived in terms of the parameters of the polarization ellipse:

$$D_x \simeq \psi_x - j\chi_x \quad \text{and} \quad D_y \simeq -\psi_y + j\chi_y , \quad (4.35)$$

where the approximations are valid for small values of the  $\chi$  and  $\psi$  parameters. Note that in Eq. (4.35),  $\psi_y$  is measured with respect to the  $y$  direction. For an ideal linearly polarized antenna,  $\chi_x$  and  $\chi_y$  are both zero, and the polarization in the  $x$  and  $y$  planes is precisely aligned with, and orthogonal to, the  $x$  direction with respect

to the antenna. Thus, for an ideal antenna,  $\psi_x$  and  $\psi_y$  are also zero. For a practical antenna, the terms in Eq. (4.35) represent limits of accuracy in the hardware, and we see that the real and imaginary parts of the leakage terms can be related to the misalignment and ellipticity, respectively.

For a pair of antennas  $m$  and  $n$ , the leakage terms allow us to express the measured correlator outputs  $R'_{xx}$ ,  $R'_{yy}$ ,  $R'_{xy}$ , and  $R'_{yx}$  in terms of the unprimed quantities that represent the corresponding correlations as they would be measured with ideally polarized antennas:

$$\begin{aligned}
 R'_{xx}/(g_{xm}g_{xn}^*) &= R_{xx} + D_{xm}R_{yx} + D_{xn}^*R_{xy} + D_{xm}D_{xn}^*R_{yy} \\
 R'_{xy}/(g_{xm}g_{yn}^*) &= R_{xy} + D_{xm}R_{yy} + D_{yn}^*R_{xx} + D_{xm}D_{yn}^*R_{yx} \\
 R'_{yx}/(g_{ym}g_{xn}^*) &= R_{yx} + D_{ym}R_{xx} + D_{xn}^*R_{yy} + D_{ym}D_{xn}^*R_{xy} \\
 R'_{yy}/(g_{ym}g_{yn}^*) &= R_{yy} + D_{ym}R_{xy} + D_{yn}^*R_{yx} + D_{ym}D_{yn}^*R_{xx} .
 \end{aligned} \tag{4.36}$$

The  $g$  terms represent the voltage gains of the corresponding signal channels. They are complex quantities representing amplitude and phase, and the equations can be normalized so that the values of the individual  $g$  terms do not differ greatly from unity. Note that Eq. (4.36) contain no small-term approximations. However, the leakage terms are typically no more than a few percent, and products of two such terms will be omitted at this point. Then, from Eqs. (4.31) and (4.32), the responses can be written in terms of the Stokes visibilities as follows:

$$\begin{aligned}
 R'_{xx}/(g_{xm}g_{xn}^*) &= I_v + Q_v[\cos 2\psi_m - (D_{xm} + D_{xn}^*) \sin 2\psi_m] \\
 &\quad + U_v[\sin 2\psi_m + (D_{xm} + D_{xn}^*) \cos 2\psi_m] - jV_v(D_{xm} - D_{xn}^*) \\
 R'_{xy}/(g_{xm}g_{yn}^*) &= I_v(D_{xm} + D_{yn}^*) - Q_v[\sin 2\psi_m + (D_{xm} - D_{yn}^*) \cos 2\psi_m] \\
 &\quad + U_v[\cos 2\psi_m - (D_{xm} - D_{yn}^*) \sin 2\psi_m] + jV_v \\
 R'_{yx}/(g_{ym}g_{xn}^*) &= I_v(D_{ym} + D_{xn}^*) - Q_v[\sin 2\psi_m - (D_{ym} - D_{xn}^*) \cos 2\psi_m] \\
 &\quad + U_v[\cos 2\psi_m + (D_{ym} - D_{xn}^*) \sin 2\psi_m] - jV_v \\
 R'_{yy}/(g_{ym}g_{yn}^*) &= I_v - Q_v[\cos 2\psi_m + (D_{ym} + D_{yn}^*) \sin 2\psi_m] \\
 &\quad - U_v[\sin 2\psi_m - (D_{ym} + D_{yn}^*) \cos 2\psi_m] + jV_v(D_{ym} - D_{yn}^*) .
 \end{aligned} \tag{4.37}$$

Note that  $\psi_m$  refers to the polarization ( $x$  or  $y$ ) indicated by the first of the two subscripts of the  $R'$  term in the same equation. Sault et al. (1991) describe Eq. (4.37) as representing the strongly polarized case. In deriving them, no restriction was placed on the magnitudes of the Stokes visibility terms, but the leakage terms of the antennas are assumed to be small. In the case where the source is only weakly polarized, the products of  $Q_v$ ,  $U_v$ , and  $V_v$  with leakage terms can be omitted.

Equation (4.37) then become

$$\begin{aligned}
 R'_{xx}/(g_{xm}g_{xn}^*) &= I_v + Q_v \cos 2\psi_m + U_v \sin 2\psi_m \\
 R'_{xy}/(g_{xm}g_{yn}^*) &= I_v(D_{xm} + D_{yn}^*) - Q_v \sin 2\psi_m + U_v \cos 2\psi_m + jV_v \\
 R'_{yx}/(g_{ym}g_{xn}^*) &= I_v(D_{ym} + D_{xn}^*) - Q_v \sin 2\psi_m + U_v \cos 2\psi_m - jV_v \\
 R'_{yy}/(g_{ym}g_{yn}^*) &= I_v - Q_v \cos 2\psi_m - U_v \sin 2\psi_m .
 \end{aligned} \tag{4.38}$$

If the antennas are operating well within the upper frequency limit of their performance, the polarization terms can be expected to remain largely constant with time since gravitational deflections that vary with pointing should be small. The instrumental gain terms can contain components due to the atmosphere, which may vary on time scales of seconds or minutes, and they also include any effects of the receiver electronics.

In the case of circularly polarized antennas, leakage terms can also be defined and similar expressions for the instrumental response derived. The leakage terms are given by the following equations:

$$v'_r = v_r + D_r v_\ell \text{ and } v'_\ell = v_\ell + D_\ell v_r , \tag{4.39}$$

where, as before, the  $v'$  terms are the measured signal voltages, the unprimed  $v$  terms are the signals that would be observed with an ideally polarized antenna, and the  $D$  terms are the leakages. The subscripts  $r$  and  $\ell$  indicate the right and left senses of rotation. Again, the relationship between the leakage terms and the orientation and ellipticity of the antenna responses is derived in Appendix 4.2. The results, which in this case require no small-angle approximations, are

$$D_r = e^{j2\psi_r} \tan \Delta\chi_r \text{ and } D_\ell = e^{-j2\psi_\ell} \tan \Delta\chi_\ell , \tag{4.40}$$

where the  $\Delta$  terms are defined by  $\chi_r = -45^\circ + \Delta\chi_r$  and  $\chi_\ell = 45^\circ + \Delta\chi_\ell$ . To derive expressions for the outputs of an interferometer in terms of the leakage terms and Stokes visibilities, the four measured correlator outputs are represented by  $R'_{rr}$ ,  $R'_{\ell\ell}$ ,  $R'_{r\ell}$ , and  $R'_{\ell r}$ . These are related to the corresponding (unprimed) quantities that would be observed with ideally polarized antennas as follows:

$$\begin{aligned}
 R'_{rr}/(g_{rm}g_{rn}^*) &= R_{rr} + D_{rm}R_{\ell r} + D_{rn}^*R_{r\ell} + D_{rm}D_{rn}^*R_{\ell\ell} \\
 R'_{r\ell}/(g_{rm}g_{\ell n}^*) &= R_{r\ell} + D_{rm}R_{\ell\ell} + D_{\ell n}^*R_{rr} + D_{rm}D_{\ell n}^*R_{\ell r} \\
 R'_{\ell r}/(g_{\ell m}g_{rn}^*) &= R_{\ell r} + D_{\ell m}R_{rr} + D_{rn}^*R_{\ell\ell} + D_{\ell m}D_{rn}^*R_{r\ell} \\
 R'_{\ell\ell}/(g_{\ell m}g_{\ell n}^*) &= R_{\ell\ell} + D_{\ell m}R_{r\ell} + D_{\ell n}^*R_{\ell r} + D_{\ell m}D_{\ell n}^*R_{rr} .
 \end{aligned} \tag{4.41}$$



Now, from the expressions in Table 4.3, the outputs in terms of the Stokes visibilities are

$$\begin{aligned}
 R'_{rr}/(g_{rm}g_{rn}^*) &= I_v(1 + D_{rm}D_{rn}^*) - jQ_v(D_{rm}e^{j2\psi_m} + D_{rn}^*e^{-j2\psi_m}) \\
 &\quad - U_v(D_{rm}e^{j2\psi_m} - D_{rn}^*e^{-j2\psi_m}) + V_v(1 - D_{rm}D_{rn}^*) \\
 R'_{r\ell}/(g_{rm}g_{\ell n}^*) &= I_v(D_{rm} + D_{\ell n}^*) - jQ_v(e^{-j2\psi_m} + D_{rm}D_{\ell n}^*e^{j2\psi_m}) \\
 &\quad + U_v(e^{-j2\psi_m} - D_{rm}D_{\ell n}^*e^{j2\psi_m}) - V_v(D_{rm} - D_{\ell n}^*) \\
 R'_{\ell r}/(g_{\ell m}g_{rn}^*) &= I_v(D_{\ell m} + D_{rn}^*) - jQ_v(e^{j2\psi_m} + D_{\ell m}D_{rn}^*e^{-j2\psi_m}) \\
 &\quad - U_v(e^{j2\psi_m} - D_{\ell m}D_{rn}^*e^{-j2\psi_m}) + V_v(D_{\ell m} - D_{rn}^*) \\
 R'_{\ell\ell}/(g_{\ell m}g_{\ell n}^*) &= I_v(1 + D_{\ell m}D_{\ell n}^*) - jQ_v(D_{\ell m}e^{-j2\psi_m} + D_{\ell n}^*e^{j2\psi_m}) \\
 &\quad + U_v(D_{\ell m}e^{-j2\psi_m} - D_{\ell n}^*e^{j2\psi_m}) - V_v(1 - D_{\ell m}D_{\ell n}^*) .
 \end{aligned} \tag{4.42}$$

Here again,  $\psi_m$  refers to the polarization ( $r$  or  $\ell$ ) indicated by the first of the two subscripts of the  $R'$  term in the same equation. The angle  $\psi_m$  represents the parallactic angle plus any instrumental offset. We have made no approximations in deriving Eq. (4.42) [in the similar Eq. (4.37), products of two  $D$  terms were omitted]. If the leakage terms are small, then any product of two of them can be omitted, as in the strongly polarized case for linearly polarized antennas in Eq. (4.37). The weakly polarized case is derived from the strongly polarized case by further omitting products of  $Q_v$ ,  $U_v$ , and  $V_v$  with the leakage terms and is as follows:

$$\begin{aligned}
 R'_{rr}/(g_{rm}g_{rn}^*) &= I_v + V_v \\
 R'_{r\ell}/(g_{rm}g_{\ell n}^*) &= I_v(D_{rm} + D_{\ell n}^*) - (jQ_v - U_v)e^{-j2\psi_m} \\
 R'_{\ell r}/(g_{\ell m}g_{rn}^*) &= I_v(D_{\ell m} + D_{rn}^*) - (jQ_v + U_v)e^{j2\psi_m} \\
 R'_{\ell\ell}/(g_{\ell m}g_{\ell n}^*) &= I_v - V_v .
 \end{aligned} \tag{4.43}$$

Similar expressions<sup>3</sup> are given by Fomalont and Perley (1989). To make use of the expressions that have been derived for the response in terms of the leakage and gain factors, we need to consider how such quantities can be calibrated, and this is discussed later.

---

<sup>3</sup>In comparing expressions for polarimetry by different authors, note that differences of signs or of the factor  $j$  can result from differences in the way the parallactic angle is defined with respect to the antenna, and similar arbitrary factors.

#### 4.7.4 Matrix Formulation

The description of polarimetry given above, using the ellipticity and orientation of the antenna response, is based on a physical model of the antenna and the electromagnetic wave, as in Eq. (4.29). Historically, studies of *optical* polarization have developed over a much longer period. A description of radio polarimetry following an approach originally developed in optics is given in Hamaker et al. (1996) and in more detail in four papers: Hamaker et al. (1996), Sault et al. (1996), Hamaker (2000), and Hamaker (2006). The mathematical analysis is largely in terms of matrix algebra, and in particular, it allows the responses of different elements of the signal path such as the atmosphere, the antennas, and the electronic system to be represented independently and then combined in the final solution. This approach is convenient for detailed analysis including effects of the atmosphere, ionosphere, etc.

In the matrix formulation, the electric fields of the polarized wave are represented by a two-component column vector. The effect of any linear system on the wave, or on the voltage waveforms of the signal after reception, can be represented by a  $2 \times 2$  matrix of the form shown below:

$$\begin{bmatrix} E'_p \\ E'_q \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \begin{bmatrix} E_p \\ E_q \end{bmatrix}, \quad (4.44)$$

where  $E_p$  and  $E_q$  represent the input polarization state (orthogonal linear or opposite circular) and  $E'_p$  and  $E'_q$  represent the outputs. The  $2 \times 2$  matrix in Eq. (4.44) is referred to as a Jones matrix (Jones 1941), and any simple linear operation on the wave can be represented by such a matrix. Jones matrices can represent a rotation of the wave relative to the antenna; the response of the antenna, including polarization leakage effects; or the amplification of the signals in the receiving system up to the correlator input. The combined effect of these operations is represented by the product of the corresponding Jones matrices, just as the effect on a scalar voltage can be represented by the product of gains and response factors for different stages of the receiving system. For a wave specified in terms of opposite circularly polarized components, Jones matrices for these operations can take the following forms:

$$\mathbf{J}_{\text{rotation}} = \begin{bmatrix} \exp(j\theta) & 0 \\ 0 & \exp(-j\theta) \end{bmatrix} \quad (4.45)$$

$$\mathbf{J}_{\text{leakage}} = \begin{bmatrix} 1 & D_r \\ D_\ell & 1 \end{bmatrix} \quad (4.46)$$

$$\mathbf{J}_{\text{gain}} = \begin{bmatrix} G_r & 0 \\ 0 & G_\ell \end{bmatrix}. \quad (4.47)$$

Here,  $\theta$  represents a rotation relative to the antenna, and the cross polarization in the antenna is represented by the off-diagonal<sup>4</sup> leakage terms  $D_r$  and  $D_\ell$ . For a nonideal antenna, the diagonal terms will be slightly different from unity, but in this case, the difference is subsumed into the gain matrix of the two channels. The gain of both the antenna and the electronics can be represented by a single matrix, and since any cross coupling of the signals in the amplifiers can be made negligibly small, only the diagonal terms are significant in the gain matrix.

Let  $\mathbf{J}_m$  represent the product of the Jones matrices required to represent the linear operations on the signal of antenna  $m$  up to the point where it reaches the correlator input. Let  $\mathbf{J}_n$  be the same matrix for antenna  $n$ . The signals at the inputs to the correlator are  $\mathbf{J}_m \mathbf{E}_m$  and  $\mathbf{J}_n \mathbf{E}_n$ , where  $\mathbf{E}_m$  and  $\mathbf{E}_n$  are the vectors representing the signals at the antenna. The correlator output is the *outer product* (also known as the Kronecker, or tensor, product) of the signals at the input:

$$\mathbf{E}'_m \otimes \mathbf{E}'_n{}^* = (\mathbf{J}_m \mathbf{E}_m) \otimes (\mathbf{J}_n^* \mathbf{E}_n^*), \quad (4.48)$$

where  $\otimes$  represents the outer product. The outer product  $\mathbf{A} \otimes \mathbf{B}$  is formed by replacing each element  $a_{ik}$  of  $\mathbf{A}$  by  $a_{ik} \mathbf{B}$ . Thus, the outer product of two  $n \times n$  matrices is a matrix of order  $n^2 \times n^2$ . It is also a property of the outer product that

$$(\mathbf{A}_i \mathbf{B}_i) \otimes (\mathbf{A}_k \mathbf{B}_k) = (\mathbf{A}_i \otimes \mathbf{A}_k) (\mathbf{B}_i \otimes \mathbf{B}_k). \quad (4.49)$$

Thus, we can write Eq. (4.48) as

$$\mathbf{E}'_m \otimes \mathbf{E}'_n{}^* = (\mathbf{J}_m \otimes \mathbf{J}_n^*) (\mathbf{E}_m \otimes \mathbf{E}_n^*). \quad (4.50)$$

The time average of Eq. (4.50) represents the correlator output, which is

$$\mathbf{R}_{mn} = \langle \mathbf{E}'_m \otimes \mathbf{E}'_n{}^* \rangle = \begin{bmatrix} R'_{mn}{}^{pp} \\ R'_{mn}{}^{pq} \\ R'_{mn}{}^{qp} \\ R'_{mn}{}^{qq} \end{bmatrix}, \quad (4.51)$$

where  $p$  and  $q$  indicate opposite polarization states. The column vector in Eq. (4.51) is known as the *coherency vector* and represents the four cross products from the correlator outputs for antennas  $m$  and  $n$ . From Eq. (4.50), it is evident that the measured coherency vector  $\mathbf{R}'_{mn}$ , which includes the effects of instrumental responses, and the true coherency vector  $\mathbf{R}_{mn}$ , which is free from such effects, are related by the outer product of the Jones matrices that represent the instrumental effects:

$$\mathbf{R}'_{mn} = (\mathbf{J}_m \otimes \mathbf{J}_n^*) \mathbf{R}_{mn}. \quad (4.52)$$

---

<sup>4</sup>The diagonal terms are those that move downward from left to right, and the off-diagonal terms slope in the opposite direction.

To determine the response of an interferometer in terms of the Stokes visibilities of the input radiation, which are complex quantities, we introduce the Stokes visibility vector

$$\mathcal{V}_{Smm} = \begin{bmatrix} I_v \\ Q_v \\ U_v \\ V_v \end{bmatrix}. \quad (4.53)$$

The Stokes visibilities can be regarded as an alternate coordinate system for the coherency vector. Let  $\mathbf{S}$  be a  $4 \times 4$  transformation matrix from Stokes parameters to the polarization coordinates of the antennas. Then we have

$$\mathbf{R}'_{mn} = (\mathbf{J}_m \otimes \mathbf{J}_n^*) \mathbf{S} \mathcal{V}_{Smm}. \quad (4.54)$$

For ideal antennas with crossed (orthogonal) linear polarization, the response in terms of Stokes visibilities is given by the expressions in Table 4.1. We can write this result in matrix form as

$$\begin{bmatrix} R_{xx} \\ R_{xy} \\ R_{yx} \\ R_{yy} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & j \\ 0 & 0 & 1 & -j \\ 1 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} I_v \\ Q_v \\ U_v \\ V_v \end{bmatrix}, \quad (4.55)$$

where the subscripts  $x$  and  $y$  here refer to polarization position angles  $0^\circ$  and  $90^\circ$ , respectively. Similarly for opposite-hand circular polarization, we can write the expressions in Table 4.3 as

$$\begin{bmatrix} R_{rr} \\ R_{r\ell} \\ R_{\ell r} \\ R_{\ell\ell} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & -je^{-j2\psi_m} & e^{-j2\psi_m} & 0 \\ 0 & -je^{j2\psi_m} & -e^{j2\psi_m} & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} I_v \\ Q_v \\ U_v \\ V_v \end{bmatrix}. \quad (4.56)$$

The  $4 \times 4$  matrices in Eqs. (4.55) and (4.56) are transformation matrices from Stokes visibilities to the coherency vector for crossed linear and opposite circular polarizations, respectively. These  $4 \times 4$  matrices are known as Mueller matrices following the terminology established in optics.<sup>5</sup> Note that these matrices depend on the particular formulation we have used to specify the angles  $\psi$  and  $\chi$ , and other factors in Fig. 4.8, which may not be identical to corresponding parameters used by other authors.

<sup>5</sup>Further explanation of Jones and Mueller matrices can be found in textbooks on optics [e.g., O'Neill (1963)].

The expression  $\mathbf{S}^{-1}(\mathbf{J}_m \otimes \mathbf{J}_n^*)\mathbf{S}$  is a matrix that relates the input and output coherency vectors of a system where these quantities are in Stokes coordinate form. As an example of the matrix usage, we can derive the effect of the leakage and gain factors in the case of opposite circular polarizations. For antenna  $m$ , the Jones matrix  $\mathbf{J}_m$  is the product of the Jones matrices for leakage and gain as follows:

$$\mathbf{J}_m = \begin{bmatrix} g_{rm} & 0 \\ 0 & g_{\ell m} \end{bmatrix} \begin{bmatrix} 1 & D_{rm} \\ D_{\ell m} & 1 \end{bmatrix} = \begin{bmatrix} g_{rm} & g_{rm}D_{rm} \\ g_{\ell m}D_{\ell m} & g_{\ell m} \end{bmatrix}. \quad (4.57)$$

Here, the  $g$  terms represent voltage gain, the  $D$  terms represent leakage, and the subscripts  $r$  and  $\ell$  indicate polarization. A corresponding matrix  $\mathbf{J}_n$  is required for antenna  $n$ . Then if we use primes to indicate the components of the coherency vector (i.e., the correlator outputs) for antennas  $m$  and  $n$ , we can write

$$\begin{bmatrix} R'_{rr} \\ R'_{r\ell} \\ R'_{\ell r} \\ R'_{\ell\ell} \end{bmatrix} = \mathbf{J}_m \otimes \mathbf{J}_n^* \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & -je^{-j2\psi_m} & e^{-j2\psi_m} & 0 \\ 0 & -je^{j2\psi_m} & -e^{j2\psi_m} & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} I_v \\ Q_v \\ U_v \\ V_v \end{bmatrix}, \quad (4.58)$$

where the  $4 \times 4$  matrix is the one relating Stokes visibilities to the coherency vector in Eq. (4.56). Also, we have

$$\mathbf{J}_m \otimes \mathbf{J}_n^* = \begin{bmatrix} g_{rm}g_{rn}^* & g_{rm}g_{rn}^*D_{rn}^* & g_{rm}g_{rn}^*D_{rm} & g_{rm}g_{rn}^*D_{rm}D_{rn}^* \\ g_{rm}g_{\ell n}^*D_{\ell n}^* & g_{rm}g_{\ell n}^* & g_{rm}g_{\ell n}^*D_{rm}D_{\ell n}^* & g_{rm}g_{\ell n}^*D_{rm} \\ g_{\ell m}g_{rn}^*D_{\ell m} & g_{\ell m}g_{rn}^*D_{\ell m}D_{rn}^* & g_{\ell m}g_{rn}^* & g_{\ell m}g_{rn}^*D_{rn}^* \\ g_{\ell m}g_{\ell n}^*D_{\ell m}D_{\ell n}^* & g_{\ell m}g_{\ell n}^*D_{\ell m} & g_{\ell m}g_{\ell n}^*D_{\ell n}^* & g_{\ell m}g_{\ell n}^* \end{bmatrix}. \quad (4.59)$$

Insertion of Eq. (4.59) into Eq. (4.58) and reduction of the matrix products results in Eq. (4.42) for the response with circularly polarized feeds. The use of matrices is convenient since they provide a format for expressions representing different effects, which can then be combined as required.

### 4.7.5 Calibration of Instrumental Polarization

The fractional polarization of many astronomical sources is of magnitude comparable to that of the leakage and gain terms that are used above to define the instrumental polarization. Thus, to obtain an accurate measure of the polarization of a source, the leakage and gain terms must be accurately calibrated. It may be necessary to determine the calibration independently for each set of observations since the gain terms may be functions of the temperature and state of adjustment

of the electronics and cannot be assumed to remain constant from one observing session to another. Making observations (i.e., measuring the coherency vector) of sources for which the polarization parameters are already known is clearly a way of determining the leakage and gain terms. The number of unknown parameters to be calibrated is proportional to the number of antennas,  $n_a$ , but the number of measurements is proportional to the number of baselines,  $n_a(n_a - 1)/2$ . The unknown parameters are therefore usually overdetermined, and a least-mean-squares solution may be the best procedure.

For any antenna with orthogonally polarized receiving channels, there are seven degrees of freedom, that is, seven unknown quantities, that must be calibrated to allow full interpretation of the measured Stokes visibilities. This applies to the general case, and the number can be reduced if approximations are made for weak polarization or small instrumental polarization. In terms of the polarization ellipses, these unknowns can be regarded as the orientations and ellipticities of the two orthogonal feeds and the complex gains (amplitudes and phases) of the two receiving channels. When the outputs of two antennas are combined, only the differences in the instrumental phases are required, leaving seven degrees of freedom per antenna. Sault et al. (1996) make the same point from the consideration of the Jones matrix of an antenna, which contains four complex quantities. They also give a general result that illustrates the seven degrees of freedom or unknown terms. This expresses the relationship between the uncorrected (measured) Stokes visibilities (indicated by primes) and the true values of the Stokes visibilities, in terms of seven  $\gamma$  and  $\delta$  terms:

$$\begin{bmatrix} I'_v - I_v \\ Q'_v - Q_v \\ U'_v - U_v \\ V'_v - V_v \end{bmatrix} = -\frac{1}{2} \begin{bmatrix} \gamma_{++} & \gamma_{+-} & \delta_{+-} & -j\delta_{-+} \\ \gamma_{+-} & \gamma_{++} & \delta_{++} & -j\delta_{--} \\ \delta_{+-} & -\delta_{++} & \gamma_{++} & j\gamma_{--} \\ -j\delta_{-+} & -j\delta_{--} & j\gamma_{--} & \gamma_{++} \end{bmatrix} \begin{bmatrix} I_v \\ Q_v \\ U_v \\ V_v \end{bmatrix}. \quad (4.60)$$

The seven  $\gamma$  and  $\delta$  terms are defined as follows:

$$\begin{aligned} \gamma_{++} &= (\Delta g_{xm} + \Delta g_{ym}) + (\Delta g_{xn}^* + \Delta g_{yn}^*) \\ \gamma_{+-} &= (\Delta g_{xm} - \Delta g_{ym}) + (\Delta g_{xn}^* - \Delta g_{yn}^*) \\ \gamma_{--} &= (\Delta g_{xm} - \Delta g_{ym}) - (\Delta g_{xn}^* - \Delta g_{yn}^*) \\ \delta_{++} &= (D_{xm} + D_{ym}) + (D_{xn}^* + D_{yn}^*) \\ \delta_{+-} &= (D_{xm} - D_{ym}) + (D_{xn}^* - D_{yn}^*) \\ \delta_{-+} &= (D_{xm} + D_{ym}) - (D_{xn}^* + D_{yn}^*) \\ \delta_{--} &= (D_{xm} - D_{ym}) - (D_{xn}^* - D_{yn}^*). \end{aligned} \quad (4.61)$$

Here, it is assumed that Eqs. (4.36) are normalized so that the gain terms are close to unity, and the  $\Delta g$  terms are defined by  $g_{ik} = 1 + \Delta g_{ik}$ . The  $D$  (leakage) terms

and the  $\Delta g$  terms are often small enough that products of two such terms can be neglected. The results, as shown in Eqs. (4.60) and (4.61), apply to antennas that are linearly polarized in directions  $x$  and  $y$ . The same results apply to circularly polarized antennas if the subscripts  $x$  and  $y$  are replaced by  $r$  and  $\ell$ , respectively, and, in the column matrices on the left and right sides of Eq. (4.60), terms in  $Q_v$ ,  $U_v$ , and  $V_v$  are replaced by corresponding terms in  $V_v$ ,  $Q_v$ , and  $U_v$ , respectively. A similar result is given by Sault et al. (1991). The seven  $\gamma$  and  $\delta$  terms defined above are subject to errors in the calibration process, so there are seven degrees of freedom in the error mechanisms.

An observation of a single calibration source for which the four Stokes parameters are known enables four of the degrees of freedom to be determined. However, because of the relationships of the quantities involved, it takes at least three calibration observations to solve for all seven unknown parameters (Sault et al. 1996). In the calibration observations, it is useful to observe one unpolarized source, but observing a second unpolarized one would add no further solutions. At least one observation of a linearly polarized source is required to determine the relative phases of the two oppositely polarized channels, that is, the relative phases of the complex gain terms  $g_{xm}g_{yn}^*$  and  $g_{ym}g_{xn}^*$ , or  $g_{rm}g_{\ell n}^*$  and  $g_{\ell m}g_{rn}^*$ . Note that with antennas on altazimuth mounts, observations of a calibrator with linear polarization, taken at intervals between which large rotations of the parallactic angle occur, can essentially be regarded as observations of independent calibrators. Under these circumstances, three observations of the same calibrator will suffice for the full solution. Furthermore, the polarization of the calibrator need not be known in advance but can be determined from the observations.

In cases in which only an unpolarized calibrator can be observed, it may be possible to estimate two more degrees of freedom by introducing the constraint that the sum of the leakage factors over all antennas should be small. As shown by the expressions for the leakage terms in Appendix 4.2, this is a reasonable assumption for a homogeneous array, that is, one in which the antennas are of nominally identical design. However, the phase difference between the signal paths from the feeds to the correlator for the two orthogonal polarizations of each antenna remains unknown. This requires an observation of a calibrator with a component of linear polarization, or a scheme to measure the instrumental component of the phase. For example, on the compact array of the Australia Telescope (Frater and Brooks 1992), noise sources are provided at each antenna to inject a common signal into the two polarization channels (Sault et al. 1996). With such a system, it is necessary to provide an additional correlator for each antenna, or to be able to rearrange correlator inputs, to measure the relative phase of the injected signals in the two polarizations.

In the case of the approximations for weak polarization, Eqs. (4.38) and (4.43) show that if the gain terms are known, the leakage terms can be calibrated by observing an unpolarized source. For opposite circular polarizations, Eq. (4.43) shows that if  $V_v$  is small, it is possible to obtain solutions for the gain terms from the outputs for the  $\ell\ell$  and  $rr$  combinations only, provided also that the number of baselines is several times larger than the number of antennas. The leakage terms

can then be solved for separately. For crossed linear polarizations, Eq. (4.38) shows that this is possible only if the linear polarization ( $Q_v$  and  $U_v$  parameters) for the calibrator have been determined independently.

Optimum strategies for calibration of polarization observations is a subject that leads to highly detailed discussions involving the characteristics of particular synthesis arrays, the hour angle range of the observations, the availability of calibration sources (which can depend on the observing frequency), and other factors, especially if the solutions for strong polarization are used. Such discussions can be found, for example, in Conway and Kronberg (1969), Weiler (1973), Bignell (1982), Sault et al. (1991), Sault et al. (1996), and Smegal et al. (1997). Polarization measurements with VLBI involve some special considerations: see, for example, Roberts et al. (1991), Cotton (1993), Roberts et al. (1994), and Kembell et al. (1995).

For most large synthesis arrays, effective calibration techniques have been devised and the software to implement them has been developed. Thus, a prospective observer need not be discouraged if the necessary calibration procedures appear complicated. Some general considerations relevant to observations of polarization are given below.

- Since the polarization of many sources varies on a timescale of months, it is usually advisable to regard the polarization of the calibration source as one of the variables to be solved for.
- Two sources with relatively strong linear polarization at position angles that do not appear to vary are 3C286 and 3C138. These are useful for checking the phase difference for oppositely polarized channels.
- For most sources, the circular polarization parameter  $V_v$  is very small,  $\sim 0.2\%$  or less, and can be neglected. Measurements with circularly polarized antennas of the same sense therefore generally give an accurate measure of  $I_v$ . However, circular polarization is important in the measurement of magnetic fields by Zeeman splitting. As an example of positive detection at a very low level, Fiebig and Güsten (1989) describe measurements for which  $V/I \simeq 5 \times 10^{-5}$ . Zeeman splitting of several components of the OH line at 22.235 GHz was observed using a single antenna, the 100-m paraboloid of the Max Planck Institute for Radio Astronomy, with a receiving system that switched between opposite circular polarizations at 10 Hz. Rotation of the feed and receiver unit was used to identify spurious instrumental responses to linearly polarized radiation, and calibration of the relative pointing of the two beams to  $1''$  accuracy was required.
- Although the polarized emission from most sources is small compared with the total emission, it is possible for Stokes visibilities  $Q_v$  and  $U_v$  to be comparable to  $I_v$  in cases in which there is a broad unpolarized component that is highly resolved and a narrower polarized component that is not resolved. In such cases, errors may occur if the approximations for weak polarization [Eqs. (4.38) and (4.43)] are used in the data analysis.
- For most antennas, the instrumental polarization varies over the main beam and increases toward the beam edges. Sidelobes that are cross polarized relative to the main beam tend to peak near the beam edges. Thus, polarization measurements



are usually made for cases in which the source is small compared with the width of the main beam, and for such measurements, the beam should be centered on the source.

- Faraday rotation of the plane of polarization of incoming radiation occurs in the ionosphere and becomes important for frequencies below a few gigahertz; see Table 14.1. During polarization measurements, periodic observations of a strongly polarized source are useful for monitoring changes in the rotation, which varies with the total column density of electrons in the ionosphere. If not accounted for, Faraday rotation can cause errors in calibration; see, for example, Sakurai and Spangler (1994).
- In some antennas, the feed is displaced from the axis of the main reflector, for example, when the Cassegrain focus is used and the feeds for different bands are located in a circle around the vertex. For circularly polarized feeds, this departure from circular symmetry results in pointing offsets of the beams for the two opposite hands. The pointing directions of the two beams are typically separated by  $\sim 0.1$  beamwidths, which makes measurements of circular polarization difficult because  $V_v$  is proportional to  $(R_{rr} - R_{\ell\ell})$ . For linearly polarized feeds, the corresponding effect is an increase in the cross-polarized sidelobes near the beam edges.
- In VLBI, the large distances between antennas result in different parallactic angles at different sites, which must be taken into account.
- The quantities  $m_\ell$  and  $m_r$ , of Eqs. (4.20) and (4.22), have Rice distributions of the form of Eq. (6.63a), and the position angle has a distribution of the form of Eq. (6.63b). The percentage polarization can be overestimated, and a correction should be applied (Wardle and Kronberg 1974).

The following points concern choices in designing an array for polarization measurements.

- The rotation of an antenna on an altazimuth mount, relative to the sky, can sometimes be used to advantage in polarimetry. However, the rotation could be a disadvantage in cases in which polarization imaging over a large part of the antenna beam is being attempted. Correction for the variation of instrumental polarization over the beam may be more complicated if the beam rotates on the sky.
- With linearly polarized antennas, errors in calibration are likely to cause  $I_v$  to corrupt the linear parameters  $Q_v$  and  $U_v$ , so for measurements of linear polarization, circularly polarized antennas offer an advantage. Similarly, with circularly polarized antennas, calibration errors are likely to cause  $I_v$  to corrupt  $V_v$ , so for measurements of circular polarization, linearly polarized antennas may be preferred.
- Linearly polarized feeds for reflector antennas can be made with relative bandwidths of at least 2 : 1, whereas for circularly polarized feeds, the maximum relative bandwidth is commonly about 1.4 : 1. In many designs of circularly polarized feeds, orthogonal linear components of the field are combined with

$\pm 90^\circ$  relative phase shifts, and the phase-shifting element limits the bandwidth. For this reason, linear polarization is sometimes the choice for synthesis arrays [see, e.g., James (1992)], and with careful calibration, good polarization performance is obtainable.

- The stability of the instrumental polarization, which greatly facilitates accurate calibration over a wide range of hour angle, is perhaps the most important feature to be desired. Caution should therefore be used if feeds are rotated relative to the main reflector or if antennas are used near the high end of their frequency range.

## 4.8 The Interferometer Measurement Equation

The set of equations for the visibility values that would be measured for a given brightness distribution—taking account of all details of the locations and characteristics of the individual antennas, the path of the incoming radiation through the Earth's atmosphere including the ionosphere, the atmospheric transmission, etc.—is commonly referred to as the *measurement equation* or the *interferometer measurement equation*. For any specified brightness distribution and any system of antennas, the measurement equation provides accurate values of the visibility that would be observed. The reverse operation, i.e., the calculation of the optimum estimate of the brightness distribution from the measured visibility values, is more complicated. Taking the Fourier transform of the observed visibility function usually produces a brightness function with physically distorted features such as negative brightness values in some places. However, starting with a physically realistic model for the brightness, the measurement equation can accurately provide the corresponding visibility values that would be observed. This provides a basis for derivation of realistic brightness distributions that represent the observed visibilities, using an iterative procedure.

The formulation of the interferometer measurement equation is based on the analysis of Hamaker et al. (1996) and further developed by Rau et al. (2009), Smirnov (2011a,b,c,d), and others. It traces the variations of the signals from a source to the output of the receiving system. Direction-dependent effects include the direction of propagation of the signals, the primary beams of the antennas, polarization effects that vary with the alignment of the polarization of the source relative to that of the antennas, and also the effects of the ionosphere and troposphere. Direction-independent effects include the gains of the signal paths from the outputs of the antennas to the correlator. It is necessary to take account of all these various effects to calculate accurately the visibility values corresponding to the source model. Several of these effects are dependent upon the types of the interferometer antennas and the observing frequencies, so the details of the measurement equation are to some extent specific to each particular instrument to which it is applied.

The variations in the signal characteristics can generally be expressed as the effects of Faraday rotation, parallactic rotation, tilting of the wavefront by propagation effects, and variations in feed responses. These are linear effects on the signal and, as noted in Sect. 4.7.4, each of them can be represented by a  $2 \times 2$  (Jones) matrix. Their effect on the signal matrix is given by a series of outer products as explained with respect to Eq. (4.48). If the original signal is represented by the vector  $I$  and the series of effects along the signal path by Jones matrices  $\mathbf{J}_1$  to  $\mathbf{J}_n$  for antenna  $p$  and  $\mathbf{J}_1$  to  $\mathbf{J}_m$  for antenna  $q$ , then the voltage at the correlator output from the pair of antennas  $m$  and  $n$  is represented by

$$\mathcal{V} = \mathbf{J}_{pn}(\dots(\mathbf{J}_{p2}(\mathbf{J}_{p1}I\mathbf{J}_{q1}^H)\mathbf{J}_{q2}^H)\dots)\mathbf{J}_{qm}^H, \quad (4.62)$$

where the superscript  $H$  indicates the Hermitian (complex) conjugate. Each of the  $\mathbf{J}_p$  terms represents a  $2 \times 2$  (Jones) matrix. This analysis is from Smirnov (2011a,b,c,d). The combination of the various corrections into a single equation is helpful in ensuring that no significant effects have been overlooked.

An alternative formulation takes each product  $\mathbf{J}_{pn} \otimes \mathbf{J}_{pn}^H$ , which results in a  $4 \times 4$  (Mueller) matrix for each of the effects to be corrected along the signal path. If the resulting matrices are represented by  $[\mathbf{J}_{pn} \otimes \mathbf{J}_{pn}^H]$ , where  $n$  indicates the physical order in which the effects are encountered in the propagation path, then the correction for the effects is obtained as a series of products:

$$\mathcal{V} = [\mathbf{J}_{pn} \otimes \mathbf{J}_{pn}^H] \dots [\mathbf{J}_{p2} \otimes \mathbf{J}_{p2}^H][\mathbf{J}_{p1} \otimes \mathbf{J}_{p1}^H]SI, \quad (4.63)$$

where  $S$  is a Fourier transform matrix that converts the Stokes visibility to brightness. Each of the  $\mathbf{J}_p \otimes \mathbf{J}_p^H$  terms represents a  $4 \times 4$  matrix. This is basically the form used by Rau et al. (2009). The details of the interferometer equation will vary for different instruments, depending upon which factors need to be included. Here, the intention is to give a general outline of how the calibration factors can be applied. Further details can be found in papers by Hamaker et al. (1996), Hamaker (2000), Rau et al. (2009), and Smirnov (2011a,b,c,d).

### 4.8.1 Multibaseline Formulation

In this chapter thus far, we have mainly considered the response of a single pair of antennas. The data gathered from a multielement array can conveniently be expressed in the form of a covariance matrix. The discussion here largely follows Leshem et al. (2000) and Boonstra and van der Veen (2003). We start from the expression for the two-element interferometer response and, for simplicity, consider the small-angle case in which the  $w$  component can be omitted, as in Eq. (3.9),

$$\mathcal{V}(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{A_N(l, m)I(l, m)}{\sqrt{1 - l^2 - m^2}} e^{-j2\pi(ul+vm)} dl dm. \quad (4.64)$$

Here,  $\mathcal{V}$  is the complex visibility, and  $u$  and  $v$  represent the projected baseline coordinates measured in wavelengths in a plane normal to the phase reference direction. We make four adjustments to the equation. (1) We assume that both the astronomical brightness function and the visibility function can each be represented by a point-source model with a number of points  $p$ . For a point  $k$ , the direction is specified by direction cosines  $(l_k, m_k)$ . We replace the integrals in Eq. (4.64) with summations over the points. (2) We replace  $A_N$  by the product of the corresponding complex voltage gain factors  $g_i(l, m)g_j^*(l, m)$ , where  $i$  and  $j$  indicate antennas. Constants representing conversion of aperture to gain, etc., can be ignored since, in practice, the intensity scale is determined by calibration. (3) We allow the factor  $\sqrt{1 - l^2 - m^2}$  to be subsumed within the intensity function  $I(l, m)$ . (4) For each antenna, we specify the components in the  $(u, v)$  plane relative to a reference point that can be chosen, for example, to be the center of the array. The  $(u, v)$  values for a pair of antennas  $i$  and  $j$  then become  $(u_i - u_j, v_i - v_j)$ . The second and fourth modifications allow the parameters involved to be specified in terms of individual antennas rather than antenna pairs. Equation (4.64) can now be written as:

$$\mathcal{V}(u_i - u_j, v_i - v_j) = \sum_{k=1}^p I_k g_i(l_k, m_k) e^{-j2\pi(u_i l_k + v_i m_k)} g_j^*(l_k, m_k) e^{j2\pi(u_j l_k + v_j m_k)}, \quad (4.65)$$

where  $I_k = I(l_k, m_k)$ . Note that  $u$  and  $v$  do not vary with the source positions within the field of view but are defined for the phase reference position (field center). Equations (4.64) and (4.65) represent the visibility as measured by a single pair of antennas.

It is useful to put Eq. (4.65) in matrix form. For an array of  $n$  antennas, we define an  $n \times p$  matrix containing terms corresponding to the first antenna gain and exponential terms of Eq. (4.65) (i.e., the terms associated with antenna  $i$ ):

$$\mathbf{A} = \begin{bmatrix} g_1(l_1, m_1)e^{-j2\pi(u_1 l_1 + v_1 m_1)} & g_1(l_2, m_2)e^{-j2\pi(u_1 l_2 + v_1 m_2)} & \dots & g_1(l_p, m_p)e^{-j2\pi(u_1 l_p + v_1 m_p)} \\ g_2(l_1, m_1)e^{-j2\pi(u_2 l_1 + v_2 m_1)} & \dots & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots \\ g_n(l_1, m_1)e^{-j2\pi(u_n l_1 + v_n m_1)} & \dots & \dots & g_n(l_p, m_p)e^{-j2\pi(u_n l_p + v_n m_p)} \end{bmatrix}. \quad (4.66)$$

The antenna index increases downward across the  $n$  rows, and the point-source index increases toward the right across the  $p$  columns.

To generate the covariance matrix, we first define a  $p \times p$  diagonal matrix containing the intensity values of the  $p$  source-model points:

$$\mathbf{B} = \begin{bmatrix} I_1 & & & \\ & I_2 & & \\ & & \ddots & \\ & & & I_p \end{bmatrix}. \quad (4.67)$$

Then we can write

$$\mathbf{R} = \mathbf{A}\mathbf{B}\mathbf{A}^H, \quad (4.68)$$

where the superscript  $H$  indicates the Hermitian transpose (transposition of the matrix plus complex conjugation).  $\mathbf{R}$  is the covariance matrix, which is Hermitian with dimensions  $n \times n$ . Each element of  $\mathbf{R}$  is of the form of the right side of Eq. (4.65), that is, the sum of responses to the  $p$  intensity points for a specific pair of antennas. For row  $i$  and column  $j$ , the element is  $r_{i,j}$ , which is equal to the right side of Eq. (4.65). The elements  $r_{i,j}$  represent the cross-correlation of signals from antennas  $i$  and  $j$ . When the gain factors  $g$  are equal to unity, the elements represent the source visibility  $\mathcal{V}$ . The diagonal elements are the  $n$  self-products ( $i = j$ ), which represent the total power responses of the antennas. Note that  $\mathbf{R}$  is Hermitian:  $r_{i,j} = r_{j,i}^*$ .  $\mathbf{R}$  contains the full set of correlator output terms for an array of  $n$  antennas for a single averaging period and a single frequency channel. These data, when calibrated as visibility, can provide a snapshot image. In cases in which the  $w$  component is important, a term of the form  $w(\sqrt{1 - l^2 - m^2} - 1)$  [as in Eq. (3.7)] with appropriate subscripts, can be included within each exponent. If the response patterns of the antennas are identical, i.e.,  $g_i = g_j$  for all  $(i, j)$ , then  $g_i g_j^* = |g|^2$ , and this (real) gain factor can be taken outside the matrix  $\mathbf{R}$ . Thus, to determine the angle of incidence ( $l, m$ ) of a signal from the covariance measurements [the  $(u, v)$  values being known], the gain factors need not be known if they are identical from one antenna to another but otherwise must be known.

The covariance matrix can also be formulated in terms of the complex signal voltages from the antennas of an array. Let the signal from antenna  $k$  be  $x_k$ , which is a function of time. For the array, the signals can be represented by a (column) vector  $\mathbf{x}$  of dimensions  $n \times 1$ , each term of which corresponds to the sum of the terms in the corresponding row of the matrix in Eq. (4.66). The outer (or Kronecker) product  $\mathbf{x} \otimes \mathbf{x}^H$  leads to a covariance matrix:

$$\mathbf{R}' = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \otimes [x_1^* \ x_2^* \ \dots \ x_n^*] = \begin{bmatrix} x_1 x_1^* & x_1 x_2^* & \dots & x_1 x_n^* \\ x_2 x_1^* & \dots & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots \\ x_n x_1^* & \dots & \dots & x_n x_n^* \end{bmatrix}. \quad (4.69)$$

The elements  $r_{ij}$  of the matrix  $\mathbf{R}$  represent the correlator outputs, which involve a time average of the signal products. If the signal products in the elements of  $\mathbf{R}'$  are similarly understood to represent time-averaged products, then  $\mathbf{R}'$  is equivalent to the covariance matrix  $\mathbf{R}$ .

An example of the application of matrix formulation in radio astronomy is provided by the discussion of gain calibration by Boonstra and van der Veen (2003). Also, the eigenvectors of the matrix can be used to identify interfering signals that are strong enough to be distinguished in the presence of the noise. Such signals can then be removed from the data, as discussed, for example, by Leshem et al. (2000).

## Appendix 4.1 Hour Angle–Declination and Elevation–Azimuth Relationships

Although the positions of cosmic sources are almost always specified in celestial coordinates, for purposes of observation, it is generally necessary to convert to elevation and azimuth. The conversion formulas between hour angle and declination ( $H, \delta$ ) and elevation and azimuth ( $\mathcal{E}, \mathcal{A}$ ) can be derived by applying the sine and cosine rules for spherical triangles to the system in Fig. 4.3. For an observer at latitude  $\mathcal{L}$ , they are, for  $(H, \delta)$  to  $(\mathcal{A}, \mathcal{E})$ ,

$$\begin{aligned}\sin \mathcal{E} &= \sin \mathcal{L} \sin \delta + \cos \mathcal{L} \cos \delta \cos H \\ \cos \mathcal{E} \cos \mathcal{A} &= \cos \mathcal{L} \sin \delta - \sin \mathcal{L} \cos \delta \cos H \\ \cos \mathcal{E} \sin \mathcal{A} &= -\cos \delta \sin H ,\end{aligned}\tag{A4.1}$$

Similarly, for  $(\mathcal{A}, \mathcal{E})$  to  $(H, \delta)$ ,

$$\begin{aligned}\sin \delta &= \sin \mathcal{L} \sin \mathcal{E} + \cos \mathcal{L} \cos \mathcal{E} \cos \mathcal{A} \\ \cos \delta \cos H &= \cos \mathcal{L} \sin \mathcal{E} - \sin \mathcal{L} \cos \mathcal{E} \cos \mathcal{A} \\ \cos \delta \sin H &= -\cos \mathcal{E} \sin \mathcal{A} .\end{aligned}\tag{A4.2}$$

Here, azimuth is measured from north through east.

## Appendix 4.2 Leakage Parameters in Terms of the Polarization Ellipse

The polarization leakage terms used to express the instrumental polarization are related to the ellipticity and orientation of the polarization ellipses of each antenna, as shown below.

### A4.2.1 Linear Polarization

Consider the antenna in Fig. 4.8, and suppose that it is nominally linearly polarized in the  $x$  direction, in which case  $\chi$  and  $\psi$  are small angles that represent engineering tolerances. A field  $E$  aligned with the  $x$  axis in Fig. 4.8a produces components  $E_{x'}$  and  $E_{y'}$  along the  $(x', y')$  axes with which the dipoles in Fig. 4.8b are aligned. Then from Eq. (4.26), we obtain the voltage at the output of the antenna (point  $A$  in Fig. 4.8b), which is

$$V'_x = E(\cos \psi \cos \chi + j \sin \psi \sin \chi) . \quad (\text{A4.3})$$

The response to the same field, but aligned with the  $y$  axis, is

$$V'_y = E(\sin \psi \cos \chi - j \cos \psi \sin \chi) . \quad (\text{A4.4})$$

$V'_x$  represents the wanted response to the field along the  $x$  axis, and  $V'_y$  represents the unwanted response to a cross-polarized field. The leakage term is equal to the cross-polarized response expressed as a fraction of the wanted  $x$ -polarization response, that is,

$$D_x = \frac{V'_y}{V'_x} = \frac{(\sin \psi_x \cos \chi_x - j \cos \psi_x \sin \chi_x)}{(\cos \psi_x \cos \chi_x + j \sin \psi_x \sin \chi_x)} \simeq \psi_x - j\chi_x , \quad (\text{A4.5})$$

where the subscript  $x$  indicates the  $x$ -polarization case. The corresponding term  $D_y$ , for the condition in which Fig. 4.8 represents the nominal  $y$  polarization of the antenna, is obtained as  $V'_x/V'_y$  by inverting Eq. (A4.5), replacing  $\psi_x$  by  $\psi_y + \pi/2$ , and replacing  $\chi_x$  by  $\chi_y$ . Then  $\psi_y$  is measured from the  $y$  axis in the same sense as  $\psi_x$  is measured from the  $x$  axis, that is, increasing in a counterclockwise direction in Fig. 4.8. Thus, we obtain

$$\begin{aligned} D_y &= \frac{V'_x}{V'_y} = \frac{[\cos(\psi_y + \pi/2) \cos \chi_y + j \sin(\psi_y + \pi/2) \sin \chi_y]}{[\sin(\psi_y + \pi/2) \cos \chi_y - j \cos(\psi_y + \pi/2) \sin \chi_y]} \\ &= \frac{(-\sin \psi_y \cos \chi_y + j \cos \psi_y \sin \chi_y)}{(\cos \psi_y \cos \chi_y + j \sin \psi_y \sin \chi_y)} \simeq -\psi_y + j\chi_y . \end{aligned} \quad (\text{A4.6})$$

Similar expressions for  $D_x$  and  $D_y$  have also been derived by Sault et al. (1991). Note that  $D_x$  and  $D_y$  are of comparable magnitude and opposite sign, so one would expect the average of all the  $D$  terms for an array of antennas to be very small. As used earlier in this chapter, subscripts  $m$  and  $n$  are added to the  $D$  terms to indicate individual antennas.

### A4.2.2 Circular Polarization

To receive right circular polarization from the sky, the antenna in Fig. 4.8b must respond to a field with counterclockwise rotation in the plane of the diagram, as explained earlier. This requires  $\chi = -45^\circ$ . In terms of fields in the  $x$  and  $y$  directions, counterclockwise rotation requires that  $E_x$  leads  $E_y$  in phase by  $\pi/2$ ; that is,  $E_x = jE_y$  for the fields as defined in Eq. (4.25). For fields  $E_x$  and  $E_y$ , we determine the components in the  $x'$  and  $y'$  directions and then obtain expressions for the output of the antenna for both counterclockwise and clockwise rotation of the incident field. For counterclockwise rotation:

$$E'_x = E_x \cos \psi + E_y \sin \psi = E_x(\cos \psi - j \sin \psi) , \quad (\text{A4.7})$$

$$E'_y = -E_x \sin \psi + E_y \cos \psi = -E_x(\sin \psi + j \cos \psi) . \quad (\text{A4.8})$$

For nominal right-circular polarization,  $\chi_r = -\pi/4 + \Delta\chi_r$ , where  $\Delta\chi_r$  is a measure of the departure of the polarization from circularity. Then from Eq. (4.26), we obtain

$$V'_r = E_x e^{-j\psi_r} (\cos \chi_r - \sin \chi_r) = \sqrt{2} E_x e^{-j\psi_r} \cos \Delta\chi_r . \quad (\text{A4.9})$$

The next step is to repeat the procedure for left circular polarization from the sky, for which we have clockwise rotation of the electric vector and  $E_y = jE_x$ . The result is

$$V'_\ell = E_x e^{j\psi_r} (\cos \chi_r + \sin \chi_r) = \sqrt{2} E_x e^{j\psi_r} \sin \Delta\chi_r . \quad (\text{A4.10})$$

The relative magnitude of the opposite-hand response of the nominally right-handed polarization state, that is, the leakage term, is

$$D_r = \frac{V'_\ell}{V'_r} = e^{j2\psi_r} \tan \Delta\chi_r \simeq e^{j2\psi_r} \Delta\chi_r . \quad (\text{A4.11})$$

For nominal left-handed polarization, the relative magnitude of the opposite-hand response is obtained by inverting the right side of Eq. (A4.11) and also substituting  $\Delta\chi_\ell + \pi/2$  for  $\Delta\chi_r$  and  $\psi_\ell - \pi/2$  for  $\psi_r$ . For the corresponding leakage term  $D_\ell$ , which represents the right circular leakage of the nominally left circularly polarized antenna, we then obtain

$$D_\ell = e^{-j2\psi_\ell} \tan \Delta\chi_\ell \simeq e^{-j2\psi_\ell} \Delta\chi_\ell . \quad (\text{A4.12})$$

Since  $-\pi/4 \leq \chi \leq \pi/4$ ,  $\Delta\chi_r$  and  $\Delta\chi_\ell$  take opposite signs. Thus, as in the case of the leakage terms for linear polarization,  $D_r$  and  $D_\ell$  are of comparable magnitude and opposite sign.



**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



## References

- Bignell, R.C., Polarization, in *Synthesis Mapping, Proceedings of NRAO Workshop No. 5*, Socorro, NM, June 21–25, 1982, Thompson, A.R., and D'Addario, L.R., Eds., National Radio Astronomy Observatory, Green Bank, WV (1982)
- Boonstra, A.-J., and van der Veen, A.-J., Gain Calibration Methods for Radio Telescope Arrays, *IEEE Trans. Signal Proc.*, **51**, 25–38 (2003)
- Born, M., and Wolf, E., *Principles of Optics*, 7th ed., Cambridge Univ. Press, Cambridge, UK (1999)
- Conway, R.G., and Kronberg, P.P., Interferometer Measurement of Polarization Distribution in Radio Sources, *Mon. Not. R. Astron. Soc.*, **142**, 11–32 (1969)
- Cotton, W.D., Calibration and Imaging of Polarization Sensitive Very Long Baseline Interferometer Observations, *Astron. J.*, **106**, 1241–1248 (1993)
- DeBoer, D.R., Gough, R.G., Bunton, J.D., Cornwell, T.J., Beresford, R.J., Johnston, S., Feain, I.J., Schinckel, A.E., Jackson, C.A., Kesteven, M.J., and nine coauthors, Australian SKA Pathfinder: A High-Dynamic Range Wide-Field of View Survey Telescope, *Proc. IEEE*, **97**, 1507–1521 (2009)
- de Vos, M., Gunst, A.W., and Nijboer, R., The LOFAR Telescope: System Architecture and Signal Processing, *Proc. IEEE*, **97**, 1431–1437 (2009)
- Fiebig, D., and Güsten, R., Strong Magnetic Fields in Interstellar Maser Clumps, *Astron. Astrophys.*, **214**, 333–338 (1989)
- Fomalont, E.B., and Perley, R.A., Calibration and Editing, in *Synthesis Imaging in Radio Astronomy*, Perley, R.A., Schwab, F.R., and Bridle, A.H., Eds., Astron. Soc. Pacific Conf. Ser., **6**, 83–115 (1989)
- Frater, R.H., and Brooks, J.W., Eds., *J. Electric. Electron. Eng. Aust.*, Special Issue on the Australia Telescope, **12**, No. 2 (1992)
- Hamaker, J.P., A New Theory of Radio Polarimetry with an Application to the Westerbork Synthesis Radio Telescope (WSRT), in *Workshop on Large Antennas in Radio Astronomy*, ESTEC, Noordwijk, the Netherlands (1996)
- Hamaker, J.P., Understanding Radio Polarimetry. IV. The Full-Coherency Analogue of Scalar Self-Calibration: Self-Alignment, Dynamic Range, and Polarimetric Fidelity, *Astron. Astrophys. Suppl.*, **143**, 515–534 (2000)
- Hamaker, J.P., Understanding Radio Polarimetry. V. Making Matrix Self-Calibration Work: Processing of a Simulated Observation, *Astron. Astrophys.*, **456**, 395–404 (2006)
- Hamaker, J.P., Bregman, J.D., and Sault, R.J., Understanding Radio Polarimetry. I. Mathematical Foundations, *Astron. Astrophys. Suppl.*, **117**, 137–147 (1996)

- IAU, *Trans. Int. Astron. Union*, Proc. of the 15th General Assembly Sydney 1973 and Extraordinary General Assembly Poland 1973, G. Contopoulos and A. Jappel, Eds., **15B**, Reidel, Dordrecht, the Netherlands (1974), p. 166
- IEEE, Standard Definitions of Terms for Radio Wave Propagation, Std. 211–1977, Institute of Electrical and Electronics Engineers Inc., New York (1977)
- James, G.L., The Feed System, *J. Electric. Electron. Eng. Aust.*, Special Issue on the Australia Telescope, **12**, No. 2, 137–145 (1992)
- Jones, R.C., A New Calculus for the Treatment of Optical Systems. I. Description and Discussion of the Calculus, *J. Opt. Soc. Am.*, **31**, 488–493 (1941)
- Kemball, A.J., Diamond, P.J., and Cotton, W.D., Data Reduction Techniques for Spectral Line Polarization VLBI Observations, *Astron. Astrophys. Suppl.*, **110**, 383–394 (1995)
- Ko, H.C., Coherence Theory of Radio-Astronomical Measurements, *IEEE Trans. Antennas Propag.*, **AP-15**, 10–20 (1967a)
- Ko, H.C., Theory of Tensor Aperture Synthesis, *IEEE Trans. Antennas Propag.*, **AP-15**, 188–190 (1967b)
- Kraus, J.D., and Carver, K.R., *Electromagnetics*, 2nd ed., McGraw-Hill, New York (1973) p. 435
- Leshem, A., van der Veen, A.-J., and Boonstra, A.-J., Multichannel Interference Mitigation Techniques in Radio Astronomy, *Astrophys. J. Suppl.*, **131**, 355–373 (2000)
- Morris, D., Radhakrishnan, V., and Seielstad, G.A., On the Measurement of Polarization Distributions Over Radio Sources, *Astrophys. J.*, **139**, 551–559 (1964)
- O’Neill, E.L., *Introduction to Statistical Optics*, Addison-Wesley, Reading, MA (1963)
- Rau, U., Bhatnagar, S., Voronkov, M.A., and Cornwell, T.J., Advances in Calibration and Imaging Techniques in Radio Interferometry, *Proc. IEEE*, **97**, 1472–1481 (2009)
- Roberts, D.H., Brown, L.F., and Wardle, J.F.C., Linear Polarization Sensitive VLBI, in *Radio Interferometry: Theory, Techniques, and Applications*, Cornwell, T.J., and Perley, R.A., Eds., Astron. Soc. Pacific Conf. Ser., **19**, 281–288 (1991)
- Roberts, D.H., Wardle, J.F.C., and Brown, L.F., Linear Polarization Radio Imaging at Milliarcsecond Resolution, *Astrophys. J.*, **427**, 718–744 (1994)
- Rowson, B., High Resolution Observations with a Tracking Radio Interferometer, *Mon. Not. R. Astron. Soc.*, **125**, 177–188 (1963)
- Sakurai, T., and Spangler, S.R., Use of the Very Large Array for Measurement of Time Variable Faraday Rotation, *Radio Sci.*, **29**, 635–662 (1994)
- Sault, R.J., Hamaker, J.P., and Bregman, J.D., Understanding Radio Polarimetry. II. Instrumental Calibration of an Interferometer Array, *Astron. Astrophys. Suppl.*, **117**, 149–159 (1996)
- Sault, R.J., Killeen, N.E.B., and Kesteven, M.J., *AT Polarization Calibration*, Aust. Tel. Tech. Doc. Ser. 39.3/015, CSIRO, Epping, New South Wales (1991)
- Seidelmann, P.K., Ed., *Explanatory Supplement to the Astronomical Almanac*, Univ. Science Books, Mill Valley, CA (1992)
- Smegal, R.J., Landecker, T.L., Vaneldik, J.F., Routledge, D., and Dewdney, P.E., Aperture Synthesis Polarimetry: Application to the Dominion Astrophysical Observatory Synthesis Telescope, *Radio Sci.*, **32**, 643–656 (1997)
- Smirnov, O.M., Revisiting the Radio Interferometer Measurement Equation. 1. A Full-Sky Jones Formalism, *Astron. Astrophys.*, **527**, A106 (11pp) (2011a)
- Smirnov, O.M., Revisiting the Radio Interferometer Measurement Equation. 2. Calibration and Direction-Dependent Effects, *Astron. Astrophys.*, **527**, A107 (10pp) (2011b)
- Smirnov, O.M., Revisiting the Radio Interferometer Measurement Equation. 3. Addressing Direction-Dependent Effects in 21-cm WSRT Observations of 3C147, *Astron. Astrophys.*, **527**, A108 (12pp) (2011c)
- Smirnov, O.M., Revisiting the Radio Interferometer Measurement Equation. 4. A Generalized Tensor Formalism, *Astron. Astrophys.*, **531**, A159 (16pp) (2011d)
- Wade, C.M., Precise Positions of Radio Sources. I. Radio Measurements, *Astrophys. J.*, **162**, 381–390 (1970)
- Wardle, J.F.C., and Kronberg, P.P., The Linear Polarization of Quasi-Stellar Radio Sources at 3.71 and 11.1 Centimeters, *Astrophys. J.*, **194**, 249–255 (1974)

- Weiler, K.W., The Synthesis Radio Telescope at Westerbork: Methods of Polarization Measurement, *Astron. Astrophys.*, **26**, 403–407 (1973)
- Weiler, K.W., and Raimond, E., Aperture Synthesis Observations of Circular Polarization, *Astron. Astrophys.*, **52**, 397–402 (1976)
- Wilson, T.L., Rohlfs, K., and Hüttemeister, S., *Tools of Radio Astronomy*, 6th ed., Springer-Verlag, Berlin (2013)