Polarization in Interferometry

Rick Perley
(NRAO-Socorro)

Tenth Synthesis Imaging Summer School
Albuquerque/Socorro NM,
June 13-20, 2006

Apologies, Up Front

- This is tough stuff. Difficult concepts, hard to explain without complex mathematics.
- I will endeavor to minimize the math, and maximize the concepts with figures and 'handwaving'.
- Many good references:
  - Born and Wolf: Principle of Optics, Chapters 1 and 10
  - Rolfs and Wilson: Tools of Radio Astronomy, Chapter 2
  - Thompson, Moran and Swenson: Interferometry and Synthesis in Radio Astronomy, Chapter 4
  - Tinbergen: Astronomical Polarimetry. All Chapters.
- Great care must be taken in studying these – conventions vary between them.

What is Polarization?

- Electromagnetic field is a vector phenomenon – it has both direction and magnitude.
- From Maxwell's equations, we know a propagating EM wave (in the far field) has no component in the direction of propagation – it is a transverse wave.

$$\mathbf{k} \bullet \mathbf{E} = 0$$

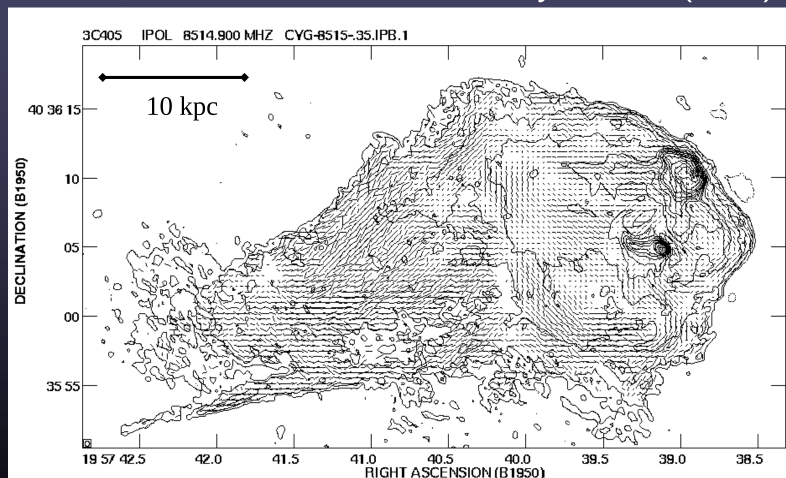
- The characteristics of the transverse component of the electric field, E, are referred to as the polarization properties.

Why Measure Polarization?

- In short – to access extra physics not available in total intensity alone.
- Examples:
  - Processes which generate polarized radiation:
    - Synchrotron emission: Up to ~80% linearly polarized, with no circular polarization. Measurement provides information on strength and orientation of magnetic fields, level of turbulence.
    - Zeeman line splitting: Presence of B-field splits RCP and LCP components of spectral lines by by 2.8 Hz/μG. Measurement provides direct measure of B-field.
  - Processes which modify polarization state:
    - Faraday rotation: Magnetoionic region rotates plane of linear polarization. Measurement of rotation gives B-field estimate.
    - Free electron scattering: Induces a linear polarization which can indicate the origin of the scattered radiation.

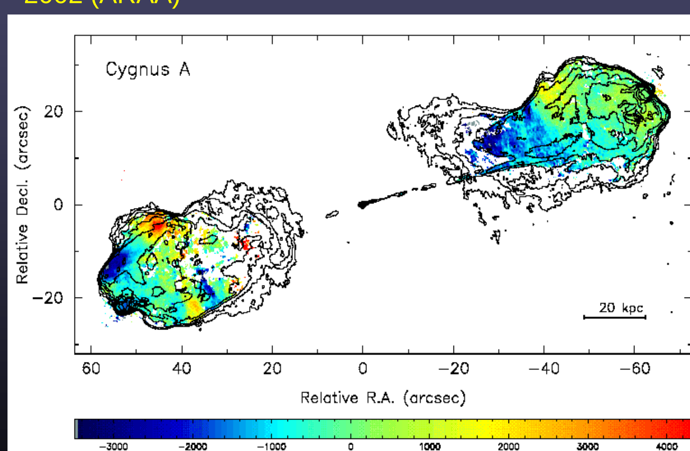
Example: Cygnus A

- VLA @ 8.5 GHz   B-vectors   Perley & Carilli (1996)

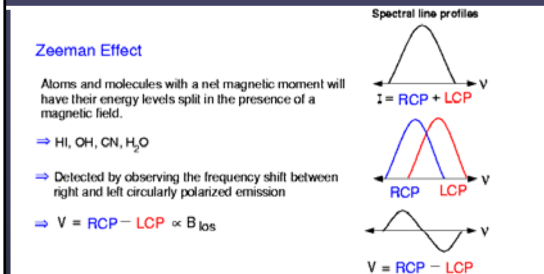


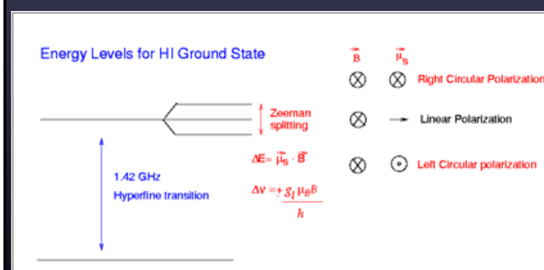
Example: more Faraday rotation

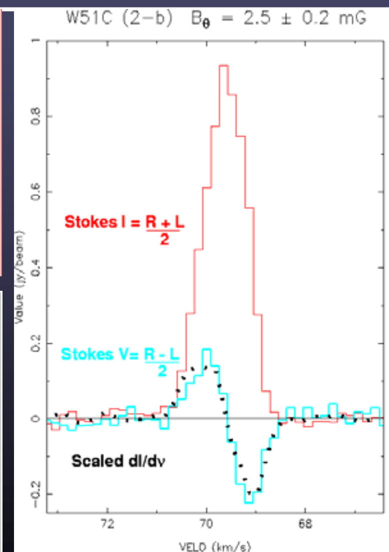
– See review of “Cluster Magnetic Fields” by Carilli & Taylor 2002 (ARAA)



Example: Zeeman effect







The Polarization Ellipse

- By convention, we consider the time behavior of the E-field in a fixed perpendicular plane, from the point of view of the receiver.
- For a monochromatic wave of frequency ν, we write

$$E_x = A_x \cos(2\pi \upsilon t + \phi_x)$$
$$E_y = A_y \cos(2\pi \upsilon t + \phi_y)$$

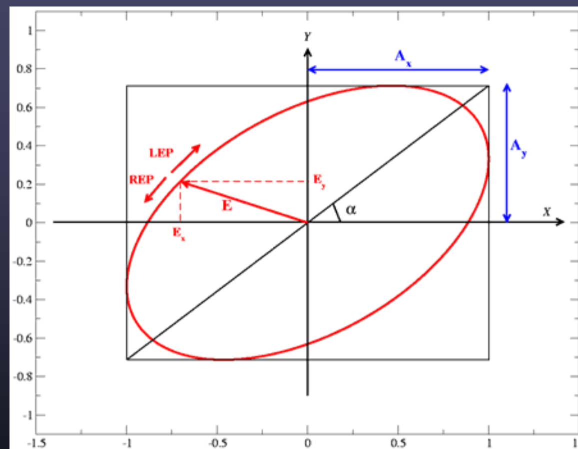
- These two equations describe an ellipse in the (x-y) plane.
- The ellipse is described fully by three parameters:
  - A_X, A_Y, and the phase difference, $\delta = \phi_Y - \phi_X$.
- The wave is elliptically polarized. If the E-vector is:
  - Rotating clockwise, the wave is 'Left Elliptically Polarized',
  - Rotating counterclockwise, it is 'Right Elliptically Polarized'.

Ellipticly Polarized Monochromatic Wave

The simplest description of wave polarization is in a Cartesian coordinate frame.

In general, three parameters are needed to describe the ellipse.

The angle $\alpha = \mathrm{atan}(A_Y/A_X)$ is used later …



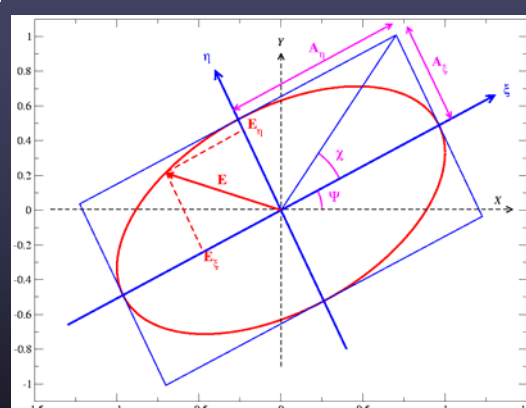
Polarization Ellipse Ellipticity and P.A.

- A more natural description is in a frame (ξ,η), rotated so the ξ-axis lies along the major axis of the ellipse.
- The three parameters of the ellipse are then:
  - A_η : the major axis length
  - $\tan\chi = A_\xi/A_\eta$: the axial ratio
  - Ψ : the major axis p.a.

$$\tan 2\Psi = \tan 2\alpha \cos\delta$$

$$\sin 2\chi = \sin 2\alpha \sin\delta$$

- The ellipticity χ is signed:
  - $\chi > 0$ => REP
  - $\chi < 0$ => LEP



Circular Basis

- We can decompose the E-field into a circular basis, rather than a (linear) cartesian one:

$$\mathbf{E} = A_R \hat{e}_R + A_L \hat{e}_L$$

  – where A_R and A_L are the amplitudes of two counter-rotating unit vectors, e_R (rotating counter-clockwise), and e_L (clockwise)

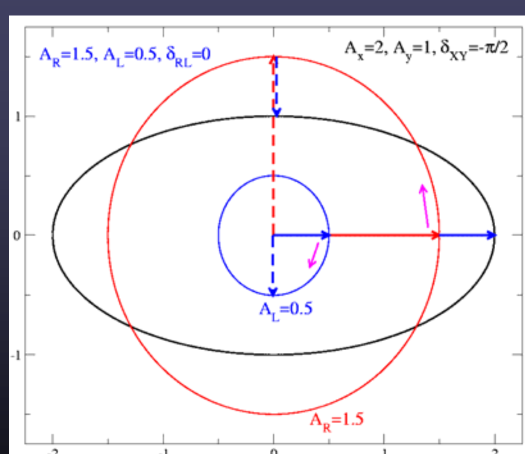
- It is straightforwards to show that:

$$A_R = \frac{1}{2}\sqrt{A_X^2 + A_Y^2 - 2A_X A_Y \sin\delta_{XY}}$$

$$A_L = \frac{1}{2}\sqrt{A_X^2 + A_Y^2 + 2A_X A_Y \sin\delta_{XY}}$$

Circular Basis Example

- The black ellipse can be decomposed into an x-component of amplitude 2, and a y-component of amplitude 1 which lags by ¼ turn.
- It can alternatively be decomposed into a counterclockwise rotating vector of length 1.5 (red), and a clockwise rotating vector of length 0.5 (blue).



Stokes' Parameters

- The three parameters already defined (major axis p.a., ellipticity, and major axis length) are sufficient for a complete description of monochromatic radiation.
- They have different units – a field amplitude, an angle, and a ratio.
- It is standard in radio astronomy to utilize the parameters defined by George Stokes (1852):

$$I = E_x^2 + E_y^2 = E_r^2 + E_l^2$$

$$Q = E_x^2 - E_y^2 = 2E_r E_l \cos\delta_{rl}$$

$$U = 2E_x E_y \cos\delta_{xy} = 2E_r E_l \sin\delta_{rl}$$

$$V = 2E_x E_y \sin\delta_{xy} = E_r^2 - E_l^2$$

- Note that $I^2 = Q^2 + U^2 + V^2$
- Thus – a monochromatic wave is 100% polarized.

Linear Polarization

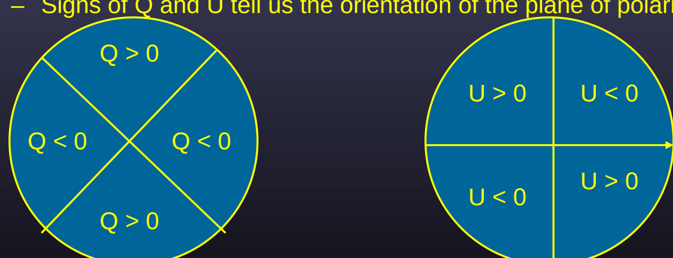
- Linearly Polarized Radiation: V = 0
  - Linearly polarized flux:

$$p = \sqrt{Q^2 + U^2}$$

  - Q and U define the plane of polarization:

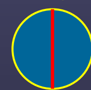
$$\tan 2\psi = U / Q$$

  - Signs of Q and U tell us the orientation of the plane of polarization:

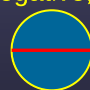


Simple Examples

- If V = 0, the wave is linearly polarized. Then,
  - If U = 0, and Q positive, then the wave is vertically polarized.



  - If U = 0, and Q negative, the wave is horizontally polarized.

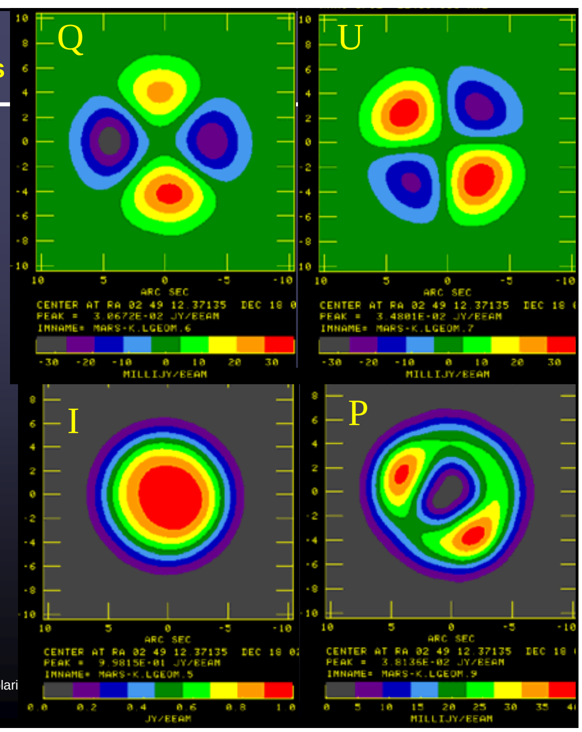


  - If Q = 0, and U positive, the wave is polarized at pa = 45 deg

  - If Q = 0, and U negative, the wave is polarized at pa = -45.

Illustrative Examples – Thermal Emission from Mars

- Mars emits in the radio as a black body.
- Shown are the I,Q,U,P images from Jan 2006 data at 23.4 GHz.
- V is not shown – all noise.
- Resolution is 3.5”, Mars’ diameter is ~6”.
- From the Q and U images alone, we can deduce the polarization is radial, around the limb.
- Position Angle image not usefully viewed in color.



Stokes Parameters

- Why use Stokes parameters?
  - Tradition
  - They have units of power
  - They are simply related to actual antenna measurements.
  - They easily accommodate the notion of partial polarization of non-monochromatic signals.
  - We can (as I will show) make images of the I, Q, U, and V intensities directly from measurements made from an interferometer.
  - These I,Q,U, and V images can then be combined to make images of the linear, circular, or elliptical characteristics of the radiation.

Non-Monochromatic Radiation, and Partial Polarization

- Monochromatic radiation is a myth.
- No such entity can exist (although it can be closely approximated).
- In real life, radiation has a finite bandwidth.
- Real astronomical emission processes arise from randomly placed, independently oscillating emitters (electrons).
- We observe the summed electric field, using instruments of finite bandwidth.
- Despite the chaos, polarization still exists, but is not complete – partial polarization is the rule.
- Stokes parameters defined in terms of mean quantities:

Stokes Parameters for Partial Polarization

$$I = \langle E_x^2 \rangle + \langle E_y^2 \rangle = \langle E_r^2 \rangle + \langle E_l^2 \rangle$$

$$Q = \langle E_x^2 \rangle - \langle E_y^2 \rangle = 2\langle E_r E_l \cos\delta_{rl} \rangle$$

$$U = 2\langle E_x E_y \cos\delta_{xy} \rangle = 2\langle E_r E_l \sin\delta_{rl} \rangle$$

$$V = 2\langle E_x E_y \sin\delta_{xy} \rangle = \langle E_r^2 \rangle - \langle E_l^2 \rangle$$

Note that now, unlike monochromatic radiation, the radiation is not necessarily 100% polarized.

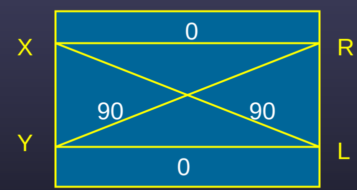
$$I^2 \geq Q^2 + U^2 + V^2$$

Antenna Polarization

- To do polarimetry (measure the polarization state of the EM wave), the antenna must have two outputs which respond differently to the incoming elliptically polarized wave.
- It would be most convenient if these two outputs are proportional to either:
  - The two linear orthogonal Cartesian components, (E_X, E_Y) or
  - The two circular orthogonal components, (E_R, E_L).
- Sadly, this is not the case in general.
- In general, each port is elliptically polarized, with its own polarization ellipse, with its p.a. and ellipticity.
- However, as long as these are different, polarimetry can be done.

An Aside: Quadrature Hybrids

- We've discussed the two bases commonly used to describe polarization.
- It is quite easy to transform signals from one to the other, through a real device known as a 'quadrature hybrid'.



- To transform correctly, the phase shifts must be exactly 0 and 90 for all frequencies, and the amplitudes balanced.
- Real hybrids are imperfect – an generate their own set of errors.

Antenna Polarization Ellipse

- We can thus describe the characteristics of the polarized outputs of an antenna in terms of its antenna polarization ellipse:

  χ_R and Ψ_R, for the RCP output

  χ_L and Ψ_L, for the LCP output

If the antenna is equipped with circularly polarized feeds,

- Or,

  χ_X and Ψ_X, for the 'X' output,

  χ_Y and Ψ_Y, for the 'Y' output

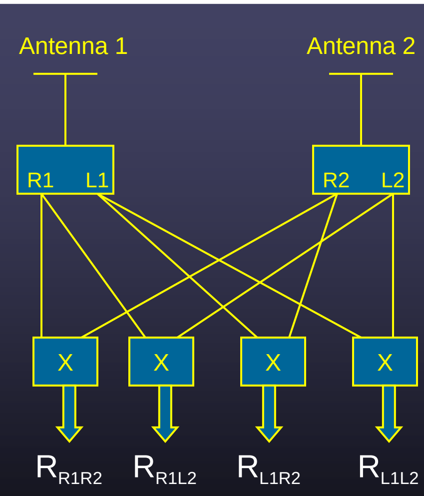
If the antenna is equipped with linearly polarized feeds.

Four Independent Outputs

- We are looking to determine the four Stokes values for the emission of interest.
- We thus need four independent quantities from which we can derive I, Q, U, and V.
- Each antenna provides two independent (differently polarized) outputs.
- We thus generate four (complex) products for each pair of antennas, and ask:
- How do these products relate to what we're looking for?

Four Complex Correlations per Pair

- Two antennas, each with two differently polarized outputs, produce four complex correlations.
- From these four outputs, we want to make four Stokes Images.



Polarization in Interferometry – Rick Perley

Interferometer Response

- DANGER! The next slide could be hazardous to your health!
- We are now in a position to show the most general expression for the output of a complex correlator, comprising imperfectly polarized antennas to wide-band partially-polarized astronomical signals.
- This is a complex expression (in all senses of that adjective), and I will make no attempt to derive, or even justify it.
- The expression is completely general, valid for a linear system.

Here it is!

$$
\begin{aligned}
R_{pq} = \tfrac{1}{2} G_{pq} \{ & I_V [\cos(\Psi_p - \Psi_q)\cos(\chi_p - \chi_q) + i\sin(\Psi_p - \Psi_q)\sin(\chi_p + \chi_q)] \\
& + Q_V [\cos(\Psi_p + \Psi_q)\cos(\chi_p + \chi_q) + i\sin(\Psi_p + \Psi_q)\sin(\chi_p - \chi_q)] \\
& - iU_V [\cos(\Psi_p + \Psi_q)\sin(\chi_p - \chi_q) + i\sin(\Psi_p + \Psi_q)\cos(\chi_p + \chi_q)] \\
& - V_V [\cos(\Psi_p - \Psi_q)\sin(\chi_p + \chi_q) + i\sin(\Psi_p - \Psi_q)\cos(\chi_p - \chi_q)] \}
\end{aligned}
$$

What are all these symbols?

R_{pq} is the complex output from the interferometer, for polarizations p and q from antennas 1 and 2, respectively.

Ψ and χ are the antenna polarization major axis and ellipticity for states p and q.

I_V, Q_V, U_V, and V_V are the Stokes Visibilities describing the polarization state of the astronomical signal.

Stokes Visibilities

- And what (you ask), are the Stokes Visibilities?
- They are the (complex) Fourier transforms of the I, Q, U, and V spatial distributions of emission from the sky.
- $I_V \Leftrightarrow I$, $Q_V \Leftrightarrow Q$, $U_V \Leftrightarrow U$, $V_V \Leftrightarrow V$
- Thus, these are what we are looking to get from the four complex outputs from the baselines of the array.
- Once we can recover the I_V, Q_V, U_V, and V_V values from the complex interferometer response, we can invert them via a Fourier transform to obtain the four spatial images.

Idealized Antennas

- We now begin an analysis of this lovely expression.
- To ease you in as painlessly as possible, let us consider the idealized situation where the antennas are perfectly polarized.
- There are two cases of interest: Linear and Circular.

Orthogonal, Perfectly Linear Feeds

- In this case, $\chi = 0$, $\Psi_V = 0$, $\Psi_H = \pi/2$. (We are presuming the antenna orientation is fixed w.r.t the sky).
- Then,

$$R_{VV} = \frac{I_V + Q_V}{2}$$

$$R_{HH} = \frac{I_V - Q_V}{2}$$

$$R_{VH} = \frac{U_V + iV_V}{2}$$

$$R_{HV} = \frac{U_V - iV_V}{2}$$

- From these, we can trivially invert, and recover the desired Stokes Visibilities.

Perfectly Circular Antennas

- So let us continue with our idealizations, and ask what the response is for perfectly circular feeds.
- Now we have: $\chi_R = -\pi/4$, $\chi_L = \pi/4$.
- Then,

$$R_{RR} = \frac{I_V + V_V}{2}$$

$$R_{LL} = \frac{I_V - V_V}{2}$$

$$R_{RL} = \frac{Q_V + iU_V}{2}$$

$$R_{LR} = \frac{Q_V - iU_V}{2}$$

- And again a trivial inversion provides our desired quantities.

Comments

1. I have assumed that the data are perfectly calibrated. In general, gain factors accompany each expression.
2. Examination of these expressions shows why in most cases, circularly polarized antennas are preferred:
   - Parallel-hand correlations of linear feeds are modulated by Q . As most of the compact sources we use for calibration are ~5% linearly polarized, but have no circular polarization, gain calibration is easier with circular feeds.
   - Derivation of Stokes Q for linear feeds requires subtraction of two large quantities, ($Q = R_{VV} – R_{HH}$) while for circular, it comes from the cross-hand response, ($Q = R_{RL} + R_{LR}$) which is independent of I.

- From this simple analysis, circularly polarized feeds allow easier calibration, and more accurate measurement of linear polarization.

Why not Circular for all?

- The VLA (and EVLA) use circularly polarized feeds.
- Most new arrays do not (e.g. ALMA, ATCA).
- Do they know something we don't?
  - Antenna feeds are natively linearly polarized. To convert to circular, a hybrid is needed. This adds cost, complexity, and degrades performance.
  - For some high frequency systems, wideband quarter-wave phase shifters may not be available, or their performance may be too poor.
  - Linear feeds can give perfectly good linear polarization performance, provided the amplifier/signal path gains are carefully monitored.
  - Nevertheless, the calibration issues remain, if linearly polarized calibrators are to be used.

Alt-Az Antennas, and the Parallactic Angle

- The prior expressions presumed that the antenna feeds are fixed in orientation on the sky.
- This is the situation with equatorially mounted antennas.
- For alt-az antennas, the feeds rotate on the sky as they track a source at fixed declination.
- The angle between a line of constant azimuth, and one of constant right ascension is called the Parallactic Angle, η :

$$\sin\eta = \sin A \cos\phi / \cos\delta$$

where A is the antenna azimuth, ϕ is the antenna latitude, and δ is the source declination.

Including Parallactic Angle variations

- Presuming all antennas view the source with the same parallactic angle (not true for VLBI!), the responses from pure polarized antennas are straightforward to derive.
- For Linear Feeds:

$$R_{VV} = \frac{I_V + Q_V \cos 2\Psi_P + U_V \sin 2\Psi_P}{2}$$

$$R_{HH} = \frac{I_V - Q_V \cos 2\Psi_P - U_V \sin 2\Psi_P}{2}$$

$$R_{VH} = \frac{U_V \cos 2\Psi_P - Q_V \sin \Psi_P + iV_V}{2}$$

$$R_{HV} = \frac{U_V \cos 2\Psi_P - Q_V \sin \Psi_P - iV_V}{2}$$

Including Parallactic Angle variations

- For circular polarized (rotating) feeds, the expressions are simpler:

$$R_{RR} = \frac{I_V + V_V}{2}$$
$$R_{LL} = \frac{I_V - V_V}{2}$$
$$R_{RL} = \frac{e^{-2i\Psi_P}(Q_V + iU_V)}{2}$$
$$R_{LR} = \frac{e^{2i\Psi_P}(Q_V - iU_V)}{2}$$

- For both systems, it is straightforward to recover the Stokes Visibilities, as the parallactic angle is known.

Imperfect (Elliptically Polarized) Antennas

- So much for perfection. In the real world, we don't get purely polarized antennas. How does reality modify our easy life?
- We must now introduce the concept of the D-terms – an alternate description of antenna polarization.

$$D_R = \tan \Delta\chi_R e^{2i\Psi_R}$$
$$D_L = \tan \Delta\chi_L e^{-2i\Psi_L}$$
where
$$\Delta\chi_R = \chi_R + \pi/4$$
$$\Delta\chi_L = \pi/4 - \chi_L$$

N.B.: The Ψ are in the antenna's frame. Parallactic angle is removed.

- These D-terms measure the departure of the antenna response from perfect circularity.
- For well-designed systems, the magnitude of the Ds is small – typically .01 to .05.

Interferometer Response, with the D-formulation

- With this substitution, we can derive an alternate general form of the interferometer response:

$$R_{R1R2} = \left[(I_V + V_V) + D_{R1}D_{R2}^*(I_V - V_V) + e^{-2i\Psi_P}D_{R2}^*(Q_V + iU_V) + e^{2i\Psi_P}D_{R1}(Q_V - iU_V)\right]/2$$

$$R_{L1L2} = \left[D_{L1}D_{L2}^*(I_V + V_V) + (I_V - V_V) + e^{-2i\Psi_P}D_{L1}(Q_V + iU_V) + e^{2i\Psi_P}D_{L2}^*(Q_V - iU_V)\right]/2$$

$$R_{R1L2} = \left[D_{L2}^*(I_V + V_V) + D_{R1}(I_V - V_V) + e^{-2i\Psi_P}(Q_V + iU_V) + e^{2i\Psi_P}D_{R1}D_{L2}^*(Q_V - iU_V)\right]/2$$

$$R_{L1R2} = \left[D_{L1}(I_V + V_V) + D_{R2}^*(I_V - V_V) + e^{-2i\Psi_P}D_{L1}D_{R2}^*(Q_V + iU_V) + e^{2i\Psi_P}(Q_V - iU_V)\right]/2$$

- Time for some approximations: |D| < .05, and |Q|, |U|, and |V| are both typically less than 5% of |I|. Thus, ignore all 2nd order products.

'Nearly' Circular Feeds

- We then get the much simpler set:

$$R_{R1R2} = \left[(I_V + V_V)\right]/2$$

$$R_{L1L2} = \left[(I_V - V_V)\right]/2$$

$$R_{R1L2} = \left[(D_{R1} + D_{L2}^*)I_V + e^{-2i\Psi_P}(Q_V + iU_V)\right]/2$$

$$R_{L1R2} = \left[(D_{L1} + D_{R2}^*)I_V + e^{2i\Psi_P}(Q_V - iU_V)\right]/2$$

- Our problem is now clear. The desired cross-hand responses are contaminated by a term of roughly equal size.
- To do accurate polarimetry, we must determine these D-terms, and remove their contribution.

Some Comments

- If we can arrange that $D_{L1} + D_{R2}^* = 0, \quad \text{and} \quad D_{R1} + D_{L2}^* = 0$ then there is no polarization leakage! (to first order).
- This condition occurs if the two antenna polarization ellipses (R1 and L2 in the first case, and L1 and R2 for the second) have equal ellipticity and are orthogonal in orientation.
- This is called the `orthogonality condition'.
- Determination of the D (leakage) terms is normally done either by:
  - Observing a source of known (I,Q,U) strengths, or
  - Multiple observations of a source of unknown (I,Q,U), and allowing the rotation of parallactic angle to separate the two terms.
- Note that for each, the absolute value of D cannot be determined – they must be referenced to an arbitrary value.

Nearly Perfectly Linear Feeds

- Sadly, the D-term formulation cannot be usefully applied to nearly-linear feeds, as the deviations from perfectly circularity are not small!
- In the case of nearly-linear feeds, we return to the fundamental set, and assume that the ellipticity is very small (χ << 1), and that the two feeds ('dipoles') are nearly perfectly orthogonal.
- We then define a *different* set of D-terms:

$$D_V = \varphi_V - i\chi_V$$
$$D_H = -\varphi_H + i\chi_H$$

- The angles φ_H and φ_V are the angular offsets from the exact horizontal and vertical orientations, w.r.t. the antenna.

This gives us …

- I'll spare you the full equation set, and show only the results after the same approximations used for the circular case are employed.

$$R_{VV} = \frac{I_V + Q_V \cos 2\Psi_P + U_V \sin 2\Psi_P}{2}$$

$$R_{HH} = \frac{I_V - Q_V \cos 2\Psi_P - U_V \sin 2\Psi_P}{2}$$

$$R_{VH} = \frac{I_V\left(D_{V1} + D_{H2}^{*}\right) + U_V \cos 2\Psi_P - Q_V \sin \Psi_P + iV_V}{2}$$

$$R_{HV} = \frac{I_V\left(D_{H1} + D_{V2}^{*}\right) + U_V \cos 2\Psi_P - Q_V \sin \Psi_P - iV_V}{2}$$

Some Comments on Linears

- The problem is the same as for the circular case – the derivation of the Q, U, and V Stokes' visibilities is contaminated by a leakage of the much larger I visibility into the cross-hand response.
- Calibration is similar to the circular case:
  - If Q, U, and V are known, then the equations can be solved directly for the Ds.
  - If the polarization is unknown, then the antenna rotation can again be used (over time) to separate the polarized response from the leakage response.

A Summary (of sorts)

- I'll put something here, once I figure it all out!