Radio Interferometry -- II

Rick Perley, NRAO/Socorro

ATNF School on Radio Astronomy

Narrabri, NSW

29 Sept – 3 Oct, 2014

Topics

- Practical Extensions to the Theory:
  - Finite bandwidth
  - Rotating reference frames (source motion)
  - Finite time averaging
  - Local Oscillators and Frequency Downconversion
- Coordinate Systems
  - Direction Cosines
  - 2d and 3d measurement space
- Example of Visibilities from Simple Sources

Review

- In the previous lecture, I set down the principles of Fourier synthesis imaging.
- I showed:

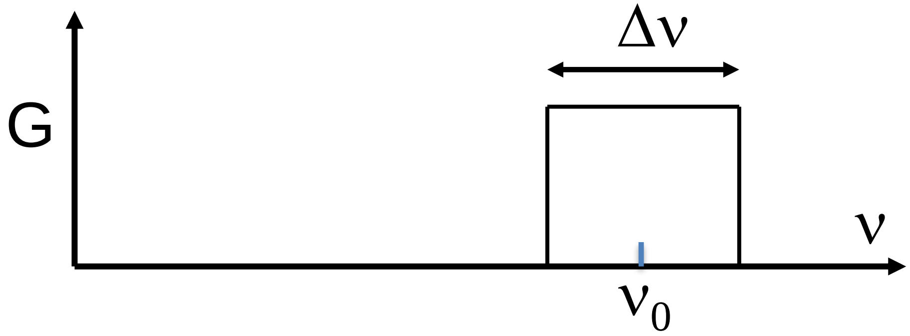
$$V_\upsilon(\mathbf{b}) = R_C - iR_S = \iint I_\nu(s) e^{-2\pi i\nu \mathbf{b}\cdot\mathbf{s}/c} d\Omega$$

  Where the intensity I_ν is a real function, and the visibility V(**b**) is complex and Hermitian.

- The model used for the derivation was idealistic – not met in practice:
  - Monochromatic
  - Stationary reference frame.
  - No time averaging
- We now relax, in turn, these restrictions.

The Effect of Bandwidth.

- Real interferometers must accept a range of frequencies. So we now consider the response of our interferometer over frequency.
- Define the frequency response function, $G(\nu)$, as the amplitude and phase variation of the signal over frequency.



- The function $G(\nu)$ is primarily due to the gain and phase characteristics of the electronics, but can also contain propagation path effects.
- In general, $G(\nu)$ is a complex function.

The Effect of Bandwidth.

- To find the finite-bandwidth response, we integrate our fundamental response over a frequency width $\Delta\nu$, centered at ν_0:

$$V = \int \left(\frac{1}{\Delta\nu} \int_{\upsilon_0-\Delta\upsilon/2}^{\upsilon_0+\Delta\upsilon/2} I(\mathbf{s},\upsilon) G_1(\upsilon) G_2^*(\upsilon) e^{-i2\pi\upsilon\tau_g} d\upsilon \right) d\Omega$$

- If the source intensity does not vary over the bandwidth, and the instrumental gain parameters G_1 and G_2 are square and identical, then

$$V = \iint I_\upsilon(\mathbf{s}) \frac{\sin(\pi\tau_g\Delta\upsilon)}{\pi\tau_g\Delta\upsilon} e^{-2i\pi\nu_0\tau_g} d\Omega = \iint I_\upsilon(\mathbf{s}) \mathrm{sinc}(\tau_g\Delta\upsilon) e^{-2i\pi\nu_0\tau_g} d\Omega$$

where the **fringe attenuation function**, sinc(x), is defined as:

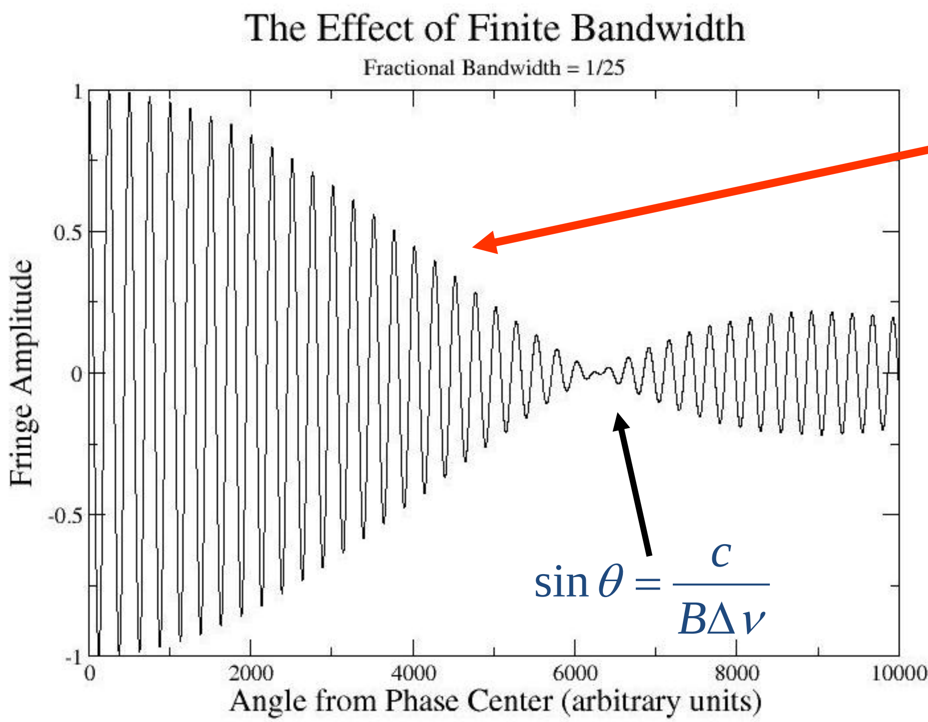
$$\mathrm{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$$

Bandwidth Effect Example

- For a square bandpass, the bandwidth attenuation reaches a null when $\tau_g \Delta\nu = 1$, or

$$\sin\theta = \frac{c}{B\Delta\nu} = \left(\frac{\lambda}{B}\right)\left(\frac{\upsilon_0}{\Delta\upsilon}\right)$$

- For the old VLA, and its 50 MHz bandwidth, and for the 'A' configuration, (B = 35 km), the null was ~1.3 degrees away.
- For the upgraded VLA, $\Delta\nu$ = 2 MHz, and B = 35 km, then the null occurs at about 27 degrees off the meridian.



Fringe Attenuation function:

$$\mathrm{sinc}\left(\frac{B}{\lambda}\frac{\Delta\upsilon}{\upsilon}\sin\theta\right) = \mathrm{sinc}\left(\frac{B\Delta\upsilon}{c}\sin\theta\right)$$

Note: The fringe-attenuation function depends only on bandwidth and baseline length – not on frequency.

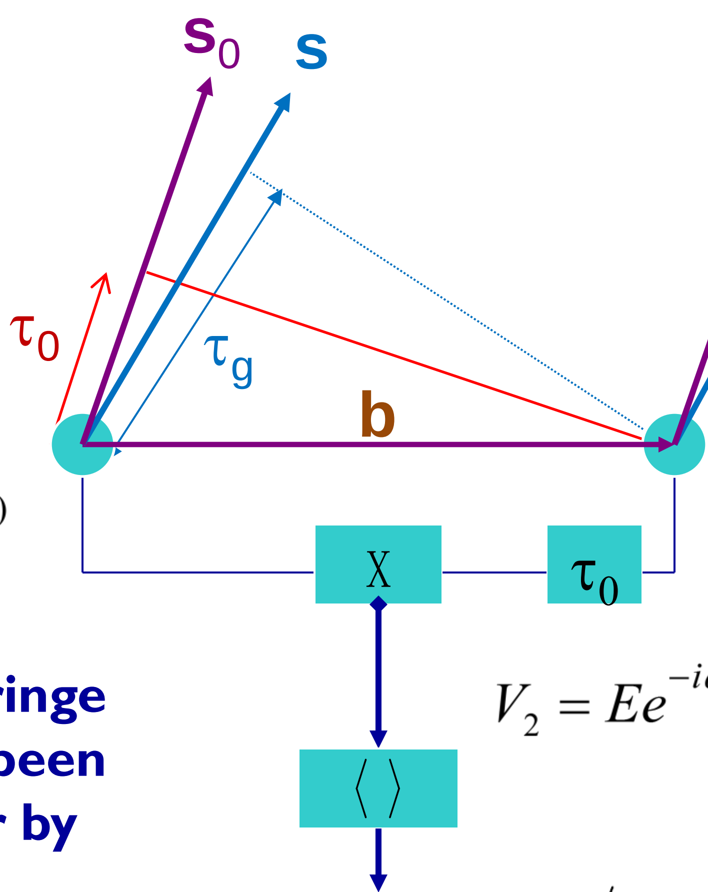
Observations off the Baseline Meridian

- In our basic scenario -- stationary source, stationary interferometer -- the effect of finite bandwidth will strongly attenuate the visibility from sources far from the meridional plane.
- Since each baseline has its own fringe pattern, the only point on the sky free of attenuation for all baselines is a small angle around the zenith (presuming all baselines are coplanar).
- Suppose we wish to observe an object far from the zenith?
- One solution is to use a very narrow bandwidth – this loses sensitivity, which can only be made up by utilizing many channels – feasible, but computationally expensive.
- Better answer: Shift the fringe-attenuation function to the center of the source of interest.
- How? By adding time delay.

Adding Time Delay

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

$$\tau_0 = \mathbf{b}\cdot\mathbf{s}_0/c$$



$\mathbf{s}_0$ = reference (delay) direction

S = general direction

$$V_1 = Ee^{-i\omega(t-\tau_g)}$$

$$V_2 = Ee^{-i\omega t}$$

$$V_2 = Ee^{-i\omega(t-\tau_0)}$$

The entire fringe pattern has been shifted over by angle

sin θ = cτ$_0$/b

$$V = \langle V_1 V_2^* \rangle = E^2 e^{-i[\omega(\tau_0-\tau_g)]}$$

$$= E^2 e^{i2\pi[\nu\mathbf{b}\cdot(\mathbf{s}-\mathbf{s}_0)/c]}$$

NRAO

Observations from a Rotating Platform

- To follow a moving source with minimal loss of coherence, we simply add in delay to match the changing geometric delay.

- To minimize bandwidth loss, the delay difference must be less than $\delta\tau << 1/\Delta\nu$. (Typically, microseconds).

- For the 'radio-frequency' interferometer we are discussing here, this will automatically track both the fringe pattern and the fringe-washing function with the source.

- To hold the phase difference to much less than a radian, a more stringent condition arises: $\delta\tau << 1/\nu$. (Typically, nanoseconds). Note that the residual phase error from an incorrect delay can be corrected for following correlation).

- By inserting the appropriate delay, a moving point source, at the reference position, will give uniform amplitude and zero phase throughout time (provided real-life things like the atmosphere, ionosphere, or geometry errors don't mess things up … ☺)

NRAO

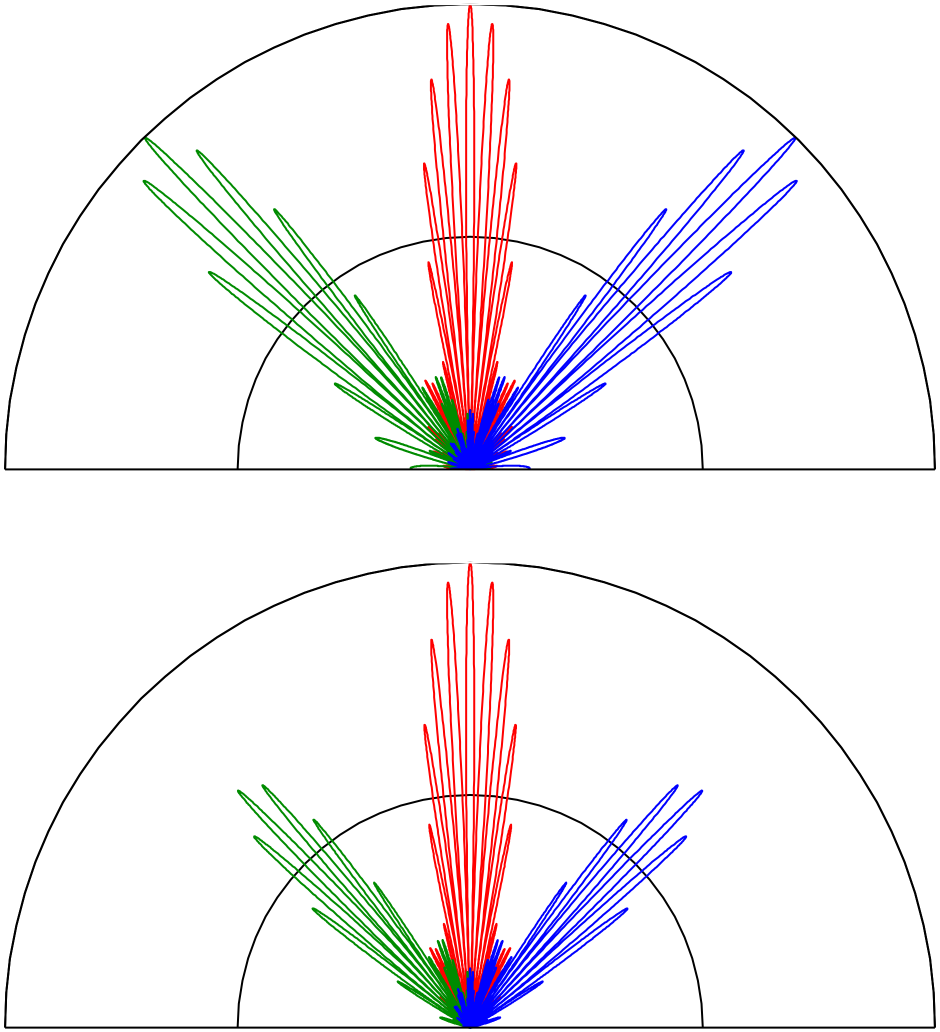
Illustrating Delay Tracking

- Top Panel:

  Delay has been added and subtracted to move the delay pattern to the source location.

- Bottom Panel:

  A cosinusoidal sensor pattern is added, to illustrate losses from a fixed sensor.



Another Justification for Delay Tracking

- There is another very good reason to 'track' the fringe pattern by adding time delay.
- The 'natural fringe rate' – due to earth's rotation, is given by

$$\upsilon_f = u\omega_e \cos\delta \quad \text{Hz}$$

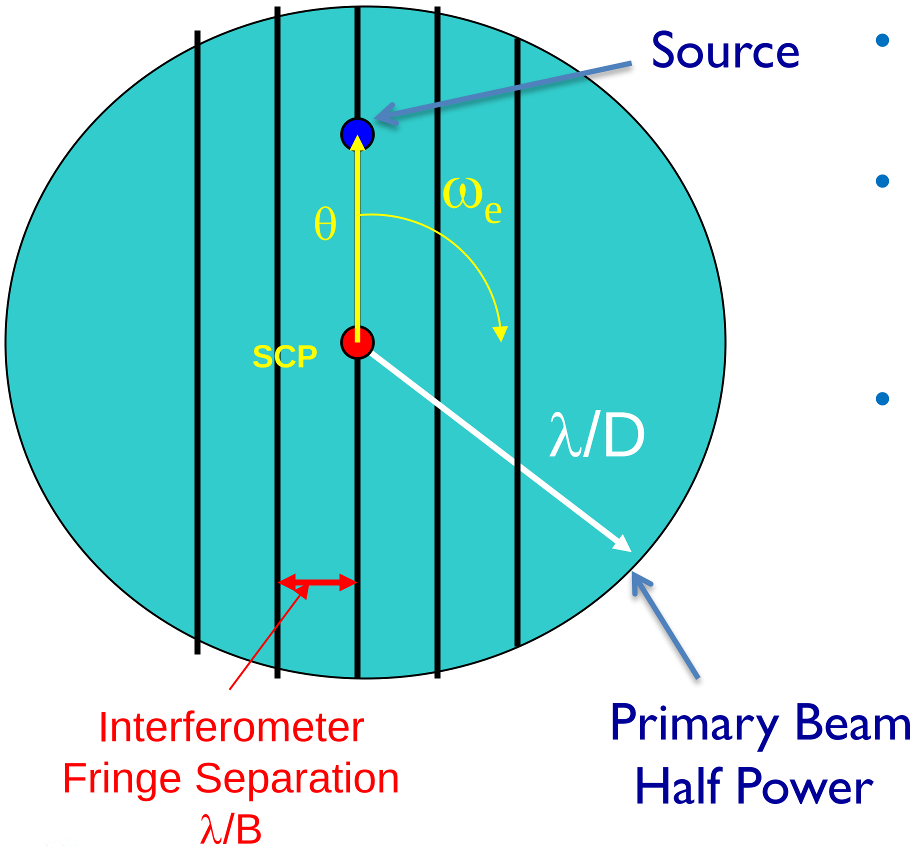
- Where $u = B/\lambda$, the (E-W) baseline in wavelengths, and $\omega_e = 7.3\times10^{-5}$ rad/s is the angular rotation rate of the earth.
- For a million-wavelength baseline, $\nu_f \sim 70$ Hz – that's fast.
- There is *no* useful information in this fringe rate – it's simply a manifestation of the platform rotation (indeed, it's a Doppler shift).
- Tracking, or 'stopping' the fringes greatly slows down the *post-correlation* data processing/archiving needs.

Time Averaging Loss

- So – we can track a moving source, continuously adjusting the delay to move the fringe pattern with the source.
- This does two good things:
  – Slows down the data recording needs
  – Prevents bandwidth delay losses.
- From this, you might think that you can increase the time averaging for as long as you please.
- But you can't – because the convenient tracking only works perfectly for the object 'in the center' – the point for which the delays have been pre-set.
- All other sources are moving w.r.t. the fringe pattern – and this is where the essential information lies…

Time-Smearing Loss Timescale

Simple derivation of fringe period, from observation at the SCP.



- Turquoise area is antenna primary beam on the sky – radius = λ/D
- Interferometer coherence pattern has spacing = λ/B
- Sources in sky rotate about NCP at angular rate:

  $\omega_e = 7.3\times10^{-5}$ rad/sec.
- Minimum time taken for a source to move by λ/B at angular distance θ is:

$$t = \frac{\lambda}{B}\frac{1}{\omega_E \sin\theta}$$

$$\approx \frac{D}{\omega_E B}$$

For sources at the half power distance

Time-Averaging Loss

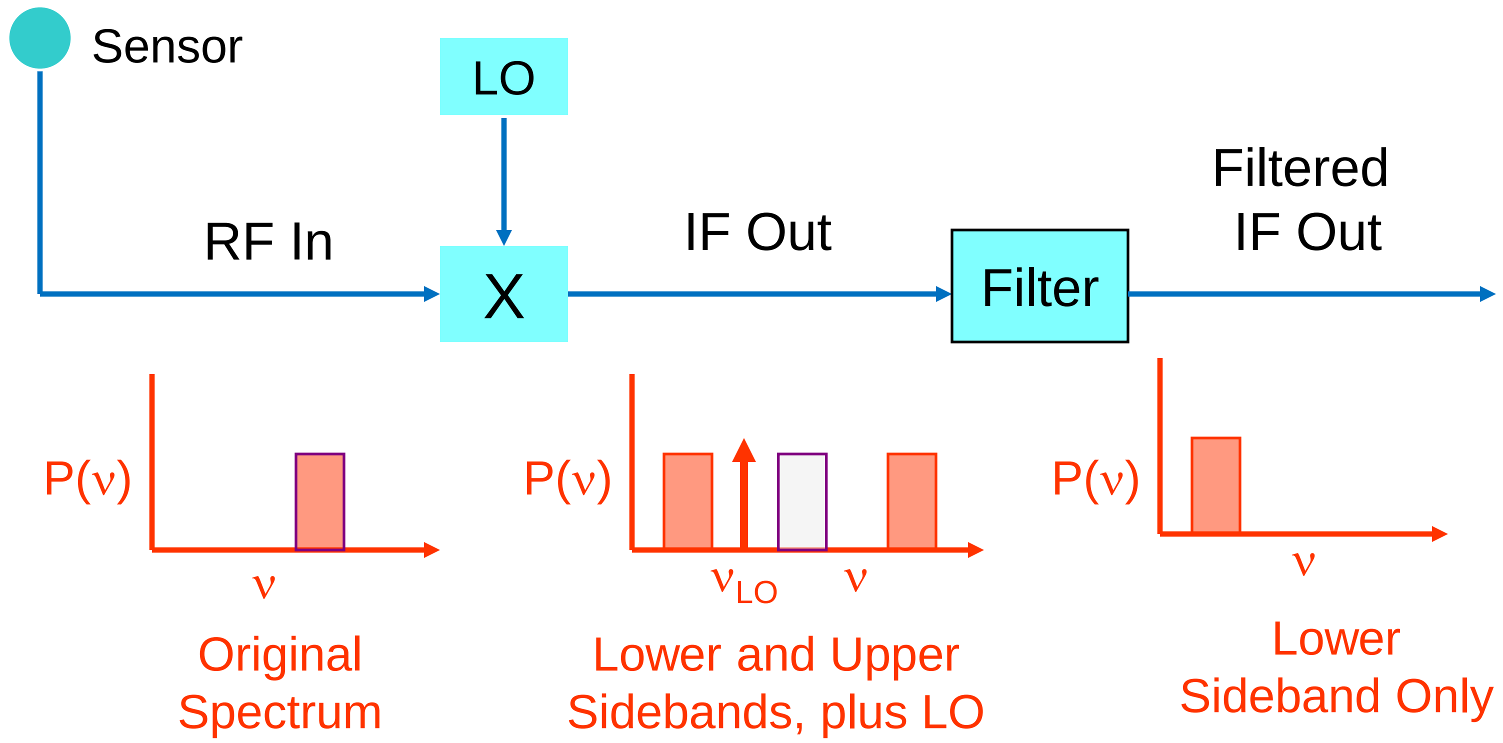
- So, what kind of time-scales are we talking about now?
- How long can you integrate before the differential motion rotates the source through the fringe pattern?
- Case A: A 25-meter parabaloid, and 35-km baseline:
  - $t = D/(B\omega_e)$ = 10 seconds. (independent of wavelength)
- Case B: Whole Hemisphere for a 35-km baseline:
  - $t = \lambda/(B\omega_e)$ sec = 83 msec at 21 cm.
- Averaging for durations longer than these will cause severe attenuation of the visibility amplitudes.
- To prevent 'delay losses', your averaging time must be much less than this.
  - Averaging time 1/10 of this value normally sufficient to prevent time loss.

The Heterodyne Interferometer: LOs, IFs, and Downcoversion

- This would be the end of the story (so far as the fundamentals are concerned) if all the internal electronics of an interferometer would work at the observing frequency (often called the 'radio frequency', or RF).
- Unfortunately, this cannot be done in general, as high frequency components are much more expensive, and generally perform more poorly than low frequency components.
- Thus, most radio interferometers use 'down-conversion' to translate the radio frequency information from the 'RF' to a lower frequency band, called the 'IF' in the jargon of our trade.
- For signals in the radio-frequency part of the spectrum, this can be done with almost no loss of information.
- But there is an important side-effect from this operation in interferometry which we now review.

Downconversion

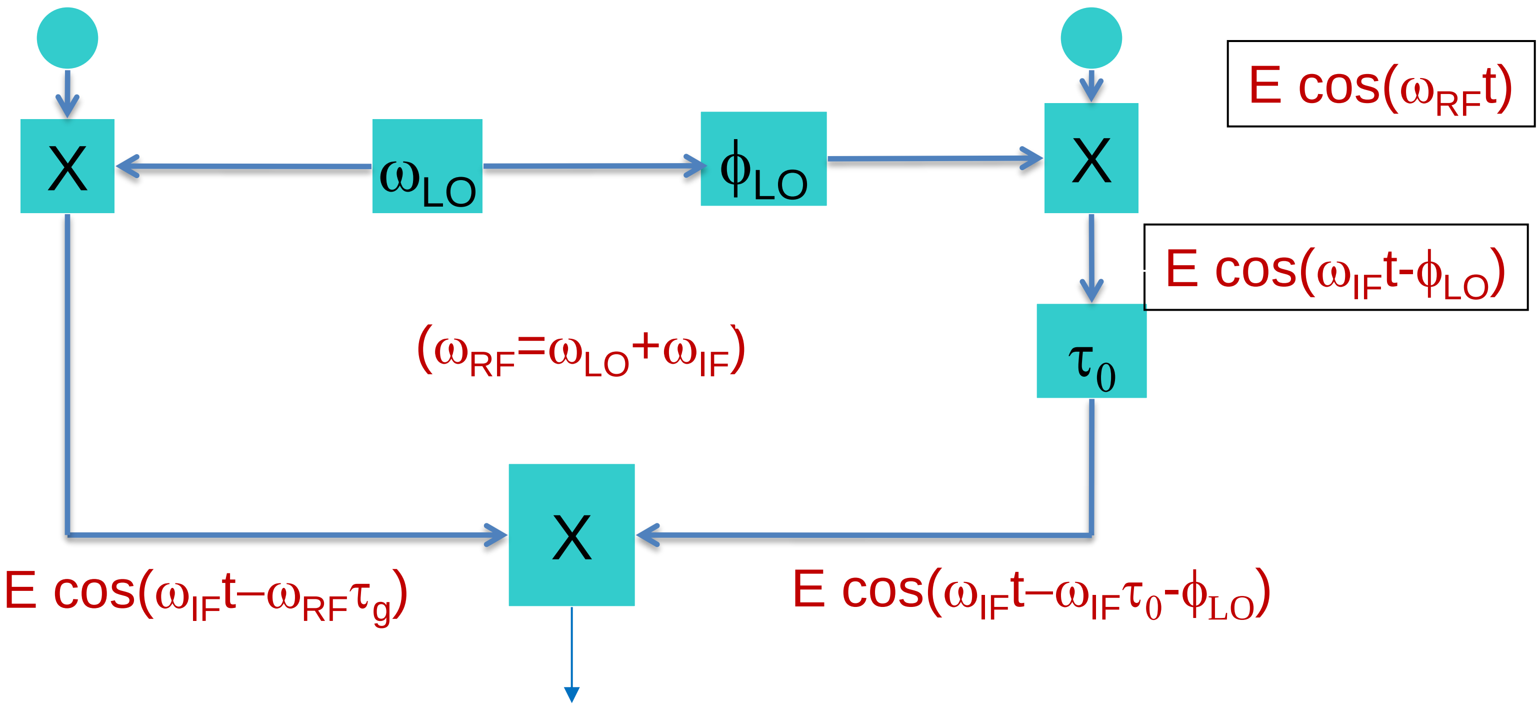
At radio frequencies, the spectral content within a passband can be shifted – with almost no loss in information, to a lower frequency through multiplication by a ‘LO’ signal.



This operation preserves the amplitude and phase relations

Signal Relations, with LO Downconversion

- The RF signals are multiplied by a pure sinusoid, at frequency ν_{LO}
- We can add arbitrary phase ϕ_{LO} on one side.



$$V = E^2 e^{-i(\omega_{RF}\tau_g - \omega_{IF}\tau_0 - \phi_{LO})}$$

Recovering the Correct Visibility Phase

- The correct phase (RF interferometer) is: $\omega_{RF}\left(\tau_g - \tau_0\right)$
- The observed phase (with frequency downconversion) is:

$$\omega_{RF}\tau_g - \omega_{IF}\tau_0 - \phi_{LO}$$

- These will be the same when the LO phase is set to:

$$\phi_{LO} = \omega_{LO}\tau_0$$

- This is necessary because the delay, τ_0, has been added in the IF portion of the signal path, rather than at the frequency at which the delay actually occurs.
- The phase adjustment of the LO compensates for the delay having been inserted at the IF , rather than at the RF.

NRAO

The Three 'Centers' in Interferometry

- You are forgiven if you're confused by all these 'centers'.
- So let's review:

1. **Beam Tracking (Pointing) Center**: Where the antennas are pointing to. (Or, for phased arrays, the phased array center position).
2. **Delay Tracking Center:** The location for which the delays are being set for maximum wide-band coherence.
3. **Phase Tracking Center:** The location for which the LO phase is slipping in order to track the coherence pattern.

- Note: Generally, we make all three the same. #2 and #3 are the same for an 'RF' interferometer. They are separable in a LO downconversion system.

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 of the source.
- At this time, we define the geometric coordinate frame for the interferometer.
- We begin with a special case: An interferometer whose antennas all lie on a single plane.

The 2-Dimensional Interferometer

To give better understanding, we now specify the geometry.

Case A: A 2-dimensional measurement plane.

- Let us imagine the measurements of $V_\nu(\mathbf{b})$ to be 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.

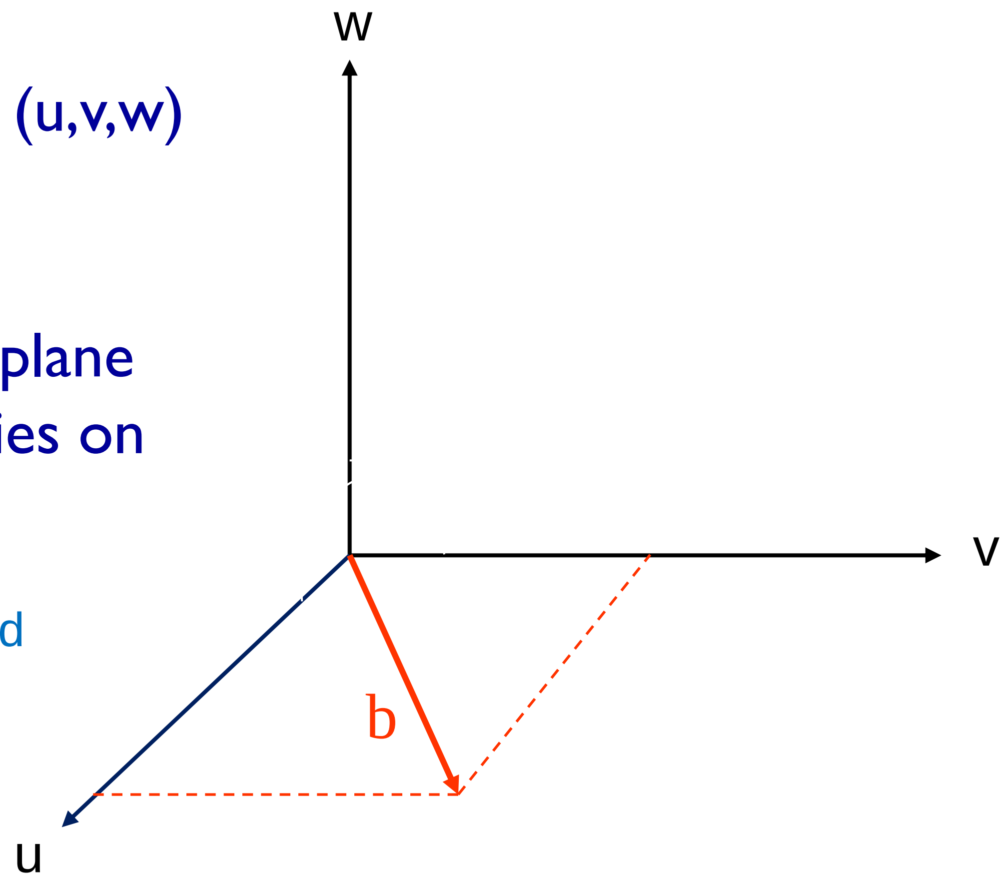
- The components of the unit direction vector, **s**, are:

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

NRAO

The (u,v,w) Coordinate System.

- Pick a coordinate system (u,v,w) to describe the antenna positions and baselines.
- Orient this frame so the plane containing the antennas lies on the plane w = 0.



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

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

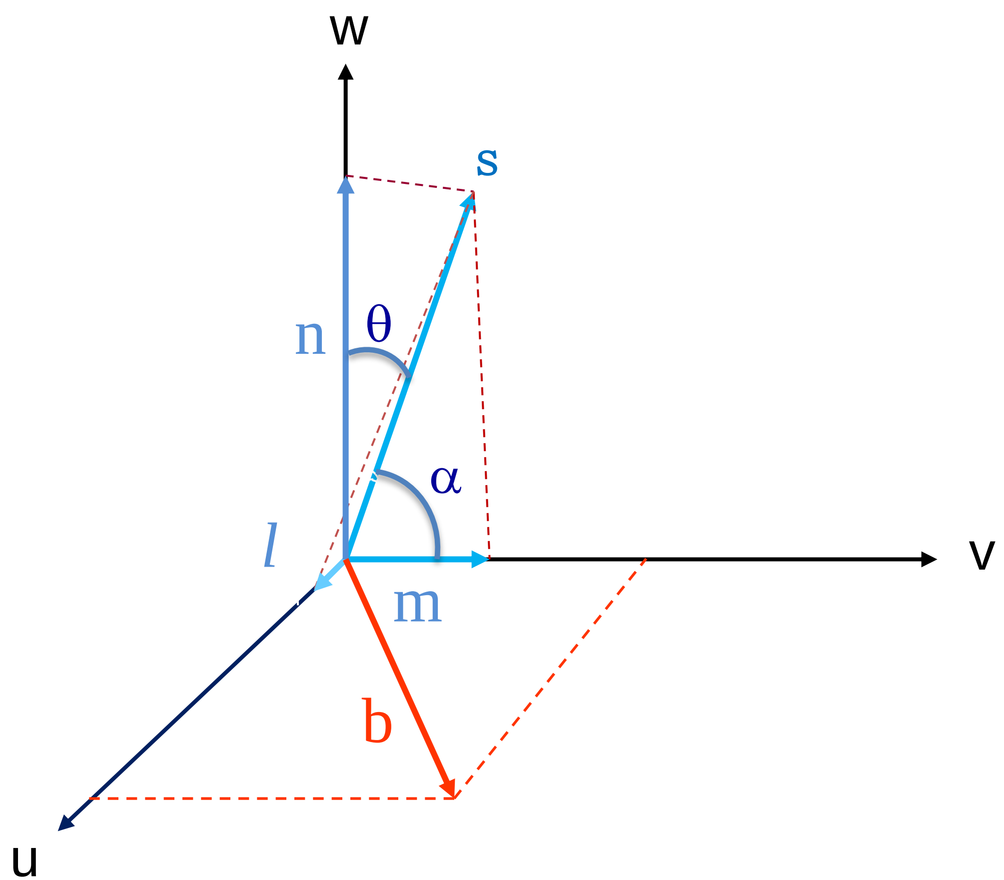
Direction Cosines – describing the source

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, $\nu \mathbf{b} . \mathbf{s}/c = ul + vm$, (since w = 0), from which we find,

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

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

$$I_\nu(l,m) \Leftrightarrow V(u,\mathrm{v})$$

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'.

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.

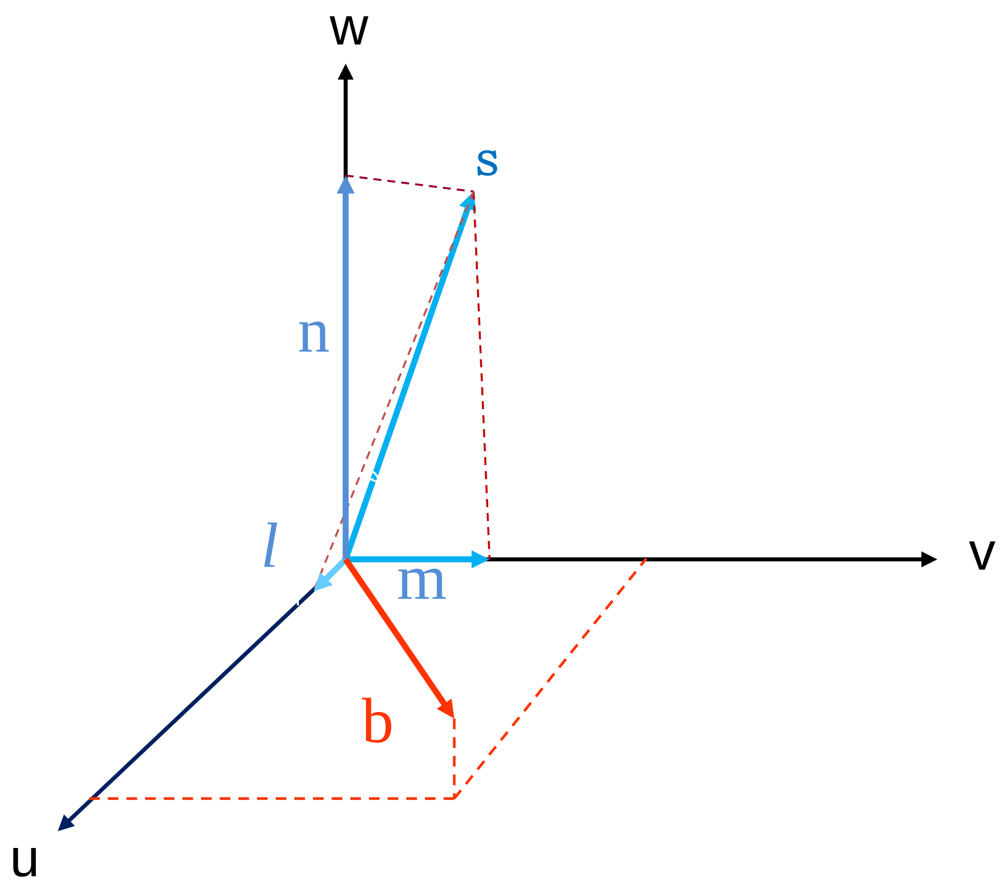
    - WSRT (Westerbork Synthesis Radio Telescope)
    - ATCA (Australia Telescope Compact Array) (before the third arm)
    - Cambridge 5km (Ryle) telescope (almost).
  - b) Any coplanar 2-dimensional array, at a single instance of time.

    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.

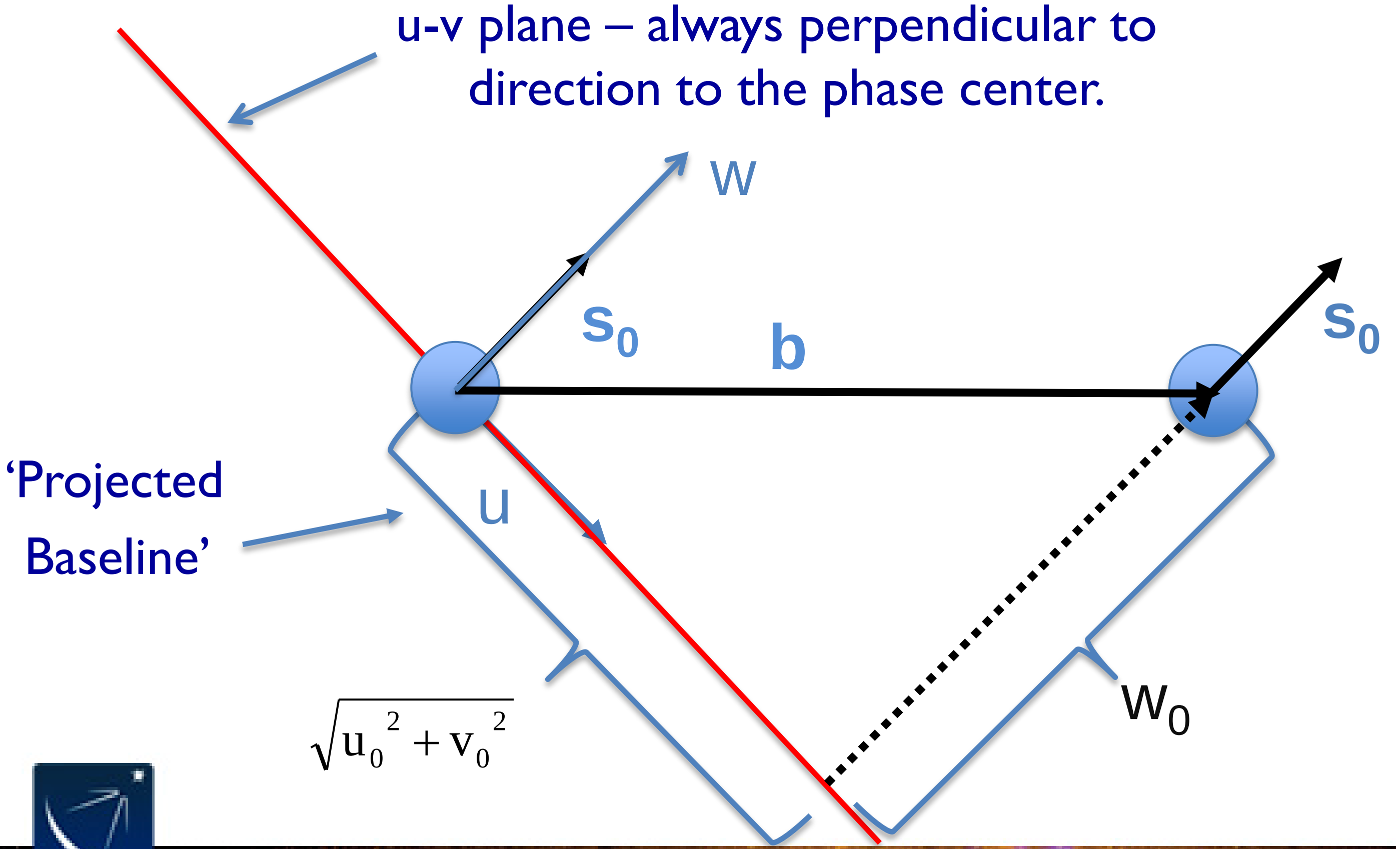
Generalized Baseline Geometry

- Coplanar arrays (like the VLA) cannot use the 2-d geometry in synthesis mode, since the plane of the array is rotating w.r.t. the source.
- The sampled region is now three-dimensional.
- In this case, we must adopt a more general geometry, where all three baseline components are to be considered.



General Coordinate System

w points to, and follows the source phase center, **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:

- What if the interferometer does not measure the coherence function on a plane, but rather does it through a volume? In this case, we adopt a different coordinate system. First we write out the full expression:

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

(Note that this is not a 3-D Fourier Transform).

- We orient the w-axis of the coordinate system to point to the region of interest. The u-axis point east, and the v-axis to the north celestial pole.
- 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:

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

3-d to 2-d

- The expression is still not a proper Fourier transform.
- We can get a 2-d FT if the third term in the phase factor is sufficient small.

- The third term in the phase can be neglected if it is much less than unity:

$$w\left(1-\sqrt{1-l^2-m^2}\right)=w(1-\cos\theta)\sim w\theta^2/2<<1$$

- This condition holds when:
(angles in radians!)

$$\theta_{\max}<\sqrt{\frac{1}{2\mathrm{w}}}\sim\sqrt{\frac{\lambda}{\mathrm{B}}}\sim\sqrt{\theta_{syn}}$$

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

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

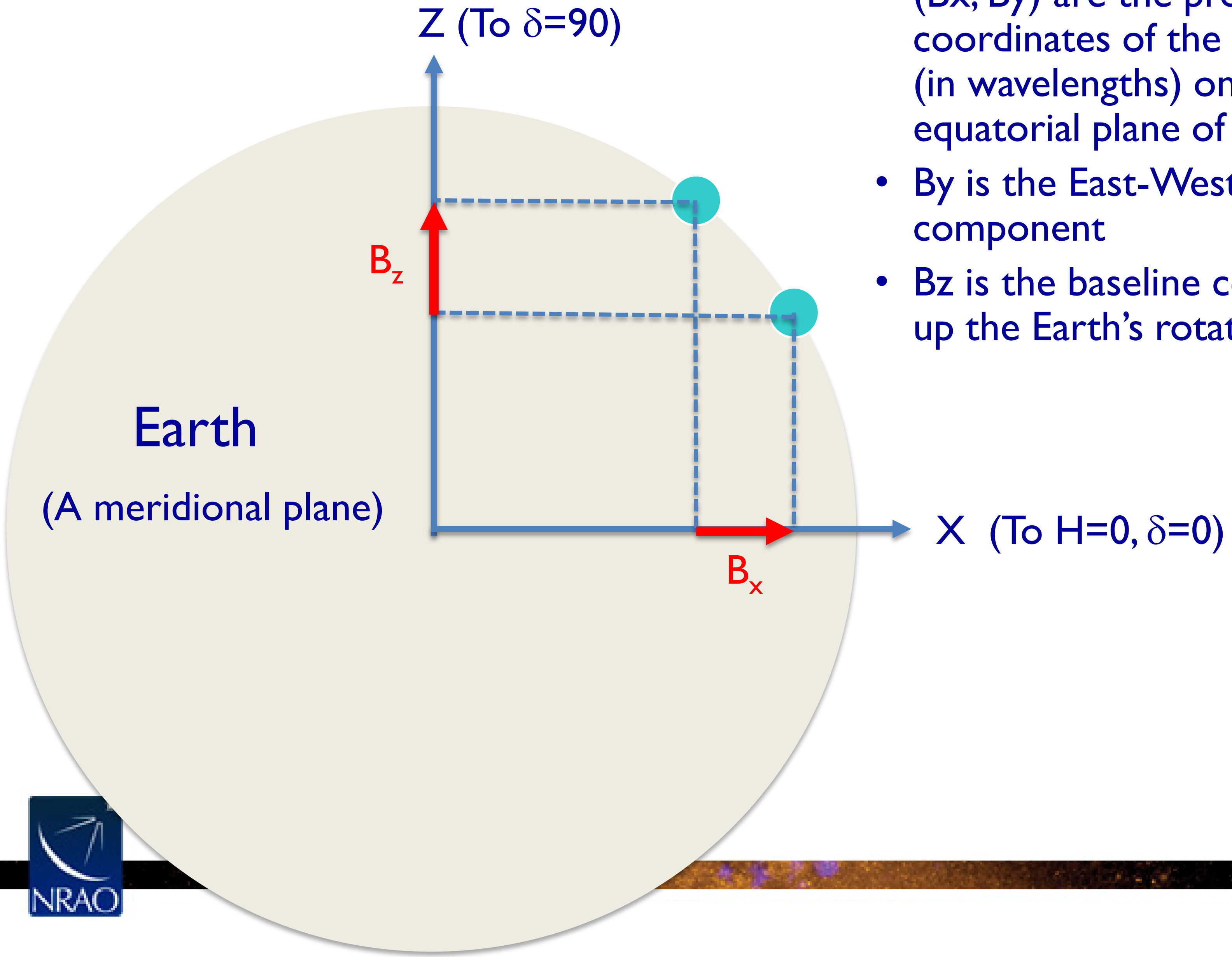
The Problem with Non-coplanar Baselines

- 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.
- The 'Clark Condition' for trouble is: $\frac{\lambda B}{D^2} > 1$
- Hence, the problem is most acute for small-diameter antennas and long wavelengths.
- The problems are not in the principles, but in the cost of the solutions. Full 3-D imaging works, but isn't cheap.
- Various solutions are available (mosaicing, w-projection, full-3D transforms), but discussion of these is beyond the scope of this talk.

Coverage of the U-V Plane

- I return now to the definition of the (u,v) plane, and discuss the 'coverage'.
- Adopt the standard geometry:
  - W points to, and tracks, the phase center
  - U points to the east, V to the north.
- To derive the values of U, V, and W, we adopt an earth-based coordinate system for describe the antenna locations.
  - 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).
- Then denote by (Bx, By, Bz) the coordinates, measured in wavelengths, of a baseline in this earth-based frame.

Array Coordinate Frame



- (Bx, By) are the projected coordinates of the baseline (in wavelengths) on the equatorial plane of the earth.
- By is the East-West component
- Bz is the baseline component up the Earth's rotational axis.

The (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 delay, τ_g is given by

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

- Its derivative, called the fringe frequency ν_F is

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

E-W Baseline – the simplest case

- For an array whose elements are oriented E-W, the geometry is especially simple:
  - $B_x = B_z = 0$, so that

$$u = B_y \cos H_0$$

$$v = B_y \sin \delta_0 \sin H_0$$

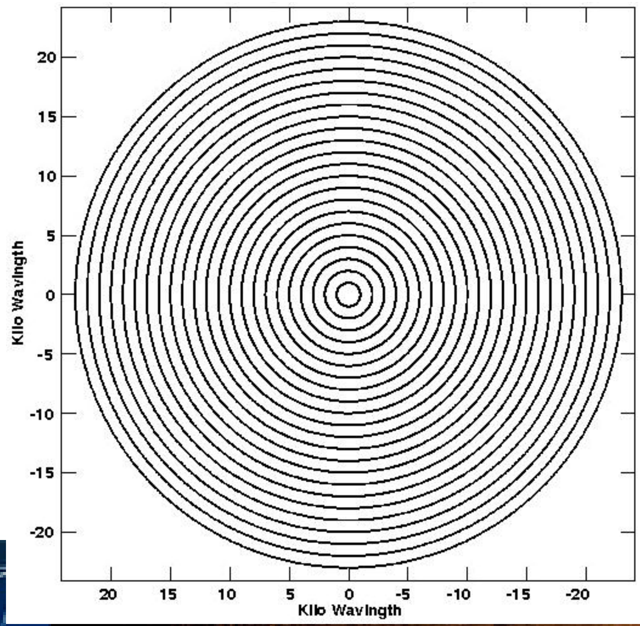
$$w = B_y \cos \delta_0 \sin H_0$$

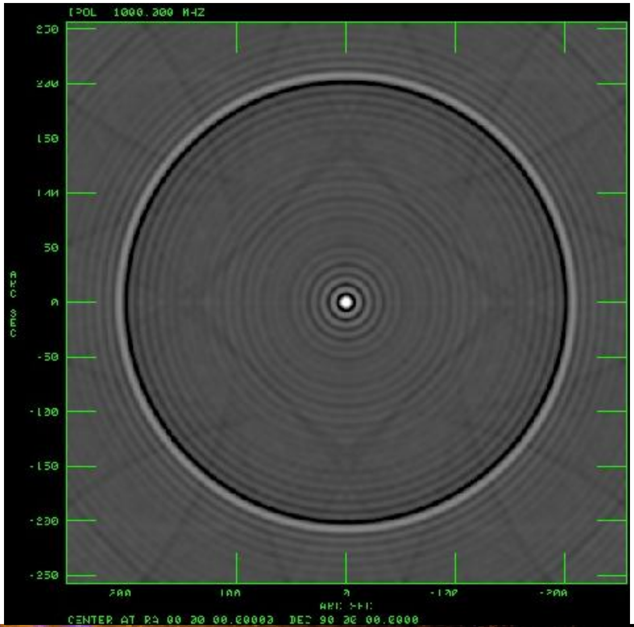
- To illustrate, I show an example of a 'minimum redundancy' E-W design.

E-W Array Coverage and Beams

- 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.

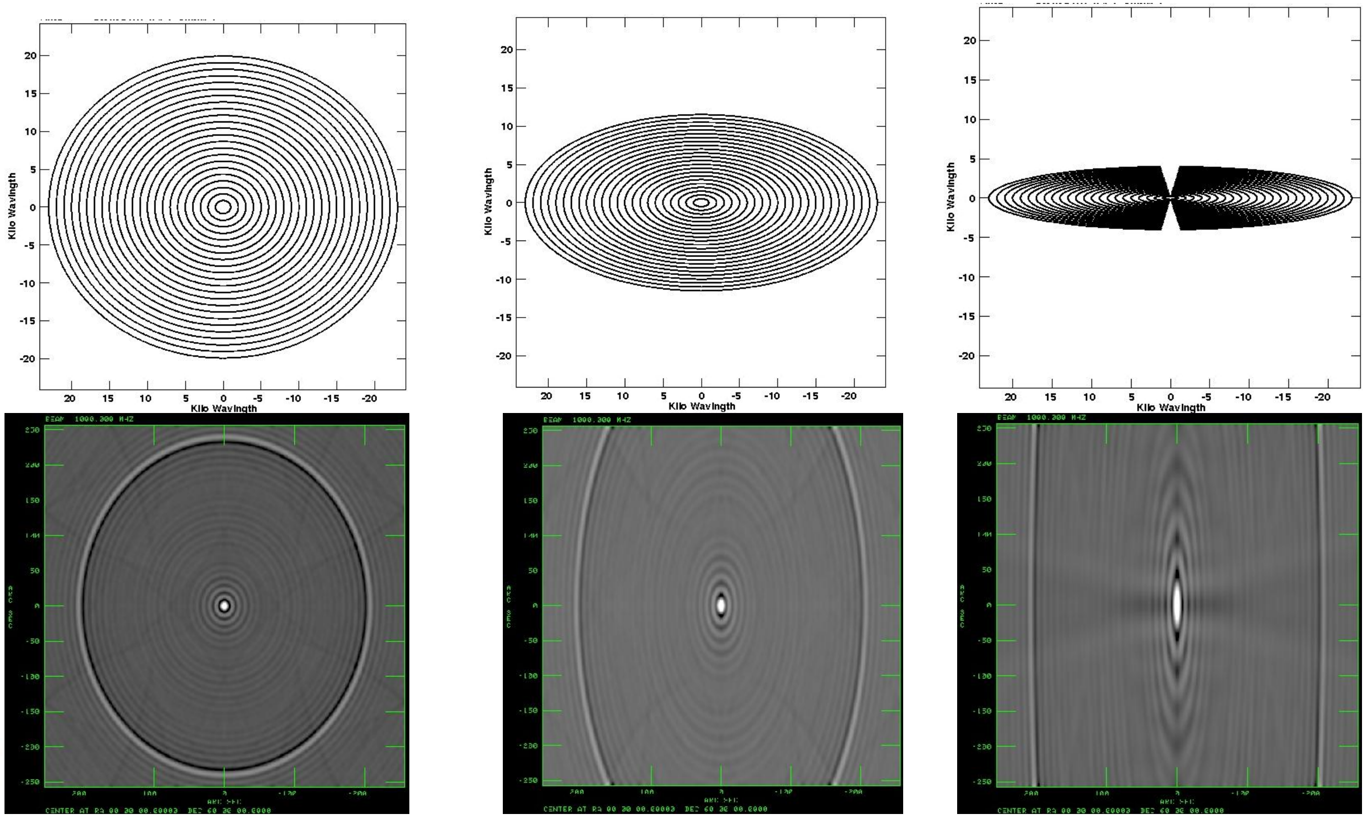




E-W Arrays and Low-Dec sources.

- 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

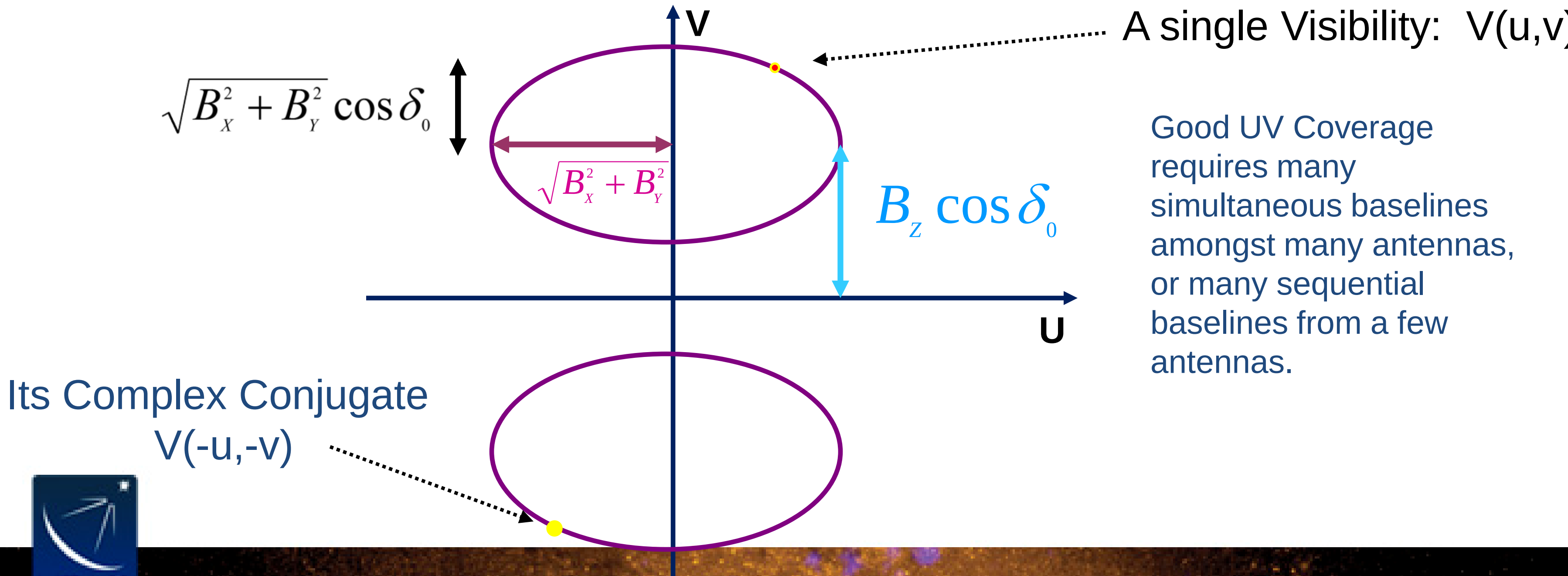


Baseline Locus – the General Case

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

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

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



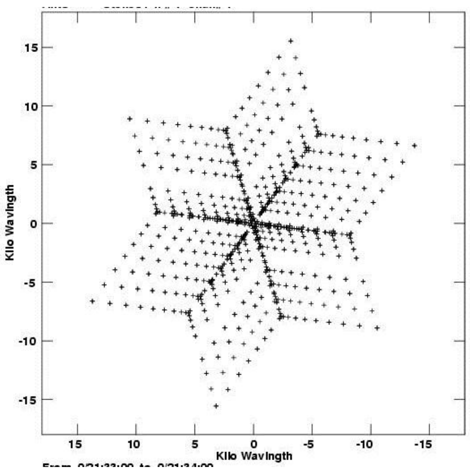
Good UV Coverage requires many simultaneous baselines amongst many antennas, or many sequential baselines from a few antennas.

Getting Good Coverage near $\delta = 0$

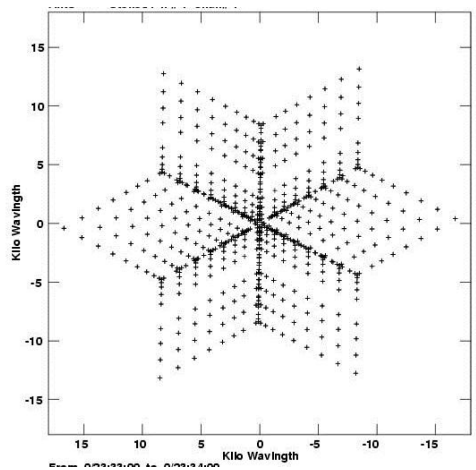
- The only means of getting good N-S angular resolution at all declinations is to build an array with N-S spacings.
- Many more antennas are needed to provide good coverage for such geometries.
- The VLA was designed to do this, using 9 antennas on each of three equiangular arms.
- Built in the 1970s, commissioned in 1980, the VLA vastly improved radio synthesis imaging at all declinations.
- Each of the 351 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.

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

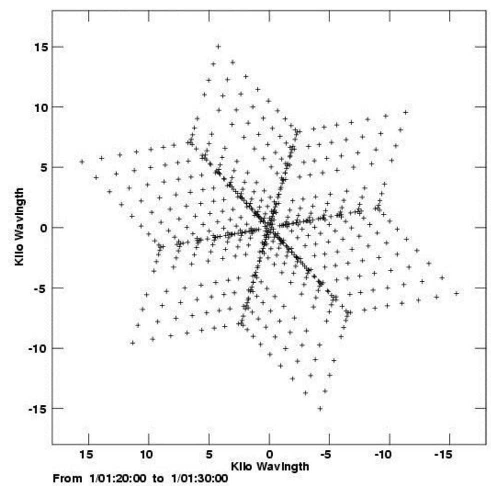
- Snapshot (u,v) coverage for HA = -2, 0, +2 (with 26 antennas).



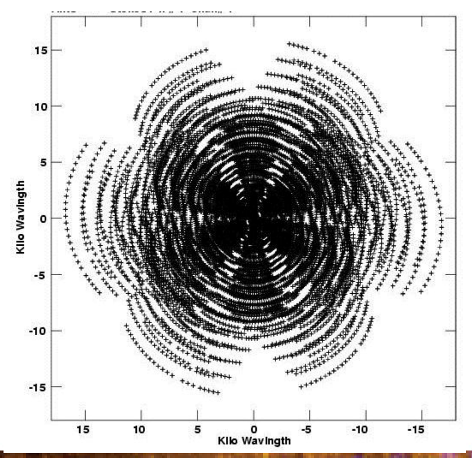
HA = -2h



HA = 0h

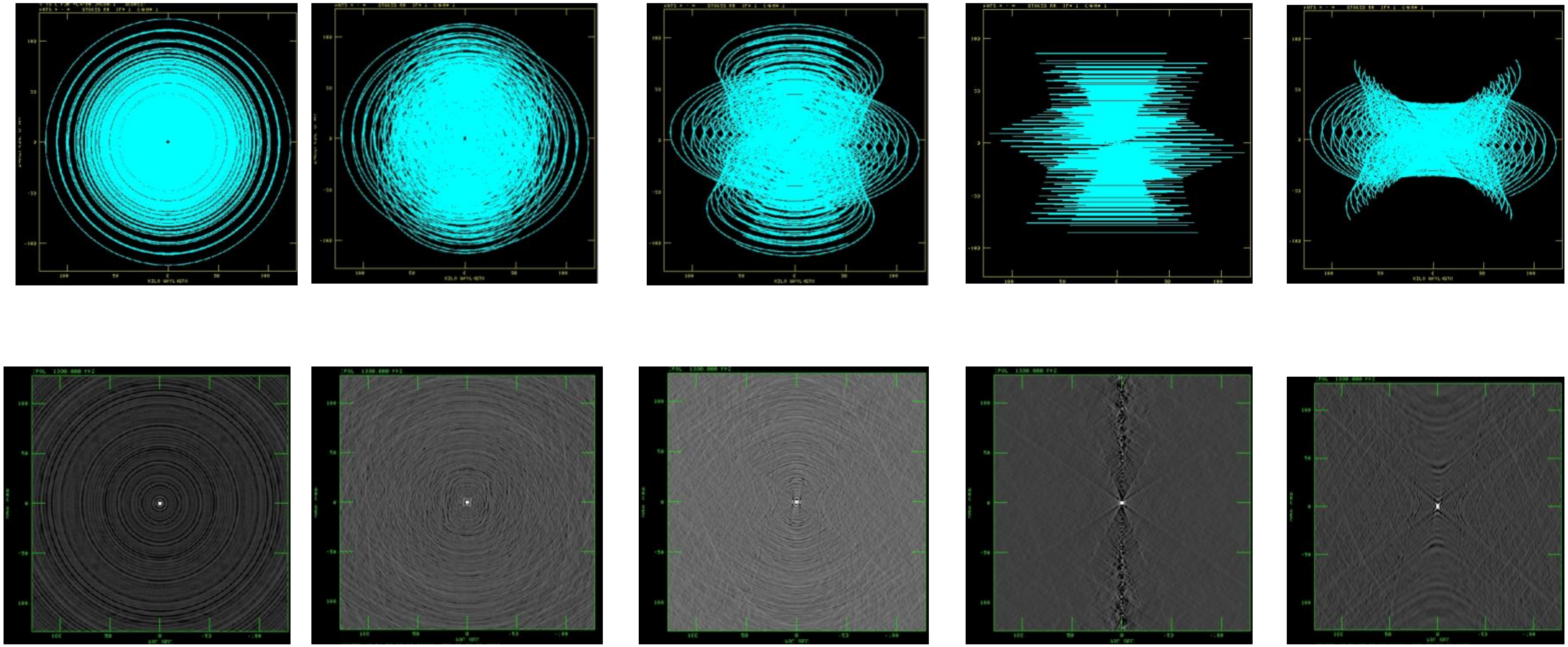


HA = 2h



Coverage over all four hours.

VLA Coverage and Beams



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

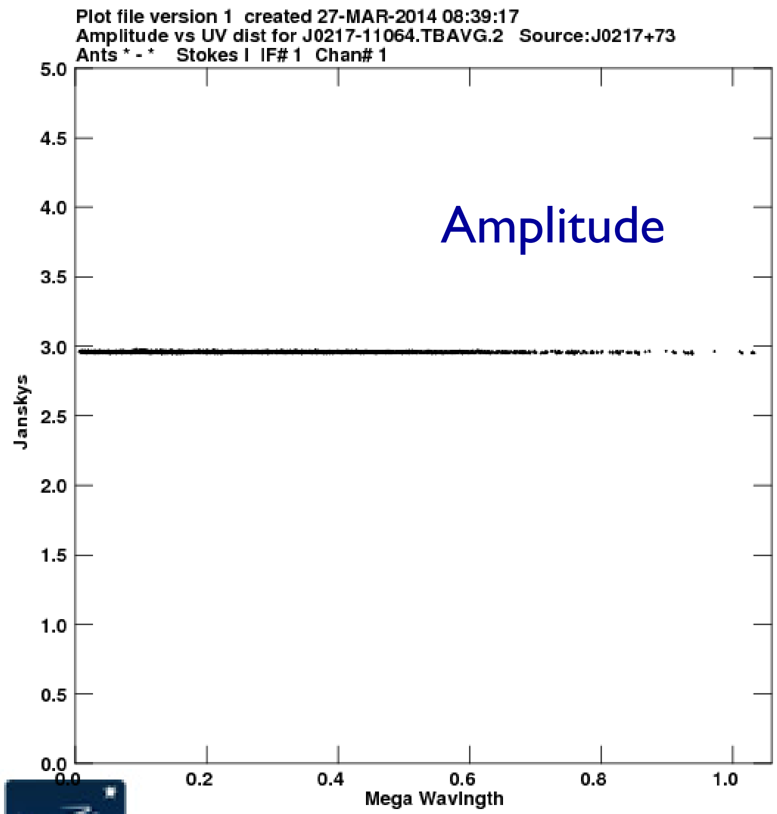
- Good coverage at all declinations, but troubles near $\delta=0$ remain.

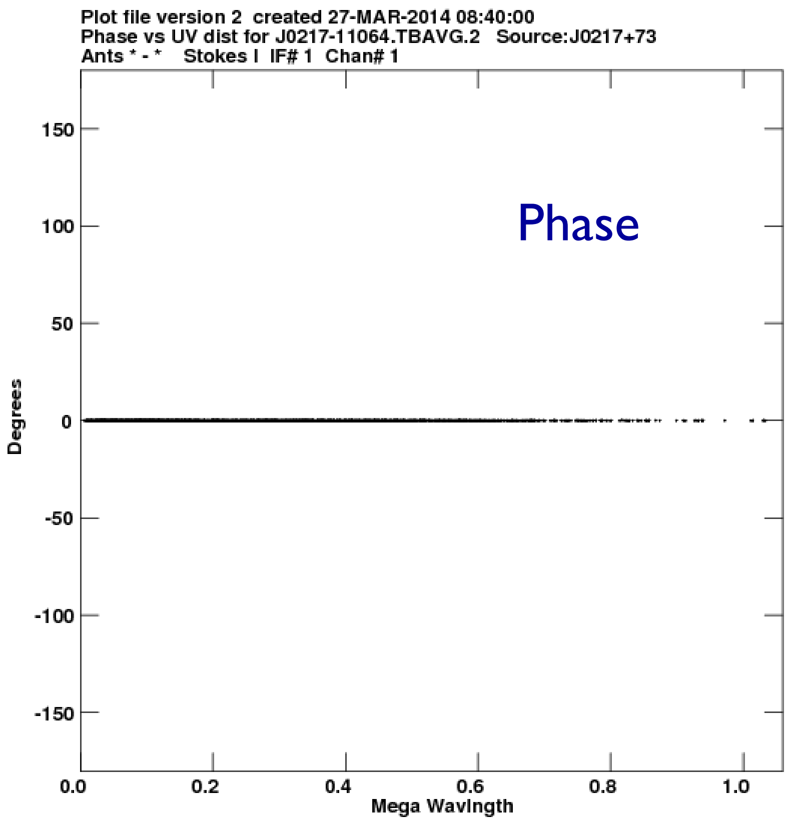
Examples of Real Visibilities from Simple Sources

- I finish with some actual visibility plots from observations of VLA calibrator sources.
- These plot the visibility amplitude or phase on the 'y' axis against the projected baseline, $\sqrt{u^2+v^2}$ on the 'x' axis.
- It is very useful to be able to interpret these plots to aid in judging quality of data and calibration.

Example 1: – A Point Source

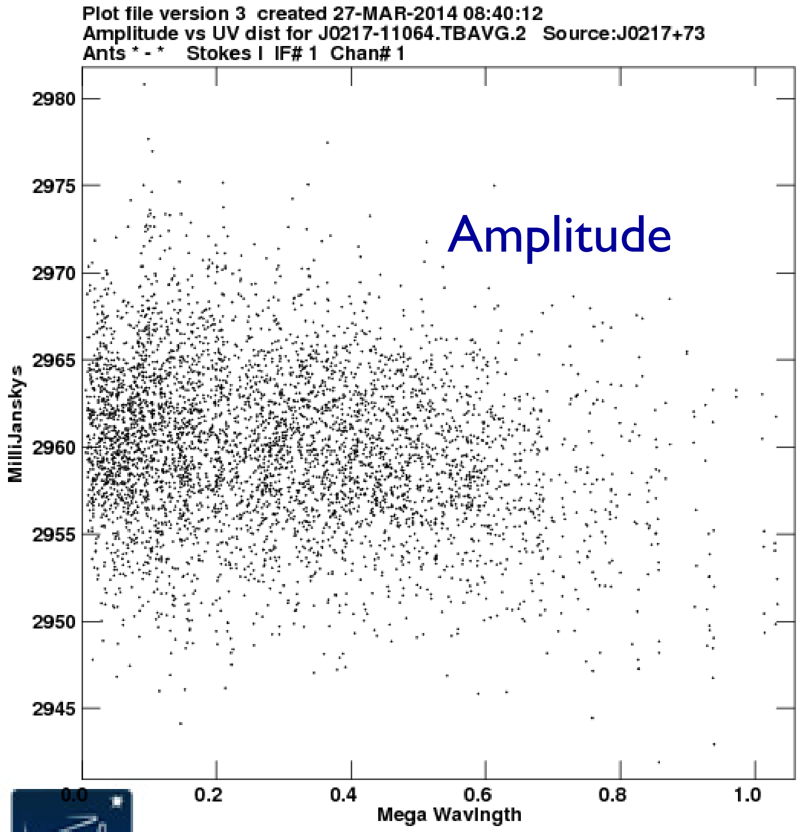
- Shown are the amplitude and phase of a strong calibrator, J0217+738.   Not very interesting on these scales.

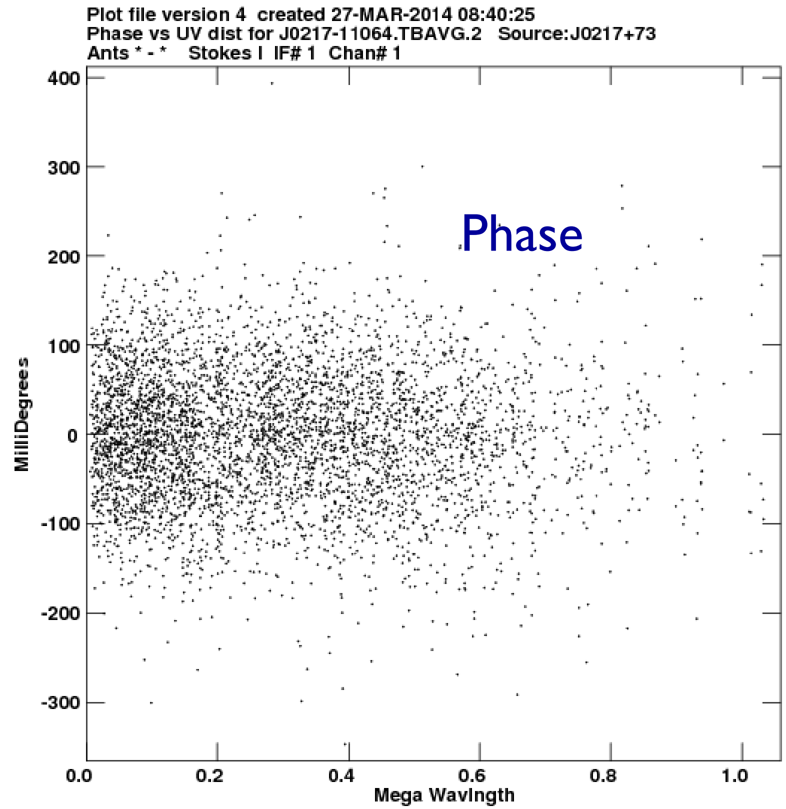




Zoom in …

- Suppose we observe an unresolved object.
- What is its visibility function?



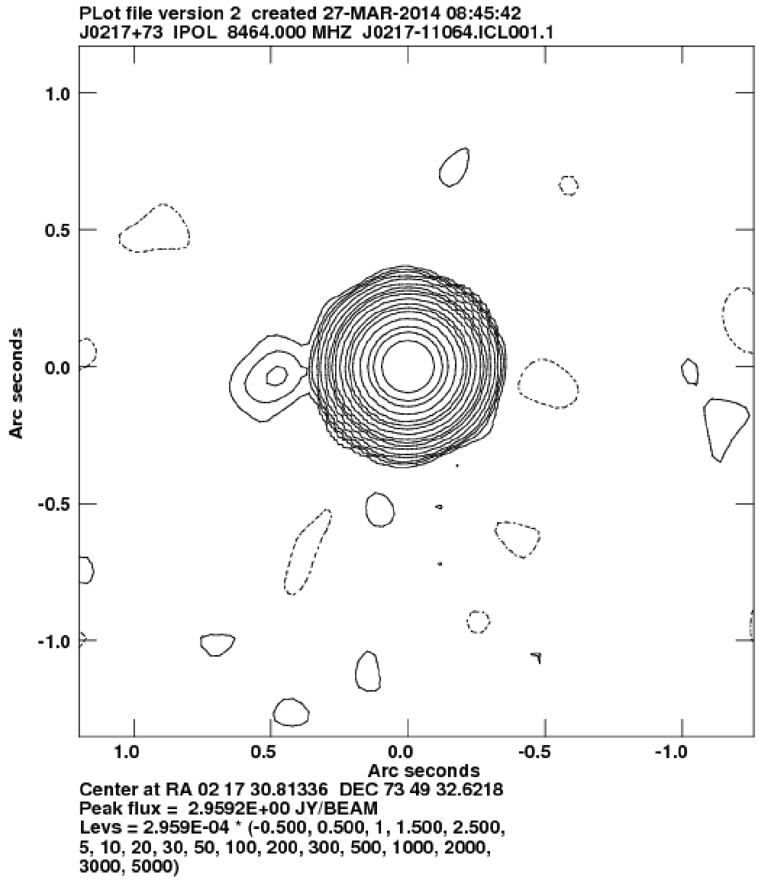


NRAO

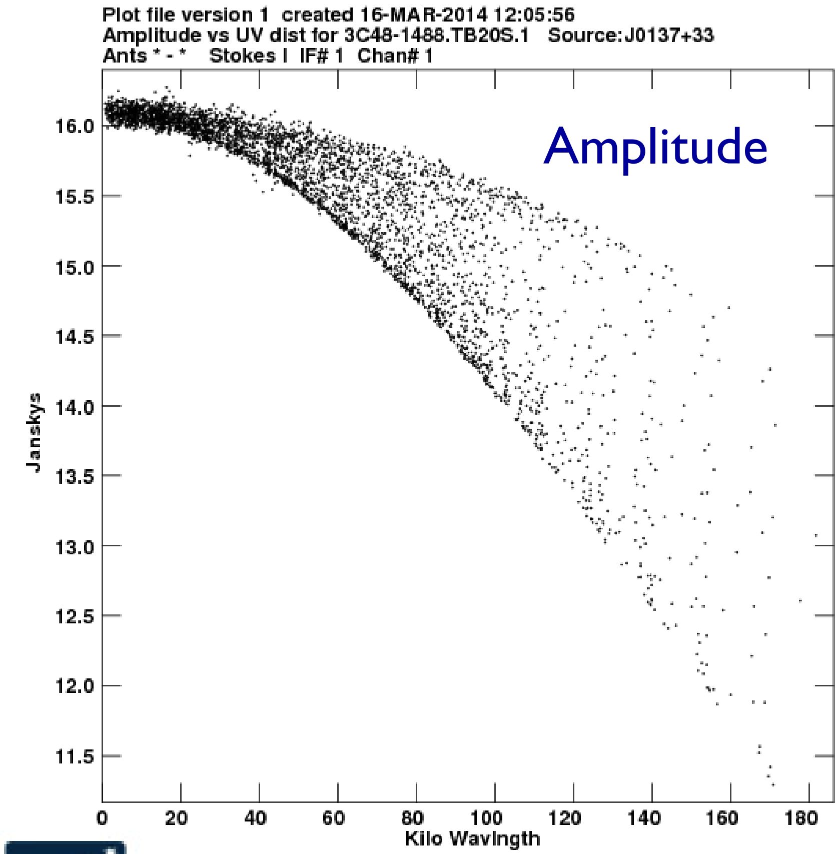
And the Map …

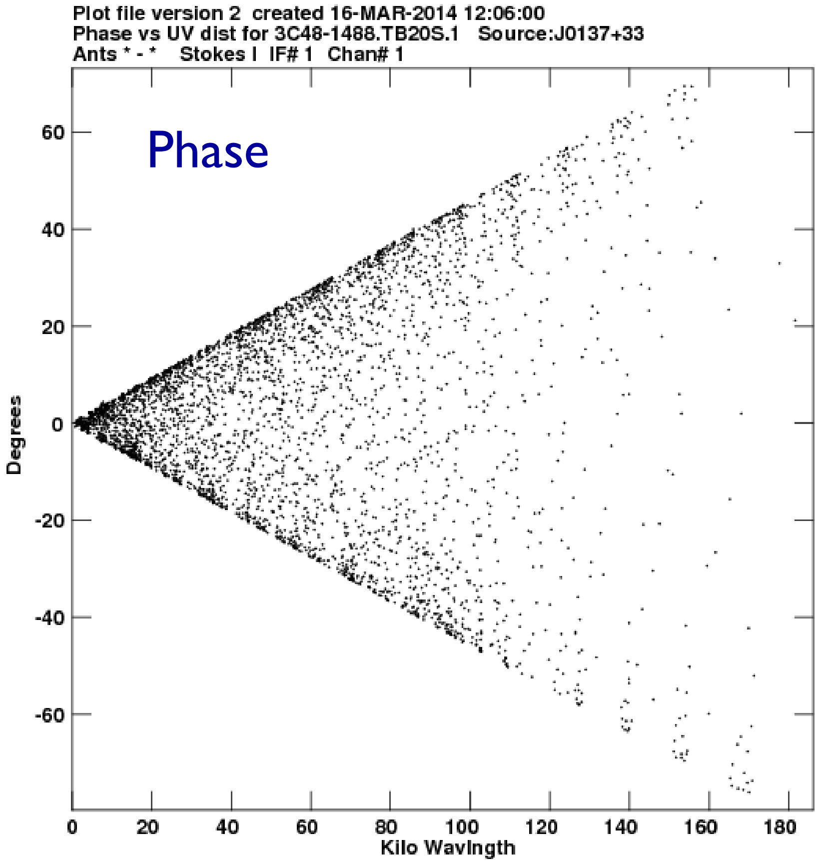
- The source is unresolved … but with a tiny background object.

- Dynamic range: 50,000:1.
- The flux in the secondary object is too small to be visible in the visibility function.



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



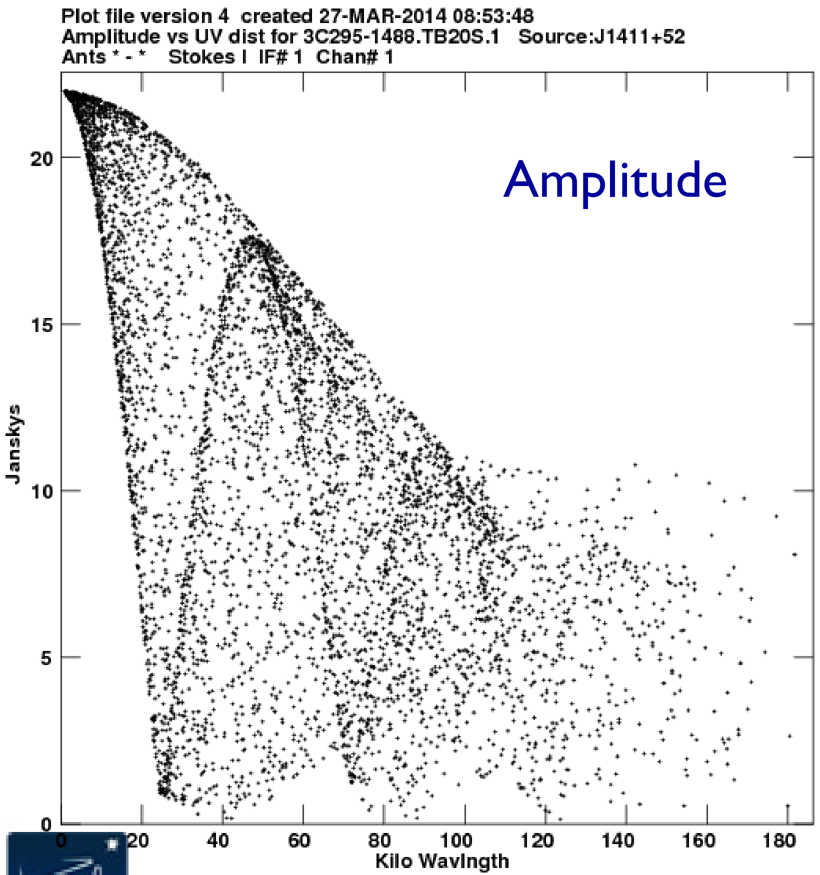


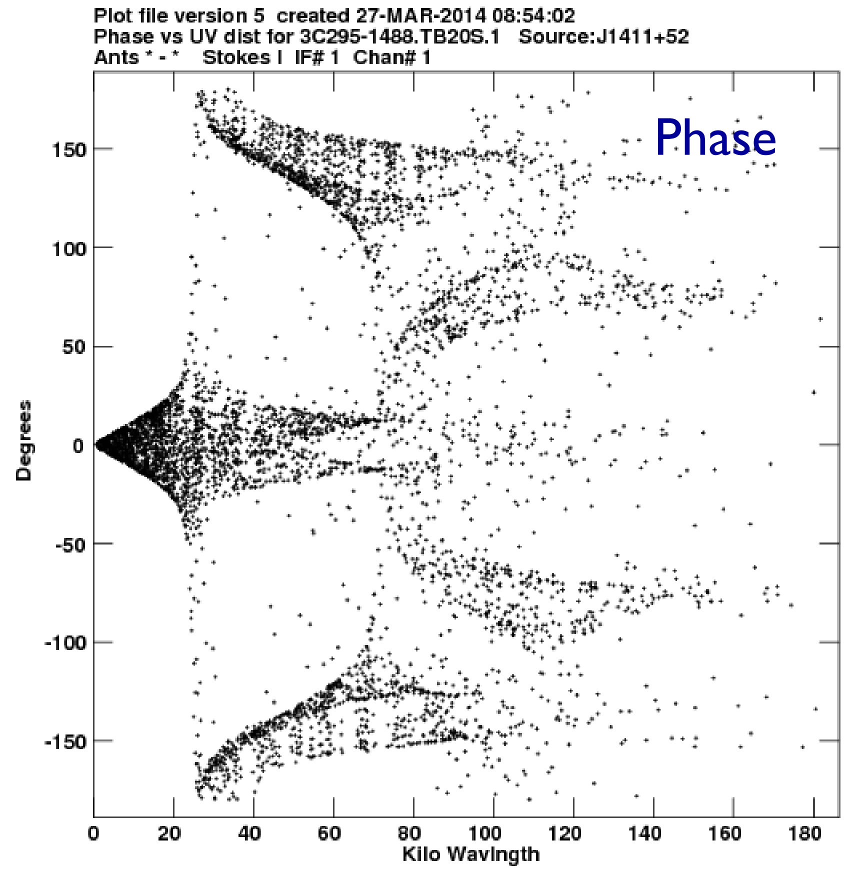
Interpreting this Visibility Function:

- The amplitude function tells us the source is roughly elliptical:
  - The 50% visibility is roughly at 200 kλ x 400 kλ, corresponding to 1" x 0.5"
- The phase slope of one turn in 850 kλ tells us that the source is offset from the phase center by ~ 0.25 arcsecond.
- But … we can't tell the angle of the offset, or the orientation of the structure from these 1-d plots.
- The few amplitude points seen above and below the smooth distribution result from *calibration errors*.

3C295 at 30 cm wavelength

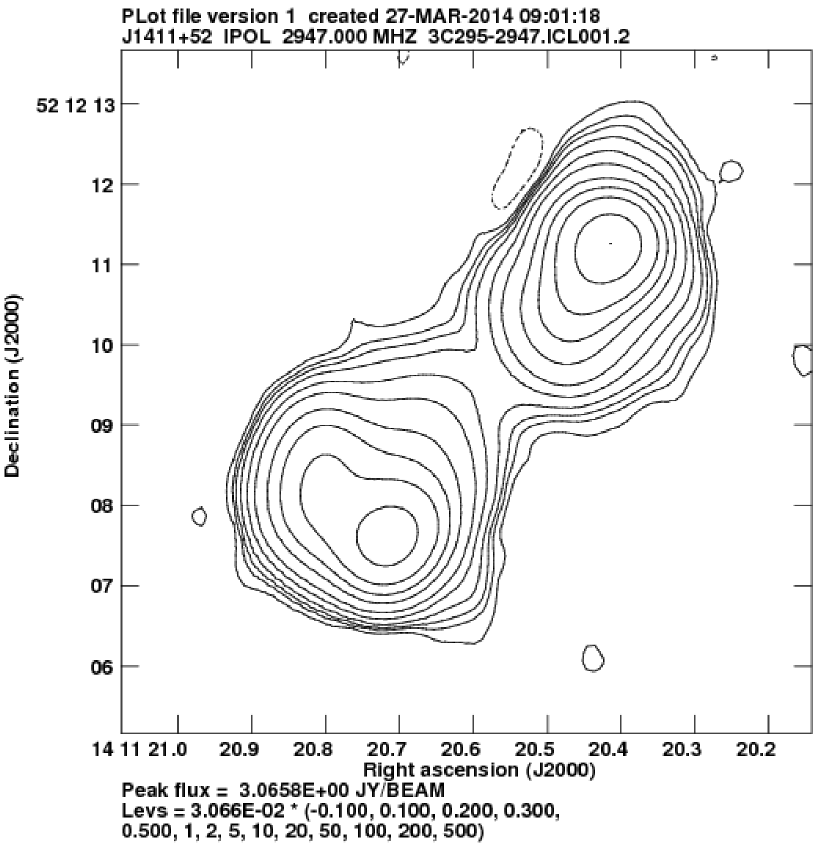
- The sinusoid of period 45 kλ tells us this source is comprised of two resolved objects, separated by 1 rad/45000 ~ 5 arcsec.





3C295 Image

- A 5-arcsecond double.
- The phase ramp in the visibilities shows the centroid of the emission is slightly off the phase center.
- Offset ~ 0.7 arcseconds.



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.