

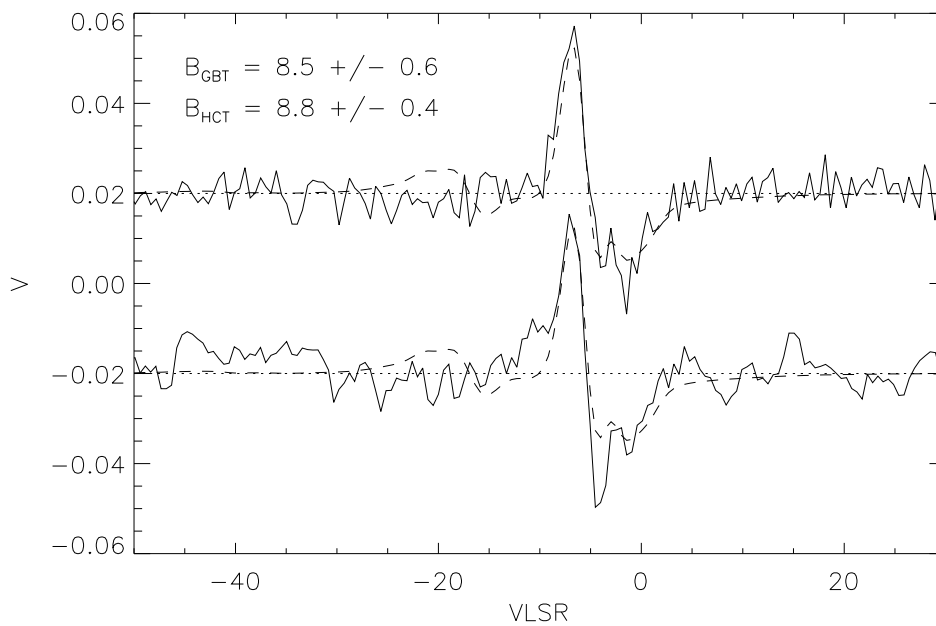
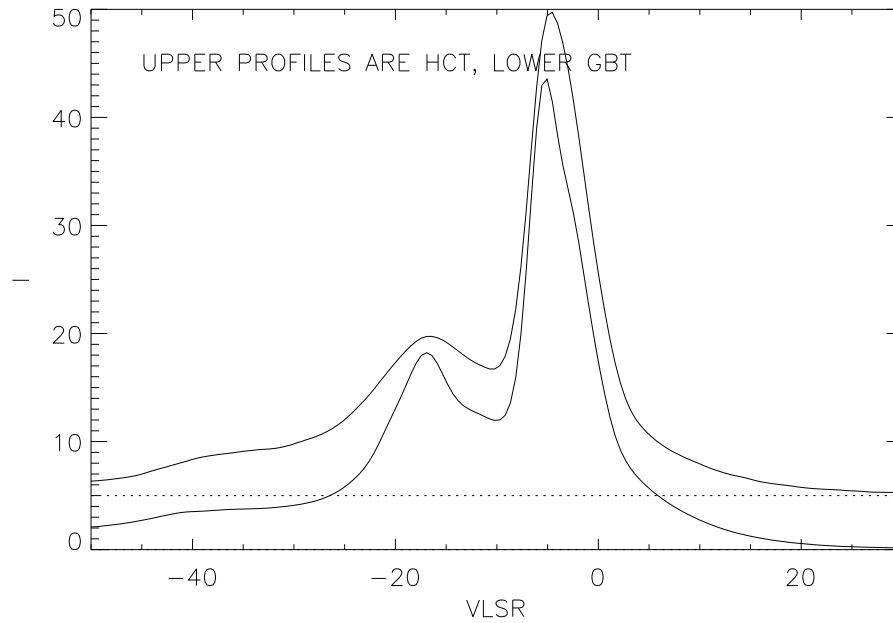
- 1 -

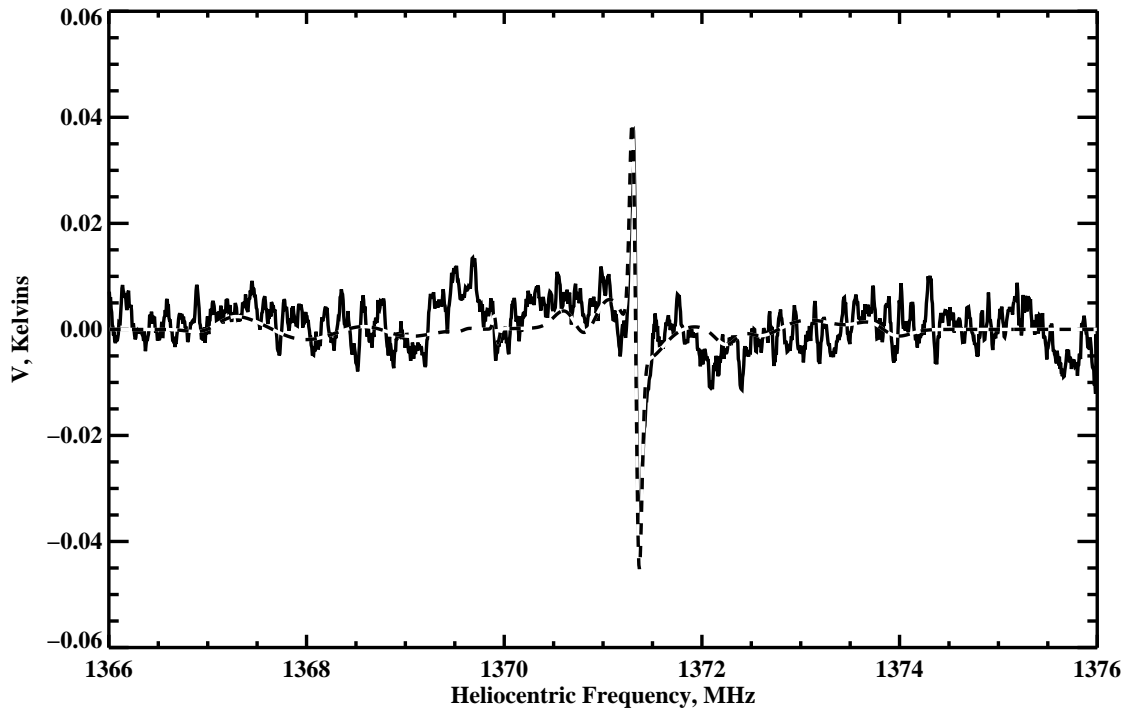
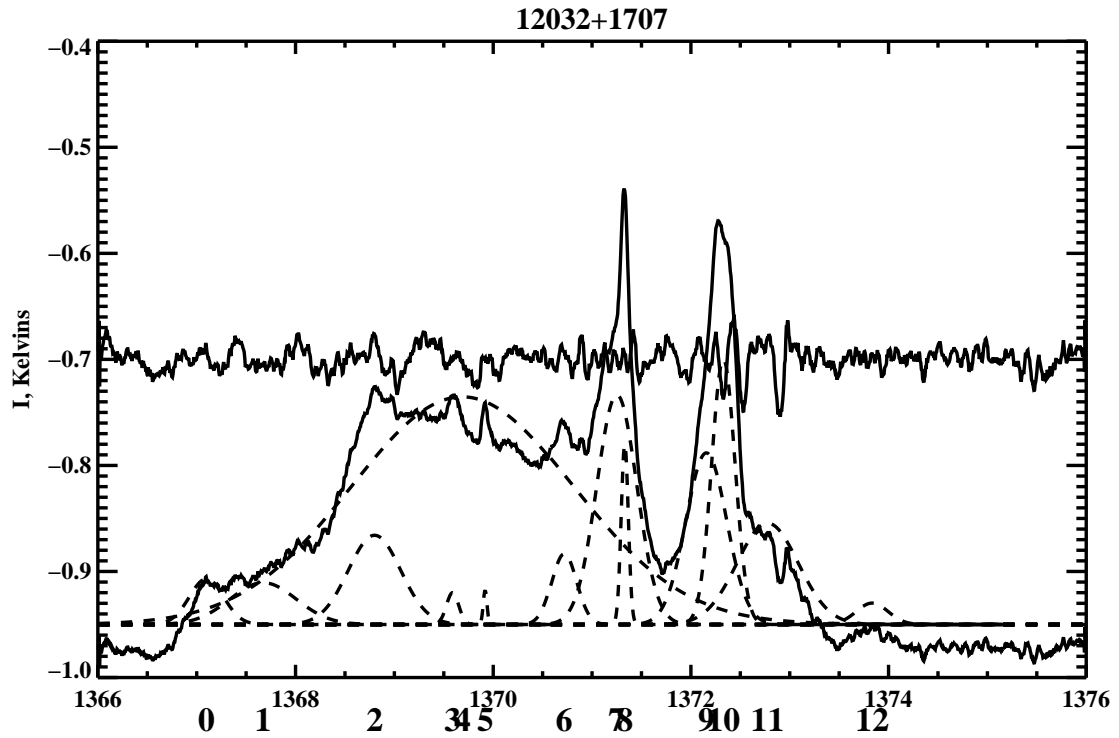
A HEURISTIC INTRODUCTION TO RADIOASTRONOMICAL POLARIZATION

CARL HEILES

Astronomy Department, UC Berkeley

POLARIZATION IS UNIQUE!!





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

J.P. Hamaker and J.D. Bregman

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

Received September 26; accepted October 22, 1995

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

Key words: methods: analytical — methods: data analysis — techniques: interferometers — techniques: polarimeters — polarization

1. Introduction

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

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

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

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

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

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

$$\begin{aligned} I &= \langle |e_x|^2 + |e_y|^2 \rangle \\ Q &= \langle |e_x|^2 - |e_y|^2 \rangle \\ U &= 2 \langle |e_x||e_y| \cos \delta \rangle \\ V &= 2 \langle |e_x||e_y| \sin \delta \rangle \end{aligned} \quad (1)$$

$$I = \langle |e_x|^2 + |e_y|^2 \rangle$$

$$Q = \langle |e_x|^2 - |e_y|^2 \rangle$$

$$U = 2 \langle |e_x| |e_y| \cos \delta \rangle$$

$$V = 2 \langle |e_x| |e_y| \sin \delta \rangle$$

POLARIZATION IS COMPLICATED ??

It's done using STOKES PARAMETERS.

There are four. They need to be measured and calibrated. Their official definitions look *awfully* complicated...

but it's not *that* complicated!

STOKES PARAMETERS: BASICS

Stokes parameters are linear combinations of power measured in *orthogonal polarizations*.

$$I = E_X^2 + E_Y^2 = E_{0^\circ}^2 + E_{90^\circ}^2$$

$$Q = E_X^2 - E_Y^2 = E_{0^\circ}^2 - E_{90^\circ}^2$$

$$U = E_{45^\circ}^2 - E_{-45^\circ}^2$$

$$V = E_{LCP}^2 - E_{RCP}^2$$

The first is total intensity. It is the sum of *any two orthogonal polarizations*.

The second two completely specify linear polarization.

The last completely specifies circular polarization.

We like to write the *Stokes vector*

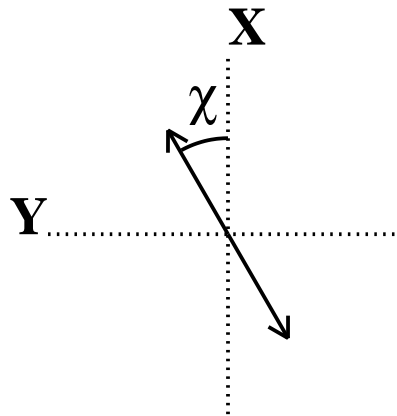
$$\mathbf{S} = \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} .$$

CONVENTIONAL LINEAR POLARIZATION PARAMETERS

AS χ CHANGES, WE HAVE

$$\frac{Q}{I} = p_{QU} \cos(2\chi)$$

$$\frac{U}{I} = p_{QU} \sin(2\chi)$$



POSITION ANGLE OF LINEAR POLARIZA-
TION:

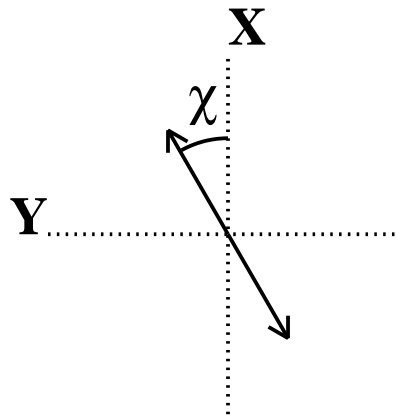
$$\chi = 0.5 \tan^{-1} \frac{U}{Q}$$

FRACTIONAL LINEAR POLARIZATION:

$$p_{QU} = \left[\left(\frac{Q}{I} \right)^2 + \left(\frac{U}{I} \right)^2 \right]^{1/2}$$

HEY!!! LINEAR POLARIZATION “DIRECTION” ??

Look at the figure again:



THERE’S NO ARROWHEAD ON THAT “VECTOR”!! That’s because it’s the angle 2χ , not χ , that’s important.

Moral of this story:

- NEVER say “linear polarization DIRECTION”.
- **INSTEAD**, always say “linear polarization ORIENTATION”.

OTHER CONVENTIONAL POLARIZATION PARAMETERS

FRACTIONAL CIRCULAR POLARIZATION:

$$p_V = \frac{V}{I}$$

TOTAL FRACTIONAL POLARIZATION:

$$p = \left[\left(\frac{Q}{I} \right)^2 + \left(\frac{U}{I} \right)^2 + \left(\frac{V}{I} \right)^2 \right]^{1/2}$$

If both p_{QU} and p_V are nonzero, then the polarization is *elliptical*.

THE (NON) SENSE OF CIRCULAR POLARIZATION

How is Right-hand Circular Polarization
defined?

- If you're a physicist: clockwise as seen by the *receiver*.
- If you're an electrical engineer: the IEEE convention, clockwise as seen by the *transmitter*. Hey!!! what does the receiver see???
- If you're a radio astronomer: the technical roots are in microwave engineering, so it's the IEEE convention. Probably!! You'd better check with your receiver engineers! Or, to be *really* sure, measure it yourself by transmitting a helix from a known vantage point (and remember that V changes sign when the signal reflects from a surface!).
- If you're an optical astronomer: you read it off the label of the camera and you have no idea (your main goal is to get grant money, so getting the science right is too much trouble).

THE (NON) SENSE OF STOKES V

OK...Now that we have RCP straight, how about Stokes V ?

- If you're a physicist: $V = RCP - LCP$. **CAUTION! This might not be correct. Currently looking into this!!**
- If you're an electrical engineer: there's no IEEE convention. Radio astronomers' convention is, historically, from Kraus (e.g. his "ANTENNAS" or his "RADIO ASTRONOMY"): $V = LCP - RCP$. *Hey! With Kraus's definition of V , do physicists and engineers agree???*
- If you're an official of the International Astronomical Union (IAU): The IAU uses the IEEE convention for RCP..., and it defines $V = RCP - LCP$, meaning that, for V , the IAU differs from both the physicist and the Kraus convention!.

IS ALL THIS PERFECTLY CLEAR?

(It's not to me! I thought I understood Stokes V definitions, but my grad student Tim Robishaw tells me I was wrong... so don't have too much faith in anything you see on this page!!)

RADIOASTRONOMICAL FEEDS

Feeds are normally designed to approximate pure linear or circular—known as *native linear* or *native circular*.

Generally speaking, native linear feeds are intrinsically accurate and provide true linear. However, *native circular feeds are less accurate and their exact polarization response is frequency dependent.*

At the *GBT*:

- Feeds below 8 GHz are native linear.
- Feeds above 8 GHz are native circular. For the 8-10 GHz receiver, the response changes from pure circular at 8 GHz to 14% elliptical at 10 GHz.

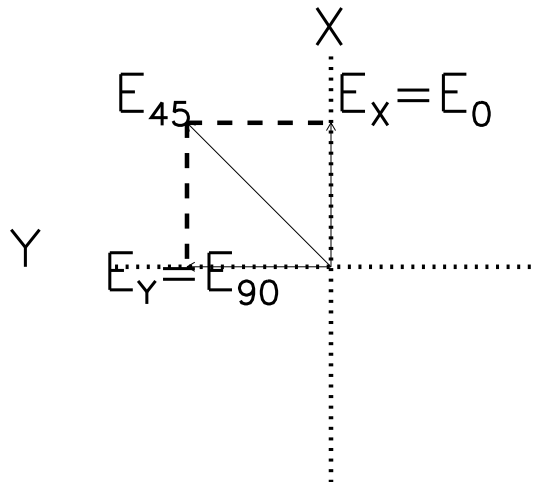
At *ARECIBO*:

- Feeds at 1-2 GHz and 4-6 GHz are native linear. In fact, most feeds are native linear.
- However, a couple are native circular. Circular is achieved with waveguide turnstile junctions, which can be tuned to produce very accurate polarization at the center frequency. However, these are narrow band devices: the feeds become increasingly elliptical, changing to linear and back again over frequency intervals ~ 100 MHz!

**REAL RADIO ASTRONOMERS MEASURE
ALL STOKES PARAMETERS
SIMULTANEOUSLY!**

Extracting two orthogonal polarizations provides *all* the information; you can synthesize *all* other E fields from the two measured ones!

Example: Sample (E_X, E_Y) and synthesize E_{45} from (E_X, E_Y) :



To generate E_{45} , add (E_X, E_Y) with no phase difference.

To generate E_{LCP} , add (E_X, E_Y) with a 90° phase difference.

CARRYING THROUGH THE ALGEBRA FOR THE TWO LINEARS ...

It's clear that

$$E_{45^\circ} = \frac{E_{0^\circ} + E_{90^\circ}}{\sqrt{2}}$$

$$E_{-45^\circ} = \frac{E_{0^\circ} - E_{90^\circ}}{\sqrt{2}}$$

Write the two linear Stokes parameters:

$$Q = E_X^2 - E_Y^2 = E_{0^\circ}^2 - E_{90^\circ}^2$$

$$U = E_{45^\circ}^2 - E_{-45^\circ}^2 = 2E_X E_Y$$

STOKES U IS GIVEN BY THE CROSS-CORRELATION $E_X E_Y$

To get V , throw a 90° phase factor into the correlation.

DOTTING THE I'S AND CROSSING THE T'S GIVES...

Carrying through the algebra and paying attention to complex conjugates and extracting the real part of the expressions yields (for the case of sampling linear polarization (X, Y)):

$$I = E_X \overline{E_X} + E_Y \overline{E_Y} \equiv \mathbf{XX}$$

$$Q = E_X \overline{E_X} - E_Y \overline{E_Y} \equiv \mathbf{YY}$$

$$U = E_X \overline{E_Y} + \overline{E_X} E_Y \equiv \mathbf{XY}$$

$$iV = E_X \overline{E_Y} - \overline{E_X} E_Y \equiv \mathbf{YX}$$

The overbar indicates the complex conjugate. These products are time averages; we have omitted the indicative $\langle \rangle$ brackets to avoid clutter.

IMPORTANT FACT for NATIVE LINEAR FEEDS::

Stokes I and Q are the sum and difference of *self-products*. These *self-products* are large—equal to the full system temperature—so Q is the difference between two large numbers, and is correspondingly inaccurate.

Stokes U and V are sums and differences of *cross products*. These *cross products* are small—in fact, in the absence of polarization they should equal *ZERO!* So...

For small polarizations, *cross products* are *much less subject to error* than *self-products*.

COROLLARY:

To accurately measure small *linear* polarization, use a dual *circular* feed (for which Q and U are cross products); to accurately measure small *circular* polarization, use a dual *linear* feed (for which V is a cross product)..

THE MUELLER MATRIX

The output terminals of a native linear feed provides voltages that sample the E-fields E_X and E_Y ; this it provides directly the Stokes parameters I (from $XX + YY$) and Q (from $XX - YY$).

Rotating the feed by 45° interchanges XX and YY , so it provides directly I and U .

A native circular feed adds a 90° phase to X (or Y) and provides directly the Stokes parameters I (from $XX+YY$) and V (from $XX-YX$). We express these interchanges of power among Stokes parameters with the *Mueller matrix* M .

$$\begin{bmatrix} XX \\ YY \\ XY \\ YX \end{bmatrix} = M \cdot \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} \quad (1)$$

Some examples of Mueller matrices:

(1) A dual linear feed: \mathbf{M} is unitary.

(2). A dual linear feed rotated 45° : Q and U interchange, together with a sign change as befits rotation:

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} . \quad (2)$$

(3). A dual linear feed rotated 90° , which reverses the signs of Q and U :

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} . \quad (3)$$

(4) The above are special instructive cases. As an alt-az telescope tracks a source, the feed rotates on the sky by the *parallactic angle* PA_{az} .

$$\mathbf{M}_{\text{SKY}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2PA_{az} & \sin 2PA_{az} & 0 \\ 0 & -\sin 2PA_{az} & \cos 2PA_{az} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (4)$$

The central 2×2 submatrix is, of course, nothing but a rotation matrix. \mathbf{M}_{SKY} doesn't change I or V .

(5). A dual circular feed, for which $V = \mathbf{X}\mathbf{X} - \mathbf{Y}\mathbf{Y}$:

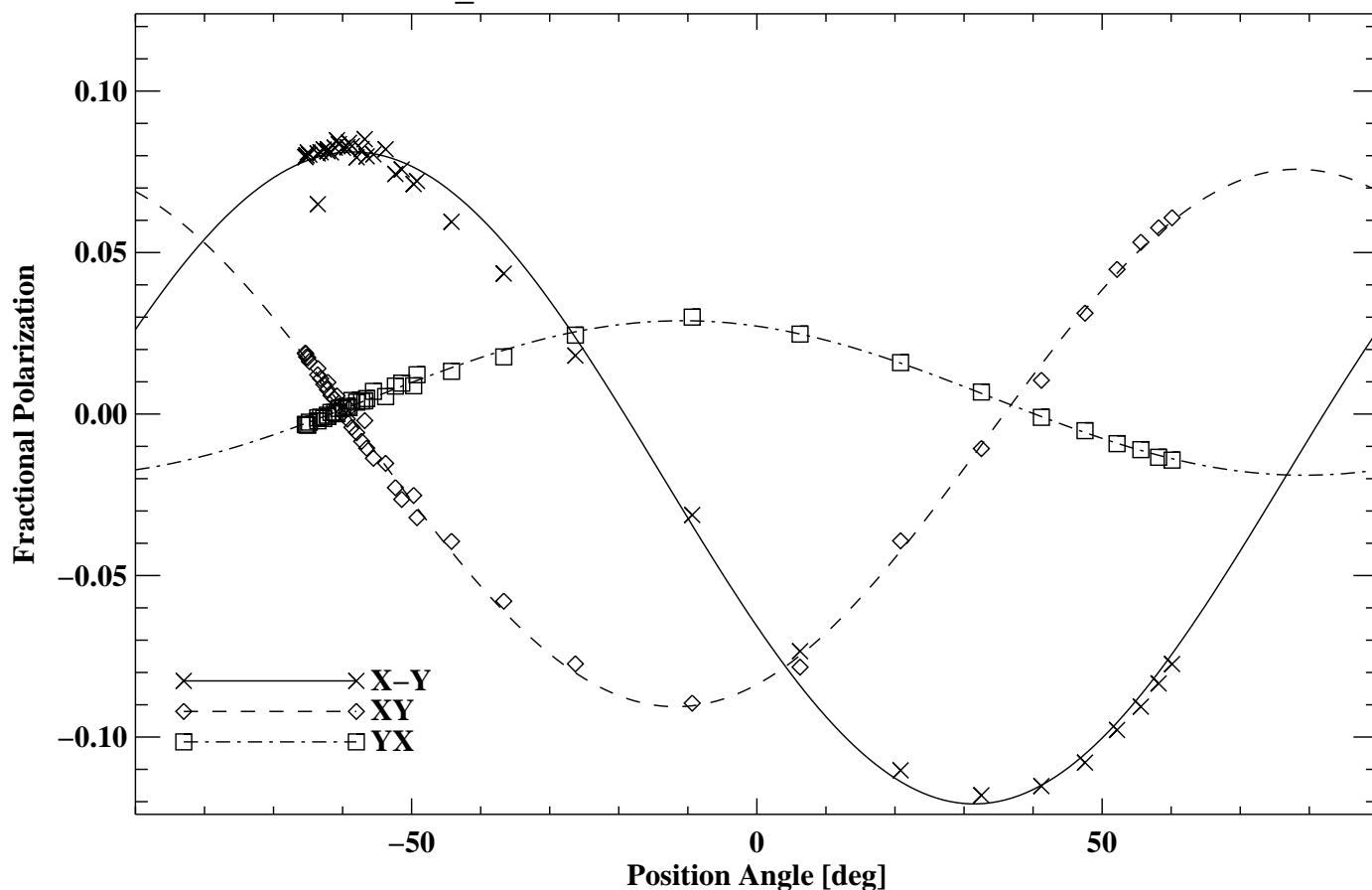
$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}. \quad (5)$$

For a perfect system, as we track a linearly polarized source across the sky the parallactic angle PA changes. This should produce:

- For $XX - YY$, $[\cos 2(PA_{AZ} + PA_{SRC})]$ centered at zero;
- For XY , $[\sin 2(PA_{AZ} + PA_{SRC})]$ centered at zero;
- For YX , zero (most sources have zero circular polarization).

REAL DATA: NATIVE LINEAR POLARIZATION

Rcvr1_2 1420 MHz BRD0 3C286 3-AUG-2003



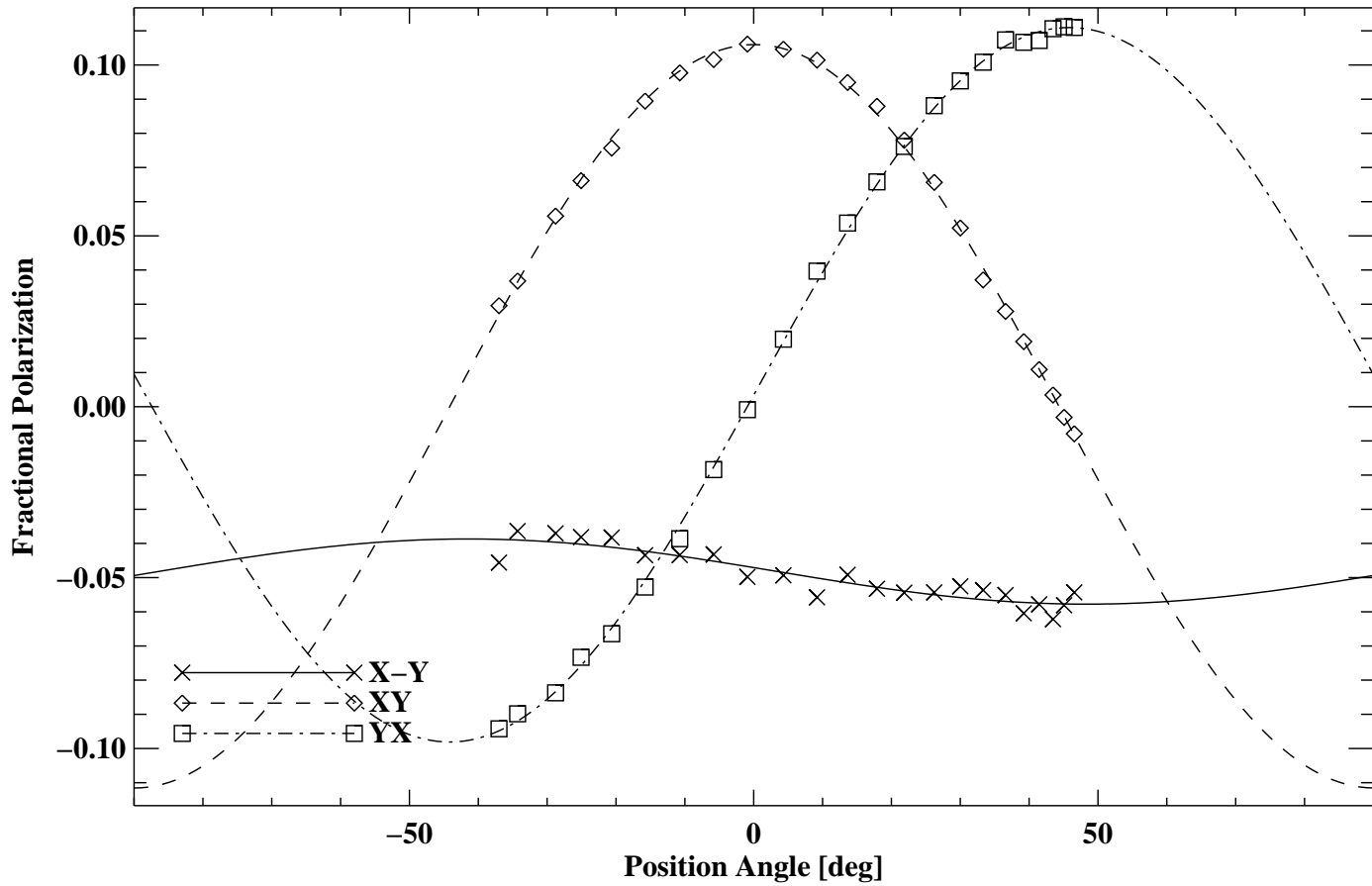
DELTA_G = -0.040 ± 0.016
PSI = -16.0 ± 4.8
ALPHA = +0.2 ± 2.4
EPSILON = 0.004 ± +0.004
PHI = +162.1 ± 50.3
Q_{SRC} = -0.040 ± 0.006
U_{SRC} = -0.085 ± 0.006
POL_{SRC} = +0.094 ± 0.000
PA_{SRC} (UNCORRECTED FOR M_{ASTRO}**) = -57.7 ± 0.0**
NR GOOD POINTS: X-Y = 41 XY = 42 YX = 42 / 42
SCAN 26

Mueller Matrix:

1.0000	-0.0198	-0.0085	0.0026
-0.0198	1.0000	0.0002	0.0075
-0.0074	-0.0021	0.9612	0.2760
0.0050	-0.0073	-0.2760	0.9611

REAL DATA: NATIVE CIRCULAR POLARIZATION

Rcvr8_10 10000 MHz BRD0 SP 3C138 12-JAN-2003



DELTA_G = -0.096 ± 0.004

PSI = +0.0 ± 0.0

ALPHA = -47.6 ± 0.5

EPSILON = +0.004 ± 0.001

PHI = +113.5 ± 18.2

Q_{SRC} = -0.002 ± 0.001

U_{SRC} = +0.107 ± 0.001

POL_{SRC} = +0.107 ± 0.001

PA_{SRC} (UNCORRECTED FOR M_{ASTRO} **) = +45.47 ± 0.38**

NR GOOD POINTS: X-Y = 23 XY = 23 YX = 23 / 23

Mueller Matrix:

1.0000	0.0107	-0.0028	0.0475
-0.0482	-0.0903	0.0001	-0.9959
-0.0028	-0.0000	1.0000	0.0000
0.0064	0.9959	0.0000	-0.0900

THE SINGLE MATRIX FOR THE RADIOASTRONOMICAL RECEIVER

The observing system consists of several distinct elements, each with its own Mueller matrix. The matrix for the whole system is the product of all of them. Matrices are not commutative, so we must be careful with the order of multiplication.

$$\mathbf{M}_{\text{TOT}} = \begin{bmatrix} 1 & (-2\epsilon \sin \phi \sin 2\alpha + \frac{\Delta G}{2} \cos 2\alpha) & 2\epsilon \cos \phi & (2\epsilon \sin \phi \cos 2\alpha + \frac{\Delta G}{2} \sin 2\alpha) \\ \frac{\Delta G}{2} & \cos 2\alpha & 0 & \sin 2\alpha \\ 2\epsilon \cos(\phi + \psi) & \sin 2\alpha \sin \psi & \cos \psi & -\cos 2\alpha \sin \psi \\ 2\epsilon \sin(\phi + \psi) & -\sin 2\alpha \cos \psi & \sin \psi & \cos 2\alpha \cos \psi \end{bmatrix}.$$

NOTE: The Mueller matrix has 16 elements, but *ONLY 7 INDEPENDENT PARAMETERS*. The matrix elements are not all independent.

ΔG is the error in relative intensity calibration of the two polarization channels. It results from an error in the relative cal values (T_{calA}, T_{calB}) .

ψ is the phase difference between the cal and the incoming radiation from the sky (equivalent in spirit to $L_X - L_Y$ on our block diagram..

α is a measure of the voltage ratio of the polarization ellipse produced when the feed observes pure linear polarization.

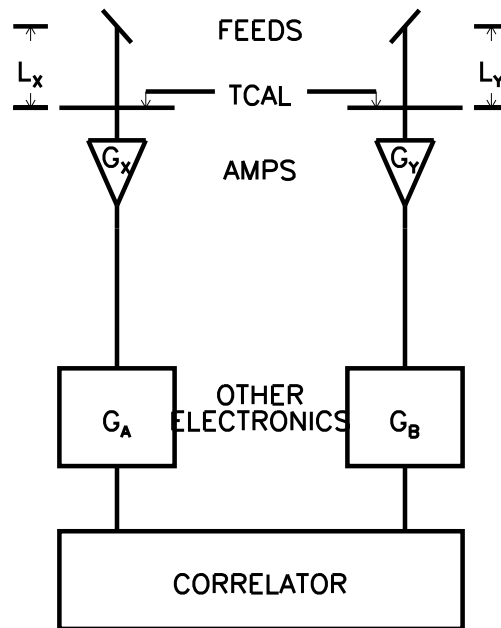
χ is the relative phase of the two voltages specified by α .

ϵ is a measure of imperfection of the feed in producing nonorthogonal polarizations (false correlations) in the two correlated outputs.

ϕ is the phase angle at which the voltage coupling ϵ occurs. It works with ϵ to couple I with (Q, U, V) .

θ_{astron} is the angle by which the derived position angles must be rotated to conform with the conventional astronomical definition.

THE RECEIVER INTRODUCES INSTRUMENTAL POLARIZATION



Cable lengths introduce phase delays. Cables are never identical!

Amplifier gains are *complex*: amplitude *and* phase.

Most amplifiers introduce a 180° phase inversion. If the two channels have different numbers of amplifiers. . .

The receiver introduces mutual coupling among all four Stokes parameters. This coupling is described:

- For the 2-element voltage (Jones) vector, by the 2×2 *Jones Matrix*;
- For the 4-element Stokes vector, by the 4×4 *Mueller Matrix*.

One must correct for this!

THE CAL...OUR CONVENIENT INTENSITY AND PHASE REFERENCE

The signal and cal share common paths after the wavy line.

If we know the relative gains and phase of the cal, we remove the system contributions.

We determine $(G, \phi)_{cal}$ by comparing the cal deflection with the deflection of a known astronomical source (e.g. 3C286). This determines the Mueller matrix parameters.

We assume that the cal remains constant. After returning the to the telescope after an interval of months, it is wise to check!

AN IMPORTANT POINT: THETA VARIES WITH FREQUENCY

The difference in path length ΔL means that the relative phase of E_x and E_y changes with frequency. Thus θ changes with frequency.

$$\frac{d\theta}{df} = \frac{2\pi \Delta L}{c} \approx 0.3 \frac{\text{rad}}{\text{MHz}}$$

This corresponds to

$$\Delta L \approx 20 \text{ m}$$

The consequences:

- You must include the frequency dependence in your least squares fit. This is a bit tricky.
- You cannot do continuum polarization over significant bandwidths without including $\frac{d\theta}{df}$.

**REMEMBER THIS # 1: AVERAGING
LINEAR POLARIZATIONS!!!**

Suppose you average two polarization observations together: Observation 1 has $p = 13.6\%$ and $\chi = 2^\circ$

Observation 2 has $p = 13.7\%$ and $\chi = 178^\circ$

**NOTE THAT THE POSITION ANGLES AGREE
TO WITHIN 4 DEGREES.**

If you average p and χ , you get $p = 13.65\%$ and $\chi = 90^\circ$.

THIS IS INCORRECT!!!!!!!!!!!!

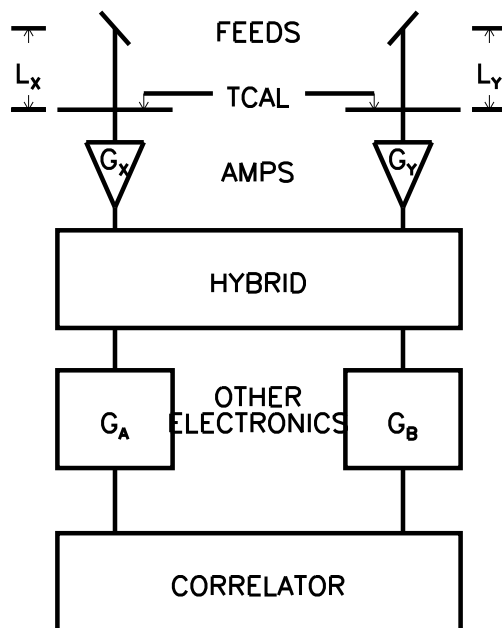
Often you find yourself needing to combine polarizations. For example, if you measure the polarization of some object several times, you need to average the results.

There is only one *proper* way to combine polarizations, and that is to use the Stokes parameters. The reason is simple: because of conservation of energy, powers add and subtract.

What you must always do is convert the fractional polarization and position angle to Stokes parameters, average the Stokes parameters, and convert back to fractional polarization and position angle.

REMEMBER THIS # 2: SHOULD YOU GENERATE CIRCULARS WITH A POST-AMP HYBRID???

Some astronomers believe that source fluxes (that's Stokes I) are better measured with circular polarization. Receiver engineers accommodate them with a post-amp hybrid:

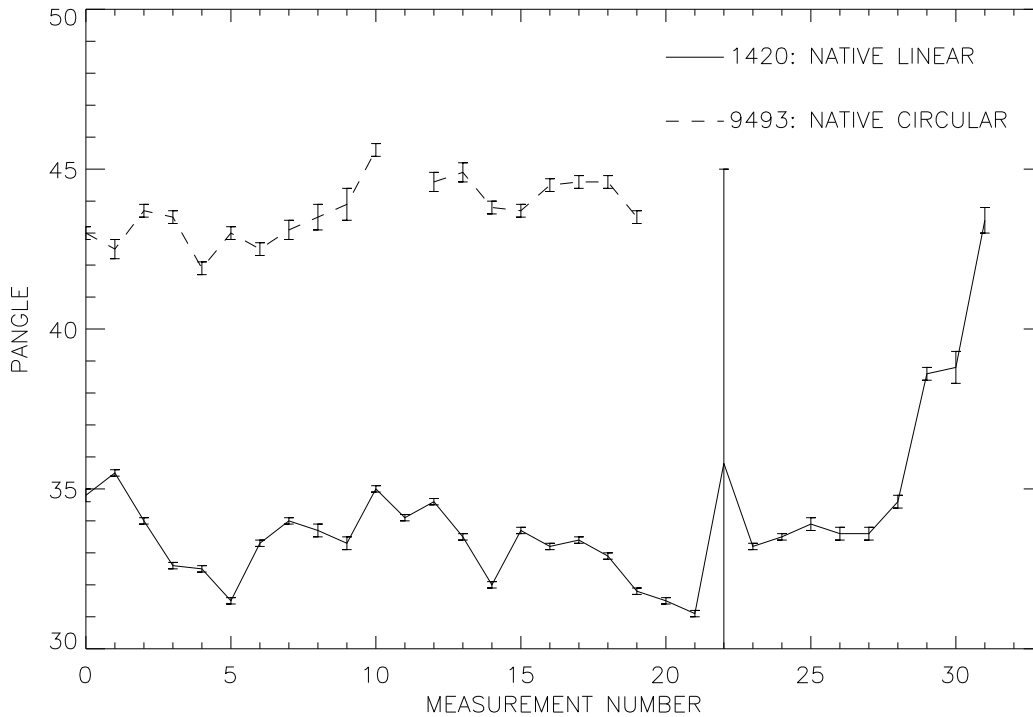


Using this system properly requires a much more complicated calibration procedure.

Example: If $G_Y = 0$, the system still appears to work!!

JUST SAY NO!!!!!!

REMEMBER THIS #3: CROSSCORRELATION VERSUS DIFFERENCING!!!



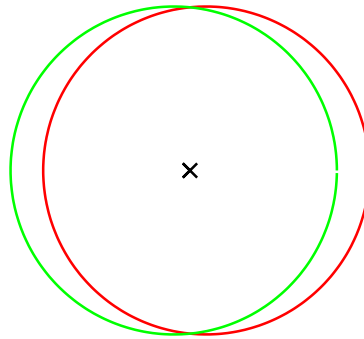
Look at the wild fluctuations for the 1420 MHz position angle measurement (native linear— differencing). This *NEVER* happens with cross-correlation (9495 MHz). Crosscorrelation is also less noisy (At 1420 MHz 3C286 is $\sim 1.5T_{sys}$; at 9493 MHz 3C286 is $\sim 0.5T_{sys}$).

POLARIZED BEAM EFFECTS: BEAM SQUINT

BEAM SQUINT

LHC

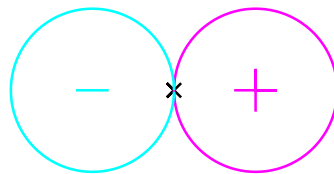
RHC



$$V = \text{LHC} - \text{RHC}$$

$$V > 0$$

$$V < 0$$

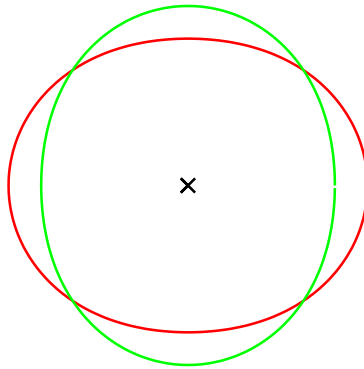


POLARIZED BEAM EFFECTS: BEAM SQUASH

BEAM SQUASH

LHC

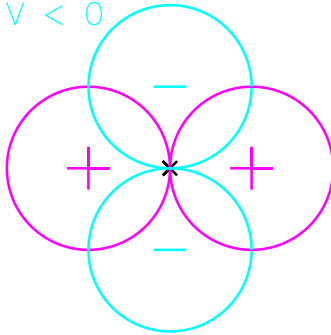
RHC



$$V = \text{LHC} - \text{RHC}$$

$$V > 0$$

$$V < 0$$



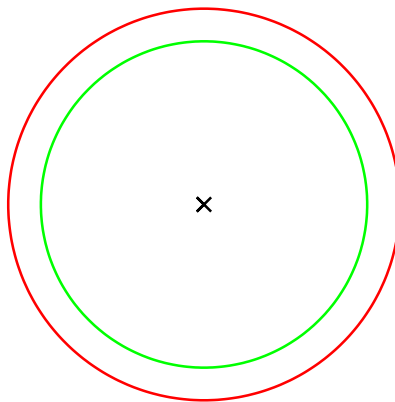
POLARIZED BEAM EFFECTS: BEAM SQUOOSH

(To be honest, we've seen tentative evidence for squoosh, and only then at the GBT)

BEAM SQUOOSH

LINEAR X

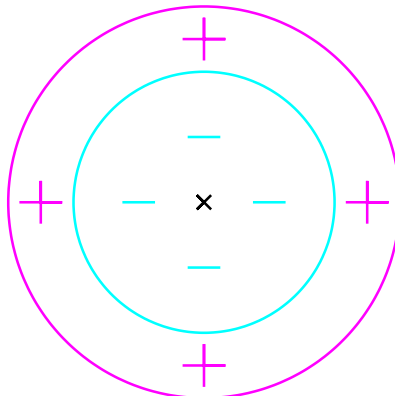
LINEAR Y

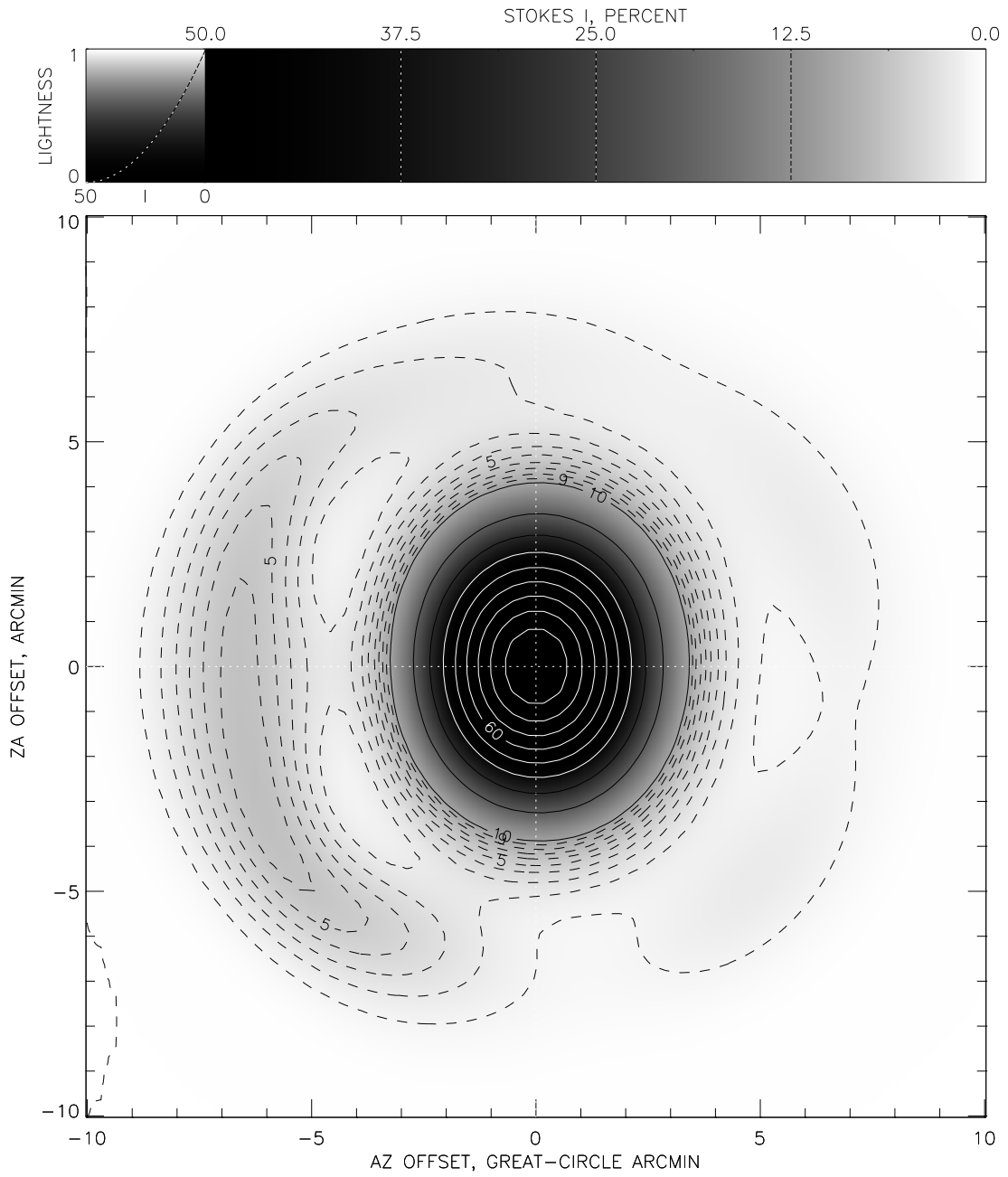


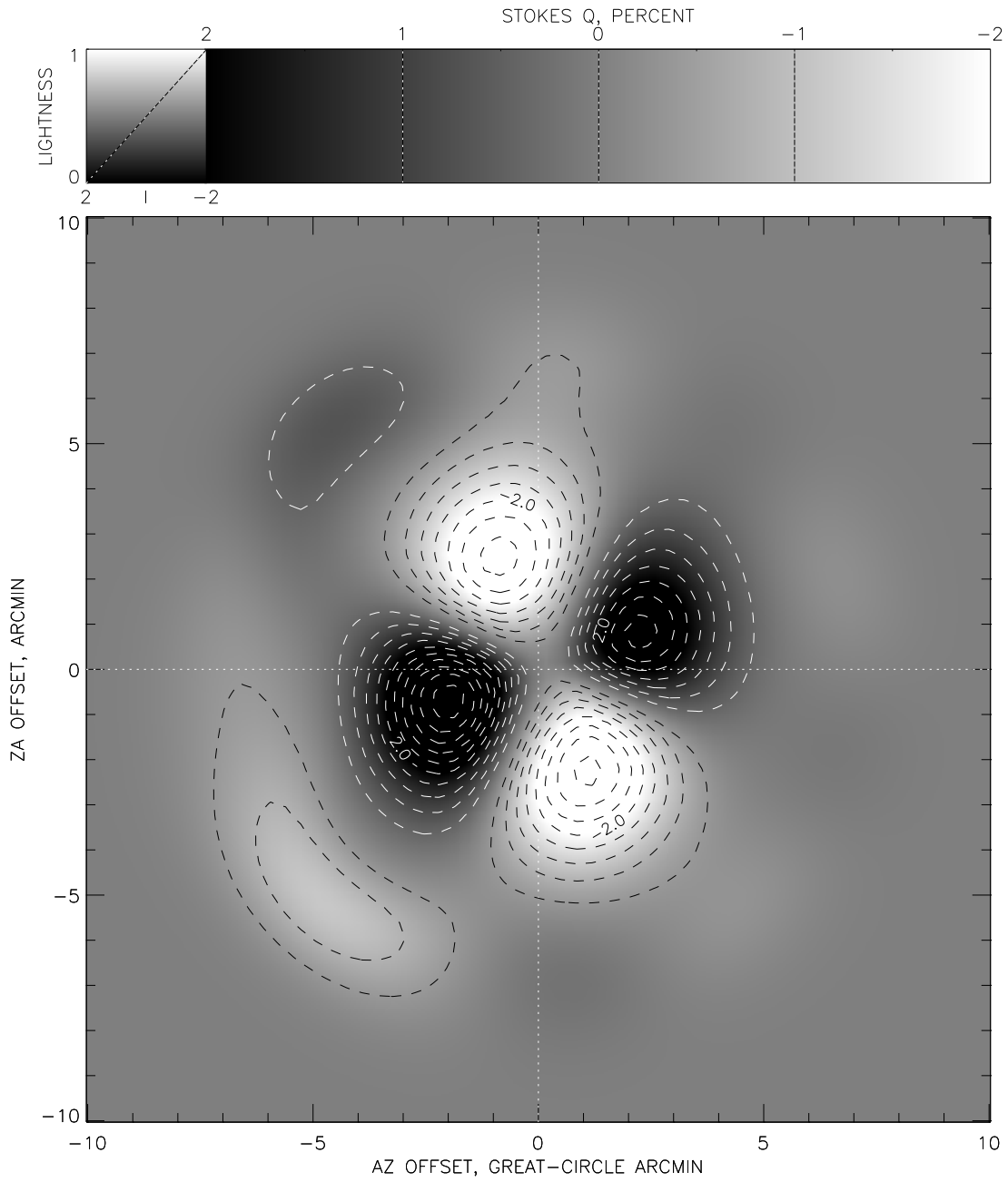
$$Q = X - Y$$

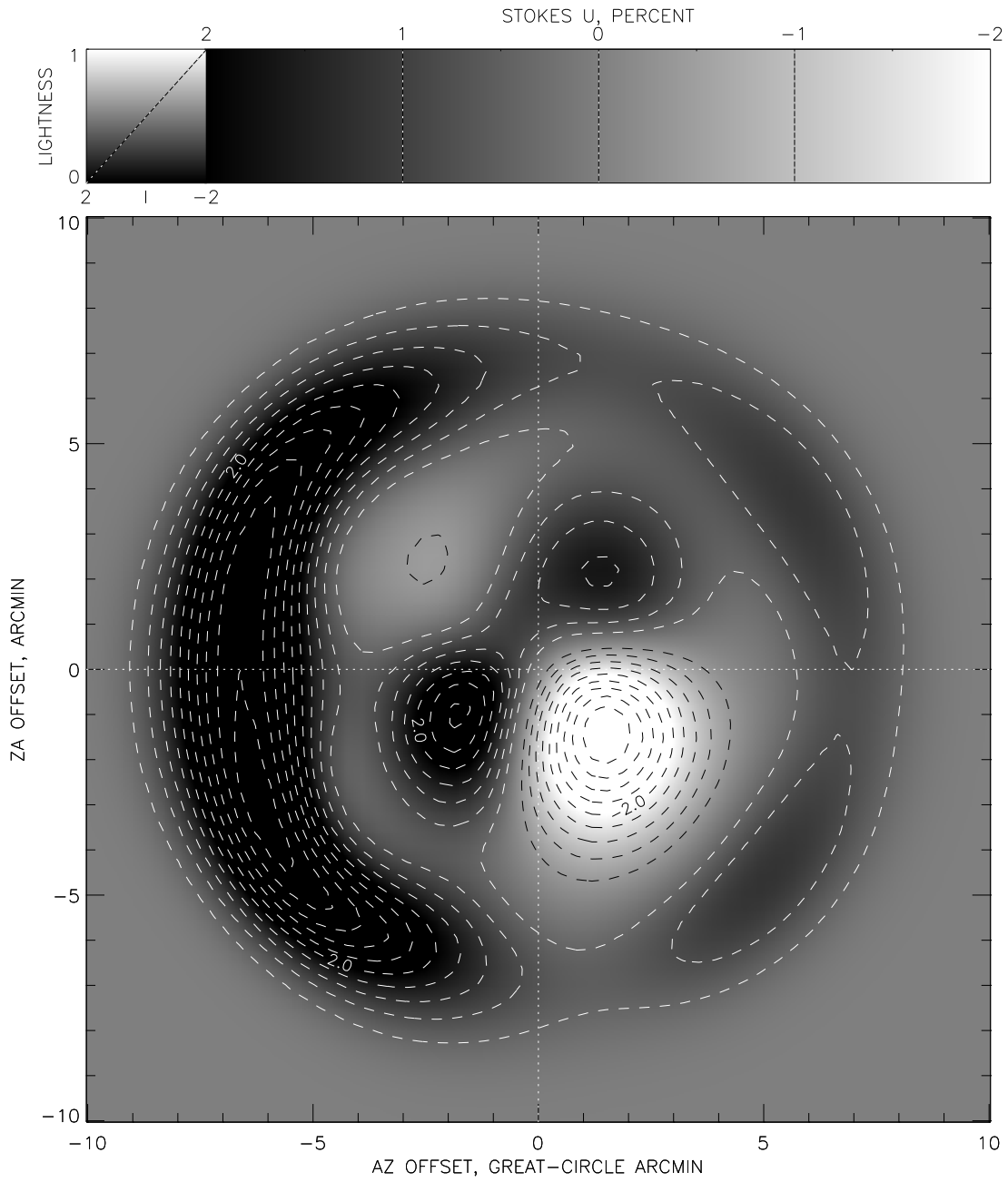
Q > 0

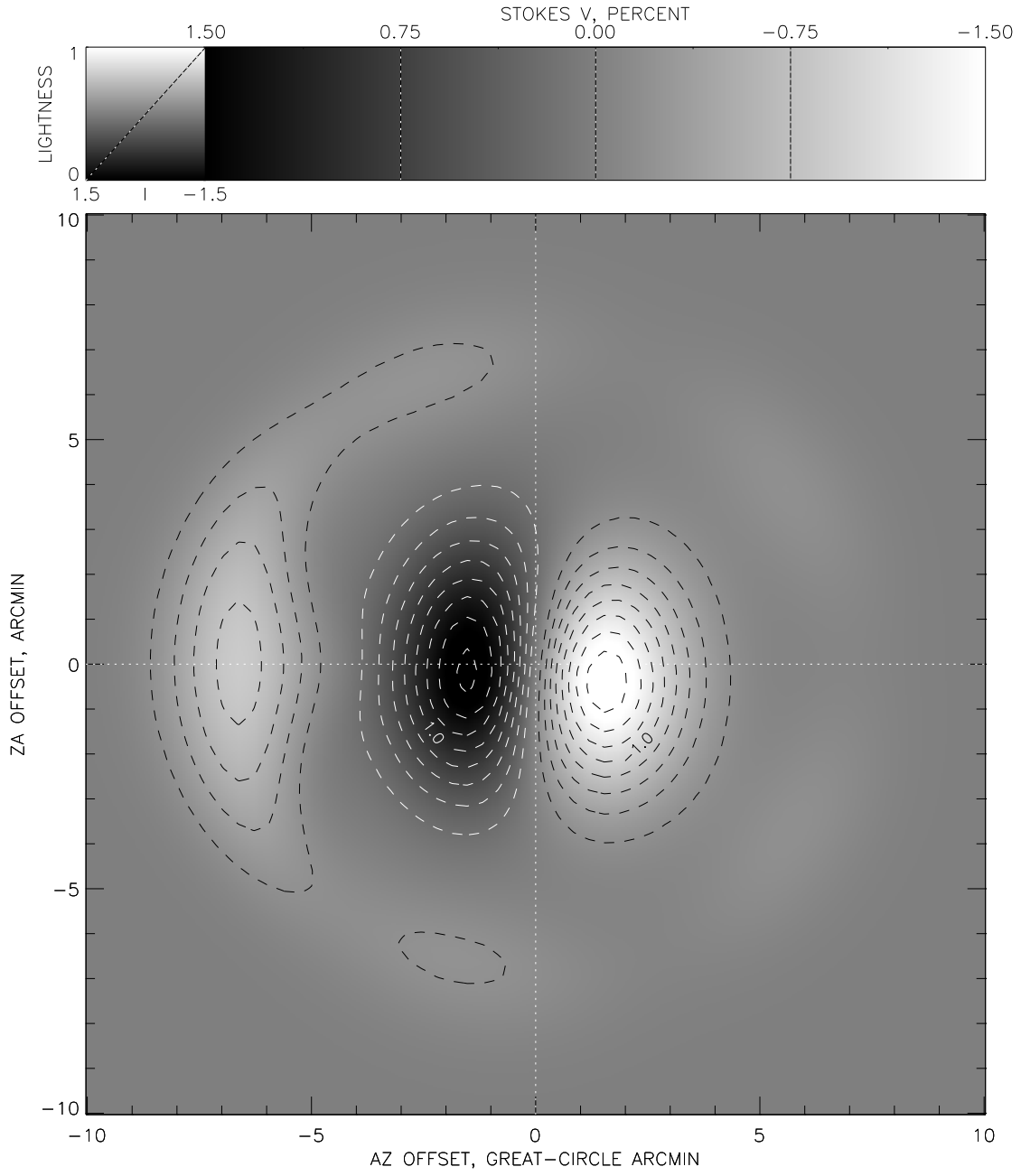
Q < 0











THE EFFECT ON ASTRONOMICAL POLARIZATION MEASUREMENTS

Large-scale features have spatial structure of *Stokes I*.

Express this variation by a two-dimensional Taylor expansion. *Beam squint* responds to the *first derivative*; *beam squash* and *beam squoos* respond to the *second*.

The polarized beam structure interacts with the Stokes *I* derivatives to produce **FAKE RESULTS** in the *polarized* Stokes parameters (Q, U, V). The effects are **exacerbated by the polarized sidelobes**, which are further from beam center.

Stokes *I* derivatives of 1 K arcmin^{-1} and second derivative 1 K arcmin^{-2} (values which are not necessarily realistic) yield fake results for Stokes $Q, U \sim 0.3 \text{ K}$. For Stokes V the contributions are about ten times smaller, $\sim 0.03 \text{ K}$.

The fractional polarization of extended emission tends to be small, so spatial gradients in I can be very serious. For example, if the central velocity of the 21-cm line has a spatial gradient $\frac{dv}{d\theta} = 1 \text{ km s}^{-1} \text{ deg}^{-1}$ we get $B_{fake} \sim 1.1 \mu\text{G}$.

Correcting for these effects at Arecibo is a complicated business because of the PA variation with azimuth and zenith angle. It is also an uncertain business, especially for (Q, U) and somewhat less so for V , because these variations are unpredictable and must be determined empirically.

It's difficult at the GBT, too. Contrary to popular opinion, the GBT is not sidelobe-free. We're still working on it...