Fundamentals of Radio Interferometry

Rick Perley, NRAO/Socorro

2019 David Dunlap Summer School

July 2019

Who Am I? And How Did I Get Here?

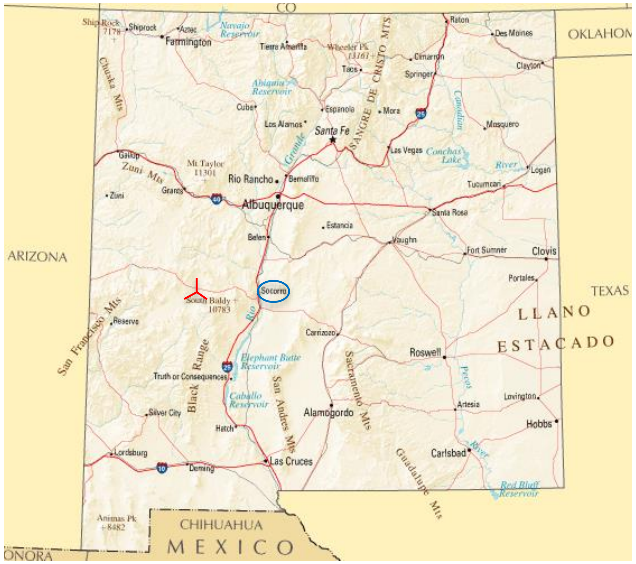
- Staff Scientist at NRAO, in Socorro, NM since 1977.
- A native of British Columbia
  - Lived in many cites,
  - Jr. High School in Prince Rupert
  - High School in Kelowna
- BSc (Physics) at UBC.
- In 1967 got NRC summer job in Ottawa
  - Discovered radio astronomy
  - Discovered wilderness canoeing (ARO, Lake Traverse, ON).
- MSc at new astronomy program at UBC, then decided I had to leave BC.
- Joined Astronomy Program at U. Md.

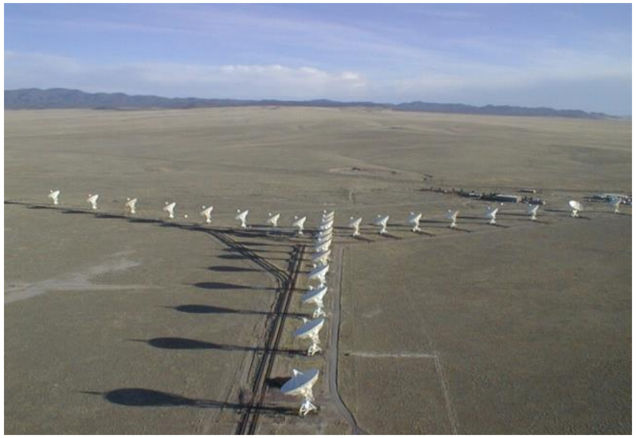
Who Am I -- PhD

- Took Fourier transforms course from Bill Erickson --- became his graduate student.
- Used the E-W arm of the \`TPT', at Clark Lake in southern California to do aperture synthesis at 20 – 73 MHz.
- 3 km 'T' array on a (usually) dry lake bed.
- One of the hottest, emptiest, and most forlorn places on earth …
- I spent a lot of time here living in a trailer…
- Finished in September, 1977.
- Got NRAO postdoc with the VLA.

Who Am I – on to NRAO and the VLA

- My CLRO experience apparently impressed NRAO staff, so …
- Postdoc at NRAO (Socorro) – first postdoc at the new VLA (Oct 1977)
- Staff scientist in 1980.
- Major research areas: radio galaxies, quasars
- Strong interest in calibration/imaging/polarimetry
  - Driven by the need to improve VLA imaging capabilities
- Was 'EVLA Project Scientist' from 1998 through 2012.
- Am now just trying to make the JVLA work as well as we designed it to be, and return to some science.

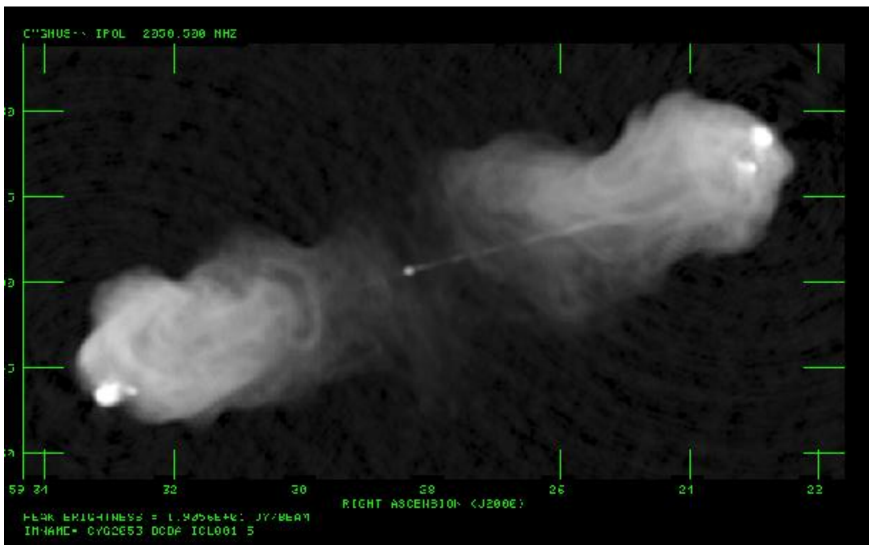


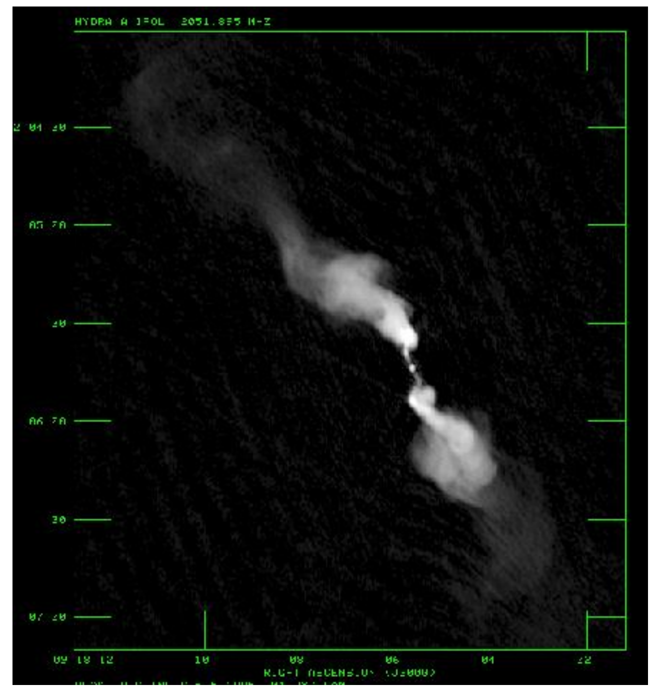


VLA in `D' configuration.

What am I doing these days?

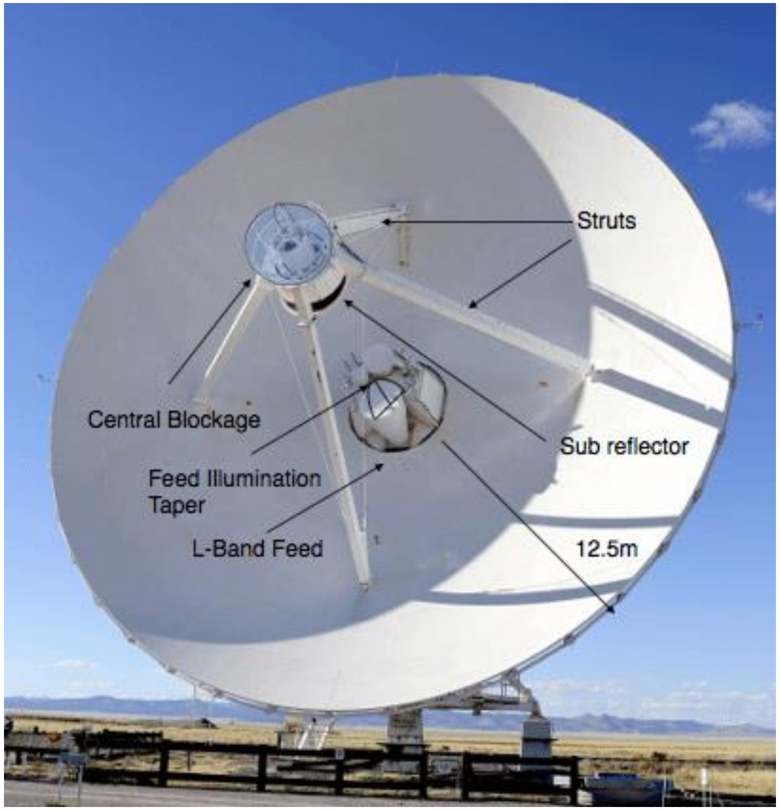
- On the science side:
  - Imaging (particularly polarimetry) of radio galaxies and quasars. Of current interest: Cygnus A and Hydra A, both of which are imbedded in dense, ionized, magnetized gas.





What am I doing these days?

- On the technical (instrumental) side:
  - Improving polarimetry with the VLA:
    - Implementing polarimetry using linearly polarized feeds on the VLA at 327 MHz.
    - Implementing higher-order corrections to improve polarimetric fidelity.
  - Improving high frequency performance (15 – 50 GHz) on the VLA
    - Implementing some corrections to offset the antenna's sag and warping
    - Using holography to correct surface deformations



Topics

- **Caveat:** The upcoming concepts are not difficult – but they are unfamiliar. Hang in there – the concepts get easier with familiarity
- Why do Interferometry?
- The Basic Interferometer
  - Simplifying Assumptions
  - Sine and Cosine Fringes
  - Response to Point and Extended Emission
  - The Complex Correlator and Visibilities
  - Relation of Visibilities to Image Structure
- Array Design Issues
  - Geometry, U-V Coverage, Array Design
  - Examples of Visibility Plots
  - Basics of Deconvolution

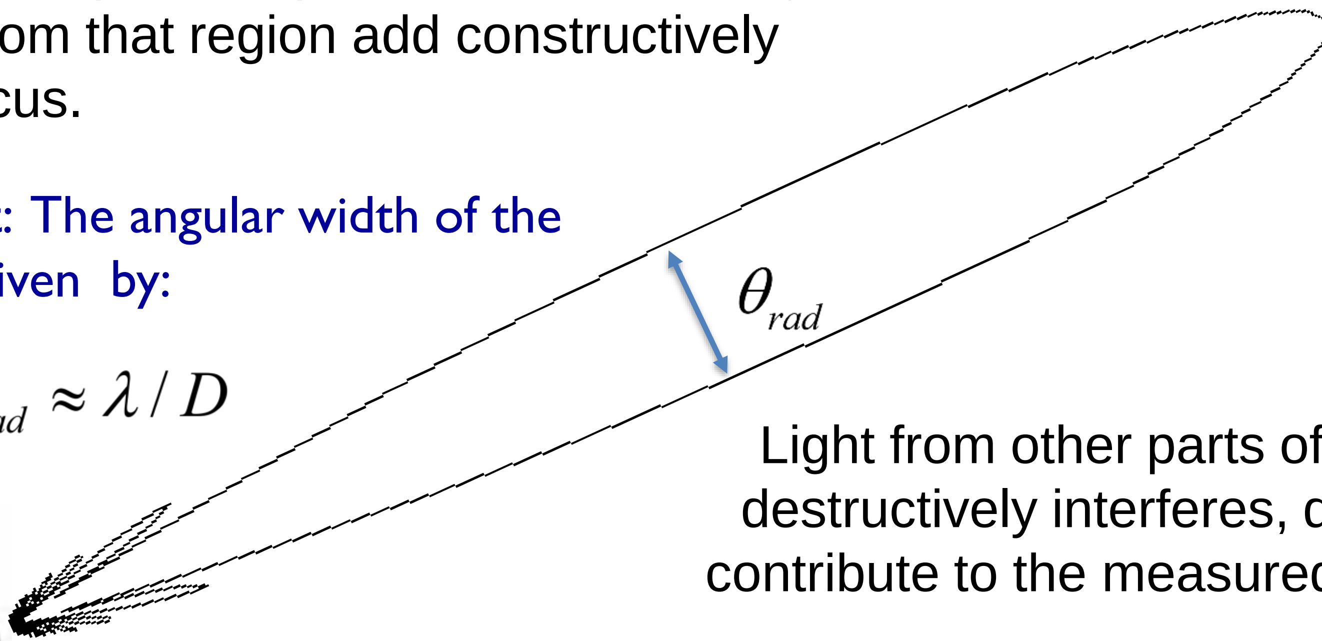
Our Goal – Mapping the Sky

- To do astronomy, we want to know the distribution of brightness $I(\nu,\mathbf{s})$ on the sky, as a function of position (**s**) frequency (ν), polarization (p), and time (t).
- To do this, we need an instrument – a telescope – which has both high spatial resolution and sensitivity (as well as polarization purity and frequency resolution).
- At the 'dawn of radio astronomy' (roughly, the 1940s and 50s), a **single dish** was sufficient.
- These are simple devices (conceptually) – they collect radiation from a small region of sky, and convert it to a signal (on a wire, say) for us to analyze.

Single Dish Antennas

Individual antennas are sensitive to emission from some small region of sky. This is called the **Primary Beam**.

Dishes focus (interfere) EM radiation so only sources from that region add constructively at their focus.



Key point: The angular width of the beam is given by:

$$\theta_{rad} \approx \lambda / D$$

Light from other parts of the sky destructively interferes, does not contribute to the measured signal.

Some new dishes employ Focal Plane Arrays, where many receivers are arranged near the focus, each seeing signals arriving from different angles.

High Resolution Needs Interferometry

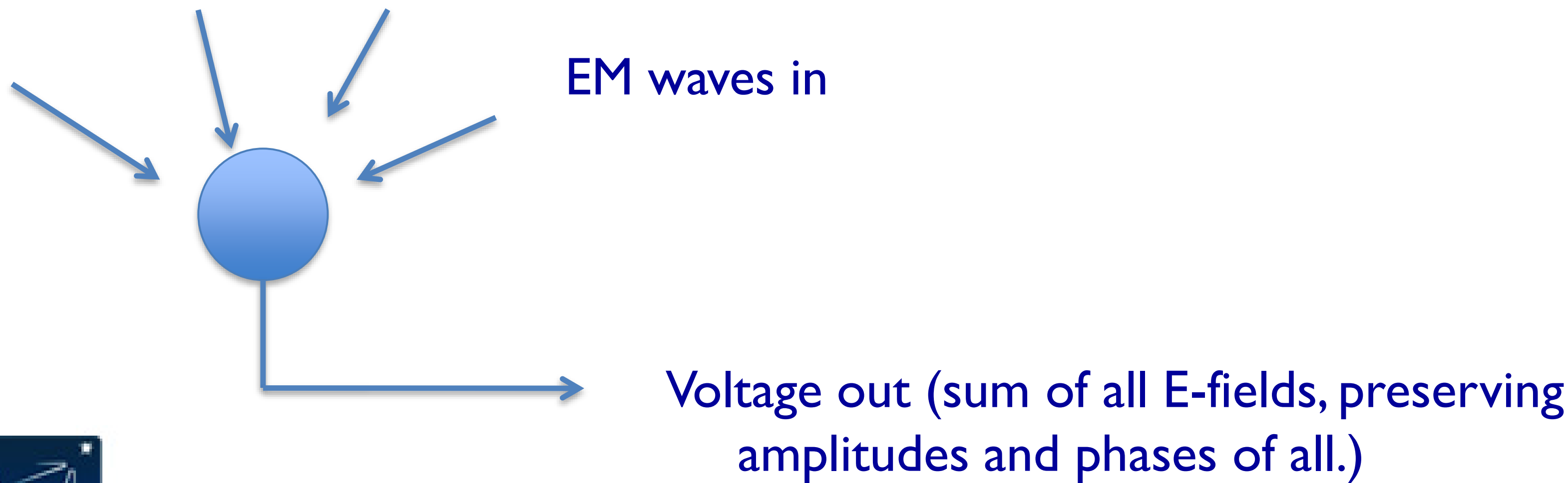
- Radio telescopes coherently sum electric fields over an aperture of size D. For this, diffraction theory applies – the angular resolution is:

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

- To obtain 1 arcsecond resolution (dime at a distance of ~3 km) at a wavelength of 21 cm, we require an aperture of ~42 km!
- It should be obvious that we're not doing that with a single dish.
- Interferometry – the process of utilizing the long spacings between pairs of antennas is the way to go.
- The technique of synthesizing a larger aperture through combinations of separated pairs of antennas is called 'aperture synthesis'.
- Developed in the 1950s and 1960s.

The Role of the Sensor (aka Antenna)

- 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.

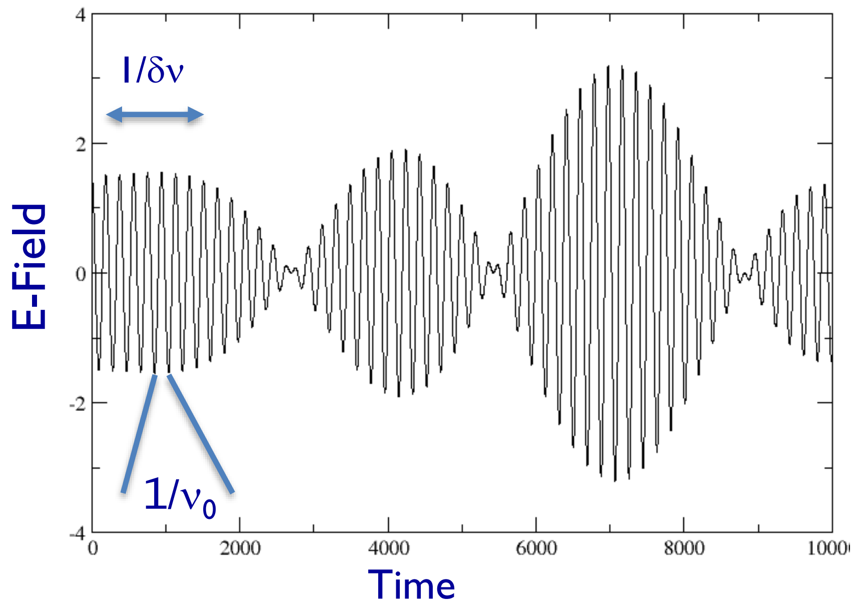


Quasi-Monochromatic Radiation

- 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.

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

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'.



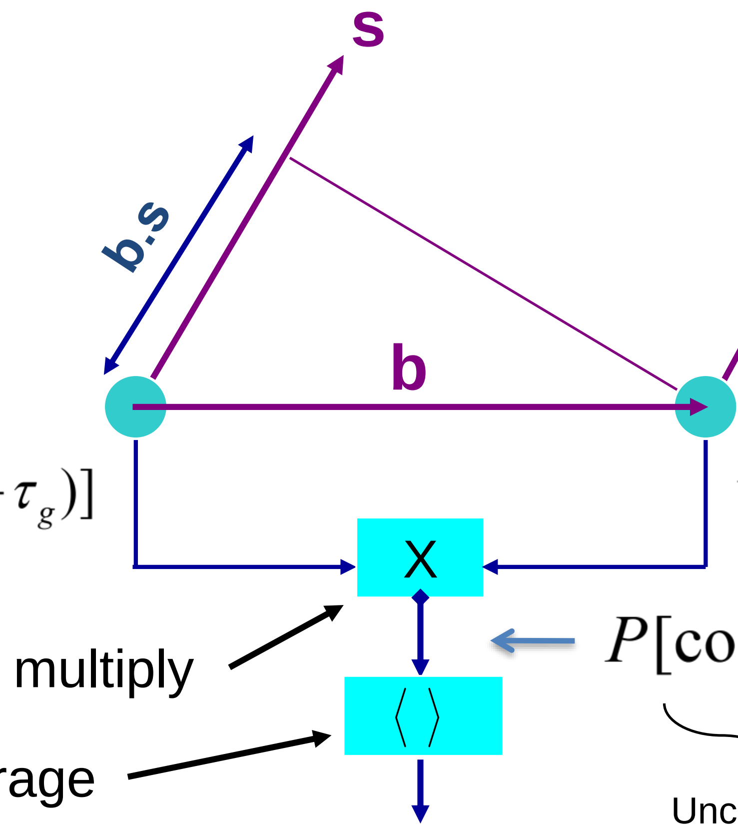
Simplifying Assumptions

- 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
  - 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

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



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

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

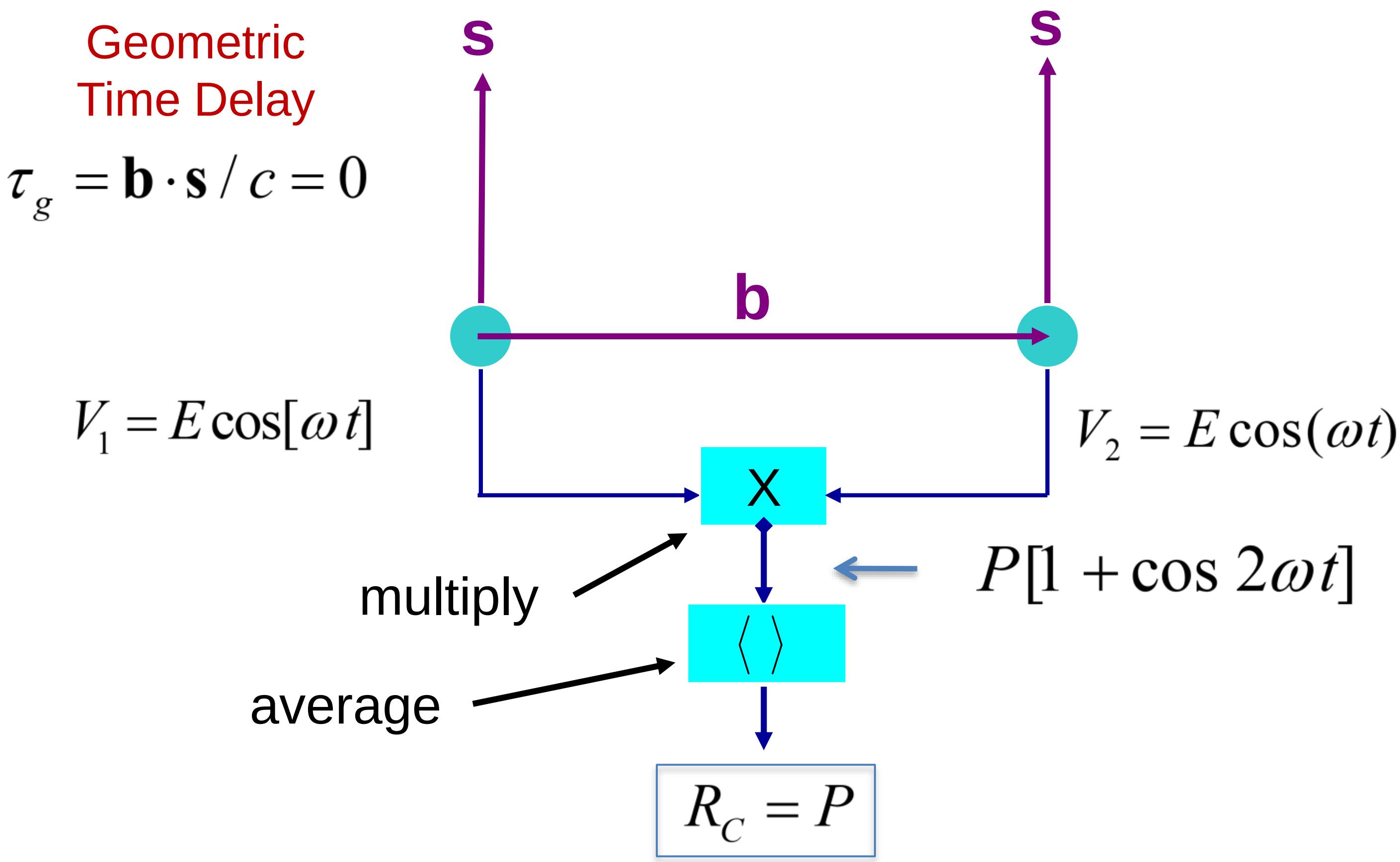
The path lengths
from sensors
to multiplier are
assumed equal!

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

$$\boxed{R_C = P\cos(\omega\tau_g)}$$

Three Simple Scenarios: A) Signals in Phase

- Consider the case when radiation come from 'straight up', (or delayed by n wavelengths)

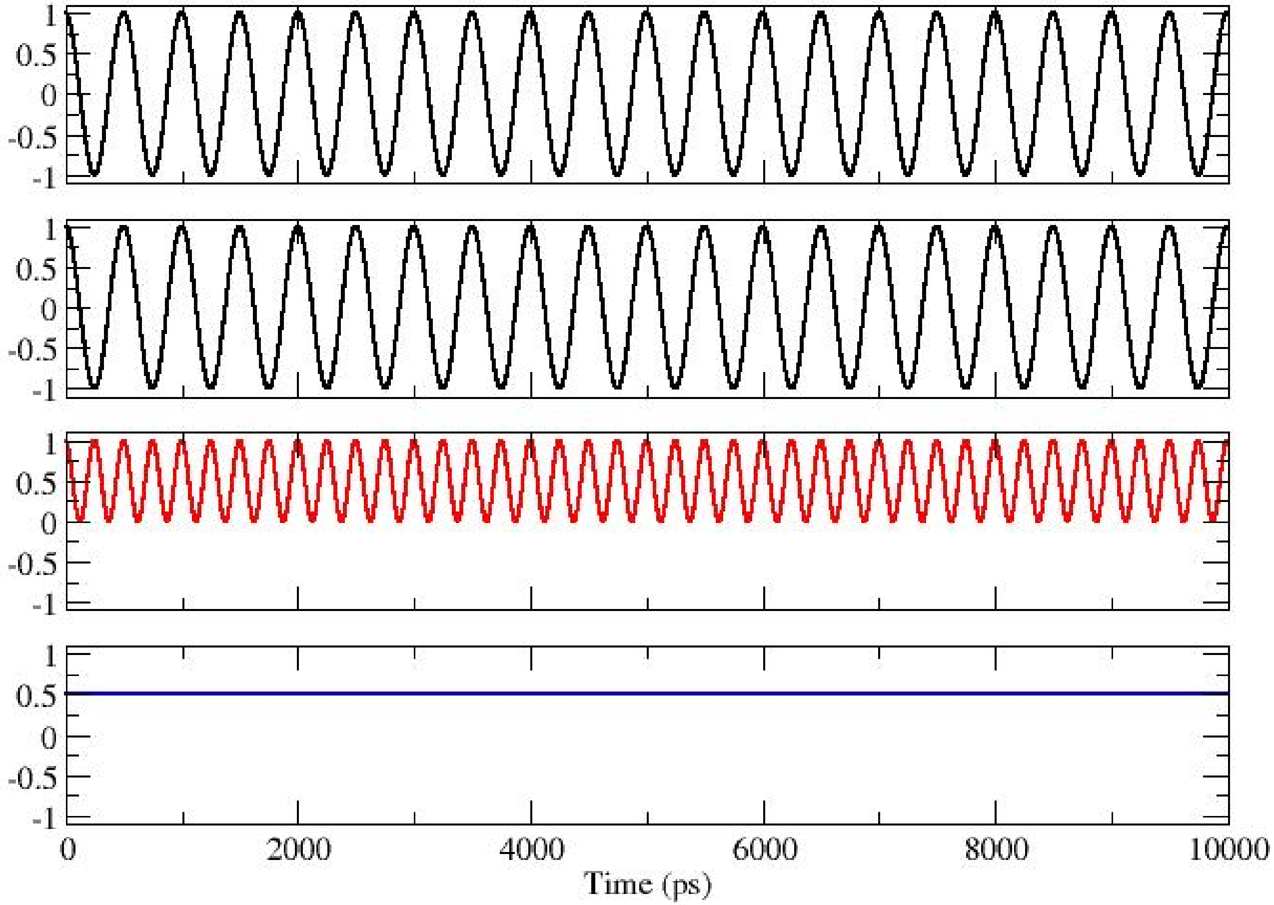


Pictorial Example: Signals In Phase

2 GHz Frequency, with voltages in phase:

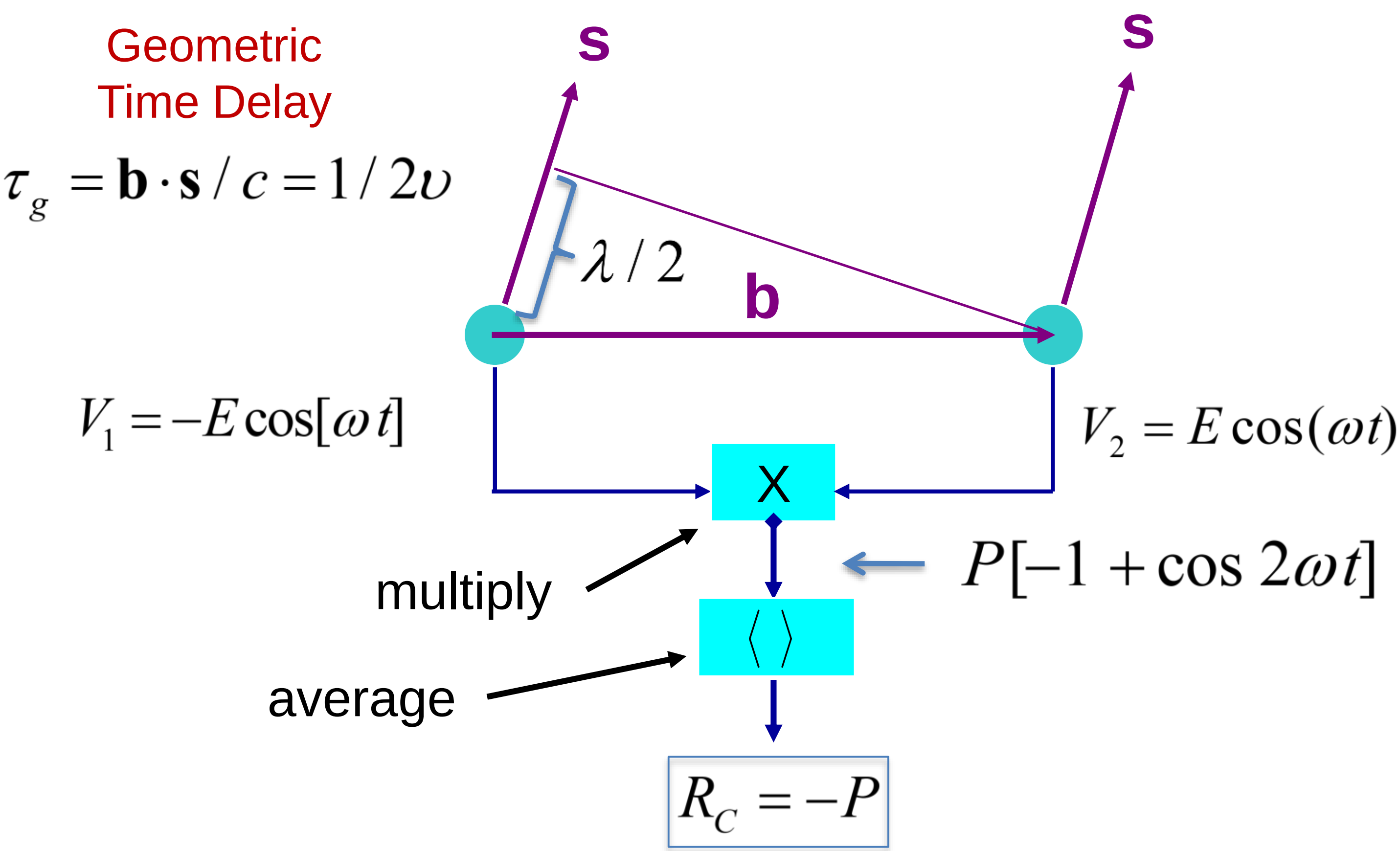
b.s = nλ, or τ_g = n/ν

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



Three Simple Scenarios:
B) Signals out of Phase

Suppose the source is at a slight angle, with b.s = $\lambda/2$. (or $3\lambda/2$, or $5\lambda/2$ or ….)

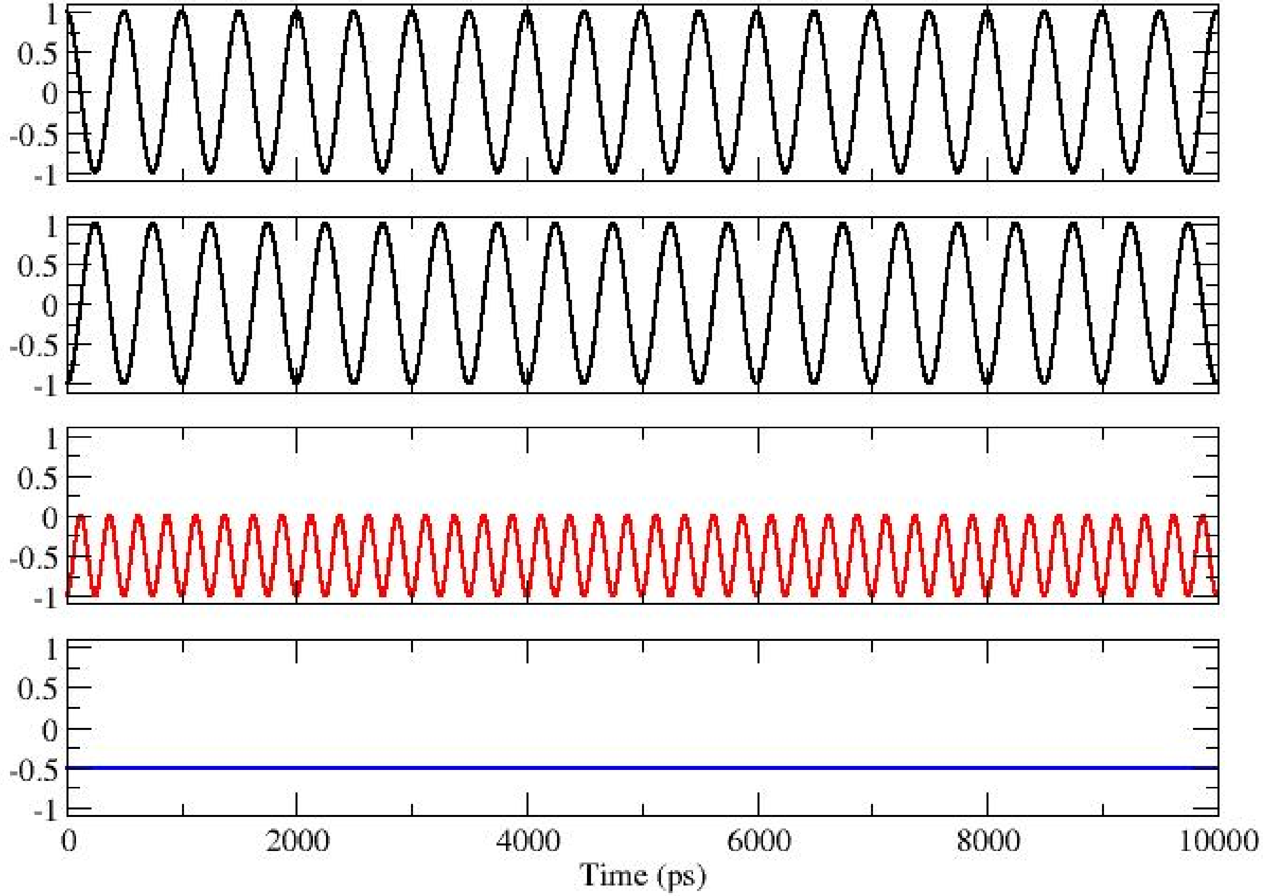


Pictorial Example: Signals out of Phase

2 GHz Frequency, with voltages out of phase:

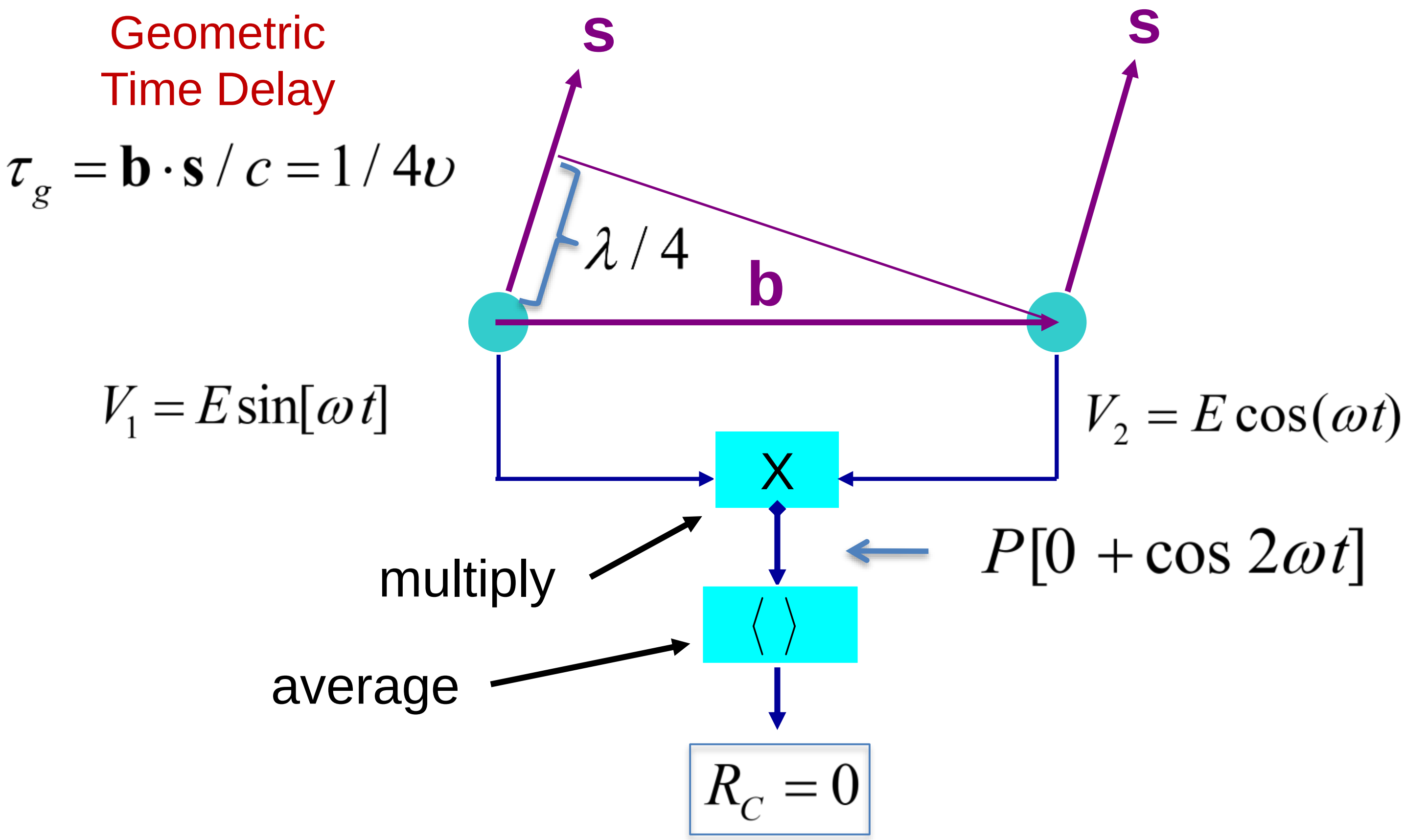
$b.s=(n \pm \tfrac{1}{2})\lambda \qquad \tau_g = (2n \pm 1)/2\nu$

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



Three Simple Scenarios: C) Signals in Phase Quadrature

Suppose the source is at a different slight angle, with b.s = λ/4 (or 3λ/4, or …)

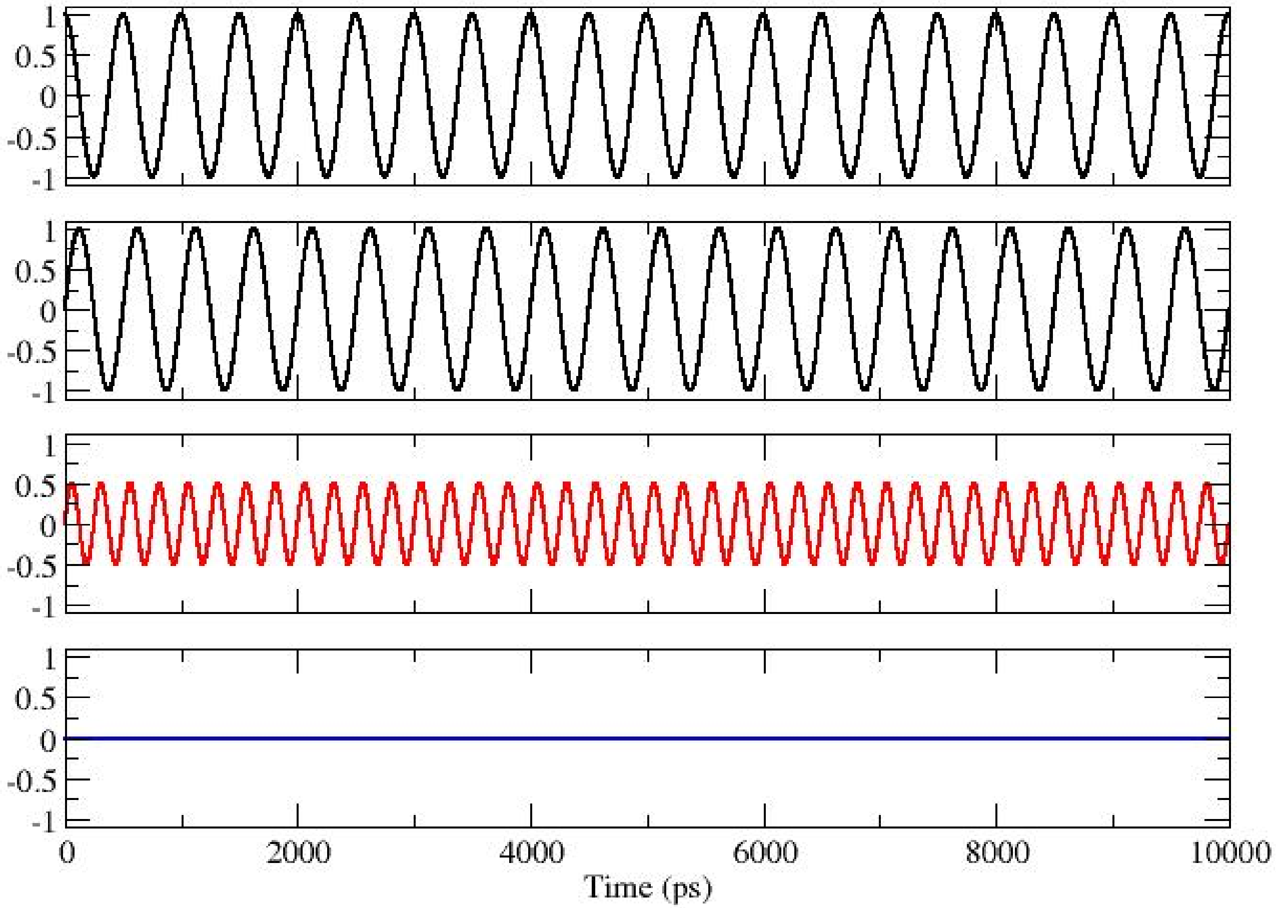


Pictorial Example: Signals in Quad Phase

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

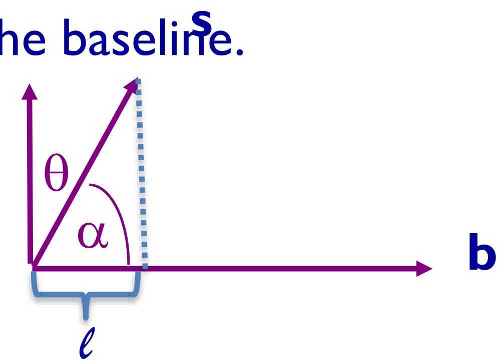


Some General Comments

- 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:

$$R_C = P\cos(2\pi\nu\tau_g) = P\cos\left(2\pi\,\frac{\mathbf{b}\cdot\mathbf{s}}{\lambda}\right) = P\cos(2\pi u\cos\alpha) = P\cos(2\pi u l)$$

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



- 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 'Cosine' Interferometer Response (to a point source)

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

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

- For u = 10

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

- For u = 25

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

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

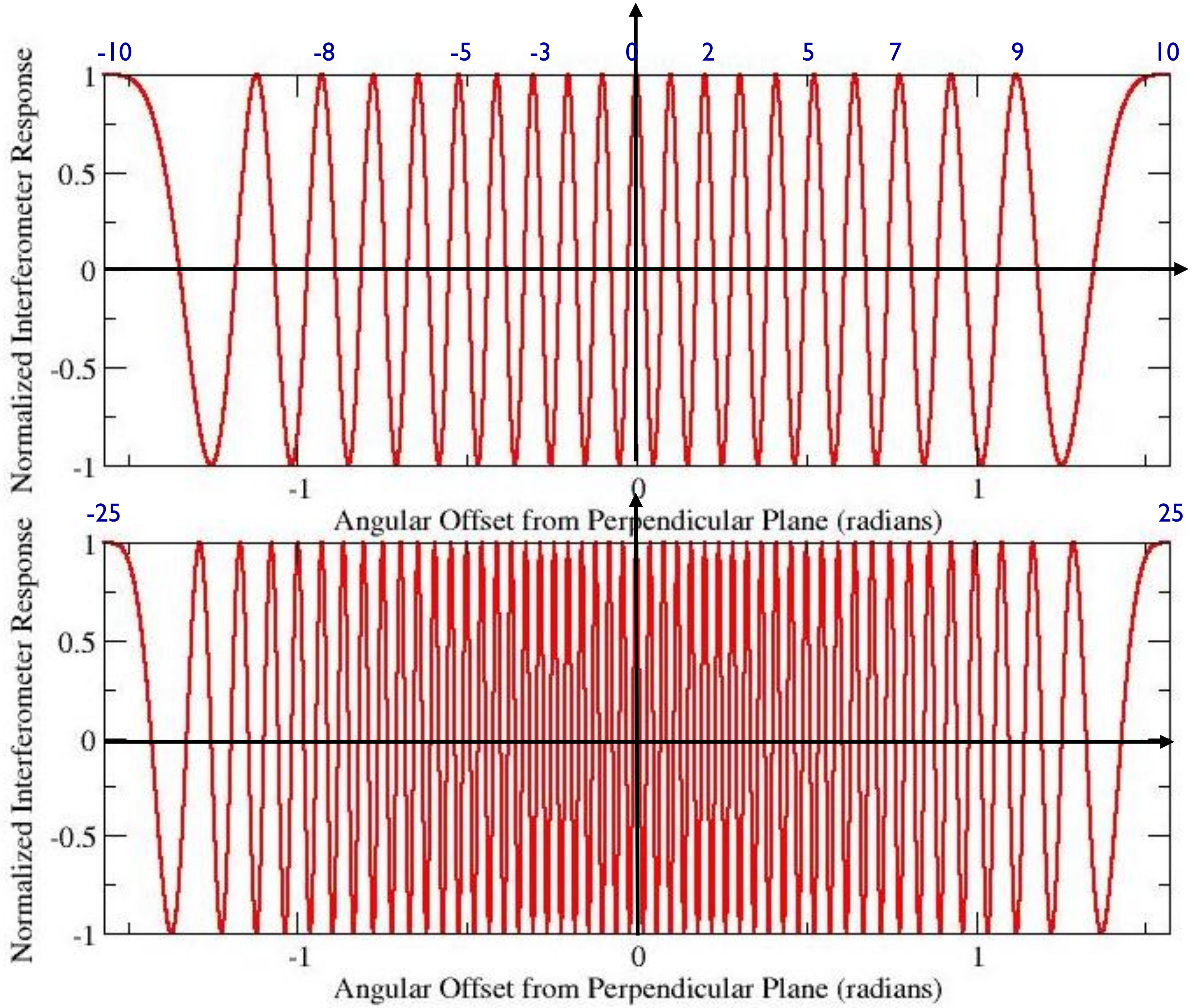
Whole-Sky Response

- Top:
  u = 10

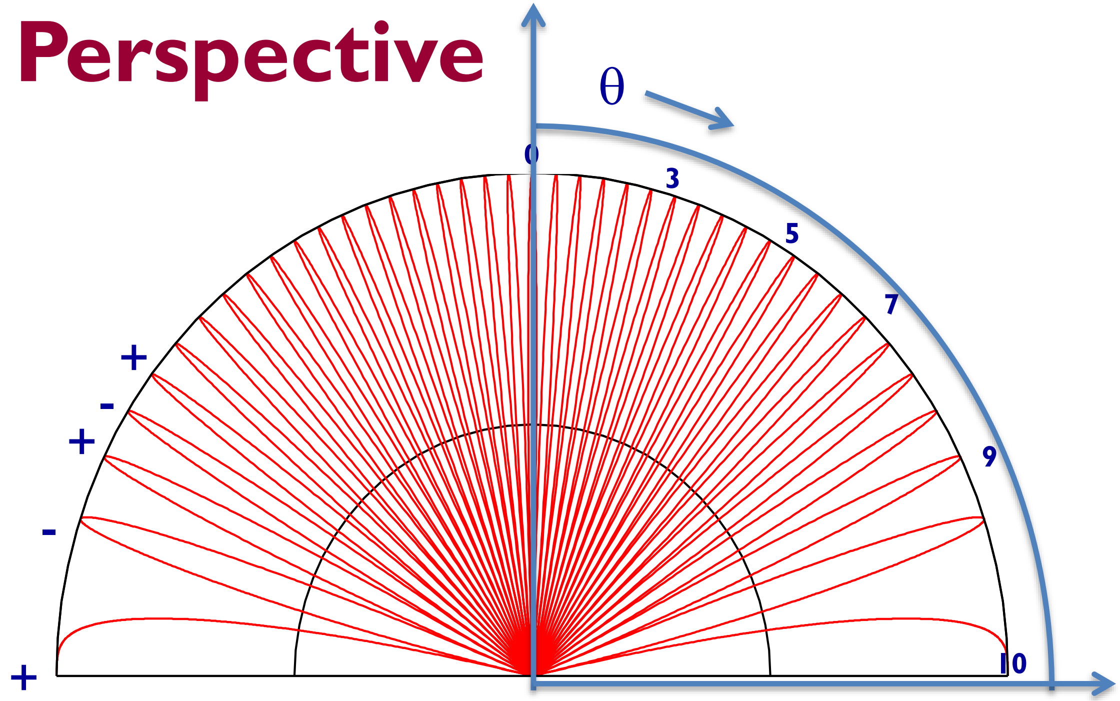
There are 21 fringe maxima, and 20 fringe minima over the hemisphere.

- Bottom:
  u = 25

There are 51 fringe maxima over the hemisphere



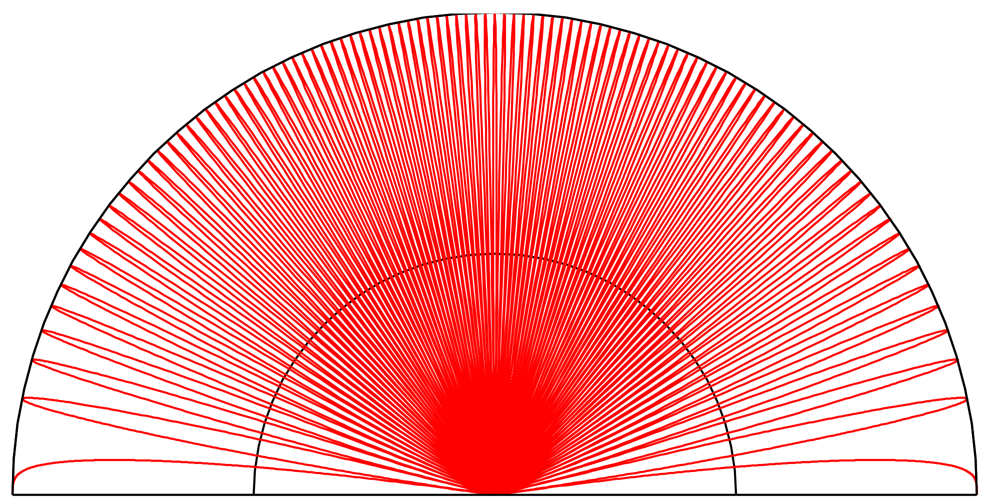
From an Angular Perspective



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

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

Hemispheric Pattern

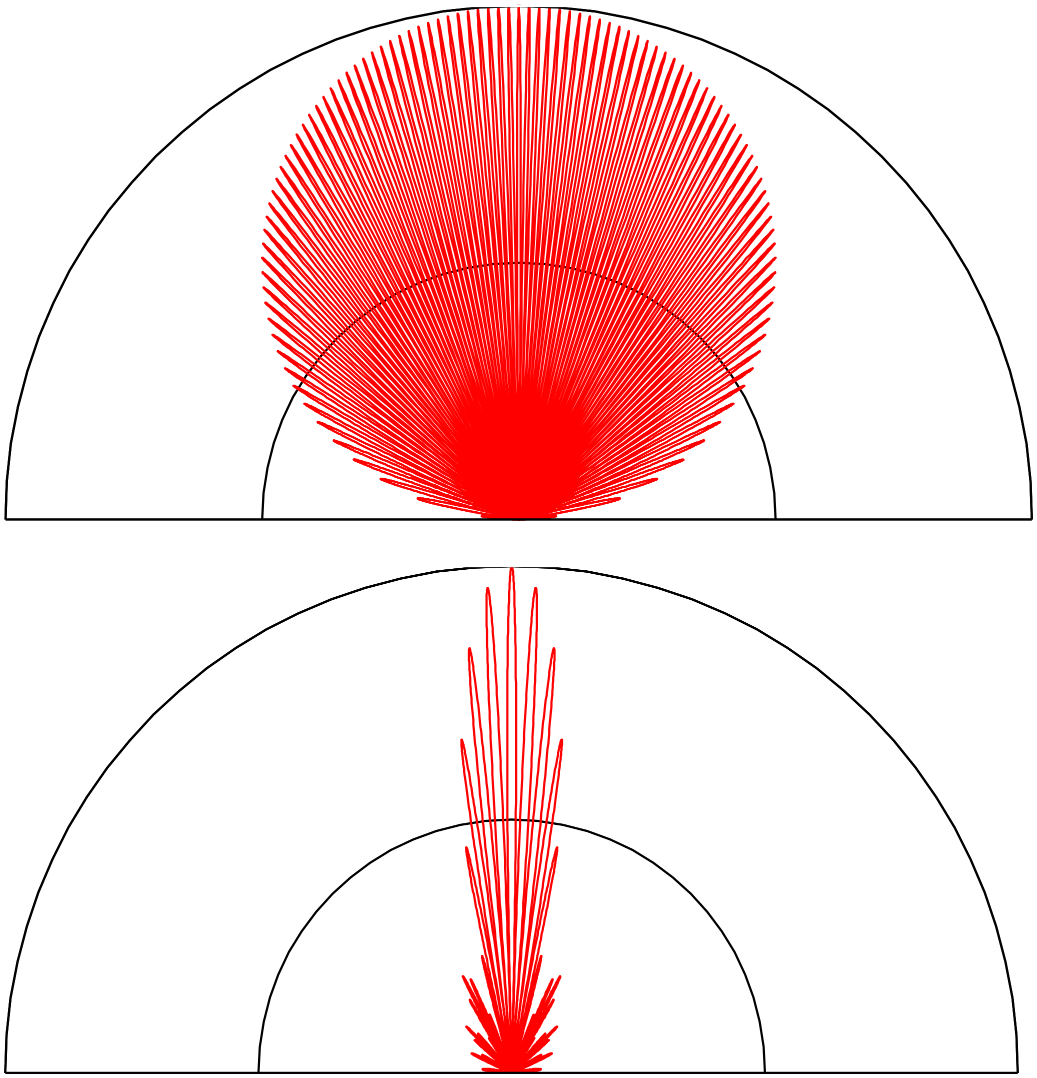
- The preceding plot is a meridional cut through the hemisphere, oriented along the baseline vector.
- In the two-dimensional space, the fringe pattern consists of a series of coaxial cones, oriented along the baseline vector.
- The figure is a two-dimensional representation when u = 4.
- As viewed along 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 ('gain') but narrow field of view --very useful for some applications.
- Small sensors have low directivity, and large field of view – useful for other 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.



Response for a Point Source

- When our (stationary) interferometer observes a (stationary) point source, what we measure is simply the value of the cosinusoid pattern in the direction of the source, multiplied by the flux of that source.
- If the source is moving, the response (over time) is a nice simple cosinusoid – looks like the preceding figures.
- But point sources are boring. We build interferometers to **resolve** objects.
- So how does our basic interferometer respond to an extended source?
  - `Extended' means angular size of source larger than fringe separation.

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

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

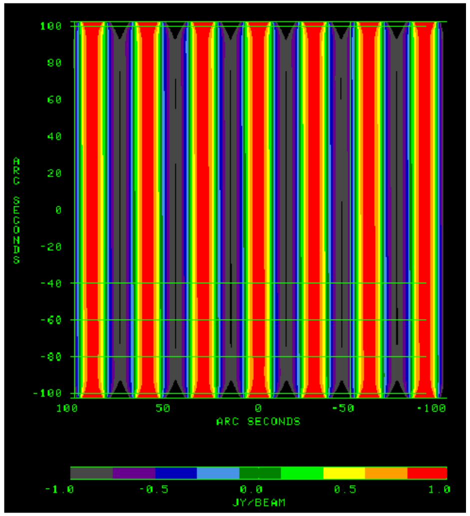
- The response is the integral of the brightness modulated by the cosinusoidal interferometer pattern.
- This expression links what we want: the brightness on the sky, $I_\nu(\mathbf{s})$, to something we can measure - R_C, the interferometer response.
- How can we recover $I_\nu(\mathbf{s})$ from observations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

A picture is worth 1000 words …

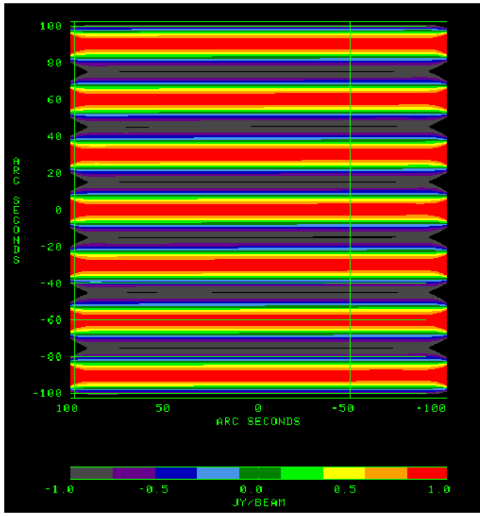
- As stated earlier, these concepts are not difficult, but are unfamiliar. We need to think in new ways, to get a deeper understanding of how all this works.
- To aid, I have generated images of interferometer fringes, of various baseline lengths and orientations.
- I then 'observe' a real source (Cygnus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

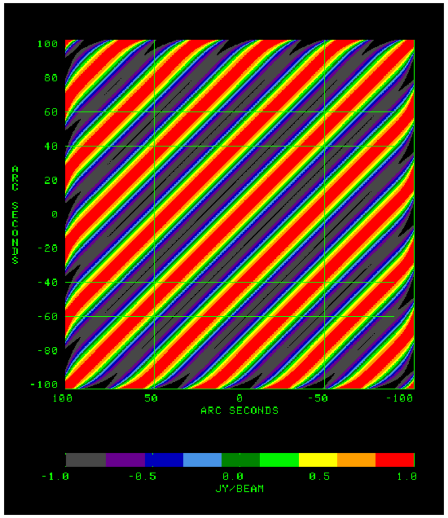
- The fringe separation given by baseline length in wavelengths, the orientation given by the orientation of the baseline.



East-West baseline makes vertical fringes



North-South baseline makes horizontal fringes



Rotated baseline makes rotated fringes

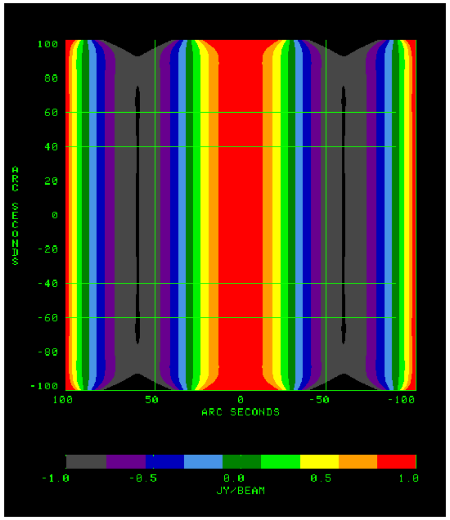
- Fringe angular spacing given by baseline length in wavelengths:

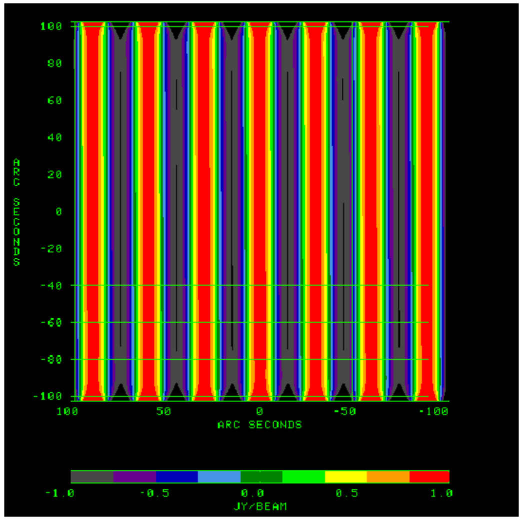
$$\Delta\theta = \lambda / B = 30.2''$$

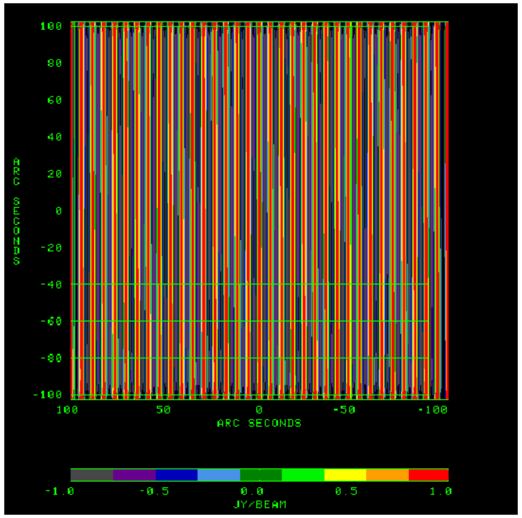
Longer Baselines => Smaller Fringes

Longer baselines make finer fringes and shorter baselines make fatter fringes.

A zero baseline makes no fringe at all…







250 meter baseline
120 arcsecond fringe

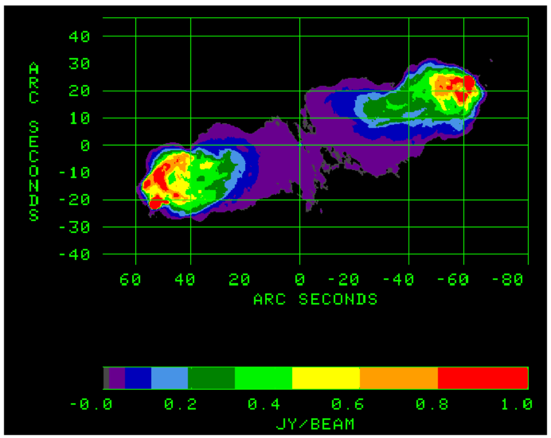
1000 meter baseline
30 arcsecond fringe

5000 meter baseline
6 arcsecond fringe

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

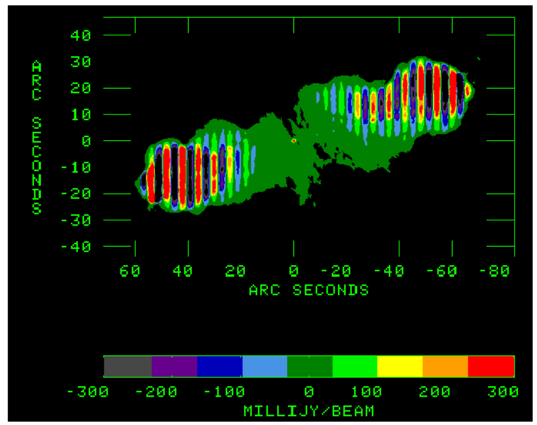
For a Real Source (Cygnus A = 3C405)

- Cygnus A is a powerful, nearby radio galaxy of 130 arcseconds extent.
- The left panel shows the actual brightness (color coded).
- The other two panels show how the 5km-baseline interferometer 'sees' it



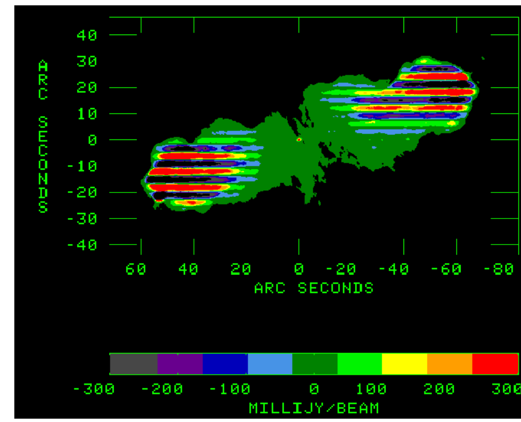
Zero-Spacing Image

Sum = 999 Jy



5 km EW spacing

Sum = 61 Jy



5 km NS spacing

Sum = -16 Jy

Remember: the interferometer doesn't 'see' anything – it integrates (sums) the brightness modified by 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 distribution?
- Time for some mathematics…. Starting with a seeming digression about odd and even functions.
- (All will be clear shortly…)

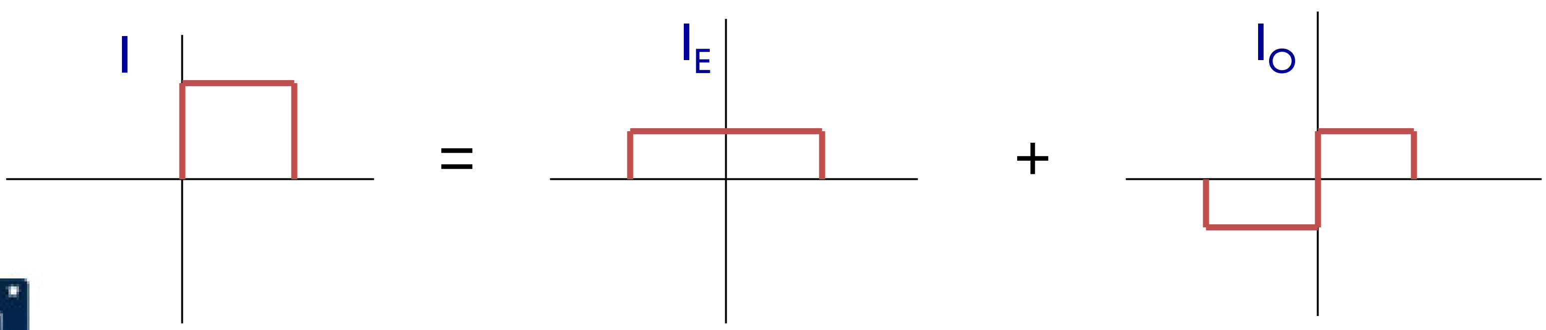
A Short Mathematics Digression – Odd and Even Functions

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

$$I(x,y) = I_E(x,y) + I_O(x,y)$$

A symmetric part: $I_E(x,y) = \dfrac{I(x,y) + I(-x,-y)}{2} = I_E(-x,-y)$

An antisymmetric part: $I_O(x,y) = \dfrac{I(x,y) - I(-x,-y)}{2} = -I_O(-x,-y)$



The Cosine Correlator is Blind to Odd Structure

- The correlator response, R_c (one dimensional for simplicity)…

$$R_C = \iint I_\nu(l)\cos(2\pi ul)\,dl$$

is not enough to recover the correct brightness. Why?

- Separate the source brightness into even and odd parts:

$$I = I_E + I_O$$

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

$$R_C = \iint I(l)\cos(2\pi ul)\,dl = \iint I_E(l)\cos(2\pi ul)\,dl$$

- The 'Cosine' interferometer only sees the symmetric components.
- 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:

$$R_C = \iint I(l)\cos(2\pi ul)dl = \iint I_E(l)\cos(2\pi ul)dl$$

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

$$R_S = \iint I(l)\sin(2\pi ul)\,dl = \iint I_O(l)\sin(2\pi ul)dl$$

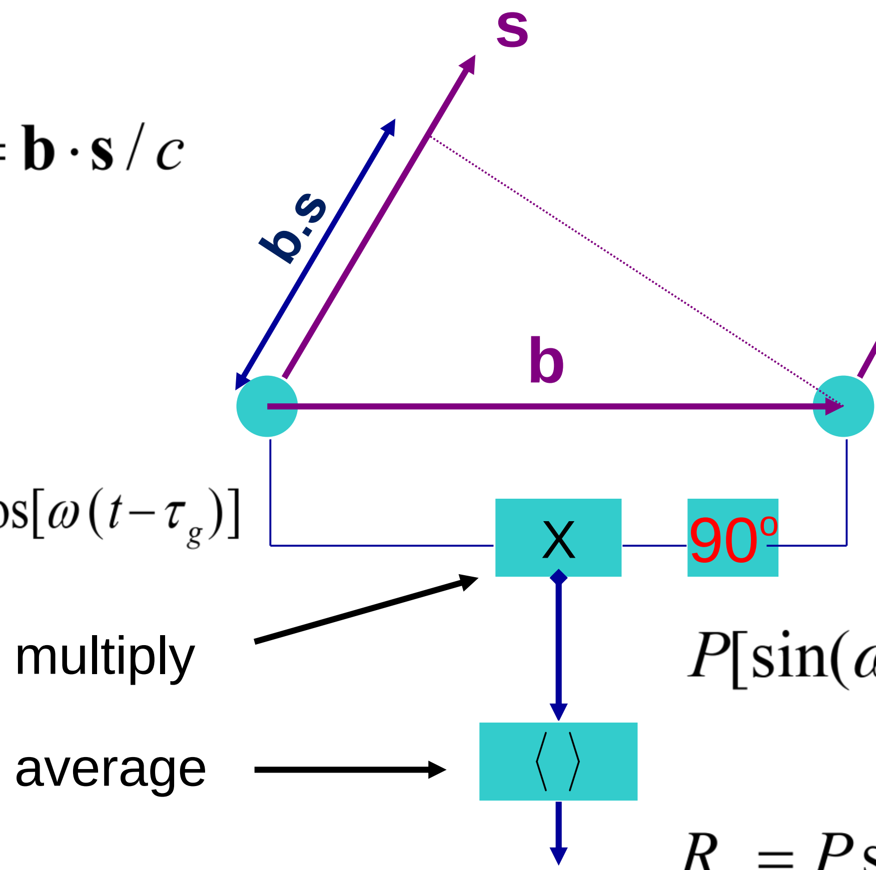
since the integral of an even times an odd function is zero.

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

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

$$A = \sqrt{R_C^2 + R_S^2} \qquad \phi = \tan^{-1}\left\{ R_S / R_C \right\}$$

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

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

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

$$\boxed{I_\upsilon(\mathbf{s}) = \iint V_\nu(u,\mathrm{v})e^{2\pi i \mathbf{b}\cdot\mathbf{s}/\lambda}du d\mathrm{v}}$$

- The problem now is to determine the function V(u,v).

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:

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

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

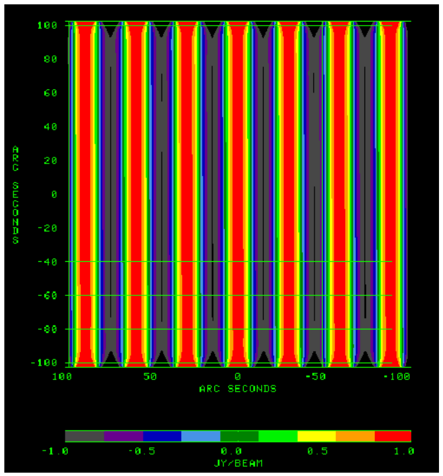
- Then:

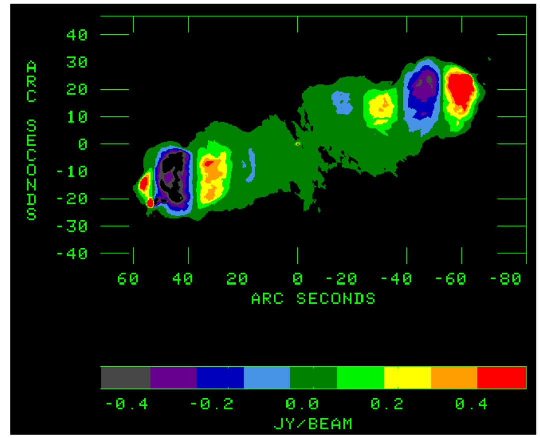
$$P_{corr} = \left\langle V_1 V_2^* \right\rangle = A^2 e^{-i\omega \mathbf{b}\bullet\mathbf{s}/c}$$

Some Pictures, to Illustrate This Point

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

COS



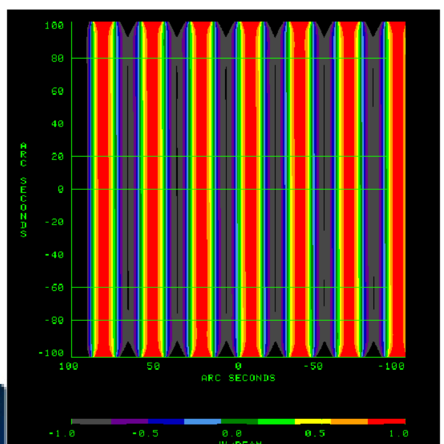


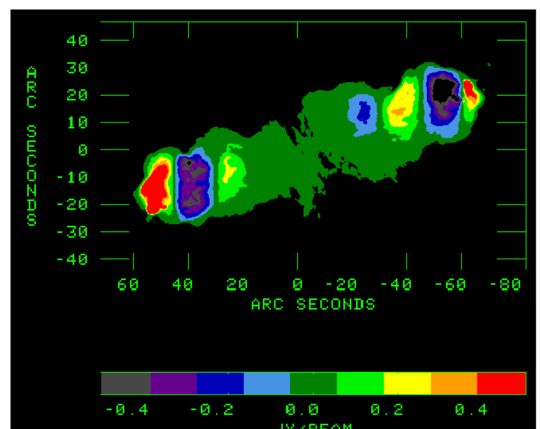
69 Jy

A=103 Jy

ϕ=48

SIN





77 Jy

Interferometer on a railcar …

- To get a better idea of the relation between the visibility and source structure, imagine a 1-dimensional interferometer, one antenna of which is on a railcar.

- Then imagine observing some distant object, while continuously moving the antenna on the rail car.
- As the baseline gets longer, more and more 'fringes' lie across the source. What happens to the visibility?

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:

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

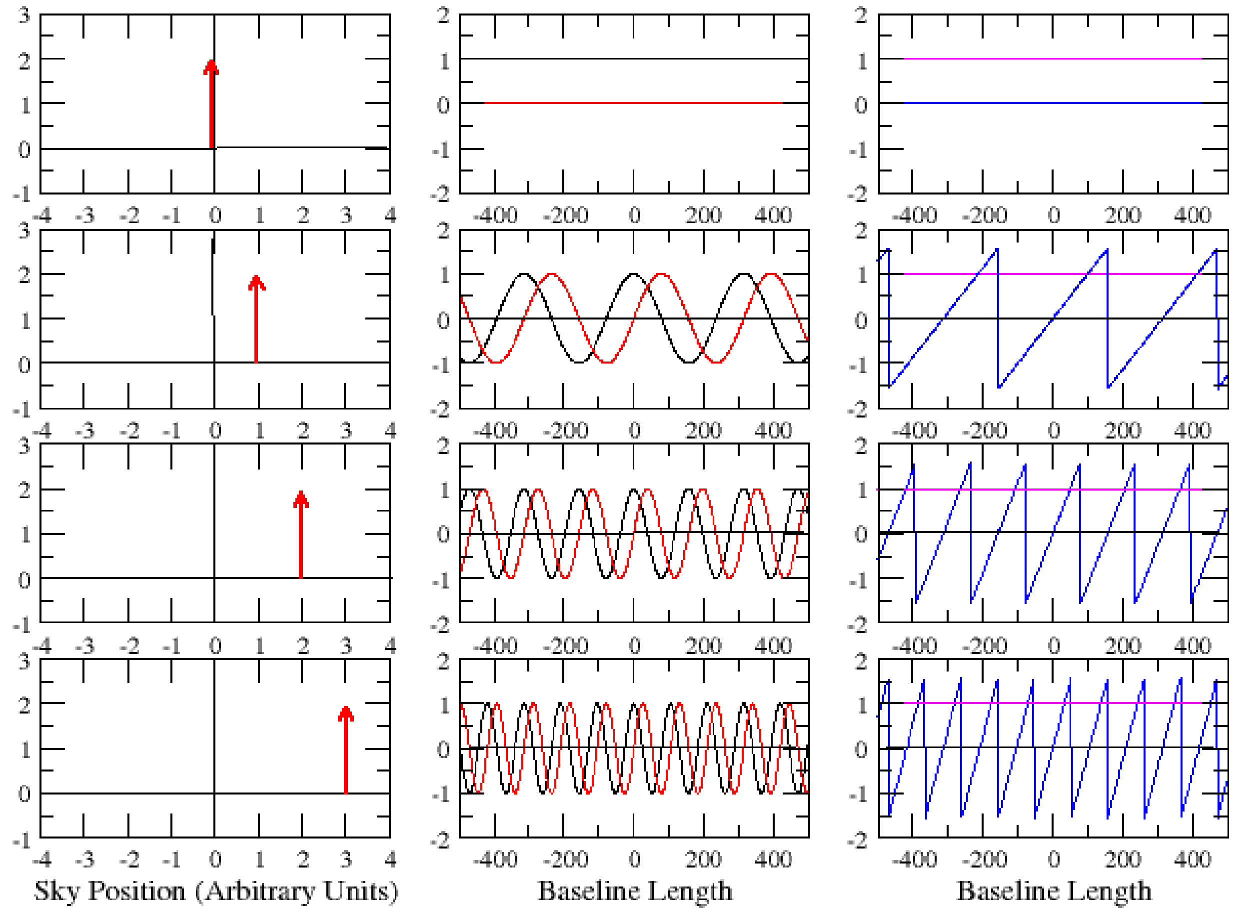
- For a source at the origin ($l_0=0$), $V(u) = 1$. (units of Jy).
- For a source off the origin, the visibility has unit amplitude, and a phase slope with increasing baseline, rotating 360 degrees every l_0^{-1} wavelengths.

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.

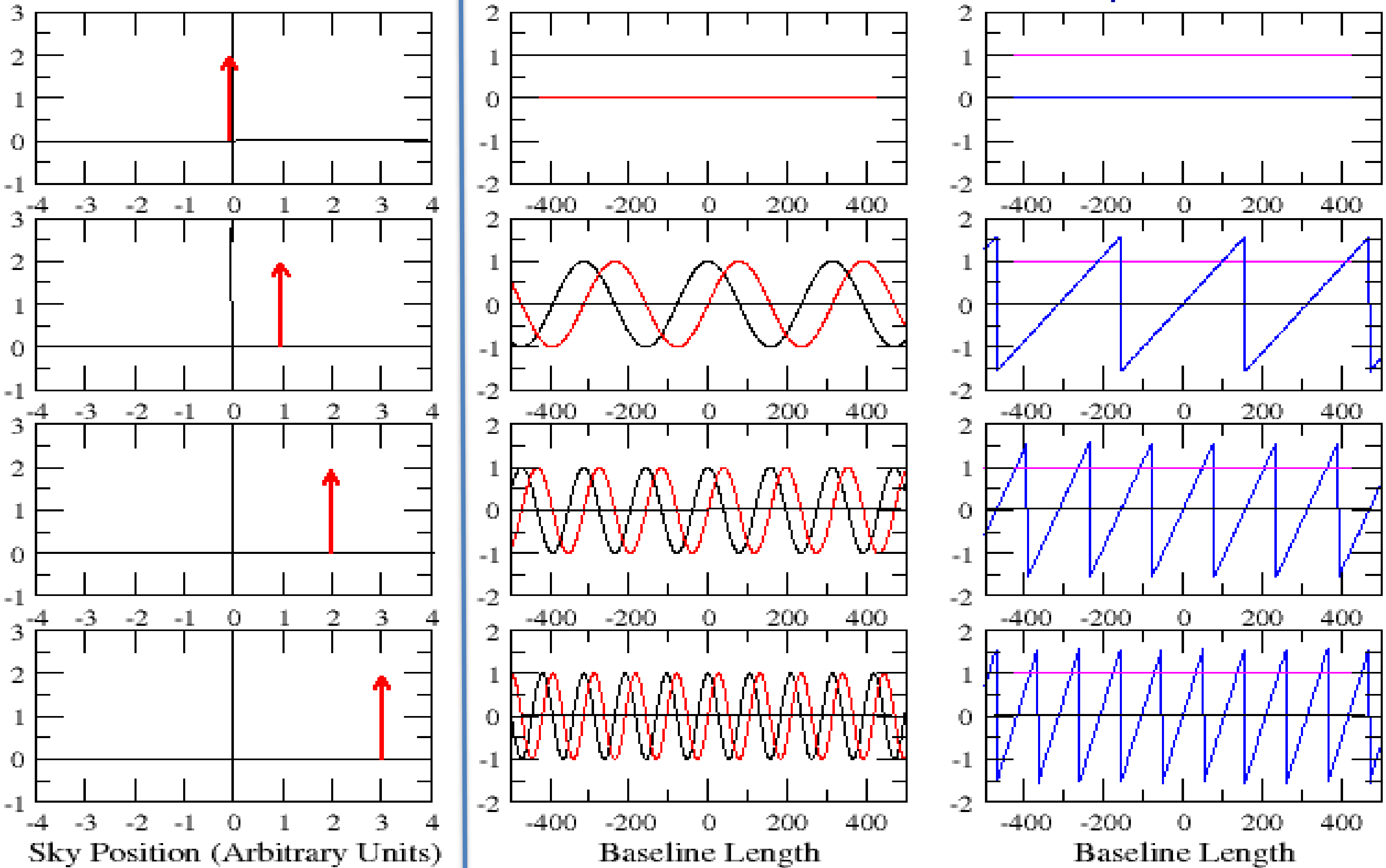


Image

Visibility

Cos and Sin

Amp and Phase

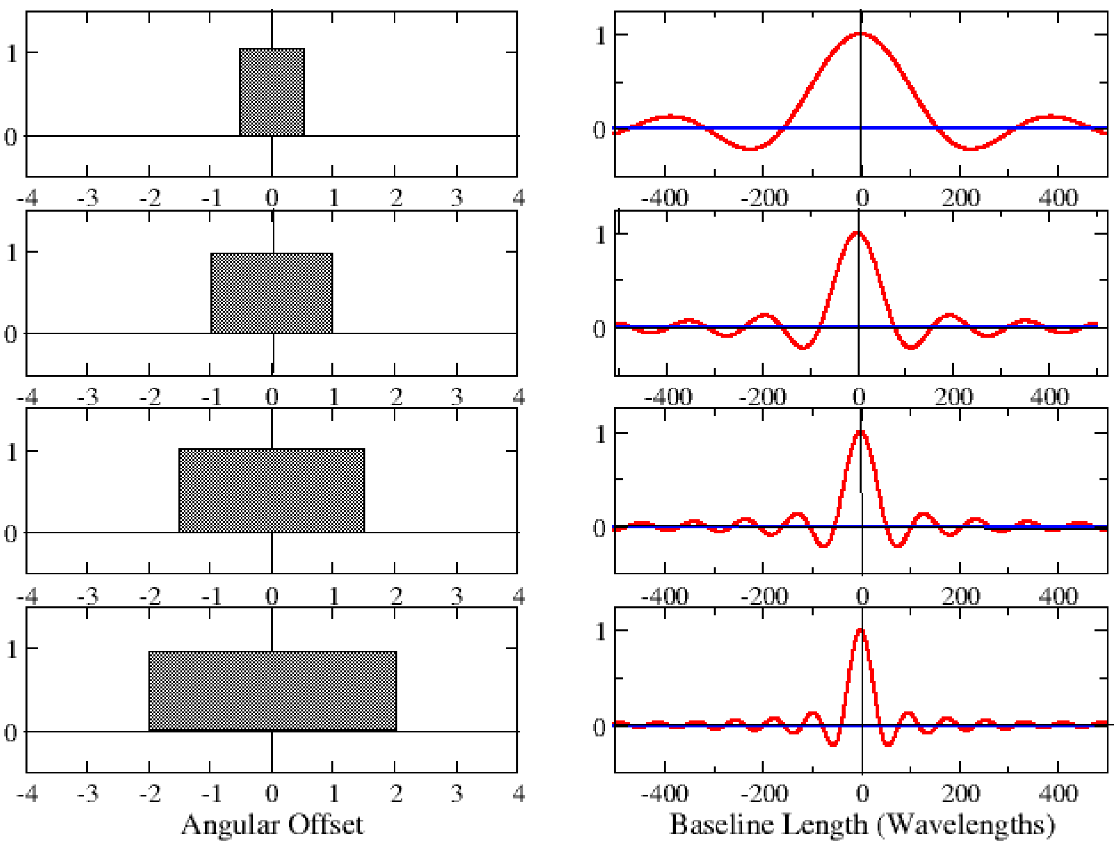


Visibility Example #2: Centered Boxes

- In this example, we have three centered (symmetric) 'box' structures. The symmetric (even) structures 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.

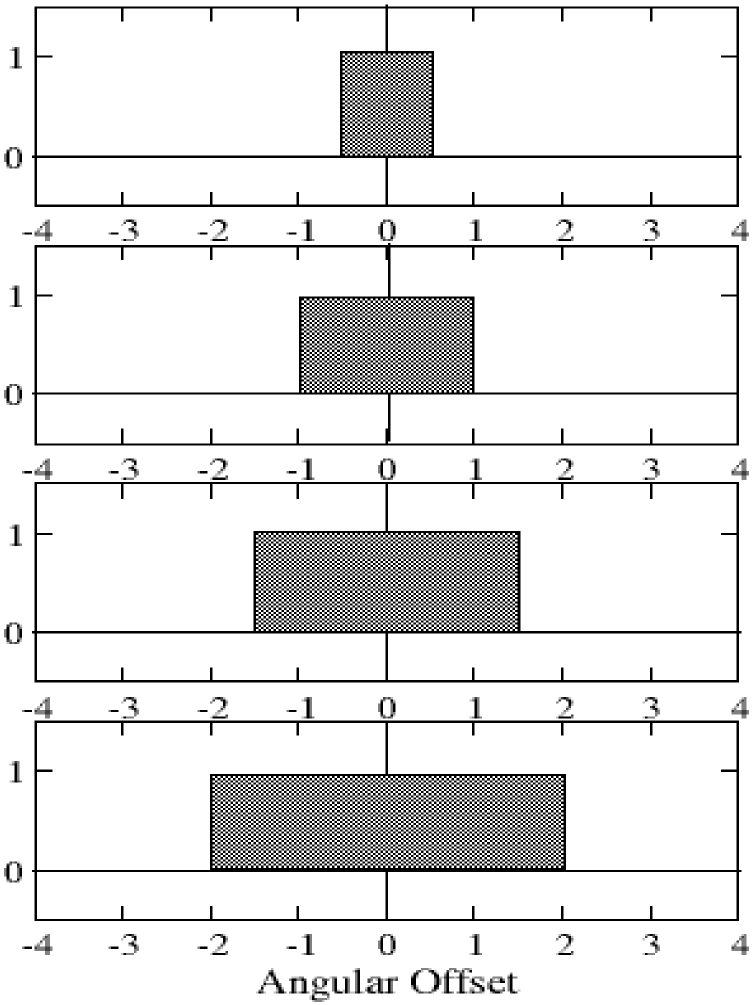


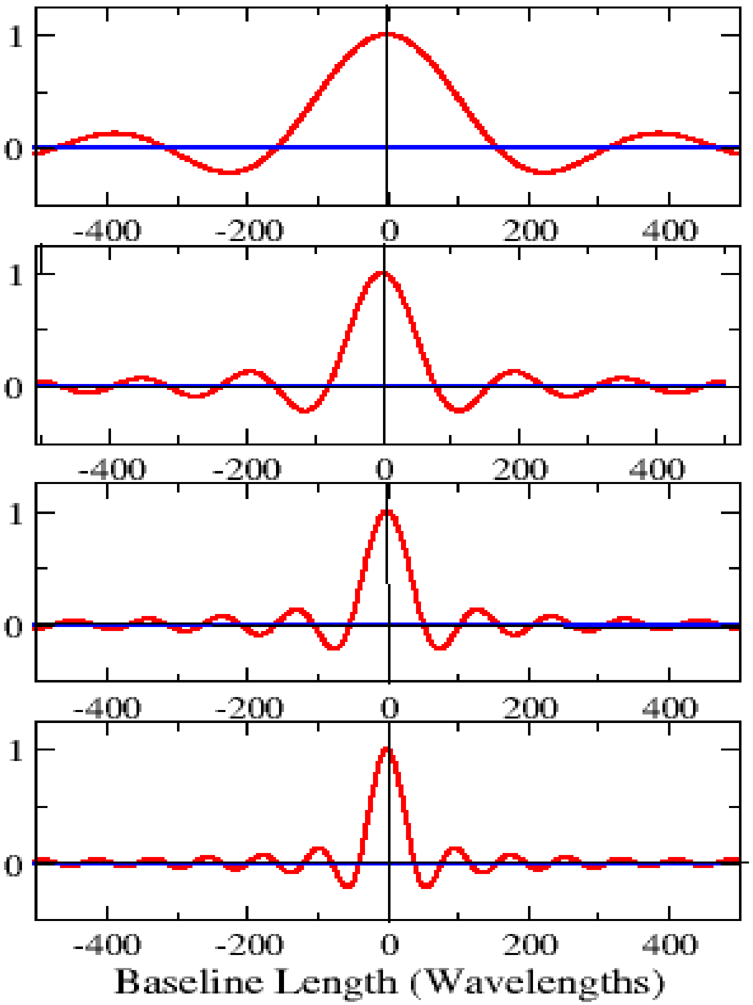
Image

Visibility

Cos and Sin

Amp and Phase





Visibility Example #3: Offset Boxes

- In this example, we show the same box, offset by increasing amounts from the phase center. Middle panel shows **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.

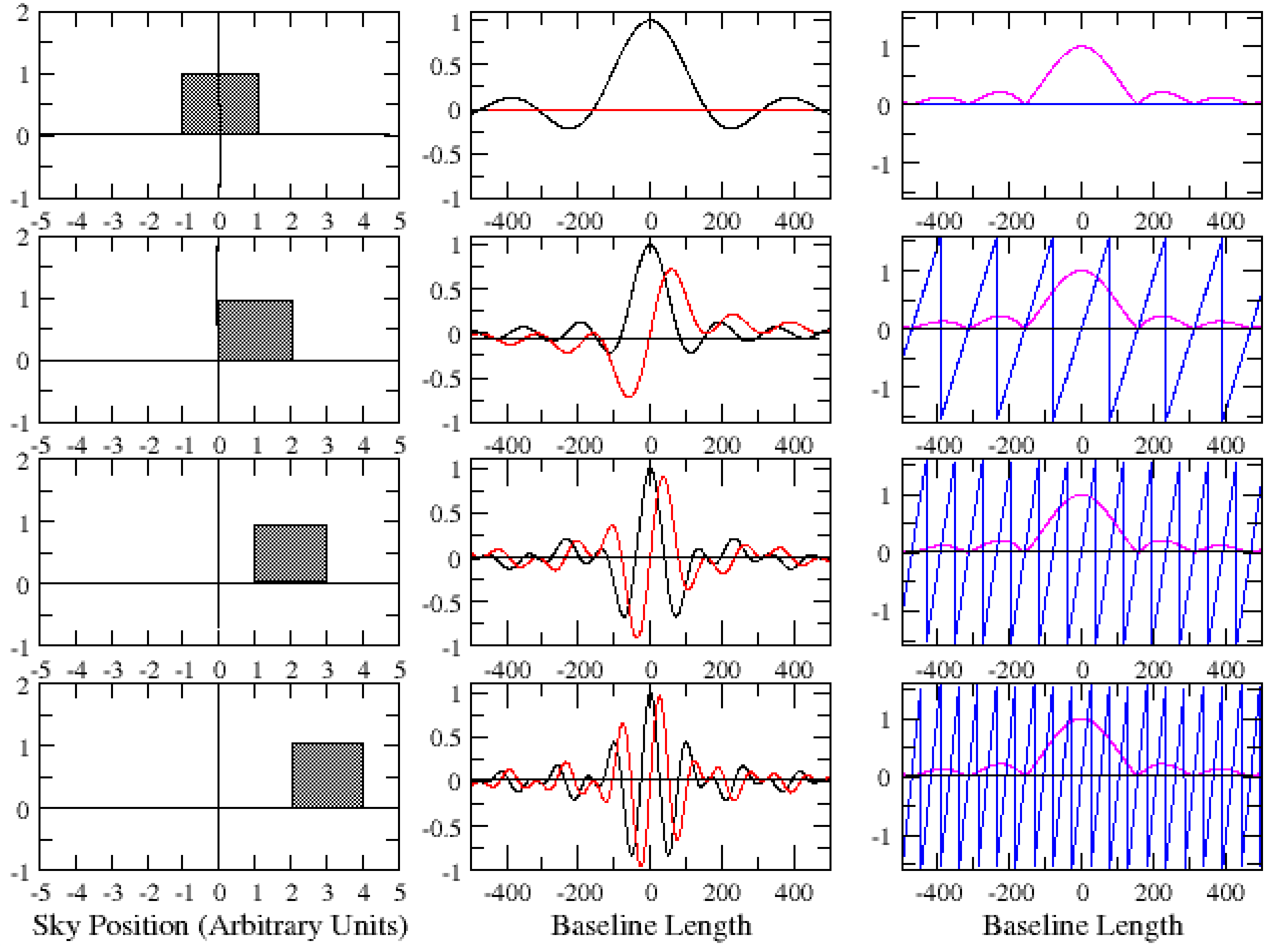
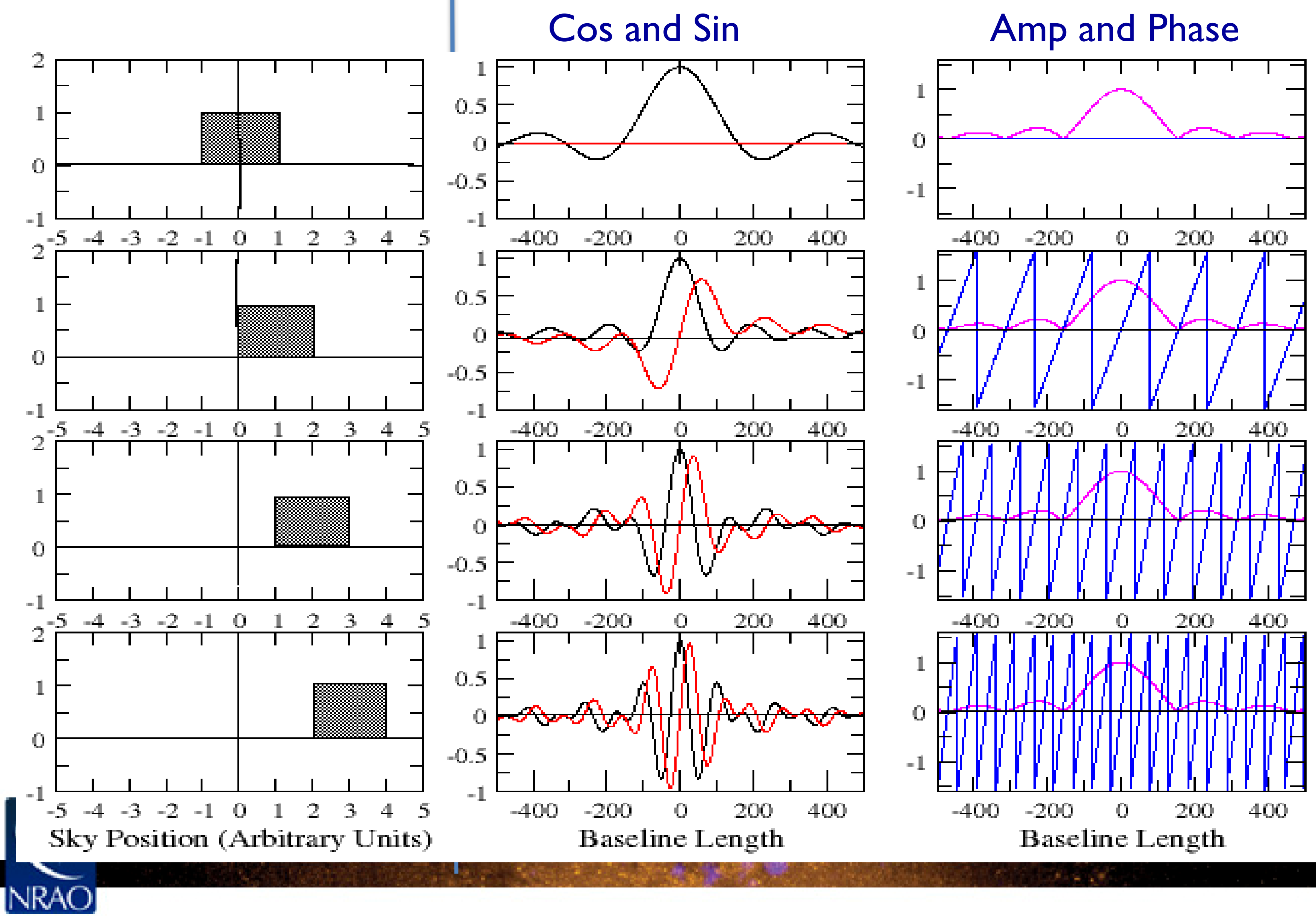
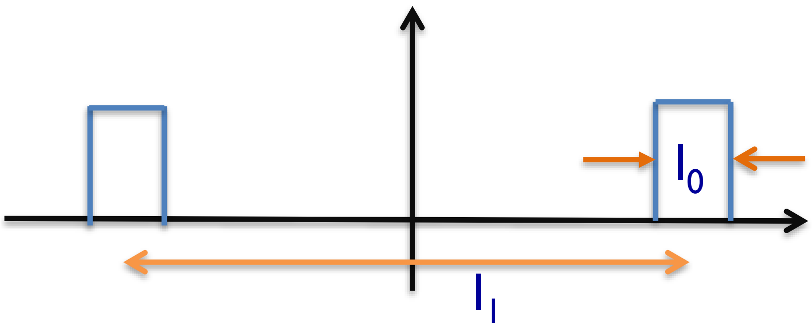


Image | Visibility



Extended Symmetric Doubles

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



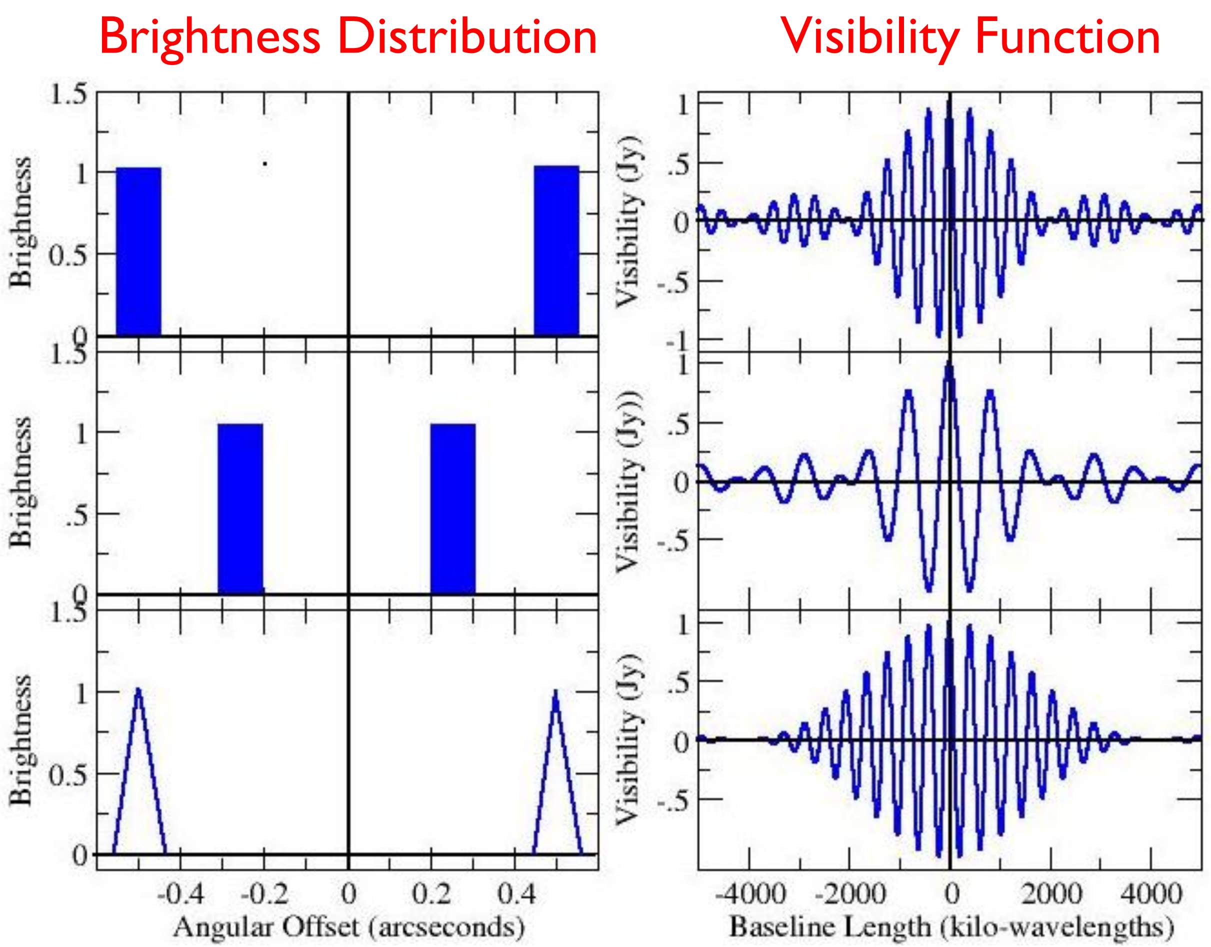
- Analysis provides: $V(u) = \dfrac{\sin(\pi u l_0)}{\pi u l_0}\cos(\pi u l_1) = \mathrm{sinc}(u\, l_0)\cos(\pi u l_1)$

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.

- Resolved Double
- Resolved Double
- Central Peaked Double



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’.

The Two 'Spaces'

- The interferometer measures the Fourier transform of sky emission.
- According to Bracewell, all physically valid brightness emission functions have a Fourier transform.
- Thus, all real sky brightnesses have a Fourier transform (the visibility function), which can be measured by an interferometer.
- The questions of how we make an interferometer which actually works, and how we convert the measurements of the visibility to an image of the brightness, are answered next.

Practical Details (left out here)

- In this shortened presentation, I skipped over many practical details:
  - Finite bandwidths
  - Finite averaging times
  - Observing from a rotating platform (i.e., the earth!)
  - Heterodyning (LO/IF downconversion).
  - Spectral resolution
  - Polarization
- All of these things modify, but don't change the essentials.
- To learn about these, look up my presentations at last year's 'NRAO synthesis summer school'.
  - Better – plan on attending next year's 'Synthesis Imaging Summer School', in Socorro.

Measuring Visibilities

- In order to generate an image, we 'lots and lots' of measurements of the visibility function.
- How many? LOTS! Usually tens of thousands, more commonly millions.
  - This subject, and the question of how the images are actually made, is the subject of another lecture.
  - So much material, so little time ..
- Two ways of getting 'enough' visibilities:
  - Build more interferometers – an array(expensive!)
  - Move the interferometers around (to change the spacing length and orientation). (less expensive, but can be complicated …)
  - Do both.
- More on these points later.

NRAO

Geometry and Imaging

- In this second half, I expand on my comment that 'under some circumstances', the basic relation is a 2-d Fourier transform.
- Coordinate systems
  - Direction cosines
  - (u,v) plane, and (u,v,w) volumes
  - 2-D ('planar') interferometers
  - 3-D ('volume') interferometers
  - Handling '3-D' imaging
- U-V Coverage, Visibilities, and Simple Structures.
- Examples – lots of them.
- A few words on imaging and deconvolution.

Interferometer Geometry

- We have not defined any geometric system for our relations.
- The response functions we defined were generalized in terms of the scalar product between two fundamental vectors:
  - The baseline '**b**', defining the direction and separation of the antennas, and
  - The unit vector '**s**', specifying the direction on the sky.
- The relationship between the interferometer's measurements, and the sky emission is:

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

- At this time, we define the geometric coordinate frame for the interferometer.

The 2-Dimensional Interferometer

Case A: A 2-dimensional measurement plane.

- Suppose the measurements of $V_\nu(\mathbf{b})$ are taken entirely on a plane.
- Then a considerable simplification occurs if we arrange the coordinate system so one axis is normal to this plane.
- Let (u,v,w) be the coordinate axes, with w normal to this plane. Then:

$$\mathbf{b} = (\lambda \mathrm{u}, \lambda \mathrm{v}, \lambda \mathrm{w}) = (\lambda \mathrm{u}, \lambda \mathrm{v}, 0)$$

u, v, and w are always measured in wavelengths.

- The components of the unit direction vector, **s**, are:

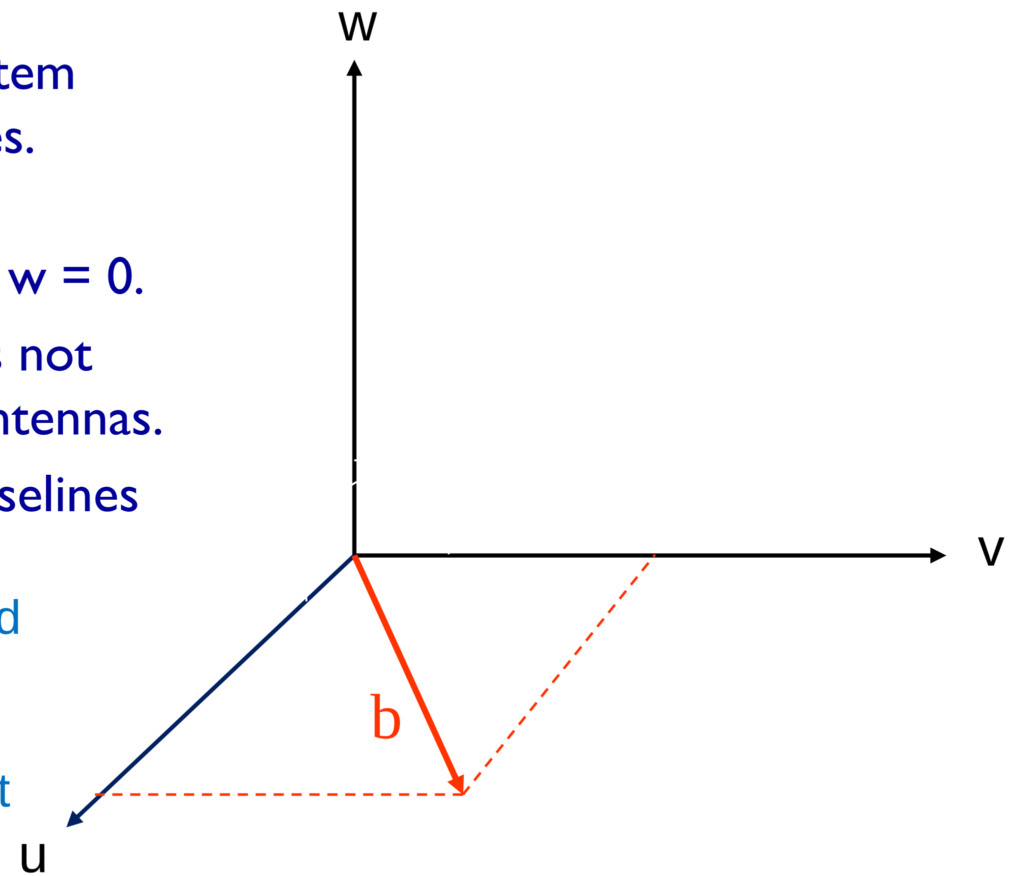
$$\mathbf{s} = (l, m, n) = \left(l, m, \sqrt{1 - l^2 - m^2}\right)$$

the simplification arises since $|\mathbf{s}|=1$. Only two coordinates are needed to specify direction.

- (l,m,n) are the **direction cosines**.

The (u,v,w) Coordinate System.

- Pick a cartesian coordinate system (u,v,w) to describe the baselines.
- Orient this frame so the plane containing the antennas lies on w = 0.
- This is a spatial frame, but does not describe the locations of the antennas.
- It measures the locus of the baselines



The baseline vector **b** is specified by its coordinates (u,v,w) (measured in wavelengths).
In the case shown, w = 0, so that

$$\mathbf{b} = (\lambda u, \lambda v, 0)$$

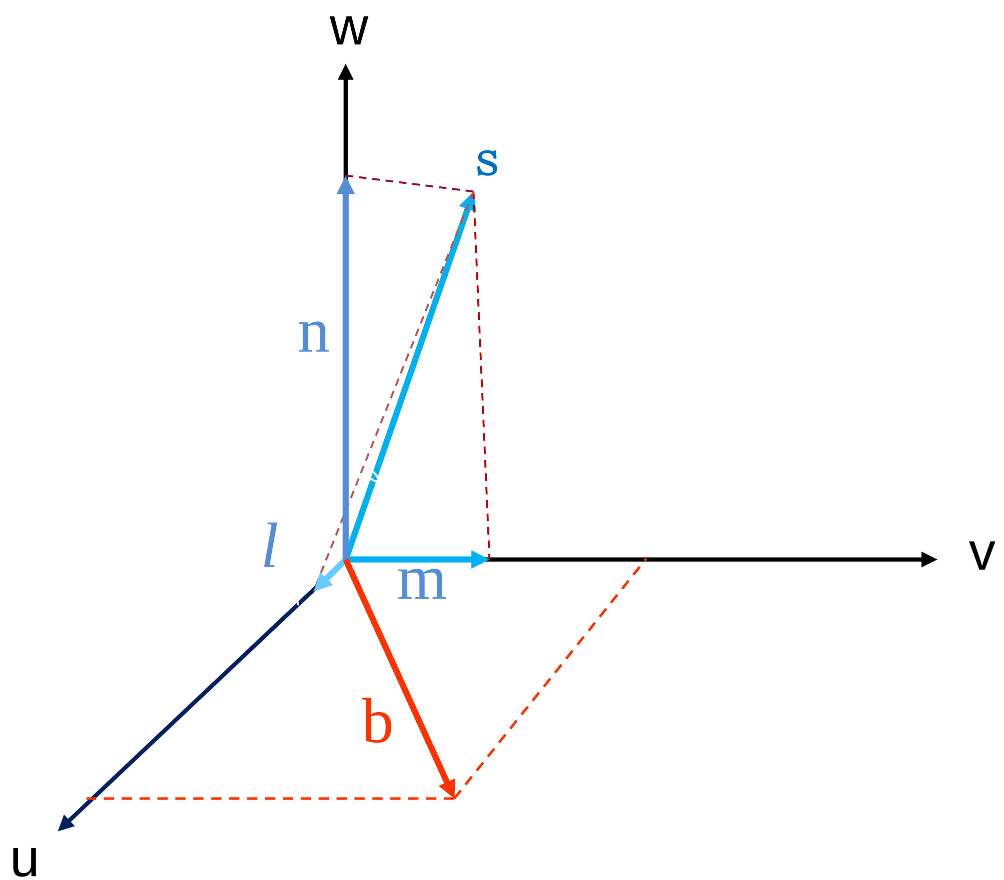
Direction Cosines – describing the source

The unit direction vector **s** is defined by its projections (l,m,n) on the (u,v,w) axes. These components are called the **Direction Cosines**.

$$l = \cos(\alpha)$$

$$m = \cos(\beta)$$

$$n = \cos(\theta) = \sqrt{1 - l^2 - m^2}$$



The angles, α, β, and θ are between the direction vector and the three axes.

The 2-d Fourier Transform Relation

Then, **b.s**/λ = u*l* + vm + wn = u*l* + vm, from which we find,

$$\boxed{V_\nu(u,\mathrm{v}) = \iint I_\upsilon(l,m) e^{-i2\pi(ul+vm)} \; dldm}$$

which is a **2-dimensional Fourier transform** between the brightness and the spatial coherence function (visibility):

$$I_\upsilon(l,m) \Leftrightarrow V_\upsilon(u,\mathrm{v})$$

And we can now rely on two centuries of effort by mathematicians on how to invert this equation, and how much information we need to obtain an image of sufficient quality.

Formally,

$$\boxed{I_\nu(l,m) = \iint V_\nu(u,\mathrm{v}) e^{i2\pi(ul+\mathrm{v}m)} du\,d\mathrm{v}}$$

In physical optics, this is known as the 'Van Cittert-Zernicke Theorem'.

How we actually do this inversion is left to the 'Imaging' lecture.

Interferometers with 2-d Geometry

- **Which interferometers can use this special geometry?**

  a) Those whose baselines, over time, lie on a plane (any plane).

  All E-W interferometers are in this group. For these, the w-coordinate points to the NCP. The (u,v) plane is the Equatorial Plane.

  - WSRT (Westerbork Synthesis Radio Telescope)
  - ATCA (Australia Telescope Compact Array) (before the third arm)
  - Cambridge 5km (Ryle) telescope (approximately).

  b) Any coplanar 2-dimensional array, at a single instance of time.

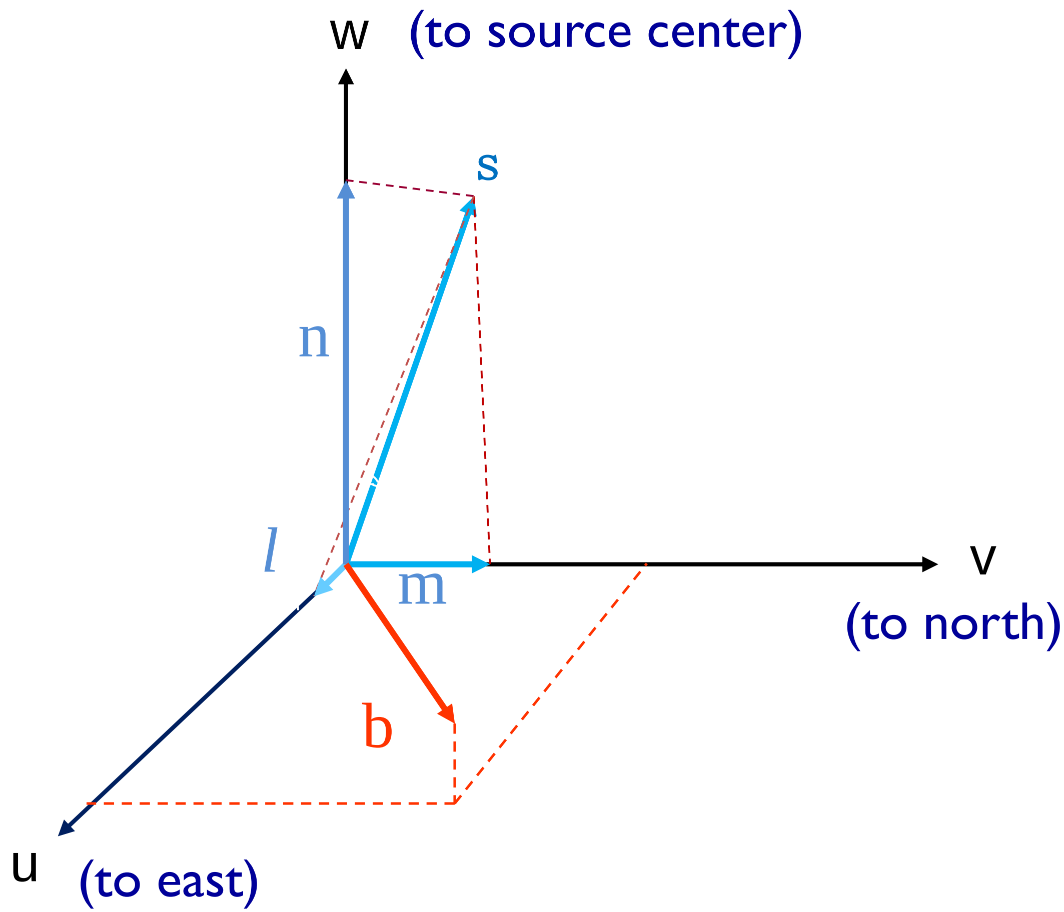
  In this case, the 'w' coordinate points to the zenith.

  - VLA or GMRT in snapshot (single short observation) mode.

- **What's the 'downside' of 2-d (u,v) coverage?**
  - Resolution degrades for observations that are not in the w-direction.
    - E-W interferometers have no N-S resolution for observations at the celestial equator.
    - A VLA snapshot of a source will have no 'vertical' resolution for objects on the horizon.

Generalized Baseline Geometry

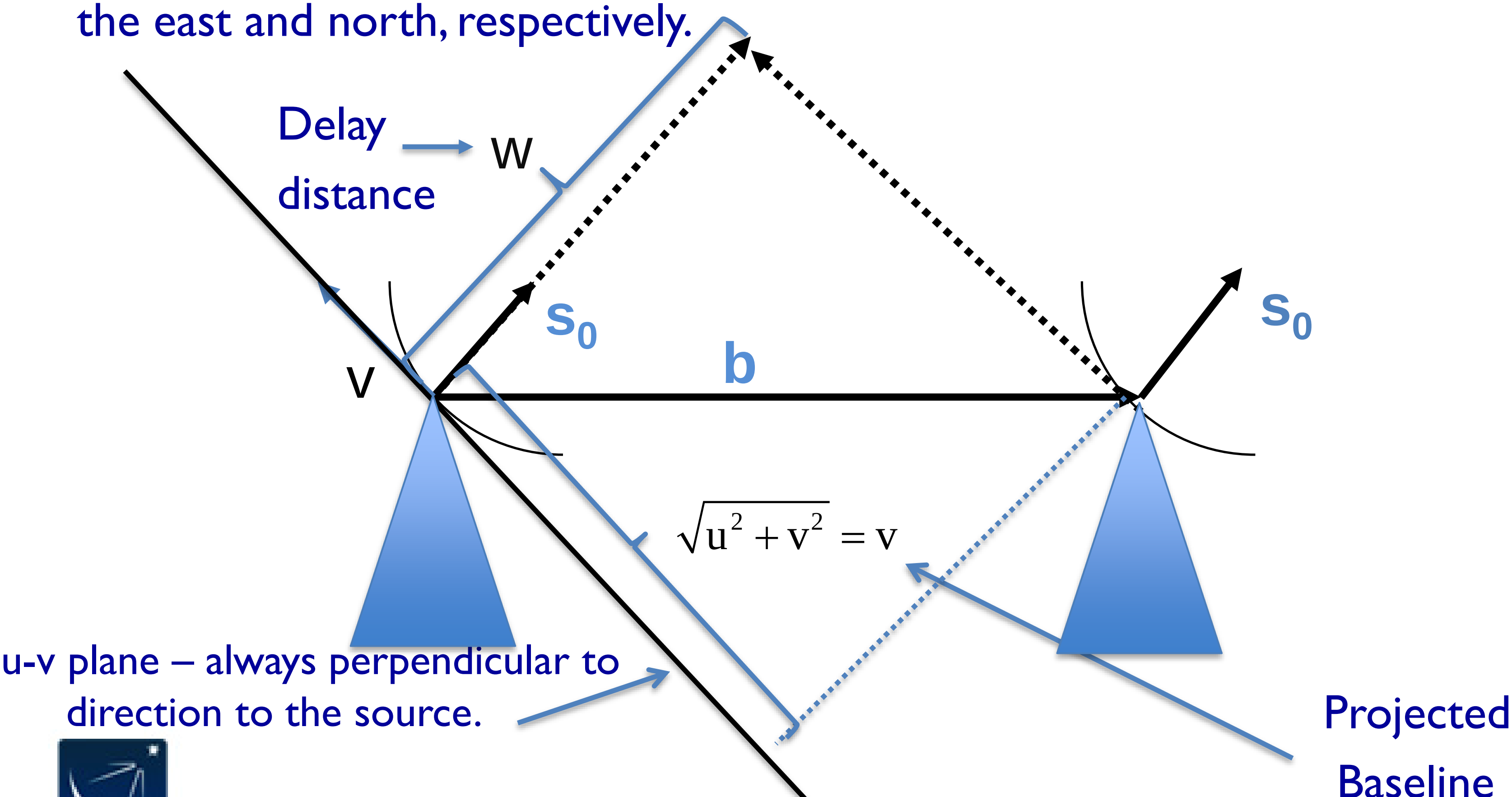
- General arrays (like the VLA) cannot use the 2-d geometry, since the antennas are not along an E-W line, and
- Over time, their baselines move through a (u,v,w) volume
- In this case, we adopt a more general geometry, where all three baseline components are to be considered.
- Arrange 'w' to point to the source (phase tracking center), and orient (u,v) plane so the 'v' axis points towards the NCP, and 'u' towards the east.



Baseline vector **b** now has three time-variable components.

General Coordinate System

- This is the coordinate system in most general use for synthesis imaging.
- **w** points to, and follows the source, **u** towards the east, and **v** towards the north celestial pole. The direction cosines *l* and m then increase to the east and north, respectively.



3-d Interferometers

Case B: A 3-dimensional measurement volume:

- The complete relation between the visibility and sky brightness is now more complicated:

$$V_\nu(\mathrm{u},\mathrm{v},\mathrm{w}) = \iint I_\nu(l,m)\, e^{-2i\pi(\mathrm{u}l+\mathrm{v}m+\mathrm{w}n)}\, dl\, dm$$

(Note that this is neither a 2-D or a 3-D Fourier Transform).

- We introduce phase tracking, so the fringes are 'stopped' for the direction l=m=0. This means we adjust the phases by $e^{2i\pi w}$
- Then, remembering that $n^2 = 1 - l^2 - m^2$ we get:

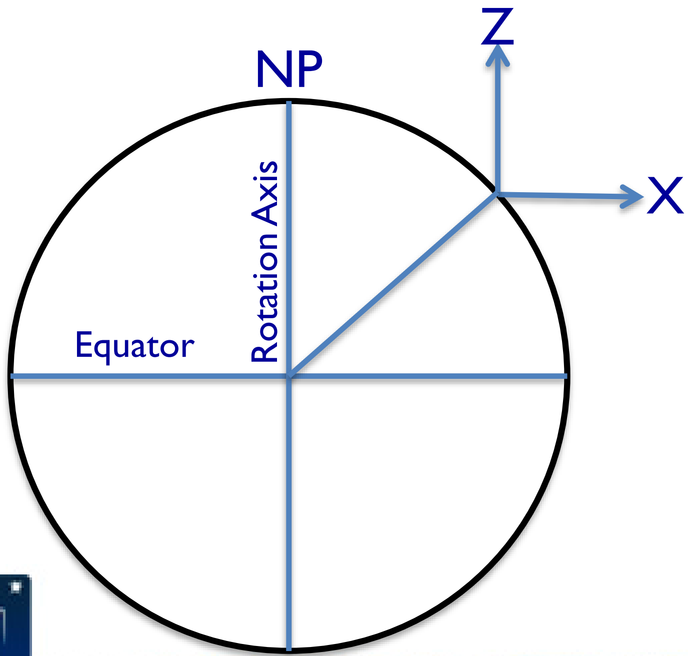
$$V_\nu(u,v,w) = \iint I_\nu(l,m)\, e^{-2i\pi[ul+vm+w(\sqrt{1-l^2-m^2}-1)]}\, dl\, dm$$

- The problem lies with the third term: $w\left(1-\sqrt{1-l^2-m^2}\right)$
- If this term is very small (<<1), then we might ignore it, in which case we return to a nice 2-D transform – **but now not exact.**

$$V_\nu(u,v,w) = \iint I_\nu(l,m)\, e^{-2i\pi(ul+vm)}\, dl\, dm$$

Coverage of the U-V Plane

- Obtaining a good image of a source requires adequate sampling ('coverage') of the (u,v) plane.
- Adopt an earth-based coordinate grid to describe the antenna positions (Bx, By, Bz)
  - X points to H=0, δ=0 (intersection of meridian and celestial equator)
  - Y points to H = -6, δ = 0 (to east, on celestial equator)
  - Z points to δ = 90 (to NCP).



- Thus, Bx, By are the baseline components in the Equatorial plane,
- Bz is the baseline component along the earth's rotation axis.
- All components in wavelengths.
- Now compute the (u,v,w) components of the baseline for a given H (hour angle) and δ.

(u,v,w) Coordinates

- Then, it can be shown that

$$\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \\ \mathbf{w} \end{pmatrix} = \begin{pmatrix} \sin H_0 & \cos H_0 & 0 \\ -\sin\delta_0 \cos H_0 & \sin\delta_0 \sin H_0 & \cos\delta_0 \\ \cos\delta_0 \cos H_0 & -\cos\delta_0 \sin H_0 & \sin\delta_0 \end{pmatrix} \begin{pmatrix} B_X \\ B_Y \\ B_Z \end{pmatrix}$$

- The u and v coordinates describe E-W and N-S components of the **projected** interferometer baseline.
- The w coordinate is the delay distance in wavelengths between the two antennas. The geometric time delay, τ_g is given by

$$\tau_g = \frac{\lambda}{c} w = \frac{w}{\upsilon}$$

- The time derivative of w, called the fringe frequency ν_F is

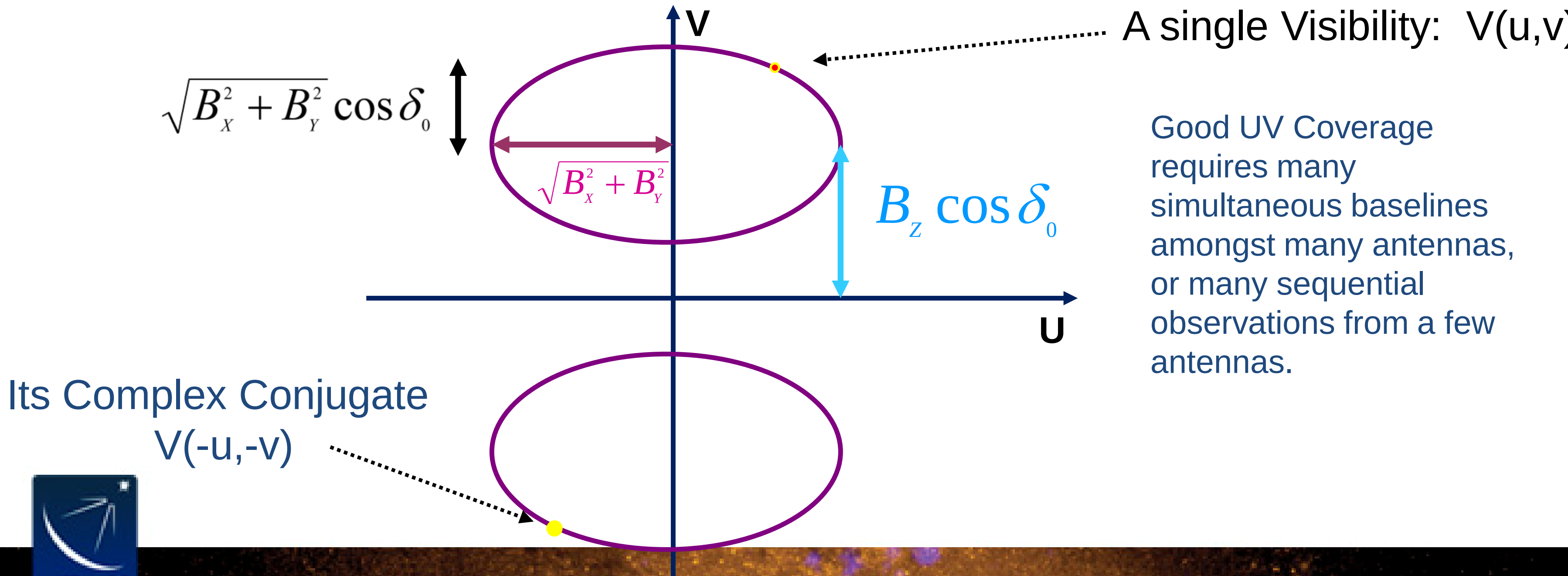
$$\nu_F = \frac{dw}{dt} = -\frac{dH}{dt} u \cos\delta_0 = -\omega_E u \cos\delta_0$$

Baseline Locus – the General Case

- Each baseline, over 24 hours, traces out an ellipse in the (u,v) plane:

$$u^2 + \left(\frac{v - B_Z \cos\delta_0}{\sin\delta_0} \right)^2 = B_x^2 + B_y^2$$

- Because brightness is real, each observation provides us a second point, where: V(-u,-v) = V*(u,v)
- E-W baselines (Bx = Bz = 0) have no 'v' offset in the ellipses.



Good UV Coverage requires many simultaneous baselines amongst many antennas, or many sequential observations from a few antennas.

E-W Array Coverage and Synthesized Beams

- The simplest case is for E-W arrays, which give coplanar coverage.
- Then, $B_x = B_z = 0$, and $B_y = B$, the baseline length.
- For this, the (u,v,w) coordinates become especially simple:

$\mathrm{u} = B\cos H_0$	E-W component
$\mathrm{v} = B\sin\delta_0 \sin H_0$	N-S component
$\mathrm{w} = -B\cos\delta_0 \sin H_0$	Delay component

- The locus in the (u,v) plane is an ellipse centered at the origin.
- At $\delta = 90$, w = 0, and the locus is a circle of radius B.
- At $\delta = 0$, v = 0, and the locus is a line of length = B.

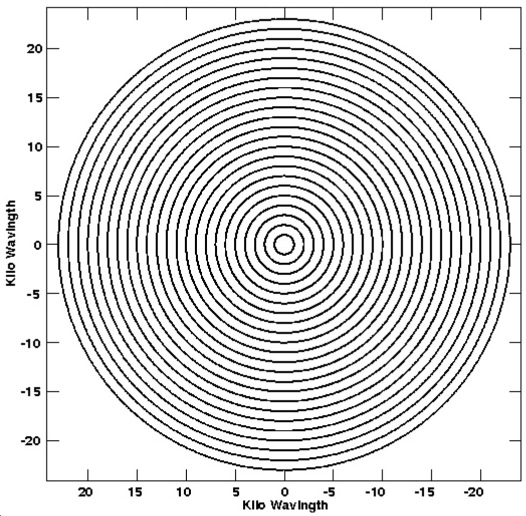
NRAO

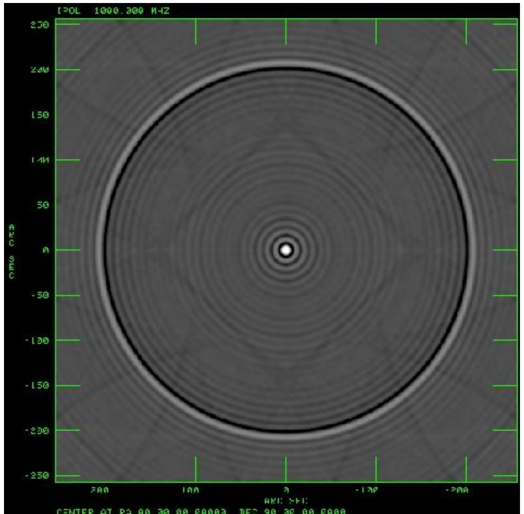
E-W Array Coverage and Beams

- The simplest case is for E-W arrays, which give coplanar coverage.
- Then, Bx = Bz = 0, and the ellipses are centered on the origin.
- Consider a 'minimum redundancy array', with eight antennas located at

  0, 1, 2, 11, 15, 18, 21 and 23 km along an E-W arm.

- Of the 28 simultaneous spacings, 23 are of a unique separation.
- The U-V coverage (over 12 hours) at δ = 90, and the synthesized beam are shown below, for a wavelength of 1m.

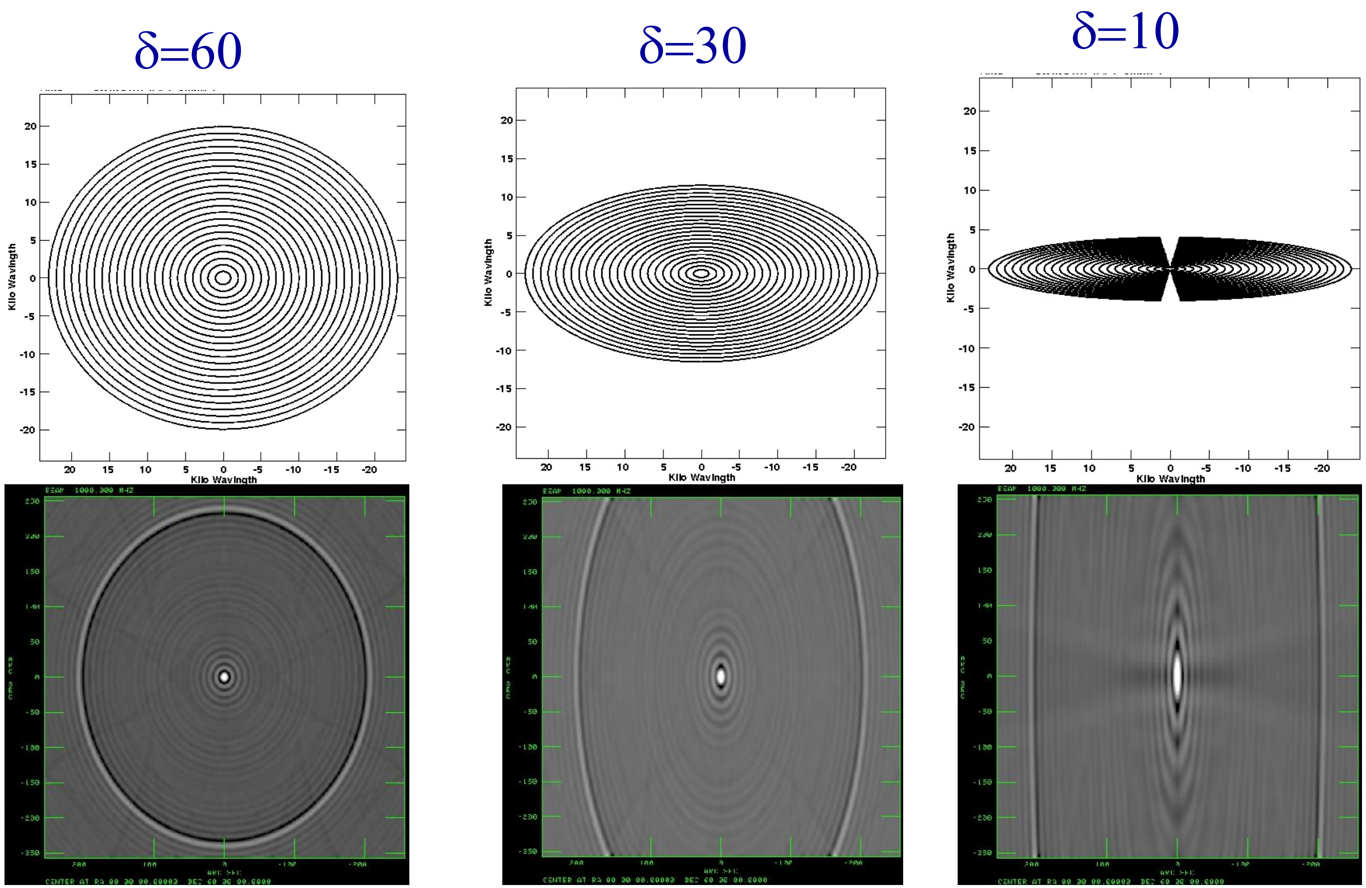




δ = 90

E-W Arrays and Low-Dec sources.

- But the trouble with E-W arrays is that they are not suited for low-declination observing.
- At δ=0, coverage degenerates to a line.



U-V Coverage

PSF

Getting Good Coverage near $\delta = 0$

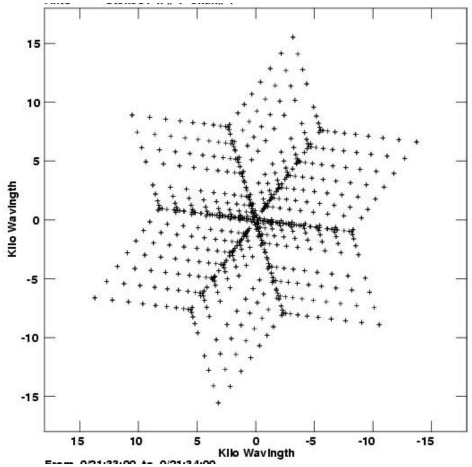
- The only means of getting good 2-d angular resolution at all declinations is to build an array with N-S spacings.
- Many more antennas are needed to provide good coverage for such geometries.
- The VLA was designed to do this, using 9 antennas on each of three equiangular arms.
- Built in the 1970s, commissioned in 1980, the VLA vastly improved radio synthesis imaging at all declinations.
- Each of the 351 (=27*26/2) spacings traces an elliptical locus on the (u,v) plane.
- Every baseline has some (N-S) component, so none of the ellipses is centered on the origin.

Advantages and Disadvantages of 2-d Arrays

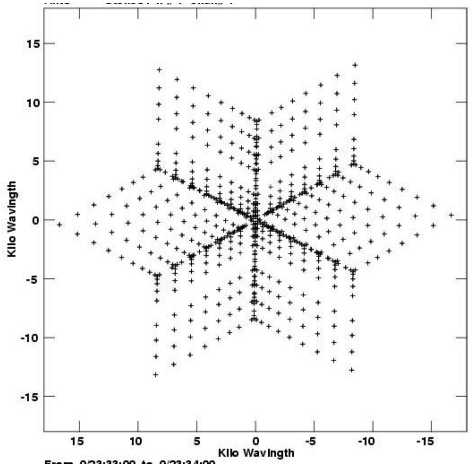
- The most obvious advantage is that the (u,v) coverage is instantaneously 2-dimensional.
  - This means that a 2-d image of the emission can – in principle – be formed from short observations.
- By contrast, the E-W interferometer must observe over a 12-hour period in order to populate the (u,v) plane.
  - A snapshot with an E-W interferometer gives a one-dimensional beam. (Not very useful).
- But --- 2-d arrays, over time, sample a 3-d (u,v,w) volume – not the desired 2-d (u,v) plane – requiring MUCH more processing in order to obtain a correct image.

Sample VLA (U,V) plots for 3C147 (δ = 50)

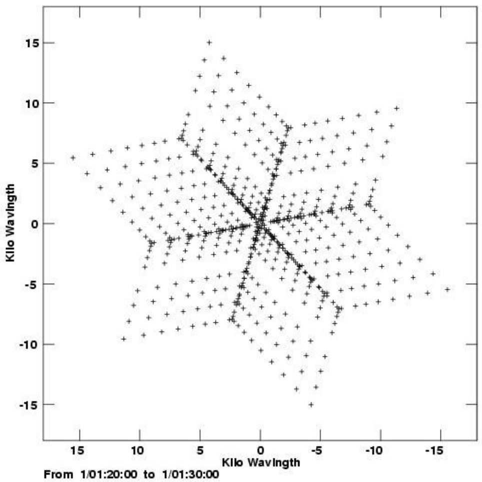
- Snapshot (u,v) coverage for HA = -2, 0, +2 (with 26 antennas).



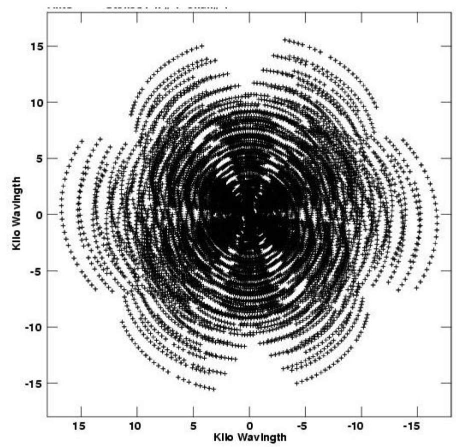
HA = -2h



HA = 0h

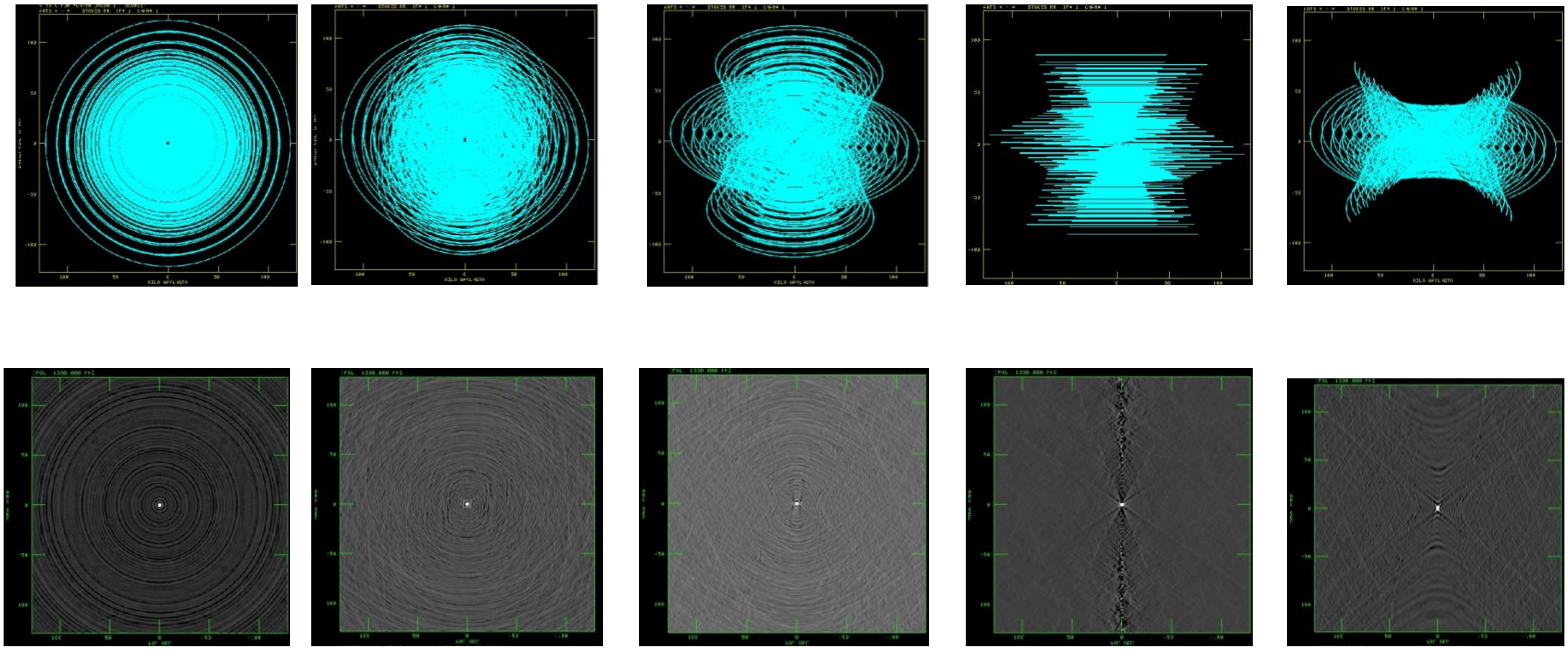


HA = 2h



Coverage over all four hours.

VLA Coverage and Beams



$\delta=90$ $\delta=60$ $\delta=30$ $\delta=0$ $\delta=-30$

- Good coverage at all declinations, but troubles near $\delta=0$ remain.

UV Coverage and Imaging Fidelity

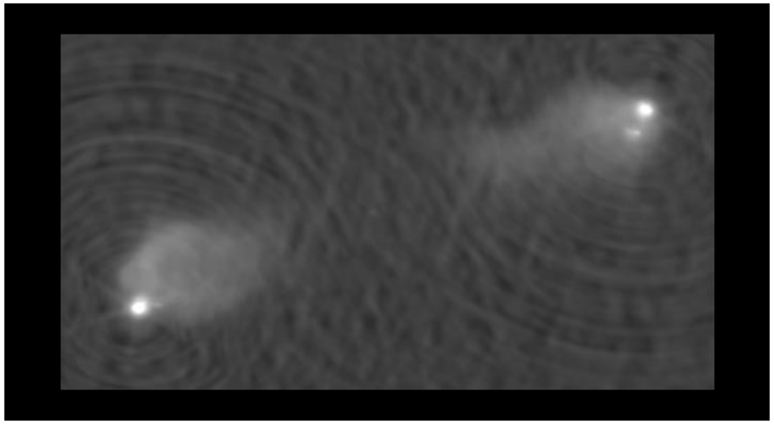
- Although the VLA represented a huge advance over what came before, its UV coverage (and imaging fidelity) is far from optimal.
- The high density of samplings along the arms (the 6-armed star in snapshot coverage) results in 'rays' in the images due to small errors.
- A better design is to 'randomize' the location of antennas within the span of the array, to better distribute the errors.
- Of course, more antennas would really help! :) .
- The VLA's wye design was dictated by its 220 ton antennas, and the need to move them. Railway tracks were the only answer.
- Future major arrays will utilize smaller, lighter elements which must not be positioned with any regularity.

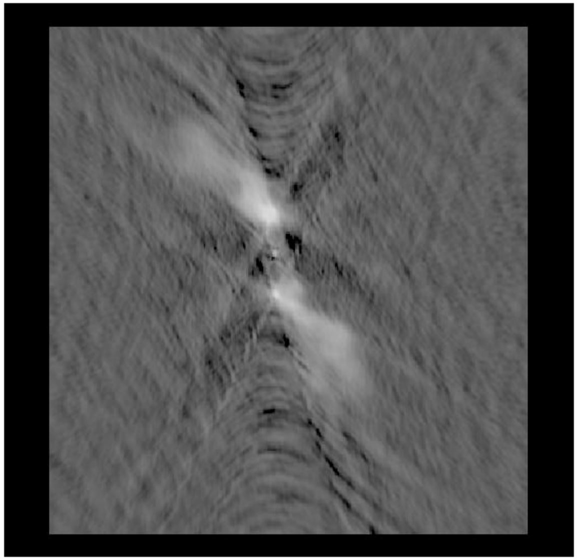
Imaging and Deconvolution

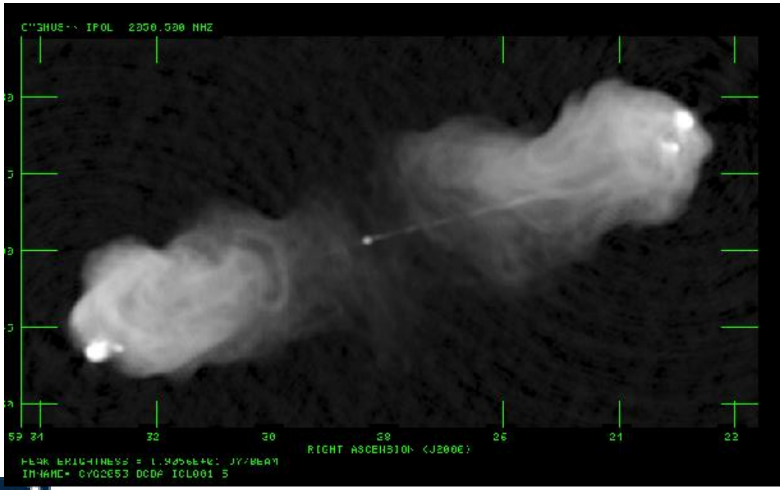
- The images made with the ‘raw’ data are truly ugly!
- They are rightfully called ‘dirty images’.
- They are ‘dirty’ because of the missing information – gaps in uv coverage.
- A major component of synthesis imaging (and the most expensive, by far) is ‘deconvolution’ – removing the sidelobes by some iterative process.
- This is a very complex subject – far beyond what I can explain here.
- ‘Deconvolution’ is an activity which falls under the topic known as ‘compressed sensing’.
- To learn more – come to Socorro next May.

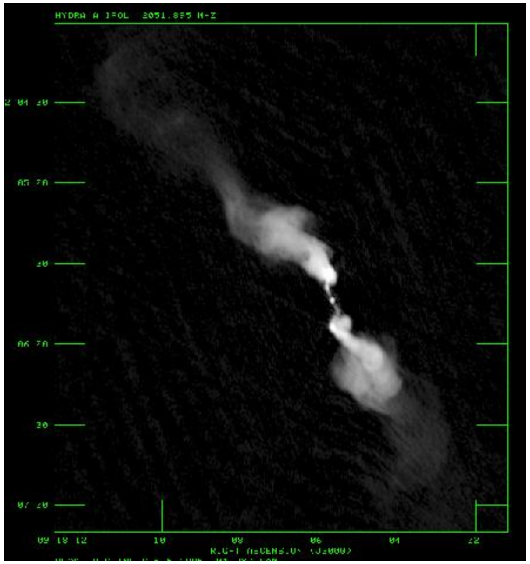
Examples of ‘Dirty’ and ‘Clean’

Cygnus A (below), Hydra A (right)









Examples with Real Data!

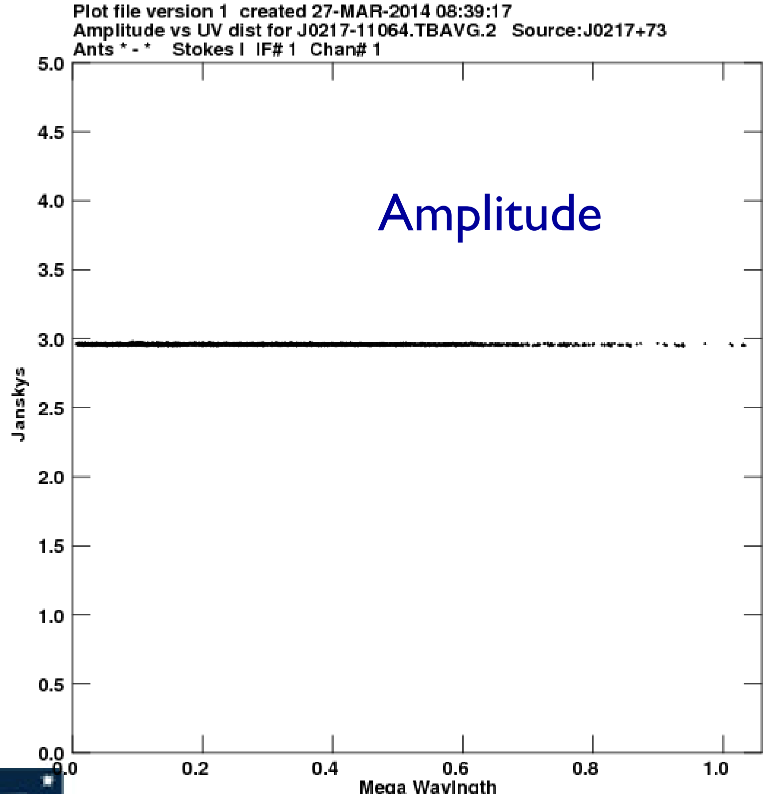
- Enough of the analysis!
- I close with some examples from real observations, using the VLA.
- These are two-dimensional observations (function of ‘u’ (EW) and ‘v’ (NS) baselines).
- Plotted are the visibility amplitudes version baseline length:

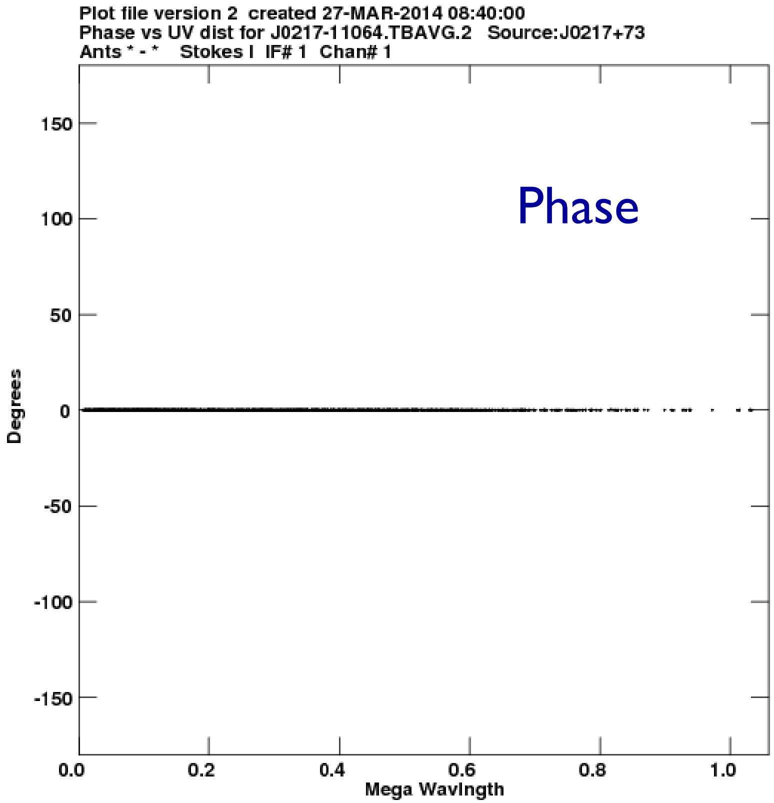
$$q = \sqrt{u^2 + v^2}$$

- 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

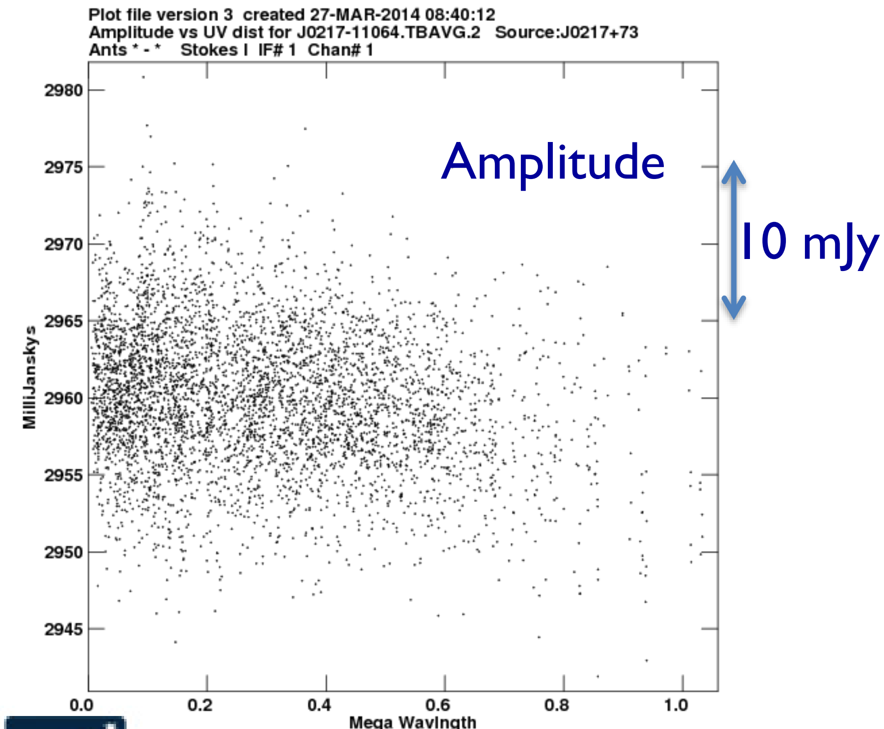
- Suppose we observe an unresolved object, at the phase center.
- What is its visibility function?

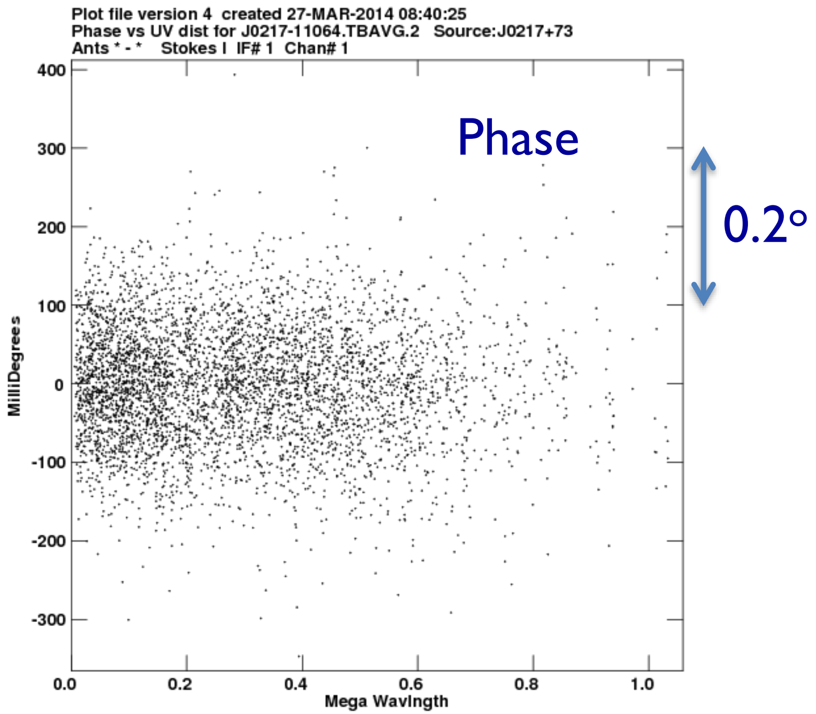




Zoom in …

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



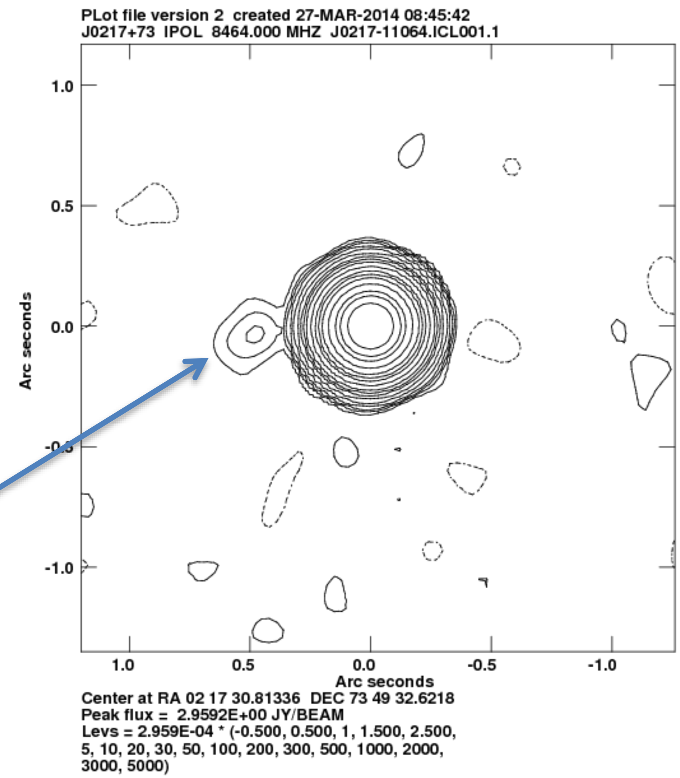


And the Map …

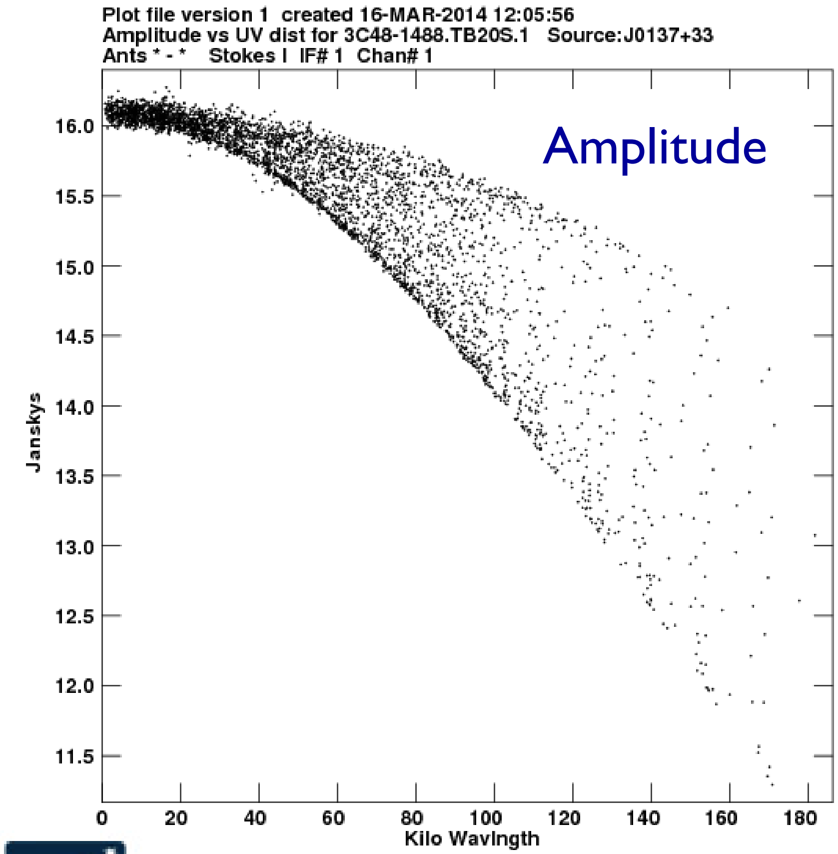
- The source is unresolved … but with a tiny background object.
- Dynamic range: 50,000:1.

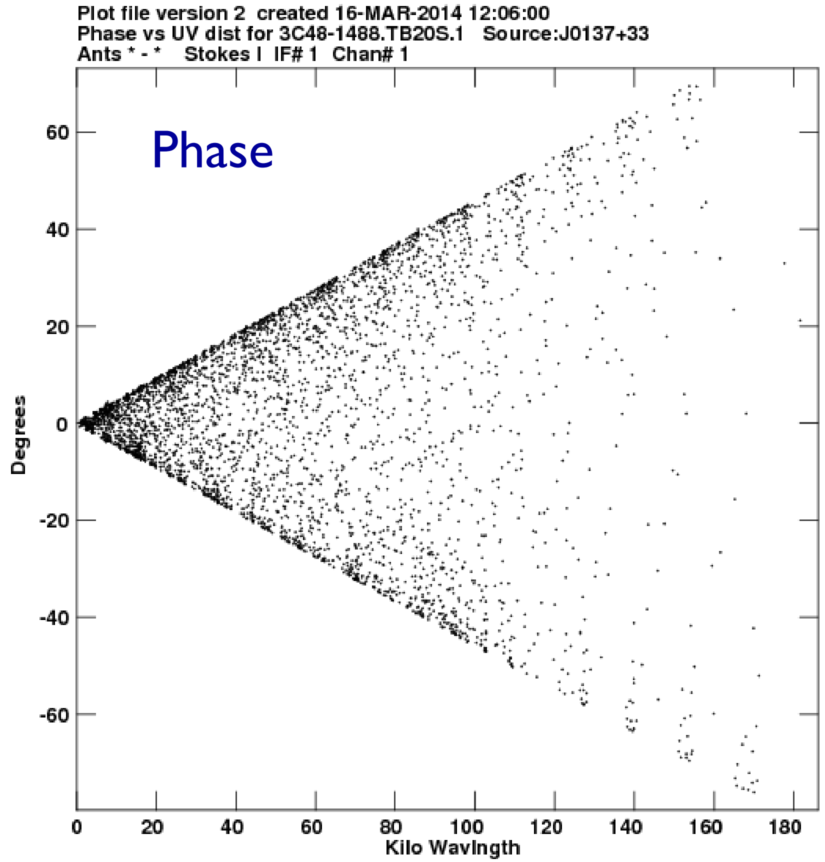
The flux in the weak nearby object is only 0.25 mJy – too low to be seen on any individual visibility.

Real!



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



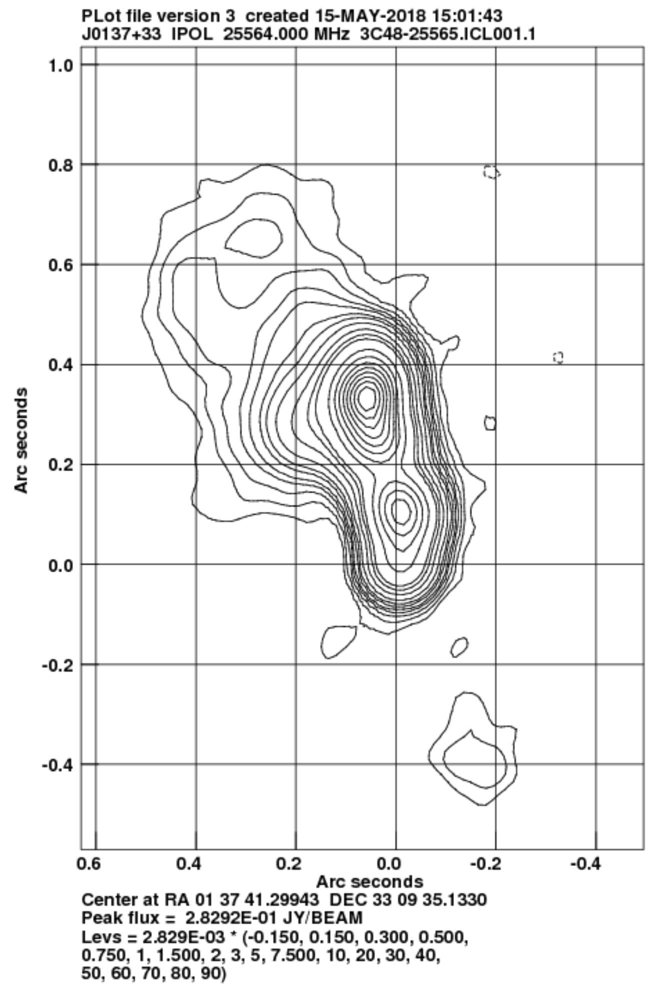


3C48 position and offset

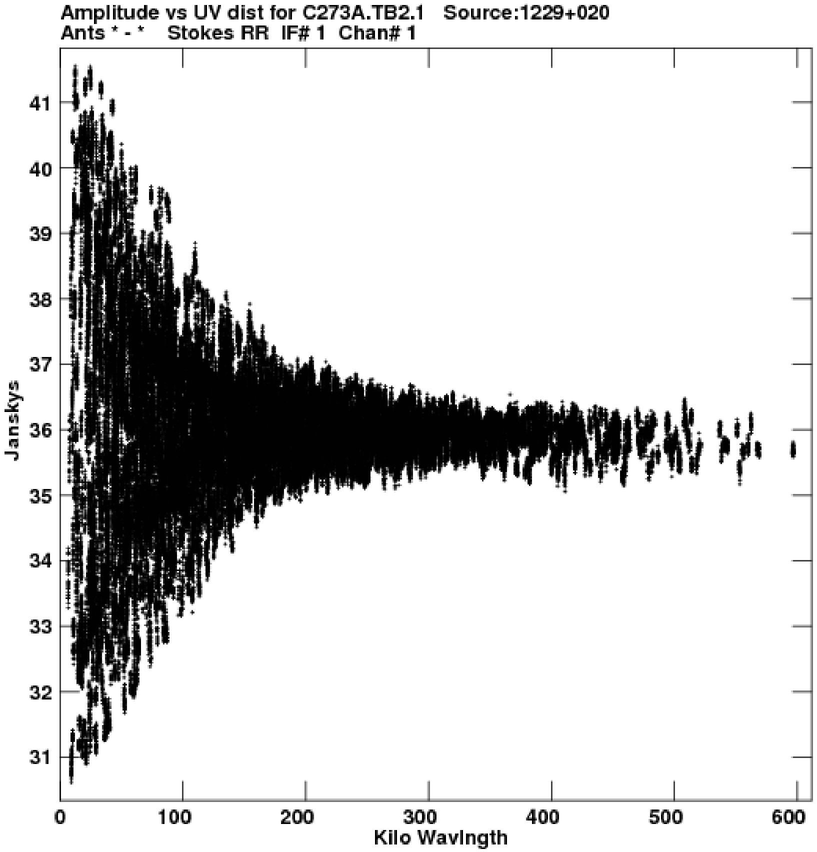
- You can learn quite a bit just by looking at the gross properties of the visibility amplitudes/phases, noting:
- **A 206265 wavelength baseline makes a 1 arcsecond fringe.**
- The linear phase slope is ~90 degrees over 250,000 wavelengths.
  - 90 degrees is ¼ of a fringe, and one fringe is one arcsecond. Thus, the source centroid is ~ 250 mas from the phase center.
- The amplitudes show slight (25%) resolution at 180,000 wavelengths. There is an upper and lower envelope.
  - The source is extended by a fraction (few tenths) of an arcsecond.
  - One axis has about a twice the size of the other.

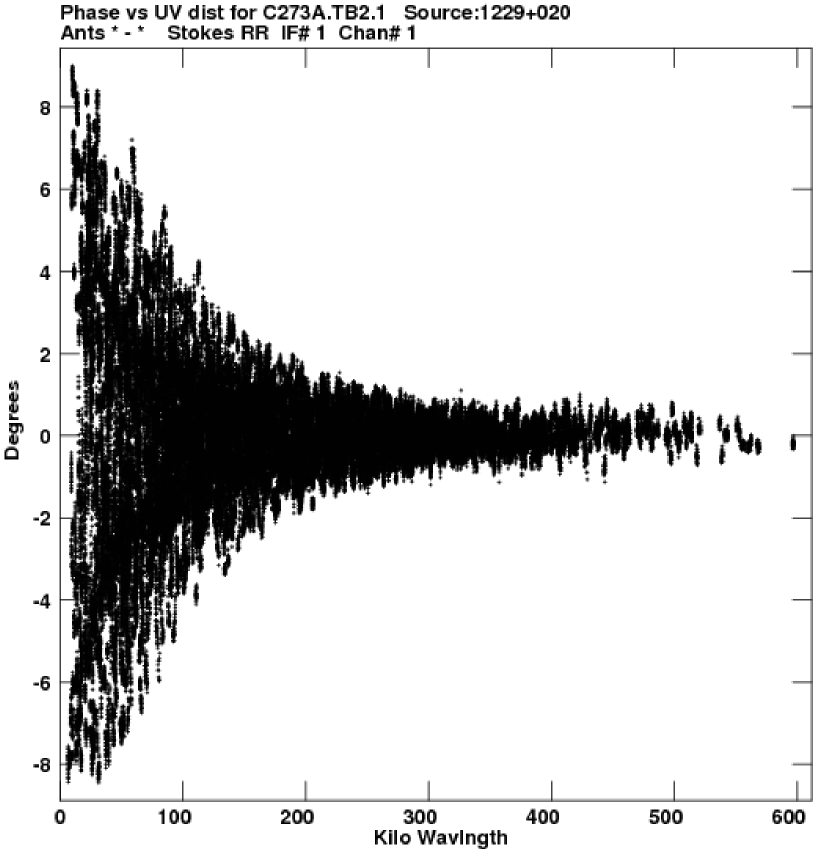
3C48 Structure (at 25 GHz …)

- The 1400 MHz image made from the data shown in the last slide doesn't show the structure well (poorly resolved).
- So here is the source at a higher frequency, where the resolution is 18 X higher (85 milliarcseconds)
- It is offset by 250 milliarcseconds from the phase center, and less than 1 arccsecond in size, with roughly a 2:1 ratio in size.



3C273 – Point source + Jet

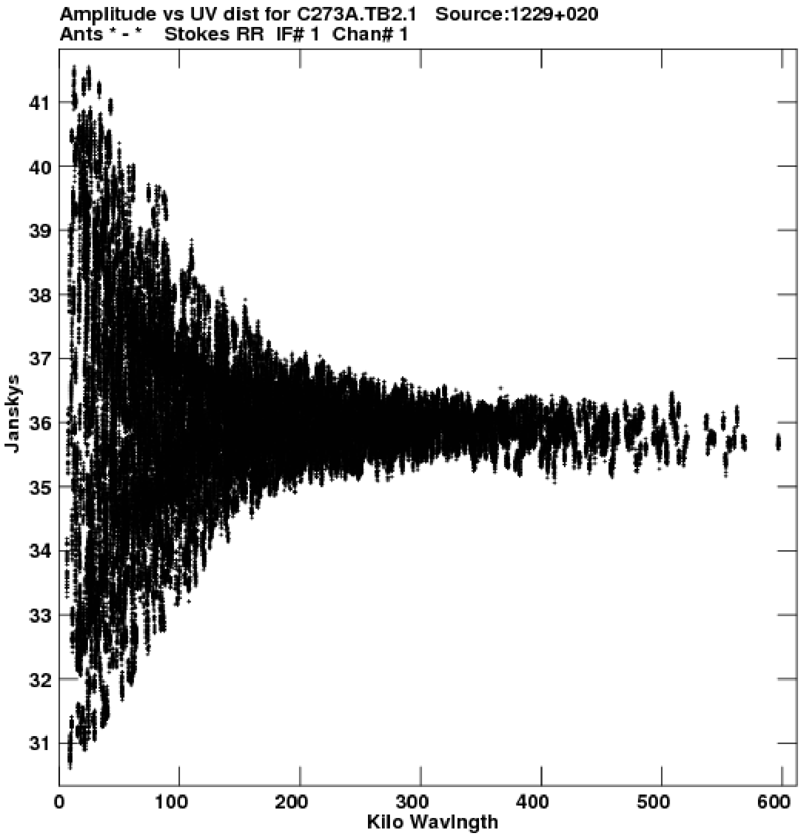




Visibility Amplitude **Visibility Phase**

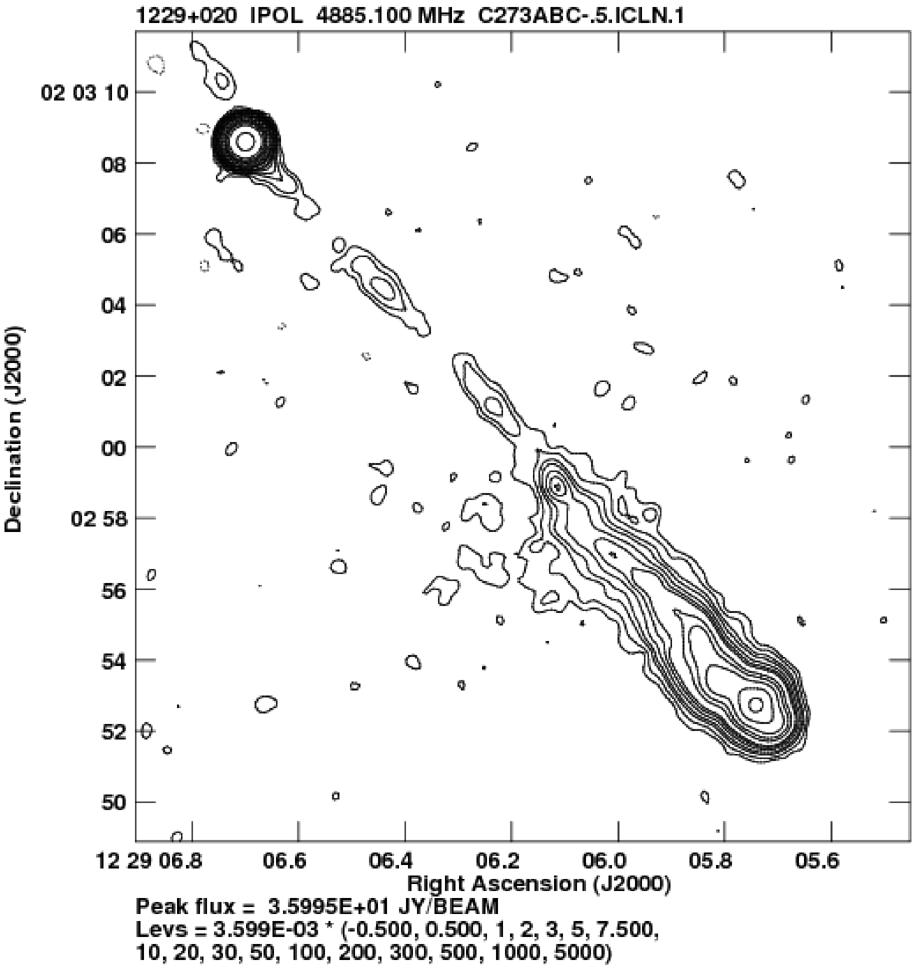
3C273 – jet size revealed …

- The narrowing of the visibility amplitudes, with no sign of resolution, tells us:
  - There a 36 Jy unresolved `point'-source.
- The absence of a phase gradient tells us:
  - The point source is at the center
- The extended emission resolves out at ~200 -- 400 Kλ
- This indicates width of ~1 –2"
- Rapid oscillations in amplitude and phase tells us:
  - A much larger extension is present.

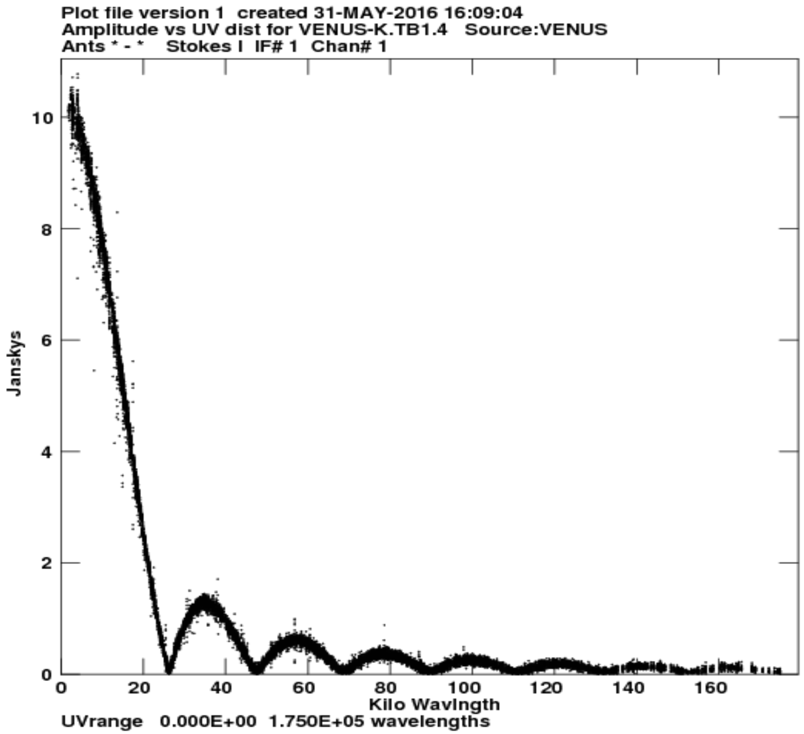


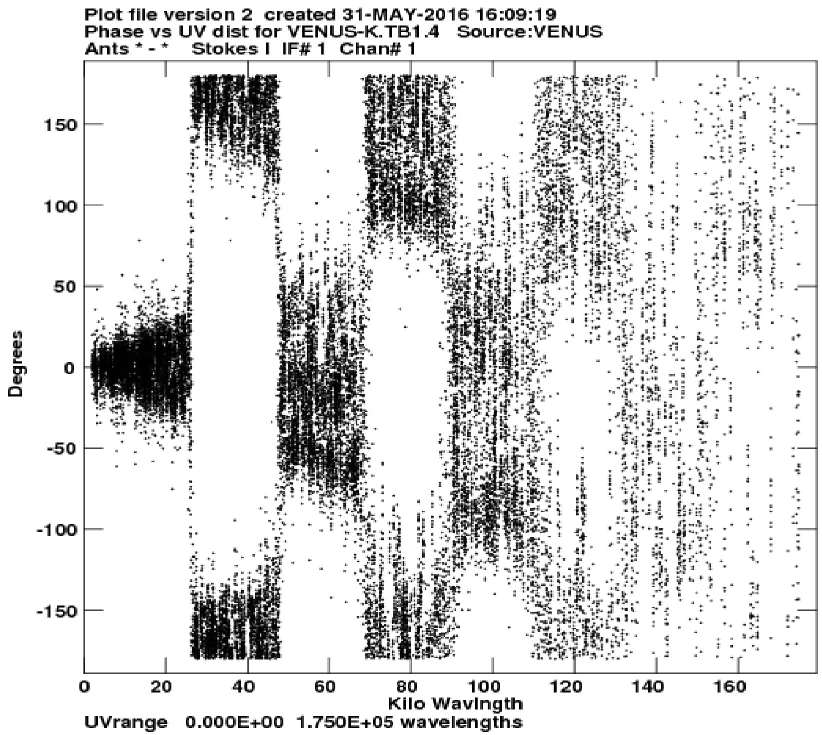
3C273 -- Image

- Actual structure revealed by making the image.
- There is a 36.0 Jy unresolved nucleus, with a one-sided jet.
- Jet width ~ 2 arcseconds
- Jet length ~ 18 arcsecond.



The Planet Venus at 19 GHz.

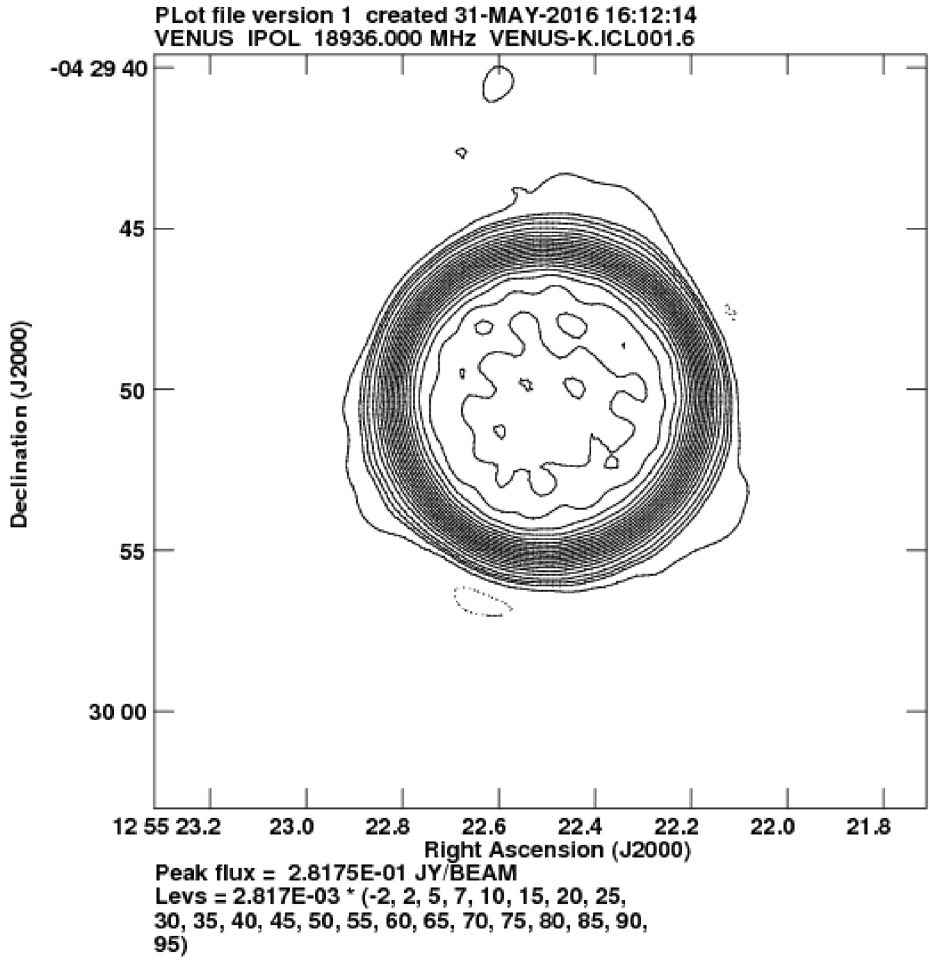




- The visibilities are circularly symmetric. The phases alternate between zero and 180 degrees.
- The source must be circularly symmetric and centered.
- The visibility null at 25 kλ indicates angular size of ~ 10 arcseconds.

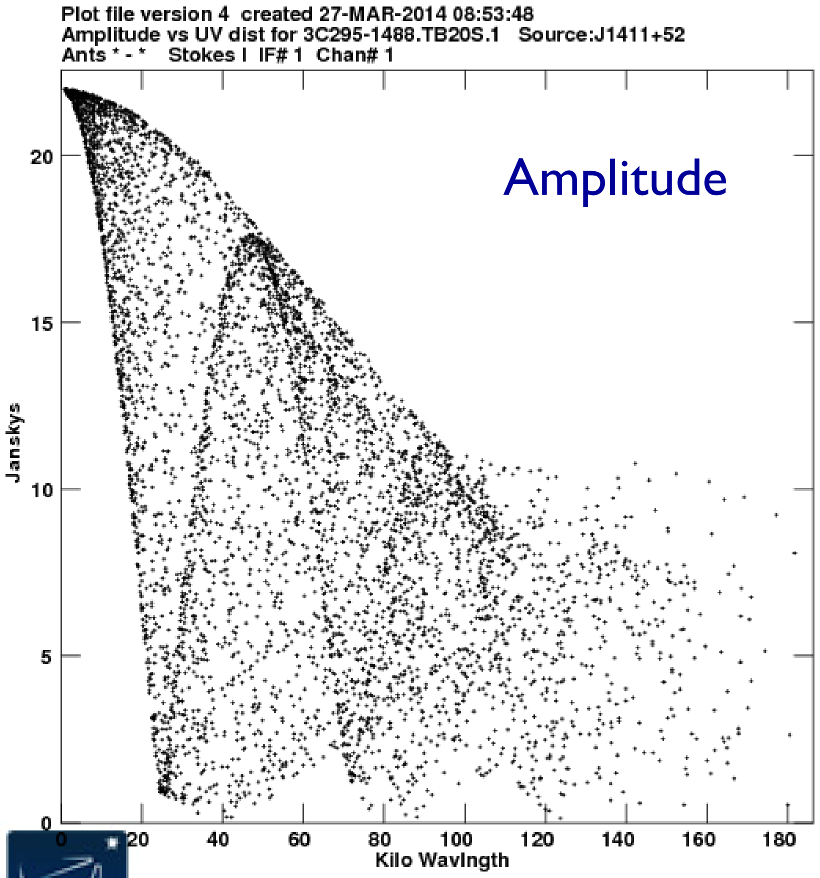
And the image looks like:

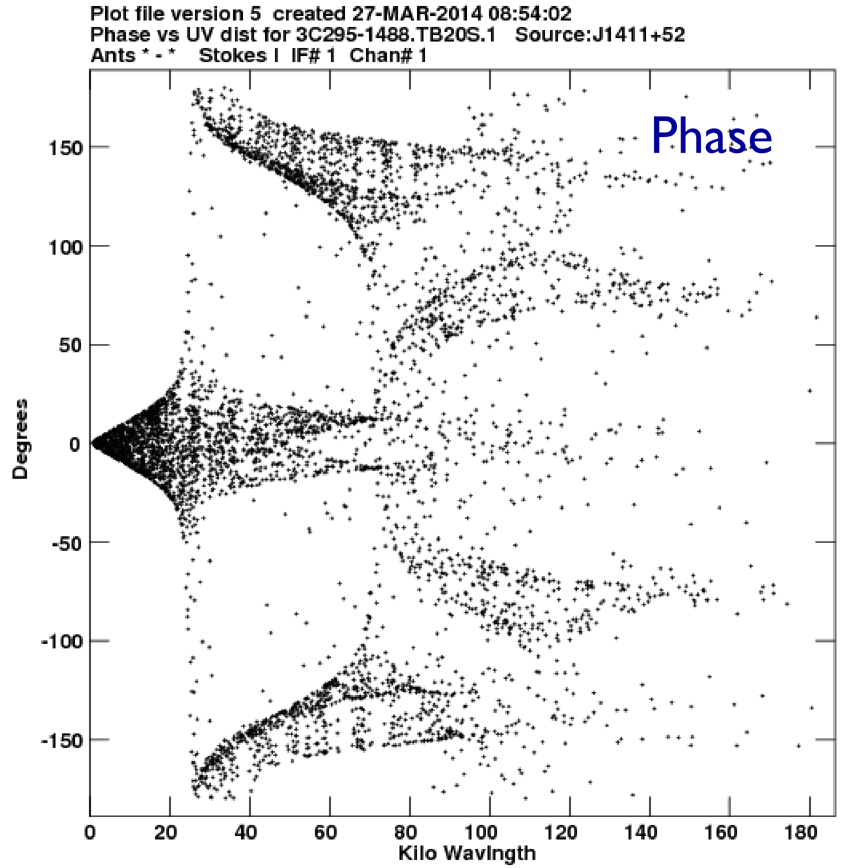
- It's a perfectly uniform, blank disk!
- The Visibility function, in fact, is an almost perfect Bessel function of zero order: $J_0(q)$.
- A perfect J_0 would arise from a perfectly sharp disk. Atmospheric opacity effects 'soften' the edge, resulting in small deviations from the J_0 function at large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



Examples of Visibilities – a Well Resolved Object

- The flux calibrator 3C295 at 1400 MHz.



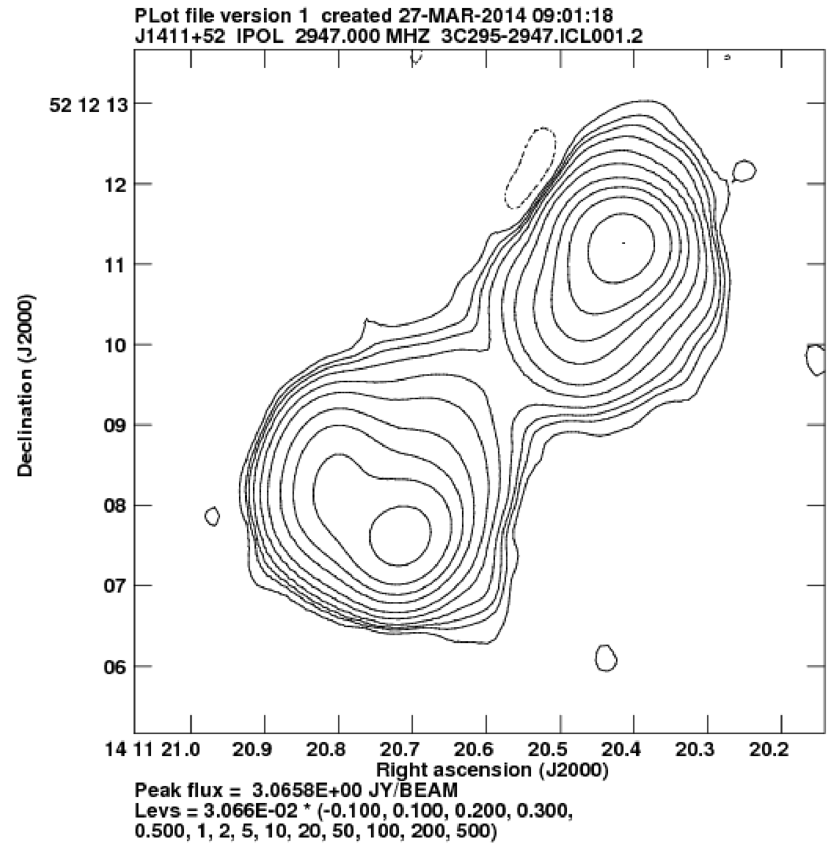


From the 1-d visibilities alone, we deduce

- The outer visibility scale of ~200 kλ says there is a 1" smaller scale.
- The cyclical variations, with period of 50kλ says there is a pair of smaller objects, separated by about 4"
- The lack of an overall phase slope tells us the object is centered on the phase center.
- Without knowing the 2-d distribution of the phases and amplitudes, we can say nothing about the orientations.

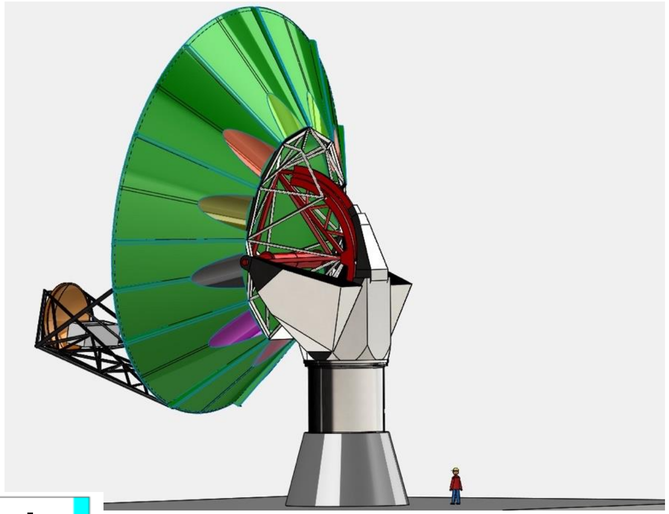
3C295 Image

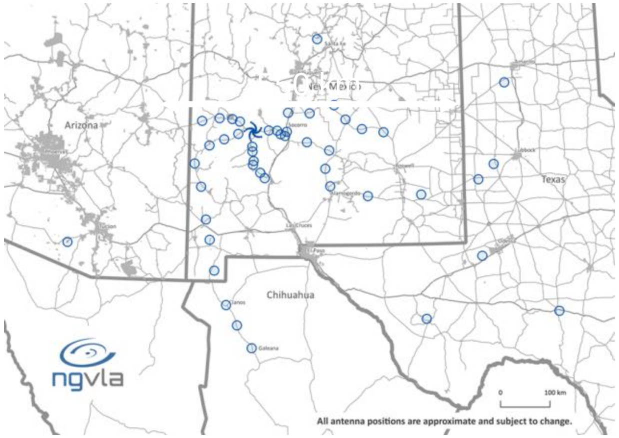
- The visibility amplitude cycles on a 60,000 wavelength period – corresponding to about 4 arcseconds extent – as shown in the image.
- The phase is too complicated to easily interpret!

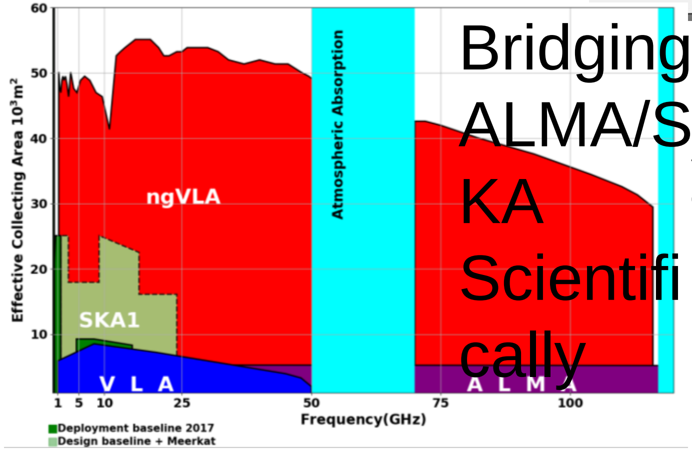


A next-generation Very Large Array (ngVLA)

- Scientific Frontier: ***Thermal imaging at milli-arcsec resolution***
- Sensitivity/Resolution Goal: ***10x sensitivity & resolution of JVLA/ALMA***
- Frequency range: ***1.2 –116 GHz***
- Located in Southwest U.S. (NM, TX, AZ) & MX, centered on VLA
- Low technical risk (reasonable step beyond state of the art)







Bridging ALMA/SKA Scientifically

Complementary suite from meter to submm arrays for the mid-21st century

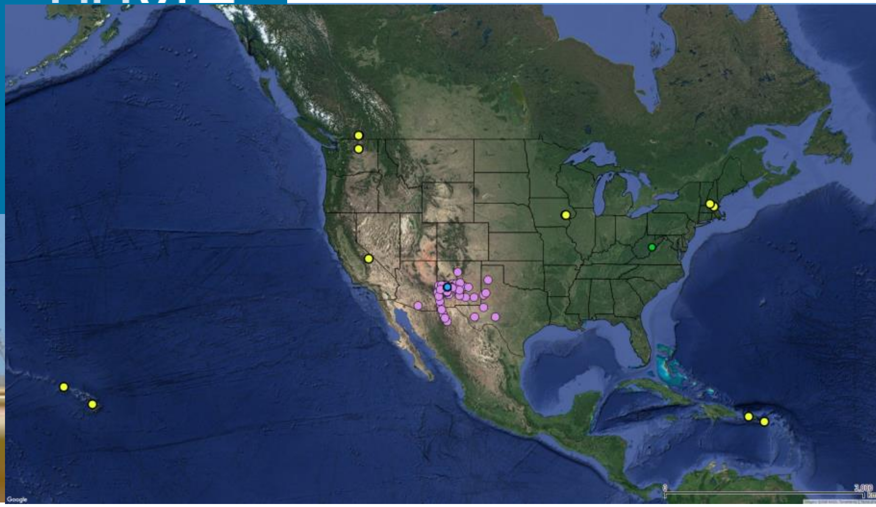
- **< 0.3 cm:** ALMA 2030
- **0.3 to 3 cm: ngVLA**
- **> 3 cm:** SKA

http://ngvla.nrao.edu

ngVLA Key Science Goals
(ngVLA memo #19)

1. *Unveiling the Formation of Solar System Analogues on Terrestrial Scales*
2. *Probing the Initial Conditions for Planetary Systems and Life with Astrochemistry*
3. *Charting the Assembly, Structure, and Evolution of Galaxies Over Cosmic Time*
4. *Using Pulsars in the Galactic Center as Fundamental Tests of Gravity*
5. *Understanding the Formation and Evolution of Stellar and Supermassive BH's in the Era of Multi-Messenger Astronomy*

ngvla.nrao.edu



- **1.2 - 116 GHz** Frequency Coverage
- **Main Array**: 214 x 18m offset Gregorian Antennas.
  - Fixed antenna locations across NM, TX, AZ, MX.
- **Short Baseline Array**: 19 x 6m offset Greg. Ant.
  - Use 4 x 18m in TP mode to fill in (*u, v*) hole.

Band #	Dewar	f_L GHz	f_M GHz	f_H GHz	f_H: f_L	BW GHz
1	A	1.2	2.35	3.5	2.91	2.3
2	B	3.5	7.90	12.3	3.51	8.8
3	B	12.3	16.4	20.5	1.67	8.2
4	B	20.5	27.3	34.0	1.66	13.5
5	B	30.5	40.5	50.5	1.66	20.0
6	B	70.0	93.0	116	1.66	46.0

- **Long Baseline Array**: 30 x 18m antennas located across continent for baselines up to 8860km.