Synthesis Imaging in Radio Astronomy Reading Note

Renkun Kuang

September 22, 2019

Contents

Preface

The Summer School on Synthesis Imaging held in Socorro, New Mexico from June 17 to June 23, 1998 was the sixth in a series held approximately every three years. This Volume contains the edited texts of lectures from the series, and succeeds the previous collection from the Third Synthesis Imaging Summer School published in 1989. It is intended for **serious students of synthesis imaging and image processing.**

Purpose of the course

The NRAO operates two of the world's most powerful synthesis radio telescopes, the Very Large Array (VLA) and the Very Long Baseline Array (VLBA), a synthesis tele- scope.
(Now in 2019, there are three telescopes, including ALMA, https://public.nrao.edu/)
The major goal of this course, like that of its predecessors beginning in 1982 was to **inform potential users of these two synthesis telescopes about the principles of these instruments' operation, about subtleties of data acquisition, calibration and processing with such instruments, and about techniques for obtaining the best results from them.**

As such, the course is aimed first at radio astronomers who need synthesis techniques and instruments for their research. Get practical experience of synthesis imaging, Fourier methods and coherence. Many of the exciting developments in data processing that have made synthesis telescopes so powerful. We therefore have set out to **discuss the subject as fully as necessary for those who wish to use the NRAO's radio synthesis instruments for their own research**. Our goal is to **discuss the subject in enough detail that the student can appreciate both the strengths and limitations of the synthesis technique**, and so begin to evaluate how much, or how little, credence to give individual synthesis images. **To exploit an instrument fully for frontier research, the user must understand it thoroughly enough to distinguish unexpected instrumental errors from unexpected discoveries about the cosmos. A firm understanding of synthesis instruments involves understanding their operating principles (and the assumptions that underlie them), their hardware and the algorithms and software used in the data reduction.** All these topics are covered here.

Although synthesis imaging is a specialized skill even within radio astronomy, it is grounded in mathematical and physical principles that have applications in other fields. Interpret images made by synthesis telescopes. It may also inter-

est researchers who **use Fourier methods or deconvolution techniques for imaging in physics, medicine, remote sensing, seismology or radar.**

Subject matter

The first segment of the course contains sixteen lectures that describe the fundamentals of synthesis imaging and which could be read as a stand-alone course by the beginning student. The second segment consists of seventeen lectures on more advanced and specialized topics.

NRAO lore

AIPS
https://public.nrao.edu/
The on-line bibliographic Abstract Service of the Astrophysics Data System at Harvard University.

This book is dedicated to the memory of Daniel S. Briggs. Dan was killed in a skydiving accident near Chicago just a few weeks after the summer school. Dan was an enthusiastic teacher and researcher who will be sorely missed. His 'Robust' weighting scheme has become the standard for synthesis imaging. At the summer school he gave an excellent lecture on imaging and deconvolution (see Lectures 7 and 8 of this volume contributed by Dan and collaborators). He entered thoroughly into the spirit of the summer school — attending all the lectures and interacting with many of the participants. Dan also influenced the style of this book by making many helpful suggestions. We are in his debt.

Chapter 1

Coherence in Radio Astronomy

In this lecture the **main principles of synthesis imaging are derived.**

1.1 Introduction

A survey of the derivation of the main principles of synthesis imaging, and the assumptions that go into them. This is because a substantial number of the lectures to follow will discuss the problems which arise when these assumptions are violated under the conditions of the observation the astronomer wants to make. At the same time, I will cast this introduction into the terminology of modern optics, in an attempt to stay abreast of current fashions in physics.

Some reference books, e.g. The alternate viewpoint on radio interferometers, from the perspective of the electrical engineers who originally developed them, is explicated in Swenson & Mathur (1968).
The interferometer in radio astronomy, IEEE, 1968

other:
SparseRI: A Compressed Sensing Framework for Aperture Synthesis Imaging in Radio Astronomy
The Sensitivity of Synthetic Aperture Radiometers for Remote Sensing Applications From Space
Basics of Interferometer, PPT
Introduction to Radio Interferometry-NRAO
Introduction to Radio Interferometers - ASTRON

1.2 Form of the Observed Electric Field

Assumptions:

1. scalar
2. far away
3. space empty
4. radiation not spatially coherent

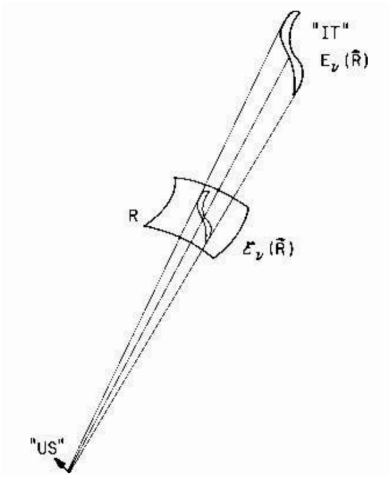


Figure 1.1: The passage of radiation from a distant source through an imaginary sphere at radius $|\mathbf{R}|$ defines the electric field distribution $\epsilon_\nu(\mathbf{R})$ on this surface. For the purposes of synthesis imaging, all astronomical sources may be considered to lie on this sphere, provided only that their real distances greatly exceed B^2/λ, where $B = |\mathbf{r_1} - \mathbf{r_2}|$ is the baseline length.

$$E_\nu(r) = \int \epsilon_\nu(\mathbf{R}) \frac{e^{2\pi i\nu|\mathbf{R}-\mathbf{r}|/c}}{|\mathbf{R}-\mathbf{r}|} dS \tag{1.1}$$

Here dS is the element of surface area on the celestial sphere.
Equation 1.1 is the general form of the **quasi-monochromatic component** of the electric field at frequency ν due to all sources of cosmic electromagnetic radiation. **This is all we have, we can measure only the properties of this field $E_\nu(\mathbf{r})$ to tell us about the nature of things at large in the universe.**

1.3 Spatial Coherence Function of the Field

Among the properties of $E_\nu(\mathbf{r})$ is the correlation of the field at two different locations.
The **correlation** of the field at points $\mathbf{r_1}$ and $\mathbf{r_2}$ is defined as the expectation of a product, namely $V_\nu(\mathbf{r_1}, \mathbf{r_2}) = \langle E_\nu(\mathbf{r_1}) E^*_\nu(\mathbf{r_2}) \rangle$, the raised asterisk indicates the complex conjugate. Use equation 1.1 to substitute for $E_\nu(\mathbf{r})$ and use assumption 4 to derive/rearrange the formula and get:

$$V_\nu(\mathbf{r_1}, \mathbf{r_2}) = \int I_\nu(\mathbf{s}) e^{-2\pi i \nu \mathbf{s}\cdot(\mathbf{r_1}-\mathbf{r_2})/c} d\Omega \tag{1.2}$$

$\mathbf{s}$ is the unit vector $\mathbf{R}/|\mathbf{R}|$, $I_\nu(\mathbf{s})$ is the observed ***intensity*** $|\mathbf{R}|^2 \langle |\epsilon_\nu(\mathbf{s})|^2 \rangle$

Observe that Equation 1.2 depends only on the separation vector $\mathbf{r_1} - \mathbf{r_2}$ of the two points, not on their absolute locations $\mathbf{r_1}$ and $\mathbf{r_2}$(两个观测台站). **Therefore, we can find out all we can learn about the correlation properties of the radiation field by holding one observation point fixed and moving the second around;** we do not have to measure at all possible pairs of points. This function V, of a single (vector) separation $\mathbf{r_1} - \mathbf{r_2}$ is called the **spatial coherence function**, or the **spatial autocorrelation function**, of the field $E_\nu(\mathbf{r})$. It is all we have to measure.

An interferometer is a device for measuring this spatial coherence function.

1.4 The Basic Fourier Inversions of Synthesis Imaging

A second interesting point about Equation 1.2 is that the equation is, within reasonable, well-defined limits, invertible. The intensity distribution of the ra- diation as a function of direction $\mathbf{s}$ can therefore be deduced in certain cases by measuring the spatial coherence function V as a function of $\mathbf{r_1} - \mathbf{r_2}$ and performing the inversion.

There are two special cases of a great deal of interest. In fact, it is usually argued that any actual case is so close to one of these two special cases that the invertibility properties (although not necessarily the effort required to perform the inversion) must be essentially similar. Since there are two forms of interest, there are two alternate forms of our fifth (and final) simplifying assumption:

1.4.1 Measurements confined to a plane

First, we could choose to make our measurements only in a plane; that is, in some favored coordinate system, the vector spacing of the separation variable in the coherence function, conveniently measured in terms of the wavelength $\lambda = c/\nu$, is $\mathbf{r_1} - \mathbf{r_2} = \lambda(u, v, 0)$.
In this same coordinate system, the components of the unit vector $\mathbf{s}$ are $(l, m, \sqrt{1 - l^2 - m^2})$. Inserting these, and explicitly showing the form, in this coordinate system, of the element of solid angle, we have

$$V_\nu(u, v, \omega \equiv 0) = \iint I_\nu(l, m) \frac{e^{-2\pi i(ul+vm)}}{\sqrt{1 - l^2 - m^2}} dldm \tag{1.3}$$

Equation 1.3 is, clearly, a Fourier transform relation between the spatial coherence function $V_\nu(u, v, \omega \equiv 0)$ (with separations expressed in wavelengths), and the modified intensity I, $I_\nu(l, m)/\sqrt{1 - l^2 - m^2}$ (with angles expressed as direction cosines). Now we are home free. **Mathematicians have devoted decades of work to telling us when we can invert a Fourier transform, and how much information it requires.**

1.4.2 All sources in a small region of sky

The alternate form of the fifth simplifying assumption is to consider the case where all of the radiation comes from only a small portion of the celestial sphere. That is, to take $s = s_0 + \sigma$, and neglect all terms in the squares of the components of σ. In particular, the statement that both s and s_0 are unit vectors implies that:

$$1 = |s| = s \cdot s \approx 1 + 2s_0 \cdot \sigma$$

i.e., s_0 and σ are perpendicular. If we again introduce a special coordinate system such that $s_0 = (0, 0, 1)$, then we have a slightly different offspring of equation 1.2:

$$V'_\nu(u, v, \omega) = e^{-2\pi i\omega} \iint I_\nu(l, m) e^{-2\pi i(ul+vm)} dldm \tag{1.4}$$

Here, the components of the vector $r_1 - r_2$ have been denoted by (u, v, ω). It is customary to absorb the factor in front of the integral in Equation 1.4 into the left hand side, by considering the quantity $V_\nu(u, v, \omega) = e^{2\pi i\omega} V'_\nu(u, v, \omega)$, which we see is independent of ω:

$$V_\nu(u, v) = \iint I_\nu(l, m) e^{-2\pi i(ul+vm)} dldm \tag{1.5}$$

$V_\nu(u,v)$ is the coherence function relative to the direction s_0, which is called the ***phase tracking center***.
Since equation 1.5 is a Fourier transform, we have the direct inversion:

$$I_\nu(l,m) = \iint V_\nu(u,v) e^{2\pi i(ul+vm)} dudv \tag{1.6}$$

1.4.3 Effect of discrete sampling

In practice the spatial coherence function V is not known everywhere but is sampled at particular places on the u-v plane. The sampling can be described by a sampling function $S(u,v)$, which is zero where no data have been taken. One can then calculate a function

$$I_\nu^D(l,m) = \iint V_\nu(u,v) S(u,v) e^{2\pi i(ul+vm)} dudv \tag{1.7}$$

Radio astronomers often refer to $I_\nu^D(l,m)$ as the **dirty image**; its relation to the desired intensity distribution $I_\nu(l,m)$ is (using the convolution theorem for Fourier transforms):

$$I_\nu^D = I_\nu * B \tag{1.8}$$

where the in-line asterisk denotes convolution
上面是时域的卷积，下面是频域的乘积：

$$V_\nu^D(u,v) = V_\nu(u,v) S(u,v)$$

and:

$$B(l,m) = \iint S(u,v) e^{2\pi i(ul+vm)} dudv \tag{1.9}$$

is the **synthesized beam** or **point spread function** corresponding to the **sampling function** $S(u,v)$. Equation 1.8 says that I^D is the **true intensity distribution** I convolved with the synthesized beam B. Lecture 8 discusses methods for undoing this convolution.

1.4.4 Effect of the element reception pattern

An additional minor alteration must be made to the above for convenience in practical calculation. In practice, the interferometer elements are not point probes which sense the voltage at that point, but are elements of finite size, which have

some sensitivity to the direction of arrival of the radio radiation. That is, there is an additional factor within the integral of Equation 1.1 (,and hence of other equations) of $A_\nu(s)$ (the **primary beam** or **normalized reception pattern** of the interferometer elements) describing this sensitivity as a function of direction. For explicitness, Equation 1.5 is rewritten below, with this factor included:

$$V_\nu(u,v) = \iint A_\nu(l,m) I_\nu(l,m) e^{-2\pi i(ul+vm)} dl dm \tag{1.10}$$

The $V_\nu(u,v)$ so defined is normally termed the **complex visibility** relative to the chosen phase tracking center.

Although the factor A_ν, looks like merely a nuisance, it is actually the reason that the second form of the final assumption (used in Section 4.2) is so acceptable in many practical cases—$A_\nu(s)$ falls rapidly to zero except in the vicinity of some s_0, the pointing center for the array elements.

1.5 Extensions to the Basic Theory

1.5.1 Spectroscopy

With current technology, it is attractive to implement the latter portions of the interferometer in digital hardware. In this technology, it is quite inexpensive to add additional multipliers to calculate the correlation as a function of lag. Admitting a range of quasi-monochromatic waves to the interferometer, we can write an expression for the correlation as a function of lag, noting that for each quasi-monochromatic wave, a lag is equivalent to a phase shift, i.e., a multiplication by a complex exponential

$$V(u,v,\tau) = \int V(u,v,\nu) e^{2\pi i \tau \nu} d\nu$$

The above is clearly a Fourier transform, with complementary variables ν and τ, and can be inverted to extract the desired $V(u,v,\nu)$ (Spectroscopy). Since, in this digital technology, one is dealing with sampled data, I give the sampled form of the inversion below:

$$V(u,v,j\Delta\nu) = \sum_k V(u,v,k\Delta\tau) e^{-2\pi i j k \Delta\nu \Delta\tau}$$

The fact that we are dealing with sampled data is of some interest, and we should stop and inquire about how the Fourier sampling theorem is to be applied. Examining the above, in its full complex form, we see that the replication interval is

$1/\Delta\tau$ in frequency, so that the band of frequencies must be limited, before sampling, to a total bandwidth of less than this, to avoid loss of information in the sampling process.

This is different from the statement one usually encounters, in which a prefiltering to $1/2\Delta\tau$ is required to preserve the information in the sampling process for a signal (actually it is usually stated, equivalently, as requiring a sampling interval of $1/2B$, where B is the prefilter bandwidth). This factor of two difference is due to the complex nature of the quantities we have been dealing with—the $V(u, v, \nu)$ are complex numbers, calculated by a complex multiplication of the complex field quantities. Complex multipliers and complex samplers require at least twice as many electronic components as devices that produce a real number, and the resulting doubling of the hardware permits us to sample a factor of two less often. (是说采样的数据点是复数的，所以Nyquist采样定理得出的最小采样间隔这时候不需要以前那么小，只需要一般情况下的2倍?)

Finally, if one derives the spectrum in this manner, one can, clearly, convert back to the single continuum channel at zero lag simply by summing the derived frequency-dependent V. This process clearly results in a complex number, even though each measurement was only a real number. The process of transforming a real function into a complex one by Fourier transforming and then transforming back on half the interval is called a **Hilbert transform**, and is an alternate method to implementing complex correlators.

1.5.2 Polarimetry

Actually, the electromagnetic field is a vector phenomenon, and the polarization properties carry interesting physical information. For the case of noise emission, one must be a bit careful about the definition of polarization. A monochromatic wave is always completely polarized, in some particular elliptical polarization, in that a single number describes the variation of the fields everywhere. For electromagnetic noise, polarization is defined by a correlation process. One picks two orthogonal polarizations and analyzes the radiation of the quasi-monochromatic waves into the components in these two polarizations. Then the polarization of the quasi-monochromatic wave is described by the 2×2 matrix of correlations between these two resolutions into orthogonal components. For instance, if we pick right and left circular polarization as the two orthogonal modes, then the matrix

$$\begin{bmatrix} RR^* & RL^* \\ LR^* & LL^* \end{bmatrix}$$

describes the polarization. This can be related to more familiar descriptions of polarization. For instance, the **Stokes parameters** have the intensity I, two linear polarization parameters Q and U, and a circular polarization parameter V related to the above numbers in simple (and more or less obvious) linear combinations:

$$\begin{bmatrix} I+V & Q+iU \\ Q-iU & I-V \end{bmatrix}$$

The complex correlation functions on the celestial sphere are preserved in the spatial coherence functions that interferometers measure. That is, one can derive, for instance, the distribution of $\langle RR^* \rangle$ on the sky by measuring the coherence function of R on the ground, and so forth for the other components of the matrix. Since the intensity is the quantity in which one is always interested, one usually forms the sum of the R and L coherence functions before transforming to the sky plane, which one can always do, since the relations are linear. One can choose to do the same with the other Stokes parameters, or one can calculate the transforms of the mutual coherence between R and L to find the distribution of $\langle RL^* \rangle$ on the sky, and later note that this is, in terms of the Stokes parameters, $Q + iU$.

Chapter 2

Fundamentals of Radio Interferometry

The practical aspects of interferometry are reviewed, starting with a two element interferometer.

2.1 Introduction

In the first lecture, it was shown that images of a distant radio source can be made by measuring the mutual coherence function of the electric fields at pairs of points in a plane normal to the direction to the source. This process can be envisioned by considering a large flat area on the Earth's surface on which antennas are located, and an electronic system including correlators to measure the coherence of the received signals. Now suppose that the Earth does not rotate, and that the source under observation is at the zenith. We can measure the coherence as a function of the spacing between pairs of antennas, independent of the absolute location of the antennas in the antenna plane. Next suppose that the Earth remains fixed relative to the sky, but that the source under investigation is not at the zenith. A wavefront from the source meets the plane in a line that progresses across the plane, and thus does not reach all antennas simultaneously. As a result, we need to include delay elements in the electronic system to ensure that the signals received by different antennas from the same wavefront arrive at the correlators at the same time. **The spacings in (u,v) coordinates are now the components of the baselines projected onto a plane normal to the direction of the source.** Finally, consider the effect of the rotation of the Earth. It is necessary for the antenna pointing and the time delays to be continuously adjusted to follow the source across the sky. Further, although at any instant the baselines lie within a plane, this plane is carried through space by the Earth, and its position relative to the source

is continuously changing. The non-coplanar distribution ot the baselines which thus generally occurs during an extended period of observation can complicate the inversion of the coherence data to obtain an image. On the other hand, the motion of the baseline vectors has the useful effect of reducing the number of antenna locations on the Earth required to measure the coherence over the necessary range of (u,v) coordinates.

2.2 Response of an Interferometer

When taking observations to make an interferometric image of a radio source, it is usual to specify a position on which the synthesized field of view is to be centered. This position is commonly referred to as the **phase tracking center or phase reference position**.

We now introduce the **visibility**, which is a measure of the coherence discussed in Lecture 1. The term visibility was first used in interferometry by Michelson (1890) to express the relative amplitude of the optical fringes that he observed. As used in radio astronomy, visibility is a complex quantity, the magnitude of which has the dimensions of spectral power flux density ($Wm^{-2}Hz^{-1}$). It can be regarded as an unnormalized measure of the coherence of the electric field, modified to some extent by the characteristics of the interferometer. The complex visibility of the source is defined as

$$V \equiv |V|e^{i\phi V} = \int_S \mathcal{A}(\sigma) I(\sigma) e^{-2\pi i \nu b \cdot \sigma / c} d\Omega \tag{2.1}$$

where $\mathcal{A}(\sigma) \equiv A(\sigma)/A_0$ is the normalized antenna reception pattern, A_0 being the response at the beam center.

$$r = A_0 \Delta\nu |V| \cos(\frac{2\pi\nu b \cdot s_0}{c} - \phi_V) \tag{2.2}$$

In the interpretation of interferometer measurements the usual procedure is to measure the amplitude and phase of the fringe pattern as represented by the cosine term in Equation 2.2, and then derive the amplitude and phase of V by appropriate calibration. The brightness distribution of the source is obtained from the visibility data by inversion of the transformation in Equation 2.1. Thus V must be measured over a sufficiently wide range of $\nu b \cdot \sigma / c$, which is the component of the baseline normal to the direction of the source and measured in wavelengths. This component can be envisaged as the baseline viewed from the direction of the source.

2.3 Effect of Bandwidth in a Two-Element Interferometer

Since the frequency of the cosine fringe term in Equation 2.2 is proportional to the observing frequency v, observing with a finite bandwidth $\Delta\nu$ results, in effect, in the combination of fringe patterns with a corresponding range of fringe frequencies. For a rectangular frequency passband, the interferometer response is:

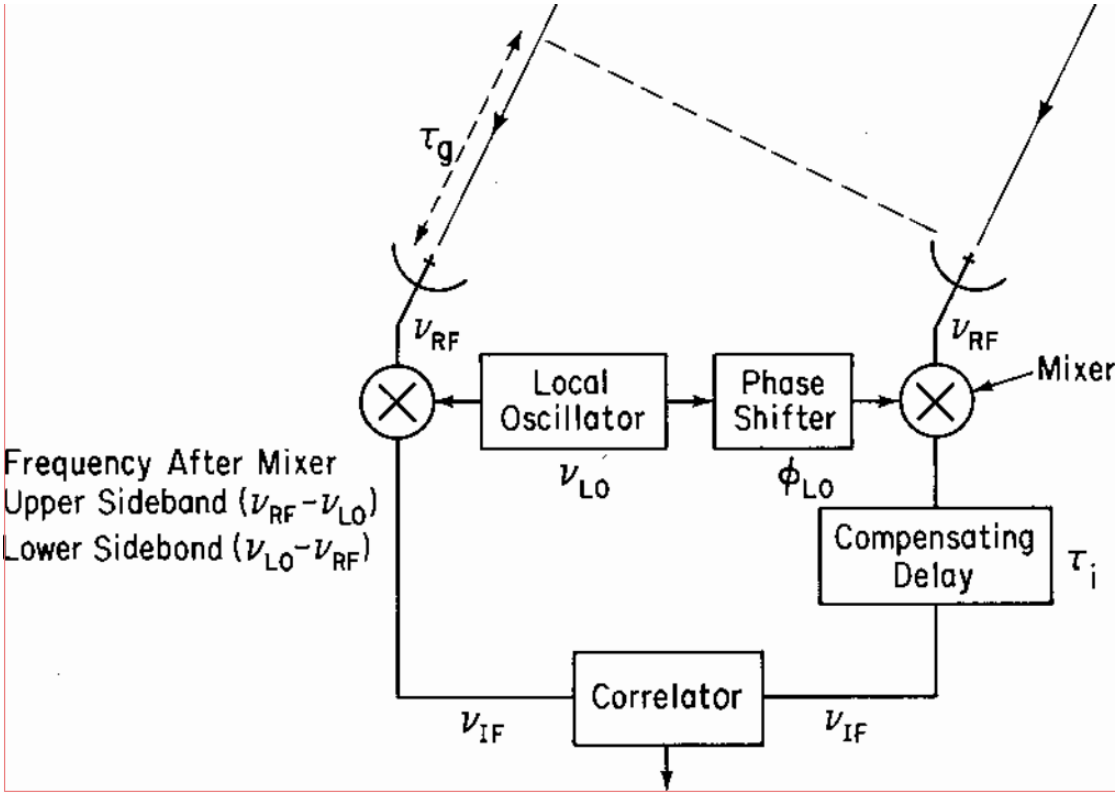
$$r = A_0|V|\Delta\nu\frac{\sin\pi\Delta\nu\tau_g}{\pi\Delta\nu\tau_g}\cos(2\pi\nu_0\tau_g - \phi_V) \tag{2.3}$$

where ν_0 is the center frequency of the observing passband. Thus in the system that we are considering the fringes are modulated by a sinc-function envelope, sometimes referred to as the **bandwidth pattern**. The full fringe amplitude is observed only when the source is in a direction normal to the baseline so that $\tau_g = 0$. The range of τ_g for which the fringe amplitude is within, say, 1% of the maximum value can be obtained by writing:

$$\frac{\sin\pi\Delta\nu\tau_g}{\pi\Delta\nu\tau_g} \approx 1 - \frac{(\pi\Delta\nu\tau_g)^2}{6} > 0.99 \tag{2.4}$$

which yields $|\Delta\nu\tau_g| < 0.078$, where the approximation in Equation 2.4 is valid for $\pi\Delta\nu\tau_g \ll 1$. **The angular range of τ_g, within this limit depends upon the length and orientation of the baseline**: for example, with $\Delta\nu = 50$ MHz and $|b| = 1$ km, the response falls by 1% when the angle θ in Fig. 2-1 is 2 arcmin. In order to observe a source over a wide range of hour-angle, it is necessary to include within the system a computer-controlled delay to compensate for τ_g.

2.4 Delay Tracking and Frequency Conversion



A block diagram of an interferometer system that includes an instrumental compensating delay is shown in Figure 2-4. Frequency conversion of the incoming signals at radio frequency ν_{RF} with a local oscillator at frequency ν_{LO} is als included. Practical receiving systems incorporate frequency conversion because it is technically more convenient to perform such functions as amplification. filtering, delaying, and cross-correlating of the signals at an intermediate frequency that is lower than vgp and remains fixed when the observing frequency is changed. The signals at the frequencies ν_{RF} and ν_{LO} are combined in a mixer which contains a non-linear element (usually a diode) in which combinations o the two frequencies are formed. The intermediate frequency ν_{IF} is related to the mixer input frequencies by

$$\nu_{RF} = \nu_{LO} \pm \nu_{IF}$$

Note that ν_{LO} is a single-valued frequency, but ν_{RF} and ν_{IF} refer to bands of width $\Delta\nu$ **For observations at frequencies up to a few tens of gigahertz the signal from each antenna is usually first applied to a low-noise amplifier to obtain high sensitivity, and then passed through a filter that**

transmits only one of the two sidebands to the mixer. The response of such a single-sideband system can be obtained by considering the phase changes ϕ_1 and ϕ_2 imposed upon the signals received by antennas 1 and 2 before reaching the correlator inputs.
(不理解这节) At frequencies approaching 100 GHz and higher, it is difficult to make low- noise amplifiers to place ahead of the mixers. Often the antenna is connected directly to the mixer input, without any filter to reject one sideband since such a filter can introduce noise unless cryogenically cooled. The result is a double- sideband system, and the response is obtained from the sum of Equation 2-16 and 2-18

Note that the delay-tracking error $\Delta\tau$ does not affect the phase of the cosine fringe term as it does in Equation 2-16 and 2-18, but here it appears in a separate cosine term that modulates the amplitude of the fringes. As a result, the double-sideband system requires more critical adjustment of the instrumental delay to maintain the visibility amplitude than does the single-sideband system. Other disadvantages of the double-sideband system include greater vulnerability to interference, and complication of spectral line observations since the spectra of the two sidebands are superimposed. Separation of the sideband responses after correlation of the signals by a technique involving periodic insertion of $\pi/2$ phase shifts in the local oscillator is used in some instruments: for a discussion see Thompson, Moran & Swenson (1986).

2.5 Fringe Rotation and Complex Correlators

The output from the correlator represented by Equation 2-16, 18 or 19 is fed to a computer which performs some form of optimal analysis to determine the amplitude and phase of the fringe oscillations. The fringe visibility V' can then be obtained by calibration of the instrumental parameters. This calibration usually involves observation of one or more sources with known positions, flux densities, and angular dimensions.

2.6 Phase Switching

Phase switching is a technique that is included in many interferometer systems to eliminate errors in the form of constant or slowly varying offsets in the correlator outputs. Such errors can result from misadjustment of the correlator circuitry, cross coupling between the signals at the correlator inputs, and various other effects.

They can be very effectively reduced by periodically reversing the phase of one of the signals at an early point in the receiving system, and synchronously reversing the sign of the multiplier output in the correlator, before the data are averaged. For the wanted component of the signal, the two reversals cancel one another, but unwanted components in the multiplier output which do not reverse sign with reversal of the phase of a received signal are averaged towards zero. In practice, the frequency of the switching is of the order of 10 or 100 Hz. This technique, known as phase switching, was first introduced by Ryle (1952) as a means of implementing the multiplicative action of a correlator using a power-linear diode detector. For a description of a more recent application of phase switching see Granlund, Thompson & Clark (1978).

2.7 Coordinate Systems for Imaging

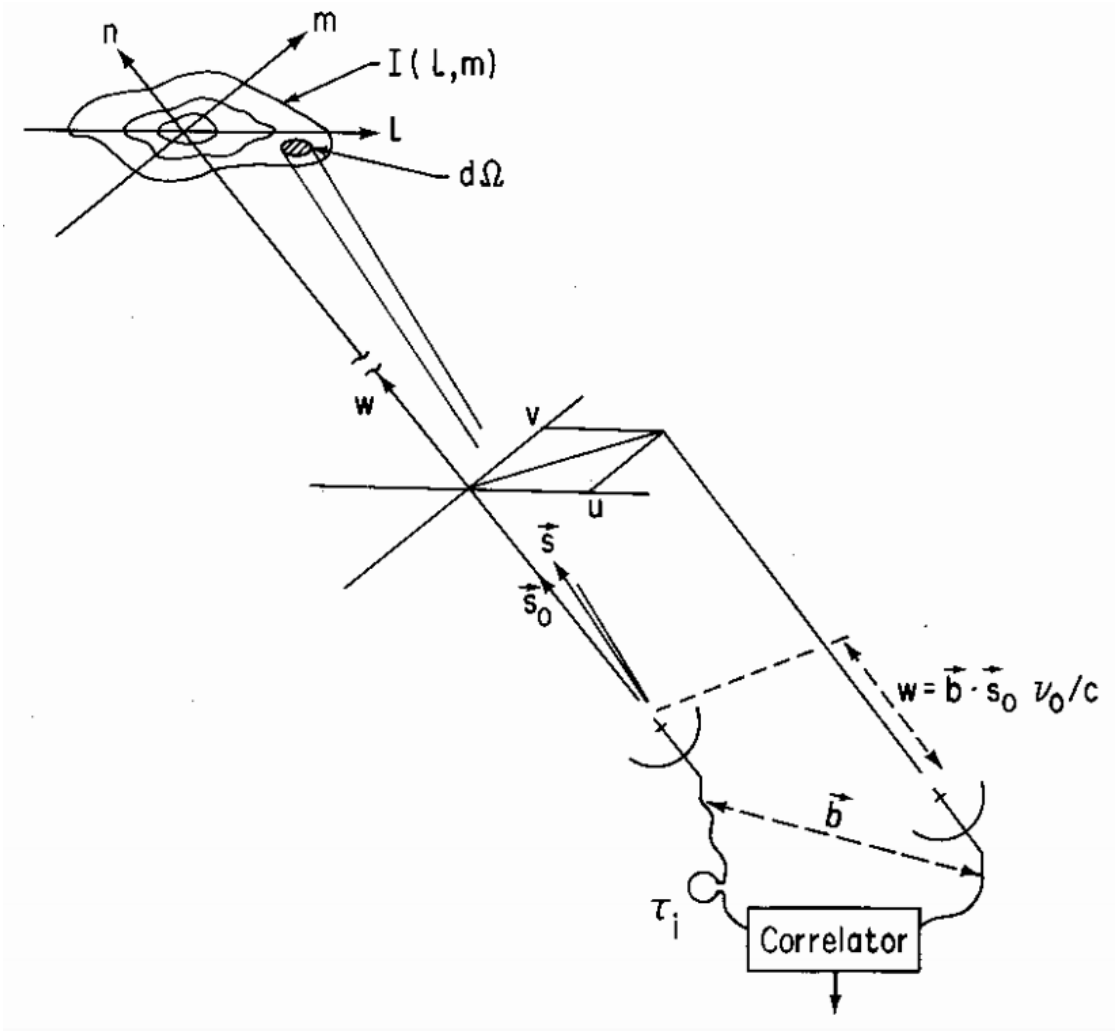


Figure 2.1: The (u,v,w) and (l, m, n) right-handed coordinate systems used to express the interferometer baselines and the source brightness distribution, respectively.

The practical application of Equation 2-7 requires the introduction of a coordinate system, and the one that is usually chosen was introduced in Lecture 1 and is shown in Figure 2-7. The baseline vector has components (u,v,w) where w points in the direction of interest, i.e., towards a position s_0 that becomes the center of the synthesized image. **Note that u, v, and w are measured in wave-lengths**

at the center frequency of the RF signal band, and in directions towards the East, the North, and the phase tracking center, respectively. Positions on the sky are defined in l and m, which are direction cosines measured with respect to the u and v axes. A synthesized image in the (l,m) plane represents a projection of the celestial sphere onto a tangent plane at the (l,m) origin. Distances in l and m are proportional to the sines of the angles measured from the origin, which is a convenient practical system. In these coordinates the parameters used in the derivation of the interferometer response in terms of visibility (Egs. 6 and 7) become

$$\frac{\nu b \cdot s}{c} = ul + vm + wm$$

$$\frac{\nu b \cdot s_0}{c} = w$$

Thus in the coordinates of Figure 2.7, Equation 2.1 becomes:

$$V(u,v,w) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathcal{A}(l,m)I(l,m)e^{-2\pi i[ul+vm+w(\sqrt{1-l^2-m^2}-1)]}\frac{dldm}{\sqrt{1-l^2-m^2}} \tag{2.5}$$

where the integrand is taken to be zero for $l^2 + m^2 > 1$. **Note that we express the complex visibility as a function of (u,v,w), since these are the coordinates that represent the spacings of the antennas with respect to the phase tracking center of the source** s_0**.** The visibility is also a function of the modified brightness distribution $\mathcal{A}I$.

To simplify the inversion of Equation 2.5, by means of which I(l,m) is obtained from the visibility, it is desirable to reduce this equation to the form of a two-dimensional Fourier transform. This can be done under two sets of conditions. The first is when the baselines are coplanar, which can be under- stood by considering the way in which the Earth's rotation carries the antennas through space. It should be evident from Figure 2-8 that the rotation causes the tip of the baseline vector to trace out a circle concentric with the Earth's rotation axis. The rising and setting of a point on the sky usually limit the range over which V can be measured to an arc of the circle. In general, for a two-dimensional array of antennas on the surface of the Earth, the circular loci resulting from the different baselines have different diameters and lie in different planes. However, for the particular case of an array of antennas in an East-West line on the Earth's surface, the components of the baseline vector parallel to the Earth's axis are zero, and the baseline-vectors are coplanar. Then, if we choose the w-axis to lie in the direction of the celestial pole, so that $w = 0$, Equation 2.5 becomes

$$V(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathcal{A}(l,m)I(l,m)e^{-2\pi i[ul+vm]}\frac{dldm}{\sqrt{1-l^2-m^2}} \tag{2.6}$$

This equation is a two-dimensional Fourier transform, the inverse of which is

$$\frac{\mathcal{A}(l,m)I(l,m)}{\sqrt{1-l^2-m^2}} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} V(u,v)e^{-2\pi i[ul+vm]}dudv \tag{2.7}$$

2.8 Antenna Spacing and (u,v,w) Components

With multiple-element antenna arrays, it is convenient to specify the antenna positions relative to some reference point measured in a Cartesian coordinate system.

Thus as the interferometer observes a point on the celestial sphere, the rotation of the Earth causes the u and v components of the baseline to trace out an elliptical locus. This ellipse is simply the projection onto the (u,v) plane of the circular locus traced out by the tip of the baseline vector, as shown earlier in Figure 2-8.

For an array of antennas the ensemble of elliptical loci is known as the **transfer function or sampling function**, S(u,v), which is a function of the declination of the observation as well as of the antenna spacings. The transfer function indicates the values of u and v at which the visibility function is sampled. Since the visibility function for a point source at the (I,m) origin is a constant in u and v, the Fourier transform of the transfer function indicates the response to a point source, i.e., the synthesized beam. In designing arrays the principal aim is to obtain transfer functions that cover the (u,v) plane as widely and as uniformly as possible. The term transfer function was introduced from an analogy with electrical filter theory. An interferometer responds to structure on the sky with spatial frequency u cycles per radian in the l-direction and v cycles per radian in the m-direction. The transfer function of an array therefore indicates its response as a spatial frequency filter.

2.9 Astronomical Data from Interferometer Observations

In synthesis imaging an interferometer or array is used to provide values of the complex visibility as a function of u and v, from which a brightness distribution can be derived. For this purpose the visibility measurements should be fairly uniformly distributed over the (u,v) plane, from the origin to some outer boundary that determines the angular resolution. The design of synthesis arrays, which we discuss below, is based largely upon these considerations. If, however, we wish to measure the positions of a series of unresolved sources, the

principal consideration is the ability to interpolate the measured visibility phase between one baseline and another, and uniformity of coverage is less important. This consideration also applies to measurements used to monitor universal time, polar motion and geodynamic variation in antenna positions.

In addition to the measurement of complex visibility, two other characteristics of the interferometer output can be used to determine astronomical data. These are principally of importance in VLBI, in which it is not always possible to calibrate the interferometer fringe phase. The first is the bandwidth pattern in Equation 2-12, which can be used to measure τ_g. This is accomplished by finding the value of the instrumental delay τ_i that maximizes the fringe amplitude. A wide receiver bandwidth, or a series of narrow bands at different frequencies, is used to minimize the width of the fringe envelope as a function of τ_i and thereby increase the accuracy. For a source at position (H_0, δ_0), τ_g is equal to ω/ν_0 where ω is given by Equation 2-30. The second characteristic that can be measured is the fringe frequency. Since the relative phase of the signal at the two antennas changes by 2π when ω changes by one (wavelength), the fringe frequency is equal to $d\omega/dt$, which can be obtained from Equation 2-30 by differentiation. A useful expression for the fringe frequency ν_F is:

$$\nu_F = \frac{d\omega}{dt} = -\omega_e u \cos\delta \tag{2.8}$$

where $\omega_e = dH_0/dt$ is the angular rotation velocity of the Earth. Equation 2.8 applies when the instrumental delay τ_i is held constant. When τ_i tracks τ_g a factor ν_0/ν_{LO}, ν_0 being the center of the receiving band, should be included for single-sideband receiving systems. In either case, ν_F goes through zero on the v-axis of the (u,v) plane. Note that a single observation of ω and $d\omega/dt$ is sufficient to determine the position of a source if the interferometer baseline is known.

2.10 Design of Synthesis Arrays

In an array of n, antennas, a total of $\frac{1}{2}n_a(n_a|1)$ pair combinations can be formed. The signal from each antenna is then divided in $n_a - 1$ ways and fed to a system of correlators. The rate at which visibility measurements can be made, relative to that for a single interferometer, is approximately proportional to n^2. **Note that since the signals are amplified before splitting there is no loss in sensitivity, as may occur in instruments for infrared or shorter wavelengths.** **The primary concern in designing the configuration of antennas is to**

obtain coverage of the (u,v) plane (ie., sampling of the visibility function) as uniformly and efficiently as possible over a range determined by the required angular resolution.

A commonly used configuration of antennas for synthesis imaging is an East—West linear array. If the various pair combinations of the antennas encompass a series of spacings which increase by a constant increment, the transfer function consists of a series of ellipses centered on the (u,v) origin with a constant increment in the major axes. The axial ratios of the ellipses are equal to $\sin \delta_0$, as in Figure 2-12, which largely determines the axial ratio of the synthesized beam. Thus, for angular distances greater than about 30° from the celestial equator, East—West linear arrays are satisfactory for two-dimensional imaging.

2.11 The effect of Bandwidth in Radio Images

2.12 The Effect of Visibility Averaging

Chapter 3

The Primary Antenna Elements

The primary antenna elements are one of the most important pieces of equipment in a synthesis telescope. Because the performance properties of the antennas can affect the quality of the synthesized images in a number of fundamental ways, the relevant antenna design and performance parameters are reviewed in this lecture.

3.1 Introduction