Chapter 5

Sensitivity and Calibration for Interferometers

Jayaram N. Chengalur

5.1 Sensitivity

As we discussed earlier, an aperture synthesis telescope can be regarded as a collection of two element interferometers. Hence, for understanding the sensitivity of such a telescope, it is easier to first start with the case of a two element interferometer. Consider such an interferometer composed of two antennas i, j, (of identical gains, but possibly different system temperatures), looking at a point source of flux density S. We assume that the point source is at the phase center[1] and hence that in the absence of noise the visibility phase is zero. Let the individual antenna gains[2] be G and system temperatures be T_{s_i} and T_{s_j}. If $n_i(t)$ and $n_j(t)$ are the noise voltages of antennas i and j respectively,then $\sigma_i^2 = \left\langle n_i^2(t) \right\rangle = \mathrm{T}_{\mathrm{s_i}}$, and $\sigma_j^2 = \left\langle n_j^2(t) \right\rangle = \mathrm{T}_{\mathrm{s_j}}$. Similarly if $v_i(t)$ and $v_j(t)$ are the voltages induced by the incoming radiation from the point source, $\left\langle v_i^2(t) \right\rangle = \left\langle v_j^2(t) \right\rangle = \mathrm{GS}$. The instantaneous correlator[3] output is given by:

$$r_{ij}(t) = (v_i(t) + n_i(t))\,(v_j(t) + n_j(t))$$

The mean[4] of the correlator output is hence:

$$\begin{aligned} \langle r_{ij}(t) \rangle &= \langle (v_i(t) + n_i(t))\,(v_j(t) + n_j(t)) \rangle \\ &= \langle v_i(t) v_j(t) \rangle \\ &= \mathrm{GS} \end{aligned} \tag{5.1.1}$$

where we have assumed that the noise voltages of the two antennas are not correlated, and also of course that the signal voltages are not correlated with the noise voltages. $r_{ij}(t)$ is hence an unbiased estimator of the true visibility.

To determine the noise in the correlator output, we would need to compute the rms of $r_{ij}(t)$ for which we need to be able to work out:

---

[1] See Chapter 4.

[2] Here the gain is taken to be in units of Kelvin per Jansky of flux in the matched polarization

[3] Here we are dealing with an ordinary correlator, not the *complex correlator* introduced in the chapter on two element interferometers.

[4] Note that the average being taken over here is *ensemble* average, and *not* an average over time.

$$\langle r_{ij}(t)r_{ij}(t)\rangle = \langle (v_i+n_i)(v_j+n_j)(v_i+n_i)(v_j+n_j)\rangle$$

where for ease of notation we have stopped explicitly specifying that all voltages are functions of time. This quantity is not trivial to work out in general. However, if we assume that all the random processes involved are Gaussian processes[5] the complexity is considerably reduced because for Gaussian random variables the fourth moment can then be expressed in terms of products of the second moment. In particular[6], if $x_1,\ x_2,\ x_3,\ \&\ x_4$ have a joint gaussian distribution then:

$$\begin{aligned}\langle x_1x_2x_3x_4\rangle &= \langle x_1x_2\rangle\langle x_3x_4\rangle + \langle x_1x_3\rangle\langle x_2x_4\rangle + \\ &\quad \langle x_1x_4\rangle\langle x_2x_3\rangle \end{aligned} \qquad (5.1.2)$$

Rather than directly computing $\langle r_{ij}(t)r_{ij}(t)\rangle$, it is instructive first to consider the more general quantity

$$\langle r_{ij}(t)r_{kl}(t)\rangle = \langle (v_i+n_i)(v_j+n_j)(v_k+n_k)(v_l+n_l)\rangle$$

viz. the cross-correlation between the outputs of interferometers (ij) and (kl). We have:

$$\begin{aligned}\langle r_{ij}(t)r_{kl}(t)\rangle &= \langle (v_i+n_i)(v_j+n_j)\rangle\langle (v_k+n_k)(v_l+n_l)\rangle + \\ &\quad \langle (v_i+n_i)(v_k+n_k)\rangle\langle (v_j+n_j)(v_l+n_l)\rangle + \\ &\quad \langle (v_i+n_i)(v_l+n_l)\rangle\langle (v_k+n_k)(v_j+n_j)\rangle \\ \\ &= (\langle v_iv_j\rangle + \langle n_i^2\rangle\delta_{ij})(\langle v_kv_l\rangle + \langle n_k^2\rangle\delta_{kl}) + \\ &\quad (\langle v_iv_k\rangle + \langle n_i^2\rangle\delta_{ik})(\langle v_jv_l\rangle + \langle n_j^2\rangle\delta_{jl}) + \\ &\quad (\langle v_iv_l\rangle + \langle n_i^2\rangle\delta_{il})(\langle v_kv_j\rangle + \langle n_k^2\rangle\delta_{kj}) \\ \\ &= (\mathrm{GS})^2 + \mathrm{GS}(\sigma_i^2\delta_{ij}+\sigma_k^2\delta_{kl}) + \sigma_i^2\delta_{ij}\sigma_k^2\delta_{kl} + \\ &\quad (\mathrm{GS})^2 + \mathrm{GS}(\sigma_i^2\delta_{ik}+\sigma_j^2\delta_{jl}) + \sigma_i^2\delta_{ik}\sigma_j^2\delta_{jl} + \\ &\quad (\mathrm{GS})^2 + \mathrm{GS}(\sigma_i^2\delta_{il}+\sigma_k^2\delta_{kj}) + \sigma_i^2\delta_{il}\sigma_k^2\delta_{kj} \end{aligned} \qquad (5.1.3)$$

The case we are currently interested in is $\langle r_{ij}(t)r_{ij}(t)\rangle$, which from eqn(5.1.3) is:

$$\begin{aligned}\langle r_{ij}(t)r_{ij}(t)\rangle &= 3(\mathrm{GS})^2 + (\sigma_i^2+\sigma_j^2)\mathrm{GS} + \sigma_i^2\sigma_j^2 \\ &= 2(\mathrm{GS})^2 + (\mathrm{GS}+\mathrm{T}_{s_i})(\mathrm{GS}+\mathrm{T}_{s_j}) \end{aligned} \qquad (5.1.4)$$

To get the variance of $r_{ij}(t)$ we need to subtract the square of the mean of $r_{ij}(t)$ from the expression in eqn(5.1.4). Substituting for $\langle r_{ij}(t)\rangle^2$ from eqn(5.1.1) we have:

$$\sigma_{ij}^2 = (\mathrm{GS})^2 + (\mathrm{GS}+\mathrm{T}_{s_i})(\mathrm{GS}+\mathrm{T}_{s_j}) \qquad (5.1.5)$$

Note that the angular brackets denote ensemble averaging. In real life of course one cannot do an ensemble average. Instead one does an average over time, i.e. we work in

[5]Recall from the discussion of sensitivity of a single dish telescope that the central limit theorem ensures that the signal and noise statistics will be well approximated by a Gaussian. This of course does not include 'systematics', like eg. interference, or correlator offsets because of bit getting stuck in the on or off mode etc.

[6]The derivation of this expression is particularly straightforward if one works with the moment generating function; see also the derivation sketched in Chapter 1.

terms of a time averaged correlator output $\bar{r}_{ij}(t)$, defined as

$$\bar{r}_{ij}(t) = \frac{1}{\mathrm{T}} \int_{t-\mathrm{T}/2}^{t+\mathrm{T}/2} r_{ij}(t^{'})dt^{'}$$

As can easily be verified, $\langle \bar{r}_{ij} \rangle = \langle r_{ij} \rangle$. However, computing the second moment, viz., $\bar{\sigma}_{ij}^2 = \langle \bar{r}_{ij}\bar{r}_{ij} \rangle - \langle \bar{r}_{ij} \rangle^2$ is slightly more tricky. It can be shown[7] that if $x(t)$ is a zero mean stationary process and that $\bar{x}(t)$ is the time average of $x(t)$ over the interval $(t-T/2, t+T/2)$, then

$$\bar{\sigma}_x^2 = \frac{1}{T} \int_{-\mathrm{T}/2}^{\mathrm{T}/2} \left(1 - \frac{|\tau|}{\mathrm{T}}\right) R_{xx}(\tau)\ d\tau \tag{5.1.6}$$

where $R_{xx}(\tau)$ is the auto-correlation function of $x(t)$, and $\bar{\sigma}$ is the variance of $x(t)$. Now, if $x(t)$ is a quasi-sinusoidal process with bandwidth $\Delta\nu$, then the integral of $R_{xx}(\tau)$ will be negligible outside the coherence time $1/\Delta\nu$. Further, if $T >> 1/\Delta\nu$, then the factor in parenthesis in eqn(5.1.6) can be taken to be ~ 1 for $\tau < 1/\Delta\nu$. Hence we have:

$$\begin{aligned} \bar{\sigma}_x^2 &\simeq \frac{1}{T} \int_{-\mathrm{T}/2}^{\mathrm{T}/2} R_{xx}(\tau)\ d\tau \quad \simeq \frac{1}{T} \int_{-\infty}^{\infty} R_{xx}(\tau)\ d\tau \\ &= \frac{1}{T}\mathrm{S}_{xx}(0) \qquad\qquad = \frac{1}{T}\frac{\sigma_x^2}{2\Delta\nu} \end{aligned} \tag{5.1.7}$$

where $\mathrm{S}_{xx}(\nu) = \sigma_x^2/2\Delta\nu$ is the power spectrum[8] of $x(t)$. From eqn(5.1.7) and eqn(5.1.5) we hence have

$$\bar{\sigma}_{ij}^2 = \frac{1}{2\mathrm{T}\Delta\nu} \left((\mathrm{GS})^2 + (\mathrm{GS} + \mathrm{T}_{s_i})(\mathrm{GS} + \mathrm{T}_{s_j}) \right) \tag{5.1.8}$$

Putting all this together we get that the signal to noise ratio of a two element interferometer is given by:

$$\mathrm{snr} = \frac{(\sqrt{2\mathrm{T}\Delta\nu}\mathrm{GS})}{\sqrt{(\mathrm{GS})^2 + (\mathrm{GS} + \mathrm{T}_{s_i})(\mathrm{GS} + \mathrm{T}_{s_j})}} \tag{5.1.9}$$

There are two special cases which often arise in practice. The first is when the source is weak, i.e. $\mathrm{GS} \ll \mathrm{T}_s$. In this case the snr becomes

$$\mathrm{snr} = \frac{(\sqrt{2\mathrm{T}\Delta\nu}\mathrm{GS})}{\sqrt{\mathrm{T}_{s_i}\mathrm{T}_{s_j}}} \tag{5.1.10}$$

For a single dish with the collecting area equal to the sum of the collecting areas of antennas i and j (i.e. with gain 2G), and with system temperature $\mathrm{T}_s = \sqrt{\mathrm{T}_{\mathrm{s_i}}\mathrm{T}_{\mathrm{s_j}}}$ the signal to noise would have been a factor of $\sqrt{2}$ better[9]. The loss of signal to noise in the two element interferometer is because one does not measure the auto-correlations of antennas i and j. Only their cross-correlation has been measured. In a sigle dish one would have effectively measured the cross-correlation as well as the auto-correlations.

---

[7] Papoulis, 'Probability, Random Variables & Stochastic Processes', Third Edition, Chapter 10

[8] Where we have made the additional assumption that $x(t)$ is a white noise process, i.e. that its spectrum is flat. The power spectrum for such processes is easily derived from noting that $\int_{-\infty}^{\infty} \mathrm{S}_{xx}(\nu)d\nu = \sigma_x^2$, and that for a quasi-sinosoidal proccess of bandwidth $\Delta\nu$, the integrand is non zero only over an interval $2\Delta\nu$ (including the negative frequencies).

[9] As you can easily derive from eqns 5.1.1 and 5.1.3 by putting $i = j = k = l$. Note that in this case eqn 5.1.1 becomes $\langle r_{ii}(t) \rangle = (v_i(t) + n_i(t))\,(v_i(t) + n_i(t)) = 2\mathrm{GS} + \mathrm{T}_s$

The other special case of interest is when the source is extremely bright, i.e. $\mathrm{GS} \gg \mathrm{T}_s$. In this case, the signal to noise ratio is:

$$\mathrm{snr} = \frac{(\sqrt{2\mathrm{T}\Delta\nu}\mathrm{GS})}{\sqrt{2(\mathrm{GS})^2}} = \sqrt{\mathrm{T}\Delta\nu} \tag{5.1.11}$$

This is as expected, because for very bright sources, one is limited by the Poisson fluctuations of the source brightness, and hence one would expect the signal to noise ratio to go as the square root of the number of independent measurements. Since one gets an independent measurement every $1/\Delta\nu$ seconds, the total number of independent measurements in a time T is just $\mathrm{T}\Delta\nu$.

Having derived the signal to noise ratio for a two element interferometer, let us now consider the case of an N element interferometer. This can be considered as ${}^NC_2$ two element interferometers. Let us take the case where the source is weak. Then from eqn(5.1.3) the correlation between $r_{12}(t)$ and $r_{13}(t)$ is given by

$$\begin{aligned} \langle r_{12}(t) r_{13}(t)\rangle &= \sigma_1^2\delta_{12}\sigma_1^2\delta_{13} + \sigma_1^2\delta_{13}\sigma_1^2\delta_{21} + \sigma_1^2\delta_{11}\sigma_2^2\delta_{23} \\ &= 0 \end{aligned} \tag{5.1.12}$$

The outputs are uncorrelated, even though these two interferometers have one antenna in common[10]. Similarly, one can show that (as expected) the outputs of two two-element interferometers with no antenna in common are uncorrelated. Since the r_{ij}'s are all uncorrelated with one another, the rms noise can simply be added in quadrature. In particular, for an N element array, where all the antennas are identical and have the same system temperature, the signal to noise ratio while looking at a weak source is:

$$\frac{\mathrm{snr} = \sqrt{\mathrm{N(N-1)T}\Delta\nu}\ \mathrm{GS}}{\mathrm{T_s}} \tag{5.1.13}$$

This is the fundamental equation[11] that is used to estimate the integration time required for a given observation. The signal to noise ratio for an N element interferometer is less than what would have been expected for a single dish telescope with area N times that of a single element of the interferometer, but only by the factor $\mathrm{N}/\sqrt{\mathrm{N(N-1)}}$. The lower sensitivity is again because the N auto-correlations have not been measured. For large N however, this loss of information is negligible. For the GMRT, $\mathrm{N} = 30$ and $\mathrm{N}/\sqrt{\mathrm{N(N-1)}} = 1.02$, hence the snr is essentially the same as that of a single dish with 30 times the collecting area of a single GMRT dish.

For a complex correlator[12], the analysis that we have just done holds separately for the cosine and sine channels of the correlator. If we call the outputs of such a correlator r_{ij}^c and r_{ij}^s then it can be shown that the noise in r_{ij}^c and r_{ij}^s is uncorrelated. Further since the time averaging can be regarded as the adding together of a large number of independent samples ($\sim \sqrt{T\Delta\nu}$), from the central limit theorem, the statistics of the noise in $\bar{r}_{ij}^c$ and $\bar{r}_{ij}^s$ are well approximated as Gaussian. It is then possible to derive the statistics of functions of $\bar{r}_{ij}^c$ and $\bar{r}_{ij}^s$, such as the visibility amplitude ($\sqrt{\bar{r}_{ij}^c + \bar{r}_{ij}^s}$) and the visibility phase ($\tan^{-1}\bar{r}_{ij}^s/\bar{r}_{ij}^c$). For example, it can be shown that the visibility amplitude has a Rice distribution[13]

[10]This may seem counter intuitive, but note that the outputs are only uncorrelated, they are not independent.

[11]In some references, an efficiency factor η is introduced to account for degradation of signal to noise ratio because of the noise introduced by finite precision digital correlation etc. This factor has been ignored here, or equivalently one can assume that it has been absorbed into the system temperature.

[12]See the chapter on two element interferometers

[13]Papoulis, 'Probability, Random Variables & Stochastic Processes', Third Edition, Chapter 6.

For an extended source, the entire analysis that we have done continues to hold, with the exception that S should be treated as the correlated part of the source flux density. For example, at low frequencies, the Galactic background is often much larger than the receiver noise and one would imagine that the limiting case of large source flux density (i.e. eqn(5.1.11) is applicable. However, since this background is largely resolved out at even modest spacings, its only effect is an increase in the system temperature.

Finally we look at the noise in the image plane, i.e. after Fourier transformation of the visibilities. Since most of the astronomical analysis and interpretation will be based on the image, it is the statistics in the image plane that is usually of interest. The intensity at some point (l, m) in the image plane is given by:

$$I(l,m) = \frac{1}{\mathrm{M}} \sum_p w_p \mathcal{V}_p e^{-i2\pi(lu_p + mv_p)}$$

where w_p is the weight[14] given to the pth visibility measurement $\mathcal{V}_p$, and there are a total of M independent measurements. The cross-correlation function in the image plane, $\left\langle I(l,m)I(l^{'},m^{'})\right\rangle$ is hence:

$$\left\langle I(l,m)I(l^{'},m^{'})\right\rangle = \frac{1}{\mathrm{M}^2} \sum_p \sum_q w_p w_q \left\langle \mathcal{V}_p \mathcal{V}_q^* \right\rangle e^{-i2\pi(lu_p+mv_p)} e^{i2\pi(l^{'}u_q+m^{'}v_q)}$$

In the absence of any sources, the visibilities are uncorrelated with one another, and hence, we have

$$\left\langle I(l,m)I(l^{'},m^{'})\right\rangle = \frac{1}{\mathrm{M}^2} \sum_m w_p^2 \sigma_p^2 e^{-i2\pi((l-l^{'})u_p+(m-m^{'})v_p)}$$

Hence in the case that all the noise on each measurement is the same, and that the weights given to each visibility point is also the same, (i.e. uniform tapering), the correlation in the map plane has exactly the same shape as the dirty beam. Further the variance in image plane would then be $\sigma_{\mathcal{V}}^2/\mathrm{M}$, where $\sigma_{\mathcal{V}}^2$ is the noise on a single visibility measurement. This is equivalent to eqn(5.1.13), as indeed it should be.

Because the noise in the image plane has a correlation function shaped like the dirty beam, one can roughly take that the noise in each resolution element is uncorrelated. The expected statistics after simple image plane operations (like smoothing) can hence be worked out. However, after more complicated operations, like the various possible deconvolution operations, the statistics in the image plane are not easy to derive.

5.2 Calibration

We have assumed till now that we have been working with calibrated visibilities, i.e. free from all instrumental effects (apart from some additive noise component). In reality, the correlator output is different from the true astronomical visibility for a variety of reasons, to do with both instrumental effects as well as propagation effects in the earth's atmosphere and ionosphere.

At low frequencies, it is the effect of the ionosphere that is most dominant. As is discussed in more detail in Chapter 16, density irregularities cause phase irregularities in the wavefront of the incoming radio waves. One would expect therefore that the image

[14] As discussed in Chapter 11, this weight is in general a combination of weights chosen from signal to noise ratio considerations and from synthesized beam shaping considerations.

of the source would be distorted in the same way that atmospheric turbulence ('seeing') distorts stellar images at optical wavelengths. To first order this is true, but for the ionosphere the 'seeing disk' is generally smaller than the diffraction limit of typical interferometers. There are two other effects however which are more troublesome. The first is 'scintillation', where because of diffractive effects the flux density of the source changes rapidly – the flux density modulation could approach 100%. The other is that slowly varying, large scale refractive index gradients cause the apparent source position to wander. At low frequencies, the source position could often wander by several arc minutes, i.e. considerably more than the synthesized beam. As we shall see below, provided the time scale of this wander is slow enough, it can be corrected for.

Let us take the case where the effect of the ionosphere is simply to produce an excess path length, i.e. for an antenna i let the excess phase[15] for a point source at sky position (l,m) be $\phi_i(l,m,t)$, where we have explicitly put in a time dependence. Then the observed visibility on a baseline (i,j) would be

$$\widetilde{\mathcal{V}}_{ij}(t) = G_{ij}(t)\int e^{-i(\phi_i(l,m,t)-\phi_j(l,m,t))}I(l,m)e^{-i2\pi(lu_{ij}+mv_{ij})} \qquad (5.2.14)$$

where $I(l,m)$ is the sky brightness distribution and we have ignored the primary beam[16]. $G_{ij}(t)$ is 'instrumental phase', i.e. the phase produced by the amplifiers, transmission lines, or other instrumentation along the signal path. If $\phi_i(l,m,t)$ were some general, unknown function of (l,m,t) it would not be possible to reconstruct the true visibility from the measured one. However, since the size scale of ionospheric disturbances is $\sim$ a few hundred kilometers, it is often the case that $\phi_i(l,m,t)$ is constant over the entire primary beam, i.e. there is no (l,m) dependence. The source is then said to lie within a single *iso-planatic patch*. In such situations, the ionospheric phase can be taken out of the integral, and eqn(5.2.14) reduces to:

$$\widetilde{\mathcal{V}}_{ij}(t) = G_{ij}(t)e^{-i(\phi_i(t)-\phi_j(t))}\int I(l,m)e^{-i2\pi(lu_{ij}+mv_{ij})} \qquad (5.2.15)$$

If it also the case that the ionospheric and instrumental gains are changing slowly, then they can be calibrated in the following manner. Suppose that close to the source of interest, there is a calibration source whose true visibility $\mathcal{V}^c_{ij}$ is known. Then one could intersperse observations of the target source with observations of the calibrator. For the calibrator, dividing the observed visibility $\widetilde{\mathcal{V}}^c_{ij}(t)$ by the (known) true visibility, $\mathcal{V}^c_{ij}(t)$ one can measure the factor $G_{ij}(t)e^{-i(\phi_i(t)-\phi_j(t))}$. This can then be applied as a correction to the visibilities of the target source. For slightly better corrections, one could interpolate in time between calibrator observations. This is the basis of what is sometimes called 'ordinary' calibration. The calibrator source is usually an isolated point source, although this is not, strictly speaking, necessary. It is sufficient to know the true visibilities $\mathcal{V}^c_{ij}(t)$. Note that if the calibrators absolute flux is also known, then this calibration procedure will also calibrate the amplitude scale of the target source[17].

In the approach outlined above, in order to calibrate the data one needs to solve for an unknown complex number per baseline, (i.e. N(N-1)/2 complex numbers for an N element interferometer). If we assume that the correlator itself does not produce any errors[18], i.e. that all the instrumental errors occur in the antennas or the transmission lines, then the

[15] by which we mean the phase difference over what would have been obtained in the absence of the ionosphere

[16] i.e. we have set the factor $B(l,m)/\sqrt{1-l^2-m^2}$ to 1.

[17] provided, as we will discuss in more detail later, that the system temperature does not differ for the target source and the calibrator

[18] which is often a good assumption for digital correlators

instrumental gain can be written out as antenna based terms, i.e.

$$G_{ij}(t) = g_i(t)g_j^*(t) \tag{5.2.16}$$

where $g_i(t)$ and $g_j(t)$ are the complex gains along the signal paths from antennas 1 and 2. But the ionospheric phase can also be decomposed into antenna based quantities (see eqn 5.2.15), and can hence be lumped together with the instrumental phase. Consequently the total unknown complex gains that have to be solved for reduces from N(N-1)/2 to N, which can be a dramatic reduction for large N. (For the GMRT it is a reduction from 435 unknowns to 30 unknowns).

However to appreciate the real power of this decomposition into antenna based gains, consider the following quantities. First let us look at the sum of the phases of the raw visibilities $\widetilde{\mathcal{V}}_{12}$, $\widetilde{\mathcal{V}}_{23}$ and $\widetilde{\mathcal{V}}_{31}$. If we call the true visibility phase $\psi_{\mathcal{V}_{ij}}$, the raw visibility phase $\psi_{\widetilde{\mathcal{V}}_{ij}}$ and the sum of the instrumental and ionospheric phases χ_i, then we have

$$\begin{aligned} \psi_{\widetilde{\mathcal{V}}_{12}} + \psi_{\widetilde{\mathcal{V}}_{23}} + \psi_{\widetilde{\mathcal{V}}_{31}} &= \chi_1 - \chi_2 + \psi_{\mathcal{V}_{12}} + \chi_2 - \chi_3 + \psi_{\mathcal{V}_{12}} + \chi_3 - \chi_1 + \psi_{\mathcal{V}_{31}} \\ &= \psi_{\mathcal{V}_{12}} + \psi_{\mathcal{V}_{23}} + \psi_{\mathcal{V}_{31}} \end{aligned} \tag{5.2.17}$$

i.e. over any triangle of baselines the sum of the phases of the raw visibilities is the true source visibility. This is called *phase closure*. Similarly it is easy to show that for any baselines 1,2,3,4, the ratio of the raw visibilities will be the same as the true visibilities, i.e.

$$\frac{|\widetilde{\mathcal{V}}_{12}||\widetilde{\mathcal{V}}_{34}|}{|\widetilde{\mathcal{V}}_{23}||\widetilde{\mathcal{V}}_{41}|} = \frac{|\mathcal{V}_{12}||\mathcal{V}_{34}|}{|\mathcal{V}_{23}|\mathcal{V}_{41}|} \tag{5.2.18}$$

This is called *amplitude closure*. For an N element interferometer, we have $1/2N(N-1) - (N-1)$ constraints on the phase and $1/2N(N-1) - N$ constraints on the amplitude. For large N, this is considerably more than the N unknown gains that one is solving for. The large number of available constraints means that the following iterative scheme would work.

1. Choose a suitable starting model for the brightness distribution. Compute the model visibilities.
2. For this model, solve for the antenna gains, subject to the closure constraints.
3. Apply these gain corrections to the visibility data, use the corrected data to make a fresh model of the brightness distribution.

For arrays with sufficient number of antennas, convergence is usually rapid. Note however, for this to work, the signal to noise ratio per visibility point[19] has to be reasonable, i.e. 2-3. This is often the case at low frequencies, and this technique of determining antenna gains (which is called *self calibration*) is usually highly successful.

Note that if one adds a phase $\chi_i = 2\pi(l_0 u_i + m_0 v_i)$ to each antenna (where $l_0,\ m_0$ are arbitrary and (u_i, v_i) are the (u,v) co-ordinates of the ith antenna), the phase closure constraints (eqn 5.2.17) continue to be satisfied. That means that in self calibration the phases can be solved only upto a constant phase gradient across the uv plane, i.e. the absolute source position is lost. Similarly, it is easy to see that the amplitude closure constraints will be satisfied even if one multiplies all the gains by a constant number, i.e. in self calibration one loses information on the absolute source flux density . The only way to determine the absolute source flux density is to look at a calibrator of known flux.

[19]Actually strictly speaking one means the signal to noise ratio over an interval for which the ionospheric phase can be assumed to be constant

Since antenna gains and system temperatures are usually stable over several hours[20], it is usually sufficient to do this calibration only once during an observing run. A more serious problem at low frequencies is that the Galactic background (whose strength varies with location on the sky) makes a significant contribution to the system temperature. Hence, when attempting to measure the source flux density, it is important to correct for the fact that the system temperature is different for the calibrator source as compared to the target source. The system temperature can typically be measured on rapid time scales by injecting a noise source of known strength at the front end amplifier.

Another related way (to selfcal) of solving for the system gains is the following. Suppose that the visibility on baselines (i, j) and (k, l) are identical. Then the ratio of the measured visibilities is directly related to the ratio of the complex instrumental gains of antennas i, j, k & l. If there are enough number of such 'redundant' baselines, one could imagine solving for the instrumental gains. Some arrays, like the WSRT have equispaced antennas, giving rise to a very large number of redundant baselines, and this technique has been successfuly used to calibrate complex sources[21] For a simple source, like a point source, all possible baselines are redundant, and this technique reduces essentially to self-calibration.

At the very lowest frequencies ($\nu < 200$ MHz, roughly for the GMRT) the assumption that the source lies within the iso-planatic patch probably begins to break down. The simple self calibration scheme outlined above will stop working in that regime. A possible solution then, is to solve (roughly speaking) for the phase changes produced by each iso-planatic patch. Often the primary beams of several antennas will pass through the same iso-planatic patches, so the extra number of degrees of freedom introduced will not be substantial, and an iterative approach to solving for the unknowns will probably converge[22].

5.3 Further Reading

1. Hamaker J. P., O'Sullivan, J. D. & Noordam, J. E., Journal of the Opt. Soc. Of America, **67**, 1122.

2. Thompson, R. A., Moran, J. M. & Swenson, G. W. Jr., 'Interferometry & Synthesis in Radio Astronomy', Wiley Interscience.

3. R. A. Perley, F. R. Schwab, & A. H. Bridle, eds., 'Synthesis Imaging in Radio Astronomy'

[20]Or change in a predictable manner with changing azimuth and elevation of the antennas

[21]see Noordam, J. E. & de Bruyn A. G., 1982, Nature **299**, 597.

[22]See Subrahmanya, C. R., (in 'Radio Astronomical Seeing', J. E. Baldwin & Wang Shouguan eds.) for more details