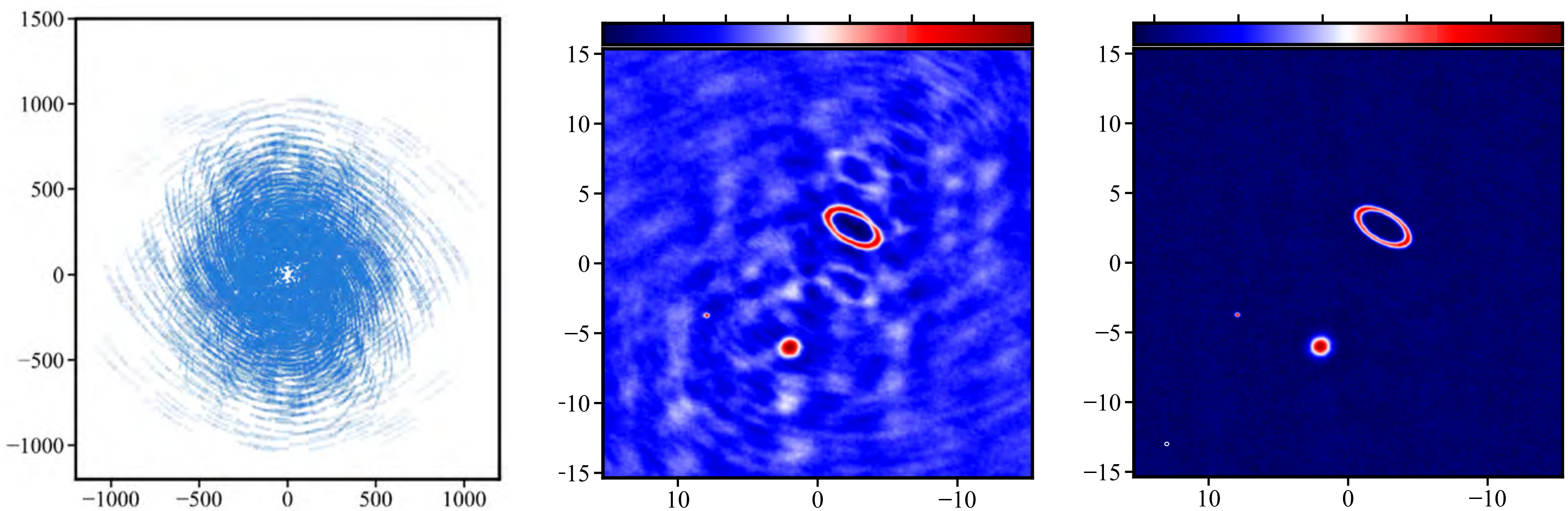
Imaging Basics

David J. Wilner

CENTER FOR ASTROPHYSICS
HARVARD & SMITHSONIAN



Overview

goals

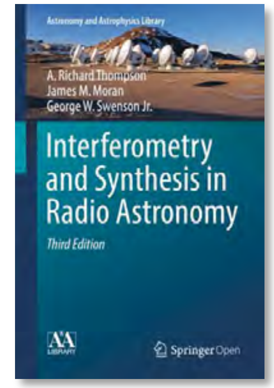
- gain intuition about interferometric imaging
- understand the need for deconvolution

topics

- get comfortable with Fourier Transforms
- review visibility concept and (u,v) plane sampling
- formal description of imaging
- imaging in practice: FFT, gridding, weighting schemes
- deconvolution and the clean algorithm

References

- Thompson, A.R., Moran, J.M. & Swensen, G.W. “Interferometry and Synthesis Imaging in Radio Astronomy” 3rd edition, 2017, Springer Open



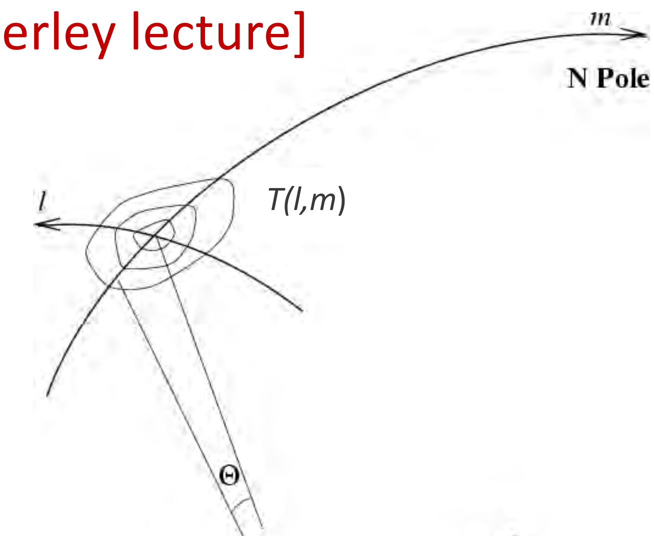
- NRAO Synthesis Workshop proceedings
  - Synthesis Imaging in Radio Astronomy II,, ASP Conference Series, Vol. 180, 1999,
    eds. Taylor, G.B., Carilli, C.L., Perley, R.A.
  - lecture slides: www.aoc.nrao.edu/events/synthesis

- IRAM 2000 Interferometry School proceedings
  - www.iram.fr/IRAMFR/IS/IS2008/archive.html
- Condon, J.J. & Ransom, S.M. 2016 “Essential Radio Astronomy”
  - science.nrao.edu/opportunities/courses/era
- Many useful pedagogical presentations available on-line
  - ALMA videos, casadocs, CASS radio school, ERIS lectures, …
- EHT M87 imaging paper, 2019, ApJ Letters, 875, L4

Visibility and Sky Brightness

For small field of view, far field, quasi-monochromatic, incoherent, etc.:

*The complex visibility function V(u,v) is the **2D Fourier Transform** of the sky brightness distribution T(l,m)* [Rick Perley lecture]

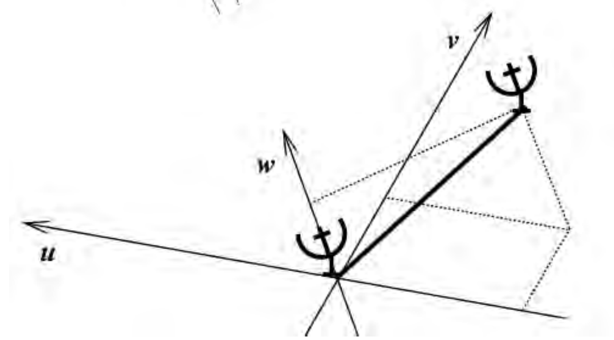


$$V(u,v) = \int\int T(l,m) e^{-i2\pi(ul+vm)} dl dm$$

(u, v) are (E-W, N-S) spatial frequencies [wavelengths]

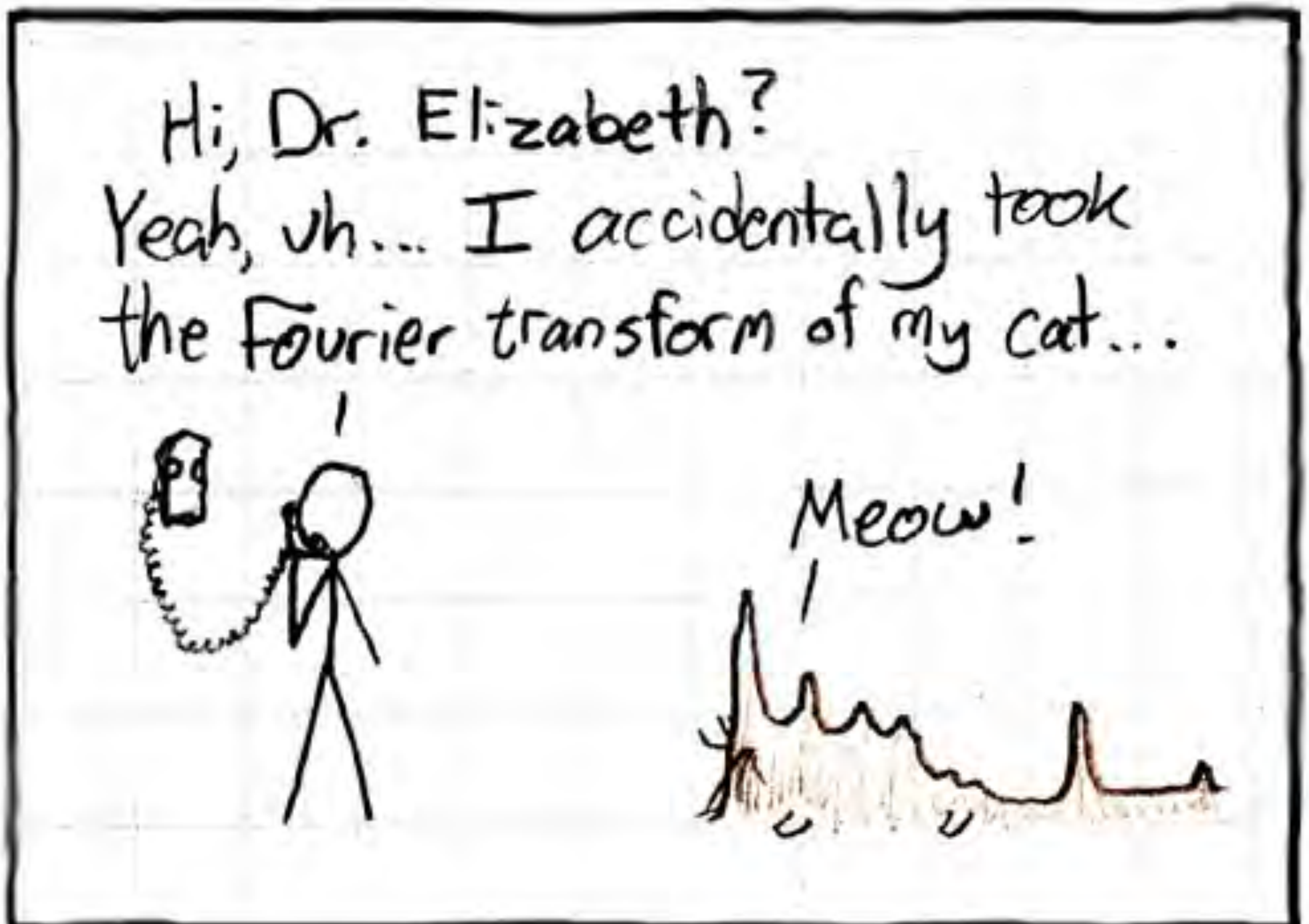
(l, m) are (E-W, N-S) angles in the tangent plane [radians]

recall $e^{ix} = \cos x + i \sin x$



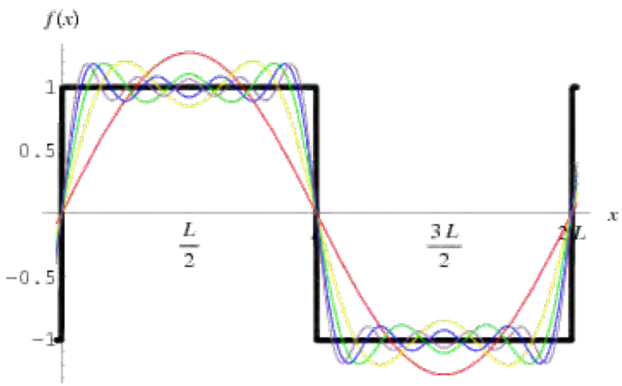
Jean Baptiste Joseph Fourier 1768-1830

xkcd.com/26/


Hi, Dr. Elizabeth?
Yeah, uh... I accidentally took the Fourier transform of my cat...
Meow!

Fourier Fundamentals

- acquire comfort with the Fourier domain…
  - functions and their Fourier transforms occupy *upper* and *lower* domains, as if “functions circulated at ground level and their transforms in the underworld” (Bracewell 1965)
- Fourier theory states that any well behaved signal can be represented as the sum of sinusoids



$$f(x) = \frac{4}{\pi} \sum_{n=1,3,5,\ldots}^{\infty} \frac{1}{n} \sin\left(\frac{n\pi x}{L}\right).$$

Wolfram Mathworld

- the Fourier transform is the mathematical tool that decomposes a signal into its sinusoidal components (frequency, amplitude, phase)
- the Fourier transform contains *all* information of the original signal

Useful Fourier Transform Properties

- **addition** in one domain is **addition** in the other

$$g(x) + h(x) = G(s) + H(s)$$

- **multiplication** in one domain is **convolution** in the other

$$g(x) = h(x) * k(x) \qquad G(s) = H(s)K(s)$$

- **"large"** in one domain is **"small"** in the other

$$g(\alpha x) = \alpha^{-1} G(s/\alpha)$$

- an **offset** in one domain is a **phase shift** in the other

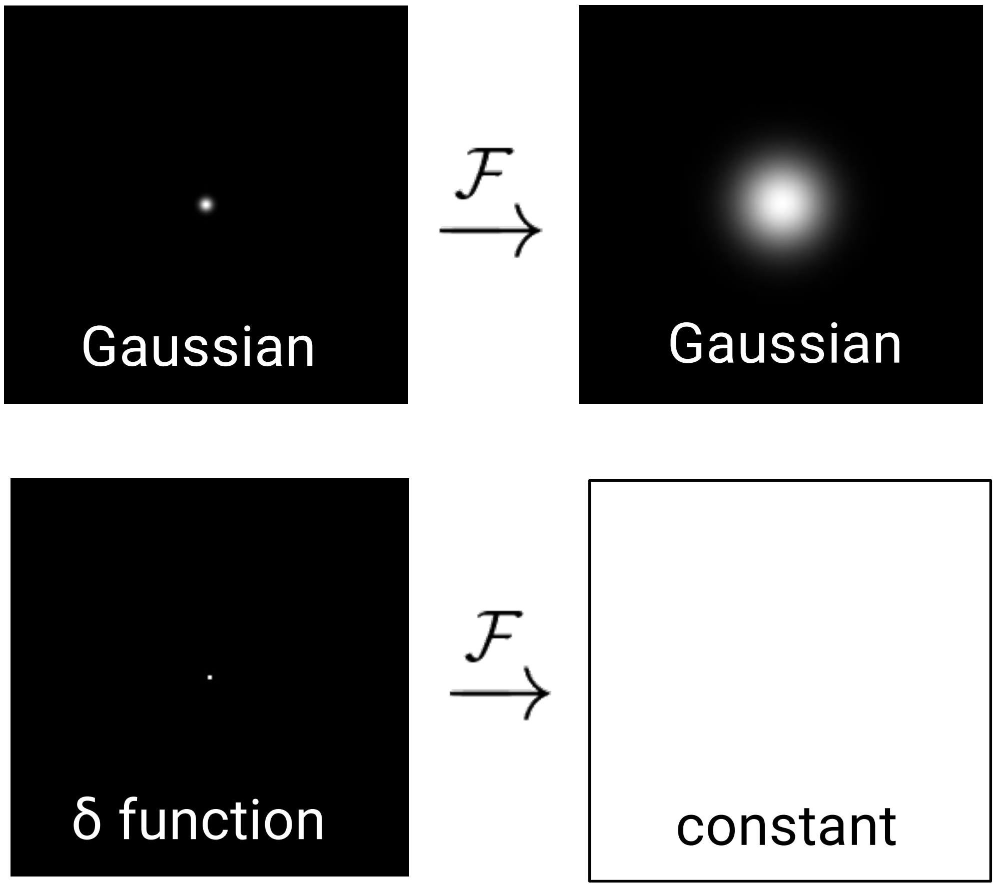
$$g(x - x_0) = G(s) e^{i2\pi x_0 s}$$

- the Nyquist-Shannon **sampling theorem**

$$g(x) \subset \Theta \text{ completely determined if } G(s) \text{ sampled at } \leq 1/\Theta$$

Some 2D Fourier Transform Pairs

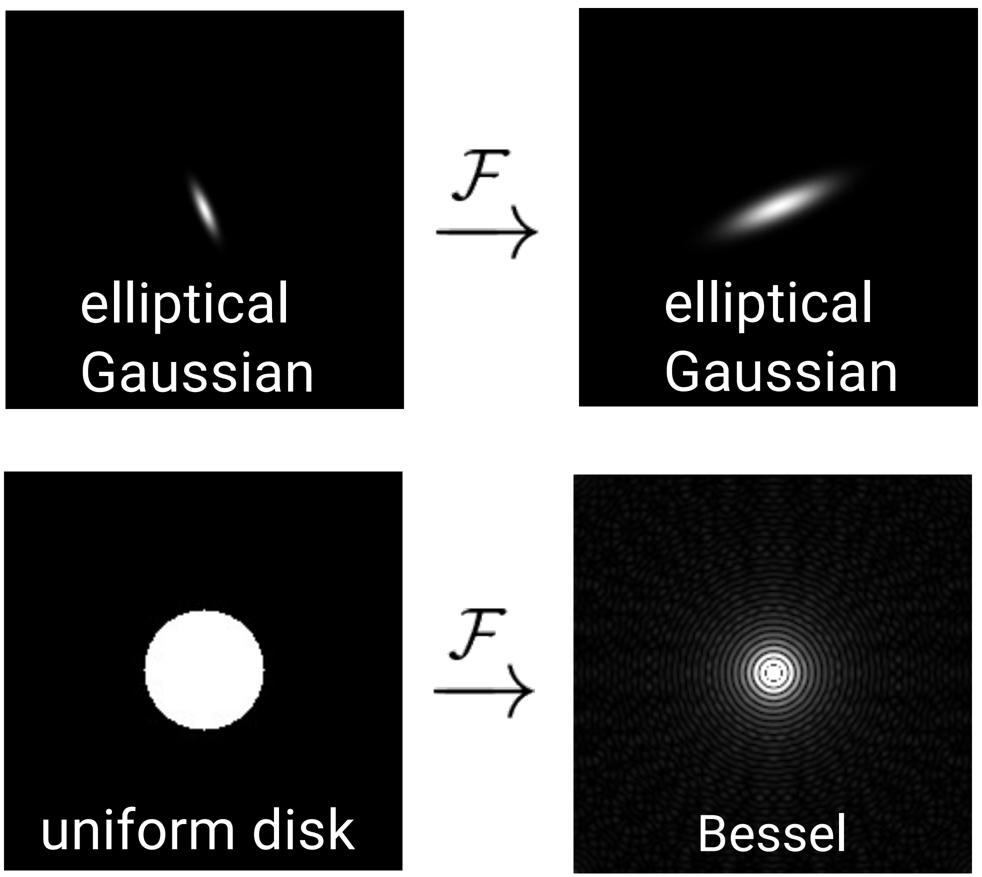
T(l,m) *V(u,v) amplitude*



narrow features transform into wide features (and vice-versa)

Some 2D Fourier Transform Pairs

T(l,m) *V(u,v) amplitude*

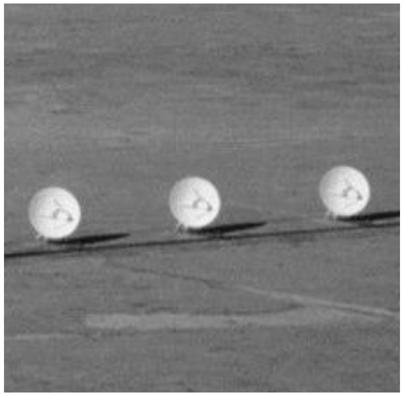


narrow features transform into wide features (and vice-versa)

sharp edges result in many high special frequencies

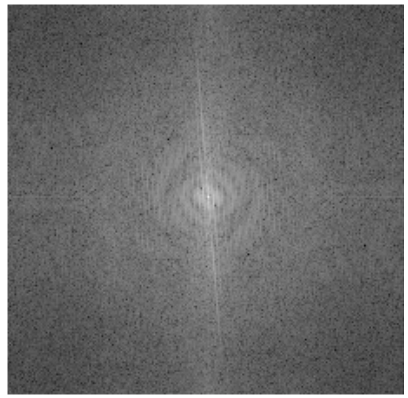
Amplitude and Phase

T(l,m)



$\xrightarrow{\mathcal{F}}$

V(u,v) amplitude



amplitude tells “how much” of a certain spatial frequency

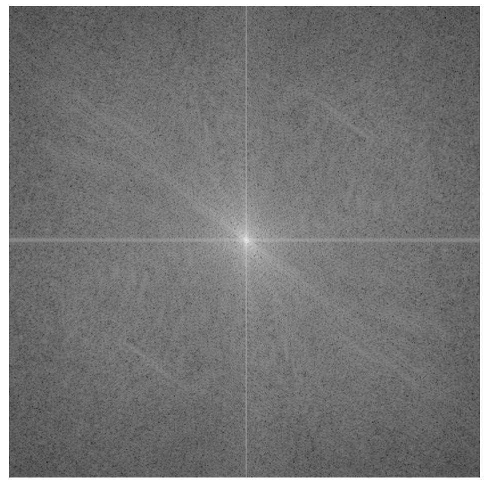
phase tells “where” the spatial frequency is located

A more complicated example

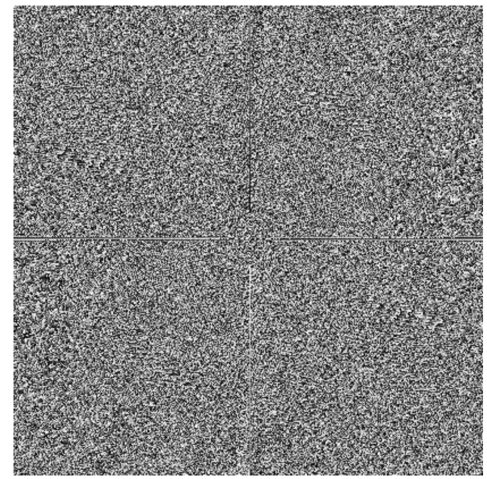
T(l,m)

$\xrightarrow{\mathcal{F}}$

V(u,v) amplitude



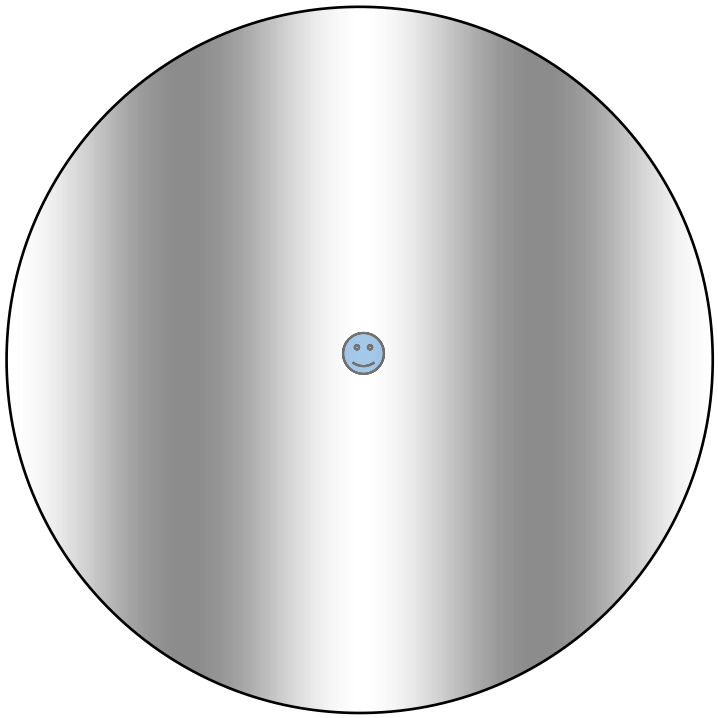
V(u,v) phase

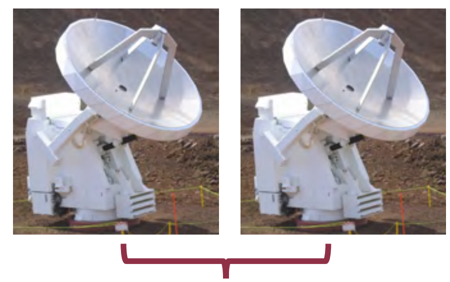


$$V(u,v) = \int\int T(l,m)e^{-i2\pi(ul+vm)}dldm$$

- ***T(l,m)* is real**: *V(-u,-v) = V*(u,v)   where * = complex conjugate*
  - get two visibilities for one measurement
- ***V(u=0,v=0)*** is the integral of *T(l,m)dldm* = total flux density

$$V(u,v) = \int\int T(l,m)e^{-i2\pi(ul+vm)}dldm$$



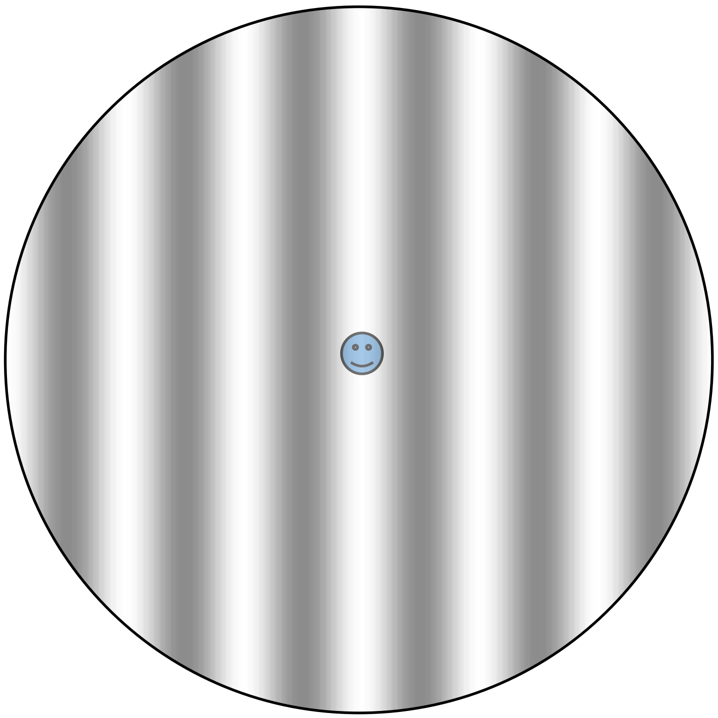


short baseline

wide fringe pattern

compact source

$$V(u,v) = \int\int T(l,m) e^{-i2\pi(ul+vm)} dl dm$$

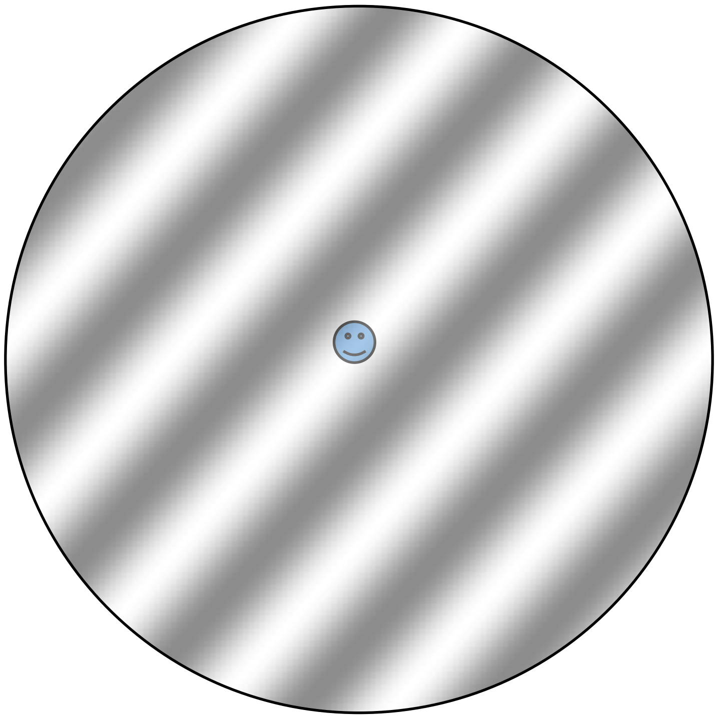


compact source

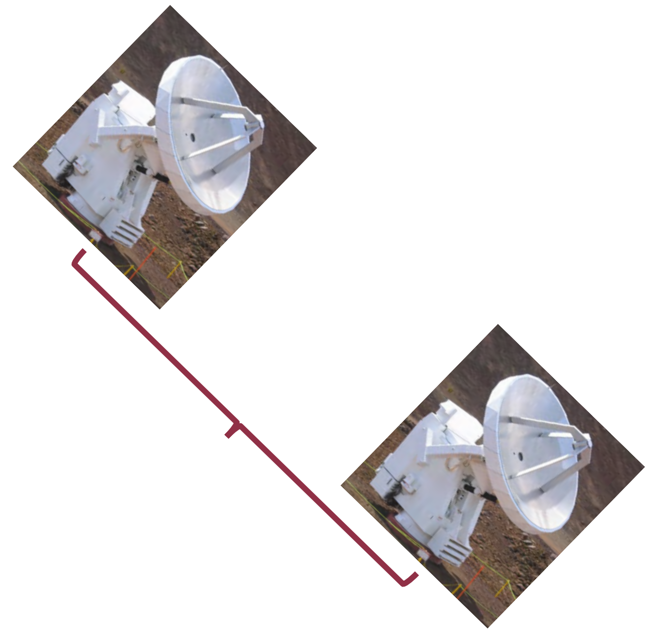
long baseline

narrow fringe pattern

$$V(u,v) = \int\int T(l,m)e^{-i2\pi(ul+vm)}dldm$$



compact source

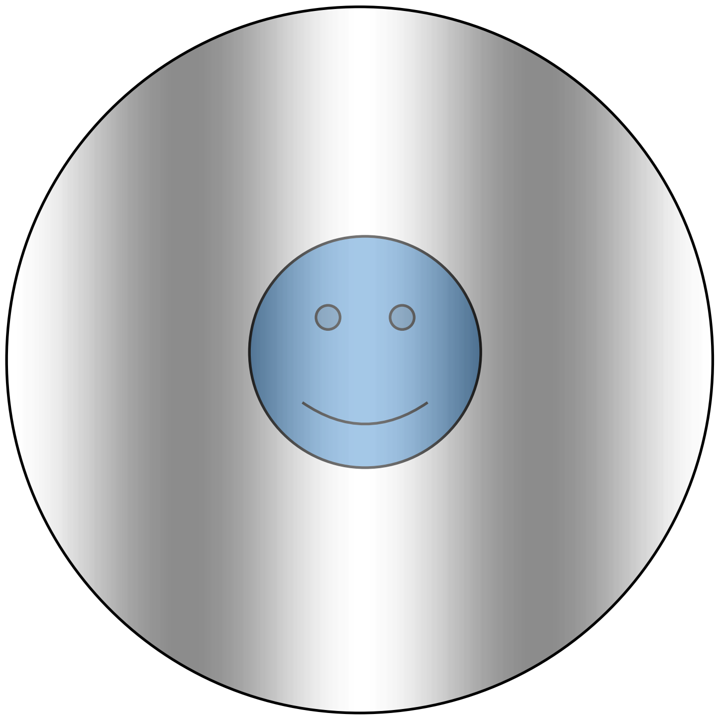


long baseline

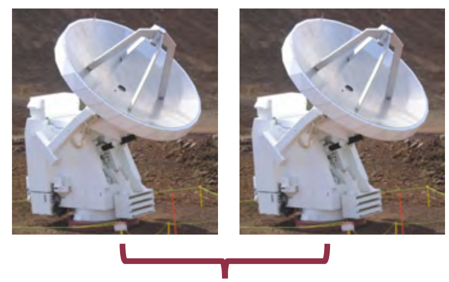
narrow fringe pattern

different orientation

$$V(u,v) = \int\int T(l,m)e^{-i2\pi(ul+vm)}dldm$$



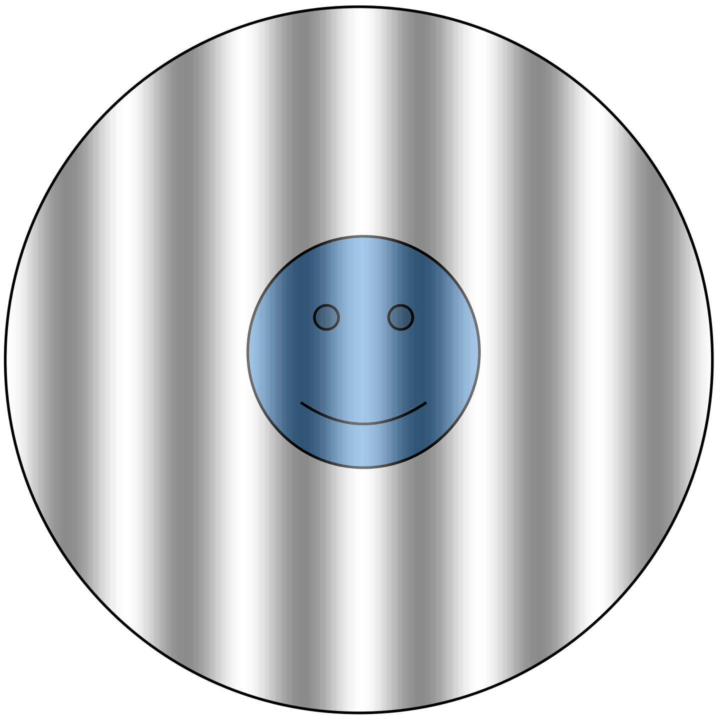
extended source



short baseline

wide fringe pattern

$$V(u,v) = \int\int T(l,m)e^{-i2\pi(ul+vm)}dldm$$

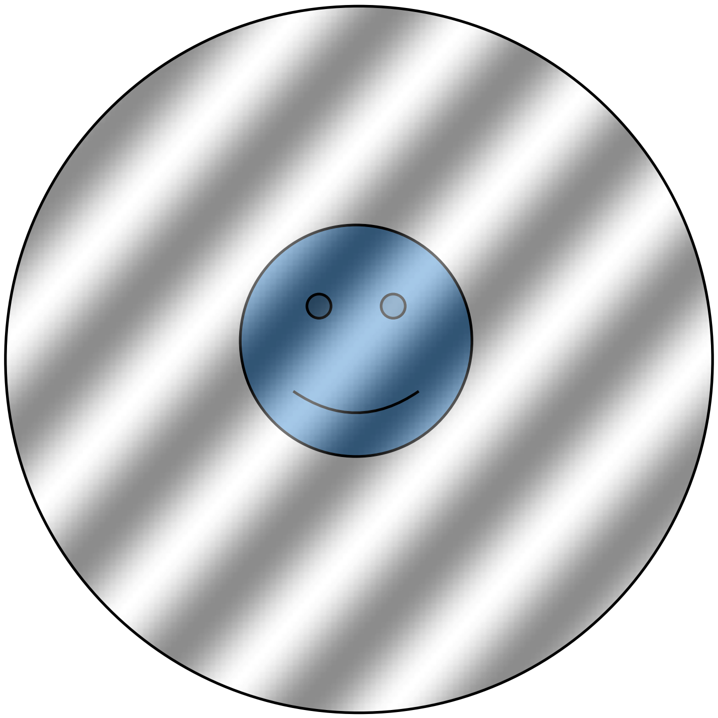


extended source

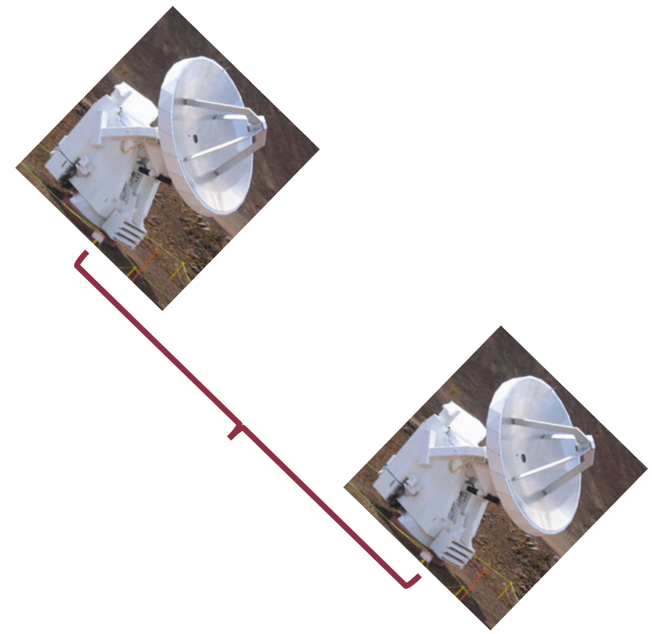
long baseline

narrow fringe pattern

$$V(u,v) = \int\int T(l,m)e^{-i2\pi(ul+vm)}dldm$$



extended source



long baseline

narrow fringe pattern

different orientation

Aperture Synthesis

basic idea: sample *V(u,v)* at enough *(u,v)* points using distributed small apertures to synthesize a large aperture of size (u_{max}, v_{max})

use **more antennas** for more samples

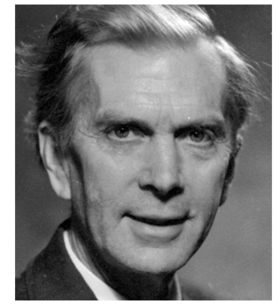
- one pair of antennas = one baseline
  = two *(u,v)* samples at one time
- N antennas = N(N-1) samples at one time
- reconfigure physical layout of N antennas for more samples (as long as source structure doesn't change with time)

use **Earth rotation** for more samples

- baseline length/orientation relative to sky change with time

use **more wavelengths** for more (continuum) samples

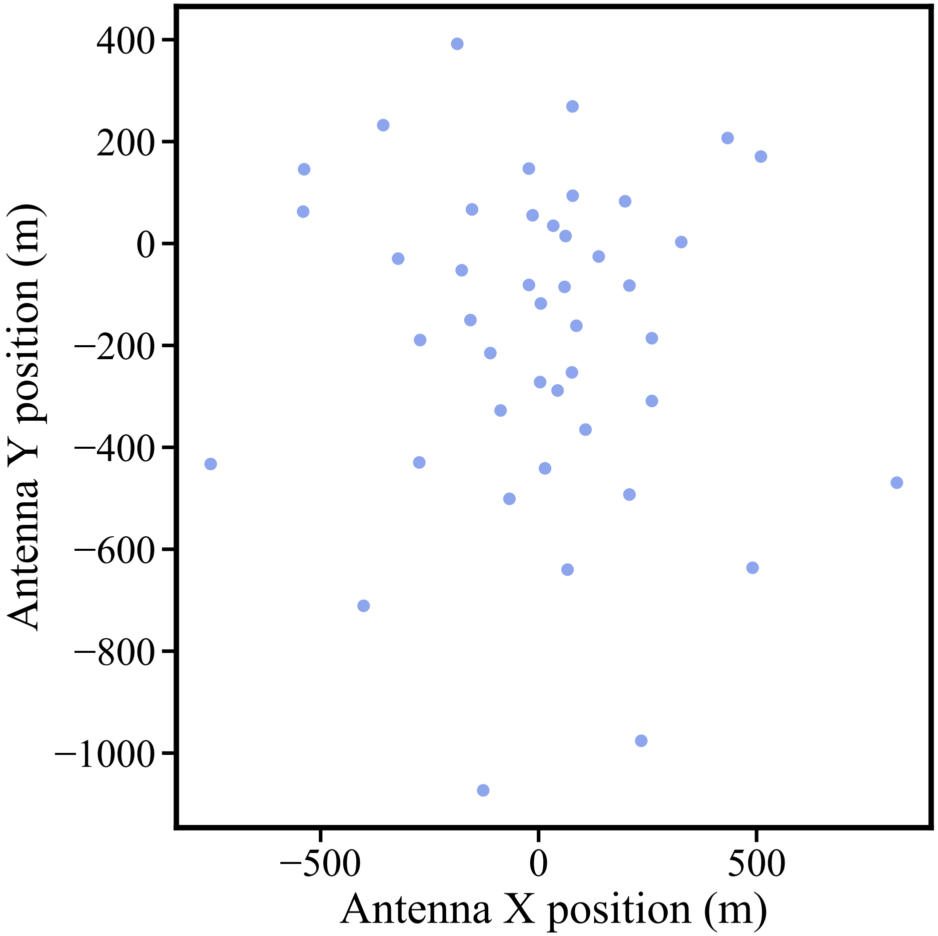
- *u* and *v* are measured in *wavelengths*
- "multi-frequency synthesis": determine structure at a fiducial wavelength together with change with wavelength



Sir Martin Ryle 1918-1984

1974 Nobel Prize in Physics

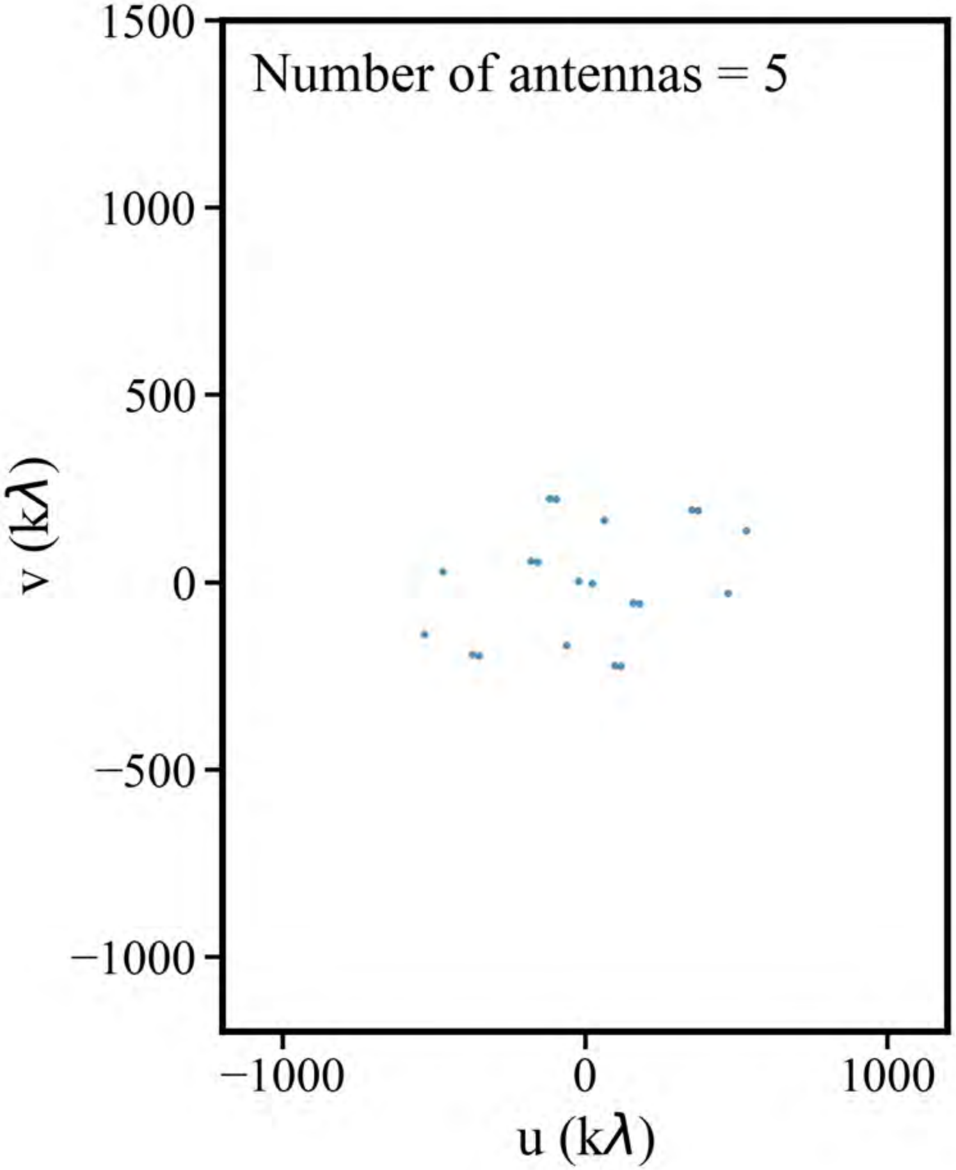
Example of *(u,v)* Plane Sampling

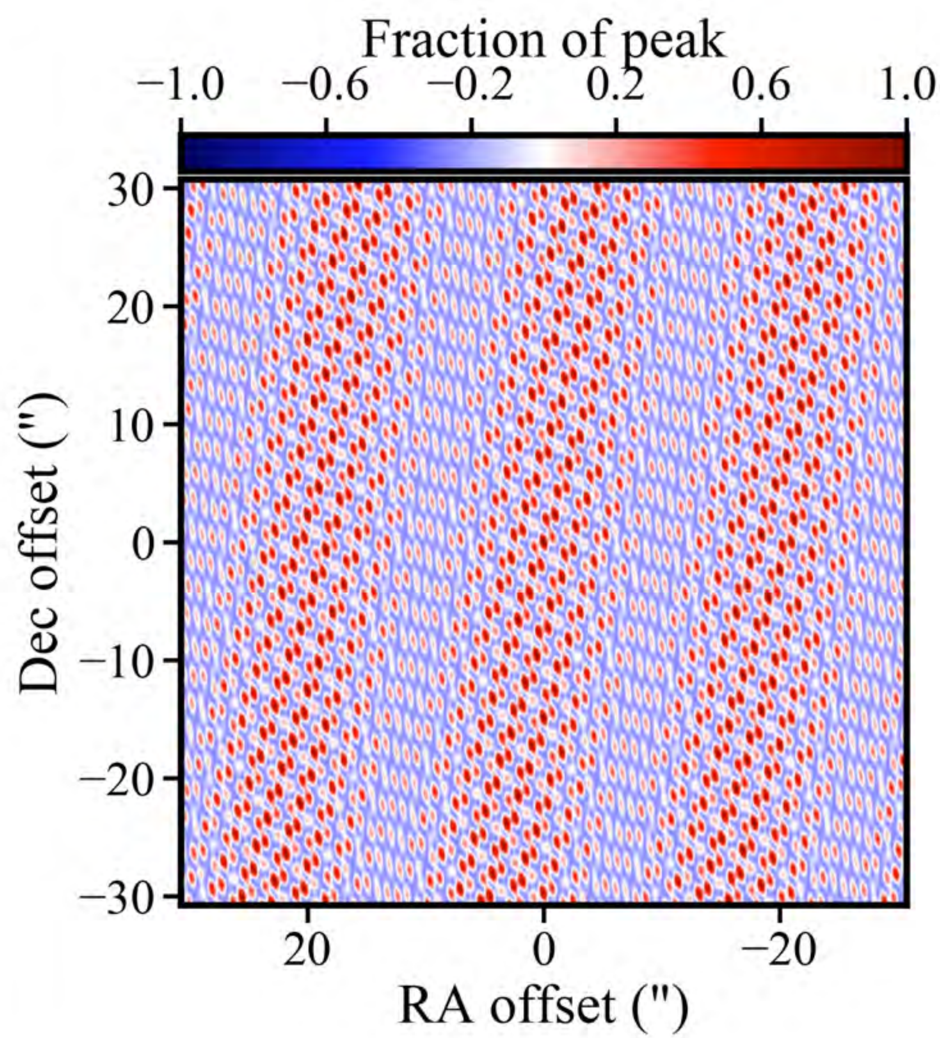


ALMA 12m antenna locations on Aug 8, 2015

[array design: Eric Murphy lecture]

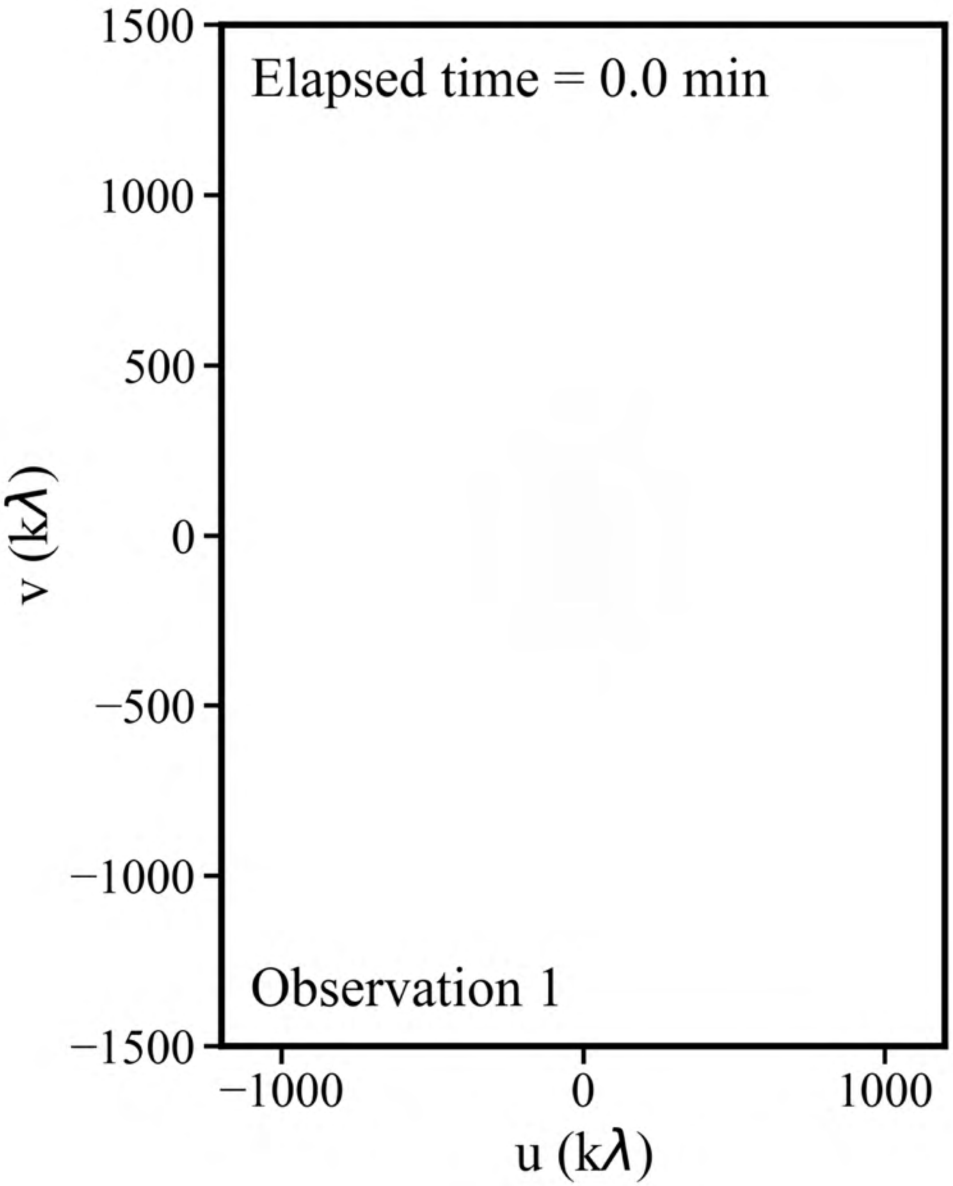
(u,v) Plane Sampling: more Antennas

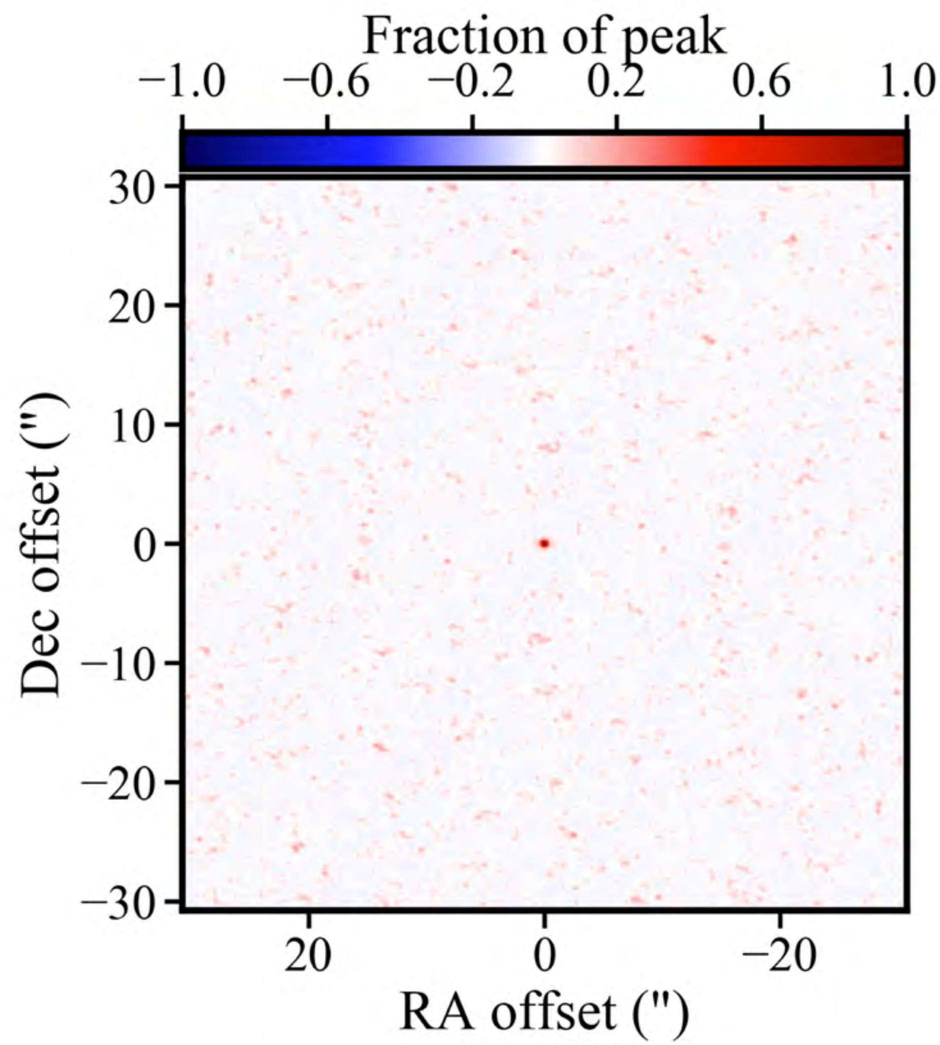




source dec -51 deg
λ = 1.3 mm

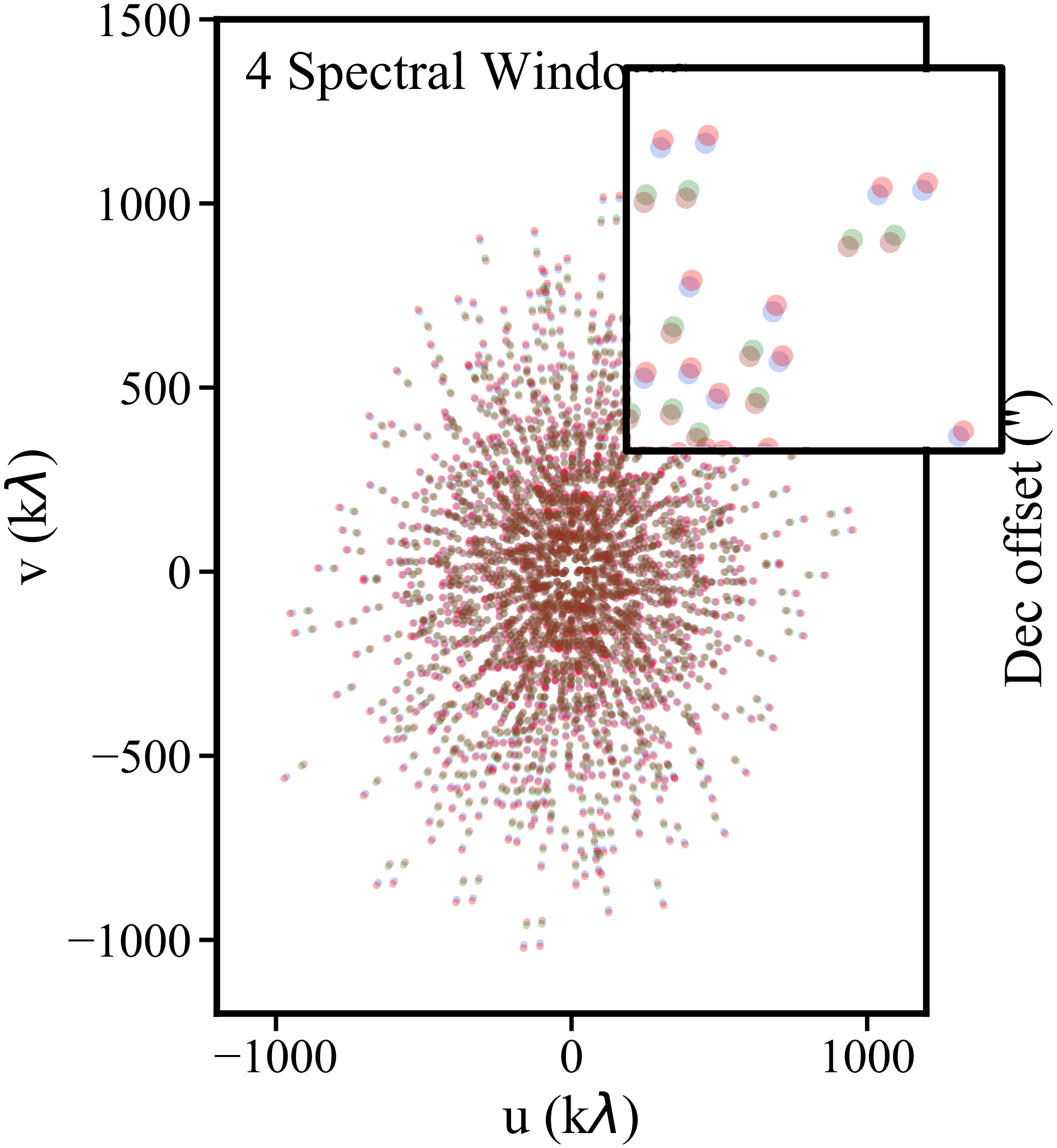
(u,v) Plane Sampling: Earth Rotation

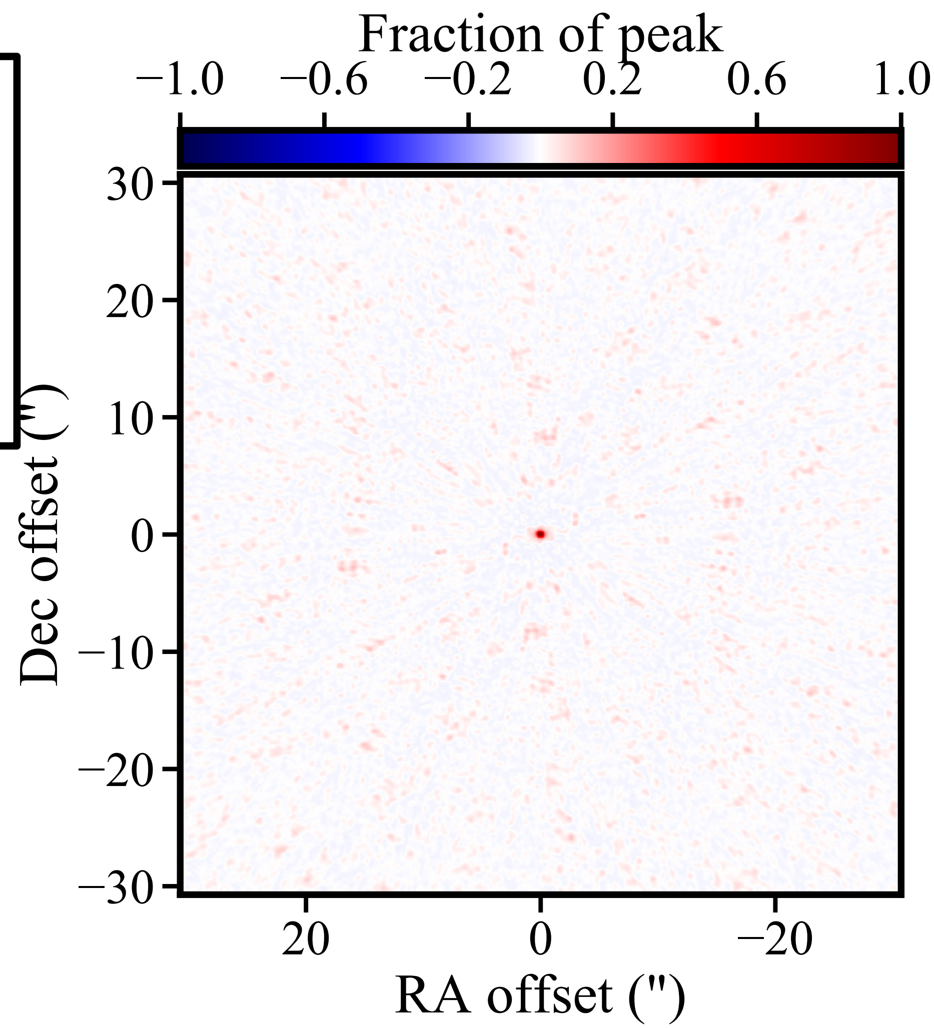




source dec -51 deg
λ = 1.3 mm

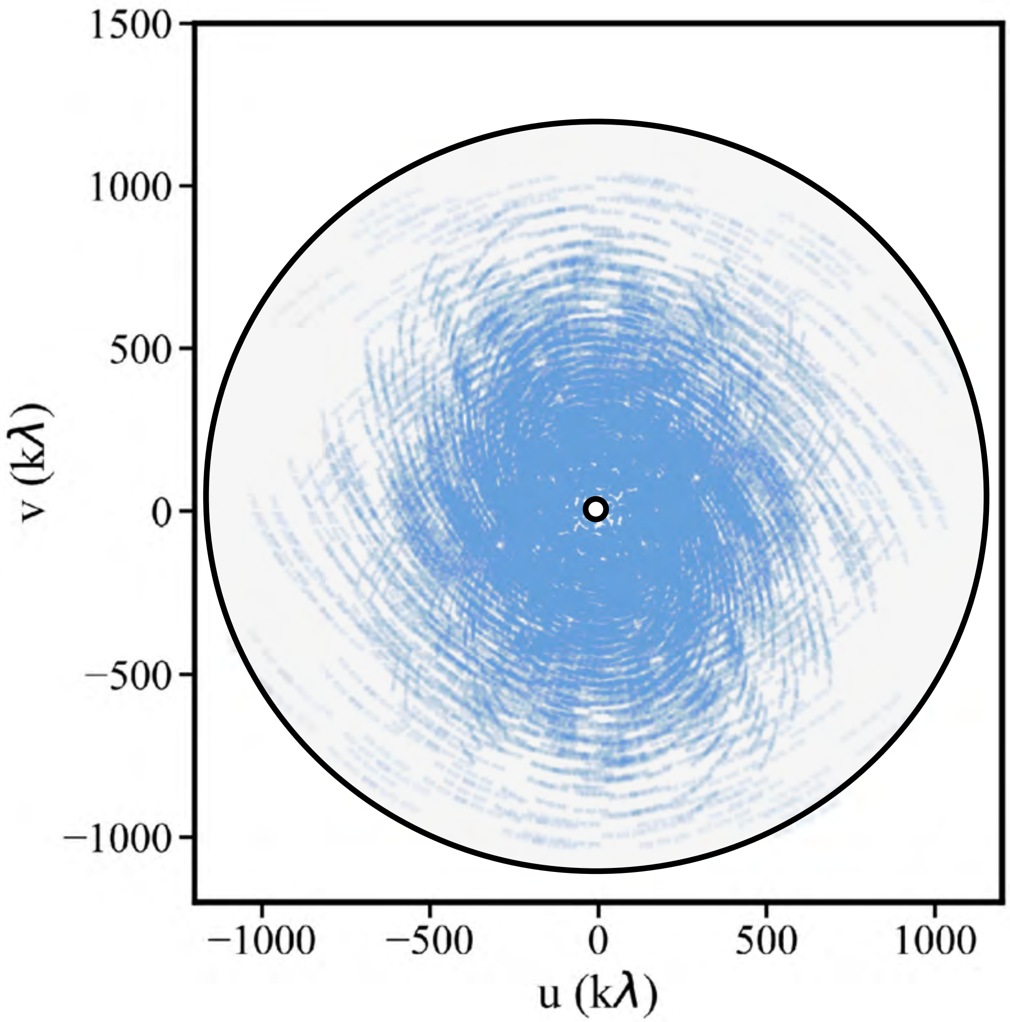
(u,v) Plane Sampling: more wavelengths





Implications of *(u,v)* Plane Sampling

samples of *V(u,v)* are limited by array and Earth-sky geometry



outer boundary

- no info on smaller scales
- resolution limit

inner boundary

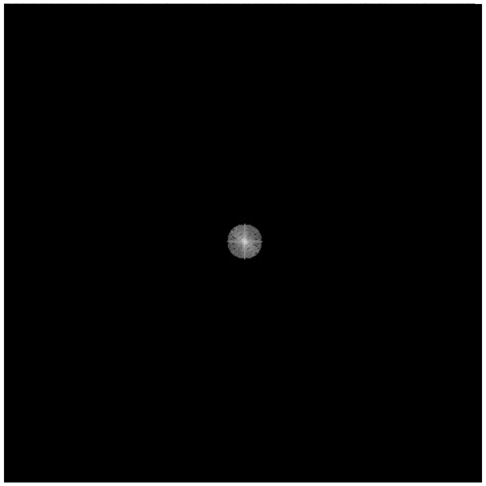
- no info on larger scales
- extended sources invisible

irregular coverage in between

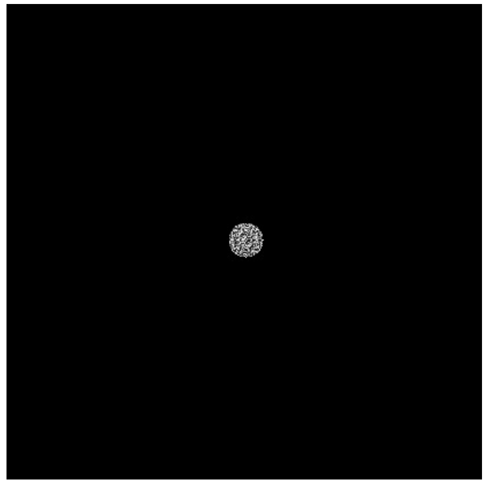
- sampling theorem violated
- information missing

Inner and Outer *(u,v)* Boundaries

V(u,v) amplitude



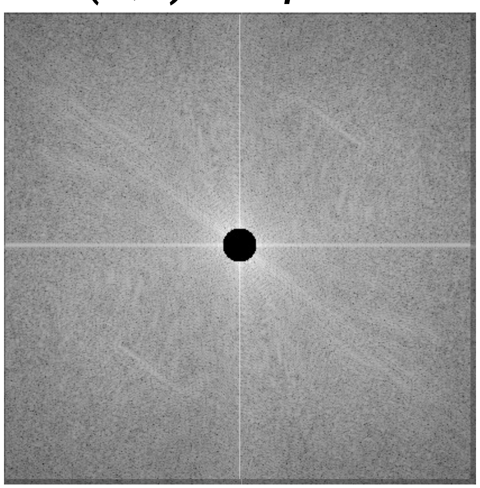
V(u,v) phase



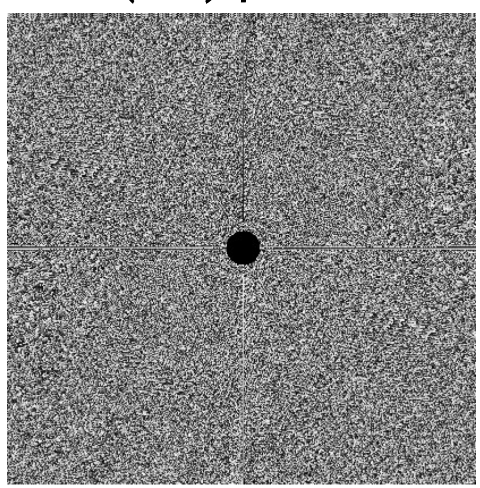
$\xrightarrow{\mathcal{F}}$

T(l,m)

V(u,v) amplitude

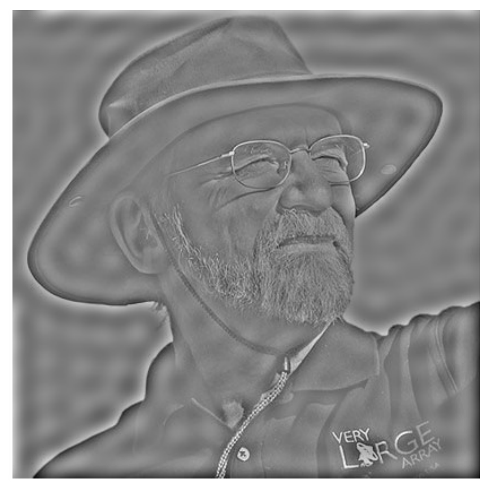


V(u,v) phase



$\xrightarrow{\mathcal{F}}$

T(l,m)



Calibrated Visibilities… Now what?

analyze *V(u,v)* samples directly by model fitting [

- good for simple structures, e.g. point sources, symmetric rings
- for a purely statistical description of sky brightness, e.g. fluctuations
- visibilities have well defined noise properties

[Dom Pesce lecture]

recover an image from incomplete, noisy samples of *V(u,v)*

- Fourier transform *V(u,v)* to create a distorted $T^D(l,m)$
- account for incomplete sampling to create a model of *T(l,m)*
- work with the model of *T(l,m)* to do science

CASA
`tclean`

Formal Description of Imaging

- sample Fourier domain at discrete points

$$S(u,v) = \sum_{k=1}^{M} \delta(u - u_k, v - v_k)$$

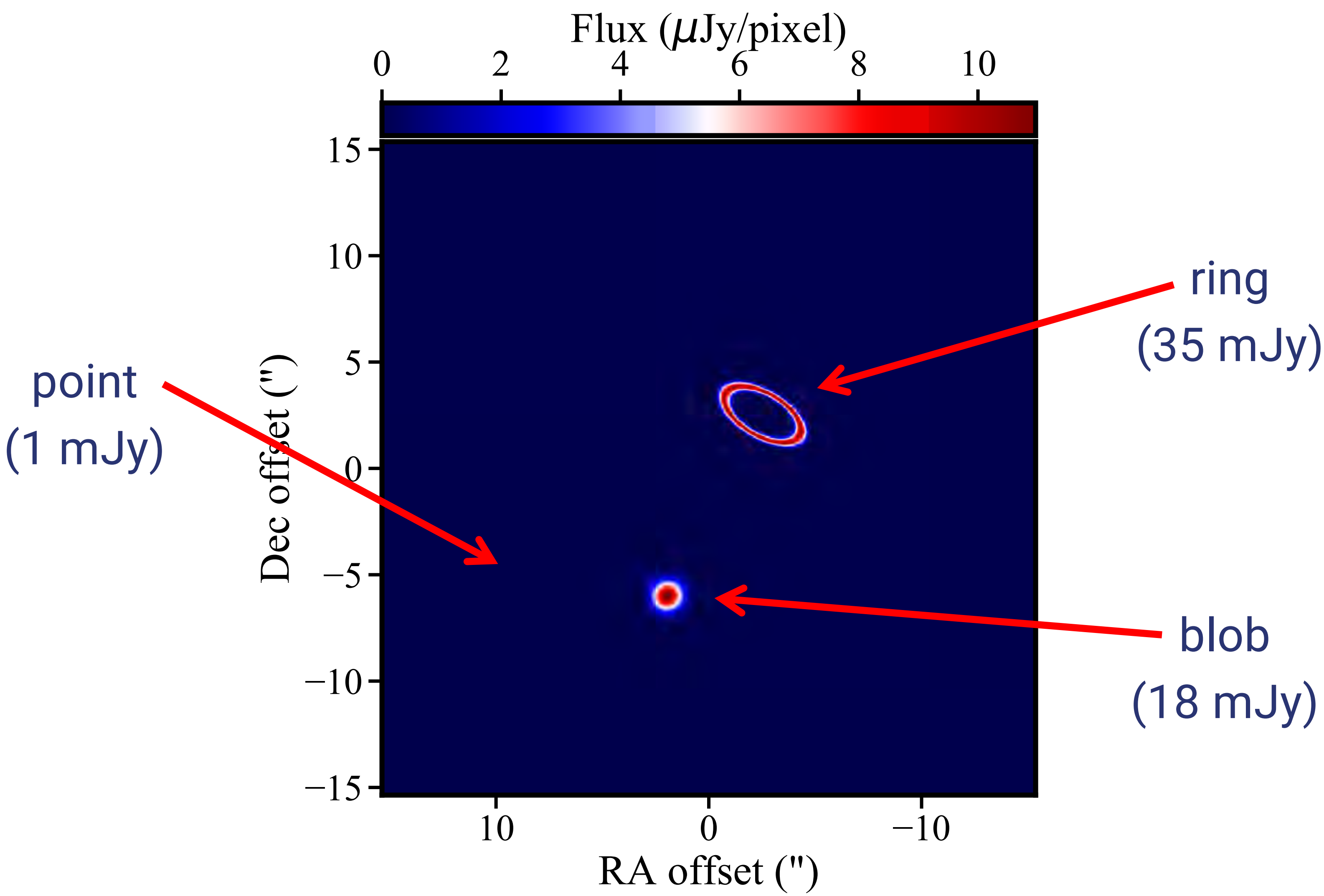
- Fourier transform sampled visibilities $V(u,v)S(u,v) \xrightarrow{\mathcal{F}} T^D(l,m)$

- apply the convolution theorem $T(l,m) * s(l,m) = T^D(l,m)$

  where the Fourier transform of the sampling pattern $s(l,m) \xrightarrow{\mathcal{F}} S(u,v)$
  is the “point spread function” or the “synthesized beam”

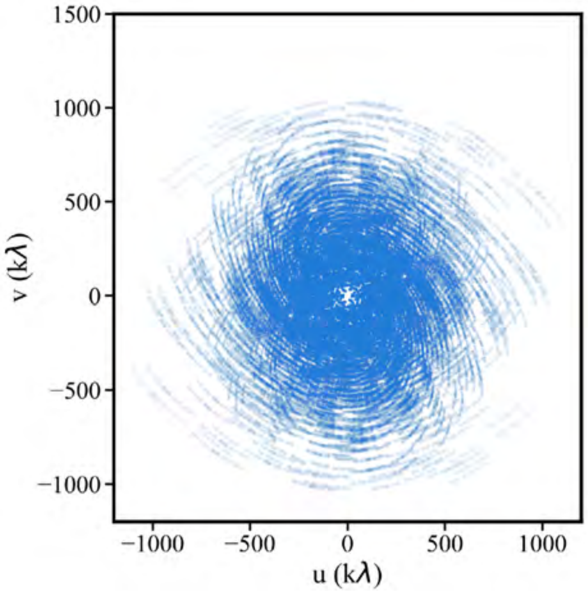
- radio astronomy jargon:
  the “dirty image” is the true image convolved with the “dirty beam”

Example model sky brightness *T(l,m)*

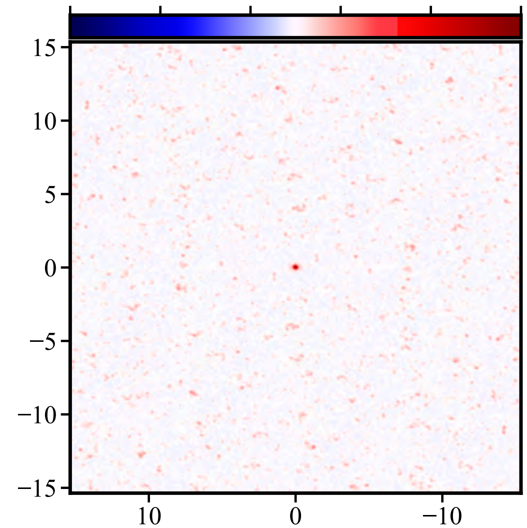


Dirty Beam and Dirty Image

S(u,v)



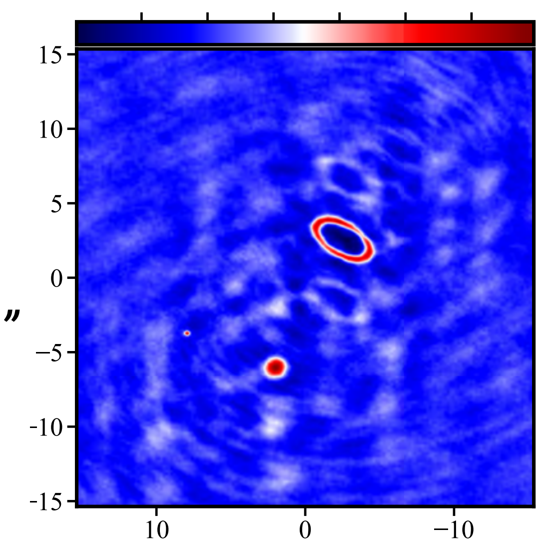
$\xrightarrow{\mathcal{F}}$



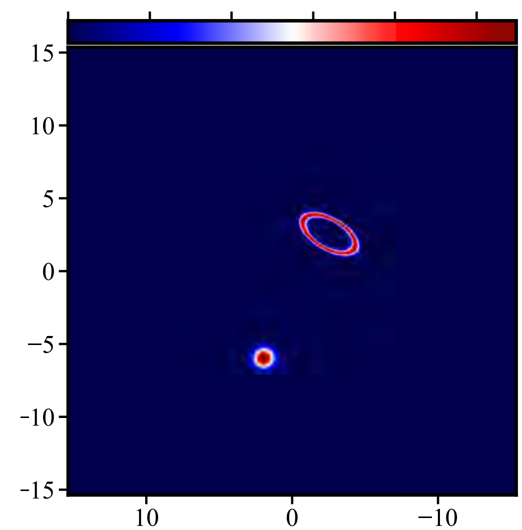
s(l,m)
'dirty beam"

$\ast$

T^D(l,m)
"dirty image"



=



T(l,m)

Image Size and Pixel Size

image size

- antenna response modifies the sky brightness distribution

  T(l,m) → T(l,m)A(l,m)

  where A(l,m) is antenna "primary beam" (fwhm ~ λ/D) [Rob Selena lecture]

- natural choice for image size is often full extent of primary beam

  e.g. ALMA 12 m antennas at 1.3 mm → image size 2 x 27 arcsec

- "primary beam correction"

  divide final image by *A(l,m)* to account for *A(l,m)* attenuation;

  raises noise away from the center

pixel size

- satisfy sampling theorem for longest baselines: $\Delta l < \frac{1}{2u_{max}}$ $\Delta m < \frac{1}{2v_{max}}$
- in practice, 3 to 5 pixels across dirty beam main lobe to aid deconvolution

  e.g. ALMA at 1.3 mm, baselines to 1 km → pixel size < 0.05 arcsec

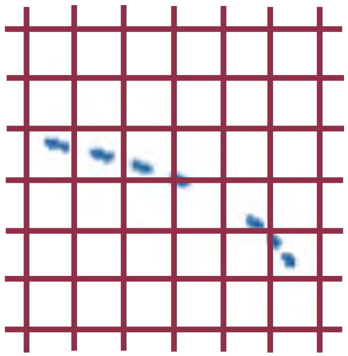
"Fourier Transform"

- direct computation of Fourier integral by summation is slow, generally not practical for modern imaging applications
- FFT = Fast Fourier Transform algorithm
- trace to Gauss' work in 1805 to interpolate orbit of asteroids Pallas and Juno; (re)discovered by Cooley & Tukey 1965
- $O(N^2 \log N)$ vs. $O(N^4)$ for N^2 image cells
- optimized codes readily available
- **FFT requires data on a regularly spaced grid**
- but aperture synthesis does not provide $V(u,v)$ samples on a regularly spaced grid, so…

Gridding

Gridding is used to resample *V(u,v)* onto a regular *(u,v)* grid for FFT

- conventional approach is to use convolution
- *(u,v)* cell size ≈ 0.5 x D, where D = antenna diameter



$$V^G(u,v) = V(u,v)S(u,v) * G(u,v)$$
$$\xrightarrow{F} T^D(l,m)g(l,m)$$

- other gridding steps may include
  - functions to apply primary beam weighting and offsets ("mosaicking")
  - functions to apply wide-field phase shifts ("W projection")
  - functions to correct for primary beam differences ("A projection")

[Brian Mason, Urvashi Rao, Preshanth Jagannathan lectures]

Visibility Weighting Schemes

Introduce weighting function *W(u,v)* into the gridding process

- W(u,v) modifies the sampling function: S(u,v) → S(u,v)W(u,v)
- changes the dirty beam shape, s(l,m)

Natural weighting: W(u,v) = $1/\sigma^2(u,v)$ where σ^2 is noise variance

- best point source sensitivity, lower resolution, more dirty beam structure

Uniform weighting: W(u,v) = $1/\delta(u,v)$ where δ is density of *(u,v)* points

- higher noise, higher resolution, less dirty beam structure

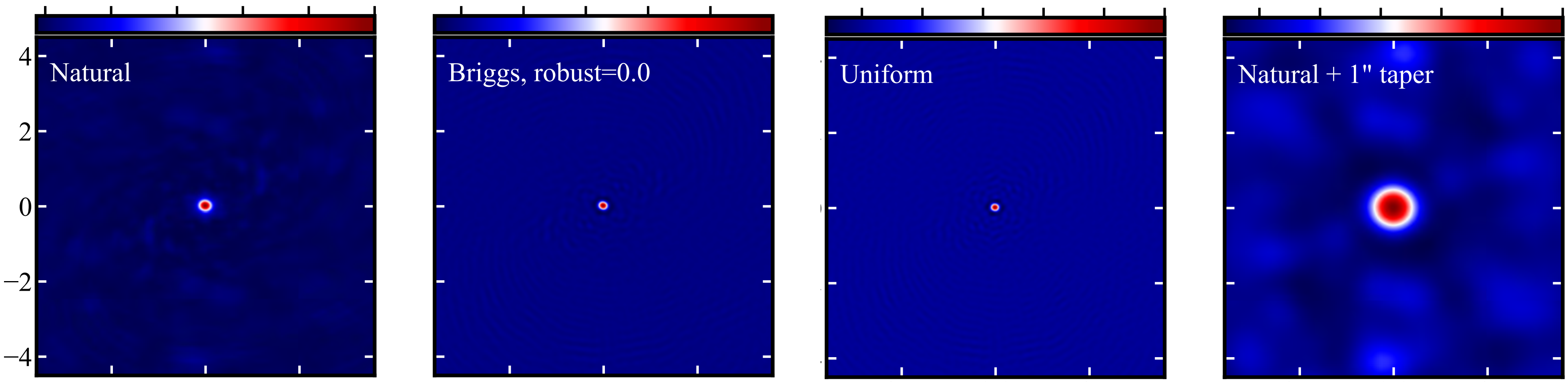
Robust ("Briggs") weighting: W(u,v) combines noise and density

- adjustable parameter allows for continuous variation between natural and uniform weighting (in CASA from -2 to 2)
- usually can obtain most of natural weight sensitivity at the same time as most of uniform weight resolution (!)

Tapering: apodize *(u,v)* sampling by a Gaussian function

- lower resolution to improve sensitivity to extended structure

Example Dirty Beams (same data!)



natural
ΔS = 10 μJy

robust=0
ΔS = 16 μJy

uniform
ΔS = 28 μJy

natural + 1” taper
ΔS = 23 μJy

- imaging parameters provide a lot of freedom
- appropriate choice depends on science goal, e.g.
  - point source detection: natural weight
  - fine detail of strong source: uniform weight
  - weak and extended emission: taper

Beyond Dirty Images: Deconvolution

- to keep you awake at night:

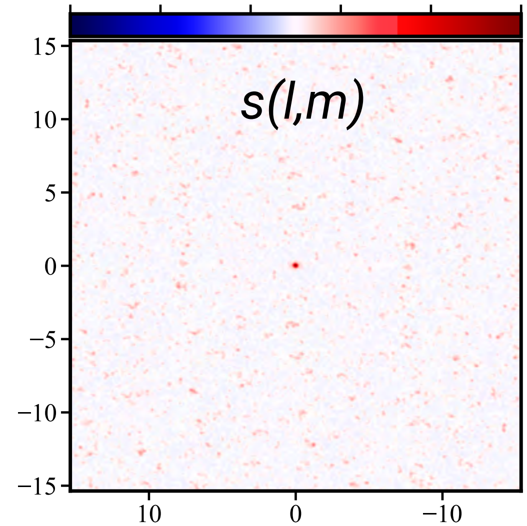
  ∃ an infinite number of *T(l,m)* compatible with sampled *V(u,v)*

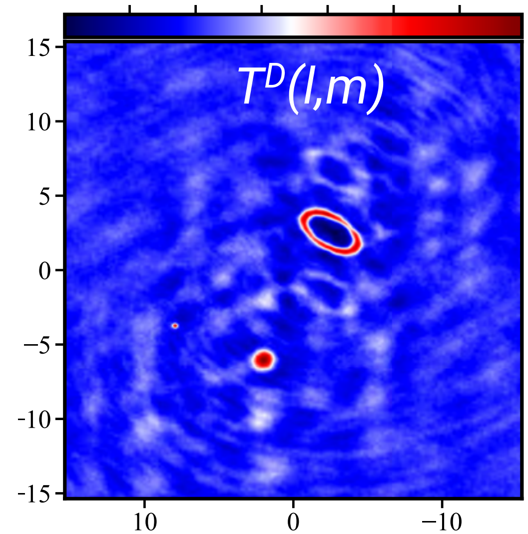
Deconvolution

- use non-linear techniques to interpolate/extrapolate *V(u,v)* samples into unsampled regions of *(u,v)* plane to find a *plausible* model of *T(l,m)*
- requires *a priori* assumptions about *T(l,m)* to fill unsampled *(u,v)* plane
- dominant deconvolution algorithm in radio astronomy is “clean”
  - *a priori* assumption: T(l,m) can be represented by point sources
  - many variants have been developed for computational efficiency and performance on extended structure (Clark, Cotton-Schwab, Multi-Scale, …)
- a very active research area!
  - see EHT paper for some examples, e.g. regularized likelihood methods

Classic Hogbom Clean Algorithm

1. find highest peak of residual image
2. subtract scaled dirty beam s(l,m) x “loop gain” from this peak
3. add this point source position and amplitude to a clean component list and to a model image
4. if peak of residual image > stopping threshold, then goto 1





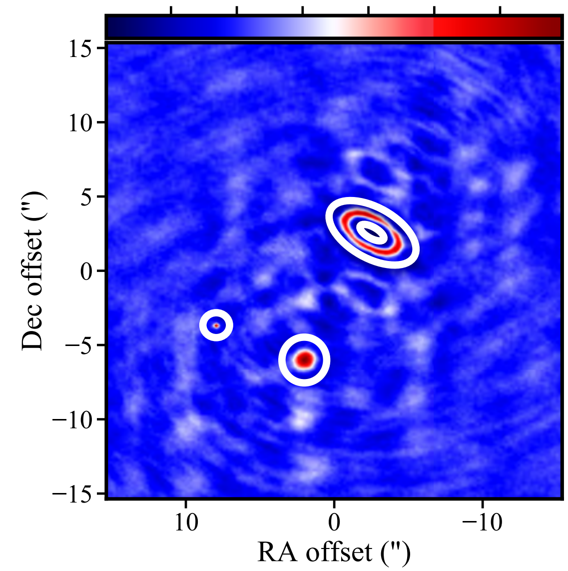
create final **restored image**

a) convolve the model image by a “clean beam” = an elliptical Gaussian fit to the main lobe of the dirty beam

b) add back residual image = noise + residual source structure

Clean Algorithm Parameters

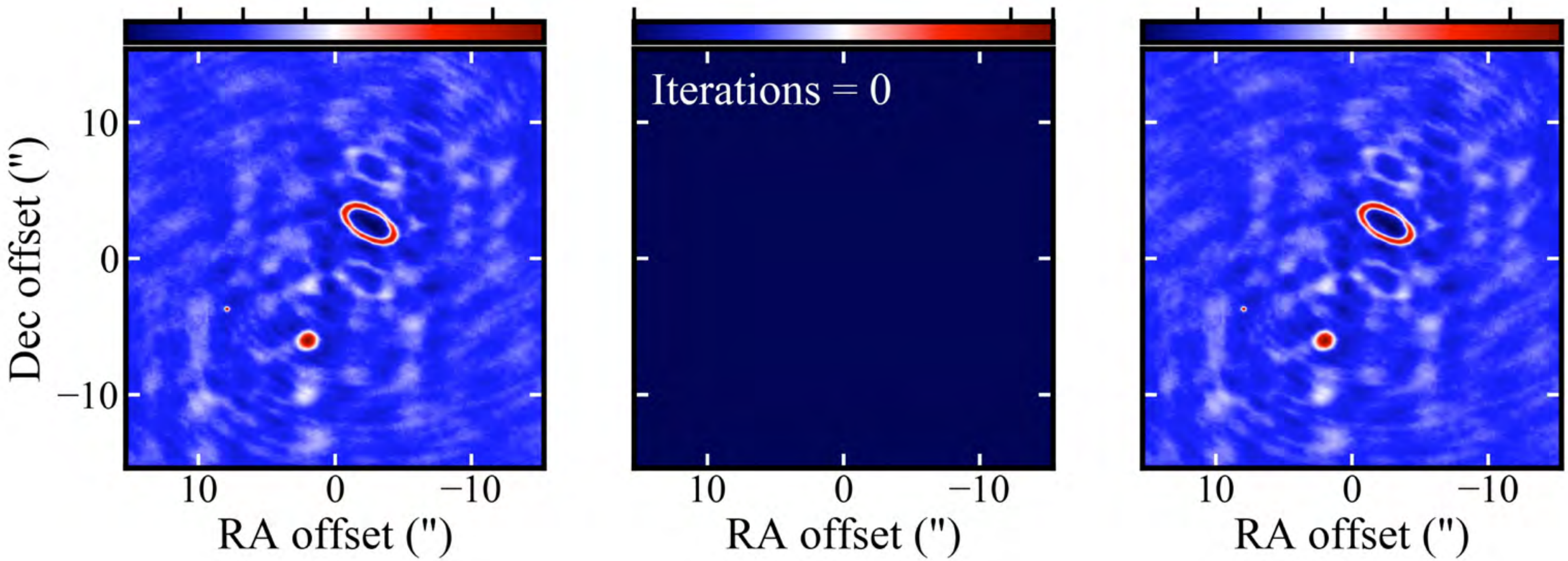
- loop gain
  - good results for 0.1 – 0.3 (CASA tclean default = 0.1)
  - lower values can help with smooth and extended emission
- stopping threshold
  - peak < threshold = multiple of theoretical rms noise, e.g. 3 x rms
  - peak < threshold = fraction of dirty map maximum (useful if strong sources prevent theoretical rms threshold from being reached)
- finite support (image masks)
  - include *a priori* information about where to search for clean components in image
  - useful, often essential, for best results
  - can be dangerous
  - can be an arduous manual process, but modern automasking algorithms work well



Clean Example

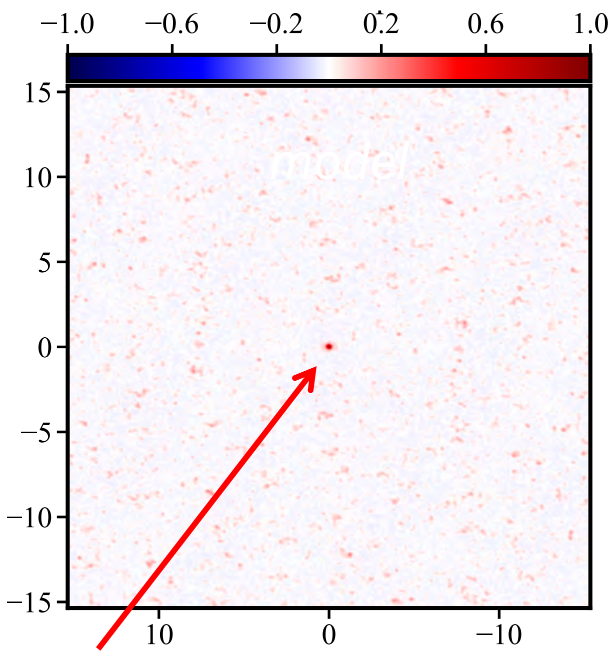
initialize

$T^D(l,m)$ *clean components* *residual image*



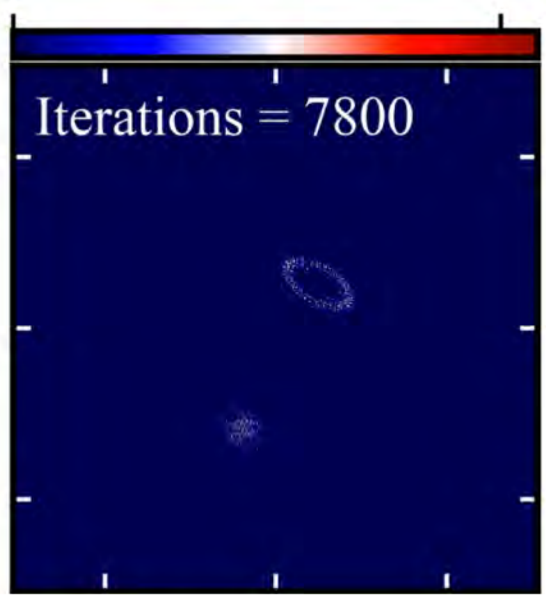
Clean Example: Restored Image

s(l,m)



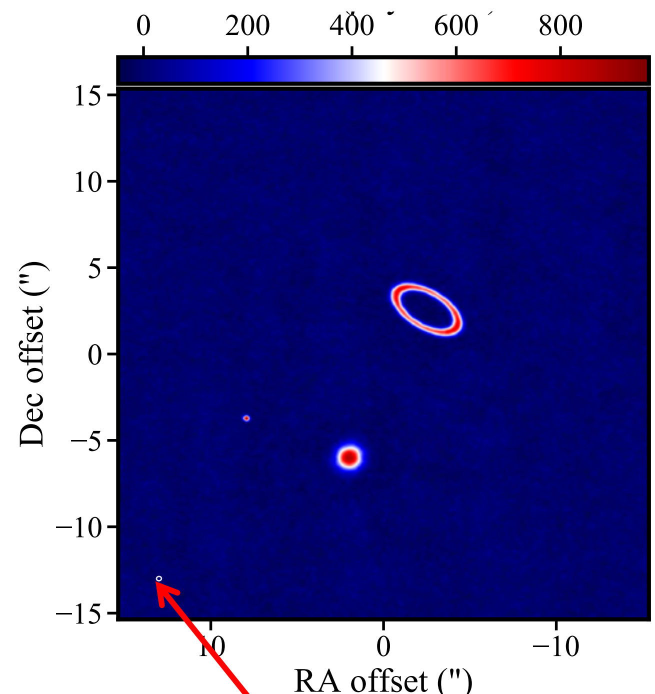
fit Gaussian to *s(l,m)*

model image



convolve model image by fitted Gaussian and add back residuals

restored image



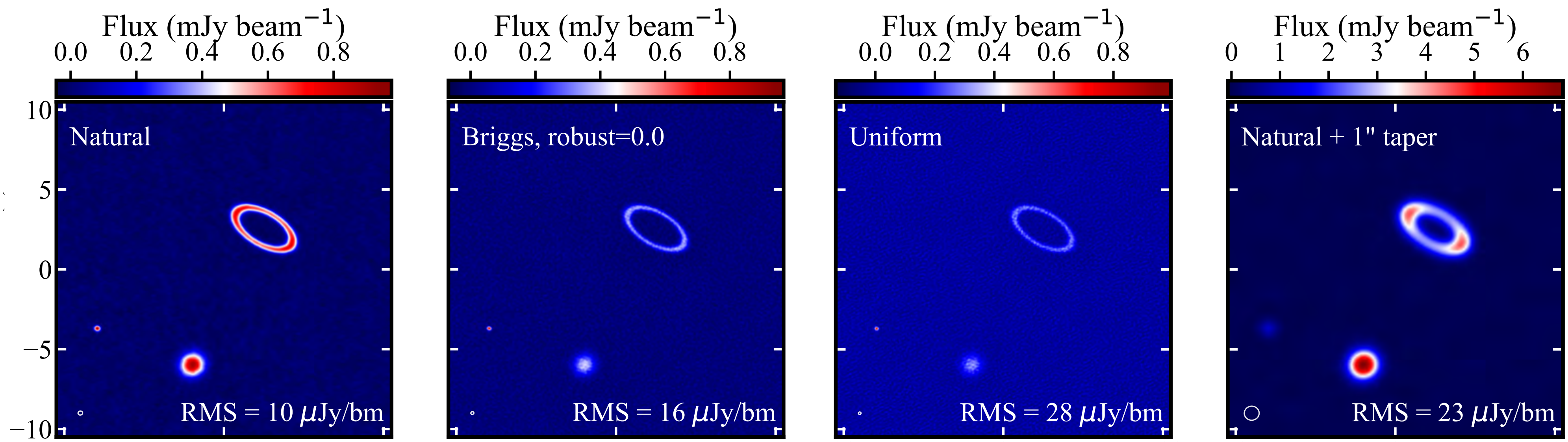
ellipse = restoring beam fwhm

final image depends on

- imaging parameters: pixel size, visibility weighting scheme, gridding, …
- deconvolution: algorithm, parameters, iterations, …

Clean: Different Weighting Schemes

- restored images emphasize different angular scales from the *V(u,v)* samples



natural
fwhm 0.29x0.25
ΔS = 10 μJy

robust=0
fwhm 0.19x0.17
ΔS = 16 μJy

uniform
fwhm 0.17x0.15
ΔS = 28 μJy

natural + 1” taper
fwhm 0.93x0.89
ΔS = 23 μJy

CASA `tclean` filename extensions

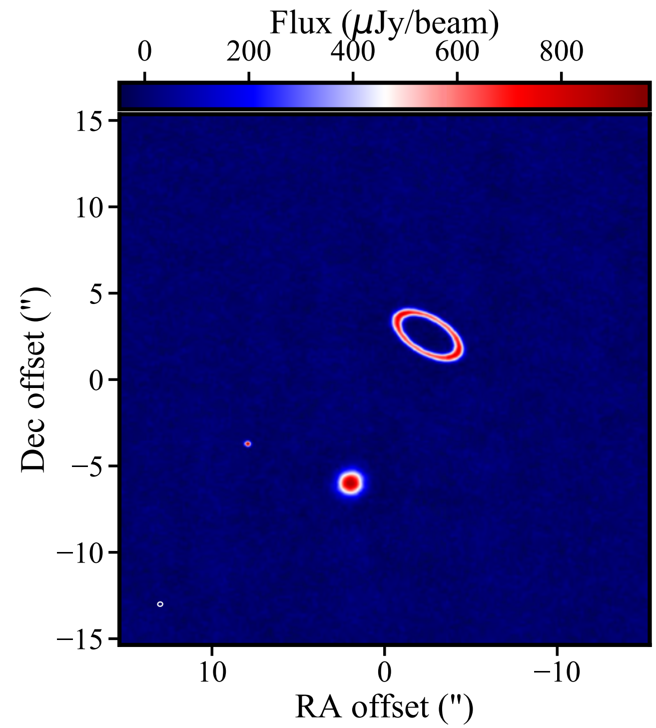
- <*imagename*>.image
  - restored image
- <*imagename*>.psf
  - point spread function (dirty beam)
- <*imagename*>.model
  - model image after deconvolution, i.e. clean components
- <*imagename*>.residual
  - residual image, i.e. after subtracting clean components
- <*imagename*>.mask
  - deconvolution mask
- <*imagename*>.pb
  - primary beam model
- <*imagename*>.pbcor
  - primary beam corrected image

Measures of Image Quality

dynamic range

- ratio of peak brightness in image to rms noise in region devoid of emission
- easy way to calculate a *lower limit* to the error in brightness in a non-empty region

e.g. peak 880 and rms 10 μJy/beam
→ dynamic range = 88



fidelity

- difference between reconstructed image and the true image
- fidelity = input model/difference = inverse of the relative error
  = model * beam / abs (model * beam − reconstructed image)
- generally much lower than implied by dynamic range

Spectral Lines and Polarization

- discussion so far applies to a single spectral channel of width $\delta\nu$
- science may require many such channels across total bandwidth $\Delta\nu$
  - spectral lines from molecular or atomic transitions
  - a significant continuum spectral slope or curvature
  - also helpful from a technical perspective
    - edit interference, avoid "bandwidth smearing" (radial smearing in image from averaging over $\Delta\nu$ that limits useable field of view)
  - each channel is imaged (and deconvolved) independently

[Ylva Pihlstrom lecture]

- science may require full Stokes parameters: I, Q, U, V
  - each Stokes parameter imaged (and deconvolved) independently
  - then combined to form, e.g. fractional linear polarization $\sqrt{(Q^2+U^2)}$ /I

[Frank Shinzel lecture]

Concluding Remarks

- interferometry samples Fourier components of sky brightness
- make images by Fourier transforming the sampled visibilities
- deconvolution attempts to correct for incomplete sampling

- remember
  - there are an infinite number of images compatible with the visibilities
  - missing (or corrupted) visibilities affect the entire image
  - astronomers must make decisions in imaging and deconvolution
- it is fun and worth the trouble $\rightarrow$ high resolution images!

many, many issues not covered in this talk,

see lectures on advanced imaging techniques and references