Fundamentals of Radio Interferometry

Rick Perley, NRAO/Socorro

Thirteenth NRAO Synthesis Imaging Summer School
Socorro, NM
29 May – 5 June, 2012
Topics

• Some Definitions: Intensity, Flux Density, etc.
• The Role of the ‘Sensor’ (a.k.a. ‘Antenna’)
• Key Properties of Antennas
• The Basic Interferometer
Spectral Flux Density and Brightness

• **Our Goal:** To measure the characteristics of celestial emission from a given direction $s$, at a given frequency $\nu$, at a given time $t$.
  – In other words: We want a map, or image, of the emission.

• **Terminology/Definitions:** The quantity we seek is called the brightness (or specific intensity): It is denoted here by $I(s, \nu, t)$, and expressed in units of: watt/(m$^2$ Hz ster).

• It is the power received, per unit solid angle $d\Omega$ from direction $s$, per unit collecting area $dA$, per unit frequency $d\nu$ at frequency $\nu$.

• Do not confuse $I$ with Spectral Flux Density, $S$ -- the integral of the brightness over a given solid angle:

\[
S = \int I(s, \nu, t) d\Omega
\]

• The units of $S$ are: watt/(m$^2$ Hz)

• Note: $1$ Jy $= 10^{-26}$ watt/(m$^2$ Hz).
Intensity and Power.

- Imagine a distant source of emission, described by brightness $I(\nu, s)$ where $s$ is a unit direction vector.
- Power from this emission is intercepted by a collector ("sensor") of area $A(\nu, s)$.
- The power, $dP$ (watts) from a small solid angle $d\Omega$, within a small frequency window $d\nu$, is
  
  $$dP = I(\nu, s)A(\nu, s)d\nu
d\Omega$$

- The total power received is an integral over frequency and angle, accounting for variations in the responses.
  
  $$P = \int\int I(\nu, s)A(\nu, s)d\nu
d\Omega$$
An Example – Cygnus A

• I show below an image of Cygnus A at a frequency of 4995 MHz.
• The units of the brightness are Jy/beam, with 1 beam = 0.16 arcsec²
• The peak is 2.6 Jy/beam, which equates to $6.5 \times 10^{-15}$ watt/(m² Hz ster)
• The flux density of the source is $370 \text{ Jy} = 3.7 \times 10^{-24}$ watt/(m² Hz)
The Role of the Sensor

- Coherent interferometry is based on the ability to correlate the electric fields measured at spatially separated locations.
- To do this (without mirrors) requires conversion of the electric field $E(r,\nu,t)$ at some place to a voltage $V(\nu,t)$ which can be conveyed to a central location for processing.
- For our purpose, the sensor (a.k.a. ‘antenna’) is simply a device which senses the electric field at some place and converts this to a voltage which faithfully retains the amplitudes and phases of the electric fields.

EM waves in

Voltage out (preserving amplitude and phase of all input fields)
Ideal vs. Real Sensor (Antennas)

- An ideal antenna converts only the fields of interest, and adds no additional signal of its own.
- One can imagine two kinds of ideal sensors:
  - An ‘all-sky’ sensor: All incoming electric fields, from all directions, are converted with no direction modulation.
  - The ‘limited-field-of-view’ sensor: Only the fields from a given direction and solid angle (field of view) are collected and conveyed.
- Sadly – neither of these is possible.
- Nor is it possible to build a sensor which does not add in additional power.
  - For even the best antennas, at cm-wave frequencies, a 1 Jy source contributes less than 0.5% of the antenna output power!
Antennas – the Single Dish

• Antennas designs span a wide range – from simple elements with nearly isotropic responses, to major mechanical structures designed for high gain and angular resolution.
• The most common antenna is a parabolic reflector – a ‘single dish’.
• Understanding how it works will help in our later discussion of interferometry.
• There are four critical characteristics of sensors (antennas):
  1. A directional gain (‘main beam’)
  2. An angular resolution given by: $\theta \sim \lambda/D$.
  3. The presence of ‘sidelobes’ – finite response at angles away from the main beam.
• A basic understanding of the origin of these characteristics will aid in understanding the functioning of an interferometer.
The Parabolic Reflector

- **Key Point:** Distance from incoming phase front to focal point is the same for all rays.
- The E-fields will thus all be in phase at the focus – the place for the receiver.
The Standard Parabolic Antenna Response

• The power response of a uniformly illuminated circular parabolic antenna of 25-meter diameter, at a frequency of 1 GHz.
• The FWHM, and the angle to the first null are both approximately $\lambda/D$ (radians).
• There must always be some response outside the main beam – these are called the ‘sidelobes’.
Beam Pattern Origin – Why is there a Null?

• An antenna’s response is a result of coherent phase summation of the electric field at the focus.

• First null will occur at the angle where one extra wavelength of path is added across the full width of the aperture:

$$\theta \sim \frac{\lambda}{D}$$

(Why?)
Specifics: First Null, and First Sidelobe

• When the phase differential across the aperture is 1, 2, 3, … wavelengths, we get a null in the total received power.
  – The nulls appear at (approximately): \( \theta = \frac{\lambda}{D}, \frac{2\lambda}{D}, \frac{3\lambda}{D}, \ldots \) radians.

• When the phase differential across the aperture is \(~1.5, 2.5, 3.5, \ldots\) wavelengths, we get a maximum in total received power.
  – But, each successive maximum is weaker than the last.
  – These maxima appear at (approximately): \( \theta = \frac{3\lambda}{2D}, \frac{5\lambda}{2D}, \frac{7\lambda}{2D}, \ldots \) radians.
Why Interferometry?

• Radio telescopes coherently sum electric fields over an aperture of size D.

• For this, diffraction theory applies – the angular resolution is:

\[ \theta_{\text{rad}} \approx \frac{\lambda}{D} \]

• In ‘practical’ units:

\[ \theta_{\text{arcsec}} \approx \frac{2 \lambda_{\text{cm}}}{D_{\text{km}}} \]

• To obtain 1 arcsecond resolution at a wavelength of 21 cm, we require an aperture of \( \sim 42 \) km!

• The (currently) largest single, fully-steerable aperture is the 100-m antennas in Bonn, and Green Bank. Nowhere big enough!

• Can we synthesize an aperture of that size with pairs of antennas?

• The methodology of synthesizing a continuous aperture through summations of separated pairs of antennas is called ‘aperture synthesis’.
We don’t need a single parabolic structure!

We can consider a series of small antennas, whose individual signals are summed in a network.

This is the basic concept of interferometry.

Aperture Synthesis is an extension of this concept.
Quasi-Monochromatic Radiation

• Analysis is simplest if the fields are perfectly monochromatic.
• This is not possible – a perfectly monochromatic electric field would both have no power ($\Delta \nu = 0$), and would last forever!
• So we consider instead ‘quasi-monochromatic’ radiation, where the bandwidth $d\nu$ is finite, but very small. Then, for times periods $dt < 1/d\nu$, we can consider the field to have a constant amplitude and phase:

$$E_{\nu}(t) = E \cos(2\pi\nu t + \phi)$$

• Consider then the electric fields from a small solid angle $d\Omega$ about some direction $\mathbf{s}$, within some small bandwidth $d\nu$, at frequency $\nu$.
• The amplitude and phase remain unchanged to a time duration of order $dt \sim 1/d\nu$, after which new values of $E$ and $\phi$ are needed.
Simplifying Assumptions

• We now consider the most basic interferometer, and seek a relation between the characteristics of the product of the voltages from two separated antennas and the distribution of the brightness of the originating source emission.

• To establish the basic relations, the following simplifications are introduced:
  – Fixed in space – no rotation or motion
  – Quasi-monochromatic
  – No frequency conversions (an ‘RF interferometer’)
  – Single polarization
  – No propagation distortions (no ionosphere, atmosphere …)
  – Idealized electronics (perfectly linear, perfectly uniform in frequency and direction, perfectly identical for both elements, no added noise, …)
The Stationary, Quasi-Monochromatic Radio-Frequency Interferometer

Geometric Time Delay

\[ \tau_g = b \cdot \frac{s}{c} \]

The path lengths from sensors to multiplier are assumed equal!

\[ V_1 = E \cos[\omega(t - \tau_g)] \]

\[ V_2 = E \cos(\omega t) \]

\[ R_C = P \cos(\omega \tau_g) \]

\[ P[\cos(\omega \tau_g) + \cos(2\omega t - \omega \tau_g)] \]

Unchanging

Rapidly varying,
Pictorial Example: Signals In Phase

2 GHz Frequency, with voltages in phase:

\[ b.s = n\lambda, \text{ or } \tau_g = n/\nu \]

- Antenna 1 Voltage
- Antenna 2 Voltage
- Product Voltage
- Average
Pictorial Example: Signals in Quad Phase

2 GHz Frequency, with voltages in quadrature phase:
\[ b.s = (n +/- \frac{\pi}{4})\lambda, \quad \tau_g = (4n +/- 1)/4\nu \]

- Antenna 1 Voltage
- Antenna 2 Voltage
- Product Voltage
- Average

[Graphs showing voltage waveforms for each case]
Pictorial Example: Signals out of Phase

2 GHz Frequency, with voltages out of phase:
\[ b.s = (n +/\ - \frac{1}{2})\lambda \quad \tau_g = (2n +/\ - 1)/2v \]
Some General Comments

• The averaged product $R_C$ is dependent on the received power, $P = E^2/2$ and geometric delay, $\tau_g$, and hence on the baseline orientation and source direction:

$$R_C = P \cos(\omega \tau_g) = P \cos \left(2\pi \frac{b \cdot s}{\lambda}\right)$$

• Note that $R_C$ is not a function of:
  – The time of the observation -- provided the source itself is not variable!
  – The location of the baseline -- provided the emission is in the far-field.
  – The actual phase of the incoming signal – the distance of the source does not matter, provided it is in the far-field.

• The strength of the product is dependent on the antenna areas and electronic gains -- but these factors can be calibrated for.
To illustrate the response, expand the dot product in one dimension:

\[ \frac{b \cdot s}{\lambda} = u \cos \alpha = u \sin \theta = ul \]

Here, \( u = \frac{b}{\lambda} \) is the baseline length in wavelengths, and \( \theta \) is the angle w.r.t. the plane perpendicular to the baseline.

\( l = \cos \alpha = \sin \theta \) is the direction cosine.

Consider the response \( R_c \), as a function of angle, for two different baselines with \( u = 10 \), and \( u = 25 \) wavelengths:

\[ R_c = \cos(2\pi ul) \]
Whole-Sky Response

• Top:
  \[ u = 10 \]

There are 20 whole fringes over the hemisphere.

• Bottom:
  \[ u = 25 \]

There are 50 whole fringes over the hemisphere.
From an Angular Perspective

**Top Panel:**

The absolute value of the response for \( u = 10 \), as a function of angle.

The ‘lobes’ of the response pattern alternate in sign.

**Bottom Panel:**

The same, but for \( u = 25 \).

Angular separation between lobes (of the same sign) is

\[ \delta \theta \sim \frac{1}{u} = \frac{\lambda}{b} \text{ radians.} \]
Hemispheric Pattern

- The preceding plot is a meridional cut through the hemisphere, oriented along the baseline vector.
- In two-dimensional space, the coherence pattern consists of a series of coaxial cones, oriented along the baseline vector.
- The figure is a two-dimensional representation when $u = 4$.
- As viewed along the baseline vector, the fringes show a ‘bulls-eye’ pattern – concentric circles.
The Effect of the Sensor

- The patterns shown presume the sensor has isotropic response.
- This is a convenient assumption, but (sadly, for some) doesn’t represent reality.
- Real sensors impose their own patterns, which modulate the amplitude and phase of the output.
- Large sensors (a.k.a. ‘antennas’) have very high directivity -- very useful for some applications.
The Effect of Sensor Patterns

- Sensors (or antennas) are not isotropic, and have their own responses.

- **Top Panel:** The interferometer pattern with a \( \cos(\theta) \)-like sensor response.

- **Bottom Panel:** A multiple-wavelength aperture antenna has a narrow beam, but also sidelobes.
The Response from an Extended Source

- The response from an extended source is obtained by summing the responses at each antenna to all the emission over the sky, multiplying the two, and averaging:

\[ R_C = \left\langle \int V_1 d\Omega_1 \int V_2 d\Omega_2 \right\rangle \]

- The averaging and integrals can be interchanged and, providing the emission is spatially incoherent, we get

\[ R_C = \iint I_v(s) \cos(2\pi \nu \mathbf{b} \cdot \mathbf{s}/c) d\Omega \]

- This expression links what we want – the source brightness on the sky, \( I_v(s) \), – to something we can measure - \( R_C \), the interferometer response.

- Can we recover \( I_v(s) \) from observations of \( R_C \)?
A Schematic Illustration in 2-D

• The correlator can be thought of ‘casting’ a cosinusoidal coherence pattern, of angular scale $\sim \lambda/b$ radians, onto the sky.

• The correlator multiplies the source brightness by this coherence pattern, and integrates (sums) the result over the sky.

• Orientation set by baseline geometry.
• Fringe separation set by (projected) baseline length and wavelength.
  • Long baseline gives close-packed fringes
  • Short baseline gives widely-separated fringes
• Physical location of baseline unimportant, provided source is in the far field.
Odd and Even Functions

• Any real function, $I(x,y)$, can be expressed as the sum of two real functions which have specific symmetries:

\[
I(x,y) = I_E(x,y) + I_O(x,y)
\]

An even part:

\[
I_E(x,y) = \frac{I(x,y) + I(-x,y)}{2} = I_E(-x,y)
\]

An odd part:

\[
I_O(x,y) = \frac{I(x,y) - I(-x,y)}{2} = -I_O(-x,y)
\]
But One Correlator is Not Enough!

- The correlator response, $R_c$:

$$R_c = \iint I_v(s) \cos(2\pi v \mathbf{b} \cdot s/c) d\Omega$$

is only sensitive to the even part of the sky brightness.

- Suppose that the source of emission has a component with odd symmetry:

$$I_o(s) = -I_o(-s)$$

- Since the cosine fringe pattern is even, the response of our interferometer to the odd brightness distribution is 0!

$$R_c = \iint I_o(s) \cos(2\pi v \mathbf{b} \cdot s/c) d\Omega = 0$$

- Hence, we need more information if we are to completely recover the source brightness.
Why Two Correlations are Needed

- The integration of the cosine response, $R_c$, over the source brightness is sensitive to only the even part of the brightness:

$$R_c = \int \int I(s) \cos(2\pi \nu b \cdot s/c) d\Omega = \int \int I_E(s) \cos(2\pi \nu b \cdot s/c) d\Omega$$

since the integral of an odd function ($I_O$) with an even function ($\cos x$) is zero.
- To recover the ‘odd’ part of the intensity, $I_O$, we need an ‘odd’ fringe pattern. Let us replace the ‘cos’ with ‘sin’ in the integral.
- We get:

$$R_s = \int \int I(s) \sin(2\pi \nu b \cdot s/c) d\Omega = \int \int I_o(s) \sin(2\pi \nu b \cdot s/c) d\Omega$$

since the integral of an even times an odd function is zero.
- Thus, the even part of the image is ‘seen’ by the COS correlation, while the odd part is ‘seen’ by the SIN correlation.
- Hence, to get the full image, we need both.
Making a SIN Correlator

• We generate the ‘sine’ pattern by inserting a 90 degree phase shift in one of the signal paths.

\[ \tau_g = b \cdot \frac{s}{c} \]

\[ V = E \cos(\omega(t - \tau_g)) \]

\[ P[\sin(\omega \tau_g) + \sin(2\omega t - \omega \tau_g)] \]

\[ R_s = P \sin(\omega \tau_g) \]
Define the Complex Visibility

• We now DEFINE a complex function, the complex visibility, \( V \), from the two independent (real) correlator outputs \( R_C \) and \( R_S \):

\[
V = R_C - iR_S = Ae^{-i\phi}
\]

where

\[
A = \sqrt{R_C^2 + R_S^2}
\]

\[
\phi = \tan^{-1}\left( \frac{R_S}{R_C} \right)
\]

• This gives us a beautiful and useful relationship between the source brightness, and the response of an interferometer:

\[
V_v(b) = R_C - iR_S = \int \int I_v(s) e^{-2\pi i b \cdot s / c} d\Omega
\]

• Under some circumstances, this is a 2-D Fourier transform, giving us a well established way to recover \( I(s) \) from \( V(b) \).
The Complex Correlator and Complex Notation

• A correlator which produces both ‘Real’ and ‘Imaginary’ parts – or the Cosine and Sine fringes, is called a ‘Complex Correlator’
  – For a complex correlator, think of two independent sets of projected sinusoids, 90 degrees apart on the sky.
  – In our scenario, both components are necessary, because we have assumed there is no motion – the ‘fringes’ are fixed on the source emission, which is itself stationary.

• The complex output of the complex correlator also means we can use complex analysis throughout: Let:

\[
V_1 = A \cos(\omega t) = \text{Re} \ Ae^{-i\omega t}
\]

\[
V_2 = A \cos[\omega (t - b \cdot s / c)] = \text{Re} \ Ae^{-i\omega (t - b \cdot s / c)}
\]

• Then:

\[
P_{\text{corr}} = \left\langle V_1 V_2^* \right\rangle = A^2 \ e^{-i\omega b \cdot s / c}
\]
Picturing the Visibility

• The source brightness is Gaussian, shown in black.
• The interferometer ‘fringes’ are in red.
• The visibility is the integral of the product – the net dark green area.

Long Baseline

Short Baseline

\[ R_C \quad R_S \]
Examples of 1-Dimensional Visibilities

- Simple pictures are easy to make illustrating 1-dimensional visibilities.

- Unresolved Doubles

- Uniform

- Central Peaked
More Examples

- Simple pictures are easy to make illustrating 1-dimensional visibilities.

- Resolved Double

- Resolved Double

- Central Peaked Double
Basic Characteristics of the Visibility

• For a zero-spacing interferometer, we get the ‘single-dish’ (total-power) response.
• As the baseline gets longer, the visibility amplitude will in general decline.
• When the visibility is close to zero, the source is said to be ‘resolved out’.
• Interchanging antennas in a baseline causes the phase to be negated – the visibility of the ‘reversed baseline’ is the complex conjugate of the original.
• Mathematically, the visibility is Hermitian, because the brightness is a real function.
Some Comments on Visibilities

• The Visibility is a unique function of the source brightness.
• The two functions are related through a Fourier transform. \( V_v(u, v) \Leftrightarrow I(l, m) \)
• An interferometer, at any one time, makes one measure of the visibility, at baseline coordinate \((u, v)\).
• Sufficient knowledge of the visibility function (as derived from an interferometer) will provide us a reasonable estimate of the source brightness.
• How many is ‘sufficient’, and how good is ‘reasonable’?
• These simple questions do not have easy answers…