

Understanding radio polarimetry

V. Making matrix self-calibration work: processing of a simulated observation*

J. P. Hamaker

ASTRON, Netherlands Foundation for Research in Astronomy, Postbus 2, 7990 AA Dwingeloo, The Netherlands
e-mail: hamaker@astron.nl

Received 6 March 2006 / Accepted 7 June 2006

ABSTRACT

Context. This is Paper V in a series on polarimetric aperture synthesis based on the algebra of 2×2 matrices.

Aims. It validates the matrix self-calibration theory of the preceding Paper IV and outlines the algorithmic methods that had to be developed for its application.

Methods. New avenues of polarimetric self-calibration opened up in Paper IV are explored by processing a simulated observation. To focus on the polarimetric issues, it is set up so as to sidestep some of the common complications of aperture synthesis, yet properly represent physical conditions. In addition to a representative collection of observing errors, the simulated instrument includes strongly varying Faraday rotation and antennas with unequal feeds. The selfcal procedure is described in detail, including aspects in which it differs from the scalar case, and its effects are demonstrated with a number of intermediate image results.

Results. The simulation's outcome is in full agreement with the theory. The nonlinear matrix equations for instrumental parameters are readily solved by iteration; a convergence problem is easily remedied with a new ancillary algorithm. Instrumental effects are cleanly separated from source properties without reference to changes in parallactic rotation during the observation. Polarimetric images of high purity and dynamic range result. As theory predicts, polarimetric errors that are common to all sources inevitably remain; prior knowledge of the statistics of linear and circular polarization in a typical observed field can be applied to eliminate most of them.

Conclusions. The paper conclusively demonstrates that matrix selfcal per se is a viable method that may foster substantial advancement in the art of radio polarimetry. For its application in real observations, a number of issues must be resolved that matrix selfcal has in common with its scalar sibling, such as the treatment of extended sources and the familiar sampling and aliasing problems. The close analogy between scalar interferometry and its matrix-based generalisation suggests that one may apply well-developed methods of scalar interferometry. Marrying these methods to those of this paper will require a significant investment in new software. Two such developments are known to be foreseen or underway.

Key words. instrumentation: interferometers – instrumentation: polarimeters – methods: data analysis – methods: observational – techniques: interferometric – techniques: polarimetric

1. Introduction

This article is the fifth in a series in which a physics-based theory of polarimetric aperture synthesis is developed. Starting from a detailed description of a single interferometer in Paper I (Hamaker et al. 1996) and a qualitative investigation of a synthesis array as a whole in Paper II (Sault et al. 1996), the main line of our work progressed to the consideration of matrix-based self-calibration¹. Paper IV (Hamaker 2000) demonstrated that such calibration is indeed possible; the paper considers in mathematical terms the ambiguities that the solution is subject to and points out the close analogies with and important differences from scalar selfcal. The theory, however, gives no answer to the many questions one must ask about a practical implementation. This paper seeks to fill that void, by studying a computer simulation in which an observation with instrumental errors of a sky field containing unpolarized and polarized sources is self-calibrated with matrix-based algorithms.

It opens, in Sect. 2, with a brief summary of the results of Paper IV that form the basis of the present work. It highlights the distinction between the initial *self-alignment*, in which Sky and Instrument Models are developed, and the subsequent *post-calibration* in which the ambiguities left after self-alignment must be resolved. The simulated Sky field and Instrument are described in Sect. 3. Section 4 considers the matrix self-alignment cycle in detail; post-calibration is discussed in Sect. 5. Various complications along with expedients to deal with them are explained and illustrated with Sky images from the simulation. A few side remarks are collected in Sect. 6. Section 7 sums up the results and their possible consequences.

We assume that all antennas see the same “apparent” Sky, i.e. they have identical primary antenna beams; this is a prerequisite for all forms of self-calibration².

* Appendices are only available in electronic form at <http://www.edpsciences.org>

¹ Paper III (Hamaker & Bregman 1996) considers the matter of Stokes-parameter definitions which is of no relevance here.

² One may question this assumption for antennas with different feed orientations as appear in the simulation to be described. I assume here that the primary beam is the product of independent scalar effects of a circularly symmetric (pair of) reflector(s) (which is the same for all antennas) and the polarimetric effects of the feeds.

Table 1. Poljargon.

polxxx	= polarization xxx
quasi-scalar	refers to the conventional linearising approach to radio polarimetry
Stokes	
<i>I</i>	intensity
polvector <i>p</i>	$\mathbf{p} = (Q, U, V)^T$
<i>pre-calibration</i>	initial correction of an observation based on prior data
self-alignment	proper name for what is traditionally called <i>self-calibration</i>
<i>post-calibration</i>	calibrating the errors that self-alignment leaves undetermined
polconversion	mutual conversion $I \leftrightarrow \mathbf{p}$
polrotation	rotation of <i>p</i> in <i>QUV</i> space
pol distortion	product of polconversion and polrotation
<i>poldeconversion</i>	correcting for polconversion
<i>polvector leveling</i>	polrotating polvectors to minimise $\sum V^2$
Sky and Instrument Models	joint solution of the self-align equations
heterogeneous array	array in which antennas have identical optics but polarimetrically different feeds

New terms introduced in this paper are italicised.

1.1. Notation and terminology

The notation for mathematical symbols is the same as in the preceding papers. Bold-faced capitals represent 2×2 matrices, bold-faced lower-case characters are column 2-vectors. For hermitian transposition or conjugation a dagger superscript is used. Roman font is reserved for constants, variables are italicized³. Unless noted otherwise, primed variables indicate estimates obtained by model fitting.

The subscripts *j, k* denote the array's antennas, in combination they indicate an interferometer. A third subscript *t* distinguishes the successive time intervals in which observed coherencies are recorded. For consistency with the coherency \mathbf{E}_{jkt} , the notation $\mathbf{B}(I)$ for the Sky brightness in Paper IV has been replaced with \mathbf{B}_I .

Paper IV rather ineptly uses the “variance” function $\text{Var } \mathbf{M}$ for what is known in linear algebra as the square of the Frobenius norm, $\|\mathbf{M}\|^2$. I now use the latter notation.

The concepts and associated terminology developed in Paper IV play a crucial role in the present work. For the reader's convenience Table 1 summarises them here; a few new terms to be coined in this paper are also included.

2. Synopsis of previous work

This section summarises those parts of Paper IV that are relevant here; for full details, refer to that paper.

2.1. Coherency matrix and Measurement Equation (M.E.)

The electric field is represented by a column vector \mathbf{e} in cartesian or circular coordinates. Depending on whether the coherency is represented as a 4-vector (Paper I) or a 2×2 matrix (Paper IV), the interferometer transfer equation can take two different forms. Here I use the *coherency matrix*

$$\mathbf{E}_{jkt} = \langle \mathbf{e}_{jt} \mathbf{e}_{kt}^\dagger \rangle. \quad (1)$$

With this definition, the interferometer equation, now widely known as the *Measurement Equation (M.E.)*, takes the form

$$\mathbf{W}_{jkt} = \mathbf{J}_{jt} \mathbf{E}_{jkt} \mathbf{J}_{kt}^\dagger, \quad (2)$$

³ Note the difference between Stokes parameters *U, V* and visibility/coherency coordinates *u, v*.

Table 2. Scalar-matrix analogies.

Scalar	Matrix
Complex gain	Jones matrix
Visibility	Coherency matrix
Brightness	Stokes <i>I</i> and <i>p</i> brightnesses
Sky image	Quartet of images, one for each of the 4 Stokes parameters

in which all symbols represent 2×2 matrices and $\{\mathbf{W}_{jkt}\}$ are the measured coherencies. This form emphasises an analogy with the conventional scalar interferometer equation of which it is a generalisation. The transfer matrices \mathbf{J} are known as *Jones* matrices.

Paper IV stresses the analogies between the algebras for scalars and matrices, which may be very helpful in developing an intuitive understanding of the latter. The most important correspondences are listed in Table 2.

By default, in this paper each Jones matrix represents a complete (dual) antenna path from the source down to the pair of voltage inputs to the correlator. It includes not only scalar gain and phase effects in parallel *X* and *Y* or *L* and *R* signal paths, but also interactions between the two, notably rotations and the cross-polarization in the feeds. (In the conventional *quasi-scalar* approach, such effects must be separately accounted for.) The set $\{\mathbf{J}_{jt}\}$ of all Jones matrices will be called *the Instrument*.

Self-calibration operates entirely on the *apparent* Sky, i.e. the Sky multiplied with the primary antenna beam; in the matrix case, the latter is a Jones-matrix function of position *I*. Correction for this is best postponed until self-calibration is complete; it requires precise knowledge of the Jones-matrix primary beam – which is beyond the scope of this paper.

2.2. Stokes parameters: intensity and polvector

We introduce the *Stokes parameters* *I, Q, U* and *V* by representing the coherency matrix in the form⁴

$$\mathbf{E} = \begin{pmatrix} I + Q & U - iV \\ U + iV & I - Q \end{pmatrix}. \quad (3)$$

⁴ The sign of *V* is given here as in Paper IV. R.J. Nijboer (priv. comm.) has noted that it deviates from the definitions given in Papers I and III (Hamaker & Bregman 1996). This is an unfortunate mistake, but it is immaterial in the present context.

Defining the *polvector* $\mathbf{p} = (Q, U, V)^T$, Paper IV introduces the symbolic notation

$$\mathbf{E} = [I + \mathbf{p}] \quad (4)$$

which summarises the well-known dichotomy of the Stokes-parameter quartet in a concise form. In practice one often considers linear polarisation separately; the corresponding *linear polvector* is $\mathbf{p}_{\text{lin}} = (Q, U)^T$.

Since the coherency \mathbf{E} and the Sky brightness \mathbf{B} are related by a Fourier transformation (van Cittert-Zernike theorem) the above two equations also apply to \mathbf{B} . I shall use the same notation for Stokes parameters, intensity and polvector in both domains; the distinction will be clear from the context.

2.3. Self-alignment

The scalar/matrix analogy suggests that the hugely successful scalar method of *self-calibration* can be generalized to the matrix domain – which is indeed the case.

Given a set of observations $\{\mathbf{W}_{jkt}\}$, self-calibration seeks to invert the M.E. to obtain values for $\{\mathbf{E}_{jkt}\}$ and the corresponding Sky brightness \mathbf{B}_l (where l are Sky coordinates). This is an ill-posed problem but, as in the scalar case, it can be regularized by the customary assumption that the Sky is “mostly empty”. Along with the *Sky Model* \mathbf{B}'_l , the solution includes an *Instrument Model* $\{\mathbf{J}'_{jt}\}$.

In the case of scalar selfcal the solution is unique except for an unknown positive multiplier⁵: in addition to the correct Sky image $\mathbf{B}'_l \equiv \mathbf{B}_l$ and the corresponding visibilities $\{\mathbf{E}_{jkt}\}$, all positive multiples of them also satisfy the M.E. provided one multiplies $\{\mathbf{J}'_{jt}\}$ with a compensating factor.

Paper IV derives the analogous property for the matrix case. The Sky-image solution that one obtains is – apart from the convolution with a scalar synthesised beam – related to the true Sky by a *poldistortion* transformation:

$$\mathbf{E}'_{jkt} = \mathbf{X} \mathbf{E}_{jkt} \mathbf{X}^\dagger \quad \forall j, k, t \quad (5)$$

$$\mathbf{B}'_l = \mathbf{X} \mathbf{B}_l \mathbf{X}^\dagger \quad \forall l, \quad (6)$$

where \mathbf{X} is an arbitrary nonsingular matrix. Using mathematical jargon, I shall refer to Eqs. (5) and (6) as *the transformation X* of \mathbf{E}_{jkt} and \mathbf{B}_l . To satisfy Eq. (2), the factor \mathbf{X} in Eq. (5) must be cancelled in the Instrument Model: $\{\mathbf{J}'_{jt}\} = \{\mathbf{J}_{jt} \mathbf{X}^{-1}\}$.

As argued in Paper IV, the solution \mathbf{B}' is not really calibrated, since it includes the unknown factor \mathbf{X} ; we use the more precise term *self-alignment* instead. The solution \mathbf{B}' self-aligns the observations to satisfy our prior assumptions about the Instrument (no interferometer-based errors) and the Sky (“mostly empty”). Note that Eq. (6) is an *in-place* transformation that operates in the same way on every point in the image.

Like scalar selfcal, matrix self-alignment eliminates all the variable errors in the Instrument and replaces them with one error factor \mathbf{X} that is constant over all antennas j and times t . This mechanism applies not only to gains and phases, but also to polarimetric errors such as polarization leakage and rotations of linear polarization as well as intentional differences between antenna feeds (*heterogeneous* array).

2.4. Polconversion and polrotation

The transformation Eq. (6) can be factored as

$$\mathbf{B}'_l = |x|^2 \mathbf{H} (\mathbf{Y} \mathbf{B}_l \mathbf{Y}^\dagger) \mathbf{H}^\dagger. \quad (7)$$

Not surprisingly, the positive scale factor $|x|^2$ reappears in the matrix case. The new transformations \mathbf{H} and \mathbf{Y} are best described in terms of the Stokes representation Eq. (4):

- The “positive hermitian” transformation \mathbf{H} represents both the well-known *leakage* of Stokes I flux into \mathbf{p} and the conversion of \mathbf{p} to I flux that is its Siamese twin. The term *polconversion* describes the combination of both.
- The “unitary” transformation \mathbf{Y} describes a *polrotation*, i.e. a rotation of \mathbf{p} in QUV space. Polrotation does not affect I and leaves $|\mathbf{p}|$ invariant.

Polrotation and polconversion operators can be interchanged: Instead of Eq. (7) we may also write

$$\mathbf{B}'_l = |x|^2 \mathbf{Y} (\mathbf{H}' \mathbf{B}_l \mathbf{H}'^\dagger) \mathbf{Y}^\dagger \quad (8)$$

where the polconversion \mathbf{H}' is different from \mathbf{H} (cf. Eq. (B.9) in the Appendix).

2.5. Post-calibration

Inverting the M.E. by self-alignment takes us only halfway; while being consistent with the observations, the Sky image \mathbf{B}' that we obtain differs from the true one. *Post-calibration* is required to eliminate the spurious transformation \mathbf{X} . Here the M.E. is to no avail: all that we can extract from it is already in the self-aligned solution \mathbf{B}' . Instead, a priori external information must be exploited to constrain \mathbf{X} .

Self-alignment and post-calibration are, in principle, mutually independent. While inevitably and uncontrollably introducing unknown \mathbf{H} and \mathbf{Y} transformations in its solutions, self-alignment is not affected by them. The latter is also true when we transform the data mid-stream – probably with the intention of making the process easier to control (cf. Sect. 5.1). In the theory, self-alignment precedes post-calibration, but in practice we may take an advance on the latter whenever we see fit to do so.

2.6. Persistence of poldistortion; pre-calibration

Self-alignment cannot eliminate poldistortion because it cannot recognise it. More generally, a proper self-aligning algorithm should have little tendency to change the poldistortion with which it is initialised. Basically, one expects it to just transmit what it is offered.

Thus poldistortion, once introduced, will be transferred to the Instrument Model, hence to the Sky image and the Sky Model extracted from it, after which the next self-alignment cycle will transfer it to the next Instrument model, etc. To keep poldistortion low, it is important to prevent it from being inadvertently introduced. It is therefore important to apply a *pre-calibration* using the best initial *nominal* Instrument Model that one can get hold of.

2.7. Practical questions

The theory is clear on what we can and cannot expect self-alignment to achieve, but it does not tell us how a solution is obtained. In the scalar case the nonlinear selfcal equations can be

⁵ As in Paper IV, we ignore the loss of absolute position in the Sky that self-calibration also entails.

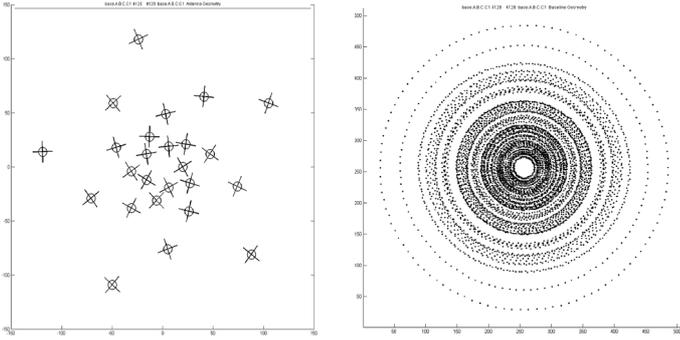


Fig. 1. Synthesis-telescope geometry. *Left:* Antenna positions and feed orientations. *Right:* uv coverage.

solved iteratively, or linearised by taking logarithms; both methods have been used successfully and the solutions are usually robust. The non-commutativity of matrix multiplication precludes the use of logarithms, so solving the nonlinear equations directly is our only option. Stability and robustness of this process must be investigated.

Once we have a solution, we must consider post-calibration: How can we bring in prior astronomical and instrumental knowledge to suppress the unknown poldistortion?

3. Processing a simulated observation

To answer these questions and develop practical algorithms, an observation with realistic errors of various sorts was simulated. The theory was used as a guideline in self-calibrating the observation.

A major advantage of a simulation is that it can be set up to focus on the polarimetric problem per se and sidestep the several distracting complications that beset aperture synthesis in the real world, such as interference, uv -convolution and aliasing, deficiencies of CLEAN, Sky curvature and projection effects. The following are the main features of the simulated observation:

- The telescope (Fig. 1) consists of 25 antennas on 5 spiral arms. Observations are made at 11 equidistant hour angles covering 2.4 h, to produce a uv coverage that is uniform in position angle; radial coverage is less regular, but this is unimportant for the point sources to be observed. The total number of uv samples is 3300.
- The antennas are equipped with linearly polarized feeds in orientations that are uniformly distributed but assigned at random, as shown in Fig. 1. The choice for a heterogeneous array was made with the aim to confirm theoretical predictions in Paper IV. We shall return to this point in Sect. 6.2. Position-angle and ellipticity errors are drawn from Gaussian distributions centered on zero with an rms spread of some 2° .
- Ionospheric Faraday rotation, identical for all antennas, varies randomly during the observation by more than a full turn with a characteristic time scale of about an hour. True gains and phases differ from the nominal ones with a peak-to-peak spread of 20% and 10° , and smooth time variations ranging over 2% and 3° , respectively. (In retrospect, the phase variations seem to be lower than one may expect in a real instrument; in a later trial run phase errors were tripled without jeopardising the self-calibration in any way.)
- In the Sky field, 130 point sources are randomly placed. The fluxes have a realistic $\log N - \log S$ distribution. 10 sources are linearly polarized, degrees being in the lower % range

with one maximum of 10%; one source is also 0.3% circularly polarized. Total I flux $\sum I = 42$ units, total absolute polarized flux $\sum |p| = 0.5$ units.

- CLEAN deficiencies are avoided by simulating only point sources at grid positions. For clusters of sources that CLEAN could not disentangle, spurious sources in the Sky Model were suppressed by jointly fitting fluxes for all positions in the cluster area where a source had been found.
- uv -convolution and image-aliasing complications are avoided by shifting antennas to grid positions for every integration time interval – as if they were “manually” moved over the Earth’s surface.
- Putting both the telescope and the source field on the North Pole avoids projection deformations. Putting the antennas in an Equatorial plane makes the van Cittert-Zernike relation between brightness and coherency functions rigorously two-dimensional.
- Gaussian noise was added at a level of 10^{-4} per image point – as would be the case for a Sky field with strong sources; it proved to be too low to noticeably affect the data processing.

Although some of these simplifications are obviously unrealistic from an operational viewpoint, they are all physically valid and do not detract from the purpose of the exercise.

4. Matrix self-alignment

4.1. Synopsis of quasi-scalar selfcal

The matrix-based self-align procedure is shown in the top half of Table 3. It corresponds to conventional scalar self-calibration, extended with a number of matrix-specific algorithms. In relating scalar and matrix self-alignment, one may utilise the correspondences listed in Table 2.

The procedure simultaneously develops *Models* of the apparent Sky \mathbf{B}'_l and the Instrument $\{\mathbf{J}'_{jl}\}$ that together satisfy Eq. (2) given the observed coherencies $\{\mathbf{W}_{jkl}\}$. The following cycle is repeated iteratively:

1. Make a *dirty image* of the Sky from the observed coherencies corrected with the best available Instrument Model.
2. Extract from this image a *Sky Model* through some deconvolution procedure; I refer to this process as CLEANing since that is the algorithm that I used. Calculate the corresponding *Model Coherencies*.
3. Assuming that the Model Coherencies were the actual input to the observation, fit the Instrument Model that provides the best match to the coherencies actually observed, per time interval t . (Some writers understand “self-calibration” to refer to this particular step rather than the iterative process as a whole.)
4. Start a new cycle, using the newly acquired Instrument Model.

4.2. Matrix extras

The elements in the self-alignment part of Table 3 that have no counterpart in scalar self-alignment are shown in boldface.

4.2.1. Instrument fit

I used the relaxation algorithm (Appendix A.1 below) of Paper IV. Alternatively, a general-purpose non-linear equation

Table 3. Self-alignment and post-calibration in the simulation. Four iteration cycles are shown. Elements that deviate from, or are absent in, scalar selfcal are set in **boldface**. Operation classes: i = making Instrument Model; m = making Sky image; p = post-calibration operation; s = making Sky Model.

Notes	Cycles				Class	Operation
	1	2	3	4		
	•	•	•	•		⇒ <i>data</i>
	•	•	•	•		⇒ <i>Raw observed coherencies</i>
						SELF-ALIGNMENT
	•	•	•	•		Enter next cycle
	•				m	Pre-calibrate: Correct using prior Jones matrices
		•	•	•	m	Correct using estimates of Jones matrices from previous cycle
		•	•	•		⇒ <i>Corrected coherencies</i>
		•	•	•	m	Convert to image in Stokes format
		•	•	•		⇒ <i>I dirty image</i>
1		•	•	•		⇒ p dirty images
	•	•	•	•	s	CLEAN Stokes <i>I</i> image
3	•	•	•	•		⇒ <i>I residual image</i>
	•	•	•	•		⇒ <i>I Sky Model: List of source positions and I fluxes</i>
1		•	•	•	s	CLEAN p images jointly at I positions (Sect. 4.2.2)
1		•	•	•		⇒ <i>p Sky Model: List including corresponding p fluxes</i>
1, 3		•	•	•		⇒ p residual images
2			•	(•)	p	Poldeconversion: Invert polconversion (Sect. 4.2.3)
			•	•	p	⇒ <i>“Leakage”-free Sky Model</i>
			•	•		Exit to post-calibration
	•	•	•			Fourier-transform to coherency matrix form
	•	•	•			⇒ <i>Coherency Model</i>
	•	•	•		i	Fit Instrument to raw coherencies and Coherency Model
	•	•	•			⇒ <i>Improved estimates of Instrument</i>
	•	•	•			Go to next cycle
						POST-CALIBRATION
2			•	(•)	p	Poldeconversion: Invert polconversion (Sect. 5.1.1)
			•	•	p	Polvector leveling (Sect. 5.1.2)
		•	•	•		⇒ <i>Sky Model with unknown rotation of p_{in}</i>
		•	•	•	p	Factor Instrument model in terms of receiver hardware (Sect. 5.2)
		•	•	•		⇒ <i>Jones matrices representing antenna-channel stages</i>
		•	•	•	p	Compare these Jones matrices with prior knowledge about stages
		•	•	•		⇒ <i>Estimated errors in Jones matrices of stages</i>
		•	•	•	p	Correct coherency model for these errors
		•	•	•	m	Convert to final image in Stokes format
		•	•	•	p	Calibrate brightness on known brightness/flux calibrators
		•	•	•	p	Verify/calibrate p_{in} on known polarization calibrators (if any)

Notes:

- 1 No attempt is made in the first cycle to process *p* data: see Sect. 4.3.2.
- 2 This is an example of a post-calibration operation interjected into the self-alignment process. See Sect. 5.1.
- 3 Except possibly for the final one, the residual images are a by-product that is only used for judging the progress of the process.

solver routine may be used. In either case convergence turns out to be a problem – to be discussed in Sect. 4.3.

Because of its simplicity, the Paper IV algorithm is very suitable for experiments of the present type and cheap in execution. However, M. Brentjens (priv. comm.) finds that it may easily get stuck in a secondary minimum when many interferometers are missing (e.g. because of interference). He patched a general-purpose solver to deal with this problem. Whether the same can be done for the Paper IV algorithm we do not know.

4.2.2. Polarization CLEAN

An obvious first step is to CLEAN the intensity image in the standard way.

For an “empty” Sky, it is a valid assumption that polarized flux can only exist where there is intensity⁶. The *Q, U, V* images are actually three components of an image of the

⁶ Recent work has revealed the existence of extended structures that may appear in interferometry to be more than 100% polarized; for such situations not only our CLEAN method, but the entire selfcal paradigm must be reconsidered.

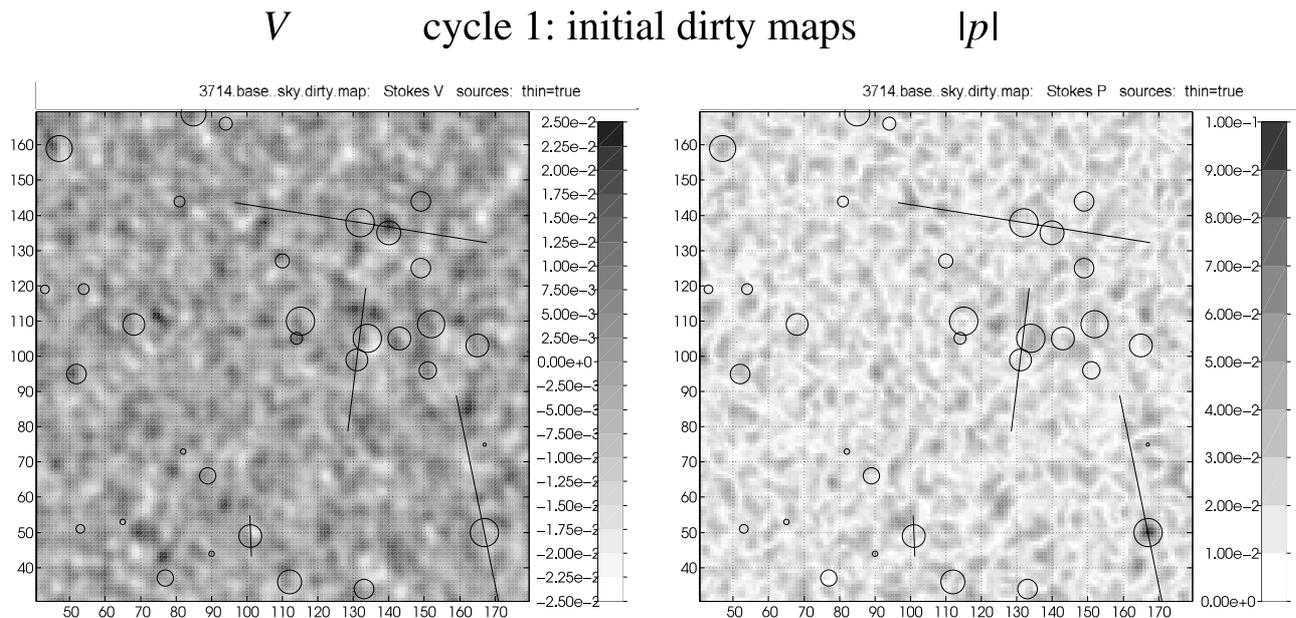


Fig. 2. Cycle-1 dirty images, made with nominal instrumental corrections; *left* V , *right* $|p|$. Circles represent positions and fluxes of true sources (some being too weak to be shown), lines their linear polarization. Levels are to be compared with the peak source flux of 1. In the $|p|$ image only one source can be identified; its high flux must be accidental, since the other p sources do not show up.

polvector p . The way to CLEAN them jointly is motivated by poldistortion theory: Apart from polconversion (to be discussed later), the principal error that remains in the final self-alignment cycles is a polrotation that leaves $|p|$ invariant. Thus, we may use a $|p|$ image to identify components, searching only the locations of I sources. Having identified a peak, we determine the individual Q , U , V fluxes and subtract components multiplied by the loop gain from the corresponding images.

This approach carries the obvious advantage that it is less sensitive to noise than searches in the separate Q , U and V images would be. It proved to be quite effective even in the early cycles, where errors other than polrotation might confuse the situation.

4.2.3. Poldeconversion

For reasons to be discussed in Sect. 5.1, one may take an advance on post-calibration by interjecting Poldeconversion (removing polconversion) in the self-alignment procedure.

4.3. Pitfalls

4.3.1. Sluggish convergence

In the Instrument fit, residuals rapidly decrease in the initial iterations. After that, the polrotation factor in the Instrument model continues to be adjusted in minute steps in an attempt to match the observed and Sky-Model polvectors; final convergence may require tens of thousands of steps. Apparently the polarized flux, totalling only about 1% of the intensity flux, is too weak to effectively drive the algorithm⁷. A complementary algorithm (Appendix A.2) addresses this problem.

⁷ S.J. Bhatnagar (priv. comm.) observed the same effect when using an entirely different algorithm in *AIPS++*. This strongly suggests that this behaviour is due to the nature of the problem rather than to a peculiarity of either algorithm.

4.3.2. Persistence of image errors

I discussed above how errors once introduced may persist in the further processing. Such errors may also be introduced through the p Sky Model when it is extracted from a poor image.

Figures 2 and 3 illustrate a case seen in practice. The pre-calibrated dirty image is of bad quality. The V image, that should be almost empty, is indicative of the error level in all three of the p images. Yet a few sources were extracted from the p image and included in the Sky Model. Several self-align cycles later, Stokes U was noted to have the wrong sign; the error was traced back to the very first p model.

In a re-run, p modeling was postponed. An I -only Sky from the first cycle served in cycle 2 to correct most of the complex-gain errors⁸. This improved the p images to the point where the major polarized sources could be reliably identified (Fig. 3).

5. Post-calibration

Post-calibration is the process of eliminating the spurious poldistortion X that appears in Eq. (6). Prior knowledge, both astrophysical and instrumental, may be harnessed for this purpose. The methods to be described rely on X being the same for the entire Sky field, as was discussed above.

5.1. Exploiting prior astrophysical information

While Stokes I is necessarily positive, p approaches zero when summed over many (independent) sources; also, usually $|V| \ll |p_{\text{in}}|$ (although important exceptions exist, see Fender & Macquart 2003). Imposing these empirical findings as prior constraints on our Sky image, we can estimate the polarimetric errors and correct the image and our Sky and Instrument Models for them.

⁸ This is similar to although not the same as the separate self-calibration of the X and Y or L and R channels with which quasi-scalar polarization processing starts up (see Paper II).

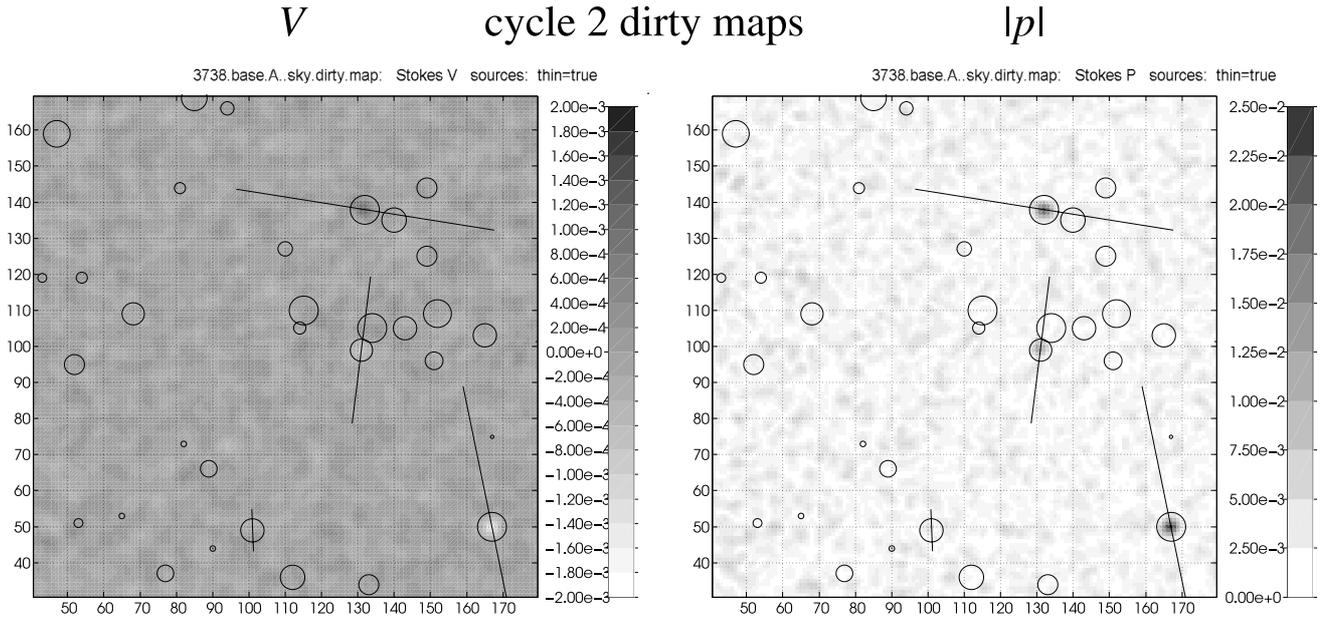


Fig. 3. Cycle-2 dirty images, corrected for instrumental errors fitted with the I -only Sky Model of cycle 1; *left* V , *right* $|p|$. Circles represent positions and fluxes of the true sources, lines their linear polarization. Note the difference in intensity scales with Fig. 2. The peak levels of .002 in the V image and .025 in the $|p|$ image are to be compared with the peak source flux of 1. Several polarized source now stand out; their fluxes are still too low by a factor 4 because polrotations at the successive hour angles are misaligned (cf. Sect. 6.3).

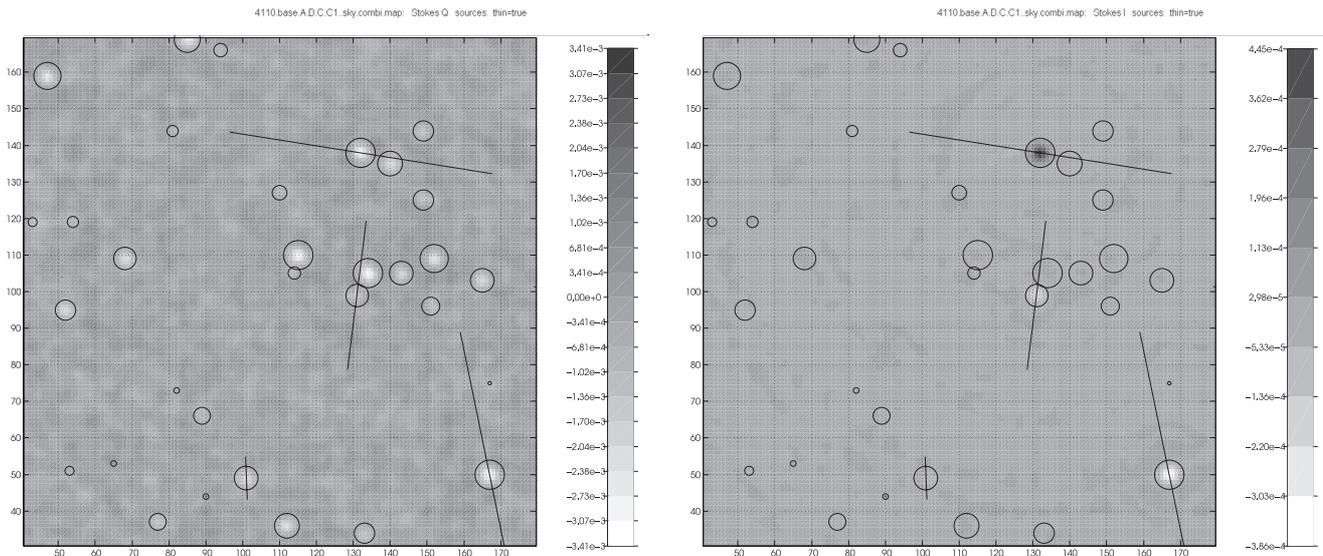


Fig. 4. Poldeconversion. Shown are the differences of Q (*left*) and I (*right*) images after poldeconversion with those before. Circles mark the true source positions with areas proportional to intensity, bars indicate orientation and strength of true linear polarization. Note the flux scales. The Q image shows that poldeconversion has removed Q flux from every source. In the I image differences exist only where the source is truly polarized; positive and negative I differences reflect the bipolar nature of Q .

5.1.1. Poldeconversion

Polconversion adds to all sources in a Sky image a polarization component with one and the same relative polarization p/I . This makes it quite conspicuous in a comparison of the I Sky Model and the CLEAN-residual p images.

Poldeconversion is the operation of correcting for the polconversion matrix H in Eq. (6). The simplest method is to keep polconverted p flux out of the Sky Model, which one may do by stopping CLEAN before it starts picking it up significantly. In the immediately following Instrument fit, the polconversion-free Sky Model will then force the Instrument model to comply.

Alternatively, one may use the Sky Model to subtract the sources identified as truly polarized from the newest dirty images. The p flux in the remainder must then be polconverted I flux and the polconversion matrix H can be determined by a comparison (Appendix B.1). An example is shown in Fig. 4.

In any method of separating instrumental artifacts from true object features, some measure of human intervention may be necessary. In matrix selfcal, true source polarization must be distinguished from polconverted intensity. Since self-alignment produces very pure p images, this was simple in my simulations; it may be less so with more complicated source

Table 4. The effect of polvector leveling. Shown is a list of the polarized sources in the final Sky Model. The true V fluxes are compared with those obtained through self-alignment before and after the leveling. Leveling reduces all of them to near-zero except for one source which is indeed the one that is truly polarized.

I flux	P flux	V flux	V flux	V flux
true	true	true	polrotated	p -leveled
0.9660	0.1323	0.0000	-0.0054	-0.0002
1.0000	0.1200	0.0000	0.0051	0.0003
0.6071	0.0685	0.0000	-0.0029	-0.0002
0.9107	0.0608	0.0000	-0.0009	-0.0002
0.6221	0.0365	-0.0030	-0.0009	-0.0025
0.2008	0.0293	0.0000	0.0011	-0.0001
0.2718	0.0266	0.0000	-0.0012	-0.0001
0.6301	0.0196	0.0000	-0.0007	0.0000
0.2319	0.0178	0.0000	-0.0000	-0.0000
0.0436	0.0069	0.0000	0.0002	-0.0001
0.0402	0.0023	0.0000	-0.0000	-0.0000
0.0493	0.0007	0.0000	0.0000	0.0000
0.0027	0.0003	0.0000	0.0000	0.0000

morphologies – but probably much simpler than with quasi-scalar methods (e.g. Leppanen et al. 1995).

Poldeconversion should be applied as early as possible, lest CLEAN mistakenly consolidates some polconverted I flux in the Sky Model.

5.1.2. Polvector leveling

The linear polvectors for the true Sky reside in the QU plane (usually represented as “horizontal”). Polrotation may rotate them in this plane but also tilt the plane as a whole, giving rise to a V component for every p_{lin} source. We may “level” all polvectors in the Sky Model by rotating the V axis to the vertical axis, minimising $\sum V^2$ (Appendix B.2). If any significant V values remain, they are probably real. Table 4 demonstrates the effect.

Out of an infinite set of possible rotations I choose the one that requires a minimal rotation angle: it achieves the required leveling with a minimal effect on p_{lin}^9 .

As noted earlier, the polconversion and polrotation operators may be represented in either order. Consequently, the order of the corrections can also be freely chosen. In Table 3 poldeconversion comes first; after that, there is no polconverted p flux left that could be affected by polvector leveling. (If there were, it could be neutralised by another poldeconversion run.)

5.2. Comparing fitted with prior instrumental parameters

Self-alignment produces an Instrument model that includes the inverse of the poldistortion matrix X . One may estimate it by comparing the model with prior knowledge of the Instrument.

A typical antenna channel can be described (Paper I) by a product of Jones matrices

$$J_j = G_j C_j D_j R_j, \quad (9)$$

⁹ This does not entirely rule out the possibility that one flips the entire QUV space “upside down”. This possibility is in fact inherent in any method that relies on zeroing V , e.g. the XY phase-difference determination in the quasi-scalar calibration of linearly polarized arrays (Paper II). This seems never to have happened in practice, which shows that the usual pre-calibration is an effective prophylactic.

where

- G_j is a diagonal matrix, $\begin{pmatrix} g_p & 0 \\ 0 & g_q \end{pmatrix}$, where the g are complex gains and p and q represent the two receiver channels;
- C_j is a known matrix describing the nominal feed configuration;
- D_j , representing the feed errors, is close to the identity matrix I ;
- R_j is a rotation matrix of the form $Y_v(\phi)$ (Paper IV).

Prior knowledge consists of pre-calibrated values for the complex gains g , the feed parameters and the rotation ϕ (apart from Faraday rotation which is often not very well known).

For a comparison, we must extract corresponding values from our Instrument model J_j . The method for doing this is outlined in Appendix B.3. By correcting those differences that one finds, one may hope to further reduce the poldistortion left by the methods of Sect. 5.1.

Faraday polrotation remains as an error that can only be calibrated from external data. Note that antennas separated by large distances may experience different polrotations; these differences are absorbed by self-alignment, but one absolute polrotation must be determined otherwise in all cases.

6. Side issues

Along with the main experiment discussed above, a number of side issues were explored to check the robustness of my approach.

6.1. Stability

The self-align algorithm was tested with various pre-calibrations of the Instrument, such as the nominal values, the model from the previous cycle or an arbitrary matrix constant such as one of the Pauli matrices (Paper IV). In all cases the residuals in the fit were the same and the speeds of convergence differed only marginally. Obviously, the poldistortion introduced did depend on the initial value set. As discussed above, one should preferably start with values that are close to those expected: Use the pre-calibrated Instrument Model in the first, and the freshly updated one in all following self-align cycles.

6.2. Heterogeneous versus homogeneous arrays

A heterogeneous array was chosen for simulation because it represents an innovative concept proposed by Paper IV. Notwithstanding its advantages, there are sound practical reasons why such an array has hitherto not been attractive. Firstly, for non-polarimetric observations a heterogeneous array needs four correlator channels per interferometer where two would suffice. Secondly, if the primary beams of the antennas have non-circular symmetry, the combination of such beams in different orientations adds to the complexity of synthesis imaging in a formidable way.

Yet the heterogeneous array is of more than academic interest. Indeed, in the context of this paper Faraday rotation is part of the Instrument, and since it is inherently variable with distance it will make any array heterogeneous that extends over “large” distances. Under such conditions, every array must observe the full coherency even for unpolarized sources.

Faraday self-alignment

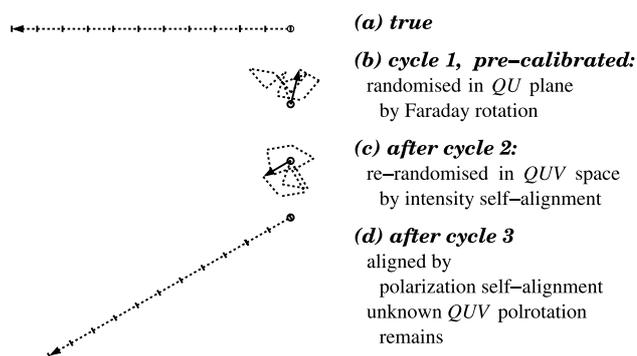


Fig. 5. Alignment of the linear polvectors observed at successive hour-angle intervals by self-alignment.

(a) Initially, the polvectors for successive hour angles are scattered in random directions by time-varying Faraday rotation. In the pre-calibrated image of cycle 1, their sum is much too small and badly misoriented. Most sources in the \mathbf{p} images are drowned in scattered radiation due to instrumental errors (Fig. 2). We satisfy ourselves with making a Sky Model for I only.

(b) In the self-alignment of cycle 2, that Sky Model serving as reference provides no clue for the alignment of the observed polvectors; they are scattered once more, in random directions. Their sum is changed but remains weak. Yet, since gain and phase errors have been much reduced, the strongest polarized sources are now clearly identifiable (Fig. 3).

(c) The \mathbf{p}_{lin} flux in the Sky Model is now strong enough to serve as reference for aligning the observed \mathbf{p}_{lin} fluxes at the end of cycle 3. The polrotation of this reference is “frozen into” the model (cf. Sect. 4.3.2). The polvectors for all hour angles are aligned to this reference with only very small errors left; $|\mathbf{p}_{\text{lin}}|$ fluxes get close to their correct values.

(d) Further cycles will reduce the remaining errors at a level that is not visible on the scale of this diagram.

The differences between a heterogeneous and a homogeneous array are amply discussed in Paper IV: The weak connection between the sets of X and Y (or L and R) channels makes the homogeneous system vulnerable to a spurious phase difference between the two, resulting in a polrotation in the Sky Model (cf. Paper II). For the linearly polarized array this is a $U \leftrightarrow V$ rotation that can be eliminated by polvector leveling. For a circularly polarized array that will not work, because the RL phase difference engenders a $Q \leftrightarrow U$ rotation on which we have no a-priori handle.

6.3. Faraday rotation

Faraday self-alignment is an important novelty. In quasi-scalar polarimetry, variable Faraday rotation can only be corrected by approximate estimates. The remaining variation mixes Q and U in such a way that effectively every polarized source is imaged with a different instrumental beam of poor quality. Matrix self-alignment suppresses this effect, makes the \mathbf{p} beams identical to the I beam and restores the images to their full dynamic range. Self-aligned Q and U images look even better than their I counterparts because there are fewer sources and hence less confusion of sidelobes; this can be seen in Fig. 3 (but the reproduction is probably not good enough to show it quite convincingly).

The mechanism of the Faraday alignment is sketched in Fig. 5.

7. Conclusion

I and \mathbf{p} fluxes in the final Sky Model agree (apart from a flux-scale factor) with the true Sky within .001 of the flux of the strongest source; having made my point, I did not try to improve this any further. I attach no great significance to this number. In the simulations, only the Instrument Model pretends to be realistic; other aspects of the observation were deliberately oversimplified in order to focus on the polarimetry issues.

The outcome of this study must be judged in relation to its purpose, which is validation of the theory of Paper IV and the development of supporting algorithms. On this count the results are entirely satisfactory. The poldistortion effects predicted have all been observed in (simulated) practice – sometimes as expected, sometimes as a surprise although they could have been expected. The latter cases in particular confirm the notion that poldistortion is an unavoidable artifact spawned by matrix self-alignment, in precisely the way that Paper IV foretells. On the practical side, a number of algorithms have been created that effectively deal with side problems which that paper hardly addresses.

Polarimetric errors in the self-aligned Sky image can be represented as a product of two poldistortion transformations. These are independent of each other and of spatial scattering by beam sidelobes, and both operate on all image points in the same way. The latter is a particularly valuable property, because it allows one to consider the statistics of the set of all (or a selected group of) Sky sources. By comparing them to the prior astrophysical knowledge that both linear and circular polarizations are zero on the average, one may estimate polconversion and polrotation and correct for them, without reference to external measurements such as a separate observation of a calibrator source. In this way, the great benefits of scalar intensity self-calibration get extended to polarimetry¹⁰. As a complement, statistical comparison of the Instrument model to prior estimates of Instrument parameters gives us another handle on the poldistortion. The latter idea is, of course, not new but is given a broader scope here.

The (pol)rotation of parallactic angle in azimuth/elevation-mounted antennas has often been suggested as a mechanism to separate instrumental effects from the observed Sky. Cotton (1993) and Roberts et al. (1994) have investigated this option at length. They show that it can be made to work to some extent, under suitable conditions, and with considerable manual labour. Matrix self-alignment and post-calibration replace such artisan methods by separating time-varying polrotation from the Sky in a clean, systematic and generally applicable way, using algorithms that require little more manual steering than simple scalar selfcal does.

The virtues of Faraday self-alignment have already been highlighted in Sect. 6.3. They are of particular importance for polarimetry toward lower frequencies since the rotation and its temporal variation increase with λ^2 . At long baselines, each antenna moreover experiences its own Faraday rotation, effectively making an array heterogeneous even if it had been built homogeneous. This possibility has not been explicitly considered here, but it is no more than a trivial departure from the static heterogeneity that has been simulated.

Notwithstanding the complications of practical aperture synthesis, scalar self-calibration has proven to be very effective with observing geometries and source morphologies much more complicated than those simulated here. With the ground rules

¹⁰ In the past, experienced polarimetrists have occasionally derived crude “manual” Faraday alignment and polvector-leveling corrections by visual inspection of their quasi-scalarly self-calibrated Sky images (A. G. de Bruyn, private comm.).

for polarimetric selfcal now firmly established, and given the close analogies between the scalar and matrix domains, one may expect matrix selfcal likewise to be applicable in real-life situations.

Polarimetry is a nasty complication to the subject of Aperture Synthesis, which is already far from simple in itself. Although many observers are using them to obtain their astrophysical results, quasi-scalar methods have certainly not helped to endear the subject to them. To some, the Jones-matrix representation appears like a way of making things even more difficult. This notion is mistaken. Rather than complicating the matter, matrix methods simplify it, if only one is willing to accept matrix algebra as a tool. (A simple but characteristic measure of this simplification is the compactness of the formulas in the papers of this series as compared to that in the typical paper on quasi-scalar polarimetry.)

The matrix formulation provides a precise description of the physics of our receiver systems. Its awkward features reflect properties that are real but overlooked by the quasi-scalar method; for example, the non-commutativity of matrix multiplication reminds us that, in a dual-polarization signal path, the order of ionosphere, antenna, feed and receiver stages cannot be arbitrarily changed in the way it can for a scalar chain. Such complexity is the inevitable consequence of very basic physics. The quasi-scalar method approaches it from the bottom up by splitting the problem into isolated morsels, treating them in mutual isolation, and ignoring the higher-order effects that do not fit. Matrix self-calibration proposes a top-down alternative that has no need for approximations: Each of the steps described above is motivated by the overall view that the Measurement Equation provides, and fits seamlessly into one coherent grand picture.

Further development in observing practice will depend on the creation of suitable software. Having endorsed the Measurement Equation as one of its foundations, *AIPS++* would seem to be the foremost place to look for that, and it indeed already includes many of the matrix software mechanisms required. However, its present implementation of matrix selfcal was created too early to benefit from the concepts that

Papers IV and V have now firmly put in place. As a first step forward, the development group is engaged in a code cleanup that will pave the way for self-calibration upgrades in the future (G. Moellenbrock, priv. comm.). Another development is to be looked for in the Netherlands, where *AIPS++* infrastructure is built into the new software for the *LOFAR* phased-array telescope. In order to deal with the unprecedented complexity of that instrument, full deployment of matrix-based data reduction methods is foreseen.

With this paper, the program of the series, *Understanding radio polarimetry*, has been successfully completed. Further development of polarimetric theory and practice will most likely take place in the context of applications.

Acknowledgements. I am much indebted to the management of ASTRON for the opportunity to conduct this long-term research project in an environment where the pressure from a host of shorter-term projects and immediate applications is continually felt. W.N. Brouw, R.J. Nijboer and J. Tinbergen kindly read the manuscript and provided useful suggestions. ASTRON is operated with financial support from the Netherlands Organisation for Scientific Research (NWO).

References

- Cotton, W. 1993, *AJ*, 106, 1241
 Fender, R., & Macquart, J.-P. 2003h Circular Polarisation from Relativistic Jet Sources, Kluwer, Reprinted from *Astrophysics and Space Science*, 288, Nos. 1–2
 Hamaker, J. P. 2000, *A&AS*, 143, 515, Also available at www.astron.nl/~hamaker
 Hamaker, J. P., & Bregman, J. D. 1996, *A&AS*, 117, 161, Also available at www.astron.nl/~hamaker
 Hamaker, J. P., Bregman, J. D., & Sault, R. J. 1996, *A&AS*, 117, 137, Also available at www.astron.nl/~hamaker
 Korn, G., & Korn, T. 1968, *Mathematical Handbook for Scientists and Engineers* (McGraw-Hill)
 Lancaster, P., & Tismenetsky, M. 1985, *The Theory of Matrices* (Academic Press)
 Leppanen, K., Zensus, J., & Diamond, P. 1995, *ApJ*, 110, 2479
 Roberts, D., Wardle, J., & Brown, L. 1994, *ApJ*, 427, 718
 Sault, R. J., Hamaker, J. P., & Bregman, J. D. 1996, *A&AS*, 117, 149, Also available at www.astron.nl/~hamaker

Online Material

Appendix A: Self-alignment algorithms

Most of the algorithms sketched below rely on the *quaternion* representation of matrices and the quaternion algebra that are introduced in Paper IV. In brief, Eq. (4) is generalised to all 2×2 matrices:

$$\mathbf{M} = [m + \mathbf{m}], \quad (\text{A.1})$$

where m is a scalar and \mathbf{m} a vector in an abstract 3-dimensional space that, in our application, coincides with QUV space. In the Appendix to Paper IV, the rules for algebraic manipulation of quaternions are derived and applied to describe polrotation and polconversion as geometric operations in this space. I shall freely use these results, mostly without further reference to their origin.

A.1. Instrument model fit

The algorithm of Paper IV, App. D.2, can be presented in a more transparent form. Given a set of observed coherencies

$$\mathbf{W}_{jk} = \mathbf{J}_j \mathbf{E}_{jk} \mathbf{J}_k^\dagger \quad (\text{A.2})$$

for a Sky Model $\{\mathbf{E}_{jk}\}$, we seek to fit a set of inverse Jones matrices $\{\mathbf{K}_j\}$ that minimises the noise power

$$S = \sum_{jk} \|\mathbf{K}_j \mathbf{W}_{jk} \mathbf{K}_k^\dagger - \mathbf{E}_{jk}\| \quad (\text{A.3})$$

at the interferometer inputs. We do this by a relaxation method in which we successively adjust the \mathbf{K}_j toward convergence:

$$\mathbf{K}'_j = \sum_k \mathbf{E}_{jk} \mathbf{K}_k \mathbf{W}_{jk}^\dagger \left(\sum_k \mathbf{W}_{jk} \mathbf{K}_k^\dagger \mathbf{K}_k \mathbf{W}_{jk}^\dagger \right)^{-1}. \quad (\text{A.4})$$

A.2. Polvector alignment

Consider two arbitrary unit polvectors \mathbf{p} and \mathbf{r} which make an angle α . Rotating one into the other can be achieved through many different rotations (all of which have their axis in the bisecting plane of \mathbf{p} and \mathbf{r}). For the axis perpendicular to both vectors the rotation angle is equal to α and minimal, which seems to be a desirable property. To find this *minimal rotation*, we take the quaternion product

$$[\mathbf{p}][\mathbf{r}] = [\mathbf{p} \cdot \mathbf{r} + i \mathbf{p} \times \mathbf{r}] = [\cos \alpha + i \mathbf{1}_{\mathbf{p} \times \mathbf{r}} \sin \alpha]. \quad (\text{A.5})$$

This is a unimodular unitary quaternion that effects a polrotation over an angle 2α with its axis along $\mathbf{p} \times \mathbf{r}$. To implement the required rotation over α we must take the square root of $[\mathbf{p}][\mathbf{r}]$. (Just as for positive hermitian matrices, a closed form can be derived for the square root of a unitary matrix.)

In the Instrument-fit procedure of Sect. 4.2, \mathbf{p} and \mathbf{r} above represent the polvectors of the Sky-Model and observed coherencies, respectively. We sum their product over all uv points to estimate the required rotation.

Appendix B: Post-calibration algorithms

If the observed field contains enough source flux that we may assume a priori to be unpolarized, we may use these sources as a reference against which polconversion can be estimated and then removed. Similarly, if there is enough linearly polarized flux we

can use it to estimate how much of it has been polrotated into circular polarization and correct for that.

We may also compare the fitted Jones matrices \mathbf{J}_{jt} to prior knowledge of the instrument. To this end we must factor them to expose the parameters by which an antenna channel is normally characterised: Rotations (Faraday, parallactic, feed), feed configuration and errors, complex receiver gains.

B.1. Poldeconversion

Polconversion is characterised by the matrix \mathbf{H} in Eq. (7). For once I revert to the quasi-scalar approximation in which only one of its effects is considered, viz. the conversion of some I flux into a polvector flux for every source k ¹¹:

$$\Delta \mathbf{p}_k = \mathbf{p}_0 I_k. \quad (\text{B.1})$$

Since the $\Delta \mathbf{p}$ are small, the best place to look for them is in the residual \mathbf{p} image after CLEANing. To measure it, we correlate this image with the dirty intensity image, using the intensity as a weight factor

$$\mathbf{p}'_0 = \frac{\sum [\mathbf{p} I]}{\sum I^2} \quad (\text{B.2})$$

where the prime indicates an estimate. To avoid contamination by sources for which CLEAN found significant \mathbf{p} components, these should first be subtracted from the I and \mathbf{p} images.

An example from the simulation is shown in Fig. 4.

B.2. Polvector leveling

Consider the set of N real column vectors $\{\mathbf{p}_k\}$ representing the polarized Sky Model, and a plane through the origin of QUV space whose normal is the column vector \mathbf{a} . The distance of a point \mathbf{p}_k to the plane is $d_k = \mathbf{a}^\text{T} \mathbf{p}_k = \mathbf{p}_k^\text{T} \mathbf{a}$. We require the value of \mathbf{a} that minimises

$$S = \sum d_k^2 = \sum_j \mathbf{a}^\text{T} \mathbf{p}_j \sum_k \mathbf{p}_k^\text{T} \mathbf{a} = \mathbf{a}^\text{T} \mathbf{P}^\text{T} \mathbf{P} \mathbf{a} \quad (\text{B.3})$$

where \mathbf{P} is an $N \times 3$ matrix whose columns are the real vectors \mathbf{p}_k . With the accessory condition $\mathbf{a}^\text{T} \mathbf{a} = 1$ this results in an eigenvalue problem

$$\mathbf{P}^\text{T} \mathbf{P} \mathbf{a} - \lambda \mathbf{a} = 0 \quad (\text{B.4})$$

$\mathbf{P}^\text{T} \mathbf{P}$ is a positive symmetric matrix and Eq. (B.3) a *quadratic form* (Lancaster & Tismenetsky 1985). The three eigenvectors of $\mathbf{P}^\text{T} \mathbf{P}$ are the normals to the *principal planes* of the set $\{\mathbf{p}_k\}$. The corresponding positive eigenvalues equal the sums Eq. (B.3); the eigenvector corresponding to the smallest of these is the normal sought for. To rotate this vector to the V axis, we can use the minimal-rotation algorithm of Sect. A.2

B.3. Reducing a Jones matrix to hardware factors

I show here the factoring procedure for a linearly polarized antenna where, in Eq. (9), $\mathbf{C} = \mathbf{I}$. For a circularly polarized one

¹¹ This use of the approximation is quite peripheral; with proper pre-calibration, the errors that it might introduce are negligible. Moreover, they affect the entire Sky rather than each source on its own, and can therefore be corrected later.

the analysis is only slightly more complicated. (The method shown here is simpler and more direct than the one suggested by Paper IV, and allows for the diagonal elements of \mathbf{D} to differ.)

The equation to be satisfied for given \mathbf{J} is Eq. (9)

$$\mathbf{J} = \mathbf{GDR} \quad (\text{B.5})$$

where \mathbf{G} is diagonal, \mathbf{R} is a $Q \leftrightarrow U$ rotation matrix, the nominal configuration matrix equals \mathbf{I} and \mathbf{D} , representing the feed errors, must be close to \mathbf{I} .

To find the factors, perform the following steps:

1. Apply the polar decomposition (Paper IV, App. B.6)

$$\mathbf{J} = |x|^2 \mathbf{H}\mathbf{Y} \quad (\text{B.6})$$

2. Factor the polrotation matrix \mathbf{Y} per Appendix B.4 into a product of three mutually orthogonal rotations (Paper IV, Eq. (27))

$$\mathbf{Y} = \mathbf{Y}_q(\phi) \mathbf{Y}_u(\epsilon) \mathbf{Y}_v(\theta) \quad (\text{B.7})$$

where the factors \mathbf{Y}_q , \mathbf{Y}_u and \mathbf{Y}_v represent polrotations with axes in the Q , U and V directions.

3. Exchange $\mathbf{Y}_q(\phi)$ with \mathbf{H} :

$$\begin{aligned} \mathbf{J} &= |x|^2 \mathbf{H}\mathbf{Y}_q(\phi) \mathbf{Y}_u(\epsilon) \mathbf{Y}_v(\theta) \\ &= |x|^2 \mathbf{Y}_q(\phi) \mathbf{H}' \mathbf{Y}_u(\epsilon) \mathbf{Y}_v(\theta) \end{aligned} \quad (\text{B.8})$$

where

$$\mathbf{H}' = \mathbf{Y}_q(-\phi) \mathbf{H} \mathbf{Y}_q(\phi). \quad (\text{B.9})$$

4. Define a unimodular matrix

$$\mathbf{M} = \begin{pmatrix} \sqrt{h_{22}/h_{11}} & 0 \\ 0 & \sqrt{h_{11}/h_{22}} \end{pmatrix} \quad (\text{B.10})$$

to obtain a unimodular matrix \mathbf{H}'' with equal diagonal elements $\sqrt{h_{11}h_{22}}$

$$\mathbf{H}'' = \mathbf{M}^{-1} \mathbf{H}'. \quad (\text{B.11})$$

5. Finally define the factors in Eq. (B.5) as

$$\begin{aligned} \mathbf{G} &= |x|^2 \mathbf{Y}_q(\phi) \mathbf{M}; \\ \mathbf{D} &= \mathbf{H}'' \mathbf{Y}_u(\epsilon); \\ \mathbf{R} &= \mathbf{Y}_v(\theta). \end{aligned} \quad (\text{B.12})$$

The factors \mathbf{G} and \mathbf{R} evidently satisfy the requirements. \mathbf{D} represents those effects that could not be accounted for by the other two factors and which we must therefore attribute to the feed. Some effects in the real feed are misinterpreted: A difference between the losses in the x and y receptors is absorbed in \mathbf{G} and a misorientation is attributed to \mathbf{R} . Clearly, \mathbf{J} does not contain enough information to distinguish them (just as in scalar self-cal one cannot distinguish various contributions to the overall phase), unless we have more precise prior information, e.g. about the feed.

B.4. Rotation decomposition

Although the problem of representing an arbitrary three-dimensional rotation in terms of mutually perpendicular ones is well known, I had trouble finding direct references. The basic information is available in handbooks (e.g. Korn & Korn 1968) but some tedious algebra is needed to extract the necessary formulae.

The product Eq. (B.7) can be written out in quaternion form as

$$\begin{aligned} \mathbf{Y} &= [y_0 + i\mathbf{y}], \\ y_0 &= \cos \phi \cos \epsilon \cos \theta + \sin \phi \sin \epsilon \sin \theta \\ y_1 &= \sin \phi \cos \epsilon \cos \theta - \cos \phi \sin \epsilon \sin \theta \\ y_2 &= \cos \phi \sin \epsilon \cos \theta + \sin \phi \cos \epsilon \sin \theta \\ y_3 &= \cos \phi \cos \epsilon \sin \theta - \sin \phi \sin \epsilon \cos \theta \end{aligned} \quad (\text{B.13})$$

It is readily seen that $y_0 y_2 - y_1 y_3 = \frac{1}{2} \sin 2\epsilon$. Further note that

$$\begin{aligned} \begin{pmatrix} y_0 \\ y_2 \end{pmatrix} &= \begin{pmatrix} \cos \epsilon & \sin \epsilon \\ \sin \epsilon & \cos \epsilon \end{pmatrix} \begin{pmatrix} \cos \phi \cos \theta \\ \sin \phi \sin \theta \end{pmatrix} \\ \begin{pmatrix} y_1 \\ y_3 \end{pmatrix} &= \begin{pmatrix} \cos \epsilon & -\sin \epsilon \\ -\sin \epsilon & \cos \epsilon \end{pmatrix} \begin{pmatrix} \sin \phi \cos \theta \\ \cos \phi \sin \theta \end{pmatrix} \end{aligned} \quad (\text{B.14})$$

from which $\phi \pm \theta$ follow.