Chapter 15

Polarimetry

Jayaram N. Chengalur

15.1 Introduction

Consider the simplest kind of electromagnetic wave, i.e. a plane monochromatic wave of frequency ν propagating along the +Z axis of a cartesian co-ordinate system. Since electro-magnetic waves are transverse, the electric field **E** must lie in the X-Y plane. Further since the wave is mono-chromatic one can write

$$\mathbf{E}(t) = E_x \cos(2\pi\nu t)\mathbf{e}_x + E_y \cos(2\pi\nu t + \delta)\mathbf{e}_y, \tag{15.1.1}$$

i.e. the X and Y components of the electric field differ in phase by a factor which does not depend on time. It can be shown[1] that the implication of this is that over the course of one period of oscillation, the tip of the electric field vector in general traces out an ellipse. There are two special cases of interest. The first is when $\delta = 0$. In this case the tip of the electric field vector traces out a line segment, and the wave is said to be *linearly polarized*. The other special case is when $E_x = E_y$ and $\delta = \pm\pi/2$. In this case the electric field vector traces out a circle in the X-Y plane, and depending on the sense[2] in which this circle is traversed the wave is called either *left circular polarized* or *right circular polarized*.

As you have already seen in chapter 1, signals in radio astronomy are not monochromatic waves, but are better described as quasi-monochromatic plane waves[3]. Further, the quantity that is typically measured in radio astronomy is not related to the field (i.e. a voltage), but rather a quantity that has units of voltage squared, i.e. related to some correlation function of the field (see chapter 4). For these reasons, it is usual to characterize the polarization properties of the incoming radio signals using quantities called *Stokes* parameters. Recall that for a quasi monochromatic wave, the electric field **E** could be considered to be the real part of a complex analytical signal $\mathcal{E}(t)$. If the X and Y components of this complex analytical signal are $\mathcal{E}_x(t)$, and $\mathcal{E}_y(t)$, respectively, then the four

[1] See for example, Born & Wolf 'Principles of Optics', Sixth Edition, Section 1.4.2

[2] Note that there is an additional ambiguity here, i.e. are you looking along the direction of propagation of the wave, or against it? To keep things interesting neither convention is universally accepted, although in principle one should follow the convention adopted by the IAU (Transactions of the IAU Vol. 15B, (1973), 166.)

[3] Recall that as all astrophysically interesting sources are distant, the plane wave approximation is a good one

Stokes parameters are defined as:

$$
\begin{array}{rcl}
I &=& <\mathcal{E}_x\mathcal{E}_x^*> + <\mathcal{E}_y\mathcal{E}_y^*> \\
Q &=& <\mathcal{E}_x\mathcal{E}_x^*> - <\mathcal{E}_y\mathcal{E}_y^*> \\
U &=& <\mathcal{E}_x\mathcal{E}_y^*> + <\mathcal{E}_x^*\mathcal{E}_y> \\
V &=& \frac{1}{i}(<\mathcal{E}_x\mathcal{E}_y^*> - <\mathcal{E}_x^*\mathcal{E}_y>)
\end{array}
\quad \text{or} \quad
\begin{array}{rcl}
<\mathcal{E}_x\mathcal{E}_x^*> &=& (I+Q)/2 \\
<\mathcal{E}_y\mathcal{E}_y^*> &=& (I-Q)/2 \\
<\mathcal{E}_x\mathcal{E}_y^*> &=& (U+iV)/2 \\
<\mathcal{E}_x^*\mathcal{E}_y> &=& (U-iV)/2.
\end{array}
\tag{15.1.2}
$$

where the angle brackets indicate taking the average value[4]. The Stokes parameters as defined in equation (15.1.2) clearly depend on the orientation of the co-ordinate system. In radio astronomy it is conventional (see chapter 10) to take the +X axis to point north and the +Y axis to point east. It is important to realize that the Stokes parameters are descriptors of the intrinsic polarization state of the electro-magnetic wave, i.e. the Stokes vector $(I\ Q\ U\ V)^T$ is a true vector. The equations (15.1.2) simply give its components in a particular co-ordinate system, the linear polarization co-ordinate system[5]. One would instead work in a circularly polarized reference frame, i.e. where the electric field is decomposed into two circularly polarized components, $\mathcal{E}_r(t)$, and $\mathcal{E}_l(t)$. The relation between these components and the Stokes parameters are:

$$
\begin{array}{rcl}
I &=& <\mathcal{E}_r\mathcal{E}_r^*> + <\mathcal{E}_l\mathcal{E}_l^*> \\
Q &=& <\mathcal{E}_r\mathcal{E}_l^*> + <\mathcal{E}_r^*\mathcal{E}_l> \\
U &=& \frac{1}{i}(<\mathcal{E}_r\mathcal{E}_l^*> - <\mathcal{E}_r^*\mathcal{E}_l>) \\
V &=& <\mathcal{E}_r\mathcal{E}_r^*> - <\mathcal{E}_l\mathcal{E}_l^*>
\end{array}
\qquad
\begin{array}{rcl}
<\mathcal{E}_r\mathcal{E}_r^*> &=& (I+V)/2 \\
<\mathcal{E}_l\mathcal{E}_l^*> &=& (I-V)/2 \\
<\mathcal{E}_r\mathcal{E}_l^*> &=& (Q+iU)/2 \\
<\mathcal{E}_r^*\mathcal{E}_l> &=& (Q-iU)/2.
\end{array}
\tag{15.1.3}
$$

Interestingly, equations (15.1.3) are formally identical to equations (15.1.2) apart from the following transformations viz. $Q^+ \to V^\odot$, $U^+ \to Q^\odot$, $V^+ \to U^\odot$, where the superscript $+$ indicates linear polarized co-ordinates and $\odot$ circular polarized co-ordinates. Although these two co-ordinate systems are the ones most frequently used, the Stokes vector could in principle be written in any co-ordinate system based on two linearly independent (but not necessarily orthogonal) polarization states. In fact, as we shall see, such non orthogonal co-ordinate systems will arise naturally when trying to describe measurements made with non ideal radio telescopes.

The degree of polarization of the wave is defined as

$$P = \frac{\sqrt{Q^2+U^2+V^2}}{I}. \tag{15.1.4}$$

From equation (15.1.2) we have

$$I^2 - Q^2 - U^2 - V^2 = 2\left(\left\langle\mathcal{E}_x^2\right\rangle\left\langle\mathcal{E}_y^2\right\rangle - \left\langle\mathcal{E}_x\mathcal{E}_y\right\rangle^2\right) \tag{15.1.5}$$

and hence from the Schwarz inequality it follows that $0 \le P \le 1$ and that $P = 1$ iff $\mathcal{E}_x = c\mathcal{E}_y$, where c is some complex constant. For a mono-chromatic plane wave (equation (15.1.1)) therefore, $P = 1$ or equivalently $I^2 = Q^2 + U^2 + V^2$, i.e. there are only three independent Stokes parameters. For a general quasi mono-chromatic wave, $P < 1$, and the wave is said to be *partially polarized*.

It is also instructive to examine the Stokes parameters separately for the special case of a monochromatic plane wave. We have (see equations (15.1.1) and (15.1.2)):

$$
\begin{array}{rclcrcl}
I &=& E_x^2 + E_y^2 & \qquad & U &=& 2E_xE_y\cos(\delta) \\
Q &=& E_x^2 - E_y^2 & \qquad & V &=& 2E_xE_y\sin(\delta),
\end{array}
$$

[4]Strictly speaking this is the ensemble average. However, as always, we will assume that the signals are ergodic, i.e. the ensemble average can be replaced with the time average.

[5]These polaraization co-ordinate systems are of course in some abstract polarization space and not real space

i.e. for a linearly polarized wave ($\delta = 0$) we have V = 0, and for a circularly polarized wave ($E_x = E_y, \delta = \pm\pi/2$) we have $Q = U = 0$. So Q and U measure linear polarization, and V measures circular polarization. This intepretation continues to be true in the case of partially polarized waves.

15.2 Polarization in Radio Astronomy

Emission mechanisms which are dominant in low frequency radio astronomy, produce linearly polarized emission. Thus extra-galactic radio sources and pulsars are predominantly linearly polarized, with polarization fractions of typically a few percent. These sources usually have no circular polarization, i.e. $V \sim 0$. Maser sources however, in particular OH masers from galactic star forming regions often have significant circular polarization. This is believed to arise because of Zeeman splitting. Interstellar maser sources also often have some linear polarization, i.e. all the components of the Stokes vector are non zero. In radio astronomy the polarization is fundamentally related to the presence of magnetic fields, and polarization studies of sources are aimed at understanding their magnetic fields.

The raw polarization measured by a radio telescope could differ from the true polarization of the source because of a number of effects, some due to propagation of the wave through the medium between the source and the telescope, (see chapter 16) and the other because of various instrumental non-idealities. Since we are eventually interested in the true source polarization our ultimate aim will be to correct for these various effects, and we will therefore find it important to distinguish between depolarizing and non-depolarizing systems. A system for which the outgoing wave is fully polarized if the incoming wave is fully polarized is called non-depolarizing. The polarization state of the output wave need not be identical to that of the incoming wave, it is only necessary that $P_{out} = 1$ if $P_{in} = 1$.

The most important propagation effect is *Faraday rotation*, which is covered in some detail in chapter 16. Here we restrict ourselves to stating that the plane of polarization of a linearly polarized wave is rotated on passing through a magnetized plasma. Faraday rotation can occur both in the ISM as well as in the earth's ionosphere. If the Faraday rotating medium is mixed up with the emitting region, then radiation emitted from different depths along the line of sight are rotated by different amounts, thus reducing the net polarization. This is called *Faraday depolarization*. If the medium is located between the source and the observer, then the only effect is a net rotation of the plane of polarization, i.e.

$$\mathcal{E}_x^{'} = \mathcal{E}_x \cos\chi + \mathcal{E}_y \sin\chi, \qquad \mathcal{E}_y^{'} = -\mathcal{E}_x \sin\chi + \mathcal{E}_y \cos\chi, \tag{15.2.6}$$

where $\mathcal{E}_x$, $\mathcal{E}_x^{'}$ are the X components of the incident and emergent field respectively and similarly for $\mathcal{E}_y$, $\mathcal{E}_y^{'}$. In terms of the Stokes paramters, the transformation on passing through a Faraday rotating medium is

$$\begin{array}{lllllll} I^{'} & = & I & \qquad & Q^{'} & = & Q\cos 2\chi + U \sin 2\chi \\ V^{'} & = & V & \qquad & U^{'} & = & -Q\sin 2\chi \ + U\cos 2\chi. \end{array} \tag{15.2.7}$$

i.e. a rotation of the Stokes vector in the (U,V) plane. The fractional polarization is hence preserved[6]. Equation (15.2.7) can also be easily obtained from equation (15.1.3)

[6]Note that non-depolarizing only means that $P_{out} = 1$ if $P_{in} = 1$, and this does not necessarily translate into conservation of the fractional polarization when $P < 1$. Pure faraday rotation is hence not only non-depolarizing, it also preserves the fractional polarization.

by noting that in a circularly polarized co-ordinate system, the effect of faraday rotation is to introduce a phase difference of 2χ between $\mathcal{E}_r$ and $\mathcal{E}_l$.

Consider looking at an extended source which is not uniformly polarized with a radio telescope whose resolution is poorer than the angular scale over which the source polarization is coherent. In any given resolution element then there are regions with different polarization characteristics. The beam thus smoothes out the polarization of the source, and the measured polarization will be less than the true source polarization. This is called *beam depolarization*. Beam depolarization cannot in principle be corrected for, the only way to obtain the true source polarization is to observe with sufficiently high angular resolution.

A dual polarized radio telescope has two voltage beam patterns, one for each polarization. These two patterns are often not symmetrical, i.e. in certain directions the telescope response is greater for one polarization than for the other. The difference in gain between these two polarizations usually varies in a systematic way over the primary beam. Because of this asymmetry, an unpolarized source could appear to be polarized, and further its apparent Stokes parameters in general depend on its location with respect to the center of the primary beam. The polarization properties of an antenna are also sharply modulated by the presence of feed legs, etc. and are hence difficult to determine with sufficient accuracy. For this reason determining the polarization across sources with dimensions comparable to the primary beam is a non trivial problem. Given the complexity of dealing with extended sources, most analysis to date have been restricted to small sources, ideally point sources located at the beam center.

Most radio telescopes measure non-orthogonal polarizations, i.e. a channel *p* which is supposed to be matched to some particular polarization p also picks up a small quantity of the orthogonal polarization q. Further, this leakage of the orthogonal polarization in general changes with position in the beam. However, for reflector antennas, there is often a leakage term that is independent of the location in the beam, which is traditionally ascribed to non idealities in the feed. For example, for dipole feeds, if the two dipoles are not mounted exactly at right angles to one another, the result is a real leakage term, and if the dipole is actually matched to a slightly elliptical (and not purely linear) polarization the result is an imaginary leakage term. For this reason, the real part of the leakage is sometimes called an *orientation* error, and the imaginary part of the leakage is referred to as an *ellipticity* error[7]. However, one should appreciate that the actual measurable quantity is only the antenna voltage beam, (i.e. the combined response of the feed and reflector) and this decomposition into 'feed' related terms is not fundamental and need not in general be physically meaningful.

The final effect that has to be taken into account has to do with the orientation of the antenna beam with respect to the source. For equitorially mounted telescopes this is a constant, however for alt-az mounted telescopes, the telescope beam rotates on the sky as the telescope tracks the source. This rotation is characterized by an angle called the parallactic angle, ψ_p, which is given by:

$$\tan\psi_p = \frac{\cos\mathcal{L}\sin\mathcal{H}}{\sin\mathcal{L}\cos\delta - \cos\mathcal{L}\sin\delta\sin\mathcal{H}}, \tag{15.2.8}$$

where $\mathcal{L}$ is the latitude of the telescope, $\mathcal{H}$ is the hour-angle of the source, and δ is the apparent declination of the source. So if one observes a source at a parallactic angle ψ_p with a telescope that is linearly polarized, the voltages that will be obtained at the

[7]Several telescopes, such as for example the GMRT, use feeds which are sensitive to linear polarization, but by using appropriate circuitry (viz a $\pi/2$ phase lag along one signal path before the first RF amplifier) convert the signals into circular polarization. Non idealities in this linear to circular conversion circuit could also produce complex leakage terms even if the feed dipoles themselves are error free.

terminals of the X and Y receivers will be

$$V_x = G_x(\mathcal{E}_x \cos\psi_p + \mathcal{E}_y \sin\psi_p), \quad V_y = G_y(-\mathcal{E}_x \sin\psi_p + \mathcal{E}_y \cos\psi_p), \tag{15.2.9}$$

where G_x and G_y are the complex gains (i.e. the product of the antenna voltage gains and the receiver gains) of the X and Y channels.

15.3 The Measurement Equation

In this section we will develop a mathematical formulation useful for polarimetric interferometry. The theoretical framework is the van Cittert-Zernike theorem, which was discussed in chapter 2 in the context of the reconstruction of the Stokes I parameter of the source. However, as can be trivially verified, the theorem holds good for any of the Stokes parameters. So, apart from the issues of spurious polarization produced by propagation or instrumental effects, making maps of the Q, U, and V Stokes parameters is in principle[8] identical to making a Stokes I map.

Not surprisingly, matrix notation leads to an elegant formulation for polarimetric interferometry[9]. Let us begin by defining a coherency vector,

$$\begin{pmatrix} <\mathcal{E}_{ap}\mathcal{E}_{bp}^*> \\ <\mathcal{E}_{ap}\mathcal{E}_{bq}^*> \\ <\mathcal{E}_{aq}\mathcal{E}_{bp}^*> \\ <\mathcal{E}_{aq}\mathcal{E}_{bq}^*> \end{pmatrix},$$

where a, b refer to the two antennas which compose any given baseline, and p, q are the two polarizations measured by the antenna. The coherency vector can be expressed as an outer product of the electric field, viz:

$$\begin{pmatrix} <\mathcal{E}_{ap}\mathcal{E}_{bp}^*> \\ <\mathcal{E}_{ap}\mathcal{E}_{bq}^*> \\ <\mathcal{E}_{aq}\mathcal{E}_{bp}^*> \\ <\mathcal{E}_{aq}\mathcal{E}_{bq}^*> \end{pmatrix} = \left\langle \begin{pmatrix} \mathcal{E}_{ap} \\ \mathcal{E}_{aq} \end{pmatrix} \otimes \begin{pmatrix} \mathcal{E}_{bp}^* \\ \mathcal{E}_{bq}^* \end{pmatrix} \right\rangle. \tag{15.3.10}$$

The Stokes vector can be obtained by multiplying the coherency vector with the Stokes matrix, (**S**). In a linear polarized co-ordinate system the components are:

$$\begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & -i & i & 0 \end{pmatrix} \begin{pmatrix} <\mathcal{E}_{ax}\mathcal{E}_{bx}^*> \\ <\mathcal{E}_{ax}\mathcal{E}_{by}^*> \\ <\mathcal{E}_{ay}\mathcal{E}_{bx}^*> \\ <\mathcal{E}_{ay}\mathcal{E}_{by}^*> \end{pmatrix}. \tag{15.3.11}$$

The component form could also be written down in the circular polarized co-ordinate system, in which case the matrix **S** would be:

$$\begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & -i & i & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} <\mathcal{E}_{ar}\mathcal{E}_{br}^*> \\ <\mathcal{E}_{ar}\mathcal{E}_{bl}^*> \\ <\mathcal{E}_{al}\mathcal{E}_{br}^*> \\ <\mathcal{E}_{al}\mathcal{E}_{bl}^*> \end{pmatrix}. \tag{15.3.12}$$

[8]apart from the fact that one has to record four correlation functions, $<\mathcal{E}_{ap}\mathcal{E}_{bp}^*>$, $<\mathcal{E}_{ap}\mathcal{E}_{bq}^*>$, $<\mathcal{E}_{aq}\mathcal{E}_{bp}^*>$, $<\mathcal{E}_{aq}\mathcal{E}_{bq}^*>$, where a, b refer to the two antennas which compose any given baseline, and p, q are the two polarizations measured by the antenna. Since Stokes I maps are often all that is required, many observatories, including the GMRT, make a trade off such that fewer spectral channels are available if you record all four correlation products, than if you recorded only the two correlation products which are required for Stokes I.

[9]Although this formulation has been in use in the field of optical polarimetry for decades, it was not appreciated until recently (Hamaker *et al.* 1996, and Sault *et al.* 1996) that it is also extendable to radio interferometric arrays.

The matrix in equation (15.3.12) is related to that in equation (15.3.11) by a simple permutation of rows, as expected.

The outer product has the following associative property, viz. for matrices, **A**,**B**, **C**, and **D**,

$$(\mathbf{AB}) \otimes (\mathbf{CD}) = (\mathbf{A} \otimes C)(\mathbf{B} \otimes D).$$

For any one antenna a, putting in all the various effects discussed in section(15.2) we can write the voltage at the antenna terminals as:

$$\begin{aligned} \mathcal{V}_a &= \mathbf{G}_a \mathbf{B}_a \mathbf{P}_a \mathbf{F}_a \mathcal{E}_a \\ &= \mathbf{J}_a \mathcal{E}_a. \end{aligned} \tag{15.3.13}$$

where,

$\mathcal{V}_a$	=	the voltage vector at the terminals of antenna a
$\mathbf{G}_a$	=	the complex gain of the receivers of antenna a
$\mathbf{B}_a$	=	the voltage beam matrix for antenna a
$\mathbf{P}_a$	=	the parallactic angle matrix for antenna a
$\mathbf{F}_a$	=	the Faraday rotation matrix for antenna a
$\mathcal{E}_a$	=	the electric field vector at antenna a
$\mathbf{J}_a$	=	the Jones matrix for antenna a

The Jones matrix has been so called because of its analogy with the Jones matrix in optical polarimetry. All of these matrices are 2×2. In the linear polarized co-ordinate system. For example, we have:

$$\begin{aligned} \mathbf{F} &= \begin{pmatrix} \cos\chi & \sin\chi \\ -\sin\chi & \cos\chi \end{pmatrix} \qquad \mathbf{P} = \begin{pmatrix} \cos\psi_p & \sin\psi_p \\ -\sin\psi_p & \cos\psi_p \end{pmatrix} \\ \mathbf{B} &= \begin{pmatrix} b_{pp}(l,m) & b_{pq}(l,m) \\ b_{qp}(l,m) & b_{qq}(l,m) \end{pmatrix} \quad \mathbf{G} = \begin{pmatrix} g_p & 0 \\ 0 & g_q \end{pmatrix}. \end{aligned} \tag{15.3.14}$$

The Jones matrix in polarimetric interferometry plays the same role as the complex gain does in scalar interferometry. Consequently one could conceive of schemes for self-calibration, since for an array with a large enough number of antennas sufficient number of closure constraints are available. However, since astrophysical sources are usually only weakly polarized, the signal to noise ratio in the cross-hand correlation products is often too low to make use of these closure constraints.

In scalar interferometry, phase fluctations caused by the atmosphere and/or ionosphere were lumped together with the instrumental gain fluctuations. In the vector formulation however, this is strictly speaking not possible, since these corrections occur at different points along the signal path, (see equations (15.3.13)) and matrices in equations (15.3.14) do not in general commute. However, for most existing radio telescopes, and for sources small compared to the primary beam, the matrices in equations (15.3.14) (apart from the Faraday rotation and Parallactic angle matrices) differ from the identity matrix only to first order (i.e. the off diagonal terms are small compared to the diagonal terms, and the diagonal terms are equal to one another to zeroth order), and consequently these matrices commute to first order. To first order hence, it is correct to lump the phase differences accumulated at different points along the signal path into the receiver gain. Alternatively, if we make the (reasonable) assumption that the complex atennuation (i.e. any absorption and phase fluctuation) produced by the atmosphere is identical for both polarizations, then it can be modeled as a constant times the identity matrix. Since the identity matrix commutes with all the other matrices, this factor can be absorbed in the receiver gain matrix, exactly as was done when dealing with interferometry of scalar fields. This is the reason why no separate matrix was introduced in equation (15.3.13) to account for atmospheric phase and amplitude fluctuations.

The matrix **B** in this formulation also deserves some attention. It simply contains the information on the relation between the electric field falling on the source and the voltage generated at the antenna terminals. It is an extension of the voltage beam in scalar field theory, and each element in the matrix depends on the sky co-ordinates (l, m). As described above in section(15.2), it is traditional to decompose it into a part which does not depend on (l, m), which is called the leakage (or in the matrix formulation, the leakage matrix "**D**"), and a part which depends on (l, m). Provided that the leakage terms are small compared to the parallel hand antenna voltage gain, it can be shown that this decomposition is unique to first order.

In terms of the Jones matrix, the measured visibility on a single baseline for a point at the phase center can be written as:

$$\begin{pmatrix} \mathcal{V}_I \\ \mathcal{V}_Q \\ \mathcal{V}_U \\ \mathcal{V}_V \end{pmatrix} = \mathbf{S}\mathbf{J}_a \otimes \mathbf{J}_b^* \mathbf{S}^{-1} \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix}. \qquad (15.3.15)$$

Note that this is a matrix equation, valid in all co-ordinate frames, i.e. it holds regardless of whether the antennas are linear polarized or circular polarized. In fact it holds even if some of the antennas are linear polarized, and the others are circular polarized.

If the point source were not at the phase center, then the visibility phase is not zero, and in equation (15.3.15), one would have to pre-multipy the Jones matrices with a matrix containing the Fourier kernel, viz. $\mathbf{K}_a(l, m)$, and $\mathbf{K}_b(l, m)$defined as:

$$\begin{aligned} \mathbf{K}_a(l, m) &= \begin{pmatrix} e^{-2\pi(u_a l + v_a m)} & 0 \\ 0 & e^{-2\pi(u_a l + v_a m)} \end{pmatrix}, \\ \mathbf{K}_b(l, m) &= \begin{pmatrix} e^{-2\pi(u_b l + v_b m)} & 0 \\ 0 & e^{-2\pi(u_b l + v_b m)} \end{pmatrix}. \end{aligned} \qquad (15.3.16)$$

To get the visibility for an extended incoherent source, one would have to integrate over all (l, m), thus recovering the vector formulation of the van Cittert-Zernike theorem. In order to invert this equation, it is necessary not only to do the inverse fourier transform, but also to correct for the various corruptions introduced, i.e. the data has to be calibrated. The rest of this chapter discusses ways in which this polarization calibration can be done.

15.4 Polarization Calibration

We restrict our attention to a point source at the phase center[10]. The visibility that we measure, averaged over all baselines is

$$\mathcal{V} = \frac{1}{N(N-1)} \mathbf{S} \sum_{a \neq b} (\mathbf{J}_a \otimes \mathbf{J}_b^*) \mathbf{S}^{-1}. \qquad (15.4.17)$$

Any system describable by a Jones matrix is non-depolarizing[11] In the general case however, the summation in equation (15.4.17) cannot be represented by a single Jones

[10]For VLBI observations this is a very good approximation, since the source being imaged is very small compared to the primary beams of any of the antennas in the VLBI array.

[11]This follows trivially from the fact that for 100% polarization we must have $\mathcal{E}_p = c\mathcal{E}_q$, where p, q are any two orthogonal polarizations, and c is some complex constant. Multiplication by the Jones matrix will preserve this relationship (only changing the value of the constant c) thus producing another 100% polarized wave.

matrix, and an interferometer is not therefore a non-depolarizing system. However, ideally, after calibration, the effective Jones matrices are all the unit matrix, and the interferometer would then be non-depolarizing.

Intuitively, it is clear that if one looks at an unpolarized calibrator source, one should be able to solve for the leakage terms, (which will produce apparent polarization) but that some degrees of freedom would remain unconstrained. Further it is also intuitive that the degrees of freedom which remain unconstrained are the following: (1) The absolute orientation of the feeds, (2) The intrinsic polarization of the feeds (i.e. for example, are they linear polarized or circular polarized?) and (3) The phase difference between the two polarizations. While one would imagine that the situation may be improved by observation of a polarized source, it turns out that this too is not sufficient to determine all the free parameters. What is required is observations of at least three differently polarized sources. For alt-az mounted dishes, the rotation of the beam with respect to the sky changes the apparent polarization of the source. For such telescopes hence, it is sufficient to observe a single source at several, sufficiently different hour angles. This is the polarization strategy that is commonly used at most telescopes. Faraday rotation due to the earth's ionosphere is more difficult to correct for. In principle models of the ionosphere coupled with a measure of the total electron content at the time of the observation can be used to apply a first order correction to the data.

We end this chapter with a brief description of the effect of calibration errors on the derived Stokes parameters. When observing with linearly polarized feeds, from equation (15.1.2) it is clear that if one observes a linearly polarized calibrator, the parallel-hand correlations will contain a contribution due to the Q component of the calibrator flux. Consequently, if one assumes (erroneously) that the calibrator was unpolarized the gain of the X channel will be overestimated and that of the Y channel underestimated. For this reason, for observations which require only measurement of Stokes I, circular feeds are preferable, since the Stokes V component of most calibrators is negligible, and consequently, measurements of the parallel hand correlations[12] are sufficient to measure the correct Stokes I flux.

It is easy to show, that (to first order) if one observes a polarized calibrator with an error free linearly polarized interferometer and solves for the instrumental parameters under the assumption that the calibrator is unpolarized, the derived instrumental parameters of all the antennas will be in error by[13]:

$$\begin{aligned} \Delta g_x &= +Q/2I & \Delta g_y &= -Q/2I \\ d_x &= (Q+iU)/2I & d_y &= -(U-iQ)/2I. \end{aligned}$$

where:

Δg_x	is the gain error of the X channel.
Δg_y	is the gain error of the Y channel.
d_x	is the leakage from the Y channel to the X channel.
d_y	is the leakage from the X channel to the Y channel.

If these calibration solutions are then applied to an unpolarized target source, then the source will appear to be polarized, with the same polarization percentage as the calibrator, but opposite sense. This again is simply the extension from scalar interferometry that if the calibrator flux is in error by some amount, the derived target source flux will be in error by the same fractional amount, but with opposite sense.

[12] recall from equations (15.1.3) that when $V = 0$, $<\mathcal{E}_r\mathcal{E}_r^*> + <\mathcal{E}_l\mathcal{E}_l^*>= I$.

[13] A similar result can of course be derived for the case of circularly polarized antennas, the only difference will be the usual transpositions of Q, U, and V.

15.5 Further Reading

1. Born, M. & Wolf, E., *'Principles of Optics'*, Cambridge University Press.
2. Hamaker, J. P., Bregman, J. D. & Sault, R. J., *'Understanding Polarimetric Interferometry I. Mathematical Foundations'*, A&A Supp. Ser., **117**,137, 1996.
3. C. Heiles, *A heuristic introduction to Radio Astronomic Polarisation*, astro-ph/0107372.
4. Sault, R. J., Hamaker, J. P. & Bregman, J. D. & *'Understanding Polarimetric Interferometry II. Instrumental Calibration of an Interferometric Array'*, A&A Supp. Ser., **117**,149, 1996.
5. Thompson, R. A., Moran, J. M. & Swenson, G. W. Jr., 'Interferometry & Synthesis in Radio Astronomy', Wiley Interscience, 1986.