Imaging and non-Imaging analysis

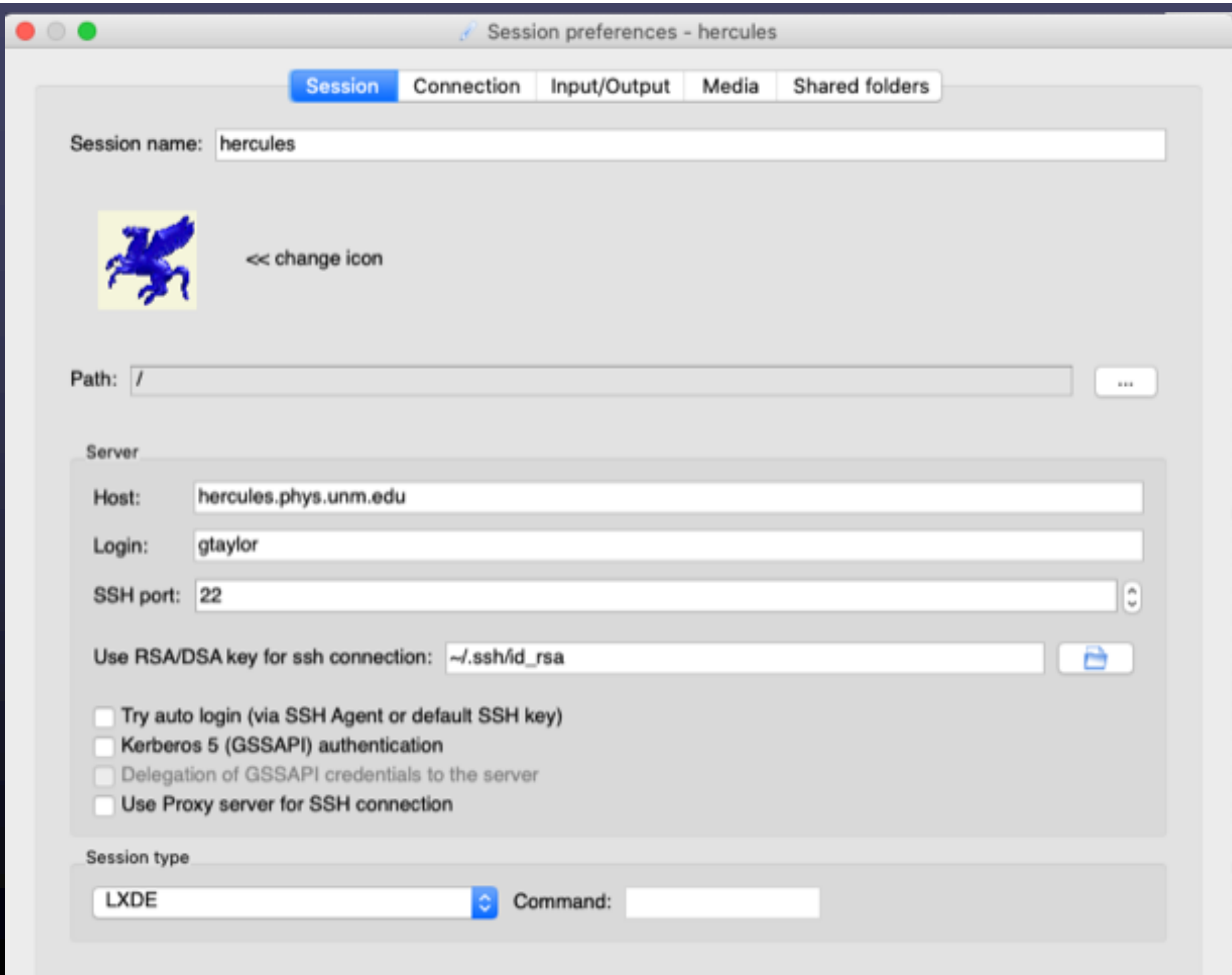
Greg Taylor

University of New Mexico

Announcements

- Test average was 84.25 - no curve
- HW5 is due on Wednesday
- Start playing around with aips and HW6, this is going to take a while
- How many people would prefer to switch to a face-to-face format for the class? See Poll
- LST for LWA1:
- https://lwalab.phys.unm.edu/OpScreen/lwa1/os2.php

Announcements

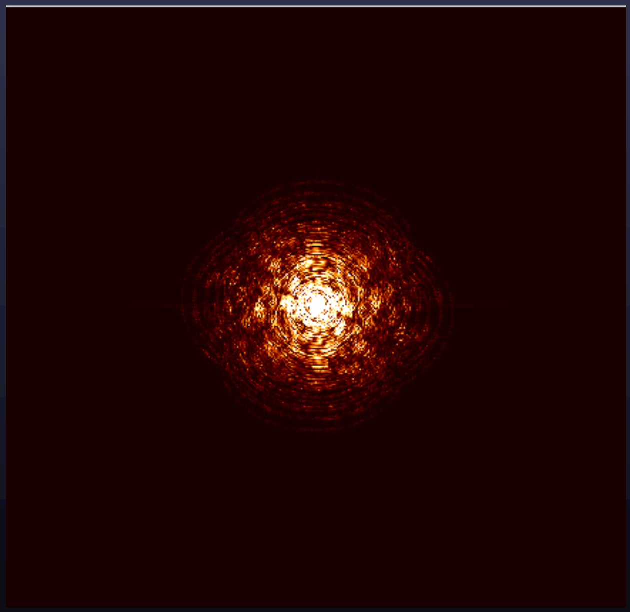


Plan for the lecture-I

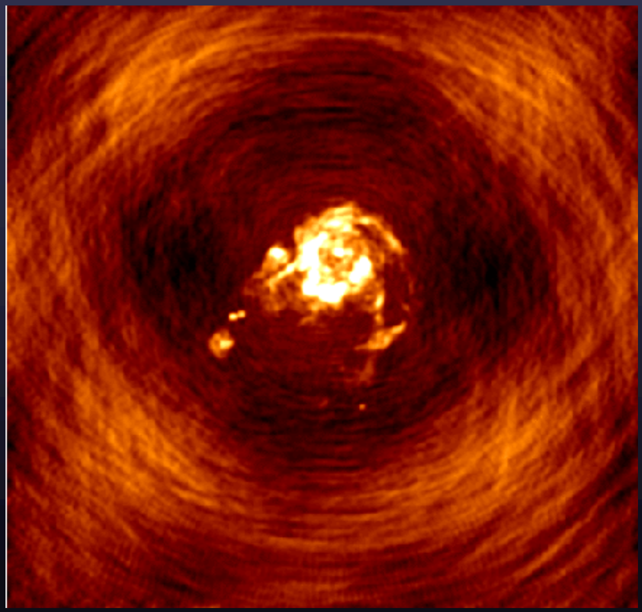
How do we go from the measurement of the coherence function (the Visibilities) to the images of the sky?

- First half of the lecture: Imaging

Measured Visibilities → Dirty Image



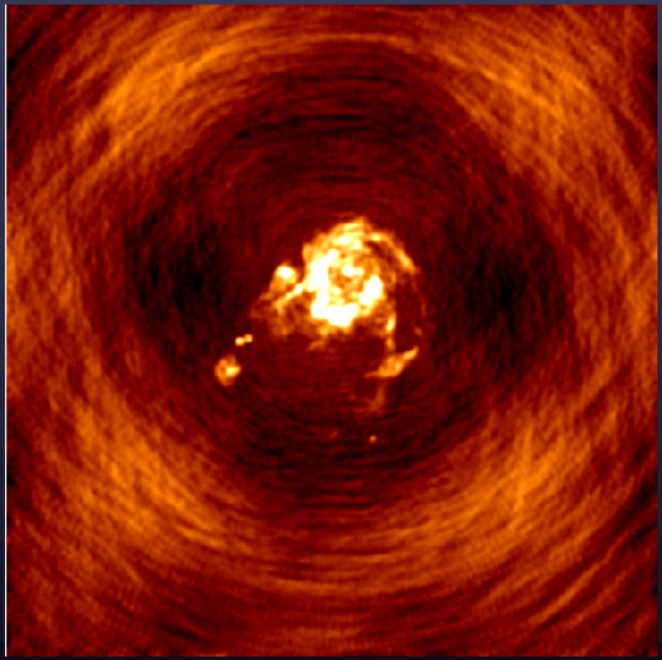
→



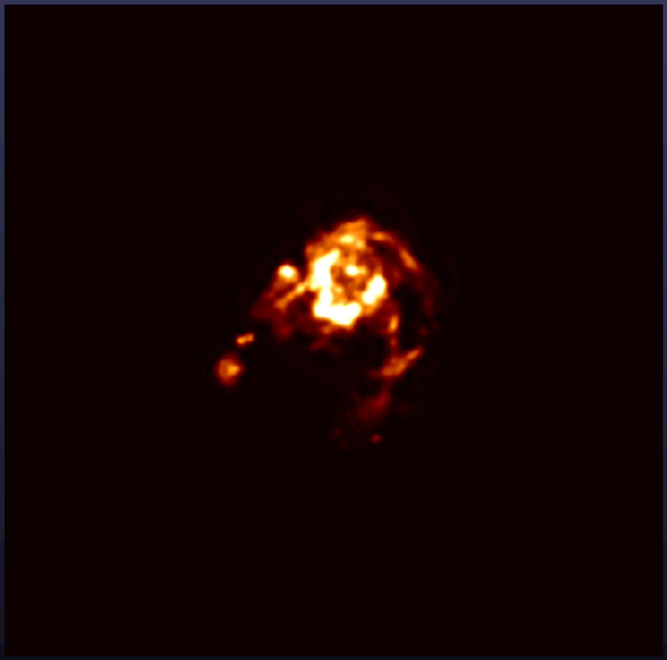
Plan for the lecture-II

- **Second part of the lecture: Deconvolution**

  Dirty image → Model of the sky (+ residuals)



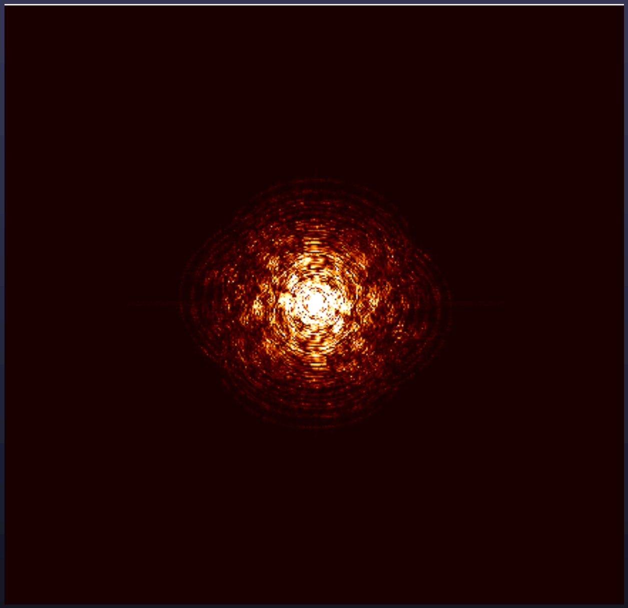
→



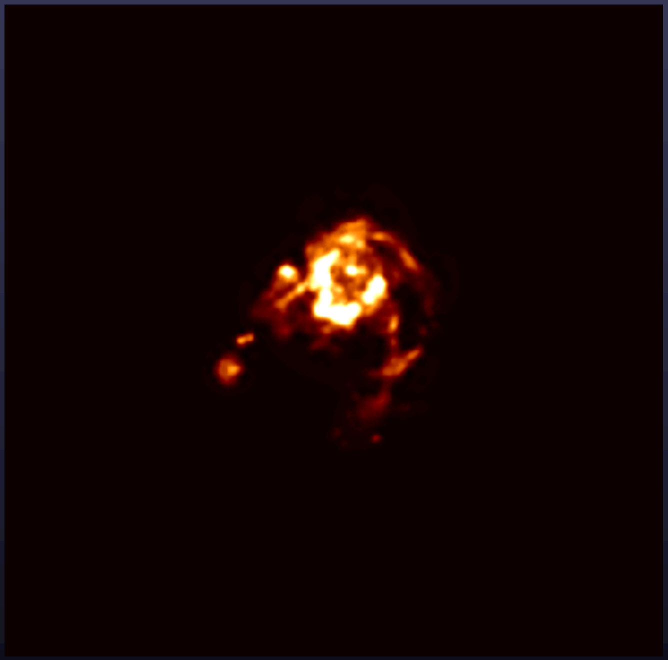
Plan for the lecture-III

- **Third part of the lecture: Modelfitting**

Measured Visibilities → Model of the sky



→



Imaging

- **Interferometers are indirect imaging devices**

$$V^{\circ}(u,v,w) = \int\int I(l,m)e^{2\pi\iota\left[ul+vm+w\left(\sqrt{1-l^2-m^2}-1\right)\right]}\frac{dl\,dm}{\sqrt{1-l^2-m^2}}$$

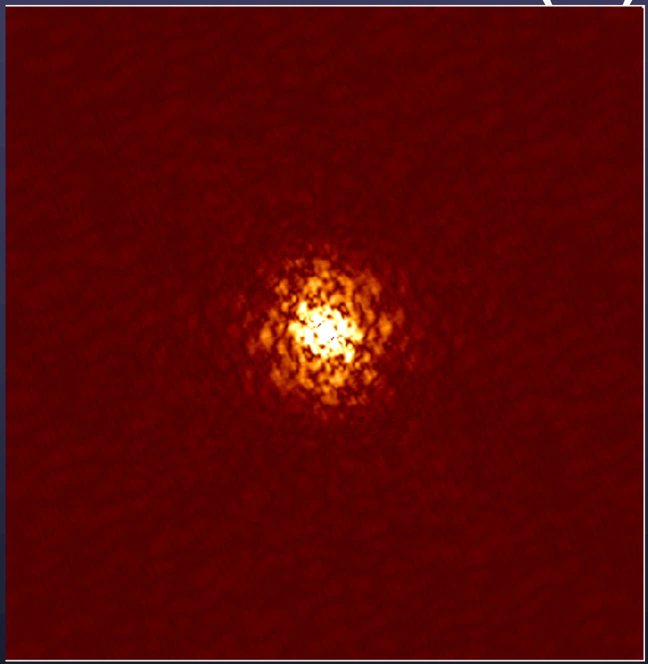
- **For small w (small max. baseline) or small field of view *($l^2 + m^2 << 1$)* *I(l,m)* is 2D Fourier transform of *V(u,v)***

$$V^{\circ}(u,v) = \int\int I(l,m)e^{2\pi\iota[ul+vm]}dl\,dm$$

$$V^{\circ}(u,v) \rightleftharpoons I(l,m)$$

Imaging: Ideal 2D Fourier relationship

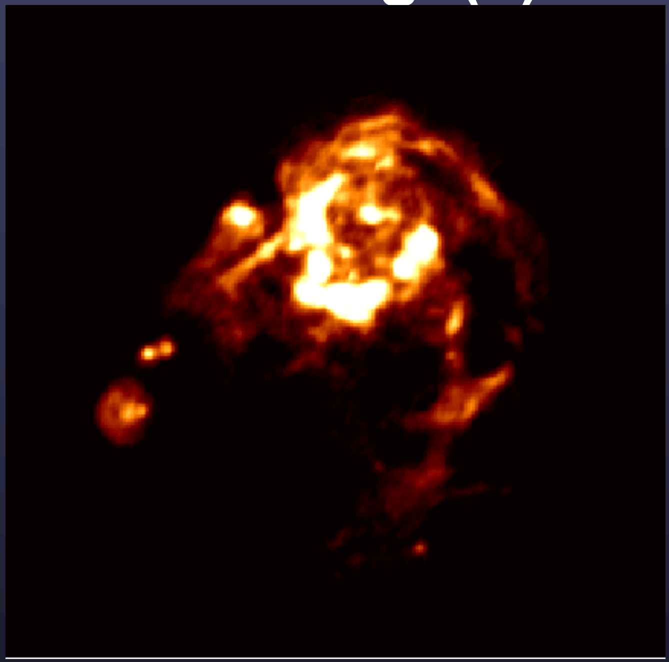
Ideal visibilities(*V*)



FT

⇔

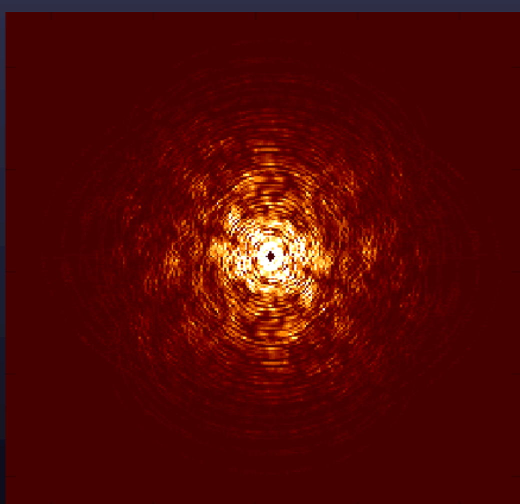
True image(*I*)



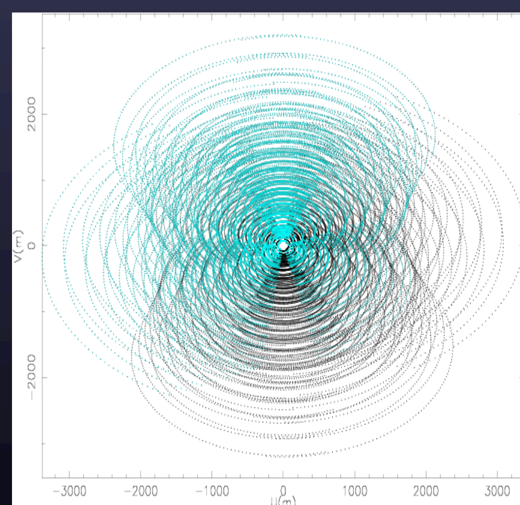
- **This is true ONLY if *V* is measured for all *(u,v)*!**

Imaging: (u,v) plane sampling

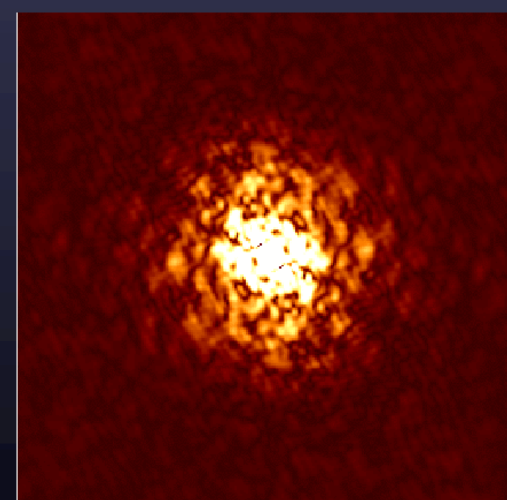
- With limited number of antennas, the (u,v) plane is sampled at discrete points:



V^M =



S x



V^O

Convolution with the PSF

- **Effect of sampling the (u,v) plane:**

$$I^d(l,m) = FT^{-1}\left[V^\circ(u,v)S(u,v)\right]$$

- **Using the Convolution Theorem:**

$$I^d(l,m) = B(l,m) \star I^\circ(l,m)$$

The Dirty Image (I^d) is the convolution of the True Image (I^o) and the Dirty Beam/Point Spread Function (B)

$$B = FT^{-1}(S)$$

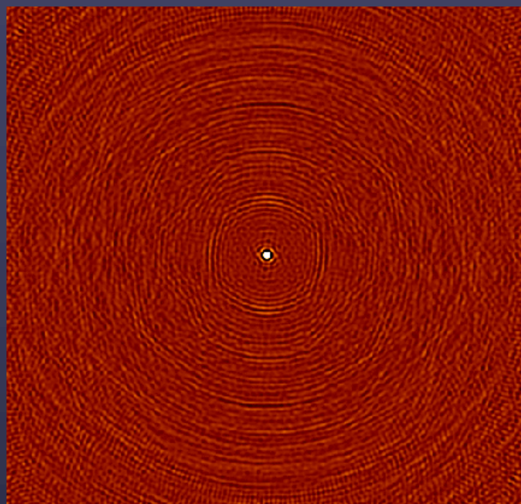
- **In practice**

$$I^d = B * I^o + B * I^N \quad \text{where} \quad I^N = FT^{-1}(\text{Vis. Noise})$$

- **To recover I^o, we must deconvolve B from I^d. The algorithm must also separate $B*I^o$ from $B*I^N$.**

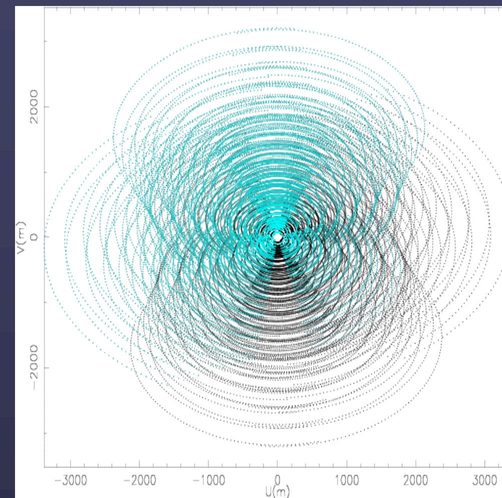
The Dirty Image

The PSF



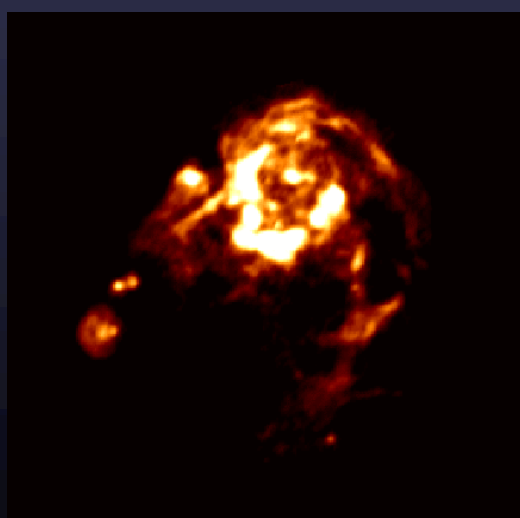
FT

↔

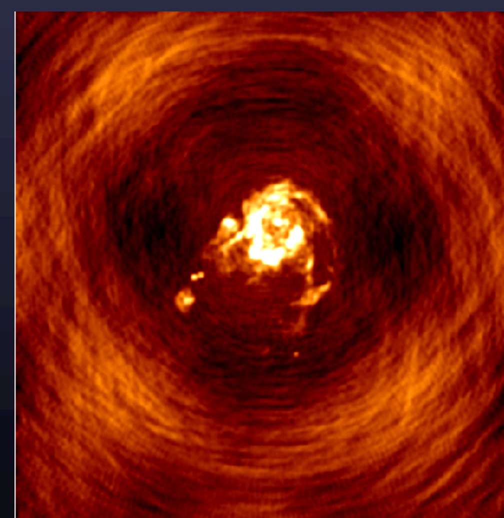


u,v coverage

*****



→



The

Dirty Image

Making of the Dirty Image

- **Fast Fourier Transform (FFT) is used for efficient Fourier transformation. It however requires regularly spaced grid of data.**
- **Measured visibilities are irregularly sampled (along u,v tracks).**
- **Convolutional gridding is used to effectively interpolate the visibilities everywhere and then re-sample them on a regular grid (the Gridding operation)**

Dirty Beam: Interesting properties

- **PSF is a weighted sum of cosines corresponding to the measured fourier components:**

$$B(l,m) = \frac{\sum_k W_k cos(u_k l + v_k m)}{\sum_k W_k}$$

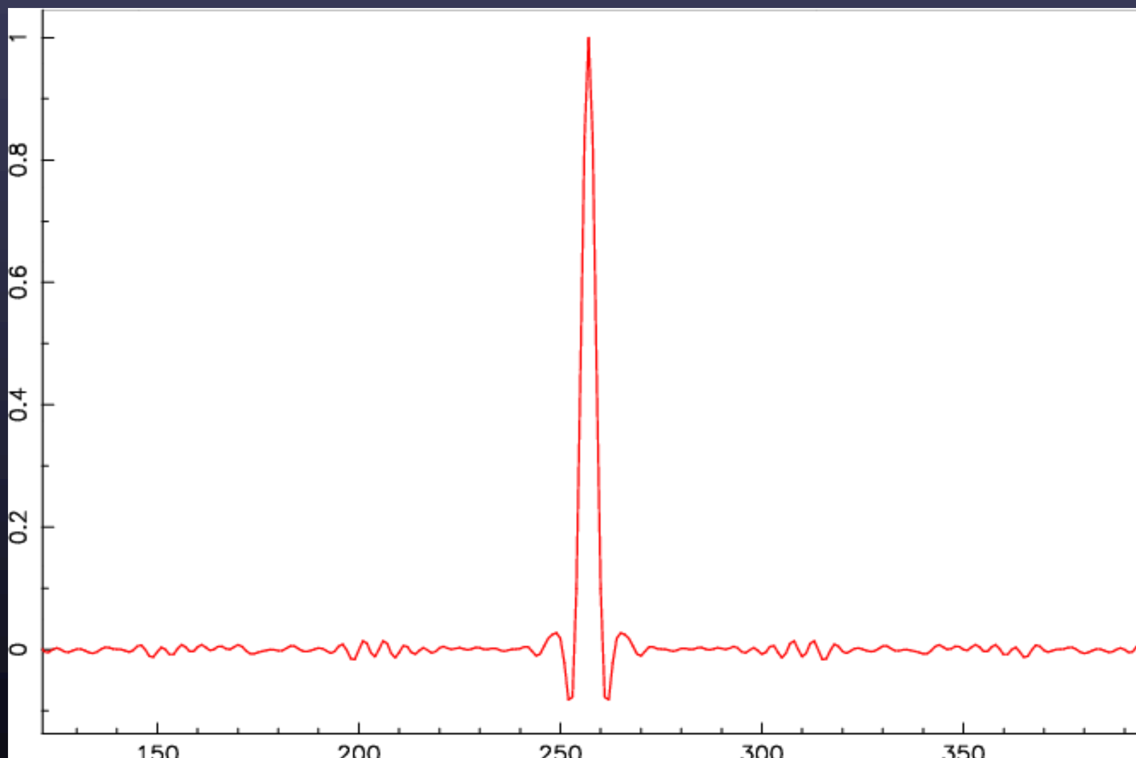
Visibility weights (w_i) are also gridded on a regular grid and FFT used to compute the Dirty Beam (aka the PSF).

- **The peak of the PSF is normalized to 1.0**
- **The 'main lobe' has a size $\Delta x \sim 1/u_{max}$ and $\Delta y \sim 1/v_{max}$**

This is the 'diffraction limited' resolution (the Clean Beam) of the telescope.

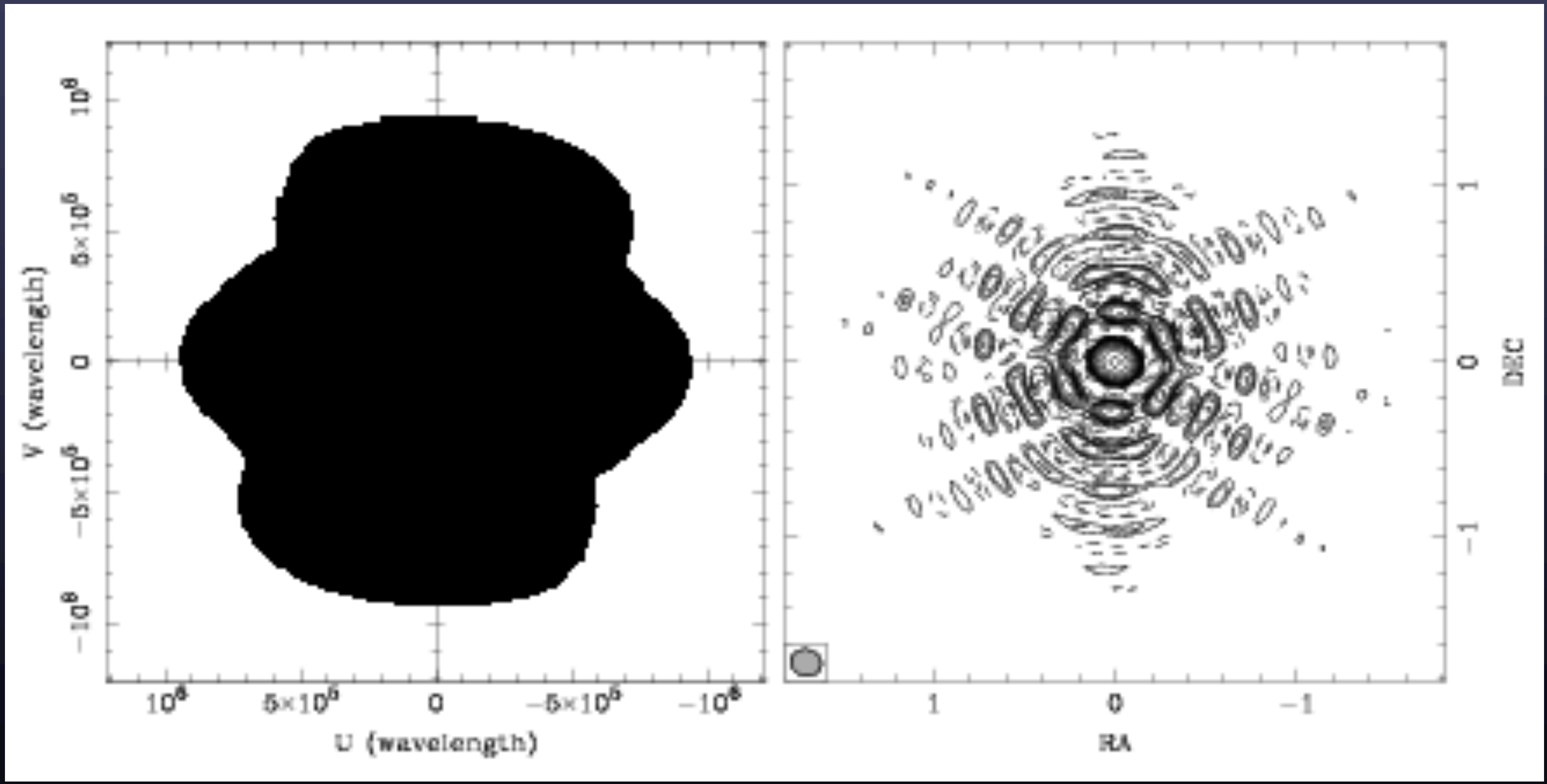
Dirty Beam: Interesting properties

- **Side lobes extend indefinitely**
- **RMS ~ *1/N*   where *N* = No. of antennas**

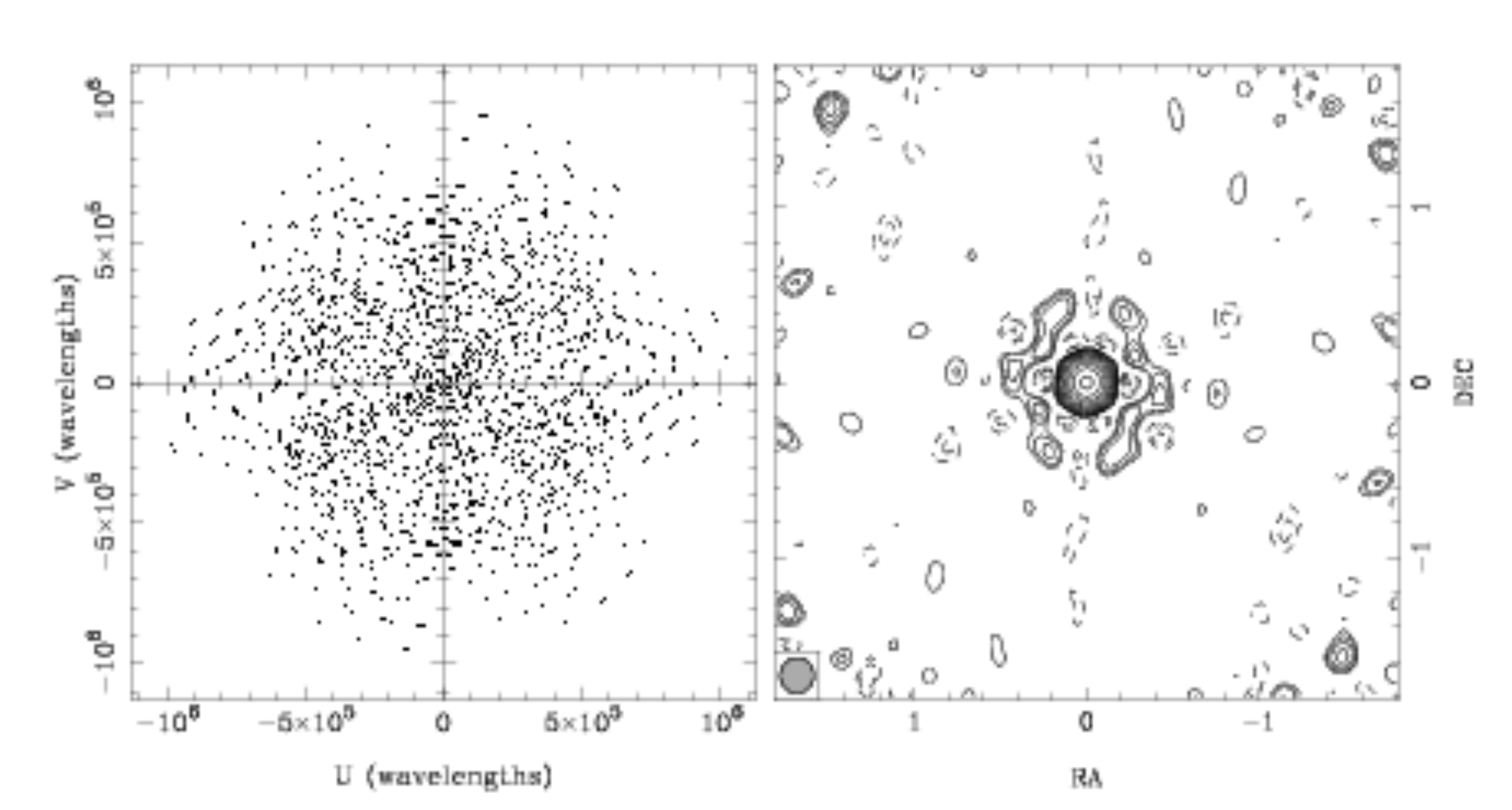


Close-in side lobes of the PSF

- Close-in sidelobes of the PSF are controlled by the (u,v) coverage envelope.



Close-in side lobes: VLA (u,v) coverage



PSF forming: Weighting...

- Weighting function (W_k) can be chosen to modify resolution and side lobes
- Natural Weighting

  $W_k=1/\sigma_k^2$ where σ_k^2 is the RMS noise

→ Best RMS across the image.

→ Large scales (smaller baselines) have higher weights.

→ Effective resolution less than the inverse of the longest baseline.

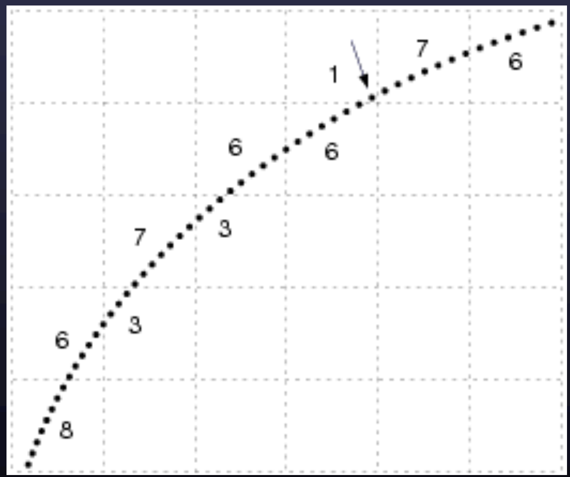
...Weighting...

- **Uniform weighting**

  $W_k = 1/\rho(u_k, v_k)$ **where $\rho(u_k,v_k)$ is the density of uv-points in the k^{th} cell.**
- **Short baselines (large scale features in the image) are weighted down.**
- **Relatively better resolution**
- **Increases the RMS noise.**
- **Super uniform weighting:**

  Consider density over larger region.

  Minimize side lobes locally.

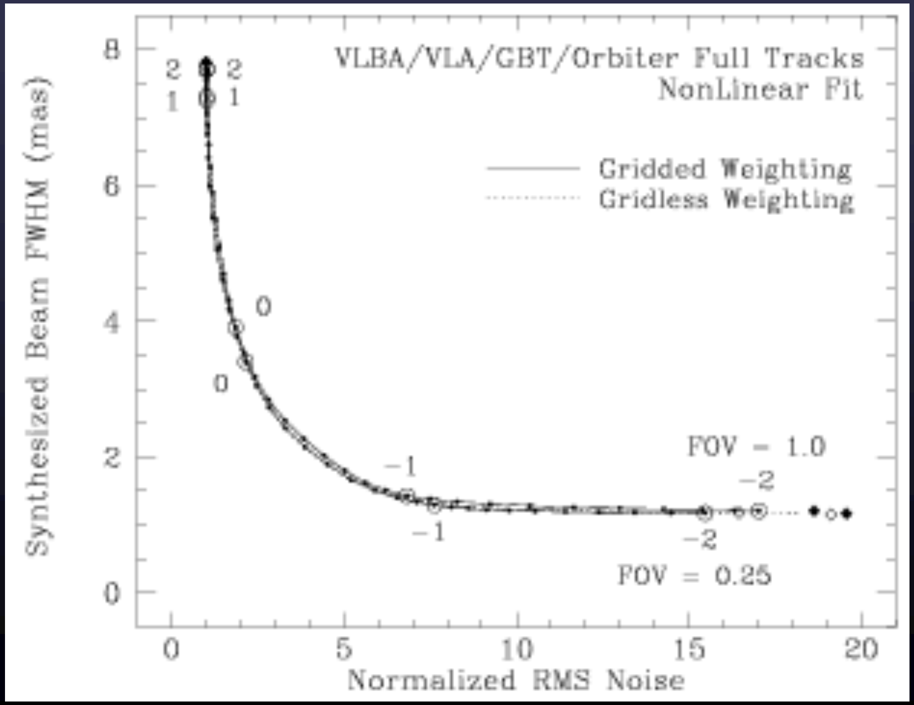


...Weighting

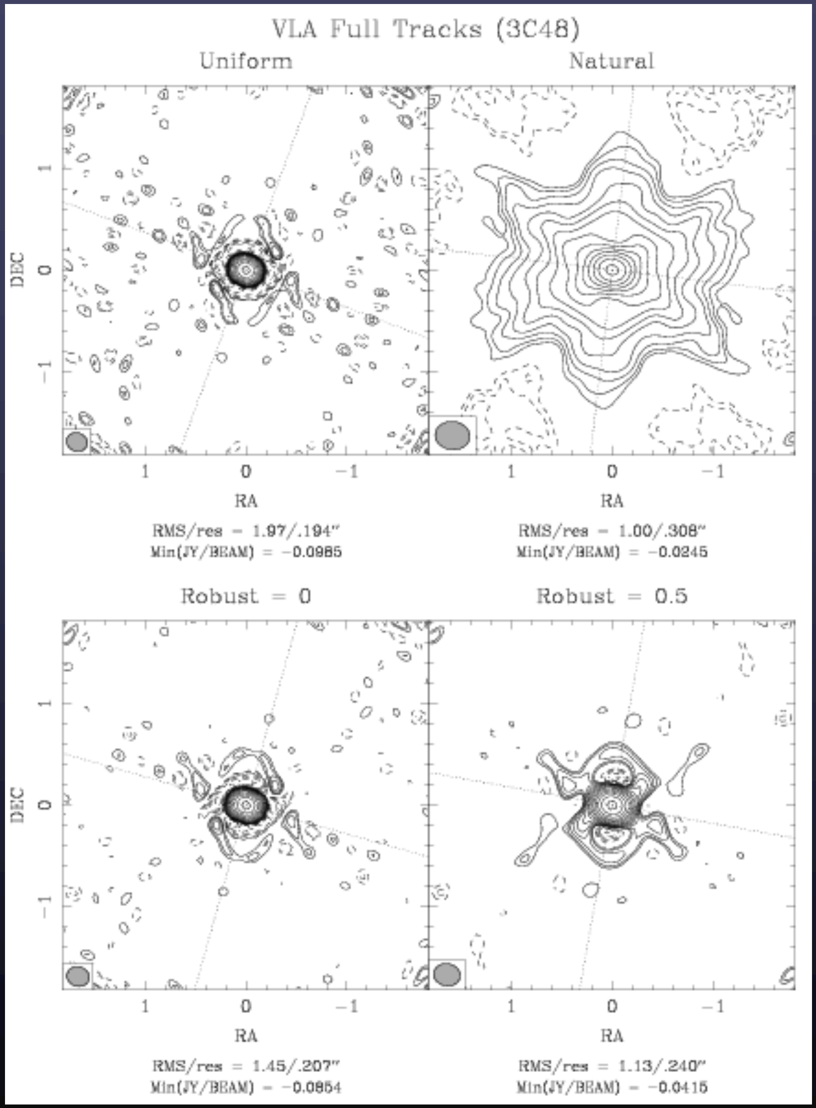
- **Robust/Briggs weighting:**

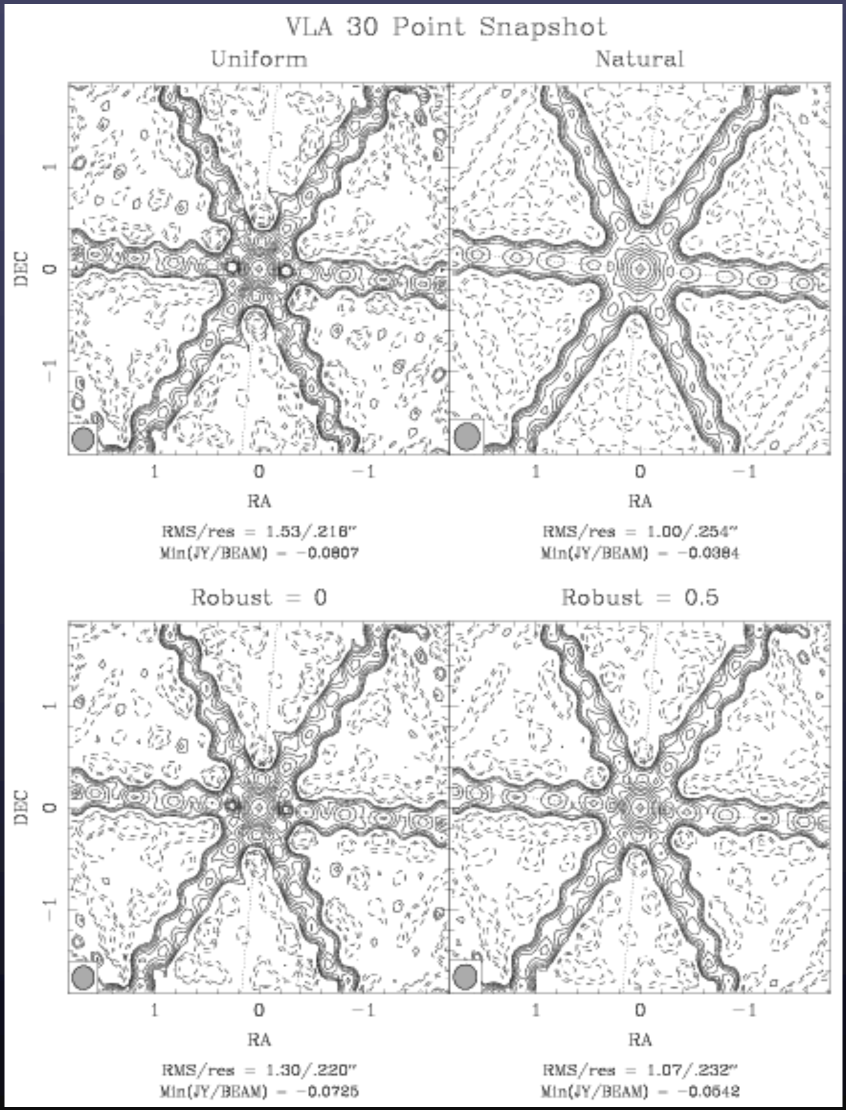
$$W_k = 1/[S.\rho(u_k, v_k) + \sigma_k^2]$$

- **Parameterized filter – allows continuous variation between optimal resolution (uniform weighting) and optimal noise (natural weighting).**



Examples of weighting





PSF Forming: Tapering

- **The PSF can be further controlled by applying a tapering function on the weights (e.g. such that the weights smoothly go to zero beyond the maximum baseline).**

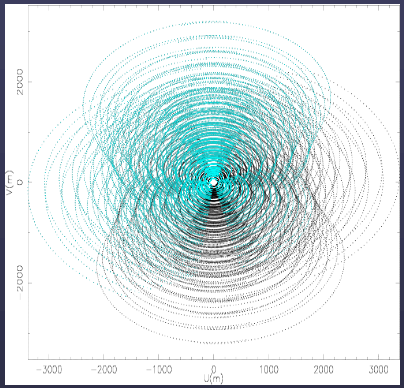
$$W'_k = T(u_k, v_k)\, W_k(u_k, v_k)$$

- **Bottom line on weighting/tapering:**

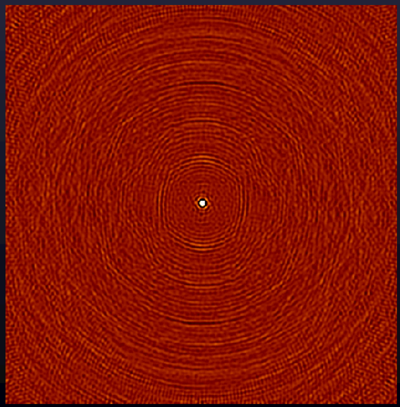
These help a bit, but imaging quality is limited by the deconvolution process!

The missing information

- As seen earlier, not all parts of the uv-plane are sampled – **the 'invisible distribution'**



1. "Central hole" below u_{min} and v_{min}:
   - Image plane effect: Total integrated power is not measured.
   - Upper limit on the largest scale in the image plane.
2. No measurements beyond u_{max} and v_{max}:
   - Size of the main lobe of the PSF is finite (finite resolution).



3. Holes in the (u,v) plane:
   - Contribute to the side lobes of the PSF.

Recovering the missing information

- For information beyond the max. baseline, one requires extrapolation. That's un-physical (unconstrained).
- Information corresponding to the "central hole": possible, but difficult (need extra information).
- Information corresponding to the (u,v) holes: requires interpolation. The measurements provide constraints – hence possible. **But non-linear methods necessary.**

Deconvolution = interpolation in the visibility plane.

Prior knowledge about the sky

- ***What can we assume about the sky emission:***
  1. ***Sky does not look like cosine waves***
  2. ***Sky brightness is positive (but there are exceptions)***
  3. ***Sky is a collection of point sources (weak assertion)***
  4. ***Sky could be smooth***
  5. ***Sky is mostly blank (sometimes justifies "boxed" deconvolution)***
- ***Non-linear deconvolution algorithms search for a model image I^M such that the residual visibilities $V^R=V^o-V^M$ are minimized, subject to the constraints given by the (assumed) prior knowledge.***

The classic Clean algorithm (Hogbom, 1974)

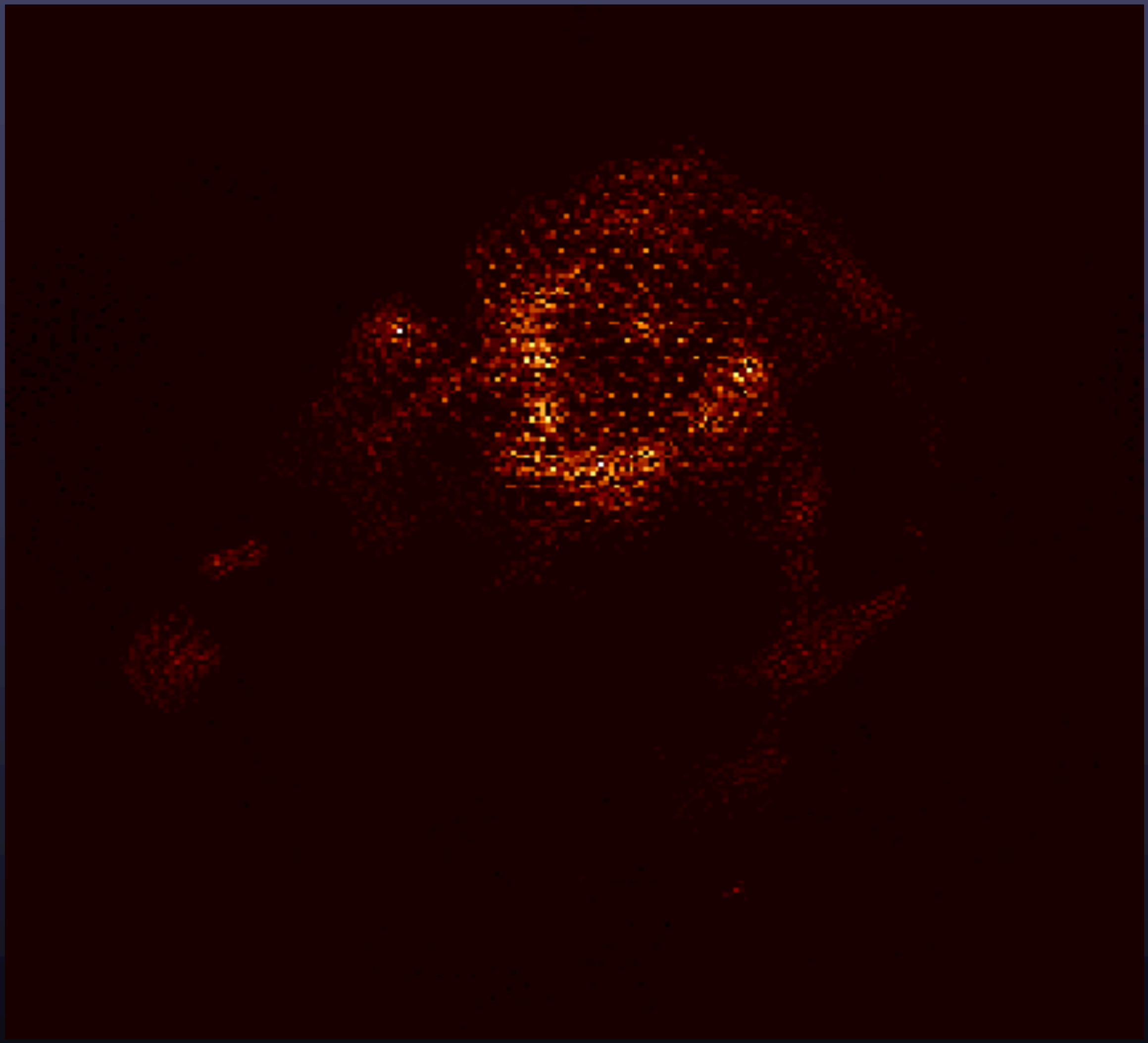
- **Prior knowledge:**
  - **- sky is composed of point sources**
  - **- mostly blank**
- **Algorithm:**
  1. **Search for the peak in the dirty image.**
  2. **Add a fraction *g* (loop gain) of the peak value to I^M.**
  3. **Subtract a scaled version of the PSF from the position of the peak.**

     $I^R_{i+1} = I^R_i - g.B.max(I^R_i)$
  4. **If residuals are not "noise like", goto 1.**
  5. **Smooth I^M by an estimate of the main lobe (the "clean beam") of the PSF and add the residuals to make the "restored image"**

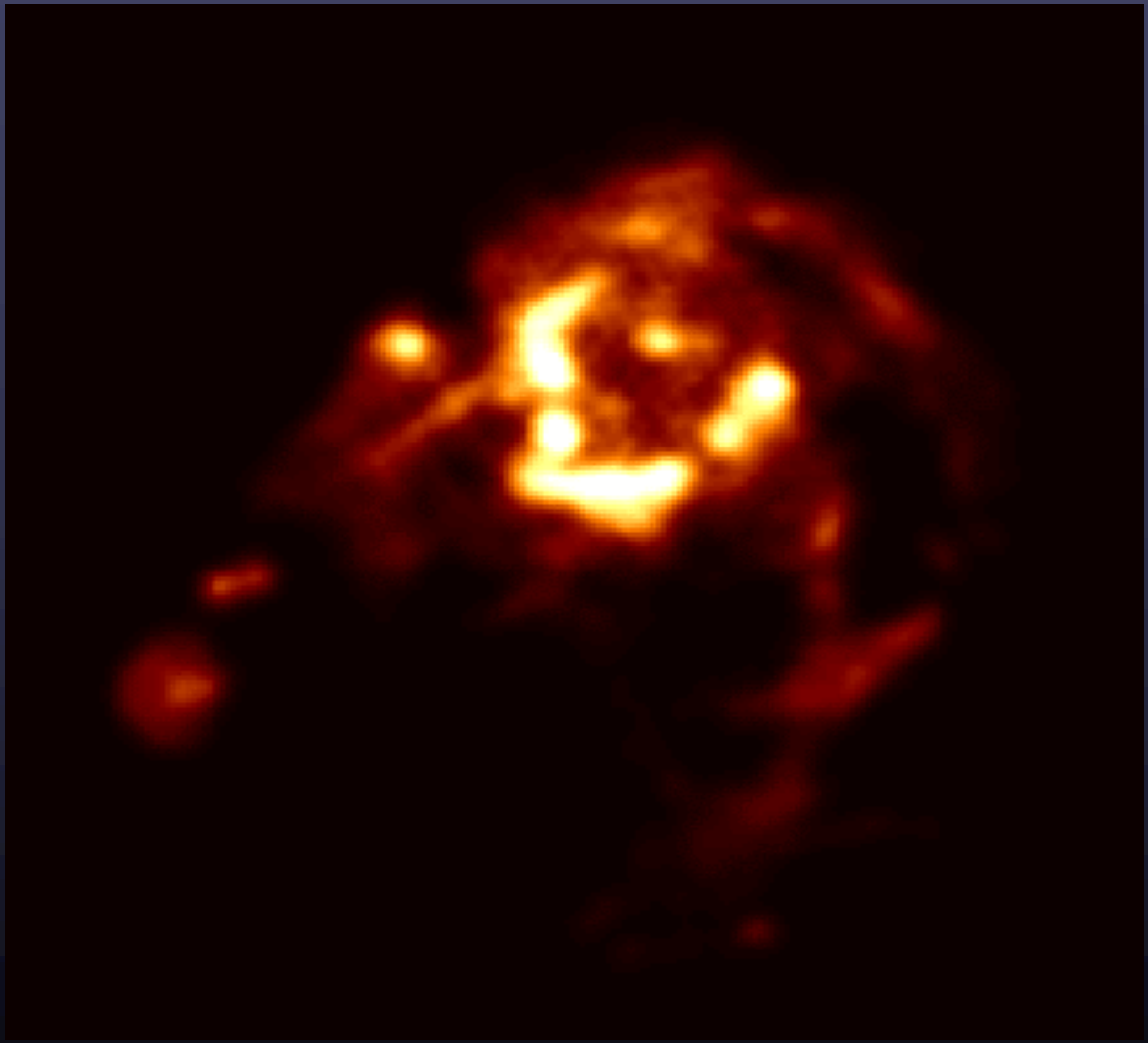
Details of Clean

- It is a steepest descent minimization.
- Model image is a collection of delta functions – a scale insensitive algorithm.
- A least square fit of sinusoids to the visibilities which is proved to converge (Schwarz 1978).
- Stabilized by keeping a small loop gain (usually g=0.1-0.2).
- Stopping criteria: either the max. iterations or max. residuals some multiple of the expected peak noise.
- Search space constrained by user defined windows.
- χ Ignores coupling between pixels (extended emission) – assumes an orthogonal search space.

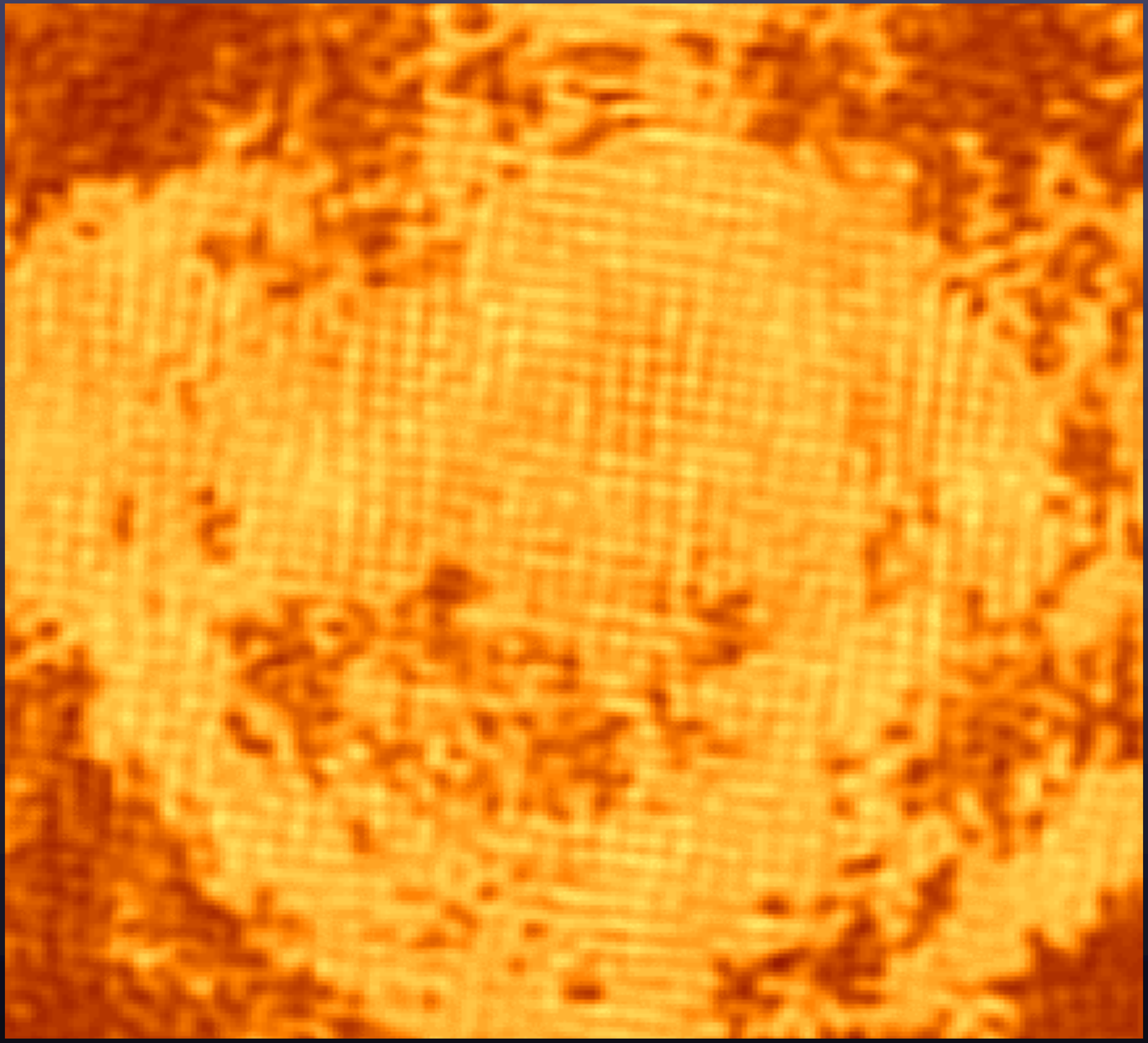
Clean: Model



Clean: Restored

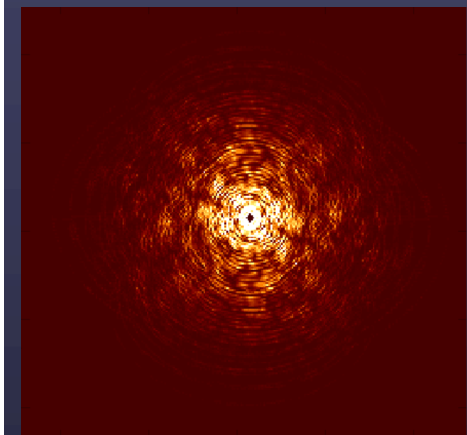


Clean: Residual

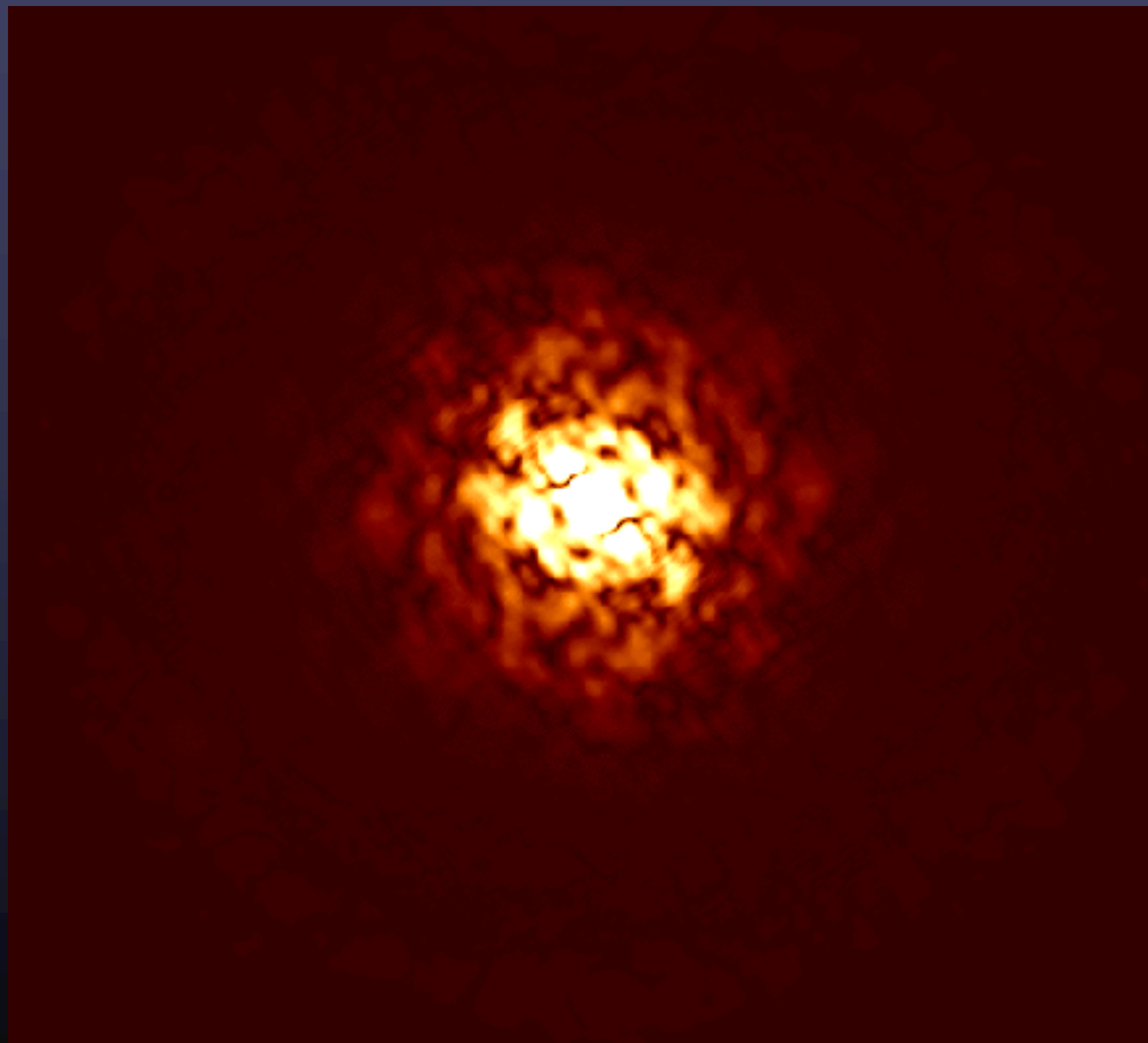


Clean: Model visibilities

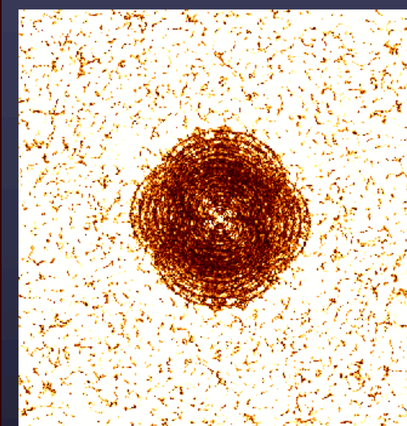
Sampled Vis



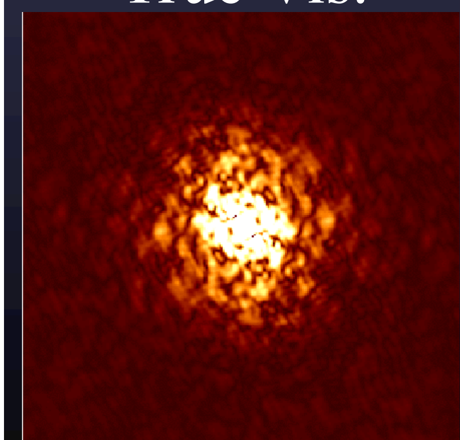
Model Vis.



Residual Vis.



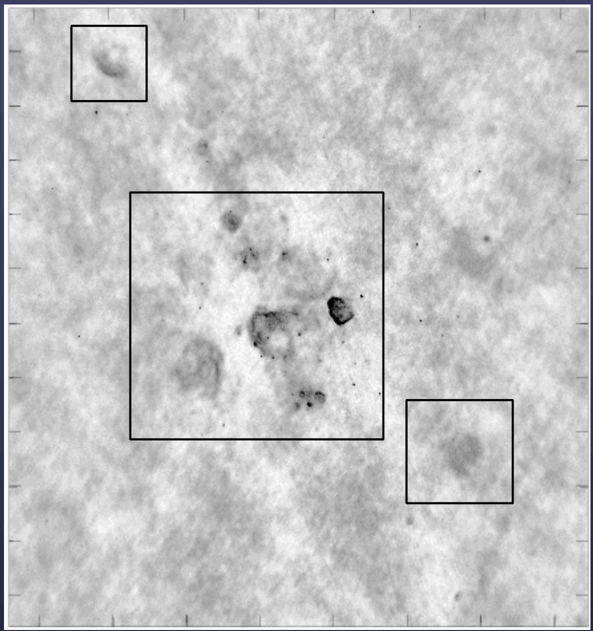
True Vis.



Role of boxes

- Limit the search for components to only parts of the image.

  A way to regularize the deconvolution process.
- Useful when small no. of visibilities (e.g. VLBI/snapshots).
- Do not over-Clean within the boxes (over-fitting).



- Deeper Clean with no/loose boxes and lower loop gain can achieve similar (more objective) results.
- Stop when Cleaning within the boxes has no global effect (insignificant coupling of pixels due to the PSF).

Inspecting Visibility Data

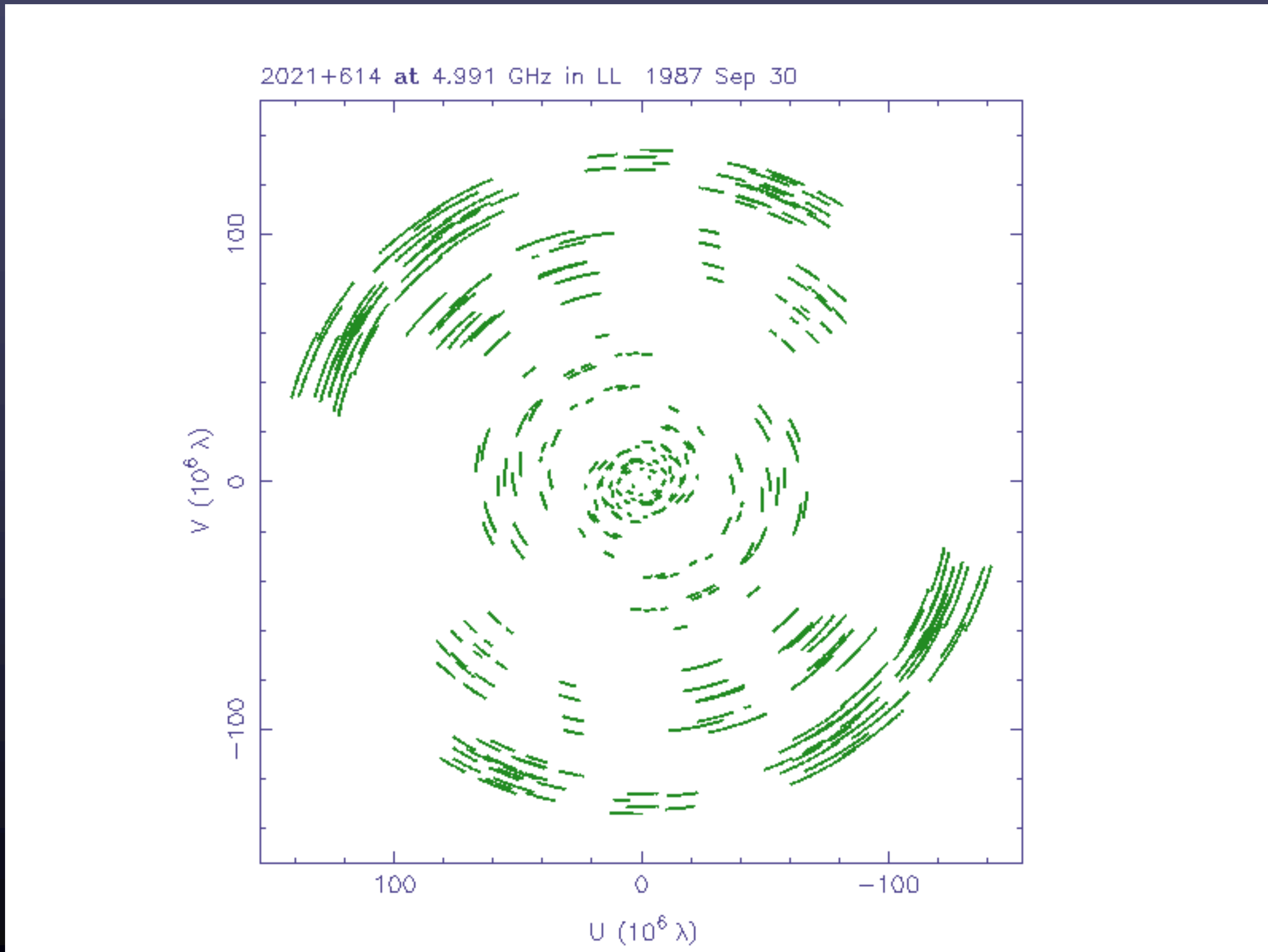
Useful displays

- Sampling of the (*u*,*v*) plane
- Amplitude and phase *vs*. radius in the (*u*,*v*) plane
- Amplitude and phase *vs*. time on each baseline
- Amplitude variation across the (*u*,*v*) plane
- Projection onto a particular orientation in the (*u*,*v*) plane

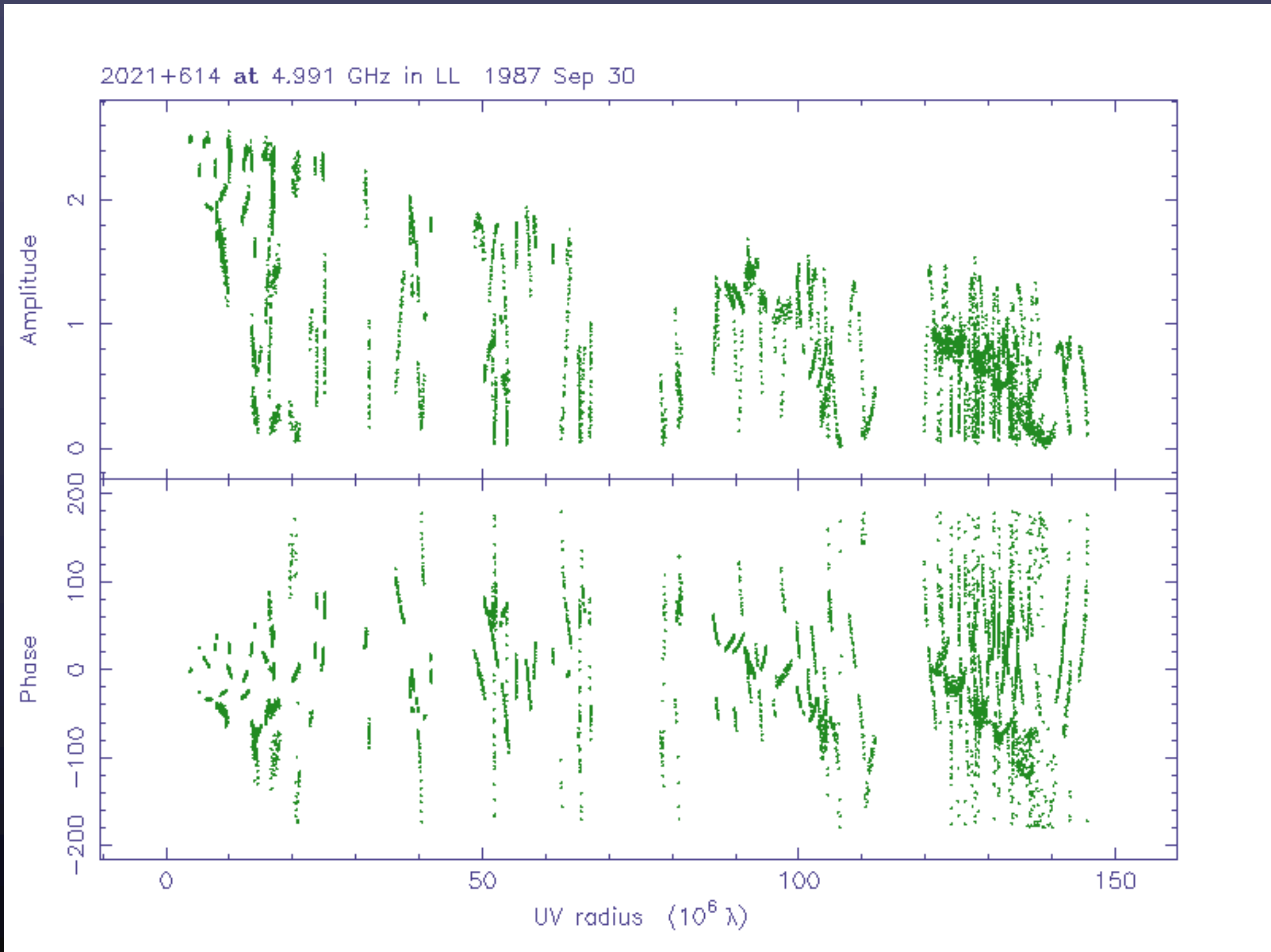
Example: 2021+614

- GHz-peaked spectrum radio galaxy at z=0.23
- A VLBI dataset with 11 antennas from 1987
- VLBA only in 2000

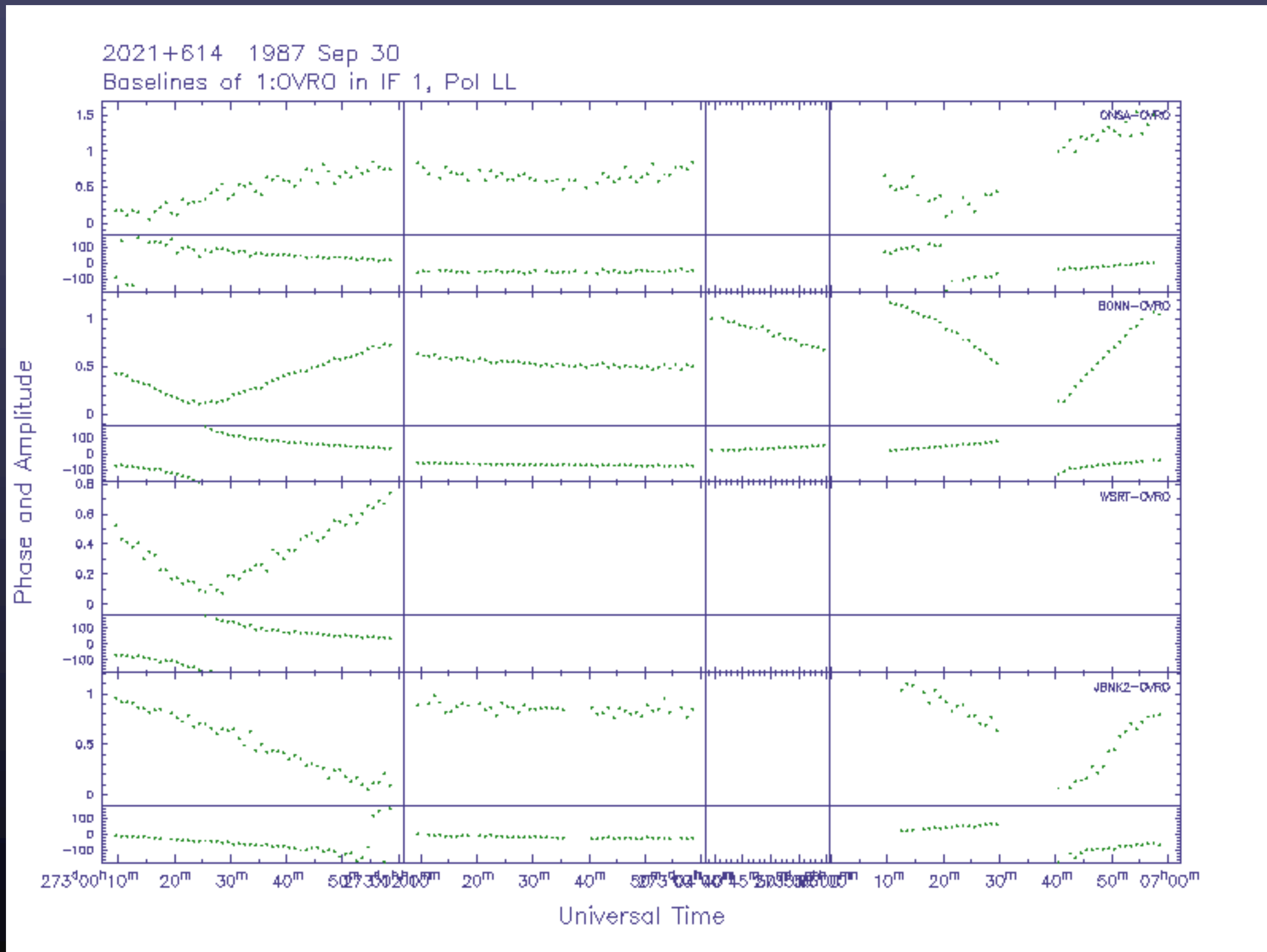
Sampling of the (u,v) plane



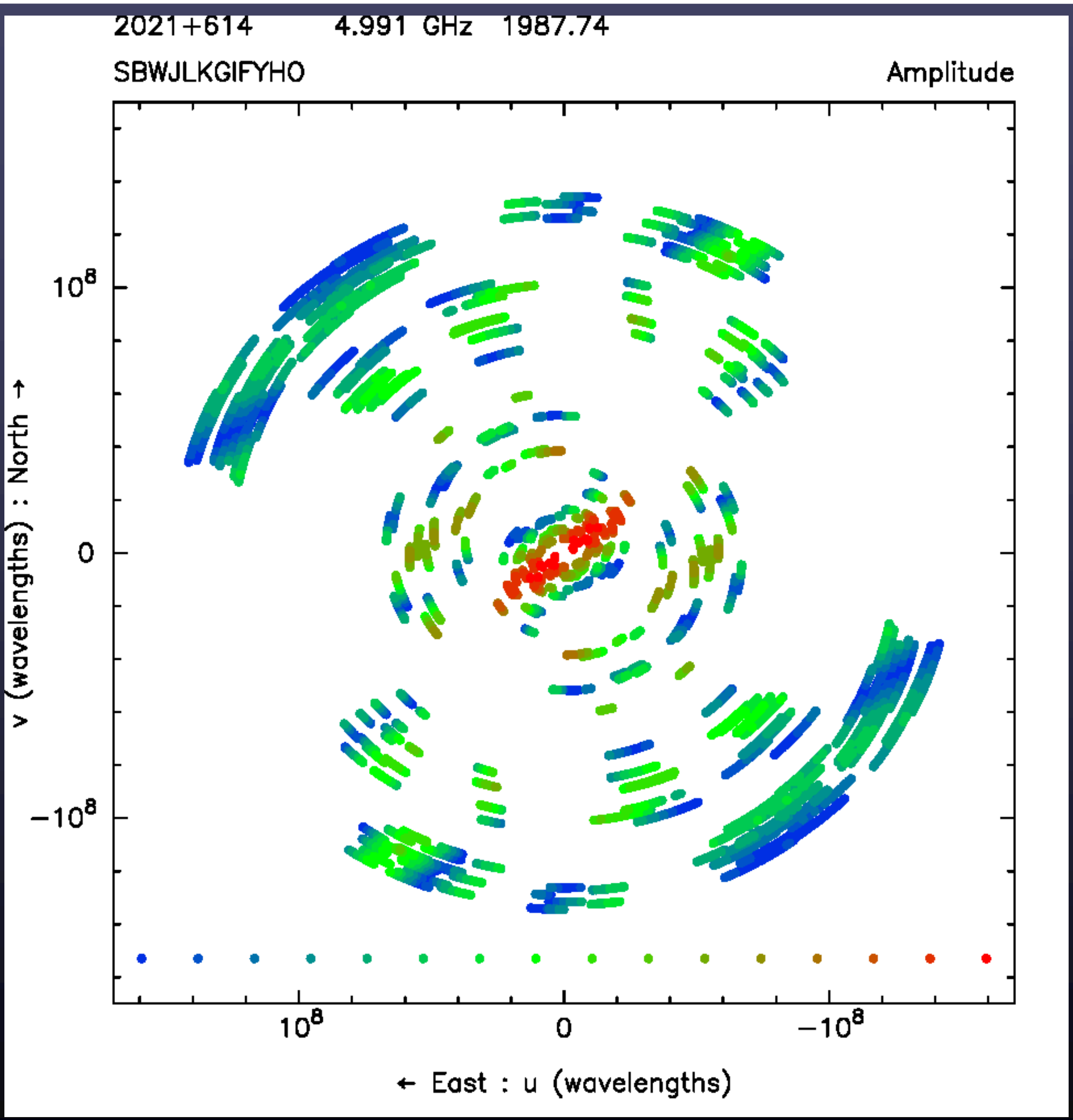
Visibility versus (u,v) radius



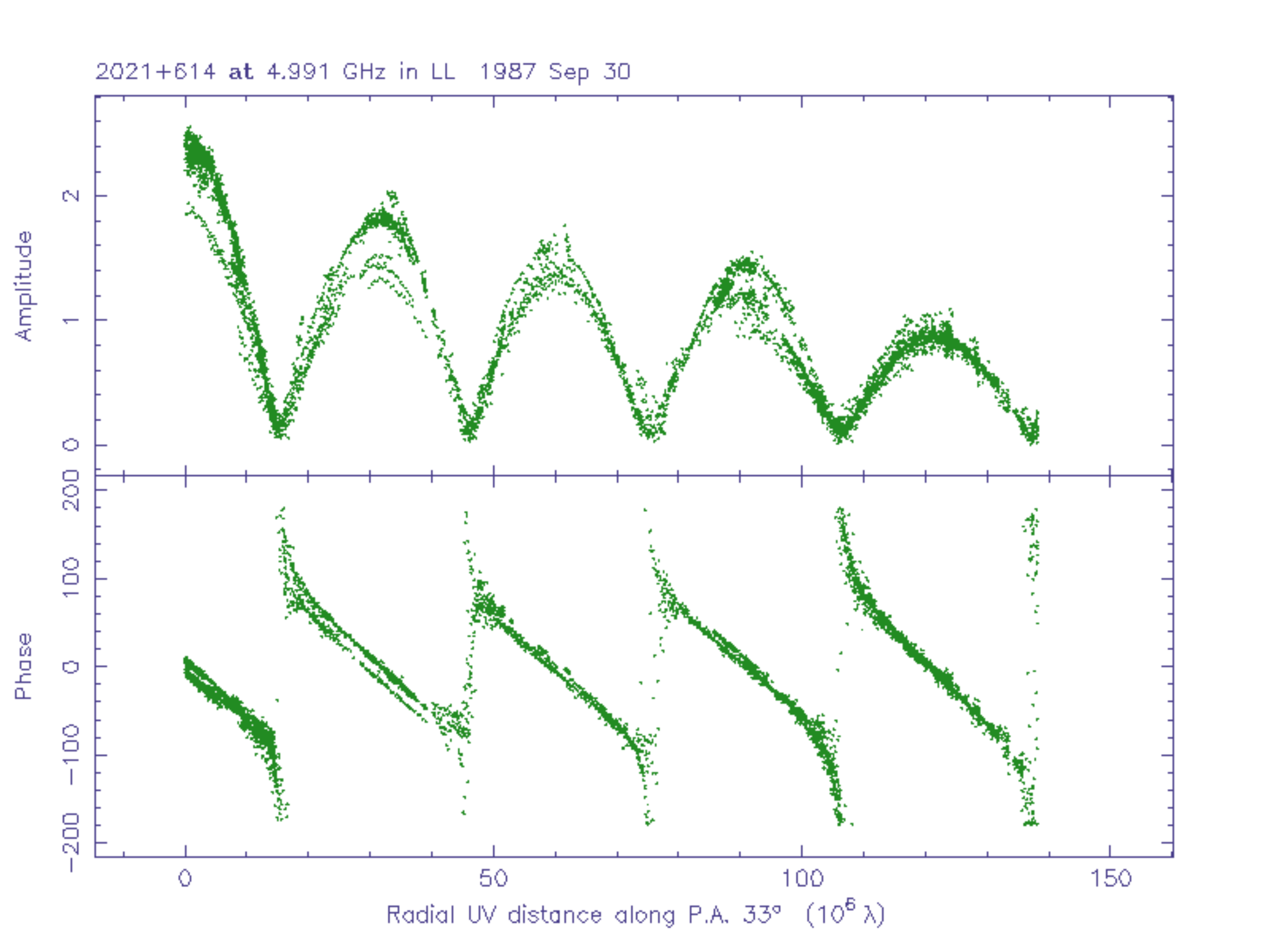
Visibility versus time

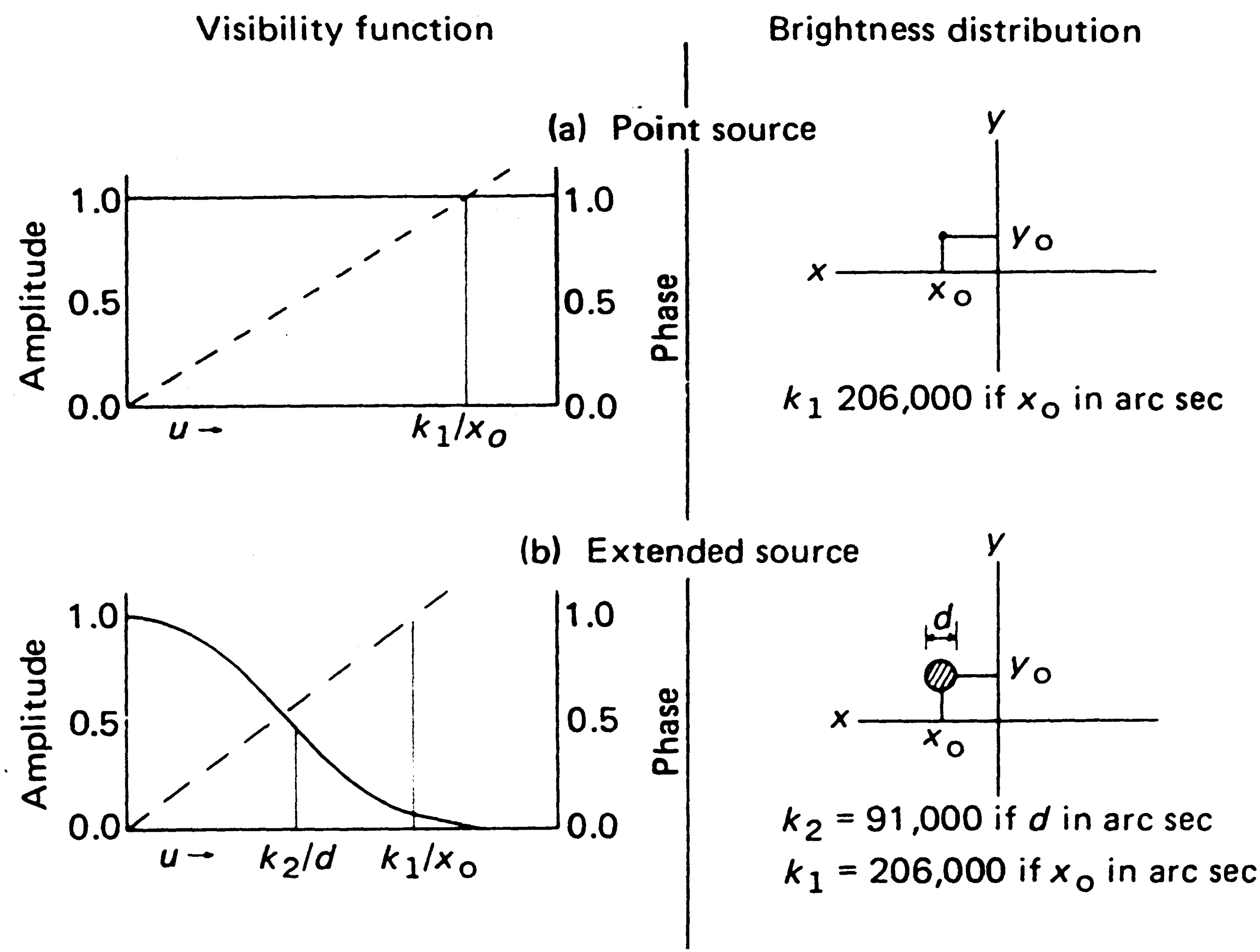


Amplitude across the (u,v) plane



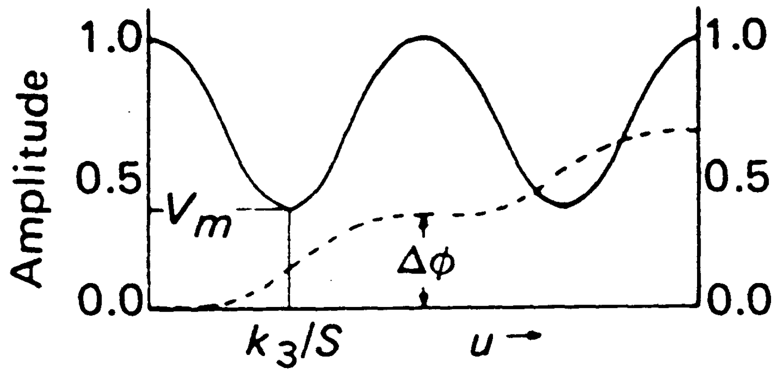
Projection in the (u,v) plane

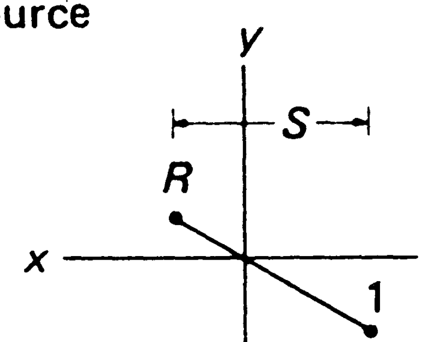




Visibility function
Brightness distribution
(a) Point source
1.0
0.5
0.0
Amplitude
Phase
u →
k1/xo
y
x
yo
xo
k1 206,000 if xo in arc sec
(b) Extended source
k2/d
d
k2 = 91,000 if d in arc sec
k1 = 206,000 if xo in arc sec

(c) Point double source

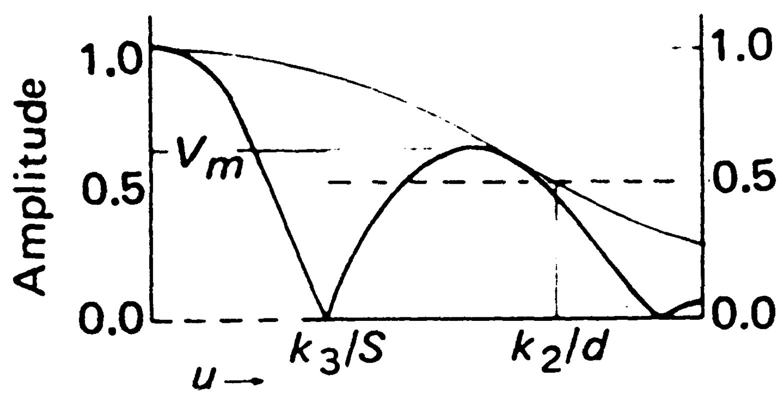


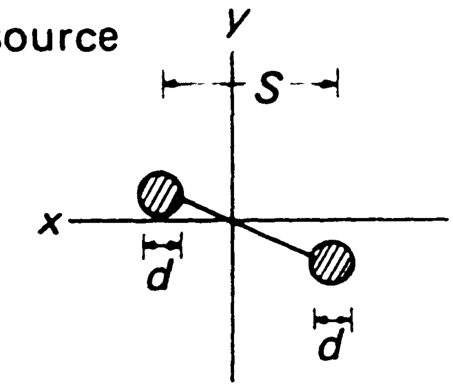


k_3 = 103,000 if S in arc sec

$$V_m = \frac{R - 1}{R + 1} \quad ; \quad \Delta\phi = \frac{1}{1 + R}$$

(d) Extended double source

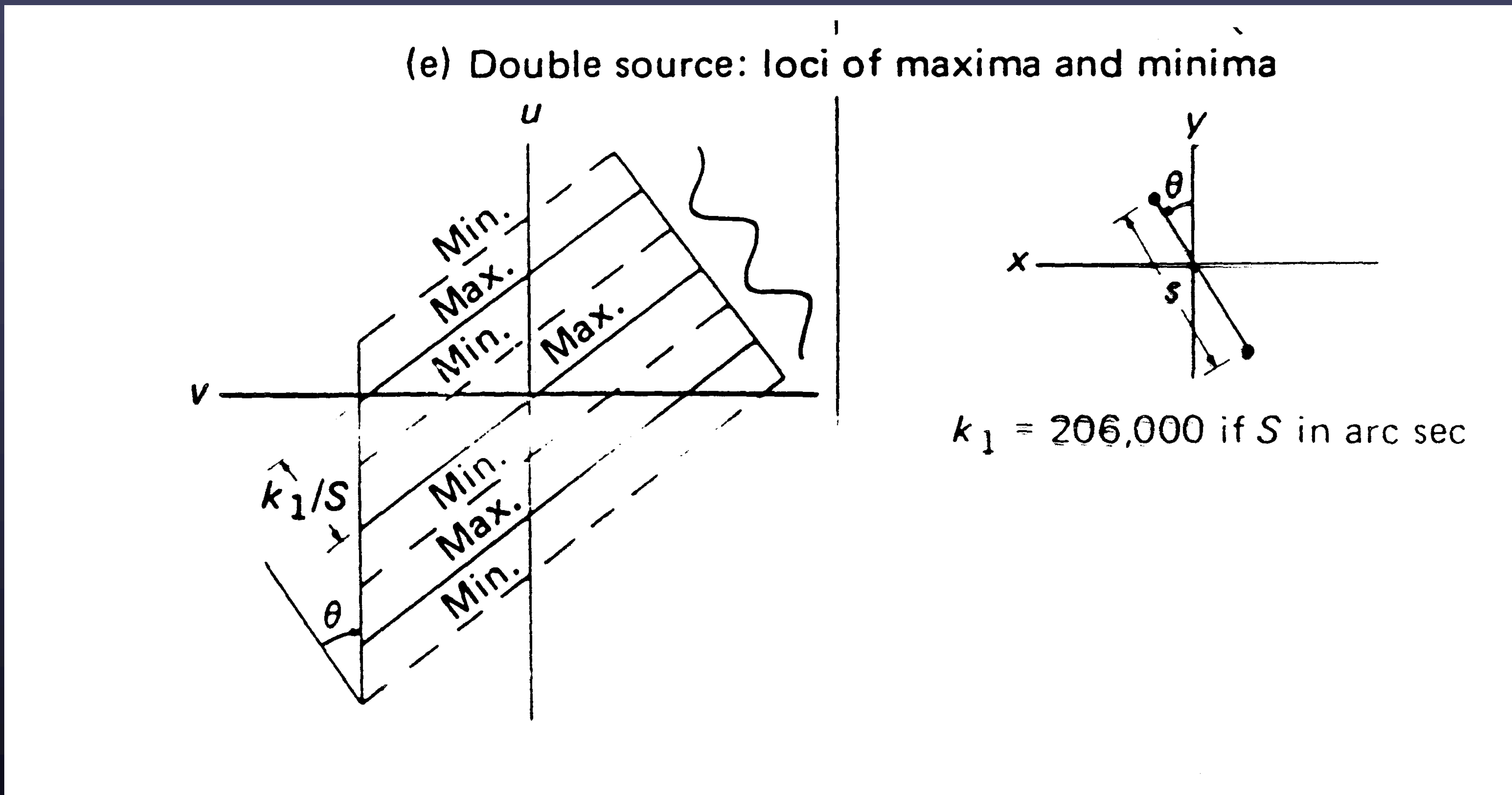




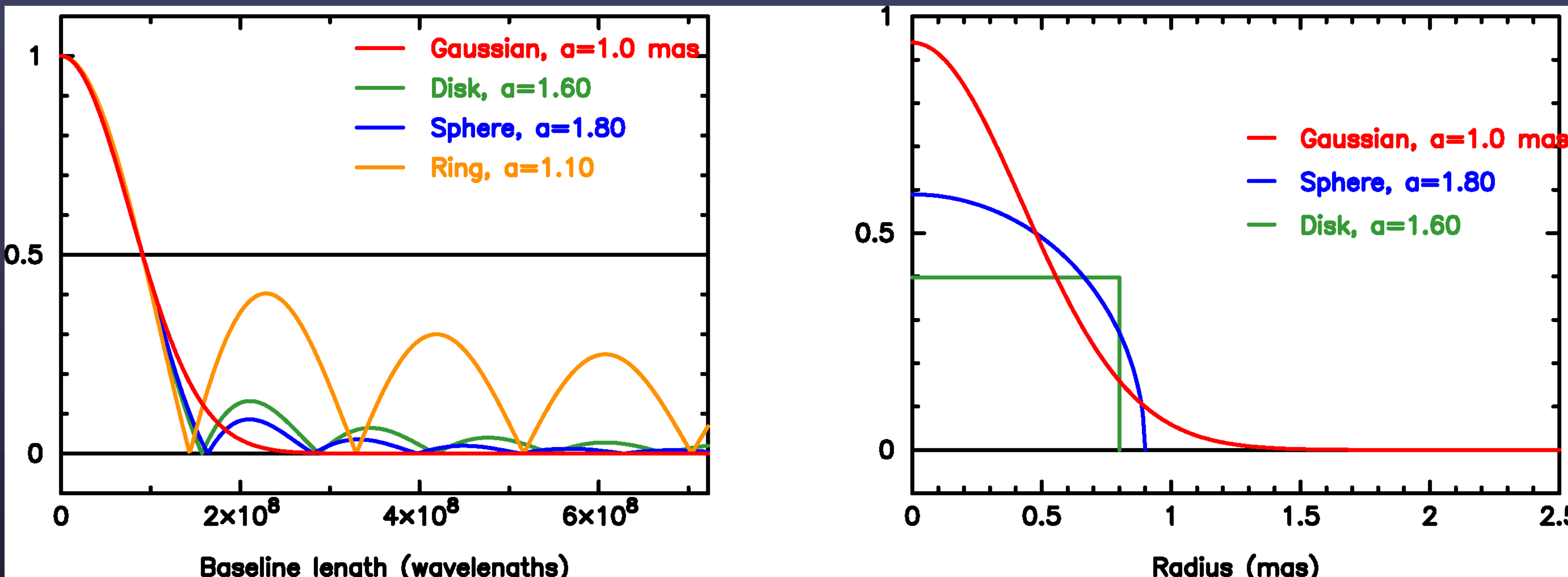
k_3 = 103,000 if S in arc sec

k_2 = 91,000 if d in arc sec

$$V_m \approx \exp\left\{-3.57 \left(\frac{d}{S}\right)^2\right\}$$


(e) Double source: loci of maxima and minima
u
v
Min.
Max.
Min.
Max.
Min.
Max.
Min.
k1/S
θ
y
x
θ
s
k1 = 206,000 if S in arc sec

Simple models



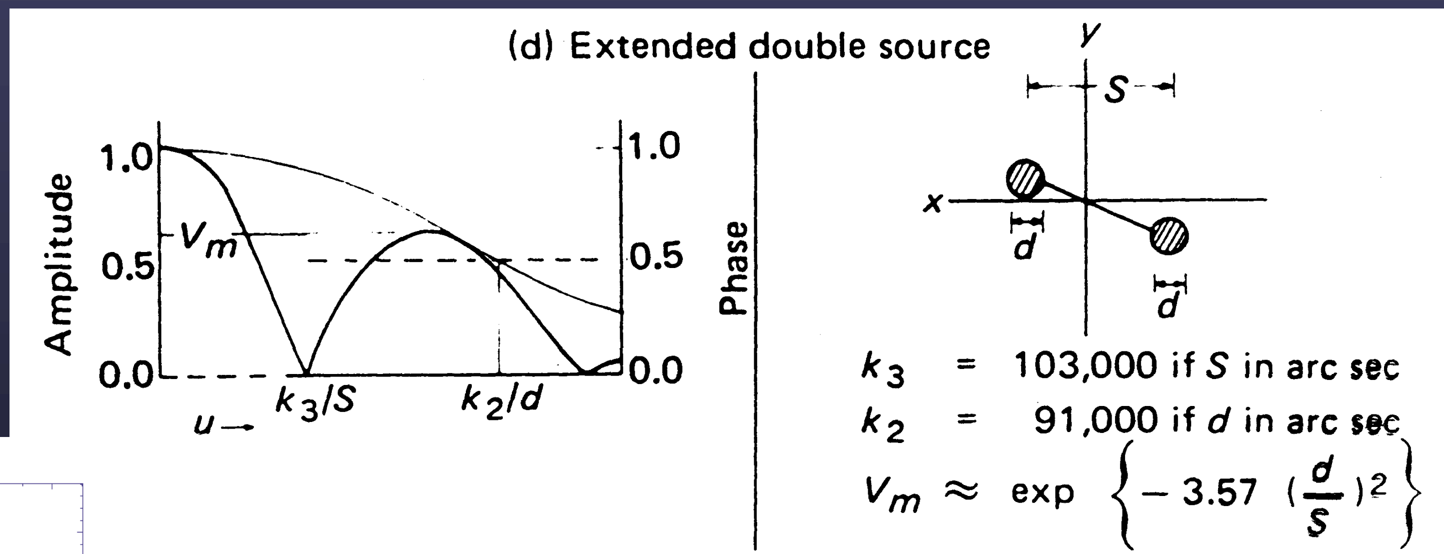
Visibility at short baselines contains little information about the profile of the source.

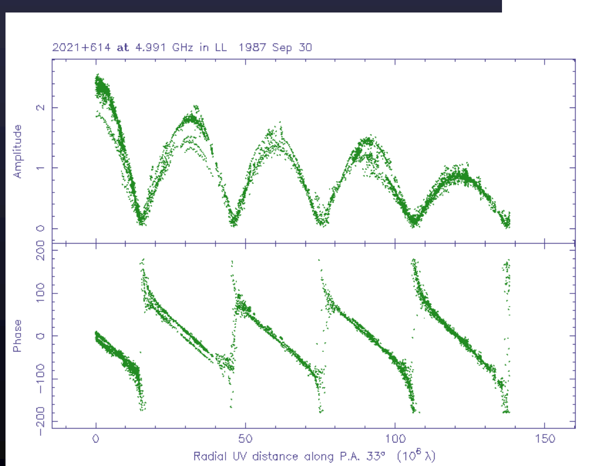
Trial model

By inspection, we can derive a simple model:

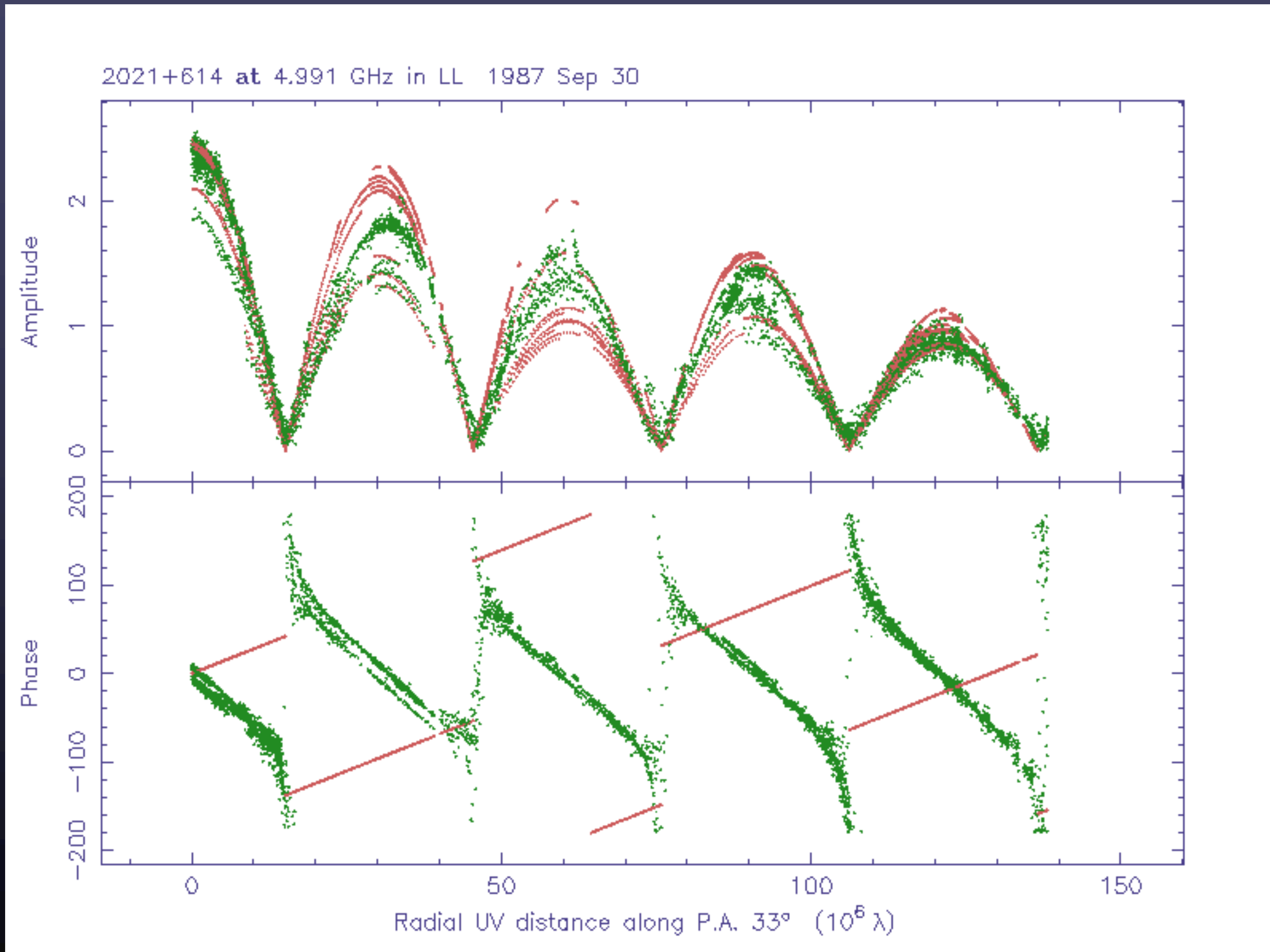
Two equal components, each 1.25 Jy, separated by about 6.8 milliarcsec in p.a. 33º, each about 0.8 milliarcsec in diameter (gaussian FWHM)

To be refined later…

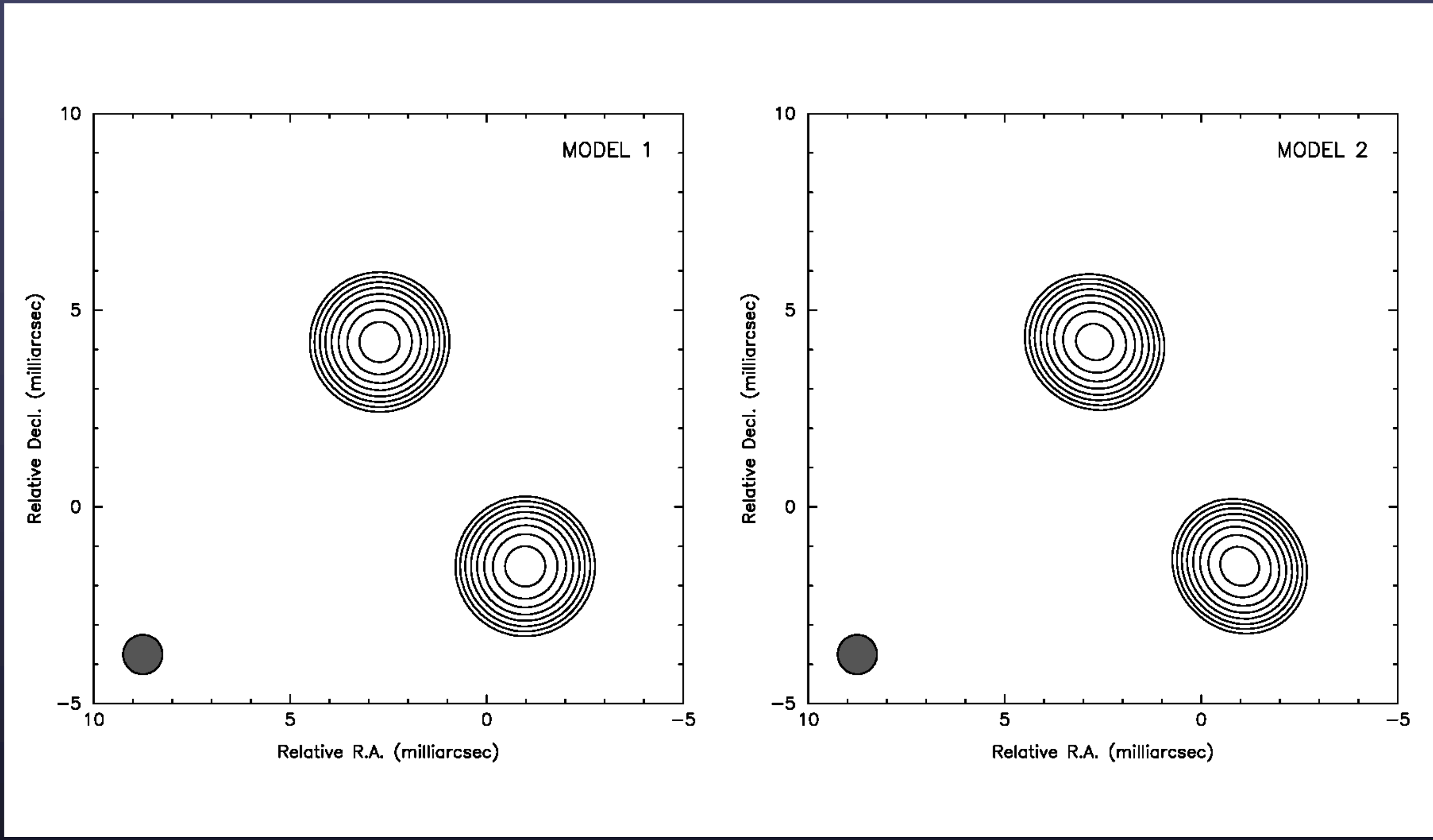




Projection in the (u,v) plane

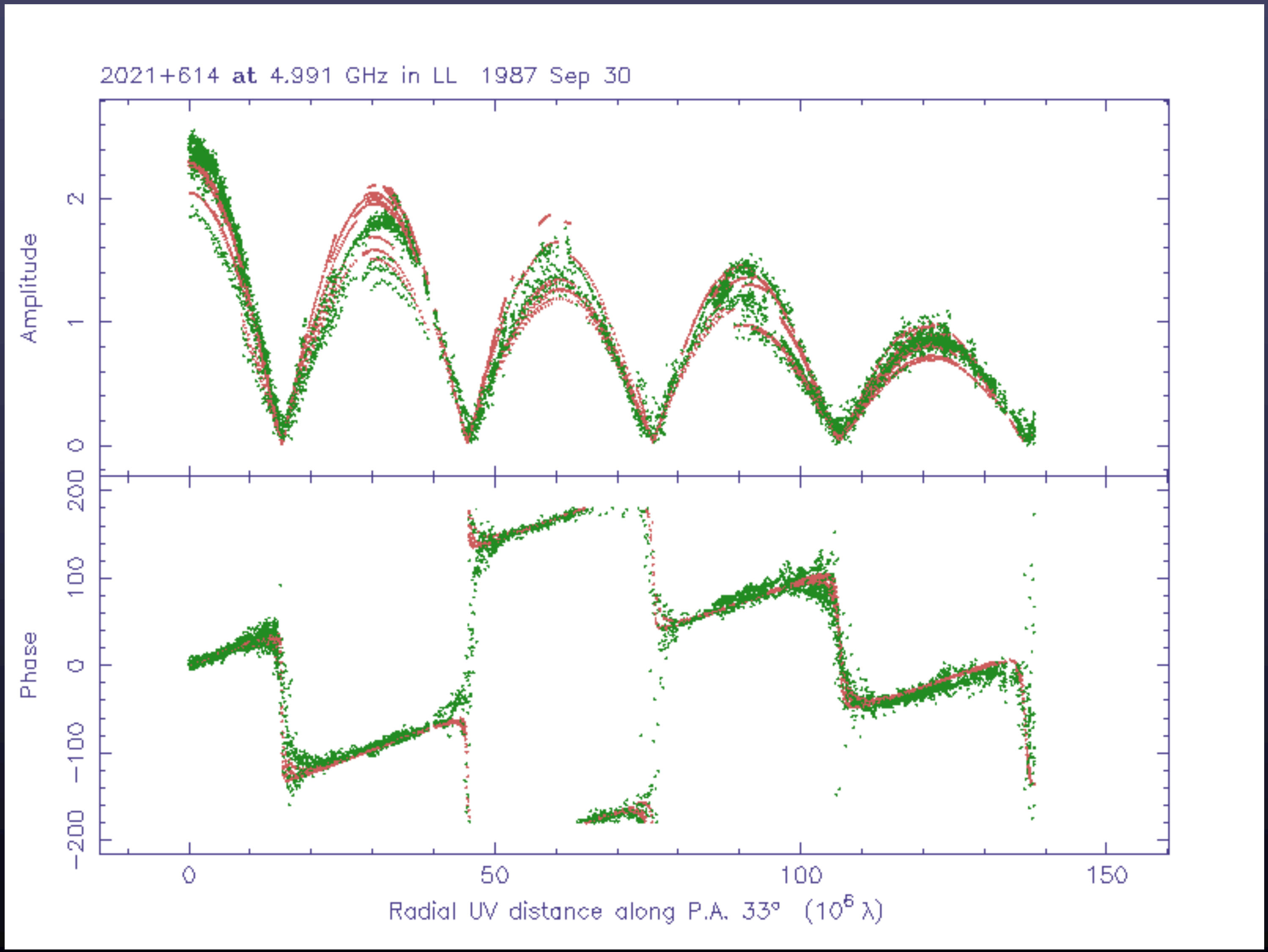


Practical model fitting: 2021

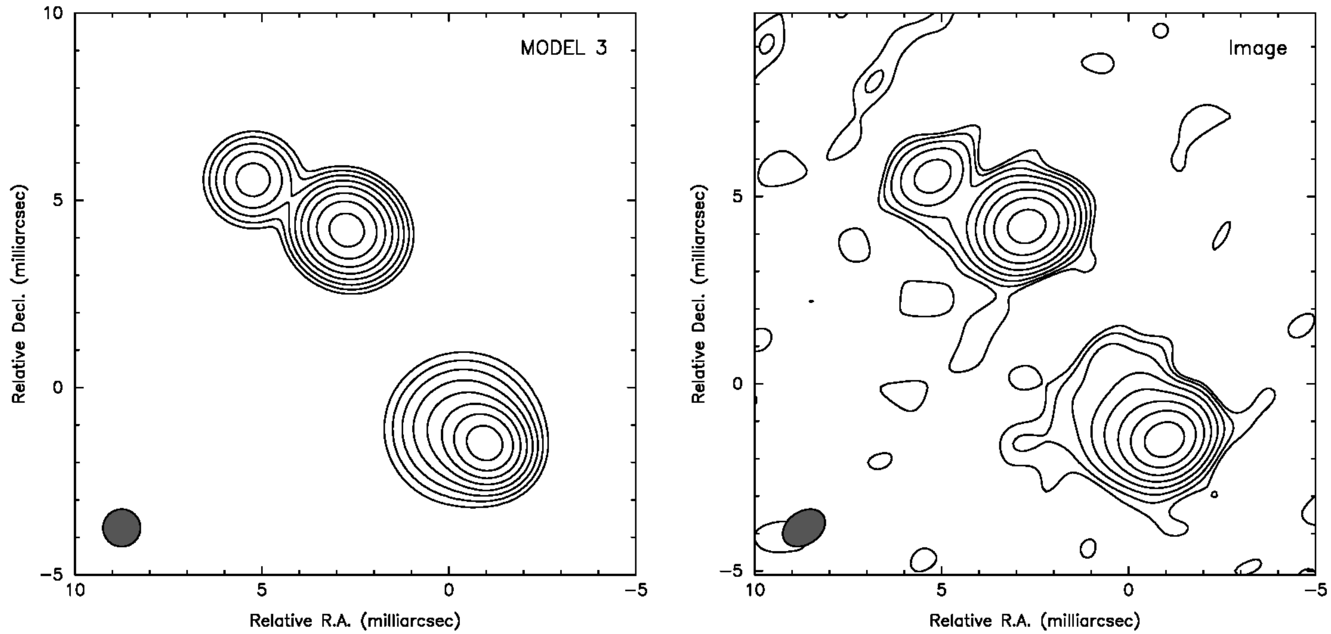


! Flux (Jy)	Radius (mas)	Theta (deg)	Major (mas)	Axial ratio	Phi (deg)	T
1.15566	4.99484	32.9118	0.867594	0.803463	54.4823	1
1.16520	1.79539	-147.037	0.825078	0.742822	45.2283	1

2021: model 2

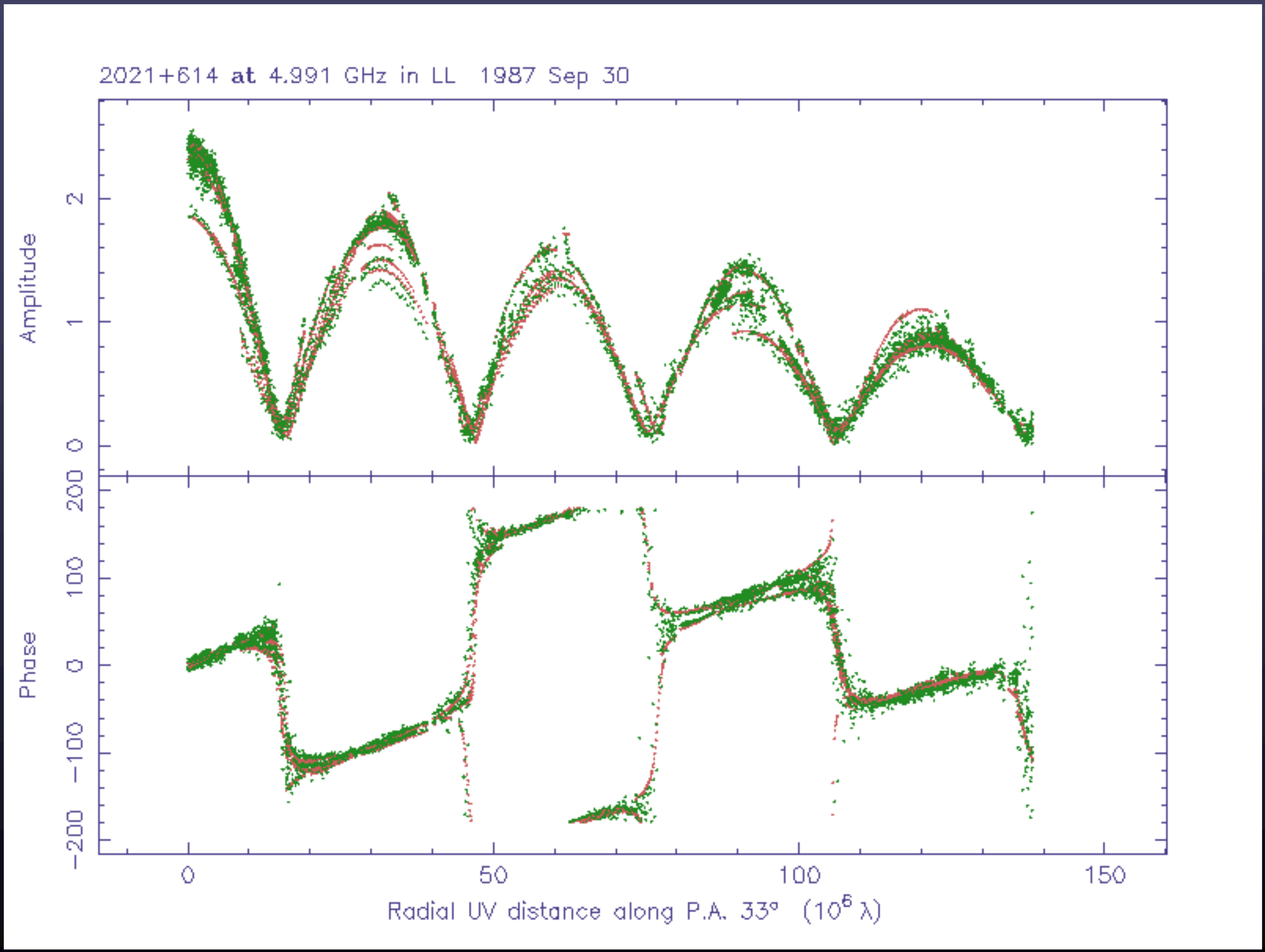

2021+614 at 4.991 GHz in LL 1987 Sep 30
Amplitude
0
1
2
Phase
200
100
0
−100
−200
0
50
100
150
Radial UV distance along P.A. 33° (10^6 λ)

Model fitting 2021



! Flux (Jy)	Radius (mas)	Theta (deg)	Major (mas)	Axial ratio	Phi (deg)	T
1.10808	5.01177	32.9772	0.871643	0.790796	60.4327	1
0.823118	1.80865	-146.615	0.589278	0.585766	53.1916	1
0.131209	7.62679	43.3576	0.741253	0.933106	-82.4635	1
0.419373	1.18399	-160.136	1.62101	0.951732	84.9951	1

2021: model 3


2021+614 at 4.991 GHz in LL 1987 Sep 30
Amplitude
Phase
Radial UV distance along P.A. 33° (10^6 λ)
0
1
2
200
100
0
−100
−200
0
50
100
150

Model fitting

Imaging as an Inverse Problem

- In synthesis imaging, we can solve the **forward problem**: given a sky brightness distribution, and knowing the characteristics of the instrument, we can predict the measurements (visibilities), within the limitations imposed by the noise.
- The **inverse problem** is much harder, given limited data and noise: the solution is rarely unique.
- A general approach to inverse problems is **model fitting**. See, e.g., Press et al., *Numerical Recipes*.
  1. Design a model defined by a number of adjustable parameters.
  2. Solve the forward problem to predict the measurements.
  3. Choose **a figure-of-merit** function, e.g., rms deviation between model predictions and measurements.
  4. Adjust the parameters to **minimize the merit function**.
- Goals:
  1. Best-fit values for the parameters.
  2. A measure of the goodness-of-fit of the optimized model.
  3. Estimates of the uncertainty of the best-fit parameters.

Uses of model fitting

Model fitting is most useful when the brightness distribution is simple.

- Checking amplitude calibration
- Starting point for self-calibration
- Estimating parameters of the model (with error estimates)
- In conjunction with CLEAN or MEM
- In astrometry and geodesy

Programs

- AIPS UVFIT
- Difmap (Martin Shepherd)

Parameters

Example

- Component position: (*x*,*y*) or polar coordinates
- Flux density
- Angular size (e.g., FWHM)
- Axial ratio and orientation (position angle)
  - For a non-circular component
- 6 parameters per component, plus a "shape"

- This is a conventional choice: other choices of parameters may be better!
- (Wavelets; shapelets* [Hermite functions])
  - * Chang & Refregier 2002, ApJ, 570, 447

Limitations of least squares

Assumptions that may be violated

- The model is a good representation of the data
  - Check the fit
- The errors are gaussian
  - True for real and imaginary parts of visibility
  - Not true for amplitudes and phases (except at high SNR)
- The variance of the errors is known
  - Estimate from T_{sys}, rms, etc.
- There are no systematic errors
  - Calibration errors, baseline offsets, etc. must be removed before or during fitting
- The errors are uncorrelated
  - Not true for closure quantities
  - Can be handled with full covariance matrix

Applications: Gravitational Lenses

Gravitational Lenses

- Single source, multiple images formed by intervening galaxy.
- Can be used to map mass distribution in lens.
- Can be used to measure distance of lens and H_0: need redshift of lens and background source, model of mass distribution, and a **time delay**.

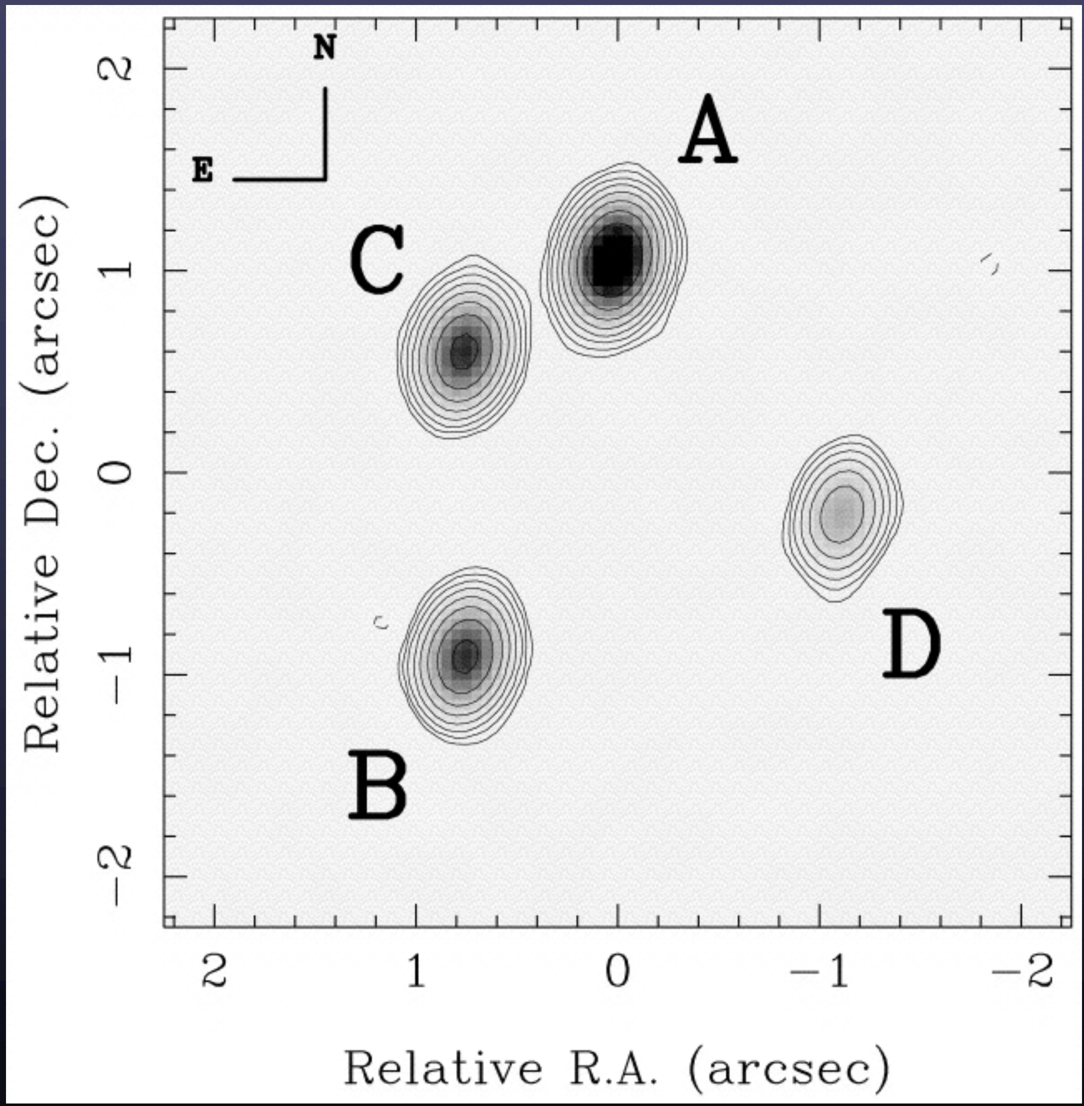
Application of model fitting

- Lens monitoring to measure flux densities of components as a function of time.
- Small number of components, usually point sources.
- Need error estimates.

Example: VLA monitoring of B1608+656 (Fassnacht et al. 1999, ApJ)

- VLA configuration changes: different HA on each day
- Other sources in the field

VLA image of 1608



1608 monitoring results

B – A = 31 days
B – C = 36 days
H_0 = 59 ± 8 km/s/Mpc