Chapter 14

Wide Field Imaging

Sanjay Bhatnagar

14.1 Introduction

It has been shown in Chapter 2 that the visibility measured by the interferometer, ignoring the phase rotation, is given by

$$V(u,v,w) = \int\int I(l,m)B(l,m)e^{-2\pi\iota(ul+vm+w(\sqrt{1-l^2-m^2}))}\frac{dldm}{\sqrt{1-l^2-m^2}}, \qquad (14.1.1)$$

where (u,v,w) defines the co-ordinate system of antenna spacings, (l,m,n) defines the direction cosines in the (u,v,w) co-ordinates system, I is the source brightness distribution (the image) and B is the far field antenna reception pattern. For further analysis we will assume $B=1$, and drop it from all equations (for typing convenience[1]!)

Eq. 14.1.1 is not a Fourier transform relation. For a small field of view ($l^2+m^2 << 1$) the above equation however can be approximated well by a 2D Fourier transform relation. The other case in which this is an exact 2D relation is when the antennas are arranged in a perfect East-West line. However often array configurations are designed to maximize the uv-coverage and the antennas are arranged in a 'Y' shaped configuration. Hence, Eq. 14.1.1 needs to be used to map full primary beam of the antennas, particularly at low frequencies. Eq. 14.1.1 reduces to a 2D relation also for non-EW arrays if the time of observations is sufficiently small (snapshot observations).

In the first part of this chapter we will discuss the implications of approximating Eq. 14.1.1 by a 2D Fourier transform relation and techniques to recover the 2D sky brightness distribution.

The field of view of a telescope is limited by the primary beams of the antennas. To map a region of sky where the emission is at a scale larger than the angular width of the primary beams, mosaicing needs to be done. This is discussed in the second part of this lecture.

[1]The same assumptuin has been made in Chapter 2

14.2 Mapping with Non Co-planar Arrays

14.2.1 Image Volume

Let $n = \sqrt{1 - l^2 - m^2}$ be treated as an independent variable. Then one can write a 3D Fourier transform of $V(u, v, w)$ with the conjugate variable for (u, v, w) being (l, m, n), as

$$F(l, m, n) = \int\int\int V(u, v, w) e^{2\pi\iota(ul+vm+wn)} du dv dw. \tag{14.2.2}$$

Substituting for $V(u, v, w)$ from Eq. 14.1.1 we get

$$F(l, m, n) =$$

$$\int\int\left\{\int\int\int \frac{I(l', m')}{\sqrt{1 - l^2 - m^2}} e^{-2\pi\iota(u(l'-l)+v(m'-m))} e^{-2\pi\iota(w(\sqrt{1-l'^2-m'^2}-n))} du dv dw\right\} dl' dm'. \tag{14.2.3}$$

Using the general result

$$\delta(l' - l) = \int e^{-2\pi\iota u(l'-l)} du, \tag{14.2.4}$$

we get

$$F(l, m, n) = \int\int \frac{I(l', m')}{\sqrt{1 - l^2 - m^2}} \delta(l' - l)\delta(m' - m)\delta(\sqrt{1 - l'^2 - m'^2} - n) dl' dm'. \tag{14.2.5}$$

This equation then provides the connection between the 2D sky brightness distribution given by $I(l, m)$ and the result of 3D Fourier inversion of $V(u, v, w)$ given by $F(l, m, n)$ referred to as the *Image volume*.

$$F(l, m, n) = \frac{I(l, m)\delta(\sqrt{1 - l^2 - m^2} - n)}{\sqrt{1 - l^2 - m^2}}. \tag{14.2.6}$$

Hereafter, I would use $I(l, m, n)$ to refer to the this *Image volume*.

In Eq.14.1.1, we have ignored the fringe rotation term $2\pi\iota w$ in the exponent. This is done here only for mathematical (and typing!) convenience. The effect of including this term would be a shift of the *Image volume* by one unit in the conjugate axis, namely n. Hence, the effect of fringe stopping is to make the top most plane of $I(l, m, n)$ tangent to the phase center position on the celestial sphere with the rest of the sphere completely contained inside the *Image volume* as shown in Fig. 14.1.

Remember that the third variable n of the *Image volume* is not an independent variable and is constrained to be $n = \sqrt{1 - l^2 - m^2}$. Eq 14.2.6 then gives the physical interpretation of $I(l, m, n)$. Imagine the celestial sphere defined by (l, m, n) enclosed by the *Image volume* $I(l, m, n)$, with the top most plane being tangent to the celestial sphere as shown in Fig. 14.1. Eq. 14.2.6 then says that only those parts of the *Image volume* correspond to the physical emission which lie on the surface of the celestial sphere. Note that since the visibility is written as a function of all the three variables (u, v, w), the transfer function will also be a volume. A little thought will then reveal that $I(l, m, n)$ will be finite away from the surface of the celestial sphere also, but that would correspond to non-physical emission in the *Image volume* due to the side lobes of the telescope transfer function (referred to by *Point spread function (PSF)* or *Dirty beam* in the literature). A 3D deconvolution using the *Dirty image-* and the *Dirty beam-volumes* will produce a *Clean image-volume*. Therefore, after deconvolution, one must perform an extra operation of projecting all points in the *image volume* along the celestial sphere onto the 2D tangent plane to recover the 2D sky brightness distribution. Fig. 14.2 is the graphical equivalent of the statements in this paragraph.

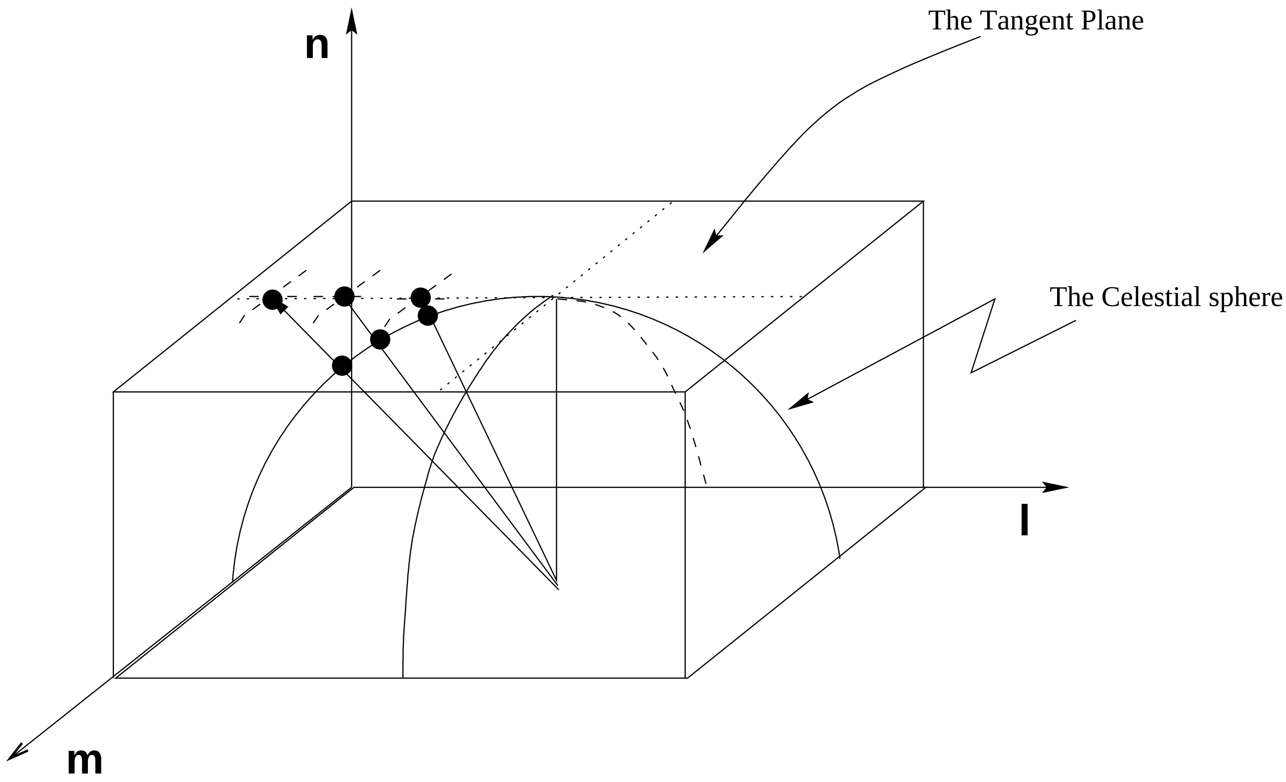

Figure 14.1: Graphical representation of the geometry of the *Image volume* and the celestial sphere. The point at which the celestial sphere touches the first plane of the *Image volume* is the point around which the 2D image inversion approximation is valid. For wider fields, emission at points along the intersection of celestial sphere and the various planes (labeled here as the celestial sphere) needs to be projected to the tangent plane to recover the undistorted 2D image. This is shown for 3 points on the celestial sphere, projected on the tangent plane, along the radial directions.

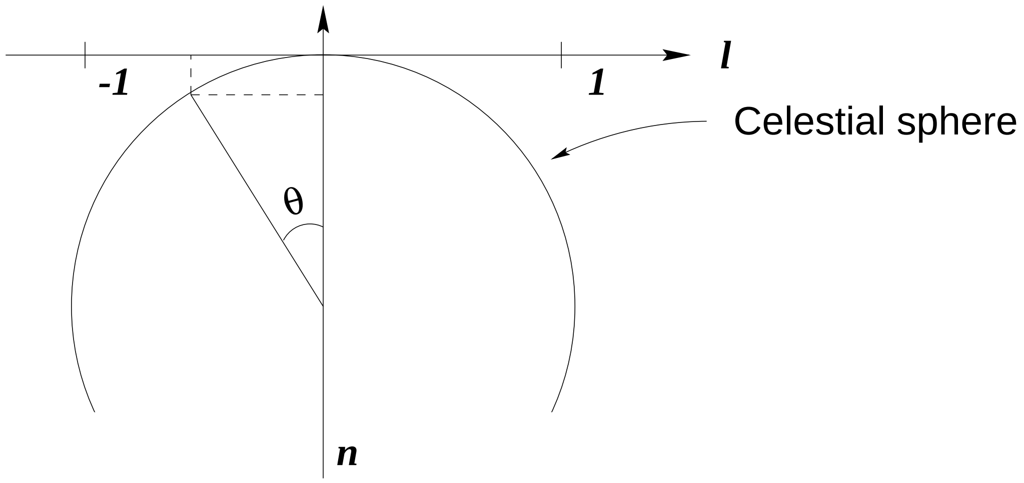

Figure 14.2: Graphical illustration to compute the distance between the tangent plane and a point in the sky at an angle of θ.

14.2.2 Interpretation of the w-term

The term $w\sqrt{1-l^2-m^2}$ is often referred to as the w-term in the literature. The origin of this term is purely geometrical and arises due to the fact that fringe rotation effectively phases the array for a point in the sky referred to as the phase center direction. A wave front originating for this direction will then be received by all antennas and the signals will be multiplied in-phase at the correlator (effectively phasing the array). The locus of all points in 3D space, for which the array will remain phased is a sphere, referred to as the celestial sphere. A wave front from a point away from the phase tracking center but on the surface of such a sphere, will carry an extra phase, not due to the geometry of the array but because of its separation from the phase center. In that sense, the phase of the wavefront measured by a properly phased array in fact carries the information about the source structure and the w-term is the extra phase due to the spherical geometry of the problem. The sky can be approximated by a 2D plane *close* to the phase tracking center and the w-term can be ignored, which is another way of saying that a 2D approximation can be made for a small field of view. However sufficiently far away from the phase center, the phase due to the curvature of the celestial sphere, the w-term, must be take into account, and to continue to approximate the sky as a 2D plane, we will have to rotate the visibility by the w-term. This will be equivalent to shifting the phase centre and corresponds to a shift of the equivalent point in the image plane. Since the w-term is a function of the image co-ordinates, this shift is different for different parts of the image. Shifting the phase centre to any *one* of the points in the sky, will allow a 2D approximation only *around* that direction and *not* for the entire image. Hence the errors arising due to ignoring the w-term cannot be removed by a constant phase rotation of all the visibilities. This is another way of understanding that, in the strict sense, the sky brightness is *not* a Fourier transform of the visibilities.

14.2.3 Inversion Of Visibilities

3D Imaging

The most straight forward method suggested by Eq. 14.2.5 for recovering the sky brightness distribution, is to perform a 3D Fourier transform of $V(u, v, w)$. This requires that the w axis be also sampled at least at Nyquist rate. For most observations it turns out that this is rarely satisfied and doing a FFT on the third axis would result into severe aliasing. Therefore in practice, the transform on third axis is usually done using the direct Fourier transform (DFT), on the un-gridded data.

For performing the 3D FT (FFT on the u and v axis and DT on the w axis) one would still need to know the number of planes needed along the n axis. This can be found using the geometry as shown in Fig. 14.2. The size of the synthesized beam in the n direction is comparable to that in the other two directions and is given by $\approx \lambda/B_{max}$ where B_{max} is the longest projected baseline length. Therefore the separation between the planes along n should be $\leq \lambda/2B_{max}$. The distance between the tangent plane and points separated by θ from the phase center is given by $1 - cos(\theta) \approx \theta^2/2$. For critical sampling then would be

$$N_n = B_{max}\theta^2/\lambda. \tag{14.2.7}$$

At 327 MHz for GMRT, $B_{max} \approx 25$ km. Therefore, for mapping 1° field of view without distortions, one would required 8 planes along the n axis. With central square alone however, one plane should be sufficient. At these frequencies it becomes important to map most of the primary beam since the number and the intensity of the background sources increase and the side lobes of these background sources limit the dynamic range in the maps. Hence, even if the source of interest is small, to get the achievable dynamic range (or close to it!), one will need to do a 3D inversion (and deconvolution).

Another reason why more than one plane would be required for very high dynamic range imaging is as follows. Strictly speaking, the only point which completely lies in the tangent plane is the point at which the tangent plane touches the celestial sphere. All other points in the image, even close to the phase center, lie slightly below the tangent plane. Deconvolution of the tangent plane then results into distortions for the same reason as the distortions arriving from the deconvolution of a point source which lies between two pixels in the 2D case. As in the 2D case, this problem can be minimized by over sampling the image and that, in this case, implies having at least 2 planes in the n axis, even if the Eq. 14.2.7 tells that 1 plane is sufficient.

Polyhedron Imaging

As mentioned above, emission from the phase center and from points close to it lie approximately in the tangent plane. Polyhedron imaging relies on exploiting this fact by approximating the celestial sphere by a number of tangent planes as shown in Fig. 14.3. The visibility data is phase rotated to shift the phase center to the tangent points of the various planes and a small region around the tangent point is then mapped using the 2D approximation. In this case however, one needs to perform a joint deconvolution involving all tangent planes since the sides lobes of a source in one plane would leak into other planes as well.

The number of planes required to map an object of size θ can be found simply by requiring that maximum separation between the tangent plane and the region around each tangent point be less than λ/B_{max}, the size of the synthesized beam. As shown earlier, the separation of a point θ degrees away from the tangent point is $\approx \theta^2$. Hence for critical sampling, the number of planes required is equal to the solid angle subtended by

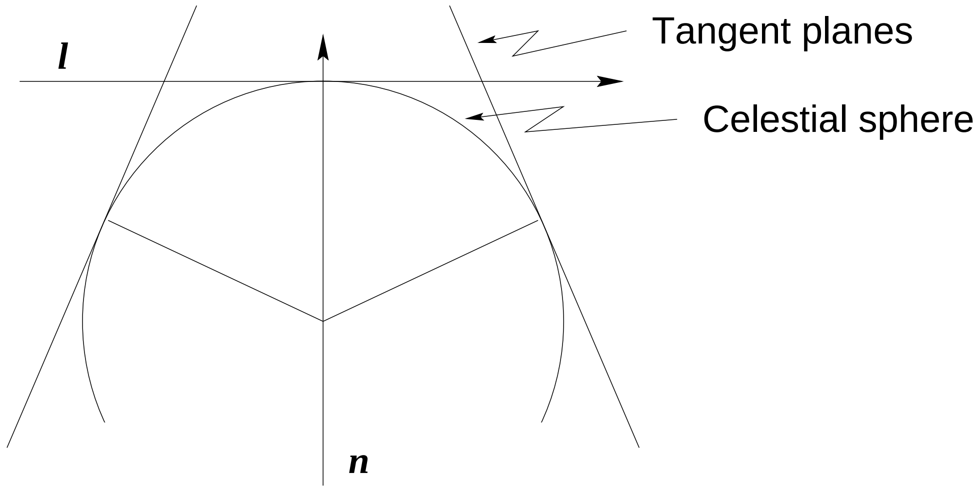

Figure 14.3: Approximation of the celestial sphere by multiple tangent planes (polyhedron imaging).

the sky being mapped (θ_f^2) divided by the solid angle of the synthesized beam (θ^2)

$$N_{poly} \;=\; 2\theta_f^2 B_{max}/\lambda \;=\; 2B_{max}\lambda/D^2 \;\; (\text{for } \theta_f \;=\; \text{full primary beam}). \tag{14.2.8}$$

Notice that the number of planes required is twice as many as the number of planes required for 3D inversion. However since a small portion around the tangent point of each plane is used, the size of each of these planes can be small, offsetting the increase in computations due to the increase in the number of planes required. Another approach which is often taken for very high dynamic range imaging is to do a full 3D imaging on each of the planes. This would effectively increase the size of the field that can be imaged on each tangent plane, thereby reducing the number of planes required.

The polyhedron imaging scheme is available in the current version of `AIPS` data reduction package and the 3D inversion (and deconvolution) is implemented in the (not any more supported) `SDE` package written by Tim Cornwell et al. Both these schemes, in their full glory, will be available in the (recently released) `AIPS++` package.

14.3 Mosaicing

The problem due to non co-planarity discussed above are for mapping the sky within the primary beam of the antennas (which are assumed to be identical). In this section we discuss the techniques used to handle the problem of mapping fields of interest which are larger than the primary beam of the antennas. The approach used is similar to that used for mapping with a single dish, namely to scan the source to be mapped. The fact that we are using an interferometer to synthesis the “lens” (or the a “single dish”) adds some more complications.

These techniques are useful for mapping with interferometers operating in the millimeter range where the size of the primary beams is less than an arcmin and at meter wavelengths where the primary beams are larger but so is the extent of emission. For example, the primary beam of GMRT antennas at 327 MHz is $\approx 1.3°$ and there are mapping projects which would benefit from mapping regions of the sky larger than this (for example, in the Galactic plane).

14.3.1 Scanning Interferometer

The co-planar approximation of Eq. 14.1.1 for a pointing direction given by (l_o, m_o) can be written as

$$V(u, v, l_o, m_o) = \int\int I(l, m)B(l - l_o, m - m_o)e^{2\pi\iota(ul+vm)}dldm. \tag{14.3.9}$$

Here we also assume that B is independent of the pointing direction and we label V with not just the (u, v) co-ordinates, but also with pointing direction since visibilities for different directions will be used in the analysis that follows. The advantage of writing the visibility as in Eq. 14.3.9 is that the pointing center (given by (l_o, m_o)) and the phase center (given by $(l, m) = (0, 0)$) are separated.

$V(0, 0, l_o, m_o)$ represents the single dish observation in the direction (l_o, m_o) and is just the convolution of the primary beam with the source brightness distribution, exactly as expected intuitively. Extending the intuition further, as is done in mapping with a single dish, we need to scan the source around (l_o, m_o) with the interferometer, which is equivalent to scanning with a single dish with a primary beam of the size of the synthesized beam of the interferometer. Then Fourier transforming $V(u, v, l_o, m_o)$ with respect to (l_o, m_o), assuming that B is symmetric, one gets, from Eq. 14.3.9

$$\int\int V(u, v, l_o, m_o)e^{2\pi\iota(u_o l_o+v_o m_o)}dl_o dm_o = b(u_o, v_o)i(u + u_o, v + v_o), \tag{14.3.10}$$

where (u_o, v_o) corresponds to the direction (l_o, m_o) and $b \rightleftharpoons B$ and $i \rightleftharpoons I$. This equation essentially tells us the following: Fourier transform of the visibility with respect to the pointing directions, from a scanning interferometer is equal to the visibility of the *entire source* modulated by the Fourier transform of the primary beams for each pointing direction. For a given direction (l_o, m_o) we can recover spatial frequency information spread around a nominal point (u, v) by an amount D/λ where D is the size of the dish. In terms of information, this is exactly same as recovering spatial information smaller than the size of the resolution of a single dish by scanning the source with a single dish. As in the case of a single dish, continuous scanning is not necessary and two points separated by half the primary beam is sufficient. In principle then, by scanning the interferometer, one can improve the short spacings measurements of V, which is crucial for mapping large fields of view.

Image of the sky can now be made using the full visibility data set (made using the Eq. 14.3.10). However, this involves the knowledge of Fourier transform of the sky brightness distribution, which in-turn is approximated after deconvolution. Hence, in practice one uses the MEM based image recovery where one maximizes the entropy given by

$$H = -\sum_k I_k ln\frac{I_k}{M_k}, \tag{14.3.11}$$

with χ^2 evaluated as

$$\chi^2 = \sum_k \frac{|V(u_k, v_k, l_{ok}, m_{ok}) - V^M(u_k, v_k, l_{ok}, m_{ok})|^2}{\sigma^2_{V(u_k, v_k, l_{ok}, m_{ok})}}, \tag{14.3.12}$$

where $V^M(u_k, v_k, l_{ok}, m_{ok})$ is the model visibility evaluated using Eq. 14.3.9. For calculation of $\Delta\chi^2$ in each iteration is estimated by the following steps:

- initialize $\Delta\chi^2 = 0$
- For all pointings
  1. Apply the appropriate primary beam correction to the current estimate of the image
  2. FT to generated V^M
  3. Accumulate χ^2
  4. Subtract from the observed visibilities
  5. Make the residual image
  6. Apply the primary beam correction to the residual image
  7. Accumulate $\Delta\chi^2$

The operation of primary beam correction on the residual image is understood by the following argument: For any given pointing, an interferometer gathers radiation within the primary beam. In the image plane then, any feature, outside the range of the primary beam would be due to the side lobes of the synthesized beam and must be suppressed before computation of $\Delta\chi^2$ and this is achieved by primary beam correction, which essentially divides the image by gaussian which represents the main lobe of the antenna radiation pattern.

This approach (rather than joint deconvolution) has several advantages.

1. Data from potentially different interferometers for different pointings can be used
2. Weights on each visibility from each pointing are used in the entire image reconstruction procedure
3. Single-dish imaging emerges as a special case
4. It is fast for extended images

The most important advantage that one gets by MEM reconstruction is that the deconvolution is done simultaneously on all points. That this is an advantage over joint-deconvolution can be seen as follows: If a point source at the edge of the primary beam is sampled by 4 different pointings of the telescope, this procedure would be able to use 4 times the data on the same source as against data from only one pointing in joint-deconvolution (where deconvolution is done separately on each pointing). This, apart from improvement in the signal-to-noise ratio also benefits from a better uv-coverage available.

Flexible software for performing Mosaic-ed observations is one of the primary motivation driving the `AIPS++` project in which algorithms to handle mosaic-ed observations would be available in full glory.

14.4 Further Reading

1. Interferometry and Synthesis in Radio Astronomy; Thompson, A. Richard, Moran, James M., Swenson Jr., George W.; Wiley-Interscience Publication, 1986.

2. Synthesis Imaging In Radio Astronomy; Eds. Perley, Richard A., Schwab, Frederic R., and Bridle, Alan H.; ASP Conference Series, Vol 6.