Fundamentals of Radio Interferometry

Rick Perley, NRAO/Socorro

Eighteenth (and Second Virtual) Synthesis Imaging Summer School

May 18 – 25, 2022

Topics

- **Introduction:**
  - Motivation: Mapping the Sky
  - Why Interferometry?
- **The Basic Interferometer**
  - Simplifying Assumptions
  - ‘Fringe’ patterns
  - Sine and Cosine Fringes
  - Response to Extended Emission
  - The Complex Correlator
  - Some Illustrations
- **Important Note:**
  - The concepts I’ll be speaking about are not difficult – but they are unfamiliar to beginners.

References

- Before launching into the material, here are five references I find most useful:

1. 'Principles of Optics': Max Born and Emil Wolf. My copy is the 6^{th} edition (1980). Pergamon Press.
2. 'Interferometry and Synthesis in Radio Astronomy.' Thompson, Moran and Swenson. 2^{nd} edition 2001. Wiley Interscience.
3. 'Synthesis Imaging in Radio Astronomy II'. Taylor, Carilli and Perley. ASP Conference Series Volume 180, 1999.
4. 'The Fourier Transform and is Applications'. Ron Bracewell. 2^{nd} edition. McGraw-Hill. (Avoid the 3^{rd} edition, which has errors).
5. Radio Astronomy. John Kraus. McGraw-Hill. (very out of date, but still useful).

Mapping (and Resolving) the Sky

- In astronomy, we wish to know the angular distribution of the emission.
  - This can be a function of frequency, polarization, and time.
- 'Angular Distribution' means we are interested in the **brightness** of the emission.
- Measuring the brightness means **resolving the emission.**
- Because our targets are so far away, the emission is extremely weak, and of very small angular size.
- Early (1950s) surveys of the radio sky employed single dishes.
- Nowadays, most (but not all!) observations are done with interferometers.
- In this lecture, we develop the relations between the sky brightness and interferometer response.

Why Interferometry?

- It's all about **Diffraction** – a consequence of the wave nature of light.
- Radio telescopes coherently sum electric fields over an aperture of size D. For this, diffraction theory applies – the angular resolution is:

$$\theta_{rad} \approx \lambda / D \qquad \text{Or, in practical units} \qquad \theta_{\mathrm{arcsec}} \approx 2\,\lambda_{\mathrm{cm}} / D_{\mathrm{km}}$$

- To obtain 1 arcsecond resolution at a wavelength of 21 cm, we require an aperture of ~42 km – not feasible.
- The (currently) largest single, fully-steerable apertures are 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 technique of synthesizing a larger aperture through combinations of separated pairs of antennas is called 'aperture synthesis'.

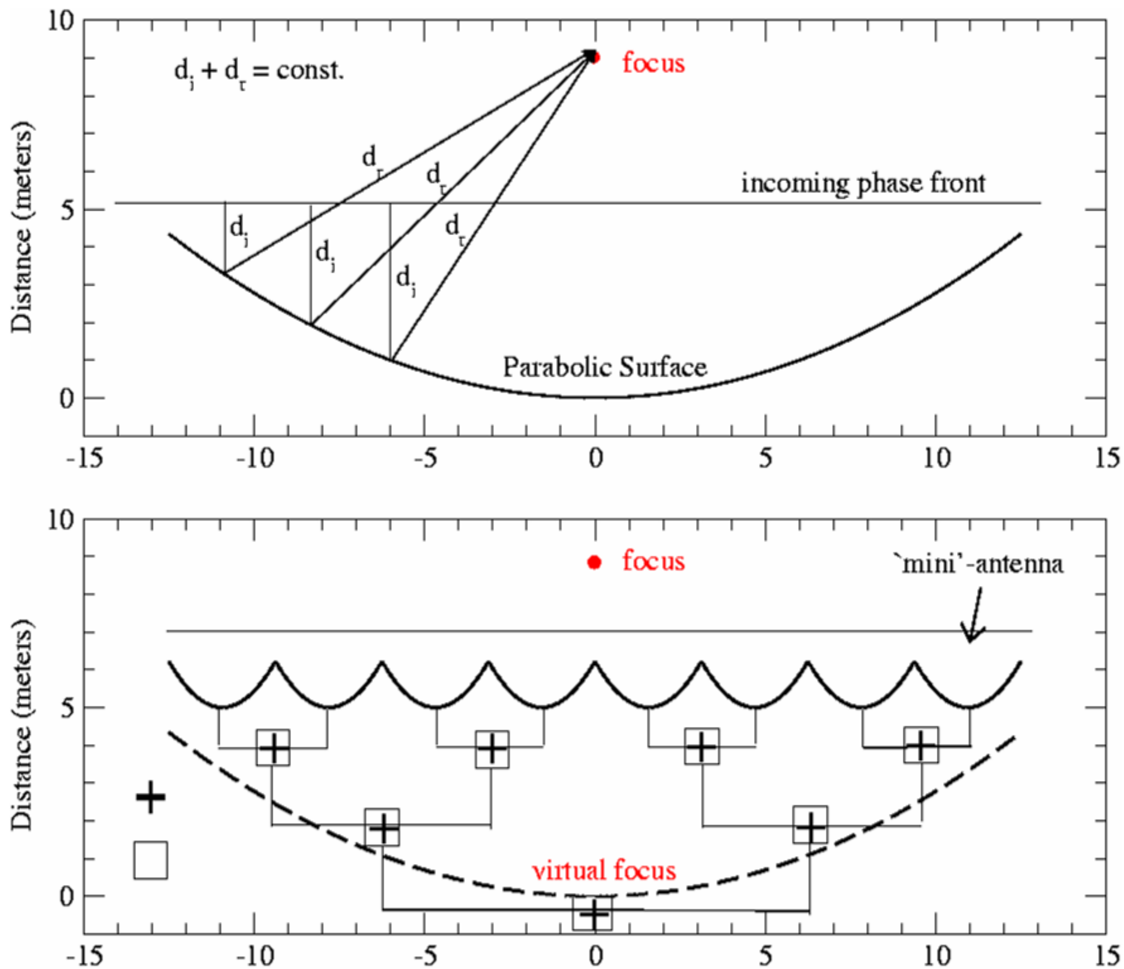
Interferometry – Basic Concept

- A parabolic dish coherently sums EM fields at the focus.

- The same result can be gotten by adding in a network of voltages from individual elements.
  - Note – they need not be adjacent.

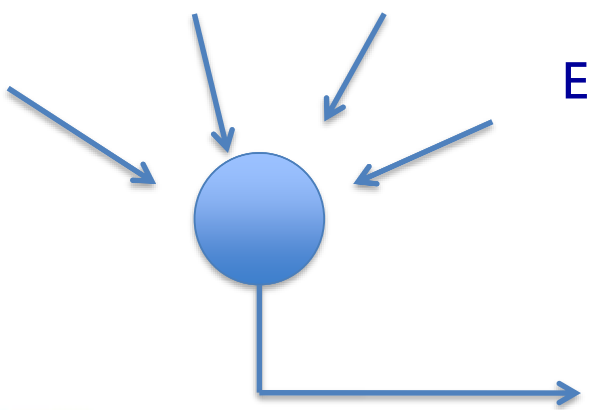
- This is the basic concept of interferometry.

- Aperture Synthesis is an extension of this concept.



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**,ν,t) at some place (**r**) 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

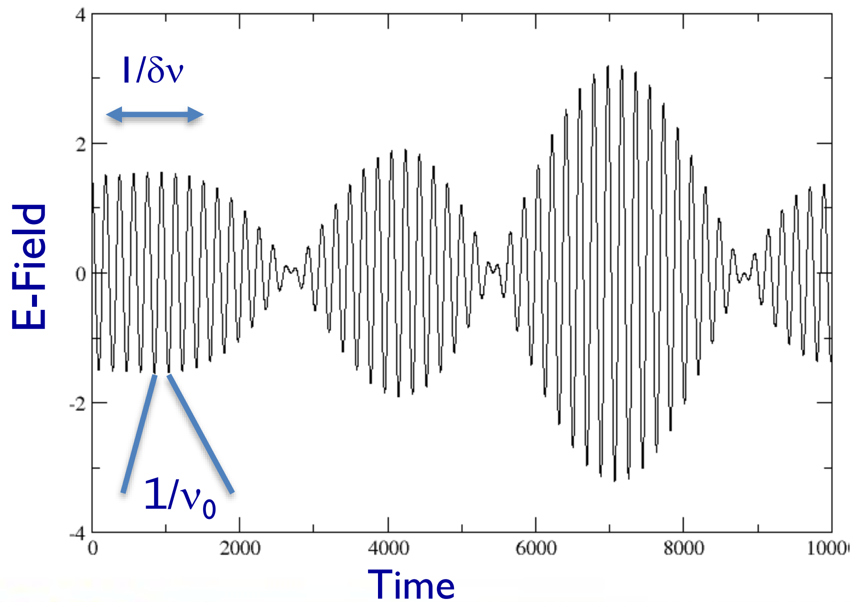
Voltage out (preserving amplitude and phase of all input fields)

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, described by

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

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'.



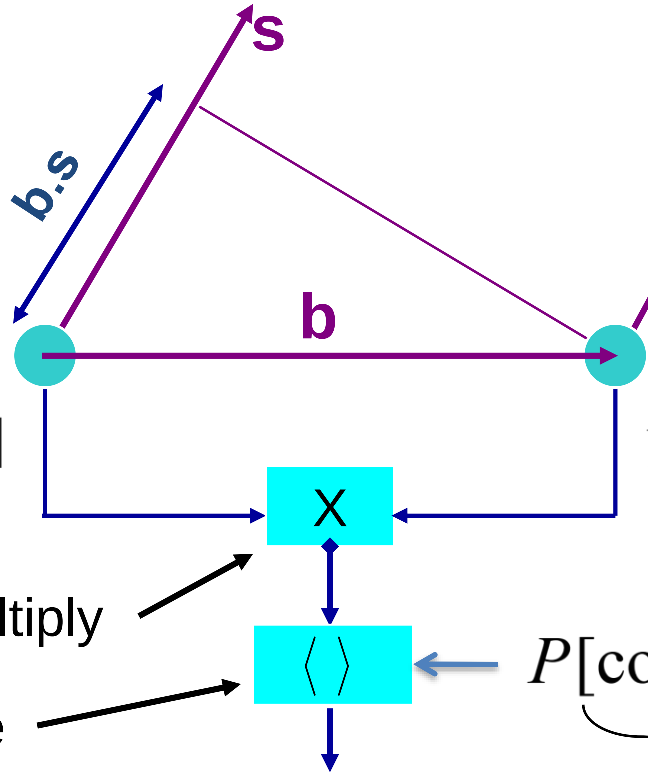
Simplifying Assumptions

- 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
  - Propagation in vacuum, without distortions (no ionosphere, atmosphere …)
  - Idealized electronics (perfectly linear, perfectly uniform in frequency and direction, perfectly identical for both elements, no added noise.
  - Isotropic sensors

The Stationary, Quasi-Monochromatic Radio-Frequency Interferometer

Geometric Time Delay

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



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

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

The path lengths from sensors to multiplier are assumed equal!

multiply

average

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

Unchanging

Rapidly varying,

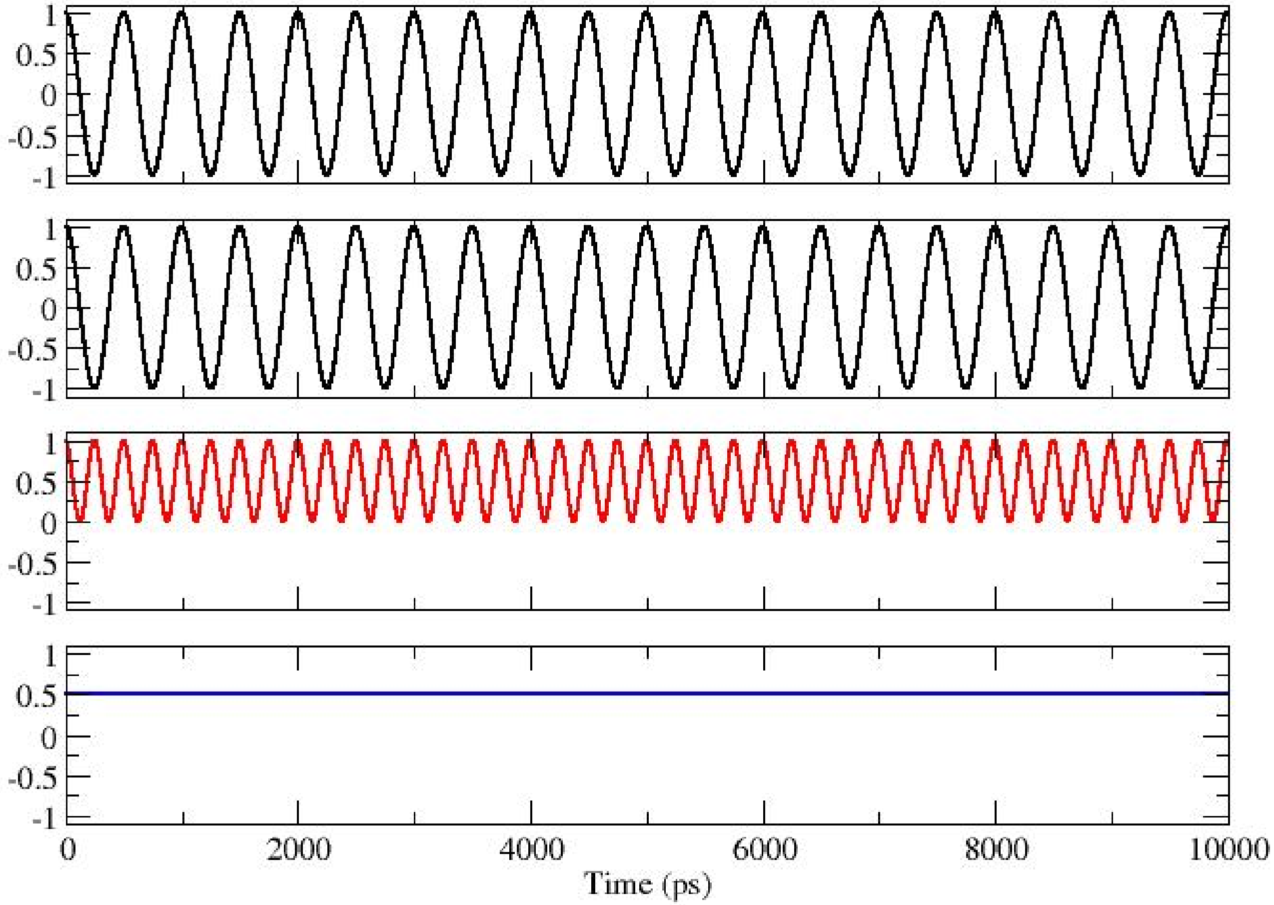
$$R_C = P\cos(\omega\tau_g)$$

Pictorial Example: Signals In Phase

2 GHz Frequency, with voltages in phase:

$\mathbf{b.s} = n\lambda$, or $\tau_g = n/\nu$

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

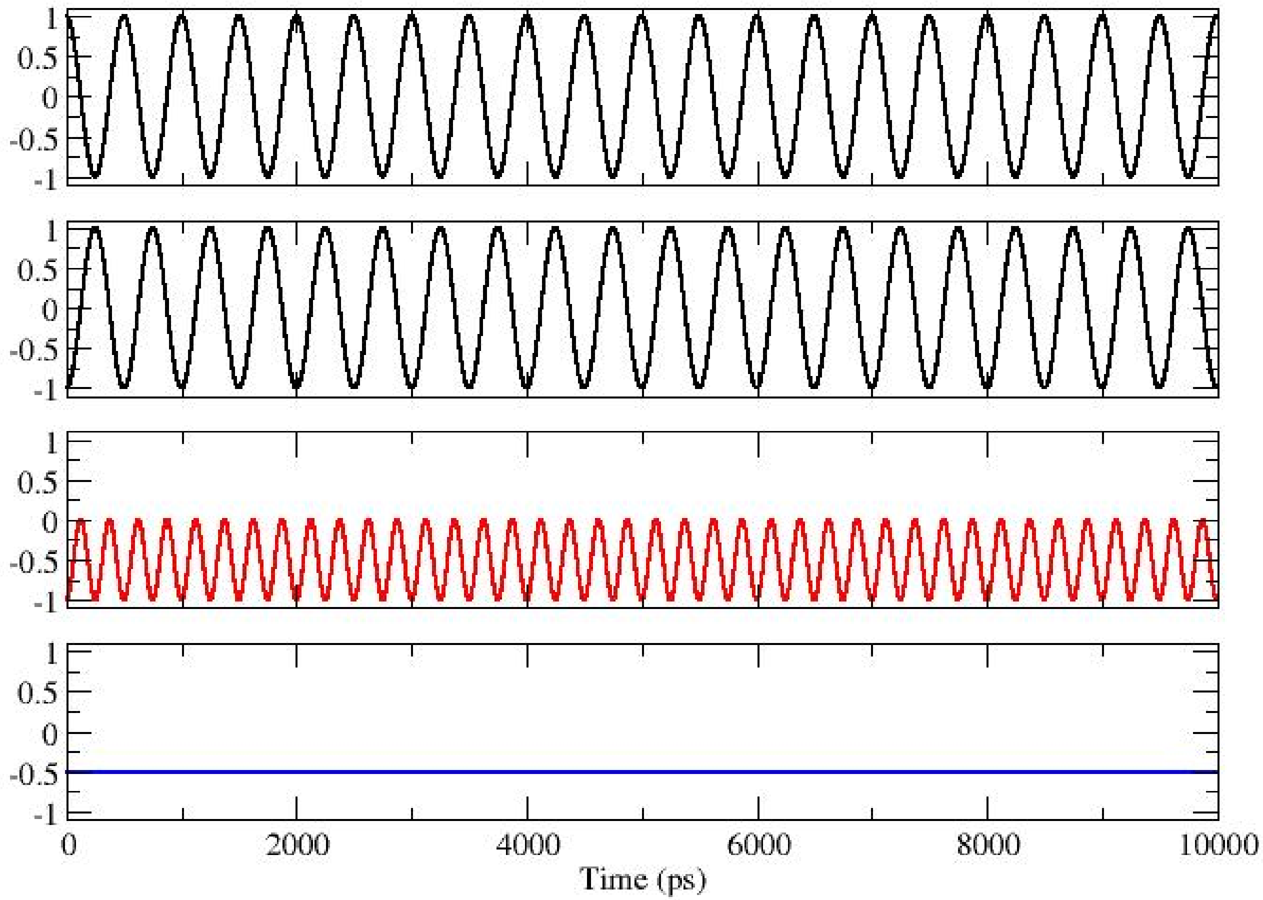


Pictorial Example: Signals out of Phase

2 GHz Frequency, with voltages out of phase:

b.s=(n +/- ½)λ τ_g = (2n +/- 1)/2ν

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

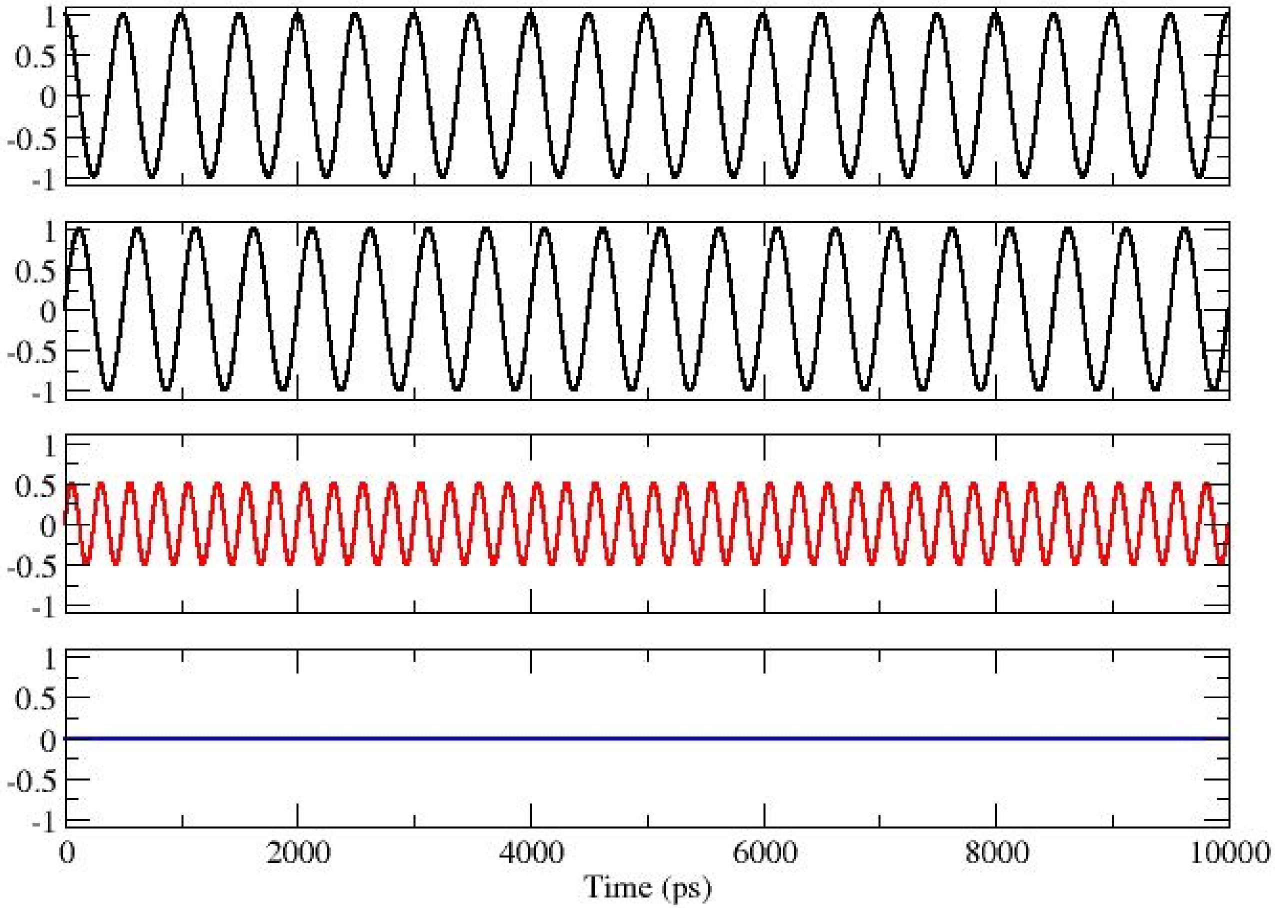


Pictorial Example: Signals in Quad Phase

2 GHz Frequency, with voltages in quadrature phase:

$b.s = (n \pm 1/4)\lambda, \; \tau_g = (4n \pm 1)/4\nu$

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



Some General Comments

- The averaged product R_C is dependent on the received power, $P = E^2/2$ and geometric delay, τ_g, and hence on the baseline orientation and source direction. From slide 9, we have:

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

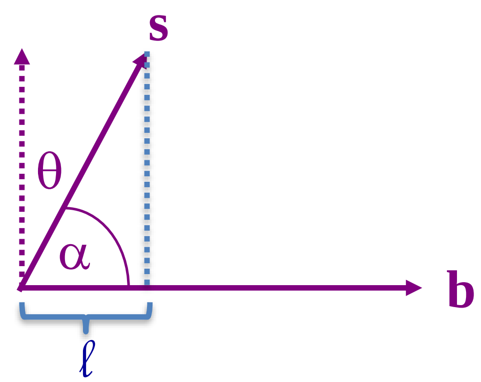
- Note that R_C is not a function of:
  - The time of the observation -- provided the source itself is not variable!
  - The location of the baseline -- provided the emission is in the far-field.
  - The actual phase of the incoming signal – the distance of the source -- provided the source is in the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

- The strength of the product is also dependent on practical factors such as the antenna sizes, electronic gains, bandwidth and time averaging – but these factors can be calibrated for.

Expansion in One Dimension.

- Consider a single baseline, and define the x-axis to extend along this baseline.
- Define $\mathbf{b} = u\hat{\mathbf{x}}$ where u =|**b**|/λ is the baseline length in wavelengths
- Define the `direction cosine' as: $l = \hat{\mathbf{x}} \bullet \mathbf{s} = \cos\alpha = \sin\theta$
- Then:

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



- So the interferometer response is:

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

Two Illustrative Examples

- Consider the response R_c, as a function of angle, for two different baselines with u = b/λ = 10, and u = 25 wavelengths.
- Since $$R_C = \cos(2\pi u l)$$
- We have, for u = 10: $$R_C = \cos(20\pi\, l)$$
- And, for u = 25: $$R_C = \cos(50\pi\, l)$$
- These are simple functions of angle on the sky.

Remember:

u = baseline length in wavelengths

ℓ = sin θ,

θ = fringe angular offset from perpendicular plane

Whole-Sky Response

- Top:

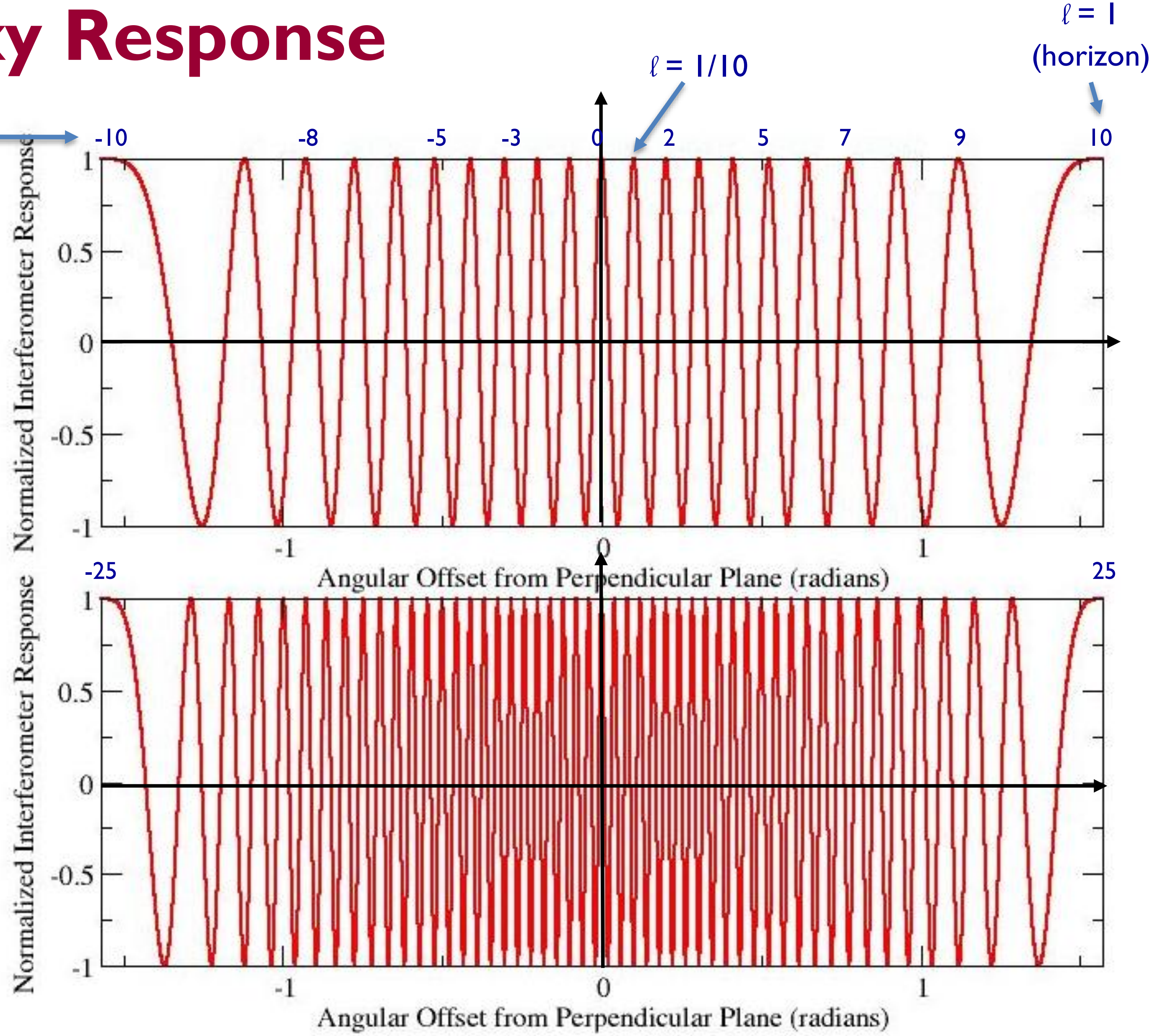
  u = 10

  There are 21 fringe maxima, and 20 fringe minima over the hemisphere.

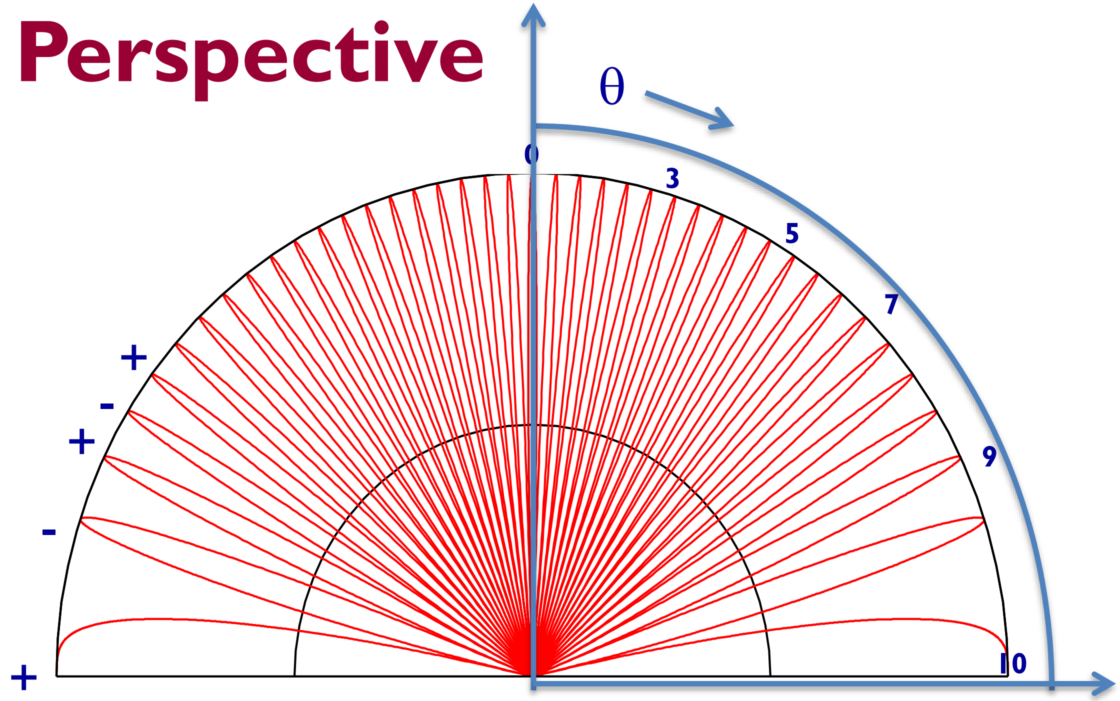
- Bottom:

  u = 25

  There are 51 fringe maxima over the hemisphere



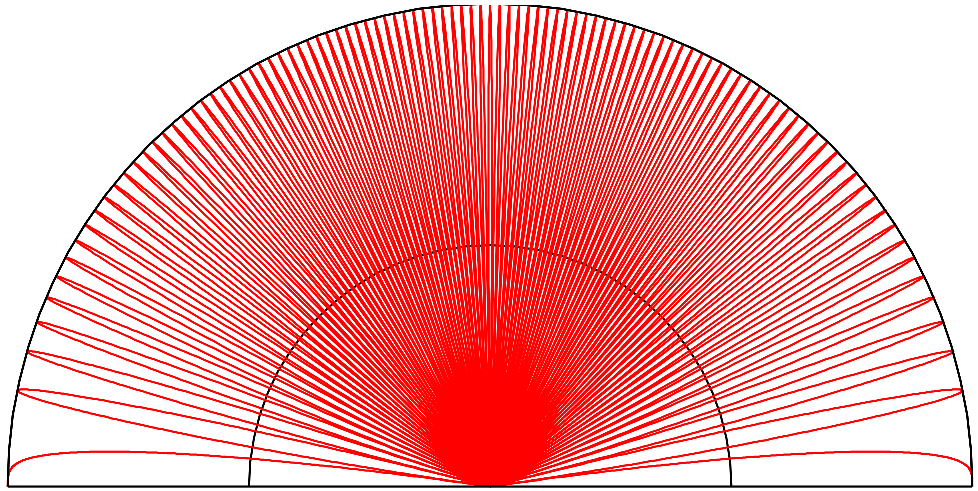
From an Angular Perspective



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

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

Hemispheric Pattern

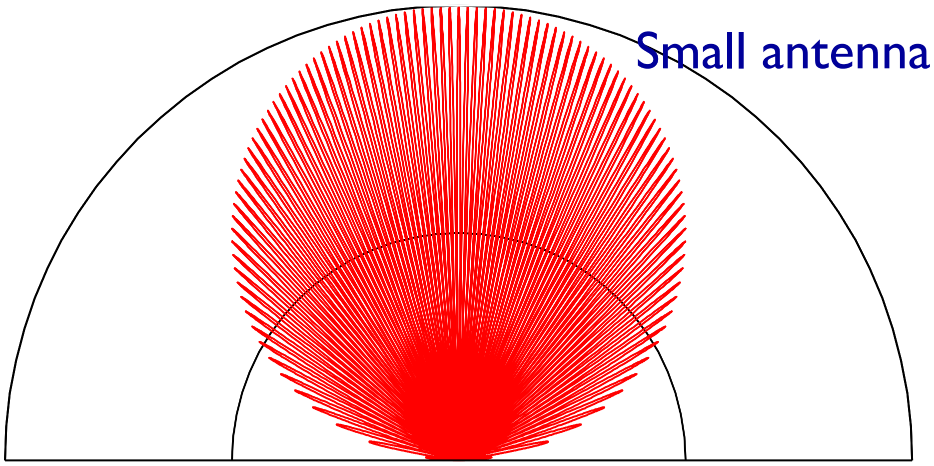
- The preceding plot is a meridional cut through the hemisphere, oriented along the baseline vector.
- In the two-dimensional space, the fringe pattern consists of a series of coaxial cones, oriented along the baseline vector.
- The figure is a two-dimensional representation when u = 4.
- As viewed along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

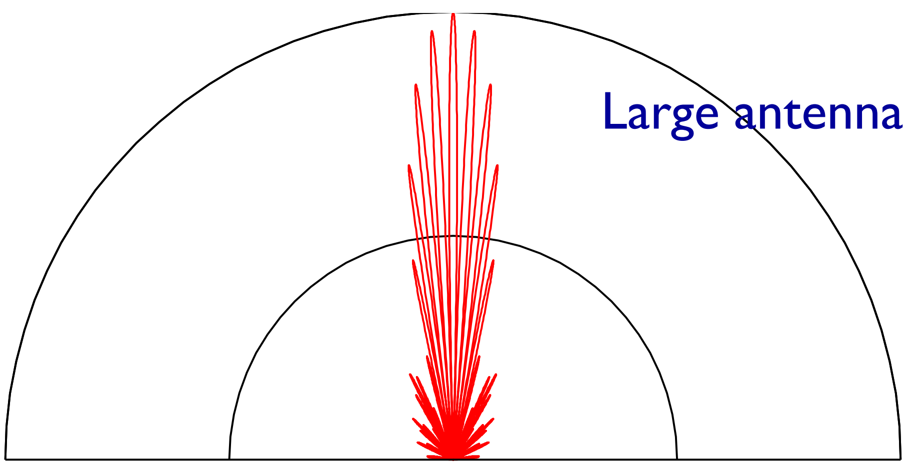
The Effect of the Sensor (aka Antenna)

- The patterns shown presume the sensor 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 sensors have good angular resolution --very useful for some applications like imaging individual objects.
- Small sensors have poor angular resolution – useful for surveys.
- Key Point: The fringe pattern is a function of the baseline length (in wavelengths) and orientation.

The Effect of Sensor Patterns

- Sensors (or antennas) are not isotropic, and have their own responses – both amplitude and phase.
- **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.
- **Emphasis:** The fringe pattern is a function of the array geometry. The angular modulation is from the sensor.





The Response from an Extended Source

- For an extended source, the voltage output is the spatial integration of the E-field emission over the primary beam:

$$V = \iint E(\mathbf{s}) d\Omega$$

- So the response becomes

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

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

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

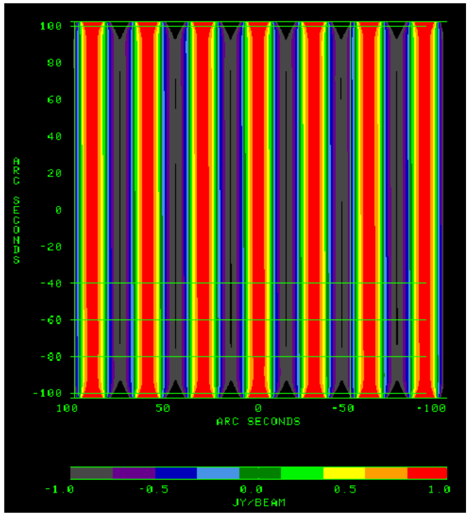
- The response is the flux (spatial integration of the brightness) modulated by the cosinusoidal interferometer pattern.
- This expression links what we want: the brightness on the sky, $I_\nu(\mathbf{s})$, to something we can measure - R_C, the interferometer response.
- Can we recover $I_\nu(\mathbf{s})$ from observations of R_C?

A picture is worth 1000 words …

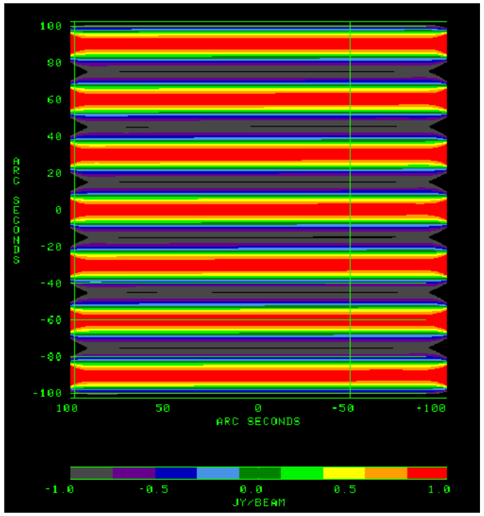
- As stated earlier, these concepts are not difficult, but are unfamiliar. We need to think in new ways, to get a deeper understanding of how all this works.
- To aid, I have generated images of interferometer fringes, of various baseline lengths and orientations.
- I then 'observe' a real source (Cygnus A, of course), to show what the interferometer actually measures.
- For all these, the 'observations' are made at 2052 MHz. The Cygnus A image is take from real VLA data.
- To keep things simple, all simulations are done at meridian transit.

‘Real’ Fringes … 1Km Baseline at 2052 MHz

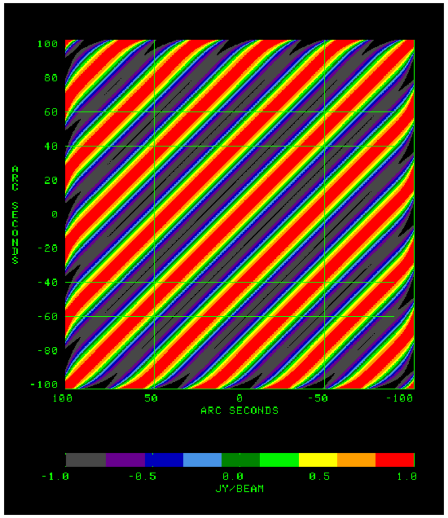
- The fringe separation given by baseline length in wavelengths, the orientation given by the orientation of the baseline.



East-West baseline makes vertical fringes



North-South baseline makes horizontal fringes



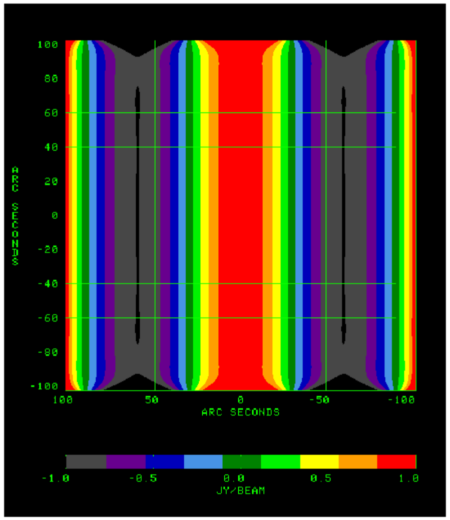
Rotated baseline makes rotated fringes

- Red = positive maximum. Black = negative maximum. Green = zero
- Fringe angular spacing given by baseline length in wavelengths:

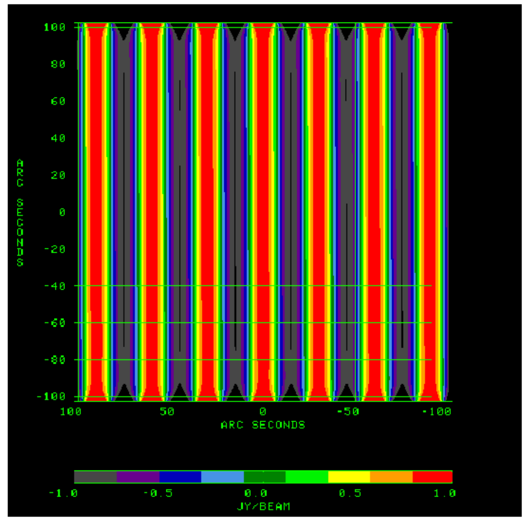
$$\Delta\theta = \lambda / B = 30.2''$$

Longer Baselines => Smaller Fringe Spacing

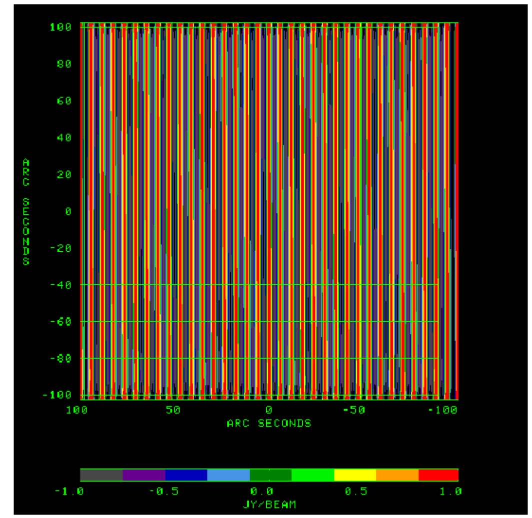
- With longer baselines (**in wavelengths!**) come finer fringes:



250 meter baseline
120 arcsecond fringe



1000 meter baseline
30 arcsecond fringe

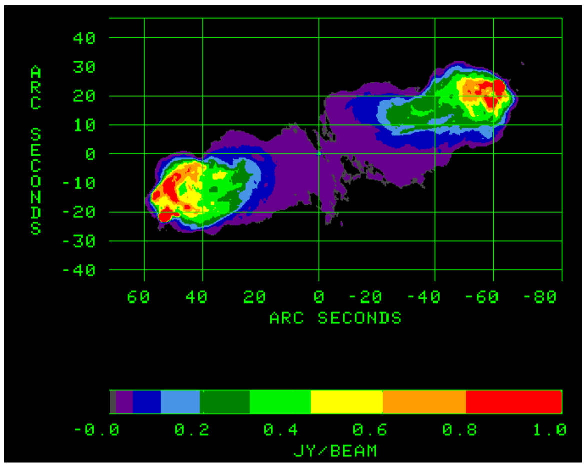


5000 meter baseline
6 arcsecond fringe

- What the interferometer measures is the integral (sum) of the product of this pattern with the actual brightness.

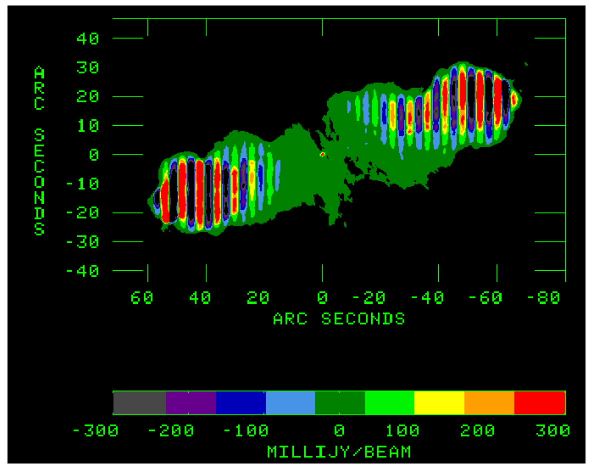
For a Real Source (Cygnus A = 3C405)

- Cygnus A is a powerful, nearby radio galaxy.
- The left panel shows the actual brightness.
- The other two panels show how the 5km-baseline interferometer 'sees' it



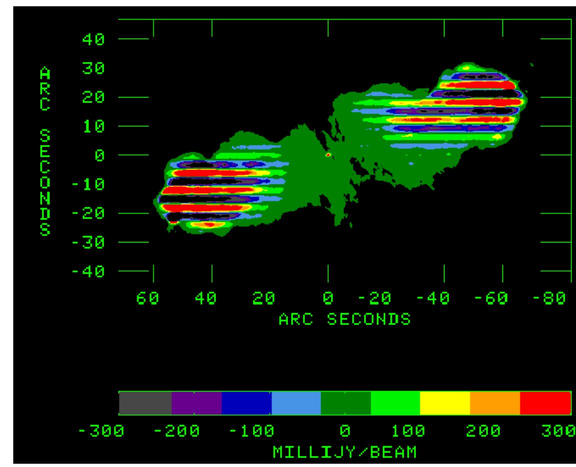
Zero-Spacing Image

Sum = 999 Jy



5 km EW spacing

Sum = 61 Jy



5 km NS spacing

Sum = -16 Jy

- Don't be alarmed by the negative flux in the third panel.
- These are `cosine' fringes – the peak of the center fringe goes through the middle of the target source.

Some Points to Ponder …

- If the target source is a 'point source', the interferometer response is the same for every baseline.
  - 'Point Source' is an object much much smaller than the fringe spacing.
- The interferometer response to a real source can be negative.
  - Although the response is proportional to source power, there is no requirement that it be positive.
- As the baseline gets longer, the response goes to zero.
  - At the point, the source is said to be 'resolved out'.
- As the baseline get shorter, the response goes to the total source flux.
  - This is termed the 'zero spacing flux'.

So … What Good is All This?

- The interferometer casts a cosinusoidal pattern on the sky, with the result that we obtain a response which is some function of the source brightness and the fringe separation and orientation.
- But in fact, something is missing. 'Cosine' fringes are not sufficient to allow recovery of the sky brightness.
- To answer why …
- Time for some mathematics…. Starting with a seeming digression about odd and even functions.
- (All will be clear shortly…)

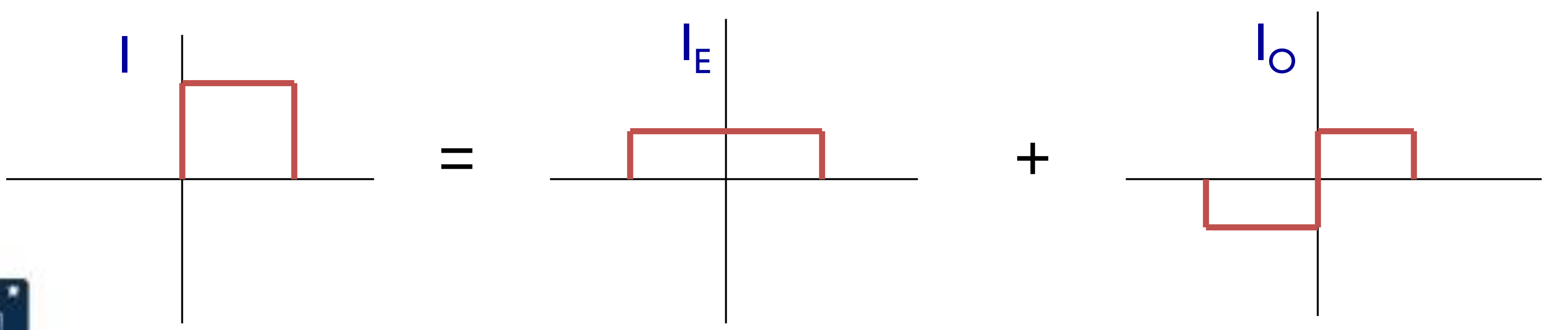
A Short Mathematics Digression – Odd and Even Functions

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



The Cosine Correlator is Blind to Odd Structure

- Suppose that the source of emission has a component with odd symmetry, for which

$$I_o(x) = -I_o(-x)$$

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

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

- Thus, the correlator response R_c:

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

is sensitive to only the even part of the source structure.

Hence, we need more information if we are to completely recover the source brightness.

Thus: 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:

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

since the integral of an odd function (I_O) with an even function (cos x) is zero.

- 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 get

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

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

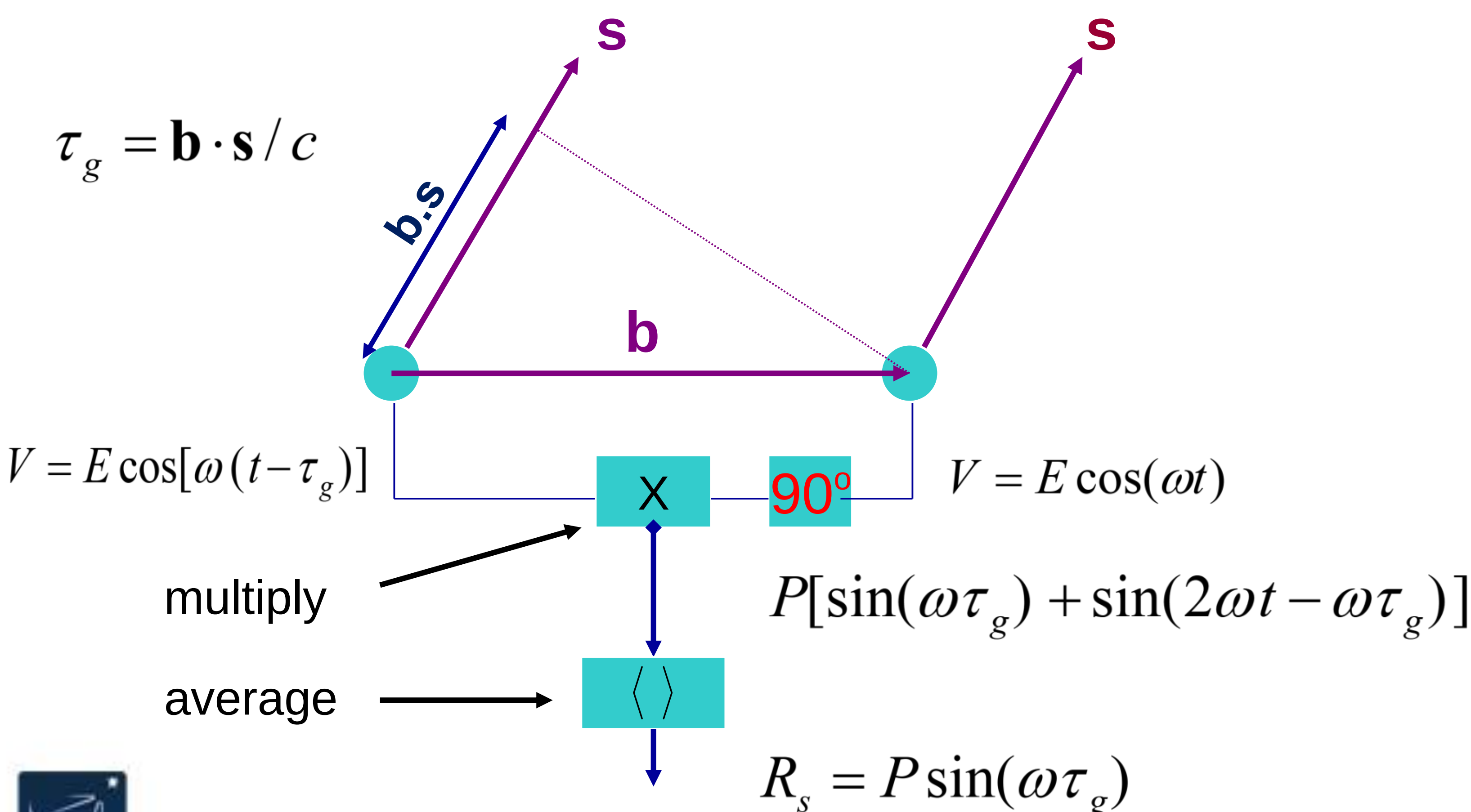
- To obtain this necessary component, we must make a 'sine' pattern.

An Additional Benefit…

- Suppose you build a 'Cos' interferometer, and observe a 'point' source, located at the phase center.
- The observed correlation will always equal the flux density, even as your 'stretch' your baseline.
- But what if your target point source is somewhere else?
- If you had only one baseline, and the source lies in the cosine's null – you won't detect the source at all.
- But … a 'Sin' correlator will…

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:

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

- Under some circumstances, this is a 2-D Fourier transform, giving us a well established way to recover $I(\mathbf{s})$ from $V(\mathbf{b})$.

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:

$$V_1 = A\cos(\omega t) \rightarrow Ae^{-i\omega t}$$

$$V_2 = A\cos[\omega(t - \mathbf{b} \bullet \mathbf{s} / c)] \rightarrow Ae^{-i\omega(t - \mathbf{b} \bullet \mathbf{s}/c)}$$

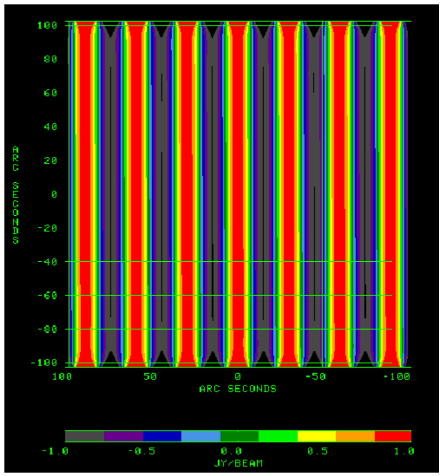
- Then:

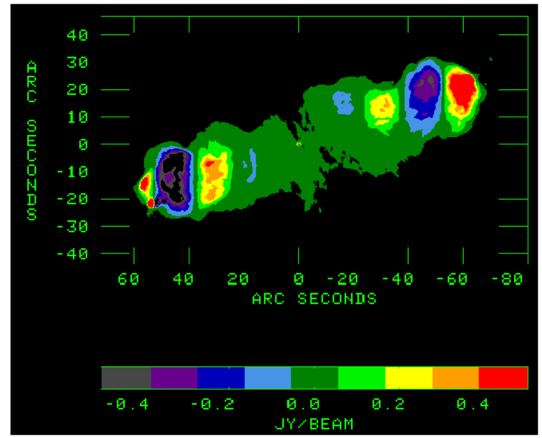
$$P_{corr} = \langle V_1 V_2^* \rangle = A^2 e^{-i\omega \mathbf{b} \bullet \mathbf{s}/c}$$

Some Pictures, to Illustrate This Point

- We now have two (real) correlators, whose patterns are phase shifted by 90 degrees on the sky:

COS



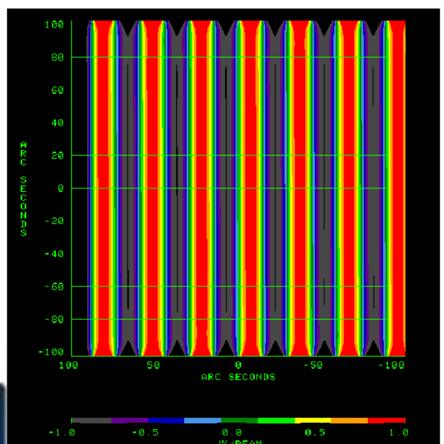


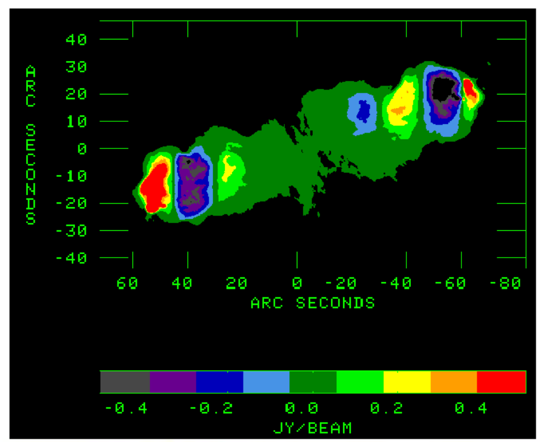
69 Jy

A=103 Jy

ϕ=48

SIN





77 Jy

More Thoughts to Ponder (at 3AM …)

- The complex visibility **amplitude** is independent of the source location*, and linearly related to source flux density.
- The complex visibility **phase** is a function of source location, and independent of source flux density.
- Reversing the elements of an interferometer ('turning it around') negates the phase of the complex visibility, and leaves the amplitude unchanged.
- For those of you familiar with Fourier transforms, the equivalent statement is that:
  - 'As the source brightness is a real function, its Fourier transform is Hermitian'.

* Not strictly true, but close enough for us now.

Some Comments on Visibilities

- The Visibility is a unique function of the source brightness.
- The two functions are related through a Fourier transform. $V_{\nu}(u,v) \Leftrightarrow I(l,m)$
- 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…

Final Comments …

- The formalism presented here presumes much … including that there is no motion between source and interferometer.
- You don't *need* a complex correlator – one can imaging a situation where the interferometer is placed on a slowly rotating platform, which 'sweeps' the fringes through the source.
- Real interferometers are on a rotating platform (the Earth), so why do we use complex correlators?
- The answer to this, and a host of other practical issues, are the subjects of my next lecture.