VLB ARRAY MEMO No. 337

Comment on the Use of Linearly Polarized Feeds for the VLBA

FRED SCHWAB
March 21, 1984

Barry Clark, in VLBA Memo. No. 320, discusses the use of linearly polarized feeds on the VLB Array elements, mentioning there that the use of linear polarization, instead of circular, can provide increased bandwidth and improved system temperature. I believe that the idea would be to use linear feeds only at 6 and 20 cm. Here I want to comment on a couple of implications of the use of linearly polarized feeds.

There would be a minor impact on self-calibration/hybrid mapping (and on global fringe search). Customarily in self-calibration—with circular feeds—the calibration parameters (the antenna/i.f. gains) for the right-hand (RCP) i.f.'s are derived independently of those for the LCP i.f.'s. In the case of linear feeds one would now need to do a simultaneous solution for the two sets of antenna/i.f. gains. Most often, the antenna gains are approximated by assuming the validity of a model $\tilde{V}_I$ of the visibility of the brightness distribution of Stokes' parameter I and by assuming the absence of circular polarization. In the case of circular feeds, the assumption that $V \equiv 0$ implies that the parallel-hand visibilities V_{RR} and V_{LL} are independent estimates of the gain-corrupted model visibility; hence, the RCP and LCP solutions can be done independently. In the case of linear feeds, one can still use the assumption that $V \equiv 0$, but the solutions cannot be done independently because the correlators that respond to Stokes' I respond to Q and U as well. (One would rather not assume the absence of linear polarization.)

This effect on self-calibration is simply a minor annoyance, in that software would have to be modified. Instead of two solutions—one for RCP and one for LCP—one would compute only one solution, for twice the number of parameters, given the data from all four correlators. There is also a minor impact on database design, in that the parallactic angle must be accessible to the self-calibration/hybrid mapping programs and to the global fringe search program. In calibrating the data taken with circularly polarized feeds, one normally doesn't need the parallactic angle, except in deriving the instrumental polarization.

Another minor consideration is that the VLBA is likely to be used frequently in conjunction with "outrigger" antennas having circularly polarized feeds. This, too, seems to present no real problem, except insofar as it complicates the software design. Since I don't know of anywhere else that they have been written down, I have written below the correlator responses for the case of linear feeds operating against circular feeds. For convenience, I also show the more familiar cases of linear–linear and circular–circular, using identical sign conventions in each instance.

The response of a correlator, assuming elliptical feed responses, can be expressed as

$$
\begin{aligned}
f(\varphi_1,\varphi_2,\theta_1,\theta_2) = \frac{\sqrt{2}}{2}\{ & (\cos(\varphi_1-\varphi_2)\cos(\theta_1-\theta_2) + i\sin(\varphi_1-\varphi_2)\sin(\theta_1+\theta_2))V_I \\
& +(\cos(\varphi_1+\varphi_2)\cos(\theta_1+\theta_2) + i\sin(\varphi_1+\varphi_2)\sin(\theta_1-\theta_2))V_Q \\
& +(\sin(\varphi_1+\varphi_2)\cos(\theta_1+\theta_2) - i\cos(\varphi_1+\varphi_2)\sin(\theta_1-\theta_2))V_U \\
& -(\cos(\varphi_1-\varphi_2)\sin(\theta_1+\theta_2) + i\sin(\varphi_1-\varphi_2)\cos(\theta_1-\theta_2))V_V\},
\end{aligned}
$$

where (φ_1,φ_2) denote the position angles of the feeds, and (θ_1,θ_2) denote the feed ellipticities, as measured by the arctangent of the axial ratio of the feed response (cf. Morris *et al.*, Ap. J., **139** (1964) 551–559). The reason for the normalization constant $\frac{\sqrt{2}}{2}$ is explained below.

In the case of circular feeds, the four correlator reponses are

$$
\begin{pmatrix} V_{\mathrm{RR}} \\ V_{\mathrm{RL}} \\ V_{\mathrm{LR}} \\ V_{\mathrm{LL}} \end{pmatrix}
= \begin{pmatrix} f(\chi_1,\chi_2,-\frac{\pi}{4},-\frac{\pi}{4}) \\ f(\chi_1,\chi_2,-\frac{\pi}{4},\frac{\pi}{4}) \\ f(\chi_1,\chi_2,\frac{\pi}{4},-\frac{\pi}{4}) \\ f(\chi_1,\chi_2,\frac{\pi}{4},\frac{\pi}{4}) \end{pmatrix}
= \frac{\sqrt{2}}{2}\begin{pmatrix} e^{-i(\chi_1-\chi_2)} & 0 & 0 & e^{-i(\chi_1-\chi_2)} \\ 0 & e^{-i(\chi_1+\chi_2)} & ie^{-i(\chi_1+\chi_2)} & 0 \\ 0 & e^{i(\chi_1+\chi_2)} & -ie^{i(\chi_1+\chi_2)} & 0 \\ e^{i(\chi_1-\chi_2)} & 0 & 0 & -e^{i(\chi_1-\chi_2)} \end{pmatrix}
\begin{pmatrix} V_I \\ V_Q \\ V_U \\ V_V \end{pmatrix},
$$

where (χ_1, χ_2) denote the parallactic angles of the feeds. The inverse relation is

$$\begin{pmatrix} V_I \\ V_Q \\ V_U \\ V_V \end{pmatrix} = \frac{\sqrt{2}}{2} \begin{pmatrix} e^{i(\chi_1-\chi_2)} & 0 & 0 & e^{-i(\chi_1-\chi_2)} \\ 0 & e^{i(\chi_1+\chi_2)} & e^{-i(\chi_1+\chi_2)} & 0 \\ 0 & -ie^{i(\chi_1+\chi_2)} & ie^{-i(\chi_1+\chi_2)} & 0 \\ e^{i(\chi_1-\chi_2)} & 0 & 0 & -e^{-i(\chi_1-\chi_2)} \end{pmatrix} \begin{pmatrix} V_{\mathrm{RR}} \\ V_{\mathrm{RL}} \\ V_{\mathrm{LR}} \\ V_{\mathrm{LL}} \end{pmatrix}.$$

It is useful to note that

$$\begin{pmatrix} V_I \\ V_Q \\ V_U \\ V_V \end{pmatrix} = \frac{\sqrt{2}}{2} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & -i & i & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} V_{\mathrm{RR}} e^{i(\chi_1-\chi_2)} \\ V_{\mathrm{RL}} e^{i(\chi_1+\chi_2)} \\ V_{\mathrm{LR}} e^{-i(\chi_1+\chi_2)} \\ V_{\mathrm{LL}} e^{-i(\chi_1-\chi_2)} \end{pmatrix},$$

so that one may correct the correlations for parallactic angle, and then safely average the data.

For one choice of the orientation/sign conventions, the linear–linear case yields the correlator responses

$$\begin{pmatrix} V_{\mathrm{VV}} \\ V_{\mathrm{VH}} \\ V_{\mathrm{HV}} \\ V_{\mathrm{HH}} \end{pmatrix} = \begin{pmatrix} f(\chi_1, \chi_2, 0, 0) \\ f(\chi_1, \chi_2 - \frac{\pi}{2}, 0, 0) \\ f(\chi_1 - \frac{\pi}{2}, \chi_2, 0, 0) \\ f(\chi_1 - \frac{\pi}{2}, \chi_2 - \frac{\pi}{2}, 0, 0) \end{pmatrix}$$

$$= \frac{\sqrt{2}}{2} \begin{pmatrix} \cos(\chi_1 - \chi_2) & \cos(\chi_1 + \chi_2) & \sin(\chi_1 + \chi_2) & -i\sin(\chi_1 - \chi_2) \\ -\sin(\chi_1 - \chi_2) & \sin(\chi_1 + \chi_2) & -\cos(\chi_1 + \chi_2) & -i\cos(\chi_1 - \chi_2) \\ \sin(\chi_1 - \chi_2) & \sin(\chi_1 + \chi_2) & -\cos(\chi_1 + \chi_2) & i\cos(\chi_1 - \chi_2) \\ \cos(\chi_1 - \chi_2) & -\cos(\chi_1 + \chi_2) & -\sin(\chi_1 + \chi_2) & -i\sin(\chi_1 - \chi_2) \end{pmatrix} \begin{pmatrix} V_I \\ V_Q \\ V_U \\ V_V \end{pmatrix}.$$

The inverse relation is

$$\begin{pmatrix} V_I \\ V_Q \\ V_U \\ V_V \end{pmatrix} = \frac{\sqrt{2}}{2} \begin{pmatrix} \cos(\chi_1 - \chi_2) & -\sin(\chi_1 - \chi_2) & \sin(\chi_1 - \chi_2) & \cos(\chi_1 - \chi_2) \\ \cos(\chi_1 + \chi_2) & \sin(\chi_1 + \chi_2) & \sin(\chi_1 + \chi_2) & -\cos(\chi_1 + \chi_2) \\ \sin(\chi_1 + \chi_2) & -\cos(\chi_1 + \chi_2) & -\cos(\chi_1 + \chi_2) & -\sin(\chi_1 + \chi_2) \\ i\sin(\chi_1 - \chi_2) & i\cos(\chi_1 - \chi_2) & -i\cos(\chi_1 - \chi_2) & i\sin(\chi_1 - \chi_2) \end{pmatrix} \begin{pmatrix} V_{\mathrm{VV}} \\ V_{\mathrm{VH}} \\ V_{\mathrm{HV}} \\ V_{\mathrm{HH}} \end{pmatrix}.$$

For the linear–circular case, one has

$$\begin{pmatrix} V_{\mathrm{VR}} \\ V_{\mathrm{VL}} \\ V_{\mathrm{HR}} \\ V_{\mathrm{HL}} \end{pmatrix} = \begin{pmatrix} f(\chi_1, \chi_2, 0, -\frac{\pi}{4}) \\ f(\chi_1, \chi_2, 0, \frac{\pi}{4}) \\ f(\chi_1 - \frac{\pi}{2}, \chi_2, 0, -\frac{\pi}{4}) \\ f(\chi_1 - \frac{\pi}{2}, \chi_2, 0, \frac{\pi}{4}) \end{pmatrix}$$

$$= \frac{1}{2} \begin{pmatrix} e^{-i(\chi_1-\chi_2)} & e^{i(\chi_1+\chi_2)} & -ie^{i(\chi_1+\chi_2)} & e^{-i(\chi_1-\chi_2)} \\ e^{i(\chi_1-\chi_2)} & e^{-i(\chi_1+\chi_2)} & ie^{-i(\chi_1+\chi_2)} & -e^{i(\chi_1-\chi_2)} \\ ie^{-i(\chi_1-\chi_2)} & -ie^{i(\chi_1+\chi_2)} & -e^{i(\chi_1+\chi_2)} & ie^{-i(\chi_1-\chi_2)} \\ -ie^{i(\chi_1-\chi_2)} & ie^{-i(\chi_1+\chi_2)} & -e^{-i(\chi_1+\chi_2)} & ie^{i(\chi_1-\chi_2)} \end{pmatrix} \begin{pmatrix} V_I \\ V_Q \\ V_U \\ V_V \end{pmatrix},$$

and the inverse relation

$$\begin{pmatrix} V_I \\ V_Q \\ V_U \\ V_V \end{pmatrix} = \frac{1}{2} \begin{pmatrix} e^{i(\chi_1-\chi_2)} & e^{-i(\chi_1-\chi_2)} & -ie^{i(\chi_1-\chi_2)} & ie^{-i(\chi_1-\chi_2)} \\ e^{-i(\chi_1+\chi_2)} & e^{i(\chi_1+\chi_2)} & ie^{-i(\chi_1+\chi_2)} & -ie^{i(\chi_1+\chi_2)} \\ ie^{-i(\chi_1+\chi_2)} & -ie^{i(\chi_1+\chi_2)} & -e^{-i(\chi_1+\chi_2)} & -e^{i(\chi_1+\chi_2)} \\ e^{i(\chi_1-\chi_2)} & -e^{-i(\chi_1-\chi_2)} & -ie^{i(\chi_1-\chi_2)} & -ie^{-i(\chi_1-\chi_2)} \end{pmatrix} \begin{pmatrix} V_{\mathrm{VR}} \\ V_{\mathrm{VL}} \\ V_{\mathrm{HR}} \\ V_{\mathrm{HL}} \end{pmatrix}.$$

The $\frac{\sqrt{2}}{2}$ normalization factor in the definition of the correlator response allows an identical constant factor ($\frac{1}{2}$ in the linear–circular case, $\frac{\sqrt{2}}{2}$ in the others) always to be factored out of a transformation matrix and its inverse. Each transformation matrix is unitary. I.e., in each case the matrix which relates the four correlator responses to Stokes' parameters has an inverse which is given simply by transposing the transformation matrix and conjugating its elements. Hence it's only necessary to memorize one matrix of each pair, not both.

Evidently, in the linear–circular case, if one wants to average the correlations, taking proper account of parallactic angle, then two averages must be computed. In the linear–linear case, four data per correlator must be carried along after averaging (the average of correlation $\times \cos(\chi_1 \pm \chi_2)$, $\times \sin(\chi_1 \pm \chi_2)$). In the circular–circular case, only one average needs to be computed. (The reason for averaging correlations before converting to Stokes' parameters is that the proper conversion to Stokes' parameters can't be made until after i.f.-dependent calibration has been applied, but often one wants to reduce the size of a database before the calibration has been completed.) Much of the time, the parallactic angle variation would be insignificant over the length of an averaging interval; but this won't always be the case, since some sources will transit nearly overhead some of the array elements.

In summary, the use of linearly polarized feeds would complicate the software design somewhat, and competent programmers would be required. The calibration software and the database probably ought to be designed to handle arbitrary mixes of feed ellipticities and feed orientations, if any of the VLBA feeds is other than nominally circular. One might argue, though, that the software ought to be able to handle the 'oddball' cases, even if the VLBA is designed with all circular feeds.