<div align="center">

VLA Scientific Memo 163

Possible Algorithms to Improve the VLA's Polarization Performance

M.A. Holdaway, C.L. Carilli, and F. Owen
National Radio Astronomy Observatory
Socorro, NM 87801

December 8, 1992

</div>

1 INTRODUCTION

Around the time the VLA was commissioned, a great deal of work was done on the polarization properties of the VLA antennas and computer algorithms to correct for the spurious instrumental polarization. In the last ten years little has been done largely due to the lack of manpower and the lack of fast computers. Algorithms which address problems in total intensity imaging have continued to develop over the years, and it has been assumed that polarization imaging can "piggyback" on these advances. In many cases, improvements in the total intensity algorithms do not improve the polarization imaging capabilities because some of the problems with polarization are unique. It is time to reexamine the issues in polarization imaging and possible algorithmic solutions to the problems we still face.

Here is a list of some outstanding problems with polarization at the VLA. Many will also be problems for the VLBA.

- Determination of the instrumental cross-polarization (D) terms. The standard approach to this problem uses the rotation of the astronomical signal with antenna parallactic angle and the non-rotation of the corrupting D terms to separate the two (Bignell, 1982). This method requires a polarization calibrator, preferably a bright point source, to be observed over a range of parallactic angle, which usually requires that the observer have a block of time of at least four hours. However, this method cannot correctly treat time dependent D terms. We will present evidence for large temporal fluctuations in the D terms.

- RCP-LCP phase drifts. Currently, the D term solutions and the polarization position angle calibration are only as good as the RCP-LCP phase stability of the reference antenna.

- Total Intensity. The instrumental D terms corrupt the total intensity visibilities as second order terms. For highly polarized sources, these second order terms may limit the dynamic range of the total intensity images.

- Beam Squint. Because of the off-axis cassegrain feed system, the left and right circular polarized beams are offset by about 6% of the FWHM beam width, greatly limiting the accuracy of circular polarization measurements. Recently, some work has been done on correcting for the beam squint in the image plane (Vourlidas, 1992). The beam squint also feeds back into a proper treatment of large field imaging of linear polarization.

- Polarization sidelobes. The polarization properties of the antenna change with position in the beam. At the present time, only the on-axis instrumental polarization is corrected for. The off-axis instrumental polarization will lead to spurious polarization features and sidelobes in the polarized image due to the spurious polarization signal changing in time. Hence, polarization measurements are often restricted to within the beam's half power point, and attempts to perform wide field polarization imaging usually meet with great difficulty (Carilli and Holdaway, 1992)

- Ionospheric Faraday Rotation. Daytime observations at L band can suffer from position angle drifts of 10 degrees or more (Bignell, 1982), and P band polarization is unexplored.

While these problems are rooted in the atmosphere or in hardware which are not likely to improve, they all have algorithmic solutions. In the past, it has been assumed that some of these algorithmic solutions were too computationally cumbersome or expensive. However, our experience with the MMA simulations in SDE indicate that such algorithmic solutions are now within reach. Now is the time to take a fresh look at how these problems affect VLA images, what improvement is possible, and how much computer power these algorithms will require.

2 D Term Variation

The most important issue in polarization calibration is the instrumental polarization, also known as the cross polarization leakage, or "D" terms. Assuming the feeds are sensitive to a linear combination of the nominal polarization as well as the orthogonal polarization, the signals detected from the R and L feeds from antenna j will be

$$R_j = E_{R_j} + D_{R_j} E_{L_j} \tag{1}$$
$$L_j = E_{L_j} + D_{L_j} E_{R_j}. \tag{2}$$

Neglecting the antenna gains, the RR correlation is

$$V_{R_i R_j} = RR e^{-i(\chi_i - \chi_j)} + RL D_{R_j}^* e^{-i(\chi_i + \chi_j)} + LR D_{R_i} e^{i(\chi_i + \chi_j)} + LL D_{R_i} D_{R_j}^* e^{i(\chi_i - \chi_j)}, \tag{3}$$

and the RL correlation is

$$V_{R_i L_j} = RL e^{-i(\chi_i + \chi_j)} + RR D_{L_j}^* e^{-i(\chi_i - \chi_j)} + LL D_{R_i} e^{i(\chi_i - \chi_j)} + LR D_{R_i} D_{L_j}^* e^{i(\chi_i + \chi_j)}, \tag{4}$$

where $RL = \tilde{Q} + i\tilde{U}$, $LR = \tilde{Q} - i\tilde{U}$, $RR = \tilde{I} + \tilde{V}$, $LL = \tilde{I} - \tilde{V}$, and $\tilde{I}$ is the Fourier transform of the I Stokes parameter image evaluated at the appropriate value of (u, v) for the i, j baseline,

D_{R_i} is a complex number indicating the leakage of LCP into the RCP feed for antenna i, χ_i is the parallactic angle for antenna i, and $V_{R_iL_j}$ is the measured correlation between the RCP signal from antenna i and the LCP signal from antenna j, not corrected for parallactic angle. For weakly polarized sources, the first three terms are all of the same magnitude, and the last term is much weaker.

Cotton (1992) also analyses the problem of feeds sensitive to general elliptically polarized radiation. The D term leakage formulation of the problem is exactly analogous to the ellipticity-orientation formulation, where the D terms are related to the ellipticity θ and the orientation ϕ as

$$D = \tan(\theta \pm \pi/4)e^{\mp i2\phi}, \tag{5}$$

where the top sign is for D_R and the bottom sign is for D_L. The main difference between the D term polarization calibration (opcode = 'APPR' in the AIPS task PCAL) and the orientation-ellipticity polarization calibration (opcode = 'ORI-' in PCAL) is that the former neglects second order terms in D (ie, $LR D_{R_i} D_{L_j}^* e^{i(\chi_i + \chi_j)}$ in Equation 4). Neglecting second order terms results in a set of linear equations which can be easily solved, but there is no fundamental reason why a nonlinear least squares optimization could not be applied to the full D term formulation to obtain the same calibration as the ellipticity-orientation formulation. The trigonometric functions in the elipticity-orientation representation may make the optimization somewhat ill-conditioned, and the full nonlinear D term optimization may yield better solutions.

What sorts of image errors result from errors in the D terms? First, consider the case of uncorrelated D terms. By assuming no circular polarization, identical parallactic angles, and neglecting the last term in Equation 4, the measured visibility is approximated by

$$V_{R_iL_j} \simeq RR\,(D_{R_i} + D_{L_j}^*) + RL\,e^{i2\chi}. \tag{6}$$

The central value theorem indicates that if the $D_R + D_L^*$ are random variables of antenna number with Gaussian statistics, but constant in time, the errors in the polarization image will be Gaussian. Applying a similar analysis as Perley (1989), the polarization dynamic range will be limited to $N/(\sqrt{2}D)$ for a snapshot and N/D for full tracks, where N is the number of antennas[1]. Because the D terms are *supposedly* constant in time, they do not integrate down like Gaussian receiver noise. Long tracks will be roughly $\sqrt{2}$ better in dynamic range because the change in parallactic angle will smear the errors out. For Gaussian D's of 0.02, the VLA should be able to observe fractional linear polarizations of 0.1% without any calibration of the D terms. Clearly, the D's are not Gaussian.

On the other hand, correlated D terms (ie, D terms which are the same for all antennas) will limit the polarization dynamic range to $1/D$, resulting in spurious polarization which copies the total intensity. A combination of correlated and uncorrelated D terms will limit the

---

[1]As total intensity images are often judged by *dynamic range*, or the peak of the total intensity image divided by the off-source rms, polarization images may be judged by the polarization dynamic range, or the peak of the total intensity image divided by the polarization image's rms. However, since some errors affecting polarization images are systematic, we may choose to evaluate the polarization image's rms error on source or off source.

polarization dynamic range to

$$\frac{I_{peak}}{P_{rms}} = \frac{1}{\sqrt{D_c^2 + (D_u/N)^2}}.$$ (7)

Data presented in Appendix A indicate that D_c is about 0.015 and D_u is about 0.02. The correlated portion of the D terms influences the final image quality more dramatically than the uncorrelated part because correlated D terms place all of the polarization error in the same place in the polarization map while uncorrelated D terms scatter the power throughout the image.

Similarly, we can estimate the effects of changing D terms on images. The polarization dynamic range will be limited to

$$\frac{I_{peak}}{P_{rms}} = \frac{1}{\sqrt{(\Delta D_c/2)^2 + (\Delta D_u/2N)^2}}.$$ (8)

It is clear that correlated changes in the D terms will affect the images more dramatically.

In principle, polarization images should be noise limited. While images of bright quasars in total intensity run up against a poorly understood dynamic range limitation of two hundred thousand to one, the polarization dynamic range is usually less than a thousand to one. This limitation in polarization dynamic range may be due to coherent changes of about 0.002 in the D terms during the observations.

Currently, the D terms can be calculated by the AIPS program PCAL by separating out the astronomical signal from the instrumental signal over the entire time range of the observations. This method assumes that the D terms are not varying in time. However, the D terms are known to vary with time. On page 6-15 of Synthesis Mapping, Bignell (1982) shows evidence for coherent variations of 0.003 over a few hours in two antennas' D terms. Breaking the solution up into smaller time ranges will often compromise the quality of the solution since full parallactic angle coverage is required for effective signal separation. The D terms may also vary with antenna azimuth and elevation, making the D term solution obtained from a distant calibrator somewhat invalid for the target source. Hence, if the time variations of the D terms limit the polarization images, PCAL cannot solve our problem.

Given a model for the source polarization, or model cross hand visibilities, and model parallel hand visibilities, it is possible to solve for the instrumental D terms with a new form of self-calibration. The uncertainty in the resulting D terms σ_D is given by

$$\sigma_D \simeq \frac{\sigma_V}{S(\sqrt{N-1}-1)}$$ (9)

where S is the total intensity flux, σ_V is the noise in the visibilities (assumed to be the same for total intensity and polarized visibilities), and N is the number of antennas. At 6 cm, 50 MHz bandwidth, and 27 antennas, observing a 1 Jy point source for 10 minutes will result in noise in the D terms at a level of 0.0005. Observations of a bright source like 3C273 will have D term noise of 0.00001 in the same amount of time, allowing us to track changes in the D terms on the order of one part in a thousand on time scales of minutes.

4

One problem with using a self-calibration scheme to determine the D terms is that a good source model must be known. Since the D terms give additive errors, a simple polarization structure with constant D terms could lead to similar visibilities as an unpolarized source with time variable D terms. Traditionally, the changing parallactic angle is used to differentiate between the astronomical polarized signal and the contribution of the D terms. It may be possible to make a first pass at the polarization imaging assuming the D terms to be constant. The resulting polarization image can then be used in the self-calibration procedure to determine the coarse time variations of the D terms, which can be applied to the polarization visibilities to yield an improved polarization image. A goodness of fit parameter for each antenna can be used as a diagnostic to indicate if the model is insufficient and if polarization structure is being absorbed into the D terms. Simple source structure, such as a polarized point source, could become confused with time dependent D terms, but complex source structure cannot be absorbed into the D terms.

We have developed a 'toy' linear D term self-calibration algorithm which is found to produce D term solutions which agree with PCAL's solutions when the solution interval is set to the entire observation (see Appendix B and Appendix E). In the linearlized model, the D terms corrupt the cross hand visibilities always in a pairwise fashion (ie, $\tilde{I}(D_{R_i} + D_{L_j}^*)$), and an arbitrary constant can be added to all the D_R and the complex conjugate of that arbitrary constant can be subtracted from all $D_{L_j}^*$ with no net effect on the polarization visibilities. This is not a problem if only polarization data is corrected, but a global offset could not be tolerated if the total intensity visibilities are to be corrected for the effects of the D terms. Turned around, we can use the constraint that the model visibilities corrupted by the instrumental D terms must be consistent with the measured total intensity visibilities to eliminate the global offset in the D terms. The best way to achieve this goal is to perform a global optimization using all parallel and cross-hand data and model visibilities to do a nonlinear least squares fit for the D.

Using the D term self-calibration program, we can see significant variations in the D terms on time scales as short as 10 s on 3C84. These unexplained variations are shown in Appendix B). Both uncorrelated and correlated variations in the D term phases and amplitudes are seen, on short and long time scales. The correlated variations are about 0.003 in at least two different observations at C and X band, which is large enough to explain the spurious time dependent linear polarization observed in the core of 3C84.

2.1 Future Work

More work needs to go into determining how to finesse time variable D terms and the correct model out of the cross hand visibilities. Alternatively, it should be determined whether the D terms vary with time, with position on the sky, or both. This question can be addressed by observing many bright calibrators with "known" polarization. If the D terms do not vary strongly with position on the sky, then the varying D terms can be solved for by observing a calibrator source of known polarization sufficiently close to the target source. Also, the full blown nonlinear optimization algorithm needs to be implemented. Solving for the *absolute* D terms rather than the D terms relative to a reference antenna will enable us to solve other

problems mentioned below.

3   RCP-LCP Phase Stability

In the instrumental polarization calibration schemes implemented in the AIPS task PCAL, data from the entire observing session are required to arrive at the D term solution. This means that any relative drift in the reference antenna's R and L phases (the AC phase drift in VLA terminology) will have a systematic effect on the D term solution, thereby limiting the quality of the polarization image.

Including the antenna gains G and the R-L phase difference for the reference antenna $\theta_{R-L}(t)$, the measured cross hand visibility will be

$$V_{R_i L_j} = e^{i\theta_{R-L}(t)} G_{R_i} G^*_{L_j} \left(RR \left(D_{R_i} + D^*_{L_j} \right) + RL\, e^{i2\chi} \right). \tag{10}$$

Bignell (1982) cites RCP-LCP phase drifts on the order of $10°$ over an 8 hour observation. The error in the solved D terms due to fluctuations in θ_{R-L} will go like

$$\Delta D \simeq \theta_{rms}(p + D_c), \tag{11}$$

where θ_{rms} is the rms of the R-L phase of the reference antenna about its mean value, p is the fractional polarization of the source and D_c is the correlated part of the instrumental D terms, and θ_{R-L} is measured in radians. This variation could be ~ 0.002, or about 10% of the D terms. The error in the D terms should be correlated, and would limit observations of fractional polarization to $\geq 0.1\%$

Currently, the D terms are solved for relative to D_R of the reference antenna. If the *absolute* D terms were solved for, as is advocated in Section 2, then (uncorrelated) AC phase drifts would result in random (ie, smaller) errors in the polarization image.

4   Total Intensity

The total intensity visibilities are corrupted by leakage from polarized signals as described by Equation 3. For an unpolarized source, only the first and last terms will contribute to the parallel hand visibility, leading to errors of order D^2 times the visibility. Fluctuating D terms will result in baseline based gains which can be self-calibrated out of the visibilities with existing software. If the source is polarized and the cross hand visibilities are p as strong as the parallel hand visibilities, then we can expect to make errors in the visibilities on the order of Dp. For the case of a point source, we can estimate the errors caused by leakage from the polarized signal. A rough estimate indicates that dynamic range of a point source will be limited to about N/pD, or about 14,000:1 for a 10% polarized source and uncorrelated D terms of 0.02. Full track VLA simulations assuming random D terms of a 10% polarized 1 Jy point source indicate artifacts in the total intensity image have an rms of about 0.0001 within 5 beams of the point source and an rms of 0.00003 far away from the point source. If the D terms are

correlated, the errors are worse. (Analysis of D terms indicates that the uncorrelated RMS of $D_{R_i}^* + D_{L_j}$ is about 0.02, and since we do not have the absolute D terms, it is not possible to determine the correlated part.) In sources with more complicated structure, the cross hand visibility will be lower due to beating between regions with different polarizations, leading to a decreased effect upon the image.

Results of high dynamic range imaging seem to contradict our assertion that the dynamic range of a polarized should be limited to a few tens of thousands to one. 3C273 has been imaged with dynamic range of 200,000:1, but the core is only weakly polarized. If a highly polarized source is observed over a small range of parallactic angle, much of the effects of the D terms on the parallel hand data could be removed by baseline based self-calibration.

The solution to the problem of total intensity visibilities corrupted by the cross hand visibilities is to take calibrated cross hand and parallel hand data visibilities or model visibilities, multiply them by the appropriate D terms which have been calculated by some method, and subtract the various terms visibility by visibility. It seems safe to say that if this procedure is employed, total intensity imaging will not be limited by residual errors in the polarization models or the D terms. This procedure should run very quickly on current computers. The AIPS task PCAL sets the D_R of the reference antenna equal to zero, and the other D terms are referenced to that antenna. If the effects of the D terms are to be removed from the parallel hand data, the D terms must be determined absolutely. The D term self-calibration scheme with explicit optimization advocated in Section 2 of this memo should produce absolute D terms, at least for high SNR data or highly polarized sources.

5 Beam Squint and Circular Polarization

The off-axis feeds on the VLA antennas result in the RR and LL beams being separated by 6% of the FWHM beam size (Clark, 1976; Thompson, 1976). This offset in RR and LL beams is known as the *beam squint*. The beam squint for off-axis feeds is explained by Chu and Turrin (1973). Since the beam squint occurs in the antenna (ie, AZ-EL) frame, the relative orientation of the RR and LL beams changes on the sky with parallactic angle. The antenna pointing is adjusted so that the the RCP and the LCP beams are symmetrically displaced from the nominal pointing position, so the measured RR and LL signals are related to the true source brightness distribution as

$$RR_{jk}(\mathbf{u}) = \int A((\mathbf{x} - \mathbf{x}_s)')(I(\mathbf{x}) + V(\mathbf{x}))e^{i2\pi\mathbf{u}\mathbf{x}}d\mathbf{x} \qquad (12)$$

$$LL_{jk}(\mathbf{u}) = \int A((\mathbf{x} + \mathbf{x}_s)')(I(\mathbf{x}) - V(\mathbf{x}))e^{i2\pi\mathbf{u}\mathbf{x}}d\mathbf{x}, \qquad (13)$$

where A is the primary beam, $\mathbf{x}_s$ is the vector by which the RCP beam is displaced from the nominal pointing center rotated by the parallactic angle, $(\mathbf{x} - \mathbf{x}_s)'$ is related to the sky coordinate $\mathbf{x}$ via a rotation about the average RR and LL pointing center by the parallactic angle, and $I(\mathbf{x})$ and $V(\mathbf{x})$ are the true brightness distributions for total intensity and circular polarized intensity.

The beam squint is a first order effect in the measurement of circular polarization. Primarily, the beam squint results in spurious fractional circular polarization which closely follows the total intensity emission. The fractional spurious polarization is equal to the difference between the RCP and LCP beams at the location of interest, averaged over time, and divided by the total intensity beam. The spurious fractional polarization can be as high as 0.01 near the primary beam center, 0.17 at the half power point, and 0.24 at the quarter power point for snapshots. Typical values for the spurious fractional circular polarization at the half power point for long integrations range between 0.02 and 0.10, depending upon the orientation of the feed relative to the emission of interest and the parallactic angle range. Long integrations lead to lower average values of the spurious fractional circular polarization, but this process introduces a secondary effect: because the effect of the beam squint is changing with parallactic angle, the spurious fractional circular polarization changes as a function of time, so even if the *average* spurious fractional circular polarization is near zero, power is scattered across the image like the RMS about the mean of the spurious fractional circular polarization times the RMS sidelobe level over the amount of time which the beam squint effects are approximately constant. Typically, the RMS about the mean of the spurious fractional circular polarization at the half power point is about 0.03, and the rms sidelobe level is about 0.10 for extended snapshots, so a point source of magnitude S at the half power point will result in scattered errors with an RMS of about $0.003S$ in circular polarization.

Fortunately, most astronomical radio sources (ie, synchrotron sources) do not display a high degree of circular polarization. However, circular polarization is observed in emission lines through Zeeman splitting, and in stars and masers. The beam squint is a serious problem when observing the sun, a group of masers scattered across the beam, or Zeeman split lines in a large HII region. For Zeeman observations, one can float the relative left and right gains by requiring that the left circular polarization integrated over frequency is equal to the right circular polarization. This solution is nonoptimal because the effects of the beam squint vary across the image, power is still scattered across the image due to the fluctuations in the beam squint with time at that particular location, and radiative transfer effects may actually make the integrated LCP and RCP signals different.

If the beam squint orientation is time variable (ie, time variable in the AZ-EL frame) or is different for each antenna, high fidelity imaging of circular polarization will require an algorithm similar to that required for imaging with known pointing errors (Holdaway, 1992b). Since this procedure requires a DFT rather than an FFT, such an algorithm would take several Sparc2 CPU *days* to produce a single full beam D array image. Since beam squint is a purely geometrical effect, the squint should be almost identical for all antennas. Appendix C bears out this assertion. The squint is so similar among antennas that the assumption of identical squint will not limit the fidelity of circular polarization imaging, and correction for the squint in software should be straightforward.

Vourlidas (1992) has developed a straightforward technique for removing the effects of the beam squint when imaging circular polarization. The data are split into several blocks of time over which the change in the parallactic angle is small, and the data from each of these blocks is imaged independently. The effects of the beam squint are removed for the mean parallactic

angle of each block, and the images are combined in a weighted average to form the final image. This is analogous to one possible solution to the non-coplanar baseline problem: the VLA is instantaneously planar, and a large dataset can be imaged piecemeal using two dimensional transforms and deconvolutions. This is also analogous to the polarization mosaicing scheme of Carilli and Holdaway (1992). Averaging of multiple images may lead to a final image which has the expected off-source noise, but because deconvolution is a nonlinear process, there will be a higher level of on-source errors in any approach which utilizes many separate deconvolutions. An example of how image quality can be degraded by cutting the data into many pieces and performing separate deconvolutions is given in Appendix D

But how can we deconvolve all the data simultaneously if the squint's RA-DEC orientation varies with time? Consider treating the squint problem as an inverse problem. Since we know the orientation of the squint in RA-DEC as a function of time, given models for the source total intensity and circular polarized brightness and distributions, we can calculate the squint-afflicted RR and LL visibilities. Comparison of model RR and LL visibilities with the data visibilities instructs us how to update the model brightness distributions so as to minimize the difference between RR and LL. The internals of such an algorithm would break the problem up into parallactic angle blocks and utilize FFTs on each. An analogous algorithm would be required for self-calibrating the RR and LL data in circular polarization imaging experiments.

This simultaneous deconvolution scheme requires the same number of FFT's as the multiple deconvolution scheme of Vourlidas (1992), it will require slightly more cpu time for explicit primary beam rotation and application, and it will require less cpu time to perform the single deconvolution rather than the multiple deconvolutions. The main advantages to this method are:

- the astronomer never has to worry about breaking the imaging up into time blocks, as this is all done internally.

- the data is deconvolved simultaneously, leading to a superior image.

- it can be extended to work with multiple pointing mosaics.

- when imaging time is dominated by the deconvolution step, this method will be much faster than separate deconvolutions.

It should be remembered that when observing in only RCP or LCP (as in spectral line observing or in some VLBI) the antennas will have a systematic, time dependent pointing error of 3% of the FWHM. This is about the same as the MMA pointing specification, which limits wide field imaging to a dynamic range of about 1000:1 and the fidelity to about 20. Wide field imaging with the VLA will be somewhat worse, since this 3% pointing error is correlated over the array, there are pointing errors on top of this systematic error, and there are fewer antennas. If necessary, imaging could be performed as suggested above. Alternatively, since we find that the VLA's spectral line images are rarely limited by the effects of pointing, we might argue that similar pointing errors will not usually limit the imaging for the MMA. In 4IF mode, both RCP and LCP are used, and the relative pointing errors will be twice as bad.

9

5.1 Future Work

In order to accurately account for the beam squint in imaging, the RR and the LL primary beams need to be well determined. For this algorithm, a two dimensional model for the primary beam will not be any more expensive to apply than a one dimensional model. The actual coding of the algorithm can piggyback on the maximum entropy algorithm of Cornwell and Evans (1985).

6 Polarization Sidelobes

Because of the way the electric field vectors reflect from the antenna's primary surface, the polarization leakage terms vary across the primary beam (Bignell, 1982). This polarization leakage pattern, called *polarization sidelobes*, is attached to the Azimuth-Elevation frame of the antenna, and rotates with parallactic angle on the sky. Each antenna has different polarization sidelobes patterns, but the general pattern is expected to be similar for all antennas. The standard (on-axis) D term calibration sets the polarization sidelobes at the center of the beam to zero. Assuming no circular polarization, the polarized visibility from a wide field as observed by AZ-EL antennas is given by

$$RL_{j,k}(\mathbf{u}) = \int A(\mathbf{x'}) \left(\left(D_{R_j}(\mathbf{x'}) + D_{L_k}^*(\mathbf{x'}) \right) I(\mathbf{x}) + (Q(\mathbf{x}) + iU(\mathbf{x})) e^{2ix} \right) e^{i2\pi \mathbf{u} \mathbf{x}} d\mathbf{x}, \qquad (14)$$

where A is the primary beam, $\mathbf{x}$ are the sky coordinates, $\mathbf{x'}$ is related to $\mathbf{x}$ via a rotation about the pointing center through the parallactic angle of the antenna, and the D functions are the polarization sidelobe patterns in AZ-EL coordinates. The linear polarization errors which result from the polarization sidelobes are similar to the circular polarization errors which result from the beam squint. A point source of S at $\mathbf{x_o}$ will cause a localized error which is equal to the time average value of $D_{R_j}(\mathbf{x'_o}) + D_{L_k}^*(\mathbf{x'_o})$ times S, and will also scatter power approximately equal to the RMS of the polarization sidelobes about the mean times the RMS of the point spread function. Both of these effects were mentioned in Carilli and Holdaway (1992).

Some relevant questions remain to be answered:

- What is the time stability of the polarization sidelobes?

- Do the polarization sidelobes have a strong dependence on elevation or azimuth?

- How similar are the different antennas' polarization sidelobe patterns? More specifically, what is the mean pattern and the dispersion about the mean of the ensemble of patterns $D_{R_i}(\mathbf{x'}) + D_{L_j}^*(\mathbf{x'})$?

If the polarization sidelobes were stable with respect to time and pointing direction, if the patterns were highly correlated (ie, similar) among antennas, and if the VLA had equatorial mount antennas, the solution to wide field polarization imaging would be as easy as multiplying the total intensity image by the polarization sidelobe pattern, Fourier transforming, and

subtracting from the measured RL visibilities as directed by Equation 14. Since the VLA's antennas have AZ-EL mounts, the easiest problem we can hope for is stable and correlated polarization patterns which will enable us to rotate the pattern to the correct orientation for a particular time, multiply the rotated polarization sidelobe pattern and the total intensity image, Fourier transform, and subtract from the RL visibilities measured during the small time interval during which the polarization sidelobe pattern is appropriate. The worst problem we could have is unstable and uncorrelated polarization sidelobe patterns, which would be possible to solve for, but unpleasant.

The data in Appendix E indicate that between a quarter and half of the power in the polarization sidelobes is coherent between antennas. At first glance, it seems that this is too low a coherent fraction for useful imaging to result from assuming that all antennas have identical patterns. Recall, however, that the most serious errors result from correlated polarization errors which produce localized image plane errors like $I(x)D_{corr}$, while uncorrelated polarization errors produce scattered image plane errors like $I(x)D_{uncorr}/N$. Hence, if we can remove the correlated parts of the polarization sidelobes, we can achieve limiting fractional polarization of about $0.03/N \simeq 0.001$ (due to the uncorrelated polarization sidelobes of about 0.03). To achieve lower errors, we would have to consider each antenna individually.

6.1 Beam Squint and Linear Polarization

The beam squint will also have an impact on wide field linear polarization imaging. The corrupted (measured) polarized correlations are given by Equation 4. While it is reasonable to assume in many cases that the *astronomical* circular polarization is zero, it cannot be assumed that $RR = LL$ in wide field imaging because of the beam squint. As mentioned above, it should be possible to determine a significant component of the D terms as a function of position in the beam. In determining the D terms across the beam, and in correcting the cross hand visibilities for the D terms, the RR and LL visibilities must be correctly calculated from a total intensity source model plus a model of the beam squint. This requires algorithms similar to the beam squint algorithm described in Section 5.

6.2 Future Work

A bright point source of known polarization such as 3C84 should be used to sample the polarization properties of the beam. The shapes of the RR and LL beams need to be determined (as opposed to the shape of the I beam). The polarization sidelobes should be determined for each antenna, as in Bignell (1982), but over a range in elevation and on different days to check for variability. After the polarization sidelobe problem is solved, polarization mosaics will become much easier to perform. If the MMA will do any linear polarization observations, this problem must be well understood. If possible, the MMA antennas must be designed to result in uniform polarization sidelobes.

11

7 Ionospheric Faraday Rotation

If time dependent ionospheric Faraday rotation occurs during an observation at low frequency, an image formed from this data will display uncleaned sidelobes due to the apparent variable nature of the source. If the Faraday rotation is strong enough, the true polarized emission might be completely canceled out. A typical strategy for dealing with time dependent Faraday rotation is to estimate the rotation in polarization position angle as a function of time through ionospheric measurements, measurements on a nearby calibrator with some polarized signal as a function of time, or by looking at a polarized feature in the source of interest as a function of time, and shifting the phase of the L (or R) antenna gains by the time dependent position angle (Fomalont, 1992). All of these methods for determining the ionospheric Faraday rotation as a function of time leave something to be desired: either we get the Faraday rotation at some position on the sky other than the source position (Faraday rotation is sky position dependent as well), or we must break the imaging up into time blocks again.

For a sufficiently bright, polarized source, it should be possible to solve for the relative Faraday rotation as a function of time as another free parameter in a self-calibration program. The polarization rotation can be interpolated and applied in the usual manner. Hence, the time variable aspect of the problem is removed, and excellent images can be formed. However, these excellent images will have no information about the true position angle. In order to obtain information about the true position angle, the atmospheric Faraday rotation through multiple lines of sight and models for the earth's magnetic field and the ionosphere must be used.

For a source which has weak polarization and cannot be detected in the time which the electric field vector changes by a radian, we may be able to use all of the data to fit to a simple parameterized model of the electric vector rotation. At P band, Fomalont (1992) finds that the electric vector rotation is almost linear in time.

In VLBI polarization experiments, the ionospheric Faraday rotation will be different above the different antennas. Again, the Faraday rotation can be solved for as a time variable, antenna dependent parameter via a new self-calibration algorithm.

8 Polarization Addendum to the Calibration Manual

The existence of an up to date catalog of many calibrator sources' polarized flux and position angle would greatly facilitate the determination of time variable D terms. We would want to choose calibrators which were hundreds of milliarcseconds in size to reduce the time variability of the polarized emission. These sources do not need to be highly polarized, they just need to have precisely known polarization properties. They could be observed every few months, and the current polarization parameters could be kept in a computer database.

VLBA polarimetry has an orthogonal requirement: in order to calibrate the absolute polarization position angle, one needs to observe a polarized calibrator in which all of the polarized emission is coming from a region a few milliarcseconds across so that the polarization position angle measured by the VLA is the same as the polarization position angle measured by the

VLBA. These compact calibrator sources, mostly BL Lacs, are highly variable, and would need to be observed quite frequently, either weekly or at the same time as the VLBA polarization observations. The Brandeis group has always used the VLA as part of its VLBI array, arguing that the VLA is required for polarization position angle calibration. Actually, the VLBA could perform polarization position angle calibration without the VLA if an accurate and up to date database of the polarization properties of a few compact calibrator sources were kept.

Obtaining accurate position angle calibration for the VLA or the VLBA at low frequencies (600 MHz and 330 MHz) will be difficult. Currently, there are no sources which have accurately known polarization position angles at low frequencies. It would be nice to know the polarization angle above the atmosphere for about 10 calibrators at the 1.4 GHz, 600 MHz, and 330 MHz bands. Currently, we know the polarization position angles of only two calibrators at high frequencies. We might be able to extrapolate the position angle above the atmosphere by using the calibrator's local rotation measure and the galactic rotation measure in that direction.

Appendices

A   Correlated and Uncorrelated Fractions of the On-Axis D Terms

To first order, the D terms at the beam center corrupt the RL visibilities as described by Equation 6, or as the parallel hand visibility times $D_{R_i} + D^*_{L_j}$. The quantity $D_{R_i} + D^*_{L_j}$ does not depend upon whether the D terms have been referenced to a particular antenna or not. In the text, we indicate that the correlated and uncorrelated parts of the D terms have very different effects upon the polarization image. We present in Figure 1 a scatter plot of the real verses imaginary parts of the quantity $D_{R_i} + D^*_{L_j}$ for two recent VLA observations. The first observation was made of 3C84 on 22 Feb 92 in X band at 50 MHz, around an LST of 4 hours, and the second observation was of the VLA calibrator 0023-263 on 15 Sep 92 in C band at 50 MHz, around an LST of 0 hours. The D terms from 0023-263 have a mean value of 0.015 and an RMS scatter about the mean of 0.023, while the D terms from 3C84 have a mean of 0.014 and an RMS scatter about the mean of 0.023. The effects of the D terms on polarization images will be dominated by the correlated part, or mean value, of $D_{R_i} + D^*_{L_j}$.

The form of the $D_R + D^*_L$ scatter plots at X and C bands are very similar, but are rotated by 57° with respect to each other. Similar scatter plots are observed for IF 1, but they are separated by a different angle. Also, the outlying points near the bottom of the C band plot are due to large D terms for antennas 15(R) and 22(L), while the outlying points near the left side of the X band plot are due to antennas 7(R) and 24(L). We have looked at C band data from two different observations separated by more than a year. The general appearance of the two scatter plots is very similar, and antennas 15 and 22 are outlyers for both observations.

B   Time Variability of On-Axis D Terms

B.1   D Term Variability Exists

The core of 3C84 is formally less than 0.03% polarized to the VLA at C band. When we perform the standard polarization calibration with a single solution interval, the brightest polarized feature is near the total intensity core, but the polarized intensity and polarization position angle varies with time (see Table 1), indicating that something is varying, and probably not the source. When a solution is performed independently upon each of the seven snapshots (about 150 s), the polarized intensity near the core decreases by a factor of 10 in each of the snapshot images. The peak in the polarized intensity is shifted away from the core and is a factor of 5 lower than the peak in the polarization image made with a single solution interval.

Paralactic Angle	P Intensity, Jy	P Position Angle
-71°	0.012	45°
-79°	0.015	-10°
-90°	0.017	-30°
-103°	0.017	20°
-134°	0.028	-20°
143°	0.021	-45°
102°	0.016	100°

Table 1: The instrumental D terms were determined for 3C84 using a single solution interval and were then applied to the 3C84 data, which was subsequently imaged in seven separate four minute snapshots separated in time by about 1.5 hours. When the seven snapshots are imaged together, the peak polarized intensity is 0.010 Jy. The fact that the core polarization is not constant indicates that the D terms really are varying.

B.2 Complex Nature of the D Term Variability

The 'toy' D self-calibration algorithm solves for the D terms relative to D_R of the reference antenna, so apparent variation in a D term includes the variation of the reference antenna. However, $D_{R_i} + D^*_{L_j}$ will not display the reference antenna's D_R variations. Both types of quantities will be studied.

The D terms behave in a bizarre fashion, and more than one cause must be invoked to explain them. First we consider the long time scale (hours) D term fluctuations:

- some D terms vary by less than 0.0005 in amplitude and a few degrees in phase (see Figure 2).

- some show amplitude drifts of 100% (0.015) and phase drifts of as much as 100° (see Figure 3).

- in many cases, the amplitude and phase changes are roughly consistent with two different complex components of the D terms beating against each other.

- in a few cases, the D terms display large amplitude fluctuations but small phase fluctuations (see Figure 4).

- in many cases, the phases of D_R and D_L vary in opposite senses (see Figure 5).

- long time scale correlated trends are seen in the quantity $D_{R_i} + D^*_{L_j}$ averaged over all baselines. IF 1 and IF 2 display similar behavior, with the amplitudes varying by about .002 and the phases varying by 5-10° (see Figure 6.

To further illustrate the nature of the long time scale fluctuations, scatter plots of the real and imaginary parts of $D_{R_i} + D_{L_j}$ are made at approximately 1 hour intervals and shown in Figure 7

It has been suggested that the long time scale uncorrelated fluctuations in the D terms are caused by time constant, frequency dependent D terms modulated by a slowly varying bandpass. Assuming the phase of the D terms for individual spectral line channels varies by 45° across a 12.5 MHz band (Ge, private communication) and the bandpass has a 5% 3 MHz time dependent ripple (Carilli, 1991), the fluctuations of the D terms due to a fluctuating bandpass will be only 1% of the D term's value at 12.5 MHz bandwidth, 5% at 3 MHz, and 15% at 0.78 MHz. Hence, the bandpass fluctuations cannot account for the types of long term fluctuations which are seen in our data.

The long time scale fluctuations which are correlated among the antennas is similar for both IFs, and can be explained by feed leg droop changing the subreflector reflection D term component. Also, the RCP-LCP phase difference can explain long term fluctuations or jumps seen in the D terms.

There is good hope for removing these long time scale D term fluctuations from the data. We need to determine how the D terms vary on the sky in order to do this routinely.

The D terms are variable on very short time scales as well. Observing 3C 84 with 10 s time resolution in the D terms, individual D terms are accurate to 0.00014, or about half a degree, and correlated D fluctuations determined using all of the data will be substantially more accurate:

- the D term phase is found to drift and oscillate by $\sim 10°$ on minute time scales (see Figure 8). The phase behavior is correlated across the array, implying that the fluctuations are due to a fluctuating RCP-LCP phase at the reference antenna or the fluctuations are truly correlated across the array.

- occasionally, the correlated part of the D terms' phase remains constant for short periods of time.

- the amplitudes of the individual D terms are consistent with noise, but the amplitude of $D_{R_i} + D_{L_j}^*$ averaged over all baselines shows disorganized fluctuations well above the noise (see Figure 9). If the 10° phase variations were caused by two D term components beating, the amplitude fluctuations would be an order of magnitude higher.

- the D term fluctuations are not the same for different IFs.

- the D term fluctuations are erratic: the fluctuations may be a periodic for several hours, and then change to a drift.

Long time scale changes dominated the correlated changes in the amplitude of $D_{R_i} + D_{L_j}$, but short time scale changes dominate the phase. Because the short term fluctuations are highly correlated across the array, we may be able to fix whatever causes them. If these fast fluctuations cannot be fixed, they will fundamentally limit polarization imaging.

16

Band	Squint/Beam Width	Orientation	Squint RMS/(Squint/Beam Width)
L	0.0577	133°.1	0.0123
C	0.0635	−29°.0	0.0152
X	0.0613	−81°.6	0.0053
U	0.0640	29°.9	0.0207
K	0.0605	−8°.1	0.0404

Table 2: Degree of stability of the beam squint among the VLA antennas at various frequencies. Some of the apparent variability in squint is due to errors in the measurement of the squint.

C   Uniformity of Beam Squint Among Antennas

Since beam squint is a purely geometrical effect, the squint should be almost identical for all VLA antennas (and for all VLBA antennas). VLA pointing data has been analyzed, and the figures in Table 2 bear out this assertion. The orientation of the observed squint is slightly different from the orientation of the feeds, and only K band shows a large discrepancy. Some of the scatter in the squint is due to measurement error. There were more squint measurements at X band than at the other bands (by a factor of 10 to 20). The 0.5% variation in squint among antennas at X band is probably close to its true variation. This does not preclude the variation from being higher at other bands. At both C and U band, the dispersion in the azimuthal component of squint values is 0.1%, but the dispersion in the elevation component of the squint is between 3% and 4%. At K band, the antennas' surface errors probably cause strange beams which might increase the scatter in the squint values. The squints of each antenna generally have a Gaussian distribution about the mean squint: after removing the mean squint, the residual squint error manifests itself as random pointing errors in RR and LL. These pointing errors range from 0.0003 to 0.0024 of the FWHM beam, and are hence much smaller than the true pointing errors at all bands. The squint is so similar among antennas that the assumption of identical squint will not limit the fidelity of circular polarization imaging, and correction for the squint in software should be straightforward.

D   Image Quality and Multiple Deconvolutions

We have performed a set of numerical simulations to illustrate the degradation of image quality which can occur when a single large dataset is split up into many small datasets which are deconvolved independently and averaged to give a final image. A single eight hour dataset was imaged with UVMAP followed by a maximum entropy deconvolution. The success of the deconvolution was gauged by the dynamic range and the fidelity index (Holdaway, 1990). The original set was then split in half and each half set was UVMAP'ed and deconvolved, the resulting images averaged to give the final image. This splitting was continued until sixteen independent images were made from 30 minute time blocks of data. The dynamic range and

N blocks	Dynamic Range	Fidelity Index
1	15000:1	165
2	10000:1	115
4	9000:1	70
8	7800:1	50
16	7700:1	45

Table 3: Relationship between image quality and number of blocks the data is split up into for independent deconvolutions, from simulated VLA data.

fidelity index for each image is recorded in Table 3. Since no errors were added to the data, any defects in the image are due to incomplete (u, v) coverage. There is some gain in dynamic range which is obtained from imaging all of the data at once. The fidelity index, which reflects the accuracy of the reconstruction on-source, is roughly proportional to the square root of the number of visibilities used in each separate deconvolution. Simpler sources will not show such a drastic improvement in the fidelity index as the number of visibilities used in each deconvolution increases. For complicated sources like the model used in these simulations, deconvolution of all of the data at once is preferred.

E Off-Axis D Terms: Correlation of the Polarization Sidelobes

In the same data which we used to demonstrate a polarization mosaicing scheme (Carilli and Holdaway, 1992), we find that the quantity $D_{R_i} + D_{L_j}^*$ is somewhat correlated among baselines at the half power point. The data consist of three pointings on NGC 253 at L band with one pointing centered on the bright nucleus and the other two pointings with the nucleus at the half power point. The central pointing indicates that the ~ 2 Jy central point source is less than 0.1% polarized. Since there was some large-scale polarized structure, only baselines longer than 1500 λ were used in the solution. After the cross hand visibilities were corrected for the leakage from the total intensity, essentially noise limited images were obtained. (Note: this method enables imaging with this data set only because the source is dominated by a single bright region, and is not a general way of producing wide field polarization images. We are primarily using this method to illustrate the statistical behavior of the off-axis polarization leakage terms.) Table 4 presents the magnitude of the average and the rms about that average of the complex quantity $D_{R_i} + D_{L_j}^*$. Between a quarter and half of the power in this quantity is coherent. It should be noted that the L band polarization sidelobes are less correlated than the C and U band polarization sidelobes (Bignell, private communication).

Azimuth	Average $D_{R_i} + D_{L_j}^*$	RMS about the Average
-56	0.026	0.042
-40	0.023	0.037
-27	0.019	0.028
-16	0.014	0.019
-13	0.011	0.013
114	0.023	0.019
131	0.015	0.012
144	0.013	0.021
157	0.022	0.036
160	0.031	0.049

Table 4: Statistics of an azimuthal slice of the polarization sidelobes: the average value and the rms about the average of $D_{R_i} + D_{L_j}^*$ evaluated at the half power point and various azimuth angles. Azimuth is defined to increase in the same direction as parallactic angle. Because a large part of the polarization sidelobes are systematic, it is possible to remove their worst effects.

Acknowledgments

19

F REFERENCES

1. Bignell, C., 1982, "Polarimetry", in *Synthesis Mapping: Proceedings of the NRAO-VLA Workshop[held in Socorro, New Mexico, ed* A.R. Thompson and L.R. D'Addario.

2. Carilli, C.L., and Holdaway, M.A., 1992, "A simple minded approach to polarization mosaics", VLA Test Memo 164.

3. Carilli, C.L. 1991, "Spectral Dynamic Range at the VLA: The 3 MHz Ripple", VLA Test Memo 158.

4. Chu and Turrin, 1973, IEEE Transactions on Antennas and Propagation, **AP-21**, 3, page 339.

5. Clark, B.G. 1976, "A Rough Measurement of the Polarization Properties of the VLA Antenna" VLA Test Memo 111.

6. Cornwell, T.J., and Evans, K.F. 1985, "A simple maximum entropy deconvolution algorithm", *A+A*, **143**, 77-83.

7. Cotton, W.D., 1992 "Polarization Calibration of VLBI DATA", AIPS Memo 79.

8. Fomalont, E., 1992 *Private Communication.*

9. Holdaway, M.A., 1990, "Imaging Characteristics of a Homogeneous Millimeter Array", MMA Memo 61.

10. Holdaway, M.A., 1992a, "Surface Accuracy Requirements for Mosaicing at Millimeter Wavelengths", MMA Memo 74.

11. Holdaway, M.A., 1992b, "Imaging With Known Pointing Errors", MMA Memo, in preparation.

12. Perley, R.A, 1989, *in 'Synthesis Imaging', Publ, A.S.P., eds. Perley, R.A., Schwab, F.R. and Bridle, A.H.*, 290.

13. Thompson, A.R., 1976, "The Effect of the Beam Offsets on Polarization Measurements" VLA Scientific Memo 125

14. Vourlidas, A., 1992 *Private communication.*

Figure 1: Scatter plot of the real and imaginary parts of $D_{R_i} + D^*_{L_j}$ for recent observations of 0023-263 (C band, 50 MHz) and of 3C84 (X Band, 50 MHz). The agreement between the difference in position angles of the mean $D_R + D^0_L$ and the difference between the C and X band feed position angles is coincidental and is not reproduced in the other IF.

Figure 2: Example D terms as a function of time obtained from X Band 3C84 observations and the 'toy' D term self-calibration algorithm. The amplitude and phase of D_R for antenna 21 vary little with respect to the reference antenna 3.

Figure 3: Example D terms as a function of time obtained from the C Band 3C84 observations. Amplitude and phase of D_L for antenna 15 vary considerably with respect to the reference antenna 3.

Figure 4: Example D terms as a function of time obtained from the C Band 3C84 observations. Amplitude of D_R for antenna 15 doubles while the phase varies by only 6°.

Figure 5: Fluctuations in the phase of D_R and D_L for antenna 24 (C Band, 3C84). Note the nearly antisymmetric behavior of the phases.

Figure 6: Plots of the amplitude and phase of the average of $D_{R_i} + D^*_{L_j}$ from 3C84. The amplitudes behave similarly for both IFs. The scatter is actually real structure rather than noise. The amplitude is dominated by long time scale variations, while the phase is dominated by short time scale variations.

Figure 7: Scatter plots of the real and imaginary parts of $D_{R_i} + D^*_{L_j}$ made at ∼1 hour intervals from 3C84.

Figure 8: Fluctuations in the phase of $D_{R_i} + D_{L_j}$ averaged over all baselines with 10 s resolution for a number of snapshots. Similar patterns are seen in the phases of the individual D terms, but averaging over all baselines removes the reference ambiguity, increases the SNR, and emphasizes the part of the D term variations which is coherent across the array.

Figure 9: Fluctuations in the amplitude of $D_{R_i} + D_{L_j}$ averaged over all baselines with 10 s resolution for a number of snapshots. Fluctuations shown here are well above the noise, but do not show a clear pattern as the phase variations do.

DR + DL*

X Band
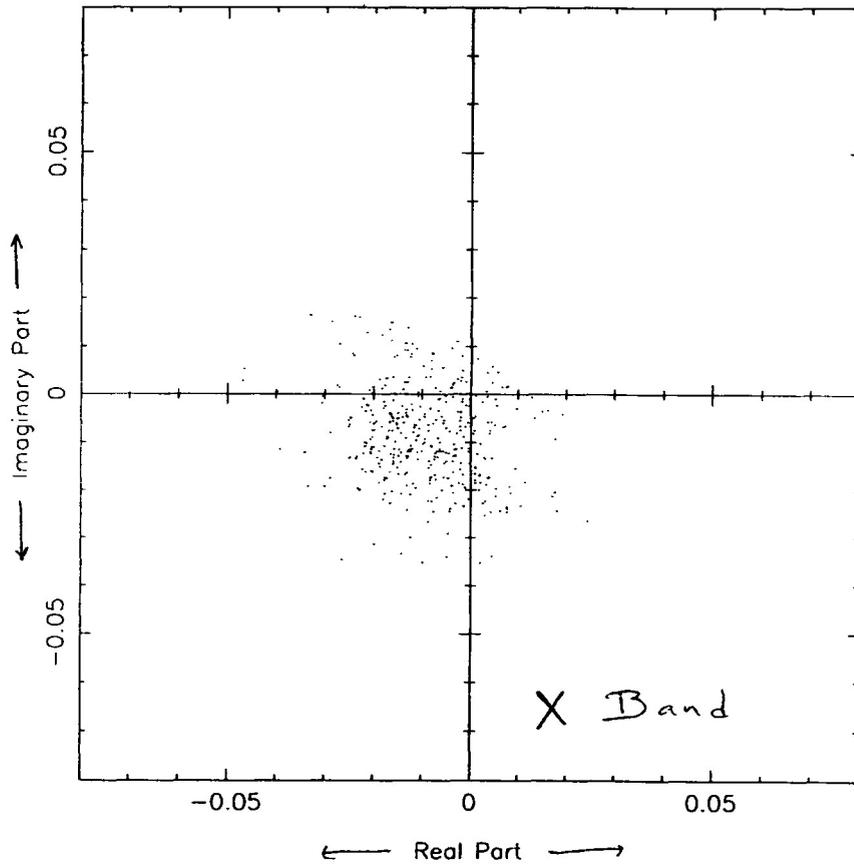
mholdawa 27-Oct-1992 15:27

C Band
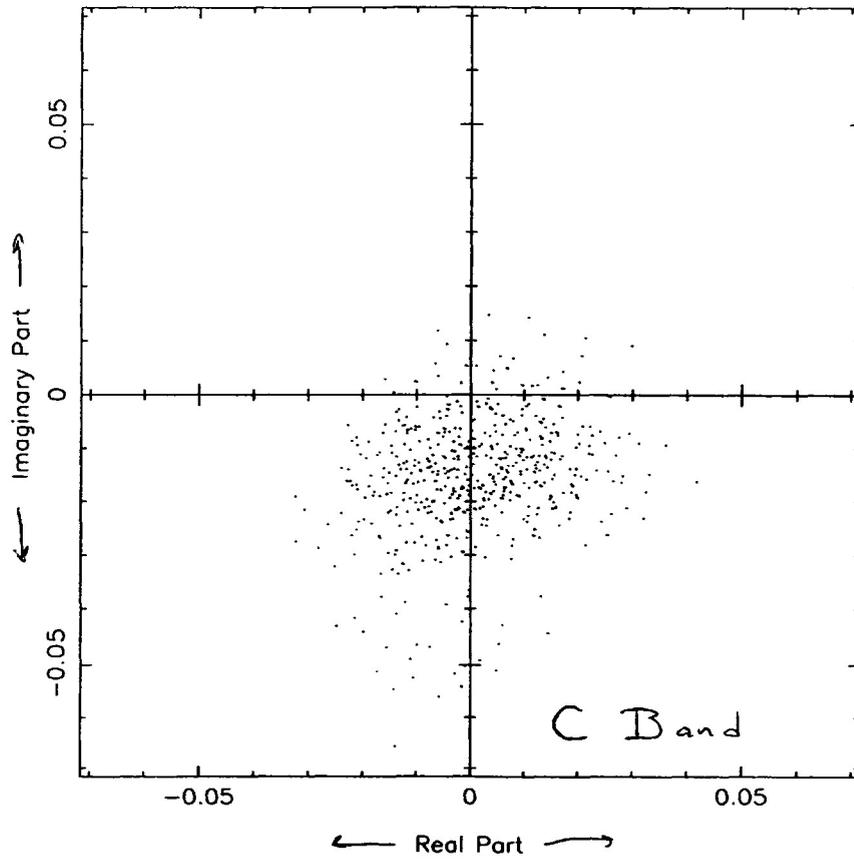
mholdawa 11-Nov-1992 16:46

Figure 1

DR for antenna 21

Figure 2

DL Amplitude

0.012

0.01

8×10^{-3}

6×10^{-3}

4×10^{-3}

0.3   0.35   0.4   0.45   0.5   0.55   0.6   0.65

Time, days
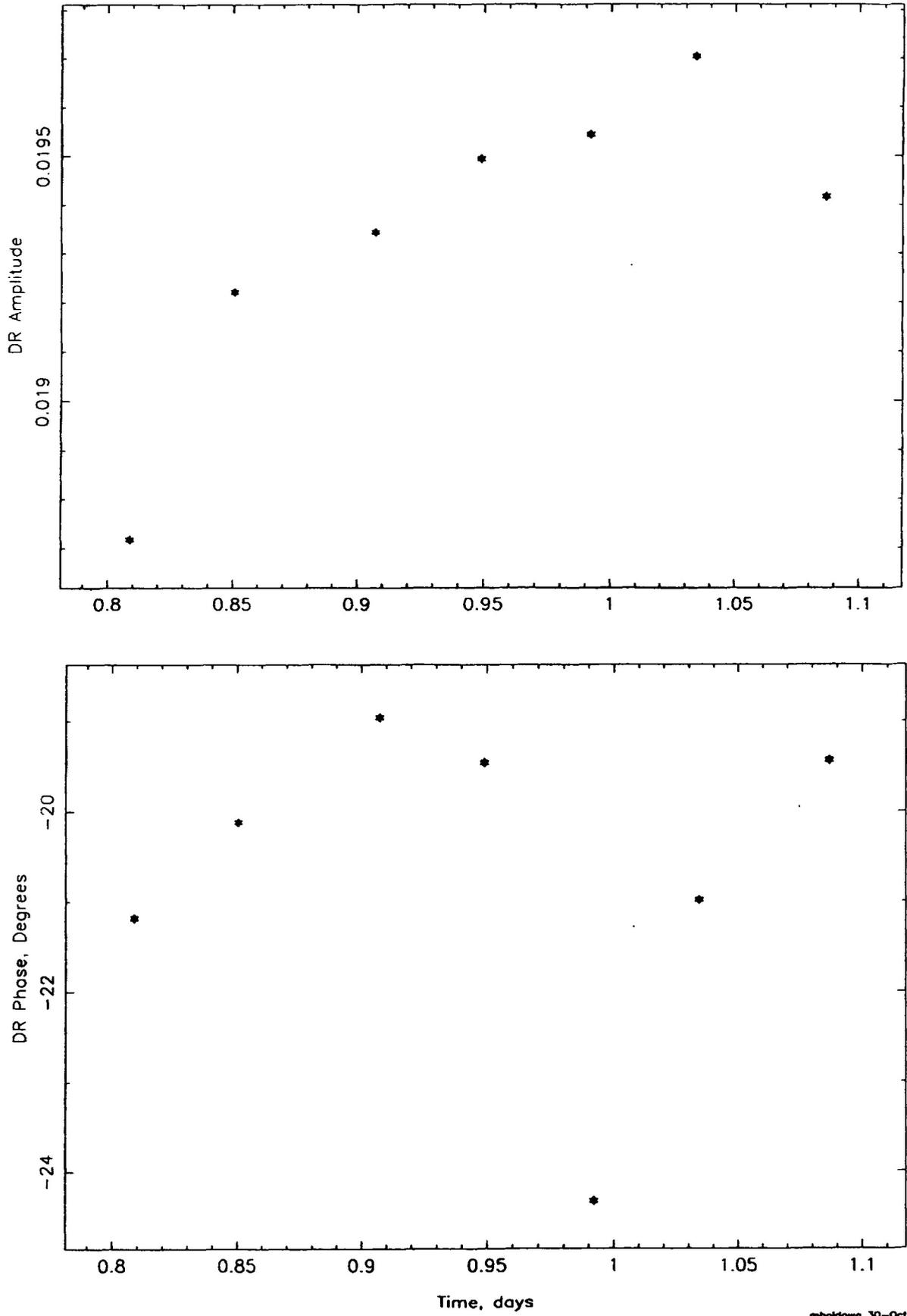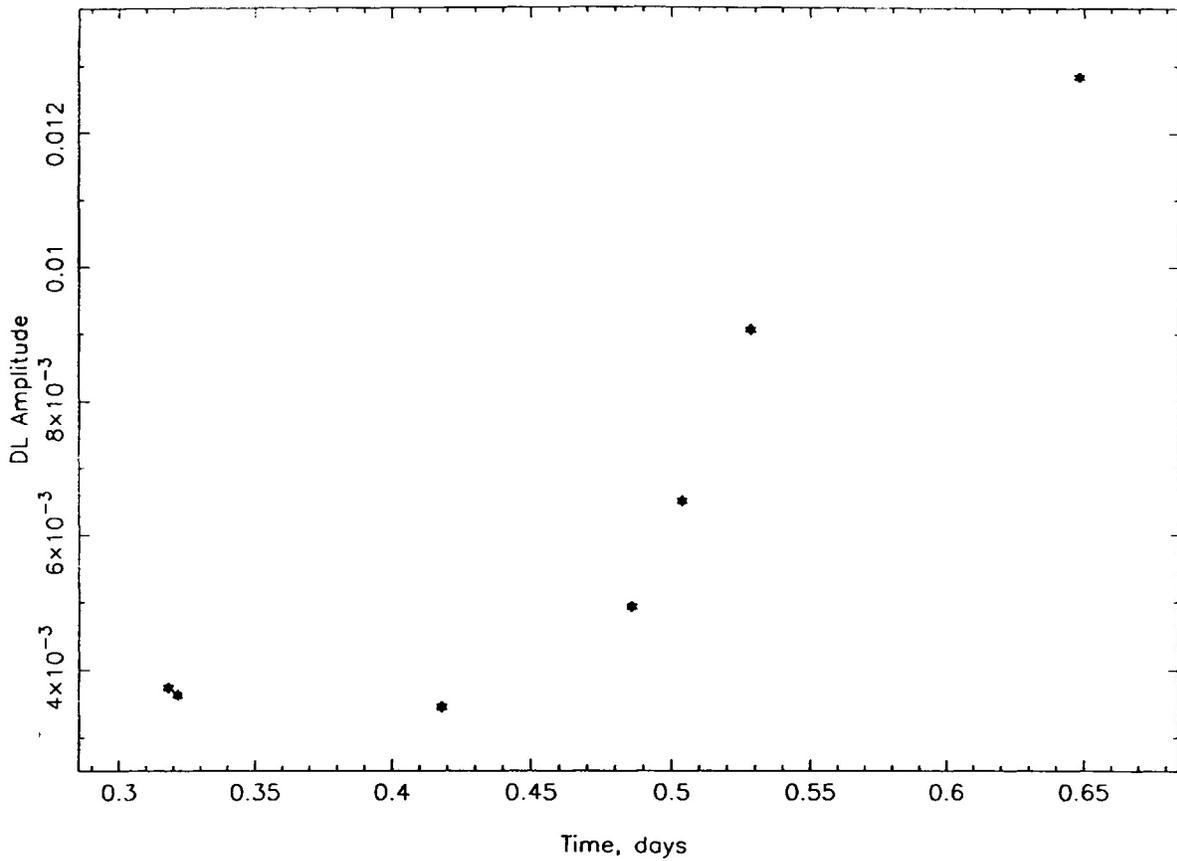
mholdawa 1-Dec-1992 07

DL for antenna 15

DL Phase, Degrees

100

50

0

0.3   0.35   0.4   0.45   0.5   0.55   0.6   0.65

Time, days
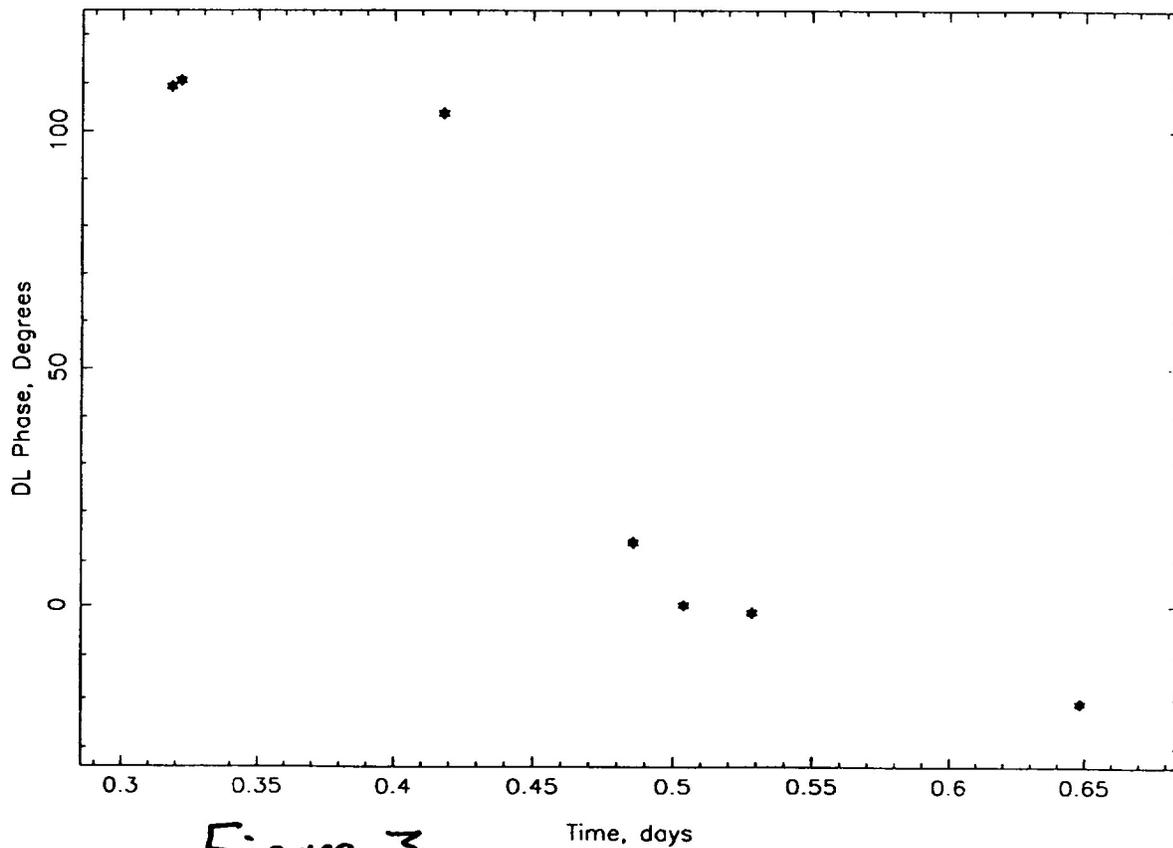
Figure 3

mholdawa 1-Dec-1992 07

DR for antenna 15
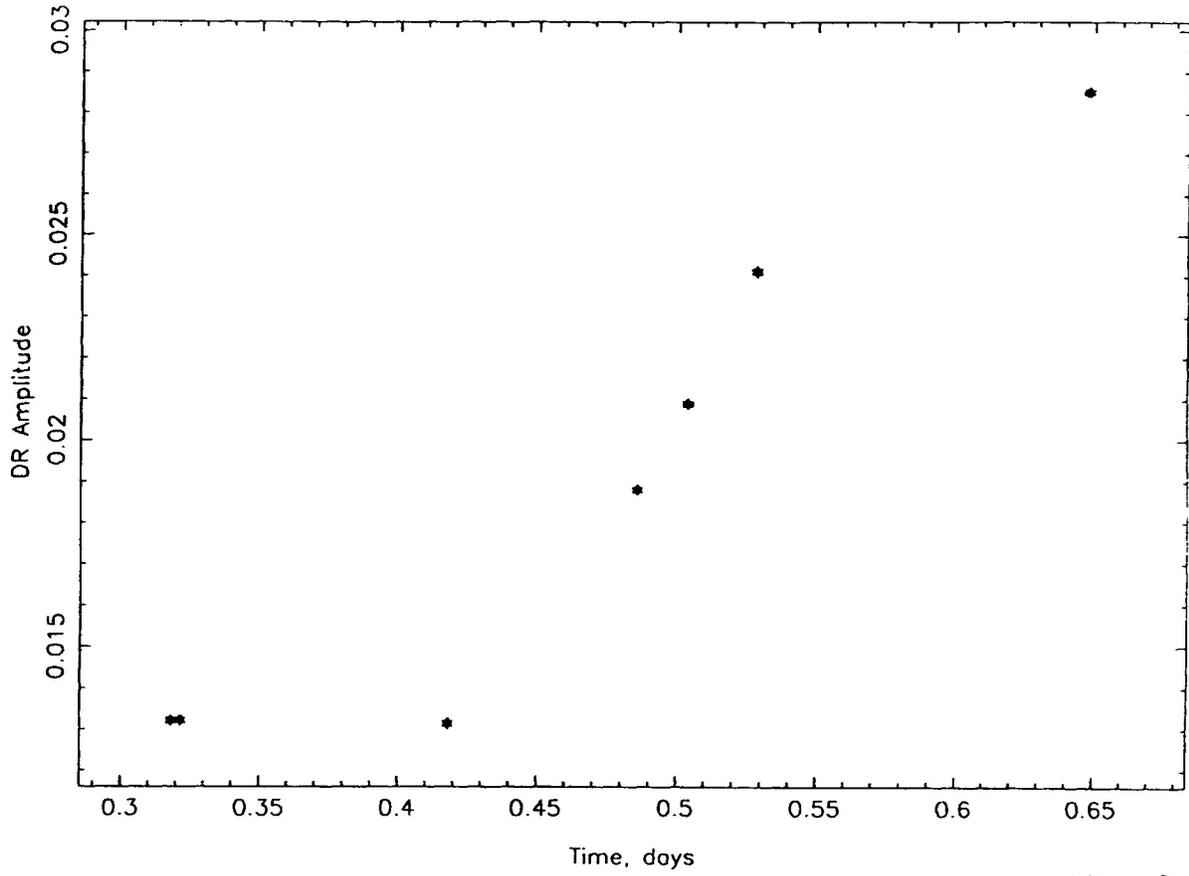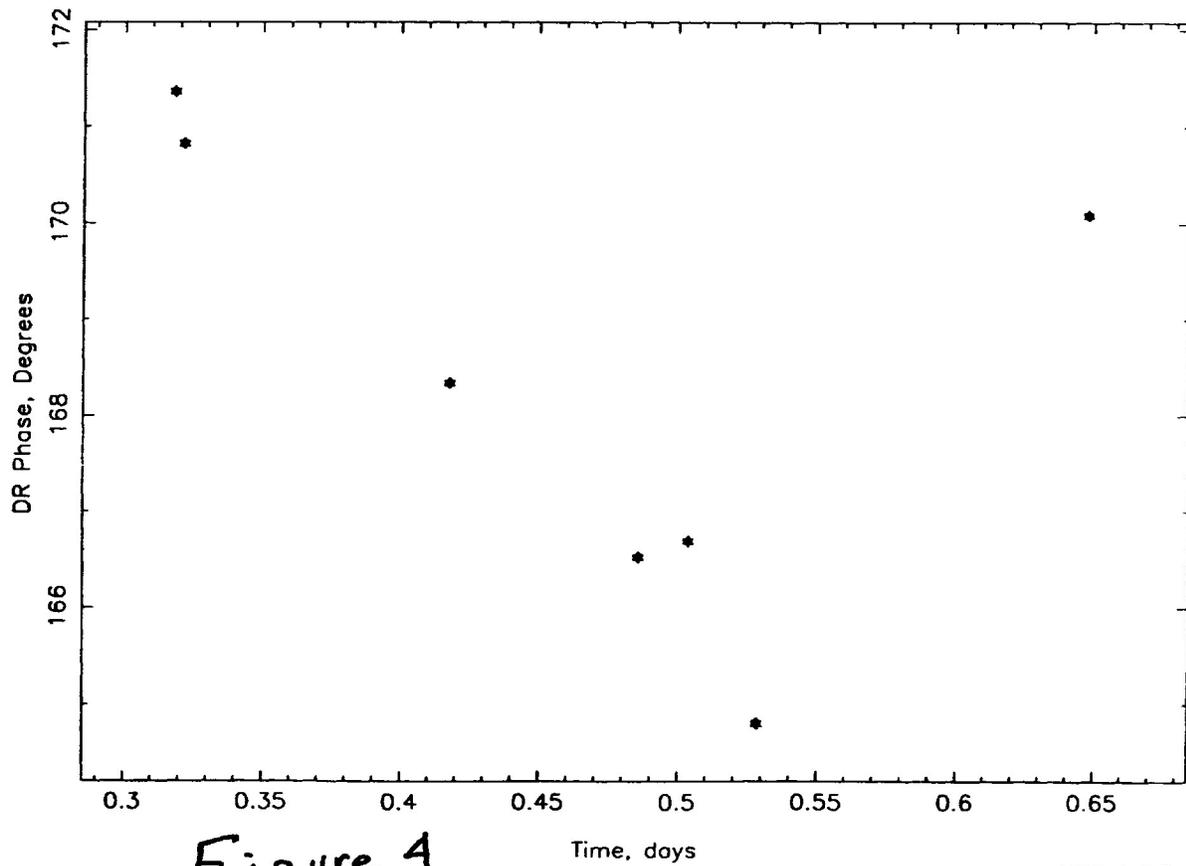
DR for antenna 15

Figure 4

DR for antenna 24
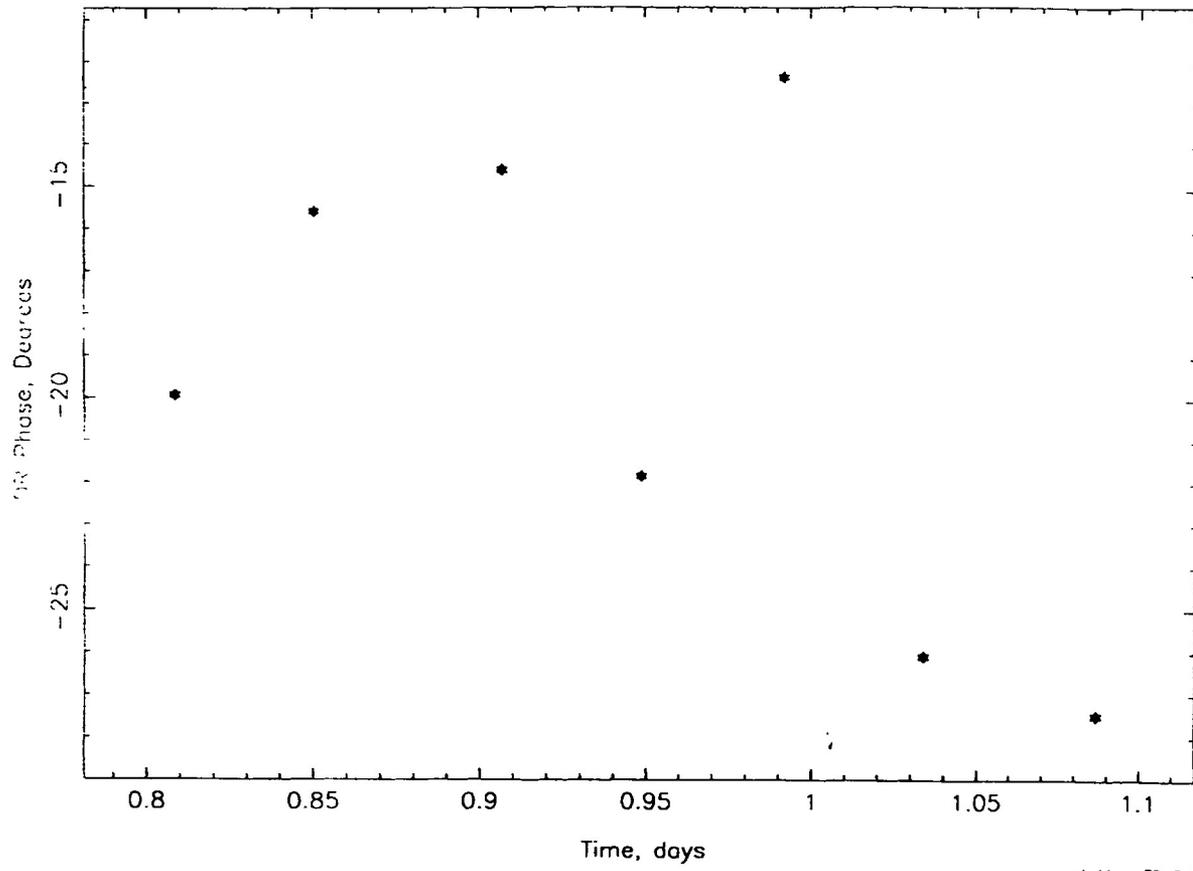
Time, days

DL for antenna 24

Time, days
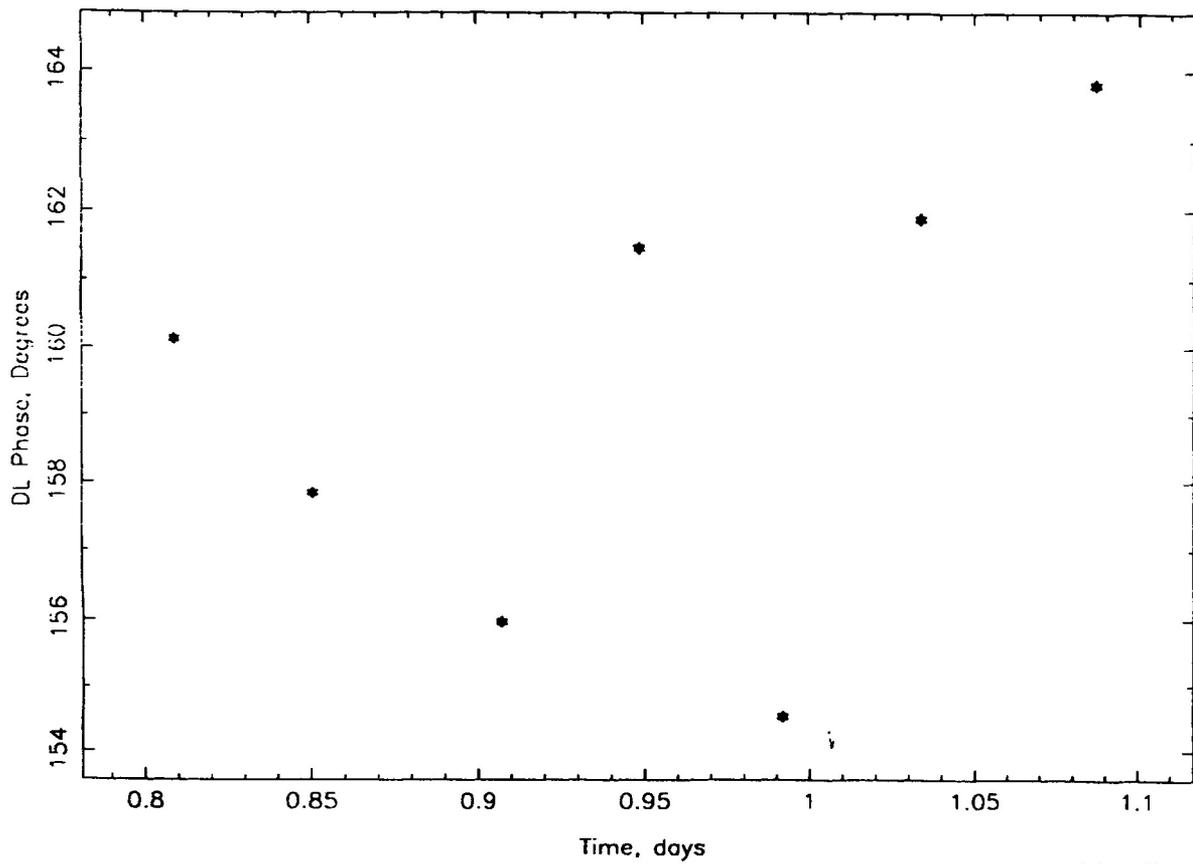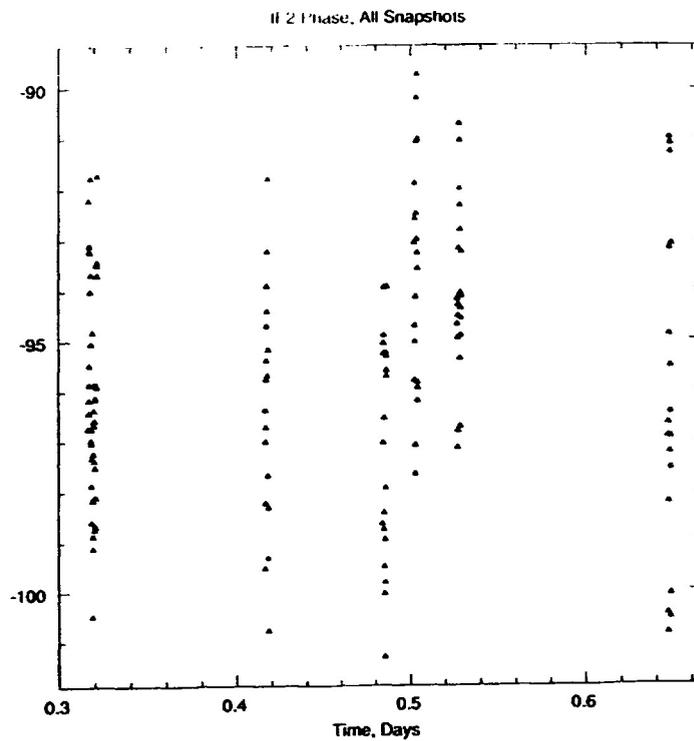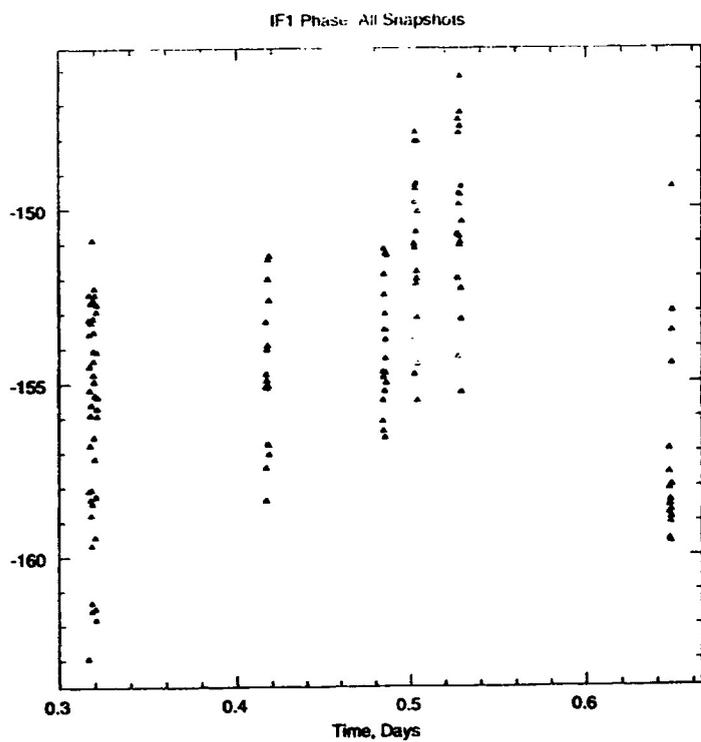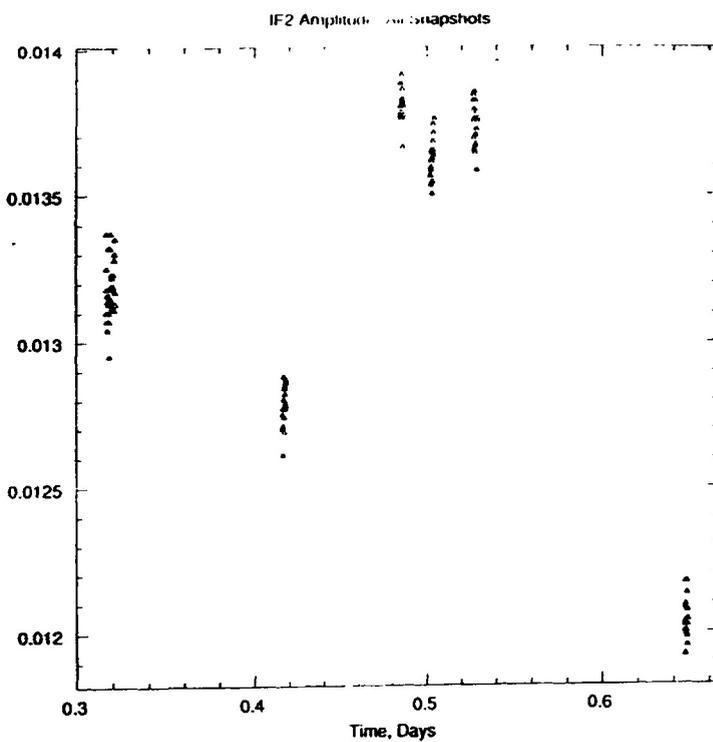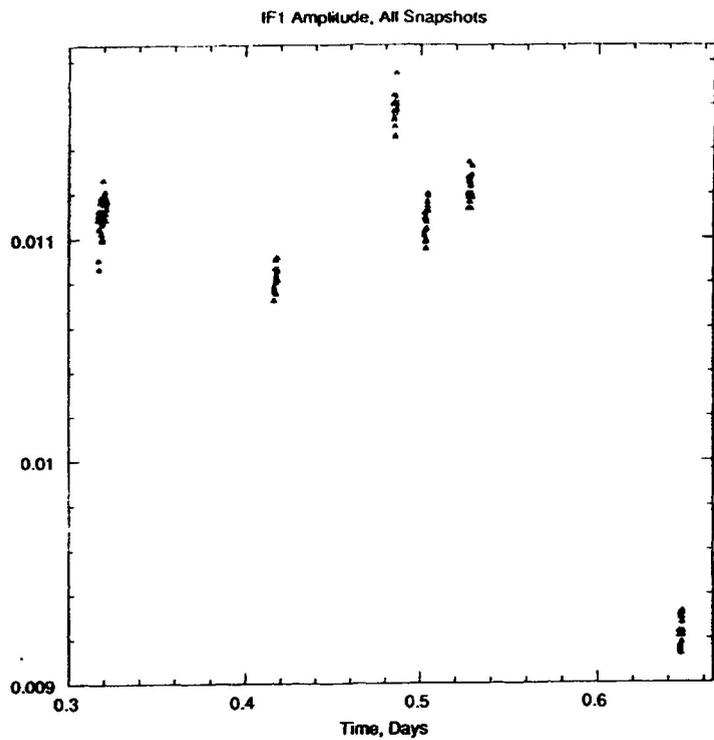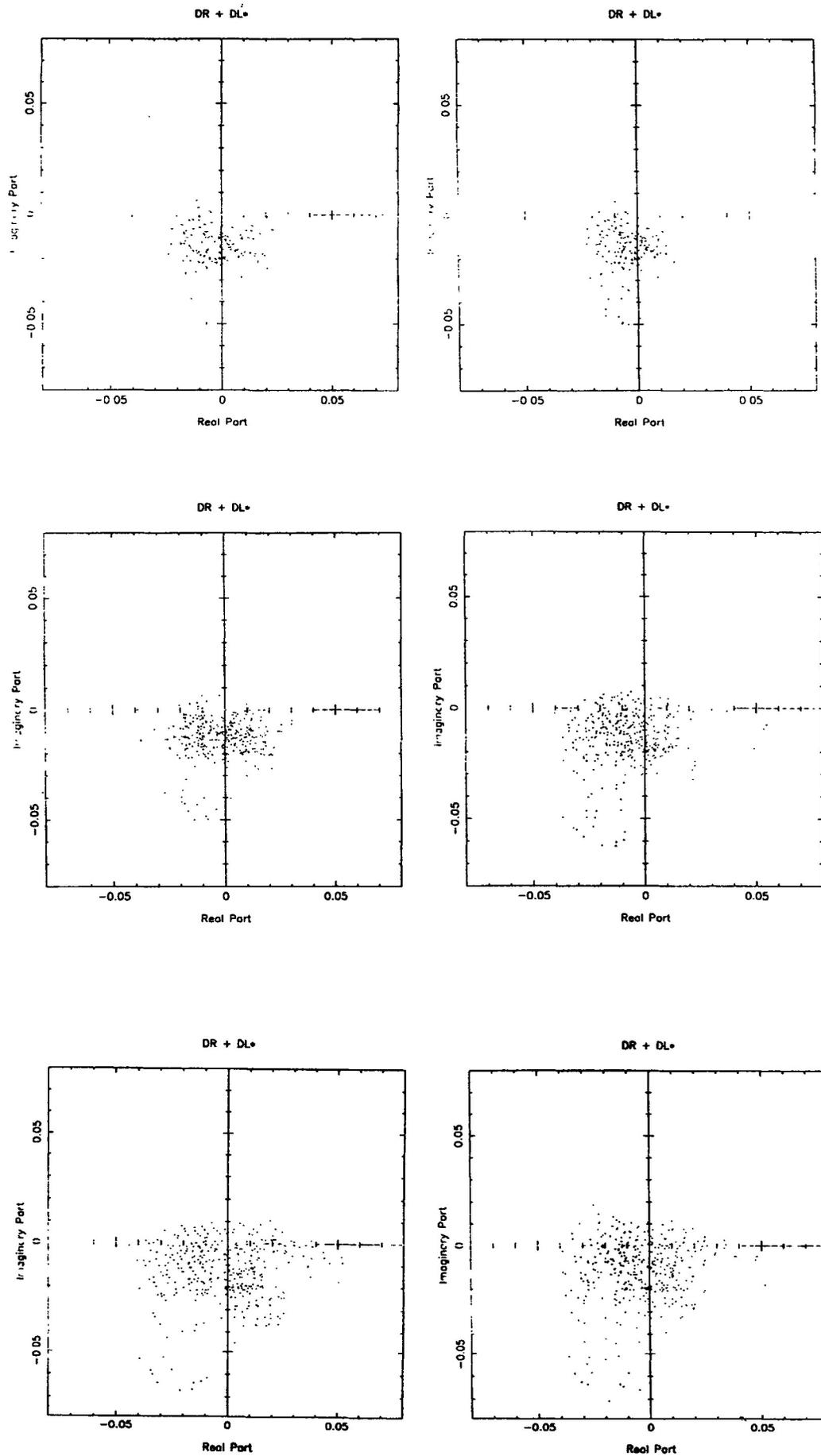
Figure 5

Figure 6

Figure 7

IF1 Phase, Snapshot 1

3 minutes

Time, Days
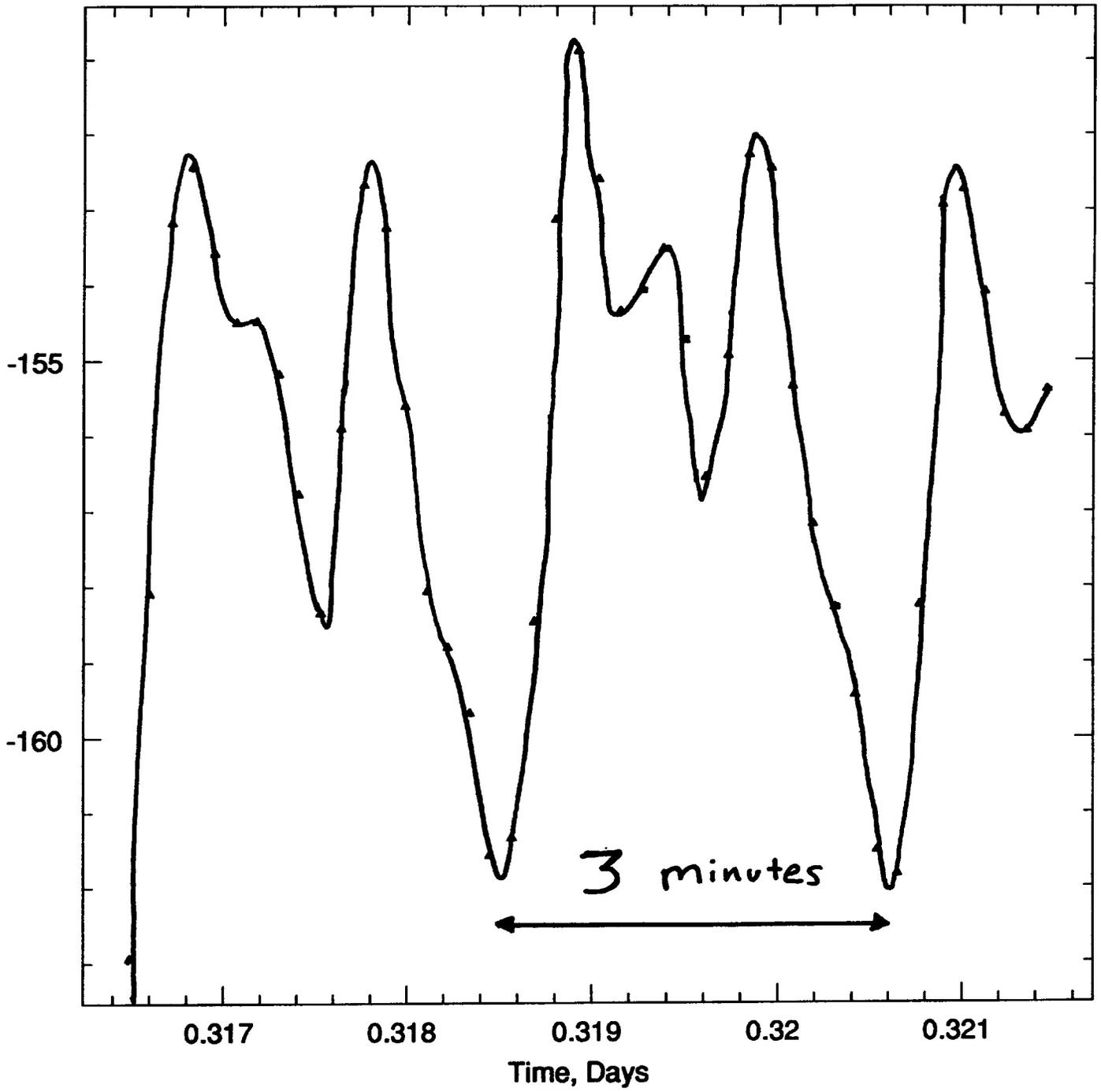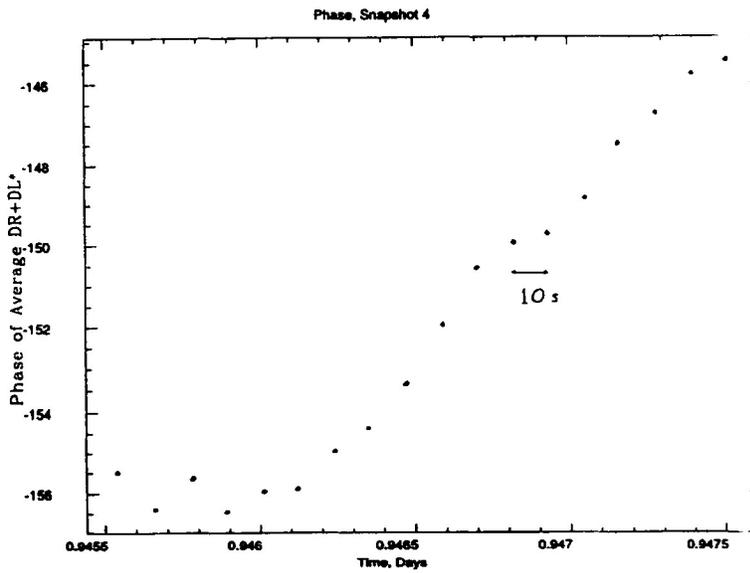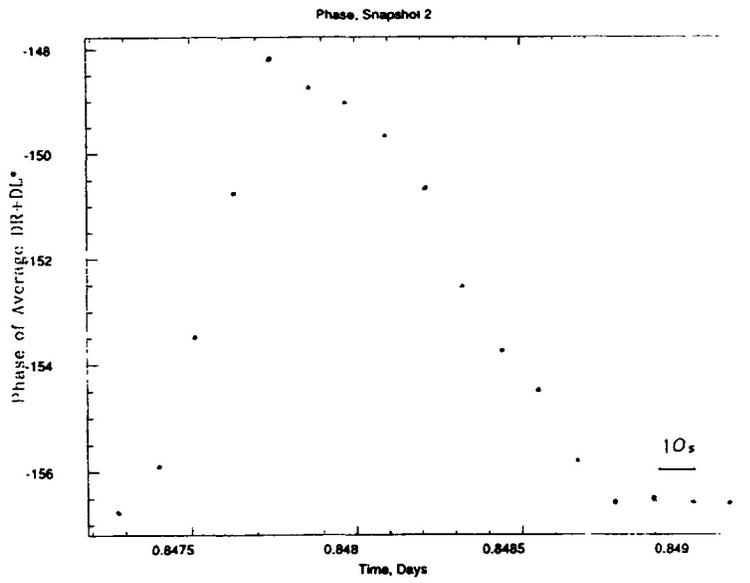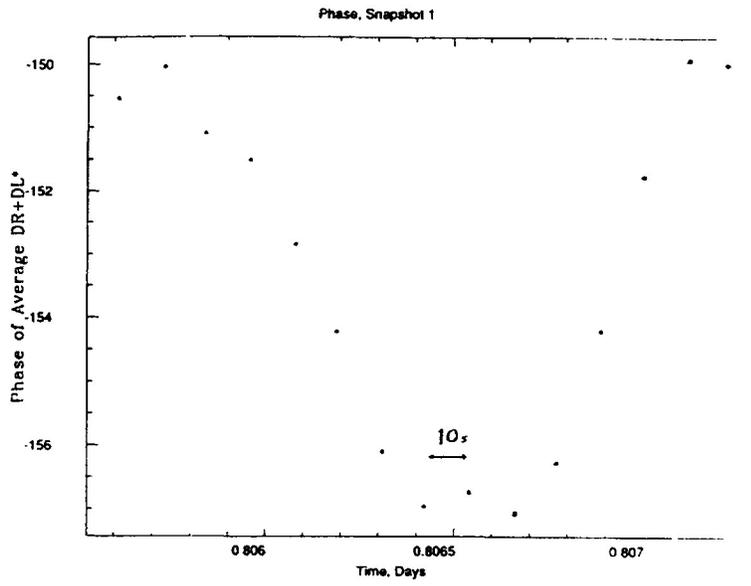
Figure 8a

Figure 8b

DR+DL* Amplitude, Snapshot 1

3 minutes

Time, Days

Figure 9