Fundamentals of Radio Interferometry

Rick Perley, NRAO/Socorro

ATNF Radio Astronomy School

Narrabri, NSW

29 Sept. – 03 Oct. 2014

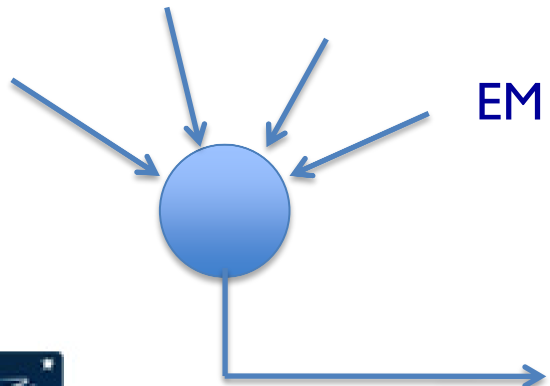
Atacama Large Millimeter/submillimeter Array
Expanded Very Large Array
Robert C. Byrd Green Bank Telescope
Very Long Baseline Array

Topics

- Introduction: Sensors, Antennas, Brightness, Power
- Quasi-Monochromatic Approximation
- The Basic Interferometer
  - Response to a Point Source
  - Response to an Extended Source
  - The Complex Correlator
  - The Visibility and its relation to the Intensity
  - Picturing the Visibility

Essentials of Sensors (Antennas)

- Coherent interferometry is based on the ability to correlate the electric fields at spatially separated locations.
- Doing this requires transport of the electric field E(**s**,ν,t), or a surrogate, at various locations **r** to a central location for analysis.
- In the radio regime, the normal practice is to convert the E-field to a voltage V(ν,t) which can be conveyed to a central location for processing. Note that information on direction **s** is lost.
- The ideal sensor is a device which responds to 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

A real sensor will modify the amplitudes and phases, but in a simple, slowly varying way.

Voltage out (preserving relative amplitude and phases of all input fields)

NRAO

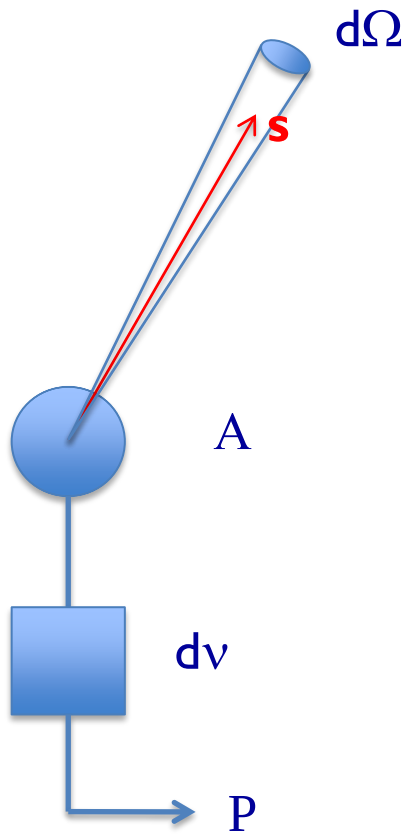
Brightness and Power.

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

$$dP = I(\upsilon,\mathbf{s})A(\upsilon,\mathbf{s})d\nu\, d\Omega$$

- The total power received is an integral over frequency and angle, accounting for variations in the responses.

$$P = \iint I(\upsilon,\mathbf{s})A(\upsilon,\mathbf{s})d\upsilon d\Omega$$



Solid Angle

Sensor Area

Filter width

Power collected

Quasi-Monochromatic Radiation

- Analysis is simplest if the fields are monochromatic.
- A perfectly monochromatic electric field ($\delta\nu = 0$), cannot exist in nature – it would both no power and would last forever.
- So we consider instead 'quasi-monochromatic' radiation, where the bandwidth $\delta\nu$ is very small, but not zero.
- Then, for a time dt $\sim 1/\delta\nu$, the electric fields will be sinusoidal, with unchanging amplitude and phase.
- Consider then the electric field from a small solid angle dΩ about some direction **s**, within some small bandwidth dν, at frequency ν.
- We can write the temporal dependence of this field as:

$$E_\upsilon(t) = A\cos(2\pi\upsilon t + \phi)$$

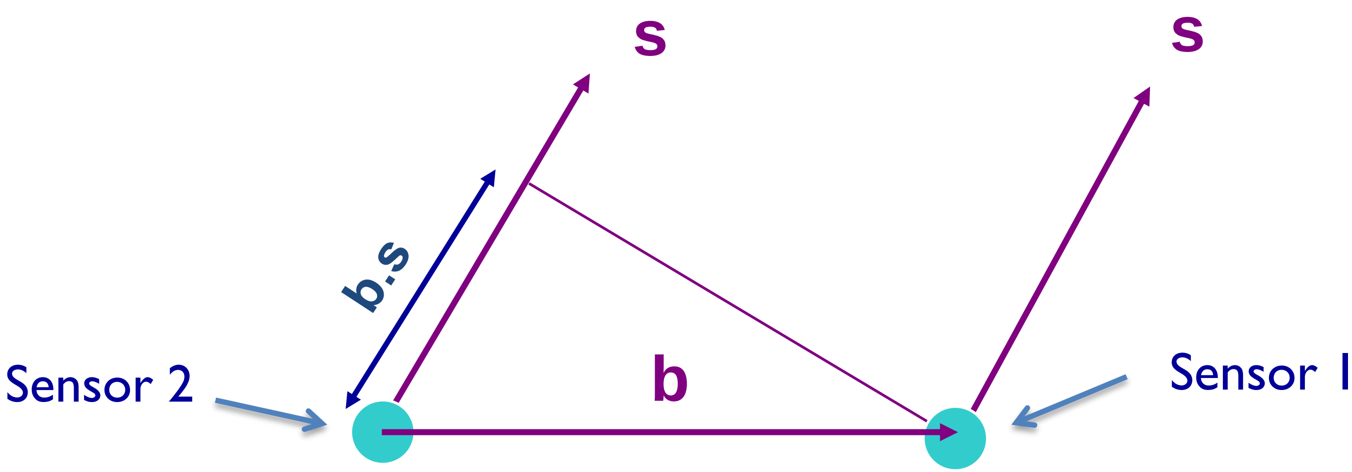
- The amplitude and phase remains unchanged to a time duration of order dt $\sim 1/\text{d}\nu$, after which new values of **A** and ϕ are needed.

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 (signals are sinusoidal)
  - No frequency conversions (an 'RF interferometer')
  - Single polarization
  - No propagation distortions (no ionosphere, atmosphere …)
  - Idealized electronics (perfectly linear, no amplitude or phase perturbations, perfectly identical for both elements, no added noise, …)

NRAO

Basic Concepts of Interferometry



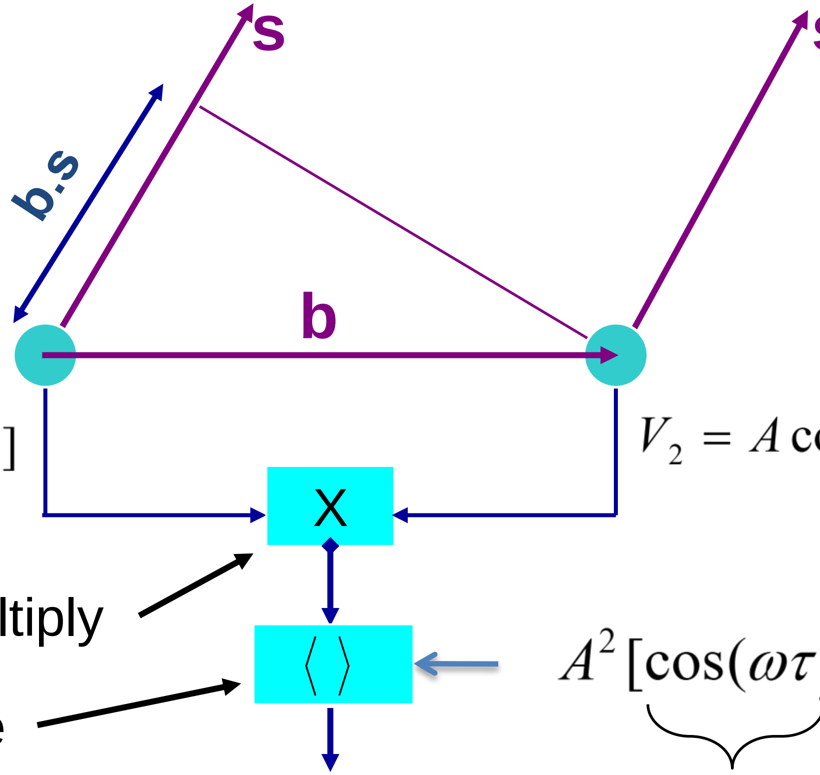
- There are two sensors, separated by vector baseline **b**
- Radiation arrives from direction **s** – assumed the same for both (far-field).
- The extra propagation path is L = **b.s**
- The time taken for this extra path is $\tau_g = \mathbf{b} \cdot \mathbf{s} / c$
- For radiation of wavelength λ, we have a phase given by:

$$\varphi = 2\pi \mathbf{b} \bullet \mathbf{s} / \lambda = 2\pi \upsilon \tau_g = \omega \tau_g \quad \text{(radians)}$$

The Stationary, Quasi-Monochromatic Radio-Frequency Interferometer

Geometric Time Delay

$$\tau_g = \mathbf{b} \cdot \mathbf{s} / c$$



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

$$V_2 = A\cos(\omega t)$$

The path lengths from sensors to multiplier are assumed equal!

multiply

average

$$A^2[\cos(\omega\tau_g) + \cos(2\omega t - \omega\tau_g)]$$

Unchanging

Rapidly varying,

$$R_C = P\cos(\omega\tau_g)$$

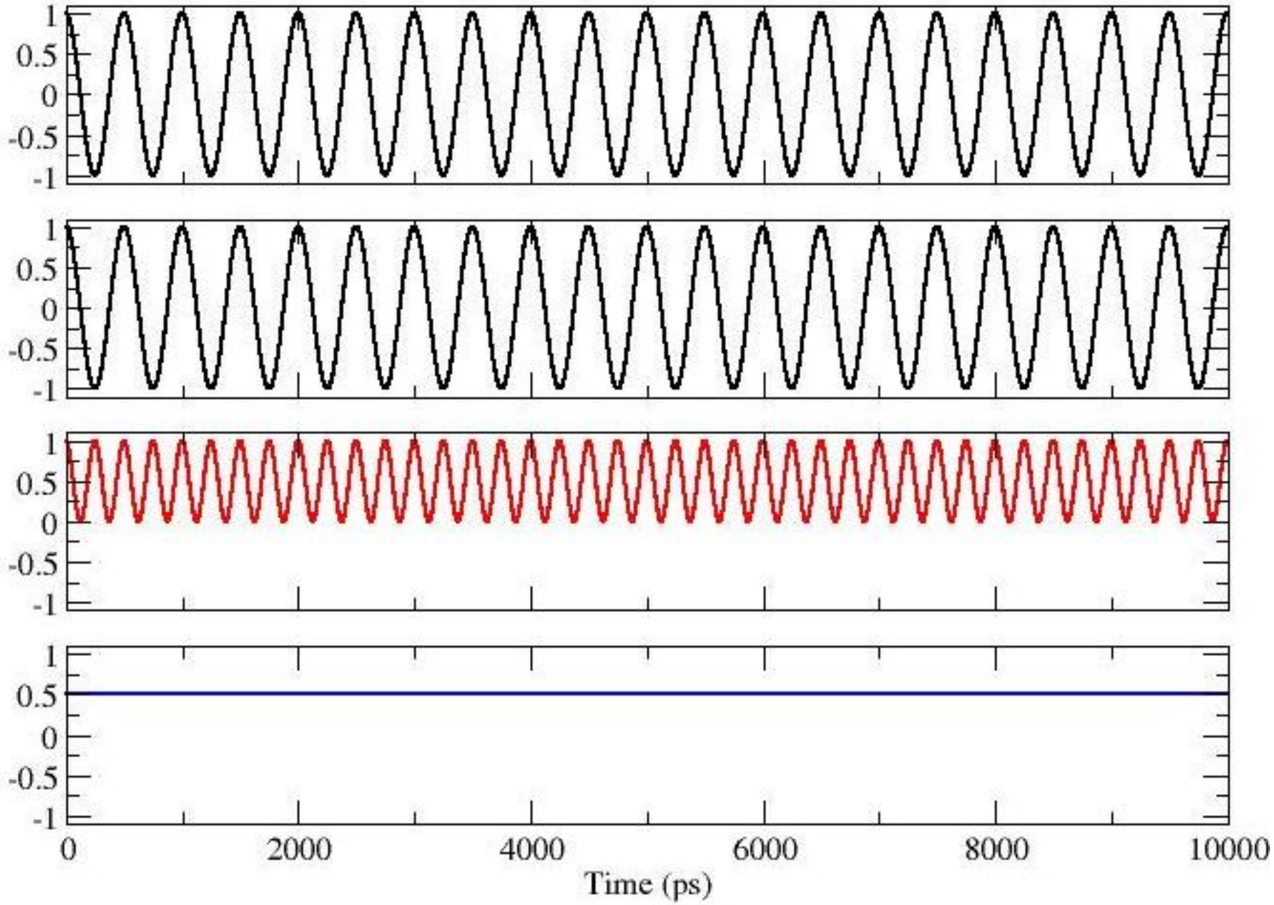
Note: R_c is not a function of time or location!

Pictorial Example: Signals In Phase

If the voltages arrive in phase:

b.s = nλ, or τ_g = n/ν (n is an integer)

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

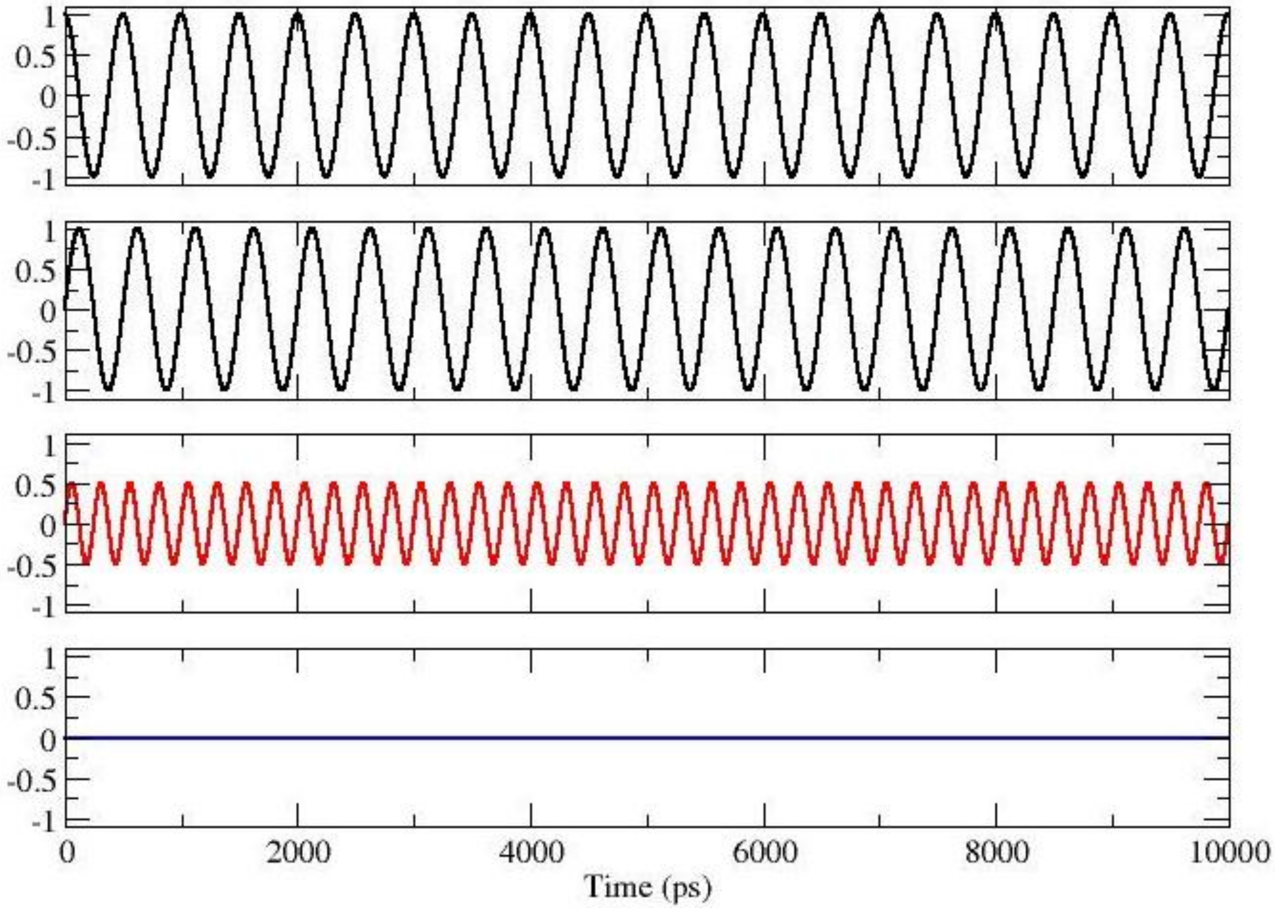


Pictorial Example: Signals in Quad Phase

If the voltages arrive in quadrature phase:

$b.s=(n \pm \frac{1}{4})\lambda, \quad \tau_g = (4n \pm 1)/4\nu$

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

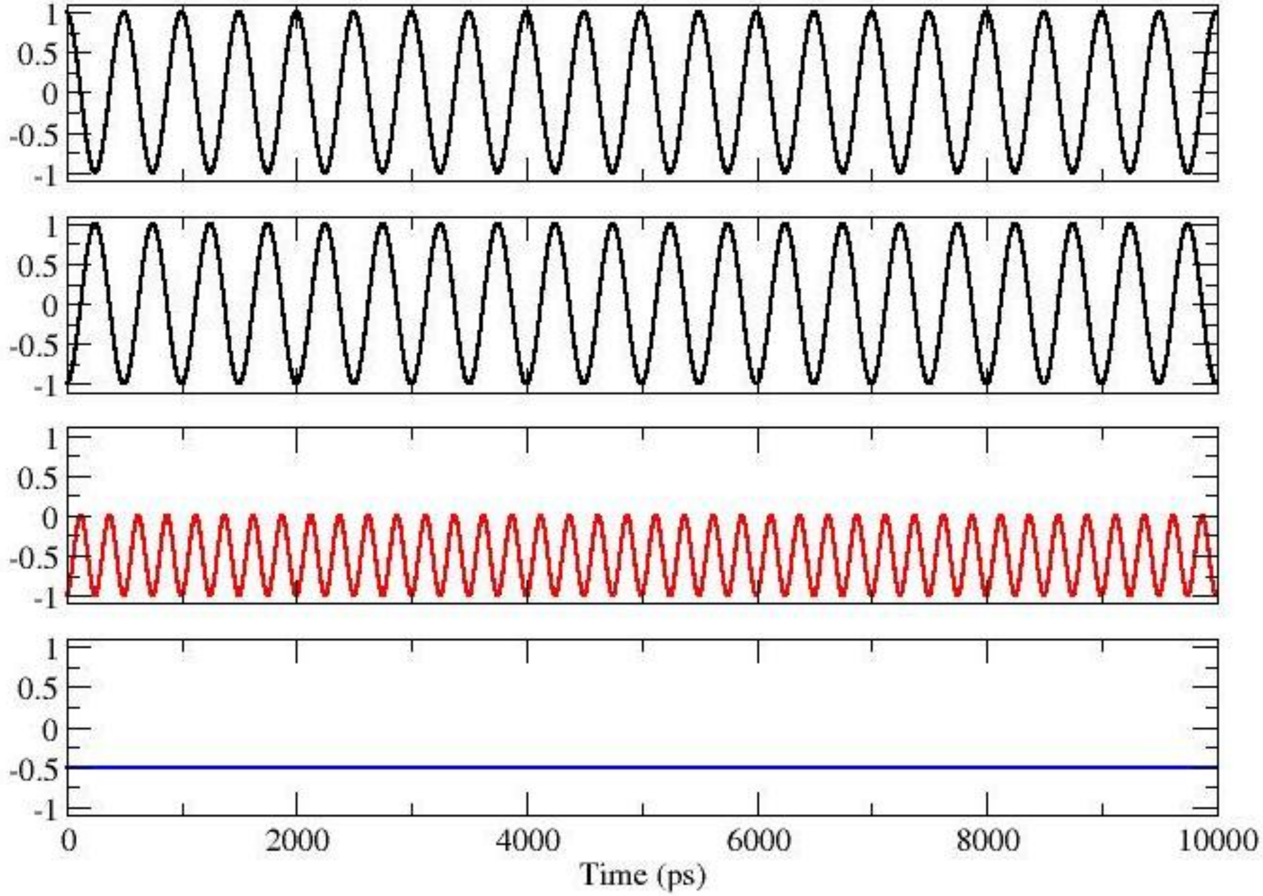


Pictorial Example: Signals out of Phase

If the signals arrive with voltages out of phase:

$b.s=(n+\frac{1}{2})\lambda \qquad \tau_g = (2n+1)/2\nu$

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



Some General Comments

- In all cases, the output is a steady voltage, with the amplitude dependent upon the signal strength, and the phase relationship.

- The averaged product R_C is dependent on the received power, $P = A^2/2$ and geometric delay, τ_g, and hence on the baseline orientation and source direction:

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

- Note that R_C is not a 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.

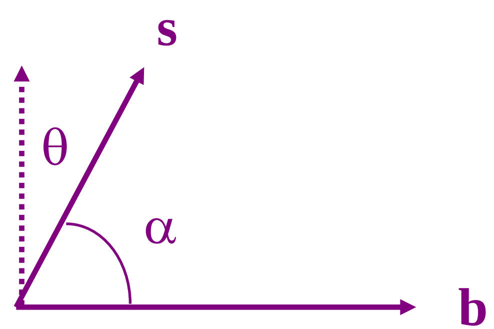
- The strength of the product is dependent on the antenna collecting areas and electronic gains – but these factors can be calibrated for.

Pictorial Illustrations

- To illustrate the response, expand the dot product in one dimension:

$$\frac{2\pi \mathbf{b} \bullet \mathbf{s}}{\lambda} = 2\pi \frac{b}{\lambda} \cos\alpha = 2\pi u \sin\theta = 2\pi u l$$

  - Where u = b/λ is the baseline length in wavelengths,
  - α is the angle w.r.t. the baseline vector
  - $l = \cos\alpha = \sin\theta$ is the direction cosine



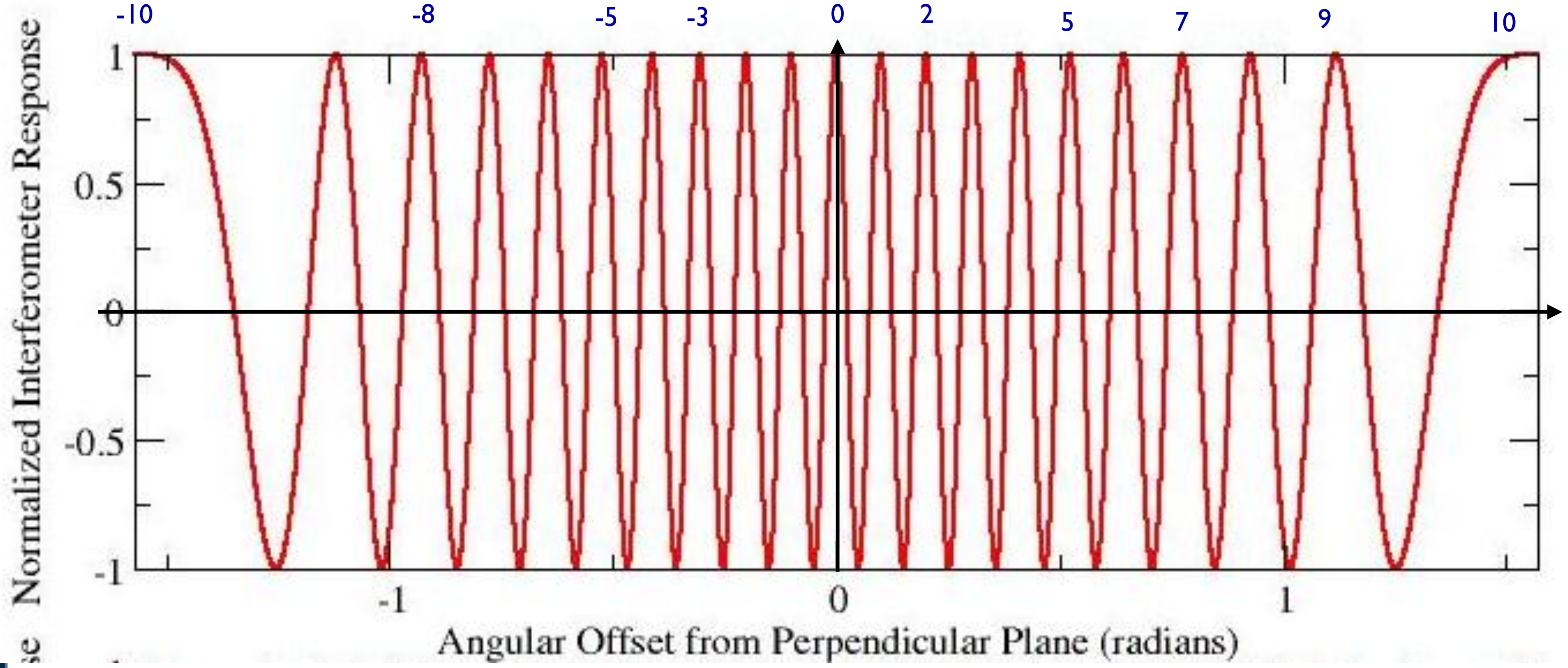
- Consider the response R_c, as a function of angle, for two different baselines with u = 10, and u = 25 wavelengths:

Whole-Sky Response for u = 10

- When u = 10 (i.e., the baseline is 10 wavelengths long), the response is

$$R_C = \cos(20\pi l)$$

- There are 21 fringe maxima over the whole hemisphere, with maxima at l = n/10 radians.
- Minimum fringe separation 1/10 radians

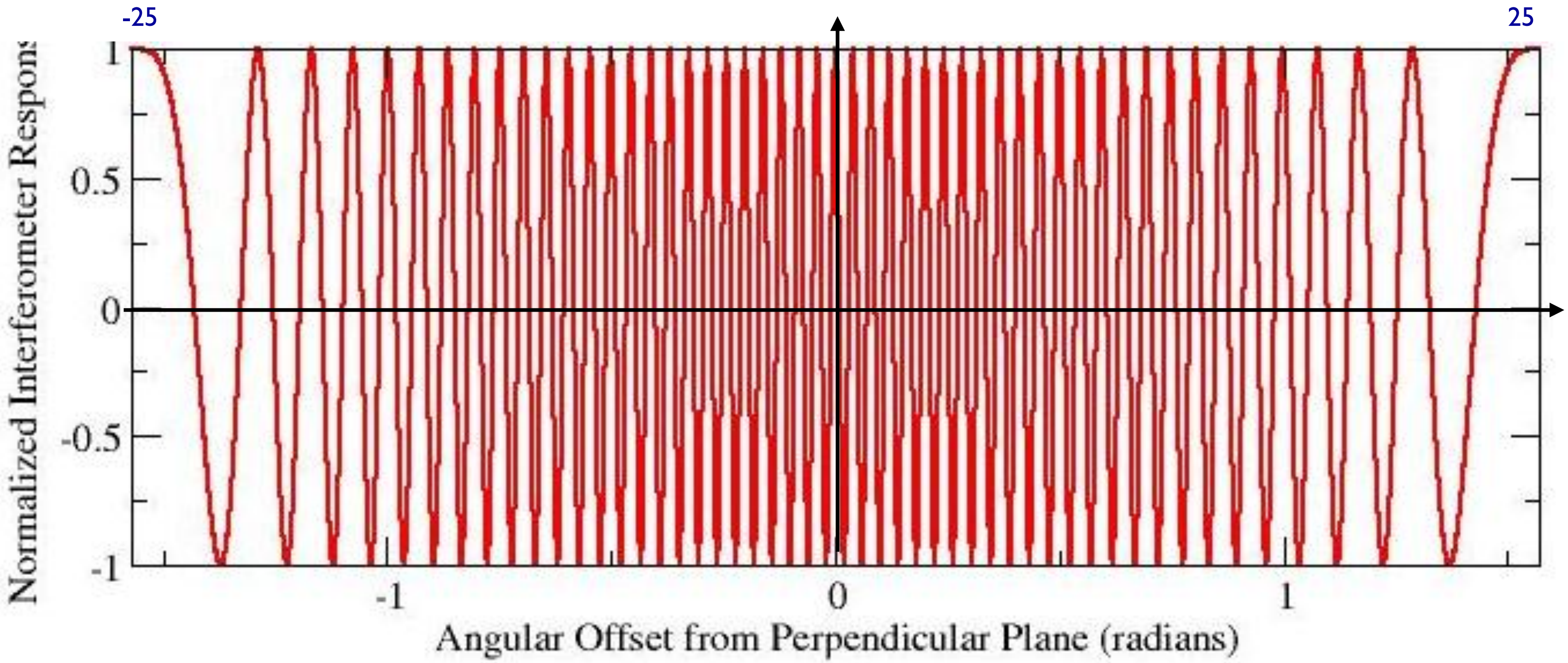


Whole-Sky Response for u = 25

For u = 25 (i.e., a 25-wavelength baseline), the response is

$$R_C = \cos(50\pi l)$$

- There are 51 whole fringes over the hemisphere.
- Minimum fringe separation 1/25 radians.



From an Angular Perspective

Top Panel:

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

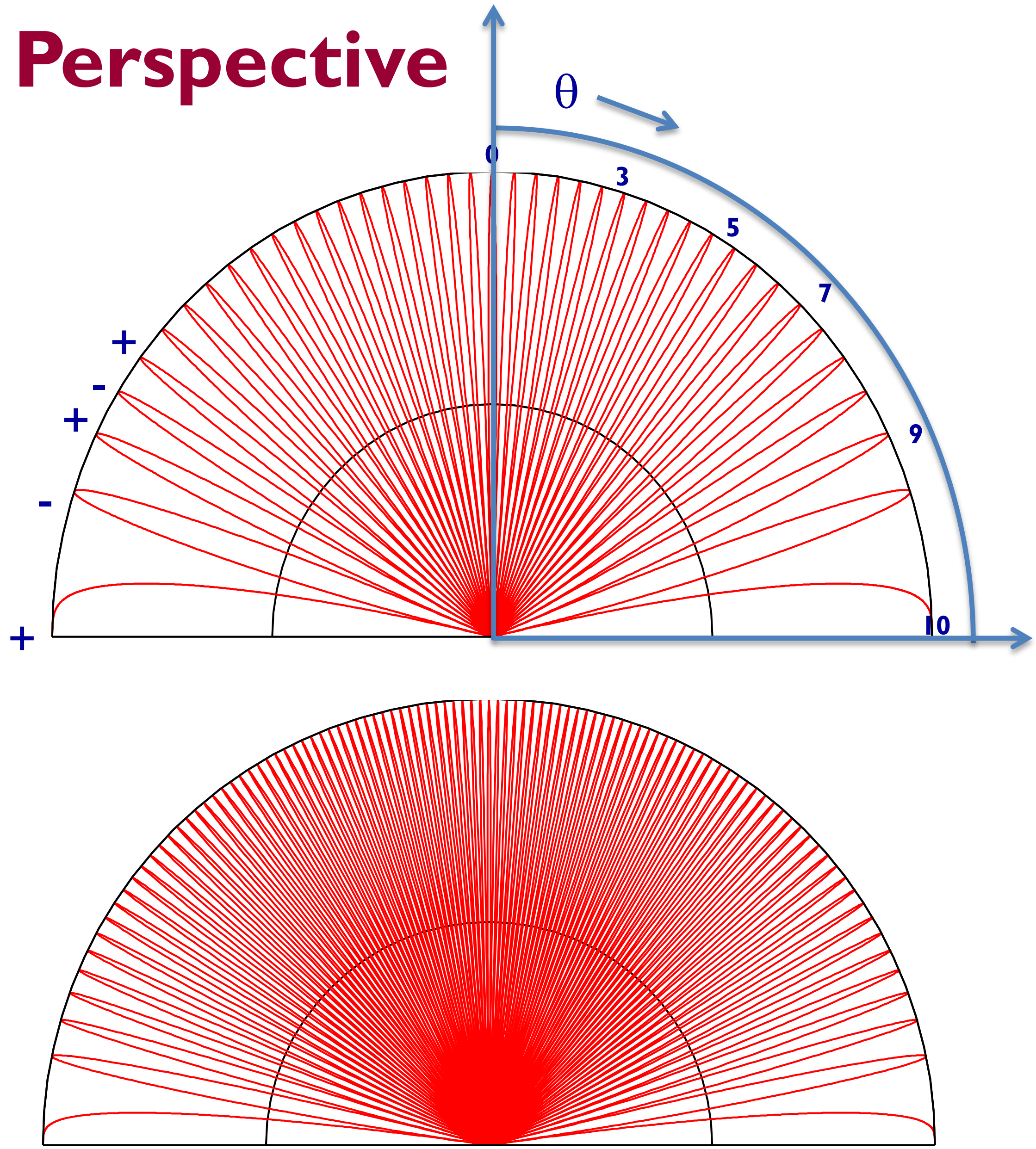
The ‘lobes’ of the response pattern alternate in sign.

Bottom Panel:

The same, but for u = 25.

Angular separation between lobes (of the same sign) is

$\delta\theta \sim 1/u = \lambda/b$ radians.



NRAO

Hemispheric Pattern

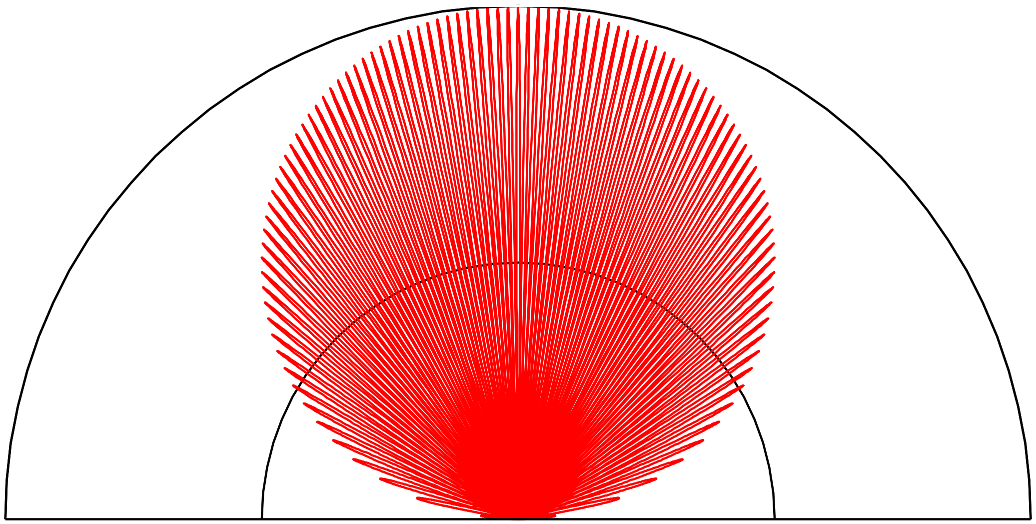
- The preceding plot is a meridional cut through the hemisphere, oriented along the baseline vector.
- In the two-dimensional space, the fringe pattern consists of a series of coaxial cones, oriented along the baseline vector.
- The figure is a two-dimensional representation when u = 4.
- As viewed along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

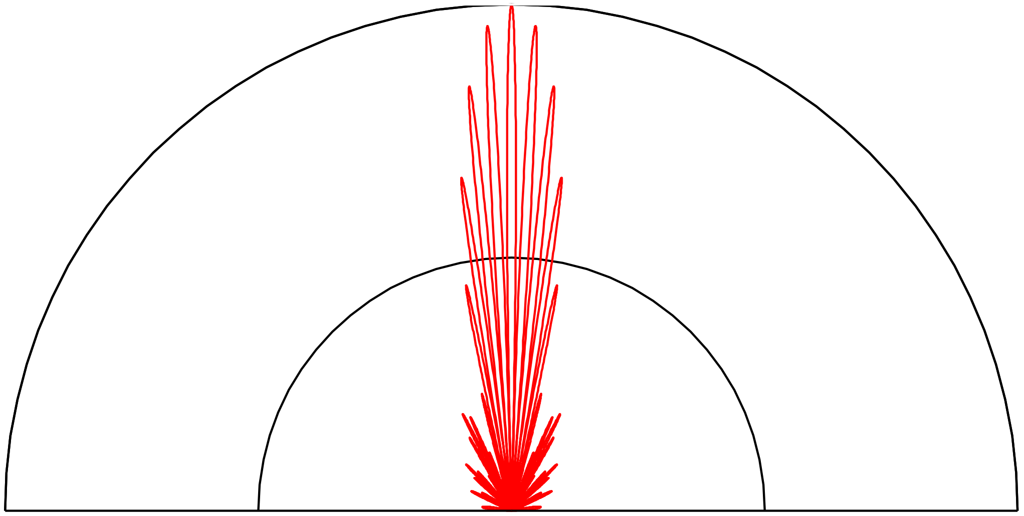
The Effect of the Sensor

- The patterns shown presume the sensor (antenna) has isotropic response.
- This is a convenient assumption, but doesn't represent reality.
- Real sensors impose their own patterns, which modulate the amplitude and phase, of the output.
- Large antennas have very high directivity -- very useful for some applications.
- Small antennas have low directivity – nearly uniform response for large angles – useful for other 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(θ)-like sensor response.
- **Bottom Panel:** A multiple-wavelength aperture antenna has a narrow beam, but also sidelobes.
- Note that the phase will also be modified.





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:

$$R_C = \left\langle \iint V_1 d\Omega_1 \times \iint V_2 d\Omega_2 \right\rangle$$

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

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

- This expression links what we want – the source brightness on the sky, $I_\nu(\mathbf{s})$, – to something we can measure - R_C, the interferometer response.
- Can we recover $I_\nu(\mathbf{s})$ from observations of R_C?
- NB I have assumed here isotropic sensors. If not, a directional attenuation function must be added.

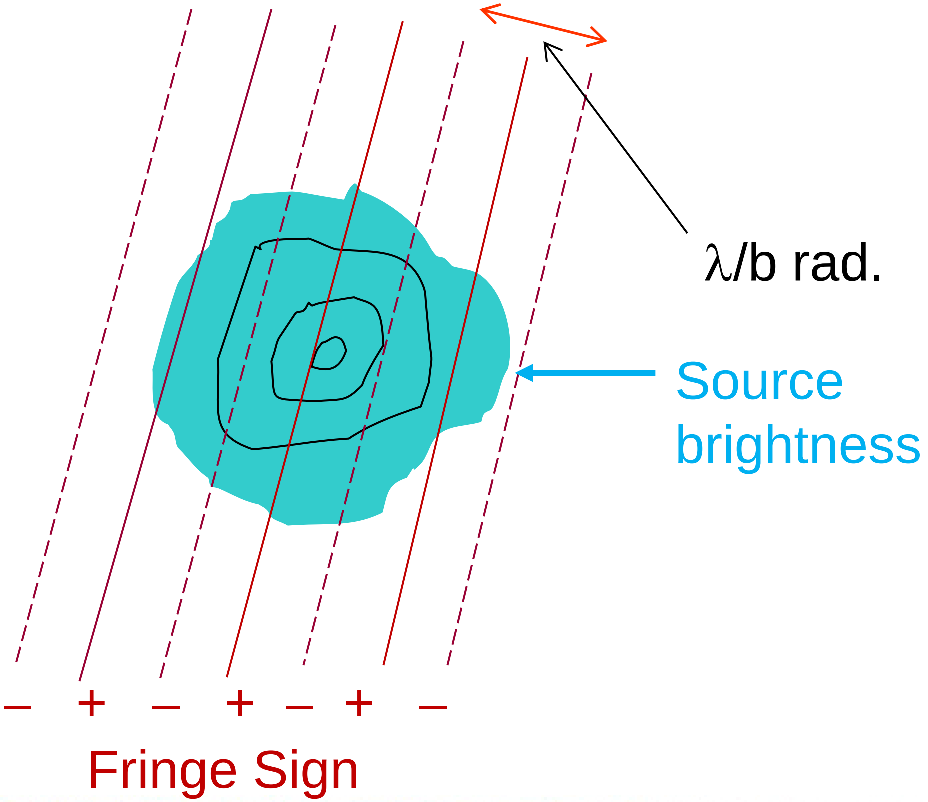
NRAO

A Schematic Illustration in 2-D

- The correlator can be thought of multiplying the actual sky brightness by a cosinusoidal coherence pattern, of angular scale ~λ/b radians.

- The correlator then integrates (adds) the modified brightness pattern over the whole sky (as weighted by the antenna response.

- Pattern 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.



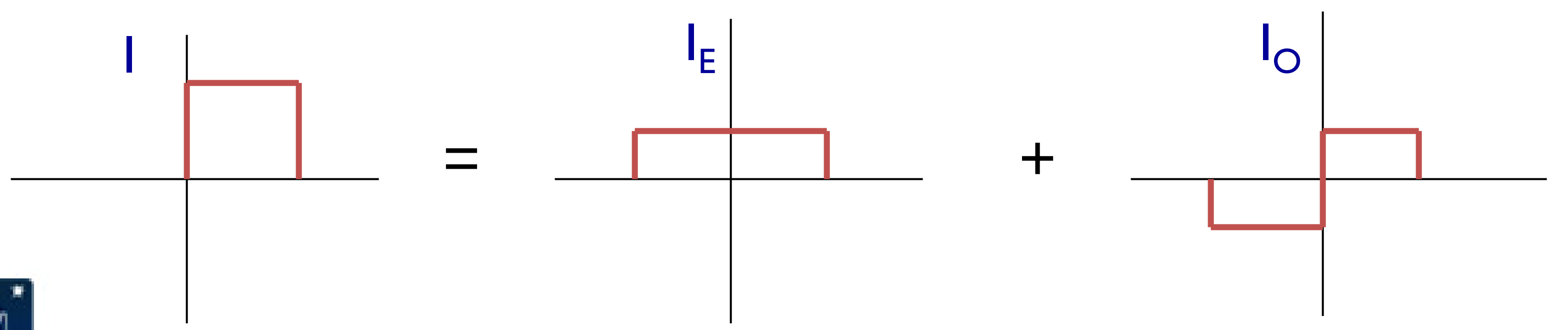
A Short Mathematics Digression – Odd and Even Functions

- 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: $$I_O(x,y) = \frac{I(x,y) - I(-x,-y)}{2} = -I_O(-x,-y)$$



Why One Correlator is Not Enough

- The correlator response, R_c:

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

is not enough to recover the correct brightness. Why?

- Express the brightness as the sum of its even and odd parts:

$$I = I_E + I_O$$

- Then form the correlation:

$$R_C = \iint I(\mathbf{s}) \cos(2\pi \nu \mathbf{b} \cdot \mathbf{s}/c) d\Omega = \iint I_E(\mathbf{s}) \cos(2\pi \nu \mathbf{b} \cdot \mathbf{s}/c) d\Omega$$

- Since the cosine fringe pattern is even, the response of our interferometer to the odd brightness distribution is 0.
- The Odd symmetric component, I_O is invisible, and is lost.
- Hence, we need more information if we are to completely recover the source brightness.

Why Two Correlations are Needed

- To recover the 'odd' part of the brightness, I_O, we need an 'odd' fringe pattern. Let us replace the 'cos' with 'sin' in the integral, to find:

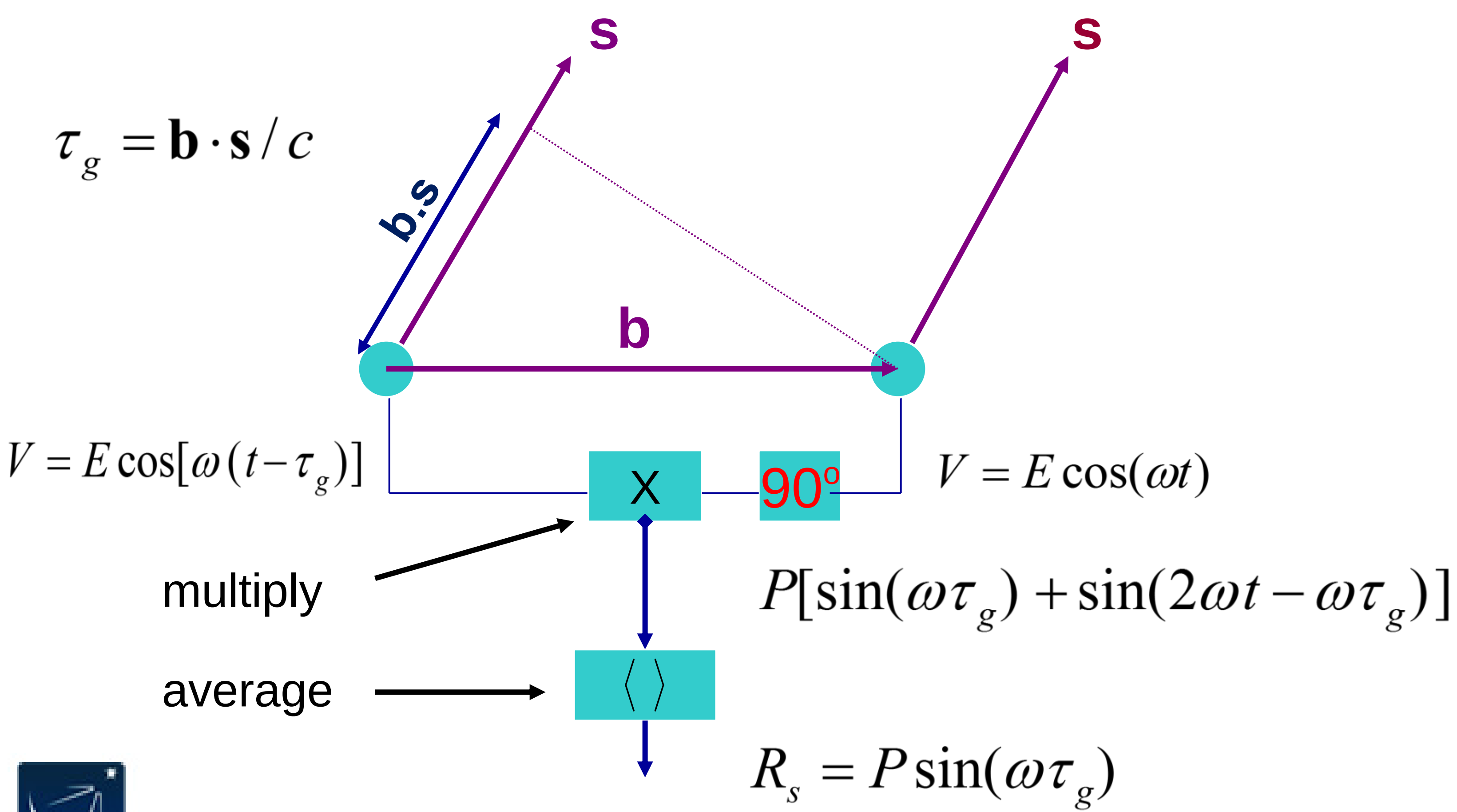
$$R_S = \iint I(\mathbf{s})\sin(2\pi\nu\mathbf{b}\cdot\mathbf{s}/c)d\Omega = \iint I_O(\mathbf{s})\sin(2\pi\nu\mathbf{b}\cdot\mathbf{s}/c)d\Omega$$

since the integral of an even times an odd function is zero.

- Thus, to provide full information on both the even and odd parts of the brightness, we require two separate correlators.
  - An 'even' (COS) and an 'odd' (SIN) correlator.
- Note that this requirement is a consequence of our assumption of no motion – the fringe pattern and the source intensity are both fixed.
- One can build a correlator which 'sweeps' its fringes across the sources – providing both fringe types.

Making a SIN Correlator

- We generate the 'sine' pattern by inserting a 90 degree phase shift in one of the signal paths.



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:

$$V_\upsilon(\mathbf{b}) = R_C - iR_S = \iint I_\nu(s)\, e^{-2\pi i \nu\, \mathbf{b}\cdot\mathbf{s}/c}\, d\Omega$$

- This is a Fourier transform – but with a quirk: The visibility distribution is in general 3-dimensional, while the brightness distribution is only 2-dimensional. More on this, later.

NRAO

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:

$$V_1 = A\cos(\omega t) = \mathrm{Re}\left(Ae^{-i\omega t}\right)$$
$$V_2 = A\cos[\omega(t - \mathbf{b} \cdot \mathbf{s}/c)] = \mathrm{Re}\left(Ae^{-i\omega(t-\mathbf{b}\bullet\mathbf{s}/c)}\right)$$

- Then:

$$P_{corr} = \left\langle V_1 V_2^* \right\rangle = A^2 e^{-i\omega\mathbf{b}\bullet\mathbf{s}/c}$$

Wideband Phase Shifters – Hilbert Transform

- For a quasi-monochromatic signal, forming a the 90 degree phase shift to the signal path is easy --- add a piece of cable $\lambda/4$ wavelengths long.
- For a wideband system, this obviously won't work.
- In general, a wideband device which phase shifts each spectral component by 90 degrees, while leaving the amplitude intact, is a Hilbert Transform.
- For real interferometers, such an operation can be performed by analog devices.
- Far more commonly, this is done using digital techniques.
- The complex function formed by a real function and its Hilbert transform is termed the 'analytic signal'.

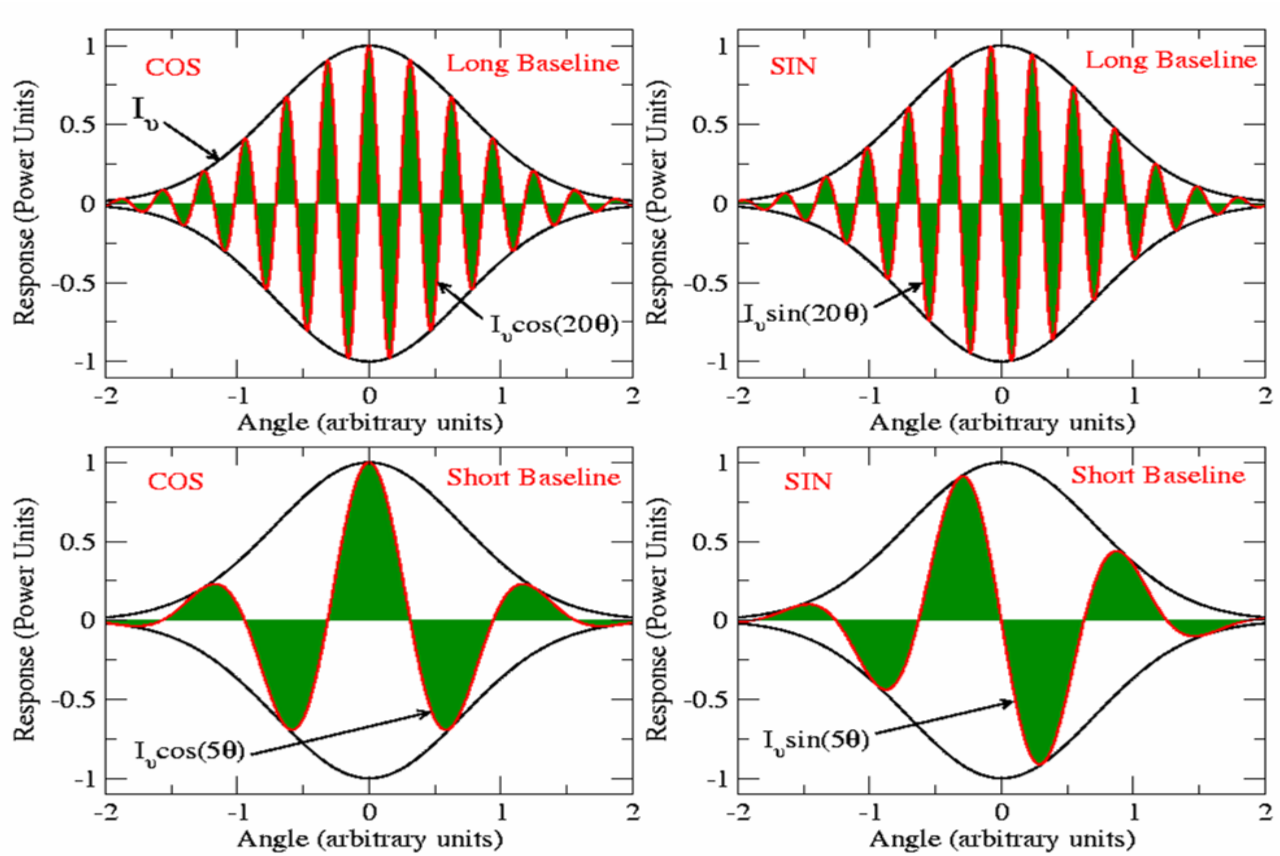
Picturing the Visibility

- 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.

R_C R_S

Long Baseline

Short Baseline

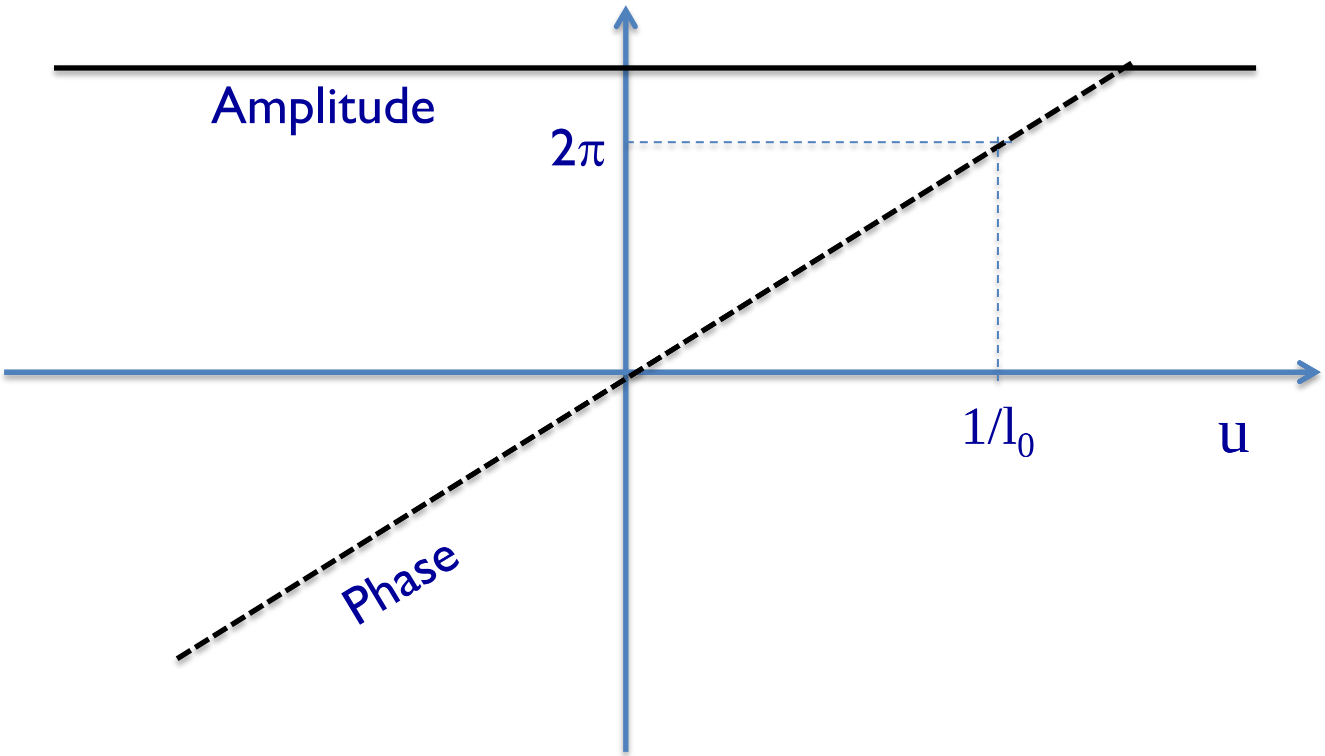


Examples of 1-dimensional Visibilities.

- Picturing the visibility-brightness relation is simplest in one dimension.
- For this, the relation becomes $V_\upsilon(u) = \int I_\upsilon(l)\, e^{-2\pi i u l}\, dl$
- Simplest example: A unit-flux point source: $I(l) = \delta(l - l_0)$
- The visibility is then: $$V(u) = e^{-2\pi i u l_0}$$
- For a source at the origin (l_0=0), V(u) = 1. (units of Jy).
- For a source off the origin, the visibility has unit amplitude, and a phase slope with baseline, rotating 360 degrees every $1/l_0$ wavelengths.

NRAO

Point Source Visibility



- For a point source, the visibility amplitude equals the source flux density
- The phase slope ($=2\pi l_0$) gives the source position.

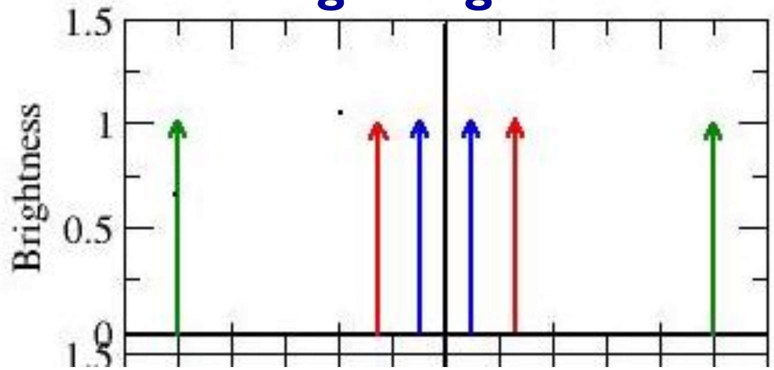
A Symmetric Double Point Source

- Mathematically, this is $I(l) = \delta(l - l_0) + \delta(l + l_0)$
- The Visibility is: $V(u) = 2\cos(2\pi u l_0)$

which is a cosinusoid of amplitude = 2, reaching its maxima at multiples of $1/l_0$. The phase = 0.

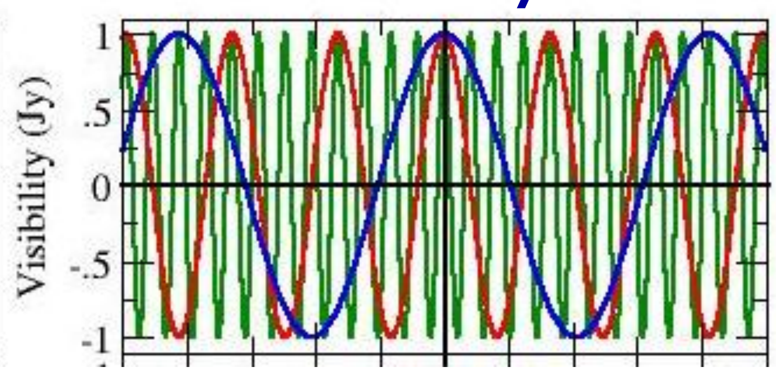
- Note the symmetry: The brightness is real and even, so the visibility is real and even.

Image Brightness



Angle Offset

Visibility



Baseline

Extended Symmetric Source

- A point source has a constant visibility amplitude for all baselines. An extended source's visibility declines with baseline.
- Consider a 'top-hat' source of width l_0. Insertion into the relation shows:

$$V(u) = \frac{\sin(\pi u l_0)}{\pi u l_o} = \mathrm{sinc}(u l_0)$$

- A 'Triangle' source of full width $2l_0$ has a visibility function

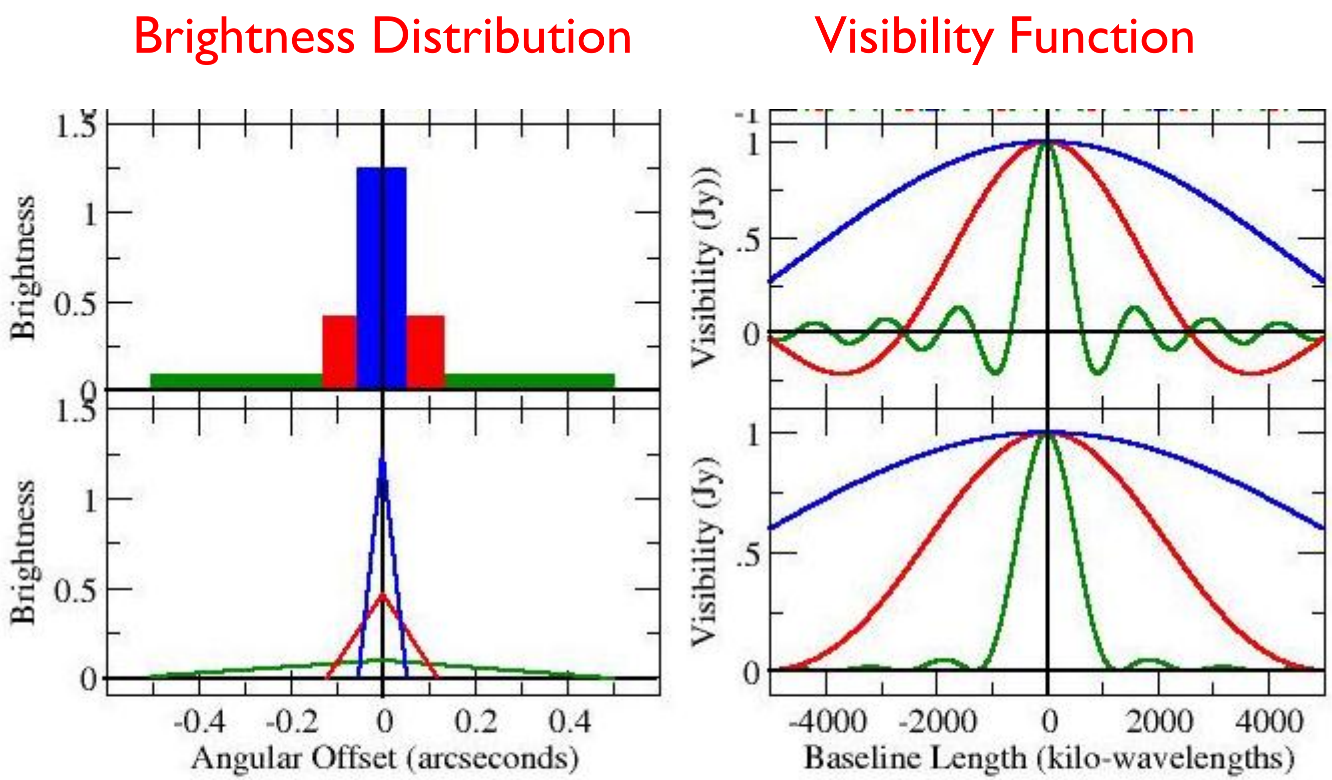
$$V(u) = \left(\frac{\sin(\pi u l_0}{\pi u l_0} \right)^2 = \mathrm{sinc}^2(u l_0)$$

- For both of these, the visibility equals 0 at all multiples of u = l_0.

Examples of 1-Dimensional Visibilities

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

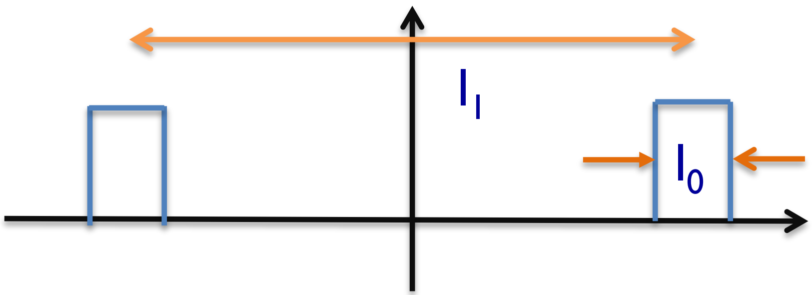
- 'Top-Hat' Sources
- Triangle Sources



For these examples, the visibility peaks are all the same (=1), reflecting the integrated flux density of the sources is the same (= 1).

Extended Symmetric Doubles

- Suppose you have a source consisting of two 'top-hat' sources, each of width l_0, separated by l_1 radians.



- Analysis provides: $$V(u) = \mathrm{sinc}(u\,l_0)\cos(\pi u l_1)$$

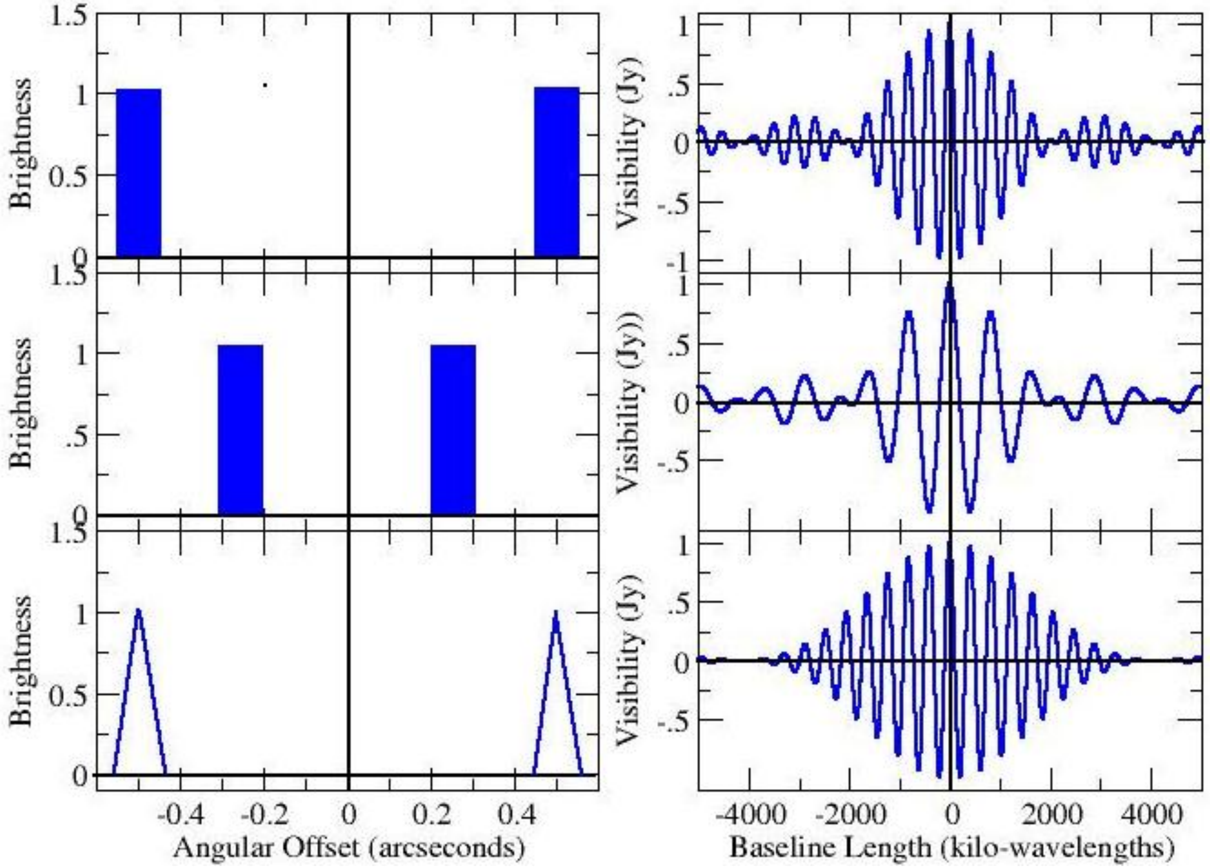
Which is an oscillatory function of period $u = 1/l_1$ attenuated by a dying oscillation of period $u = 1/l_0$.

More Examples

- Simple pictures illustrating 1-dimensional visibilities.

Brightness Distribution | Visibility Function

- Resolved Double
- Resolved Double
- Central Peaked Double



Another Way to Conceptualize …

- Return to the generalized definition of the visibility:

$$V_\upsilon(\mathbf{b}) = \iint I_\nu(s) e^{-2\pi i \nu \mathbf{b}\cdot\mathbf{s}/c} d\Omega$$

- The interferometer casts a **phase slope** across the brightness distribution.
  - The phase slope becomes steeper for longer baselines, or higher frequencies.
  - The phase slope is zero for zero baseline. (V(0) = S)
  - The phase is zero at the phase origin.
  - The amplitude response is unity (ignoring the primary beam) throughout.
- The Visibility is the complex integral (sum) of the brightness multiplied by the phase ramp.

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. (Why?)
- Mathematically, the visibility is Hermitian. ($V(u) = V^*(-u)$).

NRAO

Some Comments on Visibilities

- The Visibility is a unique function of the source brightness.
- The two functions are related through a Fourier transform. $V_\upsilon(u,v) \Leftrightarrow I(l,m)$
- An interferometer, at any one time, makes one measure of the visibility, at baseline coordinate (u,v).
- `Sufficient knowledge' of the visibility function 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…