Principles of Interferometry

Hans-Rainer Klöckner

IMPRS Black Board Lectures 2014

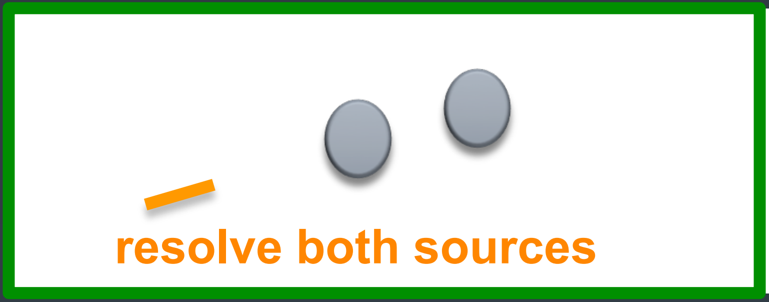
acknowledgement

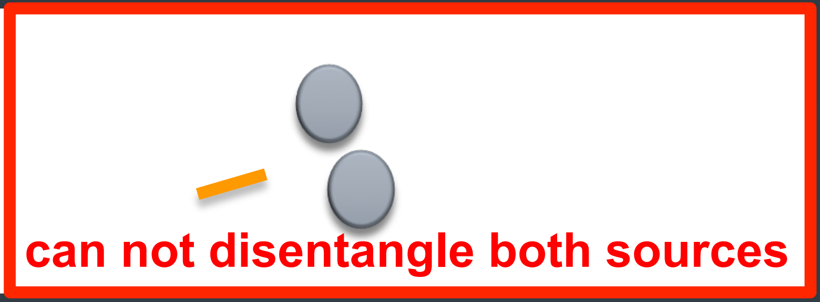
- Mike Garrett lectures
- James Di Francesco crash course lectures NAASC

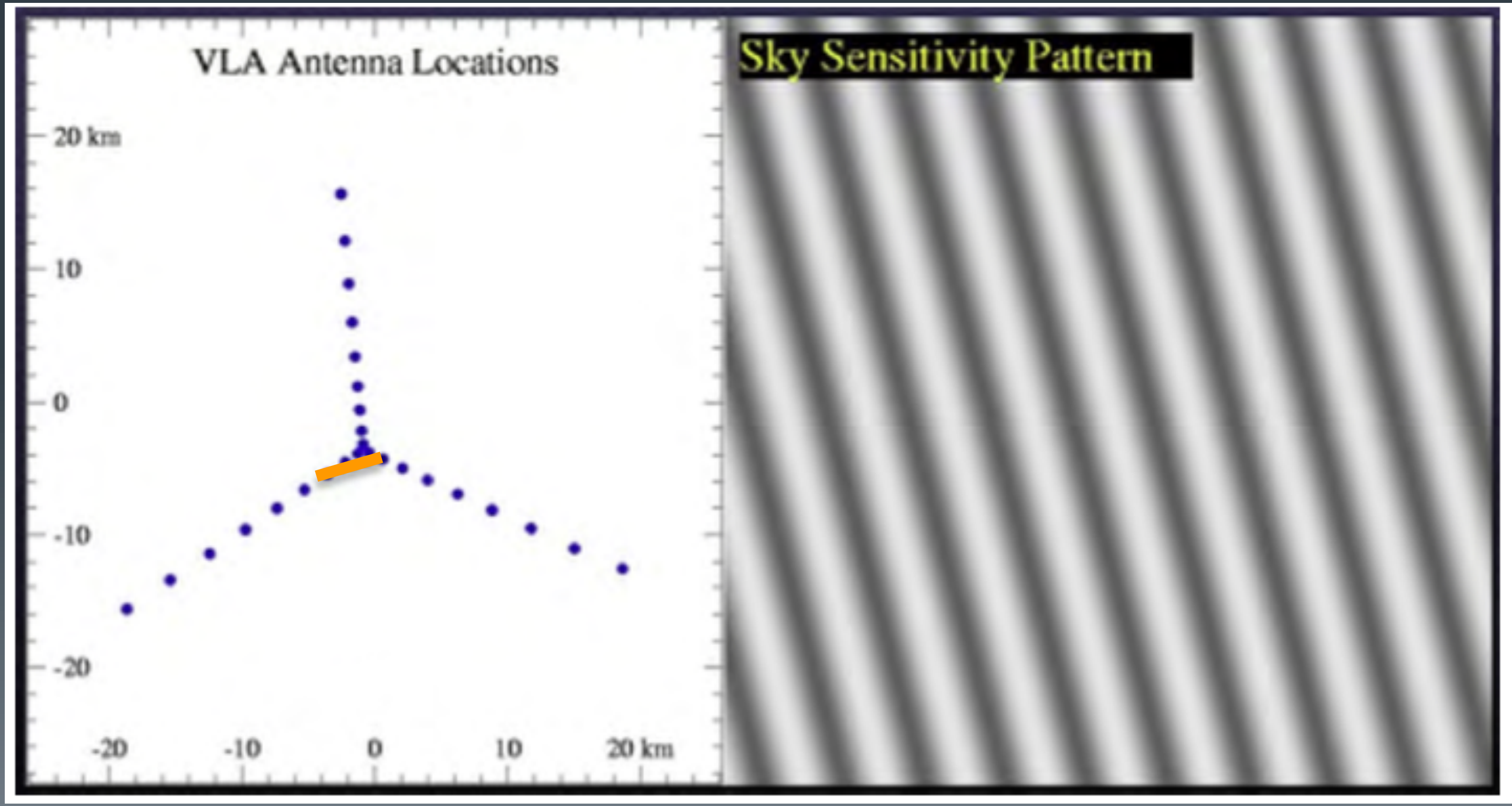
Lecture 5

- calibration
- image reconstruction
- self-calibration
- measurement equation

structure traced by baseline

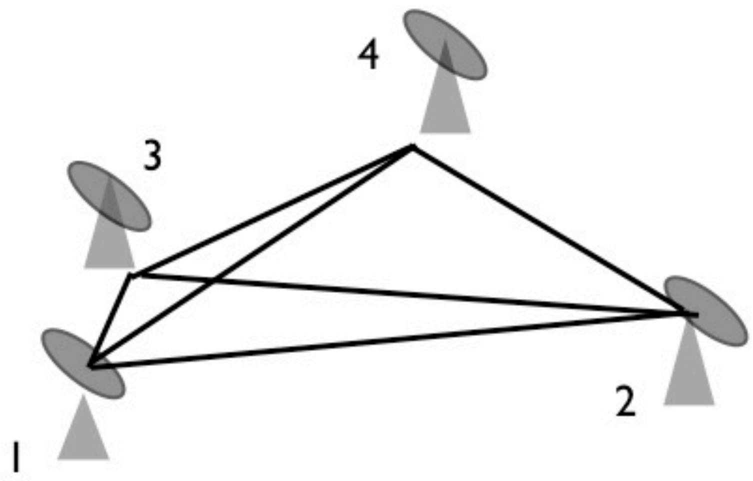






number of sampled baselines

A multi-element interferometer (say with N antennas) produces N(N-1)/2 unique responses:



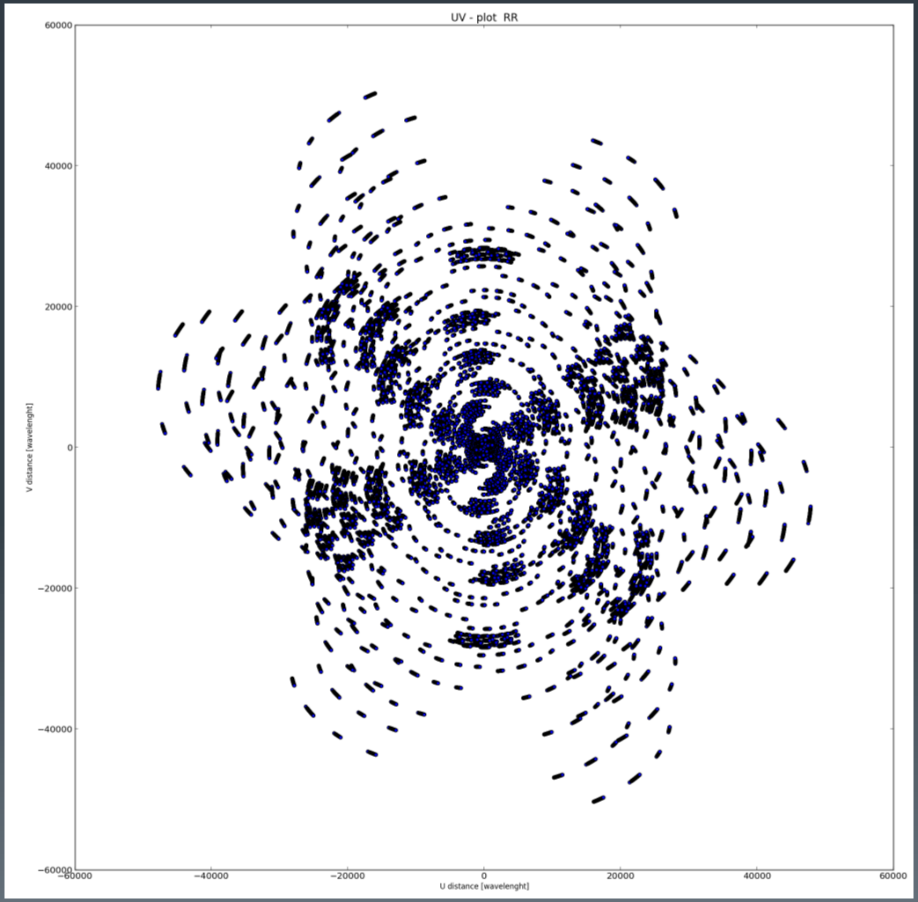
For N=4, 6 baselines responses are measured:

r_{12}, r_{13}, r_{14}, r_{23}, r_{24}, r_{34}.

VLA	(27*26) /2 = 351
WSRT	91
GMRT	435

uv - coverage

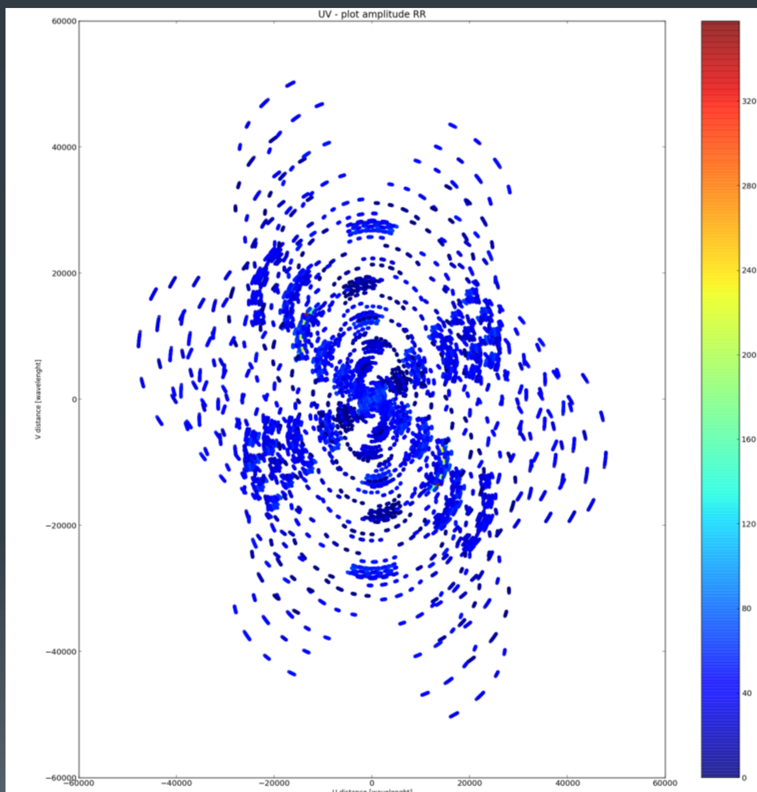
GMRT



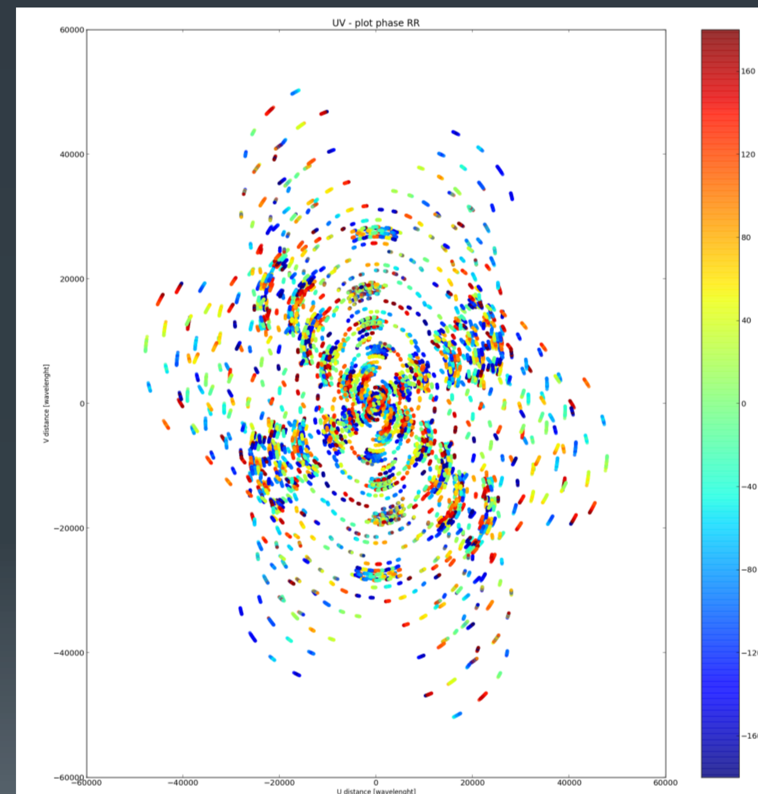
- geometric quantity
- there is nothing to be done if the model is correct
- unless you do Geodesy

what needs calibration

UV coverage amplitude (colour coded)



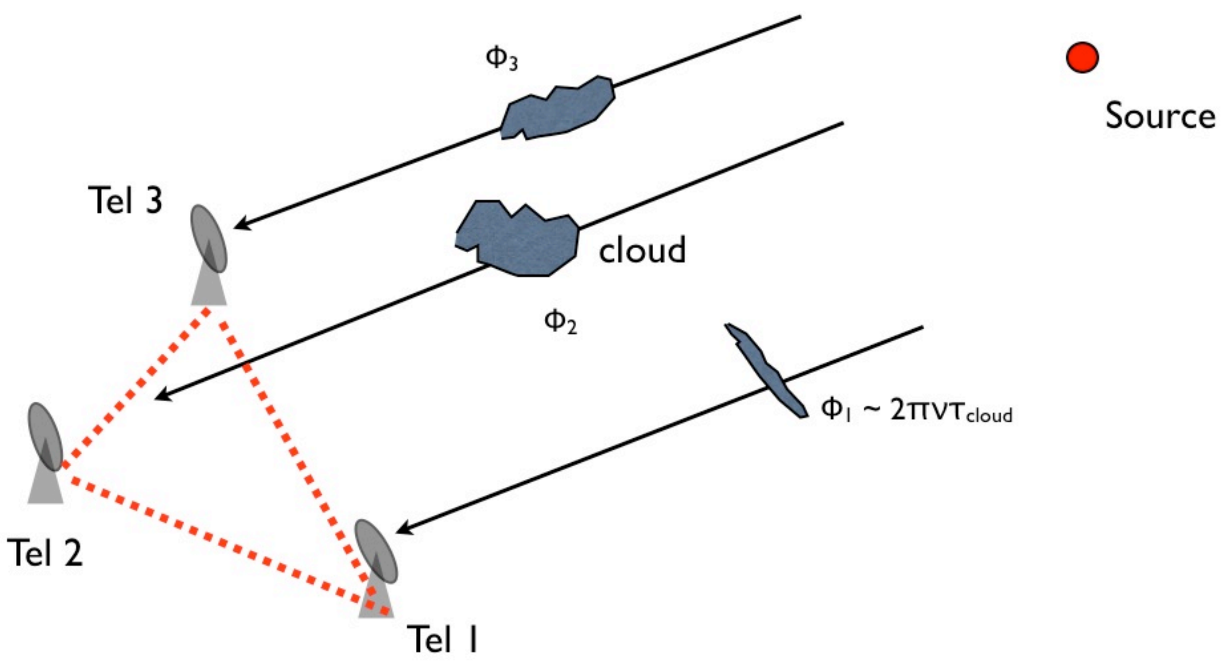
UV coverage phase (colour coded)



$$V_{ij}^{true}(t) = A_{ij}(t)\ e^{i\,\varphi ij(t)}$$

closure relation - phase

antenna based errors introduce phase and amplitude error



The measured visibility phase on baseline "12" is then: $\Phi_{12} = \varphi_{12} + \phi_1 - \phi_2$ [1]

where φ_{12} is the true source visibility on baselines 1-2, Φ_1 are phase offsets introduced by the clouds above each telescope.

closure relation - phase

So:

$$\begin{aligned}\Phi_{12} &= \varphi_{12} + \phi_1 - \phi_2 \\ \Phi_{23} &= \varphi_{23} + \phi_2 - \phi_3 \\ \Phi_{31} &= \varphi_{31} + \phi_3 - \phi_1\end{aligned} \qquad [2a]$$

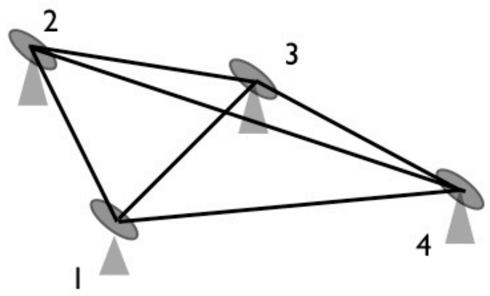
Clearly if we add these relations together:

$$\begin{aligned}\underbrace{\Phi_{12} + \Phi_{23} + \Phi_{31}}_{\textit{closure phase}} &= \varphi_{12}+\varphi_{23} + \varphi_{31} +(\phi_1 - \phi_1) +(\phi_2 - \phi_2)+(\phi_3 -\phi_3) \\ &= \varphi_{12}+\varphi_{23} + \varphi_{31}\end{aligned} \qquad [2b]$$

This formulation of adding together the observed visibility phases together of any 3 telescopes is known as forming a "closure triangle". [2b] is known as the closure phase for these 3 telescopes.

N.B. the important thing to note is that the closure phase contains information only on the true visibility of the source itself, i.e. its brightness distribution - ALL other telescope based errors cancel out (e.g. atmosphere, cable lengths, electronics etc.).

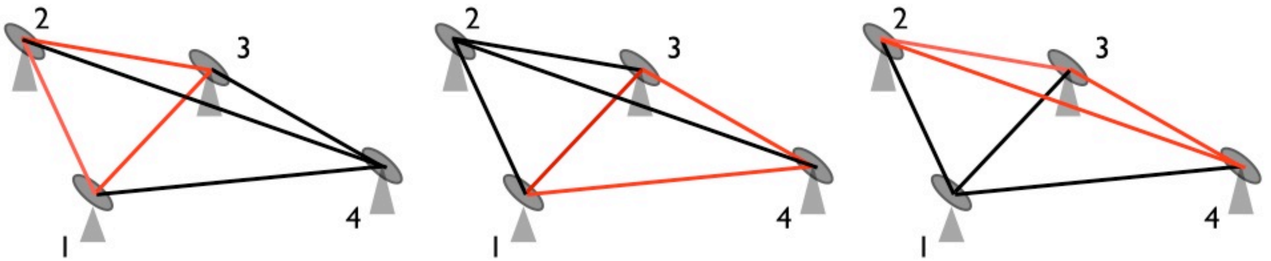
closure relation - phase



For a given array of N telescopes, there are:

(N-1)(N-2)/2 independent closure phases [3]

e.g. for N=4 there are 3 *independent* closure relations.



known are needed to solve the equations

So from the closure relations we have (N-1)(N-2)/N good observables (measurements).

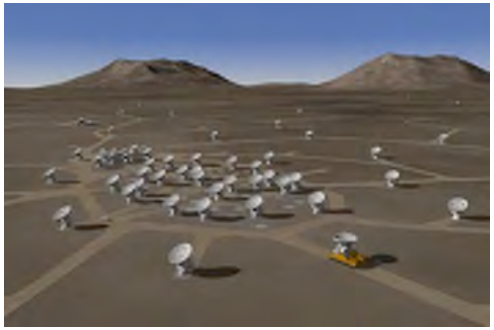
However, there are N telescope unknowns. We can reduce this to (N-1) unknowns if we make one of the telescopes the "reference antenna" (i.e. set the phase error to zero for this telescope).

Note that the ratio of *good observables/unknowns* (see eqn[3]) is then just:

$$(N-2)/N \quad [4]$$

So for N=3 we have only 33% of the information we need.

But for N=27 (the case of the VLA) we have 93% of the information we need.



==> *in general the reliability of Interferometric images favours large-N telescope arrays (e.g. see SKA, ALMA and SKA pathfinders) - calibration is more robust and uv-coverage more complete.*

closure relation – general

Telescope errors do not only effect the phase of the visibility. The amplitude can also be degraded. However, phase errors usually dominate (at least at cm wavelengths where attenuation by the atmosphere is a relatively small effect).

In order to consider how self-calibration can be used to correct for amplitude errors, we must use a complex formalism:

$$V_{ij}^{obs}(t) = g_i(t)\, g^*_j(t)\, V_{ij}^{true}(t) \qquad [7]$$

where V_{ij} are the measured and true visibilities, and $g_i(t)\, g^*_j(t)$ are known as the complex gains of the telescopes i,j

The gains contain corrections to both the amplitude and phase of the visibility:

e.g. $g_i(t)=a_i(t)e^{i\phi_i(t)}$

In this formalism the observed and true Visibility can be written as:

$$V_{ij}^{obs}(t) = a_i(t)\, a_j(t)e^{i(\phi_i - \phi_j)}\, A_{ij}(t)e^{i\,\varphi ij(t)} \qquad [8] \qquad **$$

$$V_{ij}^{true}(t) = A_{ij}(t)\, e^{i\,\varphi ij(t)} \qquad [9]$$

closure relation - general

Note that by taking the ratios of eqns such as [7] we arrive at the "closure quantities". e.g.

$$V_{12}^{obs}(t) = g_1(t)\, g_2(t)\, V_{12}^{true}(t) = a_1(t)e^{i\phi_1(t)}\, a_2(t)e^{-i\phi_2(t)}\, A_{12}(t)\, e^{i\,\varphi_{12}(t)}$$

$$V_{23}^{obs}(t) = g_2(t)\, g_3(t)\, V_{23}^{true}(t) = a_2(t)e^{i\phi_2(t)}\, a_3(t)e^{-i\phi_3(t)}\, A_{23}(t)\, e^{i\,\varphi_{23}(t)}$$

$$V_{13}^{obs}(t) = g_1(t)\, g_3(t)\, V_{13}^{true}(t) = a_1(t)e^{i\phi_1(t)}\, a_3(t)e^{-i\phi_3(t)}\, A_{13}(t)\, e^{i\,\varphi_{13}(t)}$$

If we consider the phase terms only and implicitly accept time dependance:

$$V_{12}^{obs}\, V_{23}^{obs} \,/\, V_{13}^{obs} = e^{i(\phi_1 - \phi_2 + \varphi_{12} + \phi_2 - \phi_3 + \varphi_{23})}\, e^{-i(\phi_1 - \phi_3 + \varphi_{13})}$$

$$= e^{i(\varphi_{12} + \varphi_{23} - \varphi_{13})} \qquad [10]$$

Note that [10] is just the equivalent of our original closure phase presented in eqn[2b]

closure relation - general

If we consider only the amplitude terms, we can see that for some combination of observed visibilities, the amplitude gains will cancel:

$$V_{12}^{obs}V_{34}^{obs}/(V_{13}^{obs}\,V_{24}^{obs}) = \frac{A_{12}A_{34}\,a_1a_2a_3a_4}{A_{13}A_{24}\,a_1a_2a_3a_4} = \frac{A_{12}A_{34}}{A_{13}A_{24}} \qquad [11]$$

Such ratios are known as “closure amplitudes” and require *at least 4* telescopes to be formed.

Like closure phases, closure amplitude is a "good observable", since it is not sensitive to measurement error. The closure amplitude and closure phase relations can be exploited in the hybrid mapping algorithm (see earlier slides).

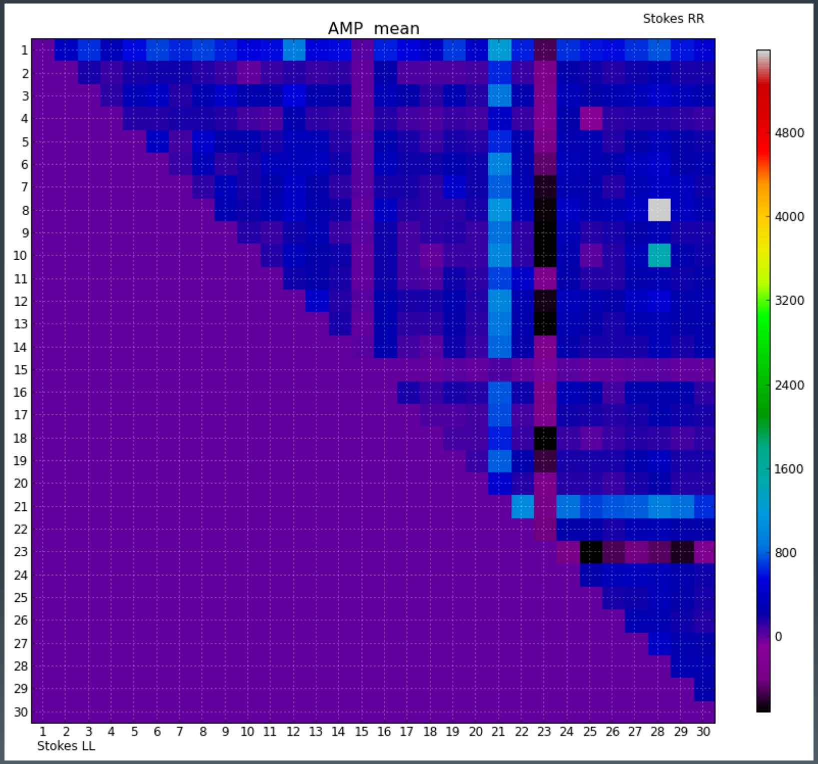
In the early days of hybrid mapping the closure phases and amplitudes were explicitly used to constrain the hybrid mapping process. In the era of the VLA it was no longer computationally efficient to calculate all the closure quantities. More sophisticated algorithms were constructed but they are all roughly equivalent to the original method. Modern algorithms seek to minimise the difference between the observed data and the predicted data:

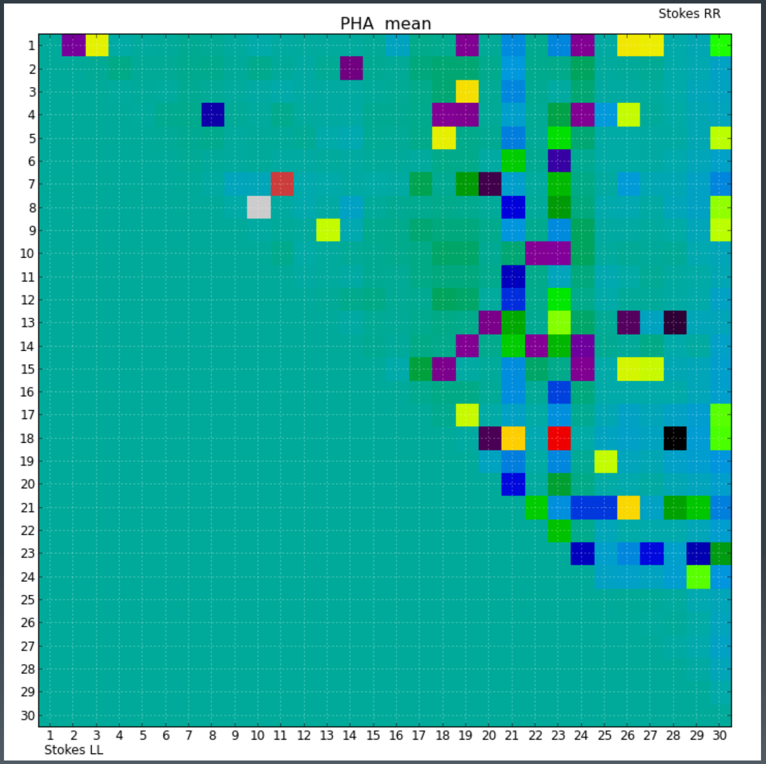
$$S = \Sigma_{ij,\, i<j} w_{ij}\, |V_{ij}^{obs} - g_i g^*_j\, V_{ij}^{true}| \qquad [12]$$

The w_{ij} reflect the fact that some data are higher weighted than other data (e.g. especially for VLBI arrays where all the telescopes have different sensitivities).

modern approach

schematic picture

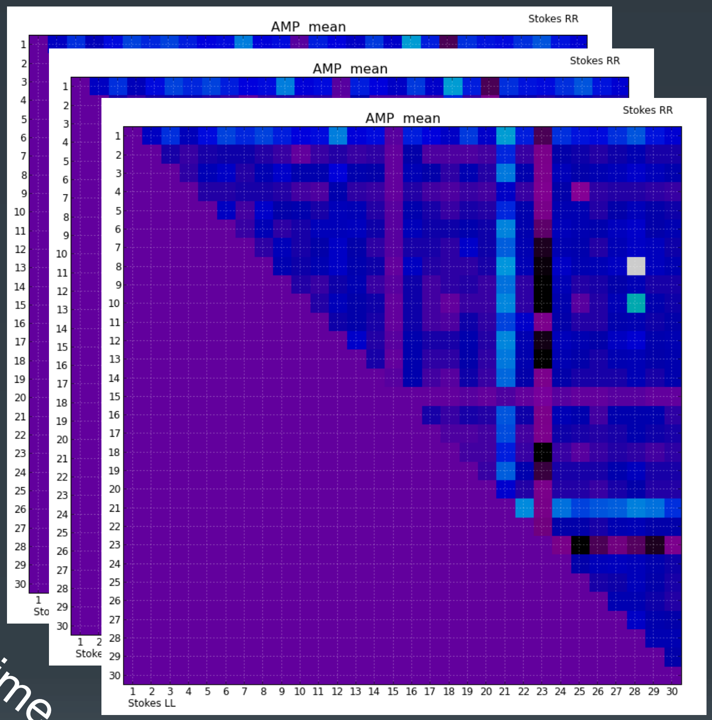




off diagonal values ONLY no autocorrelation data

$$S = \Sigma_{ij,\, i<j} w_{ij} \,|\, V_{ij}^{obs} - g_i g^*_j V_{ij}^{model} |$$

calibrate data – single source



time

actually a off diagonal complex matrix per integration time

$$V_{ij}^{model}(t) = A_{ij}(t)\ e^{i\ \varphi ij(t)}$$

can only be solved with a trick

assume a calibration source of 5 Jy at phase centre: $A_{ij} = 5$ & $\varphi_{ij}(t) = 0$

divide all the off diagonal terms of the matrix

add ones on the diagonal term

solve the matrix with the Gaussian elimination method to get diagonal matrix

$$\begin{pmatrix} \mathbf{A}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{A}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{A}_n \end{pmatrix}$$

the complex values on the diagonal are the complex gains for the antennas

information on calibrators

calibrators

4 absolute amplitude calibrator know

3C147, 3C48, 3C286 (~few percent polarized), 1934-638

There are initiatives to increase the number of absolute amplitude calibrators

phase-calibrator should be a point source !

calibrator data bases

VLA - http://www.aoc.nrao.edu/~gtaylor/csource.html

NVSS - http://www.cv.nrao.edu/nvss/

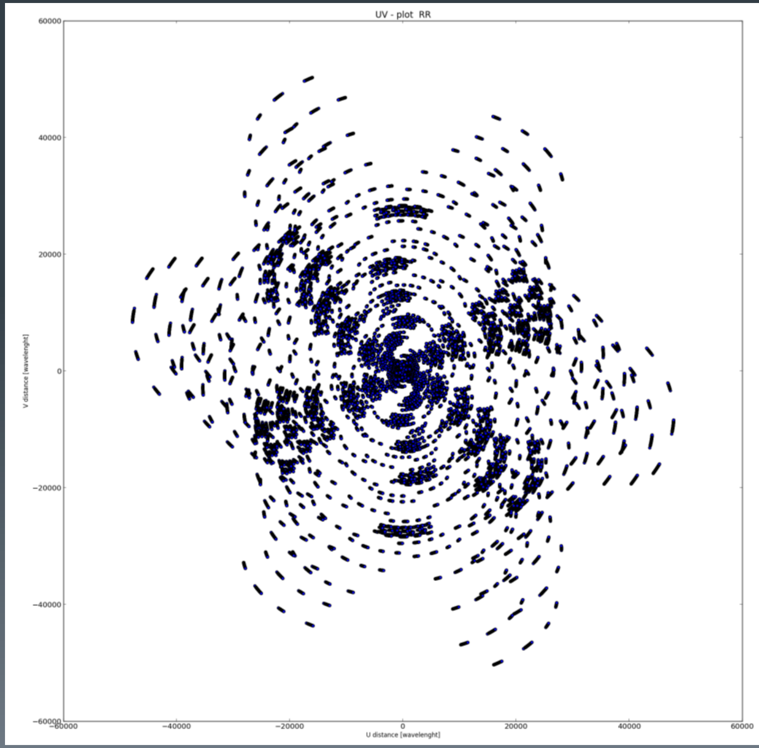
VLBA - http://www.vlba.nrao.edu/astro/calib/index.shtml

fring finder - http://www.aoc.nrao.edu/~analysts/vlba/ffs.html

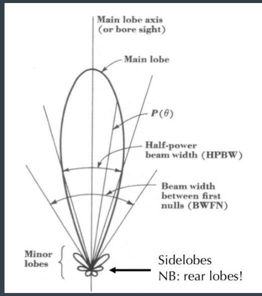
imaging

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

$$I_\nu^D(l,m) = FT(\Pi(u,v)) \bigotimes I_\nu(l,m)$$



$A_\nu(l,m)$ primary beam



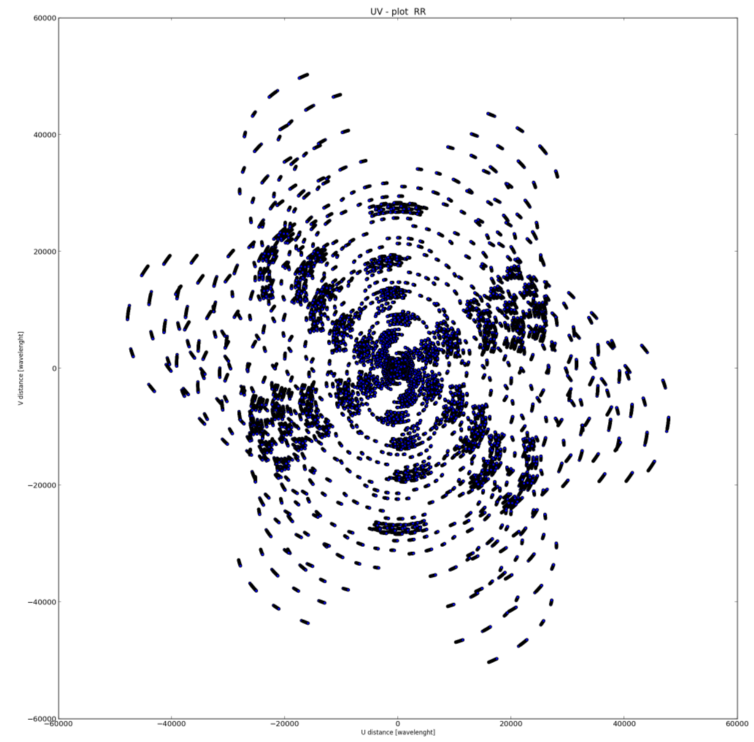
UV coverage defines when

$$V_\nu(u,v,w) = 0$$

$$V_\nu(u,v,w) = 1$$

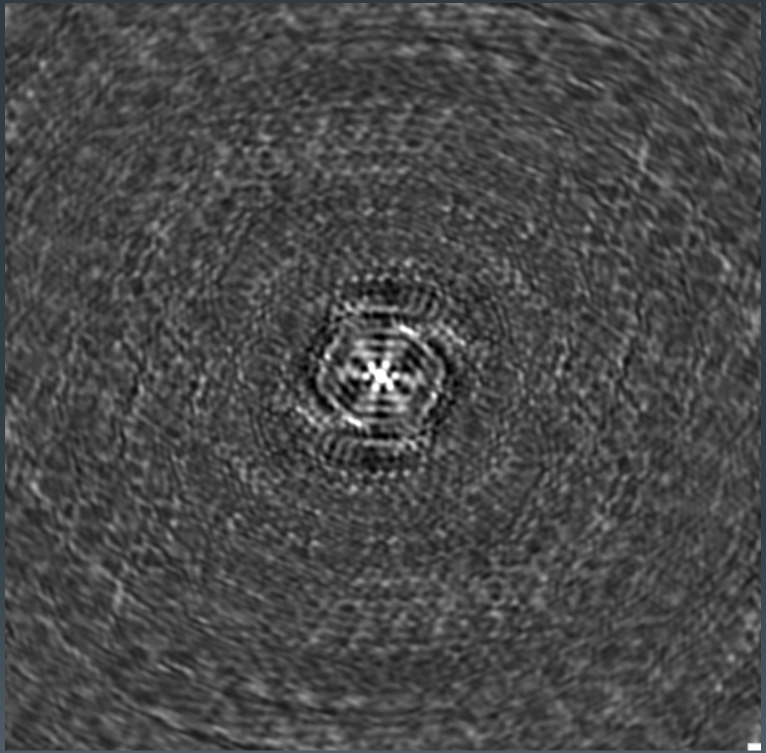
UV coverage sampling function of the sky

UV coverage



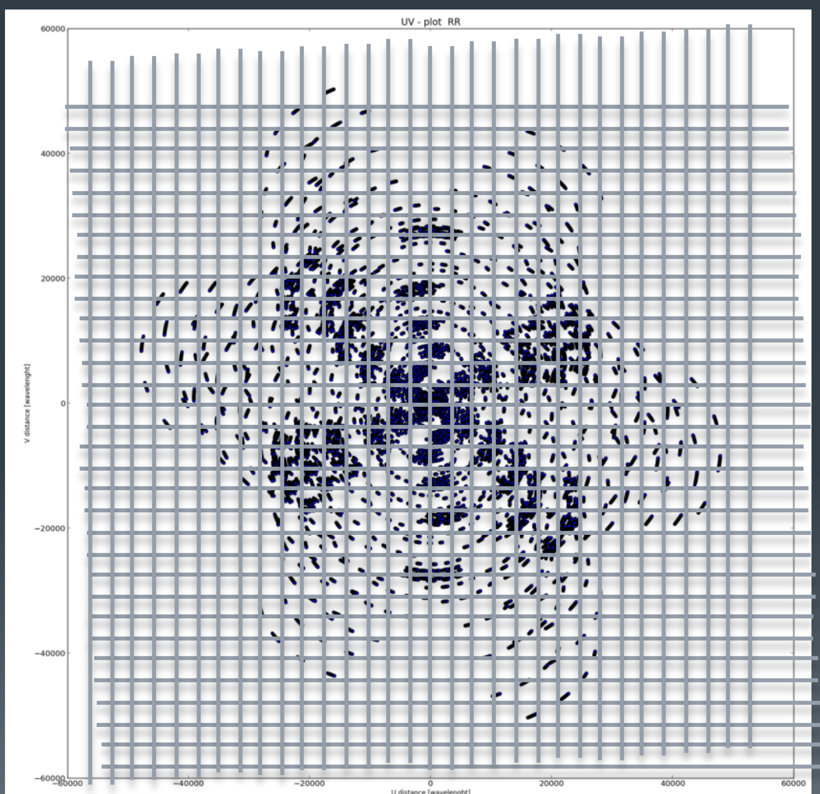
Fourier Transformation

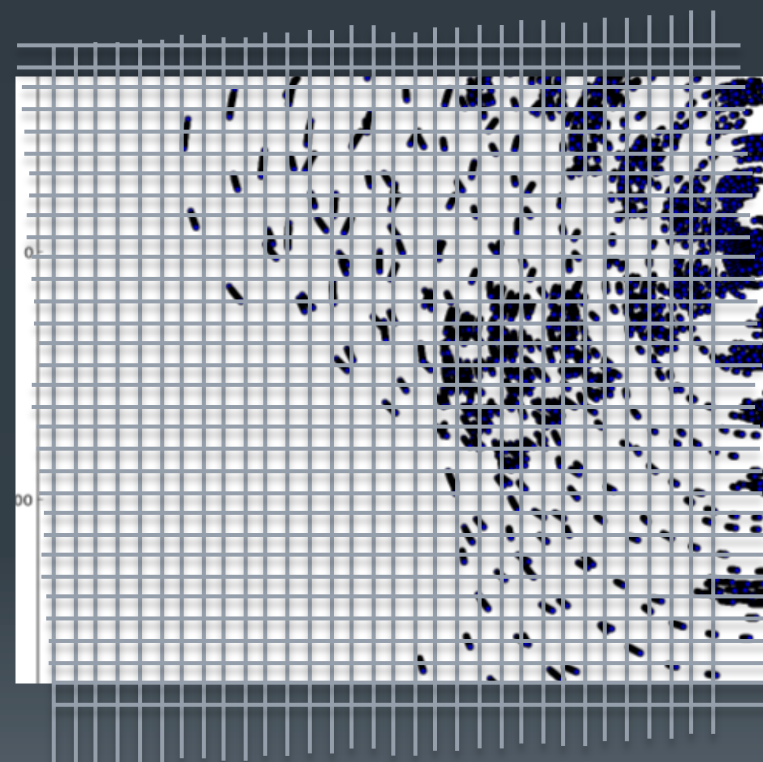
synthesized beam (dirty beam)



data weighting

using FFT and gridding the data can use different weighting



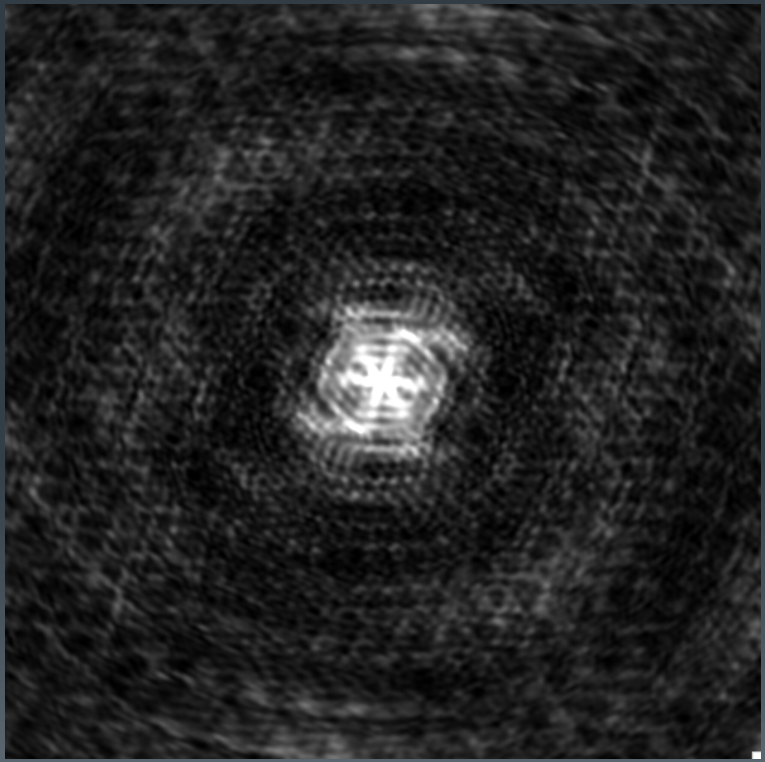


natural weighting use numbers of visibility points per bin

uniform weighting use only 1 points per bin

weighting

natural weighting



uniform weighting

example with the synthesized beam (dirty beam)

analyse interferometer data

UV plane analysis

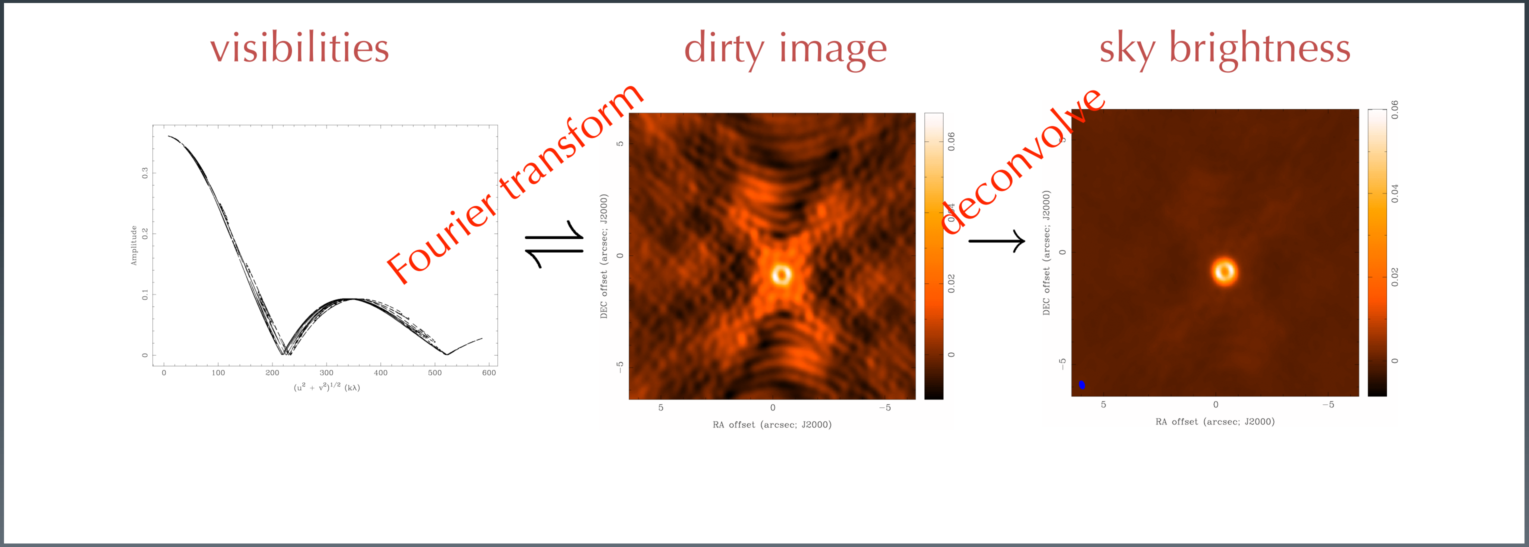
good for simple sources , point sources, doubles sources, disks

image plane analysis

difficult to do science on the dirty image

deconvolve $I^D_\nu(l,m)$ with dirty beam to determine model of $I_\nu(l,m)$

$$I^D_\nu(l,m) = FT(\Pi(u,v)) \bigotimes I_\nu(l,m)$$



deconvolution

- Deconvolution:
  - uses non-linear techniques effectively interpolate/extrapolate samples of $V(u,v)$ into unsampled regions of the (u,v) plane
  - aims to find a **sensible** model of $I(x,y)$ compatible with data
  - requires *a priori* assumptions about $I(x,y)$

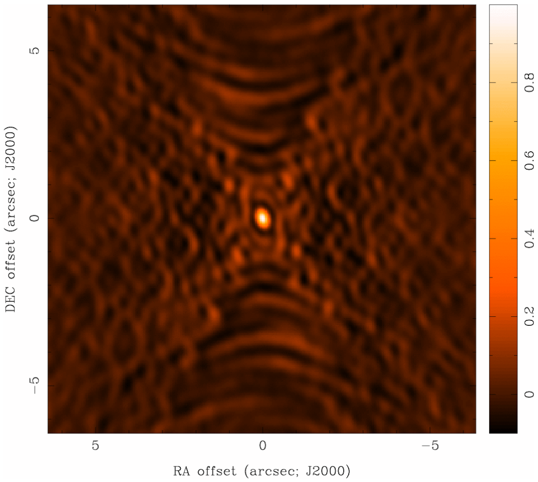
- CLEAN (Högbom 1974) is most common algorithm in radio astronomy
  - *a priori* assumption: $I(x,y)$ is a collection of point sources
  - variants for computational efficiency, extended structure

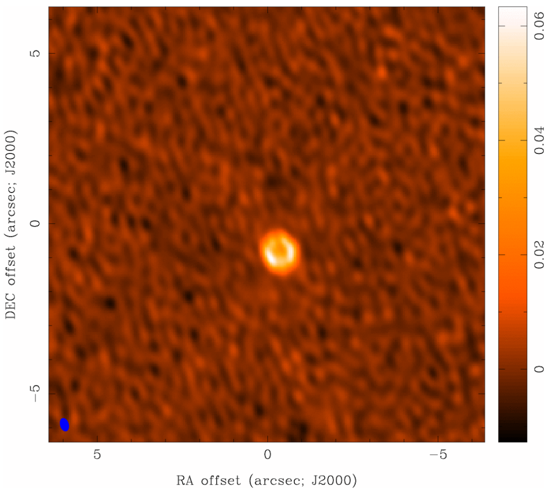
- deconvolution requires knowledge of beam shape and image noise properties (usually OK for aperture synthesis)
  - atmospheric seeing can modify effective beam shape
  - deconvolution process can modify image noise properties

basic clean algorithm

cleaning is chipping the dirty brightness distribution

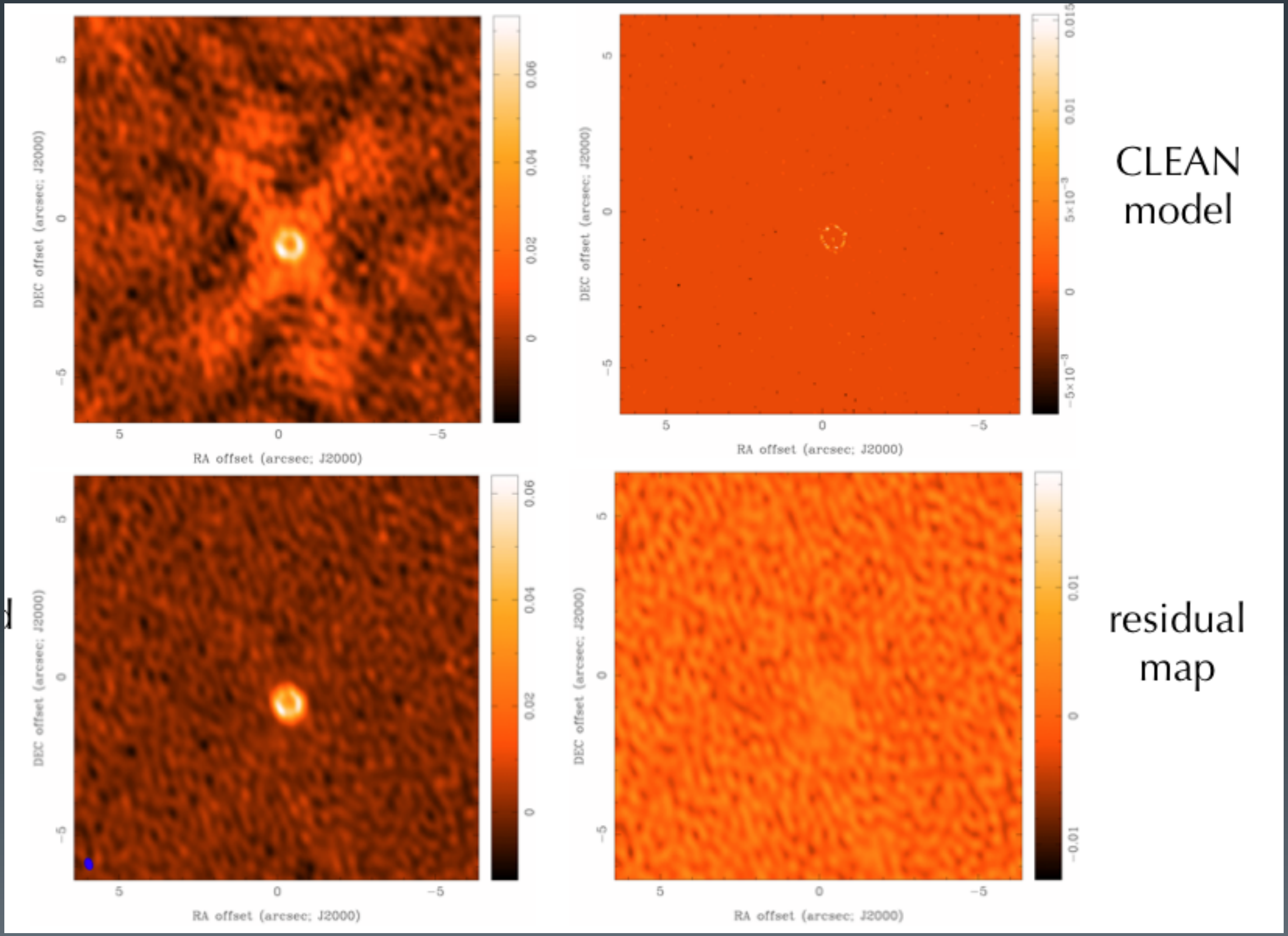
1. Initialize
   - a *residual* map to the dirty map
   - a *CLEAN component* list
2. Identify strongest feature in *residual* map as a point source
3. Add a fraction *g* (the loop gain) of this point source to the clean component list (*g* ~ 0.05-0.3)
4. Subtract the fraction *g* times *b*(*x*,*y*) from *residual* map
5. If stopping criteria* not reached, go back to step 2 (an iteration), or…
6. Convolve CLEAN component (cc) list with an estimate of the main dirty beam lobe (i.e., the "CLEAN beam") and add *residual* map to make the final "restored" image





* Stopping criteria = N x rms (if noise limited), or I^{max}/N (if dynamic range limited), where N is some arbitrarily chosen value

deconvolution



sensitivity image / baseline

[Wrobel & Walker; Chapter 9, Synthesis Imaging in RADIO ASTRONOMY II]

System Equivalent Flux Density

$$SEFD = \frac{T_{\mathrm{sys}}}{K} \qquad K = (\eta_{\mathrm{a}}\, A)\,/\,(2\, k_{\mathrm{B}})$$

equ. (9-5)

A = Area
K_b= Boltzmann
η_a = efficiency
η_s = losses in electronics
τ_{acc} = integration time [s]
$\Delta\nu$ = bandwidth [Hz]

Baseline sensitivity for one polarization

$$\Delta S_{\mathrm{ij}} = \frac{1}{\eta_{\mathrm{s}}} \sqrt{\frac{T_{\mathrm{sysi}}\, T_{\mathrm{sysj}}}{2\, \Delta\nu\, \tau_{\mathrm{acc}}\, K_{\mathrm{i}}\, K_{\mathrm{j}}}} \qquad (9\text{–}13)$$

or in terms of the $SEFD$s defined in Equation 9–5:

$$\Delta S_{\mathrm{ij}} = \frac{1}{\eta_{\mathrm{s}}} \sqrt{\frac{SEFD_{\mathrm{i}}\, SEFD_{\mathrm{j}}}{2\, \Delta\nu\, \tau_{\mathrm{acc}}}} \qquad (9\text{–}14)$$

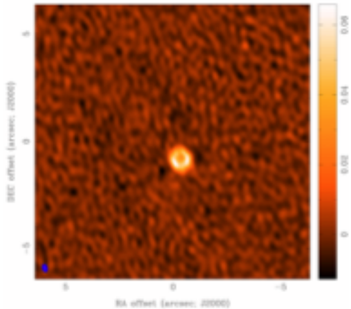
Image sensitivity for one polarization

$$\Delta I_{\mathrm{m}} = \frac{1}{\eta_{\mathrm{s}}} \frac{SEFD}{\sqrt{N\,(N-1)\, \Delta\nu\, t_{\mathrm{int}}}}. \qquad (9\text{–}23)$$

Divide by square root 2 for 2 polarization !

image quality measures

- *"dynamic range"*
  - ratio of peak brightness to rms noise in a region void of emission (common in astronomy)
  - an easy to calculate lower limit to the error in brightness in a non-empty region



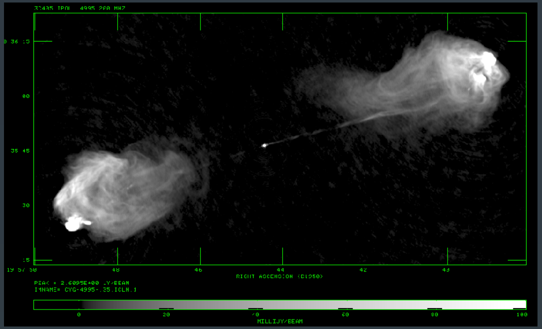
- *"fidelity"*
  - difference between any produced image and the correct image
  - a convenient measure of how accurately it is possible to make an image that reproduces the brightness distribution on the sky
  - need a priori knowledge of correct image to calculate

  - fidelity image = input model / difference
  - fidelity is the inverse of the relative error

model

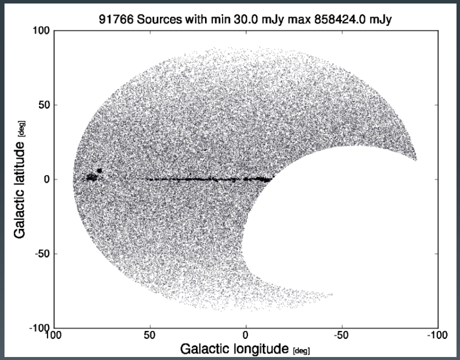
easy brightness distribution modelled by point sources or Gaussian

complicated brightness distribution modelled by number of Gaussian or use wavelet components



large fields of view one need the local sky model

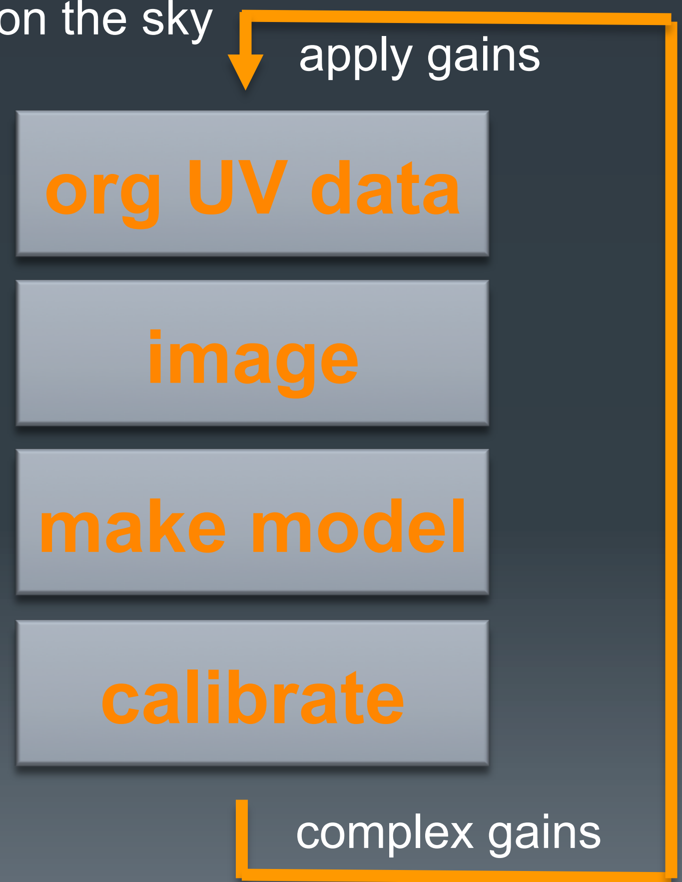
In case you use catalogued source from e.g. the NVSS you need to decrease the flux densities with respect to the phase centre or in other word careful the interferometer sees the local sky convolved with the primary beam



to detect a model component on a single baseline assume 6-8 sigma with respect to the rms of the baseline

self-calibration

self-calibration is an iterative procedure to determine the complex gains to calibrate the visibilities by iteratively improving the brightness distribution model on the sky



the model is based on the cleaned image

number of clean components

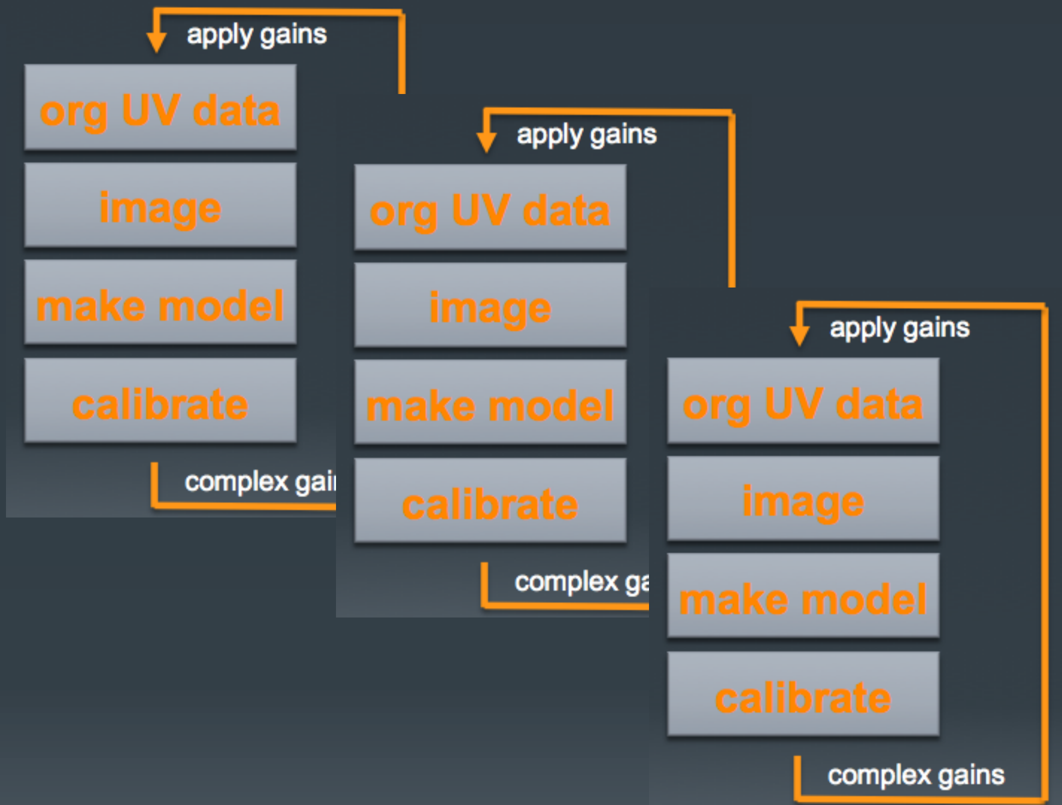
averaged time interval
to determine gain solutions

phase	use minimum 3 antennas
amplitude	use minimum 4 antennas

use dynamic range or rms as criteria to stop the self-calibration process

Caution lose absolute phase from calibrators and therefore the position

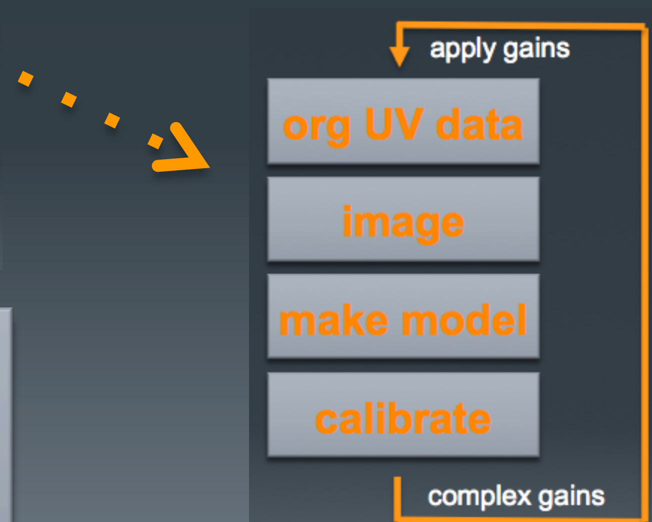
self-calibration cycles



30 clean components

calibrate amplitude

1 minutes
integration time



100 clean components

calibrate phase

60 minutes
integration time

80 clean components

calibrate phase

1 minutes
integration time

calibration generations

closure phases and amplitudes

calibration assume antenna based errors

directional dependent calibration

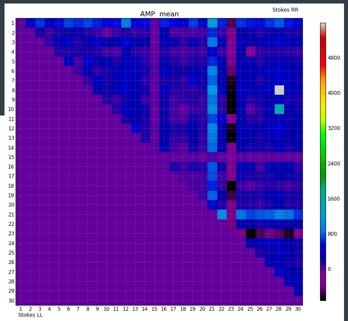
- peeling – subtract all sources except the most strongest one, self-cal on source, use final model of this source to subtract out of the database
- measurement equation – MeqTree or CASA

measurement equation

directional dependent calibration

Need of a good model of the sources brightness distribution within the LSM and of the directional dependent parameter of the interferometer and the single antennas

all can be written as a matrix



$$V_{pq} = G_p \left(\sum_{k=1}^{N} E_{pk} X_k E_{qk}^{\dagger} \right) G_q^{\dagger}$$

where V_{pq} is the 2×2 visibility (also called *coherency*, or *uv-data*) matrix measured by the interferometer formed by stations p and q. The sum is taken over the contributions X_k from N discrete sources in the field, at positions l_k, m_k.'

Hamaker et al. 1996

measurement equation

In the 2×2 **signal domain**, the electric field vector $\vec{E}$ of the incident plane wave can be represented either in a linear polarisation coordinate frame (x, y) or a circular polarisation coordinate frame (r, l). Jones matrices are linear operators in the chosen frame:

$$\vec{V}_{\mathsf{i}}^{+} = \begin{pmatrix} \mathsf{v_{ip}} \\ \mathsf{v_{iq}} \end{pmatrix} = \mathsf{J}_{\mathsf{i}}^{+} \begin{pmatrix} \mathsf{e_x} \\ \mathsf{e_y} \end{pmatrix} \qquad or \qquad \vec{V}_{\mathsf{i}}^{\odot} = \mathsf{J}_{\mathsf{i}}^{\odot} \begin{pmatrix} \mathsf{e_r} \\ \mathsf{e_l} \end{pmatrix} \tag{13}$$

For linear polarisation coordinates, equation 1 becomes:

$$\vec{V}_{\mathsf{ij}}^{+} = (\mathsf{J}_{\mathsf{i}}^{+} \otimes \mathsf{J}_{\mathsf{j}}^{+*})\,(\vec{E} \otimes \vec{E}^{*}) = (\mathsf{J}_{\mathsf{i}}^{+} \otimes \mathsf{J}_{\mathsf{j}}^{+*}) \begin{pmatrix} \mathsf{e_x e_x^*} \\ \mathsf{e_x e_y^*} \\ \mathsf{e_y e_x^*} \\ \mathsf{e_y e_y^*} \end{pmatrix} = (\mathsf{J}_{\mathsf{i}}^{+} \otimes \mathsf{J}_{\mathsf{j}}^{+*})\ \mathsf{S}^{+}\ \vec{I}(\mathsf{l}, \mathsf{m}) \tag{14}$$

and there is a similar expression for circular polarisation coordinates. Thus, as emphasised in [2], the Stokes vector $\vec{I}(\mathsf{l}, \mathsf{m})$ and the coherency vector $\vec{V}_{\mathsf{ij}}$ represent the same physical quantity, but in different abstract coordinate frames. A 'Stokes matrix' S is a coordinate transformation matrix in the 4×4 **coherency domain**: S^{+} transforms the *representation* from Stokes coordinates (I,Q,U,V) to linear polarisation coordinates $(\mathsf{xx}, \mathsf{xy}, \mathsf{yx}, \mathsf{yy})$. Similarly, $\mathsf{S}^{\odot}$ transforms to circular polarisation coordinates $(\mathsf{rr}, \mathsf{rl}, \mathsf{lr}, \mathsf{ll})$. Following the convention of [4], we write:[3]

$$\mathsf{S}^{+} = \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & i \\ 0 & 0 & 1 & -i \\ 1 & -1 & 0 & 0 \end{pmatrix} \qquad \mathsf{S}^{\odot} = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & i & 0 \\ 0 & 1 & -i & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix} \tag{15}$$

Noordam 1996 aips++ Memo 185

measurement equation

$$\mathsf{J_i} \;=\; \mathsf{G_i}\;[\mathsf{H_i}]\;[\mathsf{Y_i}]\;\mathsf{B_i}\;\mathsf{K_i}\;\mathsf{T_i}\;\mathsf{F_i} \;=\; \mathsf{G_i}\;[\mathsf{H_i}]\;[\mathsf{Y_i}]\;(\mathsf{D_i}\;\mathsf{E_i}\;\mathsf{P_i})\;\mathsf{K_i}\;\mathsf{T_i}\;\mathsf{F_i}$$

in which

$\mathsf{F_i}(\vec{\rho}, \vec{r_i})$	ionospheric Faraday rotation
$\mathsf{T_i}(\vec{\rho}, \vec{r_i})$	atmospheric complex gain
$\mathsf{K_i}(\vec{\rho}.\vec{r_i})$	factored Fourier Transform kernel
$\mathsf{P_i}$	projected *receptor* orientation(s) w.r.t. the sky
$\mathsf{E_i}(\vec{\rho})$	voltage primary beam
$\mathsf{D_i}$	position-independent *receptor* cross-leakage
$[\mathsf{Y_i}]$	commutation of *IF-channels*
$[\mathsf{H_i}]$	hybrid (conversion to circular polarisation coordinates)
$\mathsf{G_i}$	electronic complex gain (*feed*-based contributions only)

CAUTION matrices do NOT commute
the order of each matrix has a physical reason

measurement equation matrices

The following matrices and vectors play a role in the Measurement Equation:

$\vec{I}$	`\vvIQUV`	Stokes vector of the source (I,Q,U,V).
$\vec{V}$, v	`\vvCoh,\vvCohEl`	Coherency vector, and one of its elements.
S	`\mmStokes`	Stokes matrix, conversion between polarisation representations.
S^{+}	`\mmStokes\ssLin`	Conversion to linear representation.
$\mathsf{S}^{\odot}$	`\mmStokes\ssCir`	Conversion to circular representation.
$\mathcal{M}$	`\mmMueller`	Mueller matrix: Stokes to Stokes through optical 'element'
X, x	`\mmXifr,\mmXifrEl`	Correlator matrix (4×4).
M, m	`\mmMifr,\mmMifrEl`	Multiplicative *interferometer*-based gain matrix (4×4).
$\vec{A}, \mathsf{a}$	`\vvAifr,\vvAifrEl`	Additive *interferometer*-based gain vector.

measurement equation matrices

The following *feed*-based Jones matrices (2 × 2) have a well-defined meaning:

J, j	\mjJones,\mjJonesEl	Jones matrix, and one of its elements.
F, f	\mjFrot,\mjFrotEl	Faraday rotation (of the plane of linear pol.)
T, t	\mjTrop,\mjTropEl	Atmospheric gain (refraction, extinction).
P, p	\mjProj,\mjProjEl	Projected *receptor* angle(s) w.r.t. x, y frame
B, b	\mjBtot,\mjBtotEl	Total *feed* voltage pattern (i.e. B = D E P.
E, e	\mjBeam,\mjBeamEl	Traditional *feed* voltage beam.
C, c	\mjConf,\mjConfEl	Feed configuration matrix (...).
D, d	\mjDrcp,\mjDrcpEl	Leakage between *receptors* a and b.
H, h	\mjHybr,\mjHybrEl	Hybrid network, to convert to circular pol.
G, g	\mjGrec,\mjGrecEl	*feed*-based electronic gain.
K, k	\mjKern,\mjKernEl	Fourier Transform Kernel (baseline phase weight)
K^0, k^0	\mjKref,\mjKrefEl	FT kernel for the fringe-stopping centre.
K', k'	\mjKoff,\mjKoffEl	FT kernel relative to the fringe-stopping centre.
Q, q	\mjQsum,\mjQsumEl	Electronic gain of tied-array *feed* after summing.

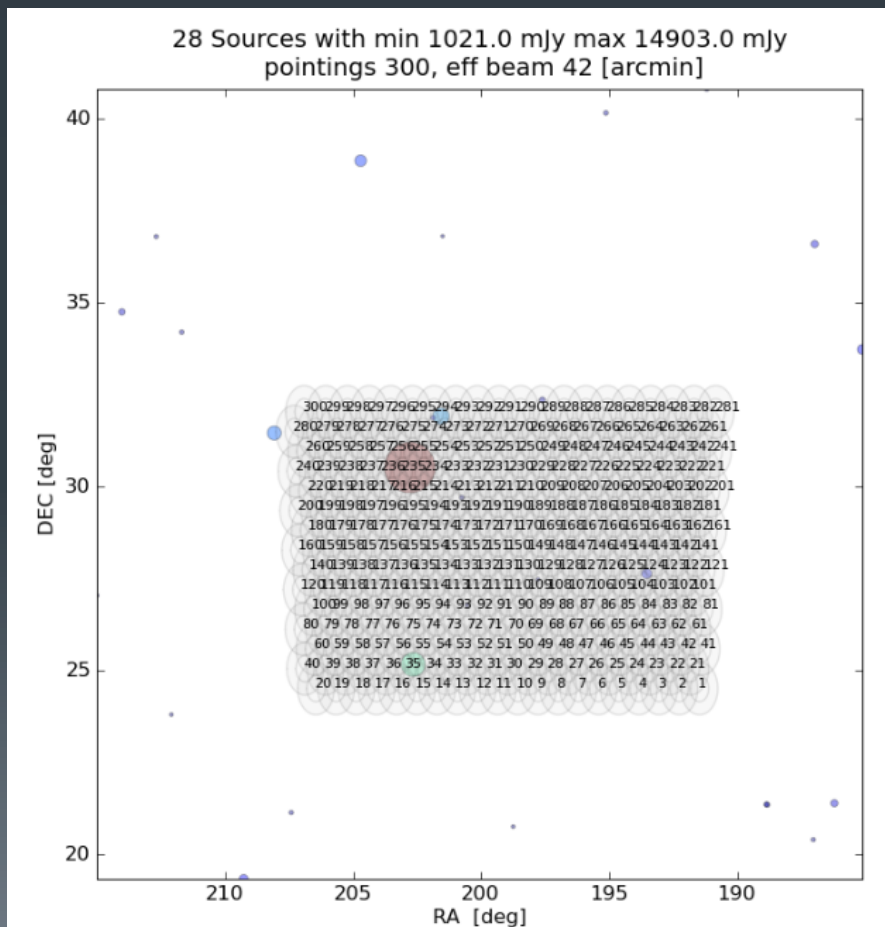
Some special matrices and vectors:

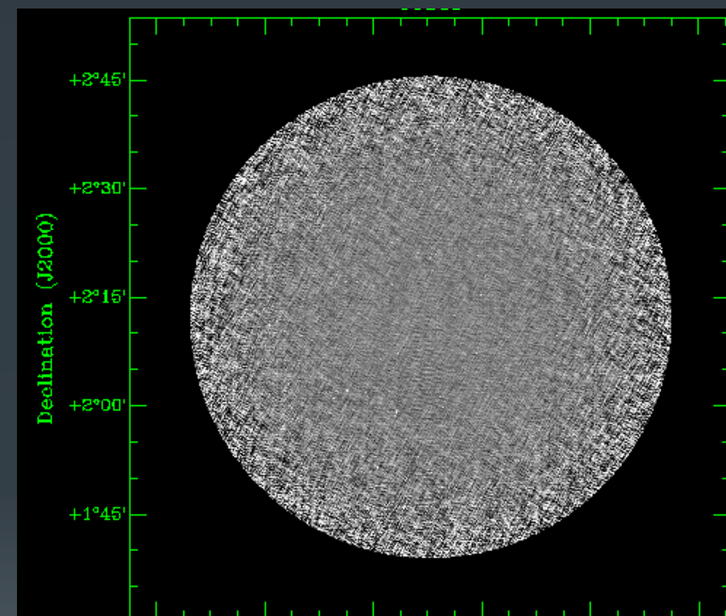
Zero	\mmZero	Zero matrix
$\vec{0}$	\vvZero	Zero vector
$\mathcal{U}$	\mmUnit	Unit matrix
Diag(a, b)	\mjDiag	Diagonal matrix with elements a, b
Mult(a)	\mjMult	Multiplication with factor a
Rot$(\alpha[, \beta])$	\mjRot	[pseudo] Rotation over an angle α, β
Ell$(\alpha[, \beta])$	\mjEll	Ellipticity angle[s] α, β

so why do we need that again

going deeper in sensitivity implies that effects need to be modelled which have been ignored so far

strong sources
far from the phase
centre



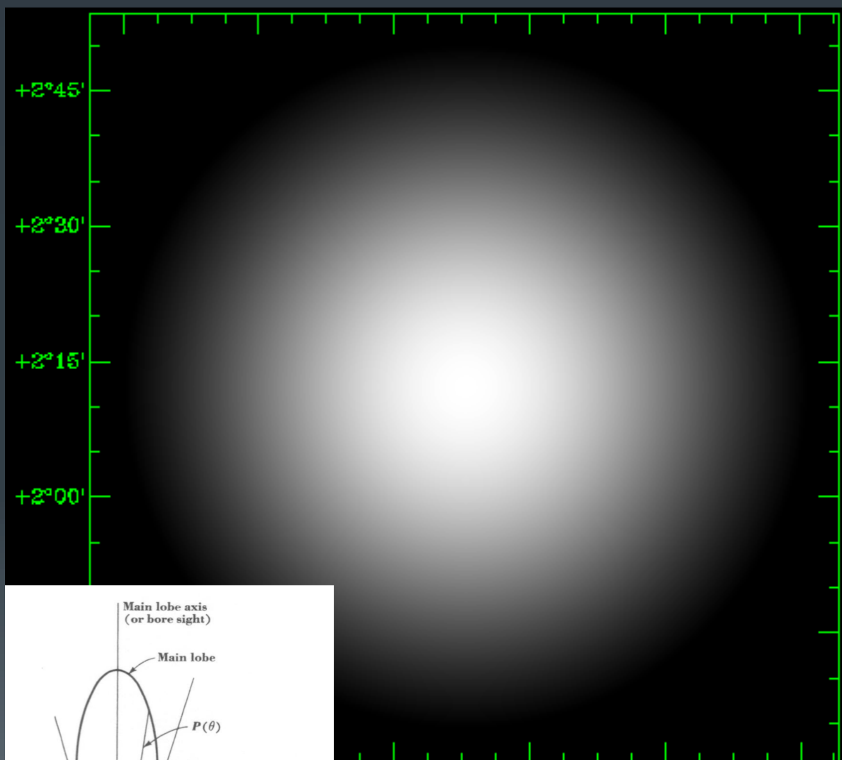


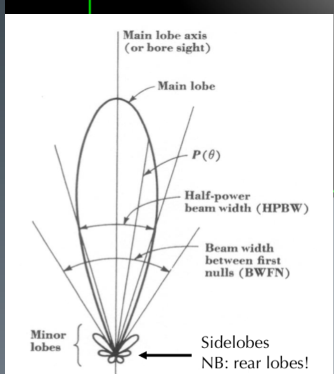
so why do we need that again

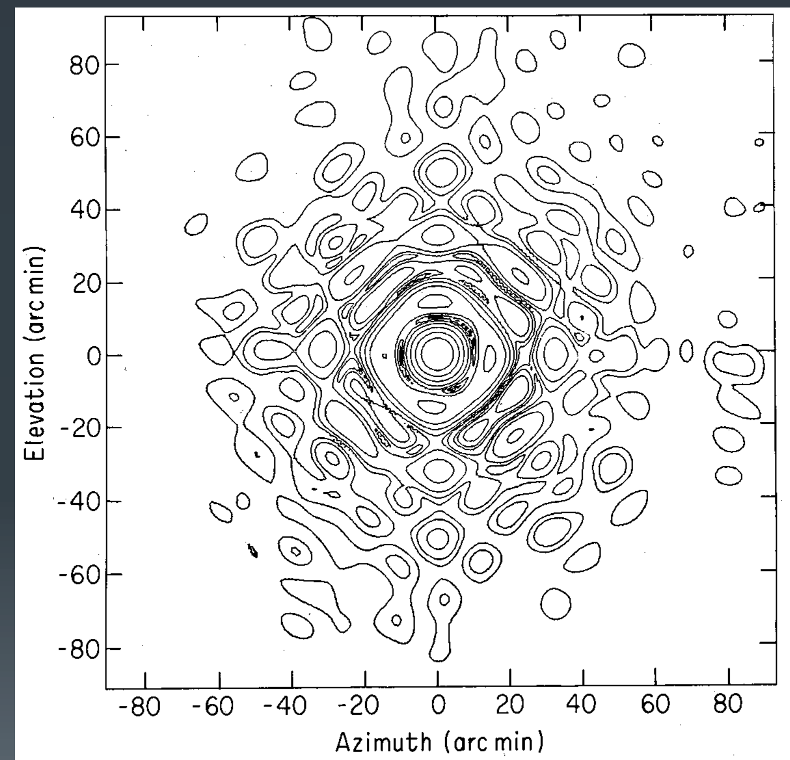
model of the interferometer needs to be more realistic

primary beam

note that the primary beam is frequency dependent
old software packages are not able to model this







The effect will change e.g. source flux densities adding a systematic error if you do surveying