Fundamentals of Radio Interferometry

Rick Perley, NRAO/Socorro

15th Synthesis Imaging School

Socorro, NM

01 – 09 June, 2016

Associated Universities, Inc.

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

NRAO

Topics

- The Need for Interferometry
- Some Basics:
  - Antennas as E-field Converters
  - Conceptual Interferometry
  - 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

Why Interferometry?

- It's all about angular resolution.
- A 'single dish' has a natural resolution given by

$$\theta_{\text{radian}} \approx \lambda / D$$

- In 'practical units', this is expressed as:

$$\theta_{\text{arcmin}} \approx 38\ \lambda_{\text{cm}} / \mathrm{D}_{\text{m}}$$

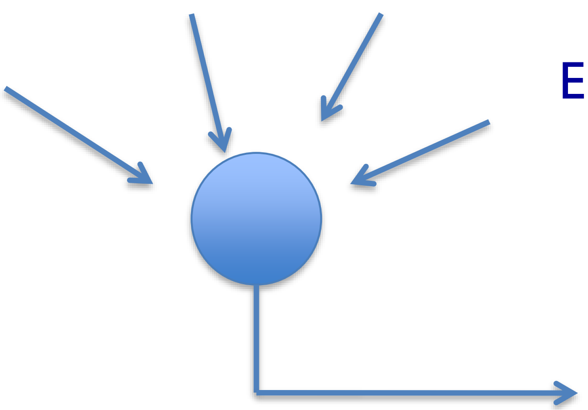
- So, for a 100 meter antenna, at λ = 20cm, the resolution is a rather modest 7 arcminutes.
- If we want 1 arcsecond* resolution, we would need an aperture ~400 times larger => an aperture 40 km in size.
- No way to build such a structure on earth – we need another way – interferometry.

* 1 arcsecond = angular size of a dime at 2.3 miles!

NRAO

The Antenna as an EM Wave Converter

- 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**,ν,t) at some place to a voltage V(ν,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.



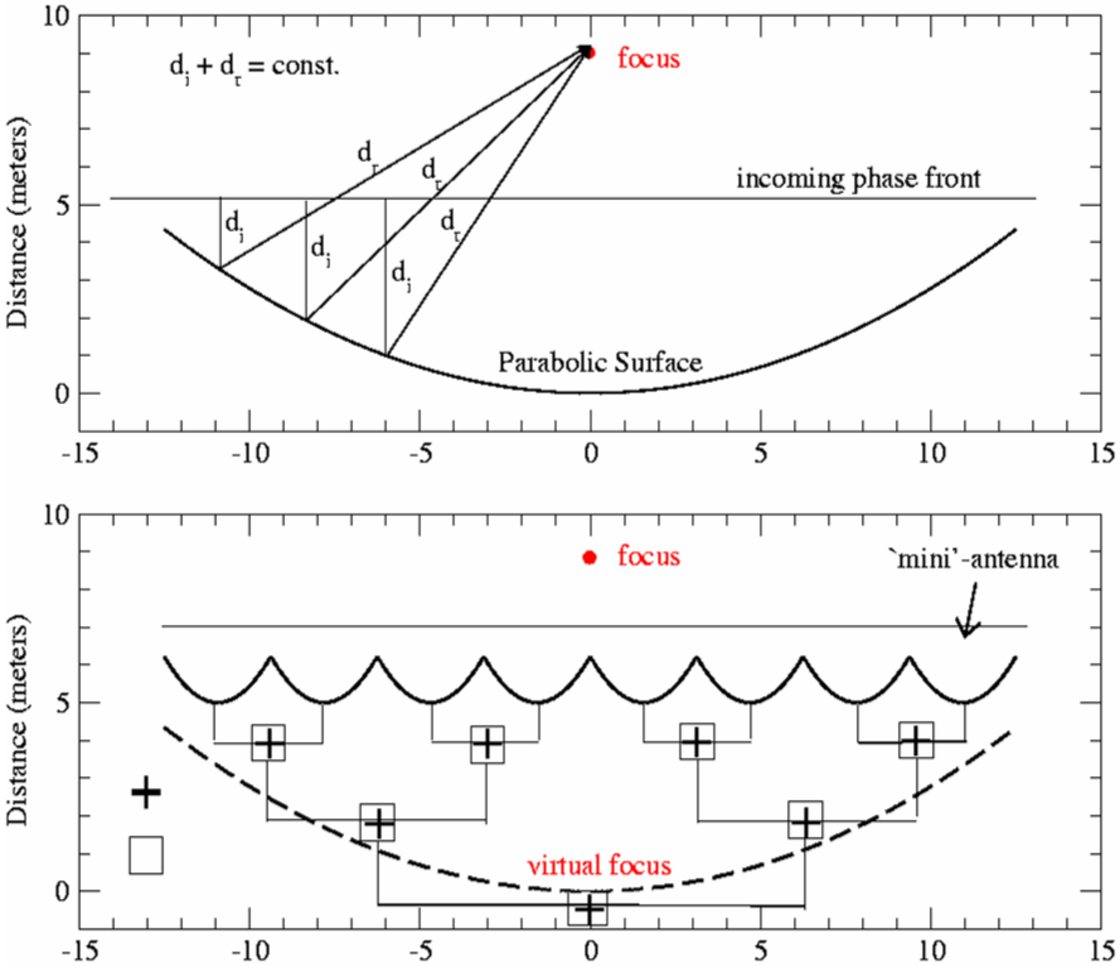
EM waves in (Volts/m)

Signal out (in Volts, preserving amplitudes and phases of all input fields)

Interferometry – Basic Concept

- A parabolic dish is actually a free-space interferometer.
- An incoming plane wave (normal to the antenna axis) is coherently summed at (and only at!) the focus.

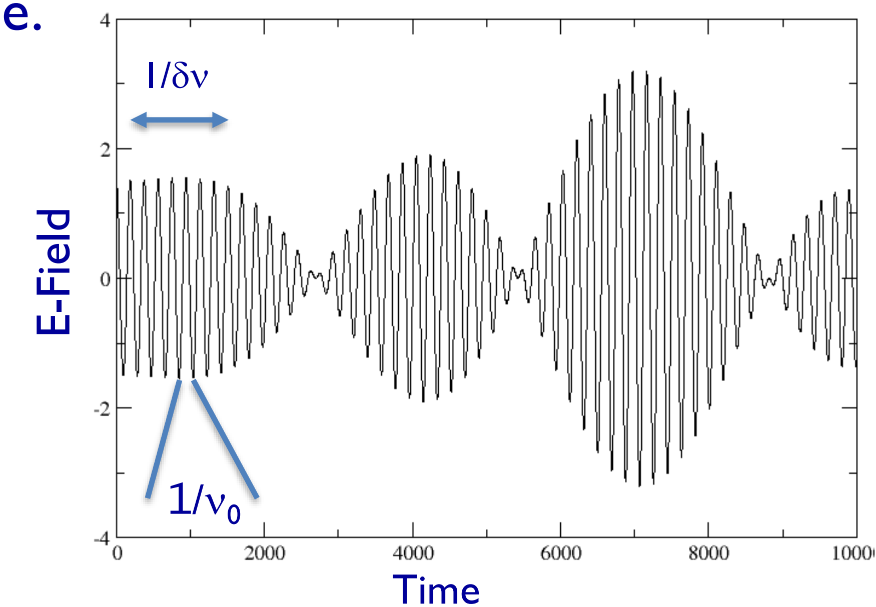
- There is another way:
- Replace the parabola with N little antennas, whose individual signals are summed in a network.
- The 'virtual focus' sum is equivalent to the free-space focus.
- The little antennas can be placed anywhere.



Quasi-Monochromatic Radiation

- Analysis is simplest if the fields are monochromatic.
- Natural radiation is never monochromatic. (Indeed, in principle, perfect monochromaticity cannot exist).
- 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.

The figure shows an 'oscilloscope' trace of a narrow bandwidth noise signal. The period of the wave is $T=1/\nu_0$, the duration over which the signal is closely sinusoidal is $T\sim 1/\delta\nu$. There are $N \sim \nu_0/\delta\nu$ oscillations in a 'wave packet'.



Representing the Electric Field

- Consider then the electric field from a small solid angle $d\Omega$ about some direction **s**, within some small bandwidth $d\nu$, at frequency ν.

- We can write the temporal dependence of this field as:

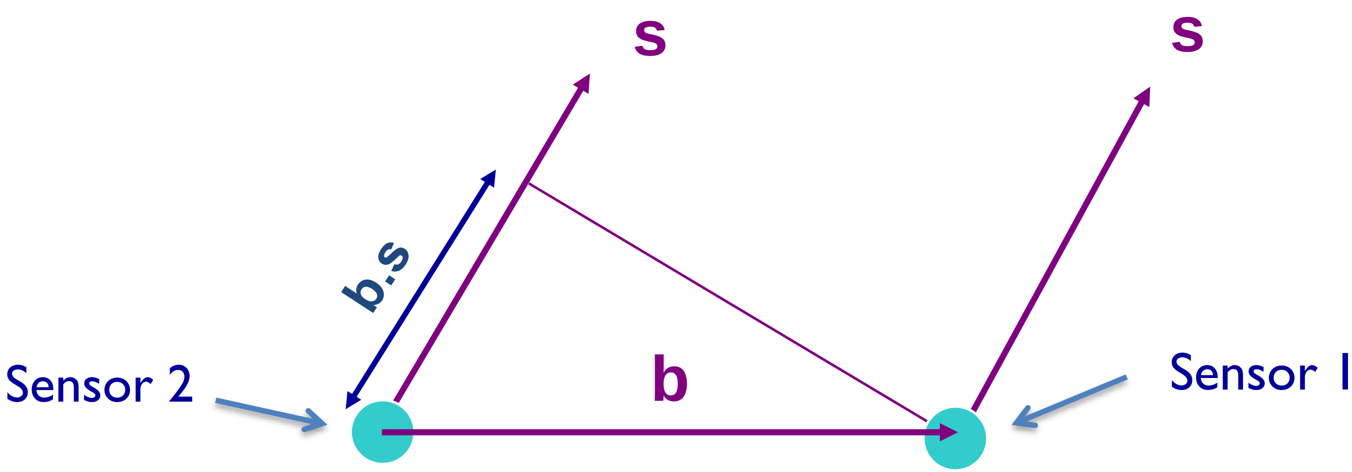
$$E_{\upsilon}(t) = A\cos(2\pi\upsilon t + \phi)$$

- The amplitude and phase remains unchanged for a time duration of order $dt \sim 1/d\nu$, 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 …)
  - Source in the far field (plane wave)
  - Idealized electronics (perfectly linear, no amplitude or phase perturbations, perfectly identical for both elements, no added noise, …)

Defining Basic Quantities



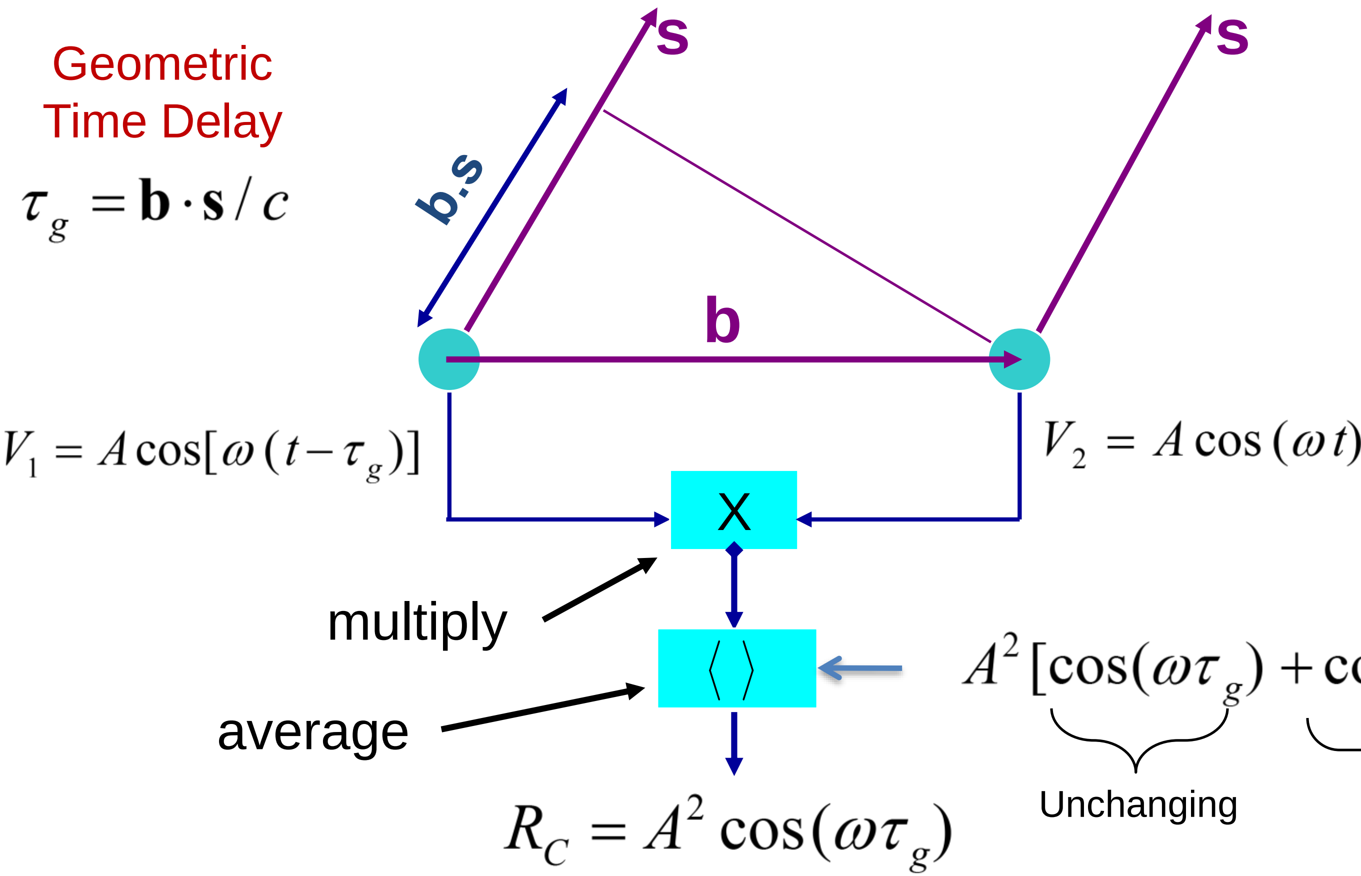
- There are two sensors, separated by vector baseline **b**
- Radiation arrives from direction **s** – assumed the same for both (far-field).
- The extra propagation path is L = b.s
- The time taken along this path is $\tau_g = \mathbf{b} \cdot \mathbf{s} / c$ Geometric delay
- For radiation of wavelength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

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

The Stationary, Quasi-Monochromatic Radio-Frequency Interferometer

Geometric Time Delay

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



The path lengths from sensors to multiplier are assumed equal.

$$A^2[\underbrace{\cos(\omega\tau_g)}_{\text{Unchanging}} + \underbrace{\cos(2\omega t - \omega\tau_g)}_{\text{Rapidly varying,}}]$$

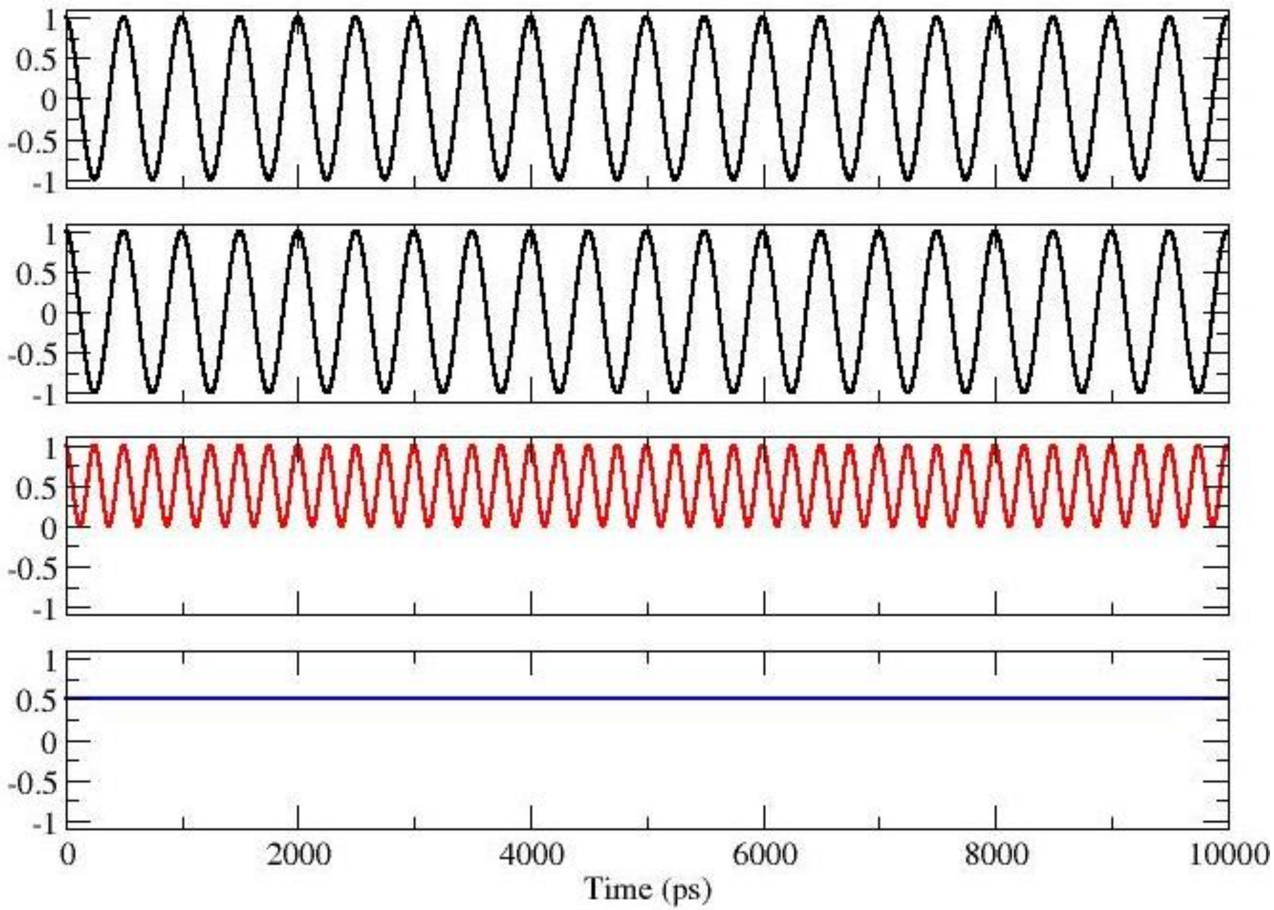
Note: R_c is not a function of time or location!

Pictorial Example: Signals In Phase

If the voltages arrive in phase:

b.s = $n\lambda$, or $\tau_g = n/\nu$ (n is an integer)

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

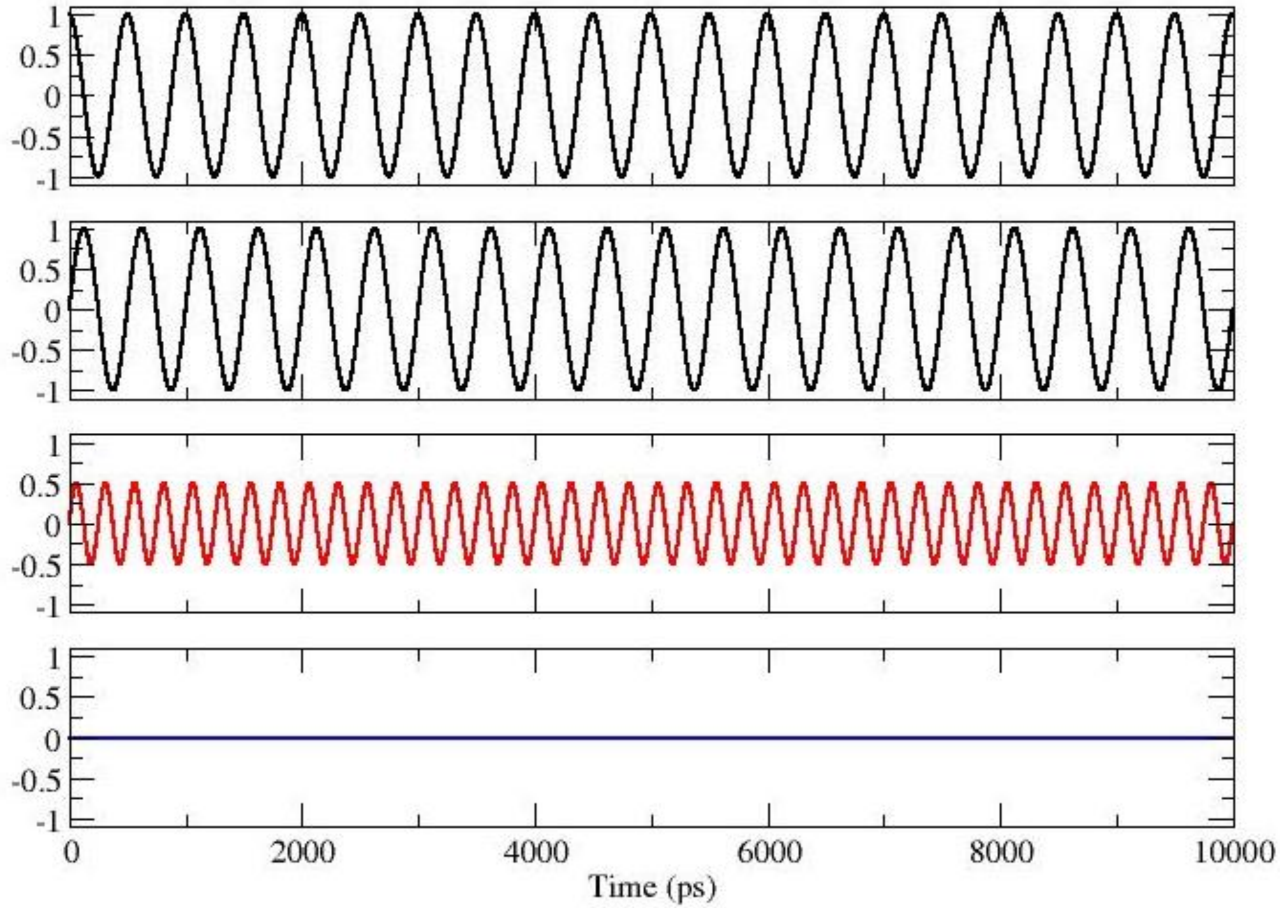


Pictorial Example: Signals in Quad Phase

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

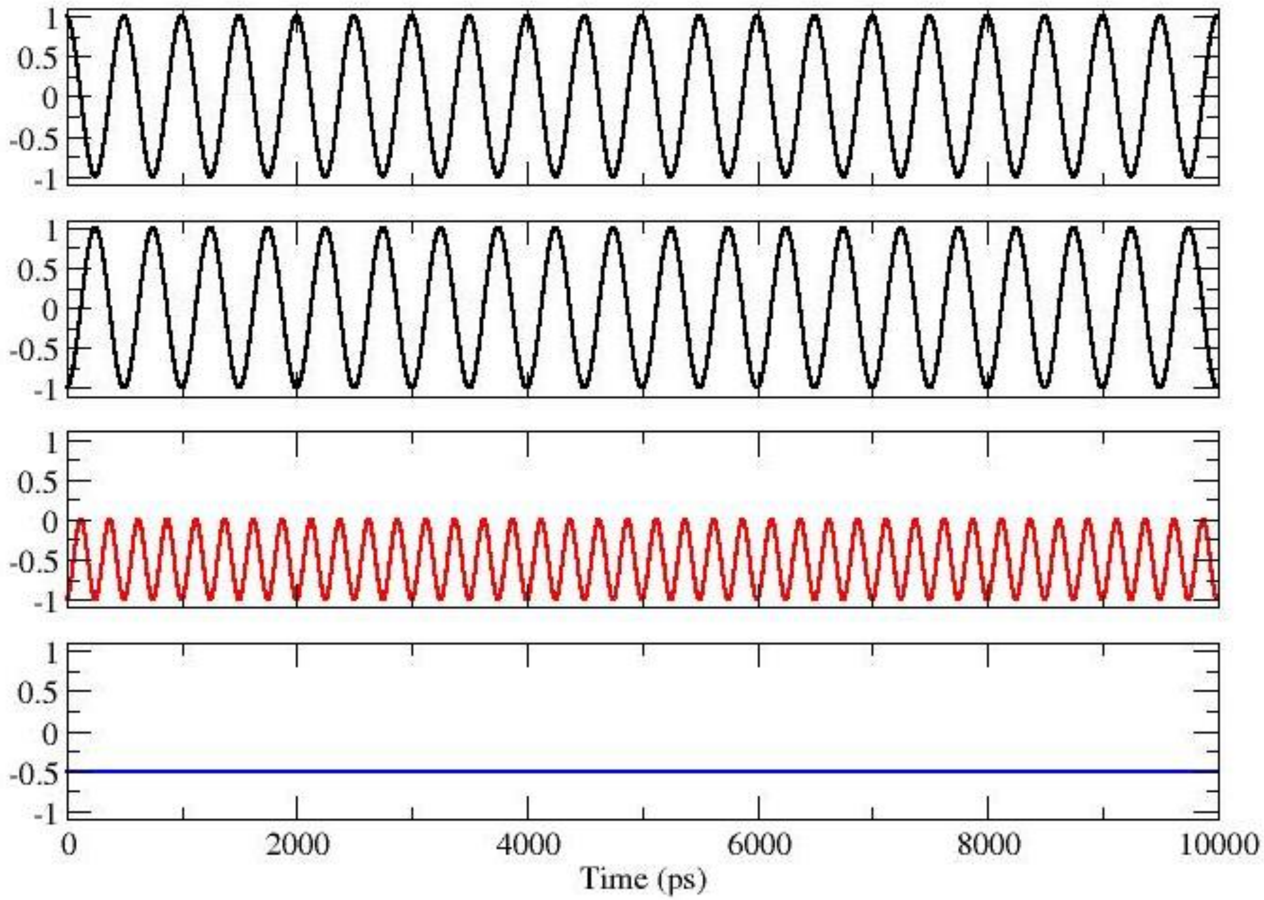


Pictorial Example: Signals out of Phase

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

- The averaged product R_C is dependent on the received power, $P = A^2/2$, frequency, 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 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.

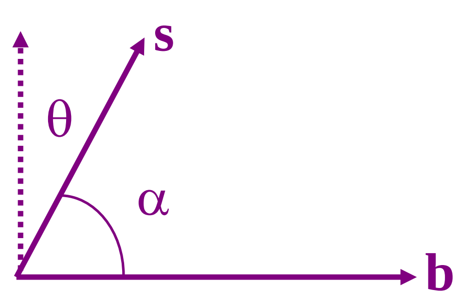
- The strength of the product is also dependent on the antenna collecting areas and electronic gains – but these factors can be calibrated for.

Nomenclature, and Direction Cosines

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

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

  - 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** for the direction **s**.

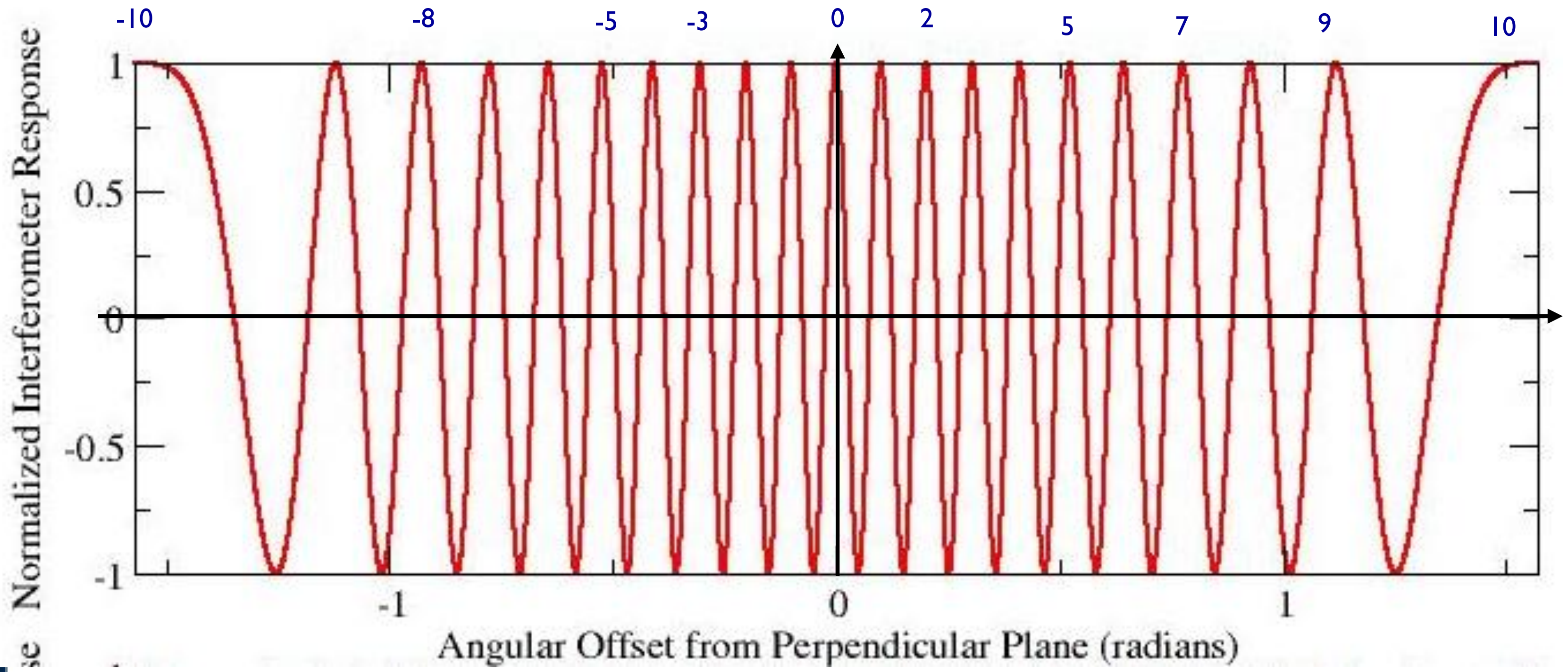


Whole-Sky Response for u = 10

- 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 radian.

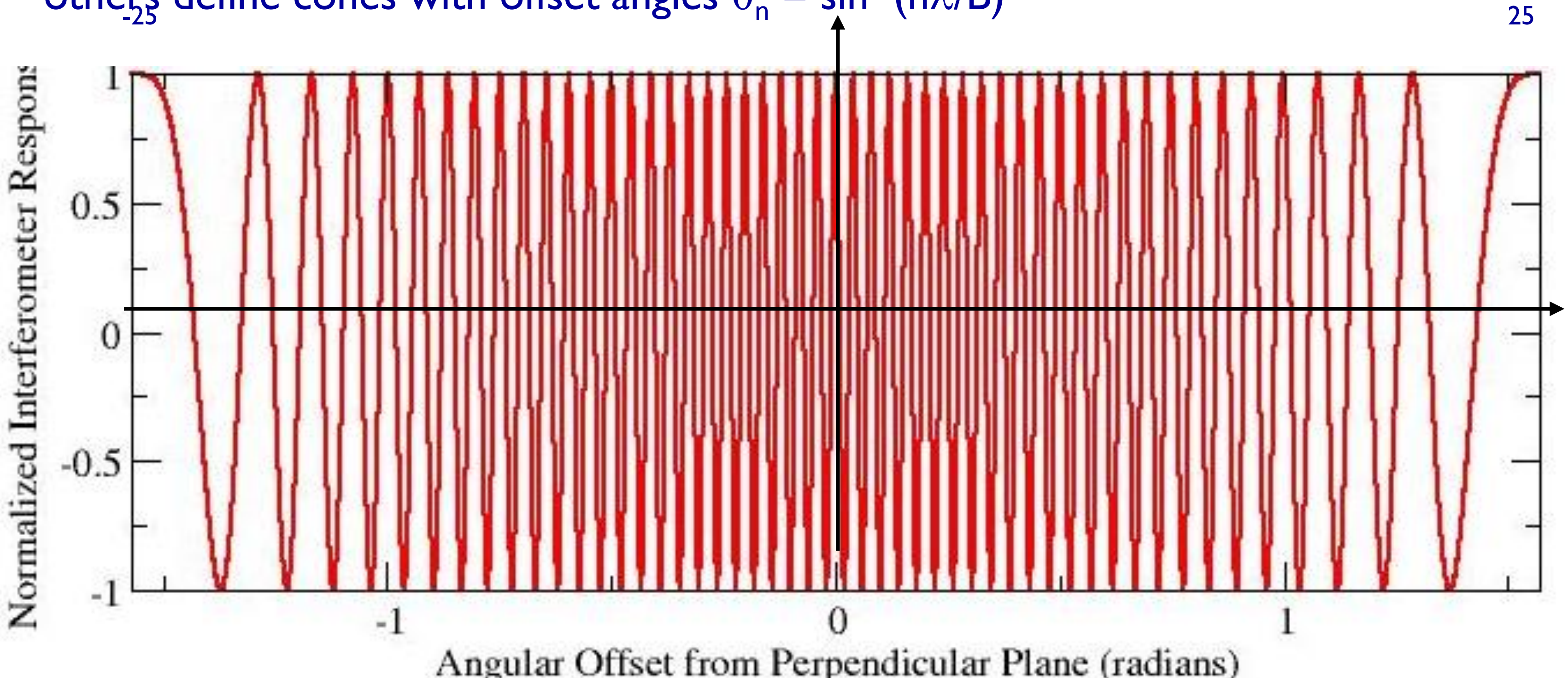


Whole-Sky Response for u = 25

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

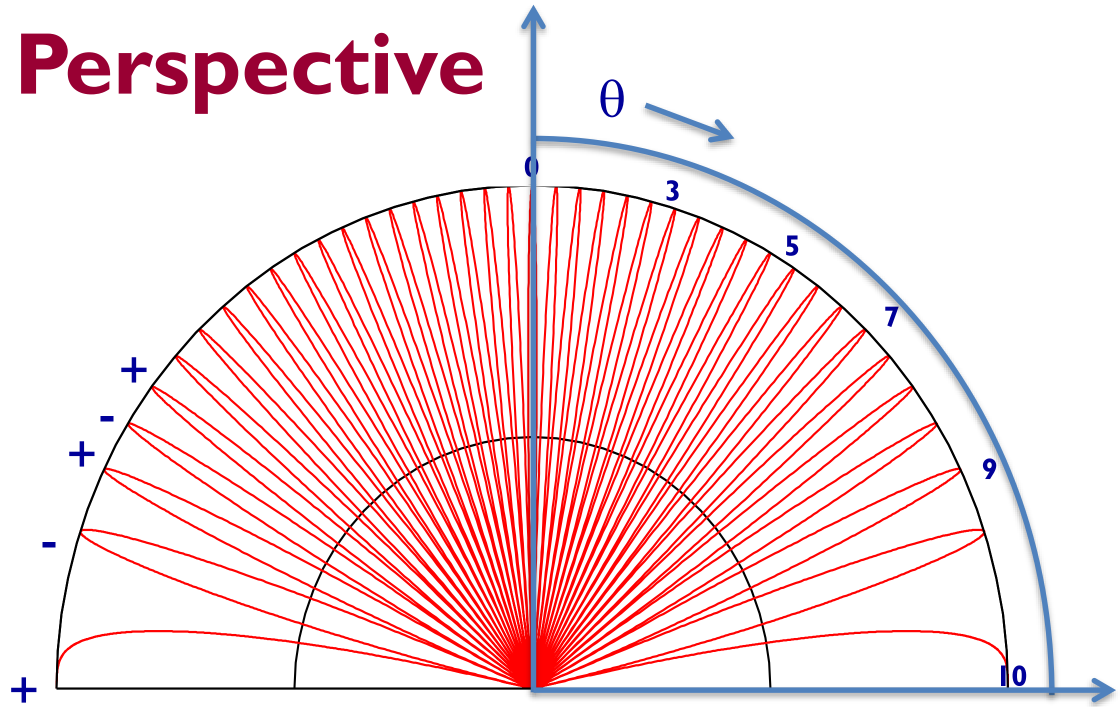
$$R_C = \cos(50\pi l)$$

- There are 51 whole fringes over the hemisphere.
- Minimum fringe separation 1/25 radians.
- The central fringe (n=0) defines a disk, perpendicular to the baseline. All the others define cones with offset angles $\theta_n = \sin^{-1}(n\lambda/B)$



NRAO

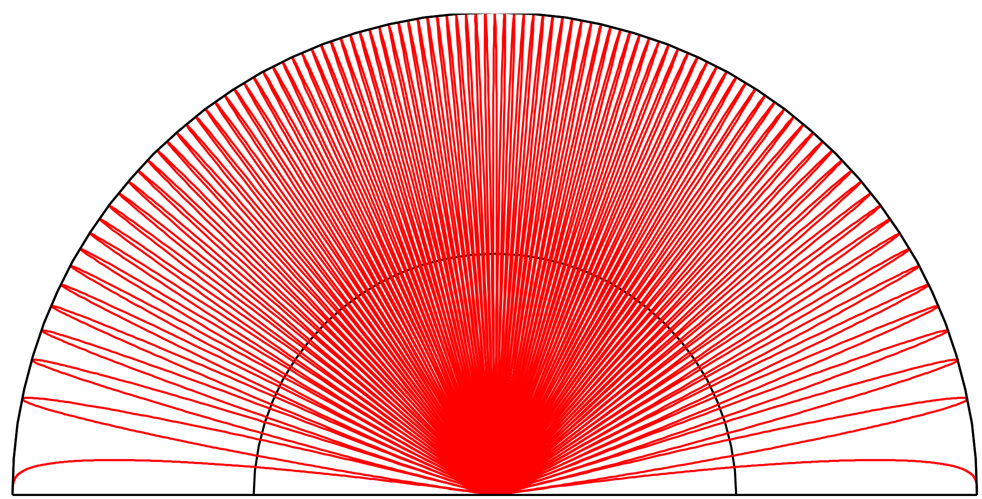
From an Angular Perspective



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

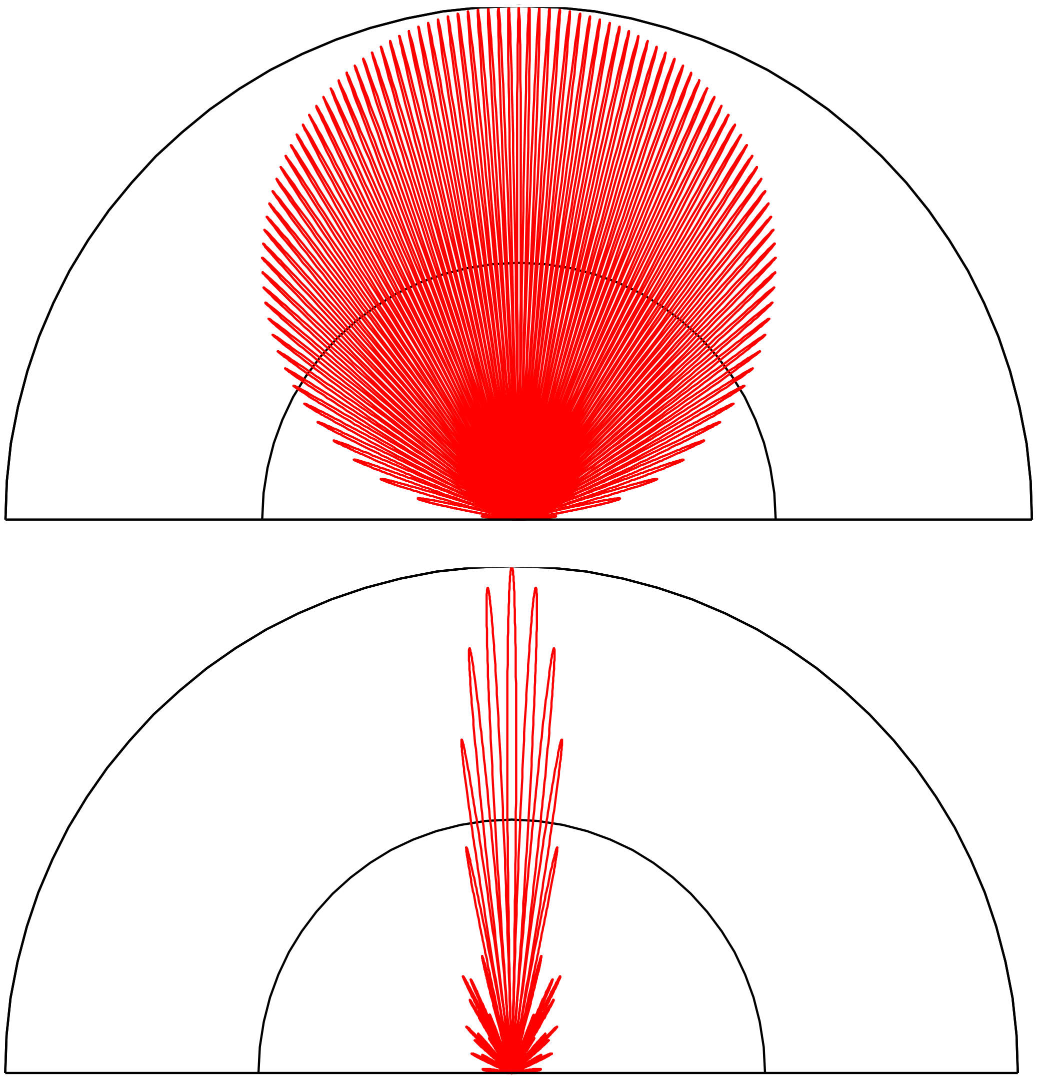
- The preceding plot is a meridional cut through the hemisphere, oriented along the baseline vector.
- In the two-dimensional space, the fringe pattern consists of a series of coaxial cones, oriented along the baseline vector.
- The figure is a two-dimensional representation when u = 4.
- As viewed along the baseline vector, the fringes show a 'bulls-eye' pattern – concentric circles.
- **Key Point: The Fringe Pattern is entirely defined by the baseline length and orientation.**

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 voltage.
- 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.
- **The baseline-defined fringe pattern is multiplied by the sensor gain pattern.**



The Response from an Extended Source

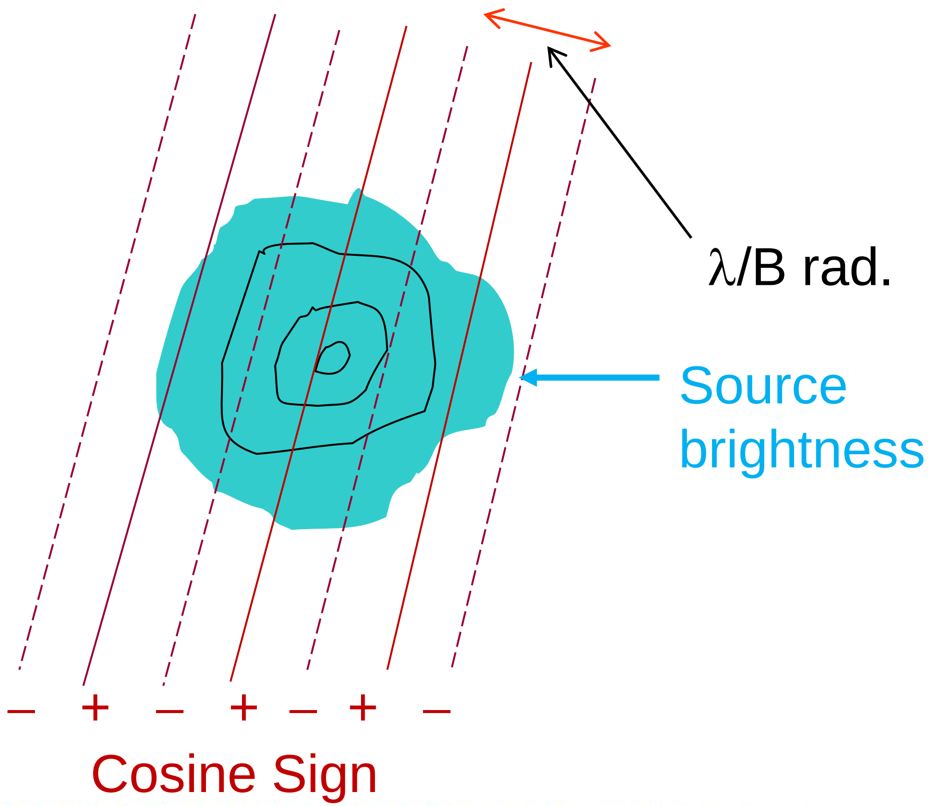
- The response from an extended source is obtained by summing the responses at each antenna to all the emission over the sky, multiplying the two, and averaging:
$$R_C = \left\langle \iint E_1 d\Omega_1 \times \iint E_2 d\Omega_2 \right\rangle$$
- The averaging and integrals can be interchanged and, **providing the emission is spatially incoherent**, we get (skipping lots of gory details)
$$R_C = \iint I_\nu(\mathbf{s}) \cos(2\pi\nu \mathbf{b}\cdot\mathbf{s}/c)\, d\Omega$$
- 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.

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.



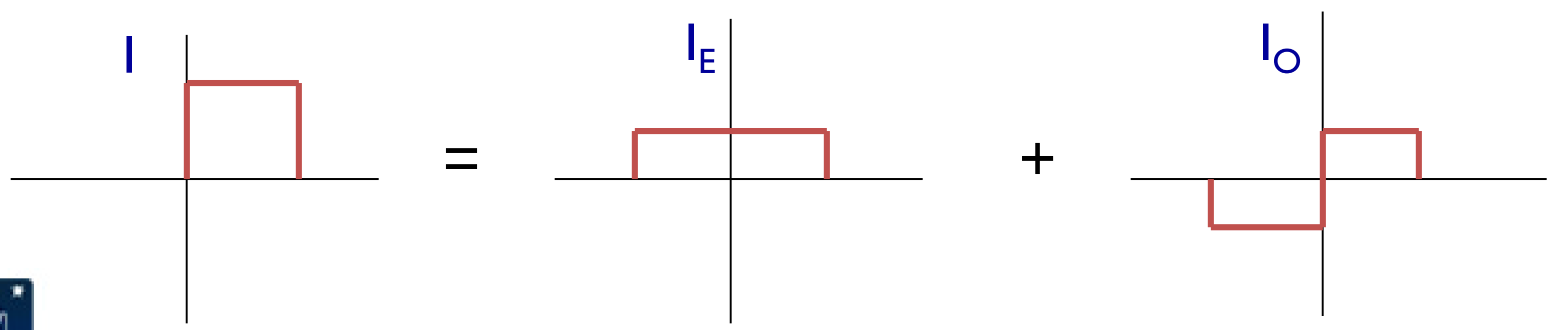
A Short Mathematics Digression – Odd and Even Functions

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



R_c is Sensitive to Even Structure Only!

- 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 response of the cosine interferometer to the odd brightness distribution is 0.

$$\iint I_O(\mathbf{s})\cos(2\pi\nu\mathbf{b}\cdot\mathbf{s}/c)\,d\Omega = 0$$

- **Only the 'even' part of the brightness is visible to the 'cosine' fringes.**

How to Recover the 'Odd' Emission?

- 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:

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

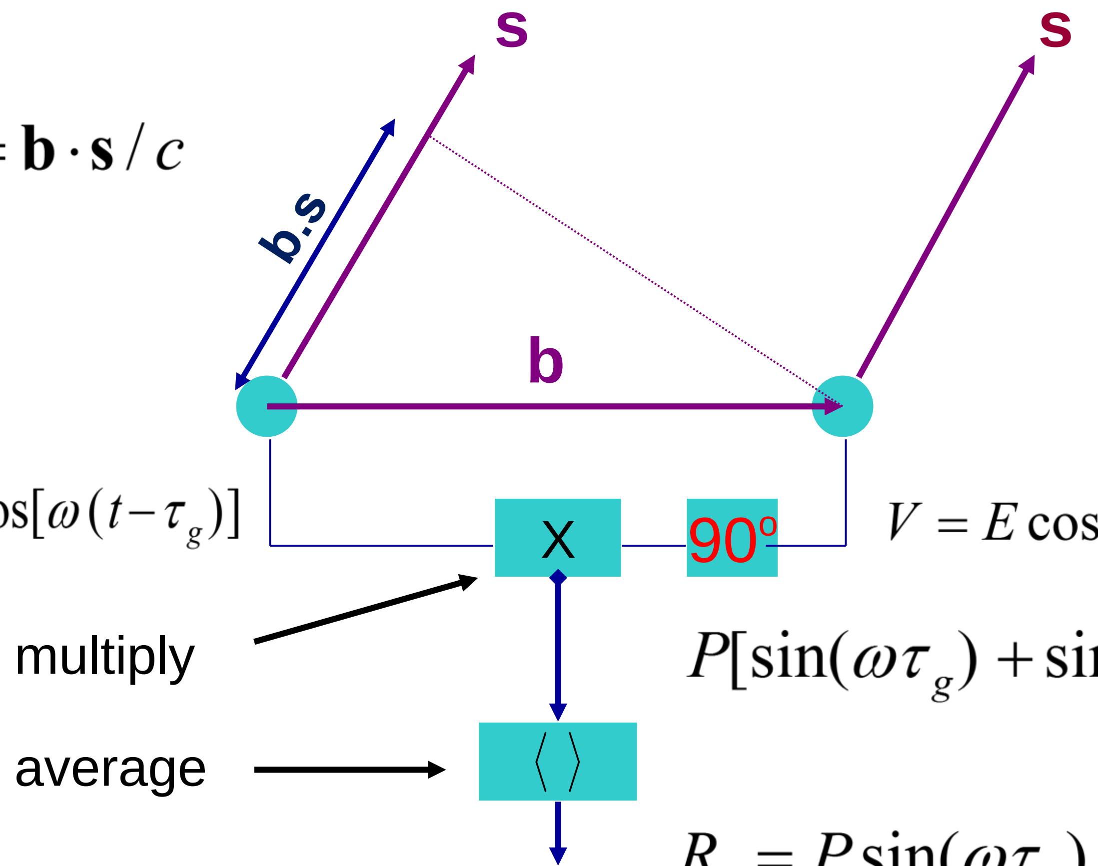
since the integral of an even times an odd function is zero.

- 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.

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



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

$$V = E\cos(\omega t)$$

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

$$R_s = P\sin(\omega\tau_g)$$

NRAO

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:

$$\mathcal{V} = R_C - iR_S = Ae^{-i\phi}$$

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:

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

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

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 amplitudes 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'.

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. Replace the real voltages with their complex analogs (the 'analytic signal'):

$$V_1 = Ae^{-i\omega t}$$

$$V_2 = Ae^{-i\omega(t - b \bullet s/c)}$$

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

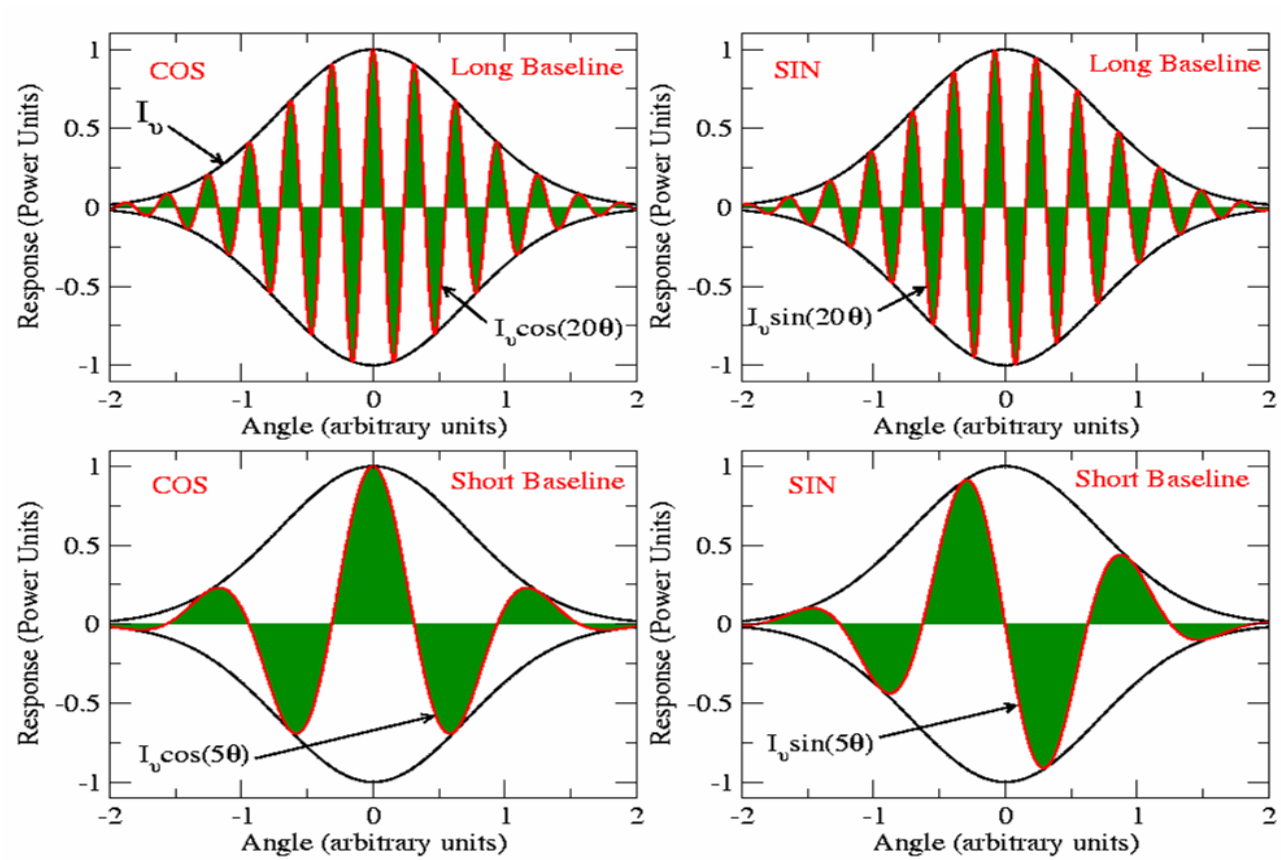
Picturing the Visibility

- 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.

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:

$$V(u) = e^{-2\pi i u l_0} = \cos(2\pi u l_0) - i \sin(2\pi u l_0)$$

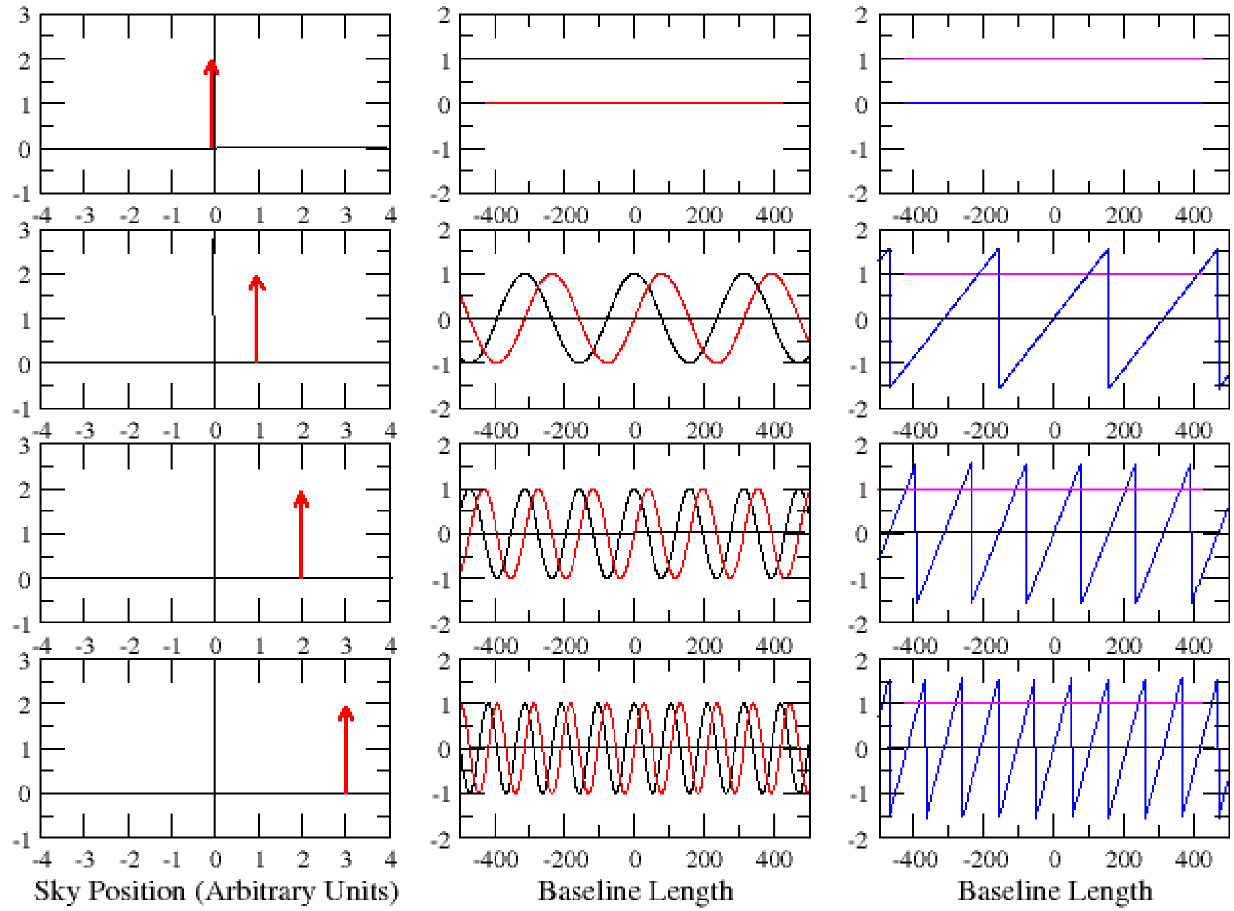
- For a source at the origin ($l_0 = 0$), $V(u) = 1$. (units of Jy).
- For a source off the origin, the visibility has unit amplitude, and a phase slope with increasing baseline, rotating 360 degrees every l_0^{-1} wavelengths.

Visibility Example #1: Point Sources

- Consider a Point Source (Red Arrow, left column), offset by 0, 1, 2, and 3 units from the phase center. The middle column shows the **Real** and **Imaginary** parts, the right column shows the **amplitude** and **phase**.

For all positions, the Amplitude is the same. The position offset information is in the phase slope.

A point source is not resolved – hence the amplitude remains constant for all baseline lengths.

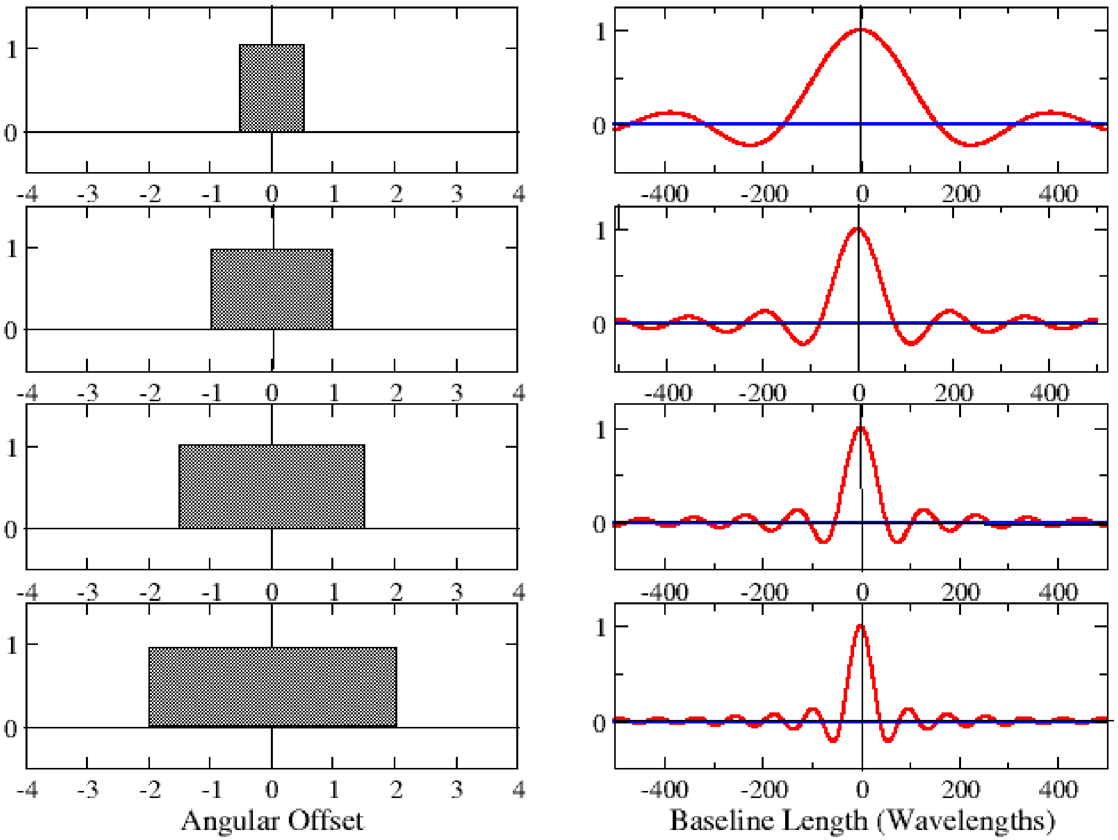


Visibility Example #2: Centered Boxes

- In this example, we have three centered (symmetric) 'box' structures. The symmetric structure ensure the imaginary component of the visibility is always zero – hence the visibility phase is zero.

The absence of a phase slope tells us the structure is centered.

The increasing size of the structure is reflected in the more rapid decrease of visibility amplitude with baseline length.

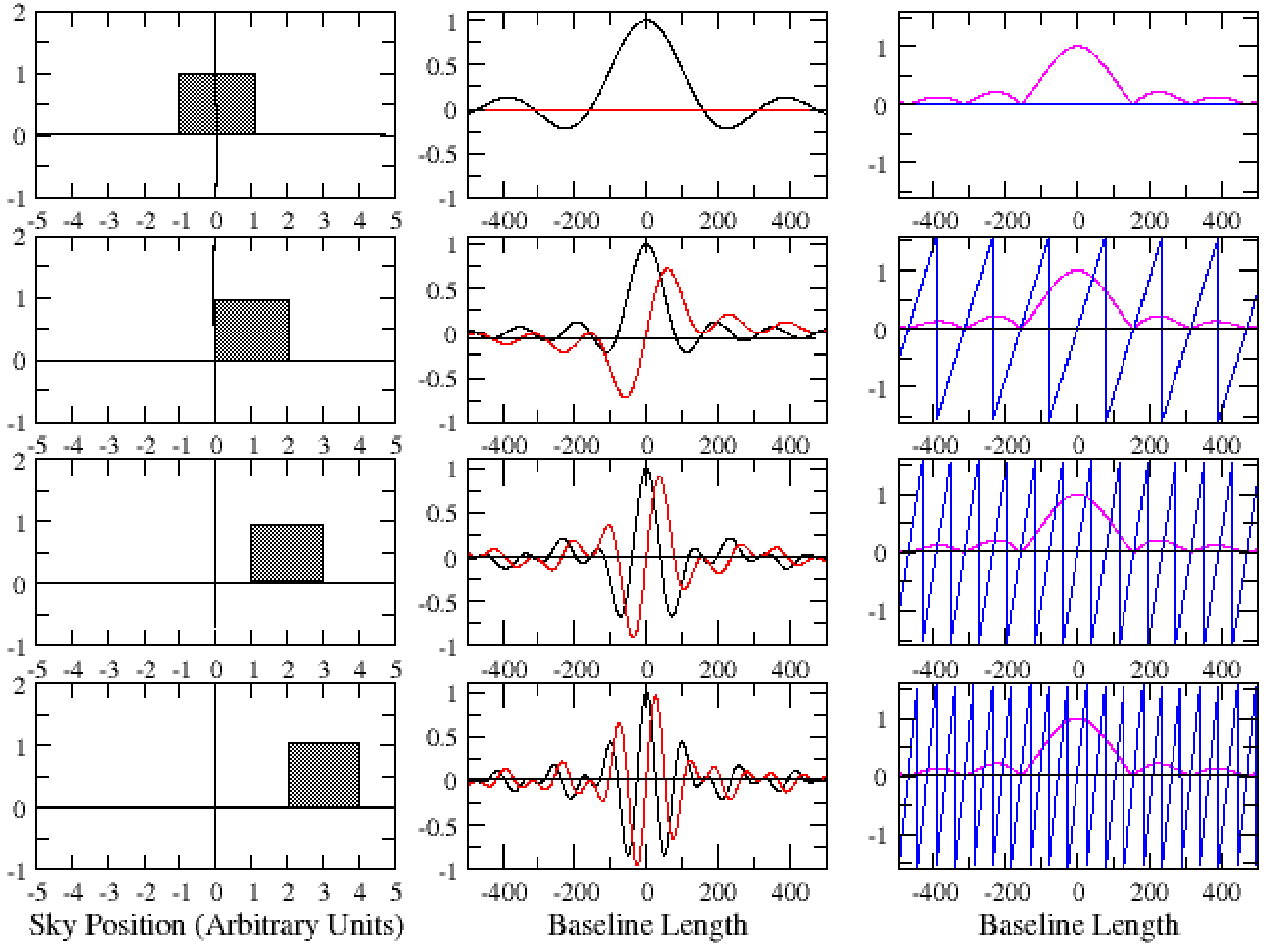


Visibility Example #3: Offset Boxes

- In this example, we show the same box, offset by increasing amounts from the phase center. Middle panel show **Real** and **Imaginary** components, Right panel shows **Amplitude** and **Phase** components.

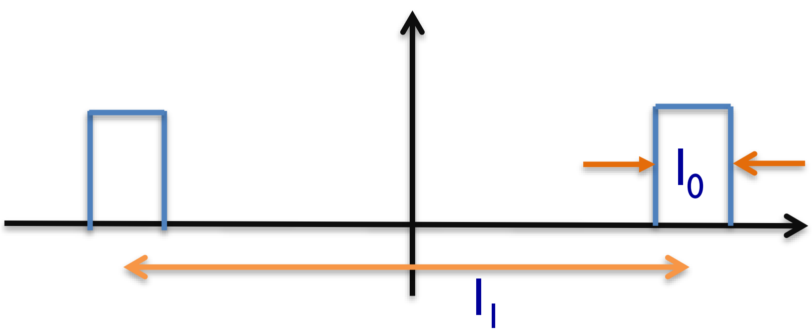
The visibility amplitude is the same for all offsets – the box is the same width.

The phase slope increases linearly with increasing offset.



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: $V(u) = \text{sinc}(u\, l_0)\cos(\pi\, u l_1)$

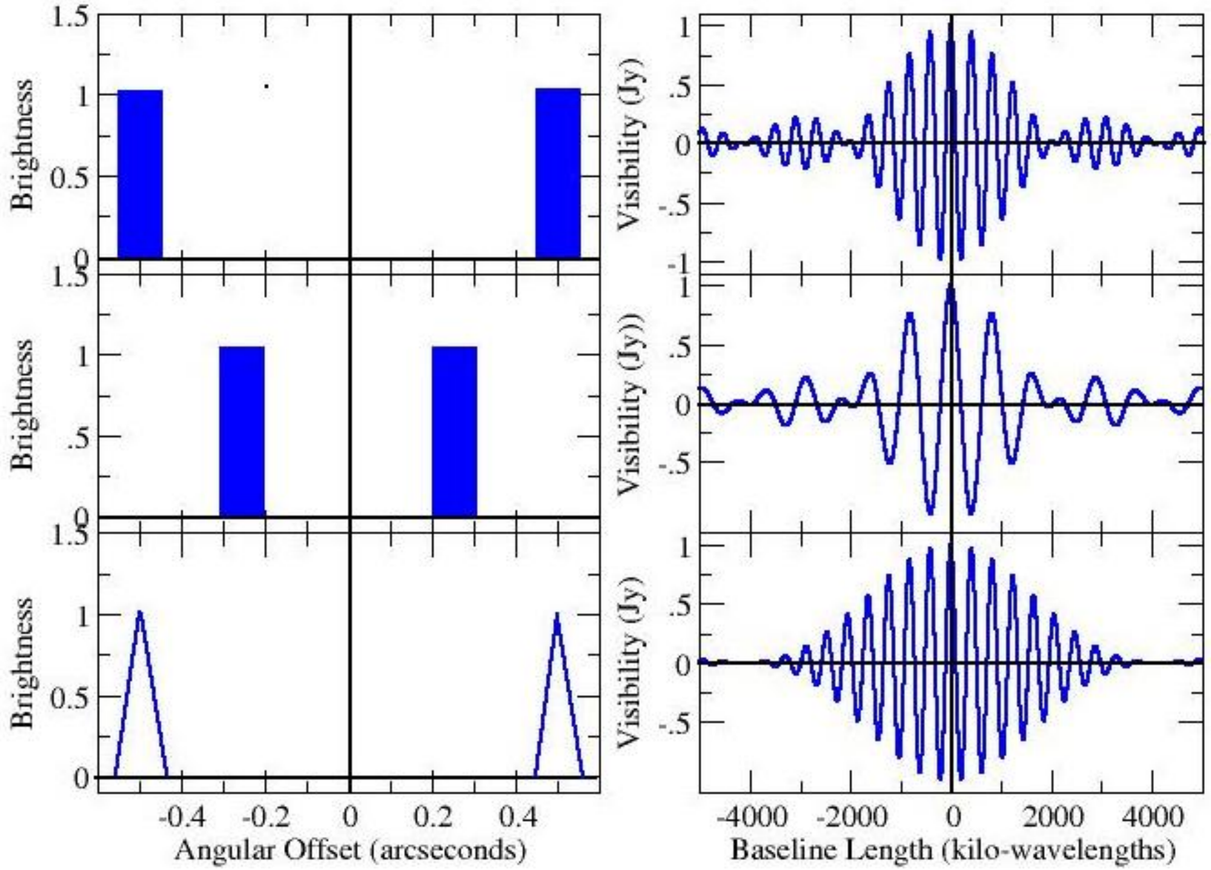
Which is an oscillatory function of period $u = 1/l_1$ attenuated by a dying oscillation of period $u = 1/l_0$.

More Examples

- Simple pictures illustrating 1-dimensional visibilities.

Brightness Distribution | Visibility Function

- Resolved Double
- Resolved Double
- Central Peaked Double



Examples with Real Data!

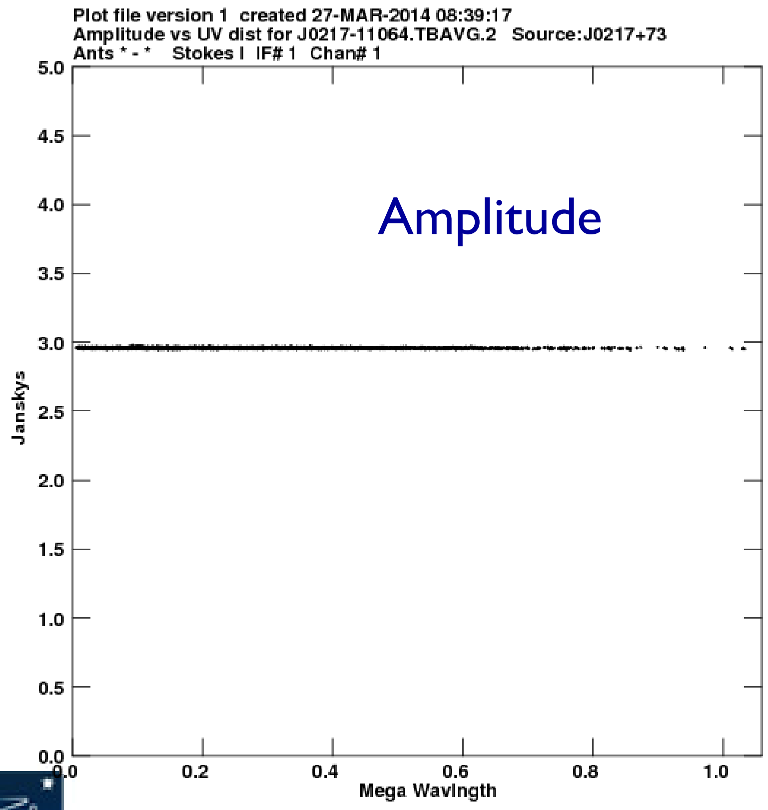
- Enough of the analysis!
- I close with some examples from real observations, using the VLA.
- These are two-dimensional observations (function of 'u' (EW) and 'v' (NS) baselines).
- Plotted are the visibility amplitudes version baseline length:

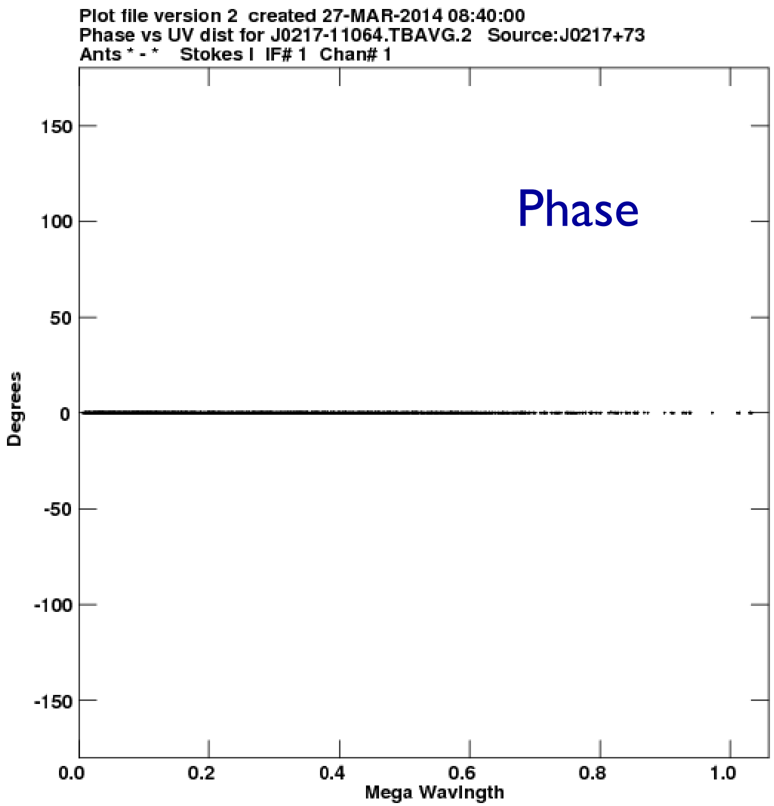
$$q = \sqrt{u^2 + v^2}$$

- Plotting visibilities in this way is easy, and often gives much information into source structure – as well as a diagnosis of various errors.

Examples of Visbilities – A Point Source

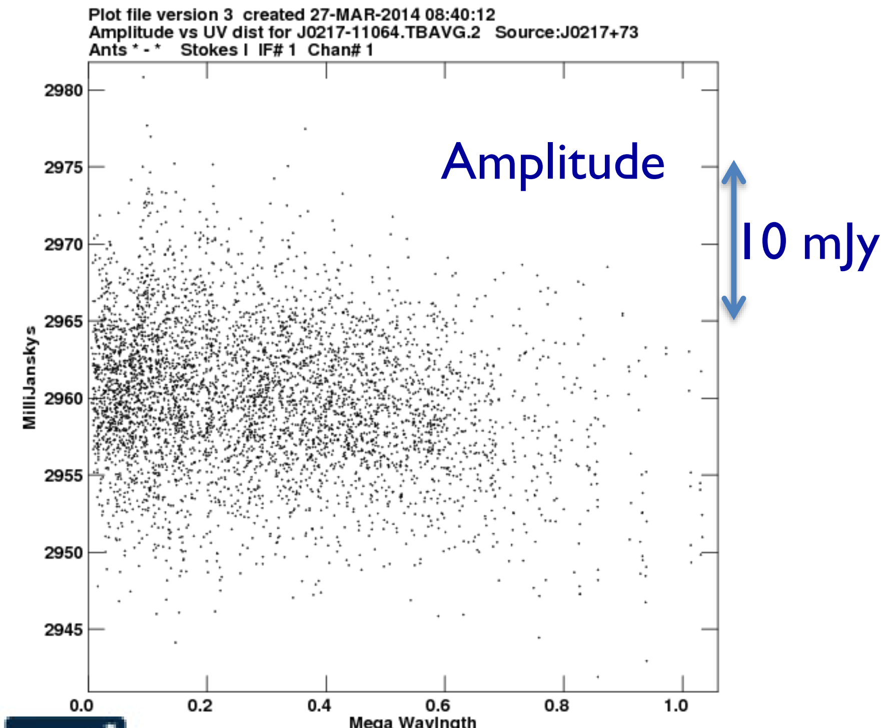
- Suppose we observe an unresolved object, at the phase center.
- What is its visibility function?

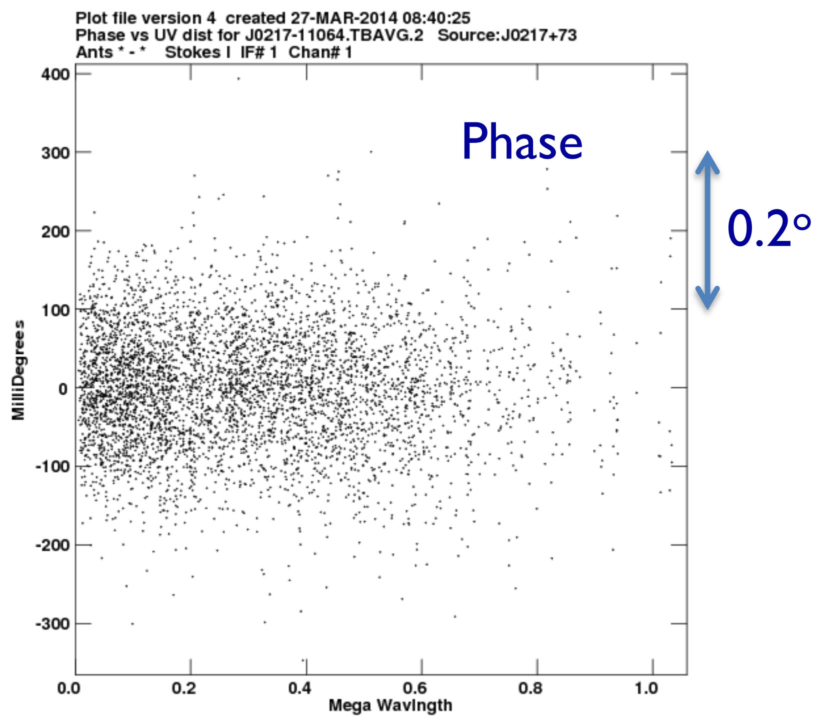




Zoom in …

- The previous plots showed consistent values for all baselines.
- Zooming in shows the noise (and, possibly, additional structure).



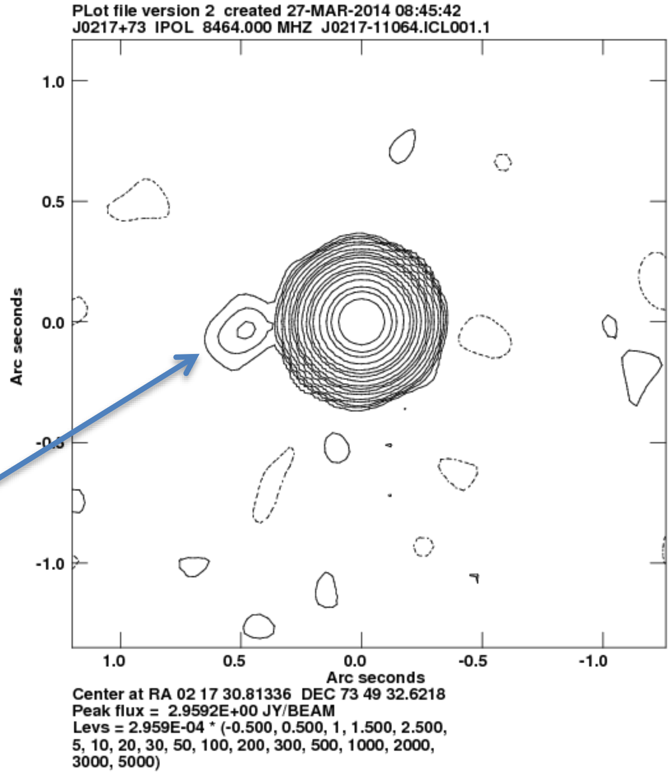


And the Map …

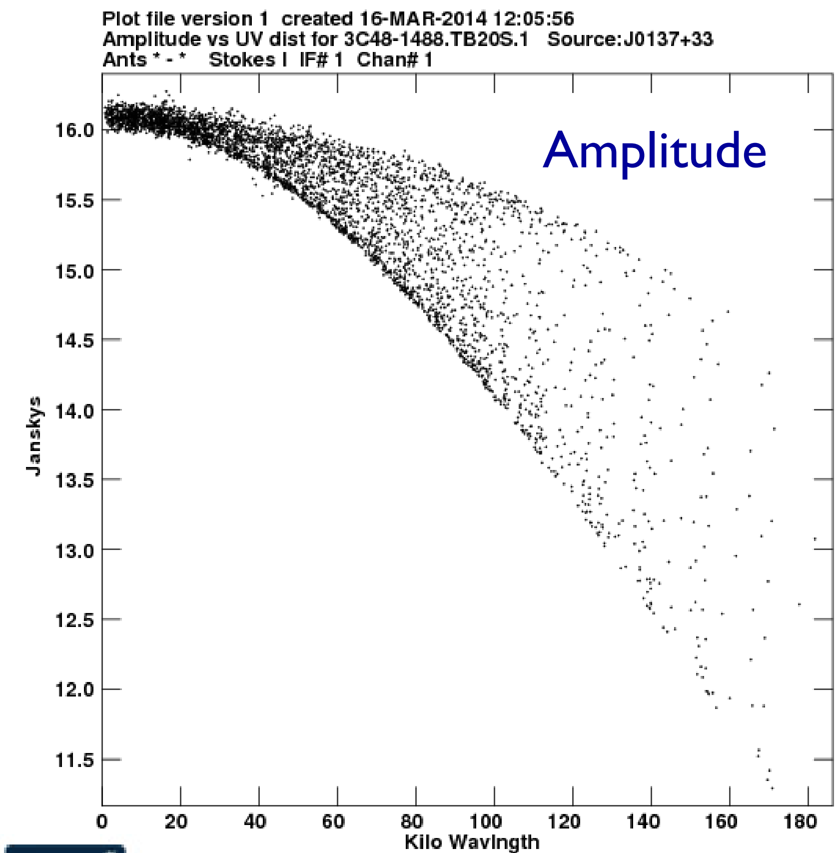
- The source is unresolved … but with a tiny background object.
- Dynamic range: 50,000:1.

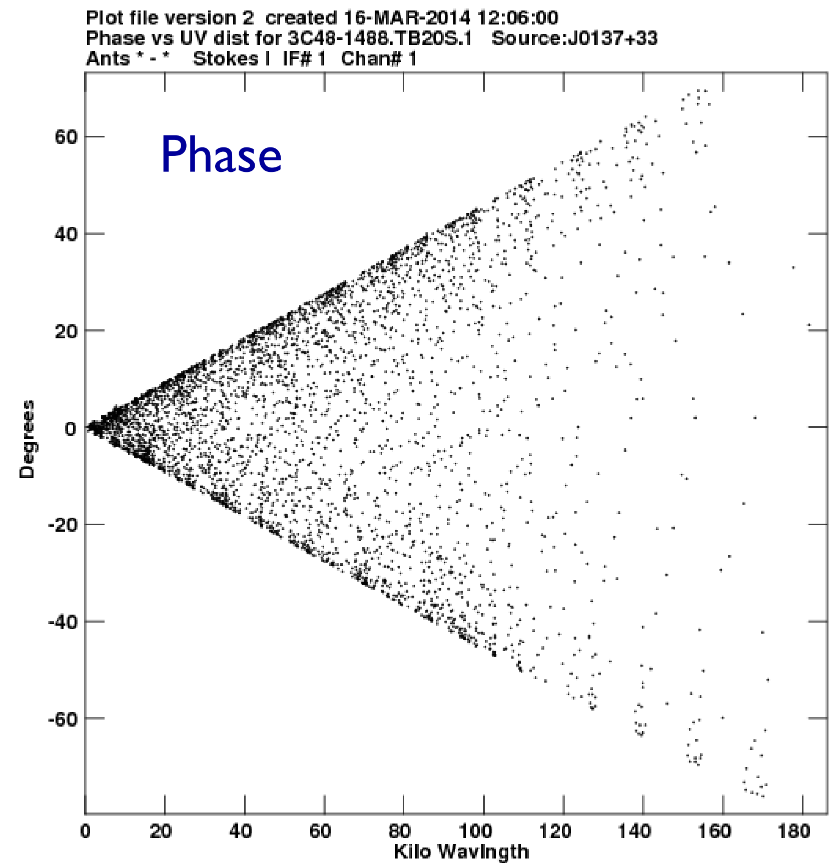
The flux in the weak nearby object is only 0.25 mJy – too low to be seen on any individual visibility.

Real!



3C48 at 21 cm wavelength – a slightly resolved object.

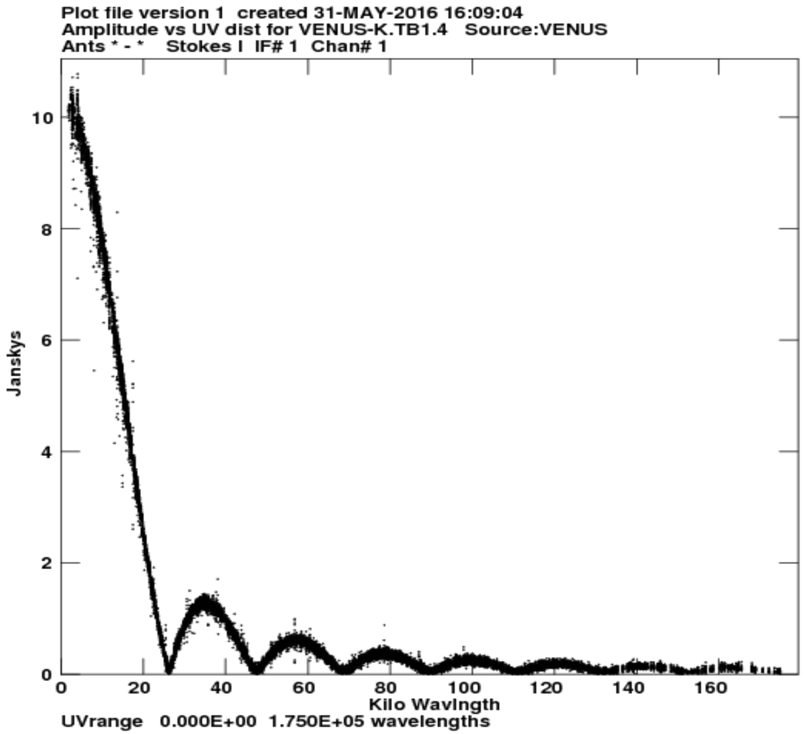


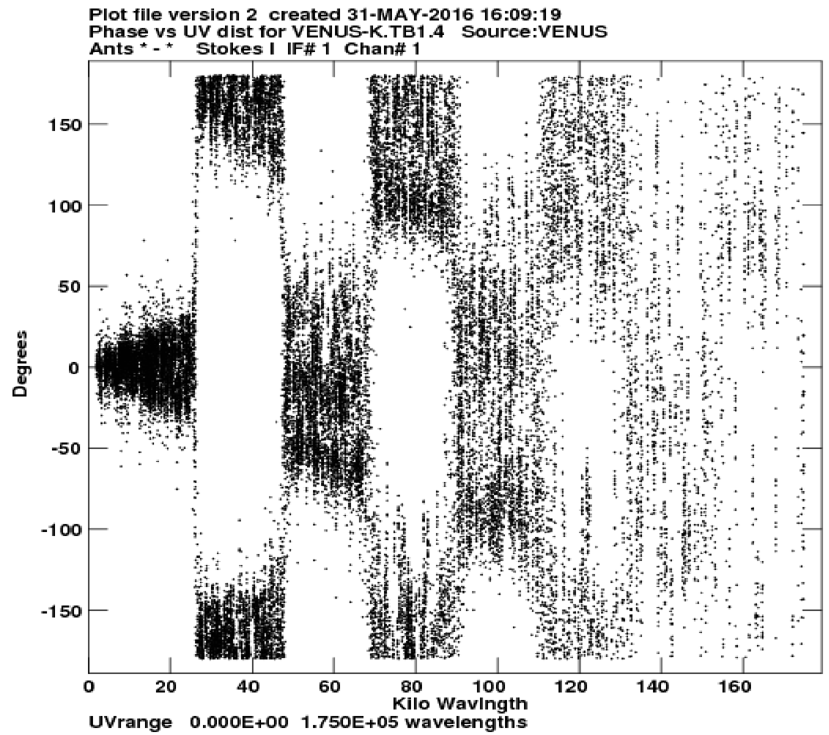


3C48 position and offset

- The linear phase slope is 90 degrees over 250,000 wavelengths.
- This is the phase slope corresponding to ¼ arcsecond – 250 milliarcsecond offset of the emission centroid from the phase center.
- The amplitudes show slight (25%) resolution at 180,000 wavelengths – consistent with extension on the sub-arcsecond scale.

The Planet Venus at 19 GHz.

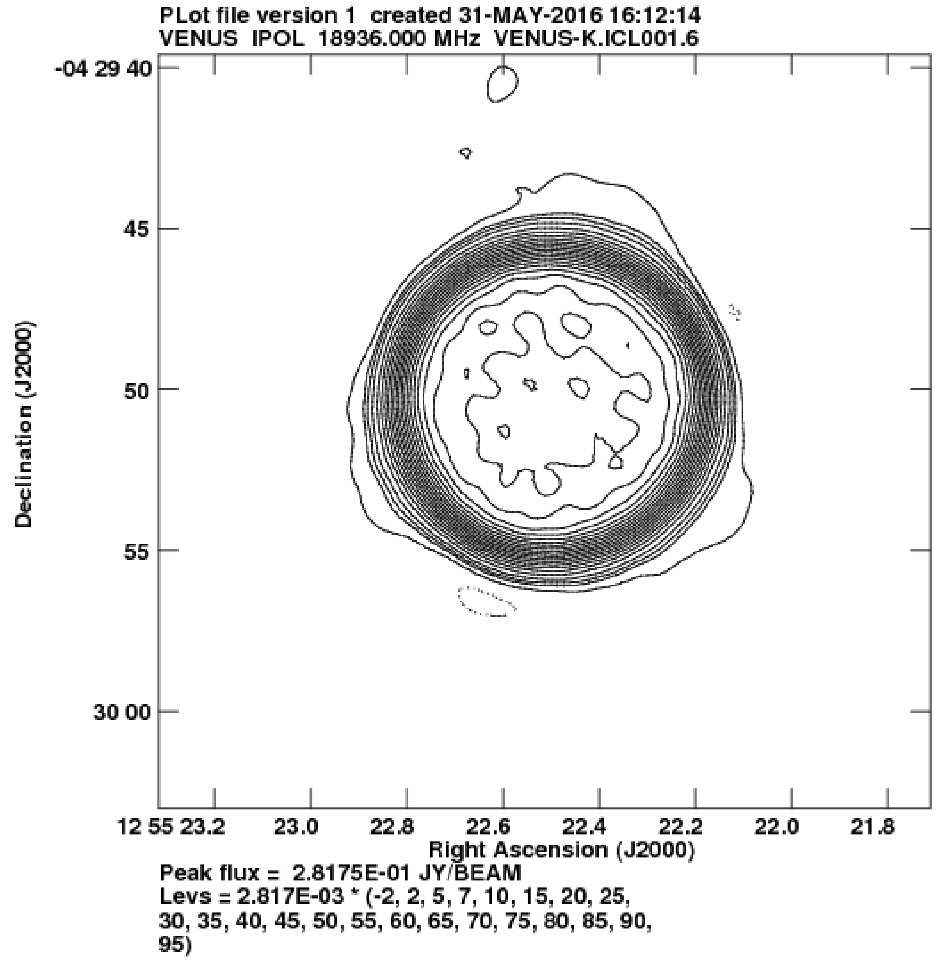




- The visibilities are circularly symmetric. The phases alternates between zero and 180 degrees. The source must be circularly symmetric.
- The visibility null at 25 kλ indicates angular size of ~ 10 arcseconds.

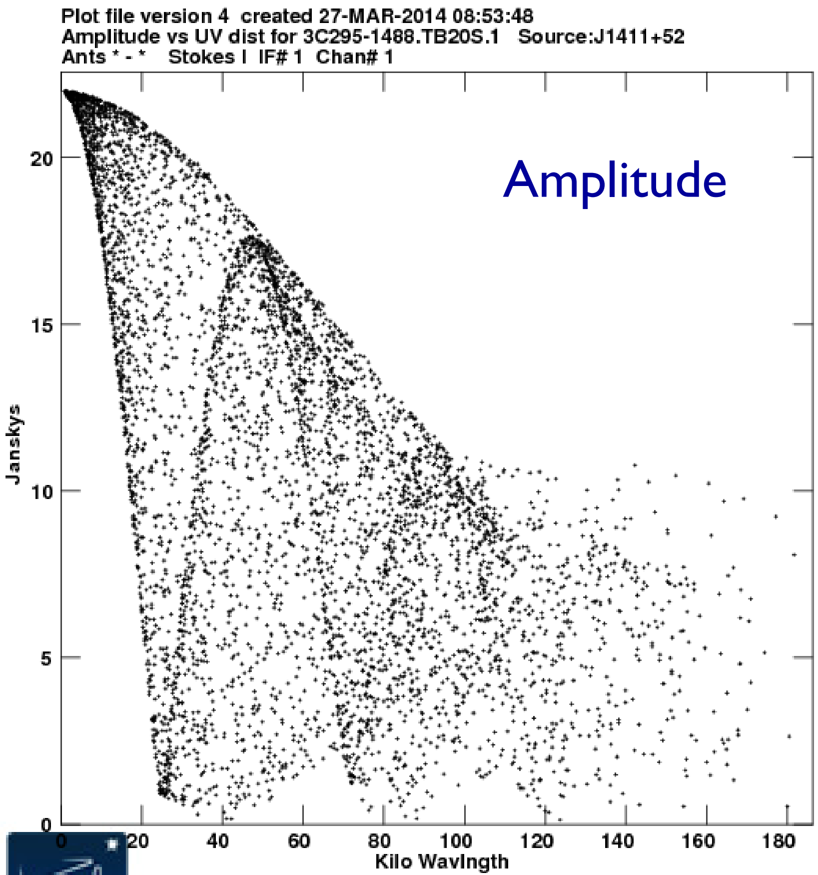
And the image looks like:

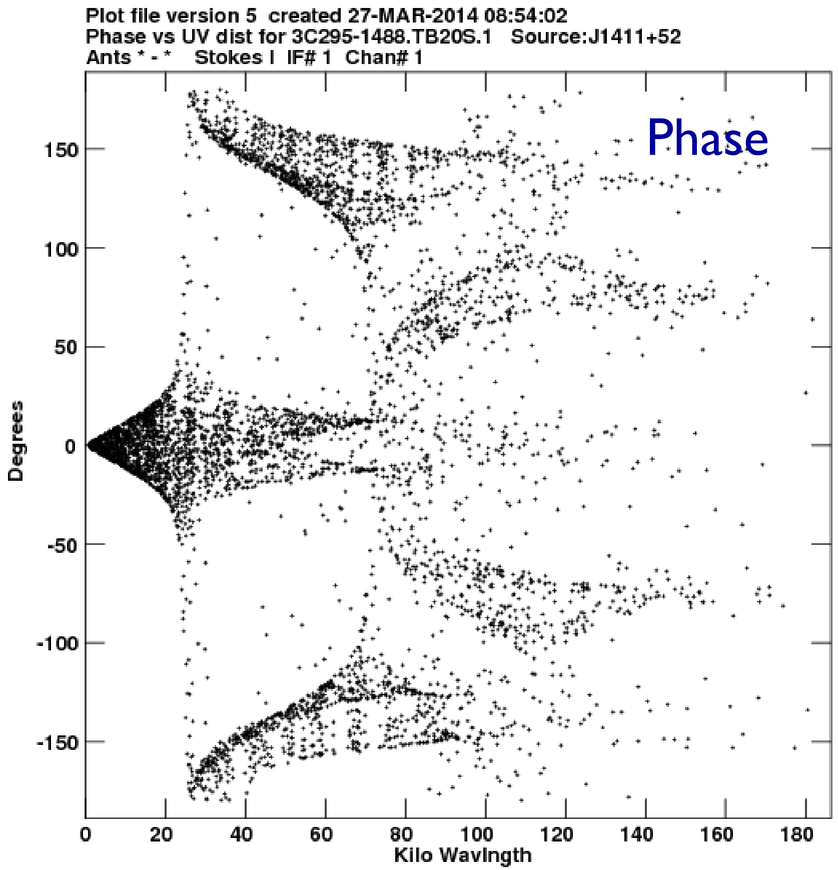
- It's a perfectly uniform, blank disk!
- The Visibility function, in fact, is an almost perfect Bessel function of zero order: $J_0(q)$.
- A perfect J_0 would arise from a perfectly sharp disk. Atmospheric opacity effects 'soften' the edge, resulting in small deviations from the J_0 function at large baselines.



Examples of Visibilities – a Well Resolved Object

- The flux calibrator 3C295





3C295 Image

- The visibility amplitude cycles on a 60,000 wavelength period – corresponding to about 4 arcseconds extent – as shown in the image.
- The phase is too complicated to easily interpret!