Very Long Baseline Interferometry and the VLBA
ASP Conference Series, Vol. 82, 1995
J. A. Zensus, P. J. Diamond, and P. J. Napier (eds.)

Printed June 13, 1995

CHAPTER 3

Imaging Concepts

TIM CORNWELL

NATIONAL RADIO ASTRONOMY OBSERVATORY

Abstract The three principal areas of imaging are discussed: Fourier-inversion, deconvolution, and self-calibration. The emphasis is on the major concepts in these areas. Other chapters in these proceedings cover the practical details of how one drives a particular imaging algorithm.

Contents

3.1 Introduction

Interferometric arrays form images of the radio sky only indirectly. An image is constructed and refined in the three steps:

Fourier inversion: Formation of an initial *dirty* image from the visibility samples collected by an interferometric array. For review material, see Thompson, Moran and Swenson (1986; hereafter TMS), or Sramek and Schwab (1989).

Deconvolution: Correction of the effect of Fourier plane sampling deficiencies on the dirty image. For review material, see TMS or Cornwell and Braun (1989).

Self-calibration: Correction of the effect of calibration errors on the deconvolved image. For review material, see Pearson and Readhead (1984), TMS, Cornwell and Fomalont (1989), or Cornwell (1989).

It is important to realize the relative status of these three steps. Fourier inversion is ordained by the measurement equation of an interferometric array and is thus on firm ground. Deconvolution involves an equation with many possible solutions and so selecting one solution as an image is fraught with difficulties and uncertainties. Self-calibration uses subtle consistency relations in the measured visibility data to further refine the deconvolved images. Nevertheless, it has been found by experience that all three steps are vital for high quality interferometric imaging in VLBI, and that with understanding of the dangers, reliable images can be produced.

3.2 Fourier Inversion: Theory

An interferometric array measures the visibility function, whereas for scientific purposes an image of the sky brightness is usually required. How do we convert from one to the other? In principle, it is quite straightforward: since the Visibility V and the Sky Brightness I form a Fourier pair:

$$V(u,v) = \int I(x,y) e^{2\pi j(ux+vy)}\, dx\, dy\,, \tag{3.1}$$

one can obtain the sky brightness by a Fourier inversion:

$$I(x,y) = \int V(u,v) e^{-2\pi j(ux+vy)}\, du\, dv\,. \tag{3.2}$$

In practice, the true visibility function V is not available and we have only noisy and perhaps poorly calibrated samples of the visibility function at discrete locations in the Fourier plane. Leaving aside the role of noise and calibration errors, we can represent this by replacing V by the *sampled* visibility function $S(u,v)V(u,v)$, where the sampling function $S(u,v)$ is:

$$S(u,v) = \sum_k w_k \delta(u-u_k)\delta(v-v_k)\,. \tag{3.3}$$

Inserting this visibility function into the Fourier inverse, we obtain not the true image $I(x,y)$ but an inferior image, the *dirty* image:

$$I^D(x,y) = \int S(u,v)V(u,v) e^{-2\pi j(ux+vy)}\, du\, dv\,. \tag{3.4}$$

We can reveal more of the structure of the dirty image by substituting for S:

$$I^D(x,y) = \sum_k V(u_k, v_k) e^{-2\pi j(u_k x + v_k y)} w_k \, . \tag{3.5}$$

Thus the dirty image is simply a collection of sinusoidals of amplitudes given by the visibility samples. To improve the approximation this gives to a given source, it pays to have more sinusoids in the collection. Hence more visibility samples yield a better image of a source. We give firmer meaning to this somewhat vague statement below.

3.3 Fourier Inversion: The Problem with I^D-Sidelobes

The *dirty* image I^D is the FT of the *sampled* visibility function:

$$I^D = \mathcal{F}^{-1}(SV) \, . \tag{3.6}$$

Applying the convolution theorem from Fourier transform theory, we find that the dirty image is the convolution of the true image I with a *Dirty* beam B:

$$I^D = B * I \, . \tag{3.7}$$

where the Dirty Beam or Point Spread Function $B = \mathcal{F}^{-1}(S)$ is given by:

$$B(x,y) = \sum_k \cos\left[2\pi\left(u_k x + v_k y\right)\right] w_k \, . \tag{3.8}$$

Hence every point source in the true image I is seen in the dirty image as an appropriately scaled and centered dirty beam pattern. The dirty beam can be thought of as composed of two parts: a compact core surrounded by a rumble of *sidelobes* far out (see fig. 3.3). The sidelobes of the dirty beam necessarily extend over the entire image plane. The RMS sidelobe level far from the center of the dirty beam, σ_B, can be roughly estimated using a random approximation for location of samples:

$$\sigma_B \sim \frac{1}{\sqrt{2N_{\text{samples}}}} \, . \tag{3.9}$$

If the RMS is to be calculated over a region with a distance L in the image plane, then one has to count the number of cells, of size $1/L$, in the Fourier plane which contain samples:

$$\sigma_B \sim \frac{1}{\sqrt{2N_{\text{cells}}}} \, . \tag{3.10}$$

The near-in sidelobe level can be calculated using a disk-like approximation for the Fourier plane coverage. The corresponding point spread function then resembles a jinc function and has close-in sidelobes of about 10-15%. The data may be tapered in the Fourier plane to reduce the effects of the sharp cutoff at the edge of the sampled region, in which case the point spread function comes to resemble the square of a Jinc function with sidelobe levels of a few percent. The disadvantage is that resolution and signal to noise are lost since the longest spacings are down weighted.

To a considerable degree, one can control the properties of the dirty image by suitable choice of the weights w_k. To minimize the noise of the dirty image, one must use *natural* weighting:

$$w_k = \frac{1}{\sigma_k^2} . \tag{3.11}$$

Another way of interpreting this form of weighting is that all other factors being equal, visibility samples are weighted by the integration time.

The chief defect of natural weighting is that regions of high density of sampling of the Fourier plane are strongly weighted in the final dirty PSF. While this is required for optimum signal-to-noise, it is less satisfactory for images of bright objects where noise is not the limiting factor. In such a case, the quality of the dirty PSF can be improved by choosing a weighting scheme that minimizes the rms sidelobes of the dirty beam. This leads to *uniform* weighting, where the local density of sampling is corrected:

$$w_k = \frac{1}{\rho(u_k, v_k)} , \tag{3.12}$$

where $\rho(u, v)$ is the local density of samples.

Recently, Briggs (1995) has considerably enhanced possibilities for weighting of data by introducing *robust* weighting. In this approach, a linear combination of the thermal noise and the sidelobe power is minimized, thus combining uniform and natural weighting. The relative importance of the two sources of noise can be tuned. By careful tuning, the sensitivity can often be improved over that for uniform weighting while also avoiding the loss of resolution of natural weighting. This tradeoff works best for inhomogeneous arrays such as the VLBA plus the EVN plus an orbiting element. However, the tradeoff is still worth investigating for even the VLBA alone. Robust weighting is available in the AIPS imaging programs.

3.4 Fourier Inversion: A Digression on Computation

It is important to understand how the Fourier Inverse is actually computed since a couple of subtle limitations of the resulting images may occur. On a first reading of this chapter, you can skip this section, but do come back to it later on.

In the seldom-used *direct Fourier summation* approach, one uses direct evaluation of a number, N, of pixels:

$$I^D(x_p, y_p) = \sum_{k=1}^{M} V(u_k, v_k) e^{-2\pi j(u_k x_p + v_k y_p)} w_k \,. \tag{3.13}$$

The work involved scales as $O(MN)$. If $N \sim M$, then this goes as M^2 and is therefore only feasible for small images (64 by 64, say). While the DFS is easy to code, the *fast Fourier transform* (FFT) is usually much more computationally efficient. If the visibility samples naturally lie on a grid then we can use periodic nature of $e^{-2\pi j(u_k x_p + v_k y_p)}$ to reduce number of operations to $O(N \log_2(N))$ and so the cost saving of an FFT over DFS is $\sim N/\log_2(N)$.

In most interferometric arrays, the visibility samples come from ellipses in the Fourier plane. There are quite a few ways of converting from such irregularly spaced samples to the regularly spaced samples needed for the FFT. The most common approach in radio-astronomy is to smooth and then re-sampled the visibility function onto a regular grid. The observed visibility function is:

$$SV \,. \tag{3.14}$$

After *smoothing* with a function C, this becomes:

$$C * (SV) \,. \tag{3.15}$$

Applying a *re-sampling* operator R (which just samples at the regularly spaced locations needed for the FFT), we have:

$$R\left[C * (SV)\right] \,. \tag{3.16}$$

The resampling operator R is Bracewell's Shah (or bed-of-nails) function which has the interesting property that it Fourier transforms to another Shah function:

$$R(u, v) = \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} \delta\left(j - \frac{u}{\Delta u}\right) \delta\left(k - \frac{v}{\Delta v}\right) \,, \tag{3.17}$$

and

$$\mathcal{F}^{-1}(R)(x, y) = \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} \delta(j - x\Delta u)\, \delta(k - y\Delta v)\, \Delta u\, \Delta v \,. \tag{3.18}$$

Thus, after Fourier Inversion, we have:

$$\mathcal{F}^{-1}(R[C * (SV)]) = \mathcal{F}^{-1}(R) * \left[\mathcal{F}^{-1}(C)\, I^D\right] \,. \tag{3.19}$$

In plain English, this means that one obtains a copy of I^D, tapered by $\mathcal{F}^{-1}(C)$, and replicated by $\mathcal{F}^{-1}(R)$. To recover a reasonable estimate of I^D, one therefore has to divide by $\mathcal{F}^{-1}(C)$. This is colloquially known as *grid correction*. Note that the replication cannot be countered in an easy way. Furthermore if the replication spacing is insufficiently large then the source can overlap (or *alias*) itself. Actually, since the Fourier transforms of both the sampling function S and the smoothing function C necessarily extend for ever in the image plane, some aliasing will always occur (this is usually called *sidelobe aliasing* to distinguish it from aliasing of the source itself). The tradeoff is therefore between the sharpness of cutoff in $\mathcal{F}^{-1}(C)$, which minimizes aliasing, and the limited size of C, which keeps the computation requirements manageable. The best choice for the convolution function is generally considered to be the prolate spheroidal wave functions (and relatives), and the support of C is usually chosen so that about seven cells in the Fourier plane are affected by any one sample (see Sramek and Schwab, 1989).

3.4.1 Deconvolution: The Convolution Equation

To re-iterate what we learned above, if we ignore the effects of noise and calibration errors then the dirty image is the convolution of the true sky brightness with a dirty beam.

$$I^D = B * I\,. \tag{3.20}$$

This convolution is quite disastrous since because B is composed of a finite number of sinusoids, there are functions $Z(x, y)$ that are invisible to B:

$$B * Z = 0\,. \tag{3.21}$$

A simple example is, e.g.,

$$Z(x, y) = \cos\left[2\pi\left(ux + vy\right)\right] \tag{3.22}$$

or

$$Z(x, y) = \sin\left[2\pi\left(ux + vy\right)\right]\,, \tag{3.23}$$

where (u, v) was not sampled. These invisible distributions devastate our ability to solve the convolution equation since if I^S is a solution of the convolution equation, then so is $I^S + \alpha Z$ for any α. This means that no *linear* convolution C will recover the invisible distributions:

$$\begin{aligned} C * I^D &= C * \left[B * \left(I^S + \alpha Z\right)\right] \\ &= C * \left[B * I^S\right]\,. \end{aligned} \tag{3.24}$$

There is nothing mysterious about the invisible distributions. To eliminate a given sinusoidal invisible distribution of the type described above, all one has to do is to measure the visibility function at the appropriate spatial frequency

(u, v). However, even after measuring as many visibility points as possible, there will always be invisible distributions. Consequently one always needs extra information to constrain the invisible distributions Z. There are many different examples of information that can be used in this way:

Finite support The emission is known or assumed non-zero only within a limited region of known shape and position.

Extrema in brightness E.g., the brightness is known to positive or the fractional polarization is known to be less than unity.

Sky = Collection of point sources : The sky is assumed to be representable by a small number of point sources:

$$I^S = \sum_{i=1}^{P} I_i^C \delta(x - x_i)\delta(y - y_i)\,. \tag{3.25}$$

Sky is smooth The sky is assumed to be smooth. This can be expressed in many different ways: e.g., the image dispersion is minimal:

$$\int I^2(x, y)\, dx\, dy\,. \tag{3.26}$$

It now becomes clear what we mean by a deconvolution algorithm such as CLEAN or the maximum entropy method: a deconvolution algorithm uses extra information to solve the convolution equation for an estimate of the true sky brightness I. We now discuss the two principal deconvolution algorithms used in radio-interferometry.

3.4.2 Deconvolution: The CLEAN Algorithm

The CLEAN algorithm, which was invented by Jan Högbom in early '70s, is the most commonly used deconvolution algorithm in radio interferometry. It assumes that the sky can be decomposed into a small number of point sources:

$$I^S = \sum_{i=1}^{P} I_i^C \delta(x - x_i)\delta(y - y_i)\,. \tag{3.27}$$

With this assumption, the dirty image is given by:

$$I^D = \sum_{i=1}^{P} I_i^C B(x - x_i, y - y_i)\,. \tag{3.28}$$

One could solve for the parameters of the point sources by a brute force least-squares optimization. Since this was not computationally feasible at the time, Högbom instead developed a more computationally feasible iterative algorithm in which one has to find a sequence of PSF shapes in the dirty image. For any one beam-shape in the sequence, the optimum estimate of the peak and position can be obtained from the cross-correlation of the dirty image and the dirty beam $I^D * B$. One then removes this peak point source and iterates. To obtain a useful image at the end of the process, the point sources are smoothed with a gaussian CLEAN beam and added to any residual emission.

Although the Högbom algorithm is simple and robust, it is also quite slow for complex objects. Hence two variants are more generally used: the Clark and Schwab-Cotton algorithms (see, e.g., Cornwell and Braun, 1989 for more details).

Although the conditions for convergence of the (Högbom) CLEAN algorithm are known (Schwarz 1978), the properties of the final solution are hard to pin down. Indeed, understanding of the algorithm in even simple cases is poor. For example, Tan (1986) shows examples of chaotic behavior of the algorithm even for very simple, three pixel images. Nevertheless, in real cases, the CLEAN algorithm seems to work well and is reasonably fast.

3.4.3 Deconvolution: The Maximum Entropy Algorithm

The next most popular deconvolution algorithm in radio-interferometry, is the maximum entropy method invented by Jaynes (1957). The first application of MEM to image deconvolution was made by (Frieden 1972). Since then many authors have made contributions to the literature, both in the form of algorithms for calculating MEM images and in interpretations of the meaning of the entropy of an image. Unfortunately the latter, less useful contributions outnumber the former by a considerable margin. Despite this, MEM images can be very useful. Narayan and Nityananda (1986) argue that MEM simply imposes the assumption that the source is a collection of peaks on an otherwise empty background. Entropy forces peaks to be sharpened, baselines to be flattened.

There are many forms of entropy which impose the Narayan-Nityananda image plane constraint, e.g.:

$$\mathcal{H}[I] = \int \log\left[I(x,y)\right] dx\, dy \tag{3.29}$$

$$\mathcal{H}[I] = \int \sqrt{[I(x,y)]}\, dx\, dy \tag{3.30}$$

$$\mathcal{H}[I] = -\int I(x,y) \log\left[I(x,y)\right] dx\, dy\,. \tag{3.31}$$

The form used in the AIPS MEM program, VTESS, is:

$$\mathcal{H}[I] = -\int I(x,y) \log\left[I(x,y)/M(x,y)\right] dx\, dy\,, \tag{3.32}$$

where $M(x, y)$ is a *default* image that represents the optimum solution in the absence of data. Usually the default image is taken to be flat, but it can also usefully be a lower resolution image.

Given a form of entropy, the MEM defines the best image as that which maximizes $\mathcal{H}[I]$ subject to the data constraints:

1. Fit the observed visibilities V_k to within some error term ϵ_k:

$$\mathcal{F}(I) = V_k + \epsilon_k \,. \tag{3.33}$$

2. Fit the total flux

$$\int I(x, y)\, dx\, dy = F \,. \tag{3.34}$$

3. The RMS error must be plausible:

$$\sum_k \frac{|\epsilon_k|^2}{\sigma_k^2} = \Omega \,. \tag{3.35}$$

This usually requires constrained optimization of $10^2 - 10^7$ variables. There are a number of published algorithms (see Narayan and Nityananda 1986 for an overview), e.g.,

1. The MEMSYS algorithm of Skilling et al. is very powerful and can handle many different variants but the code is proprietary and expensive (Skilling and Bryan 1985; Gull and Skilling 1991).
2. The Cornwell-Evans algorithm (1985) used in AIPS was designed to be simple and easy to implement. It works well for many cases and has the great advantage that the code is public domain.
3. The Wilczek-Drapatz (1985) algorithm is well-suited to high-accuracy deconvolution of a small number of data points.

3.5 Deconvolution: Comparison of the CLEAN and MEM Algorithms

To demonstrate the practical differences between CLEAN and MEM deconvolutions, I show the results of applying both algorithms to simulated VLBA observations of an M87-like object (fig. 3.1). The model for this object was supplied by John Conway and the simulations were performed by Dan Briggs. The declination of the source was 50 degrees, and the frequency was 1.6 GHz. Thermal noise was added but the calibration was assumed to be perfect.

The following steps were performed:

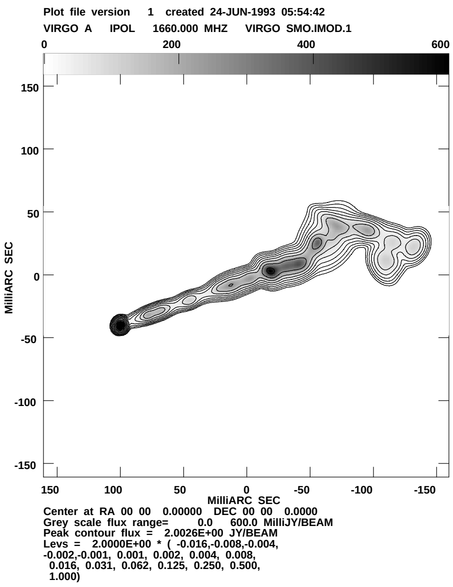

Figure 3.1: Simulation of an M 87-like object

1. Calculate sampling function $S(u, v)$ (fig. 3.2).

2. Fourier invert the sampling pattern S to obtain the PSF $B(x, y) = \mathcal{F}^{-1}(S(u, v))$ (fig. 3.3.) Note the core, the close-in sidelobes and the far-out rumble.

3. Evaluate Fourier transform of model at sample points (fig. 3.2).

4. Fourier invert to obtain the Dirty image $I^D(x, y)$ The basic structure of the object is still apparent but the fine details are swamped by the sidelobes, especially those from the bright point source.

5. Calculate the CLEAN deconvolution, with and without CLEAN boxes (i.e., support constraints). 20,000 CLEAN components were subtracted with a loop gain of 0.1. Note the large improvement in quality of deconvolution with the addition of the support constraint (fig. 3.4).

6. Calculate the MEM deconvolution, with and without support constraints. Eighty iterations were required. This nicely demonstrates the well known feature of MEM that it does rather poorly when a point source is seen close to extended emission: the sidelobes from the point source are seen embedded in the extended emission. This can be cured by subtracting the point source prior to deconvolution. (fig. 3.5).

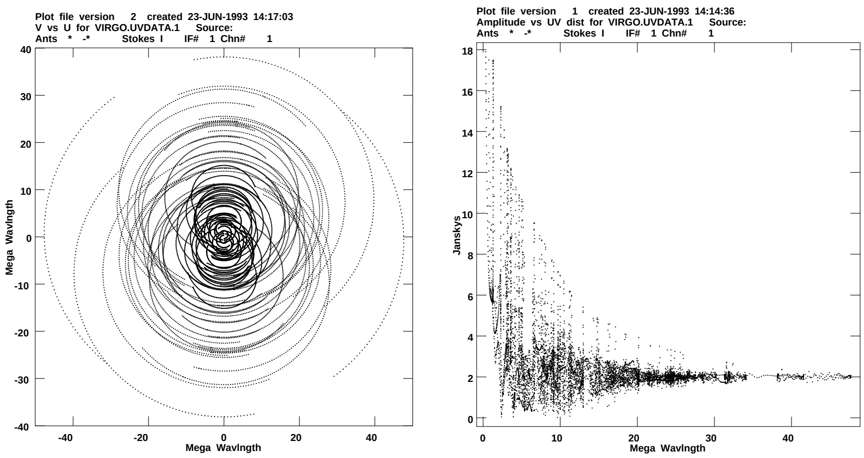

Figure 3.2: *Left:* The Fourier plane sampling function $S(u, v)$. *Right:* Fourier transform of model at sampled points

3.5.1 Deconvolution: Non-Negative Least Squares

In the discussion above of CLEAN, it was mentioned that one could solve for the parameters of the point sources by a brute force least-squares optimization. This is now computationally feasible for images having non-zero flux in about 5000–6000 pixels. Briggs (1995) has applied a non-negatively constrained least-squares solution algorithm (NNLS) to solve the convolution equation for compact objects. This approach allows very high dynamic range imaging of VLBA observations. Using this algorithm, Cornwell, Kemball and Benson (1995) show an image of DA 193 observed with the VLBA at λ6cm having dynamic range (peak:rms) of about 120,000. While such dynamic range requires that the VLBA correlator have low closure errors, it is also true that NNLS is vital, since CLEAN alone, when used normally, leaves residual errors off-source that mimic calibration errors but that are really due to deconvolution errors. This so-called *Briggs'* effect is now thought to be potentially present in a considerable number of VLA and VLBI images. A more detailed description of this important effect is to be found in Briggs' Ph. D. thesis (1995). Currently the only implementation of the NNLS algorithm is in the experimental *SDE* package (see Cornwell, Kemball and Benson 1995). CLEAN can be driven to avoid the Briggs effect by the simple expedient of cleaning much deeper than is typically done (about a tenth of the noise level is adequate for VLBA images). However, it is hard to make this strategy work with that required for self-calibration (see next section). By contrast, NNLS works well with self-calibration procedures.

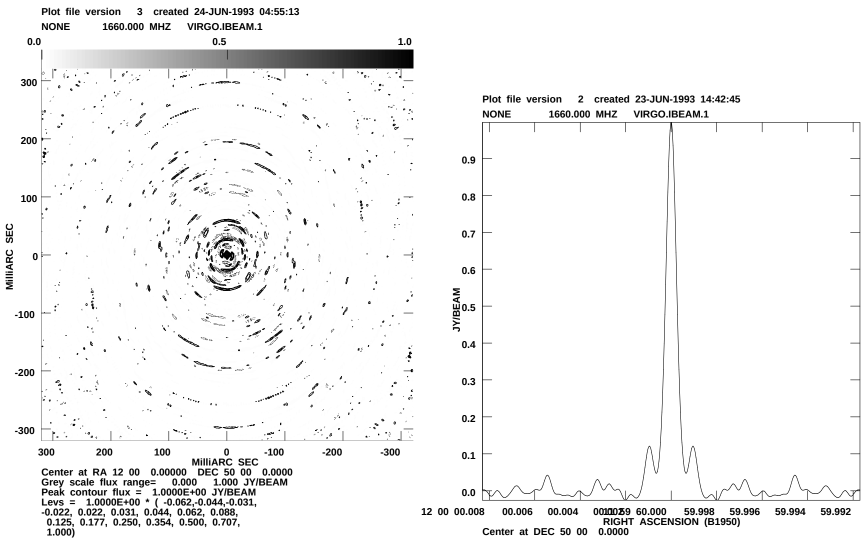

Figure 3.3: *Left:* the point spread function $B(x, y)$. *Right:* Slice through point spread function.

3.6 Self-calibration: Origin of Phase Errors

In VLBI, the imaging performance of an array is usually limited by two factors: the Fourier plane coverage, and the inability to provide accurately calibrated data. The deconvolution algorithms described in the previous sections address the Fourier coverage limitations. The calibration problem is addressed by self-calibration algorithms.

Before discussing how self-calibration works, I will first briefly mention the types of calibration errors that occur. The errors are mainly in the phase of the measured visibility.

The major sources of phase error are:

1. Poor instrumental stability/knowledge, e.g., clocks, antenna positions, source position),
2. Troposphere water vapor fluctuations,
3. Ionospheric electron content fluctuations.

The latter two saturate for connected-element interferometers such as the VLA and MERLIN so the major cause of phase error in VLBI tends to be the first class of error.

In the presence of phase errors, the measured visibility function is:

$$V_{i,j}(u, v) = e^{j(\phi_i - \phi_j)} \int I(x, y) e^{2\pi j(ux+vy)}\, dx\, dy\,. \tag{3.36}$$

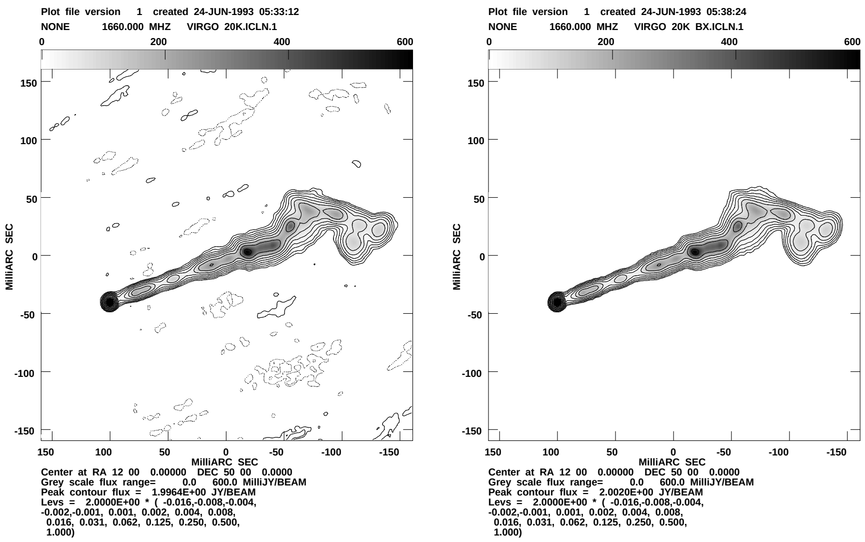

Figure 3.4: CLEAN deconvolution, 20000 iterations, loop gain 0.1. *Left:* without support box. *Right:* with support box.

Calibration by time-multiplexing to a nearby source will remove part of the phase errors. Imaging quality then depends upon the rms residual phase error σ_ϕ:

$\sigma_\phi \ll 1$: Image Dynamic range is limited to $\frac{N}{\sigma_\phi}$ where N is number of antennas.

$\sigma_\phi \gg 1$: Imaging is impossible.

3.7 Self-calibration: Phase Closure

In the case where time-multiplexing to a nearby source is not possible, imaging would seem to be impossible. However, Jennison (1958) noticed that the (antenna based) phase error ϕ vanishes in a sum of the coherence phase around a closed loop:

$$\begin{aligned}\Phi_{123}^{CL} &= \theta_{12}^M + \theta_{23}^M + \theta_{31}^M \\ &= \theta_{12} + (\phi_1 - \phi_2) + \theta_{23} + (\phi_2 - \phi_3) + \theta_{31} + (\phi_3 - \phi_1) \\ &= \theta_{12} + \theta_{23} + \theta_{31}\ . \end{aligned} \tag{3.37}$$

Thus, this *closure* phase provides a good observable even in the presence of severe (antenna-based) phase errors, provided that the SNR is sufficient to define the visibility phase in a phase coherence time.

An array of N antennas has $(N-1)(N-2)/2$ independent closure phases. For large N this is close to the number of coherence phases $N(N-1)/2$ and so the loss of information about the object would seem to be small.

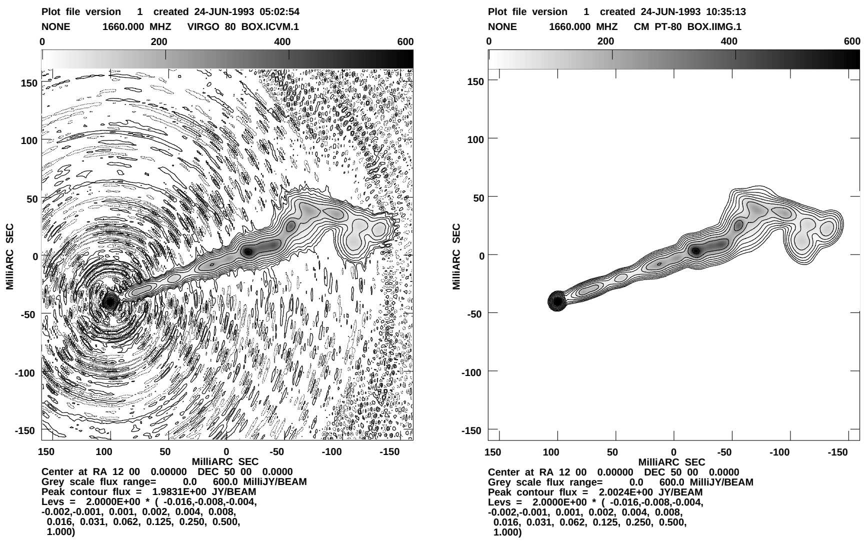

Figure 3.5: MEM deconvolution, 80 iterations, with support box: *Left:* with point source present. *Right:* point source removed prior to deconvolution.

The closure phase can be used in two ways:

1. If the array spacings are redundant then the closure phase can be used to predict all coherence phases up to a linear phase slope across the array. This requires a special geometry and redundancy of array (Noordam and de Bruyn, 1982) but no assumptions about the source structure are necessary.

2. As a constraint in an iterative image estimation procedure. Image plane constraints such as positivity and finite support make up for extra degrees of freedom (Readhead and Wilkinson, 1978, Pearson and Readhead, 1984). A typical hybrid mapping algorithm proceeds as follows:

   (a) Choose an initial model I^{Model}, e.g., a point source
   (b) Predict visibilities for the model $V^{\mathrm{Model}} = \mathcal{F}\left(I^{\mathrm{Model}}\right)$.
   (c) Keep observed visibility amplitudes
   (d) Select new visibility phases by modifying model visibility phases to be consistent with the observed closure phases
   (e) Form new dirty image from new visibilities
   (f) Form a new model I^{Model} by e.g., CLEANing the new dirty image.
   (g) Go back to step 1 unless the process has converged.

   The various algorithms differ in how new visibility phases are assigned (see, e.g., Pearson and Readhead, 1984).

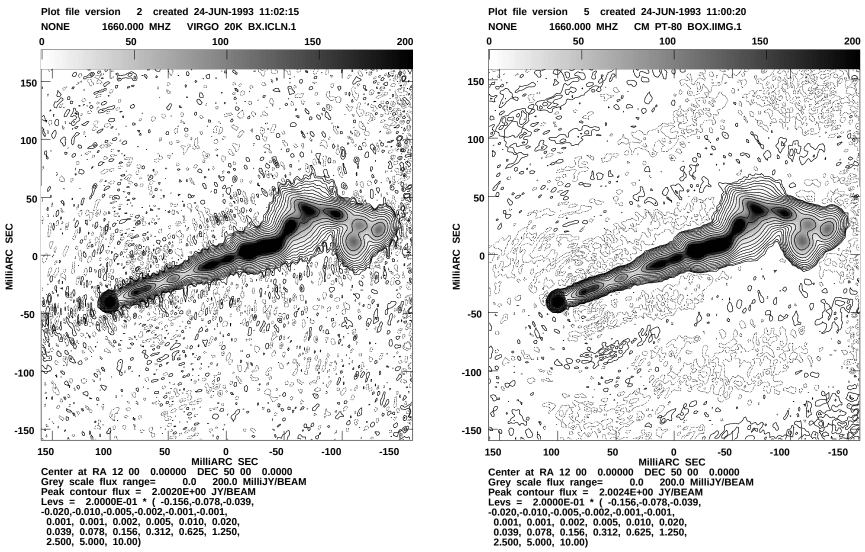

Figure 3.6: Comparison of best CLEAN and MEM deconvolution contoured at lower level. *Left:* CLEAN *Right:* MEM.

3.8 Self-calibration: Self-calibration

Closure phase methods cancel any antenna-based phase errors by the formation of the closure phase. In contrast, self-calibration estimates values for the phase errors and corrects the coherence samples accordingly.

Self-calibration can be performed in two ways:

1. If the array spacings are redundant then the antenna phase errors ϕ_i can be estimated from an over-determined set of linear equations (up to a linear phase slope across the array).

2. Deriving the phase errors as explicit degrees of freedom in an iterative image estimation procedure incorporating deconvolution. Image plane constraints make up for extra degrees of freedom in the phase errors. A generic self-calibration algorithm proceeds as follows:

   (a) Choose an initial model I^{Model}, e.g., a point source

   (b) Predict visibilities for the model $V^{\mathrm{Model}} = \mathcal{F}\left(I^{\mathrm{Model}}\right)$.

   (c) Keep observed visibility amplitudes

   (d) Select new visibility phases by modifying model visibility phases to be consistent with the observed closure phases

(e) Solve for the antenna phase errors using a least squares criterion:

$$\sum_{i,j} \left|\left|V_{ij}^M - e^{j(\phi_i - \phi_j)} V_{ij}^{model}\right|\right|^2 \; w_{ij} \,. \tag{3.38}$$

(f) Correct the observed visibilities for the antenna phase errors

(g) Form new dirty image from new visibilities

(h) Form a new model I^{Model} by, e.g., CLEANing the new dirty image.

(i) Go back to step 1 unless the process has converged.

The signal to noise ratio required is 2-3 per array antenna in the atmospheric coherence time.

3.9 Self-calibration: Comparison of Phase Closure and Self-calibration

Phase closure and self-calibration are two sides of the same coin.

1. Phase closure eliminates antenna phase errors from explicit consideration,
2. Self-calibration derives antenna phase errors explicitly.

One advantage of antenna-based self-calibration is that it is considerably simpler to program and to control than closure phase algorithms, especially for large numbers of antennas (Schwab 1980; Cornwell and Wilkinson 1981) More importantly, it also leads to generalization in a way that closure phase does not. For example, once derived, the phase error estimate can be applied to observations of another (nearby) object (*phase referencing*). Self-calibration can also be extended to phase derivatives with respect to time and frequency (*global fringe fitting*, see Schwab and Cotton 1983). Expanding the phase as a Taylor series in time and frequency, we have:

$$\phi_i(t,\nu) = \phi_i(t_0,\nu_0) + \frac{\partial \phi_i}{\partial t} \,.\, (t - t_0) + \frac{\partial \phi_i}{\partial \nu} \,.\, (\nu - \nu_0) \tag{3.39}$$

The first derivative is usually called the *fringe rate*, and the second is proportional to the *delay*. In global fringe fitting, these derivatives are estimated in the same way that the antenna-based phase is estimated. The first-order derivatives can then be removed and the visibility data averaged in time and frequency.

The closure phase has one main advantage over self-calibration: the closure phase is the phase of the *bispectrum* or *triple product* which is also immune to (antenna based) phase errors:

$$\begin{aligned} B_{123} &= V_{12}^M V_{23}^M V_{31}^M \\ &= e^{j(\phi_1-\phi_2)} V_{12} e^{j(\phi_2-\phi_3)} V_{23} e^{j(\phi_3-\phi_1)} V_{31} \\ &= V_{12} V_{23} V_{31} \,. \end{aligned} \tag{3.40}$$

For weak objects, the bispectrum can be integrated to build up SNR. However, for very weak objects, the gain in detectability goes as the sixth root of time (Cornwell 1987). Consequently, bispectrum methods tend to be useful only for weak objects with roughly constant visibility across the entire sampled Fourier plane.

For more on the physical connections of closure phase with other methods, see Cornwell and Fomalont (1989). For more on the mathematical structure of closure phase and self-calibration, see Lannes (1990).

References

Briggs, D. S. 1995. *High Fidelity Deconvolution of Moderately Resolved Sources.* Ph. D. thesis, New Mexico Institute of Mining and Technology.

Cornwell, T. J. 1987. "Radio-interferometric imaging of weak objects in conditions of poor phase stability: the relationship between speckle masking and phase closure methods". *Astron. Astrophys.* **180**, 269.

Cornwell, T. J., & Brown, R. 1989. "Deconvolution". See Perley, Schwab and Bridle (1989), p. 167.

Cornwell, T. J., & Evans, K. F. 1985. "A simple maximum entropy deconvolution algorithm". *Astron. Astrophys.* **143**, 77.

Cornwell, T. J., & Fomalont, E. 1989. "Self-calibration". See Perley, Schwab and Bridle (1989), p. 185.

Cornwell, T. J., Kemball, A. J., & Benson, J. M. 1995. "High dynamic range imaging with the VLBA". VLBA Scientific Memo. In preparation.

Cornwell, T. J., & Wilkinson, P. N. 1981. "A new method for making maps with unstable radio interferometers". *Monthly Notices Roy. Astron. Soc.* **196**, 1067.

Corwell, T. J., & Fomalont, E. B. 1989. "Self-calibration". See Perley, Schwab and Bridle (1989), p. 185.

Frieden, B. R. 1972. "Restoring with maximum likelihood and maximum entropy". *J. Opt. Soc. Am.* **72**, 511.

Gull, S. G., & Skilling, J. 1991. *MemSys5 quantified maximum entropy user's manual.*

Jaynes, E. T. 1957. "Informaton theory and statistical mechanics". *Phys. Rev.* **106**, 620.

Jennison, R. C. 1958. "A phase sensitive interferometer technique for the measurement of the Fourier transforms of spatial brightness distributions of small angular extent". *Monthly Notices Roy. Astron. Soc.* **118**, 276.

Lannes, A. 1990. "Remarkable algebraic structures of phase-closure imaging and their algorithmic implications in aperture synthesis". *J. Opt. Soc. Am.* **7**, 500.

Narayan, R., & Nityananda, R. 1986. "Maximum entropy image restoration in astronomy". *Ann. Rev. Astron. Astrophys.* **24**, 127.

Noordam, J., & de Bruyn, G. 1982. "High dynamic range mapping of strong radio sources, with application to 3C84". *Nature* **299**, 597.

Pearson, T. J., & Readhead, A. C. S. 1984. "Image formation by self-calibration in radio astronomy". *Ann. Rev. Astron. Astrophys.* **22**, 97.

Perley, R. A., Schwab, F. R., & Bridle, A. H. (Eds.) 1989. *Synthesis Imaging in Radio Astronomy.* San Francisco: Astronomical Society of the Pacific.

Readhead, A. C. S., & Wilkinson, P. N. 1978. "The mapping of compact radio sources from VLBI data". *Astrophys. J.* **223**, 25.

Schwab, F. 1980. "Adaptive calibration of radio interferometric data". *Proc. S. P. I. E.* **231**, 18.

Schwab, F. R., & Cotton, W. D. 1983. "Global fringe search techniques for VLBI". *Astron. J.* **88**, 688.

Schwarz, U. J. 1978. "Mathematical-statistical description of the iterative beam removing technique (method CLEAN)". *Astron. Astrophys.* **65**, 345.

Skilling, J., & Bryan, R. K. 1985. "Maximum entropy image reconstruction: General algorithm". *Monthly Notices Roy. Astron. Soc.* **211**, 111.

Sramek, R. A., & Schwab, F. R. 1989. "Imaging". See Perley, Schwab and Bridle (1989), p. 117.

Tan, S. 1986. "An analysis of the properties of CLEAN and smoothness stabilized CLEAN—some warnings". *Monthly Notices Roy. Astron. Soc.* **220**, 971.

Thompson, A. R., Moran, J. M., & Swenson, G. W., Jr. 1986. *Interferometry and Synthesis in Radio Astronomy*. New York: Wiley-Interscience. First (1991) and second (1994) reprintings by Krieger Pub. Co., Malabar (Florida).

Wilczek, R., & Drapatz, S. 1985. "A high accuracy algorithm for maximum entropy image restoration in the case of small data sets". *Astron. Astrophys.* **142**, 9.