

RADIO INTERFEROMETRY AND FOURIER SYNTHESIS IMAGING

T. S. BASTIAN

National Radio Astronomy Observatory
Charlottesville, VA

1 Introduction

These notes summarize the essentials of radio interferometry and Fourier synthesis imaging. They are based on several sources, notably *Synthesis Imaging in Radio Astronomy*, a collection of summer school lectures edited by Perley, Schwab, & Bridle, and *Interferometry and Synthesis in Radio Astronomy* by Thompson, Moran, & Swenson.

Radio interferometry dates back to the mid-1940s when the first radio interferometric observations were performed on the Sun. McCready, Pawsey, & Payne-Scott (1947) used a single antenna located on a high coastal cliff in Australia to point east as the Sun rose. A fringe pattern was recorded from the interference between the direct signal and that reflecting from the ocean surface below. This “sea interferometer” finds its analog at visible wavelengths in a Lloyd’s mirror. The paper of McCready et al. is the first to note the Fourier transform relationship between the quantity measured by an interferometer and the radio brightness distribution, a relationship which forms the basis of all modern Fourier synthesis telescopes.

The remainder of these notes is devoted to outlining the principles and practice of radio synthesis imaging, and to pointing out some of the difficulties and limitations associated with using these techniques. In §2 I introduce the fundamental relationship between the spatial coherence function and the sky brightness which forms the basis of synthesis imaging. In §3, I show how interferometry exploits this relationship in practice. In §4, I discuss image formation and reconstruction. In §5 I point out several limiting factors, and in §6 I take up two special topics: mosaicing and frequency synthesis.

2 Coherence and the van Cittert–Zernike Theorem

In general, an electromagnetic disturbance at a point \mathbf{r} and time t is described by a vector field, $\mathbf{E}(\mathbf{r}, t)$, which satisfies Maxwell’s equations. Although the field

varies rapidly and randomly in time, its fluctuations are ergodic in the sense that time-averaged measurements of the field are well-defined and are similar for all such measurements. We ignore the polarization properties of the electric field here and treat it as a scalar for the remainder of the discussion.

We define a “signal” as the response of a probe to the radiation field. The signal $v(\mathbf{r}, t)$ is proportional to $\mathbf{E}(\mathbf{r}, t)$ and, ignoring constants of proportionality, the intensity of the radiation $I(\mathbf{r})$ can be measured as

$$I(\mathbf{r}) = \langle v(\mathbf{r}, t)v^*(\mathbf{r}, t) \rangle,$$

where the asterisk denotes the complex conjugate.

Coherence theory, originally developed to describe the phenomenon of interference, is a statistical description of electromagnetic radiation. Coherence is a term which refers to the degree of correlation between two measurements of the radiation field. The *mutual coherence function* is defined by

$$\Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau) = \langle v(\mathbf{r}_1, t)v^*(\mathbf{r}_2, t + \tau) \rangle,$$

which is nothing more than the correlation function between random signals measured at two points \mathbf{r}_1 and \mathbf{r}_2 at times differing by τ . The *degree of coherence* is given by

$$\gamma(\mathbf{r}_1, \mathbf{r}_2, \tau) = \frac{\Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)}{[I(\mathbf{r}_1)I(\mathbf{r}_2)]^{1/2}}.$$

The radiation is said to be coherent if $|\gamma| = 1$, incoherent if $|\gamma| = 0$, and partially coherent if $|\gamma|$ lies somewhere in between. If the points \mathbf{r}_1 and \mathbf{r}_2 coincide, $\Gamma(\mathbf{r}, \mathbf{r}, \tau)$ is called the *autocorrelation function*. By the Wiener–Khinchin Theorem, the autocorrelation function is the Fourier transform of the power spectrum of the radiation field. If, on the other hand, the mutual coherence function is measured with $\tau = 0$, then $\Gamma(\mathbf{r}_1, \mathbf{r}_2, 0)$ is called the *spatial coherence function*. The spatial coherence of two points illuminated by a quasi-monochromatic source of incoherent radiation is given by the van Cittert–Zernike theorem, which dates back to the 1930s. It forms the basis of Fourier synthesis imaging.

Two assumptions are required: 1) that the source is far enough away that they are, for all intents and purposes, pasted onto the surface of the celestial sphere – that is, we have no hope of inferring anything about the depth of the emission and must be content with surface brightness. 2) That the emission is spatially *incoherent* at the source (spatial coherence at $(\mathbf{r}_1, \mathbf{r}_2)$ acquired from propagation). The van Cittert–Zernike Theorem then states that the spatial coherence function is related to the sky brightness as follows:

$$\Gamma(\mathbf{r}_1, \mathbf{r}_2, 0) = \int I(\mathbf{s}) e^{-2\pi i \mathbf{s} \cdot (\mathbf{r}_1 - \mathbf{r}_2) / c} d\Omega.$$

Here \mathbf{s} is a unit vector directed toward the source, $I(\mathbf{s})$ is the intensity, and $d\Omega$ an element of solid angle. Note that a measurement of the spatial coherence function depends only on the relative difference between \mathbf{r}_1 and \mathbf{r}_2 .

3 Interferometry in Practice

3.1 The Response of an Interferometer

The basic measuring device in synthesis mapping is the interferometer. An interferometer is composed of two antennas (presumed identical) separated by a distance \mathbf{b} . Let the two antennas point toward a distant radio source in a direction indicated by the unit vector \mathbf{s} . In general, the plane wave incident on the interferometer arrives at one antenna first, and then at the other. The time delay between the two is $\tau_g = \mathbf{b} \cdot \mathbf{s} / c$ and is called the *geometrical delay*. The signals from the antennas pass through amplifiers and filters which select the frequency of interest with bandwidth $\Delta\nu$. The voltage signal responses produced by each antenna are multiplied together and time-averaged in a device called a correlator. That is, for input voltages from the two antennas, $V_1(t)$ and $V_2(t)$, the correlator output is proportional to $\langle V_1(t)V_2(t) \rangle$. Representing the two voltage signals as

$$V_1(t) = v_1 \cos 2\pi\nu(t - \tau_g); \quad V_2(t) = v_2 \cos 2\pi\nu t$$

the correlator output is then

$$r(\tau_g) = v_1 v_2 \cos 2\pi\nu\tau_g.$$

Note that compensation occurs for the geometrical delay before the correlation is performed. The result is the interferometer fringe pattern. The correlator output can be recast in terms of the radio brightness integrated over the sky. Let $I(\mathbf{s})$ be the sky brightness in the direction \mathbf{s} at the frequency ν . If $A(\mathbf{s})$ is the effective collecting area of an antenna in direction \mathbf{s} , the signal power received by each antenna over a bandwidth $\Delta\nu$ in a solid angle element $d\Omega$ is $A(\mathbf{s})I(\mathbf{s})\Delta\nu d\Omega$. So the correlator signal per solid angle element $d\Omega$ is

$$dr = A(\mathbf{s})I(\mathbf{s})\Delta\nu d\Omega \cos 2\pi\nu\tau_g,$$

Integrating over the celestial sphere, we obtain

$$r = \Delta\nu \int A(\mathbf{s})I(\mathbf{s}) \cos 2\pi\nu \frac{\mathbf{b} \cdot \mathbf{s}}{c} d\Omega.$$

In practice, the angular response of the antenna elements falls rapidly to small values outside of a narrow angular width defined by their diameter. It is usually more convenient to refer measurements to a reference position \mathbf{s}_o , commonly referred to as the *phase tracking center*. Then we have $\mathbf{s} = \mathbf{s}_o + \sigma$ and

$$r = \Delta\nu \cos\left(2\pi\nu \frac{\mathbf{b} \cdot \mathbf{s}_o}{c}\right) \int A(\sigma)I(\sigma) \cos \frac{2\pi\nu \mathbf{b} \cdot \sigma}{c} d\Omega$$

$$- \Delta\nu \sin\left(2\pi\nu \frac{\mathbf{b} \cdot \mathbf{s}_o}{c}\right) \int A(\sigma)I(\sigma) \sin \frac{2\pi\nu \mathbf{b} \cdot \sigma}{c} d\Omega.$$

Defining the complex *visibility* as

$$V \equiv |V|e^{i\phi_V} = \int A'(\sigma)I(\sigma)e^{-2\pi i\nu \mathbf{b} \cdot \sigma/c} d\Omega, \quad (1)$$

where $A'(\sigma) \equiv A(\sigma)/A_o$ is the normalized antenna beam pattern, we obtain

$$r = A_o \Delta\nu |V| \cos\left(\frac{2\pi\nu \mathbf{b} \cdot \mathbf{s}_o}{c} - \phi_V\right). \quad (2)$$

It is now clear that an interferometer is a machine for measuring the visibility, which is nothing more than the spatial coherence function with a different normalization. By measuring the amplitude and phase of the fringe pattern in eqn. (3), the amplitude and phase of the visibility is determined after application of suitable calibration. Eqn.(2) may then be inverted.

3.2 Coordinate System for Imaging

In order to make use of eqn.(2) in practice, a convenient coordinate system must be introduced. A commonly used (but not universal) system is one where the baseline vector is specified in a coordinate system by (u, v, w) where w is chosen toward the phase tracking center (\mathbf{s}_o), u is toward the east and v toward the north. The coordinates (u, v, w) are measured in wavelengths. Positions on the sky are defined in l and m , the direction cosines measured with respect to the u and v axes. Thus a synthesized image in the $l - m$ plane represents a projection of the celestial sphere onto a plane tangent to the $l - m$ origin. In these coordinates we have

$$\frac{\nu \mathbf{b} \cdot \mathbf{s}}{c} = ul + vm + wn, \quad \frac{\nu \mathbf{b} \cdot \mathbf{s}_o}{c} = w$$

$$d\Omega = \frac{dl \, dm}{n} = \frac{dl \, dm}{\sqrt{1-l^2-m^2}}$$

so that eqn.(2) can be rewritten as

$$V(u, v, w) = \int \int A'(l, m) I(l, m) e^{-2\pi i[ul+vm+w(\sqrt{1-l^2-m^2}-1)]} \frac{dl \, dm}{\sqrt{1-l^2-m^2}} \quad (3)$$

This reduces to a two-dimensional Fourier transform relation when $|l|$ and $|m|$ are sufficiently small, since

$$w(\sqrt{1-l^2-m^2}-1) \approx -\frac{1}{2}(l^2+m^2)w \approx 0.$$

Then eqn.(4) reduces to

$$V(u, v) = \int \int A'(l, m) I(l, m) e^{-2\pi i(ul+vm)} dl \, dm. \quad (4)$$

This can be inverted to give

$$A'(l, m) I(l, m) = \int \int V(u, v) e^{2\pi i(ul+vm)} du \, dv. \quad (5)$$

To summarize, an interferometer is a device for measuring the amplitude and phase of the complex visibility function, a function which is nothing more than the spatial coherence function (albeit with a different normalization). By the van Cittert-Zernike theorem, the visibility is related to the sky brightness. If the measurement of the visibility function is confined to a plane, or if only a small region of the sky is considered, $V(u, v)$ and $I(l, m)$ reduce to a Fourier transform pair, as given by eqns. 5 & 6.

4 Synthesis Imaging

4.1 Image Formation

The fundamental result of the previous section is that, under a not-too-burdensome variety of constraints and assumptions, there is a Fourier transform relationship between the sky brightness, I , and the visibility function V . Furthermore, it has been shown that an interferometer can be used to measure the visibility function.

I will avoid a detailed discussion of how visibility measurements are calibrated, saying only that instrumental parameters are generally determined by observing simple sources (point sources) of known flux density and position, although there

are exceptions. For cosmic sources, V is usually calibrated in units of Janskys ($1 \text{ Jy} = 10^{-26} \text{ W m}^{-2} \text{ Hz}^{-1}$ and, in the image domain, I has units of flux density/beam volume (flux density per unit of solid angle). In the case of solar observations, V is usually calibrated in units of solar flux units ($1 \text{ sfu} = 10^4 \text{ Jy}$).

In practice, antenna arrays are used to measure the visibility function on many baselines and therefore, at many spatial frequencies. However, since one is always limited to a finite number of antennas and a finite amount of time over which to make the observation, there will always be gaps or holes in the measurement of the visibility function. Furthermore, measured visibilities are always corrupted by noise (e.g., thermal noise due to the receivers or, in the case of the Sun, due to the source itself). I designate the noisy visibility measurement by $V'(u, v)$. The set of spatial frequencies at which visibility measurements were made can be written in terms of a sampling function $S(u, v)$. Then eqn. (6) can be recast in the form

$$I^D(l, m) = \int \int S(u, v) V'(u, v) e^{2\pi i(ul+vm)} du dv \quad (6)$$

where $I^D(l, m)$ is referred to as the “dirty image”. Note that I have absorbed the antenna response $A(l, m)$ into $I^D(l, m)$. Giving the sampling function a more explicit form,

$$S(u, v) = \sum_{k=1}^M \delta(u - u_k, v - v_k),$$

the sampled visibility function can then be written

$$V^S(u, v) = \sum_{k=1}^M \delta(u - u_k, v - v_k) V'(u, v).$$

Letting \mathbf{F} denote the (inverse) Fourier transform operator. Eqn. (7) can then be written

$$I^D = \mathbf{F}(V^S) = \mathbf{F}(SV') = \mathbf{F}(S) \star \mathbf{F}(V') = \mathbf{F}(S) \star I'$$

where \star denotes a convolution. Hence the dirty image is the convolution of the inverse Fourier transform of the sampling function with the “true” (albeit noise corrupted) brightness distribution; $\mathbf{F}(S)$ is commonly referred to as the “dirty beam” or the point spread function since, when the source is a point of unit amplitude at the phase tracking center, the visibility function is unity everywhere.

In analogy to the sampling function, a weighting function can be written as

$$W(u, v) = \sum_{k=1}^M T_k D_k \delta(u - u_k, v - v_k),$$

and the weighted and sampled visibility function V^W is then

$$V^W(u, v) = \sum_{k=1}^M T_k D_k \delta(u - u_k, v - v_k) V'(u, v).$$

I have ignored a possible instrumental weighting of the visibility data and only consider weighting which can be controlled during image formation. Both the density weight D_k and the tapering function T_k are used to control the shape of the dirty beam, which is now written $B = \mathbf{F}(W)$. It is common to use a simple Gaussian for the tapering function. Its purpose is to smoothly weight down the highest spatial frequencies in order to suppress smallscale sidelobes, although high-resolution information is also suppressed as a result.

The density weighting is used to emphasize or de-emphasize one part of the sampling function over another. The need for this is as follows: many synthesis arrays sample many more short baselines than long baselines. Short baselines therefore receive relatively more weight than long baselines in forming the image. It is often desirable to have all parts of the sampled visibility function contribute uniformly to I^D (the best angular resolution is achieved in this way). This may be achieved by letting $D_k = 1/N_s(k)$, where N_s is the number of visibility samples within a specified area around (u_k, v_k) . In densely sampled parts of the visibility domain, the data are weighted down relative to sparsely sampled regions. This kind of weighting is called *uniform weighting*. The opposite extreme is no density-based weighting ($D_k = 1$), which is called *natural weighting*. Uniform weighting yields higher angular resolution and, often, a more pleasing dirty beam. In contrast, natural weighting yields a broader dirty beam, but higher sensitivity since all data receive full weight. Optimum weighting usually lies somewhere between these two extremes.

The dirty image is commonly calculated in one of two ways, both involving a discrete representation of $I^D(u, v)$. A discrete representation involves evaluating $I^D(l, m)$ on a uniform grid of pixels. One method is to compute the “direct Fourier transform” at each grid point of an $N \times N$ matrix by evaluation of the sum

$$\frac{1}{M} \sum_{k=1}^M V'(u_k, v_k) e^{2\pi i(u_k l + v_k m)},$$

where M is the number of visibilities. Obviously, this method is impractical if either N or M becomes large. The other method – and by far the most common – is

to interpolate the visibility measurements onto a regular grid so the Fast Fourier Transform (FFT) algorithm can be used. The process of interpolating the data onto a uniform grid is referred to as *gridding* the data. It involves convolving the weighted, sampled data with some suitable function C . The convolution is then resampled on a uniform grid. Let the resampling function be written

$$R(u, v) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \delta(j - u/\Delta u, k - v/\Delta v),$$

where Δu and Δv define the separation between grid points (R is Bracewell's "sha" function). The convolution of C with the weighted visibility is written $C \star V^W$ and resampling is written as

$$V^R = R(C \star V^W) = R(C \star (WV')).$$

Since V^R is sampled on a regular grid, $\mathbf{F}V^R$ can be computed using the FFT algorithm. From this we obtain an estimate of I^D :

$$\tilde{I}^D = \mathbf{F}R \star [(\mathbf{F}C)(\mathbf{F}V^W)] = \mathbf{F}R \star [(\mathbf{F}C)(\mathbf{F}W \star \mathbf{F}V')].$$

Appropriate normalization is usually required, and correction for the convolution function C – that is, \tilde{I}^D is divided by $\mathbf{F}C$. The dirty beam is computed in a similar fashion, with

$$\tilde{B}^D = \mathbf{F}R \star [(\mathbf{F}C)(\mathbf{F}W)].$$

Because of the presence of the resampling function R , some care must be exercised in choosing the convolution function C . The reason is that by resampling the visibility data, one effectively makes \tilde{I}^D a periodic function of l and m with a period $1/\Delta u$ in l and $1/\Delta v$ in m . If $\mathbf{F}C$ does not fall to zero sufficiently rapidly outside the primary field of view ($1/\Delta u, 1/\Delta v$), parts of the sky brightness that lie outside of primary field of view are aliased back into the field of view.

4.2 Image Reconstruction

The general problem of image reconstruction is summarized by the *measurement equation*

$$\mathbf{I}^D = \mathbf{B} \star \mathbf{I} + \mathbf{n}$$

where $\mathbf{I}^{\mathbf{D}}$ is the dirty map, \mathbf{I} is the ideal or “true” image, \mathbf{B} represents the instrumental point spread function or dirty beam, and \mathbf{n} represents additive noise. As before, \star indicates a convolution. Image reconstruction involves recovering \mathbf{I} through various deconvolution schemes.

Two aspects of the measurement equation make the process of image deconvolution a difficult one. First, there is additive noise. One might suppose that one could simply Fourier transform the measurement equation, divide both sides by the Fourier transform of \mathbf{B} , and then compute the inverse Fourier transform to recover \mathbf{I} . For filled apertures, such linear techniques are in fact used, although in the presence of noise, division of the Fourier transform of the measurement equation amplifies the noise and introduces high-frequency artefacts into \mathbf{I} . Even so, linear deconvolution techniques have been developed (e.g., Wiener filtering), which optimize the deconvolution in the presence of noise.

No such techniques can be used to invert the measurement equation for synthesis mapping because \mathbf{B} is singular. In other words, the Fourier transform of \mathbf{B} , which represents the sampling function, has many spatial frequencies where no measurement of the visibility function was made. The *principle solution* assumes that the amplitude of the visibility function is zero at all unsampled points. Let \mathbf{z} be a distribution of brightness which contains only the unmeasured spatial frequencies; then $\mathbf{B} \star \mathbf{z} = 0$. If \mathbf{I} is a solution of the measurement equation, then so too is $\mathbf{I} + \alpha \mathbf{z}$. Many solutions therefore exist to the measurement equation. The measurement equation is formally “ill-posed” and nonlinear deconvolution techniques must be employed.

Deconvolution in these circumstances involves estimating the visibility function at those spatial frequencies where it was not measured. Of the many solutions $\mathbf{I} + \alpha \mathbf{z}$ possible, the trick is to identify that which is somehow “best”. Two deconvolution algorithms have been prevalent over the years: the Högbom CLEAN algorithm and maximum entropy method (MEM) deconvolution, and related techniques.

4.2.1 The CLEAN Algorithm

The dirty map $\mathbf{I}^{\mathbf{D}}$ is a convolution between the “true” brightness distribution and the dirty beam. Each point in the field of view has been multiplied by the dirty beam, and the dirty map may be regarded as a superposition of point sources, each multiplied by the dirty beam. The Högbom CLEAN is an algorithm for dividing each point by the dirty beam. However, because \mathbf{B} is singular, it must do so in a nonlinear fashion. Just as multiplication is repeated addition, division is repeated subtraction. The essence of the Högbom CLEAN is repeated “shift-and-subtraction” of the dirty beam.

Inherent to the Högbom CLEAN algorithm is the assumption that radio sources are relatively spatially confined and are thus well-approximated by an appropriately weighted superposition of delta functions. In practice, the method works remarkably well. The deconvolution proceeds as follows: 1) the pixel with the maximum

brightness is located in the map; 2) the dirty beam is multiplied by the product of the maximum brightness and a gain factor $g < 1$; 3) the weighted dirty beam is shifted to the location of the maximum and subtracted from the map; 4) the algorithm loops back to (1). The process is iterative and continues for a specific number of iterations or until the residual map reaches some specified level. At this point, the ensemble of weighted delta functions are added back to the residual map. The addition of a δ function in the image domain is equivalent to adding a sine wave of some amplitude and phase to the visibility function. In this way, “holes” in the uv domain are filled. Usually, each delta function is first multiplied by a “clean beam”, a Gaussian with a FWHM similar to the intrinsic resolution of the instrument. In the uv domain, this is equivalent to multiplying the visibility function by a broad gaussian with a width corresponding to the size of the array.

The appeal of the Högbom CLEAN is its simplicity and the fact that it often works! However, for extended, complicated sources, it can break down – it can become unstable and introduce artefacts into the reconstruction. Alternatives must be considered. One group of alternatives are maximum entropy or maximum-entropy-like deconvolution schemes.

4.2.2 The MEM Algorithm

Under some circumstances, the source brightness distribution is poorly described as a superposition of weighted delta functions. An attractive alternative to the CLEAN algorithm is image regularization through optimization of a suitably constrained functional of the data. One such functional is the configurational “entropy” of the image. The configurational entropy, H , of an N pixel estimate $\{b_i | i = 1, 2, \dots, N\}$ of \mathbf{I} , relative to some measure $\{m_i | i = 1, 2, \dots, N\}$ is

$$H = - \sum_{i=1}^N b_i \log(b_i/m_i).$$

There has been much confusion and anxiety over the use of this function and the term used to describe it. Suffice it to say that it is a convenient function in a practical sense because it 1) enforces positivity of the solution; 2) minimizes the dispersion in pixel values (i.e., enforces maximum “smoothness” of the image - detail does not appear unless it is found in the data); 3) it allows the introduction of a *priori* information about the source via m . It should be pointed out, however, that there are other functional forms which serve this purpose equally well. In practice, MEM deconvolution is nothing more than a model/data comparison algorithm wherein a model is iteratively constructed which 1) maximizes some entropy measure, and 2) matches the data to within the noise \mathbf{n} . A requirement that the flux F in the model match the zero-spacing flux is also usually enforced. Hence, a functional of the form

$$J = H - \alpha\chi^2 - \beta F$$

is employed, where α and β are Lagrange multipliers.

A key advantage of MEM deconvolution techniques is that they enable one to introduce independent information about the source. For example, data from more than one instrument can be used to jointly constrain the brightness distribution (e.g., single dish plus interferometric observations), or from more than one pointing (mosaicing – see §7.1). In the case of the solar imaging, MEM deconvolution has been used to reconstruct full disk images at 21 cm using data from both the VLA and the Arecibo 305 m telescope.

4.3 Sensitivity

In concluding this section on imaging, it is worth making a few points concerning the sensitivity of an instrument. In particular, it's worth distinguishing between point-source sensitivity and surface-brightness sensitivity, or the sensitivity of the instrument to extended emission.

4.3.1 Point Source

The sensitivity of an array to a point source of emission is given by

$$\Delta I = \frac{\sqrt{2}k_B T_{sys}}{\eta_a \eta_c A \sqrt{N_b N_{IF} T \Delta\nu}} \quad (7)$$

where ΔI is the rms variation in the brightness I , k_B is Boltzmann's constant, η_a is the aperture efficiency, η_c is the correlator efficiency, and A is the antenna area (presumed identical for all antennas), $N_b = N(N - 1)/2$ is the number of antenna baselines, N_{IF} is the number of independent IF channels, T is the total observing time on source, and $\Delta\nu$ is the bandwidth. The system temperature T_{sys} embodies many sources of noise, including that due to the receivers, ground spillover, the atmosphere (at high frequencies), the galactic background (at low frequencies), the source itself, and others. When cosmic sources are observed, the receiver noise usually dominates T_{sys} , (although this is becoming less true as receivers have steadily improved). When observing the Sun with a telescope of even minimal forward gain, the radiation flux from the Sun itself dominates T_{sys} . If the noise were completely uncorrelated at all antennas, the system could be regarded as just having especially poor receivers. However, the situation is complicated by the additional contribution to the noise produced by the correlated flux S_C on each baseline (“self-noise”). Under many circumstances, the correlated flux is negligible compared to the total. When this is not the case (e.g., for the shortest baselines in compact antenna configurations or during a strong, compact microwave burst) one can no longer ignore the contribution to the noise due to the correlated flux on each baseline.

4.3.2 Extended Emission

It is important to emphasize the difference between the sensitivity of a synthesis array to a point source, and that to an extended source of emission. The brightness of a source is usually expressed in terms of flux per synthesized beam. A point source has a constant flux per synthesized beam, regardless of the size (i.e., resolving power) of that beam. The signal-to-noise ratio (SNR) is therefore independent of the beam size. For an extended source which is resolved by the synthesized beam the situation is different. If the surface brightness of the source is constant, say I , the flux density per synthesized beam, Ω_s , is $I\Omega_s$ and the SNR is $I\Omega_s/\Delta I$. The SNR increases with Ω_s , which is desirable so long as Ω_s is smaller than the extended structure of interest. The angular resolution of an array tends to increase linearly with the size of the array, but its sensitivity to extended emission *decreases as the square of the array size*. Clearly, the best sensitivity to extended emission is achieved by matching the resolution of the instrument to the angular scale of the emission. Short of moving antennas to a configuration matched to a particular source, this can be achieved through a variety of schemes which weight the visibility data in various ways (e.g., through tapering of the visibility function and/or weighting the visibilities according to their distribution in the uv domain).

5 Limiting Factors in Synthesis Imaging

When one or more of the assumptions inherent to synthesis imaging are violated, artefacts of varying severity are introduced into the image. I'll mention a few.

5.1 Bandwidth Smearing

So far we've considered interferometry and synthesis mapping in terms of quasi-monochromatic radiation. In reality, radio telescopes observe radiation with a finite bandwidth $\Delta\nu$ centered on some frequency ν_o . One of the consequences of this reality is *bandwidth smearing*, or chromatic aberration. The spatial frequency coordinates (u_o, v_o) are typically computed with reference to ν_o . However, a range of spatial frequencies are in fact measured, with $(u, v) = (u_o\nu_o/\nu, v_o\nu_o/\nu)$. Since the process of correlation and Fourier transformation are linear, we can therefore consider a synthesized image obtained with finite bandwidth as a sum of contributions from all parts of the frequency bandpass. The effect on the image can be shown through an application of the similarity theorem of Fourier transforms:

$$V\left(\frac{\nu_o}{\nu}u, \frac{\nu_o}{\nu}v\right) \Rightarrow \left(\frac{\nu}{\nu_o}\right)^2 I\left(\frac{\nu}{\nu_o}l, \frac{\nu}{\nu_o}m\right).$$

What this expression shows is that the image is scaled radially by a factor ν/ν_o for each frequency within the band, and the amplitude is multiplied by $(\nu/\nu_o)^2$ (which

conserves flux). The net effect of a finite bandwidth, then, is to smear the image in the radial direction. The degree of smearing increases with angular distance from the phase tracking center; that is, the smearing goes as $(\Delta\nu/\nu)\sqrt{l^2 + m^2}$.

An approximate expression for the reduction in brightness due to bandwidth smearing is

$$\frac{I}{I_0} \approx 1 - \frac{1}{3} \left(\frac{\gamma \Delta\nu \theta}{2\nu_0 \theta_S} \right)^2 - \frac{1}{10} \left(\frac{\gamma \Delta\nu \theta}{2\nu_0 \theta_S} \right)^4 \quad (8)$$

where $\gamma = 2\sqrt{\ln 2} = 1.665$, θ is the angular offset from the phase tracking center, θ_S is the angular width of the synthesized beam. For example, suppose we observe the full disk of the Sun with an array of 5 km extent at a frequency of 10 GHz ($\theta_S \approx 0.8''$). A bandwidth of only 1.2 MHz yields a 2% reduction in brightness due to bandwidth smearing at the solar limb ($\theta \approx 16'$). Since relatively large instantaneous bandwidths are being considered for modern radioheliographs (~ 500 MHz) it appears that it will need to be channelized prior to correlation.

5.2 Temporal Smearing

In §4.2 it was shown that an interferometer baseline traces out an ellipse in the uv plane. The rate at which u and v change with time is given by

$$\begin{aligned} \frac{du}{dt} &= \frac{1}{\lambda} (L_X \cos H - L_Y \sin H) \frac{dH}{dt} \\ \frac{dv}{dt} &= \frac{1}{\lambda} (L_X \sin \delta \sin H + L_Y \sin \delta \cos H) \frac{dH}{dt}. \end{aligned}$$

Hence, for a finite integration time, a long baseline traces out a longer uv track than a short baseline. A single (u, v) point is assigned to the time-averaged measurement, however. The effect of time averaging is therefore equivalent to a distorted azimuthal smearing; that is, the time-averaged image may be regarded as the sum of snapshot maps acquired at each instant of time included in the integration. As the array geometry rotates slightly during the integration, so do the snapshot maps. The reduction in brightness due to time-averaging is

$$\frac{I}{I_0} \approx 1 - \frac{\pi^2}{6} \left(\frac{du}{dt} l + \frac{dv}{dt} m \right)^2$$

which can be written approximately (under certain assumptions which are not listed here) as

$$\frac{I}{I_0} \approx 1 - \frac{\gamma^2}{12} \omega_e^2 \tau_a^2 \left(\frac{\theta}{\theta_S} \right)^2$$

Here, ω_e is the angular rotation rate of the Earth and τ_a is the averaging time. A convenient “rule-of-thumb” is to choose an averaging time which yields the same degree of smearing as does the finite bandwidth. This may be expressed as

$$\tau_{\Delta\nu} \approx \frac{\Delta\nu}{\omega_e \nu_o} = 1.375 \times 10^4 \frac{\Delta\nu}{\nu_o}. \quad (9)$$

For a bandwidth $\Delta\nu = 1.2$ MHz at a frequency of 10 GHz, the integration time should be no more than about 1 sec.

5.3 The w term

So long as synthesis imaging is restricted to a region of small angular extent, the w term can be neglected (see §4.2), and a two dimensional Fourier transform relationship between $I(l, m)$ and $V(u, v)$ holds. How small a region is small enough?

By neglecting w , a phase error of order $\pi w \theta^2$ is introduced. For a coplanar array (such as the VLA), w is a linear function of u and/or v and the phase error increases linearly across the $/uv/$ plane. It therefore manifests itself as an apparent shift of the source position (which is dependent on the zenith angle and azimuth of the source). Since the array geometry, as viewed from the source, changes with time, the apparent source position also changes. The time-integrated sampling is no longer coplanar. Again, the net result is a distortion across the field of view. The magnitude of the apparent position shift in arcsec for a coplanar array is given approximately by

$$\Delta\theta \approx \frac{\theta^2}{4.125 \times 10^5} \sin z \quad (10)$$

where z is the instrumental zenith angle. For a source on the zenith, the error is negligible. If the source is at $z \sim 60^\circ$, and $\theta = 16'$, the position error amounts to $\approx 2''$.

It is important to note that for a snapshot map, an array like the VLA is coplanar and a correction can be made for the distortion introduced by neglect of the w term. When earth-rotation aperture synthesis is performed, the baselines are no longer coplanar, and corrections for neglect of the w term cannot be easily made. Instead, the w term must be included explicitly and a three-dimensional Fourier transformation must be formulated.

5.4 Limited uv Coverage

The difficulties presented by limited uv coverage are twofold. The first problem, the “hole” at the center of the uv plane whose radius is defined by the shortest (projected) interferometer baseline, is not specific to solar observing. The origin of the hole is in the existence of a minimum antenna spacing, which corresponds to

the maximum angular scale measured. The problem can be particularly acute for solar imaging because most of the power emitted by the (nonflaring) Sun resides in the largest angular scales. Unless these scales are adequately measured or otherwise estimated, artifacts are present in the resulting maps, and the observed brightness distribution is in error.

At the VLA this problem has sometimes been circumvented through the use of the powerful image reconstruction and deconvolution algorithms now available. Maximum entropy algorithms, briefly described in §5.2.2 provide a convenient way to introduce measurements made with more than one instrument to the image deconvolution problem. Hence, programs that depend on information contained in the shortest baselines may acquire it from a separate, single-dish observation of the Sun (e.g., by using the 305 m telescope at Arecibo).

For the remainder of this section I'll describe some problems which have a more direct bearing on solar imaging. One is closely related to the limitations of uv coverage, which I now describe.

5.5 Spatial Confusion

For single dish observation, confusion limits the instrumental sensitivity via contributions to T_{sys} from background sources in the beam. In synthesis imaging, the limitation to sensitivity imposed by confusion is the result of uncleanable sidelobe "clutter" in the image. The role of confusion in limiting the sensitivity of synthesis arrays has long been recognized as a problem when imaging cosmic sources at low frequencies, where the number of background sources in the primary beam is largest. Perley and Erickson (1984) have shown that if the baseline of a given interferometer is much larger than the element size (so that a large number of fringes cross the field of view), the interferometer response to a large number of point sources in the beam can be regarded as a random variable. It can then be shown that the confusion "noise", σ_c , is given by

$$\sigma_c = \rho \sqrt{\int A^2(\Omega) d\Omega \int S^2 n(S) dS}$$

where A is the primary beam response, $n(S)$ is the source distribution function (i.e., the number of sources with flux densities between S and $S + dS$), and ρ is the rms fluctuation of the synthesized beam. In other words, the confusion "noise" may be expressed as an appropriately weighted, incoherent sum of the sidelobe responses to all sources in the primary beam. The confusion noise can be reduced to the extent that sidelobe responses to individual sources in the primary beam can be identified and removed.

The situation with regard to the Sun is complicated by the fact that structure is present in the beam on a wide range of angular scales. Assigning a source distribution

function to the brightness distribution function is problematic. Nevertheless, the importance of the expression for σ_c is that it depends linearly on ρ . Confusion noise can be reduced to the extent that the rms fluctuations of the synthesized beam can be reduced. Improved uv coverage reduces ρ and increases dynamic range. Often, observers rely on long scans for improved uv coverage. Unfortunately, in the case of solar observations, the Sun's radio brightness distribution *varies* in time due to (1) the Sun's rotation and (2) intrinsic source variability. Another problem that must be confronted is therefore source variability.

5.6 Source Variability and Temporal Confusion

The improved dynamic range that one might hope to obtain through improved uv coverage may be sabotaged by variations in the source brightness on a timescale less than the duration of the observation. First, consider the Sun's rotation. Since the Sun's equatorial rotation velocity is $\sim 2 \text{ km s}^{-1}$, a source at the center of the Sun's disk (e.g., an active region) will apparently move away from the center of the disk at $9.3 \text{ arcsec hr}^{-1}$. On the other hand, since the Sun is spherical, sources near the limb apparently move very little due to rotation. In other words, the brightness distribution changes differentially.

The success with which one can eliminate the effects of solar rotation depends on the type of imaging program. If the field of interest is small compared to the radius of the Sun, one may choose to explicitly correct for the rotation of the Sun by tracking a particular feature, as is often done at the VLA. As the field becomes large, however, parts of the image will be smeared and/or distorted by differential motion across the field. Ideally, the region should be imaged in a time that is short compared to that for significant degradation of the image by the Sun's rotation. I return to this point below.

I now turn to the problem of intrinsic variability in the radio brightness distribution. Consider a discrete source at some location (l, m) with brightness $S \text{ Jy/beam}$. Let its lifetime be τ but suppose the observation is of duration $T > \tau$. Then while it is true that, limitations imposed by confusion aside for the moment, the rms fluctuations on the image are reduced by $\sqrt{T/\tau}$ relative to those expected from an integration of duration τ , it is also true that the brightness at (l, m) is averaged down to $\tau S/T$. The net signal-to-noise ratio at (l, m) is thus reduced by $\sqrt{\tau/T}$. In other words, the temporal smearing of shortlived contributions to the Sun's radio brightness distribution may lead to a net reduction in effective dynamic range.

A further complication is that if the radio brightness distribution on the sky varies as a function of time, due to either solar rotation or intrinsic variability, the convolution relation between the brightness distribution and the instrument response function is no longer valid. The situation is further exacerbated by the fact that the temporal variability in turn depends on position. The consequent breakdown of deconvolution algorithms again limits the final dynamic range.

What is required, then, is improved uv coverage over time scales that are short in comparison with the lifetime of the phenomenon of interest, or compared to the time for solar rotation to have a significant effect. I return to this point in §7.2

6 Special Topics

6.1 Mosaicing

When the object one wishes to image is larger than the field of view of the instrument, one must resort to *mosaicing* techniques. As the term implies, mosaicing involves observing many fields, which are then patched together to form an image of large angular size. One way to do this is through “linear mosaicing” wherein observations acquired at many different pointings are separately mapped, deconvolved, and then linearly combined with appropriate weighting. While linear mosaicing is sufficient for some purposes, it does not recover the brightness distribution on large angular scales.

An alternative is “nonlinear mosaicing”. The technique is well-developed and has rigorous underpinnings. It finds its basis in work of Ekers and Rots (1979), but its present-day form, motivated by synthesis observations at mm wavelengths, is largely due to Cornwell and collaborators.

As shown in §2, an interferometer is a device which measures $V(u, v)$, the visibility function. The visibility function is related to the sky brightness, $I(l, m)$, through a Fourier transform:

$$V(u, v) = \iint A(l, m)I(l, m)e^{-2\pi i(ul+vm)} dl dm,$$

where $A(l, m)$ is the primary beam pattern. $A(l, m)$, of course, defines the field of view and, as such, is a limiting factor for imaging large objects. Let us rewrite the above equation with the pointing (l_o, m_o) explicitly labeled:

$$V(u, v; l_o, m_o) = \iint A(l, m; l_o, m_o)I(l, m)e^{-2\pi i(ul+vm)} dl dm.$$

This form implicitly separates the pointing from the phase and delay tracking center. If we assume that the primary beam is independent of pointing, we can then write

$$V(u, v; l_o, m_o) = \iint A(l - l_o, m - m_o)I(l, m)e^{-2\pi i(ul+vm)} dl dm.$$

Note that when $(u, v) = (0, 0)$, this becomes

$$V(0, 0; l_o, m_o) = \iint A(l - l_o, m - m_o) I(l, m) dl dm,$$

which represents a single-dish measurement of the total power coming from the region of sky within the primary beam.

To carry things further, let $i(u, v)$ be the Fourier transform of $I(l, m)$, $a(u, v)$ the Fourier transform of the primary beam, and (u_o, v_o) be the variables conjugate to (l_o, m_o) . Then the Fourier transform of $V(u, v; l_o, m_o)$ with respect to (l_o, m_o) is

$$\iint V(u, v; l_o, m_o) e^{-2\pi i(u_o l_o + v_o m_o)} dl_o dm_o = a(u_o, v_o) i(u + u_o, v + v_o).$$

The physical meaning of this equation is the following: suppose we scan the interferometer in pointing position (l_o, m_o) . Then the Fourier transform of the measured visibilities with respect to (l_o, m_o) will yield the Fourier transform of the sky brightness distribution, modulated by the Fourier transform of the primary beam. The spatial frequencies included in the Fourier transform of the primary beam extend from 0 to D/λ , where D is the antenna diameter and λ is the *wavelength* observed. We therefore recover spatial frequencies around a particular (u, v) out to a radius D/λ . To put this more simply, consider a pair of antennas, each with a diameter D , separated by a distance B . In one dimension, the range of spatial frequencies to which the interferometer is sensitive is then $(B - D)/\lambda$ to $(B + D)/\lambda$. An observation at a single pointing yields a visibility measurement which is a linear combination of these spatial frequencies. An observation at a different pointing yield a different linear combination of these spatial frequencies. Given a sufficient number of pointings, the individual spatial frequencies can be recovered via the Fourier transform with respect to pointing. In practice, one only needs to sample in the image domain at regular intervals of $\lambda/2D$ or less.

Cornwell (1989) has shown that nonlinear mosaicing is conveniently implemented in a MEM-like joint deconvolution scheme where an extremum of the modified functional

$$J = H - \alpha \sum_{i=1}^M \chi_i^2 - \beta F$$

is found, where M is the number of discrete pointings. Mosaicing will play an important role in imagin with the NRAO's Millimeter Array (MMA), which will have a rather small field of view. While only small antennas are likely needed for the SRT (2-3 m), the highest frequencies will still suffer from a limited field of view and mosaicing techniques may become necessary.

6.2 Frequency Synthesis

The spatial frequency measured by an antenna baseline \mathbf{b} is b/λ . While antenna locations are generally fixed, the spatial frequency sampled by a particular baseline can nevertheless be changed by varying the wavelength observed. This is the essence of frequency synthesis, wherein several discrete frequencies are observed to improve the uv coverage of the array. In the face of problems like spatial and temporal confusion, described in §6.5 and §6.6, frequency synthesis is an attractive way to greatly improve the uv coverage on a short timescale.

The idea dates back to the late-1950s but has not been applied extensively in practice because of problems discussed below. It has been employed by the OVRO solar array for the purposes of source identification, and by the MERLIN array in the UK. The technique has been described in detail by Conway, Cornwell, and Wilkenson (1990). These authors were motivated by the fact that the uv coverage of the MERLIN synthesis array is quite sparse. Frequency synthesis over a bandwidth $\Delta\nu/\nu \approx \pm 20\%$ yielded an enormous improvement in the uv coverage.

The cost of frequency synthesis lies in the complexity introduced by the fact that the source structure varies with frequency. Unless one can successfully deconvolve both the instrumental point spread function and the source spectrum from the dirty map, the result of frequency synthesis mapping is once again limited in dynamic range, this time by *spectral confusion*. However, since the radio brightness distribution is highly correlated in adjacent frequencies, joint estimation of the brightness distribution and its spectrum is possible. Conway et al. have analyzed the case for which

$$I(l, m; \nu_i) = I(l, m; \nu_o) \left(\frac{\nu_i}{\nu_o} \right)^{-\alpha(l, m)}$$

Designating the Fourier transform operator as \mathbf{F}' , the visibilities sampled at frequency ν_i can be written

$$V(u, v; \nu_i) = \mathbf{F}' \left[I(l, m; \nu_o) \left(\frac{\nu_i}{\nu_o} \right)^{\alpha(l, m) + \alpha_o} \right] S(u, v; \nu_i)$$

Assuming $\alpha(l, m)$ can be written $\alpha_o + \alpha'(l, m)$, with α_o representing some overall spectral index and α' position-dependent departures from α_o . Dividing by the overall brightness distribution, one obtains

$$V'(u, v; \nu_i) = \mathbf{F}' \left[I(l, m; \nu_o) \left(\frac{\nu_i}{\nu_o} \right)^{\alpha(l, m)} \right] S(u, v; \nu_i)$$

The frequency synthesized visibility data, divided by the overall brightness distribution, are then the sum over all frequencies sampled:

$$V'(u, v) = \sum_{i=1}^{N_F} \mathbf{F}'[I'(l, m; \nu_i)]S(u, v; \nu_i)$$

where $I'(l, m; \nu_i) = I(l, m; \nu_o)(\nu_i/\nu_o)^{-\alpha'(l, m)}$. $I'(l, m; \nu_i)$ can be expanded in a Taylor series which is either linear or logarithmic in ν_i/ν_o . Because the dependence on α is a power-law, a logarithmic expansion is perhaps more natural.

Conway et al. show that, by doing so, one can represent the dirty image as a composite:

$$I^D = (I \star B_o) + (I\alpha' \star B_1) + \frac{1}{2}(I\alpha'^2 \star B_2) + \dots \quad (11)$$

where the dirty beams are given by

$$B'_n = \mathbf{F}' \left[\sum_{i=1}^{N_F} [\ln(\nu_o/\nu_i)]^n S(u, v; \nu_i) \right]$$

In cases considered by Conway et al. , $\alpha' \lesssim 0.1 - 0.2$. One can imagine case of interest on the Sun where this is also true over bandwidths of $\pm 25\%$ or so. Because of the presence of a weighting of α'^n , the higher order terms of the expansion become negligible and in practice the deconvolution might be confined to only the first, or the first and second terms of eqn. (11).

The above analysis is meant to be illustrative. Practical application of the technique to the Sun is in its fledgling stage. A great deal of work on the spectral forms encountered in practice over the bandwidths of interest remains to be done, as well as further exploration of techniques for performing the double (spectral and spatial) deconvolution of frequency-synthesized maps.

REFERENCES

- Allisandrakis, C. E., Lantos, P., and Nicolaidis, E. 1985, *Solar Phys.*, 97, 267.
- Bregman, J. D. & Felli, M. 1976, *A&A*, 46, 41.
- Christiansen, W. N. & Warburton, J. A. 1955, *Aust. J. Phys.*, 8, 474.
- Conway, J. E., Cornwell, T. J., & Wilkenson, P. N. 1990, *MNRAS*, 246, 490.
- Ekers, R. D., & Rots, A. H. 1979, in *Proc. IAU Coll. 49, Image Formation from Coherence Functions in Astronomy*, C. van Schooneveld (ed.), D. Reidel:Dordrecht, p. 61.
- Hurford, G. J., Read, R. B., and Zirin, H. 1984, *Solar Phys.*, 94, 413.
- McCready, L. L., Pawsey, J.L., & Payne-Scott, R. 1947, *Proc. R. Soc. London, Ser. A.*, 190, 357.
- Nakajima, H., et al. 1994, *Proc. IEEE*, 82, 705.
- Perley, R. A. and Erickson, W. C. 1984, VLA Scientific Memorandum No. 146.
- Perley, R. A., Schwab, F. R., & Bridle, A. H. 1989, *Synthesis Imaging in Radio Astronomy*, ASP Conf. Series v. 6
- Thompson, A. R., Moran, J. M., and Swenson, G. W., Jr. 1986, *Interferometry and Synthesis in Radio Astronomy*, John Wiley & Sons, New York

Fig. 1

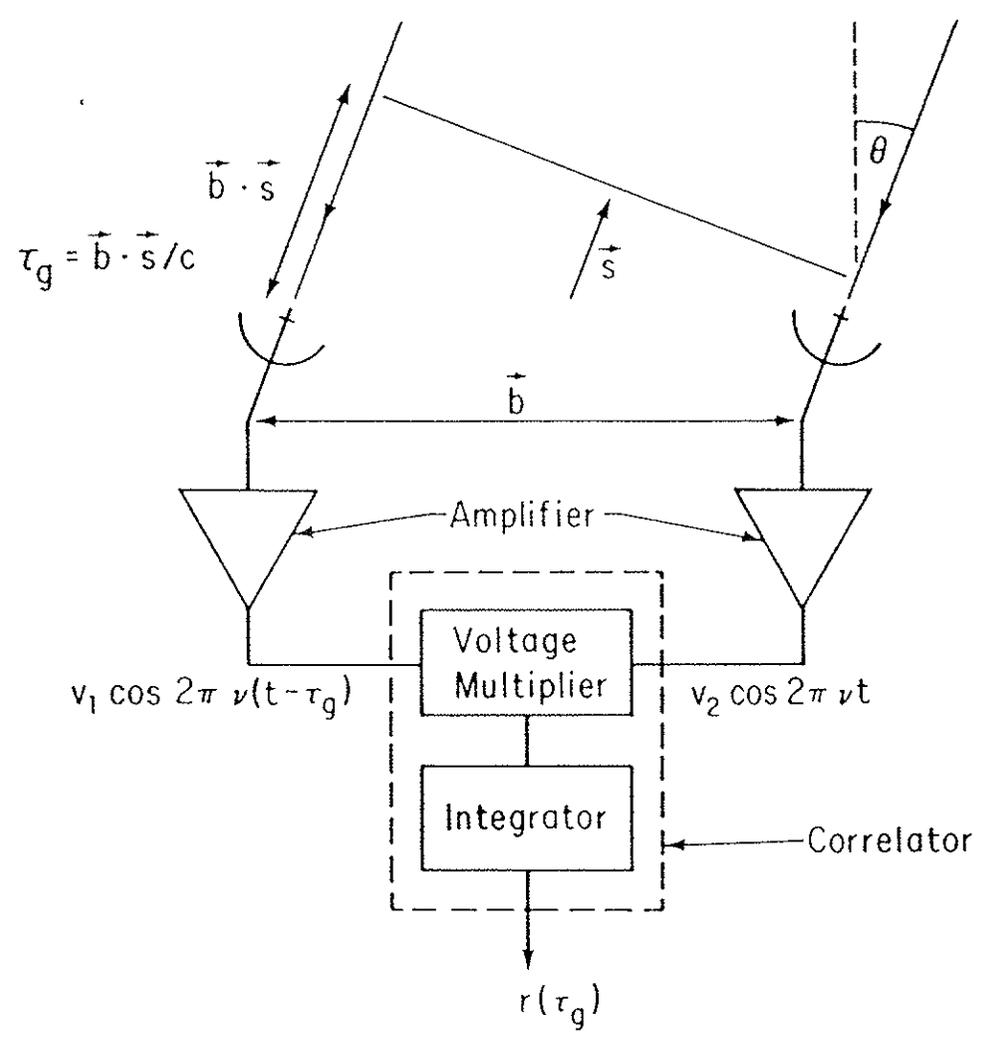
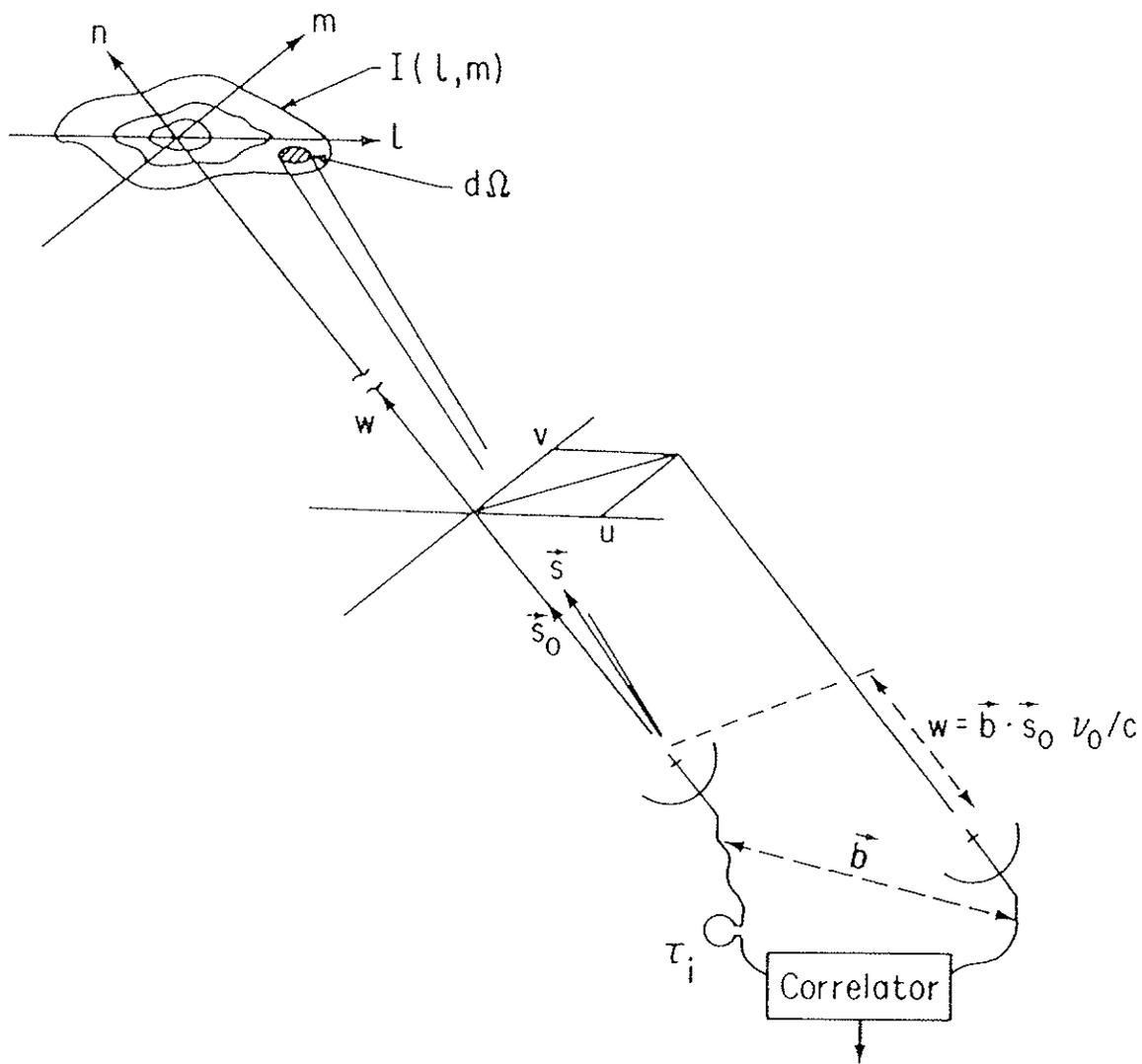
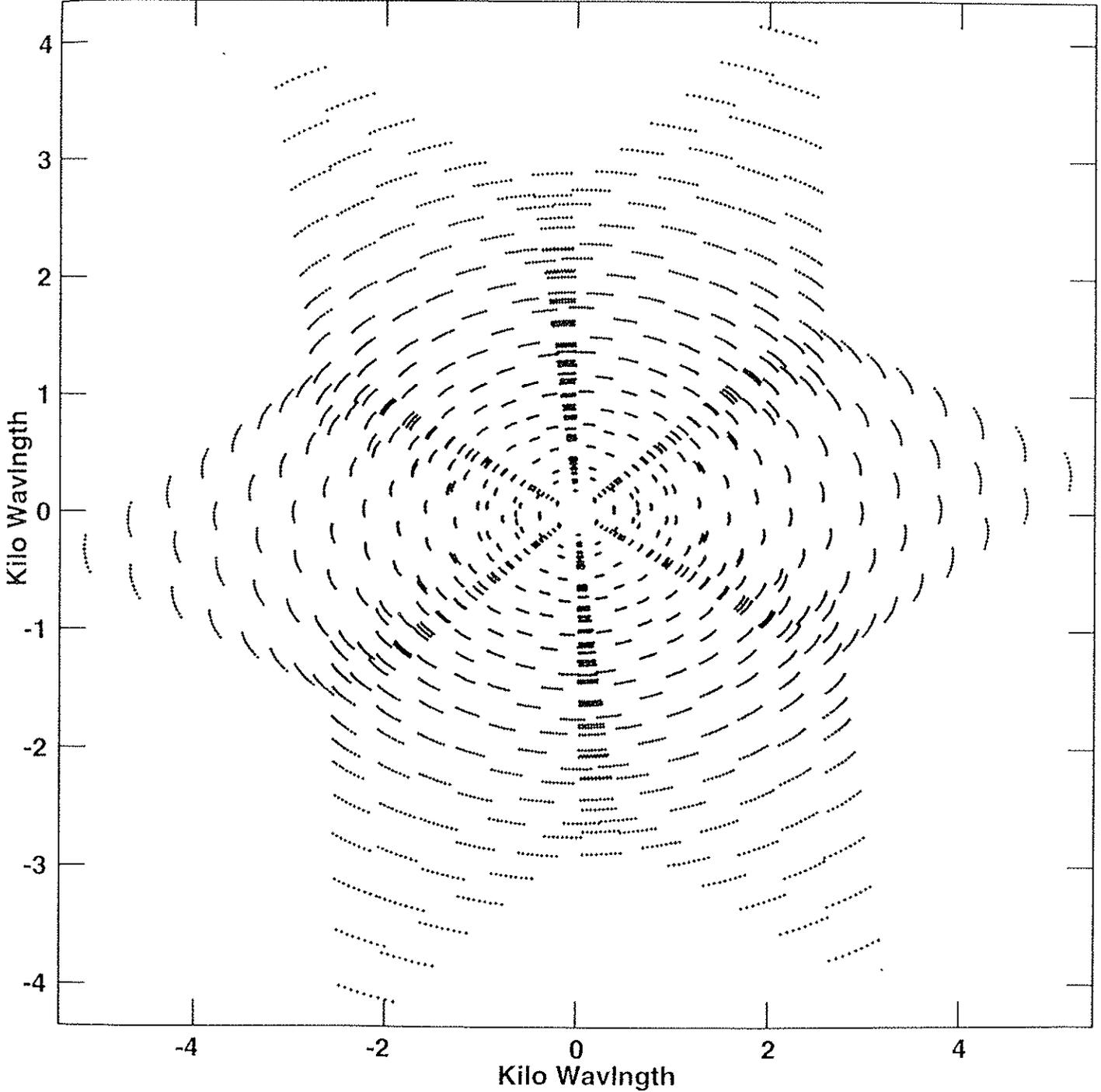


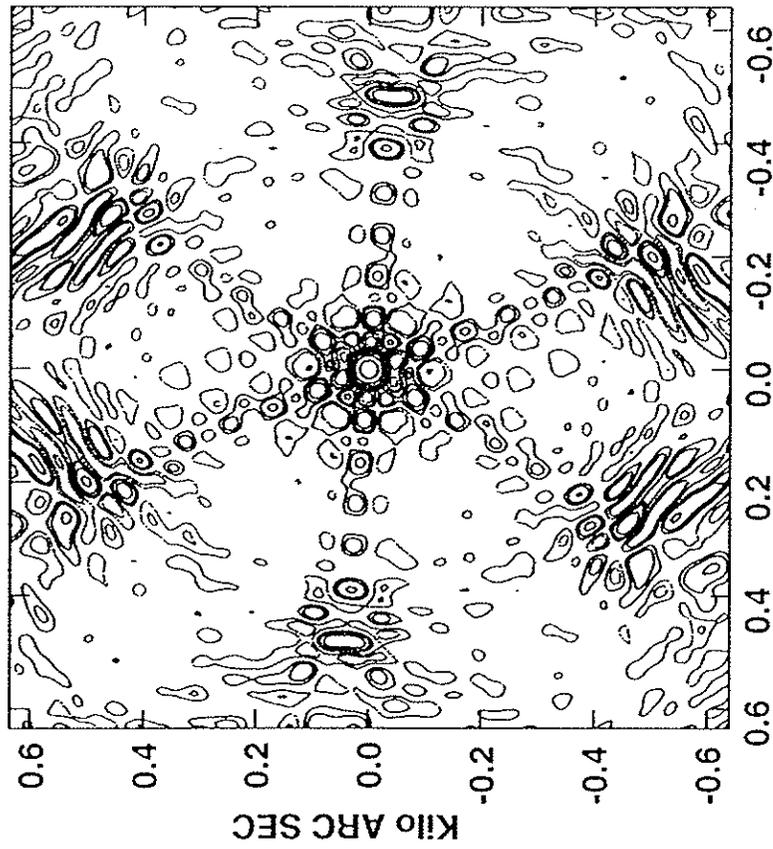
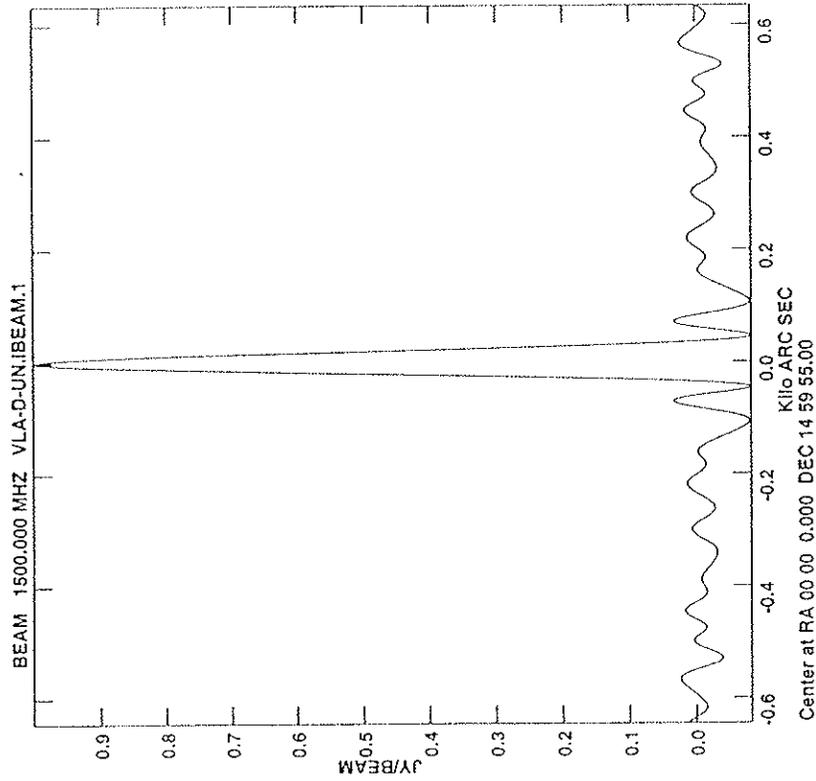
Fig. 2



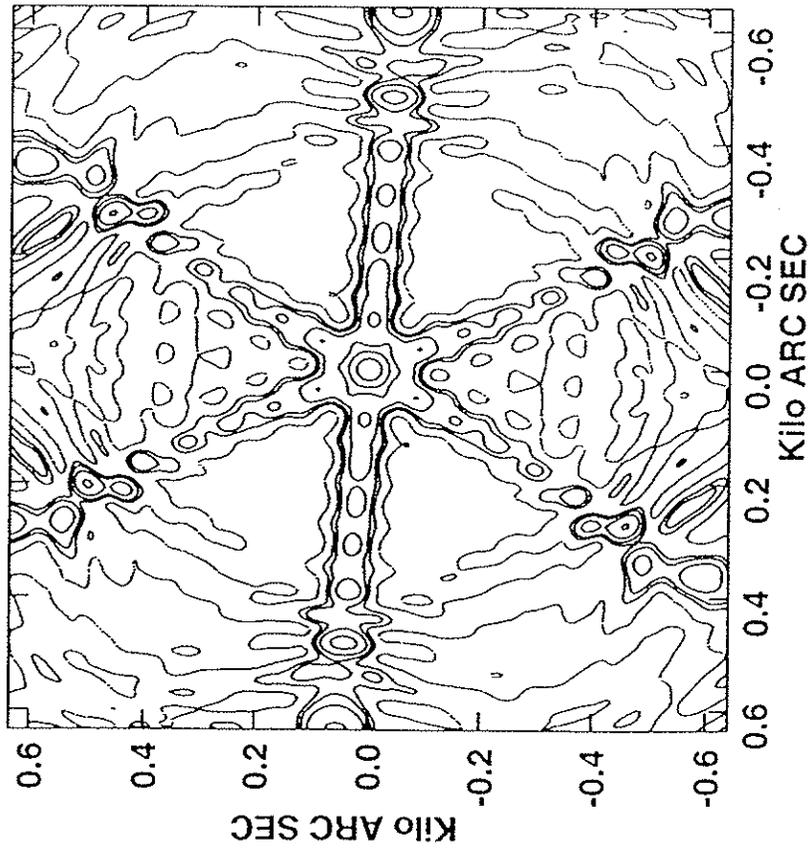
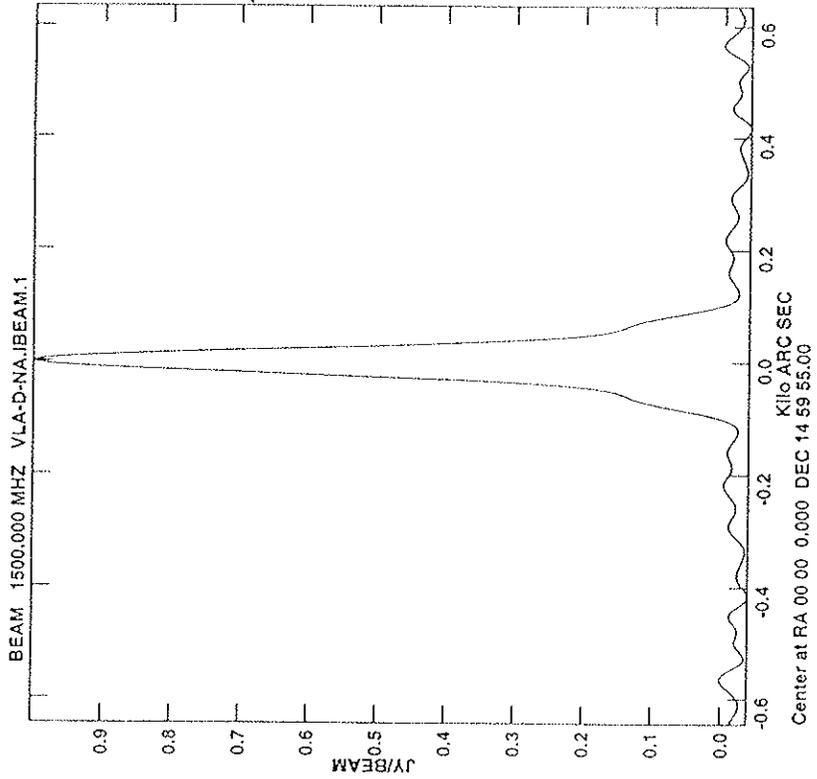
Plot file version 2 created 12-APR-1995 16:23:45
V vs U for VLA-D-SIG.ARMOD.1 Source:
Ants * -* Stokes I IF# 1 Chn# 1



"Sampling function"

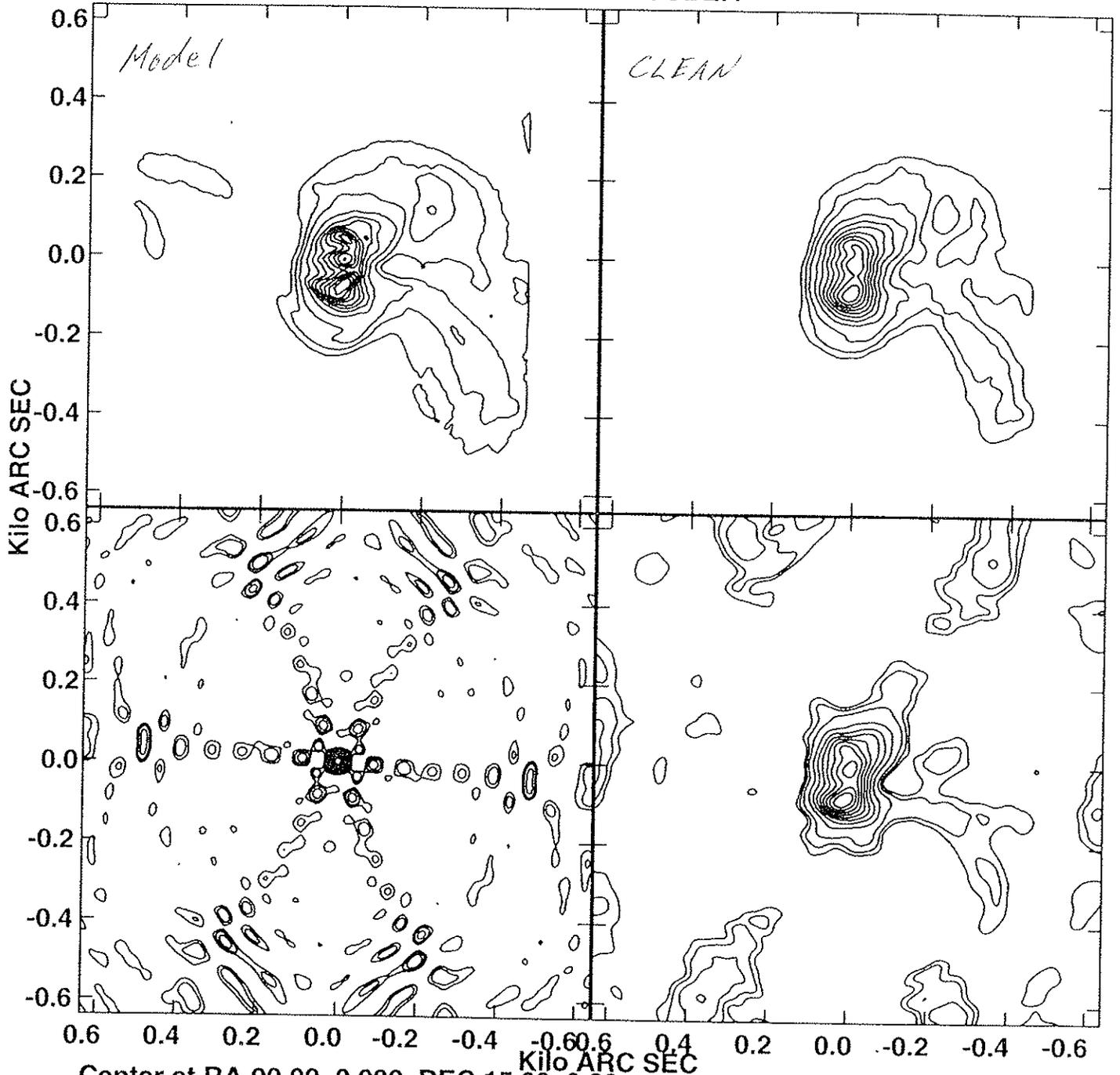


Uniformly-weighted beam



Naturally-weighted beam

SEQ.NUM. 1.000 BEAM VLA-D-UN.CUBE.4



Center at RA 00 00 0.000 DEC 15 00 0.00
Peak flux = 1.0000E+00 JY/PIX
Levs = 1.0000E-02 * (2.000, 5.000, 10.00,
20.00, 30.00, 40.00, 50.00, 60.00, 70.00,
80.00, 90.00)

