ALL-STOKES SINGLE DISH DATA with the RHSTK_2021 (Robishaw/Heiles SToKes) SOFTWARE PACKAGE February 20, 2021

Carl Heiles[1], Tim Robishaw[2]

[1] *Department of Astronomy, UC Berkeley, B-20 Hearst Field Annex, Berkeley, CA 94720, USA; heiles@astro.berkeley.edu*

[2] *National Research Council, Herzberg Institute of Astrophysics, DRAO, Penticton, BC V2A 6J9, Canada; tim.robishaw@nrc-cnrc.gc.ca*

February 20, 2021

ABSTRACT

Major changes from previous version:
change name of package to RHSTK_2021.
TEX file resides in .../dzd4/heiles/rhstk_2021/docs/rhstk/rhstk_4.0.tex
Include pol calibrator source lists in Tables 1 and 2
fitting for Mueller matrix params for 3 types of telescope: traditional alt/az, equatorial (eg GALT) and weird (FAST)
3 versions of Mueller fitting programs for the 3 telescope types
define 2 stages of calibration, Stage 1 and Stage 2
match notation in IDL procs to notation in this document
use onoffs_to_pacoeffs instead of stripfit_to_coeffs
better discussion of example in section 11.

==========

We discuss practical aspects of all-Stokes spectropolarimetric calibration for single-dish radio telescopes using digital spectrometers that produce auto- and cross-correlation products. We describe our polarization-related software package, which we call Robishaw/Heiles SToKes (RHSTK_2021), and provide an example for reducing position-switched data and deriving Mueller matrices.

We begin with the basics—theory and practice—in §describe the quagmire of standards and conventions for polarization, and even magnetic field direction, in §2. We then attempt to describe how to use RHSTK_2021 for determining the Mueller matrix coefficients and using them to determine source polarization properties.

Contents

1. ABOUT RHSTK_2021

1.1. History and Philosophy

Our Robishaw-Heiles SToKes (RHSTK_2021) software, written in IDL, originates from work with the NRAO 140-foot and, particularly, Arecibo in the late 1990's and early 2000's. A few years

ago we began to organize and generalize this RHSTK_2021 restructured version, which is generally applicable and straightforwardly maintainable. 'Generally applicable' means that it can deal with all kinds of telescope mount—alt/az, equatorial, FAST—and inputs to procedures are variables and spectra whose format is simple, not tied to any particular telescope, so that the data can be manipulated by standard IDL commands. Earlier versions of RHSTK included more extensive discussions of Arecibo and the GBT; we have eliminated those in RHSTK_2021 because Arecibo no longer exists and the hardware at GBT has totally changed. The RHSTK_2021 software is applicable at any single-dish radio telescope.

1.2. Continuum Observers!!

This memo is written with multichannel spectra in mind. Nevertheless, *even if you are only interested in the continuum*, you must perform the spectral reduction and add up all the spectral points *after* the calibration. You must *not* add them all up beforehand. The reason is the phase change with frequency discussed in §4.3, which is produced by, primarily, the unequal cable lengths between the feed and the spectrometer input. You must correct for this before summing the cross-spectral points.

1.3. ¡¡¡Important Comment on Mueller Correcting for Native Linear Polarization!!!

For native linear polarization, the measured Stokes $Q_{\rm meas} = XX - YY$ and is the difference between two large numbers. The XX and YY spectra have independent gain calibrations which are not perfect, for three reasons: (1) the DIODEON/DIODEOFF deflection has some noise; (2) the assumed Diode values are not perfect, and the actual ones certainly depend on frequency and might even depend on time; and (3) the X and Y receiver gains almost certainly change a little with time while observing after the intensity calibration was performed. Thus, for native linear polarization, the $Q_{\rm meas}$ *is not accurate*. However, the $U_{\rm meas}$ *is* accurate because it is determined by the cross-product XY, for which gain uncertainties and fluctuations are unimportant.

When you correct for parallactic angle by applying $\mathbf{M}_{\rho\mathbf{x}}$ as in equation 15, you apply a rotation matrix to obtain the *source* Stokes parameters $(Q_{\rm tel}, U_{\rm tel})$ from the *measured* ones $(Q_{\rm meas}, U_{\rm meas})$. This means that *both* $Q_{\rm tel}$ and $U_{\rm tel}$ contain the *inaccurately measured* $Q_{\rm meas}$. Accordingly, when Mueller-correcting native linear polarizations, you might well do better by *not* correcting for parallactic angle. You can then do a least-squares fit of the parallactic angle variation of the accurately-measured $U_{\rm meas}$ using equation 26.

1.4. ¡¡¡Important Comment on Observing polarization of extended sources!!!

Polarization measurements of extended sources are greatly complicated by the telescope sidelobes, which tend to be strongly polarized. Even for an unpolarized source, the angular structure of the source is sampled by the angular structure of the polarized sidelobes and this can produce a spurious observed polarization. This is particularly serious for 21-cm line emission, where these instrumental effects are usually the limiting factor in deriving the weak polarization produced by Zeeman splitting.

The GBT is a particularly difficult case. At the GBT, the *L*-band system (1-2 GHz) uses the secondary focus. There are severe sidelobes, particularly from spillover around the secondary reflector, and because the Galactic 21-cm line exists everywhere on the sky, these sidelobes always see it. This contaminates the measured 21-cm line spectra. Robishaw & Heiles (2009) developed an approximate model of these sidelobes. They found three components: (1) spillover around the secondary reflector, which is the most serious; (2) a component caused by reflection from the screen, which is located near the secondary; and (3) the Arago spot.[1] RHSTK_2021 contains an IDL procedure called `predict_gbt_sidelobes_rhstk` to estimate the sidelobe contributions for Stokes I.

1.5. Basics and Fundamental Assumptions

For the purposes of description, we implicitly assume native linear polarization. We therefore designate self-products by XX and YY and cross-products by XY and YX, as described in §4. After phase and amplitude calibration of the self- and cross-products, the measured Stokes parameters are given by equation 11. If the system has native circular polarization, then almost everything in this memo remains the same except for some obvious changes.

1.6. Source Selection for calibrating linear and circular polarization

Classically, one derives the Mueller matrix from a series of scans on a linearly polarized calibration source such as 3C286. This series should cover a substantial range in parallactic angle; the larger the range, the more accurate the result. So request telescope time and plan your observations accordingly!

The polarization memo by Heiles and Fisher (1999) contains lists of suitable single-dish polarization calibrators at 1.4, 5, 8, and 14 GHz. These lists are reproduced here in Tables 1 and 2. *Caveats:*

[1] After Robishaw & Heiles (2009) appeared, the NRAO staff published a more detailed and accurate model of the sidelobes; see Boothroyd et al. (2011).

1. Data for these tables were obtained before the beginning of the 21^{st} century. Sources can be time-variable. The polarization properties of some sources have been redetermined with more recent data.

2. The Earth's ionosphere introduces time-variable Faraday rotation which can change position angles by $\sim 20^{\circ}$ for the HI line.

Table 1. POLARIZATION RESULTS: SOME SINGLE-DISH CALIBRATORS AT 1420 MHZ

Source	α_{1950}	δ_{1950}	S, Jy	P_s, Jy	%Pol	θ_s
3C27	00 52 44.9	68 06 06	6.65	0.93	7.02 ± 0.03	131.9 ± 0.5
3C27	00 52 44.9	68 06 06	6.65	0.82	6.15 ± 0.35	131.8 ± 1.8
3C27 sp	00 52 44.9	68 06 06	6.65	0.52	3.9± ···	122.3± ···
3C29	00 55 00.7	–01 40 30	5.21	1.15	11.01 ± 0.65	171.6 ± 0.4
3C29	00 55 00.7	–01 40 30	5.21	1.10	10.52 ± 0.28	171.6 ± 0.4
3C33	01 06 14.2	13 03 37	12.53	1.81	$7.09 \pm 0.42^*$	$68.6 \pm 0.9^*$
3C33	01 06 14.2	13 03 37	12.53	1.76	$6.92 \pm 0.35^*$	$67.4 \pm 0.7^*$
3C41 AO	01 23 54.7	32 57 36	3.5	0.42	6.0	48.9
3C98	03 56 11.0	10 17 41	10.25	1.05	5.10 ± 0.15	72.0 ± 0.8
3C98	03 56 11.0	10 17 41	10.25	0.96	4.70 ± 0.10	71.0 ± 0.6
3C123	04 33 55.2	29 34 14	45.16	0.45	$0.50 \pm 0.19^*$	$140.2 \pm 5.5^*$
3C123	04 33 55.2	29 34 14	45.16	0.33	$0.36 \pm 0.13^*$	$153.4 \pm 5.1^*$
3C138	05 18 16.5	16 35 26	8.88	1.11	6.81 ± 0.41	176.2 ± 1.1
3C138	05 18 16.5	16 35 26	8.88	1.05	6.48 ± 0.11	176.9 ± 0.3
3C144-TAU	05 31 31.0	21 59 17	895.50	14.68	0.82 ± 0.08	87.9 ± 1.6
3C144-TAU	05 31 31.0	21 59 17	895.50	13.92	0.75 ± 0.09	86.0 ± 2.5
ORION-A	05 32 44.0	–05 24 54	389.65	0.86	0.11 ± 0.10	···
ORION-A	05 32 44.0	–05 24 54	389.65	0.94	0.12 ± 0.14	···
3C147.1	05 39 11.0	–01 55 36	57.78	0.69	$0.60 \pm 0.30^*$	$84.0 \pm 7.2^*$
3C147.1	05 39 11.0	–01 55 36	57.78	0.68	$0.59 \pm 0.10^*$	$84.5 \pm 2.4^*$
P0736+01 AO	07 36 42.6	01 44 00	2.9	0.48	8.2	103.4
3C227	09 45 07.8	07 39 09	7.18	0.74	5.14 ± 0.37	142.4 ± 1.0
3C227	09 45 07.8	07 39 09	7.18	0.66	4.56 ± 0.06	144.8 ± 0.7
3C227 sp	09 45 07.8	07 39 09	7.18	0.47	3.3± ···	146.7± ···

Table 1—Continued

Source	α_{1950}	δ_{1950}	S, Jy	P_s, Jy	%Pol	θ_s
3C270	12 16 51.2	06 06 13	17.20	2.62	7.62 ± 0.30	122.1 ± 2.3
3C270	12 16 51.2	06 06 13	17.20	2.63	7.63 ± 0.17	122.3 ± 1.5
3C270 sp	12 16 51.2	06 06 13	17.20	2.22	6.3± ···	128.1± ···
3C273	12 26 32.9	02 19 39	49.77	1.23	1.24 ± 0.04	149.7 ± 0.9
3C273	12 26 32.9	02 19 39	49.77	1.17	1.17 ± 0.06	151.0 ± 0.7
3C274	12 28 17.8	12 39 50	213.56	4.27	1.00 ± 0.05	144.5 ± 5.6
3C274	12 28 17.8	12 39 50	213.56	3.89	0.91 ± 0.09	144.5 ± 6.3
3C274.1	12 32 57.0	21 37 06	2.64	0.77	$14.52 \pm 1.70^*$	$149.9 \pm 1.7^*$
3C274.1	12 32 57.0	21 37 06	2.64	0.61	$11.61 \pm 0.37^*$	$152.6 \pm 0.5^*$
3C274.1 AO	12 32 57.0	21 37 06	2.64	0.67	12.8	158.6
3C286	13 28 49.7	30 46 02	14.78	2.74	9.52 ± 0.16	27.4 ± 0.1
3C286	13 28 49.7	30 46 02	14.78	2.60	9.04 ± 0.13	27.4 ± 0.2
3C286 AO	13 28 49.7	30 46 02	14.78	2.86	9.7	28.8
3C286 sp	13 28 49.7	30 46 02	14.78	2.34	7.9± ···	34.3± ···
P1414+11	14 14 27.3	11 02 16	4.14	0.82	9.89 ± 0.22	25.4 ± 3.1
P1414+11	14 14 27.3	11 02 16	4.14	0.78	9.45 ± 0.15	26.4 ± 0.2
3C336 AO	16 22 33.5	23 52 06	2.7	0.15	2.7	29.1
3C348	16 48 40.1	05 04 28	43.69	1.37	$1.57 \pm 0.17^*$	$57.0 \pm 1.5^*$
3C348	16 48 40.1	05 04 28	43.69	1.27	$1.45 \pm 0.09^*$	$57.1 \pm 0.8^*$
3C348 sp	16 48 40.1	05 04 28	43.69	1.31	1.5± ···	68.1± ···
M17	18 17 33.0	40 35 02	558.25	9.60	0.86 ± 0.02	81.4 ± 1.3
M17	18 17 33.0	40 35 02	558.25	8.49	0.76 ± 0.02	82.8 ± 0.7
W43	18 44 57.0	–01 56 36	140.50	3.04	$1.08 \pm 0.15^*$	$85.1 \pm 1.9^*$
W43	18 44 57.0	–01 56 36	140.50	2.78	$0.99 \pm 0.12^*$	$86.5 \pm 1.7^*$
3C399.1 AO	19 14 00.0	30 14 23	2.7	0.56	10.3	53.6
3C405-CYG	19 57 44.5	40 35 02	1654.90	16.88	0.51 ± 0.05	175.7 ± 6.3
3C405-CYG	19 57 44.5	40 35 02	1654.90	17.54	0.53 ± 0.06	178.3 ± 1.5

Table 1—Continued

Source	α_{1950}	δ_{1950}	S, Jy	P_s, Jy	%Pol	θ_s
3C433 sp	21 21 30.6	24 51 18	11.68	1.67	7.1± ···	136.0± ···
CTA102 sp	22 30 07.7	11 28 23	6.37	0.65	5.1± ···	105.3± ···
3C452 sun	22 43 33.0	39 25 28	14.40	2.74	9.52 ± 0.16	27.4 ± 0.1
3C452 sun	22 43 33.0	39 25 28	14.40	2.60	9.04 ± 0.13	27.4 ± 0.2
3C452 sp	22 43 33.0	39 25 28	9.71	1.22	6.3± ···	13.9± ···
3C454.3	22 51 29.4	15 52 56	13.56	2.09	$7.69 \pm 0.55^*$	$67.8 \pm 1.0^*$
3C454.3	22 51 29.4	15 52 56	13.56	1.81	$6.67 \pm 0.22^*$	$68.0 \pm 0.5^*$
3C454.3 AO	22 51 29.4	15 52 56	13.56	1.57	5.8	70.6
CAS-A	23 21 07.0	58 33 48	2032.00	10.97	0.27 ± 0.18	4.2 ± 8.8
CAS-A	23 21 07.0	58 33 48	2032.00	7.72	0.19 ± 0.06	4.5 ± 2.7

Note. — Column 1 is the source name, column 2 and 3 the 1950 equatorial coordinates, column 4 the flux density S in Jy ($S = \frac{\text{Stokes } I}{2}$), column 5 the polarized flux density $P_s = (Q_s^2 + U_s^2)^{1/2}$ in Jy, column 6 the percent polarization (defined as $\frac{100 P_s}{2S}$), column 7 the position angle.

Note. — Sources listed twice were observed during Jan 99. The first listing was derived from the Stokes Q (the difference between orthogonal linear polarizations) and the second from the Stokes U (the cross-correlation of the two linears). The letters "sp" means that the source was observed during the Spring 1998 observing period; uncertainties are not available and the data somewhat less accurate than in Jan 1999. "sun" means that the observations were severely affected by the Sun and should not be trusted, and in particular that the errors are almost certainly underestimates.

Note. — Most results were derived from at least two 12 position angle datasets at the 140-foot telescope. For them, the quoted errors are derived from the differences among those datasets. For results whose errors have the superscript * there was only a single 12 position

angle dataset and the quoted errors are too small.

Note. — Sources followed by "AO" were observed at Arecibo during Feb99; their results should be more accurate than the 140-foot results.

Table 2. SOME SINGLE-DISH CALIBRATORS AT 4.8, 8.0, and 14.5 GHZ

Source	α_{1950}	δ_{1950}	$S_{4.8}$	$S_{8.0}$	$S_{14.5}$	$P_{4.8}$	$P_{8.0}$	$P_{14.5}$	$\theta_{4.8}$	$\theta_{8.0}$	$\theta_{14.5}$
NRAO5	00 03 40.3	–06 40 17	2.2	2.7	2.3	3.5	3.5	5.0	40	20	20
3C10	00 22 32.0	63 51 42	15.5	8.0	1.7	0.5	1.0	2.5	110	140	110?
3C48	01 34 49.8	32 54 21	5.40	3.56	1.84	4.02	5.17	6.62	104.0	113.7	115.2
3C58	02 01 52.0	64 35 06	29.3	28.0	23.1	5.6	5.5	5.5	163	177	7
3C66B	02 20 01.9	42 45 54	3.26	1.78	0.81	3.63	3.6	2.5	70.5	79.	94.
P0218+35.7	02 18 04.1	35 42 32	1.3	1.2	1.2	2.5	2.0	7.0	20	40	50
3C83.1	03 15 00.0	41 41 12	1.8	1.1	0.6	5.5	6.0	6.5	110	115	120
3C84	03 16 29.6	41 19 52	22	21	19	0.05	0.05	0.05	—	—	—
NRAO140	03 33 22.6	32 08 37	1.6	1.4	1.7	4.0	4.5	3.0	60	45	50
3C93	03 40 51.6	04 48 22	0.87	0.59	0.32	7.5	7.4	10.6	139.2	131.	134.
4C76.03	04 03 58.6	76 48 54	2.8	2.2	1.5	0.5	2?	3.0	0	100?	50
3C138	05 18 16.5	16 35 27	3.8	2.8	1.5	10.5	11.0	11.0	170	170	170
P0521–36.5	05 21 12.9	–36 30 17	8.0	7.5	5.0	3.5	2.2	2.0	75	70	70
3C144	05 31 31.0	21 59 17	596	560	430	5.0	6.8	9.9	141	146	152
3C147	05 38 43.5	49 49 43	7.5	5.5	2.8	0.3	1.0	3.0	0	150	50
3C153	06 05 44.5	48 04 49	1.32	0.81	0.40	3.94	5.1	5.2	52.3	50.	54.
3C196	08 09 59.4	48 22 07	4.3	2.6	1.2	2.3	2.0	2.0	120	150	160
P0836+71.0	08 36 21.5	71 04 22	2.3	2.6	2.3	7.0	4.8	4.0	100	105	125
3C207	08 38 01.8	13 23 06	1.3	1.3	1.3	3.0	3.0	2.0	25	20	15
3C216	09 06 17.3	43 05 59	1.6	1.3	1.1	1.5	1.5	2.0	90	0	150
3C219	09 17 50.7	45 51 44	2.4	1.4	0.8	3.0	4.0	2.5	145	140	130
3C245	10 40 06.1	12 19 15	1.61	1.33	0.98	8.38	7.00	5.10	33.0	27.9	29.4
P1127–14.5	11 27 35.7	–14 32 55	3.8	3.3	2.6	3.5	3.5	3.0	160	160	160
3C273	12 26 33.2	02 19 43	37	44	48	3.3	3.5	3.0	170	155	148
3C274	12 28 17.6	12 40 02	71	49	29	0.48	1.6	2.9	40	87	53
3C280	12 54 41.4	47 36 32	1.66	1.07	0.54	7.64	8.2	11.2	44.2	51.2	53.2
3C286	13 28 49.7	30 45 59	7.37	5.53	3.53	11.09	11.46	11.82	33.21	33.13	35.21
3C330	16 09 16.2	66 04 30	2.24	1.32	0.64	3.59	4.2	4.0	131.5	132.	108.
MK-501	16 52 11.8	39 50 25	1.6	1.6	1.3	2.7	3.0	3.0	10	5	0
3C353	17 17 55.6	–00 55 54	22.2	15.5	—	5.2	4.4	—	87	79	—
3C390.3	18 45 45.5	79 42 45	4.4	3.1	1.5	6.0	7.0	4.0	25	25	25
3C395	19 01 02.2	31 55 12	1.5	1.5	1.3	4.0	3.5	3.5	65	45	30
P2005+40.3	20 04 13.1	40 20 34	2.7	2.2	1.8	4.5	2.5	2.5	20	50	60
P2014+37.0	20 14 34.6	37 05 03	3.5	2.5	0.9	7.0	7.0	7.0	103	125	135
3C452	22 43 32.8	39 25 28	3.14	1.82	0.63	7.14	7.0	5.9	121.4	159.6	173.

Note. — Column 1 is the source name, column 2 and 3 the 1950 equatorial coordinates, column 4-6 the flux density in Jy ($S = \frac{\text{Stokes } I}{2}$), columns 7-9 the percentage polarization (defined as $\frac{100\times\text{polarized intensity}}{2S}$), columns 10-12 the position angle in degrees. Subscripts indicate frequency in

GHz.

Note. — We perused the University of Michigan catalog (Aller et al 1996) and include all reasonably strong sources whose polarization properties have not varied terribly much over the period 1990 $\rightarrow$ 1999.0. Sources with names in **boldface** are stable and have typical uncertainties equal to a few or less in the last quoted decimal place; these values were kindly provided by Hugh Aller. For the others, values were estimated by eye from graphs and uncertainties are at least a few 0.1% in S, a few 0.1% in P, a few degrees in θ; some sources exhibit distinct time variability and the user should check with recent observations. Most positions are from Kuhr et al (1981); some are from BDFL and Simbad. The Michigan database is on the web: *http://www.astro.lsa.umich.edu:80/obs/radiotel/radiotel.html*.

You can use our IDL routine `find_gbt_polcal` to plot the parallactic angle swing of the Heiles & Fisher calibrators as a function of LST (read the documentation for details). Figure 1 shows the parallactic angle swing for a set of calibrators over the LST range 8–16 hours.

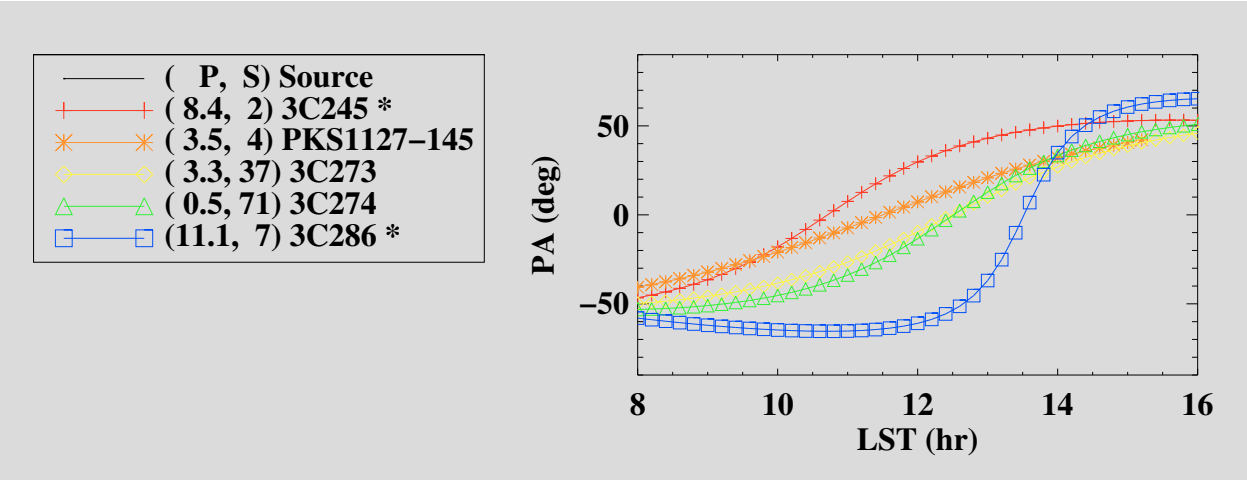


Fig. 1.— Parallactic angle swing for some polarization calibrator s over the LST range 8–16 hours. The legend lists the calibrators along with their total flux density S in mJy and the percentage polarization P at 4.8 GHz. This plot was made by `find_gbt_polcal`.

At cm wavelengths, there are no continuum sources having strong Stokes V. This is not a problem for determining Mueller matrix parameters, because all but one of them can be obtained from the linearly polarized calibration sources. The errant parameter is the sign of Stokes V, for which one requires either a circularly-polarized calibration source or a locally-produced test signal[2]. The only astronomical sources with strong circular polarization are pulsars and OH masers. For spectropolarimetry, pulsars present the additional complication of time-resolving the spectrum. OH masers are much more convenient, but offer only highly restricted frequency coverage. Nevertheless, combining their Stokes V sign at one frequency with the linearly-polarized continuum results for a broader range of frequencies works well. Problems with using OH masers include their time variability and their clumpy structure on scales of arcminutes to degrees. Time variability only occasionally affects the sign of Stokes V. The clumpy structure makes it difficult to compare spectra from large and small telescopes; the larger beam of a small telescope means more OH maser clumps contribute to its OH maser spectrum, making it more complicated and harder to interpret. This is particularly the case for W49.

[2]When using locally-produced test signals, remember that the sign of V changes upon reflection!

1.7. Use more than one calibration source

After deriving the Mueller matrix, you should check a *second* linearly-polarized calibration source by observing, calibrating, and Mueller-correcting it. This observation can be quick—just a single on/off will do. The reason is to make sure that the 'handedness' of the position angle of linear polarization is defined properly, i.e., that position angle increases from North towards the East; this can occur if the cables for the two polarizations are interchanged (equations 11 and 12).

2. STANDARDS AND SIGN CONVENTIONS: A BIG MESS

The magnificent review of the Zeeman-splitting literature of radio astronomy by Robishaw (2008) makes clear the rampant confusion regarding definitions and signs of Stokes V and the derived magnetic fields. Accordingly, we provide here a very brief summary.

The International Astronomical Union has adopted definitions for the two circular polarizations and, also, Stokes V (IAU 1974). The IAU definitions for the polarizations follow the IEEE definitions, which are opposite to those conventionally used by physicists. The IEEE definition of RCP is: as viewed from the transmitter, the E-vector rotates clockwise. The IAU definition of Stokes V is: $V = \mathrm{RCP} - \mathrm{LCP}$, where the polarizations follow the IEEE definition. This is opposite to that historically used by many radio astronomers, who followed the definition of Kraus (1966): $V = \mathrm{LCP} - \mathrm{RCP}$, where the polarizations follow the IEEE definition.

We now always use the IAU-sanctioned definitions: RCP rotates clockwise as viewed by the transmitter, and $V = RCP - LCP$ (IEEE definition).

The 1665 MHz OH maser profile for W49 and the Zeeman splitting of the 21-cm line in absorption against Cas A, shown in Figure 2, are easily detectable and act as a good system and sanity check. Coles & Rumsey (1970), used the IAU definition—even though their work pre-dates the IAU definition!). To our knowledge, the Cas A profile in Figure 2 is the only publicly-available one that follows the IAU definition for Stokes V. Robishaw (2008) stresses that *all* published Stokes V profiles for Cas A either do not follow the IAU convention or are ambiguously or incorrectly defined.

With these definitions, and in addition the common Zeeman convention that *positive B points away from the observer*, the signs of the derived fields are most straightforwardly given by consulting Table 2.2 of Robishaw (2008), which is reproduced here as Figure 3. For emission lines, this figure can be summarized succinctly: *For positive B, RCP lies at a smaller frequency than LCP.*

It is worth more than a footnote to mention here that the convention for magnetic field sign used by radio astronomers making Faraday rotation measurements is exactly opposite to the Zeeman convention. Faraday rotation experts will tell you that a positive field points towards the observer. This convention was arbitrarily chosen by Manchester (1972) because he preferred that a positive

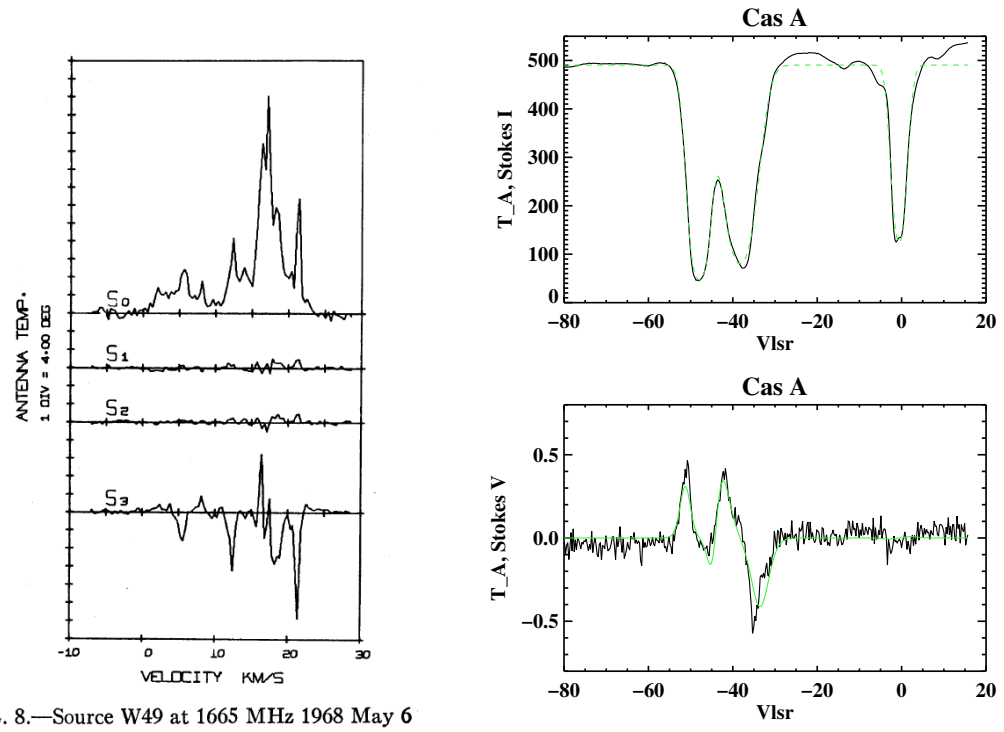


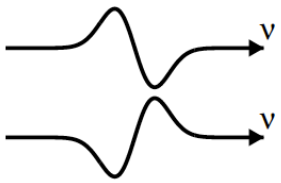
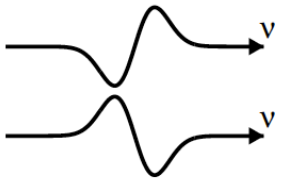
FIG. 8.—Source W49 at 1665 MHz 1968 May 6

Fig. 2.— Left panel: The four Stokes parameters for W49 (Coles & Rumsey 1970). Right panel: Stokes I (top) and V (bottom) for Cas A, from decades-old observations with the Hat Creek 85 foot. In both, Stokes V is plotted according to the IAU convention.

rotation measure also correspond to a positive magnetic field (he even points out that this is opposite to the Zeeman convention!). The Zeeman convention has a nice analogue: velocities are also defined to be positive when pointing away from the observer. This convention had been firmly established since the early 1900s, and the early Faraday rotation papers (e.g. Burn 1966) suggest that a positive field also points away from the observer, so it was a questionable decision to have so cavalierly changed the established convention. Clearly for the worse, we are left with this duality in the polarization world and not many astronomers are aware of this. It should be clear that one must be very careful to state one's conventions.

We show two examples of properly-defined Stokes V from real observations. Figure 2 shows the 21-cm line absorption line spectrum for Cas A in Stokes I and Stokes V. The dashed green lines are Gaussian fits with three components. For the three components with velocities $[-48.0, -38.1, -0.7]$ km/s, the derived line-of-sight magnetic field strengths are $[+7.8, +18.1, -0.0]\pm$ $[0.5, 0.6, 0.8]$ μG, respectively. We fitted these components with optical depth profiles using our IDL procedure `tbgfitflex_exp.pro` and we obtained field strengths with `zgfit_selfabs.pro`, both found in `rhstk_2021/procs/zeeman`. Equivalent results can be obtained by fitting negative Gaussians (instead of optical depth profiles) using `gfit.pro` and `zgfit.pro`. These programs give

TABLE 2.2
MAGNETIC FIELD DIRECTION DETERMINED FROM THE ZEEMAN EFFECT

Field Direction[a]	$\nu_0 - \Delta\nu_Z$	$\nu_0 + \Delta\nu_Z$	Stokes V Profile[b]
Emission Lines			
$B > 0$	RCP	LCP	
$B < 0$	LCP	RCP	
Absorption Lines			
$B > 0$	LCP	RCP	
$B < 0$	RCP	LCP	

[a]A positive magnetic field points away from the observer by convention.
[b]V = RCP – LCP by IAU convention (IAU 1974).

Fig. 3.— Obtaining the sign of B from Stokes V, when V is plotted against *frequency*. From Robishaw (2008), Table 2.2.

the correct signs of derived fields and return the splitting in units of the horizontal axis, e.g. km/s. One must then convert these velocity-splitting units to Hz and then, for the 21-cm line, divide by the Zeeman splitting coefficient[3] for the 21-cm transition, $b = 2.8$ Hz $\mu\mathrm{G}^{-1}$.

3. ALL-STOKES CALIBRATION IN A NUTSHELL: STAGE 1, STAGE 2

Figure 4 is a block diagram showing a telescope with two signal paths, one for each polarization, labeled 'X' and 'Y'. These might be either approximately orthogonal linear polarizations or circular polarizations. The essence of all-Stokes observing is to measure the time-average self-products XX and YY together with the cross-products XY and YX. These products are computed by the digital spectrometer, labeled 'CORRELATOR' in Figure 4. The spectrometer uses either the XF or FX technique and provides four spectra (see §4.1), which refer to the signals presented to the correlator input. The correlator inputs differ from the astronomical electromagnetic waves because they have gone through the telescope system, which includes the feed, electronics, and cables. Inside the correlator, these analog signals are digitized.

[3]The Zeeman splitting coefficient for a radiative transition depends on the Landé g-factor for that transition and the Bohr magneton; see Heiles et al. (1993) for more details.

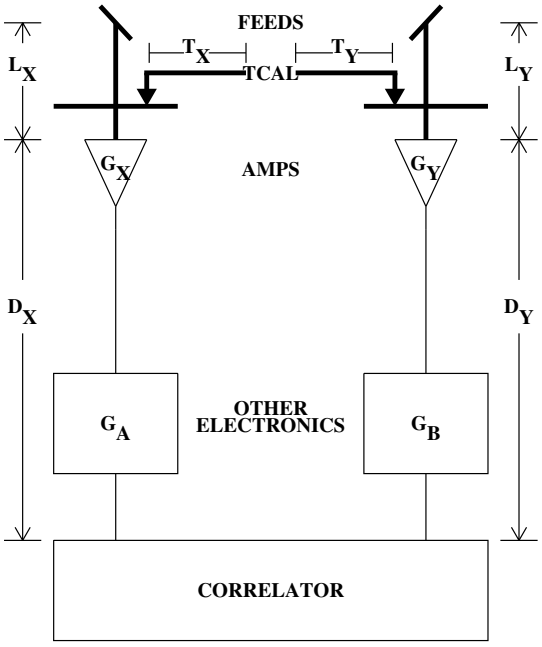


Fig. 4.— Block diagram of a dual-polarized single-dish radio telescope. Thick lines represent time-independent aspects, which are calibrated by the feed's Mueller matrix. Thin lines represent time-variable aspects, which are calibrated frequently using the Noise Diode: the r.f. amplifier gains (G_X, G_Y), the 'other electronics' gains (G_A, G_B), and the delays caused by cable length and electronics represented by the lengths (L_X, L_Y); all of these quantities are time-variable. This diagram implies there is zero coupling (crosstalk) between the two signal paths, which is not the case in real life; crosstalk is represented in the feed's Mueller matrix.

We are interested in the astronomical signals, so we need to remove the effects of the telescope system. In other words, we need to calibrate the measured correlation products and turn them into the four Stokes parameters. This is done in two explicit stages:

1. *Stage 1* calibrates the self- and cross-products to units of Kelvins and calibrates the phase difference between the two polarization channels to its reference value. These calibrations use a standard noise source called the 'Noise Diode' whose intensity and relative phase injected into the two polarization channels are constant in time. We measure the Diode deflection (DIODEON – DIODEOFF) with the correlator, which tells us the gains and relative phase of the thin lines in Figure 4, which delineate the telescope system after the point where the Diode is injected.

The thin-line portion of Figure 4 is *time-variable*. Consider the relative phase of the two signal paths. This depends on at least two things: the cable length difference and the phase delays produced by active circuit elements such as amplifiers. The cables can be very long (measured in hundreds of meters at Arecibo, and thousands at the GBT), so thermal expansion plays a role in their lengths; moreover, there is no guarantee that individual cables used for each signal path don't happen to get interchanged, either by a technician or by automatic assignments of optical fibers by the system. The gain and phase delay of an amplifier depend on applied voltages and temperature, and moreover an individual amplifier usually produces a 180° phase jump; the number of amplifiers can change as gains are automatically adjusted.

2. *Stage 2* combines these calibrated products and either derives or applies the system Mueller matrix and the source polarization properties. In stage 2, we can either (1) apply system Mueller matrix to derive the source polarization parameters, (2) determine the system Mueller matrix, or sometimes (3) both(!). The Mueller matrix calibrates the thick lines in Figure 4, which delineate the telescope system before the Diode is injected. This includes imperfections of the feed and the path length difference $D_Y - D_X$.

   We regard the thick-line portion of Figure 4 as *constant in time*; this part is embodied in the feed's Mueller Matrix $\mathbf{M}_{\rho_\mathbf{X}}$ (§5.3). This portion includes the feed and Noise Diode circuitry. These components are mechanical structures and don't change without good reason. What's important for *polarimetry* are the *dimensions and cable lengths*, so even if the Noise Diode itself is replaced without changing cables, the thick-line portion remains unchanged for polarimetry. Of course, a different Noise Diode will produce different noise power, which changes the overall intensity calibration, but not the polarization calibration. An astronomical observer is usually given the Diode strength in Kelvin, which is its equivalent antenna temperature, and one normally assumes it to be reliable and constant with time; this is usually a good approximation.

4. STAGE 1 BASICS: INTENSITY AND PHASE CALIBRATION OF THE MEASURED SELF- AND CROSS-PRODUCTS

4.1. Measuring Self- and Cross-Products with Digital Methods

The time-averaged voltage products are derived from digital samples in one of two ways. Historically, the XF correlation technique[4] prevailed because of its simpler hardware requirements. With XF, one uses a correlation spectrometer, which produces time-averaged auto- and cross-correlation functions (ACFs and CCFs, respectively). These are Fourier transformed, usually in a general-purpose computer, to produce power spectra. Each ACF is computed for N positive lags; negative lags are unnecessary because autocorrelations are symmetric with respect to lag. The

[4]The "X" represents correlation and the "F" represents a Fourier transform.

ACFs are averaged over time and the Fourier transform (FT) of the resulting average ACF gives the self-power spectrum. Because the ACF is symmetric with respect to lag, its Fourier transform is real and symmetric with frequency, so each self-product power spectrum has N independent channels. Symbolically, for the self-product spectra we write

$$XX = \mathrm{FT}\langle \mathrm{ACF}(V_X)\rangle\,, \tag{1a}$$

$$YY = \mathrm{FT}\langle \mathrm{ACF}(V_Y)\rangle\,. \tag{1b}$$

where V_X means voltage for polarization X, etc. The cross-correlation of the two polarizations is not symmetric with lag, so it must be computed both for N positive and N negative lags. Its FT is complex with Hermitian symmetry, so the cross-power spectrum can be regarded as consisting of a real and imaginary part, each with N independent channels. Symbolically, for the cross-product spectra we write

$$XY = \mathrm{Re}(\mathrm{FT}\langle \mathrm{CCF}(V_X V_Y)\rangle)\,, \tag{2a}$$

$$YX = \mathrm{Im}(\mathrm{FT}\langle \mathrm{CCF}(V_X V_Y)\rangle)\,. \tag{2b}$$

Today, the FX technique is favored because of the heavy computing ability of FPGAs and GPUs. With FX, each polarization is sampled at rate t_s over time interval $2T$, providing $2N = \frac{2T}{t_s}$ samples. This block of data is Fourier transformed, producing a complex transform of $2N$ channels with Hermitian symmetry having N positive-frequency and N negative-frequency channels. The self-product power spectrum is this FT times its complex conjugate, and because of the Hermitian symmetry, it is real with the N negative- and positive-frequency portions identical. Thus, it is a power spectrum with N independent channels. Similarly, one calculates cross-product power spectra by multiplying the Fourier transforms of the two polarizations with both possibilities of complex conjugate (eq. 4a). This produces a complex cross-power spectrum having $2N$ independent channels, split between negative and positive frequencies. This cross-power spectrum does not have Hermitian symmetry, so it has a real part and an imaginary part, each with N independent channels. Thus, we have four spectra of length N. Symbolically, for the self-product spectra we write

$$XX = \langle \mathrm{FT}(V_X)\overline{\mathrm{FT}(V_X)}\rangle\,, \tag{3a}$$

$$YY = \langle \mathrm{FT}(V_Y)\overline{\mathrm{FT}(V_Y)}\rangle\,, \tag{3b}$$

and for the cross-product spectra

$$XY = \left[\, \langle \mathrm{FT}(V_X)\overline{\mathrm{FT}(V_Y)}\rangle + \langle \overline{\mathrm{FT}(V_X)}\mathrm{FT}(V_Y)\rangle \,\right]\,, \tag{4a}$$

$$YX = i\left[\, \langle \mathrm{FT}(V_X)\overline{\mathrm{FT}(V_Y)}\rangle - \langle \overline{\mathrm{FT}(V_X)}\mathrm{FT}(V_Y)\rangle \,\right]\,, \tag{4b}$$

For both the XF and FX techniques we obtain two self-product and two cross-product spectra. For the cross-products, we can speak of the Real and Imaginary parts of $XY(f)$. Above, and in the rest of this document, we use the shorthand notation

$$XY \equiv Re(XY(f)) \tag{5a}$$

$$YX \equiv Im(XY(f)) \tag{5b}$$

With this, the phase difference between X and Y is

$$\Delta\psi = \text{atan}\left(\frac{YX}{XY}\right) \tag{6}$$

4.2. The power gains of X and Y signal paths

Most observers are not concerned with polarization information, so they treat the X and Y signal paths independently. They turn on and off the Diode and measure its deflection in terms of correlator counts; knowing the Diode's strength in Kelvin, this provides the 'Counts per Kelvin', which we call CpK. This is a power gain, not a voltage gain; that is, CpK relates the equivalent system temperatures to the counts of the self-products XX and YY. The numerical value of CpK for each signal path depends on the electronics, in particular the gains of the amplifiers in the chain (G_X and G_A for the X-path in Figure 4). Being active elements, some of which are cryogenically cooled, these gains cannot be considered constant with time, and they are calibrated at intervals that may be separated by order of minutes. The calibration procedure uses the Diode deflection, i.e., the comparison of DIODEON and DIODEOFF measurements made close in time.

The gain, i.e. CpK, is a very strong function of frequency within the observing band. Band-limiting filters define the spectral bandpass that enters the digital spectrometer, and their gains vary between almost zero and the maximum within the bandpass. The black and magenta curves of Figure 5 show typical examples (top left panel for Arecibo's Interim Correlator, top right for the GBT's Spectral Processor). We characterize this frequency-dependent gain by two concepts. One is a representative value of CpK over the bandpass, and the other is the 'bandpass shape'. We define the former by averaging over a specified spectral channel range in the output spectrum, which normally includes most of the bandpass except for the ends.

4.3. The Relative Phase Delay Between X and Y Signal Paths

To derive polarization, we must measure the cross-products of the X and Y signals. This cross-correlation depends not only on the above-defined gains, but also on their relative phase delay $\Delta\psi$. We need to use the Diode for calibrating this phase difference. If each signal path has its own Diode, so that the X and Y Diode voltages are uncorrelated, then the XY product for the Diode deflection is zero, so we cannot measure the relative phase delay. However, we arrange that the Diode deflections are correlated. To accomplish this, one can use a single Noise Diode and a power splitter (as is done in Figure 4, and at the GBT and Arecibo); or one can insert the noise in the feed, or at the paraboloid's vertex, by radiating the Diode's power with a probe at 45° to the orthogonal X and Y feed probes (as was done with the Hat Creek 85-footer).

The relative phase of the X and Y Diode deflections depends on, at minimum, the following four factors. The first is constant in time and, consequently, is embodied in the Mueller Matrix correction; the others need to be calibrated concurrently.

1. The phase difference between the correlated Diodes as seen by the two first amplifiers. If the Diode is injected with a waveguide probe or vertex radiator, this phase difference should depend only on the length difference of the cables that connect the feed to the amplifiers. If it is injected by cables, using a power splitter and directional couplers as in Figure 4, then this phase difference also results from cable length differences in the cables connecting the Diode to the two feeds, and in addition there might be a phase offset introduced by the power splitter and directional couplers. As depicted by the thick lines in Figure 4, this component of phase offset depends on mechanical structures and should be stable and constant with time.

2. The phase difference introduced by the different cable lengths D_X and D_Y, which can be time-variable because the cables are long and environmental conditions (temperature) come into play, as mentioned above. The relative phase delay ψ from the cable length difference $\Delta D = D_Y - D_X$ varies linearly with frequency:

$$\frac{d\psi}{df} = \frac{2\pi\ \Delta D}{v_{ph}}\,, \tag{7}$$

   where v_{ph} is the phase velocity in the cable. Typical values at both Arecibo and GBT are a few tenths of a radian per MHz, which leads to $\Delta D \sim$ several meters.

3. The phase difference introduced by components in the X and Y signal paths. Some of these are active circuit elements and can change with time.

4. At some point in the receiver chain one always has a band-limiting filter. Frequency-dependent gains automatically introduce phase delays, the minimum values of which can be calculated from the Kramers-Kronig relations. The exact formula for electrical circuits is equation (2) of Bode (1940)[5]:

$$[Phase\ Shift](f_c) = -\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{d\mathcal{G}}{du}\ \ln\left[\coth\left(\frac{|u|}{2}\right)\right]du\,, \tag{8}$$

   where $\mathcal{G}(u)$ is the logarithmic filter power gain in nepers, $u = \ln\left(\frac{\nu}{\nu_c}\right)$, and ν is frequency. The weighting function $\ln\left[\coth\left(\frac{|u|}{2}\right)\right]$ is sharply peaked at $f = f_c$, so a good approximation eliminates the integral and uses only the local derivative (equation 22 of O'Donnell, Jaynes, & Miller 1981)

$$\phi(\nu_c) = -\frac{\pi}{2}\frac{d\mathcal{G}(u)}{du}\bigg|_{u=0}\,. \tag{9}$$

[5]You would miss a lot if you pass up the opportunity to read this paper, particularly the first six pages. Go to http://www.alcatel-lucent.com/bstj/ .

Thus, a bandpass filter centered at f_{cntr} with width Δf introduces a phase shift proportional to $\frac{f_{cntr}}{\Delta f}$: sharp, narrow filters produce more phase shift.

If the filters in the two polarizations are not perfectly matched, a frequency-dependent phase difference between the two polarization channels ensues. This can be particularly serious when the filters have significant gain changes within the usable portion of the band, and in real life these phase changes are, in fact, significant. For example, the 'radar backend' (Margot 2021) baseband low-pass filters at the GBT have a roughly Gaussian shape and the delay of a single filter near the half-power point, relative to that at the center frequency, is $\sim 600°$ (!). Those filters are imperfectly matched and, as a result, produce a *relative* phase delay $\Delta\phi \sim 150°$.

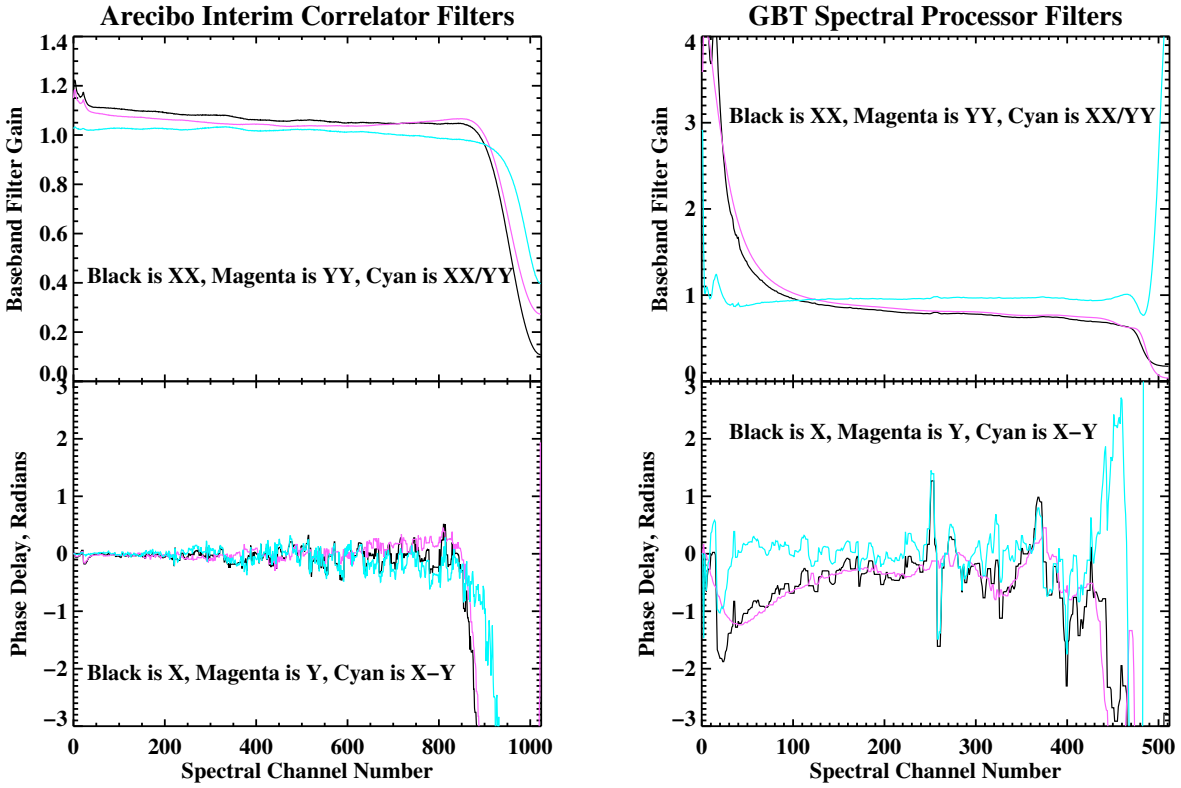


Fig. 5.— Filter shapes and their theoretical phase delays for the Arecibo Interim Correlator (left panel) and the GBT Spectral Processor (right panel).

Our work at Arecibo has been done with the interim correlator, for which the baseband low-pass filters are digitally defined and are remarkably flat; Figure 5, left-hand panel, shows the positive-frequency half of the X and Y signal-path filters for bandwidth 12.5 MHz (cutoff frequency 6.25 MHz). Phase shifts occur only at the end, where the response of filters drops precipitously.

In contrast, the GBT Spectral Processor filters are far from flat. For example, the 0.625 MHz bandwidth filter is shown in the right-hand panel of Figure 5. The phase difference between the X and Y signal paths is the cyan line in the bottom half, and there is a clear frequency-dependent phase delay difference.

5. STAGE 2 BASICS: DERIVING THE MUELLER MATRIX

5.1. Basic Definitions of Reference Frames

We need to turn the four measured intensity- and phase-calibrated self- and cross-products into the four Stokes parameters. The four *measured* Stokes parameters depend not only on on the polarization properties of the source, but also the details of the telescope, its mounting, and the electronics. Our final result needs to be the Stokes parameters of the source, with the telescope details removed, and they must be expressed according to the IAU definitions, §2: **(1)** the position angle of linear polarization increases from North towards the East, and (**2**) the sign of the circularly polarized Stokes V is correct. The measured Stokes parameters do not necessarily follow any of these conventions.

We specify the two reference frames of our Stokes parameters:

1. The subscripts *meas* and *tel* refer to the *measured* Stokes parameters, which are in the *telescope* reference frame. In particular, $\mathbf{S_{tel}}$ is the set of 4 Stokes parameters of the source that is derived from the measured Stokes parameters.

2. Subscripts *IAU* and *src-IAU* refer to the IAU reference frame. $\mathbf{S_{IAU}}$ represents the source Stokes parameters as defined by the IAU.

We use Mueller matrices to convert from one frame to another, using subscripts to indicate the frames. For example, the Mueller matrix $\mathbf{M_{tel-IAU}}$ converts from the telescope frame to the IAU frame (more detail below).

5.2. Deriving the Measured Stokes Parameters from the Measured Voltage Products

We convert the 4 measured voltage products to the measured Stokes parameters by applying a series of transformations. These transformations are best accomplished by considering each set of four quantities to be a 4-vector and using matrix methods. We write the four source Stokes

parameters in the telescope frame as the Stokes vector $\mathbf{S_{tel}}$

$$\mathbf{S_{tel}} = \begin{bmatrix} I_{tel} \\ Q_{tel} \\ U_{tel} \\ V_{tel} \end{bmatrix} . \tag{10}$$

Similarly, the four *measured* correlator outputs are

1. For native linear feeds:

$$\mathbf{S}_{\rm meas} = \begin{bmatrix} I_{\rm meas} \\ Q_{\rm meas} \\ U_{\rm meas} \\ V_{\rm meas} \end{bmatrix} = \begin{bmatrix} XX + YY \\ XX - YY \\ 2XY \\ \pm 2YX \end{bmatrix} . \tag{11}$$

2. For native circular feeds:

$$\mathbf{S}_{\rm meas} = \begin{bmatrix} I_{\rm meas} \\ Q_{\rm meas} \\ U_{\rm meas} \\ V_{\rm meas} \end{bmatrix} = \begin{bmatrix} RR + LL \\ \pm 2LR \\ 2RL \\ RR - LL \end{bmatrix} . \tag{12}$$

The sign ambiguities above arise because one can't be sure that the cables carrying the two polarizations are not interchanged, nor how the native polarizations are treated inside the correlator. Resolving these sign ambiguities is part of the calibration process.

The transformation matrices are called Mueller matrices. We can regard a Mueller matrix as the transfer function associated with an individual device or operation, or a set of them; the Mueller matrix for the set is the matrix product of the matrices for the individual components. So we have

$$\mathbf{S_{meas}} = \mathbf{M_{tel-meas}} \cdot \mathbf{S_{tel}} \ . \tag{13}$$

$\mathbf{M_{tel-meas}}$ is the product of 2 individual Mueller matrices associated with the 2 devices that the source's electromagnetic radiation sequentially encounters on its voyage from the source to the correlator input: $\mathbf{M}_{\rho\mathbf{X}}$, the matrix involving the telescope pointing, and $\mathbf{M_{RX}}$, the matrix involving the receiving system:

$$\mathbf{M_{tel-meas}} = \mathbf{M_{RX}} \cdot \mathbf{M}_{\rho\mathbf{X}} \ . \tag{14}$$

5.3. $\mathbf{M_{\rho X}}$

$\mathbf{M_{\rho X}}$ has one parameter, ρ_X, which is the position angle on the celestial sphere of the X-polarization probe.

$$\mathbf{M_{\rho X}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\rho_X & \sin 2\rho_X & 0 \\ 0 & -\sin 2\rho_X & \cos 2\rho_X & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} . \tag{15}$$

The central 2×2 submatrix is, of course, nothing but a rotation matrix. For an alt-az mounted telescope such as Arecibo, $\rho_X = \chi$, the parallactic angle, plus an offset angle that depends on how the feed is mechanically mounted.

5.4. $\mathbf{M_{RX}}$

M_{RX} has six parameters, briefly summarized below; one of these (χ in 4 below) is fixed to zero. For a detailed discussion, see Heiles et al. (2001b). The first three are the most important because they describe zeroth-order effects. The others describe imperfections such as nonorthogonality in the polarizations, so for high-quality feeds (the typical case) they are very small. The parameters, listed in order of their typical importance in real life, are:

1. $\boldsymbol{\Delta G}$ 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}). Our expansion currently takes terms in ΔG to first order only, so if the relative cal intensities are significantly incorrect then the other parameters will be affected. The relative cal values should be modified to make $\Delta G = 0$, keeping their sum the same. To accomplish this, make $T_{calX,modified} = T_{calX}\left(1 - \frac{\Delta G}{2}\right)$ and $T_{calY,modified} = T_{calY}\left(1 + \frac{\Delta G}{2}\right)$.

2. $\boldsymbol{\psi}$ is the phase difference between the cal and the incoming radiation from the sky. It redistributes power between (U, V) for a native linear feed and between (Q, U) for a native circular feed. For these two cases:

   (a) Native linear polarization. If the cables connecting the correlated Diode to the feed are of identical length and the r.f. paths through the feed are identical for the two polarizations (the thick lines in Figure 4), then $\psi = 0$. If the cable lengths differ, then $\psi \neq 0$. Moreover, in this case we expect $\psi \propto \nu$, where ν is the r.f. frequency.

   (b) Native circular polarization. In this case, the phase difference ψ determines the angle of linear polarization for Stokes Q and U. This is indeterminant from our least-squares fit for $\mathbf{M_{RX}}$. Accordingly, for native circular we force $\psi = 0$ and don't fit for it; we use $\mathbf{M_{tel-IAU}}$ to empirically obtain the correct position angle of linear polarization.

Because ψ depends linearly on frequency, so too will this correction to the position angle for native circular.

3. $\boldsymbol{\alpha}$ is a measure of the voltage ratio of the polarization ellipse produced when the feed observes pure linear polarization. Generally, the electric vector traces an ellipse with time; $\tan\alpha$ is the ratio of major and minor axes of the voltage ellipse. Thus, $\tan^2\alpha$ is the ratio of the powers. If a source having fractional linear polarization $P_{src} = \sqrt{Q_{src}^2 + U_{src}^2}$ is observed with a native circular feed that has $\alpha = \frac{\pi}{4} + \delta\alpha$, with $\delta\alpha \ll 1$, then the measured Stokes V will change sinusoidally with $2PA_{az}$ and have peak-to-peak amplitude $4\delta\alpha$. α might also be a function of frequency, particularly for native circular.

4. $\boldsymbol{\chi}$ is the relative phase of the two voltages specified by α. Our analysis assumes $\chi = 90°$. This hardwired specification is compensated for by $\Delta\rho$ (see §5.5 below), so this incurs no loss of generality.

5. $\boldsymbol{\epsilon}$ is a measure of imperfection of the feed in producing nonorthogonal polarizations (false correlations in the two correlated outputs). Our expansion takes ϵ to first order only. The only astronomical effect of nonzero ϵ is to contaminate the polarized Stokes parameters (Q, U, V) by coupling Stokes I into them at level $\sim 2\epsilon$; the exact coupling depends on the other parameters. For weakly polarized sources, this produces false polarization; for strongly polarized sources such as pulsars, it also produces incorrect Stokes I. If ϵ is large enough so that its first-order approximation is insufficient, ϵ itself needs to be modified appropriately (consult Heiles et al. (2001) for details).

6. $\boldsymbol{\phi}$ is the phase angle at which the voltage coupling ϵ occurs. It works with ϵ to couple I with (Q, U, V).

Putting all these together, we follow Heiles et al. 2001b): We set $\chi = 90°$, we ignore second order terms in the imperfection amplitudes $(\epsilon, \Delta G)$, and we retain all orders in α. This gives

$$\mathbf{M_{RX}} = \begin{bmatrix} 1 & (-2\epsilon\sin\phi\sin 2\alpha + \frac{\Delta G}{2}\cos 2\alpha) & 2\epsilon\cos\phi & (2\epsilon\sin\phi\cos 2\alpha + \frac{\Delta G}{2}\sin 2\alpha) \\ \frac{\Delta G}{2} & \cos 2\alpha & 0 & \sin 2\alpha \\ 2\epsilon\cos(\phi+\psi) & \sin 2\alpha\sin\psi & \cos\psi & -\cos 2\alpha\sin\psi \\ 2\epsilon\sin(\phi+\psi) & -\sin 2\alpha\cos\psi & \sin\psi & \cos 2\alpha\cos\psi \end{bmatrix} . \quad (16)$$

The terms in the top row make $I \neq 1$ for a polarized source. If one derives fractional polarization, for example Q/I, then it will be in error by amounts comparable to $[(\epsilon, \Delta G) \times (Q, U, V)]$. For the frequent case of a weakly polarized source, these products are second order and therefore are of no concern. However, for a strongly polarized source such as a pulsar or OH maser, these terms are first order.

5.5. Deriving the Source Stokes Parameters from the Measured Stokes Parameters

The above shows how to convert the voltage products to measured Stokes parameters $\mathbf{S_{meas}}$. To obtain $\mathbf{S_{tel}}$ from $\mathbf{S_{meas}}$ we apply the inverse matrix to equation 13:

$$\mathbf{S_{tel}} = \mathbf{M^{-1}_{tel-meas}} \cdot \mathbf{S_{meas}} \ . \tag{17}$$

In real life, we obtain $\mathbf{S_{meas}}$ and want to know the source Stokes parameters in the IAU frame. This final transformation converts $\mathbf{S_{tel}}$ to the astronomical frame $\mathbf{S_{IAU}}$

$$\mathbf{S_{IAU}} = \mathbf{M_{tel-IAU}} \cdot \mathbf{S_{tel}} \ . \tag{18}$$

or

$$\mathbf{S_{IAU}} = \mathbf{M_{tel-IAU}} \cdot \mathbf{M^{-1}_{tel-meas}} \cdot \mathbf{S_{meas}} \ . \tag{19}$$

$\mathbf{M_{tel-IAU}}$ has 2 parameters:

1. $\boldsymbol{\Delta\rho}$ is the angle by which the derived position angles must be rotated to conform with the conventional astronomical definition. In principle, $\Delta\rho$ depends only on the angle of the mechanically-mounted feed and is a known quantity; in practice, we check it out empirically.

2. $\mathbf{V_{fctr}}$, the factor (equal to ± 1) to convert the derived Stokes V to the astronomically-defined one.

$$\mathbf{M_{tel-IAU}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\Delta\rho & \sin 2\Delta\rho & 0 \\ 0 & -\sin 2\Delta\rho & \cos 2\Delta\rho & 0 \\ 0 & 0 & 0 & V_{fctr} \end{bmatrix} \ . \tag{20}$$

6. USING OBSERVATIONS TO DERIVE THE MUELLER MATRIX PARAMETERS

To derive the Mueller matrix coefficients from observations one needs to observe standard polarization calibration sources such as 3C286. Usually there are two kinds of observations:

1. Most commonly, one is using an alt-az mounted telescope, for which the angle $\rho_X = \chi$, the parallactic angle, and one observes a single source, such as the gold-standard calibrator 3C286, over a large range of parallactic angles. Closely related is the case of multiple sources, *all* having the *same* set of parallactic angles; in real life, this is useful for the case of polarized masers with multiple spectral components, each with its own unique polarization properties.

2. Alternatively, one observes a set of different calibration sources, each with its own ρ_X or its own set of ρ_X angles. This method is more cumbersome, but is the only option for equatorially mounted telescopes like GALT and for the unique case of FAST.

We discuss the least-squares fitting for the two cases separately, because the approaches are totally different. The former technique in **(1)** above is the classical one, because of the prevalence of alt-az mounted telescopes and the ease with which one samples a range of ρ_X.

6.1. Fitting for the Classical (Alt-Az Mount) Case

Classically, we evaluate the Mueller matrix parameters in equation 16 using observations of a polarized calibration source tracked over a wide range of parallactic angle χ. As a bonus, the least-squares fit can simultaneously obtain not only the Mueller matrix parameters but also the Stokes parameters of the source. The source is described by $\mathbf{S_{tel}}$ and the Mueller matrix for the radiation entering the feed by $\mathbf{M_{\rho X}}$. The full Mueller matrix is $\mathbf{M_{tel-meas}} = \mathbf{M_{RX}} \cdot \mathbf{M_{\rho X}}$. The product of this matrix with $\mathbf{S_{tel}}$ results in a set of of four equations, one for each element of the observed $\mathbf{S_{tel}}$ vector. Expressed in terms of matrices, these equations are

$$\begin{bmatrix} I_{meas} \\ Q_{meas} \\ U_{meas} \\ V_{meas} \end{bmatrix} = \mathbf{M_{RX}} \cdot \mathbf{M_{\rho X}} \cdot \begin{bmatrix} I_{tel} \\ Q_{tel} \\ U_{tel} \\ V_{tel} \end{bmatrix} . \tag{21}$$

In practice, we cannot reliably measure the ρ dependence of I_{meas} because it is rendered inaccurate by small gain errors, either from the electronics or the zenith-angle gain dependence of the telescope. Thus, in practice we use fractional correlator outputs and fractional source polarization.

$$\mathbf{S'_{meas}} = \frac{\mathbf{S_{meas}}}{\mathbf{I_{meas}}} \tag{22}$$

and

$$\mathbf{S'_{tel}} = \frac{\mathbf{S_{tel}}}{\mathbf{I_{tel}}} \tag{23}$$

We also take $I_{meas} = I_{tel}$ instead of $I_{meas} \approx I_{tel}$. Note that the division by I_{meas} produces errors in the elements of $\mathbf{S'_{meas}}$, but these errors are second order because they are products of ΔG and/or ϵ with quantities such as Q_{tel} that are already first order. Our whole treatment neglects second order products, so we can neglect these errors.

We now rewrite equation 21, expressing $\mathbf{M_{RX}}$ in the gory detail of all its 16 elements. To save space, we omit the subscript RX from each individual matrix element, i.e. we write m_{QU} instead

of $m_{RX,QU}$:

$$\begin{bmatrix} 1 \\ Q'_{meas} \\ U'_{meas} \\ V'_{meas} \end{bmatrix} = \begin{bmatrix} m_{II} & m_{IQ} & m_{IU} & m_{IV} \\ m_{QI} & m_{QQ} & m_{QU} & m_{QV} \\ m_{UI} & m_{UQ} & m_{UU} & m_{UV} \\ m_{VI} & m_{VQ} & m_{VU} & m_{VV} \end{bmatrix} \cdot \mathbf{M}_{\rho \mathbf{X}} \cdot \begin{bmatrix} 1 \\ Q'_{tel} \\ U'_{tel} \\ V'_{tel} \end{bmatrix} . \tag{24}$$

This matrix equation produces four ordinary equations in which the the four elements of $\mathbf{S_{meas}}$ are expressed as complicated functions of **(1)** the 5 parameters defined above in §5.4, **(2)** the 3 fractional source Stokes parameters, and **(3)** the angle ρ_X. We have N sets of 4 equations like this, each for a different value of the parallactic angle χ. Given the N observed values of S_{meas} and χ, we could solve this set of equations using least squares. However, this approach would be complicated and cumbersome because the known values of parallactic angle χ, together with the 8 unknown parameters, are buried together within the 4N equations.

There is a simpler formulation. All of the information lies in the parallactic-angle dependence of the measured Stokes parameters $\mathbf{S'_{meas}}$. Moreover, the functional dependences of Q'_{meas}, U'_{meas}, V'_{meas} upon χ involve only the functions $\cos(2\chi)$ and $\sin(2\chi)$. Accordingly, it is straightforward to regroup and express $\mathbf{S_{meas}}$ in terms of the functions of the parallactic angle:

$$\begin{bmatrix} Q'_{meas} \\ U'_{meas} \\ V'_{meas} \end{bmatrix} = \begin{bmatrix} (m_{QI} + V'_{tel} m_{QV}) & (Q'_{tel} m_{QQ} + U'_{tel} m_{QU}) & (-Q'_{tel} m_{QU} + U'_{tel} m_{QQ}) \\ (m_{UI} + V'_{tel} m_{UV}) & (Q'_{tel} m_{UQ} + U'_{tel} m_{UU}) & (-Q'_{tel} m_{UU} + U'_{tel} m_{UQ}) \\ (m_{VI} + V'_{tel} m_{VV}) & (Q'_{tel} m_{VQ} + U'_{tel} m_{VU}) & (-Q'_{tel} m_{VU} + U'_{tel} m_{VQ}) \end{bmatrix} \cdot \begin{bmatrix} 1 \\ \cos(2\rho) \\ \sin(2\rho) \end{bmatrix} . \tag{25}$$

This provides 3 equations, one for each measured Stokes parameter; for example,

$$Q'_{meas} = A_{Q'} + B_{Q'} \cos 2\rho + C_{Q'} \sin 2\rho \, . \tag{26}$$

The coefficients (A, B, C) in equation 26 are expressed in the three columns of equation 25; for example, $B_{Q'} = (Q'_{tel} m_{IQ} + U'_{tel} m_{IU})$. We solve for (A, B, C) in each of the 3 equations separately. We refer to the set of these three separate fits as 'the *first* least-squares fit'. Figure 6 shows the data and the fits for a real-life example. One aspect of this solution requires discussion, which we provide in §6.2.

With three coefficients for each of the three observed Stokes parameters we have a total of 9 coefficients. These coefficients are complicated functions of the 5 parameters $(\alpha, \epsilon, \phi, \Delta G, \psi)$ and, also, the three source Stokes parameters $(Q_{tel}, U_{tel}, V_{tel})$. The *second* least-squares fit uses the 9 coefficients as the input data for a nonlinear least squares fit to determine the 5 Mueller matrix coefficients and, if desired, the 3 fractional Stokes parameters of the source.

6.2. A Detail of the least-squares fit to equation 26

Equation 26 fits the parallactic-angle dependence of the fractional Stokes parameters, e.g. $Q'_{meas} = Q_{meas}/I_{meas}$, making the implicit assumption that the uncertainty in I_{meas} is small

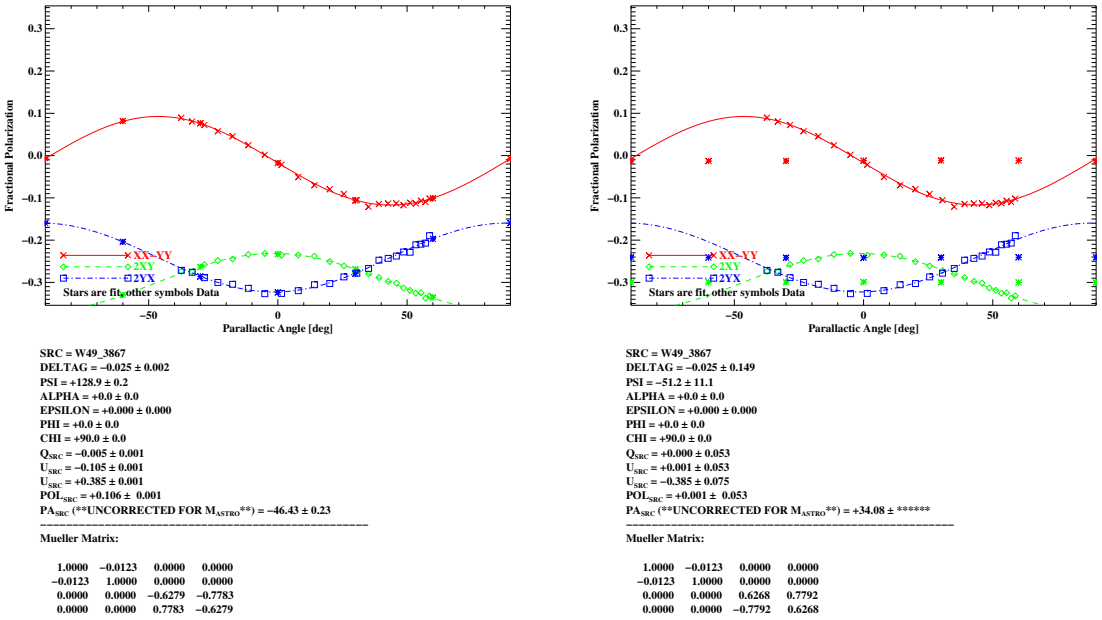


Fig. 6.— *Left,* a successful fit for channel 3867 of W49—the guessed value for α was close enough. *Right panel,* Not a good enough guess!

enough that the uncertainty in Q'_{meas} is dominated by the uncertainty in Q_{meas}, i.e. that $\sigma_{Q'_{meas}} = \sigma_{Q_{meas}}/I_{meas}$. This is not necessarily the case, particularly for the case of high fractional polarization.

Fitting for (A, B, C) using equations of condition $Q_{meas}/I_{meas} = A + B\cos(2PA) + C\sin(2PA)$ is bad practice in this case because points with abnormally low I_{meas} caused by noise receive improperly large weights in the fit. The solution to this problem is to fit $Q_{meas} = I_{meas}(A + B\cos(2PA) + C\sin(2PA))$. This fix regards I_{meas} as a known quantity having no noise, so its noise becomes embedded in the uncertainties of the derived quantities (A, B, C). But the improper weighting of noise-induced abnormally small values of I_{meas} is gone.

Our IDL procedure `onoffs_to_pacoeffs.pro` incorporates this fix. An older version, `stripfit_to_pacoeffs.p` does not. Unless Stokes I is noisy, both give almost identical results.

6.3. Fitting for Telescopes with Limited Range of ρ_X

For telescopes with limited range of ρ_X, we must replace the wide range of parallactic angle for a single source by using a set of sources having a wide range of polarization position angle. For the least-squares fit we would use the more cumbersome approach, mentioned above, and formulate equations of condition based on equations 24.

7. THE SUBTLETIES OF THE NONLINEAR FIT FOR THE MUELLER MATRIX PARAMETERS

7.1. Too many unknowns in the fits

The second fit for alt-az mounted telescopes in §6.1 uses the 9 outputs of the first fit (equation 26) as inputs to derive the Mueller matrix parameters. For the second fit, the full suite of unknowns comprises 8: 5 Mueller matrix parameters ($\Delta G,\ \psi,\ \alpha,\ \epsilon,\ \phi$); and 3 source parameters ($Q_{tel},\ U_{tel},\ V_{tel}$). If you solve for all 8 parameters, there is only one degree of freedom! This can make the solution unstable because of covariances. You can bypass this problem by reducing the number of unknowns. There are two straightforward approaches:

1. If you know any of the 3 source polarization parameters you can exclude them from the fit. This is particularly useful for continuum polarization calibrators such as 3C286, because they have $V = 0$; excluding only V_{tel} from the fit generally produces excellent fit results.

2. If you think the feed is good, with no crosscoupling between the feed outputs, then exclude ϵ and ϕ from the fit. And, especially for native linear feeds, exclude α also; setting it to zero is usually an excellent approximation. For native circular, however, you may need to include α (using a guessed value of $\pi/4$) because native circular feeds often have a substantial response to linear polarization—which is often significantly frequency-dependent.

7.2. Nonlinear Least-Square fits require reasonably good initial guesses

The two panels of Figure 6 show the fitting results for an elliptically-polarized OH maser component in W49, for two different guesses for α. On each figure, the $(XX - YY)$ curve is in red, with crosses); the XY curve is in green, with diamonds); and the YX curve is in blue, with squares. These curves and their points are the first fit above. The asterisks, which are not connected by lines, come from the second fit above.

Figure 6, left panel, is a successful second fit because the asterisks lie near their associated curves from the first fit. In contrast, Figure 6, left panel, shows an unsuccessful fit because the asterisks don't lie near first-fit curves. This is one of those cases where the fit converged not to a *minimum* of the sum-of-squares $\Sigma\sigma^2$ of the residuals, but a *maximum!* This can happen because the least-squares fit technique finds solutions for which the first derivative of $\Sigma\sigma^2$ with respect to the parameters is zero—whether or not this represents a minimum or a maximum.

The former second fit is successful because the guessed value of α, used as input for the fit, is close enough to the true value; the latter, not. By experimenting, we found that obtaining a successful fit required the guess for α to be within $\pm 1.06 \times \pi/2$ radians of its correct value (which is 2.25 radians). Whether or not this range for success is general is unknown. The message: *in all nonlinear fits, look at the results* and if they are unsatisfactory, try different input guesses!

7.3. Commentary on some Fundamentals

Nonlinear least squares fitting is often plagued by multiple minima, and the present case of the second fit is no exception. Heiles et al. (2001b) discuss these in some detail; here we only summarize some of the degeneracies.

1. For native linear polarization, one cannot distinguish between the two cases $(\alpha, \psi) = (0°, \psi_0)$ and $(\alpha, \psi) = (90°, \psi_0 + 180°)$. For these two cases, the signs of (Q, U) change, which is equivalent to rotating the feed by 90°. For a conventional linear feed, loosely described as two E-field probes in a circular waveguide, the combination $(\alpha, \psi) = (90°, 180°)$ is physically unreasonable.

2. For native circular polarization, the position angle of the source $2PA_{src}$ and ψ are inextricably connected. We can determine only their sum (for $\alpha = 45°$) or their difference. In particular, the two solutions $\alpha = 45°$, $\alpha = -45°$ are degenerate; the two solutions have different signs for Q_{src}, thereby rotating the derived PA_{src} by 90°).

   The physical interpretation of this degeneracy is straightforward: for a pure circular feed, the phase of a linearly polarized source rotates with 2ρ and its absolute value depends both on the system phase ψ and the source position angle PA_{src}. There is no substitute for an independent calibration of the linearly polarized position angle, either with a test radiator or with a source of known polarization.

3. ϵ is the quadrature sum of the ρ-independent portions of the two correlated outputs (AB, BA). This power is distributed between those outputs according to $(\phi + \psi)$. In the near-linear case, ψ can change by 180° by changing the choice for α, and this also produces a 180° change in ϕ.

4. Consider a high-quality standard linearly polarized feed that has a correlated cal connected by equal-length cables. Such a feed has $\alpha \approx 0°$ and the equal-length cables mean that $\psi \approx 0°$. However, the solution yields $\psi \approx 180°$ if the sign of $(XX - YY)$ is incorrect, which can easily happen if one interchanges cables carrying the two polarizations; this is equivalent to reversing the handedness of the position angle.

5. We have adopted the following procure for phase calibration. If there is a correlated cal then we measure ψ_{cal} and fit it to a constant plus a slope $\frac{d\psi}{df}$; we subtract this fit from the source phase and produce corrected versions of (XY, YX). Thus, the only component left in the correlated products is the difference between source and cal phase, which is the same as ψ.

6. Finally, at low frequencies, one must include the effects of terrestrial ionospheric Faraday rotation, which is time variable at a level ~ 1 rad m^{-2}.

8. USING RHSTK_2021 SOFTWARE—STAGE 1

8.1. Flagging Bad Data

Always begin by purging your dataset of bad data using visual inspection. The flagging process for all-Stokes data is best performed by looking at all four Stokes parameters simultaneously, because interference or bad data might show up in only one of the four auto- and cross-product spectra. We constructed a procedure that displays a set of all four Stokes spectra as four grayscale images. This makes it easy to examine all four simultaneously. It is always our philosophy that even if only one of the Stokes parameters is suspect, all four should be discarded.

This program is called `flag_rfi.pro`. It is best run with a large screen monitor because the grayscale images are displayed side-by-side.[6] It can be employed at any stage in the process: uncalibrated auto- and cross-product spectra, calibrated ones, Mueller-uncorrected Stokes spectra, Mueller-corrected Stokes spectra.

8.2. Stage 1 Basics

For the purposes of description, we implicitly assume native linear polarization. We therefore designate self-products by XX and YY and cross-products by XY and YX, as described in §5.2. After phase and amplitude calibration of the self- and cross-products, the Stokes parameters are given by equation 11. If the system has native circular polarization, then almost everything in this memo remains the same except for some obvious changes, such as (1) X and Y should be replaced by R and L in your mind's eye, and (2) the Stokes parameters are given by equation 12 instead of equation 11.

We assume you have 'signal' and 'reference' spectra (often called 'SRCON' and 'SRCOFF', or 'SIG' and 'REF'). It's the source *deflection*—SIG–REF—that matters. For continuum observations, the SRCON and SRCOFF spectra are nearly always position differences. For spectral line observations, they can be either position differences or frequency differences. We also assume that you have calibration data: DIODEON and DIODEOFF spectra. As for the source, it's the Diode *deflection* that matters.

The digital spectrometer produces uncalibrated self- and cross-products (spectra), which need

[6] We had a difficult time making `flag_rfi.pro` work on a MacBook Pro running OS X Lion. First, we need to tell OS X to allow IDL windows to get focus when clicked, so at a command line, type `defaults write com.apple.x11 wm_click_through -bool true` (this only needs to be done once ever). Then, if using Xquartz X11, go to Preferences and select "Emulate three button mouse" in the Input tab; in the Windows tab, select "Click-through Inactive Windows". These latter steps allow you to hold Option while clicking to emulate a middle mouse click, or Command while clicking to emulate a right mouse click. If using some other X11, there should be similar options in the Preferences for that version of X11.

to be turned into calibrated spectra. This is done with the Stage 1 software. There are three essential calibration quantities:

1. A representative value for CpK, which is an average over a portion of the bandpass specified by the spectral channels to include.
2. The linear relative phase XY gradient.[7]
3. The bandpass shape. After correcting for the bandpass shape, spectra should be flat unless there are spectral features.

8.3. Obtaining Counts per Kelvin (CpK) and relative phase between X and Y

`init_cal.pro` uses two IDL procedures to derive items 1 and 2 above, the system gain and relative phase:

1. `intensitycal_self.pro` derives the correlator counts per Kelvin (`CpK`) averaged over a set of spectral channels (the array `gainchnls`) from the Diode deflection for a single self-product, e.g. for XX. This `CpK` factor is what's usually considered as the 'system gain'. You must invoke this separately for XX and YY.
2. `phasecal_cross.pro` derives the relative phase between XY and YX diode deflections, which varies linearly with frequency as represented by the parameters `zero`, `slope`, and the array of frequencies of the spectral channels. These frequencies can be the RF or baseband (BB) frequencies. In the past we used BB, but we now recommend using RF frequencies because of the wide bandwidths of today's digital spectrometers.

   We least-squares fit the cross-product spectra for the slope and phase of the diode deflection. This is a nonlinear fit, so you must enter a reasonably accurate guessed value for the slope (`dpdf`). If your guess is too far off the fit will not converge to the correct value, so you need to plot the results and check! To check, use `phaseplot_08.pro`.

8.4. Applying the System Gain and Phase to produce calibrated voltage products

We need to apply the above-derived system gain and phase response to the ONSRC/OFFSRC data to produce intensity- and phase-calibrated self- and cross-product spectra that are bandpass-corrected. We use the following three procedures.

[7]We have not yet implemented the filter phase correction of §4.3, item (4).

1. `self_uncal_to_cal` derives intensity-calibrated and bandpass-corrected source deflection spectra, by applying the above-determined `CpK` to uncalibrated spectra. The input spectra are arrays of uncalibrated SRCON and SRCOFF spectra, with dimensions `[nchnls, nONspectra]` and `[nchnls, nOFFspectra]`.

   The SRCON and SRCOFF spectra can be paired, or the SRCOFFs can be averaged or medianed. Some people have the erroneous impression that averaging the SRCOFFs produces less noise than using paired SRCONs and SRCOFFs, because then each SRCON is divided by a much less noisy SRCOFF spectrum. However, when you average all the data, this isn't true: the noise in the average is the same (to first order). Generally, using paired SRCONs and SRCOFFs is better, because the SRCOFF spectra might change significantly with time for various and sundry reasons. Hence, we recommend using paired SRCONs and SRCOFFs. You do this by setting the keyword `mean_med_off = 0`.

   The calibration that this procedure performs includes (1) converting correlator counts to Kelvins, which is done with `CpK`; and (2) dividing by the bandpass. You must invoke this procedure separately for XX and YY. If you are doing in-band frequency switching, then for each spectral path (like XX) you must invoke this twice, once with (LO_1, LO_0) as the (SRCON, SRCOFF), and again with (LO_0, LO_1) as the (SRCON, SRCOFF).

2. `cross_uncal_to_cal` derives `cal_cross`, the intensity-calibrated and bandpass-corrected spectra for a cross-product, e.g. for XY. The inputs are the array of uncalibrated XY spectra, `gainchnls`, the XX and YY `CpK` (from `init_cal.pro`), and the XX and YY bandpasses (from `self_uncal_to_cal`). The process of calibration includes (1) converting correlator counts to Kelvins, which is done with the geometric mean of the XX and YY values of `CpK`; and (2) dividing by the geometric mean of the XX and YY bandpasses.

   The calibrated cross spectra `cal_cross` are *not source-deflection spectra.* Rather, they are the phase-corrected versions of the input cross-product spectra, which are otherwise uncorrected.

3. `phasecorr_xyyx` removes the linear phase gradient *determined from the diode deflection* from *each individual pair* of the *original* XY and YX spectra. Inputs are the set of all phase-uncorrected XY and YX spectra, the frequency array, and the derived zero and slope from `phasecal_cross.pro`; the outputs are are the set of all phase-corrected XY, YX products (where 'phase-corrected' means the relative phase of the diode deflection between X and Y is zero).

   Because these phase-corrected cross product spectra have zero relative phase delay between the X and Y signal paths for the diode signal, it is as if *all 3 of the following* were true:

   (a) the length difference of the cables connecting TCAL in Figure 4 were zero, and

   (b) the cable length difference ($D_X - D_Y$) in Figure 4 were zero.
   These two mean that

   (c) the relative phase delay between the X and Y signal paths for the *astronomical soource* depends only on **(1)** the cable length difference ($L_X - L_Y$) minus **(2)** the length difference

between the cables connecting TCAL in Figure 4, and **(3)** the relative phase difference in the feed itself. These 3 phase differences lie in the thick lines in Figure 4, meaning that they are stable in time and therefore are suitably incorporated in the M_{RX} parameter ψ.

4. `products_to_stokes_2021.pro` creates the $\mathbf{S_{meas}}$ Stokes parameters from the calibrated voltage products using equation 11.

We use the run-time procedure `stg1.idlprc.pro` to perform all the necessary calculations for Stage 1; see §11.

9. USING RHSTK_2021 SOFTWARE—STAGE 2: DERIVING $\mathbf{M_{RX}}$ AND APPLYING THE MUELLER MATRICES

Stage 2 uses the output from Stage 1. There are two actions we can take: **(1)** use observations of a linearly polarized source to *derive* $\mathbf{M_{RX}}$; and use observations of any source to *apply* the Mueller matrix transformations to obtain the polarization properties of that source. Given observations of polarized calibration sources, we have 3 IDL procedures that fit the data to *derive* the Mueller matrix elements. They are useful for 3 different types of observing.

9.1. Deriving the Mueller matrix coefficients using `mmfit_2016.pro` or `mmfit_2016_chisq.pro`

Given a single polarized source observed over a range of parallactic angles, this least squares fits the data to provide the Mueller matrix parameters. The two versions are identical except one does a least-squares fit and the other a chi-square fit; the latter automatically accounts for the increased noise in Stokes Q over Stokes U. This is our classic fitting procedure, modified to allow nonzero Stokes V. It is intended for the very common case where you have well-sampled parallactic angle coverage of a single polarization calibration standard source. As of Aug2016, this nonlinear fit is done with `mmfit_2016.pro` or `mmfit_2016_chisq.pro`. The major improvements over previous versions are (1) we now include the possibility of a nonzero Stokes V for the source(s) being fit, which is necessary when using OH masers and pulsars for Mueller matrix calibration, and (2) we allow you to exclude any combination of parameters from the fit, which is handy for high-quality feeds and sparse measurements.

The input data consist of the ρ dependences of each of the 3 observed fractional Stokes parameters, which are expressed by 3 coefficients in equation 26 for each observed Stokes parameter. The coefficients (A, B, C) are fitted to the observed fractional Stokes parameters by `onoffs_to_pacoeffs`. Thus, there are 9 input numbers (3 sets of (A, B, C), one for each of the three fractional Stokes parameters).

These 9 coefficients then serve as the inputs to the nonlinear fit for the Mueller matrix parameters and source Stokes parameters. There are up to 8 parameters to fit: 5 Mueller matrix parameters and 3 source Stokes parameters (Q', U', V'). You can choose any combination of the 8 parameters to fit. For example, if you know the source Stokes parameters, you specify that they not be fit. Because the number of fitting parameters can be as high as 8, and there are only 9 inputs, the more parameters you can fix, the better. As we discussed in §7.1, when you fit for all 8 parameters there is only 1 degree of freedom, and the degeneracies (covariances) might be strong and the fit can become unstable. It is probably best that you do not include all 8 parameters in the fit. In our classic fits, we forced the Stokes V parameter to equal zero, because our calibration sources were classic continuum sources with very little V; this worked well. Now, however, we might want to use OH masers or pulsars, which can have strong Stokes V.

The nonlinear least squares fit requires initial guesses for the parameters to be solved for. It is OK to set the initial guesses for the source Stokes parameters to zero. However, if you set the initial guesses for 5 the Mueller matrix parameters to zero and fit for all 8 parameters, the fit will not properly converge. The initial guesses for the 5 matrix parameters need to depart a bit from zero; our default is 0.001. These are generated by the program `mm_coeffs_in_setup.pro`. Alternatively, you can exclude one or more parameters from the fit, setting its value by hand equal to a constant.

With nonzero V, you should eliminate other parameters from the fit. The most important Mueller matrix parameters are ΔG and ψ, and you always need to solve for them. Because a reasonably well-engineered feed will have small imperfections, the parameters that express these imperfections can be eliminated from the fit. These parameters are ϵ and its associated phase angle ϕ; and α and its associated angle χ. Almost always, ϵ is the least important parameter because it is the smallest. For any native polarization you can set $\epsilon = 0$ and its angle to anything and not solve for them. This should be enough to obtain a good fit. You can reduce covariance problems further by fixing α: for native dual linear you set $\alpha = 0°$ and $\chi = 90°$; for dual circular, $\alpha = 45°$ and $\chi = 90°$. The program never allows a fit for χ, which should always be set to 90°.

We use the run-time procedure `stg2.idlprc.pro` or `stg2_2016_chisq.pro` to perform all the necessary calculations for Stage 2, which uses `mmfit_2016.pro` or `mmfit_2016_chisq.pro`; see §11.

9.1.1. `mmfit_2016_multiplesources.pro`

This is like `mmfit_2016.pro`, but treats the case when you have observations of more than one source and want to derive Mueller matrix parameters by including the data from all those sources. You can also obtain the Stokes parameters for all, or any combination of, the sources.

The most important use for this is when you have well-sampled ρ coverage of a polarized OH maser, e.g. W49, for which you can regard each spectral channel as an independent source. The program then provides not only the Mueller matrix parameters, but also the 3 Stokes parameter spectra. For each source (or spectral channel), the input data consist of the ρ dependence of each of

the 3 observed fractional Stokes parameters, which is expressed by 3 parameters for each observed Stokes parameter—a total of 9 data points for each source. The total number of input datapoints is then $9N_{src}$.

If you don't know the source polarizations, then you are fitting for 5 Mueller matrix elements plus 3 Stokes parameters per source. The total number of fitting parameters is $5 + 3N_{src}$. The number of degrees of freedom is the difference between these two numbers, which is $6N_{src} - 5$. So having only just two sources instead of one has a big effect because it changes the number of degrees of freedom from 1 to 7. Nevertheless, if the feed is good you still might want to not fit for ϵ and ϕ, and maybe also α. We have little experience for real-life data, so explore.

A caution for large N_{src} (e.g., hundreds or thousands of channels in a single spectrum): The number of parameters being solved for is $5 + 3N_{src}$. Suppose $N_{src} = 1000$, i.e. you have a three 1000-channel spectra (one for each fractional Stokes parameter)—so $3N_{src} = 3000$. The covariance matrix is then a 3005 x 3005 matrix, and the program has to evaluate its inverse. The computer time required for matrix inversion goes as $(3N_{src} + 5)^3$. !!CUBED!! When N_{src} exceeds a few hundred, the time can become excessive. You can deal with this in a two-step fashion: step 1, rebin the spectra so that you have fewer channels, and do the fit at lower spectral resolution; step 2, apply the derived Mueller matrix to the spectra with full resolution.

9.1.2. `mparamsfit.pro`

This program determines Mueller matrix parameters when you have observations of 2 or more sources and little or no parallactic angle coverage. You have to know the polarizations of the observed sources. For example, if you have observations of several continuum polarization standard calibrators, each at 1 or more parallactic angles, this will provide the Mueller matrix parameters. But it cannot provide the source Stokes parameters; you have to know them and provide them as inputs.

The FAST and GALT telescopes need this program because of their limited or zero coverage in parallactic angle. The sources can be continuum polarization calibrators or a known OH maser spectrum; the latter is especially convenient because you don't have to do position switching.

For each observed source, the input data are its 3 observed fractional polarized Stokes parameters, so the number of input datapoints is $3N_{src}$. There are five Mueller matrix parameters to solve for, so with 2 sources there is 1 degree of freedom (but see paragraphs below). As in §7.1, this might not be enough, so you might need to exclude ϵ and ϕ, and maybe also α, from the fit. The more sources, the better; the fewer parameters to solve for, the better.

9.2. Applying the Mueller Matrix

When you observe a source with the intent of measuring its polarization, you want to convert $\mathbf{S_{meas}}$ to either the telescope frame ($\mathbf{S_{tel}}$) or the IAU frame ($\mathbf{S_{IAU}}$). This is done with `mmcorr.pro`, which uses equation 13 to convert the measured ones to the astronomical ones by applying the inverse, $\mathbf{M^{-1}_{tel-meas}}$:

```
mmcorr, mm_rx, mm_tel_iau, parallactic_deg, stk_meas, stk_corr, $
        /m_rho_x, /m_tel_iau
```

Note the keywords `m_rho_x` and `m_tel_iau`. Setting them applies $\mathbf{M_{\rho x}}$ and $\mathbf{M_{tel-IAU}}$, respectively. Not setting `m_rho_x` removes the correction for position angle, which keeps Stokes Q (the difference $XX - YY$) and U (the cross product XY) separate; this can be useful when trying to detect very weak polarization because the cross products are unaffected by gain errors. Not setting $\mathbf{M_{tel-IAU}}$ provides results in the telescope frame. The defaults are *not* to set these keywords.

10. AN ILLUSTRATIVE TEXTBOOK EXAMPLE: POSITION-SWITCHED DATA

Let's see how the above procedures are used together to determine the polarized Stokes parameters for on-off observations of an astronomical source.

10.1. Stage 1

First, deal with the Diode deflections. Suppose we have 10 DIODEON and DIODEOFF spectra, each with 32 channels, in each of two polarizations `xx, yy`; these are the outputs from either a native-linear or native-circular feed. The `freq` array has 32 elements. Then we have the self-product arrays `diodeonxx[32,10], diodeoffxx[32,10]` and `diodeonyy[32,10], diodeoffyy[32,10]`; and we have the cross-product arrays `diodeonxy[32,10], diodeoffxy[32,10]` and `diodeonyx[32,10], diodeoffyx[32,10]`. We will use all 32 channels in the calibration. We have Diode values `tcalxx, tcalyy`. We invoke `init_cal`, which calculates the factors `CpK` and phase correction from the cal deflection using the 3 basic procedures in §8:

```
init_cal, gainchnls, tcalxx, tcalyy, freq, dpdf_guess, $
        diodeonxx, diodeoffxx, diodeonyy, diodeoffyy, $
        diodeonxy, diodeoffxy, diodeonyx, diodeoffyx, $
        cts_per_k_xx, cts_per_k_yy, $
        phase_zero, phase_slope, phase_cntr, rf=0
```

Next, apply these derived calibration parameters to the data. Suppose we have 60 on-source and off-source self-product arrays in each of two polarizations `srconxx[32,60], srcoffxx[32,60]` and `srconyy[32,60], srcoffyy[32,60]`; and we have the cross-product arrays `srconxy[32,60], srcoffxy[32,60]` and `srconyx[32,60], srcoffyx[32,60]`.

Finally, we have 10 values of `cts_per_k_xx` that apply to the 60 source measurements, so we need a 60-element integer array telling which `cts_per_k_xx` element to use for each src measurement; call this 60-element array `cal_indx`.

We invoke `products_to_stokes_2021` to use the results from `init_cal` to calculate the measured Stokes parameters $\mathbf{S_{meas}}$:

```
products_to_stokes_2021, gainchnls, cal_indx, cts_per_k_xx,, cts_per_k_yy, $
        phase_zero, phase_slope, freq, $
        srconxx, srcoffxx, srconyy, srcoffyy, $
        srconxy, srcoffxy, srconyx, srcoffyx, $
        stki_meas, stkq_meas, stku_meas, stkv_meas
```

Note that $\mathbf{S_{meas}}$ is the Stokes array version of the 4 Stokes parameters `stki_meas`, ...

10.2. Stage 2

Stage 2 uses the output of Stage 1, $\mathbf{S_{meas}}$, to either (1) derive the Mueller matrix coefficients, or (2) to apply the Mueller matrix to obtain $\mathbf{S_{tel}}$ or $\mathbf{S_{IAU}}$. For the latter, we use

```
stk_meas= [stki_meas, stkq_meas, stku_meas, stkv_meas]
mmcorr, mm_rx, mm_tel_iau, parallactic_deg, stk_meas, stk_iau, $
        /m_rho_x, /m_tel_iau
```

This converts from the measured values to the telescope system by applying the inverse of `mm_tel_meas`. Setting the keyword `m_rho_x` applies the $\mathbf{M_{\rho x}}$ correction; default is *not* to set it. Setting the keyword `m_tel_iau` converts from the telescope system to the IAU system by applying `mm_tel_iau`; again, default is *not* to set it.

11. A REAL-LIFE EXAMPLE: POSITION-SWITCHED DATA on 3C286

We illustrate the use of RHSTK_2021 software for both deriving and applying the Mueller matrix coefficients for a straightforward set of ON/OFF data on the linearly-polarized calibration source 3C286 observed with the GBT at multiple parallactic angles. This was an attempt to

measure Zeeman splitting of the $z = 0.692$ HI absorption line against 3C286. The initial publication incorrectly claimed a detection (Wolfe et al. 2008Natur.455..638W), but the polarization was linear instead of circular and therefore not Zeeman splitting (Wolfe et al. 2011ApJ...733...24W); the problem was an error in an earlier version of the RHSTK software.

We will take 3 steps: **(1)** use the 3C286 data to derive $\mathbf{M_{RX}}$; **(2)** apply the matrix corrections to determine the polarized Stokes parameters of 3C286; **(3)** re-derive $\mathbf{M_{RX}}$ using the Mueller-corrected original data. Using the Mueller-corrected data for **(3)** means that the re-derived $\mathbf{M_{RX}}$ parameters should be zero, and the re-derived $\mathbf{M_{RX}}$ itself should be unitary, so this serves as a check on our techniques.

This example resides in `rhstk/rhstk_examples/pswitch/artiedata`. The datafile contains two structures, `src` (58 elements) and `cal` (15 elements). The `src` data are position switched, with alternate scan numbers being on and off source. Thus, `src[0]` is on source, `src[1]` is off source, etc. The `cal` data are interspersed with the `src` data. The datastream begins with a Cal measurement, which consists of two spectra, one with Diode on and one with Diode off; this has `cal.scannum`=1. Every fifth scannum is a Cal measurement: thus, `src.scannum`=[2,3,4,5,7,8,9,10,12,13,14,15,...] and `cal.scannum`=[1,6,11,...]. Thus, the observing consists of groups of five scan numbers, the first of which is a Cal measurement and the next four are source measurements. Each `cal` measurement has a single subscan, which contains both the DIODEON and DIODEOFF measurements. Each `src` measurement has 237 subscans, all of which have identical observing parameters (such as being on source or off source).

You can distinguish between on source and off source scans by the `src.subscan.cra2000`, which is the right ascension in hours; on-source measurements are at higher right ascension. The idea of this position-switched measurement was to cover the same hour-angle range for the on-source and off-source measurements within each pair. Each source scan lasts 4 minutes of time, so the on- and off-source right ascensions are separated by one degree. Within each on/off pair, the on-source data were taken first, so covering the same hour angle range within each pair would require the off-source right ascension to be higher than the on-source one; the opposite occurred, which was a mistake in setting up the observing file.

The calibrated Stokes parameters refer to the source deflection. Thus there are half as many calibrated spectra as there are on- and off-source uncalibrated spectra. The output of this Stokes calibration calculation consists of 29 switched, calibrated sets of 237 spectra apiece—a total of 6873 position-switched spectra. These are derived from 29 pairs of on- and off-source `src` spectra.

11.1. The Stage 1 IDL commands for this example

Our goal in this example is **(1)** we determine the $\mathbf{M_{RX}}$ coefficients using `mmfit_2016.pro`; **(2)** we use the derived $\mathbf{M_{RX}}$ in `mmcorr.pro` to obtain the 3C286 Stokes parameters (using the very same 3C286 data). This illustrative example serves as a check on the accuracy of the derived

$\mathbf{M_{RX}}$. We do these for both `mmfit_2016.pro`. and `mparamsfit.pro`.

Run the IDL session in `rhstk_2021/rhstk_examples/mueller/artiedata`. Look at the README file, which tells how to invoke the associated command files and run-time files; it also provides some commentary. More detailed documentation is in the run-time procedures (such as `stg0.idlprc.pro`) and the procedures (such as `mmcorr.pro`).[8]

The following IDL commands sum the measured Stokes parameters over channels to create a 'continuum' version and divides the 3 polarized Stokes parameters by Stokes I to create fractional continuum Stokes parameters. It uses them to derive the Mueller matrix parameters, and plots the input data and results in Figure 7.

```
IDL> .run stg0.idlprc   ;reads the datafile and defines some
        quantities using info in the datafile.

IDL> @init_params.idl   ;defines parameters needed for running the stage
        1 and 2 software.

IDL> .run stg1.idlprc   ;does all the Stage 1 calculations
```

11.2. The Stage 2 IDL commands, using `mmfit_2016.pro` or `mmfit_2016_chisq.pro`

In Stage 2 we first derive the Mueller matrix coefficients using `mmfit_2016.pro`.

```
IDL> .run stg2.idlprc   ;does all the Stage 2 calculations using least
;        squares; see plot in Figure 7 left panel.

or

IDL> .run stg2\_chiaq.idlprc   ;does all the Stage 2 calculations
;       using chi-square; see plot in Figure 7 right panel.
```

We then apply the Mueller matrices to obtain the calibrated Stokes parameters. finally, we check the results by re-deriving the Muller matrix coefficients form the Mueller-corrected data; the derived corrections should be zero. First, we rerun to produce left panel of Figure 8:

```
IDL> m_rho_x=0 ;setup for runniing stg3.idlprc below; tells mmcorr
```

[8]A command file with name `command.idl.pro` is initiated in IDL by typing `@command.idl`; a run-time file with name `runtime.idlprc.pro` is initiated by typing `.run runtime.idlprc`.

```
        to apply mm_RX but not mm_rho_x.
IDL> .run stg3.idlprc   ;applies Mueller matrix corrections (using
        mmcorr.pro) and sets up inputs to rerun stg2.idlprc using
        the Mueller-corrected data as input.
IDL> .run stg2.idlprc   ;rerun stg2.idlprc and produce the figure.
```

The resulting Figure 8, left panel, plots $\mathbf{M_{RX}^{-1}} \cdot \mathbf{S_{meas}}$, which are the actual Stokes parameters of 3C286 with position angle rotated by the parallactic angle. The $\mathbf{M_{RX}}$-corrected measured Stokes parameters are amplitude- and phase-corrected, so when Q and U are plotted vs. parallactic angle the amplitudes should be identical and they should be 45° apart; also, V should be zero. Such is the case.

Finally, we produce the right panel of Figure 8:

```
IDL> m_rho_x=1 ;in stg3.idlprc below, tells mmcorr
        to apply mm_RX and also mm_rho_x.
IDL> .run stg3.idlprc   ;applies Mueller matrix corrections (using
        mmcorr.pro) and sets up inputs to rerun stg2.idlprc using
        the Mueller-corrected data as input.
IDL> .run stg2.idlprc   ;rerun stg2.idlprc and produce the figure.
```

The resulting Figure 8, right panel, plots the fully-corrected Stokes parameters $\mathbf{S_{tel}}$. These values should be independent of parallactic angle, which is the case. Evidently our calibration scheme works!

11.3. The Stage 2 IDL commands, using `mparamsfit.pro`

We do
`IDL> .run domparamsfit.idlprc.pro`,
which produces the output structures `mmcoeffs_out` and `sigmmcoeffs_out`. These are substantially the same as the output structures of `mmfit_2016.pro`.

Acknowledgments: This work was supported in part by NSF grant AST-0908572. Support for this work was also provided by the NSF to T. R. through awards GSSP 05-0001, 05-0004, and 06-0003 from the NRAO, and to A. K. through GBT student awards GSSP07-0001, GSSP07-0002, GSSP07-0003, GSSP07-0019. Portions of this work were performed when A. K was employed by University of Wisconsin, Madison, and University of Virginia, Charlottesville.

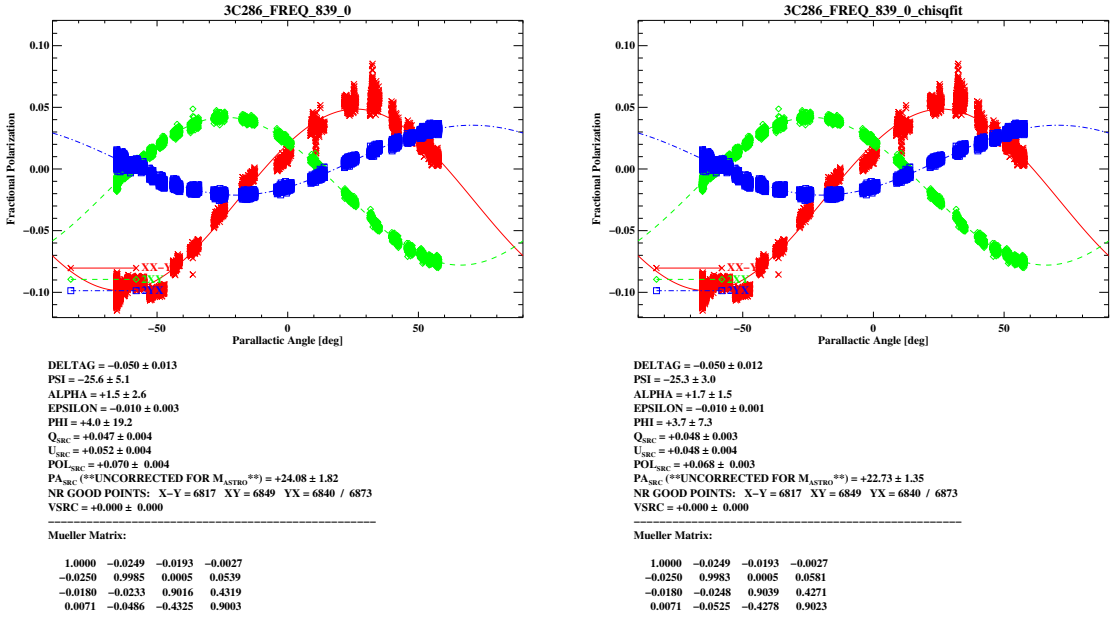


Fig. 7.— The 3 polarized components of $\mathbf{S_{meas}}$ for 3C286 versus parallactic angle. As in equation 11, these components are the intensity- and phase-calibrated voltage products $XX-YY$ (red), $2XY$ (green), and 2YX (blue). The least-squares derived Mueller matrix parameters, which determine $\mathbf{M_{RX}}$, are listed below the figure. *Left panel:* from `mmfit_2016.pro`. *Right panel:* from `mmfit_2016_chisq.pro`.

REFERENCES

Aller, H.D., Aller, M.F., Hughes, P.A. 1996, in Blazar Continuum Variability, Astronomical Society of the Pacific Conference Series, 1996, 110, 208.

Bode, H. W. 1940, "Relations Between Attenuation and Phase in Feedback Amplifier Design", Bell Sys. Tech. J., 19, 421

Boothroyd, A. I., Blagrave, K., Lockman, F. J., Martin, P. G., Pinheiro Gonalves, D., & Srikanth, S. 2011, "Accurate galactic 21-cm H I measurements with the NRAO Green Bank Telescope", A&A, 536, 81

Burn, B. J. 1966, "On the depolarization of discrete radio sources by Faraday dispersion", MNRAS, 133, 67

Coles, W. A., & Rumsey, V. H. 1970, "Stokes Parameters for OH Sources", ApJ, 159, 247

Heiles, C., & Fisher, R. 1999, "Calibrating the GBT for Spectral Polarimetry Using Cross Correlation", NRAO electronics Division Internal Reports 309. Unfortunately, this is not available on NRAO's web page. For an electronic copy, see http://astro.berkeley.edu/~heiles/handouts/handouts_radio.html

Heiles, C., Goodman, A. A., McKee, C. F., & Zweibel, E. G. 1993, in Protostars and Planets III, ed. E. H. Levy & J. I. Lunine (Tucson: Univ. Arizona Press), 279

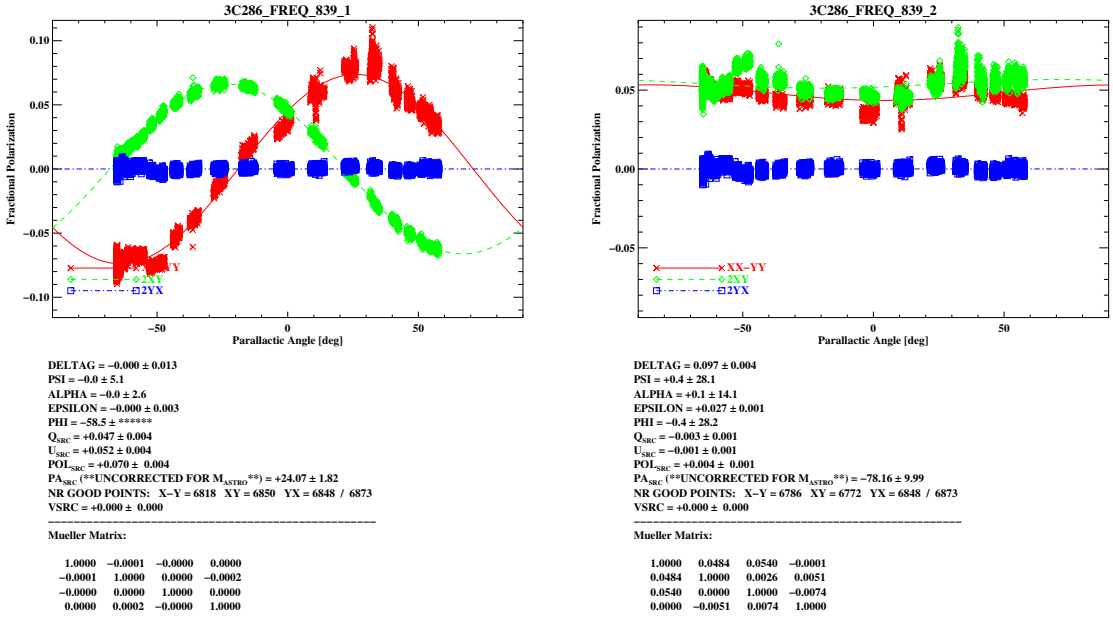


Fig. 8.— The 3 polarized components of $\mathbf{M_{RX}^{-1}} \cdot \mathbf{S_{meas}}$ for 3C286 versus parallactic angle. *Left panel:* As in equations 13 and 14, these are equal to $\mathbf{M}_{\rho_\mathbf{X}} \cdot \mathbf{S_{tel}}$, which are the actual Stokes parameters of 3C286 with position angle rotated by the parallactic angle. *Left panel:* As in equations 13 and 14, these are equal to $\mathbf{S_{tel}}$, which are the actual Stokes parameters of 3C286 in the telescope frame.

Heiles, C., Perillat, P., Nolan, M., Lorimer, D., Bhat, R., Ghosh, T., Howell, E., Lewis, M., O'Neil, K., Salter, C., & Stanimirović, S. 2001a, "All-Stokes Parameterization of the Main Beam and First Sidelobe for the Arecibo Radio Telescope", PASP, 113, 1247

Heiles, C., Perillat, P., Nolan, M., Lorimer, D., Bhat, R., Ghosh, T., Lewis, M., O'Neil, K., Salter, C., & Stanimirović, S. 2001b, "Mueller Matrix Parameters for Radio Telescopes and Their Observational Determination", PASP, 113, 1274

Heiles, C., Robishaw, T., Troland, T., & Roshi, A. 2003, "Calibrating the GBT at L, C, and X Bands", available at http://www.gb.nrao.edu/~rmaddale/GBT/Commissioning/index.html

IAU 1974, in Transactions of the IAU, Vol. XVB 1973, Proceedings of the Fifteenth General Assembly, ed. G. Contopoulos & A. Jappel (Dordrecht: Reidel), 165

Kraus, J. D. 1966, Radio Astronomy (1st ed.; New York: McGraw-Hill)

Kuhr, H., Witzel, A., Pauliny-Toth, I.I.K., Nauber, U. 1981, A&A Suppl, 45, 367.

Manchester, R. N. 1972, "Pulsar Rotation and Dispersion Measures and the Galactic Magnetic Field", ApJ, 172, 43

Margot, Jean-Luc 2021, 'A Data-Taking System for Planetary Radar Applications', Journal of Astronomical Instrumentation, Volume 10, Issue 1, id. 2150001

O'Donnell, M., Jaynes, E. T., & Miller, J. G. 1981, "Kramers-Kronig relationship between ultrasonic attenuation and phase velocity", J. Acoust. Soc. Am., 69, 696

Robishaw, T. 2008, "Magnetic fields near and far: Galactic and extragalactic single-dish radio observations of the Zeeman effect", Ph.D. Thesis, University of California, Berkeley

Robishaw, T., & Heiles, C. 2009, "On Measuring Accurate 21 cm Line Profiles with the Robert C. Byrd Green Bank Telescope", PASP, 121, 272

Wolfe et al. 2008Natur.455..638W

Wolfe et al. 2011ApJ...733...24W