# Fundamentals of Radio Interferometry

Rick Perley, NRAO/Socorro

Sixteenth Synthesis Imaging Workshop

16-23 May 2018

# Topics

- **Goals:**
  - Mapping – and Resolving -- the Sky
  - Why Interferometry?
- **The Basic Interferometer**
  - Simplifying Assumptions
  - 'Fringe' patterns
  - Sine and Cosine Fringes
  - Response to Extended Emission
  - The Complex Correlator
- **Illustrative Examples**
  - Pictures (lots of pictures)

# Mapping (and Resolving) the Sky

- In astronomy, we wish to know the angular distribution of EM emission.
  - This can be a function of frequency, polarization, and time.
- 'Angular Distribution' means we are interested in the **brightness** of the emission, not just the **flux**.
  - Flux is the spatial integration of the brightness over solid angle.
- Measuring the brightness means making a map.
- 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.

# 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:

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

- To obtain 1 arcsecond resolution at a wavelength of 21 cm, we require an aperture of ~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 technique of synthesizing a larger aperture through combinations of separated pairs of antennas is called 'aperture synthesis'.

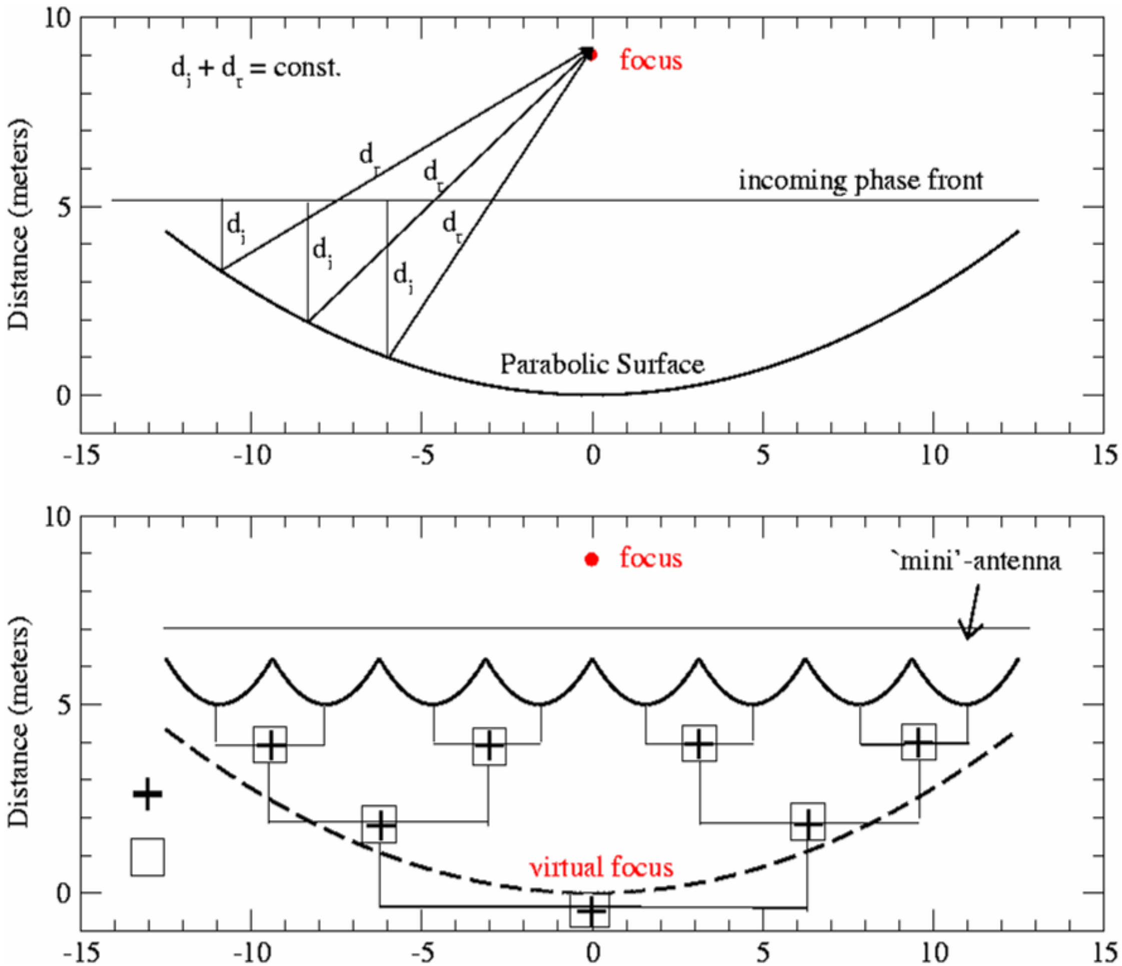
# Interferometry – Basic Concept

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

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

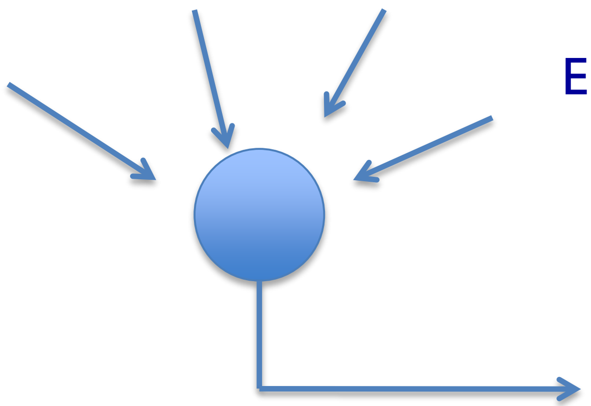
- This is the basic concept of interferometry.

- 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 $E(\mathbf{r},\nu,t)$ at some place ($\mathbf{r}$) 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.



# 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 compared to the frequency: $d\nu << \nu$
- Consider then the electric fields from a small sold angle $d\Omega$ about some direction **s**, within some small bandwidth $d\nu$, at frequency $\nu$.
- We can write the temporal dependence of this field as:

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

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

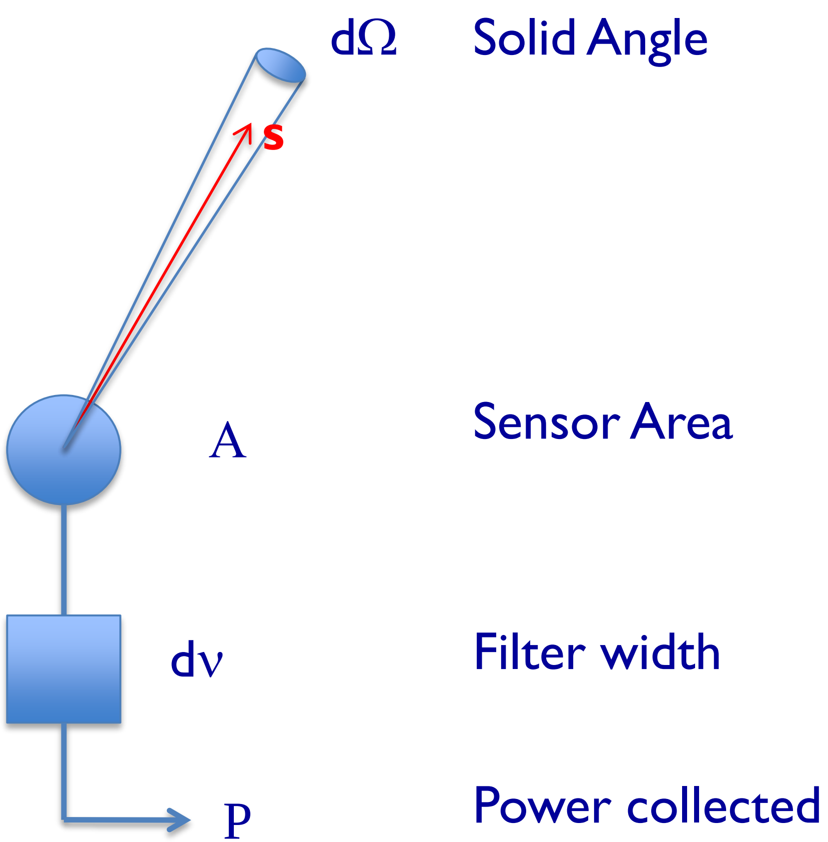
# Analysis Fundamentals

- Imagine a distant source of emission, described by brightness I(ν,**s**) where **s** is a unit direction vector.
- Power from this emission is intercepted by a collector (`sensor') of area A(ν,**s**).
- The power, dP (watts) from a small solid angle dΩ, within a small frequency window dν, 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.

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



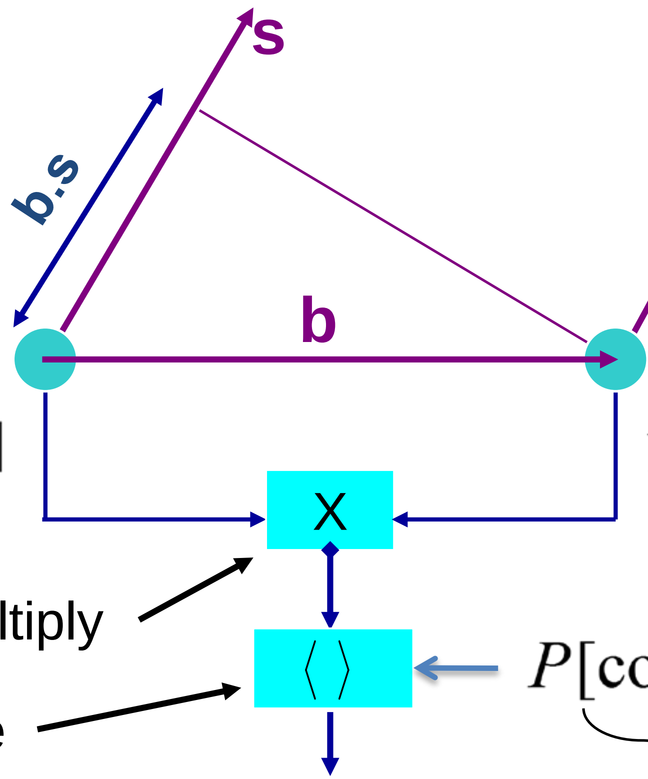
# Simplifying Assumptions

- 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

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




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

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

The path lengths from sensors to multiplier are assumed equal!

multiply

average

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

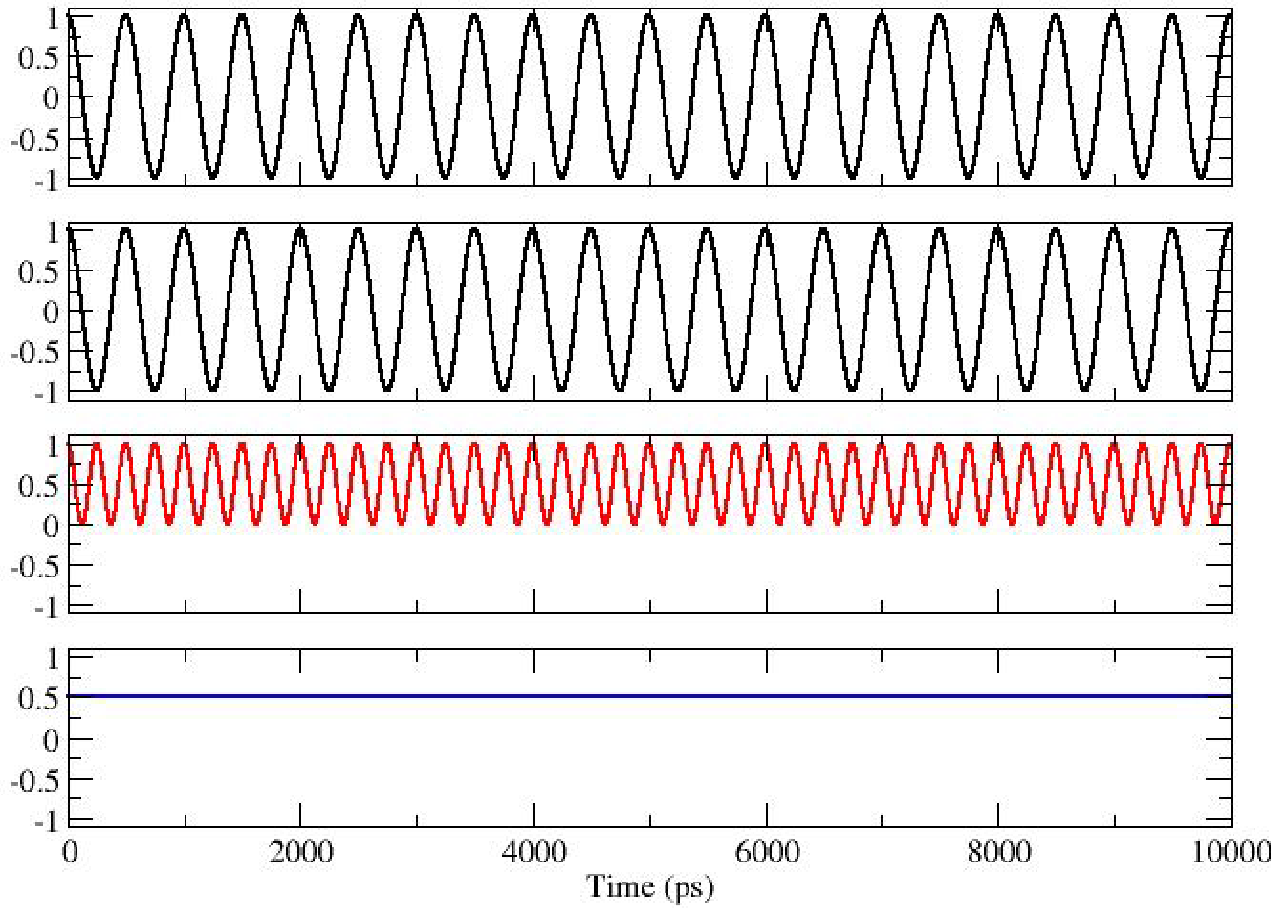
$$R_C = P \cos(\omega\tau_g)$$

# Pictorial Example: Signals In Phase

2 GHz Frequency, with voltages in phase:

**b.s** = n$\lambda$, or $\tau_g = n/\nu$

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

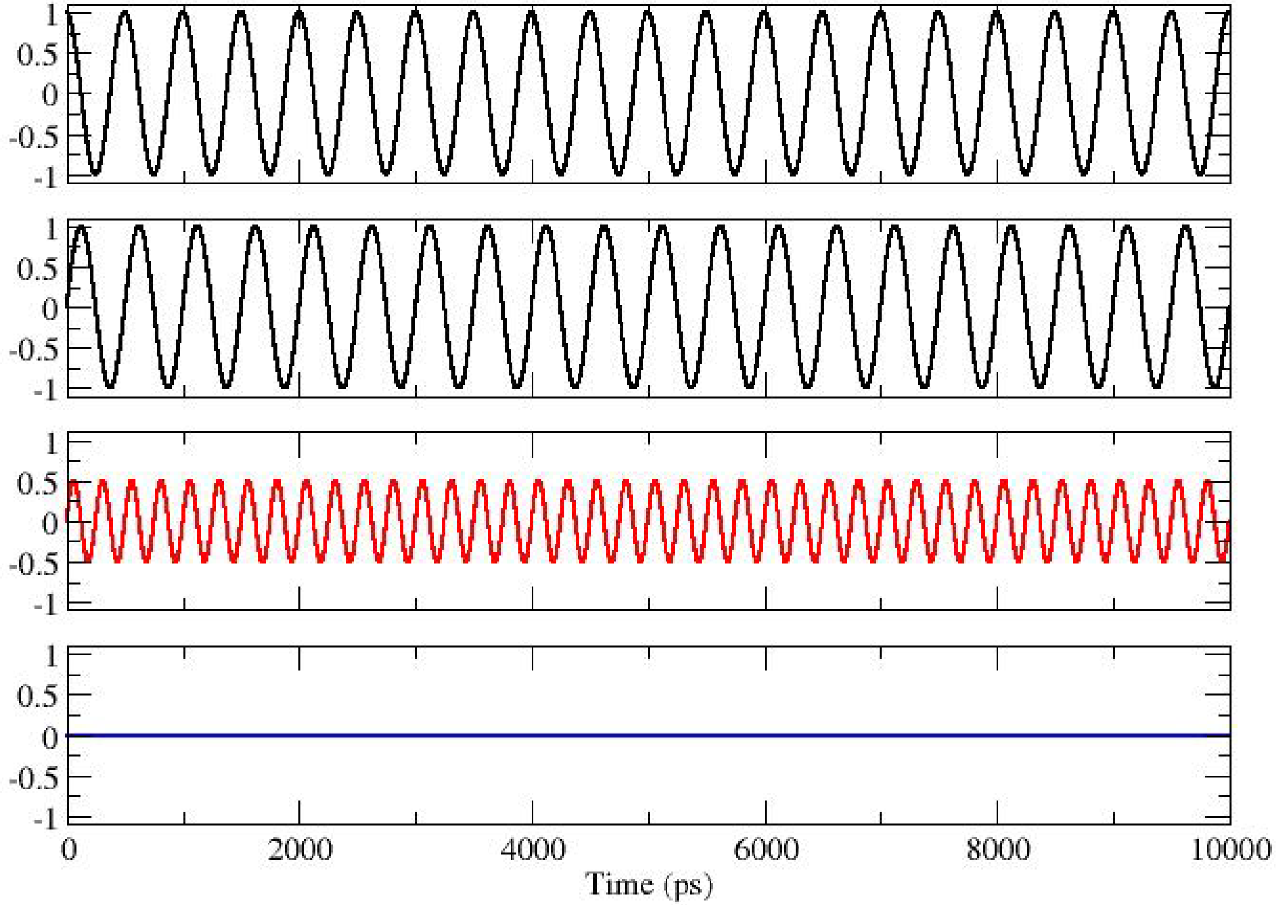


# Pictorial Example: Signals in Quad Phase

2 GHz Frequency, with voltages in quadrature phase:

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

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

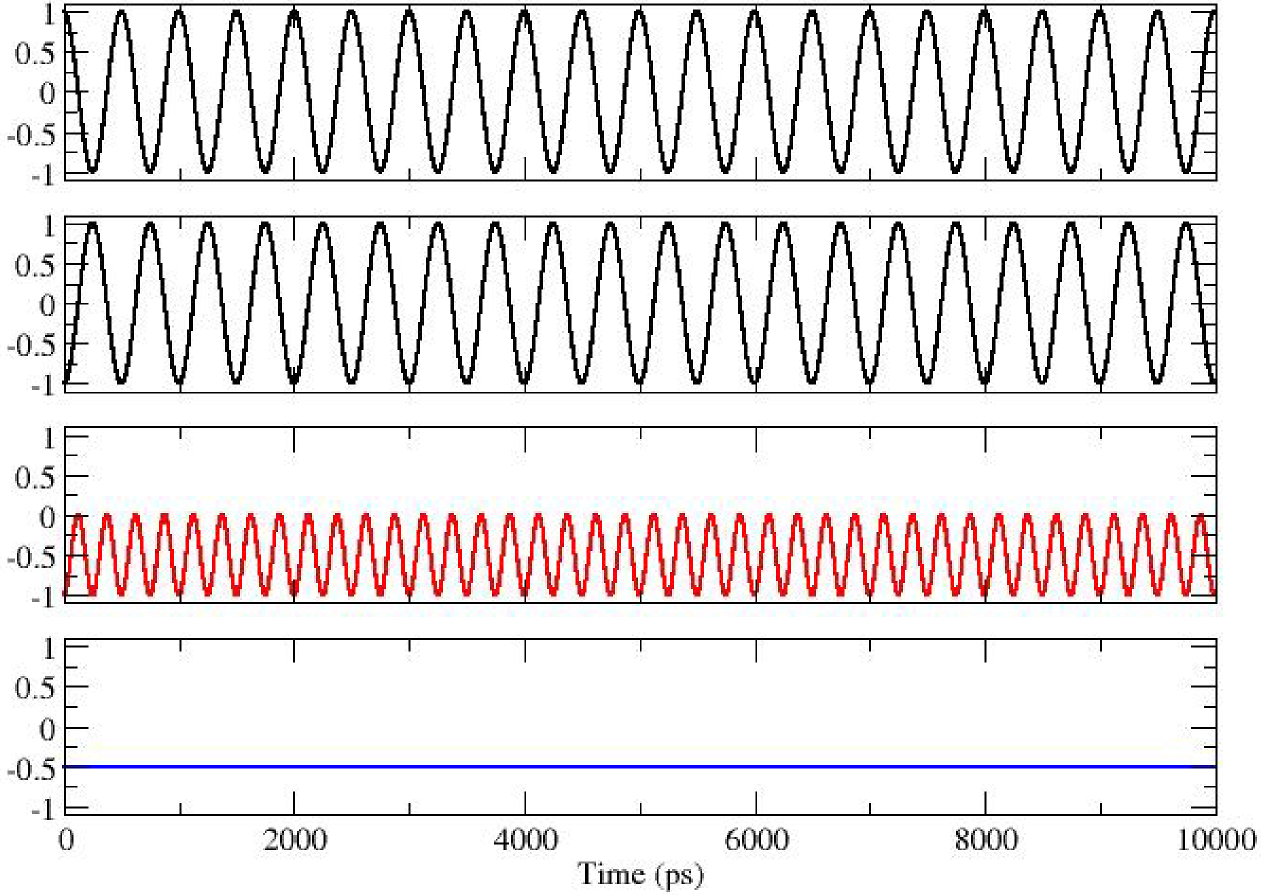


# Pictorial Example: Signals out of Phase

2 GHz Frequency, with voltages out of phase:

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

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



# Some General Comments

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

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

- 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 dependent on the antenna sizes and electronic gains – but these factors can be calibrated for.

# Pictorial Illustrations

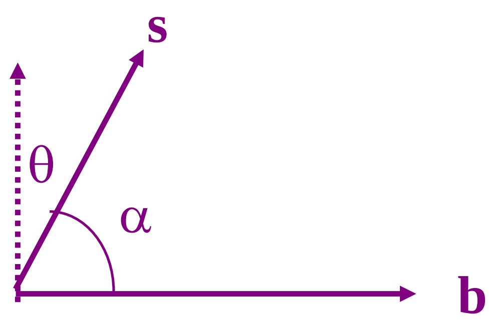
- To illustrate the response, expand the dot product in one dimension:

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

- Thus, the 'cosine' response can now be written

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

- Here, **u = b/λ is the baseline length in wavelengths**, and θ is the angle w.r.t. the plane perpendicular to the baseline.
- And $l = \cos\alpha = \sin\theta$ is the **direction cosine**




- So, what do these patterns 'look like' on the sky?

# The 'Cosine' Interferometer Response

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

$$R_C = \cos(2\pi u l)$$

- We have:

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

- And

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

- These are simple functions of angle on the sky.
- Some illustrations should help.

# Whole-Sky Response

- Top:

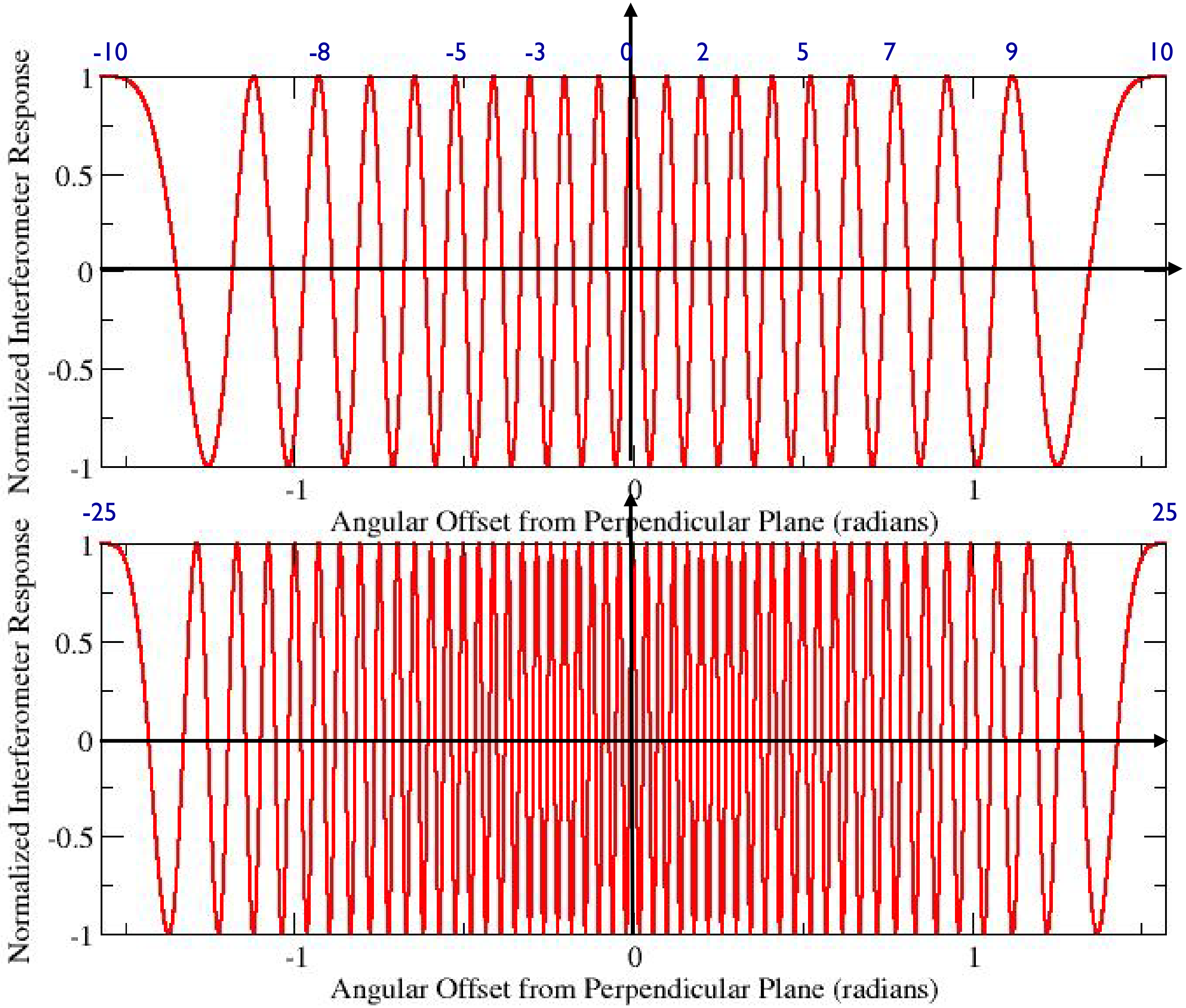
  u = 10

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.

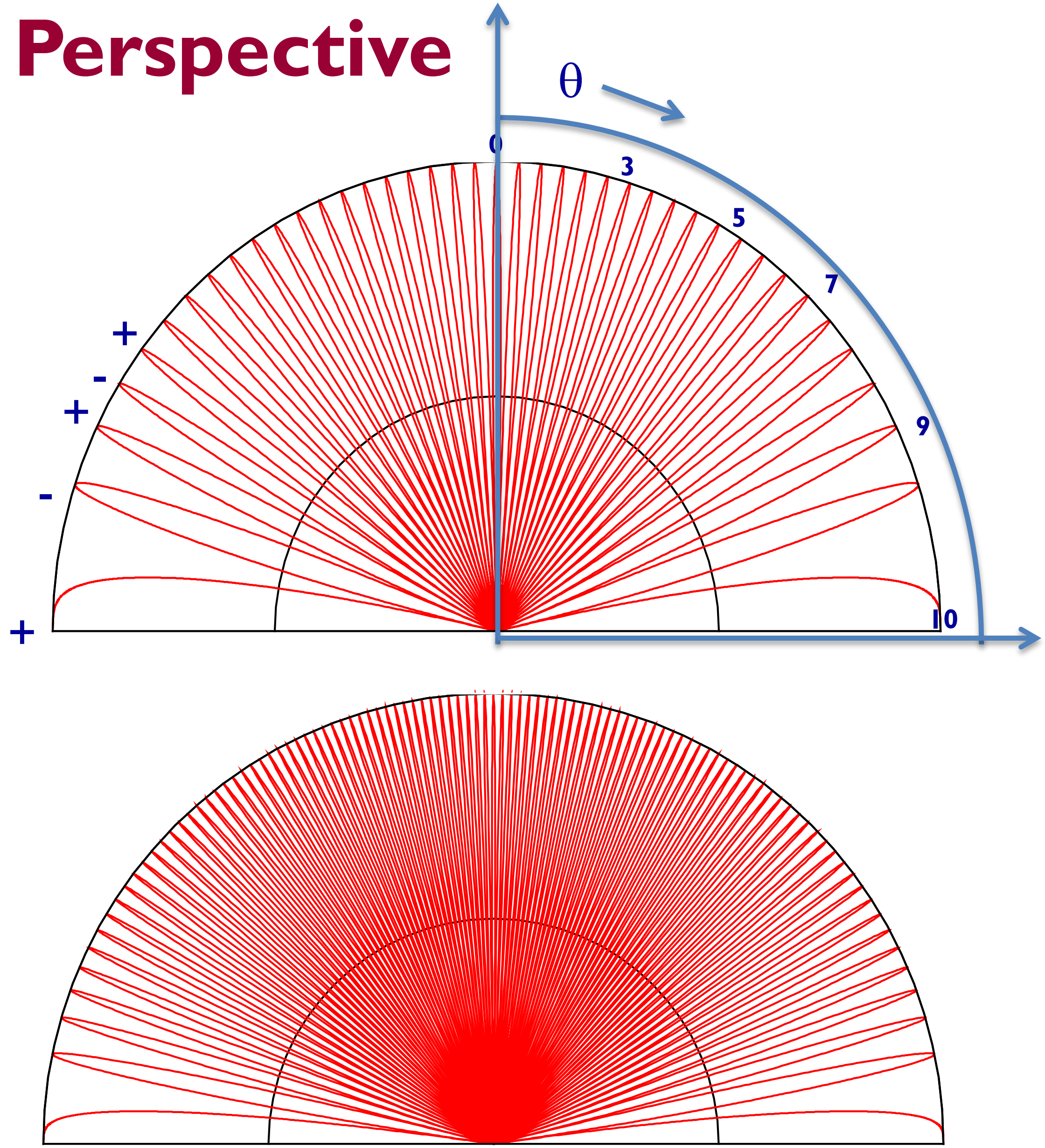
The ‘lobes’ of the response pattern alternate in sign.

**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

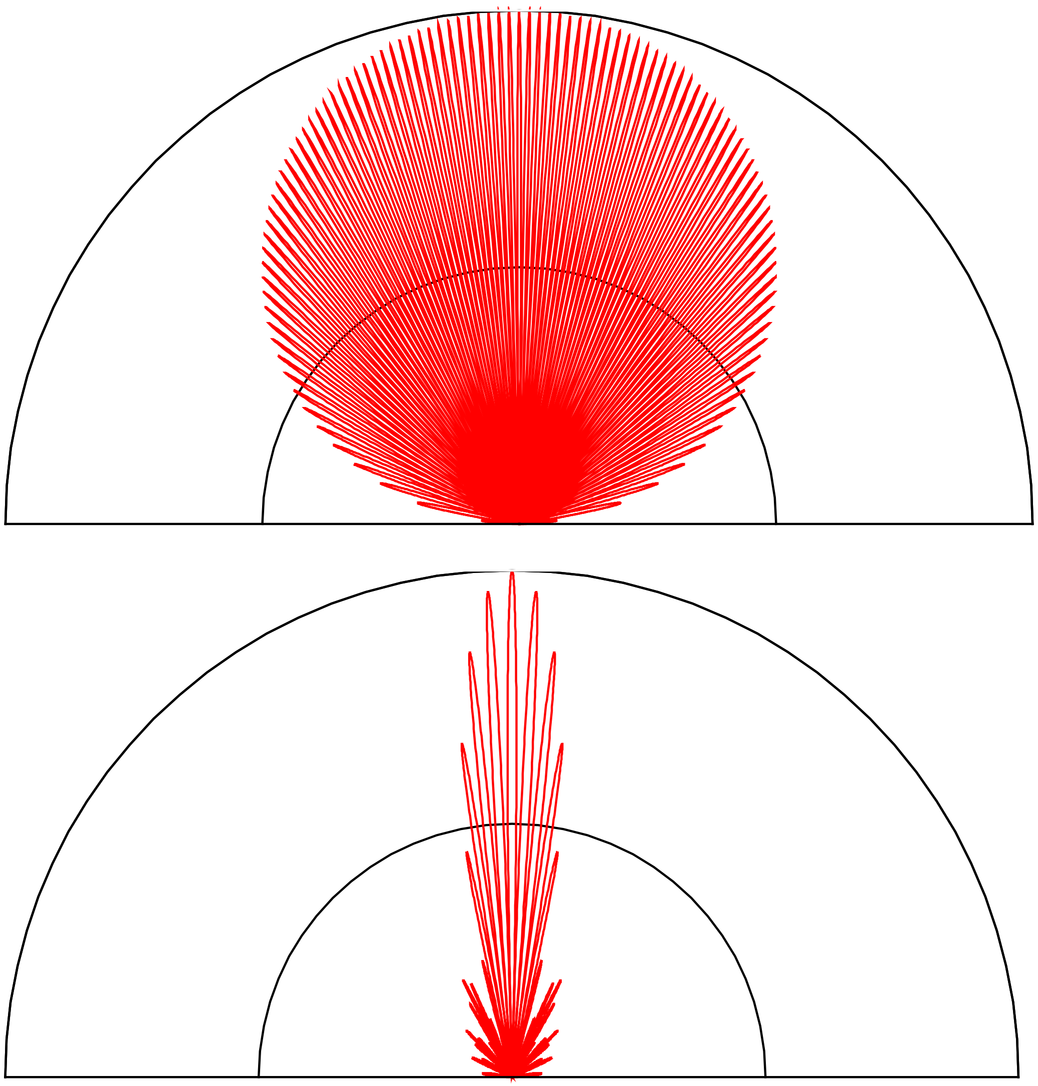
- The preceding plot is a meridional cut through the hemisphere, oriented along the baseline vector.
- In the two-dimensional space, the fringe pattern consists of a series of coaxial cones, oriented along the baseline vector.
- The figure is a two-dimensional representation when u = 4.
- As viewed along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

# The Effect of the Sensor

- The patterns shown presume the sensor has isotropic response.
- This is a convenient assumption, but (sadly, in some cases) 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 – 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.



# 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. Formal details are complicated, but in summary

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

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

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

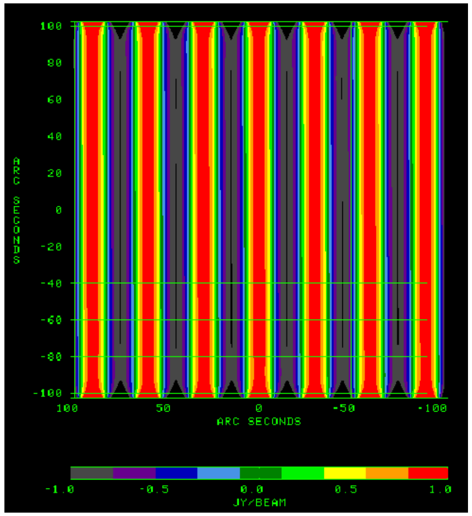
- The response is the flux (spatial integration of the brightness) modulated by the cosinusoidal interferometer pattern.
- This expression links what we want: the brightness on the sky, $I_\nu(\mathbf{s})$, to something we can measure - $R_C$, the interferometer response.
- Can we recover $I_\nu(\mathbf{s})$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

# A picture is worth 1000 words …

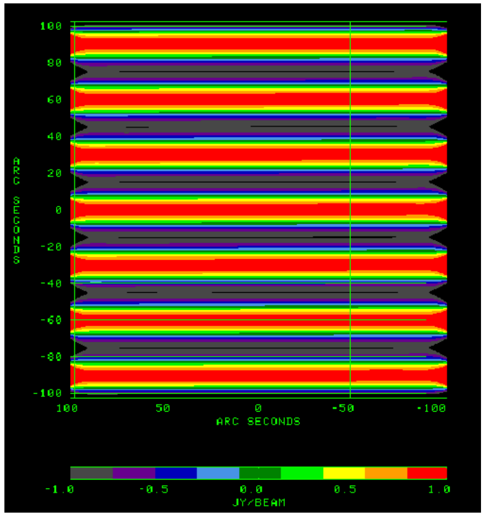
- As stated earlier, these concepts are not difficult, but are unfamiliar. We need to think in new ways, to get a deeper understanding of how all this works.
- To aid, I have generated images of interferometer fringes, of various baseline lengths and orientations.
- I then 'observe' a real source (Cygnus 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

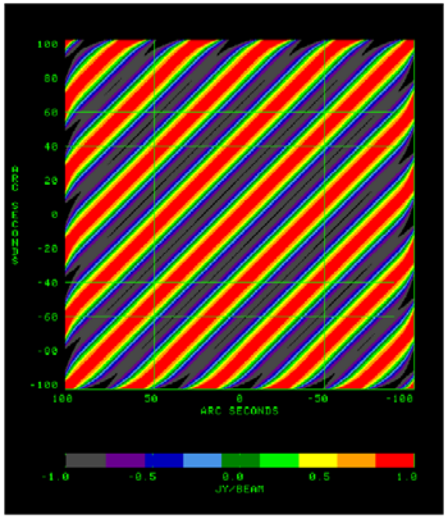
- The fringe separation given by baseline length in wavelengths, the orientation given by the orientation of the baseline.




East-West baseline makes vertical fringes




North-South baseline makes horizontal fringes



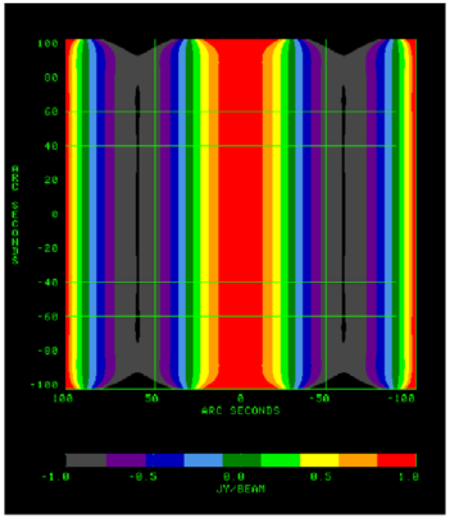

Rotated baseline makes rotated fringes

- Fringe angular spacing given by baseline length in wavelengths:

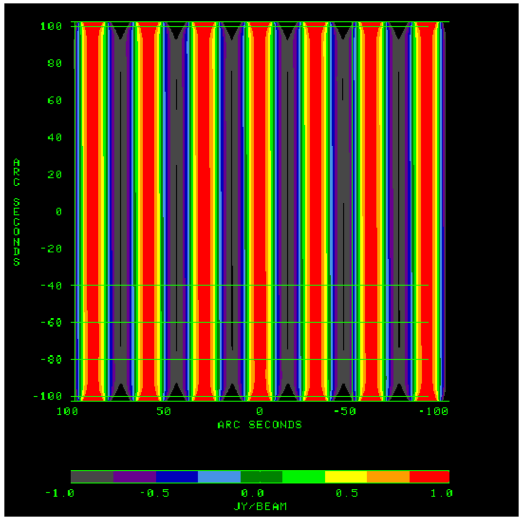
$$\Delta\theta = \lambda / B = 30.2''$$

# Longer Baselines => Smaller Fringes

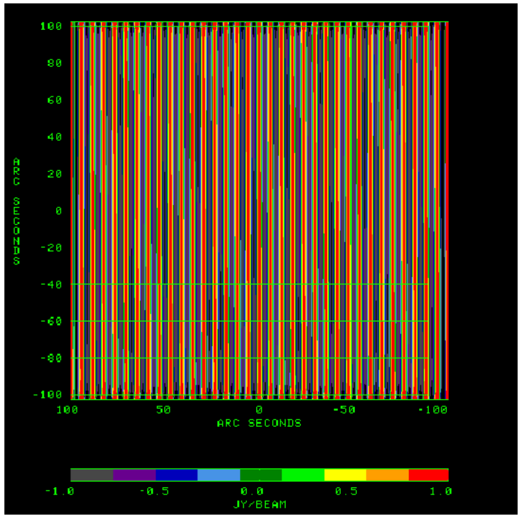
- With longer baselines (**in wavelengths!**) come finer fringes:




250 meter baseline
120 arcsecond fringe




1000 meter baseline
30 arcsecond fringe

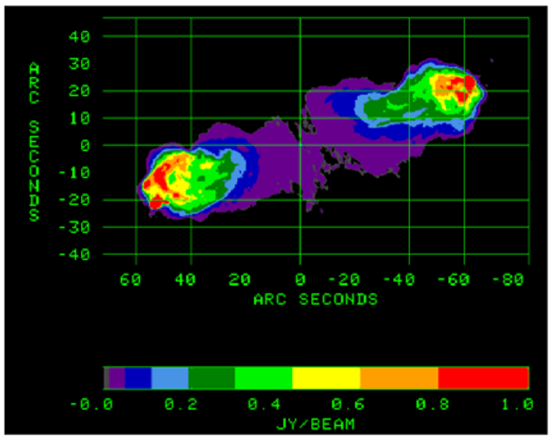



5000 meter baseline
6 arcsecond fringe

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

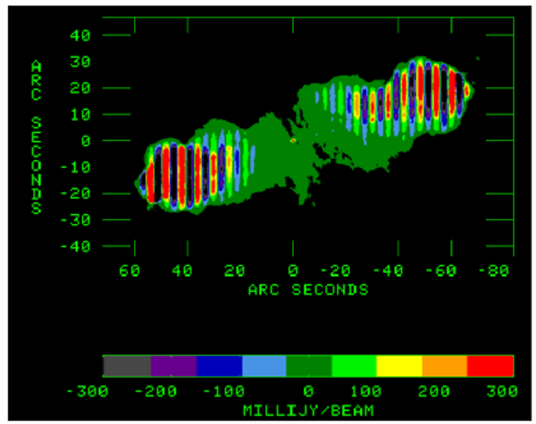
# For a Real Source (Cygnus A = 3C405)

- 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



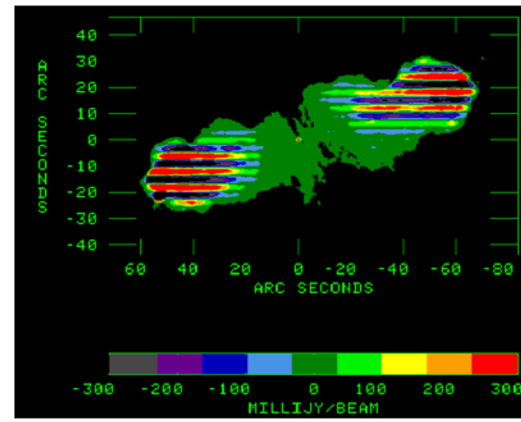

Zero-Spacing Image

Sum = 999 Jy




5 km EW spacing

Sum = 61 Jy




5 km NS spacing

Sum = -16 Jy

# 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'.
- What do we do with the case where the source lies in the null of the cosinusoidal pattern?

# 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.
- How does that get us to our goal of determining the actual brightness?
- Time for some mathematics…. Starting with a seeming digression about odd and even functions.
- (All will be clear shortly…)

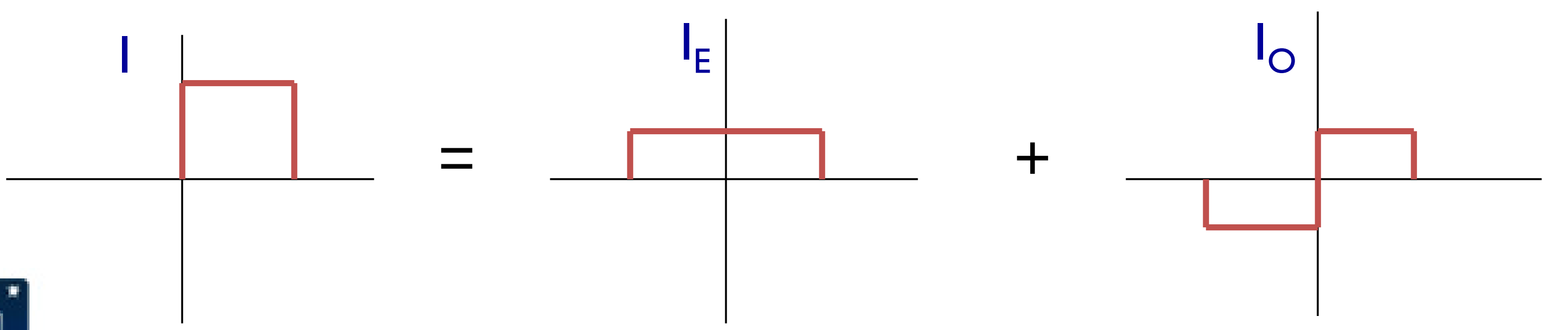
# A Short Mathematics Digression – Odd and Even Functions

- 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

- 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?

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

$$I_o(\mathbf{s}) = -I_o(-\mathbf{s})$$

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

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

- 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:

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

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

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

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

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

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

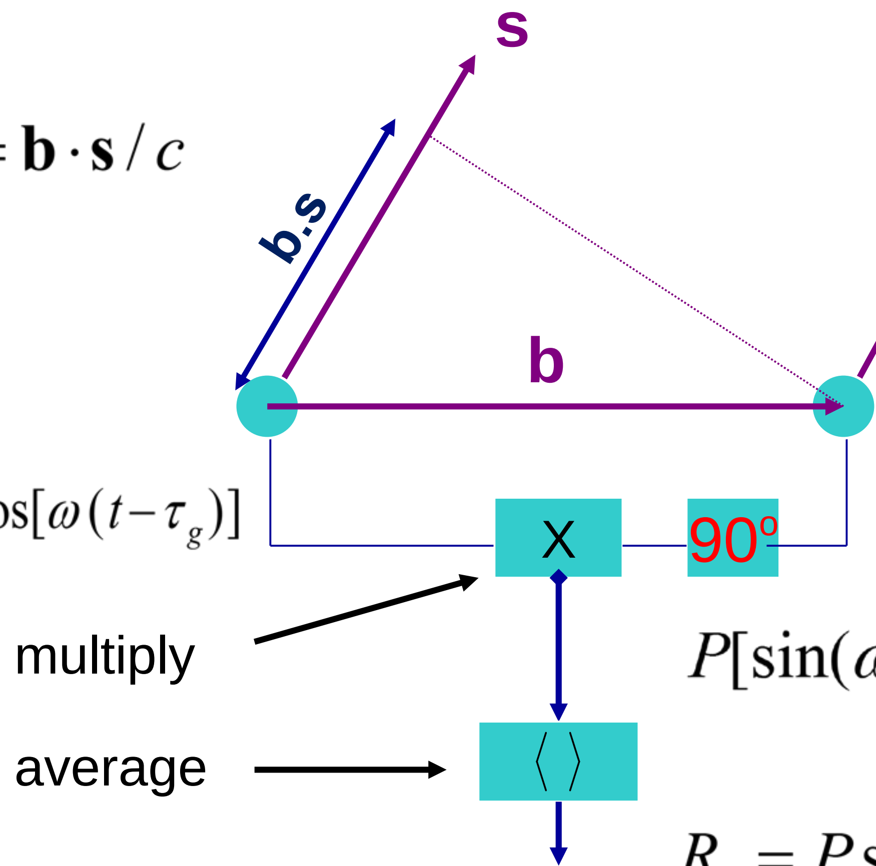
# Another Way to Think About This …

- 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?
- Then, as the baseline gets longer, the response from this point source will oscillate sinusoidally as the baseline lengthens.
- 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.

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




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

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

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

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

# 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:

$$V_1 = A\cos(\omega t) = \mathrm{Re}\left(Ae^{-i\omega t}\right)$$

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

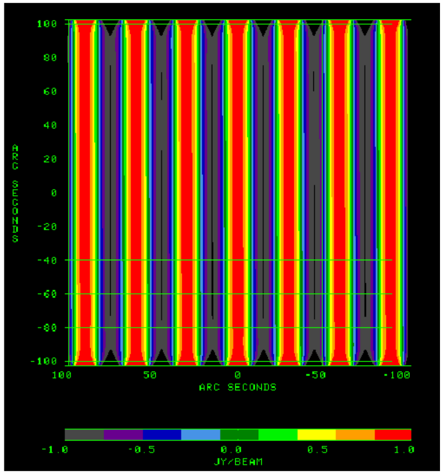
- Then:

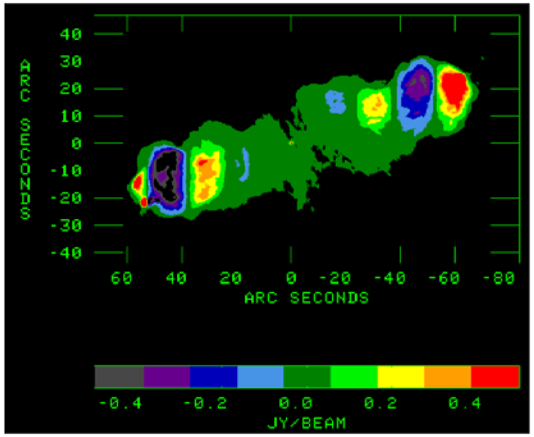
$$P_{corr} = \left\langle V_1 V_2^* \right\rangle = A^2 e^{-i\omega \mathbf{b}\bullet\mathbf{s}/c}$$

# Some Pictures, to Illustrate This Point

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

COS



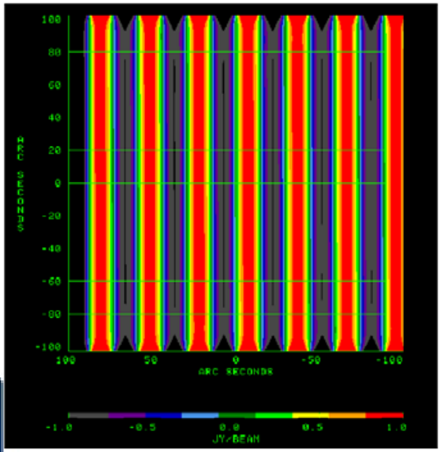



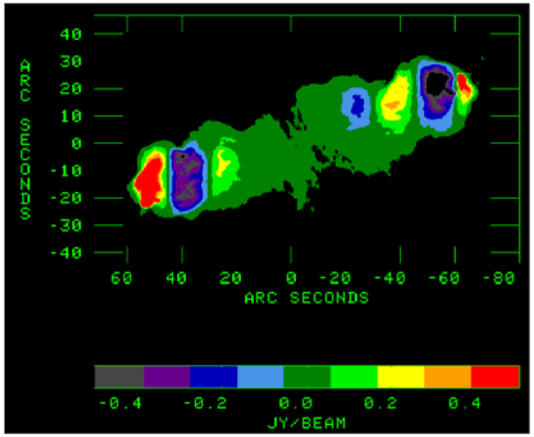

69 Jy

A=103 Jy

$\phi$=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'.

# 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 (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 …

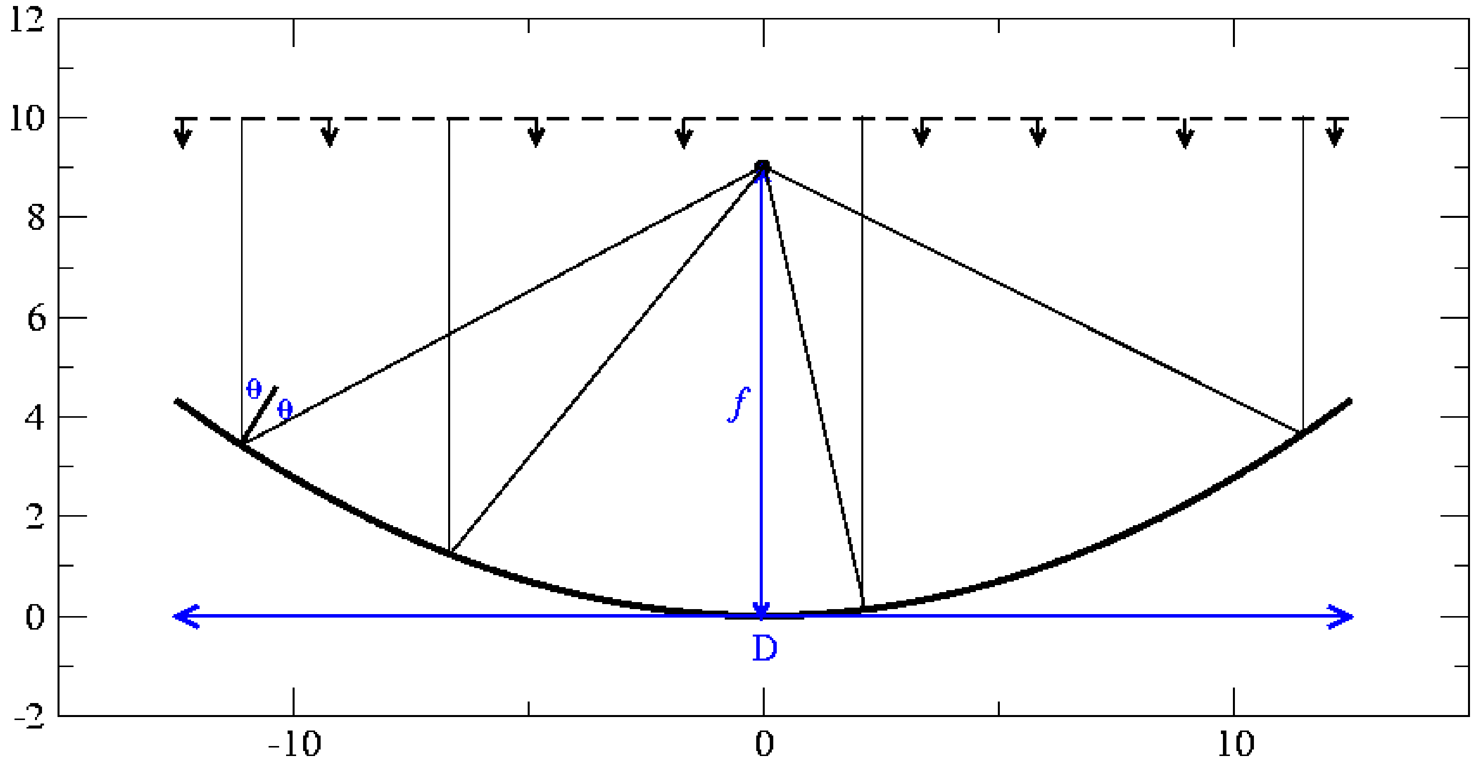
- The formalism presented here presumes much … including that there is no motion between source and interferometer.
- You don’t *need* a complex correlator – one can imaging a situation where the interferometer is placed on a slowly rotating platform, which ‘sweeps’ the fringes through the source.
- 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.

# Antennas – the Single Dish

- Antennas 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: $\theta \sim \lambda/D$ (radians)
  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

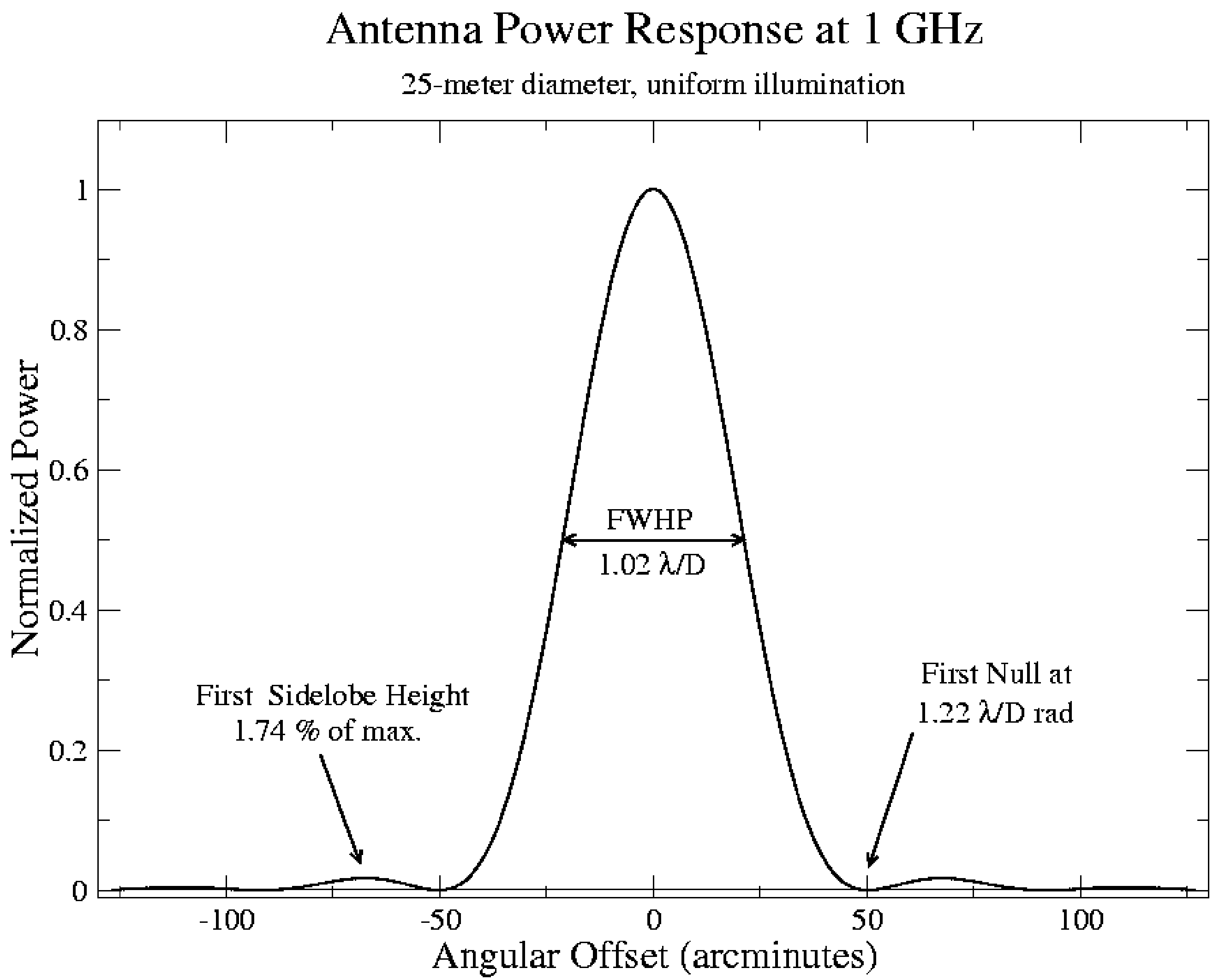
- Key Point: Distance from incoming phase front to focal point is the same for all rays.
- The E-fields will thus all be in phase at the focus – the place for the receiver.



# The Standard Parabolic Antenna Response

The angular power response of a uniformly illuminated circular parabolic antenna of 25-meter diameter,
at a frequency of 1 GHz.

Uniform weighting over the surface is assumed.

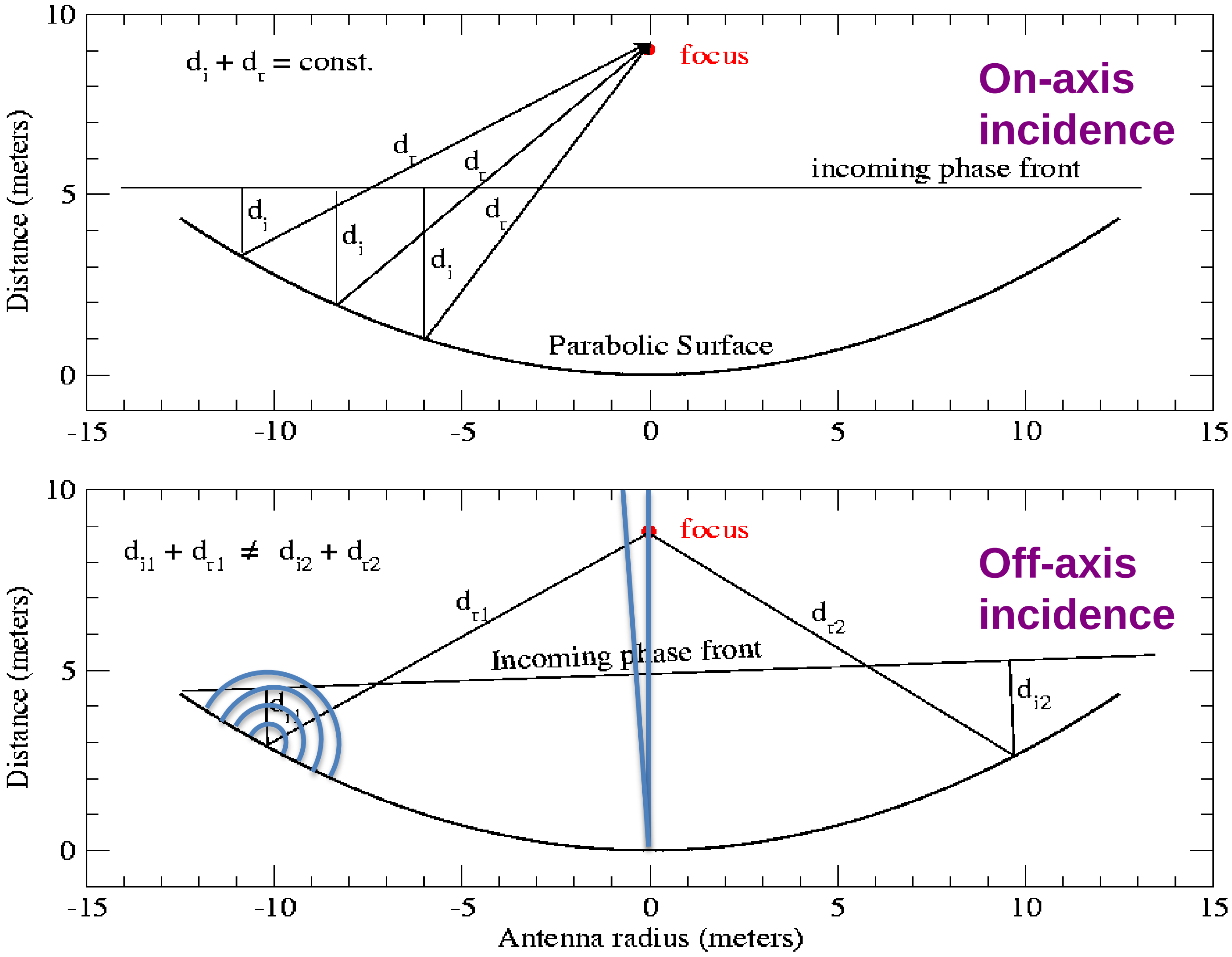


# Beam Pattern Origin
## (1-Dimensional Example)

- 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:

$\theta \sim \lambda/D$

(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): $\theta = \lambda/D, 2\lambda/D, 3\lambda/D, \ldots$ radians.
- When the phase differential across the full aperture is ~1.5, 2.5, 3.5, … wavelengths, we get a maximum in total received power.
  - These are the 'sidelobes' of the antenna response.
  - But, each successive maximum is weaker than the last.
  - These maxima appear at (approximately): $\theta = 3\lambda/2D, 5\lambda/2D, 7\lambda/2D, \ldots$ radians.