Basic Radio Interferometry – Geometry

Rick Perley, NRAO/Socorro

Eighteenth (and Second Virtual) Synthesis Imaging Summer School

May 18 – 25, 2022

Topics

- Coordinate systems
  - Direction cosines
  - (u,v) planes, 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 of Visibilities– lots of them.

Interferometer Geometry

- We have not defined any geometric system for our relations.
- The visibility functions we defined were generalized in terms of the scalar product between two fundamental vectors:
  - The baseline '**b**', defining the orientation 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 the baseline's components are

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

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

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

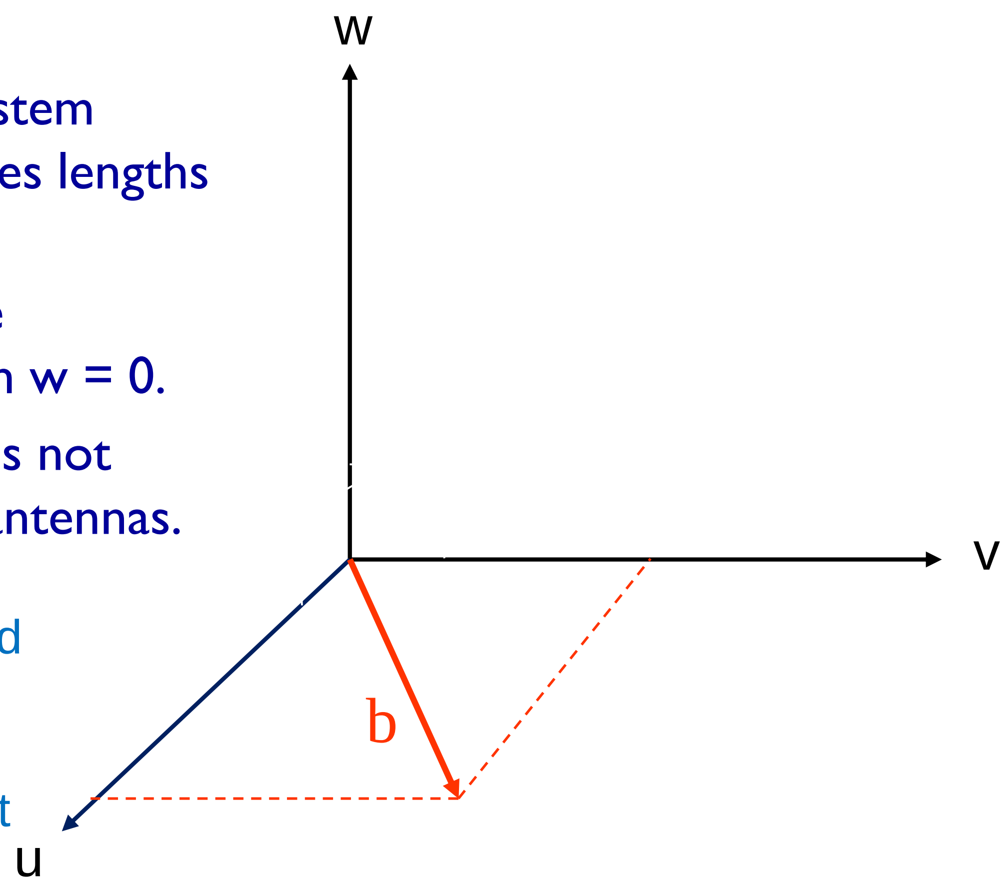
$$\mathbf{s} = (l, m, n) = \left(l, m, \sqrt{1 - l^2 - m^2}\right)$$

the simplification arises since |**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 lengths and orientations.
- Orient this frame so the plane containing the baselines lies on w = 0.
- This is a spatial frame, but does not describe the locations of the antennas.



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

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

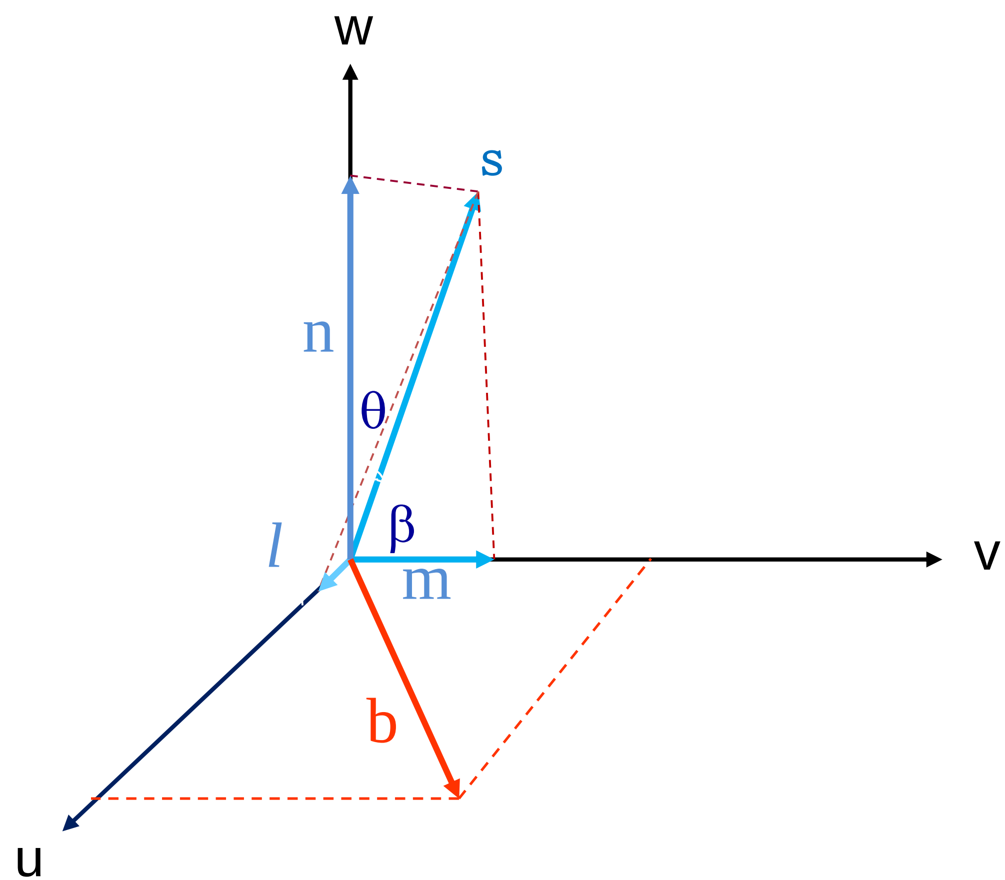
Direction Cosines – describing the source direction

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,   (because w = 0) from which we find,

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

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,

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

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.

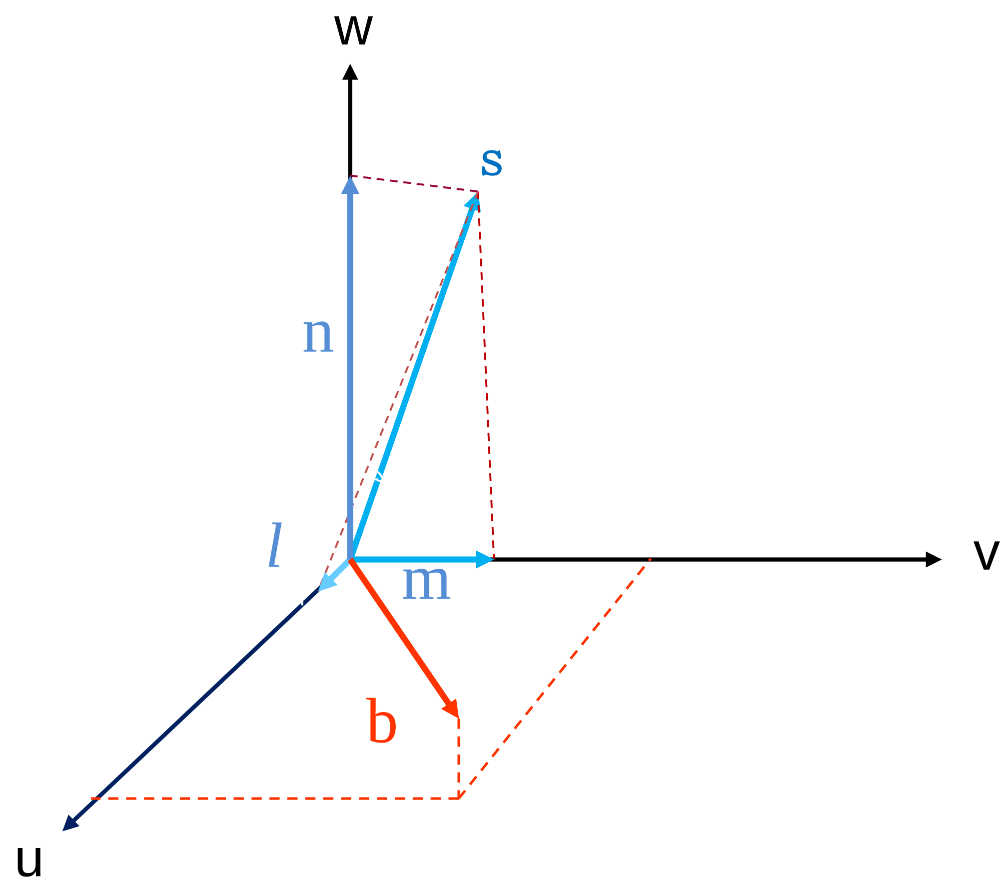
  In this case, the 'w' coordinate points to the zenith.

  - VLA, MeerKAT, and 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 near the horizon will have no 'vertical' resolution.

Generalized Baseline Geometry

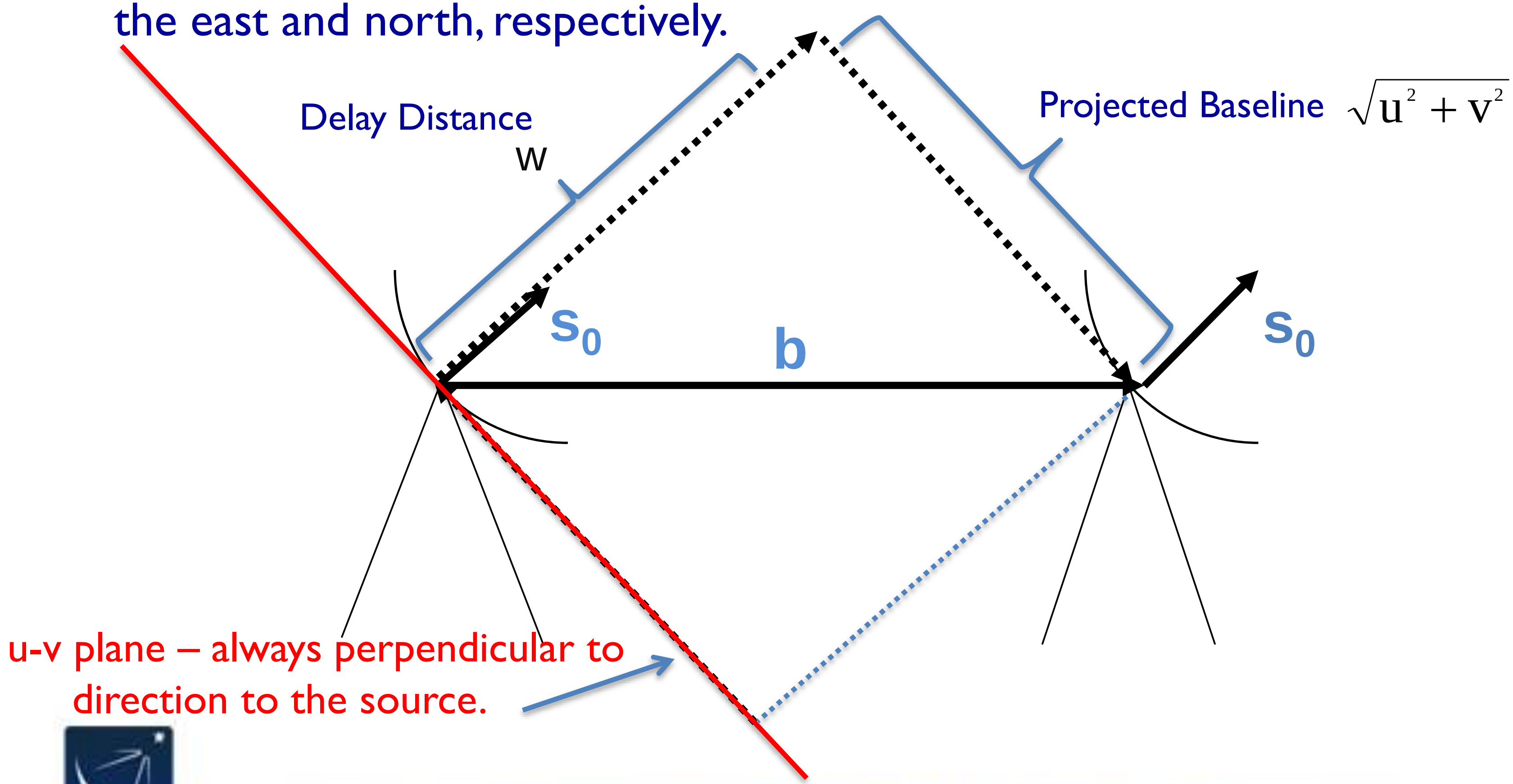
- General arrays (like the VLA) cannot use the 2-d geometry, since the antennas are not along an E-W line.
- Over time, their baselines move through a (u,v,w) volume
- In this case, we adopt a more general description, where all three baseline components are to be considered.
- Arrange 'w' to point to 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:

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

(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:

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

- 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.
- This is practical, in the FOV is sufficiently limited.

3-d to 2-d

- We can get a 2-d FT if the third term in the phase factor is sufficient small:

$$w\left(1-\sqrt{1-l^2-m^2}\right) << 1$$

- Noting that w < B/λ, and if $l^2+m^2=\theta^2 << 1$ one can write the condition for effective coplanarity (i.e., 2-d) as:

$$\theta_{\max} < \sqrt{\frac{2}{\mathrm{w}}} \sim \sqrt{\frac{2\lambda}{\mathrm{B}}} \sim \sqrt{2\theta_{syn}}$$

  - If this condition is met, then the relation between the Intensity and the Visibility again becomes a 2-dimensional Fourier transform:

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

  - But note now that this is no longer an exact relation.

Maximum FOV and Mapsize for given resolution

PSF	FOV	Ratio
0.1"	3.4'	2030
0.3"	5.9'	1172
1.0"	10.7'	642
3.0"	18.6'	371
10"	34'	203
30"	59'	117
100"	108'	64

NRAO

The Problem with Non-coplanar Baselines

- Use of the 2-D transform for non-coplanar interferometer arrays (like the VLA, when used over time) always results in an error in the images.
- We can derive an easy criterion for significant errors by noting the maximum angle for imaging is limited by the primary beam to:

$$\theta_{\max} \sim \lambda / D$$

- This gives us the 'Clark' condition: $$\frac{\lambda B}{D^2} > 1$$
- If this quantity is greater than one, you're in trouble …
- Hence, the problem is most acute for small-diameter antennas (D small) long baselines (B large), and long wavelengths (λ large)
- The problems are not in the principles, but in the cost of the solutions.
- Four solutions (or approximations) to this problem are known:

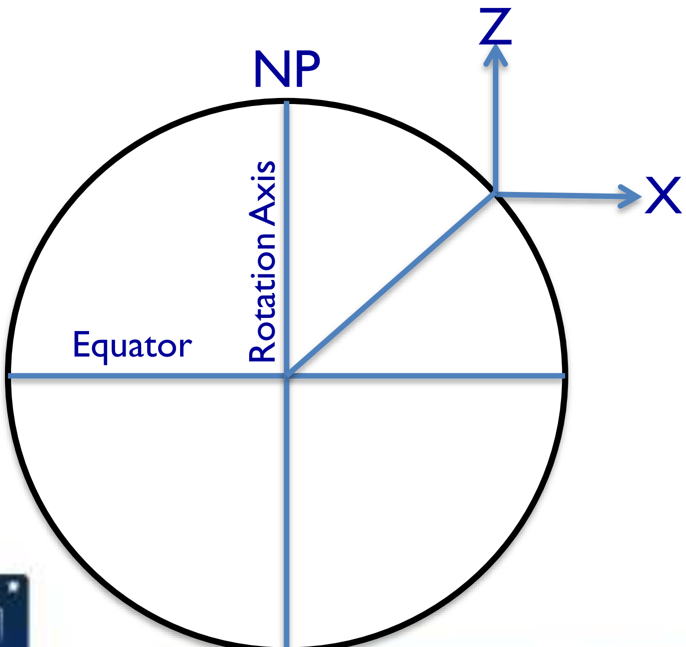
Four Solutions for Non-Coplanar Arrays

1. Do a 3-D transform. This is the **only** correct solution.
   - Grid data in 3-dimensions, and transform to a 3-d cube.
   - The sky appears on a sphere of unit radius. (Cool!)
   - Not a practical solution (~90% of computed cells are empty).
2. If the array is instantaneously coplanar, sum up a series of snapshots.
   - Requires re-projection of each image's coordinates
   - Deconvolution problematic
3. Do facetted imaging – partition the field of view into lots of little images, each of which meets the small-angle criterion.
   - Requires phase offsets and recomputation of baselines for each facet
   - Can apply 'local' calibration for each facet.
4. Project the visibilities onto the w = 0 plane. ('W-Projection').
   - Effectively makes the array fully coplanar.

- Last two approaches are implemented in various imaging packages.

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:
  - 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 and do not change in time.
- 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

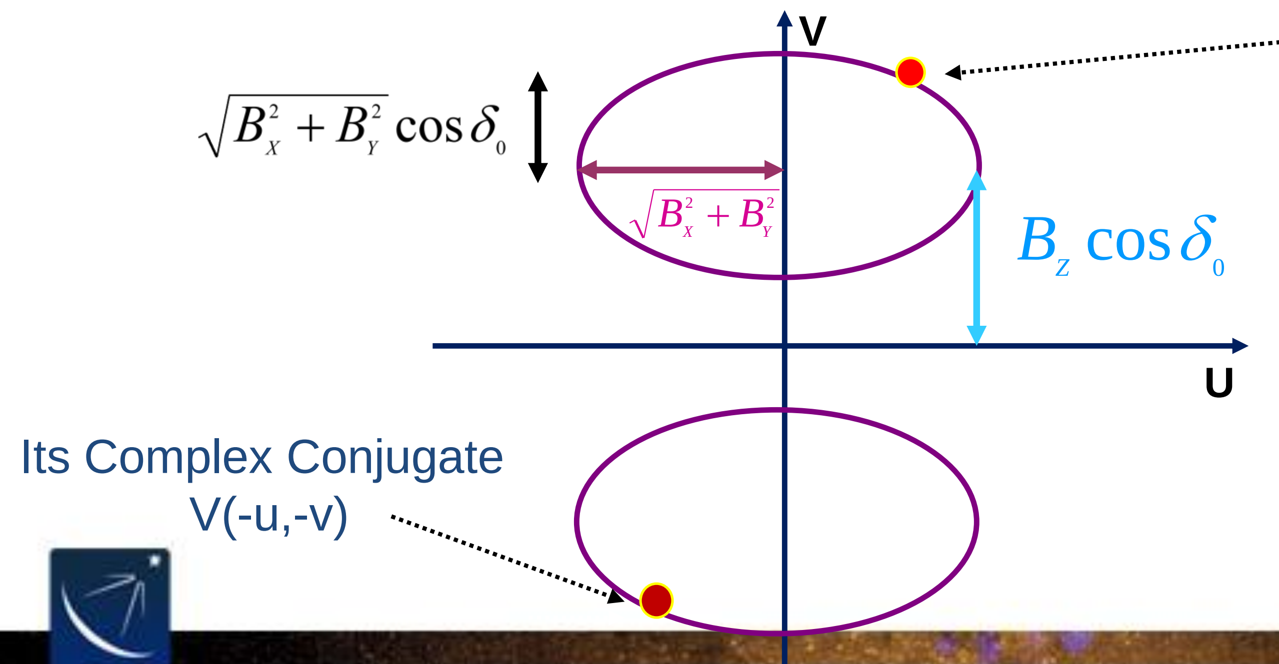
$$\nu_F = \frac{dw}{dt} = -\frac{dH}{dt} u \cos\delta_0 = -\omega_E u \cos\delta_0$$

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 baselines from a few antennas.

E-W Array Coverage and Synthesized Beams

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

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

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

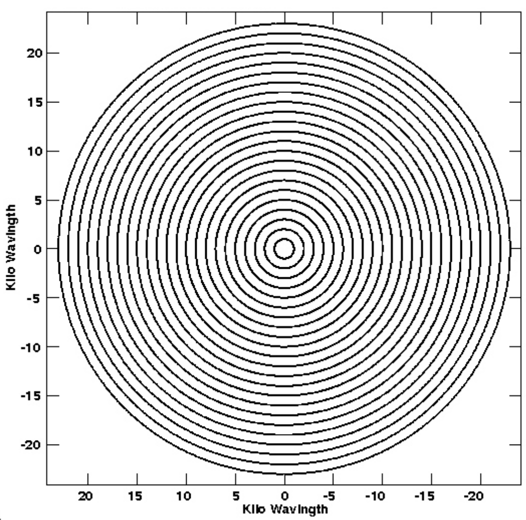
NRAO

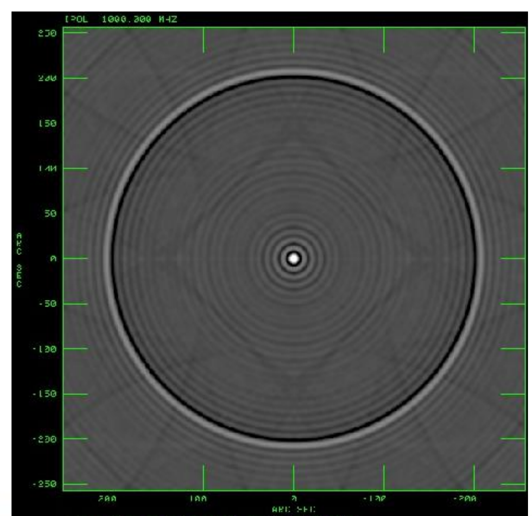
E-W Array Coverage and Beams

- 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.



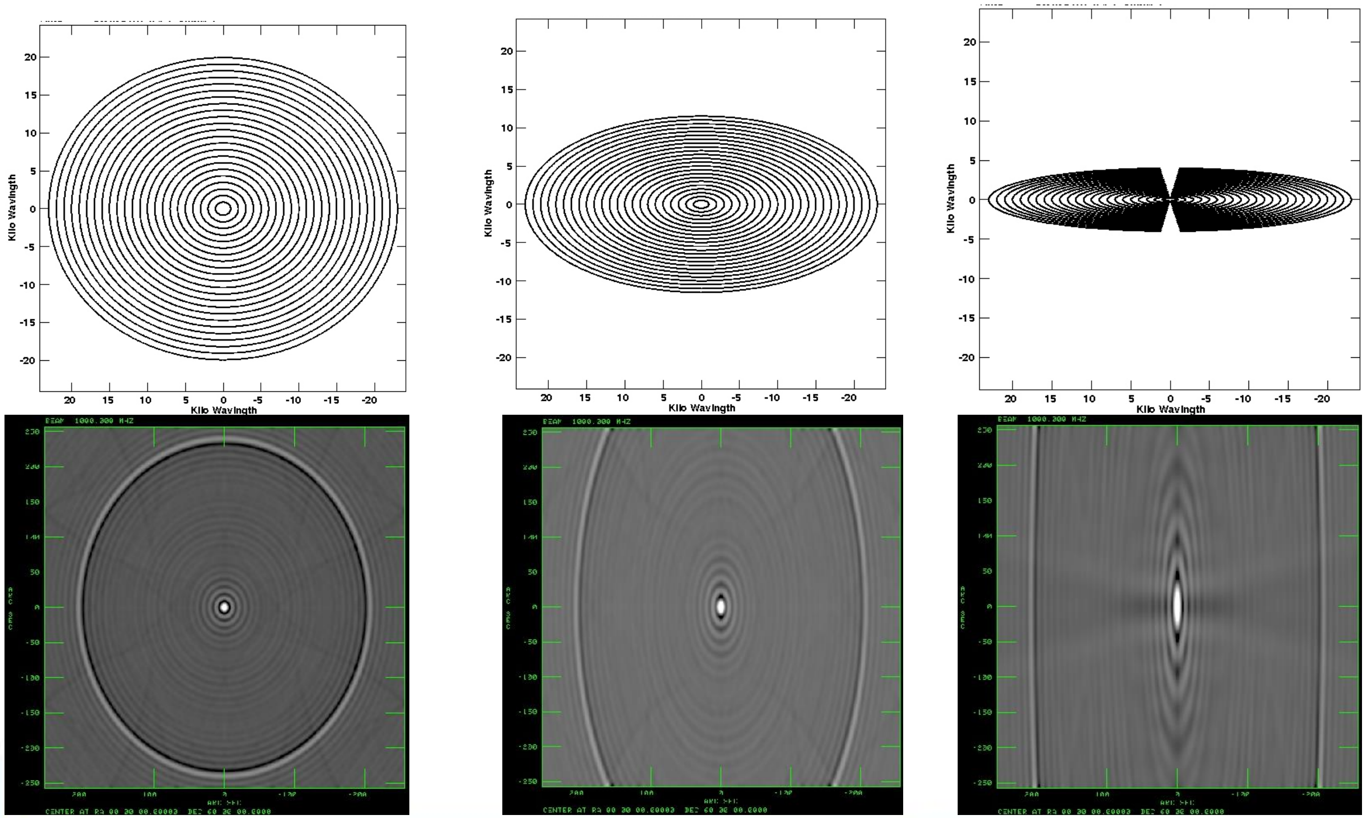


δ = 90

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.

δ=60 δ=30 δ=10



U-V Coverage

PSF

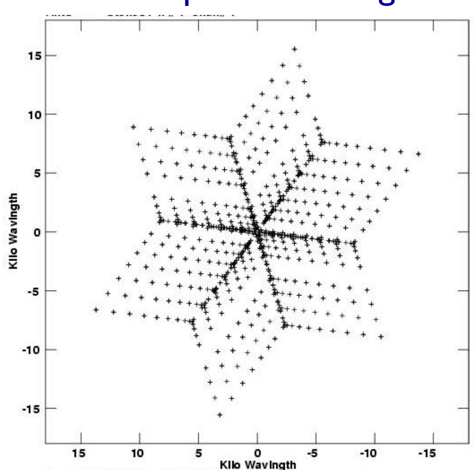
Getting Good Coverage near $\delta = 0$

- 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.

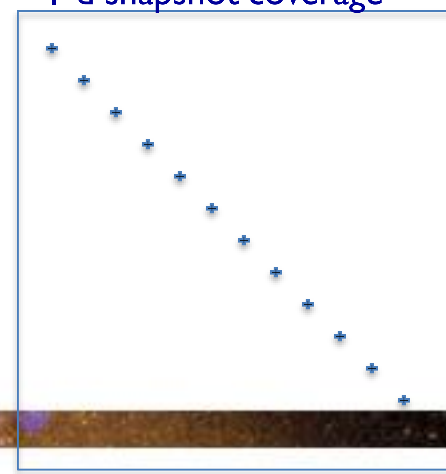
Advantages of 2-d arrays

- 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(`1-d') 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).

2-d snapshot coverage

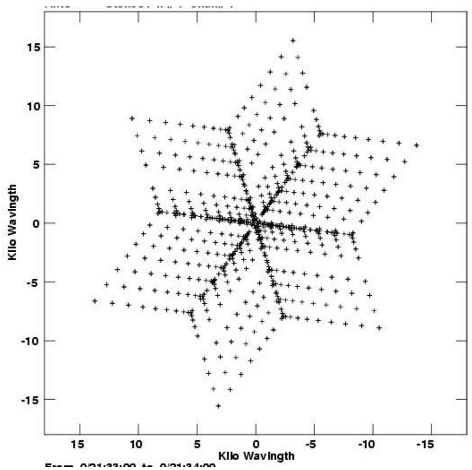


1-d snapshot coverage

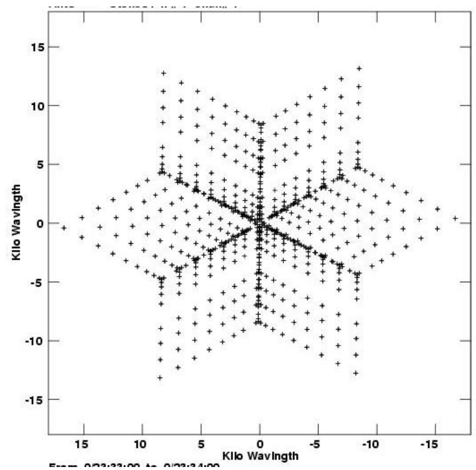


Sample VLA (U,V) plots for 3C147 (δ = 50)

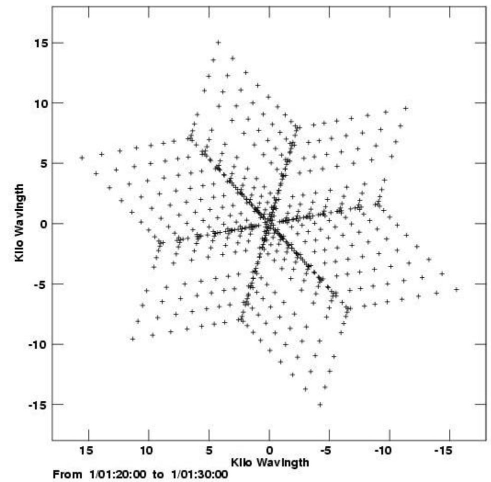
- Snapshot (u,v) coverage for HA = -2, 0, +2 (with 26 antennas).



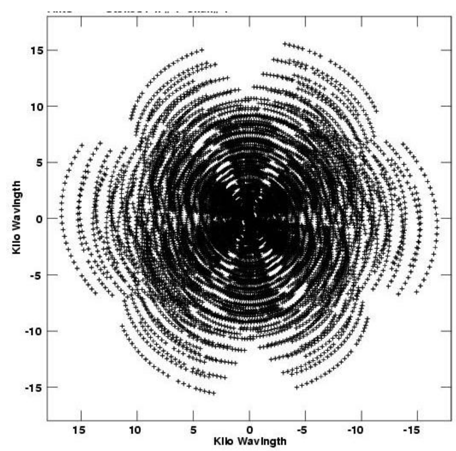
HA = -2h



HA = 0h



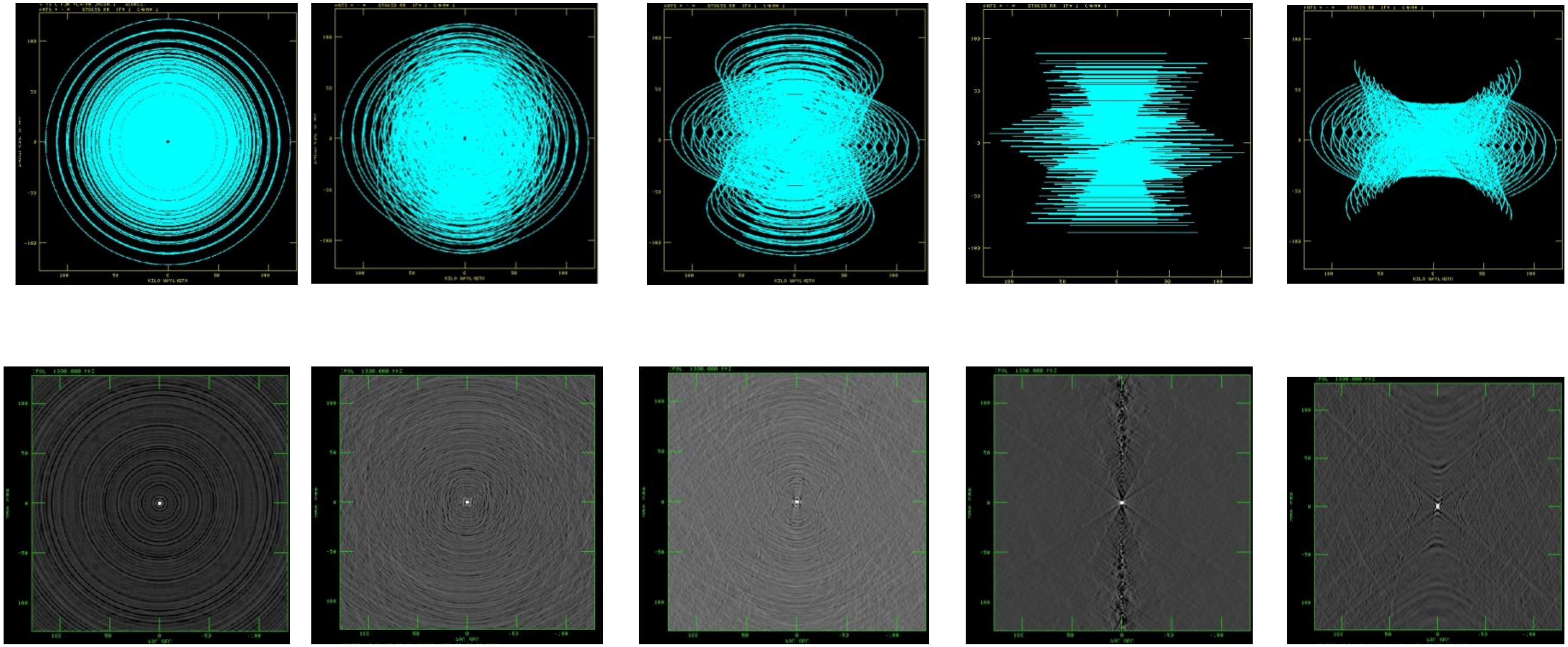
HA = 2h



Coverage over all four hours.

NRAO

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 linear regularity.

Examples with Real Data!

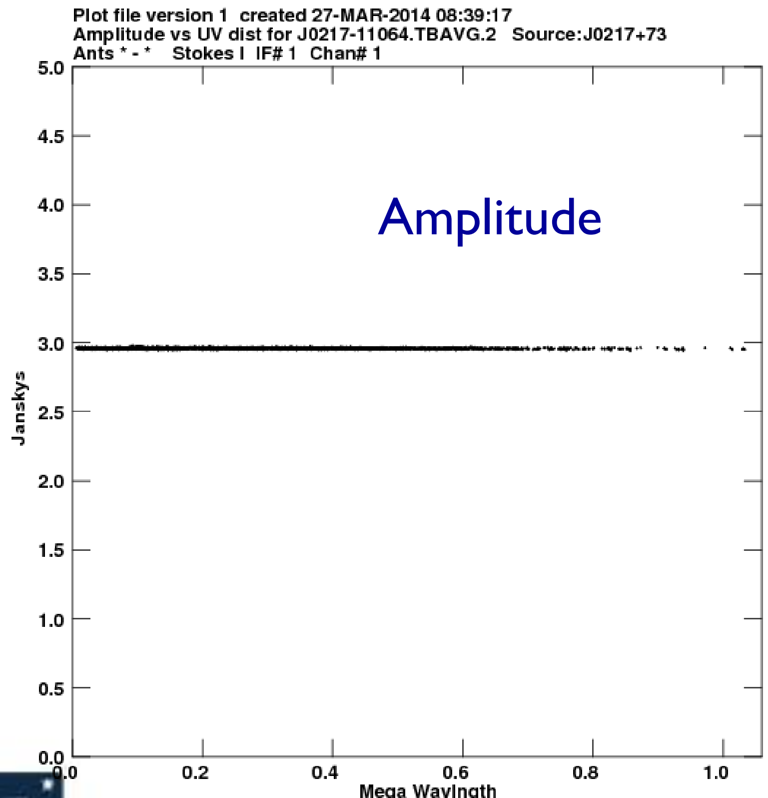
- Enough of the analysis!
- I close with some examples from real observations, using the VLA.
- These are two-dimensional observations (function of 'u' (EW) and 'v' (NS) baselines).
- Plotted are the visibility amplitudes version baseline length:

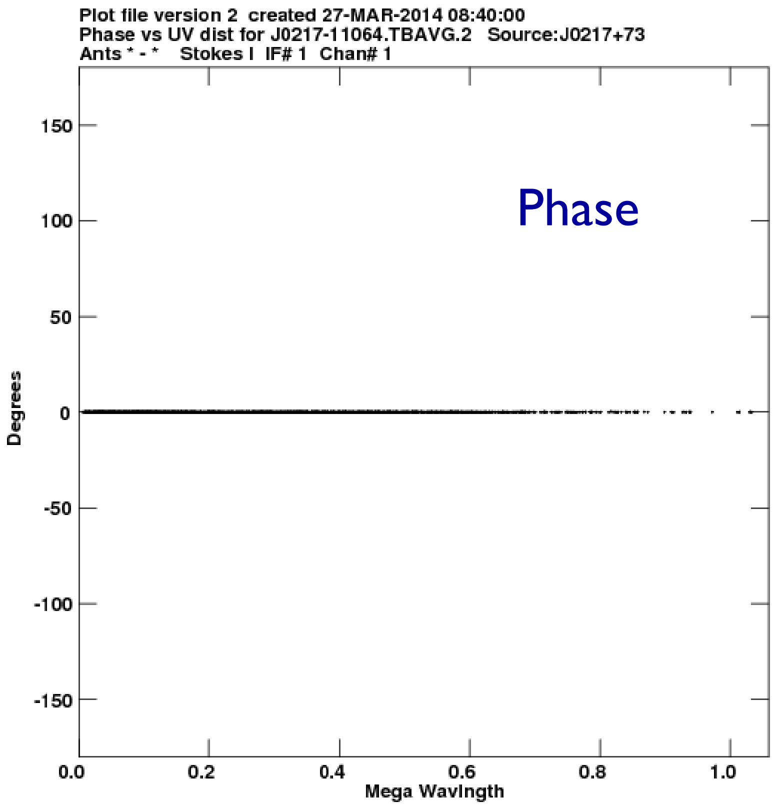
$$q = \sqrt{\mathrm{u}^2 + \mathrm{v}^2}$$

- 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

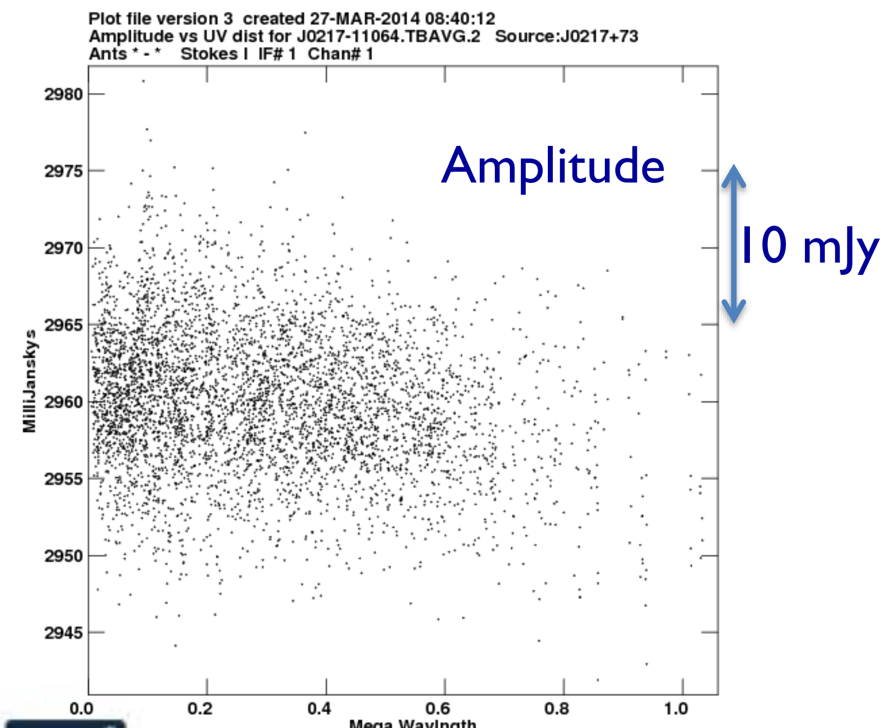
- Suppose we observe an unresolved object, at the phase center.
- What is its visibility function?

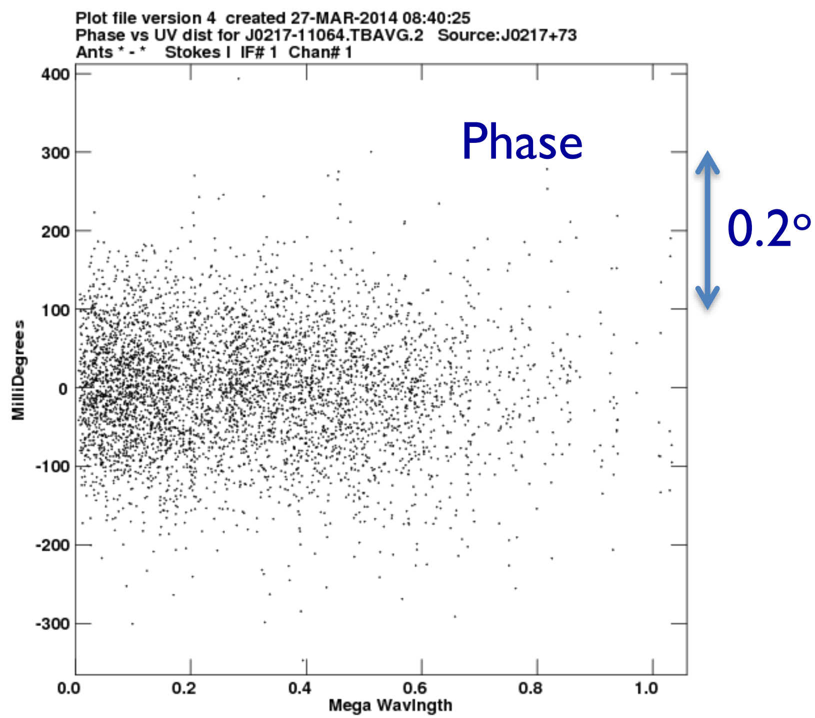




Zoom in to see the noise …

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

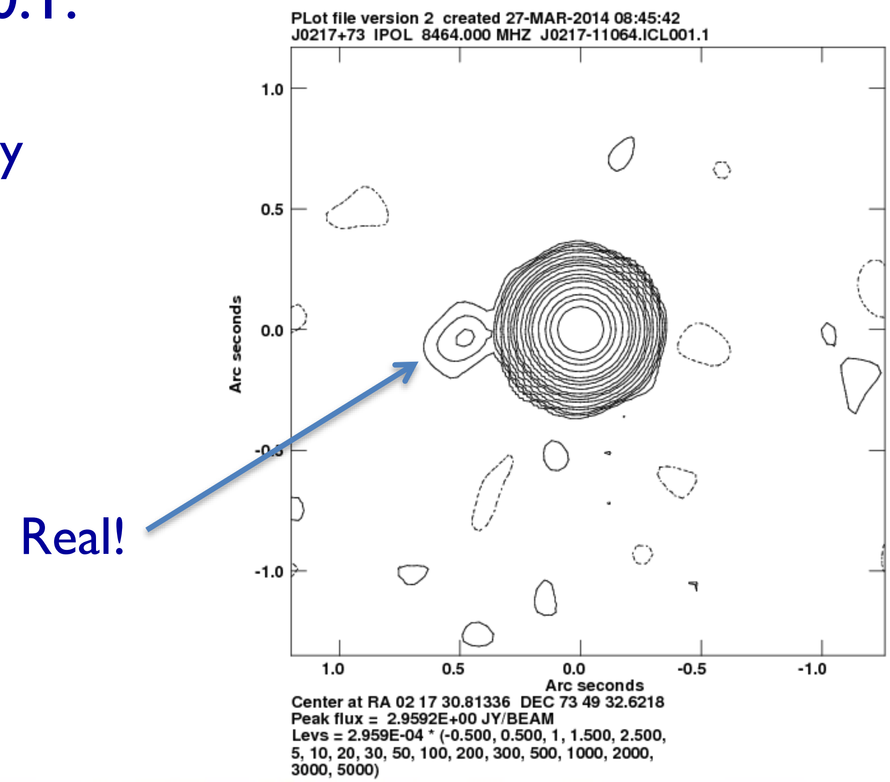




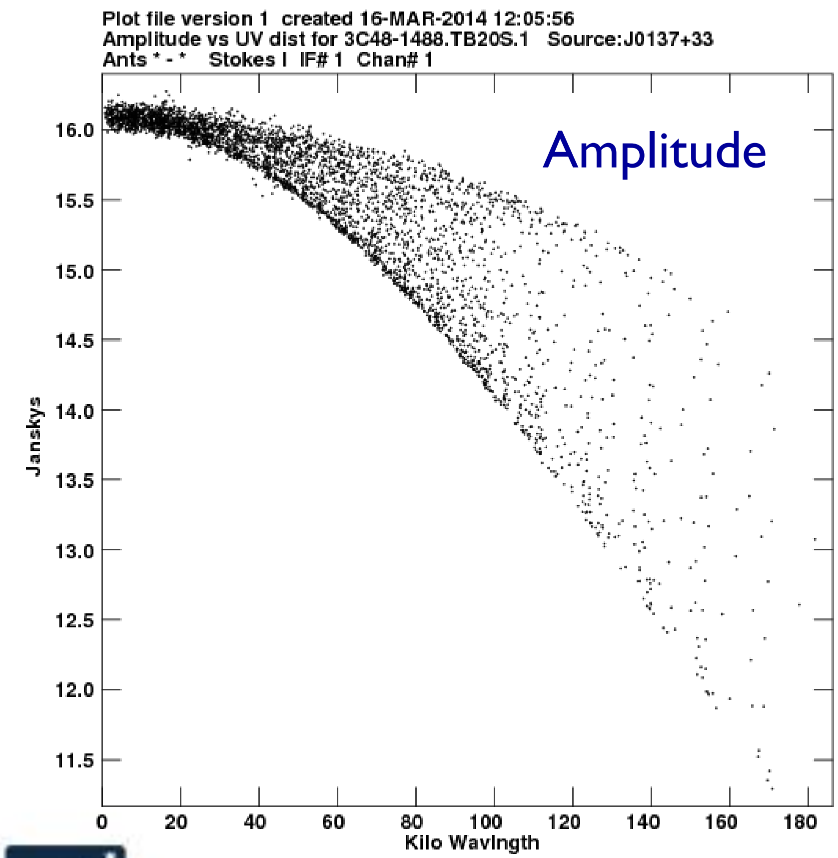
And the Map …

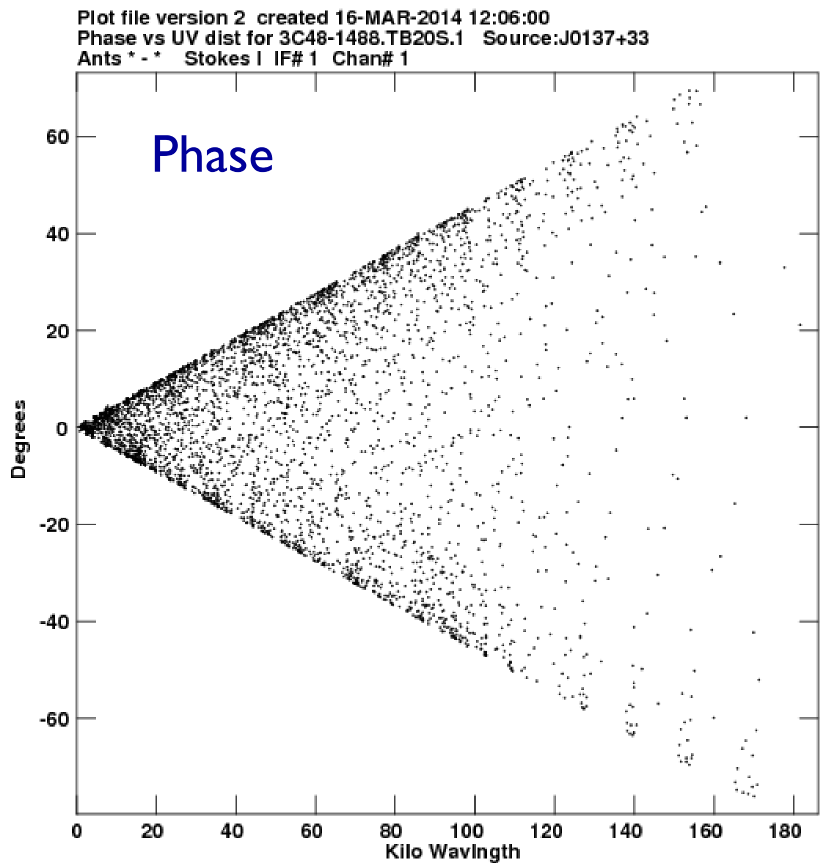
- The source is unresolved … but with a tiny background object.
- Dynamic range: 50,000:1.

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



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



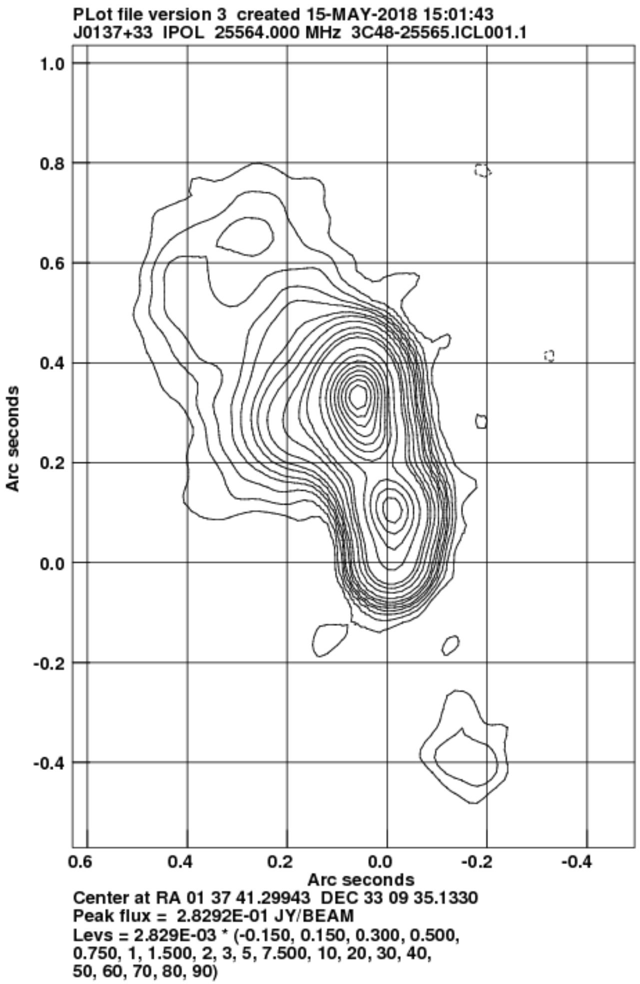


3C48 position and offset

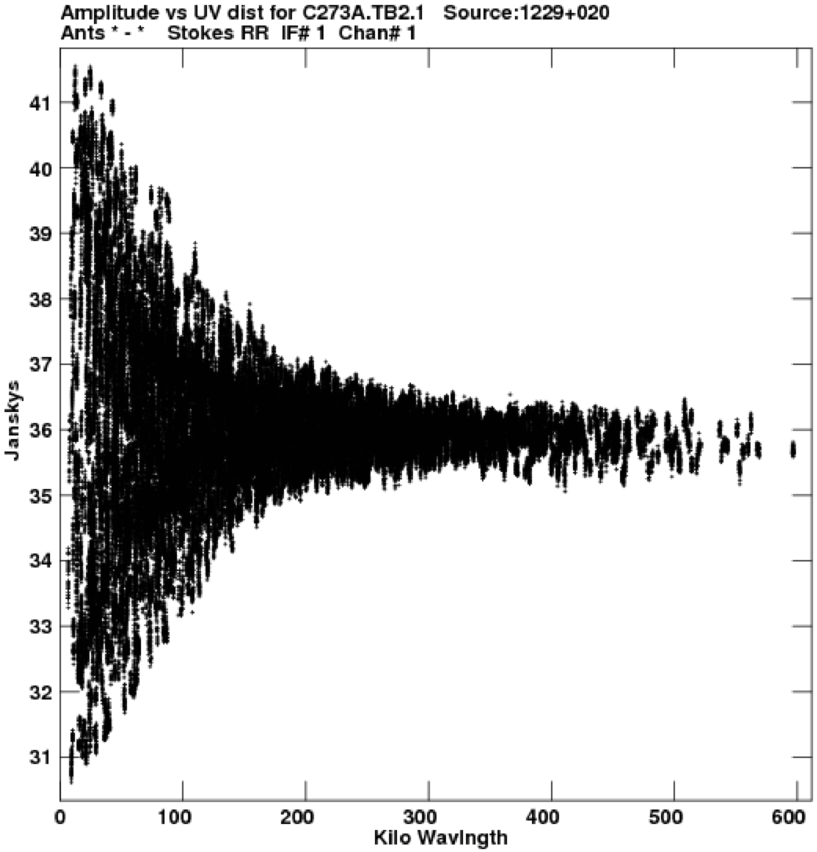
- You can learn quite a bit just by looking at the gross properties of the visibility amplitudes/phases, noting:
- **A 206265 wavelength baseline makes a 1 arcsecond fringe.**
- The linear phase slope is ~90 degrees over 180,000 wavelengths.
  - 90 degrees is ¼ of a fringe, and one fringe is one arcsecond. 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.

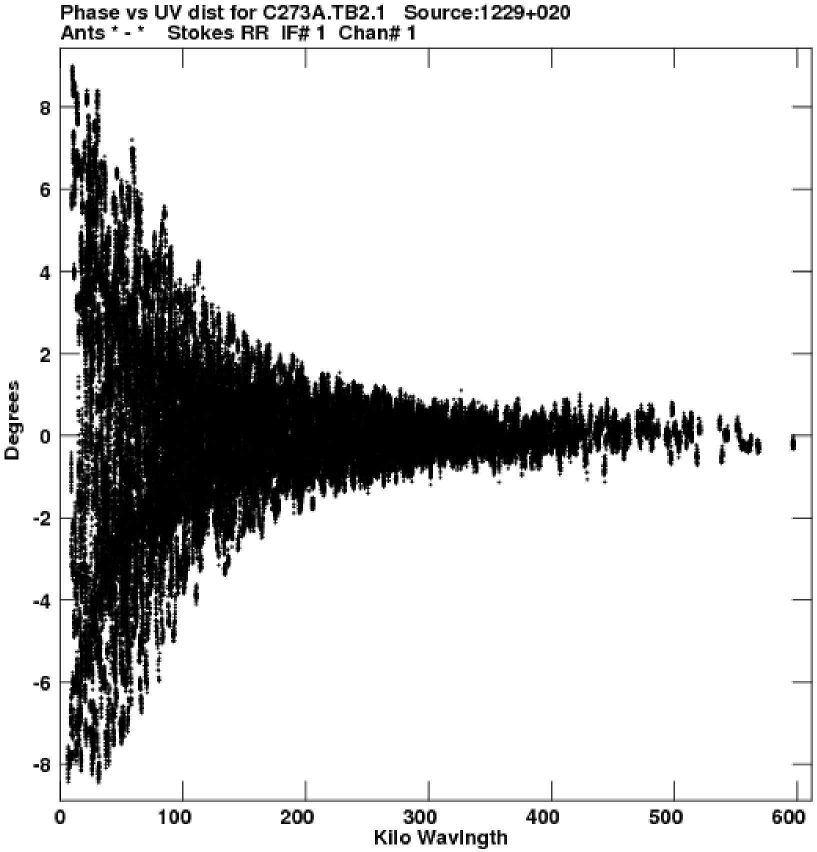
3C48 Structure (at 25 GHz …)

- 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

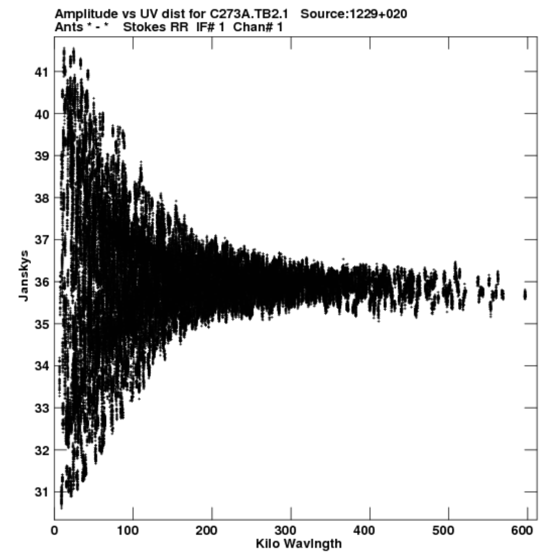


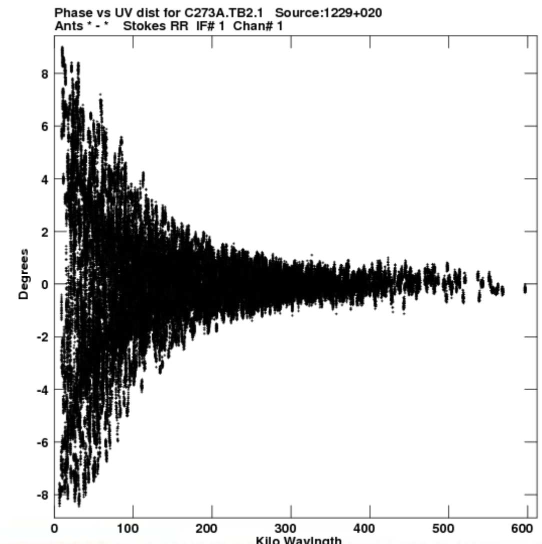


Visibility Amplitude **Visibility Phase**

3C273 – jet size revealed …

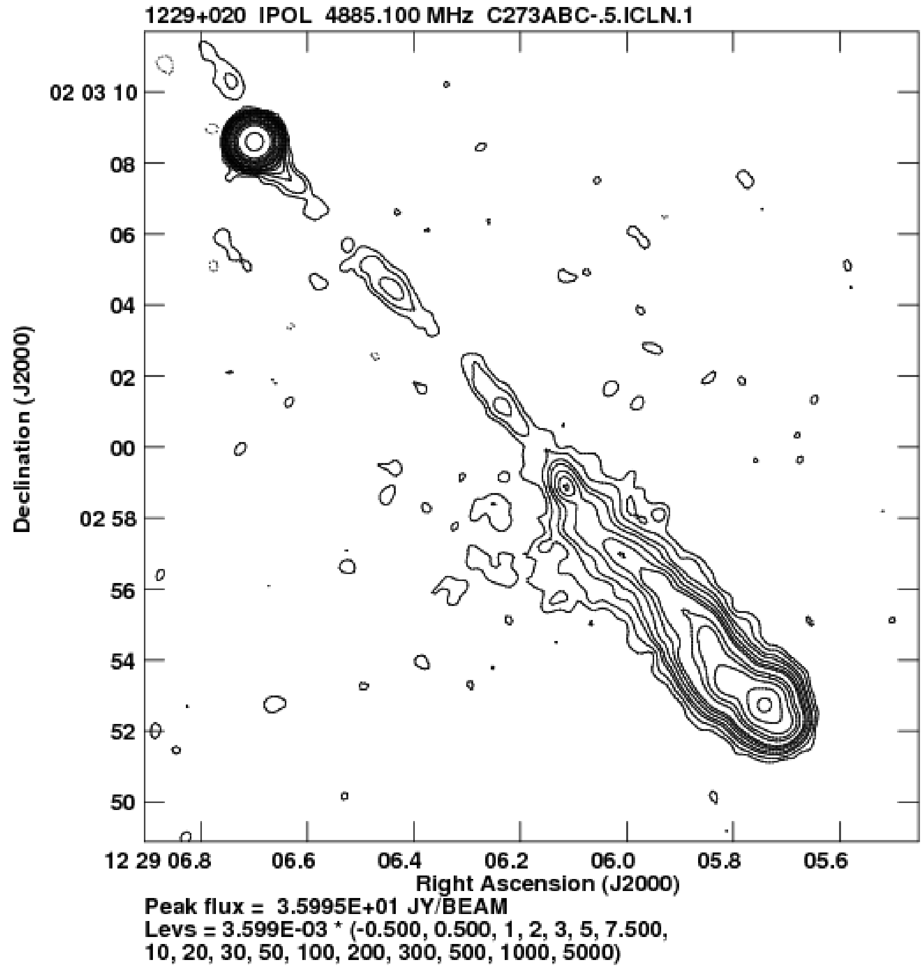
- The narrowing of the visibility amplitudes, with no sign resolution on the long spacings, tells us:
  - There a 36 Jy unresolved `point'-source.
- The absence of a phase gradient on the long spacings tells us:
  - The point source is at the center
- The extended emission resolves out at ~100 -- 200 Kλ
- This indicates an extended structure with width of ~1 to 2”
- Rapid oscillations in amplitude and phase tells us:
  - A much larger elongation is also present



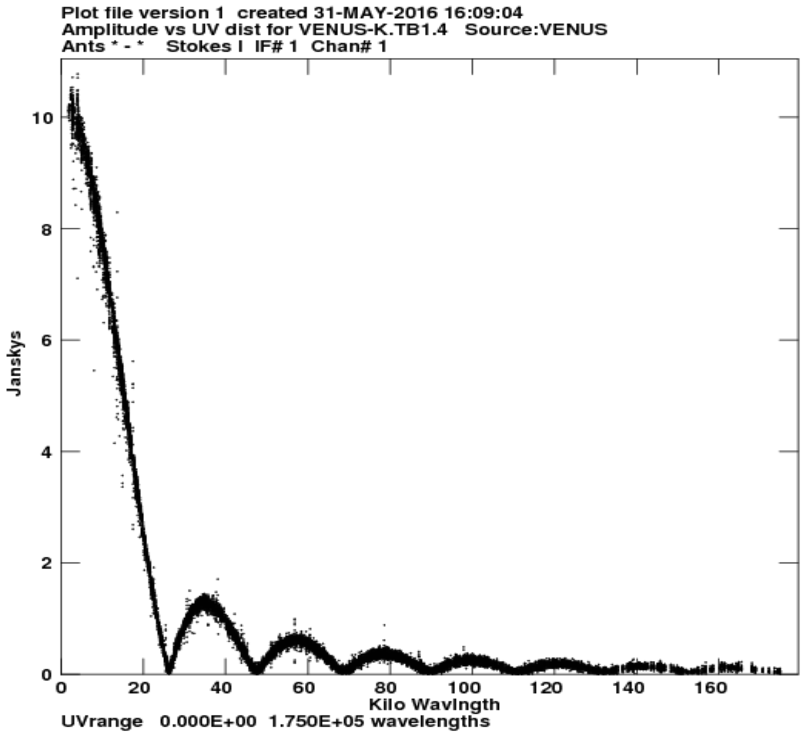


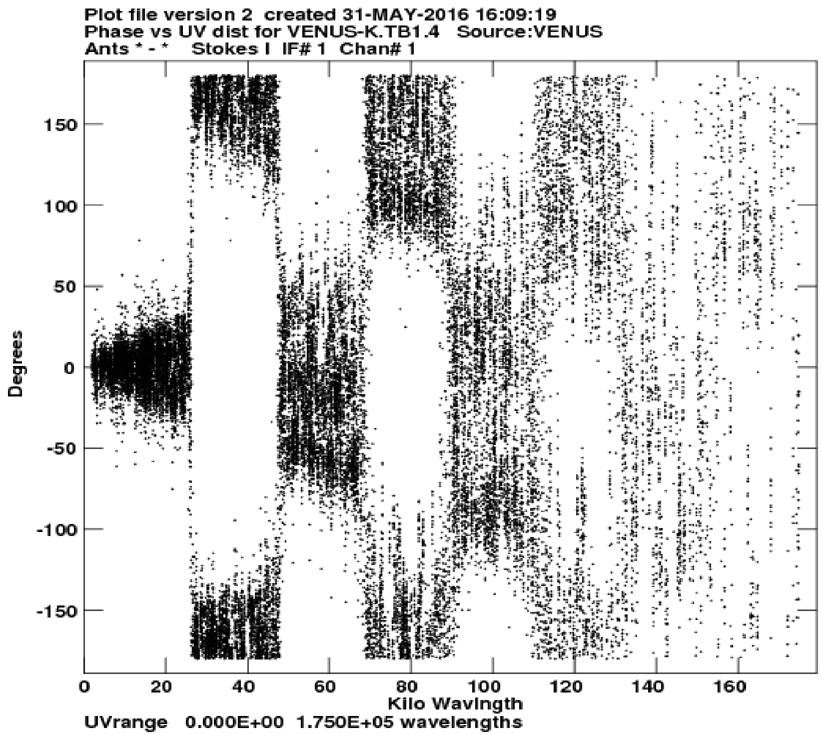
3C273 -- Image

- 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.

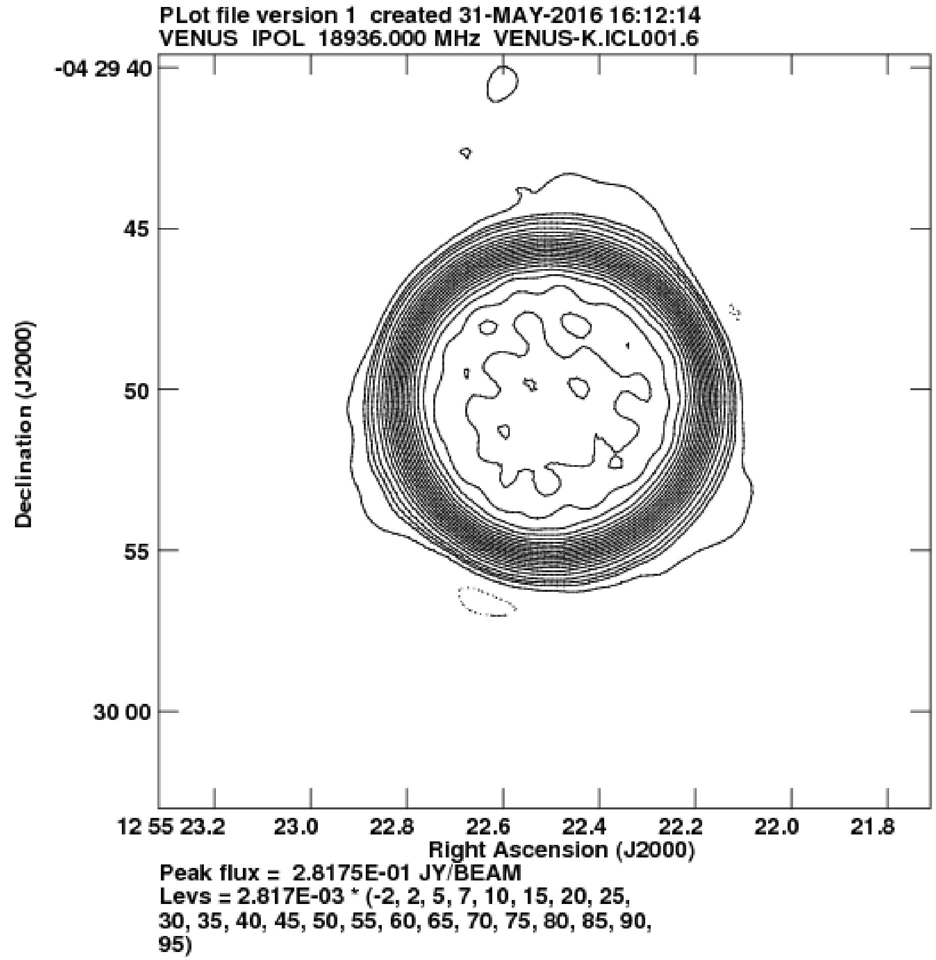




- The visibilities are circularly symmetric. The phases alternate between zero and 180 degrees.
- The source must be circularly symmetric and centered.
- The visibility null at 25 kλ indicates angular size of ~ 10 arcseconds.

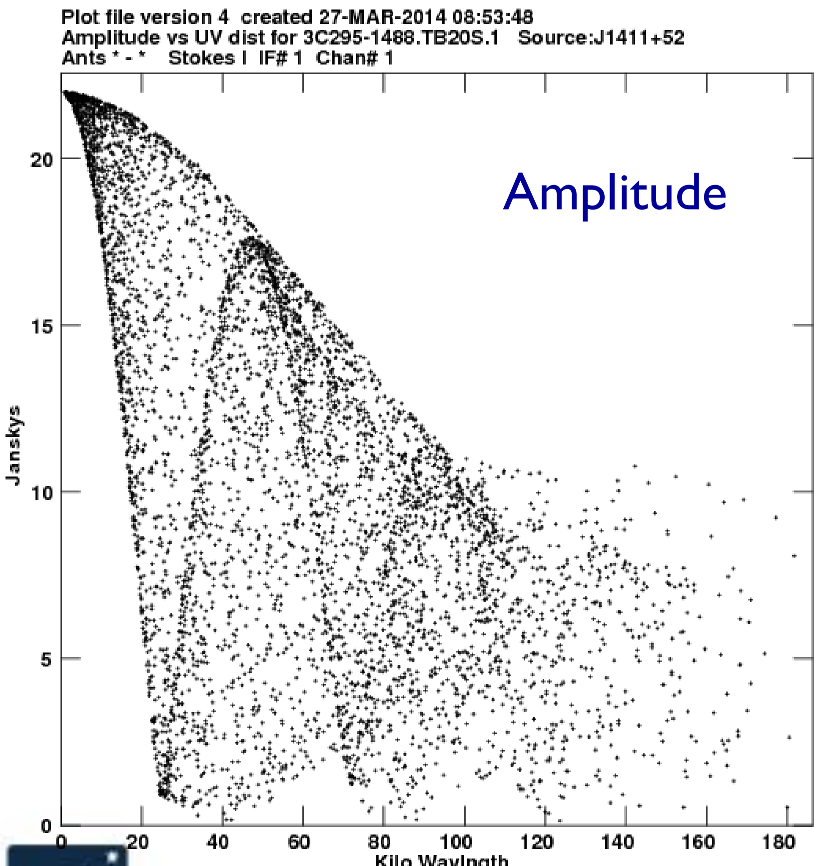
And the image looks like:

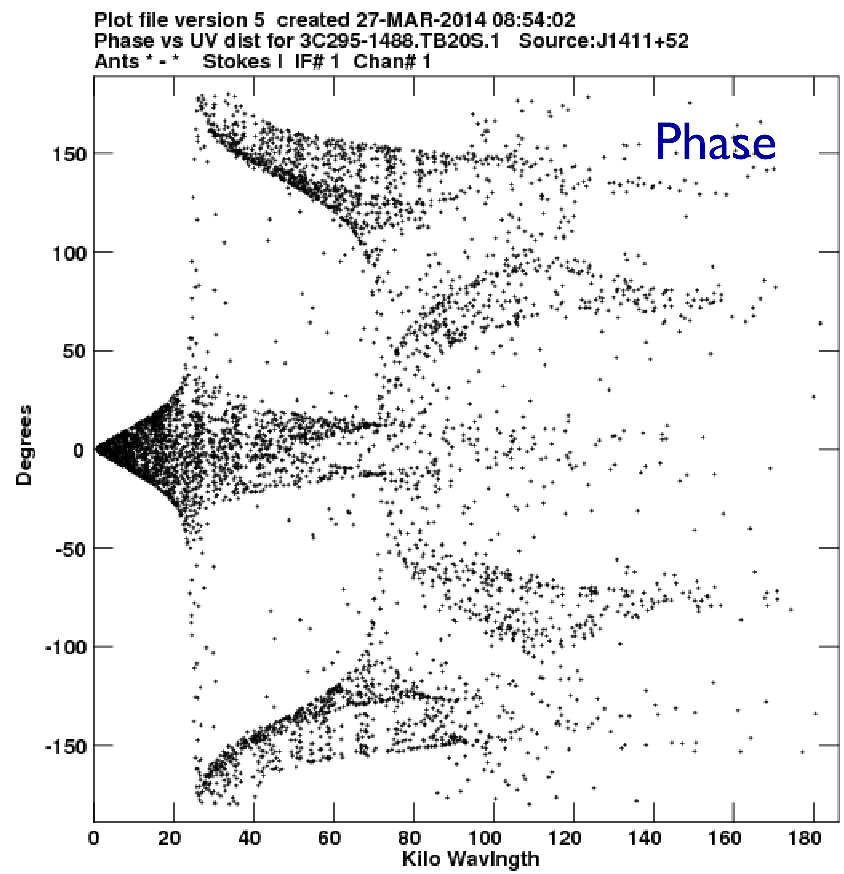
- It's a perfectly uniform, blank disk!
- The Visibility function, in fact, is an almost perfect Bessel function of zero order: $J_0(q)$.
- A perfect J_0 would arise from a perfectly sharp disk. Atmospheric opacity effects 'soften' the edge, resulting in small deviations from the J_0 function at large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



Examples of Visibilities – a Well Resolved Object

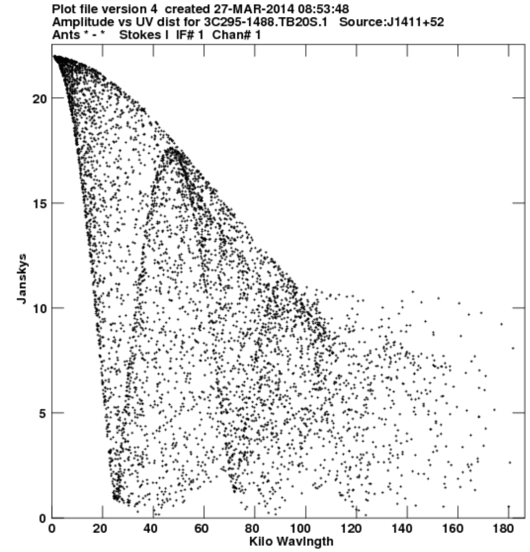
- The flux calibrator 3C295 at 1400 MHz.

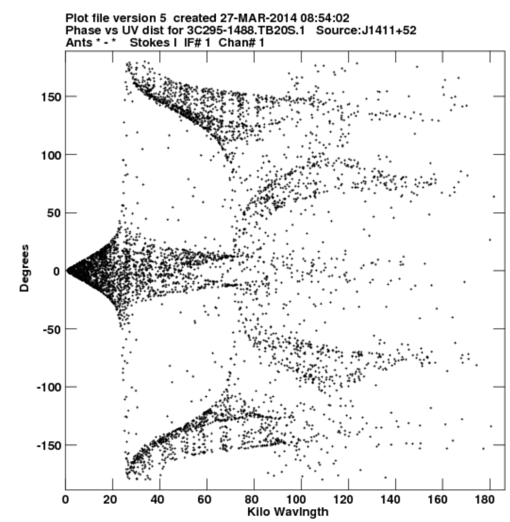




From the 1-d visibilities alone, we deduce

- 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.
- An unresolved component *may* be present, as the longest spacing visibility amplitudes remain high.





3C295 Image

- A double source, with separation of about four arceconds.
- A width of 1 – 2 arcseconds.
- The resolution (psf) is not high enough to detect an unresolved component.
- (But, higher resolution data show the presence of compact 'hot spots').