The MEASUREMENT EQUATION
of a generic radio telescope
AIPS++ Implementation Note nr 185

J.E.Noordam
(jnoordam@nfra.nl)

15 February 1996, version 2.0

File: /aips++/nfra/185.latex Symbols File: /aips++/nfra/megi-symbols.tex

Abstract: This note is a step towards an 'official' AIPS++ description of the Measurement Equation, based on an agreed set of names and conventions. The latter have been defined in a separate TeX file, and can (should) be used in subsequent AIPS++ documents to ensure consistency.

Contents

1 INTRODUCTION

The matrix-based Measurement Equation (ME) of a Generic Radio Telescope was developed by Hamaker, Bregman and Sault [2] [3], based on earlier work by Bregman [1]. After discussion by Noordam [5] and Cornwell [6] [7] [8] [9] [10] [11], the M.E. has been adopted as the generic foundation of the uv-data calibration and imaging part of AIPS++. In the not too distant future, an 'official' AIPS++ description of the ME will be needed, with agreed conventions and nomenclature (see Appendix A). This note is a step towards that goal.

The heart of the M.E. is formed by the 2×2 *feed*-based 'Jones' matrices, which describe the effects of various parts of the observing instrument on the signal. The main section of this document is devoted to describing the basic form of the Jones matrices in linear and circular polarisation coordinates. Another section discusses the conditions under which their *order* may be modified (matrices do not always commute).

It is expected that the details of the M.E. (and of this note) will be refined during the first few iterations of design and implementation of AIPS++. But the *structure* of the M.E. formalism as presented here appears to be rich enough to accomodate all existing and planned radio *telescope*s. This includes 'exotic' ones like cylindrical mirrors, phased arrays, and interferometer arrays with very dissimilar antennas. Further refinements should only require the addition of new Jones matrices, or devising new expressions for existing matrix elements.

In order to test this bold assertion, the various institutes might endeavour to model their own *telescope*s in terms of the precise and common language of the M.E., using this note as a reference. The following 'rules' are probably good ones:

- In modelling an instrument, stay as close to the actual physical situation as possible. Violations of this principle, for whatever reasons, will lead to problems sooner or later.
- It is counterproductive to try and simplify the M.E. to make it 'look more tractable'. This practice introduces hidden assumptions, which tend to be forgotten by the programmer, and unknown to the user.
- Use the suggested nomenclature and conventions.

It is also good to realise that there are two basic forms of ME, which should not be confused: In the *physical* form, each instrumental effect is modelled separately by its own matrix. This is useful for simulation purposes. In the *mathematical* form, effects are 'lumped together' if they cannot be solved for separately. Example: the various contributions to the receiver gain, and tropospheric gain.

Acknowledgements: The author has greatly benefited from detailed discussions with Jayaram Chengalur, Jaap Bregman, Johan Hamaker, Tim Cornwell, Wim Brouw and Mark Wieringa.

2 THE M.E. FOR A SINGLE POINT SOURCE

For the moment, it will be assumed that there is a single point source at an arbitrary position (direction) $\vec{\rho} = \vec{\rho}(\mathsf{l},\mathsf{m})$ w.r.t. the fringe-tracking centre, and that observing bandwidth and integration time are negligible. Multiple and extended sources, and the effects of non-zero bandwidth and integration time will be treated for the Full Measurement Equation in section 3.

For a given *interferometer*, the measured visibilities can be written as a 4-element 'coherency vector' $\vec{V}_{\mathsf{ij}}$, which is related to the so-called 'Stokes vector' $\vec{I}(\mathsf{l},\mathsf{m})$ of the observed source by a matrix equation,

$$\vec{V}_{\mathsf{ij}} \;=\; \begin{pmatrix} \mathsf{v}_{\mathsf{ip\,jp}} \\ \mathsf{v}_{\mathsf{ip\,jq}} \\ \mathsf{v}_{\mathsf{iq\,jp}} \\ \mathsf{v}_{\mathsf{iq\,jq}} \end{pmatrix} \;=\; (\mathsf{J}_{\mathsf{i}} \otimes \mathsf{J}_{\mathsf{j}}^{*})\;\mathsf{S}\; \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix}_{\mathsf{l,m}} \tag{1}$$

The subscripts i and j are the labels of the two *feed*s that make up the *interferometer*. The subscripts p and q are the labels of the two output *IF-channel*s from each *feed*.[1]

The 'Stokes matrix' S is a constant 4×4 coordinate transformation matrix. It is discussed in detail in section 4 below. The real heart of the M.E. is the 'direct matrix product' $\mathsf{J}_{\mathsf{i}} \otimes \mathsf{J}_{\mathsf{j}}^{*}$ of two 2×2 *feed*-based Jones matrices.

The 'Stokes-to-Stokes' transmission of a Stokes vector through an 'optical' element may be described by multiplication with a 4×4 *Mueller matrix* $\mathcal{M}_{\mathsf{ij}}$ [2] [3]. Using equation 1:

$$\vec{I}^{out}(\mathsf{l},\mathsf{m}) \;=\; \mathsf{S}^{-1}\,\vec{V}_{\mathsf{ij}} \;=\; \mathsf{S}^{-1}\,(\mathsf{J}_{\mathsf{i}} \otimes \mathsf{J}_{\mathsf{j}}^{*})\;\mathsf{S}\;\vec{I}^{in}(\mathsf{l},\mathsf{m}) \;=\; \mathcal{M}_{\mathsf{ij}}(\mathsf{l},\mathsf{m})\;\vec{I}^{in}(\mathsf{l},\mathsf{m}) \tag{2}$$

Mueller matrices are useful in simulation, when studying the effect of instrumental effects on a test source $\vec{I}(\mathsf{l},\mathsf{m})$. They can be easily generalised to the full M.E. (see section 3).

2.1 The *feed*-based instrumental Jones matrices

It will be assumed (for the moment) that all instrumental effects can be factored into *feed*-based contributions, i.e. any *interferometer*-based effects are assumed to be negligible (see section 3). The 4×4 *interferometer* response matrix J_{ij} then consists of a 'direct matrix product'[2] $\mathsf{J}_{\mathsf{i}} \otimes \mathsf{J}_{\mathsf{j}}^{*}$ of two 2×2 *feed*-based response matrices, called 'Jones matrices'. The reader will note that this factoring is the polarimetric generalisation of the familiar 'Selfcal assumption', in which the (scalar) gains are assumed to be *feed*-based rather than *interferometer*-based.

The 2×2 Jones matrix J_{i} for *feed* i can be decomposed into a product of several 2×2 Jones matrices, each of which models a specific *feed*-based instrumental effect in the signal path:

[1] The generic *IF-channel* labels p and q are known as X and Y for WSRT and ATCA, and R and L for the VLA. They should *not* be confused with the two *receptor*s a and b, since the signal in an *IF-channel* may be a linear combination of the *receptor* signals.

[2] Also called the outer matrix product, or tensor product, or Kronecker product. See [2].

$$\mathsf{J_i} \;=\; \mathsf{G_i}\;[\mathsf{H_i}]\;[\mathsf{Y_i}]\;\mathsf{B_i}\;\mathsf{K_i}\;\mathsf{T_i}\;\mathsf{F_i} \;=\; \mathsf{G_i}\;[\mathsf{H_i}]\;[\mathsf{Y_i}]\;(\mathsf{D_i}\;\mathsf{E_i}\;\mathsf{P_i})\;\mathsf{K_i}\;\mathsf{T_i}\;\mathsf{F_i} \tag{3}$$

in which

$\mathsf{F_i}(\vec{\rho}, \vec{r_i})$	ionospheric Faraday rotation
$\mathsf{T_i}(\vec{\rho}, \vec{r_i})$	atmospheric complex gain
$\mathsf{K_i}(\vec{\rho}.\vec{r_i})$	factored Fourier Transform kernel
$\mathsf{P_i}$	projected *receptor* orientation(s) w.r.t. the sky
$\mathsf{E_i}(\vec{\rho})$	voltage primary beam
$\mathsf{D_i}$	position-independent *receptor* cross-leakage
$[\mathsf{Y_i}]$	commutation of *IF-channel*s
$[\mathsf{H_i}]$	hybrid (conversion to circular polarisation coordinates)
$\mathsf{G_i}$	electronic complex gain (*feed*-based contributions only)

Matrices between brackets ([ ]) are not present in all systems. $\mathsf{B_i}$ is the 'Total Voltage Pattern' of an arbitrary *feed*, which is usually split up into three sub-matrices: $\mathsf{D_i}\;\mathsf{E_i}\;\mathsf{P_i}$. Jones matrices that model 'image-plane' effects depend on the source position (direction) $\vec{\rho}$. Some also depend on the antenna position $\vec{r_i}$. Of course most of them depend on time and frequency as well. The various Jones matrices are treated in some detail in section 5.

Since the Jones matrices do not always commute with each other, their order is important. In principle, they should be placed in the 'physical' order, i.e. the order in which the signal is affected by them while traversing the instrument. In practice, this is not always possible or desirable. Section 6 discusses the implications of choosing a different order.

2.2 The Jones matrix of a Tied Array *feed*

The output signals from the two *IF-channel*s of a 'tied array' is the weighted sum of the *IF-channel* signals from n individual *feed*s. A tied array is itself a *feed* (see definition in appendix A), modelled by its own Jones matrix. For a single point source, we get:

$$\mathsf{J_i}^{tied\ array} = \mathsf{Q_i} \sum_n w_{\mathsf{i}n}\;\mathsf{J}_{\mathsf{i}n} \tag{4}$$

and for an *interferometer* between two tied arrays i and j with n and m constituent *feed*s respectively:

$$\mathsf{J_{ij}} = (\mathsf{J_i} \otimes \mathsf{J_j}^*) \;=\; (\mathsf{Q_i} \otimes \mathsf{Q_j}^*) \sum_n \sum_m w_{\mathsf{i}n}\;w_{\mathsf{j}m}\;(\mathsf{J}_{\mathsf{i}n} \otimes \mathsf{J}_{\mathsf{j}m}{}^*) \tag{5}$$

See also section 6.4. The matrix $\mathsf{Q_i}$ models electronic gain effects on the *added* signal of the tied array *feed* i. The $\mathsf{Q_i}$ can be solved by the usual Selfcal methods, in contrast to instrumental errors in the constituent *feeds* *before* adding. The latter will often cause decorrellation, and thus closure errors in an *interferometer*.

Since a tied array *feed* can be modelled by a Jones matrix, it can be combined with *any* other type of *feed* to form an *interferometer*. Examples are the use of WSRT and VLA as tied arrays in VLBI arrays. Note that this is made possible by factoring the Fourier Transform kernel $\mathsf{K}_{\mathsf{ij}}(\vec{u}_{\mathsf{ij}}.\vec{\rho})$ into $\mathsf{K}_{\mathsf{i}}(\vec{r}_{\mathsf{i}}.\vec{\rho})$ and $\mathsf{K}_{\mathsf{j}}(\vec{r}_{\mathsf{j}}.\vec{\rho})$, and including the latter in the Jones matrices of the individual *feeds* (see equ 28).

Obviously, the primary beam of a tied array can be rather complicated, but it is fully modelled by equ 4. Moreover, the contributing *feeds* in a tied array are allowed to be quite dissimilar. It is nor even necessary for their *receptors* (dipoles) to be aligned with each other! Thus, equation 4 can also be used to model 'difficult' *telescopes* like Ooty or MOST, or an element of the future Square Km Array (SKAI). This puts the crown on the remarkable power of the Measurement Equation.

2.3 Jones matrices for multiple beams

Using the definition in appendix A, each beam in a multiple beam system should be treated like a separate logical *feed*, modelled by its own Jones matrix. Any communality between them can be modelled in the form of shared parameters in the expressions for the various matrix elements.

3 THE FULL MEASUREMENT EQUATION

3.1 Summing and averaging

For k 'real' incoherent sources, observed with a 'real' *telescope*, equ 1 becomes:

$$\vec{V}_{\mathsf{ij}} = \frac{1}{\Delta\mathsf{t}\Delta\mathsf{f}} \int d\mathsf{t} \int d\mathsf{f} \sum_k \frac{1}{\Delta\mathsf{l}\Delta\mathsf{m}} \int d\mathsf{l}\, d\mathsf{m}\ \mathsf{J}_{\mathsf{i}} \otimes \mathsf{J}_{\mathsf{j}}^{*}\ \mathsf{S}\ \vec{I}(\mathsf{l},\mathsf{m}) \tag{6}$$

The visibility vector $\vec{V}_{\mathsf{ij}}$ is integrated over the extent of the sources ($\int d\mathsf{l}\, d\mathsf{m}$), over the integration time ($\int d\mathsf{t}$) and over the channel bandwidth ($\int d\mathsf{f}$). Integration over the aperture ($\int d\mathsf{u}\, d\mathsf{v}$) is taken care of by the primary beam properties.

There are only four integration coordinates, whose units are determined by the flux density units in which $\vec{I}$ is expressed: $energy/sec/Hz/beam$. These coordinates define a 4-dimensional 'integration cell'. If the variation of $\vec{V}(\mathsf{f},\mathsf{t},\mathsf{l},\mathsf{m})$ is *linear* over this cell, integration is not necessary:

$$\vec{V}_{\mathsf{ij}} = \sum_k \vec{V}_{0k}(\mathsf{f}_0,\mathsf{t}_0,\mathsf{l}_0,\mathsf{m}_0) \tag{7}$$

in which $\vec{V}_{0k}$ is the value for source k at the centre of the cell, for $\Delta\mathsf{f} = 1\ Hz$ and $\Delta\mathsf{t} = 1\ sec$. If the variation of $\vec{V}(\mathsf{f},\mathsf{t},\mathsf{l},\mathsf{m})$ over the cell can be approximated by a polynomial of order ≤ 3, then it is sufficient to calculate only the 2nd derivative(s) at the centre of the cell:

$$\vec{V}^{int} \;=\; \sum_k \vec{V}_{0k} \;+\; \frac{1}{12}\,(\frac{\partial^2 \vec{V}_{0k}}{\partial \mathsf{f}^2}\,(\Delta \mathsf{f})^2 \;+\; \frac{\partial^2 \vec{V}_{0k}}{\partial \mathsf{t}^2}\,(\Delta \mathsf{t})^2 \;+\; \frac{\partial^2 \vec{V}_{0k}}{\partial \mathsf{l}^2}\,(\Delta \mathsf{l})^2 \;+\; \frac{\partial^2 \vec{V}_{0k}}{\partial \mathsf{m}^2}\,(\Delta \mathsf{m})^2) \quad (8)$$

Here it is assumed that the 2nd derivatives are be constant over the cell, i.e. the cross-derivatives $\frac{\partial \vec{V}_0}{\partial p_1 \partial p_2}$ are zero.

3.2 *interferometer*-based effects

Until now, we have assumed that all instrumental effects could be factored into *feed*-based contributions, i.e. we have ignored any *interferometer*-based effects. This is justified for a well-designed system, provided that the signal-to-noise ratio is large enough (thermal noise causes *interferometer*-based errors, albeit with a an average of zero). However, if systematic errors *do* occur, they can be modelled:

$$\vec{V}'_{\mathsf{ij}} \;=\; \mathsf{X}_{\mathsf{ij}}\;(\vec{A}_{\mathsf{ij}} \;+\; \mathsf{M}_{\mathsf{ij}}\;\vec{V}_{\mathsf{ij}}) \quad (9)$$

The 4×4 diagonal matrix X, the 'Correlator matrix', represents *interferometer*-based corrections that are applied to the uv-data *in software* by the on-line system. Examples are the Van Vleck correction. In the newest correlators, it approaches a constant (x).

$$\mathsf{X}_{\mathsf{ij}} \;=\; \begin{pmatrix} \mathsf{x}_{\mathsf{ip\,jp}} & 0 & 0 & 0 \\ 0 & \mathsf{x}_{\mathsf{ip\,jq}} & 0 & 0 \\ 0 & 0 & \mathsf{x}_{\mathsf{iq\,jp}} & 0 \\ 0 & 0 & 0 & \mathsf{x}_{\mathsf{iq\,jq}} \end{pmatrix} \;\approx\; \mathsf{x}\,\mathcal{U} \quad (10)$$

The 4×4 diagonal matrix M represents *multiplicative interferometer*-based effects.

$$\mathsf{M}_{\mathsf{ij}} \;=\; \begin{pmatrix} \mathsf{m}_{\mathsf{ip\,jp}} & 0 & 0 & 0 \\ 0 & \mathsf{m}_{\mathsf{ip\,jq}} & 0 & 0 \\ 0 & 0 & \mathsf{m}_{\mathsf{iq\,jp}} & 0 \\ 0 & 0 & 0 & \mathsf{m}_{\mathsf{iq\,jq}} \end{pmatrix} \;\approx\; \mathcal{U} \quad (11)$$

The 4-element vector $\vec{A}_{\mathsf{ij}}$ represents *additive interferometer*-based effects. Examples are receiver noise, and correlator offsets.

$$\vec{A}_{\mathsf{ij}} \;=\; \begin{pmatrix} \mathsf{a}_{\mathsf{ip\,jp}} \\ \mathsf{a}_{\mathsf{ip\,jq}} \\ \mathsf{a}_{\mathsf{iq\,jp}} \\ \mathsf{a}_{\mathsf{iq\,jq}} \end{pmatrix} \;\approx\; \vec{0} \quad (12)$$

In some cases, *interferometer*-based effects can be calibrated, e.g. when they appear to be constant in time. It will be interesting to see how many of them will disappear as a result of better modelling with the Measurement Equation. In any case, it is desirable that the cause of *interferometer*-based effects is properly understood (simulation!).

4 POLARISATION COORDINATES

In the 2×2 **signal domain**, the electric field vector $\vec{E}$ of the incident plane wave can be represented either in a linear polarisation coordinate frame (x, y) or a circular polarisation coordinate frame (r, l). Jones matrices are linear operators in the chosen frame:

$$\vec{V}_{\mathsf{i}}^{+} \;=\; \begin{pmatrix} \mathsf{v}_{\mathsf{ip}} \\ \mathsf{v}_{\mathsf{iq}} \end{pmatrix} \;=\; \mathsf{J}_{\mathsf{i}}^{+} \begin{pmatrix} \mathsf{e}_{\mathsf{x}} \\ \mathsf{e}_{\mathsf{y}} \end{pmatrix} \qquad or \qquad \vec{V}_{\mathsf{i}}^{\odot} \;=\; \mathsf{J}_{\mathsf{i}}^{\odot} \begin{pmatrix} \mathsf{e}_{\mathsf{r}} \\ \mathsf{e}_{\mathsf{l}} \end{pmatrix} \tag{13}$$

For linear polarisation coordinates, equation 1 becomes:

$$\vec{V}_{\mathsf{ij}}^{+} \;=\; (\mathsf{J}_{\mathsf{i}}^{+} \otimes \mathsf{J}_{\mathsf{j}}^{+*})\;(\vec{E} \otimes \vec{E}^{*}) \;=\; (\mathsf{J}_{\mathsf{i}}^{+} \otimes \mathsf{J}_{\mathsf{j}}^{+*}) \begin{pmatrix} \mathsf{e}_{\mathsf{x}}\mathsf{e}_{\mathsf{x}}^{*} \\ \mathsf{e}_{\mathsf{x}}\mathsf{e}_{\mathsf{y}}^{*} \\ \mathsf{e}_{\mathsf{y}}\mathsf{e}_{\mathsf{x}}^{*} \\ \mathsf{e}_{\mathsf{y}}\mathsf{e}_{\mathsf{y}}^{*} \end{pmatrix} \;=\; (\mathsf{J}_{\mathsf{i}}^{+} \otimes \mathsf{J}_{\mathsf{j}}^{+*})\;\mathsf{S}^{+}\;\vec{I}(\mathsf{l}, \mathsf{m}) \tag{14}$$

and there is a similar expression for circular polarisation coordinates. Thus, as emphasised in [2], the Stokes vector $\vec{I}(\mathsf{l}, \mathsf{m})$ and the coherency vector $\vec{V}_{\mathsf{ij}}$ represent the same physical quantity, but in different abstract coordinate frames. A 'Stokes matrix' S is a coordinate transformation matrix in the 4×4 **coherency domain**: S^{+} transforms the *representation* from Stokes coordinates (I,Q,U,V) to linear polarisation coordinates $(\mathsf{xx}, \mathsf{xy}, \mathsf{yx}, \mathsf{yy})$. Similarly, $\mathsf{S}^{\odot}$ transforms to circular polarisation coordinates $(\mathsf{rr}, \mathsf{rl}, \mathsf{lr}, \mathsf{ll})$. Following the convention of [4], we write:[3]

$$\mathsf{S}^{+} \;=\; \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & i \\ 0 & 0 & 1 & -i \\ 1 & -1 & 0 & 0 \end{pmatrix} \qquad\qquad \mathsf{S}^{\odot} \;=\; \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & i & 0 \\ 0 & 1 & -i & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix} \tag{15}$$

S-matrices are almost unitary, i.e. except for a normalising constant: $(\mathsf{S})^{-1} = 2(\mathsf{S})^{*T}$. S cannot be factored into *feed*-based parts. The two Stokes matrices are related by:

$$\mathsf{S}^{\odot} \;=\; (\mathcal{H} \otimes \mathcal{H}^{*})\;\mathsf{S}^{+} \qquad\qquad \mathsf{S}^{+} \;=\; (\mathcal{H}^{-1} \otimes (\mathcal{H}^{-1})^{*})\;\mathsf{S}^{\odot} \tag{16}$$

with[4]

$$\mathcal{H} \;=\; \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \qquad\qquad \mathcal{H}^{-1} \;=\; \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \tag{18}$$

[3] In one influential book [12], the factor 0.5 is omitted from $\mathsf{S}^{\odot}$. This is clearly incorrect, since a single receptor can never measure more than one half of the total flux of an unpolarised source.

[4] One might argue that a more consistent form of $\mathcal{H}$ would be an expression in terms of the $\pm\pi/4$ *ellipticities* that are intrinsic to a circular *receptor*:

$$\mathcal{H}^{alternative} \;=\; \mathsf{Ell}(\pi/4, -\pi/4) \;=\; \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} \;=\; \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \mathcal{H} \tag{17}$$

However, a choice for a different $\mathcal{H}$ should not be made lightly, since it would affect the deeply entrenched form of the Stokes matrices.

Most Jones matrices will have the same form in both polarisation coordinate frames. But if a Jones matrix is expressed in terms of parameters *that are defined in one of the two frames*, it will have two different but related forms. This is the case for Faraday rotation $\mathsf{F_i}$, *receptor* orientation $\mathsf{P_i}$, and *receptor* cross-leakage $\mathsf{D_i}$, in which the orientation w.r.t. the x, ccY frame plays a role. The two forms of a Jones matrix A can be converted into each other by the coordinate transformation matrix $\mathcal{H}$ and its inverse:

$$A^{\odot} \;=\; \mathcal{H}\, A^{+}\, \mathcal{H}^{-1} \qquad\qquad A^{+} \;=\; \mathcal{H}^{-1}\, A^{\odot}\, \mathcal{H} \tag{19}$$

The conversion may be done by hand, using (the elements a, b, c, d may be complex):

$$\mathcal{H} \begin{pmatrix} a & c \\ d & b \end{pmatrix} \mathcal{H}^{-1} \;=\; 0.5 \begin{pmatrix} (a+b)-i(c-d) & (a-b)+i(c+d) \\ (a-b)-i(c+d) & (a+b)+i(c-d) \end{pmatrix} \tag{20}$$

$$\mathcal{H}^{-1} \begin{pmatrix} a & c \\ d & b \end{pmatrix} \mathcal{H} \;=\; 0.5 \begin{pmatrix} (a+b+c+d) & i(a-b-c+d) \\ -i(a-b+c-d) & (a+b-c-d) \end{pmatrix} \tag{21}$$

Applying these general expressions to rotation $\mathsf{Rot}(\alpha)$ and ellipticity $\mathsf{Ell}(\alpha,-\alpha)$ matrices (see Appendix for their definition), the conversions are:

$$\begin{aligned} \mathcal{H}\ \ \mathsf{Rot}(\alpha)\ \mathcal{H}^{-1} &= \mathsf{Diag}(\exp^{i\alpha}, \exp^{-i\alpha}) \\ \mathcal{H}\ \ \mathsf{Rot}(\alpha,\beta)\ \mathcal{H}^{-1} &= \textit{see equation } 34 \\ \mathcal{H}\ \ \mathsf{Ell}(\alpha,-\alpha)\ \mathcal{H}^{-1} &= \mathsf{Rot}(\alpha) \end{aligned} \tag{22}$$

$$\begin{aligned} \mathcal{H}^{-1}\ \ \mathsf{Rot}(\alpha)\ \mathcal{H} &= \mathsf{Ell}(\alpha,-\alpha) \\ \mathcal{H}^{-1}\ \ \mathsf{Ell}(\alpha,-\alpha)\ \mathcal{H} &= \mathsf{Diag}(\exp^{i\alpha}, \exp^{-i\alpha}) \end{aligned} \tag{23}$$

Usually, all matrices in a 'Jones chain' will be defined in the same coordinate frame. An exception is the case where linear dipole receptors are used in conjunction with a 'hybrid' $\mathsf{H_i}$ to create pseudo-circular receptors:

$$\begin{aligned} \mathsf{J_i} &= \mathsf{G_i^{\odot}}\ \mathsf{H_i}\ \mathsf{D_i^{+}}\ \mathsf{E_i^{+}}\ \mathsf{P_i^{+}}\ \mathsf{K_i^{+}}\ \mathsf{T_i^{+}}\ \mathsf{F_i^{+}} & (\textit{using}\ \mathsf{S}=\mathsf{S^{+}}) \\ &= \mathsf{G_i^{\odot}}\ (\mathsf{H_i}\ \mathsf{D_i^{+}}\ \mathcal{H}^{-1})\ \mathcal{H}\ \mathsf{E_i^{+}}\ \mathsf{P_i^{+}}\ \mathsf{K_i^{+}}\ \mathsf{T_i^{+}}\ \mathsf{F_i^{+}} & (\textit{using}\ \mathsf{S}=\mathsf{S^{+}}) \\ &= \mathsf{G_i^{\odot}}\ \mathsf{D_i^{\odot}}\ \mathcal{H}\ \mathsf{E_i^{+}}\ \mathsf{P_i^{+}}\ \mathsf{K_i^{+}}\ \mathsf{T_i^{+}}\ \mathsf{F_i^{+}} & (\textit{using}\ \mathsf{S}=\mathsf{S^{+}}) \\ &= \mathsf{G_i^{\odot}}\ \mathsf{D_i^{\odot}}\ \mathsf{E_i^{\odot}}\ \mathsf{P_i^{\odot}}\ \mathsf{K_i^{\odot}}\ \mathsf{T_i^{\odot}}\ \mathsf{F_i^{\odot}}\ \mathcal{H} & (\textit{using}\ \mathsf{S}=\mathsf{S^{+}}) \\ &= \mathsf{G_i^{\odot}}\ \mathsf{D_i^{\odot}}\ \mathsf{E_i^{\odot}}\ \mathsf{P_i^{\odot}}\ \mathsf{K_i^{\odot}}\ \mathsf{T_i^{\odot}}\ \mathsf{F_i^{\odot}} & (\textit{using}\ \mathsf{S}=\mathsf{S^{\odot}}) \end{aligned} \tag{24}$$

in which $\mathsf{H_i}$ represents an electronic implementation of the coordinate transformation matrix $\mathcal{H}$. All these expressions are equivalent in the sense that, in conjunction with the indicated Stokes matrix, they produce a coherency vector in circular polarisation coordinates. The choice of which expression to use depends on whether one wishes to model the *feed* explicitly in terms of its physical (dipole) properties, or whether one wishes to regard is as a 'black box' circular *feed* with unknown internal structure.

5 GENERIC FORM OF JONES MATRICES

In this section, the 'generic' form of various 2×2 *feed*-based instrumental Jones matrices (operators) will be treated in some detail.

It will be noted that for each matrix, the 4 elements have been given an 'official' name (e.g. $\mathsf{f_{ixx}}$). The (possibly naive) idea is that, if the *structure* of the Measurement Equation is more or less complete, these 'standard' matrix elements could be referred to explicitly by their official names in other AIPS++ documents (and code), for instance to replace them with specific expressions for particular *telescope*s or purposes.

The subscript convention is as follows: $\mathsf{y_{ibp}}$ is an element of matrix Y for *feed* i, which models the 'coupling factor' for the signal going *from receptor* b *to IF-channel* p. Where possible, the expressions have been reduced to matrices like the diagonal matrix (Diag), rotation matrix (Rot) etc. These are defined in the Appendix.

5.1 Ionospheric Faraday rotation ($\mathsf{F_i}(\vec{\rho}, \vec{r_i})$)

The matrix $\mathsf{F_i^+}$ represents (ionospheric) Faraday rotation of the electric vector over an angle χ_i w.r.t. the celestial $\mathsf{x,y}$-frame. Since χ_i is defined in one of the polarisation coordinate frames, there will be two different forms for $\mathsf{F_i}$ (see also section 4). For linear polarisation coordinates:

$$\mathsf{F_i^+}(\vec{\rho}, \vec{r_i}) \;=\; \begin{pmatrix} \mathsf{f_{ixx}} & \mathsf{f_{iyx}} \\ \mathsf{f_{ixy}} & \mathsf{f_{iyy}} \end{pmatrix} \;=\; \mathsf{Rot}(\chi_\mathsf{i}) \tag{25}$$

In circular polarisation coordinates, the matrix $\mathsf{F_i^\odot}$ is a diagonal matrix which introduces a phase difference, or rather a delay difference. It expresses the fact that ionospheric Faraday rotation is caused by a (strongly frequency-dependent) difference in propagation velocity between right-hand and left-hand circularly polarised signals when travelling through a charged medium like the ionosphere. In terms of the Faraday rotation angle χ_i (see above), we get:

$$\mathsf{F_i^\odot}(\vec{\rho}, \vec{r_i}) \;=\; \begin{pmatrix} \mathsf{f_{irr}} & \mathsf{f_{ilr}} \\ \mathsf{f_{irl}} & \mathsf{f_{ill}} \end{pmatrix} \;=\; \mathcal{H}\; \mathsf{F_i^+}\; \mathcal{H}^{-1} \;=\; \mathsf{Diag}(\exp^{i\chi_\mathsf{i}}, \exp^{-i\chi_\mathsf{i}}) \tag{26}$$

In principle, the Faraday rotation angle is a function of source direction and *feed* position: $\chi_\mathsf{i} = \chi_\mathsf{i}(\vec{\rho}, \vec{r_i})$. However, Faraday rotation is a large-scale effect, so it will usually have the same value for all sources in the primary beam: $\chi_\mathsf{i} = \chi(\vec{r_i})$. For arrays smaller than a few km, the rotation angle will usually also be the same for all *feeds*: $\chi_\mathsf{i} = \chi(t)$. These assumptions reduce the number of independent parameters considerably.

5.2 Atmospheric gain ($\mathsf{T_i}(\vec{\rho}, \vec{r_i})$)

The matrix $\mathsf{T_i^+}$ represents complex atmospheric gain: refraction, extinction and perhaps non-isoplanaticity. Since $\mathsf{T_i^+}$ does not depend on a polarisation coordinate frame, there is only one form:

$$\mathsf{T}_i^+ \;=\; \mathsf{T}_i^{\odot} \;= \mathsf{T}_i(\vec{\rho},\vec{r}_i) \;\approx\; \begin{pmatrix} \mathsf{t}_i & 0 \\ 0 & \mathsf{t}_i \end{pmatrix} \;=\; \mathsf{Mult}(\mathsf{t}_i) \tag{27}$$

The matrix is diagonal because the atmosphere does is not supposed to cause cross-talk. The diagonal elements are assumed to be equal, because the atmosphere is not supposed to affect polarisation.

Atmospheric effects in the 'pupil-plane' (i.e. originating directly above the *feeds*) can be modelled with a complex gain. It is less clear how to deal with effects that originate higher up in the atmosphere, i.e. between pupil plane and image plane.

A phase screen over the array can be modelled as $\mathsf{t}_i = \exp^{i\psi_i}$ in which the phase is assumed to be a low-order 2D polynomial as a function of the *feed* position $\vec{r}$: $\psi_i \;=\; a_0(t) + a_1(t)\,\vec{r}_i + a_2(t)\,\vec{r}_i^{\,2} + \cdots$

5.3 Fourier Transform kernel ($\mathsf{K}_i(\vec{r}_i.\vec{\rho})$)

The matrix K_i represents the Fourier Transform kernel, which can also be seen as a phase weight factor). *It is factored into feed-based parts in order to be able to model a tied array (see section 2.2).* Since K_i does not depend on the polarisation coordinate frame, there is only one form:

$$\mathsf{K}_i^+ \;=\; \mathsf{K}_i^{\odot} \;=\; \mathsf{K}_i(\vec{r}_i.\vec{\rho}) \;=\begin{pmatrix} \mathsf{k}_{iaa} & 0 \\ 0 & \mathsf{k}_{ibb} \end{pmatrix} \;=\begin{pmatrix} \mathsf{k}_i & 0 \\ 0 & \mathsf{k}_i \end{pmatrix} \;=\; \mathsf{Mult}(\mathsf{k}_i) \tag{28}$$

in which $\mathsf{k}_i = \frac{1}{\sqrt{\mathsf{n}}}\exp^{i2\pi\vec{r}_i.\vec{\rho}/\lambda}$, which depends on the *projected feed* position $\vec{r}_i$ and the source direction $\vec{\rho} = \vec{\rho}(\mathsf{l},\mathsf{m})$ w.r.t. the fringe tracking centre $\vec{\rho}_{ftc}$, and $\mathsf{n} = \sqrt{1-\mathsf{l}^2-\mathsf{m}^2} \approx 1 - 0.5(\mathsf{l}^2 + \mathsf{m}^2)$.

If $\mathsf{k}_{iaa} \;=\; \mathsf{k}_{ibb}$, the *interferometer* matrix $\mathsf{K}_{ij} \;=\; (\mathsf{K}_i \otimes \mathsf{K}_j{}^*)$ is a 4×4 diagonal matrix with equal elements. This is equivalent to a multiplicative factor of the familiar form $\mathsf{k}_{ij} = \mathsf{k}_i\mathsf{k}_j^* = \frac{1}{\mathsf{n}}\exp^{i2\pi(\vec{r}_i-\vec{r}_j).\vec{\rho}/\lambda} = \frac{1}{\mathsf{n}}\exp^{i\vec{u}_{ij}.\vec{\rho}}$, i.e. the Fourier Transform kernel or 'phase weight' for the baseline $\vec{u}_{ij}$. For small fields, $\mathsf{n} \approx 1$, so $\vec{u}_{ij}.\vec{\rho} = (\mathsf{ul}+\mathsf{vm}+\mathsf{w}(\mathsf{n}-1)) \approx (\mathsf{ul}+\mathsf{vm})$ becomes a 2D FT.

The *receptors* of a *feed* are practically always co-located, i.e. they have the same phase-centre: $\vec{r}_{ia} = \vec{r}_{ib} = \vec{r}_i$, so $\mathsf{k}_{iaa} = \mathsf{k}_{ibb} = \mathsf{k}_i$. But note that it is possible to model a *receptors* that are *not* co-located, i.e. $\vec{r}_{ia} \neq \vec{r}_{ib}$. It is not immediately obvious why one would want to do such a thing, but it is good to know that the formalism allows it.

5.4 Projection matrix (P_i) if $\gamma_{xa} = \gamma_{yb}$

The 'Projection matrix' models the *projected* orientation of the *receptors* w.r.t. the electrical x,y frame on the sky, as seen from the direction of the source (see also section 5.6 below). Since the orientations are defined in one of the polarisation coordinate frames, there will be two different forms for P_i (see section 4). For linear polarisation coordinates:

$$\mathsf{P}_i^+ \;=\; \begin{pmatrix} \mathsf{p}_{ixa} & \mathsf{p}_{iya} \\ \mathsf{p}_{ixb} & \mathsf{p}_{iyb} \end{pmatrix} \;\equiv\; \begin{pmatrix} \cos\gamma_{xa} & -\sin\gamma_{xa} \\ \sin\gamma_{xa} & \cos\gamma_{xa} \end{pmatrix} \;=\; \mathsf{Rot}(\gamma_{xa}) \tag{29}$$

in which γ_{xa} is the projected angle between the positive x-axis and the orientation of *receptor* a (see also Appendix A.3). There is an implicit assumption here that the *feed* has **perpendicular** *receptor*s and is **fully steerable**, which is the case for the majority of existing *telescope*s. See the next section for the case where the projected orientations are not perpendicular ($\gamma_{\mathsf{xa}} \neq \gamma_{\mathsf{yb}}$).

For circular polarisation coordinates:

$$\mathsf{P}_{\mathsf{i}}^{\odot} = \begin{pmatrix} \mathsf{p_{ira}} & \mathsf{p_{ila}} \\ \mathsf{p_{irb}} & \mathsf{p_{ilb}} \end{pmatrix} = \mathcal{H}\ \mathsf{P}_{\mathsf{i}}^{+}\ \mathcal{H}^{-1} = \mathsf{Diag}(\exp^{i\gamma_{\mathsf{xa}}}, \exp^{-i\gamma_{\mathsf{xa}}}) \tag{30}$$

It is sometimes useful to introduce an intermediate coordinate frame, attached to the *feed* i. In that case: $\gamma_{\mathsf{xa}} = \gamma_{\mathsf{xi}} + \gamma_{\mathsf{ia}} = \beta + \gamma_{\mathsf{ia}}$. The 'offset' angle γ_{ia} between *receptor* a and the frame of *feed* i will be zero in most cases. The angle β is the *parallactic angle*, i.e. the angle between two great circles through the source, and through the celestial North Pole and the local zenith respectively. This parallactic angle is zero for an equatorial *feed*, and varies smoothly with $HA(\mathsf{t})$ for an alt-az *feed*:

$$\begin{aligned} \sin\beta &= \cos LAT\ \sin HA \\ \cos\beta &= \cos DEC\ \sin LAT\ -\ \sin DEC\ \cos LAT\ \cos HA \end{aligned} \tag{31}$$

5.5 Projection matrix (P_{i}) if $\gamma_{\mathsf{xa}} \neq \gamma_{\mathsf{yb}}$

The M.E. formalism must also be able to deal with more 'exotic' *antenna*s like parabolic cylinders (Arecibo, MOST) or horizontal dipole arrays (SKAI). In those cases, the projected angles of the two *receptor*s will generally not be equal, i.e. $\gamma_{\mathsf{xa}} \neq \gamma_{\mathsf{yb}}$.

NB: The angle γ_{yb} of *receptor* b is defined w.r.t. the y-axis rather than the x-axis. This ensures that $\gamma_{\mathsf{yb}} = \gamma_{\mathsf{xa}}$, so that matrix $\mathsf{P}_{\mathsf{i}}^{+}$ reduces to a simple rotation $\mathsf{Rot}(\gamma_{\mathsf{xa}})$, in the common case described in section 5.4 above.

For linear polarisation coordinates $\mathsf{P}_{\mathsf{i}}^{+}$ becomes a 'pseudo-rotation' (compare with equ 29 above):

$$\mathsf{P}_{\mathsf{i}}^{+} = \begin{pmatrix} \mathsf{p_{ixa}} & \mathsf{p_{iya}} \\ \mathsf{p_{ixb}} & \mathsf{p_{iyb}} \end{pmatrix} \equiv \begin{pmatrix} \cos\gamma_{\mathsf{xa}} & -\sin\gamma_{\mathsf{xa}} \\ \sin\gamma_{\mathsf{yb}} & \cos\gamma_{\mathsf{yb}} \end{pmatrix} = \mathsf{Rot}(\gamma_{\mathsf{xa}}, \gamma_{\mathsf{yb}}) \tag{32}$$

For circular polarisation coordinates:

$$\begin{aligned} \mathsf{P}_{\mathsf{i}}^{\odot} &= \begin{pmatrix} \mathsf{p_{ira}} & \mathsf{p_{ila}} \\ \mathsf{p_{irb}} & \mathsf{p_{ilb}} \end{pmatrix} = \mathcal{H}\ \mathsf{P}_{\mathsf{i}}^{+}\ \mathcal{H}^{-1} \\ &= 0.5 \begin{pmatrix} \cos\gamma_{\mathsf{xa}} + \cos\gamma_{\mathsf{yb}} + i(\sin\gamma_{\mathsf{xa}} + \sin\gamma_{\mathsf{yb}}) & \cos\gamma_{\mathsf{xa}} - \cos\gamma_{\mathsf{yb}} - i(\sin\gamma_{\mathsf{xa}} - \sin\gamma_{\mathsf{yb}}) \\ \cos\gamma_{\mathsf{xa}} - \cos\gamma_{\mathsf{yb}} + i(\sin\gamma_{\mathsf{xa}} - \sin\gamma_{\mathsf{yb}}) & \cos\gamma_{\mathsf{xa}} + \cos\gamma_{\mathsf{yb}} - i(\sin\gamma_{\mathsf{xa}} + \sin\gamma_{\mathsf{yb}}) \end{pmatrix} \end{aligned} \tag{33}$$

The future large radio *telescope*s may have *feed*s in the form of dipole arrays, possibly tilted over an angle α towards the South w.r.t. the local horizontal plane. In that case, the *projected* angle γ_{xa} between a North-South (NS) dipole and the x-axis differs from the *projected* angle γ_{yb} between an East-West (EW) dipole and the y-axis (*I hope this is correct now*):

$$
\begin{aligned}
cos\gamma_{\mathsf{xa}} &= \cos HA\ \sin DEC\ \cos(LAT-\alpha)\ -\ \cos DEC\ \sin(LAT-\alpha) \\
sin\gamma_{\mathsf{xa}} &= -\sin HA\ \cos(LAT-\alpha) \\
cos\gamma_{\mathsf{yb}} &= \cos HA \\
sin\gamma_{\mathsf{yb}} &= -\sin HA\ \sin DEC
\end{aligned} \tag{34}
$$

5.6 Voltage primary beam ($\mathsf{E_i}(\vec{\rho})$)

The effects of the primary beam are ignored by [2], which deals implicitly with on-axis sources observed by *feed*s with fully steerable parabolic mirrors. The AIPS++ M.E. must of course deal with the general case, including 'exotic' *telescope*s like Arecibo, MOST and SKAI. To this end, we define a *total voltage pattern* matrix $\mathsf{B_i}$, which fully describes the conversion of the incident electric field (V/m) into two voltages (V):

$$
\mathsf{B_i^+}(\vec{\rho}) = \begin{pmatrix} \mathsf{b_{ixa}} & \mathsf{b_{iya}} \\ \mathsf{b_{ixb}} & \mathsf{b_{iyb}} \end{pmatrix} \qquad \mathsf{B_i^\odot}(\vec{\rho}) = \begin{pmatrix} \mathsf{b_{ira}} & \mathsf{b_{ila}} \\ \mathsf{b_{irb}} & \mathsf{b_{ilb}} \end{pmatrix} \tag{35}
$$

NB: Since the Jones matrix $\mathsf{J_i}$ is feed-based, it deals with voltage beams. The power beam for interferometer ij is modelled by $\mathsf{B_i} \otimes \mathsf{B_j}^*$. Note that the formalism deals implicitly with interferometers between feeds with quite dissimilar primary beams.

In practice, it is often convenient to split the matrix $\mathsf{B_i}$ into a chain of sub-matrices:

- It is always possible to split off a projection matrix $\mathsf{P_i}$:
  $\mathsf{B_i} = (\mathsf{B_i}\ \mathsf{P_i^{-1}})\ \mathsf{P_i} = \mathsf{E_i'}\ \mathsf{P_i}$
  See sections 5.4 and 5.5 above.
- It is always possible to split off a *position-independent* leakage matrix $\mathsf{D_i}$:
  $\mathsf{E_i'}\ \mathsf{P_i} = \mathsf{D_i}\ (\mathsf{D_i^{-1}}\ \mathsf{E_i'})\ \mathsf{P_i} = \mathsf{D_i}\ \mathsf{E_i}\ \mathsf{P_i}$
  See section 5.7 below.

This is most useful in the common case of a fully steerable parabolic *antenna*. The voltage patterns of its *feed*(s) have a fixed shape, which are rotated and translated w.r.t. the sky when pointing the *antenna* in different directions. What remains after splitting off $\mathsf{P_i}$ and $\mathsf{D_i}(\vec{\rho})$ is an (approximately) real and diagonal matrix $\mathsf{E_i}$ which decsribes the position-dependent primary beam attenuation and the *position-dependent* leakage (see also equation 38 below):

$$
\mathsf{E_i^+}(\vec{\rho}) = \mathsf{E_i^\odot}(\vec{\rho}) = \mathsf{E_i}(\vec{\rho}) = \begin{pmatrix} \mathsf{e_{iaa}} & \mathsf{e_{iba}} \\ \mathsf{e_{iab}} & \mathsf{e_{ibb}} \end{pmatrix} \approx \mathsf{Diag}(\mathsf{e_{iaa}}, \mathsf{e_{ibb}}) \tag{36}
$$

As an example, the diagonal elements of $\mathsf{E_i^+}$ for an idealised axially symmetric gaussian beam and dipole *receptor*swould look like:

$$
\begin{aligned}
\mathsf{e}_{\mathsf{iaa}} &= \exp -[(\frac{\mathsf{l}''_{\mathsf{ia}}}{\sigma_{\mathsf{a}}(1+\epsilon_{\mathsf{a}})})^2 \;+\; (\frac{\mathsf{m}''_{\mathsf{ia}}}{\sigma_{\mathsf{a}}(1-\epsilon_{\mathsf{a}})})^2] \\
\mathsf{e}_{\mathsf{ibb}} &= \exp -[(\frac{\mathsf{l}''_{\mathsf{ib}}}{\sigma_{\mathsf{b}}(1+\epsilon_{\mathsf{b}})})^2 \;+\; (\frac{\mathsf{m}''_{\mathsf{ib}}}{\sigma_{\mathsf{b}}(1-\epsilon_{\mathsf{b}})})^2]
\end{aligned} \tag{37}
$$

Note that the two *receptor* beams are each described in their own coordinate frame $\mathsf{l}''_{\mathsf{ia}}, \mathsf{m}''_{\mathsf{ia}}$ and $\mathsf{l}''_{\mathsf{ib}}, \mathsf{m}''_{\mathsf{ib}}$ projected on the sky (see Appendix A). *The projection matrix* P_{i} *only takes care of electrical rotation, but not of the rotation of the voltage beam on the sky!.*

Equation 37 illustrates that the voltage beam of a dipole *receptor* will be slightly elongated in the direction of the dipole by a factor $(1+\epsilon)$, even if the mirror is perfectly circular and symmetrical. Obviously, the two asymmetric voltage beams of a *feed* will not coincide, because they are oriented differently. The resulting position-dependent difference is one cause of off-axis instrumental polarisation.

In reality, things will be more complicated, especially for off-axis sources. For instance, standing waves between the primary mirror and the frontend box, or scattering off support legs, may cause position-dependent leakage terms. Since these cannot be part of D_{i}, they must be modelled as off-diagonal elements of E_{i} itself.

In general, E_{i} will be more complicated for *antenna*s with less symmetry. In some exotic cases, it may not be very useful to split off D_{i} or even P_{i}, *although it is always allowed.* In any case, the M.E. formalism offers a framework for the ful description of the primary beam of any radio *telescope* that can be conceived.

5.7 Position-independent *receptor* cross-leakage (D_{i})

The off-diagonal elements $\mathsf{e}_{\mathsf{iba}}$ and $\mathsf{e}_{\mathsf{iab}}$ of E'_{i} describe 'leakage' between *receptor*s, i.e. the extent to which each *receptor* is sensitive to the radiation that is supposed to be picked up by the other one.

It is customary to split off the *position-independent* part $\mathsf{e}'_{\mathsf{iba}}$ and $\mathsf{e}'_{\mathsf{iab}}$ of this leakage into a separate matrix D_{i}:

$$
\begin{aligned}
\mathsf{E}'_{\mathsf{i}}(\vec{\rho}) &= \begin{pmatrix} \mathsf{e}_{\mathsf{iaa}} & \mathsf{e}_{\mathsf{iba}} + \mathsf{e}'_{\mathsf{iba}} \\ \mathsf{e}_{\mathsf{iab}} + \mathsf{e}'_{\mathsf{iab}} & \mathsf{e}_{\mathsf{ibb}} \end{pmatrix} \\
&\approx \begin{pmatrix} 1 & \mathsf{e}'_{\mathsf{iba}}/\mathsf{e}_{\mathsf{ibb}} \\ \mathsf{e}'_{\mathsf{iab}}/\mathsf{e}_{\mathsf{iaa}} & 1 \end{pmatrix} \begin{pmatrix} \mathsf{e}_{\mathsf{iaa}} & \mathsf{e}_{\mathsf{iba}} \\ \mathsf{e}_{\mathsf{iab}} & \mathsf{e}_{\mathsf{ibb}} \end{pmatrix} \\
&= \begin{pmatrix} \mathsf{d}_{\mathsf{iaa}} & \mathsf{d}_{\mathsf{iba}} \\ \mathsf{d}_{\mathsf{iab}} & \mathsf{d}_{\mathsf{ibb}} \end{pmatrix} \begin{pmatrix} \mathsf{e}_{\mathsf{iaa}} & \mathsf{e}_{\mathsf{iba}} \\ \mathsf{e}_{\mathsf{iab}} & \mathsf{e}_{\mathsf{ibb}} \end{pmatrix} \;=\; \mathsf{D}_{\mathsf{i}}\;\mathsf{E}_{\mathsf{i}}(\vec{\rho})
\end{aligned} \tag{38}
$$

Usually, the position-dependent leakage coefficients $\mathsf{e}_{\mathsf{iba}}$ and $\mathsf{e}_{\mathsf{iab}}$ are assumed to be zero, but that is not always justified.

If the leakage coefficients are determined *empirically* by calibration, it is not necessary to know the details of the leakage mechanism. It is sufficient to solve for the elements of $\mathsf{D_i}$. In that case, there is only one form:

$$\mathsf{D_i^+} \;=\; \mathsf{D_i^\odot} \;=\; \mathsf{D_i} \;=\; \begin{pmatrix} \mathsf{d_{iaa}} & \mathsf{d_{iba}} \\ \mathsf{d_{iab}} & \mathsf{d_{ibb}} \end{pmatrix} \tag{39}$$

But in many cases, position-independent leakage can be physically explained by deviations ϕ from the *nominal receptor* position angles (see $\mathsf{P_i}$), and by deviations θ from *nominal receptor* 'ellipticities' θ. For linear polarisation coordinates:

$$\begin{aligned} \mathsf{D_i^+} \;=\; \begin{pmatrix} \mathsf{d_{iaa}} & \mathsf{d_{iba}} \\ \mathsf{d_{iab}} & \mathsf{d_{ibb}} \end{pmatrix} &= \mathsf{Ell}(\theta_{\mathsf{ia}}, \theta_{\mathsf{ib}})\ \mathsf{Rot}(\phi_{\mathsf{ia}}, \phi_{\mathsf{ib}}) \\ &\approx \mathsf{Ell}(\theta_{\mathsf{ia}}, -\theta_{\mathsf{ia}})\ \mathsf{Rot}(\phi_{\mathsf{ia}}) \end{aligned} \tag{40}$$

The $\approx$ sign gives the approximation for a well-designed system. Often the two *receptors* are mounted in a single unit, so position angle deviations caused by mechanical bending of the *feed* structure are the same for both: $\phi_{\mathsf{ia}} = \phi_{\mathsf{ib}}$. One might also argue that ellipticity should be a reciprocal effect, so that $\theta_{\mathsf{ib}} = -\theta_{\mathsf{ia}}$. This is roughly consistent with WSRT experience, and these two assumptions are implicit in equ 27 of [3]. However, for high accuracy polarisation measurements, the parameters for each *receptor* should be at least partly independent.

For circular polarisation coordinates (see equ 22):

$$\begin{aligned} \mathsf{D_i^\odot} \;=\; \mathcal{H}\ \mathsf{D_i^+}\ \mathcal{H}^{-1} &= (\mathcal{H}\ \mathsf{Ell}(\theta_{\mathsf{ia}}, \theta_{\mathsf{ib}})\ \mathcal{H}^{-1})\ \ (\mathcal{H}\ \mathsf{Rot}(\phi_{\mathsf{ia}}, \phi_{\mathsf{ib}})\ \mathcal{H}^{-1}) \\ &\approx \mathsf{Rot}(\theta_{\mathsf{ia}}) \quad \mathsf{Diag}(\exp^{i\phi_{\mathsf{ia}}}, \exp^{-i\phi_{\mathsf{ia}}}) \end{aligned} \tag{41}$$

Again, the $\approx$ sign gives the approximation for $\phi_{\mathsf{ia}} = \phi_{\mathsf{ib}}$ and $\theta_{\mathsf{ib}} = -\theta_{\mathsf{ia}}$. See equation 34 for an expression for $(\mathcal{H}\ \mathsf{Rot}(\phi_{\mathsf{ia}}, \phi_{\mathsf{ib}})\ \mathcal{H}^{-1})$ where $\phi_{\mathsf{ia}} \neq \phi_{\mathsf{ib}}$. The expression for $(\mathcal{H}\ \mathsf{Ell}(\theta_{\mathsf{ia}}, \theta_{\mathsf{ib}})\ \mathcal{H}^{-1})$ with $\theta_{\mathsf{ib}} \neq -\theta_{\mathsf{ia}}$ is similar, but with real coefficients, as expected for circular polarisation coordinates.

5.8 Commutation ($\mathsf{Y_i}$)

In some systems, the *receptor* signals can be switched (commuted) between *IF-channels* for calibration.

$$\mathsf{Y_i} \;=\; \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \textit{or} \qquad \mathsf{Y_i} \;=\; \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{42}$$

5.9 Hybrid ($\mathsf{H_i}$)

In some cases, circularly polarised *receptors* consist of linearly polarised dipoles, followed by a 'hybrid'. The latter is an electronic implementation of the coordinate transformation matrix $\mathcal{H}$ from linear to circular polarisation coordinates:

$$\mathsf{H_i} \;\approx\; \mathcal{H} \tag{43}$$

See equation 18 for the definition of $\mathcal{H}$. If no hybrid is present, $\mathsf{H_i}$ is the unit matrix. Any gain effects in these electronic components are ignored, or rather they are assumed to be 'absorbed' by the gain matrix $\mathsf{G_i}$.

5.10 Electronic gain ($\mathsf{G_i}$)

The matrix $\mathsf{G_i}$ represents the product of all complex *electronic* gain effects per output *IF-channel* p and q. It models the effects of all *feed*-based electronics (amplifiers, mixers, LO, cables etc). (The correlator causes *interferometer*-based effects, which are discussed in section 3).

$$\mathsf{G_i^+} = \mathsf{G_i^\odot} = \begin{pmatrix} \mathsf{g_{ipp}} & \mathsf{g_{iqp}} \\ \mathsf{g_{ipq}} & \mathsf{g_{iqq}} \end{pmatrix} \approx \begin{pmatrix} \mathsf{g_{ip}} & 0 \\ 0 & \mathsf{g_{iq}} \end{pmatrix} \;=\; \mathsf{Diag}(\mathsf{g_{ip}}, \mathsf{g_{iq}}) \tag{44}$$

The $\approx$ sign indicates that electronic cross-talk is assumed to be absent in well-designed systems, i.e. $\mathsf{g_{ipq}} = \mathsf{g_{iqp}} = 0$. Since this kind of crosstalk is not necessarily reciprocal, $\mathsf{g_{ipq}} \neq \mathsf{g_{iqp}}$.

In reality, $\mathsf{G_i}$ will be a product of many electronic gain matrices, one for each *linear* electronic component in the system: $\mathsf{G_i} \;=\; \mathsf{G_i}^{LNA}\ \mathsf{G_i}^{mixers}\ \mathsf{G_i}^{cables}\ \mathsf{G_i}^{IF-system}\ \cdots$ Although a solver will not be able to distinguish these different effects from each other, but it is useful for **simulation** of instrumental effects.

5.11 Do we need a configuration matrix ($\mathsf{C_i}$)?

NB: This section is a little polemical, and should disappear when things are more settled.

There has been some debate about the concept of a 'configuration matrix' $\mathsf{C_i}$, as proposed by [2], which models the nominal feed configuration. It represents an idealised coordinate transformation 'from the frame of the rotating antenna mount to the electronic voltage frame'. It models any rotation of the *receptor*s w.r.t. 'the *antenna* mount', which must be added to the 'parallactic' rotation $\mathsf{P_i}$ of the *antenna* w.r.t. the sky. $\mathsf{C_i}$ also models the hybrid $\mathsf{H_i}$ if present, but it ignores the primary beam $\mathsf{E_i}$. Any deviations from this idealised behaviour are covered by the 'leakage' matrix $\mathsf{D_i}$.

However, the proposed $\mathsf{C_i}$ is most suitable for the special case of fully steerable parabolic *antenna*s. The introduction of an intermediate *antenna* coordinate frame seems an unnecessary complication in those cases where the mirror is not steerable, or is absent entirely (like in a dipole array). Moreover, $\mathsf{C_i}$ violates the rules of modelling by lumping together two effects that have nothing to do with each other, and do not even occur at the same point in the signal path.

In principle it is a good idea to have one matrix that models the transition from electric fields (V/m) to electric voltages (V), and this is precisely what $\mathsf{B_i}$ does. This very general matrix can be split up *if relevant* into sub-matrices like $\mathsf{P_i}$, $\mathsf{E_i}$ and $\mathsf{D_i}$. The matrix $\mathsf{H_i}$ has no part in this, since it represents a rearranging of electronic signals (V), just like $\mathsf{Y_i}$ (and will come *after* $\mathsf{Y_i}$ if present!). The projection matrix $\mathsf{P_i}$ takes care of the entire orientation angle of the *receptor*s w.r.t. the sky, which is the only thing that really counts.

6 THE ORDER OF JONES MATRICES

The Jones matrices in equation 3 generally do not commute, so their order is important. In principle, the matrices must be placed in the 'physical' order, i.e. the order of the signal propagation path. But in the equations that are enshrined in existing reduction packages, this is often not the case. This begs the question why these 'wrong' equations seem to produce so many good (even spectacular) results. The question is especially important since a different order often results in considerable gains in computational efficiency.

The answer is that, for existing (arrays of) circularly symmetric parabolic *feed*s, many Jones matrices can be approximated by matrices that *do* commute with at least some of the others.

6.1 Overview of commutation properties

We will analyse this in terms of those special matrices (see Appendix for their definition), whose commutation properties are:

- **Unit** matrices $\mathcal{U}$ commute with all matrices.
- **Multiplication** matrices $\mathsf{Mult}(a)$, i.e. diagonal matrices with equal elements a, are equivalent to a multiplicative factor. Therefore, they commute with all matrices.
- **Diagonal** matrices $\mathsf{Diag}(a,b)$ with *unequal* elements a, b commute with each other.
- **Pure rotation** matrices $\mathsf{Rot}(\alpha)$ commute with each other.
- **Pseudo rotation** matrices $\mathsf{Rot}(\alpha,\beta)$ do not commute wit each other or with pure rotation matrices $\mathsf{Rot}(\alpha)$. Moreover, there should only be *one* pseudo rotation matrix in the chain, and it should be to the left of (i.e. after) all other rotation matrices: $\mathsf{Rot}(\alpha,\beta)\mathsf{Rot}(\gamma) = \mathsf{Rot}(\alpha+\gamma,\beta+\gamma) \neq \mathsf{Rot}(\gamma)\mathsf{Rot}(\alpha,\beta)$.
- **Ellipticity** matrices $\mathsf{Ell}(\alpha,\beta)$ do not commute with each other , except when $\beta = -\alpha$. Moreover: $\mathsf{Ell}(\alpha,\beta)\mathsf{Ell}(\gamma,-\gamma) = \mathsf{Ell}(\alpha+\gamma,\beta-\gamma) \neq \mathsf{Ell}(\gamma,-\gamma)\mathsf{Ell}(\alpha,\beta)$.

In order to study the general implications of changing the order of multiplication, we take the two products $m.M$ and $M.m$ of two general matrices (whose elements may be complex):

$$\begin{pmatrix} a & c \\ d & b \end{pmatrix} \begin{pmatrix} A & C \\ D & B \end{pmatrix} = \begin{pmatrix} aA+cD & aC+cB \\ dA+bD & dC+bB \end{pmatrix}$$

$$\begin{pmatrix} A & C \\ D & B \end{pmatrix} \begin{pmatrix} a & c \\ d & b \end{pmatrix} = \begin{pmatrix} aA+dC & cA+bC \\ aD+dB & cD+bB \end{pmatrix} \tag{45}$$

The difference (i.e. commutation error) between the two matrix products can be expressed as a matrix Δ:

$$m\,M \;=\; M\,m \;+\; \Delta \;=\; M\,m \;+\; \begin{pmatrix} cD-dC & -c(A-B)+C(a-b) \\ d(A-B)-D(a-b) & -(cD-dC) \end{pmatrix} \tag{46}$$

Thus, by taking the wrong matrix order, one makes the following *fractional* errors of the following order in the result:
- in the diagonal elements: of the order of c/a, i.e. the ratio of non-diagonal and diagonal elements of the original matrices (which is often small).
- in the off-diagonal elements: in the order of $(a-b)/a$, i.e. they will be smaller as the diagonal elements of the original matrices are more equal.

If one of the two matrices is diagonal, e.g. $c = d = 0$ then this reduces to:

$$m\ M \ = \ M\ m \ + \ \begin{pmatrix} 0 & C(a-b) \\ D(b-a) & 0 \end{pmatrix} \tag{47}$$

The (not very surprising) conclusion is that the error caused by taking the wrong matrix order is smaller when one of the matrices is diagonal, and the values of its diagonal elements are almsot equal.

6.2 Overview of Jones matrix forms

It is sufficient to discuss the commutation properties of the *feed*-based Jones matrices because, if A_{i} commutes with B_{i} and A_{j} with B_{j}, then $(A_{\mathsf{i}} \otimes A_{\mathsf{j}}{}^*)$ commutes with $(B_{\mathsf{i}} \otimes B_{\mathsf{j}}{}^*)$:

$$(\mathsf{J_i} \otimes \mathsf{J_j}^*) \ = (A_{\mathsf{i}} B_{\mathsf{i}} \cdots Z_{\mathsf{i}}) \otimes (A_{\mathsf{j}} B_{\mathsf{j}} \cdots Z_{\mathsf{j}})^* \ = \ (A_{\mathsf{i}} \otimes A_{\mathsf{j}}{}^*)\ (B_{\mathsf{i}} \otimes B_{\mathsf{j}}{}^*) \ \cdots \ (Z_{\mathsf{i}} \otimes Z_{\mathsf{j}}{}^*) \tag{48}$$

Inspecting the various Jones matrices separately:

$$
\begin{array}{ll}
\mathsf{F_i^+} & = \text{pure rotation } \mathsf{Rot}(\chi_{\mathsf{i}}) \\
\mathsf{F_i^\odot} & = \text{diagonal matrix } \mathsf{Diag}(\exp^{i\chi_{\mathsf{i}}}, \exp^{-i\chi_{\mathsf{i}}}) \\
\mathsf{T_i^+}, \mathsf{T_i^\odot} & = \text{multiplication } \mathsf{Mult}(\mathsf{t_i}) \\
\mathsf{K_i} & = \text{multiplication } \mathsf{Mult}(\exp^{i\vec{\rho}.\vec{r}_{\mathsf{i}}}) \quad \text{if } \vec{r}_{\mathsf{ia}} = \vec{r}_{\mathsf{ib}} \text{ (virtually always the case)} \\
\mathsf{P_i^+} & = \text{pure rotation } \mathsf{Rot}(\gamma_{\mathsf{xa}}) \quad \text{if } \gamma_{\mathsf{xa}} = \gamma_{\mathsf{yb}} \\
\mathsf{P_i^\odot} & = \text{diagonal matrix } \mathsf{Diag}(\exp^{i\gamma_{\mathsf{xa}}}, \exp^{-i\gamma_{\mathsf{xa}}}) \quad \text{if } \gamma_{\mathsf{xa}} = \gamma_{\mathsf{yb}} \\
\mathsf{P_i^+} & = \text{pseudo-rotation } \mathsf{Rot}(\gamma_{\mathsf{xa}}, \gamma_{\mathsf{yb}}) \quad \text{if } \gamma_{\mathsf{xa}} \neq \gamma_{\mathsf{yb}} \\
\mathsf{P_i^\odot} & = \text{A general matrix} \quad \text{if } \gamma_{\mathsf{xa}} \neq \gamma_{\mathsf{yb}} \\
\mathsf{E_i^+}, \mathsf{E_i^\odot} & = \text{diagonal matrix } \mathsf{Diag}(\mathsf{e_{iaa}}, \mathsf{e_{ibb}}) \quad \text{if no cross-leakage } (\mathsf{e_{iab}} = \mathsf{e_{iba}} = 0) \\
 & = \text{multiplication } \mathsf{Mult}(\mathsf{e_i}) \quad \text{if also } \mathsf{e_{iaa}} = \mathsf{e_{ibb}} \text{ for all } \vec{\rho} \\
\mathsf{D_i^+}, \mathsf{D_i^\odot} & \approx \text{unit matrix } \mathcal{U} \quad \text{if small leakage, i.e. } (\mathsf{d_{iab}} \approx \mathsf{d_{iba}} \approx 0) \\
\mathsf{D_i^+} & = \mathsf{Ell}(\theta_{\mathsf{ia}}, \theta_{\mathsf{ib}})\ \mathsf{Rot}(\phi_{\mathsf{ia}}, \phi_{\mathsf{ib}}) \\
 & \approx \mathsf{Ell}(\theta_{\mathsf{ia}}, -\theta_{\mathsf{ia}})\ \mathsf{Rot}(\phi_{\mathsf{ia}}) \quad \text{if } \theta_{\mathsf{ib}} = -\theta_{\mathsf{ia}} \text{ and } \phi_{\mathsf{ib}} = \phi_{\mathsf{ia}} \\
\mathsf{D_i^\odot} & = (\mathcal{H}\ \mathsf{Ell}(\theta_{\mathsf{ia}}, \theta_{\mathsf{ib}})\ \mathcal{H}^{-1})\ (\mathcal{H}\ \mathsf{Rot}(\phi_{\mathsf{ia}}, \phi_{\mathsf{ib}})\ \mathcal{H}^{-1}) \\
 & \approx \mathsf{Rot}(\theta_{\mathsf{ia}})\ \mathsf{Diag}(\exp^{i\phi_{\mathsf{ia}}}, \exp^{-i\phi_{\mathsf{ia}}}) \quad \text{if } \theta_{\mathsf{ib}} = -\theta_{\mathsf{ia}} \text{ and } \phi_{\mathsf{ib}} = \phi_{\mathsf{ia}} \\
[\mathsf{Y_i}] & = \text{anti-diagonal matrix: a problem, if present....} \\
[\mathsf{H_i}] & = \text{effectively hidden if present, see equation 24} \\
\mathsf{G_i} & = \text{diagonal matrix } \mathsf{Diag}(\mathsf{g_{ipp}}, \mathsf{g_{iqq}}) \quad \text{if no cross-talk}
\end{array}
$$

Problems are caused predominantly by matrices with non-zero off-diagonal elements like $\mathsf{D_i}$, $\mathsf{Y_i}$, and $\mathsf{P_i}$ if $\gamma_{\mathsf{xa}} \neq \gamma_{\mathsf{yb}}$. Of these, only $\mathsf{D_i}$ is present in all *telescope*s. $\mathsf{P_i}$ will be a problem for SKAI, bacause $\gamma_{\mathsf{xa}} \neq \gamma_{\mathsf{yb}}$.

6.3 Allowable changes of order

The following changes in the order of Jones matrices is allowed, but only under the indicated conditions. *NB: Some Jones matrices will commute if it can be assumed that the observed source is compact, dominating, unpolarised and near the centre of the field. This is often the case.*

- If the Faraday angle does not vary over the primary beam, $\mathsf{F_i}$ might be applied in the uv-plane. $\mathsf{F_i}$ will in general commute with $\mathsf{P_i}$ *except when* $\mathsf{P_i}$ *is a pseudo-rotation ($\gamma_{\mathsf{xa}} \neq \gamma_{\mathsf{yb}}$).* $\mathsf{F_i^{\odot}}$ is diagonal, and will commute with $\mathsf{E_i}$ if it is diagonal. But $\mathsf{F_i^{+}}$ will only commute with $\mathsf{E_i}$ if the latter is a multiplication. If there is appreciable cross-leakage, $\mathsf{F_i}$ should stay to the right of $\mathsf{D_i}$, which means that in that case $\mathsf{F_i^{\odot}}$ cannot be lumped with $\mathsf{G_i}$ as is often done.

- $\mathsf{T_i}$ is a multiplication, which commutes with everything. If it does not vary over the primary beam, it can be lumped with $\mathsf{G_i}$.

- If the two *receptor*s of a *feed* are located at the same position (which is virtually always the case), the FT kernel matrix $\mathsf{K_i}(\vec{\rho}.\vec{r}_{\mathsf{i}})$ reduces to a multiplication $\mathsf{k_i}(\vec{\rho}.\vec{r}_{\mathsf{i}})$. This means that the FT can be performed at any desired place in the chain, even to the right of the Stokes matrix. *NB: If $\vec{r}_{\mathsf{ia}} \neq \vec{r}_{\mathsf{ib}}$, it would not be trivial to figure out what the correct position of* $\mathsf{K_i}$ *should be.*

- If the map centre $\vec{\rho}_{mc}$ is different from the fringe tracking centre $\vec{\rho}_{ftc}$, the FT kernel may be split into a product: $\mathsf{K_i}(\vec{\rho}.\vec{r}_{\mathsf{i}}) \;=\; \mathsf{K^0}_{\mathsf{i}}(\vec{\rho}_{mc}.\vec{r}_{\mathsf{i}})\; \mathsf{K'}_{\mathsf{i}}((\vec{\rho}-\vec{\rho}_{mc}).\vec{r}_{\mathsf{i}})$. Since $\vec{\rho}_{mc}$ does not depend on source position, $\mathsf{K^0}_{\mathsf{i}}(\vec{\rho}_{mc}.\vec{r})$ may be moved to the leftmost part of the chain, i.e. to the uv-plane part.

- If $\mathsf{E_i}(\vec{\rho}) = \mathsf{E_j}(\vec{\rho}) = \mathsf{Mult}(\mathsf{e}(\vec{\rho}))$, i.e. if all voltage patterns are identical, then $\mathsf{E_{ij}} = (\mathsf{E_i} \otimes \mathsf{E_j}^*)$ commutes with the Stokes matrix S and may be applied directly to the Stokes vector $\vec{I}$ in the image plane. This condition is more likely to occur near the beam centre. *NB: Because* $\mathsf{E_{ij}}$ *does definitely not commute with* S *if* $\mathsf{e_{iaa}} \neq \mathsf{e_{ibb}}$*, the justification for the practice of applying off-axis instrumental polarisation to $\vec{I}$ seems a little doubtful.*

- $\mathsf{P_i}$ may be moved to the left of $\mathsf{E_i}$ if they are both diagonal matrices, or if $\mathsf{E_i}$ is a multiplication. Since $\mathsf{P_i^{\odot}}$ is diagonal and $\mathsf{P_i^{+}}$ is not (except for equatorial mounts), this appears to be an argument in favour of the use of circular polarisation coordinates. If $\mathsf{E_i}$ is diagonal and almost a multiplication (i.e. $\mathsf{e_{iaa}} \approx \mathsf{e_{ibb}}$), $\mathsf{P_i^{+}}$ may be moved to the left of $\mathsf{E_i}$ at the cost of a small error of the order $(\mathsf{e_{iaa}} - \mathsf{e_{ibb}})/\mathsf{e_{iaa}}$ (see equation 47).

- If $\mathsf{P_i}$ and $\mathsf{E_i}$ do not commute at all, one can still move $\mathsf{P_i}$ to the left of $\mathsf{E_i}$ by using $\mathsf{E_i}\;\mathsf{P_i} \;=\; \mathsf{P_i}\;(\mathsf{P_i^{-1}}\;\mathsf{E_i}\;\mathsf{P_i}) \;=\; \mathsf{P_i}\;\mathsf{E_i''}$
  Since this re-introduces time-dependent off-diagonal elements into $\mathsf{E_i''}$, it is not clear how useful this is.

6.4 VisJones and SkyJones

The Jones matrices may split up in two groups: $\mathsf{J_i} \;=\; \mathsf{J_i^{\mathit{vis}}}\;\mathsf{J_i^{\mathit{sky}}}$. In these terms, the full M.E. (ignoring normalisation factors, see equ 6) becomes:

$$\vec{V}_{\mathsf{ij}} = \int d\mathsf{t} \int d\mathsf{f} \; (\mathsf{J}_{\mathsf{i}}^{vis} \otimes \mathsf{J}_{\mathsf{j}}^{vis*}) \sum_k \int d\mathsf{l}\, d\mathsf{m} \; (\mathsf{J}_{\mathsf{i}}^{sky} \otimes \mathsf{J}_{\mathsf{j}}^{sky*}) \; \mathsf{S} \; \vec{I}_k \tag{49}$$

We now see the reason for placing the integration over f and t to the left of the sum over k sources. Since it is computationally advantageous to minimise the number of Jones matrices that operate in the image plane, it must be investigated whether Jones matrices that do not depend on the source position can be moved to the left in the chain, using the rules in section 6.3 above. Depending on the chosen coordinate system, (*and always keeping in mind the conditions for re-ordering Jones matrices*), the following split appears to be the maximum obtainable:

$$\mathsf{J}_{\mathsf{i}}^{vis} = \mathsf{K}^0{}_{\mathsf{i}} \; (\mathsf{G}_{\mathsf{i}} \; \mathsf{T}_{\mathsf{i}}) \; \mathsf{D}_{\mathsf{i}}^{+} \; \mathsf{P}_{\mathsf{i}}^{+} \; \mathsf{F}_{\mathsf{i}}^{+} \qquad (using\ \mathsf{S} = \mathsf{S}^{+}) \tag{50}$$

$$= \mathsf{K}^0{}_{\mathsf{i}} \; (\mathsf{G}_{\mathsf{i}} \; \mathsf{T}_{\mathsf{i}} \; \mathsf{F}_{\mathsf{i}}^{\odot}) \; \mathsf{D}_{\mathsf{i}}^{\odot} \; \mathsf{P}_{\mathsf{i}}^{\odot} \qquad (using\ \mathsf{S} = \mathsf{S}^{\odot}) \tag{51}$$

$$\mathsf{J}_{\mathsf{i}}^{sky} = \mathsf{E}_{\mathsf{i}} \; \mathsf{K}'{}_{\mathsf{i}} \tag{52}$$

This is what is done implicitly in some existing reduction packages.

6.4.1 Tied Array

For a tied array (ignoring integration and weight factors for the moment), equation 5 becomes:

$$\vec{V}_{\mathsf{ij}} = (\mathsf{Q}_{\mathsf{i}} \otimes \mathsf{Q}_{\mathsf{j}}^*) \sum_n \sum_m \; (\mathsf{J}_{\mathsf{i}n}^{vis} \otimes \mathsf{J}_{\mathsf{j}m}^{vis*}) \sum_k \; (\mathsf{J}_{\mathsf{i}nk}^{sky} \otimes \mathsf{J}_{\mathsf{j}mk}^{sky\,*}) \; \mathsf{S} \; \vec{I}_k \tag{53}$$

Under extremely favourable conditions, i.e. if:
- individual *feed* beams per tied array are identical.
- Faraday rotation is the same for an entire tied array
- All *receptor*s of a tied array have the same orientation.
- *receptor* cross-leakages are small.
- tied array *feed* signals are corrected before adding.
- there are no delay errors.

then equation 53 can be reduced to:

$$\vec{V}_{\mathsf{ij}} = (\mathsf{Q}_{\mathsf{i}} \otimes \mathsf{Q}_{\mathsf{j}}^*) \; (\mathsf{P}_{\mathsf{i}} \otimes \mathsf{P}_{\mathsf{j}}^*) \; (\mathsf{F}_{\mathsf{i}} \otimes \mathsf{F}_{\mathsf{j}}^*) \sum_k \; (\mathsf{E}_{\mathsf{i}k} \otimes \mathsf{E}_{\mathsf{j}k}^*) \sum_n \sum_m \; (\mathsf{K}_{\mathsf{i}nk} \otimes \mathsf{K}_{\mathsf{j}mk}^*) \; \mathsf{S} \; \vec{I}_k \tag{54}$$

References

[1] **J.D.Bregman, J.E.Noordam** *Matrix formalism for Interferometric Polarisation Calibration.* **Internal proposal to AIPS++ project, April 1993.**

[2] **J.P.Hamaker, J.D.Bregman, R.J. Sault** *Understanding Radio Polarimetry I: Mathematical foundations.* **Accepted by Astronomy and Astrophysics, Sept 1995. (***For a preprint, see http:://www.nfra.nl/∼hamaker***).**

[3] **R.J.Sault, J.P.Hamaker, J.D.Bregman** *Understanding Radio Polarimetry II: Instrumental calibration of an interferometer array.* **Accepted by Astronomy and Astrophysics, Sept 1995.** (*For a preprint, see http:://www.nfra.nl/∼hamaker*).

[4] **J.P.Hamaker, J.D.Bregman** *Understanding Radio Polarimetry III: Interpreting the IAU/IEEE definitions of the Stokes parameters* **Submitted to Astronomy and Astrophysics, Oct 1995.** (*For a preprint, see http:://www.nfra.nl/∼hamaker*).

[5] **J.E.Noordam** *Some practical aspects of the matrix-based Measurement Equation of a generic radio telescope.* **AIPS++ Implementation note 182 (June 1995)**

[6] **T.J.Cornwell** *Calibration and Imaging using the Measurement Equation for the Generic Interferometer.* **AIPS++ Implementation note 183 (July 1995)**

[7] **T.J.Cornwell** *The Generic Interferometer I: Overview of Calibration and Imaging* **AIPS++ Implementation note 183 (August 1995)**

[8] **T.J.Cornwell** *The Generic Interferometer II: Image Solvers* **AIPS++ Implementation note ... (revised version, Aug 1995) developing**

[9] **T.J.Cornwell** *The Generic Interferometer III: Analysis of Calibration and Imaging* **AIPS++ Implementation note ... (Nov 1995) developing**

[10] **T.J.Cornwell, M.H.Wieringa** *The Generic Interferometer IV: Design of Calibration and Imaging* **AIPS++ Implementation note ... (Dec 1995) developing**

[11] **T.J.Cornwell** *The Generic Interferometer V: Specification of Calibration and Imaging* **AIPS++ Implementation note ... (Sept 1995) developing**

[12] **A.R.Thompson, J.M.Moran, G.W.Swenson** *Interferometry and Synthesis in Radio Astronomy.* **John Wiley and Sons (1986)**

[13] **R.A.Perley, F.R.Schwab, A.H.Bridle** *Synthesis Imaging in Radio Astronomy.* **Astronomical Society of the Pacific Conference Series, Vol 6 (1989)**

A APPENDIX: CONVENTIONS

A consistent nomenclature and precise definitions are extremely important for a software package like AIPS++, which aspires to be a 'world reduction package', and to which workers with a large spacetime separation are supposed to contribute. One of the most sensitive areas in this respect is the Measurement Equation, which underlies the central subject of uv-calibration and imaging.

However, it is not easy to *define*, *adopt* and *enforce* the use of a suitable set of conventions. This appendix is a hopefully useful step in that process. It proposes coordinate conventions and some definitions (notably the one for *feed*!), and lists symbols that have been defined in a separate TeX file (referred to as `\include(megi-symbols)` in this LaTeX document). The TeX syntax is shown in small print (e.g. `\FeedI`), for easy reference.

A.1 Some definitions

The following definitions are displayed in a distinctive font throughout the text of this document in order to emphasize that they have been defined explicitly.

- A *receptor* (`\Receptor`) converts the incident electric field into a voltage.
- An *IF-channel* (`\IFchannel`) is one of the two output signals of a *feed*, one for each 'polarisation'. *NB: The signals in a pair of IF-channels may be a linear combination of the signals of the two receptors.*
- A *feed* (`\Feed`) is the most fundamental concept of the M.E. formalism, since Jones-matrices are *feed*-based. Although a *feed* may sometimes have only one *receptor*, it usually has two, which is necessary and sufficient to fully sample the incident e.m. field. Each *feed* is modelled by its own Jones matrix. *NB: A feed is a logical concept. Thus, the same physical feed may be involved in several logical feeds, e.g. for different beams in a multi-beam instrument, or for different spectral windows.*
- An *antenna* (`\Antenna`) is a physical grouping of *feeds*. *NB: As a concept, it tends to play a rather confusing role in the M.E. discussions.*
- An *interferometer* (`\Interferometer`) is the combination of two *feeds*. Its output is a *visibility* of 1-4 spectra, depending on the number of *IF-channels* per *feed*. *NB: Sometimes the combination of two individual IF-channels is also called an interferometer. In that case, its output is a single spectrum.*
- A *telescope* (`\Telescope`) is an entire instrument. It can be a single dish (e.g. GBT) or an aperture synthesis array (e.g. ATCA).
- A *projected* (`\Projected`) angle is an angle projected on the plane perpendicular to the propagation direction (the `z`-axis).

A.2 Labels, sub- and super-scripts

i, j	`\FeedI,\FeedJ`	*feed* labels
a, b	`\RcpA,\RcpB`	*receptor* labels, two per *feed*.
p, q	`\IFP,\IFQ`	*IF-channel* labels,two per *feed*.
r, l	`\RPol,\LPol`	circular polarisation (right, left)
x, y	`\XPol,\YPol`	linear polarisation (N-S, E-W)
$A^{+}, A^{\odot}$	`A\ssLin,A\ssCir`	superscripts for linear and circular polarisation
$A_{\mathsf{i}}, A_{\mathsf{ij}}$	`A\ssI,A\ssIJ`	*feed* subscripts

The subscript convention of matrix elements is as follows: $\mathsf{Y}_{\mathsf{ibp}}$ refers to a matrix element of matrix Y for *feed* i, which models the coupling of the signal going *from receptor* b *to IF-channel* p.

A.3 Coordinate frames

Fig 1 gives an overview of the coordinate system(s) used. All angles on the Sky are measured counter-clockwise, i.e. in the direction North through East. When relevant, 'axis' means 'positive axis' (e.g. the positive x-axis). It is important to make a distinction between:

The beam frame(s): In order to calculate the effects of the primary beam on the signal of a source in direction $\vec{\rho}(\mathsf{l}, \mathsf{m})$, the shape and position of the *voltage* beams of each *receptor* on the Sky has to be calculated. For fully steerable parabolic *antenna*s, which have constant beamshapes, this can be done most conveniently in coordinate frames defined by the projected position angles of the *receptor*s. To allow for the fact that the two beams of a *feed* are closely coupled, an intermediate *feed*-frame is defined also.

The electrical frame: For the polarisation of the signal, the *only* relevant parameters are the *projected* angles w.r.t. the 'electrical' axes x and y defined by the IAU.

NB: In order to see that two frames are needed, consider that Faraday rotation rotates the electric vector, but not the beam on the sky.

Frame of the entire *telescope* (single dish or array):

$\vec{r}$	`\vvAntPos`	Projected *feed* (*receptor*?) position vector
$\mathsf{u}, \mathsf{v}, \mathsf{w}$	`\ccU,\ccV,\ccW`	Projected baseline coordinates
$\vec{u}$	`\vvUVW`	Projected baseline vector $\vec{u}(\mathsf{u}, \mathsf{v}, \mathsf{w})$

Electrical frame on the sky (IAU definition):

x, y	`\ccX,\ccY`	IAU electrical frame on the sky.
z	`\ccZ`	propagation direction of incident field.
γ_{xy}	`\aaXY`	Angle from x-axis to y-axis ($= \pi/2$)
x, y	`\ccXPol,\ccYPol`	linear polarisation coordinates.
r, l	`\ccRPol,\ccLPol`	circular polarisation coordinates.

Sky frame (w.r.t. fringe stopping centre):

$\mathsf{l}, \mathsf{m}, \mathsf{n}$	`\ccL,\ccM,\ccN`	Coordinates (direction cosines)

$\vec{\rho}$	`\vvLMN`	Source direction vector $\vec{\rho}(\mathsf{l},\mathsf{m})$
$\vec{\rho}_{ftc}$	`\vvFTC`	Fringe Tracking Centre $\vec{\rho}_{ftc}(RA, DEC, \mathsf{f})$
$\vec{\rho}_{mc}$	`\vvMC`	Map Centre $\vec{\rho}_{ftc}(\mathsf{l},\mathsf{m})$
γ_{lm}	`\aaLM`	Angle from l-axis to m-axis ($=\pi/2$)
γ_{lx}	`\aaLX`	Angle from l-axis to x-axis ($=\pi/2$)

Coordinate frame of *feed* i, projected on the sky:

$\mathsf{l}_{\mathsf{i}}^{'}, \mathsf{m}_{\mathsf{i}}^{'}$	`\ccLI,\ccMI`	Coordinates
$\mathsf{l}_{\mathsf{i}0}, \mathsf{m}_{\mathsf{i}0}$	`\ccLI0,\ccMI0`	Origin (l,m) of *feed*-frame.
γ_{li}	`\aaLI`	Angle from l-axis to $\mathsf{l}_{\mathsf{i}}^{'}$-axis
γ_{xi}	`\aaXI`	Angle from x-axis to $\mathsf{l}_{\mathsf{i}}^{'}$-axis ($=-\gamma_{\mathsf{lx}}+\gamma_{\mathsf{li}}$)

Coordinate frame of *receptor* a of *feed* i, projected on the sky:

$\mathsf{l}_{\mathsf{ia}}^{''}, \mathsf{m}_{\mathsf{ia}}^{''}$	`\ccLIA,\ccMIA`	Coordinates
$\mathsf{l}_{\mathsf{i\,a}0}^{'}, \mathsf{m}_{\mathsf{i\,a}0}^{'}$	`\ccLIA0,\ccMIA0`	Origin $(\mathsf{l}_{\mathsf{i}}^{'},\mathsf{m}_{\mathsf{i}}^{'})$ of *receptor*-frame.
γ_{ia}	`\aaIA`	Angle from $\mathsf{l}_{\mathsf{i}}^{'}$-axis to $\mathsf{l}_{\mathsf{ia}}^{''}$-axis
γ_{xa}	`\aaXA`	Angle from x-axis to $\mathsf{l}_{\mathsf{ia}}^{''}$-axis ($=-\gamma_{\mathsf{lx}}+\gamma_{\mathsf{li}}+\gamma_{\mathsf{ia}}$)

Coordinate frame of *receptor* b of *feed* i, projected on the sky:

$\mathsf{l}_{\mathsf{ib}}^{''}, \mathsf{m}_{\mathsf{ib}}^{''}$	`\ccLIB,\ccMIB`	Coordinates
$\mathsf{l}_{\mathsf{i\,b}0}^{'}, \mathsf{m}_{\mathsf{i\,b}0}^{'}$	`\ccLIB0,\ccMIB0`	Origin $(\mathsf{l}_{\mathsf{i}}^{'},\mathsf{m}_{\mathsf{i}}^{'})$ of *receptor*-frame.
γ_{ib}	`\aaIB`	Angle from $\mathsf{l}_{\mathsf{i}}^{'}$-axis to $\mathsf{l}_{\mathsf{ib}}^{''}$-axis
γ_{yb}	`\aaYB`	Angle from y-axis (!) to $\mathsf{l}_{\mathsf{ib}}^{''}$-axis ($=-\gamma_{\mathsf{xy}}-\gamma_{\mathsf{lx}}+\gamma_{\mathsf{li}}+\gamma_{\mathsf{ib}}$)

The coordinates $\mathsf{l}_{\mathsf{ia}}^{''}, \mathsf{m}_{\mathsf{ia}}^{''}$ and $\mathsf{l}_{\mathsf{ib}}^{''}, \mathsf{m}_{\mathsf{ib}}^{''}$ of the frames of *receptor*s a and b in equ 37 are related to the celestial coordinate frame l, m in a two-step process. First we define an intermediate *feed*-frame $\mathsf{l}_{\mathsf{i}}^{'}, \mathsf{m}_{\mathsf{i}}^{'}$ for *feed* i, projected on the Sky:

$$\begin{pmatrix} \mathsf{l}_{\mathsf{i}}^{'} \\ \mathsf{m}_{\mathsf{i}}^{'} \end{pmatrix} = \mathsf{Rot}(\gamma_{\mathsf{li}}) \begin{pmatrix} \mathsf{l}-\mathsf{l}_{\mathsf{i}0} \\ \mathsf{m}-\mathsf{m}_{\mathsf{i}0} \end{pmatrix} \tag{55}$$

in which $(\mathsf{l}_{\mathsf{i}0},\mathsf{m}_{\mathsf{i}0})$ is the Pointing Centre of *feed* i, and $\mathsf{Rot}(\gamma_{\mathsf{li}})$ is a rotation over the *projected* angle γ_{li} between the positive l-axis of the Sky frame and the $\mathsf{l}_{\mathsf{i}}^{'}$-axis of the *feed*-frame.

The voltage beams themselves are best modelled in a *receptor*-frame (see equ 37), again projected on the Sky. For *receptor* a we have:

$$\begin{pmatrix} \mathsf{l}_{\mathsf{ia}}^{''} \\ \mathsf{m}_{\mathsf{ia}}^{''} \end{pmatrix} = \mathsf{Rot}(\gamma_{\mathsf{ia}}) \begin{pmatrix} \mathsf{l}_{\mathsf{i}}^{'}-\mathsf{l}_{\mathsf{i\,a}0}^{'} \\ \mathsf{m}_{\mathsf{i}}^{'}-\mathsf{m}_{\mathsf{i\,a}0}^{'} \end{pmatrix} \tag{56}$$

The matrix $\mathsf{Rot}(\gamma_{\mathsf{ia}})$ represents a rotation over the angle γ_{ia} between the positive $\mathsf{l}_{\mathsf{i}}^{'}$-axis of the *feed*-frame and the $\mathsf{l}_{\mathsf{ia}}^{''}$-axis of the relevant *receptor*-frame. For *receptor* b:

$$\begin{pmatrix} \mathsf{l}_{\mathsf{ib}}^{''} \\ \mathsf{m}_{\mathsf{ib}}^{''} \end{pmatrix} = \mathsf{Rot}(\gamma_{\mathsf{ib}}) \begin{pmatrix} \mathsf{l}_{\mathsf{i}}^{'}-\mathsf{l}_{\mathsf{i\,b}0}^{'} \\ \mathsf{m}_{\mathsf{i}}^{'}-\mathsf{m}_{\mathsf{i\,b}0}^{'} \end{pmatrix} \tag{57}$$

Figure 1: Label: fig-coord File: fig-coord.eps

A (rather crowded) overview of the various coordinate frames for the Measurement Equation. See also the text. The origin of the Sky frame (l, m) *is defined by the* fringe stopping centre. *The origin of the* feed*-frame* $(\mathsf{l}_\mathsf{i}^{'}, \mathsf{m}_\mathsf{i}^{'})$ *is defined by the pointing centre of* feed i. *The 'pointing centres' of the voltage beams of* receptors a *and* b *(marked with a and b) define the origins of the* receptor*-frames* $(\mathsf{l}_\mathsf{ia}^{''}, \mathsf{m}_\mathsf{ia}^{''})$ *and* $(\mathsf{l}_\mathsf{ib}^{''}, \mathsf{m}_\mathsf{ib}^{''})$. *The shapes and position offsets of these voltage beams are exaggerated, in order to emphasise that they do not necessarily coincide.*

$(\mathsf{l}'_{\mathsf{i\,a0}}, \mathsf{m}'_{\mathsf{i\,a0}})$ and $(\mathsf{l}'_{\mathsf{i\,b0}}, \mathsf{m}'_{\mathsf{i\,b0}})$ represent pointing offsets of *receptor* a and b respectively. These can be used to model 'beam-squint' of *feeds* that are not axially symmetric.

A.4 Matrices and vectors

The following matrices and vectors play a role in the Measurement Equation:

$\vec{I}$	`\vvIQUV`	Stokes vector of the source (I,Q,U,V).
$\vec{V}, \mathsf{v}$	`\vvCoh,\vvCohEl`	Coherency vector, and one of its elements.
S	`\mmStokes`	Stokes matrix, conversion between polarisation representations.
S^{+}	`\mmStokes\ssLin`	Conversion to linear representation.
$\mathsf{S}^{\odot}$	`\mmStokes\ssCir`	Conversion to circular representation.
$\mathcal{M}$	`\mmMueller`	Mueller matrix: Stokes to Stokes through optical 'element'
X, x	`\mmXifr,\mmXifrEl`	Correlator matrix (4×4).
M, m	`\mmMifr,\mmMifrEl`	Multiplicative *interferometer*-based gain matrix (4×4).
$\vec{A}, \mathsf{a}$	`\vvAifr,\vvAifrEl`	Additive *interferometer*-based gain vector.

The following *feed*-based Jones matrices (2×2) have a well-defined meaning:

J, j	`\mjJones,\mjJonesEl`	Jones matrix, and one of its elements.
F, f	`\mjFrot,\mjFrotEl`	Faraday rotation (of the plane of linear pol.)
T, t	`\mjTrop,\mjTropEl`	Atmospheric gain (refraction, extinction).
P, p	`\mjProj,\mjProjEl`	Projected *receptor* angle(s) w.r.t. x, y frame
B, b	`\mjBtot,\mjBtotEl`	Total *feed* voltage pattern (i.e. $\mathsf{B} = \mathsf{D}\ \mathsf{E}\ \mathsf{P}$.
E, e	`\mjBeam,\mjBeamEl`	Traditional *feed* voltage beam.
C, c	`\mjConf,\mjConfEl`	Feed configuration matrix (...).
D, d	`\mjDrcp,\mjDrcpEl`	Leakage between *receptors* a and b.
H, h	`\mjHybr,\mjHybrEl`	Hybrid network, to convert to circular pol.
G, g	`\mjGrec,\mjGrecEl`	*feed*-based electronic gain.
K, k	`\mjKern,\mjKernEl`	Fourier Transform Kernel (baseline phase weight)
$\mathsf{K}^0, \mathsf{k}^0$	`\mjKref,\mjKrefEl`	FT kernel for the fringe-stopping centre.
K', k'	`\mjKoff,\mjKoffEl`	FT kernel relative to the fringe-stopping centre.
Q, q	`\mjQsum,\mjQsumEl`	Electronic gain of tied-array *feed* after summing.

Some special matrices and vectors:

Zero	`\mmZero`	Zero matrix
$\vec{0}$	`\vvZero`	Zero vector
$\mathcal{U}$	`\mmUnit`	Unit matrix
$\mathsf{Diag}(a, b)$	`\mjDiag`	Diagonal matrix with elements a, b
$\mathsf{Mult}(a)$	`\mjMult`	Multiplication with factor a
$\mathsf{Rot}(\alpha[, \beta])$	`\mjRot`	[pseudo] Rotation over an angle α, β
$\mathsf{Ell}(\alpha[, \beta])$	`\mjEll`	Ellipticity angle[s] α, β

$\mathcal{H}$	`\mjLtoC`	Signal conversion from linear to circular.
$\mathcal{H}^{-1}$	`\mjCtoL`	Signal conversion from circular to linear.

Definitions of some special matrices:

$$\mathsf{Diag}(a,b) \equiv \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \qquad \mathsf{Diag}(a,a) = \mathsf{Mult}(a) = a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{58}$$

A 'pure' rotation $\mathsf{Rot}(\alpha)$ is a special case of a 'pseudo rotation' $\mathsf{Rot}(\alpha,\beta)$:

$$\mathsf{Rot}(\alpha,\beta) \equiv \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\beta & \cos\beta \end{pmatrix} \qquad \mathsf{Rot}(\alpha) \equiv \mathsf{Rot}(\alpha,\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \tag{59}$$

Ellipticity:

$$\mathsf{Ell}(\alpha,\beta) \equiv \begin{pmatrix} \cos\alpha & i\sin\alpha \\ -i\sin\beta & \cos\beta \end{pmatrix} \qquad \mathsf{Ell}(\alpha) \equiv \mathsf{Ell}(\alpha,-\alpha) = \begin{pmatrix} \cos\alpha & i\sin\alpha \\ i\sin\alpha & \cos\alpha \end{pmatrix} \tag{60}$$

A.5 Miscellaneous parameters

β	`\ppParall`	Parallactic angle, form North pole to zenith
HA	`\ppHA`	Hour Angle
RA	`\ppRA`	Right Ascension
DEC	`\ppDEC`	Declination
LAT	`\ppLAT`	Latitude on Earth
t	`\ccT`	Time
f	`\ccF`	Frequency
χ	`\ppFarad`	Faraday rotation angle
a	`\ppAmpl`	Amplitude
ψ	`\ppPhase`	Phase
ζ	`\ppPhaseZero`	Phase zero
ϕ	`\ppRcpPosDev`	Dipole position angle error
θ	`\ppRcpEllDev`	*receptor* ellipticity