Some practical aspects of the matrix-based MEASUREMENT EQUATION of a generic radio telescope AIPS++ implementation note nr 182

J.E.Noordam

May 29, 2015

File: /aips++/nfra/matform.latex

Abstract: This paper describes the full Measurement Equation of a generic radio telescope in matrix form, covering both uv-plane and image-plane effects. It includes a 'catalog' of practical examples of its application to Consortium telescopes (WSRT, AT, VLA, VLBI, GBT, SKAI etc). The purpose is to pave the way towards making this formalism the basis for calibration and imaging in AIPS++. There is also an outline of a 'generalised Solver' for parameters of the Measurement Model and the Sky Model.

Contents

1 INTRODUCTION

Over the last few years, various people have suggested that the Measurement Equation of a radio telescope (which includes interferometers and single dishes) can best be described in the form of matrices. In 1993, Bregman [Bregman93] and later Hjellming [Hjellming93] suggested that such a 'universal' formalism is essential for AIPS++, which is supposed to deal with all the Consortium telescopes, both by themselves and in combination with each other (e.g. VLBI). Thus, a good formalism should be adopted from the start. They provided an initial formulation based on 2×2 instrumental response matrices, and a 2×2 coherency matrix. Later that year, Hamaker realised that the 'direct matrix product' provided an elegant way to link a 4-dimensional coherence response matrix of an interferometer with the 2×2 response matrices of a single antenna. This also opened the way to a uniform treatment of radio and optical polarimetry, where Stokes and coherence 4-vectors are related by such interferometer matrices. For a full account, see the two papers by Hamaker et al [Hamaker95], and Sault et al [Sault95].

Since the work mentioned above limits itself to uv-domain effects, while ignoring image-plane effects, it is strictly speaking only valid for the case of a point source in the centre of the field. However, it is a good approximation for the many actual cases where a compact source dominates the field. It correctly describes most of the important differences between the various Consortium telescopes, like alt-az and equatorial mounts, and linearly and circularly polarised feeds. Therefore, this *uv-domain Measurement Equation* is used in section 3 to demonstrate the generic nature of the formalism by giving a 'catalog' of practical applications to existing telescopes.

The *full Measurement Equation*, which includes image-plane effects, was developed by Bregman during an AIPS++ workshop in Dwingeloo in June 1995. Section 4 gives a first description of it for the benefit of AIPS++ developers. It will be published in the open literature next year [Bregman96].

It should be emphasised that some of the matrix elements in this paper may not have reached their final form (particularly in the description of the antenna beam). But the overall structure of the formalism seems solid enough for AIPS++ to start building on. Some implementation issues are discussed. There is also an outline of a 'universal' Solver for parameters of the Measurement Model and Sky Model.

2 THE CASE OF A CENTRAL POINT SOURCE

For a given interferometer, the measured visibilities *for a point source in the centre of the field!* can be described by a 4-element 'coherency vector' $\vec{v}$, which is related to the so-called 'Stokes visibility vector' $\vec{s}_{ij} = (I, Q, U, V)^{\mathrm{T}}_{ij}$ of the observed source by a matrix operation,

$$\vec{v}_{ij} = \begin{pmatrix} v_{\mathrm{pp}} \\ v_{\mathrm{pq}} \\ v_{\mathrm{qp}} \\ v_{\mathrm{qq}} \end{pmatrix}_{ij} = \vec{M}_{ij}\ (\vec{J}_i \otimes \vec{J}_j^*)\ \vec{S} \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix}_{ij} + \vec{A}_{ij} \ \approx\ (\vec{J}_i \otimes \vec{J}_j^*)\ \vec{S} \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix}_{ij} \tag{1}$$

Here subscripts p and q represent the two polarisation channels measured by each antenna. (NB: They are named X and Y for WSRT and ATCA, and R and L for the VLA). The subscripts i and j represent the antenna numbers, and $\otimes$ represents the matrix *direct* product (also called the tensor product, or Kronecker product).

The Stokes visibility vector depends on the brightness distribution and on the length and orientation of baseline ij. The 4×4 matrix $\vec{S}$ converts it into a coherency vector (ignoring instrumental effects). $\vec{S}$ is unitary, except for a normalising constant: $\vec{S}^{-1} = 2\vec{S}^{*T}$. It cannot be decomposed into antenna-based parts.

$$\vec{S} = \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & i \\ 0 & 0 & 1 & -i \\ 1 & -1 & 0 & 0 \end{pmatrix} \tag{2}$$

The 4-element vector $\vec{A}_{ij}$ represents *additive* interferometer-based effects. Examples are receiver noise, and correlator offsets. The 4×4 diagonal matrix $\vec{M}_{ij}$ represents *multiplicative* interferometer-based effects, which cannot be factored into antenna-based contributions. Examples are decorrelations, which usually give diagonal matrices with identical elements. Fortunately, the elements of $\vec{A}_{ij}$ and $\vec{M}_{ij}$ tend to be small or close to unity in practice, and will be ignored here.

Thus it is assumed that, in the case of a central point source, all instrumental effects can be factorised into antenna-based contributions. The 4×4 interferometer response matrix $\vec{J}_i \otimes \vec{J}_j^*$ then consists of a *direct matrix product* of two 2×2 antenna-based response matrices.[1] The reader will note that this is the polarimetric generalisation of the familiar 'Selfcal assumption'.

This antenna response matrix $\vec{J}_i$ can be decomposed into a product of matrices, each of which models a specific instrumental effect in the signal path.[2]

$$\vec{J}_i \;=\; \vec{G}_i\; \vec{D}_i\; \vec{C}_i\; \vec{B}_i\; \vec{P}_i\; \vec{F}_i \tag{3}$$

in which,

- $\vec{G}_i$ represents complex gain (amplitude, phase) per IF channel. It includes the effects of all the electronics after the feed (amplifiers, mixers, LO, cables etc), excluding the correlator of course. It is usually also assumed to includes atmospheric gain effects, even though $\vec{G}_i$ does not necessarily commute with $\vec{D}_i$ (see below).

$$\vec{G}_i = \begin{pmatrix} g_{i\mathrm{p}} & 0 \\ 0 & g_{i\mathrm{q}} \end{pmatrix} \tag{4}$$

  NB: It is wrong to identify IF-channels with receptors, like the X and Y dipoles at WSRT or ATCA. In a VLA antenna, the circularly polarised IF-signals are a combination of the signals of the two linear dipoles. Another example: in the ATCA and the new WSRT frontends, the signals from the dipoles may be rapidly switched between IF-channels, for calibration purposes.

- $\vec{D}_i$ represents the signal 'leakage' between the two receptors in a feed, due to deviations from nominal. The real part corresponds roughly to a dipole alignment error, and the imaginary part to ellipticity.

$$\vec{D}_i = \begin{pmatrix} 1 & d_{i\mathrm{p}} \\ -d_{i\mathrm{q}} & 1 \end{pmatrix} \tag{5}$$

  NB: Other processes may also contribute to leakage. For example, cross-talk between the two IF-channels can be described by adding a non-zero mutual coupling factor c_i to the off-diagonal terms.

- $\vec{C}_i$ represens the nominal feed configuration, which includes rotation of the dipoles w.r.t. the antenna (e.g. WSRT) , and conversion from linear to circular polarisation (e.g. VLA). See section 3 for examples.

[1] The 4×4 interferometer response matrices, which describe the transmission of a Stokes vector through an optical element, are called 'Mueller' matrices in the optical polarisation literature. The 2×2 antenna-based response matrices, which describe the transmission of the instantaneous vector *amplitude* through an optical element, are called 'Jones' matrices. Mueller matrices cannot always be factorised into Jones matrices.

[2] For completeness it should be noted that one of the properties of the direct matrix product is that: $(\vec{J}_i \otimes \vec{J}_j^*) = (\vec{G}_i \otimes \vec{G}_j^*)\ (\vec{D}_i \otimes \vec{D}_j^*)\ (\vec{C}_i \otimes \vec{C}_j^*)\ (\vec{P}_i \otimes \vec{P}_j^*)\ (\vec{F}_i \otimes \vec{F}_j^*)$. However, it seems better to work with 2×2 matrices as much as possible.

- $\vec{B}_i$ represents the antenna beam response. For the important case of a dominating compact source in the centre of the field (and an axially symmetric antenna-feed combination), it is the unit matrix. See also section 4.
- $\vec{P}_i$ represents parallactic angle ϕ, i.e. rotation of the antenna w.r.t. the sky. For an *equatorial* antenna like the WSRT, it is a unit matrix. For an *alt-az* antenna, ϕ varies smoothly with Hour Angle, as a function of Latitude and Declination: ($\phi(t) = \arctan[\cos LAT \sin HA \cos DEC \;-\; \sin LAT \cos HA \sin DEC]$ (or something like that):

$$\vec{P}_i^{alt-az} = \begin{pmatrix} \cos\phi(t) & \sin\phi(t) \\ -\sin\phi(t) & \cos\phi(t) \end{pmatrix} \tag{6}$$

- $\vec{F}_i$ represents (ionospheric) Faraday rotation.

$$\vec{F}_i = \begin{pmatrix} \cos\psi & \sin\psi \\ -\sin\psi & \cos\psi \end{pmatrix} \tag{7}$$

It should be noted that these matrices generally do not commute, so that their order is important! For instance, a diagonal matrix like $\vec{G}_i$ does *not* commute with a matrix with non-zero elements off the diagonal, like $\vec{D}_i$. Thus, one should be a little careful in sweeping *all* complex gain effects into a single matrix $\vec{G}_i$, irrespective of where they occur in the signal path.

Matrices $\vec{G}_i$,$\vec{P}_i$ and $\vec{F}_i$ are Cartesian coordinate transforms, and thus unitary: $\vec{A}^{-1} = \vec{A}^{*T}$. Matrix $\vec{C}_i$ is not unitary if it represents a conversion from linear to circular polarisation (..?). *Do they commute??*

3 APPLICATION TO EXISTING TELESCOPES

To allow the reader to get some intuitive 'feeling' for the matrix formalism, it will be made explicit in this section for some practical cases. This is done for the simplified case of a central point source (rather than for the full Measurement Equation described in section 4), because:

- the uv-domain Measurement Equation clearly shows how the main differences between the various telescopes can be described generically.
- the case of a dominating compact source in the centre of the field is an important one in practice, for instance for calibration.

It should be noted that the formalism makes it very easy to calculate the response of an interferometer that consists of two antenna's which can each be described by an antenna-based response matrix $\vec{J}_i$, *even if the constituent antenna's are quite different*. This is particularly important for VLBI, with its collection of often quite dissimilar antenna's. Another potentially important example is the integration of one or two prototype antenna's for the Square Km Array (SKAI) in the WSRT. An example of a type of antenna that cannot be easily described by its own $\vec{J}_i$ seems to be a tied array (see section 3.2).

3.1 WSRT, VLA, ATCA, GMRT

An (equatorial) **WSRT antenna**, with its linear dipoles rotated w.r.t. the antenna over α *degrees*. Special cases are 'parallel' (+) dipoles, i.e. rotated over $\alpha = 0$, and 'crossed' (X) dipoles, i.e. rotated over $\alpha = -45$.

$$\vec{J}_i^{WSRT} = \vec{G}_i\ \vec{D}_i \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \vec{F}_i \tag{8}$$

The resulting WSRT interferometer response matrix is of course equivalent to the equations derived by Weiler [Weiler72]. In the new WSRT frontends (due in 1997/8), the dipoles cannot be rotated anymore, so they will always be 'parallel' ($\alpha = 0$):

$$\vec{J}_i^{WSRT+} = \vec{G}_i \; \vec{D}_i \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \vec{F}_i \tag{9}$$

An (alt-az) **ATCA antenna**, with fixed linear dipoles. The parallactic angle in matrix $\vec{P}_i^{alt-az}$ (see equ 6) varies smoothly with Hour-Angle (time):

$$\vec{J}_i^{ATCA} = \vec{G}_i \; \vec{D}_i \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \vec{P}_i^{alt-az} \; \vec{F}_i \tag{10}$$

An (alt-az) **VLA antenna**, with circularly polarised receptors. The C-matrix descibes the transformation of the signals from the fixed linear dipoles into Right and Left circularly polarised R and L signals, by means of a 'hybrid'. Any instrumental (e.g. phase) effects in this hybrid could be modelled too. The matrix $\vec{B}_i$ is *not* a unit matrix, even for a central point source, because the system is not axially symmetric and produces some instrumental polarisation even in the centre of the field.

$$\vec{J}_i^{VLA} = \vec{G}_i \; \vec{D}_i \; \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \vec{B}_i \; \vec{P}_i^{alt-az} \; \vec{F}_i \tag{11}$$

For a central point source (i.e. when the beam $\vec{B}_i$ reduces to a unit matrix), an (alt-az) **GMRT** antenna is mathematically equivalent to an ATCA antenna (see expression 10 above):

$$\vec{J}_i^{GMRT} = \vec{J}_i^{ATCA} \tag{12}$$

The same is true for the (alt-az) **single dish** Effelsberg radio telescope, with fixed linear dipoles:

$$\vec{J}_i^{Effelsberg} = \vec{J}_i^{ATCA} \tag{13}$$

The **GBT** is not axially symmetric, so that the beam $\vec{B}_i$ has to be taken into account even for a central point source.

3.2 More exotic cases

A **dipole array**, with North-South and East-West linear dipoles in the horizontal plane. This is one of the possible concepts for an element of the Square Km Array (SKAI) planned by NFRA. *It is assumed that the dipole array is internally calibrated, i.e. that it behaves like a single antenna.* The parallactic rotation of the EW dipoles w.r.t. the sky differs from that of the NS dipoles: $\phi_{EW}(t) = \arctan(\tan HA \sin DEC)$, and $\phi_{NS}(t) = f(HA, DEC, latitude) \cdots ?$. The beam properties described by $\vec{B}_i$ change as a function of the same parameters.

$$\vec{J}_i^{SKAI} = \vec{G}_i \; \vec{D}_i \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \vec{B}_i(t) \begin{pmatrix} \cos\phi_{NS}(t) & \sin\phi_{NS}(t) \\ -\sin\phi_{EW}(t) & \cos\phi_{EW}(t) \end{pmatrix} \vec{F}_i \tag{14}$$

The description of the **Arecibo** antenna resembles that of a dipole array $\cdots$:

$$\vec{J}_i^{Arecibo} = \vec{G}_i \; \vec{D}_i \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \vec{B}_i(t) \begin{pmatrix} \cos\phi_{NS}(t) & \sin\phi_{NS}(t) \\ -\sin\phi_{EW}(t) & \cos\phi_{EW}(t) \end{pmatrix} \vec{F}_i \tag{15}$$

A free-floating **space VLBI** antenna can be rotated over an arbitrary parallactic angle ϕ, and does not suffer from ionospheric Faraday rotation. Assuming that it has fixed linear dipoles and is axially symmetric, we get:

$$\vec{J}_i^{Space} = \vec{G}_i \, \vec{D}_i \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{16}$$

A **tied array**, in which the signals of a number of identical WSRT (+) antenna's are added coherently. This case is important for VLBI, where arrays like WSRT and VLA are used as tied arrays. It does not seem possible to describe such a system with a single Jones antenna-based matrix $\vec{J}_i$. Therefore, we write down the *interferometer response* equation of an interferometer between two tied arrays. The indices i run over the antenna's that make up the first tied array, and the j run over antenna's of the second. Thus, the overall output is the sum of all the constituetnt interferometers:

$$\begin{pmatrix} v_{\mathrm{pp}} \\ v_{\mathrm{pq}} \\ v_{\mathrm{qp}} \\ v_{\mathrm{qq}} \end{pmatrix}_{ij} = \sum_{i=1}^{n} \sum_{j=1}^{m} (\vec{J}_i \otimes \vec{J}_j^*) \, \vec{S} \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix}_{ij} \tag{17}$$

4 THE FULL MEASUREMENT EQUATION

The formalism discussed hitherto has been called the *uv-domain Measurement Equation* because it ignores image-plane effects, i.e. instrumental effects that depend on the position in the field. This is only valid for a dominating compact source in the centre of the field, and then only in the case of axially symmetric systems. Only then does the matrix $\vec{B}_i$, which describes effects like primary beam and instrumental polarisation, reduce to the unit matrix.

In this section, the formalism will be generalised to the *full Measurement Equation*, which describes the behaviour of a real instrument observing an arbitrary brightness distribution, and thus includes image-plane effects. For k (point) sources within the primary beam, we get:

$$\vec{v}_{ij} = \begin{pmatrix} v_{\mathrm{pp}} \\ v_{\mathrm{pq}} \\ v_{\mathrm{qp}} \\ v_{\mathrm{qq}} \end{pmatrix}_{ij} = (\vec{G}_i \otimes \vec{G}_j^*) \, (\vec{D}_i \otimes \vec{D}_j^*) \, (\vec{C}_i \otimes \vec{C}_j^*) \, \Sigma \tag{18}$$

in which the matrices $\vec{G}$ (complex gain), $\vec{D}$ (leakage) and $\vec{C}$ (nominal feed configuration) are typical uv-domain effects in the sense that they affect all the sources in the same way. The matrix Σ describes image-plane effects, which depend on the position ρ_k in the field, and therefore have to be applied to each source independently:

$$\Sigma = \sum_k \sum_t \sum_f (\vec{B}_i \otimes \vec{B}_j^*) \, (\vec{P}_i \otimes \vec{P}_j^*) \, (\vec{F}_i \otimes \vec{F}_j^*) \, \vec{S} \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix}_k \exp^{-i \, \vec{u}_k \cdot \vec{\rho}_k} \tag{19}$$

in which the sum is taken over all k sources in the field. The sums over time t and frequency f take care of the effects of finite integration time and bandwidth. The effects of integrating over a finite antenna size are indistinguishable from the primary beam, and are thus part of $\vec{B}$ (see below).

Note that the matrices $\vec{P}$ for parallactic angle and $\vec{F}$ for ionospheric Faraday rotation are now behind the sum, even though they are (usually) the same for all sources in the field. This is because they do not commute with $\vec{B}$. The same is true for tropospheric gain effects (e.g. extinction and refraction), which should not really be lumped with the receiver gain $\vec{G}$. *It is clear that current calibration practices could do with some critical scrutiny!*

The matrix $\vec{B}_i$ describes the primary beam response (including pointing errors!) of antenna i. The matrix elements are functions of the difference vector $\Delta\vec{\rho_k}$ between the direction vector $\vec{\rho}_i$ of the optical axis of antenna i, and the direction vector $\vec{\rho_k}$ of a point source k,

$$\vec{B}_i = \vec{B}_i(\vec{\rho}_i - \vec{\rho}_k) = \cdots \tag{20}$$

Instrumental polarisation is an artifact caused by asymmetries in $\vec{B}_i$, and thus in $\vec{B}_i \otimes \vec{B}_j^*$. Ionospheric Faraday rotation does *not* affect it, of course. There is a growing consensus that at least two effects are involved: the linear shape of the dipoles, which cause the familiar 'clover-leaf' patterns in Q and U, and a standing wave in the front-end support legs. The latter effect is strongly frequency-dependent. *NB: Expressions for the elements of $\vec{B}_i$ for the WSRT will emerge over the summer.*

Figure 1: Label: matform File: matform.eps

Schematic overview of the full Measurement Model (MM), and its place in the overall AIPS++ scheme. Note that the MM models the actual *instrument, including errors. The MM can be regarded as a matrix $A(p)$, which contains instrumental parameters (p), and which transforms the Sky Model 'vector' $X(q)$ into a 'uv-model' $A(p)X(q)$. It should be noted that the inverse operation is not possible in general. If the Sky Model is complete, and the MM parameters are known, the uv-model should be equal to the measured data 'vector' M. If not, the difference vector $\mid M - A(p).X(q) \mid$ is the input to a Solver which estimates improved values for the parameter sets p and q. The convenient mathematical form of the matrix formalism should make it possible to design a generalised Solver, which can solve for any subset of parameters, given sufficient constraints.*

5 IMPLEMENTATION IN AIPS++

Figure 4 gives a schematic overview of how the formalism described here could be implemented in AIPS++. Obviously, the OO experts will determine what actually happens, but perhaps the following points are worth making:

- The Measurement Equation is enshrined in the Measurement Model (MM), which now described the *actual* instrument, including corrections. This is different from the Green Bank model, where the MM described an idealised instrument (and we did not know what to do about image-plane effects). I have no idea how the MM should be implemented, and *whether it should contain the Correctors.*
- The MM does not really have an inverse. All instrumental effects can be applied by corrupting the Sky Model (SM) and the uv-model, but generally not the other way around. Calibration consists of fiddling the parameters of the MM and the SM until the uv-model is equal to the measured uv-data. Any non-zero differences are inputs to a Generalised Solver (see below), which estimates improved values for instrumental *and Sky Model* parameters.
- We do not (yet) need a Source Model. The Sky Model projection will do for the moment.
- I would be interested in a clear idea about the precise relation between MM and Measurement Set (MS). Upon reflection, it does not seem such a good idea to put uv-data from different telescopes (or even observations) into a single MS. Separate MS's would make it easier to associate an MM with an MS. But of course it should be possible to include a collection of MS's into a calibration and/or imaging process.

6 A GENERALISED SOLVER

The elegance of the matrix formalism opens the way for a generalised non-linear *complex* Solver for the parameters of the Measurement Model (MM) and the Sky Model (SM). Such a solver can be used for any telescope that can be described with the formalism.

The full Measurement Equation expressed by 18 and 19 can be written as:

$$\begin{pmatrix} v_{\rm pp} \\ v_{\rm pq} \\ v_{\rm qp} \\ v_{\rm qq} \end{pmatrix} = \vec{H} \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} \tag{21}$$

The matrix coefficients h_{ij} of the 'system matrix' $\vec{H}$ are functions $h_{ij}(x_1, x_2, x_3, ..., x_n)$ of the (complex) variables x_k. In general, these functions will be non-linear. For small variations Δx_k:

$$\Delta h_{ij} = \sum_{k=1}^{n} \frac{\partial h_{ij}}{\partial x_k} \Delta x_k \tag{22}$$

Assuming that our Sky Model is correct, the uv-data are the result of multiplication with the *true* values of the h_{ij}:

$$v_{\rm pp}^{dat} = h_{11}^{true} I^{mod} + h_{12}^{true} Q^{mod} + h_{13}^{true} U^{mod} + h_{14}^{true} V^{mod} \tag{23}$$

The uv-model values have been calculated from the Sky Model by multiplying them with the *current* values of the h_{ij}, i.e. the values that are obtained by using the best available values of the variables x_k:

$$v_{\rm pp}^{mod} \;=\; h_{11}^{curr}\, I^{mod} \;+\; h_{12}^{curr}\, Q^{mod} \;+\; h_{13}^{curr}\, U^{mod} \;+\; h_{14}^{curr}\, V^{mod} \tag{24}$$

So, a 'Selfcal equation' can be written in terms of a linear combination of the increments Δx_k, which would move the current values of the variables x_k closer to the true values:

$$v_{\rm pp}^{dat} \;-\; v_{\rm pp}^{mod} \;=\; (h_{11}^{true} - h_{11}^{curr})\, I^{mod} \;+\; \cdots \;=\; \sum_{k=1}^{n} c_k \Delta x_k \tag{25}$$

in which,

$$c_k \;=\; \frac{\partial h_{11}}{\partial x_k}\, I^{mod} \;+\; \frac{\partial h_{12}}{\partial x_k}\, Q^{mod} \;+\; \frac{\partial h_{13}}{\partial x_k}\, U^{mod} \;+\; \frac{\partial h_{14}}{\partial x_k}\, V^{mod} \tag{26}$$

The 'sensitivity' of h_{ij} for a variation in a variable x_k can be approximated by calculating the change in h_{ij} as result of a small 'trial' variation δx_k:

$$\frac{\partial h_{ij}}{\partial x_k} \;\approx\; \frac{h_{ij}(x_1, x_2, \cdots, x_k + \delta x_k, \cdots, x_n)}{h_{ij}(x_1, x_2, \cdots, x_k, \cdots, x_n)} \tag{27}$$

These sensitivities will be different for different telescopes, due to the different values of the elements of the instrumental matrices etc. But given these matrices, they can be readily calculated.

The procedure can be further generalised to the case where better values have to be estimated for the parameters of the Sky Model.

The solving **procedure** would be as follows:

1. Make sure that the parametrised instrumental matrices G,D,C,B,P,F and a parametrised Sky Model (SM) are available.
2. Choose which of the parameters are *variables* to be solved, and which are assumed to be known values. The latter may also be expressions, like the parallactic angle.
3. Supply extra constraints on the variables, in the form of extra equations. Example: solve for phase gradients over the array, rather than individual antenna phases.
4. Feed it all to a *symbolic processor* which sets up the solution in an efficient way. It is important that most of the overhead that is caused by the generality goes into setting up the solver, and is not repeated for each MS row.
5. Solve for the specified variables x_k, or rather their increments Δx_k. NB: The Solver should utilise a least-squares method (like SVD) that can deal automatically with situations in which there are too few constraints for a solution.
6. After each solution, the estimated values Δx_k are *added* to the values of the x_k in the relevant Correctors. These are used to create the starting point of new solutions.
7. Iterate until some criterion is met. Iteration can also mean that one the Solver alternates between two sets of variables, or that an improved Sky Model is obtained inside the loop.

References

[Weiler72] **K.W.Weiler:** *The Synthesis Radio Telescope at Westerbork: Methods of Polarization Measurement.* **Astronomy and Astrophysics, 1973, 26, 404-407**

[Bregman93] **J.D.Bregman, J.E.Noordam:** *A matrix formalism for AIPS++* **Submitted to the AIPS++ project, May 1993**

[**Hjellming93**] **R.M.Hjellming:** *A proposal for a Framework for Mathematical Objects and Methods.* **Submitted to the AIPS++ project, November 1993**

[**Cornwell93**] **T.J.Cornwell:** *Some suggestions for the AIPS++ Imaging Model.* **Submitted to the AIPS++ project, May 1993**

[**Hamaker95**] **J.P.Hamaker, J.D.Bregman, R.J. Sault:** *Understanding Radio Polarimetry I: Mathematical foundations.* **Submitted to Astronomy and Astrophysics, July 1995.**

[**Sault95**] **R.J.Sault, J.P.Hamaker, J.D.Bregman:** *Understanding Radio Polarimetry II: Instrumental calibration of an interferometer array.* **Submitted to Astronomy and Astrophysics, July 1995.**

[**Bregman96**] **J.D.Bregman et al:** *The full Measurement Equation of a generic radio telescope.* **In preparation.**