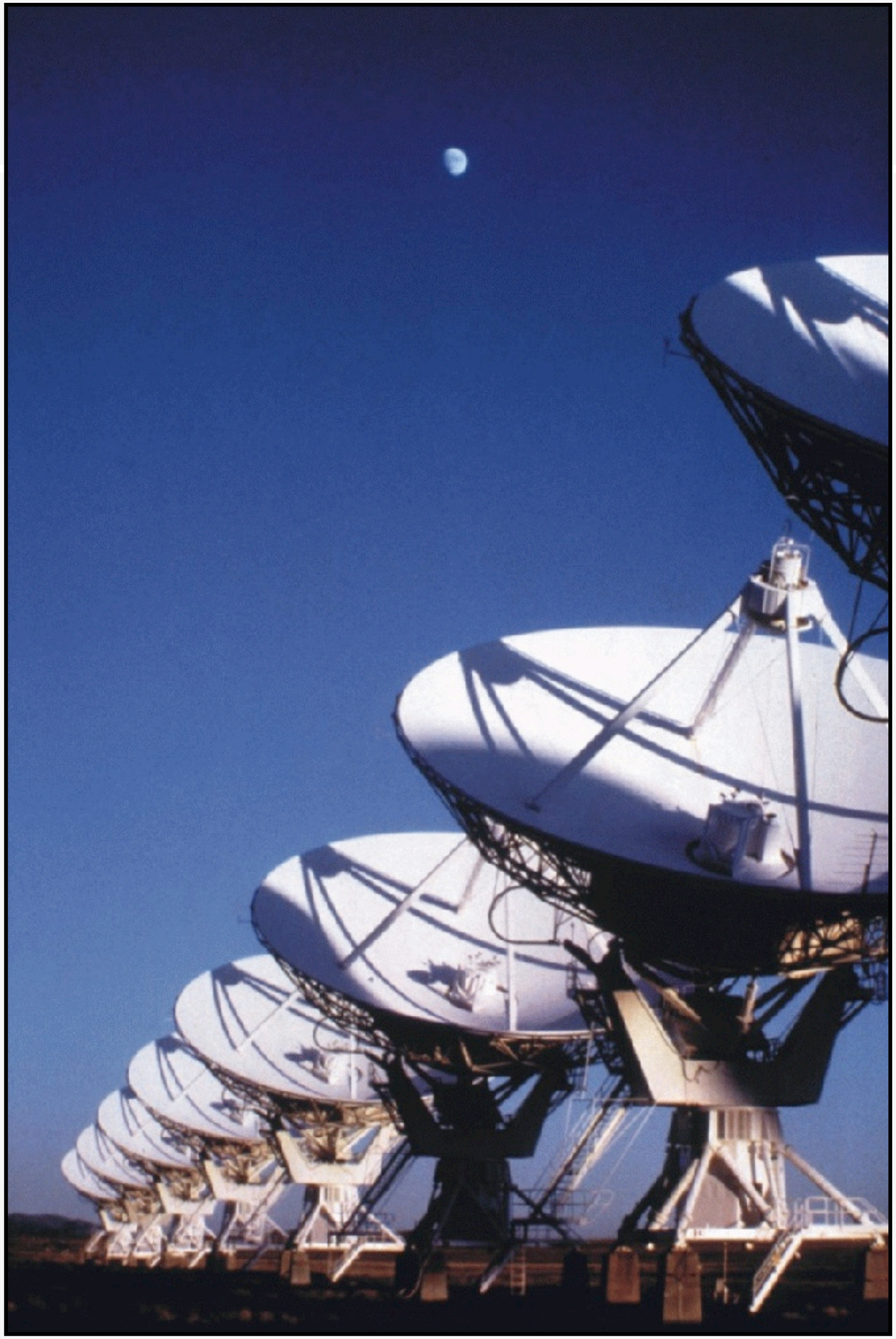
Polarization in Interferometry

Rick Perley
(NRAO-Socorro)

Fourth INPE Advanced Course on Astrophysics:

Radio Astronomy in the 21st Century



My Apologies, Up Front

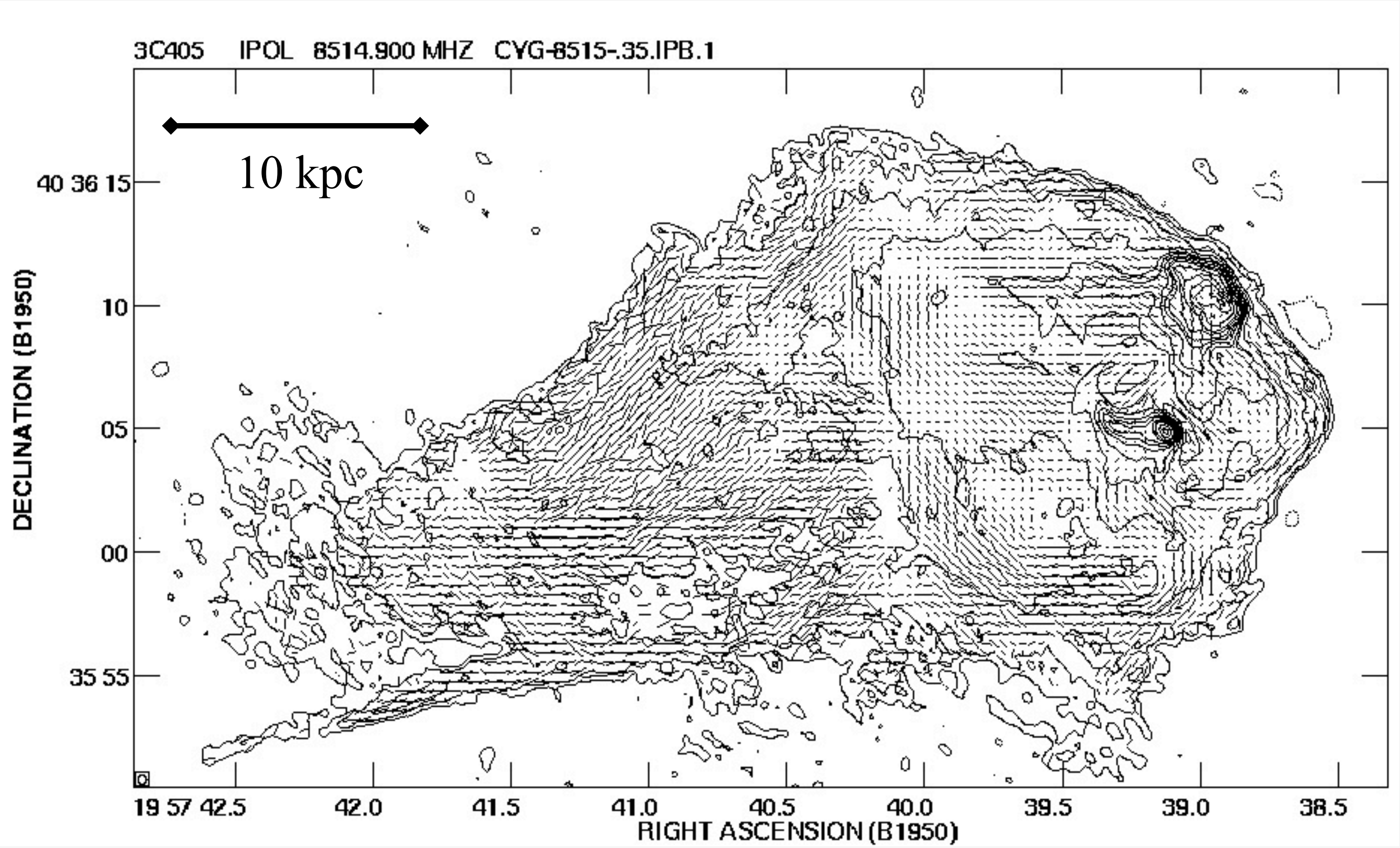
- This lecture contains some difficult material.
- It is *impossible* to adequately convey this material in 75 minutes!
- So this lecture will only hit the ‘high points’, and makes extensive use of figures and diagrams.
- Many good references:
  - Born and Wolf: ‘Principle of Optics’, Chapters 1 and 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.

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 2.8 Hz/μG. Measurement provides direct measure of B-field.
  - Processes which modify polarization state:
    - **Faraday rotation:** Magneto-ionic 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.

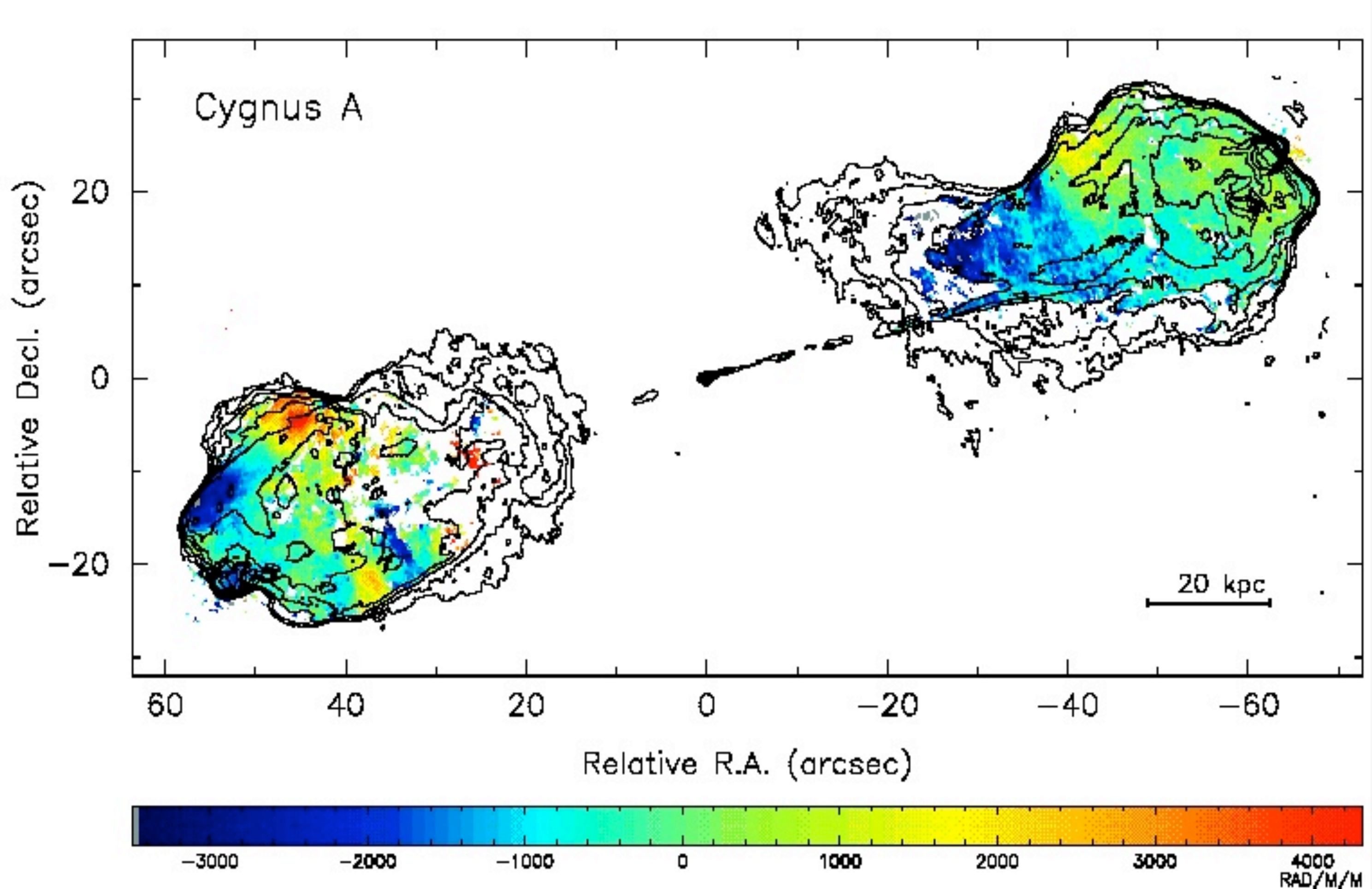
Example: Linear Polarization of Cygnus A

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



Example: Faraday Rotation for Cygnus A

– Color-coded map of the Rotation Measure – the slope of Faraday Rotation vs. λ^2.
– RM is proportional to the density-weighted longitudinal B-field, imbedded in the cluster gas surrounding the radio source.



Dreher, Carilli and Perley (1987))

Example: Zeeman effect

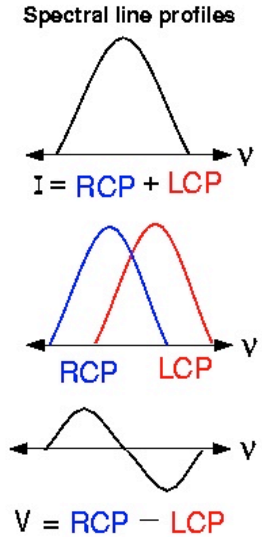
Zeeman Effect

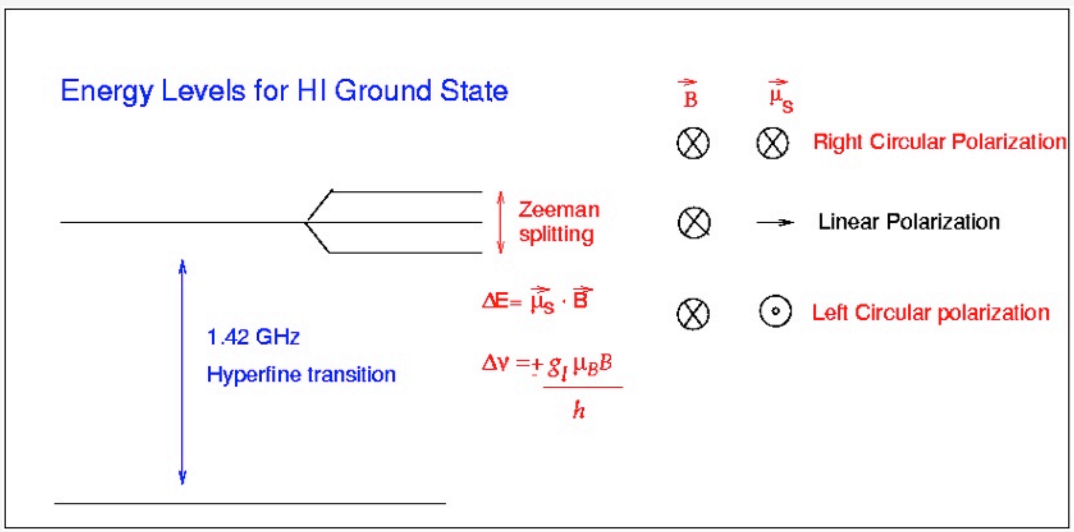
Atoms and molecules with a net magnetic moment will have their energy levels split in the presence of a magnetic field.

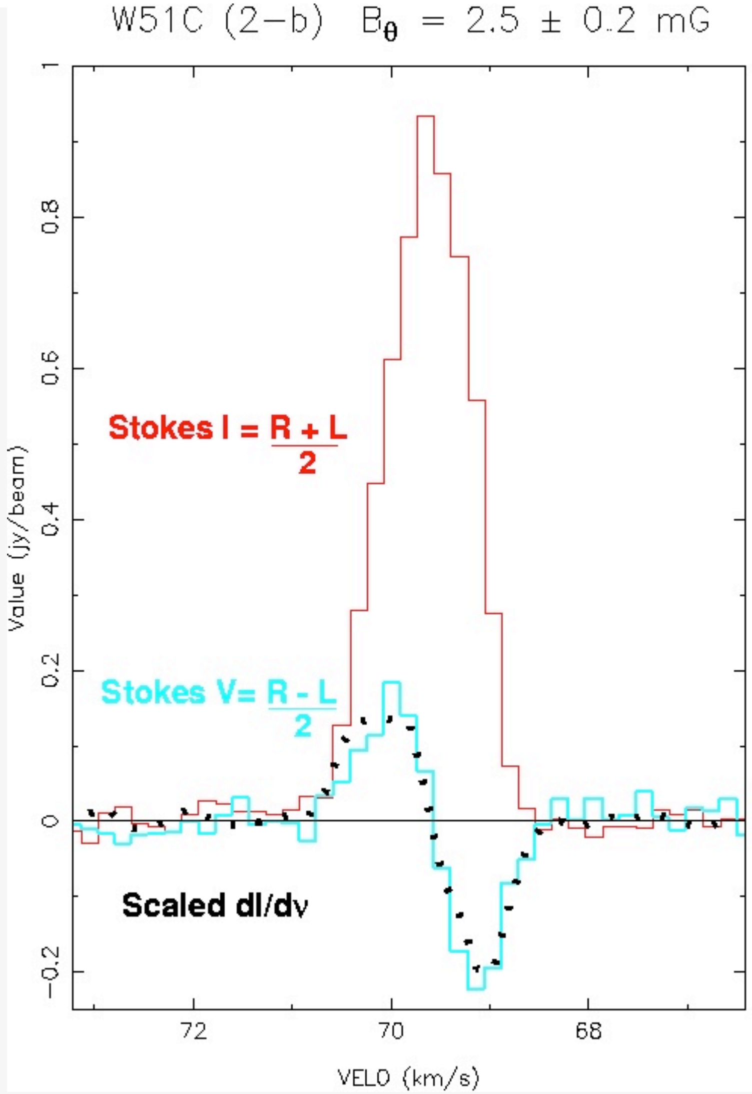
⇒ HI, OH, CN, H_2O

⇒ Detected by observing the frequency shift between right and left circularly polarized emission

⇒ V = RCP − LCP $\propto B_{los}$







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:

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

- Hence, in a coordinate frame with the (z) axis oriented along the direction of propagation, $E_z = 0$.
- The characteristics of the independent transverse components (E_x, E_y) of the electric field are referred to as the polarization properties.

The Polarization Ellipse

- We consider the time behavior of the E-field in a fixed perpendicular plane, with the z-axis directed along the direction of propagation towards the observer.
- For a monochromatic wave of frequency ν, we can write

$$E_x = A_x \cos(2\pi\upsilon\, t + \delta_x)$$
$$E_y = A_y \cos(2\pi\upsilon\, t + \delta_y)$$

- These two equations describe an ellipse in the (X-Y) plane:

$$\left(\frac{E_x}{A_x}\right)^2 + \left(\frac{E_y}{A_y}\right)^2 - 2\frac{E_x}{A_x}\frac{E_y}{A_y}\cos\delta = \sin^2\delta$$

- The ellipse is described by three parameters:
  - The amplitudes A_x, A_y, and
  - The phase difference, $\delta = \delta_y - \delta_x$

Elliptically Polarized Monochromatic Wave

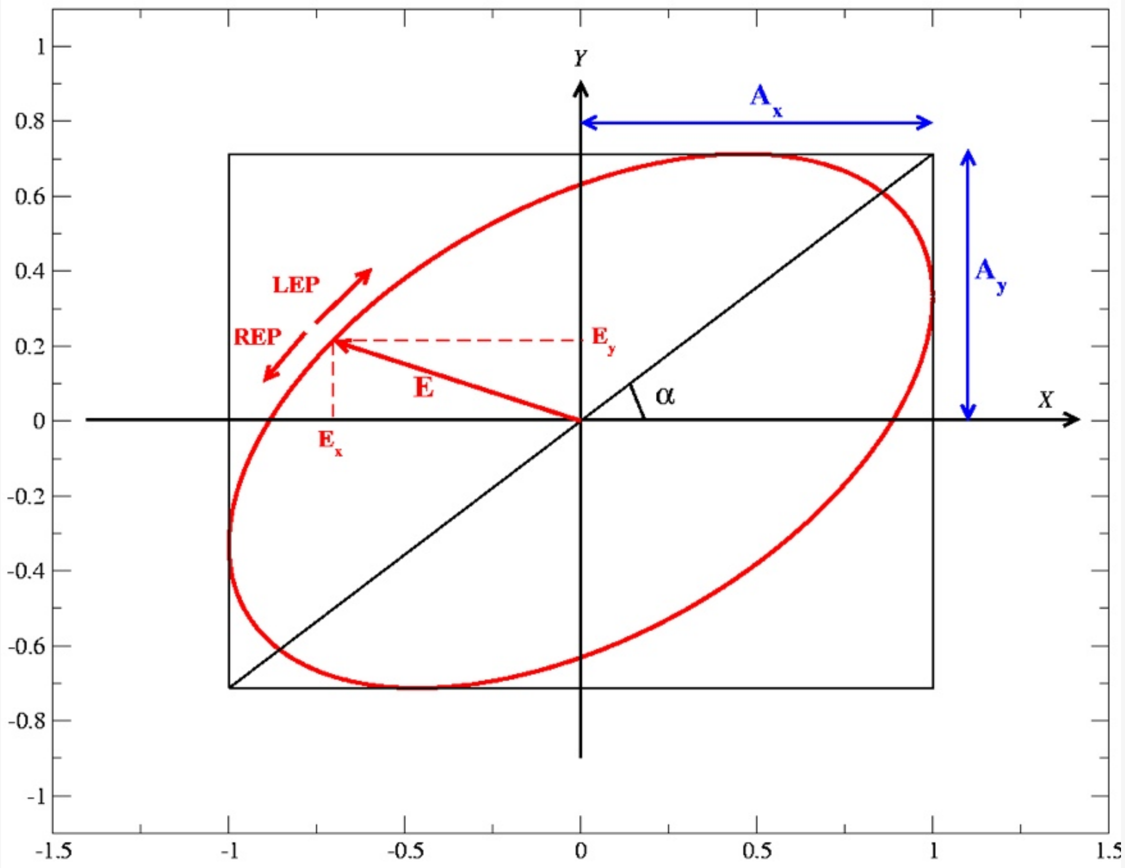
According to Maxwell's equations, an EM wave is elliptically polarized.

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

- A_x – X-axis amplitude max
- A_y – Y-axis amplitude max
- $\alpha = \text{atan}(A_y/A_x)$ – an angle describing the orientation

If the E vector is rotating:

- Clockwise, the wave is Left Elliptically Polarized:
- Anti-clockwise, the wave is Right Elliptically Polarized.



In a more Natural Reference Frame

- 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
  - Ψ: the position angle of this major axis, and
  - $\tan\chi = A_\xi/A_\eta$: the axial ratio
- It can be shown that:

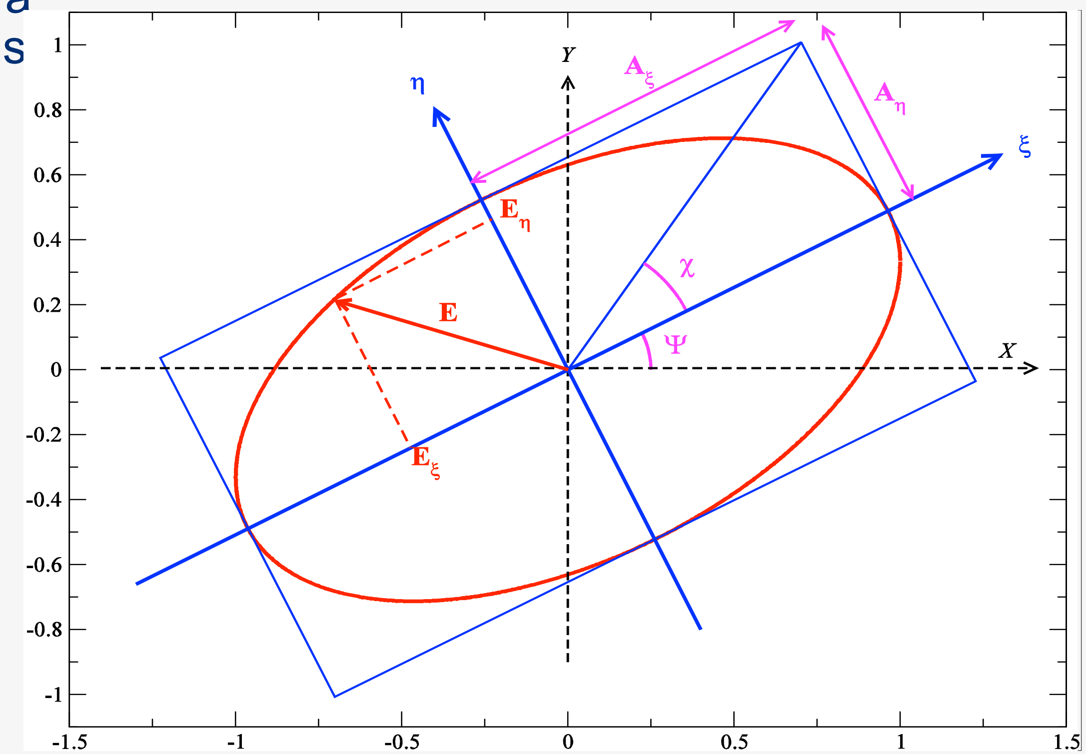
$$\tan 2\Psi = \tan 2\alpha \cos\delta$$

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

- The ellipticity χ is signed:

  $\chi > 0$ => LEP (clockwise)

  $\chi < 0$ => REP (anti-clockwise)



Alternate Description – The Circular Basis

- We can decompose the E-field into a circular basis, rather than a 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 (rotating clockwise)

- The polarization ellipse is again described by three parameters:

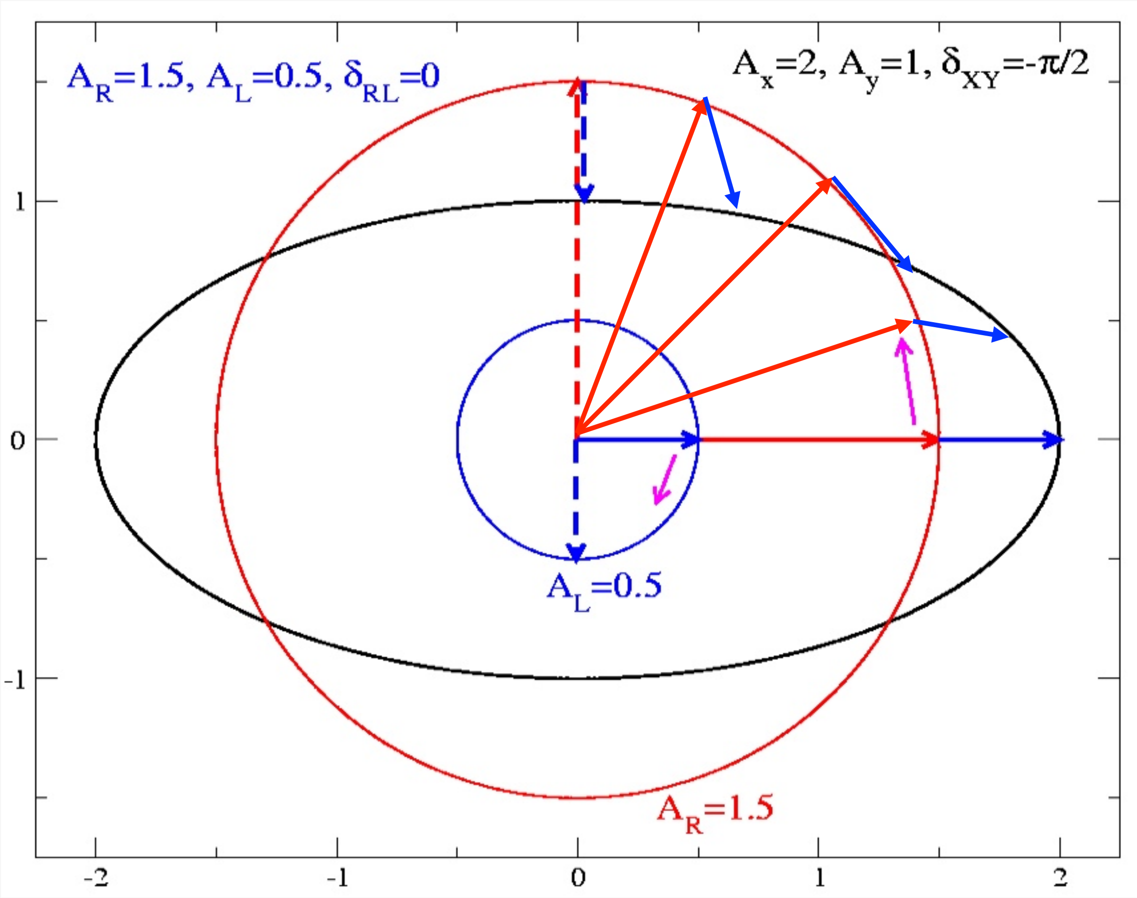
$$A_R = \frac{1}{2}\sqrt{A_X^2 + A_Y^2 - 2A_X A_Y \sin\delta_{XY}}$$

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

$$\tan\delta_{RL} = \frac{2A_x A_y \cos\delta_{XY}}{A_x^2 - A_y^2}$$

Circular Basis Example

- The polarization ellipse (black) 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 (RCP) rotating vector of length 1.5 (red), and a clockwise rotating (LCP) vector of length 0.5 (blue).



Stokes Parameters

- Three parameters are sufficient to describe the monochromatic EM wave properties.
- It is most convenient to have the three parameters share the same units, and have easily grasped physical meanings.
- It is standard in radio astronomy to utilize the parameters defined by George Stokes (1852), and introduced to astronomy by Chandrasekhar (1946):

$$
\begin{aligned}
I &= A_X^2 + A_Y^2 &&= A_R^2 + A_L^2 \\
Q &= A_X^2 - A_Y^2 &&= 2A_R A_L \cos\delta_{RL} \\
U &= 2A_X A_Y \cos\delta_{XY} &&= 2A_R A_L \sin\delta_{RL} \\
V &= 2A_X A_Y \sin\delta_{XY} &&= A_R^2 - A_L^2
\end{aligned}
$$

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

Interpreting the Meaning of I, V, Q, and U

$$I = A_x^2 + A_y^2 = A_R^2 + A_L^2$$

- I is the total power.

$$V = A_R^2 - A_L^2$$

- V is the difference between the power in RCP and LCP components.

$$Q = A_X^2 - A_Y^2$$

- Q is the difference between vertical and horizontal power.

$$U = 2A_X A_Y \cos\delta_{XY}$$

- U is the difference between orthogonal components in a frame rotated by 45 degrees.

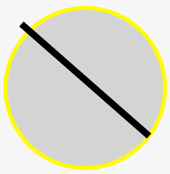
- If V = 0, the wave is 100% Linearly Polarized
- If Q=U=0, the wave is 100% Circularly Polarized
- In general, the wave is Elliptically Polarized

Pure Linear Polarization: V = 0

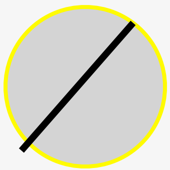
- Presume a linearly polarized wave, of unit power.
- Then, if:
  - The wave is vertically polarized:
    - Q = 1   U = 0

  - The wave is horizontally polarized:
    - Q = -1   U = 0

  - The wave is polarized at pa = 45 deg:
    - Q = 0   U = 1



  - The wave is polarized at pa = -45.
    - Q = 0   U = -1



Linear Polarization (Q, U)

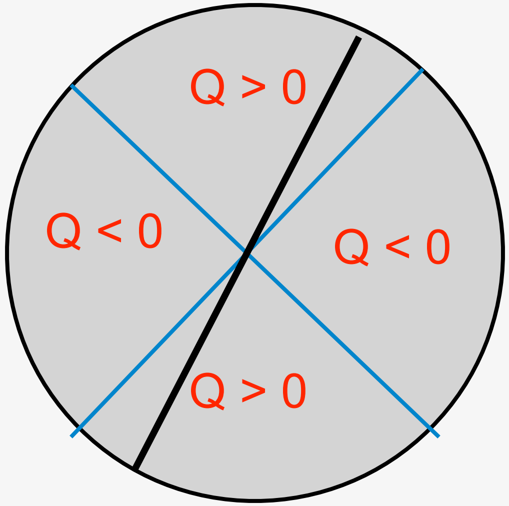
- In general, we define:
  - The Linearly Polarized Intensity:

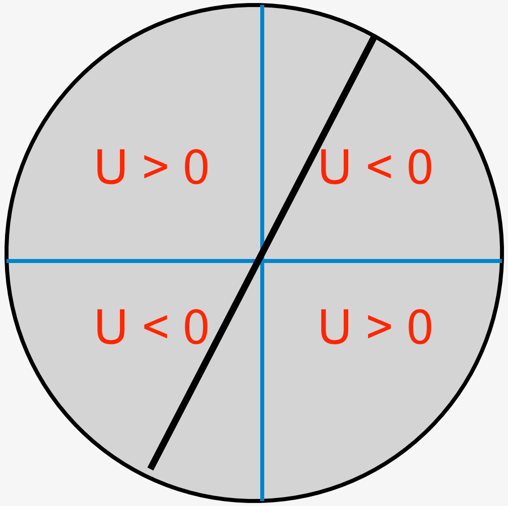
$$p = \sqrt{Q^2 + U^2}$$

  - The orientation of the plane of polarization:

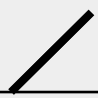
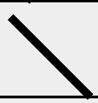
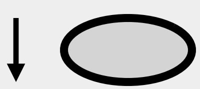
$$\tan 2\psi = U / Q$$

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





In General: Some Illustrative Examples

Pol Ellipse	I	V	Q	U
	1	1	0	0
	1	-1	0	0
	1	0	1	0
	1	0	-1	0
	1	0	0	-1
	1	0	0	1
	1	$1/\sqrt{2}$	$-1/\sqrt{2}$	0
	1	$-1/\sqrt{2}$	0	$1/\sqrt{2}$

Non-Monochromatic Radiation, and Partial Polarization

- Monochromatic radiation is nonphysical.
- No such condition can exist (although it can be closely approximated).
- In real life, radiation has a finite bandwidth $\Delta\nu$ – the signals are quasi-sinusoidal, specified by an amplitude and phase for a limited time $t \sim 1/\Delta\nu$.
- Because the amplitudes and phases of the received signals are now variable, there are no unique, stationary values of amplitude and phase.
- We must then use suitably defined averages of these quantities to define the polarization characteristics.

Stokes Parameters

- In the quasi-monochromatic approximation, the incoming EM wave can be described, for a period $\Delta t \sim 1/\Delta\nu$, by two amplitudes, A_p and A_q, and a phase difference, δ_{pq}.
- The two amplitudes describe the electric field amplitudes of the two independent orthogonal states of the radiation.
- For the orthogonal linear, and opposite circular bases, we have:

$$
\begin{aligned}
I &= \left\langle A_X^2 \right\rangle + \left\langle A_Y^2 \right\rangle &&= \left\langle A_R^2 \right\rangle + \left\langle A_L^2 \right\rangle \\
Q &= \left\langle A_X^2 \right\rangle - \left\langle A_Y^2 \right\rangle &&= \left\langle 2 A_R A_L \cos\delta_{RL} \right\rangle \\
U &= \left\langle 2 A_X A_Y \cos\delta_{XY} \right\rangle &&= \left\langle 2 A_R A_L \sin\delta_{RL} \right\rangle \\
V &= \left\langle 2 A_X A_Y \sin\delta_{XY} \right\rangle &&= \left\langle A_R^2 \right\rangle - \left\langle A_L^2 \right\rangle
\end{aligned}
$$

- The angle brackets <> denote an average over a time much longer than the coherence time $1/\Delta\nu$.
- These four real numbers are a complete description of the polarization state of the incoming radiation.
- They are a function of frequency, position, and time.

Stokes Visibilities

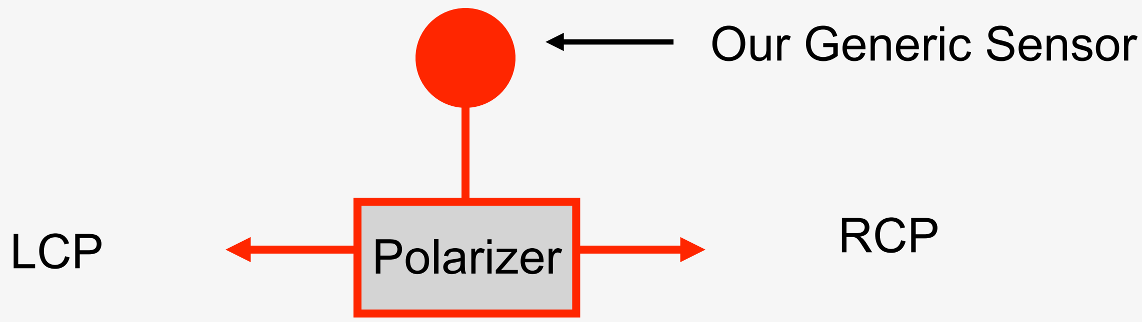
- Recall the earlier lectures, where we defined the Visibility, V(u,v), and showed its relation to the sky brightness:

$$V(u,v) \Leftrightarrow I(l,m) \quad \text{(a Fourier Transform Pair)}$$

- In our derivation, we were deliberately vague about what this brightness was.
- We will now be more formal, and consider the true brightness distributions for I, Q, U, and V.
- Define the **Stokes Visibilities** I, Q, U, and V, to be the Fourier Transforms of these brightness distributions.
- Then, the relations between these are:
- I $\Leftrightarrow$ I, Q $\Leftrightarrow$ Q, U $\Leftrightarrow$ U, V $\Leftrightarrow$ V
- Stokes Visibilities are complex functions of (u,v), while the Stokes Images are real functions of (l,m).
- Our task is now to measure these Stokes visibilities.

Polarimetric Interferometry

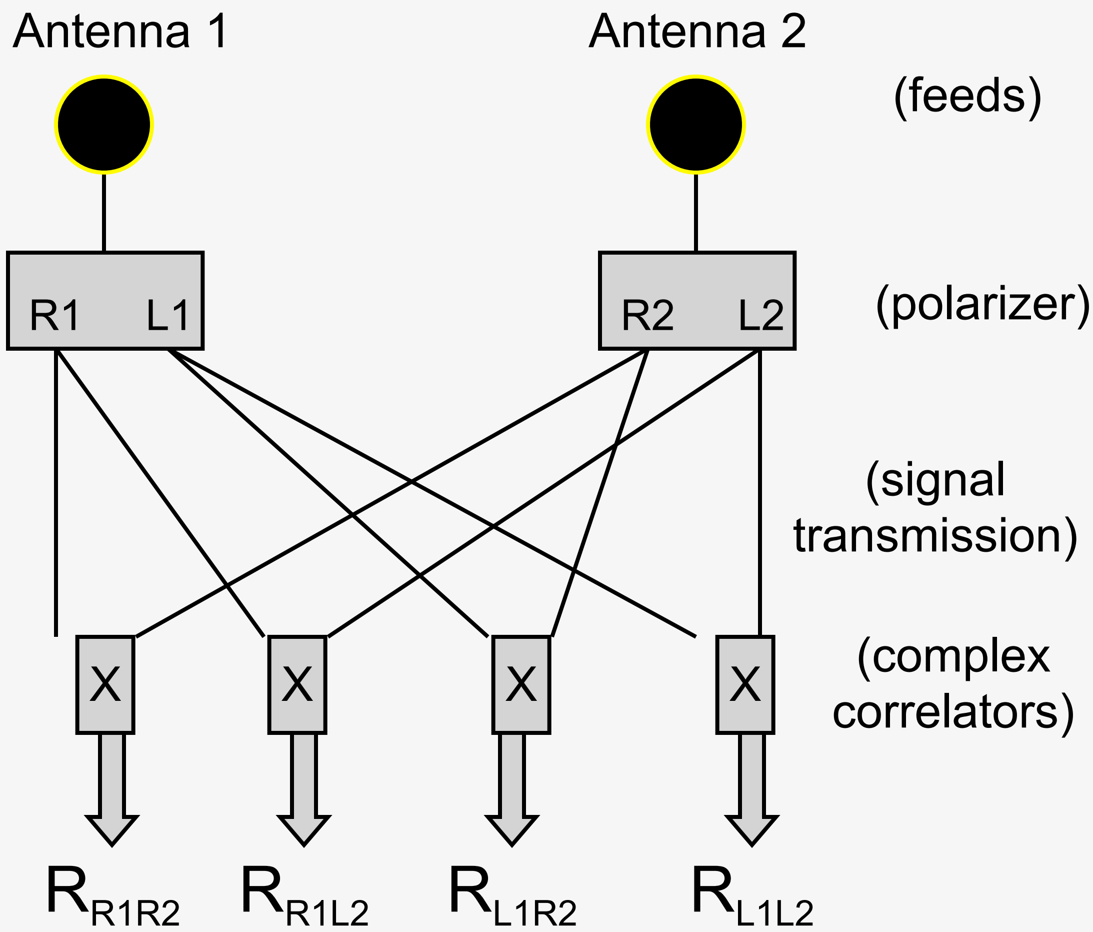
- Polarimetry is possible because antennas are polarized – their output is not a function of I alone.
- It is highly desirable (but not required) that the two outputs be sensitive to two orthogonal modes (i.e. linear, or circular).



- In interferometry, we have two antennas, each with two differently polarized outputs.
- We can then form four complex correlations.
- What is the relation between these four correlations and the four Stokes' parameters?

Four Complex Correlations per Pair of Antennas

- Two antennas, each with two differently polarized outputs, produce four complex correlations.
- From these four outputs, we want to generate the four complex visibilities, I, Q, U, and V



Relating the Products to Stokes' Visibilities

- Let E_{R1}, E_{L1}, E_{R2} and E_{L2} be the complex representation (phasors) of the RCP and LCP components of the EM wave which arrives at the two antennas.
- We can then utilize the definitions earlier given to show that the four complex correlations between these fields are related to the desired visibilities by (ignoring gain factors):

$$R_{R1R2} = \langle E_{R1}E^{*}_{R2} \rangle = (\mathrm{I} + \mathrm{V})/2$$

$$R_{L1L2} = \langle E_{L1}E^{*}_{L2} \rangle = (\mathrm{I} - \mathrm{V})/2$$

$$R_{R1L2} = \langle E_{R1}E^{*}_{L2} \rangle = (\mathrm{Q} + i\mathrm{U})/2$$

$$R_{L1R2} = \langle E_{L1}E^{*}_{R2} \rangle = (\mathrm{Q} - i\mathrm{U})/2$$

- So, if each antenna has two outputs whose voltages are faithful replicas of the EM wave's RCP and LCP components, then the simple equations shown are sufficient.
- (I've ignored gain factors here!)

Solving for Stokes Visibilities

- The solutions are straighforward:

$$
\begin{aligned}
\mathrm{I} &= R_{R1R2} + R_{L1L2} \\
\mathrm{V} &= R_{R1R2} - R_{L1L2} \\
\mathrm{Q} &= R_{R1L2} + R_{L1R2} \\
\mathrm{U} &= -i(R_{R1L2} - R_{L1R2})
\end{aligned}
$$

- Normally, Q, U, and V are much smaller than I (low polarization).
- Thus, the amplitudes of the cross-hand correlations are much less than the parallel hand correlations.
- V is formed from the difference of two large quantities, while Q and U are formed from the sum and difference of small quantities.
- If calibration errors dominate (and they often do), the circular basis favors measurements of linear polarization.

For Linearly Polarized Antennas …

- We can go through the same exercise with perfectly linearly polarized feeds and obtain (presuming they are oriented with the vertical feed along a line of constant HA, and again ignoring issues of gain):

$$R_{V1V2} = \left\langle E_{V1} E^*_{V2} \right\rangle = (\mathrm{I} + \mathrm{Q})/2$$

$$R_{H1H2} = \left\langle E_{H1} E^*_{H2} \right\rangle = (\mathrm{I} - \mathrm{Q})/2$$

$$R_{V1H2} = \left\langle E_{V1} E^*_{H2} \right\rangle = (\mathrm{U} + i\mathrm{V})/2$$

$$R_{H1V2} = \left\langle E_{H1} E^*_{V2} \right\rangle = (\mathrm{U} - i\mathrm{V})/2$$

- For each example, we have four measured quantities and four unknowns.
- The solution for the Stokes visibilities is easy.

Stokes' Visibilities for Pure Linear

- Again, the solution in straightforwards:

$$
\begin{aligned}
\mathrm{I} &= R_{V1V2} + R_{H1H2} \\
\mathrm{Q} &= R_{V1V2} - R_{H1H2} \\
\mathrm{U} &= R_{V1H2} + R_{H1V2} \\
\mathrm{V} &= -i(R_{V1H2} - R_{H1V2})
\end{aligned}
$$

- We wish life were only so simple …
- We have ignored two realities of life in polarimetry:
  - Antennas rotate on the sky (commonly), and
  - Antennas are not perfectly polarized (always)

Antenna Rotation -- Circular

- I give (without derivation) how antenna rotation affects the results for the situation when all antennas are rotated by an angle Ψ_P w.r.t. the sky:
- For perfectly circularly polarized antennas:

$$R_{R1R2} = (\mathrm{I} + \mathrm{V})/2 \qquad \mathrm{I} = R_{R1R2} + R_{L1L2}$$

$$R_{L1L2} = (\mathrm{I} - \mathrm{V})/2 \qquad \mathrm{V} = R_{R1R2} - R_{L1L2}$$

$$R_{R1L2} = (\mathrm{Q} + i\mathrm{U})e^{i2\psi_P}/2 \qquad \mathrm{Q} = R_{R1L2}e^{i2\Psi_P} + R_{L1R2}e^{-i2\Psi_P}$$

$$R_{L1R2} = (\mathrm{Q} - i\mathrm{U})e^{-i2\psi_P}/2 \qquad \mathrm{U} = i(R_{L1R2}e^{-i2\Psi_P} - R_{R1L2}e^{i2\Psi_P})$$

- The effect of antenna rotation is to simply rotate the RL and LR visibilities.

Antenna Rotation, Linear

- For perfect linearly polarized antennas, rotated at an angle Ψ_P:

$$R_{V1V2} = (\mathrm{I} + \mathrm{Q}\cos 2\Psi_P + \mathrm{U}\sin 2\Psi_P)/2$$
$$R_{H1H2} = (\mathrm{I} - \mathrm{Q}\cos 2\Psi_P - \mathrm{U}\sin 2\Psi_P)/2$$
$$R_{V1H2} = (-\mathrm{Q}\sin 2\Psi_P + \mathrm{U}\cos 2\Psi_P + i\mathrm{V})/2$$
$$R_{H1V2} = (-\mathrm{Q}\sin 2\Psi_P + \mathrm{U}\cos 2\Psi_P - i\mathrm{V})/2$$

- With easy solution:

$$\mathrm{I} = R_{V1V2} + R_{H1H2})$$
$$\mathrm{Q} = \left(R_{V1V2} - R_{H1H2}\right)\cos 2\Psi_P - \left(R_{V1H2} + R_{H1V2}\right)\sin 2\Psi_P$$
$$\mathrm{U} = \left(R_{V1V2} - R_{H1H2}\right)\sin 2\Psi_P - \left(R_{V1H2} + R_{H1V2}\right)\cos 2\Psi_P$$
$$\mathrm{V} = i(R_{H1V2} - R_{V1H2})$$

Circular vs. Linear

- One of the ongoing debates is the advantages and disadvantages of Linear and Circular systems.
- Point of principle: For full polarization imaging, both systems must provide the same results. Advantages/disadvantages of each are based on points of practicalities.

Circular System	Linear System
$I = R_{R1R2} + R_{L1L2}$	$I = R_{V1V2} + R_{H1H2}$
$V = R_{R1R2} - R_{L1L2}$	$V = i\left(R_{H1V2} - R_{V1H2}\right)$
$Q = e^{i2\Psi_P} R_{R1L2} + e^{-i2\Psi_P} R_{L1R2}$	$Q = \left(R_{V1V2} - R_{H1H2}\right)\cos 2\Psi_P - \left(R_{V1H2} + R_{H1V2}\right)\sin 2\Psi_P$
$U = i\left(e^{-i2\Psi_P} R_{L1R2} - e^{i2\Psi_P} R_{R1L2}\right)$	$U = \left(R_{V1V2} - R_{H1H2}\right)\sin 2\Psi_P + \left(R_{V1H2} + R_{H1V2}\right)\cos 2\Psi_P$

- For both systems, Stokes ‘I’ is the sum of the parallel-hands.
- Stokes ‘V’ is the difference of the crossed hand responses for linear, (good) and is the difference of the parallel-hand responses for circular (bad).
- Stokes ‘Q’ and ‘U’ are differences of cross-hand responses for circular (good), and differences of parallel hands for linear (bad).

Circular vs. Linear

- Both systems provide straightforward derivation of the Stokes' visibilities from the four correlations.
- Making sense of differences of large numbers requires good stability and/or good calibration.
- To do good circular polarization using circular system, or good linear polarization with linear system, we need special care and special methods to ensure good calibration.
- But there are practical reasons to use linear:
  – Antenna polarizers are natively linear – extra components are needed for circular. This hurts performance.
  – These extra components are also generally of narrower bandwidth – it's harder to build circular systems with really wide bandwidth.
  – At mm wavelengths, the needed phase shifters are not available.
- One important practical reason for circular:
  – Nearly all of our calibrator sources are linearly polarized – making calibration of linear systems much more compllicated.

Calibration Troubles …

- To understand this last point, note that for the linear system:

$$R_{V1V2} = G_{V1}G^*_{V2}(\mathrm{I} + \mathrm{Q}\cos 2\Psi_P + \mathrm{U}\sin 2\Psi_P)/2$$
$$R_{H1H2} = G_{H1}G^*_{H2}(\mathrm{I} - \mathrm{Q}\cos 2\Psi_P - \mathrm{U}\sin 2\Psi_P)/2$$

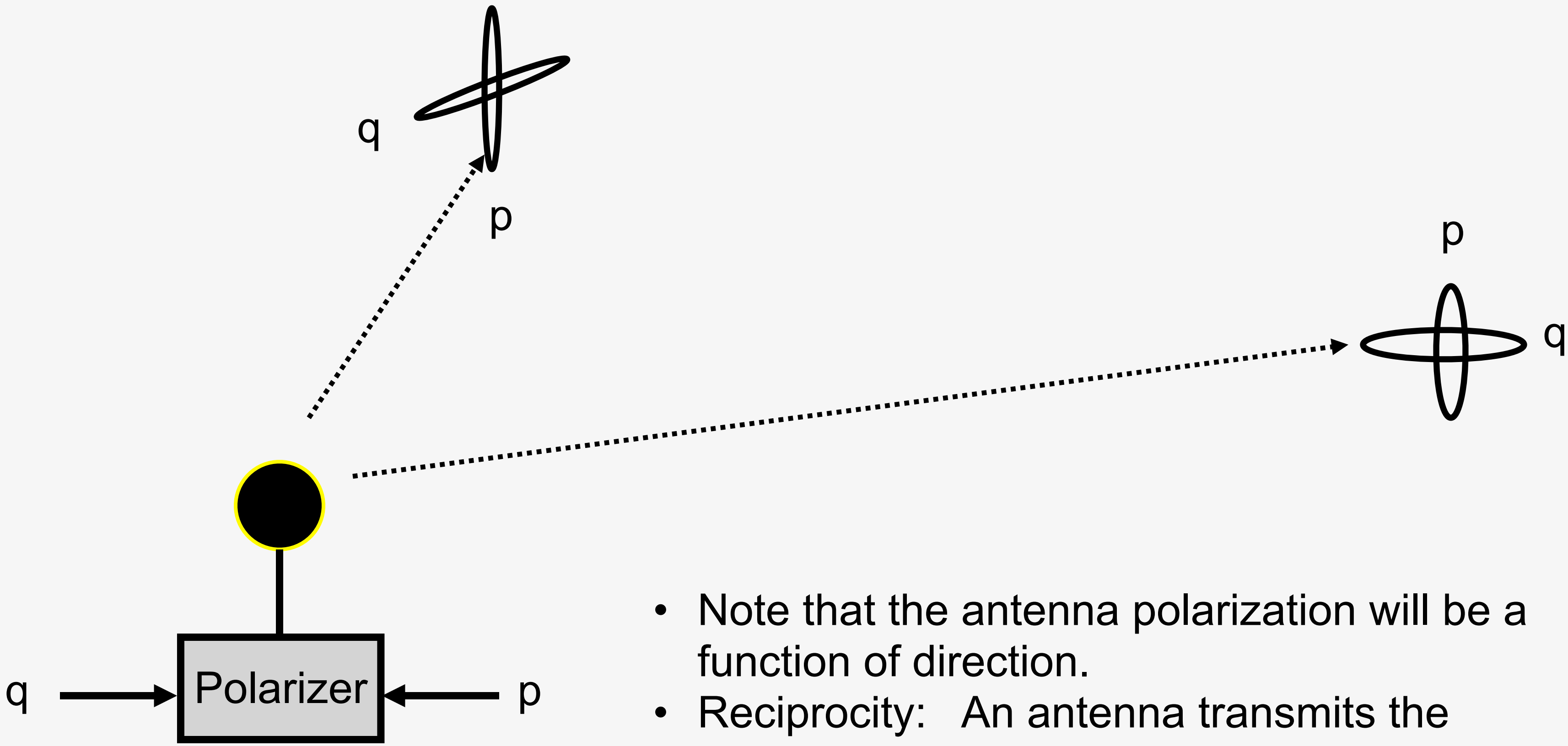
- To calibrate means to solve for the G_V and G_H terms.
- Easy if you know in advance Q and U – (and best if the source has no Q or U at all!). But often you don't know these.
- Meanwhile, for circular:

$$R_{R1R2} = G_{R1}G^*_{R2}(\mathrm{I} + \mathrm{V})/2$$
$$R_{L1L2} = G_{L1}G^*_{L2}(\mathrm{I} - \mathrm{V})/2$$

- Now we have *no* sensitivity to Q or U (good!). Instead, we have a sensitivity to V.
- But as it turns out – V is nearly always negligible for the 1000-odd sources that we use as standard calibrators.

Polarization of Real Antennas

- Unfortunately, antennas never provide perfectly orthogonal outputs.
- In general, the two outputs from an antenna are elliptically polarized.



- Note that the antenna polarization will be a function of direction.
- Reciprocity: An antenna transmits the same polarization that it receives.

Relating Output Voltages to Input Fields

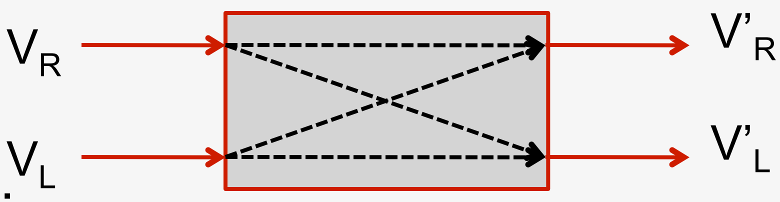
- The Stokes visibilities we want are defined in terms of the complex cross-correlations (coherencies) of electric fields: e.g. $<E_{R1}E^*_{R2}>$
- The quantities provided by the antenna are voltages, so what we get from our correlator are quantities like: $<V_{R1}V^*_{R2}>$
- Furthermore, in a real system, V_R isn't uniquely dependent upon E_R – it's a function of both polarizations and some gain factors:

$$V_R = G_R\left(C_{RR}E_R + S_{LR}E_L\right)$$

- We now develop a formalism to handle this general case.

Jones Matrix Algebra

- The analysis of how a real interferometer, comprising real antennas and real electronics, is greatly facilitated through use of Jones matrices.
- In this, we break up our general system into a series of 4-port components, each of which is presumed to be linear.
- We consider each component to have two inputs and two outputs:



- And write:

$$\begin{pmatrix} V'_{R'} \\ V'_L \end{pmatrix} = \begin{pmatrix} G_{RR} & G_{LR} \\ G_{RL} & G_{LL} \end{pmatrix} \begin{pmatrix} V_R \\ V_L \end{pmatrix}$$

- Or, in shorthand **V' = JV**
- The four G components of the Jones matrix describe the linkages within the 'grey box'.

Example Jones Matrices

- Each component of the overall system, including propagation effects, can be represented by a Jones matrix.
- These matrices can then be multiplied to obtain a 'system Jones' matrix.
- Examples (in a circular basis):
  - Faraday rotation by a magnetized plasma: $\begin{pmatrix} e^{i\phi_R} & 0 \\ 0 & e^{i\phi_L} \end{pmatrix}$
  - Atmospheric attenuation and phase retardation: $\begin{pmatrix} \alpha e^{i\phi} & 0 \\ 0 & \alpha e^{i\phi} \end{pmatrix}$
  - Antenna rotated by angle Ψ_P $\begin{pmatrix} e^{-i\Psi_P} & 0 \\ 0 & e^{i\Psi_P} \end{pmatrix}$
  - An imperfect polarizer (components are complex) $\begin{pmatrix} C_{RR} & S_{LR} \\ S_{RL} & C_{LL} \end{pmatrix}$
  - Post-polarizer electronic gains (complex): $\begin{pmatrix} G_R & 0 \\ 0 & G_L \end{pmatrix}$

The System Jones Matrix

- Now imagine a simple model, comprising of an antenna oriented at some angle Ψ_P to the sky, feeding an imperfect polarizer, followed by post-polarizer electronic gains.
- For this system, the output voltage (column vector) is related to the input electric fields by:

$$\mathbf{V} = \mathbf{J_G J_{pol} J_{rot} E} = \mathbf{J_{ant} E}$$

- Multiplying the various Jones matrices, we find

$$\begin{pmatrix} V_R \\ V_L \end{pmatrix} = \begin{pmatrix} G_R C_{RR} e^{-i\Psi_P} & G_R S_{LR} e^{i\Psi_P} \\ G_L S_{RL} e^{-i\Psi_P} & G_L C_{LL} e^{i\Psi_P} \end{pmatrix} \begin{pmatrix} E_R \\ E_L \end{pmatrix}$$

- We can now perform the complex cross-multiplies, and express the result in terms of the Stokes visibilities.
- One could do this serially (four products, with 16 combinations of the coefficients), or we can utilize matrix algebra.
- This operation, applied to matrices, is called the ‘outer product’.

Definition of the Outer (Kronecker) Product

- Each element of the first matrix is expanded to four elements, formed from multiplication with the four elements of the second:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \otimes \begin{pmatrix} b_{11}^* & b_{12}^* \\ b_{21}^* & b_{22}^* \end{pmatrix} = \begin{pmatrix} a_{11}b_{11}^* & a_{11}b_{12}^* & a_{12}b_{11}^* & a_{12}b_{12}^* \\ a_{11}b_{21}^* & a_{11}b_{22}^* & a_{12}b_{21}^* & a_{12}b_{22}^* \\ a_{21}b_{11}^* & a_{21}b_{12}^* & a_{22}b_{11}^* & a_{22}b_{12}^* \\ a_{21}b_{21}^* & a_{21}b_{22}^* & a_{22}b_{21}^* & a_{22}b_{22}^* \end{pmatrix}$$

- Similarly, for row vectors, we have:

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \otimes \begin{pmatrix} b_1^* \\ b_2^* \end{pmatrix} = \begin{pmatrix} a_1 b_1^* \\ a_1 b_2^* \\ a_2 b_1^* \\ a_2 b_2^* \end{pmatrix}$$

When applied to our simple model:

- We have $R = V_1 \otimes V_2^* = (J_{ant1} E_1) \otimes (J_{ant2}^* E_2^*)$

- This is, from a property of outer products:

$$R = (J_{G1} \otimes J_{G2}^*)(J_{pol1} \otimes J_{pol2}^*)(J_{\Psi 1} \otimes J_{\Psi 2}^*)(E_1 \otimes E_2^*)$$

- Which I write as:

$$\mathbf{R} = \mathbf{G} \times \mathbf{P} \times \text{Ø} \times \mathbf{S}$$

Where **R** = the response vector – the correlator output.

G = the gain matrix – effect of post-polarizer amplifiers

P = the polarization mixing matrix (Mueller matrix)

Ψ = the antenna rotation matrix (can include propagation)

S = the Stokes vector – what we want.

The various terms are:

- Response Vector, R:

$$\mathbf{R} = \begin{pmatrix} \langle V_{R1} V_{R2}^{*} \rangle \\ \langle V_{R1} V_{L2}^{*} \rangle \\ \langle V_{L1} V_{R2}^{*} \rangle \\ \langle V_{L1} V_{L2}^{*} \rangle \end{pmatrix}$$

- Gain Matrix, G:

$$\mathbf{G} = \begin{pmatrix} G_{R1} G_{R2}^{*} & 0 & 0 & 0 \\ 0 & G_{R1} G_{L2}^{*} & 0 & 0 \\ 0 & 0 & G_{L1} G_{R2}^{*} & 0 \\ 0 & 0 & 0 & G_{L1} G_{L2}^{*} \end{pmatrix}$$

- Polarization Matrix, P:

$$\mathbf{P} = \begin{pmatrix} C_{RR1} C_{RR2}^{*} & C_{RR1} S_{LR2}^{*} & S_{LR1} C_{RR2}^{*} & S_{LR1} S_{LR2}^{*} \\ C_{RR1} S_{RL2}^{*} & C_{RR1} C_{LL2}^{*} & S_{LR1} S_{RL2}^{*} & S_{LR1} C_{LL2}^{*} \\ S_{RR1} C_{RR2}^{*} & S_{RR1} S_{LR2}^{*} & C_{LL1} C_{RR2}^{*} & C_{LL1} S_{LR2}^{*} \\ S_{RR1} S_{RL2}^{*} & S_{RR1} C_{LL2}^{*} & C_{LL1} S_{RL2}^{*} & C_{LL1} C_{LL2}^{*} \end{pmatrix}$$

Terms, continued …

- Rotation Matrix, Ψ:

$$\text{Ø} = \begin{pmatrix} e^{-i(\Psi_{R1}-\Psi_{R2})} & 0 & 0 & 0 \\ 0 & e^{-i(\Psi_{R1}+\Psi_{L2})} & 0 & 0 \\ 0 & 0 & e^{i(\Psi_{L1}+\Psi_{R2})} & 0 \\ 0 & 0 & 0 & e^{i(\Psi_{L1}-\Psi_{L2})} \end{pmatrix}$$

- Stokes Vector, **S**:

$$\mathbf{S} = \begin{pmatrix} (\mathrm{I}+\mathrm{V})/2 \\ (\mathrm{Q}+i\mathrm{U})/2 \\ (\mathrm{V}-i\mathrm{U})/2 \\ (\mathrm{I}-\mathrm{V})/2 \end{pmatrix}$$

- <Whew!> Almost there.
- It gets easier from here …

Inverting the Polarization Equation

- We have, for the relation between the correlator output and the Stokes visibility:

$$\mathbf{R} = \mathbf{G} \times \mathbf{P} \times \mathbf{\O} \times \mathbf{S}$$

- The solution for S is trivial to write:

$$\mathbf{S} = \mathbf{\O}^{-1} \times \mathbf{P}^{-1} \times \mathbf{G}^{-1} \times \mathbf{R}$$

- The inverses for the rotation and gain matrices are trivial.
- More interesting is $\mathbf{P}^{-1}$:

$$\mathbf{P}^{-1} = K \begin{pmatrix} C_{LL1}C^*_{LL2} & -C_{LL1}S^*_{LR2} & -S_{LR1}C^*_{LL2} & S_{LR1}S^*_{LR2} \\ -C_{LL1}S^*_{RL2} & C_{LL1}C^*_{RR2} & S_{LR1}S^*_{RL2} & -S_{LR1}C^*_{RR2} \\ -S_{RL1}C^*_{LL2} & S_{RL1}S^*_{LR2} & C_{RR1}C^*_{LL2} & -C_{RR1}S^*_{LR2} \\ S_{RL1}S^*_{RL2} & -S_{RL1}C^*_{RR2} & -C_{RR1}S^*_{RL2} & C_{RR1}C^*_{RR2} \end{pmatrix}$$

Where K is a normalizing factor: $K = \dfrac{1}{\left(C_{RR1}C_{LL1} - S_{LR1}S_{RL1}\right)\left(C^*_{RR2}C^*_{LL2} - S^*_{LR2}S^*_{RL2}\right)}$

Obtaining the Stokes Visibilities

- All this shows that – in principle – the four complex outputs from an interferometer can be easily inverted to obtain the desired Stokes visibilities.
- Sadly, it's not quite that easy. To correctly invert, we need to know all the factors in the Jones matrices.
- In fact we do not, because …
  - Atmospheric gains are continually changing.
  - System gains change (but hopefully more slowly).
  - Antennas rotate on the sky (but we think we know this in advance …)
  - Antenna polarization may change (but probably very slowly)
  - Standard calibration techniques do not provide the correct values of C and S, but rather values relative to one antenna.

The Physical Meaning …

- To understand the meaning of the C and S terms, consider the antenna in 'transmission' mode.
- One can show (problem for the student!) that the elements in the polarization matrix are determined by the antenna's polarization, with:

$$C_R = \cos\beta_R e^{-i\varphi_R}$$
$$C_L = \cos\beta_L e^{i\varphi_L}$$
$$S_R = \sin\beta_R e^{i\varphi_R}$$
$$S_L = \sin\beta_L e^{-i\varphi_L}$$

$$\beta_R = \chi_R + \pi/4$$
$$\beta_L = \pi/4 - \chi_L$$

  - The β term is the deviation of the antenna polarization ellipse from perfectly circular.
  - The χ term is the antenna's ellipticity
  - The ϕ term is the position angle of the antenna's polarization ellipse, in the antenna frame.
- You can, by substituting the terms above into the polarization matrix, and including the antenna rotation terms, show that:

The response of one of the four correlations:

$$
\begin{aligned}
R_{pq} = G_{pq}\{&[\cos(\Psi_p - \Psi_q)\cos(\chi_p - \chi_q) + i\sin(\Psi_p - \Psi_q)\sin(\chi_p + \chi_q)]\mathrm{I}\;/2 \\
&+[\cos(\Psi_p + \Psi_q)\cos(\chi_p + \chi_q) + i\sin(\Psi_p + \Psi_q)\sin(\chi_p - \chi_q)]\mathrm{Q}/2 \\
&-i[\cos(\Psi_p + \Psi_q)\sin(\chi_p - \chi_q) + i\sin(\Psi_p + \Psi_q)\cos(\chi_p + \chi_q)]\mathrm{U}/2 \\
&-[\cos(\Psi_p - \Psi_q)\sin(\chi_p + \chi_q) + i\sin(\Psi_p - \Psi_q)\cos(\chi_p - \chi_q)]\mathrm{V}/2\}
\end{aligned}
$$

This is the famous expression derived by Morris, Radhakrishnan and Seielstad (1964), relating the output of a single complex correlator to the complex Stokes visibilities, where the antenna effects are described in terms of the polarization ellipses of the two antennas.

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 polarizations p and q.

I,Q, U, and V are the Stokes Visibilities

G_{pq} is a complex gain, including the effects of transmission and electronics

Application: Nearly Perfect Antennas

- I finish up with a description of how to handle imperfectly polarized antennas.
- First consider circularly polarized systems, and assume our engineers can produce polarizers which are 'nearly perfect'.
- Then, the `C' terms are of nearly unit amplitude, and are very steady in time.
- We can then factor them out of the Mueller matrix, and consider them as part of the gain calibration.
- If we define the D-term as: D = C/S, then we a form very familiar to many 'old hands':

Slightly Imperfect Circularly Polarized Antennas

$$\begin{pmatrix} R_{R1R2} \\ R_{R1L2} \\ R_{L1R2} \\ R_{L1L2} \end{pmatrix} = \begin{pmatrix} 1 & D^{*}_{LR2} & D_{LR1} & D_{LR1}D^{*}_{LR2} \\ D^{*}_{RL2} & 1 & D_{LR1}D^{*}_{RL2} & D_{LR1} \\ D_{RL1} & D_{RL1}D^{*}_{LR2} & 1 & D^{*}_{LR2} \\ D_{RL1}D^{*}_{LR2} & D_{RL1} & D^{*}_{RL2} & 1 \end{pmatrix} \begin{pmatrix} (\mathrm{I} + \mathrm{V})/2 \\ e^{-2i\Psi_P}(\mathrm{Q} + i\mathrm{U})/2 \\ e^{2i\Psi_P}(\mathrm{Q} - i\mathrm{U})/2 \\ (\mathrm{I} - \mathrm{V})/2 \end{pmatrix}$$

Where:

$$D_R = \tan\beta_R e^{i2\varphi_R}$$

$$D_L = \tan\beta_L e^{-i2\varphi_L}$$

- If |D|<<1, we can then ignore D*D products.
- Furthermore, as |Q| and |U| << |I|, we can ignore products between them and the Ds.
- And V can be safely assumed to be zero.
- These (very reasonable) approximations then give us:

‘Nearly’ Circular Feeds (small D approximation)

- We get:

$$R_{R1R2} = \mathrm{I}\ /2$$

$$R_{L1L2} = \mathrm{I}\ /2$$

$$R_{R1L2} = \left[\left(D_{R1} + D_{L2}^{*}\right)\mathrm{I}\ + e^{-2i\Psi_P}\left(\mathrm{Q} + i\mathrm{U}\right)\right]/2$$

$$R_{L1R2} = \left[\left(D_{L1} + D_{R2}^{*}\right)\mathrm{I}\ + e^{2i\Psi_P}\left(\mathrm{Q} - i\mathrm{U}\right)\right]/2$$

- Our problem is now clear. The desired cross-hand responses are contaminated by a term proportional to ‘I’.
- Stokes ‘I’ is typically 20 to 100 times the magnitude of ‘Q’ or ‘U’.
- If the ‘D’ terms are of order a few percent (and they are!), we must make allowance for the extra terms.
- To do accurate polarimetry, we must determine these D-terms, and remove their contribution.
- Knowing the D-terms, one can easily modify the Rs to their correct values.

Nearly Perfectly Linear Feeds

- In this case, 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:

$$D_X = \varphi_X - i\chi_X$$

$$D_Y = -\varphi_Y + i\chi_Y$$

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

$$R_{V1V2} = (\mathrm{I} + \mathrm{Q}\cos 2\Psi_P + \mathrm{U}\sin 2\Psi_P)/2$$

$$R_{H1H2} = (\mathrm{I} - \mathrm{Q}\cos 2\Psi_P - \mathrm{U}\sin 2\Psi_P)/2$$

$$R_{V1H2} = [\mathrm{I}\left(D_{V1} + D^*_{H2}\right) - \mathrm{Q}\sin 2\Psi_P + \mathrm{U}\cos 2\Psi_P + \mathrm{iV}]/2$$

$$R_{H1V2} = [\mathrm{I}\left(D_{H1} + D^*_{V2}\right) - \mathrm{Q}\sin 2\Psi_P + \mathrm{U}\cos 2\Psi_P - \mathrm{iV}]/2$$

- The situation is the same as for the circular system.

Measuring Cross-Polarization

- Correction of the X-hand response for the 'leakage' is important, since the leakage amplitude is comparable to the fractional polarization.
- There are two ways to proceed:
  1. Observe a calibrator source of known polarization (preferably zero!)
  2. Observe a calibrator of unknown polarization for a 'long time'.
- First case (with polarization = 0).

$$R_{V1V2} = I/2$$
$$R_{H1H2} = I/2$$
$$R_{V1H2} = I\left(D_{V1} + D^*_{H2}\right)/2$$
$$R_{H1V2} = I\left(D_{H1} + D^*_{V2}\right)/2$$

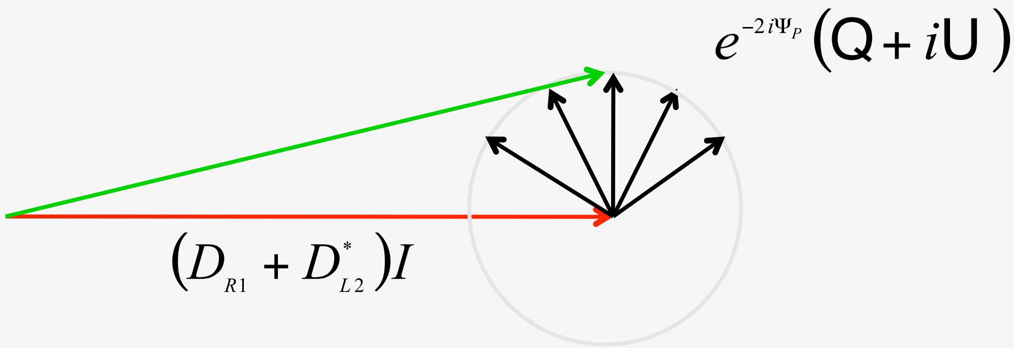
- Then a single observation should suffice to measure the leakage terms.
- This is not actually correct – because the cross-hand visibility is always the sum of two terms, the 'D' values must be referenced to an assumed value (D_{V1} = 0, for example).

Determining Source and Antenna Polarizations

- You can determine both the (relative) D terms and the calibrator polarizations for an alt-az antenna by observing over a wide range of parallactic angle. (Conway and Kronberg invented this)

$$R_{L1R2} = \left[\left(D_{L1} + D^*_{R2} \right) \mathrm{I} + e^{2i\Psi_P} \left(\mathrm{Q} - i\mathrm{U} \right) \right] / 2$$
$$R_{R1L2} = \left[\left(D_{R1} + D^*_{L2} \right) \mathrm{I} + e^{-2i\Psi_P} \left(\mathrm{Q} + i\mathrm{U} \right) \right] / 2$$

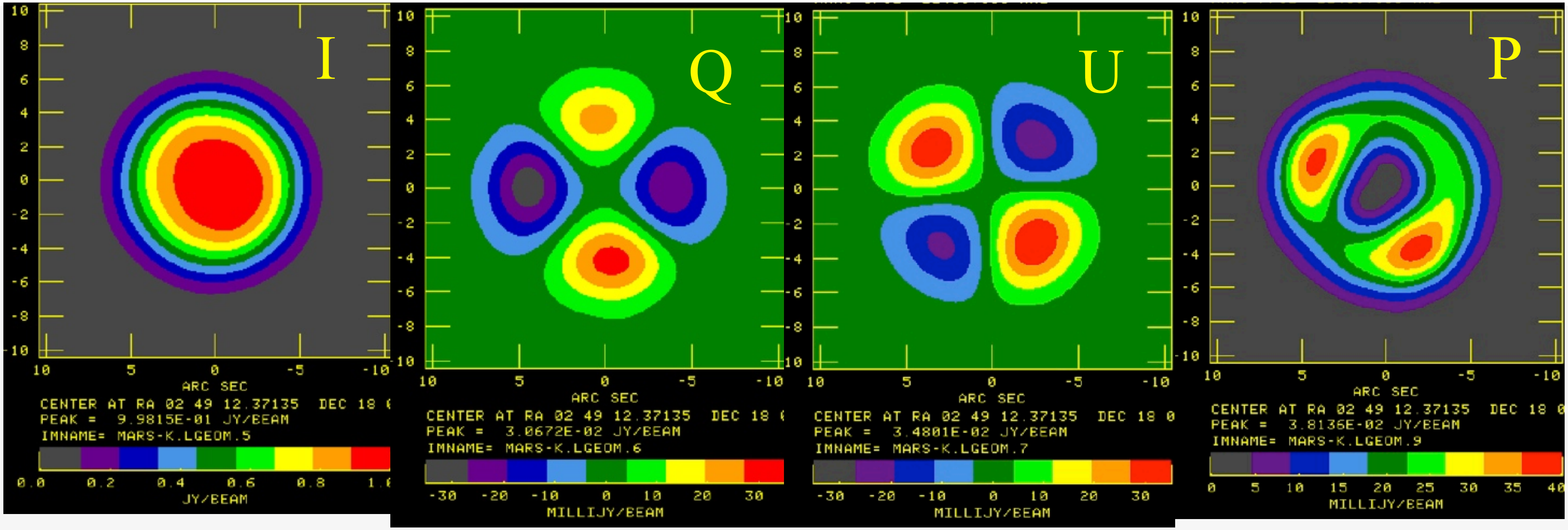
- As time passes, Ψ_P changes in a known way.
- The source polarization term then rotates w.r.t. the antenna leakage term, allowing a separation.



Relative vs. Absolute D terms

- For both linear and circular systems, the standard methodology only provides a ‘relative’ D term.
- This is O.K. for most polarimetry, using the linear approximations employed here to simplify the equations.
- For highly polarized sources, or highly polarized antennas, this methodology will fail.
- Absolute D terms will be needed for accurate polarimetry.
- Obtaining these is not easy – the best method is to rotate one antenna in the array by 90 degrees about an axis pointing to an unpolarized source. (See EVLA Memo 131 for details).
- For VLA, we can physically rotate the feed at some bands.
- ASKAP can rotate the whole antenna upon demand! (Whoever designed this in deserves a star award!).
- With absolute D terms, one can properly invert the full mixing matrix.

Illustrative Example – Thermal Emission from Mars



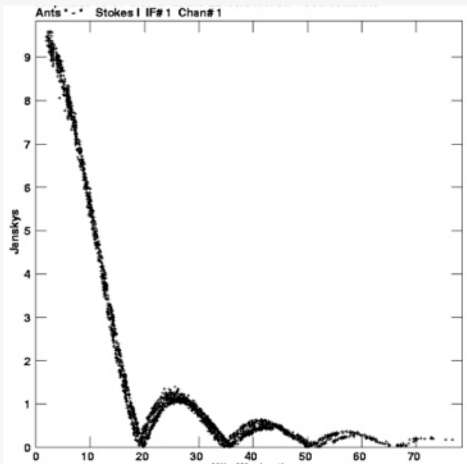
- Mars emits in the radio as a black body.
- Shown are false-color coded I,Q,U,P images from Jan 2006 data at 23.4 GHz.
- V is not shown – all noise – no circular polarization.
- Resolution is 3.5”, Mars’ diameter is ~6”.
- From the Q and U images alone, we can deduce the polarization is radial, around the limb.
- Position Angle image not usefully viewed in color.

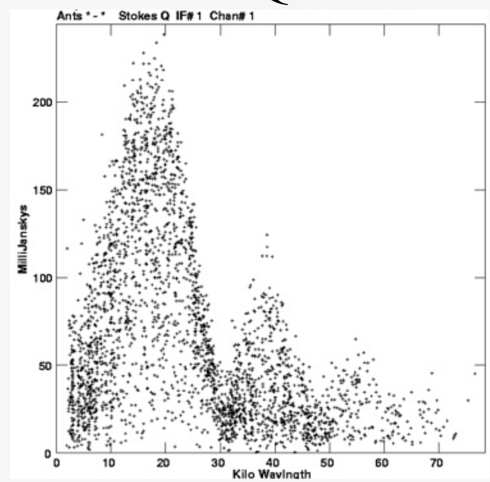
I,Q,U,V Visibilities

- It's useful to look at the visibilities which made these images.

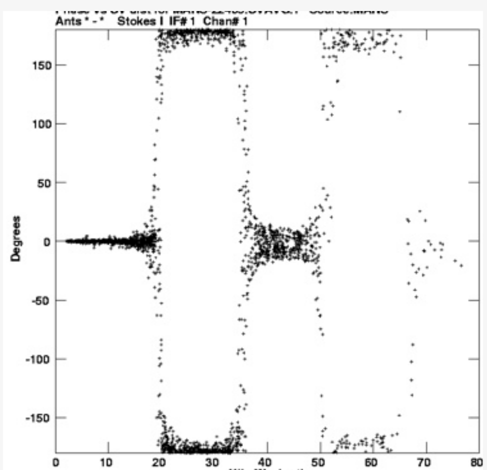
I

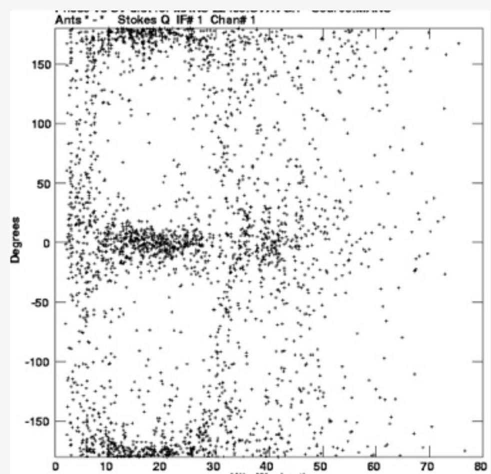
Q





Amplitude

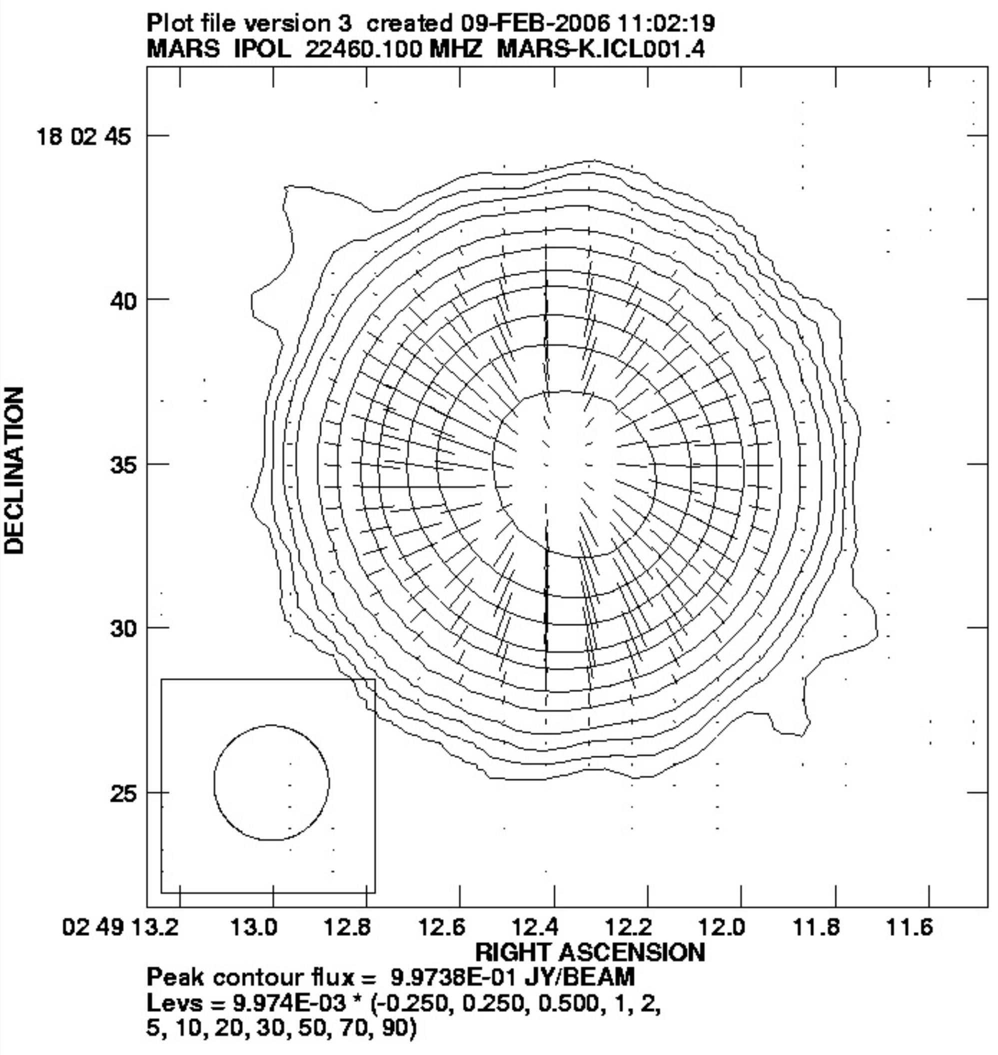




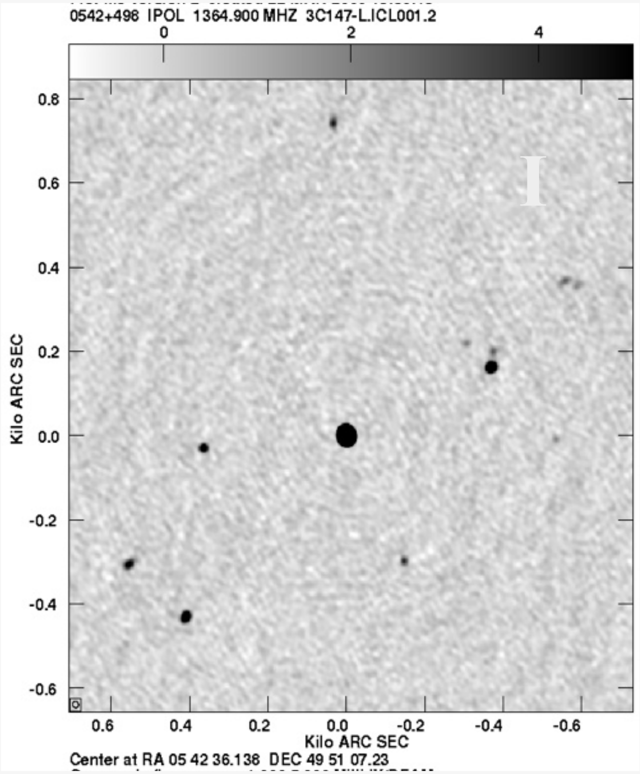
Phase

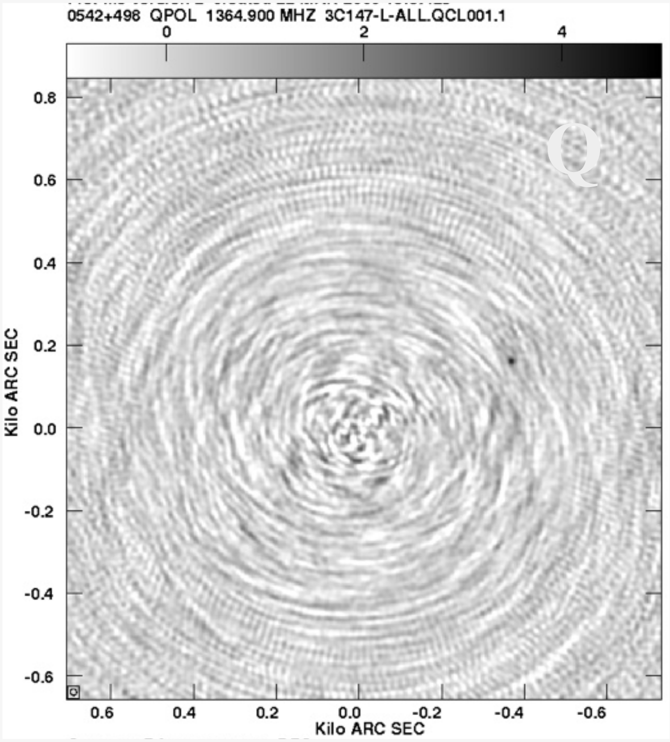
Mars – A Traditional Representation

- Here, I, Q, and U are combined to make a more realizable map of the total and linearly polarized emission from Mars.
- The dashes show the direction of the E-field.
- The dash length is proportional to the polarized intensity.
- One could add the V components, to show little ellipses to represent the polarization at every point.



How Well Does This Work?
3C147, a strong unpolarized source …





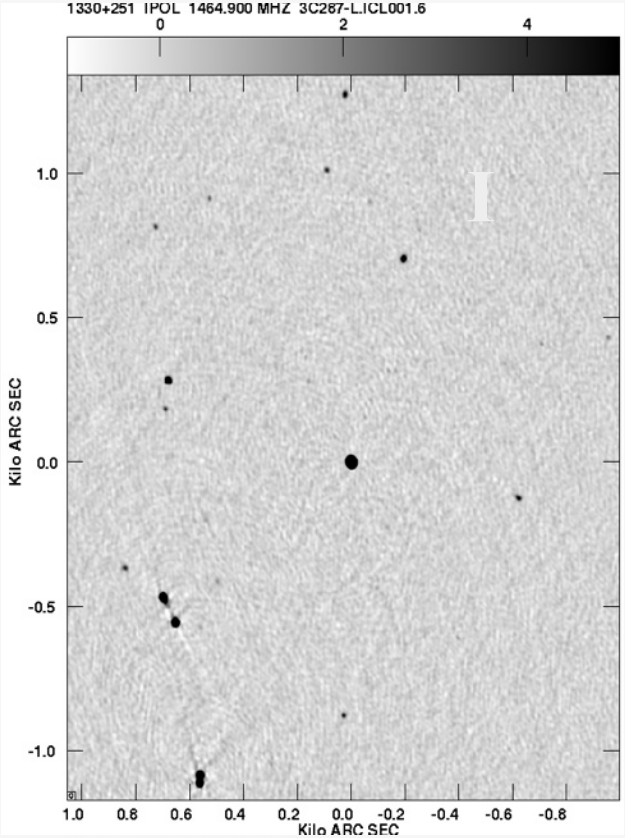
Peak = 21241 mJy, σ = 0.21 mJy

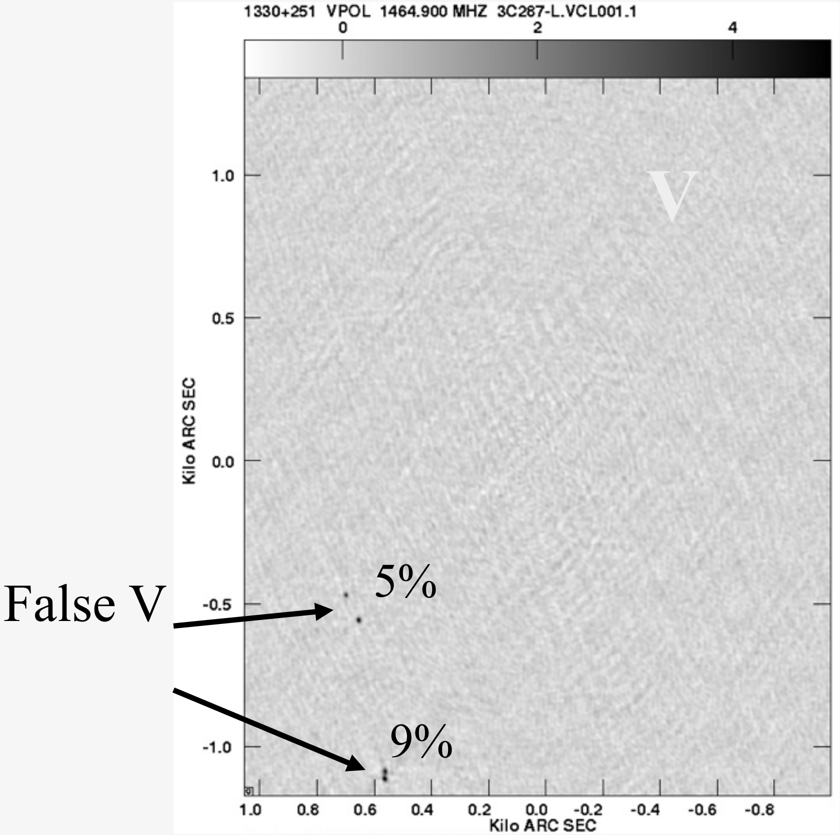
Max background object = 24 mJy

Peak = 4 mJy, σ = 0.8 mJy

Peak at 0.02% level – but not noise limited!

3C287 at 1465 MHZ
I and V with the VLA





Peak = 6982 mJy, σ = 0.21 mJy
Max Bckg. Obj. = 87 mJy

Peak = 6 mJy, σ = 0.16 mJy
Background sources falsely polarized.

A Summary

- Polarimetry is a little complicated.
- But, the polarized state of the radiation gives valuable information into the physics of the emission.
- Well designed systems are stable, and have low cross-polarization, making correction relatively straightforward.
- Such systems easily allow estimation of polarization to an accuracy of 1 part in 10000.
- Beam-induced polarization can be corrected in software – development is under way.