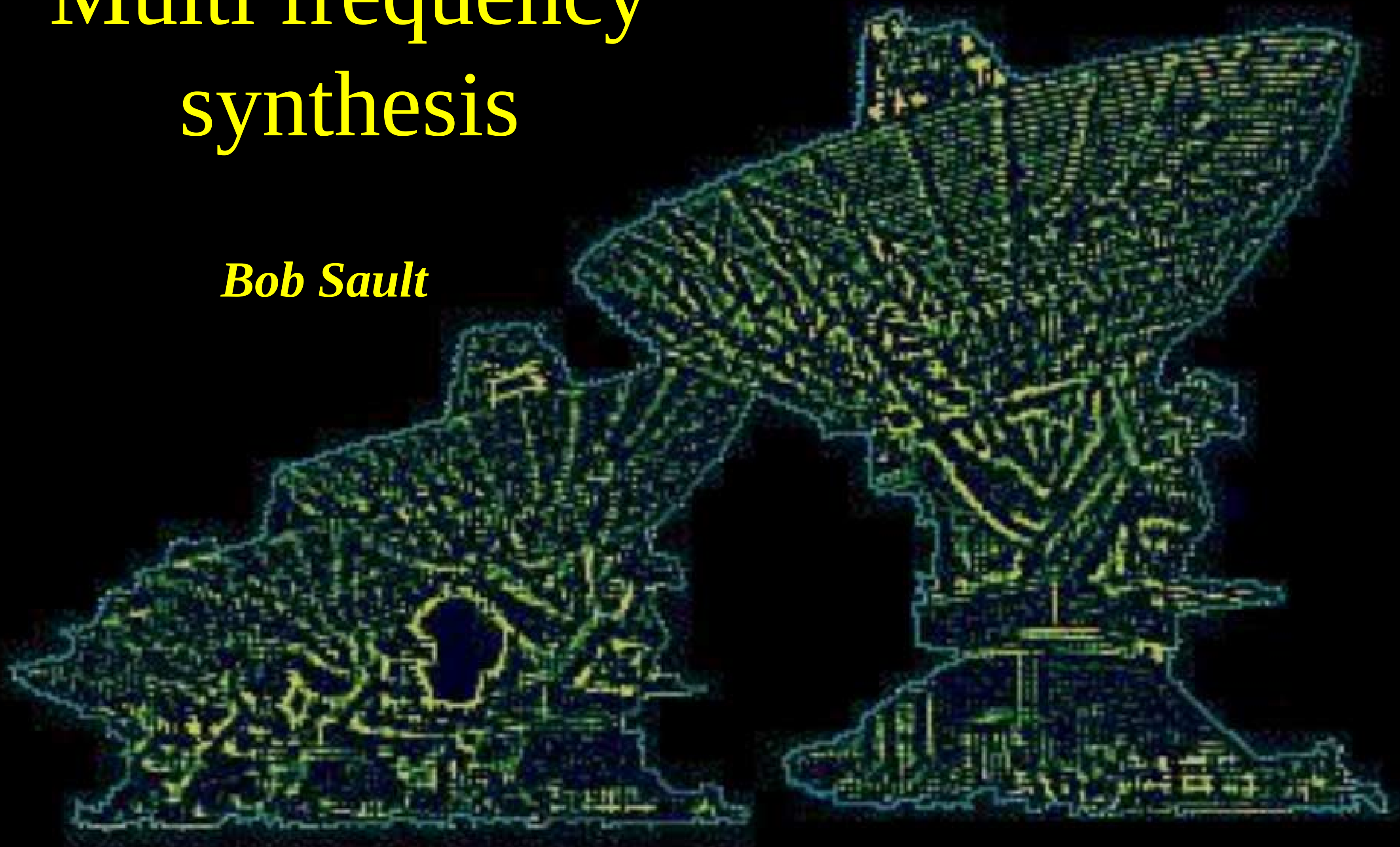
Deconvolution and Multi frequency synthesis

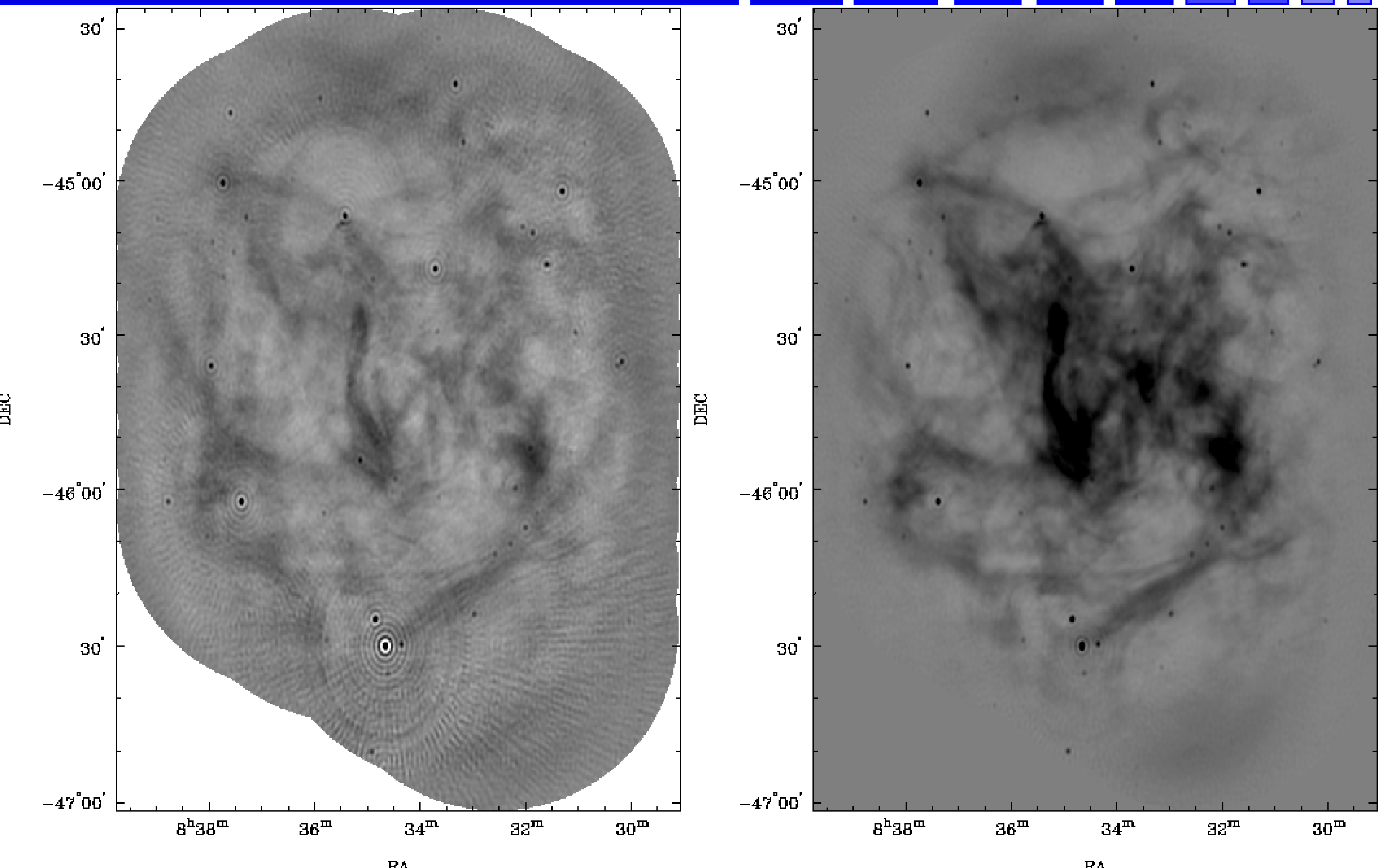
Bob Sault



Deconvolution

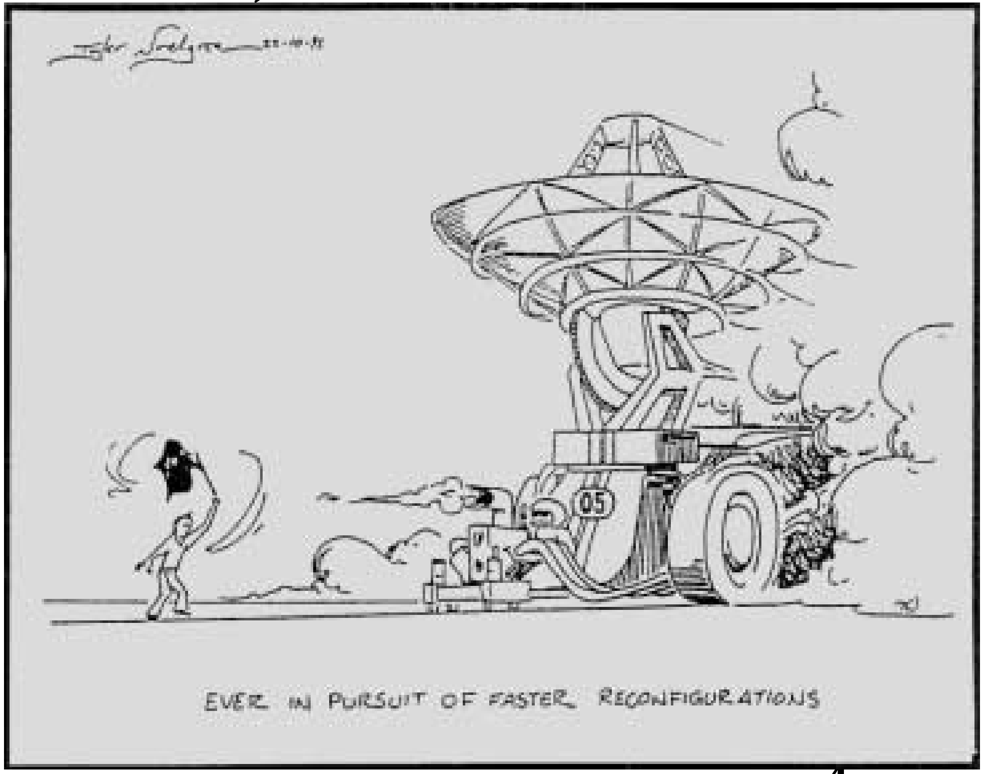
- Basics (again!)
- Multi-frequency synthesis
- Characteristics of the dirty beam
- Linear deconvolution
- Constraints
- CLEAN
- Maximum entropy
- Restoration
- Multi-frequency deconvolution

An example of deconvolution



Filling the Fourier plane

- Use many antennas (6 antennas or more)
- Use Earth rotation (12 h observations)
- Physically move antennas



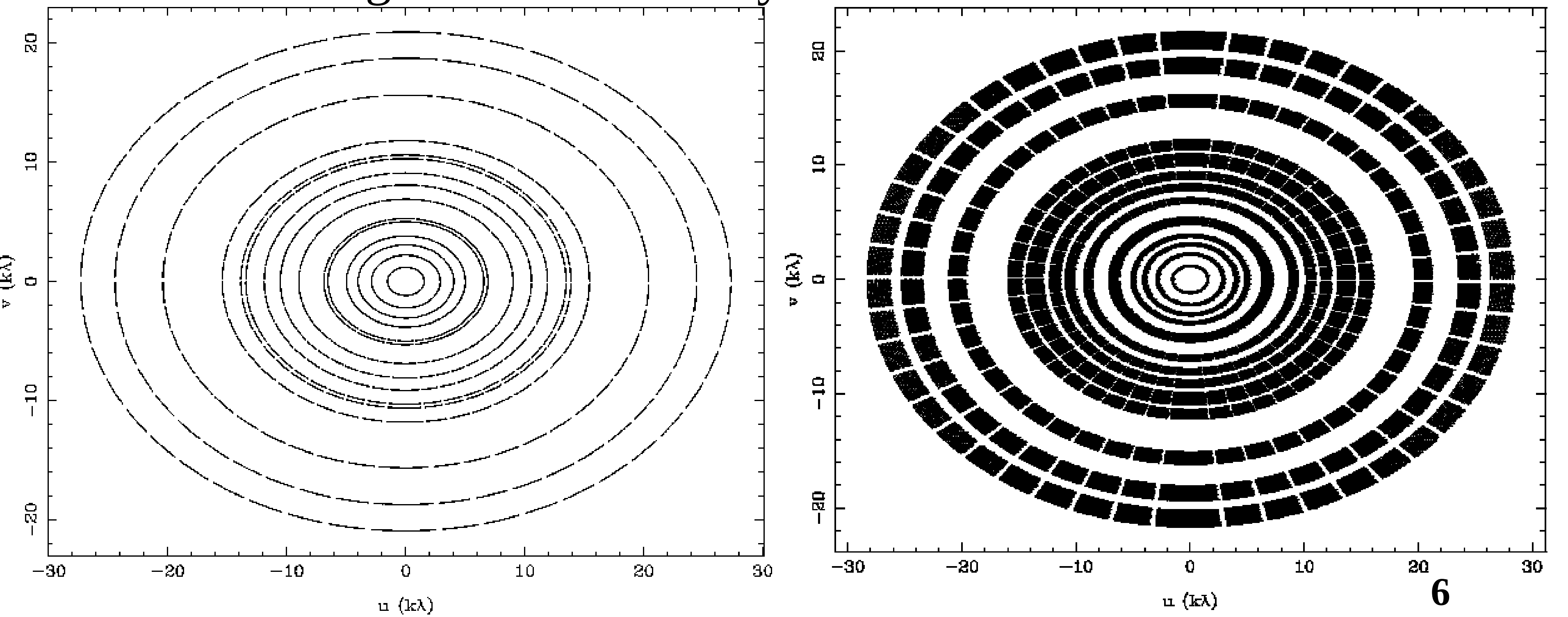
Unfilled Fourier Plane

But the aperture is ***NEVER*** completely filled

- Limited observing time
- Limited number of antennas
- Various interruptions to the observation
- Min and max baselines

Multi-frequency synthesis

- As *(u,v)* coordinate is measured in wavelengths, another way of filling the Fourier plane is to observed at multiple wavelengths simultaneously.



Basic imaging relationship

$$V(u,v) = \int I(\ell,m) \exp(i2\pi(u\ell + vm)\, dl\, dm$$

Using a "direct Fourier transform" we produce the dirty image

$$\begin{aligned} I_{\mathrm{D}}(\ell,m) &= \sum_k w_k V(u_k,v_k) \exp(-2\pi(u_k\ell + v_k m) \\ &= \mathcal{F}\left[\left(\sum_k w_k \delta(u-u_k, v-v_k)\right) \cdot V(u,v)\right] \\ &= \mathcal{F}\left[S(u,v) \cdot V(u,v)\right] \end{aligned}$$

Convolution relationship

Fourier theory tells us that

$$F(u)G(u) \Leftrightarrow f(\ell) * g(\ell)$$

so

$$I_{\mathrm{D}}(\ell, m) = B(\ell, m) * I(\ell, m)$$

where

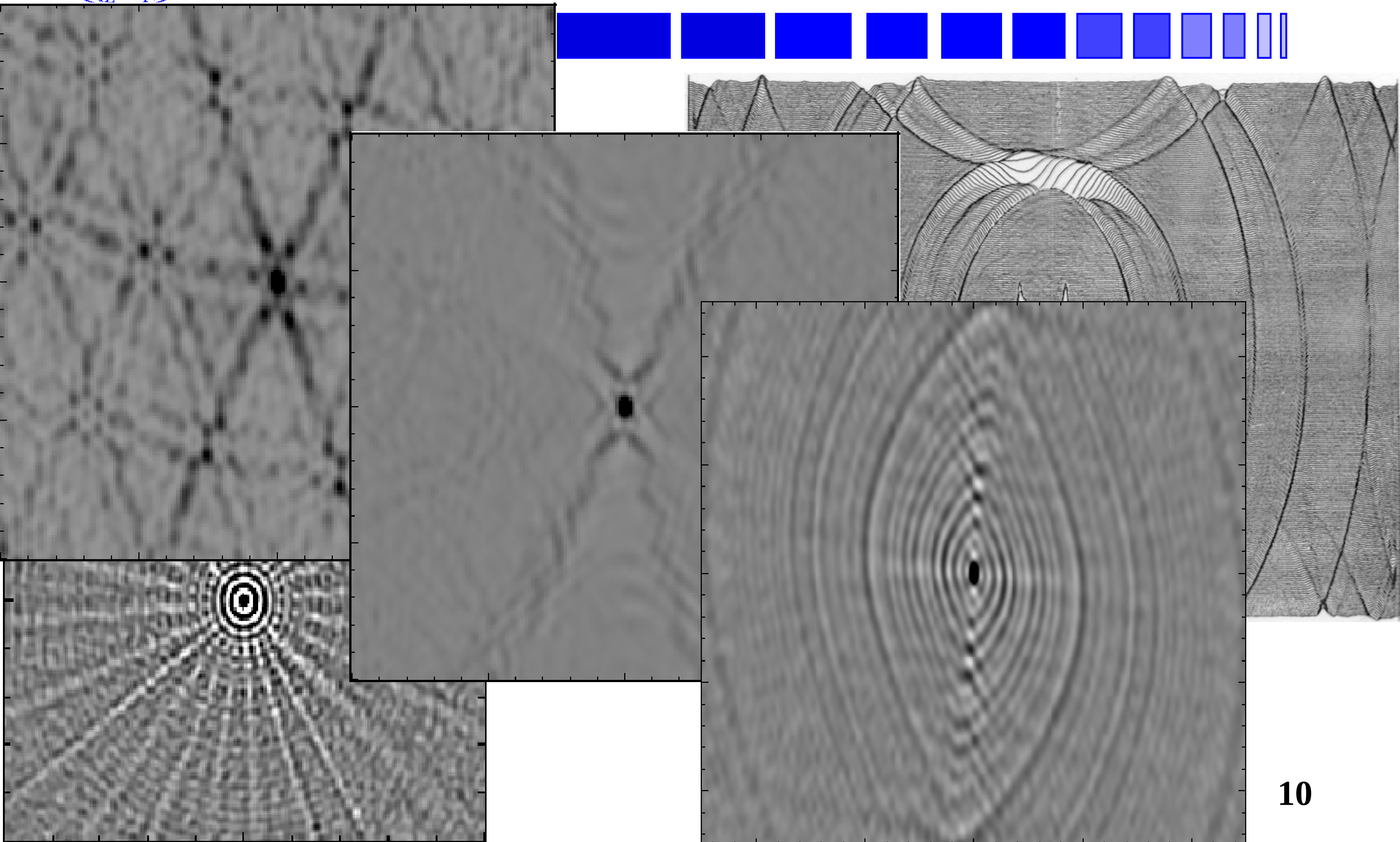
$$B(\ell, m) = \mathcal{F}\left[\sum_k w_k \delta(u - u_k, v - v_k)\right]$$

Jargon: The point-spread function is usually called the “beam”.

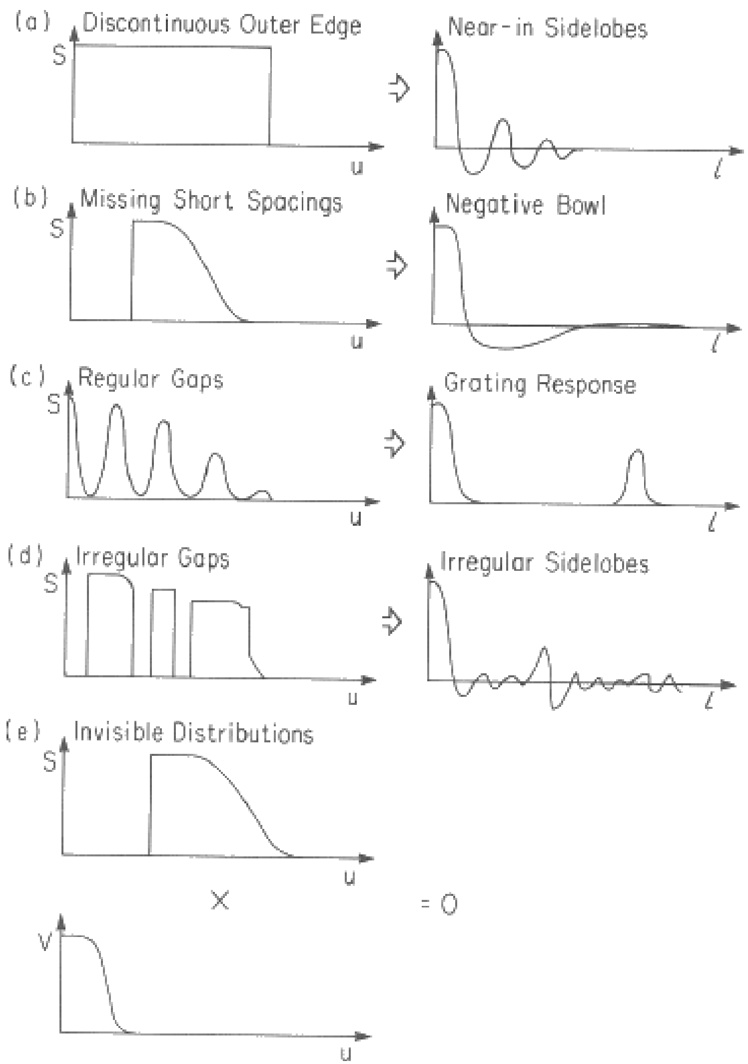
Deviations from convolution relationship

- Wide-field effects are usually neglected. These include
  - Time and bandwidth smearing
  - Primary beam effects
  - So-called non-coplanar baseline effects
- Convolution relationship strictly applies only for continuous functions (not a sampled grid of pixels).
- "Aliasing" in the imaging process is also not accounted for.
- Finite extent assumption.

Dirty beams



Dirty beam characteristics



Differing “holes” in the Fourier plane lead to a wide variety of sidelobe structure

Linear deconvolution

- Inverse filter

$$I_{\mathrm{D}} = B * I$$

then

$$\mathcal{F}[I] = \mathcal{F}[I_{\mathrm{D}}] / \mathcal{F}[B]$$

- Wiener filters

$$\mathcal{F}[I] \approx \mathcal{F}[I_{\mathrm{D}}] \cdot \frac{1}{\mathcal{F}[B] + \mathrm{SNR}^{-1}}$$

Linear deconvolution …

- Noise properties are well understood
- Generally non-iterative and computationally cheap

But

- It does a very poor job
- Rarely used in practical radio interferometry

Non-linear deconvolution

- Linear deconvolution is fundamentally unable to extrapolated unmeasured spatial frequencies.
- A function which is non-zero only in the unsampled part of the Fourier plane is called an *invisible distribution*.
- A good non-linear deconvolution algorithm is one that picks plausible invisible distributions to fill in the Fourier plane.

Prior Information or Assumptions

- Bounded support ("CLEAN boxes").
- Positivity
- The sky is mostly empty
- Use a goodness measure to pick "reasonable" solutions.

But also use extra data

- Joint deconvolution of multiple pointings (mosaicing).
- Joint deconvolution of multiple polarisations.

CLEAN Algorithm (Högbom, 1974)

- Assumes that the sky can be modelled as a collection of point sources.
- Iteratively decomposes the sky into a collection of point sources.
- In principle, CLEAN is guaranteed to converge, although in practice it can become unstable if pushed too far.
- Generally it is quite a robust algorithm.

CLEAN algorithm

1. Search for the largest peak in the residual image
2. Assume this is a result of a point source – a component!
3. Subtract off some fraction (“damping factor” or “loop gain”) of the point source.
4. Add that fraction of the point source to a component list.
5. *Iterate*

Iteration stops when the residual is below some cut-off, when a negative component is encountered, or when a fixed number of components are found.

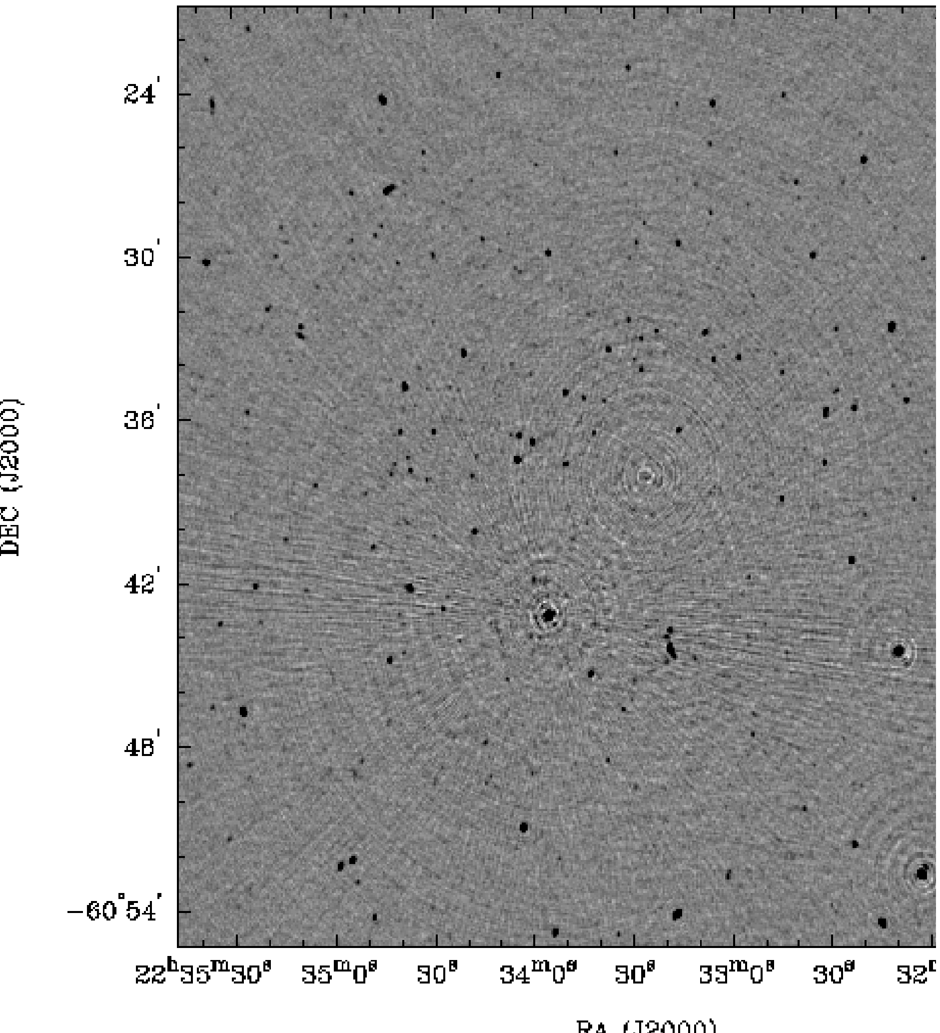
CLEAN implementations

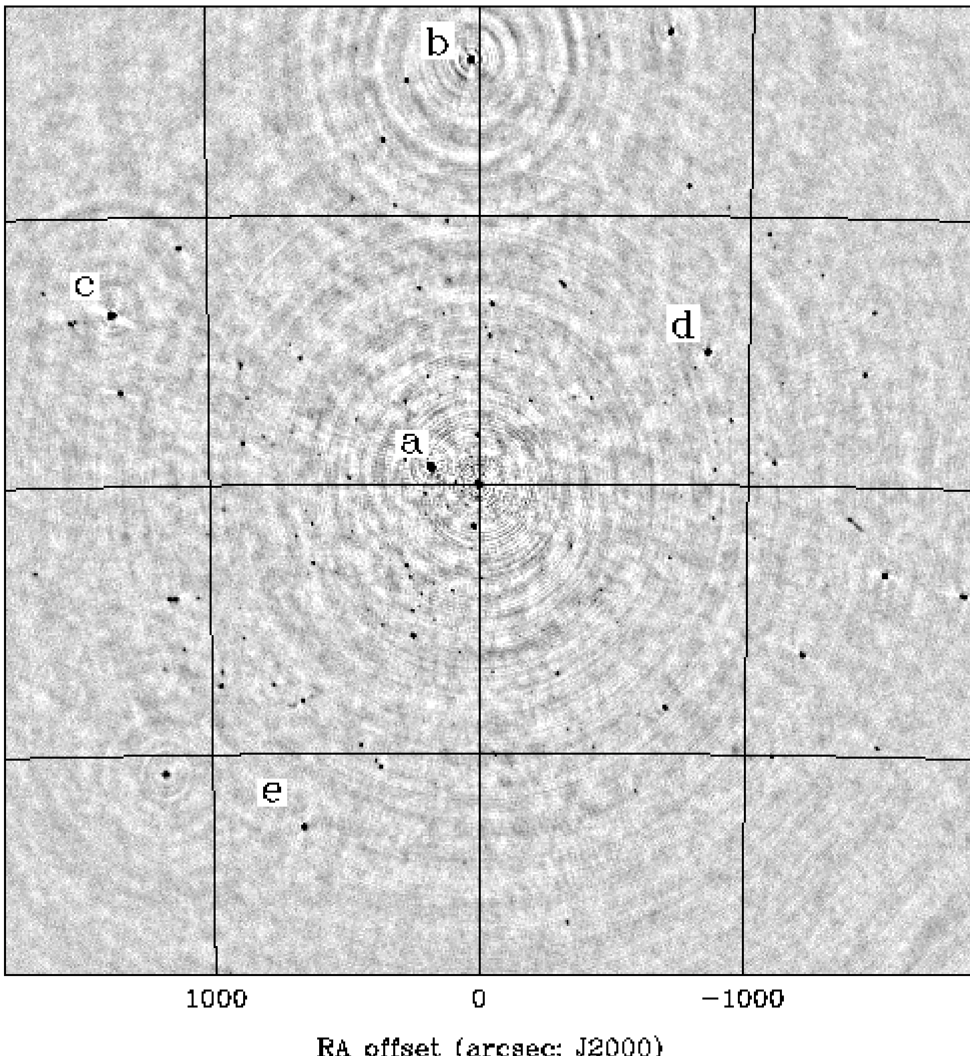
- There are different implementations of the algorithms (with their individual strengths and weaknesses):
  - Högbom algorithm – the classical one
  - Clark algorithm – faster for large images with many point sources.
  - Cotton-Schwab ("MX") algorithm – works partially in the visibility domain. Able to cope with extra artefacts. Can be slow.
  - Steer Dewdney Ito algorithm – works best for very extended objects.

Strengths/weaknesses

- CLEAN is good for fields of sources which are unresolved or just resolved.
- Generally quite robust in the face of many defects.
- CLEAN is very poor for very extended objects:
  - Slow!
  - Corrugation instability.
  - CLEAN poorly estimates broad structure (short spacings). The result is the so-called “negative bowl” effect.
- CLEAN’s procedural definition makes it difficult to analyse.

Examples of CLEANed images





Bayesian Statistics and Maximum Entropy

- Two basic views of probability theory:
  - Views probability distribution function as a measure of the relative frequency of an outcome.
  - Views probability distribution function as a reflection of our uncertainty.
- Principle of maximum entropy:
  Of all the possible probability distributions which are consistent with the available information, the one that has the maximum entropy is most likely the correct one.
- Maximum entropy image deconvolution:
  Of all the possible images consistent with the observed data, the one that has the maximum entropy is most likely to be the correct one.

Maximum entropy

- Of all the possible images, pick that one which maximises some goodness measure called "entropy".

$$\mathcal{H} = \sum f(I(\ell, m))$$

- The most popular choice is the entropy function

$$f(I) = -I \log(I/eM)$$

Maximum entropy

- The solution is generally constrained so that a χ^2 measure is consistent: i.e. the χ^2 measure is consistent with the expected noise level.
- Integrated flux constraint can be included.
- "CLEAN box" constraint is readily added.
- The default image, *M*, can be chosen to be a uniform value, or can be set to some prior expectation of the source.
- Solution image must be positive-valued.

Strengths/weaknesses

- Fourier extrapolation tends to be more conservative than CLEAN.
- Tends to work better for images with a large amount of extended emission.
- Tends to be faster for large images (> 1024x1024 pixels).
- Susceptible to analysis.
- Depends more critically on its control parameters (e.g. noise variance and integrated flux).
- More likely to blow up on poorly calibrated data, or data that violates the convolution relationship in some way.
- Poorly deconvolves point sources.

CLEAN vs MEM

The answer is image dependent:

- "High quality" data, extended emission, large images
  $\Rightarrow$ Maximum entropy
- "Poor quality" data, confused fields, point sources
  $\Rightarrow$ CLEAN

"Restoration" Step

- CLEAN and MEM "super-resolve", and the high spatial frequencies can be of poor quality (particularly CLEAN).

  Solution: Downweight the high spatial frequencies by convolving with a gaussian "CLEAN beam".

$$\hat{I}_{\mathrm{R}} = \hat{I} * B_{\mathrm{C}} + \underbrace{(\hat{I}_{\mathrm{D}} - \hat{I} * B)}_{\text{residuals}}$$

- The CLEAN beam usually has the same FWHM as the main lobe of the dirty beam.

Why include the residuals?

- The residuals give an easy way of seeing how believable the features in an image are.
- The residuals still contain emission from sources that have not been CLEANed out.

Multi-frequency deconvolution

- Multi-frequency synthesis uses observations at many frequencies to prove the Fourier plane coverage.
  Problem: *Source structure is a function of frequency*.

- For modest spread in fractional bandwidth (< 15%), and modest dynamic range (< 500), the errors caused by source structure varying with frequency can be ignored.

- When this is not the case, a multi-frequency deconvolution algorithm can be used to eliminate the resultant errors.

Multi-frequency deconvolution algorithm

- The algorithm models the spectral variation at each pixel as a constant and a linearly varying component with frequency.
- The response to the constant part of this variation is just the normal dirty beam.
- The response to the linearly-varying component can be represented by a second response function. The dirty image is the sum of the responses to the constant and varying components.
- A joint deconvolution, simultaneously solving for the two components can be performed.

In *Miriad*

Clean algorithms	CLEAN
Maximum entropy	MAXEN
Multi-frequency Clean	MFCLEAN
Mosaicing (max entropy, clean)	MOSMEM, MOSSDI
Joint polarimetric (single pointing or mosaicing)	PMOSMEM