Radio Astronomy Lecture 2: Interferometry

Cameron Van Eck
19 Mar 2020

- Review of Fourier transforms
- correlation of signals
- complex visibility, the uv-plane; Earth-rotation synthesis and snapshot synthesis
- imaging, ~~weights, deconvolution (CLEAN), sensitivity, missing scales~~
- ~~calibration~~
- ~~Actual interferometers~~

This took more time than expected, so lecture 2 covered up to and including imaging. Remaining topics will be moved to lecture 3 (replacing some of the applications of radio astronomy)

Fourier transform

$$X(\omega) = \mathcal{F}\{x(t)\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x(t) e^{-i\omega t} dt$$

$$x(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} X(\omega) e^{i\omega t} d\omega$$

- Fourier transforms pop up all over in radio astronomy (for example, beam of a dish/phased array is Fourier transform of 'illumination function'). The heart of interferometry is the Fourier transform.
- Many properties of Fourier transforms can be used to understand behaviour of the system.

There are variations on the FT: using frequency instead of angular frequency, with different conventions for where to put the 2pi factors. None of it really matters except for scaling.

Fourier properties

- Always two variables, a Fourier (variable) pair, with opposite units. Time and frequency, position and wavenumber, angular position and number of wavelengths, wavelength squared and Faraday depth.
- In general, input and output are both complex. Certain symmetries affect this.
- Can be extended to multiple dimensions. The sky is 2D and has a corresponding 2D Fourier transform.*

*: when only dealing with small parts of the sky, then we can treat the sky as effectively flat and normal Fourier formalism is fine. Going to larger scales starts to mess with this, in ways that can be dealt with (w-projection, etc). Looking at the whole sky becomes problematic, in that it makes the math a lot harder.

Fourier properties

Domain A	Domain B	Description
x(t) + y(t)	X(ω) + Y(ω)	Linearity
a*x(t)	a*X(ω)	Linearity
Real	Hermitian, X(ω)=X(-ω)*	Symmetry: to make a real function all imaginary parts of FT must have equal and opposite counterpart
Real and even	Real and even	Even symmetry implies made of cosines only (sine/imaginary components all equal 0)

All of these theorems also apply in reverse.

Fourier properties

Domain A	Domain B	Description
x(a*t)	X(ω/a) / |a|	Scaling: narrower in one domain makes other domain wider
x(t-τ)	$X(\omega)e^{-i\omega\tau}$	Shift theorem: Shift in one domain results in phase shifts in other domain.
x(t) ★ y(t)	X(ω)Y(ω)	Convolution theorem: Convolution of two functions corresponds to product of Fourier transforms.

Common Fourier (transform) pairs

Domain A	Domain B	Description
δ(t)	1	Delta function has equal signal on all frequencies. OR: constant signal corresponds to zero frequency.
cos(at)	(δ(ω-a)+ δ(ω+a)) /2	Cosine is equal parts positive and negative frequency
rect(t)	sin(ω)/ω	Tophat/rectangle transforms to sinc function
Circular aperature	Airy disk (Bessel function)	Classic circular mirror/dish beam.
Gaussian(t)	Gaussian(ω)	Gaussian is it's own FT (subject to scaling, shift theorems) Also applies in higher dimensions.

Interferometry: Motivation

- A phased array (or dish) requires you to choose what direction to have peak sensitivity in, then returns the flux density in that direction (sky convolved with beam). Imaging the sky means shifting the peak direction all over the sky.
- Also, emission from off-target acts as noise (contributing flux density in sidelobes).
- What if you could have a system that just told you where the emission was, as an output rather than as an input? And treated all directions as signal, simultaneously?
- An interferometer gives you an (incomplete) answer to where the emission is, for each baseline (pair of antennas); combining many baselines improves the localization.

Correlation of signals

- Consider the (monochromatic) voltage signals coming off of two telescopes, with some arbitrary phase difference:

$$V_1 = V\cos(\omega t) \qquad V_2 = V\cos(\omega t + \phi)$$

- Correlate them together (multiply and average/integrate):

$$\frac{1}{T}\int V_1 V_2 dt = \frac{1}{T}\int \frac{V^2}{2}\left(\cos(2\omega t) + \cos(\phi)\right) dt = \frac{V^2}{2}\cos(\phi)$$

Two-element interferometer

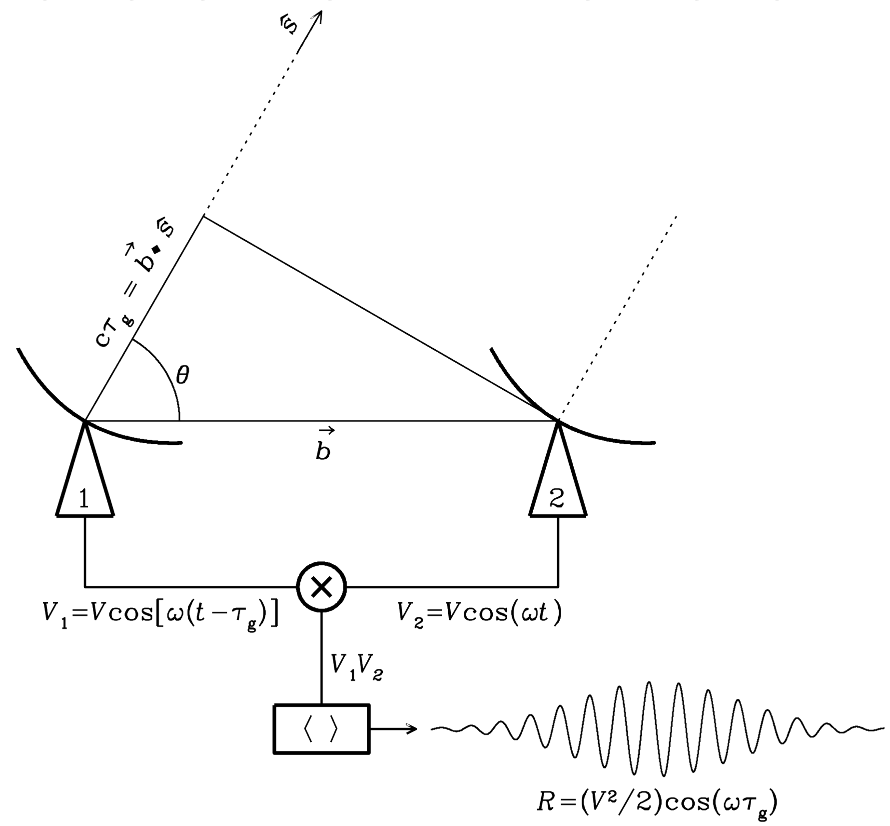


Image credit: Fig 3.41 from Essential Radio Astronomy

Mixing (quasi-monochromatic) signals from two antennas yields a result proportional to voltage squared (power), and depending on relative phase (from the geometric delay). Long delays move outside the quasi-monochromatic correlation time, so the correlation drops to zero. Delay lines correct for this, to put it back at the peak of the function, but delay line corrects for **known/input** position (more on this later). In-phase/anti-phase produce strong positive/negative output; 90deg out of phase gives zero voltage; thus the output tells you the fraction of a wavelength but not the integer number of wavelengths.

Two-element interferometer

- Output voltage encodes information on power (flux) and position, but there are ambiguities
- Add pi/2 phase shift to one of the signals, then correlate again: cos(ϕ) becomes sin(ϕ)
- Do both, and combine as a complex number:

$$\mathcal{V} = \frac{V^2}{2}(\cos(\phi) - i\sin(\phi)) = A\,e^{-i\phi}$$

$$\phi = 2\pi(\vec{b}\cdot\hat{s})/\lambda \quad \mathrm{mod}\ 2\pi$$

Ambiguities: every even multiple of pi gives positive correlation, every odd gives negative correlation, and every half-integer gives zero.
The bottom equation defines the **complex visibility**, the combined output of a cosine and a sine correlator, which contains the full information on the correlation of the two waves.

Two-element interferometer

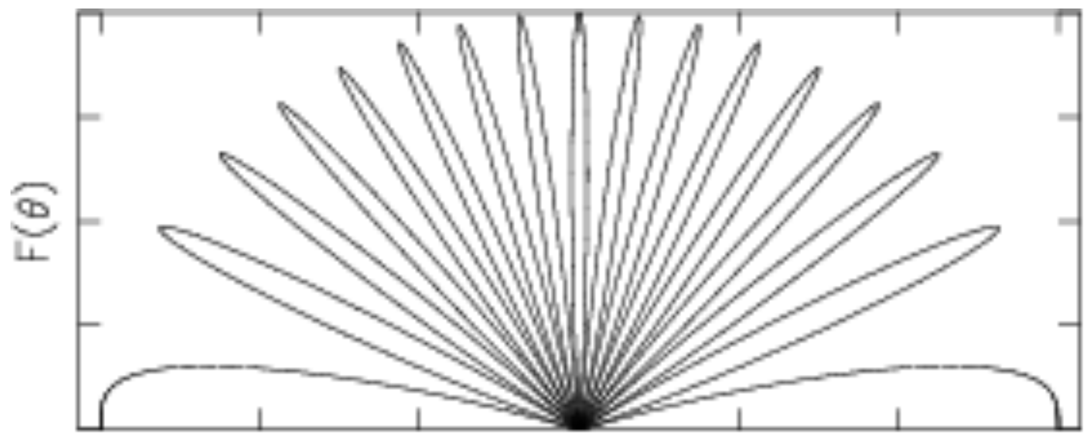


Image credit: Fig 2.2 from Rodrigo Parra's website

Consider a measurement with zero phase difference (ϕ = 0). It could be that the path difference is zero (at the zenith). Or it could be ±2π. Or ±4π. Or 6π. And so on. You end up with this kind of distribution of locations on the sky. The spacing of the fringes depends on the length of the baseline (in wavelengths).
Note that the spacing is narrower towards the zenith compared to the horizon: this will be significant soon.

(3D) Two-element interferometer

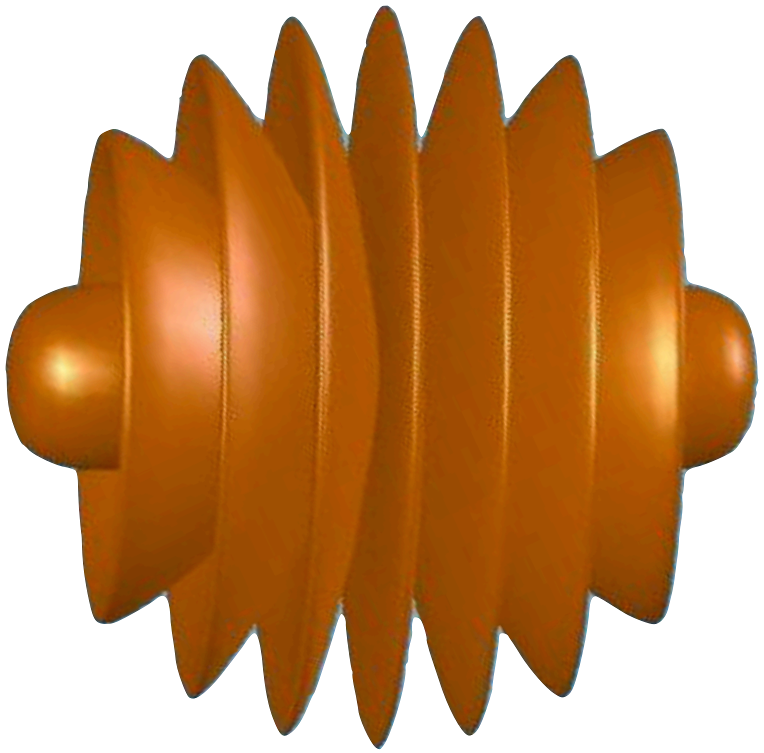
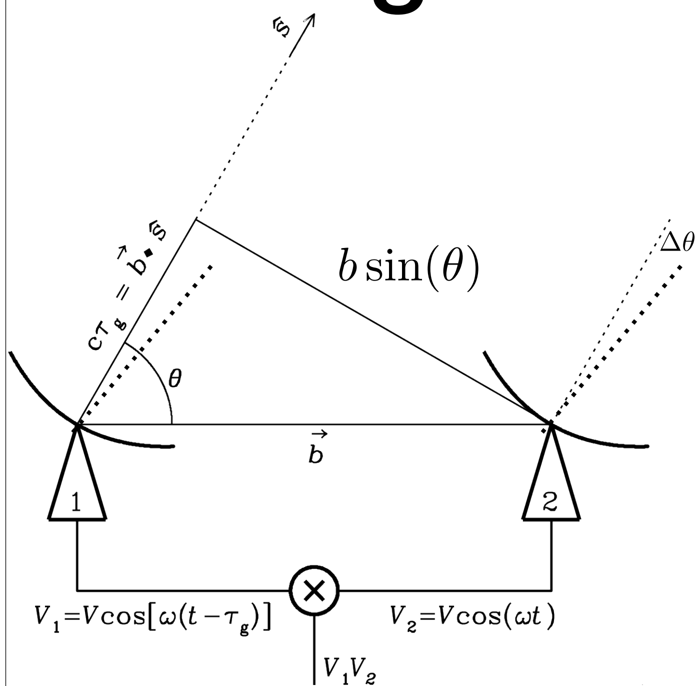


Image credit: Rick Perley (NRAO), 14th Synthesis Imaging workshop lectures

Bear in mind that the world is not 2D. Each possible location is actually a ring perpendicular to the baseline vector.
How to resolve the position ambiguity? Have multiple baselines, with different lengths and orientations. Then only one position will line up in all baselines.

Angular resolution



$$\phi = 2\pi \vec{b} \cdot \hat{s}/\lambda = 2\pi \frac{b}{\lambda} \cos(\theta)$$

$$\Delta\phi = -2\pi \frac{b}{\lambda} \sin(\theta) \Delta\theta$$

- Spacing between two allowed locations:

$$\Delta\theta = \frac{\lambda}{b\sin(\theta)} = \frac{\lambda}{\text{“}D\text{”}}$$

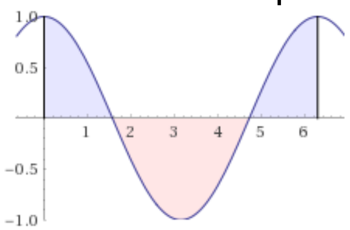
Image credit: Fig 3.41 from Essential Radio Astronomy

We can define the resolution in terms of the spacing between the ‘fan blades’ in the previous diagram: how does the phase change with small changes in angle? This leads to the slightly counterintuitive result that the best sensitivity is at the zenith (perpendicular to the baseline). We can rewrite this into a resolution equation with an effective ‘diameter’ equal to the projected baseline length.

Extending to the whole sky

$$\mathcal{V} = \int I(\hat{s}) e^{-i2\pi \frac{\vec{b}}{\lambda} \cdot \hat{s}} d\Omega$$

- Simply integrate over position in the sky
- Start to see weird behaviour: what happens when a source is large enough to have emission that is both in-phase and anti-phase?



- Emission features larger than Δθ are **resolved out**: produce no output signal on that baseline

Full 3D formulation

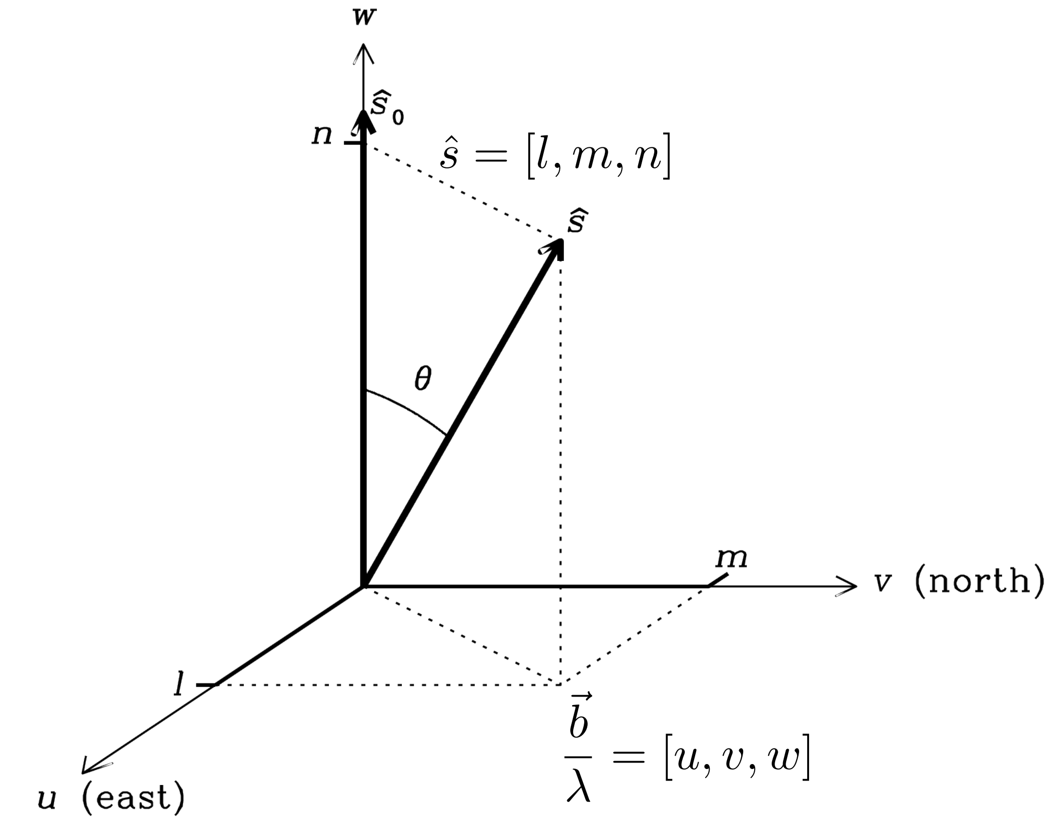


Image credit: Fig 3.45 from Essential Radio Astronomy

Choose some reference direction, s_0, and set up a coordinate system with *w* towards that direction and *u,v* orthogonal to that. These axes will be the space of (possible) baselines, normalized in units of wavelengths (\vec{b}/λ).

We'll want to integrate over all directions, s-hat, which will have components *l,m,n* in these coordinates.

Full 3D formulation

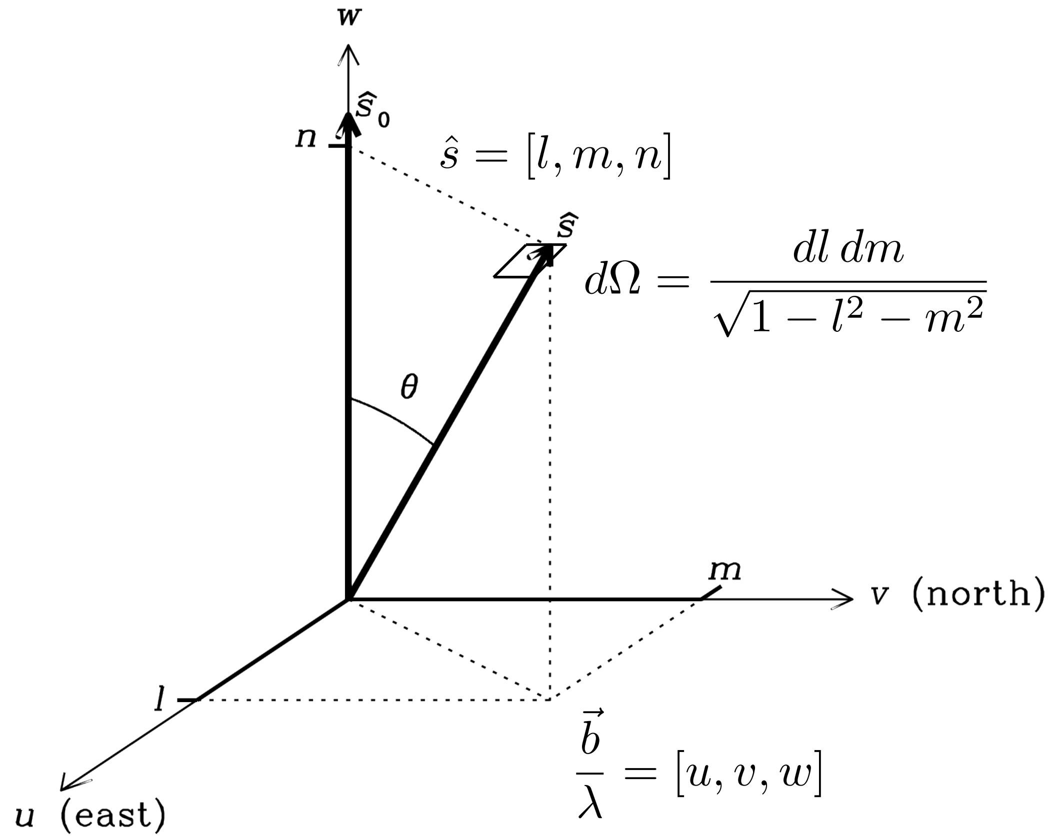


Image credit: Fig 3.45 from Essential Radio Astronomy

Let's take the unusual approach of thinking of the sky as an infinite plane. The infinitesimal solid angle, *d*Ω, in the direction s-hat, can be expressed like this.

Full 3D formulation

$$\mathcal{V} = \int I(\hat{s}) e^{-i2\pi \frac{\vec{b}}{\lambda} \cdot \hat{s}} d\Omega$$

$$\mathcal{V}(u,v,w) = \int\int \frac{I(l,m)}{\sqrt{1-l^2-m^2}} e^{-i2\pi(ul+vm+wn)} dl\, dm$$

$$\mathcal{V}(u,v,w) = \int\int \frac{I(l,m)}{\sqrt{1-l^2-m^2}} e^{-i2\pi(w\sqrt{1-l^2-m^2})} e^{-i2\pi(ul+vm)} dl\, dm$$

- The $\sqrt{1-l^2-m^2}$ terms are essentially corrections for trying to force a Cartesian coordinate system on what is fundamentally a spherical system. If we consider only a small field of view then we can neglect this term.

Start with equation from 2D, expand out the solid angle and dot-product.

Has the form of the 2D Fourier transform (with Fourier pairs *u* and *l*, and *v* and *m*), except for that pesky exp(*wn*) term.

There are methods (*w*-projection) that try to work around this term without completely neglecting it.

Full 3D formulation

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

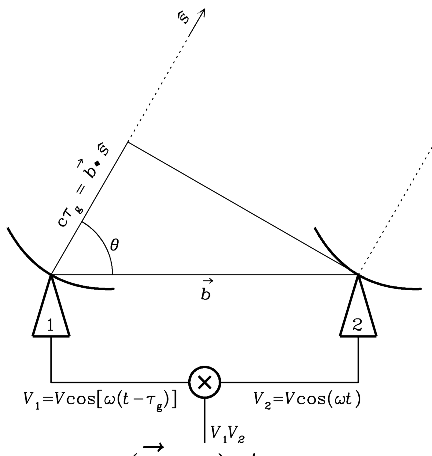
$$I(l,m) = \int\int \mathcal{V}(u,v,w) e^{i2\pi w} e^{i2\pi(ul+vm)} du\, dv$$

- The fundamental nature of radio interferometry: each **visibility** is a measurement of a single component of the **Fourier transform** of the sky, with the **(angular) frequency** corresponding to the length and orientation of the **baseline**. **Inverting** your measured visibilities returns the **image** of the sky.

We did it! We have a method of producing images of the sky from correlated measurements. Well, yes, but no. Note that to get the image we need to integrate over u and v, the space of possible (projected) baselines. So we'd need to have infinitely many baselines in order for this to work.

Symmetry

- Instead of correlating V_1 with V_2, what if we correlated V_2 with V_1?



$$\phi = 2\pi(\vec{b} \cdot \hat{s})/\lambda \mod 2\pi$$

$$\vec{b} \to -\vec{b}$$

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

- *I(l,m)* is real, therefore $\mathcal{V}$ must be Hermitian:

$$\mathcal{V}(-u,-v) = \mathcal{V}(u,v)^*$$

- For every baseline, the opposite baseline has the opposite phase

(u,v) plane and Earth rotation synthesis

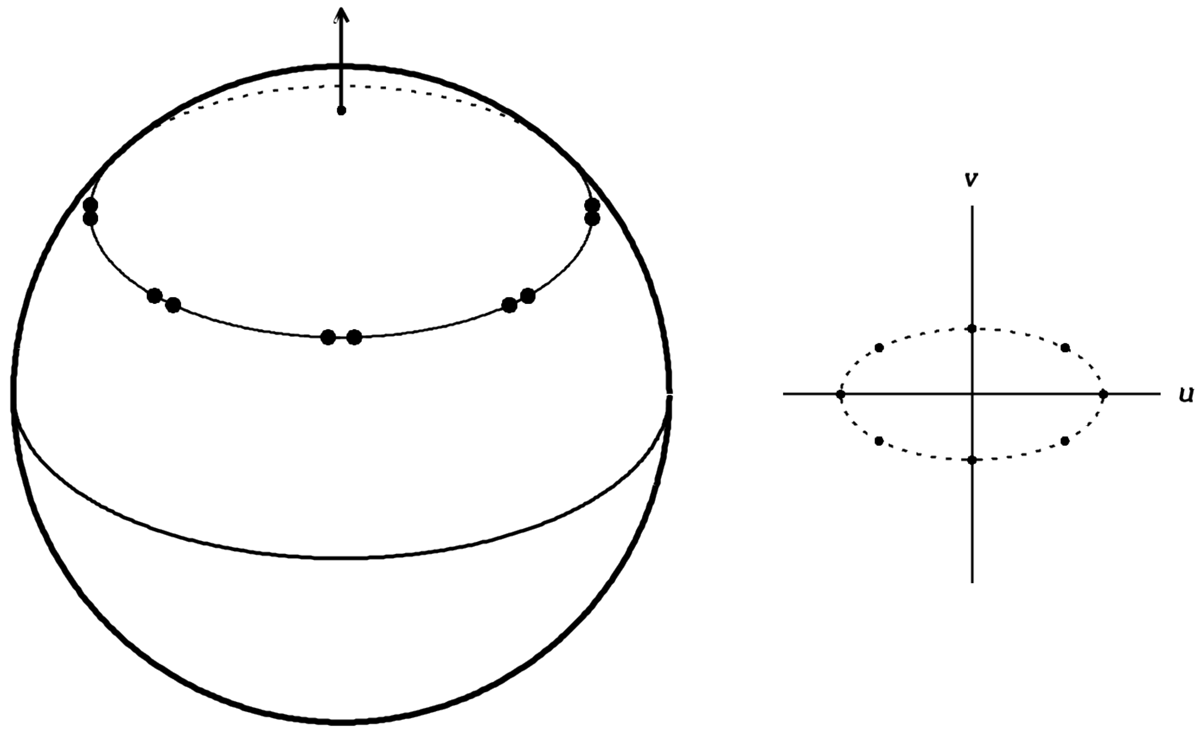


Image credit: Fig 3.44 from Essential Radio Astronomy

The *u,v* plane is the parameter space of baselines (ignoring *w*). Each baseline maps to two positions in the plane (for both ‘directions’ of each baseline). The Earth’s rotation changes the orientation of the baseline (relative to the source) over time (depending on declination); this can be used to ‘fill in’ different parts of the *u,v* plane over time.

Fun fact: 1974 Nobel prize in physics was half for aperture synthesis, including Earth rotation synthesis

Effects of adding baselines

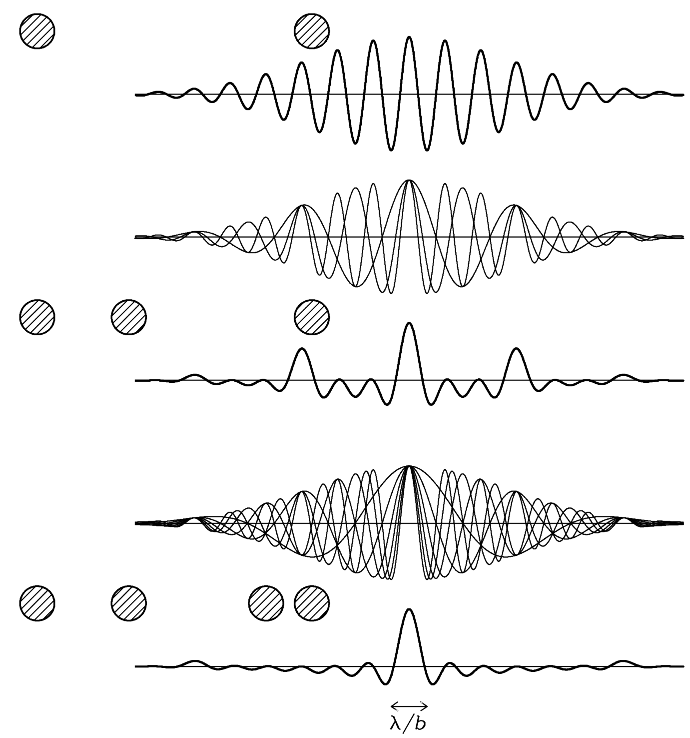


Image credit: Fig 3.42 from Essential Radio Astronomy

With 2 antennas, you get 1 baseline; with 3 you get 3 baselines; with 4 you get 6 baselines. Adding more improves the reconstruction and gives better position accuracy.

Number of baselines

- The number of baselines scales as $$N_{\rm b} = \frac{N_{\rm ant}(N_{\rm ant}-1)}{2}$$
- This increases really fast: 7 antennas gives 21 baselines, 27 gives 351 baselines, 63 antennas gives 1953 baselines.
- If two baselines are the same (same length and orientation), then they add no new information (but this can be useful for calibration). Building an array pattern with no redundancy takes a bit of work.

For large N_ant, this scales as N_a^2.

Adding realism

- We don't measure all possible baselines. How does that affect our ability to reconstruct the image of the sky?
- We can define a measurement/weight function, *W(u,v)*, which takes on non-zero values at the position of every measured baseline:

$$W(u,v) = \sum_i \sum_{j \neq i} W_{ij} \delta(u - u_{ij}) \delta(v - v_{ij})$$

i and *j* iterate over antennas, so this covers all baselines

$$I(l,m) = \mathcal{F}\{\mathcal{V}(u,v)\}$$

$$\mathcal{V}_{\text{measured}} = W(u,v)\mathcal{V}(u,v)$$

$$\mathcal{F}\{\mathcal{V}_{\text{measured}}\} = \mathcal{F}\{W(u,v)\mathcal{V}(u,v)\}$$

$$\mathcal{F}\{\mathcal{V}_{\text{measured}}\} = \mathcal{F}\{W(u,v)\} \circledast \mathcal{F}\{\mathcal{V}(u,v)\}$$

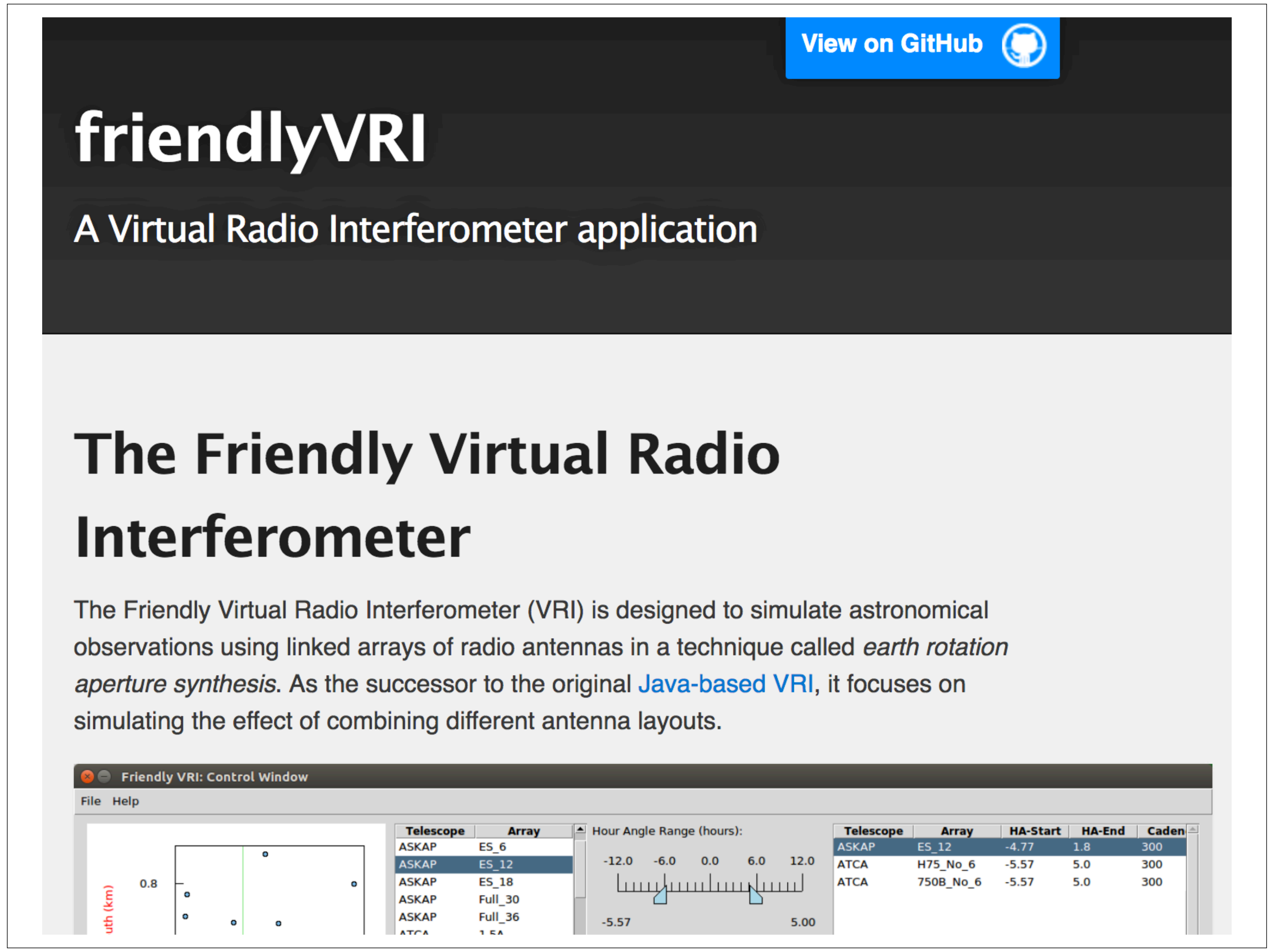
$$\mathcal{F}\{\mathcal{V}_{\text{measured}}\} = \mathcal{F}\{W(u,v)\} \circledast I(l,m)$$

$$I(l,m)_{\text{measured}} = b_0(l,m) \circledast I(l,m)$$

- The **synthesized beam**, b_0, acts as a point-spread function, convolving the actual sky to give us the measured image of the sky brightness.

$$b_0(l,m) = \mathcal{F}\{W(u,v)\}$$

Using script-F to denote Fourier transforms, and applying the Fourier product/convolution theorem, we get the synthesized beam.
Typo on this slide! Should be inverse Fourier transforms. Sky to visibilities is forward transform, visibilities to sky is inverse transform.



View on GitHub

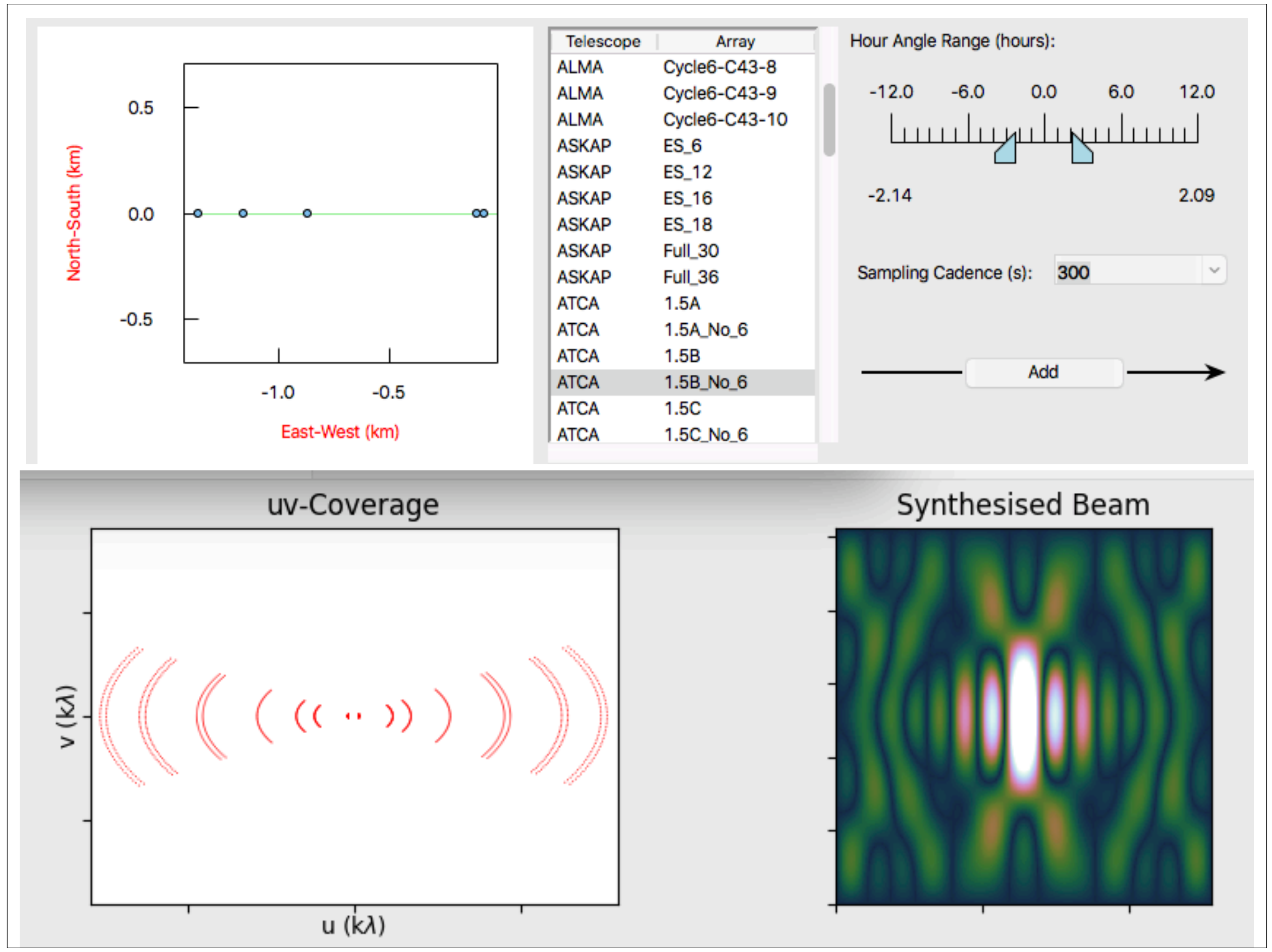
friendlyVRI

A Virtual Radio Interferometer application

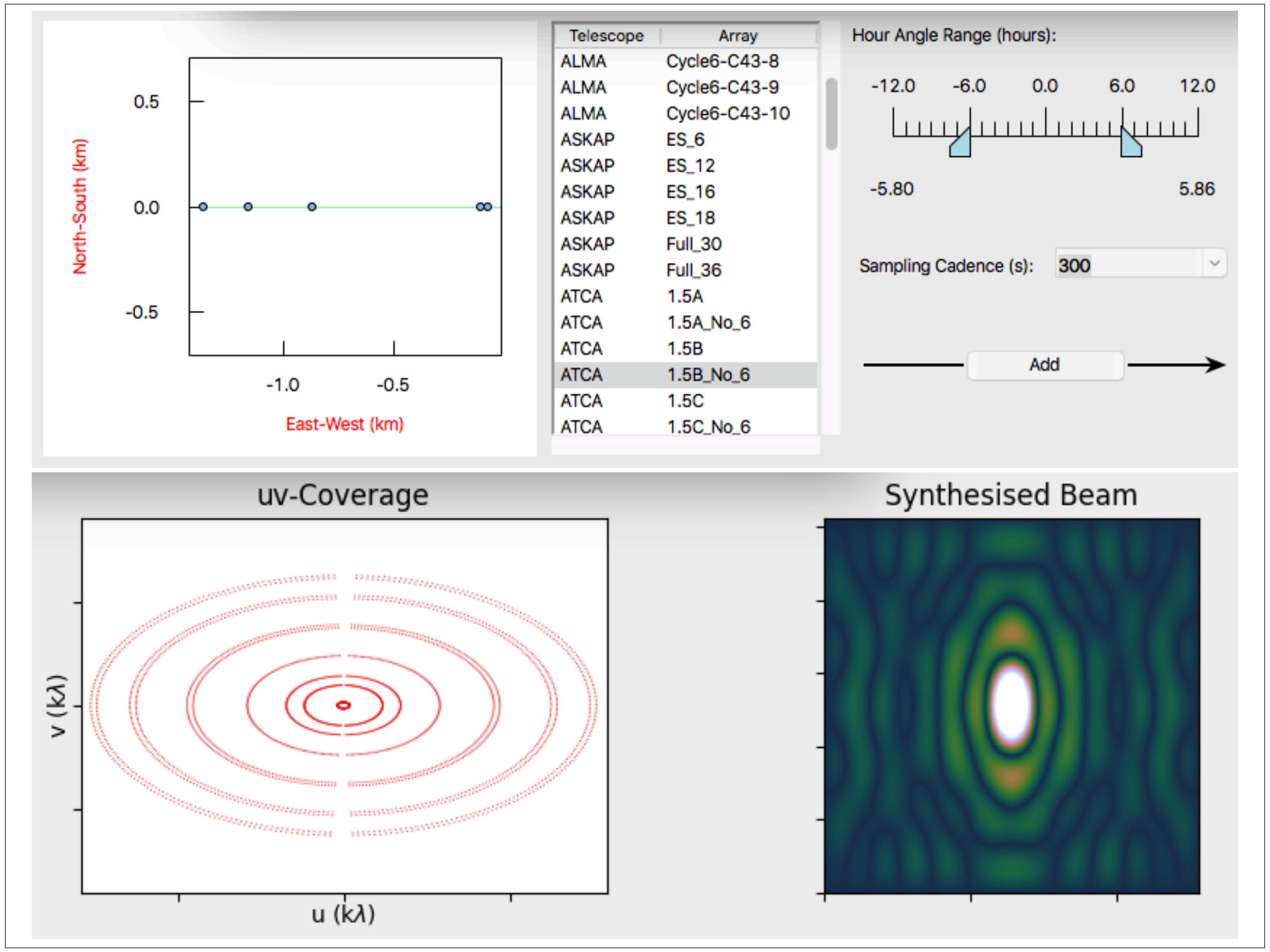
The Friendly Virtual Radio Interferometer

The Friendly Virtual Radio Interferometer (VRI) is designed to simulate astronomical observations using linked arrays of radio antennas in a technique called *earth rotation aperture synthesis*. As the successor to the original Java-based VRI, it focuses on simulating the effect of combining different antenna layouts.

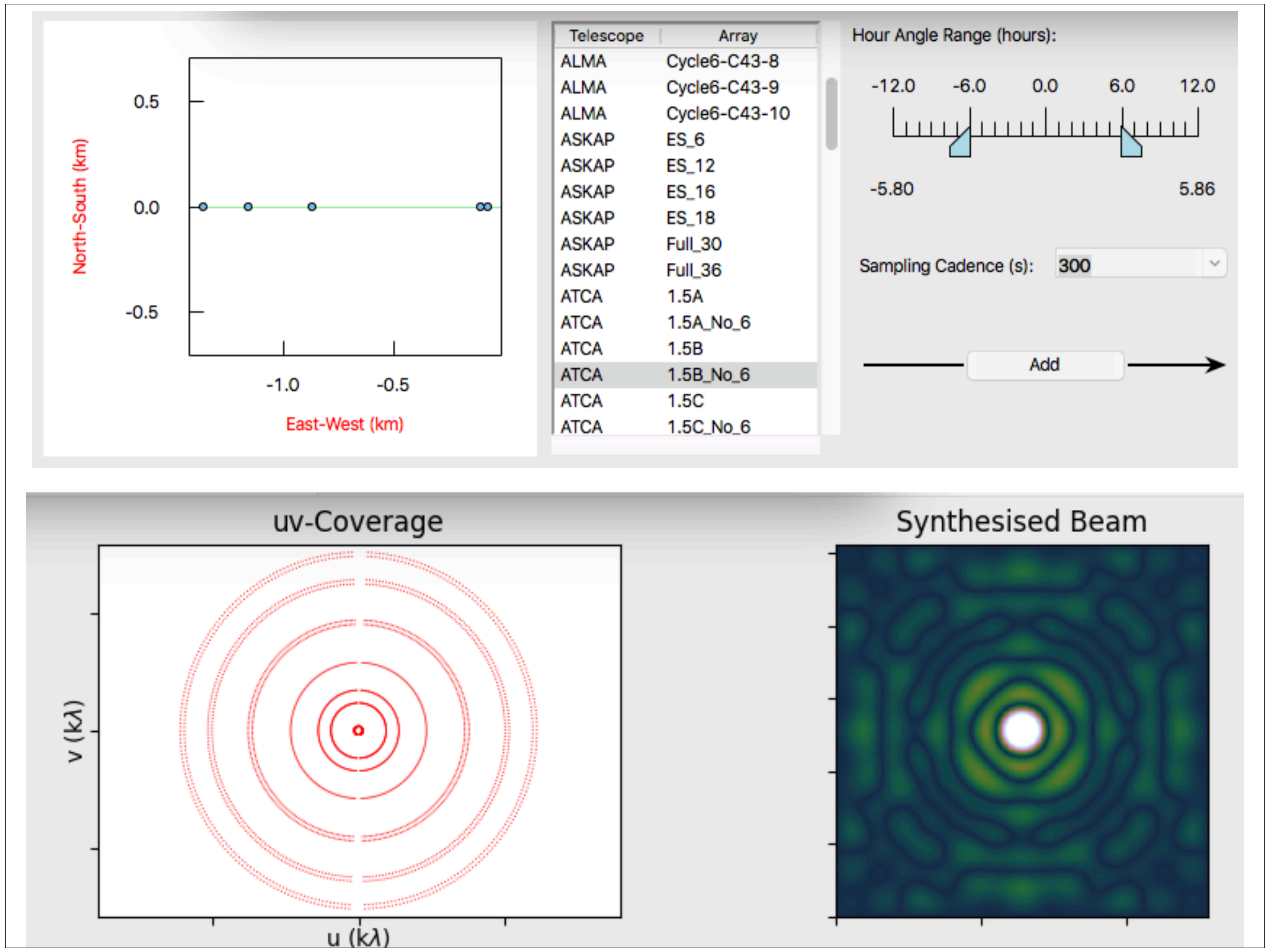
Images on the following slides from friendlyVRI, by Cormac Purcell. https://crpurcell.github.io/friendlyVRI/



An East-west array (ATCA), with a 4-hour observation. Note that there are 5 antennas, giving 10 baselines (each with a track in the u-v plane). Try to match each baseline to it's two telescopes!

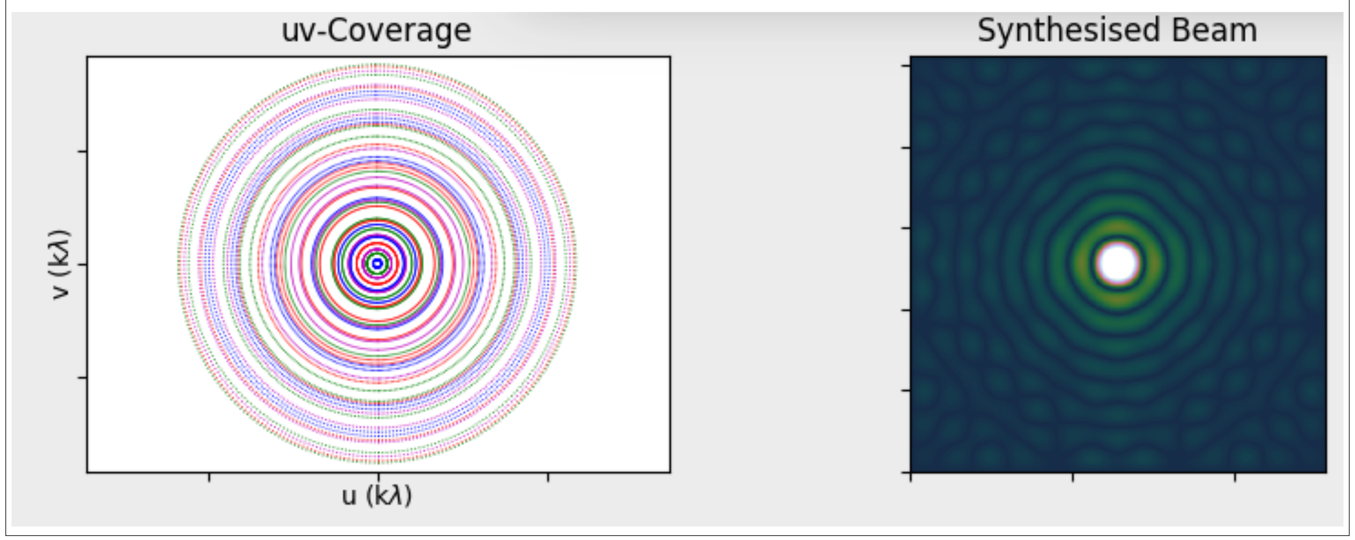


The same antenna configuration, but a 12 hour observation, giving full elliptical tracks. Note the difference between the *u* and *v* scales (from the tick marks): *v* is smaller, causing the beam to be elliptical.



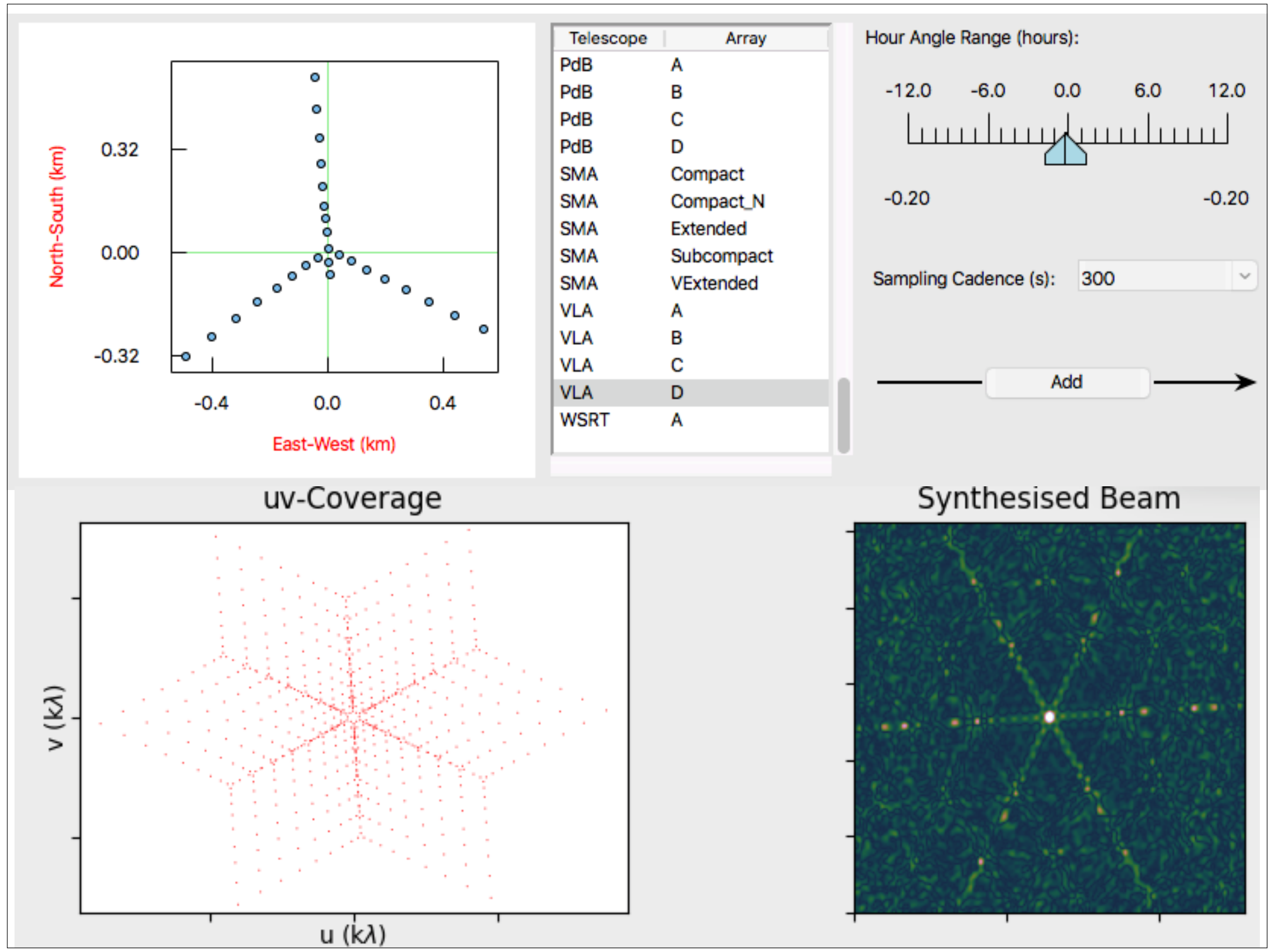
Same array configuration, but I've moved the source to the (south) pole, making the *u,v* tracks perfectly circular, which makes the beam circular.

Telescope	Array	HA-Start	HA-End	Cadence
ATCA	1.5A_No_6	-6.03	5.97	300
ATCA	1.5B_No_6	-6.03	5.97	300
ATCA	1.5C_No_6	-6.03	5.97	300
ATCA	1.5D_No_6	-6.03	5.97	300

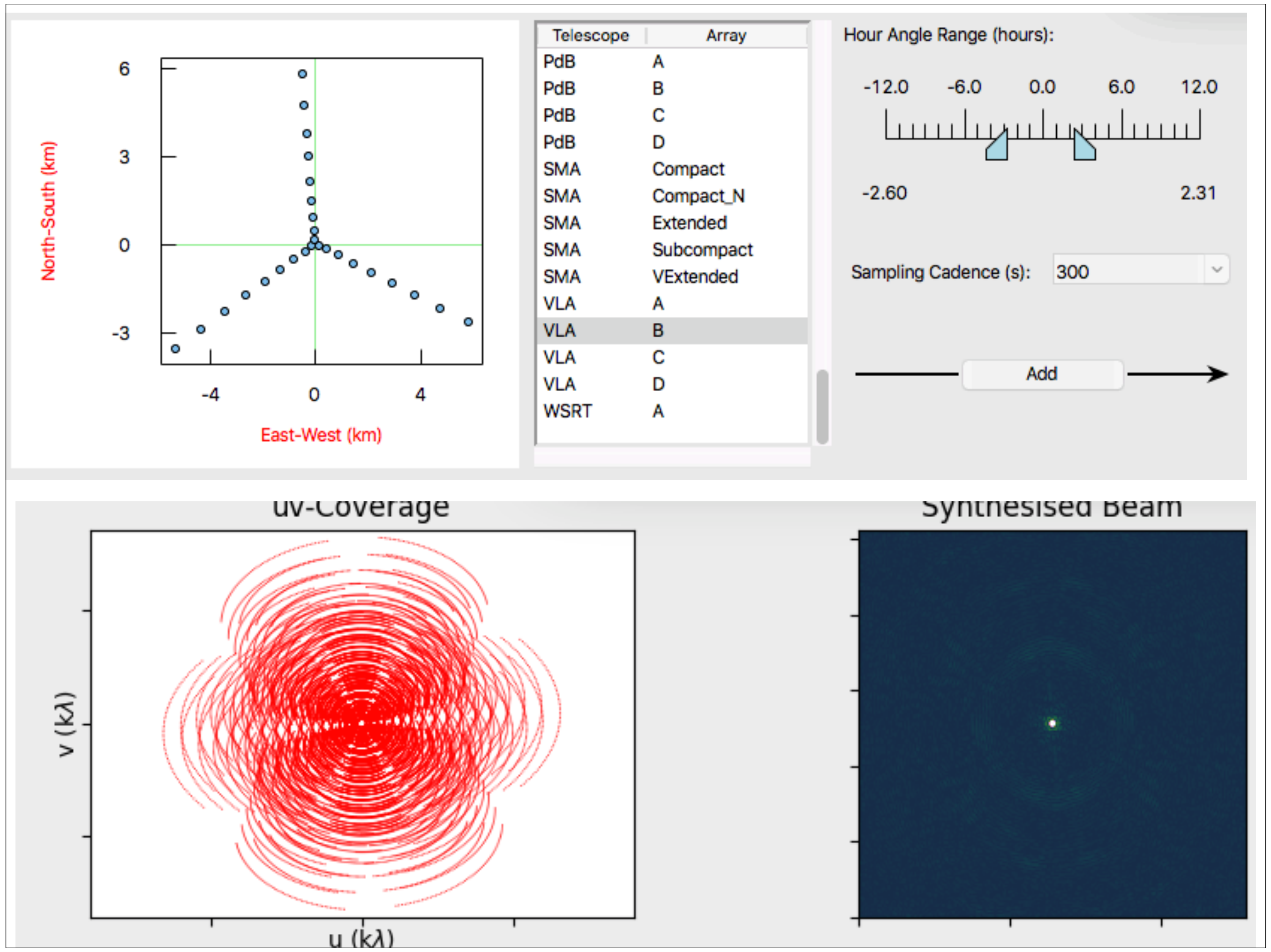


Moving the antennas in an Earth-rotation synthesis telescope lets you fill in the missing gaps to some degree. With 4 configurations of antennas, you get 40 baseline tracks, which improves the beam significantly.

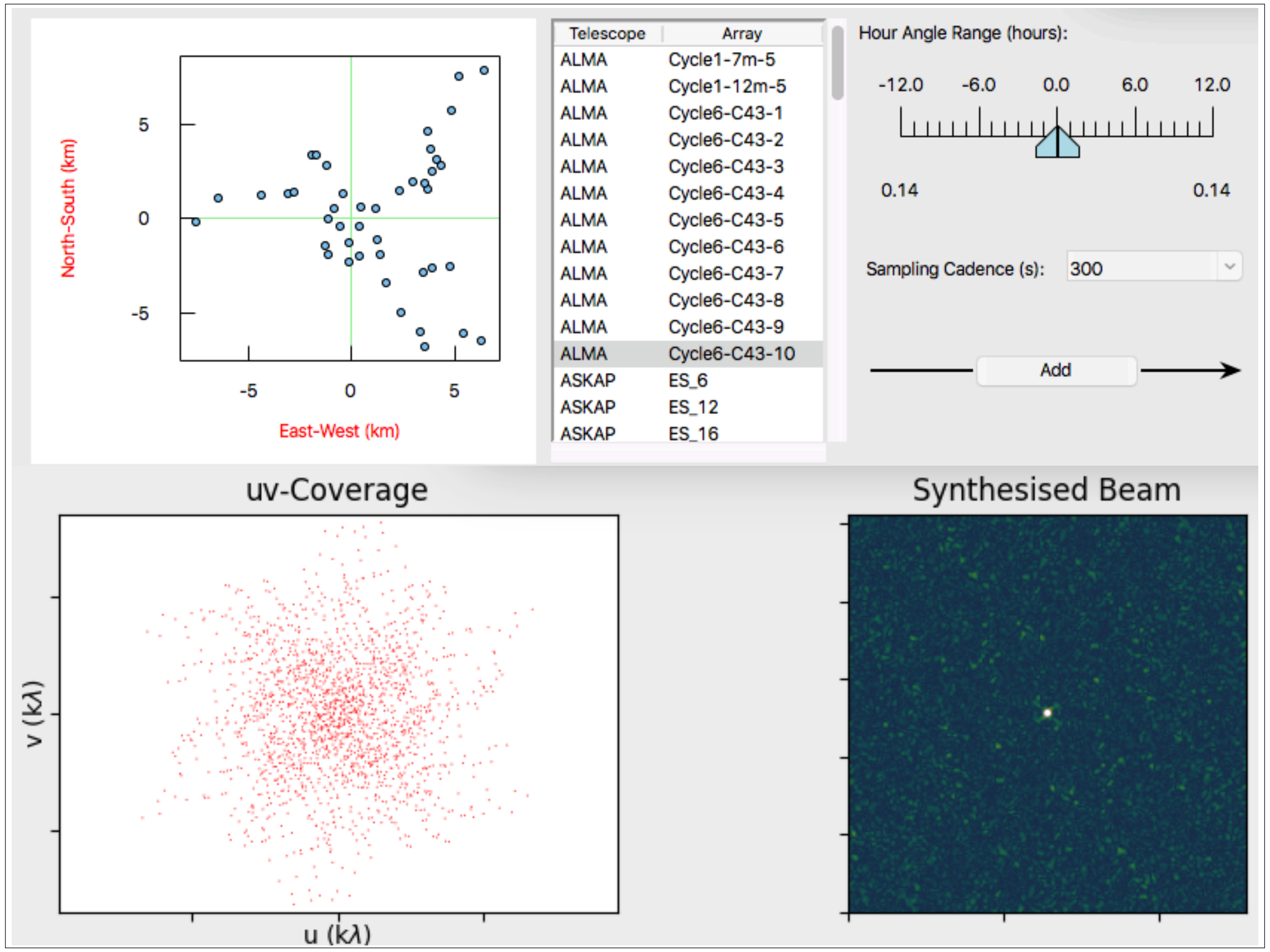
Also, not simulated here: multifrequency synthesis, where you can use visibilities at different frequencies to fill in the *u,v* plane somewhat. (More on that later, maybe?)



Here's a VLA snapshot. Note the 3 arms means there's 3 axes along which you get baselines within an arm. You then get 3 'petals' associated with baselines between different arms. See how you get a synthesized beam with spokes? Those spokes are aligned with the axes where there are the fewest visibilities: the image is least constrained along those axes.



VLA 5 hour observation. Earth rotation synthesis fills in the gaps, producing a more circular beam.



Arrays with many antennas can have a more complicated layout, which improves the instantaneous synthesized beam. Very good for shorter observations!

Primary beam effects

- The **primary beam** is the beam of the individual elements that are being correlated (which doesn't have to be the same for all antennas).
- The primary beam modifies the signal received by the antenna, and the resulting voltage. For a single source:

$$V \rightarrow \sqrt{A(l,m)}V$$

Interferometers can have elements with different beams, such as in VLBI where different sized/shaped dishes are used.

The reason for the square root is because then the beam is defined in intensity/power units, rather than voltage units. In principle there's also the integral over directions to get the voltage in the antenna, but for now let's just work through for single source (since everything's linear, right?).

$$\mathcal{V} = \int I(\hat{s}) e^{-i2\pi \frac{\vec{b}}{\lambda}\cdot\hat{s}} d\Omega$$

$$\mathcal{V}_{ij} = \int \frac{V_i(l,m)V_j(l,m)}{2} e^{-i2\pi(ul+vm)} d\Omega$$

$$\mathcal{V}_{ij} = \int \sqrt{A_i(l,m)A_j(l,m)} I(l,m) e^{-i2\pi(ul+vm)} d\Omega$$

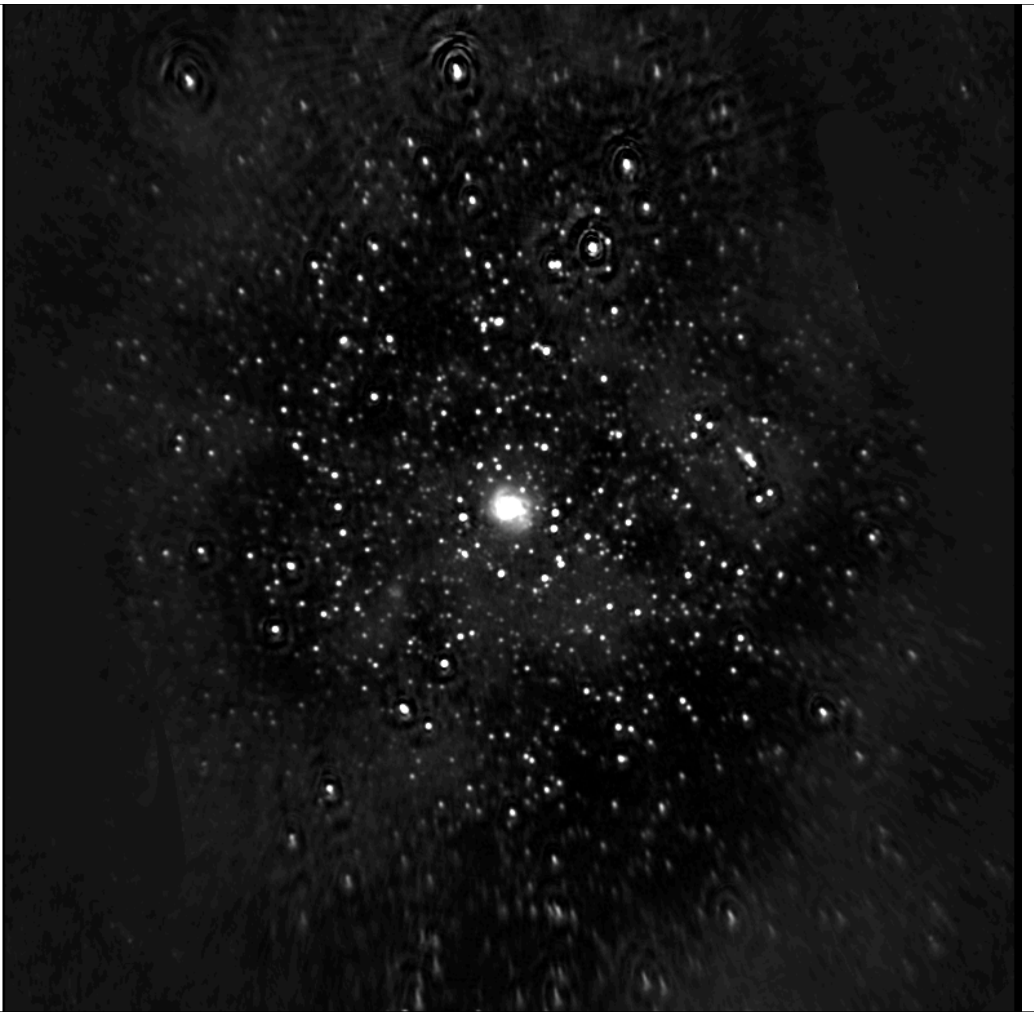
$$\mathcal{V}_{ij} = \mathcal{F}\left\{\sqrt{A_i(l,m)A_j(l,m)} I(l,m)\right\}$$

$$\mathcal{F}\{\mathcal{V}_{ij}\} = \sqrt{A_i(l,m)A_j(l,m)} I(l,m)$$

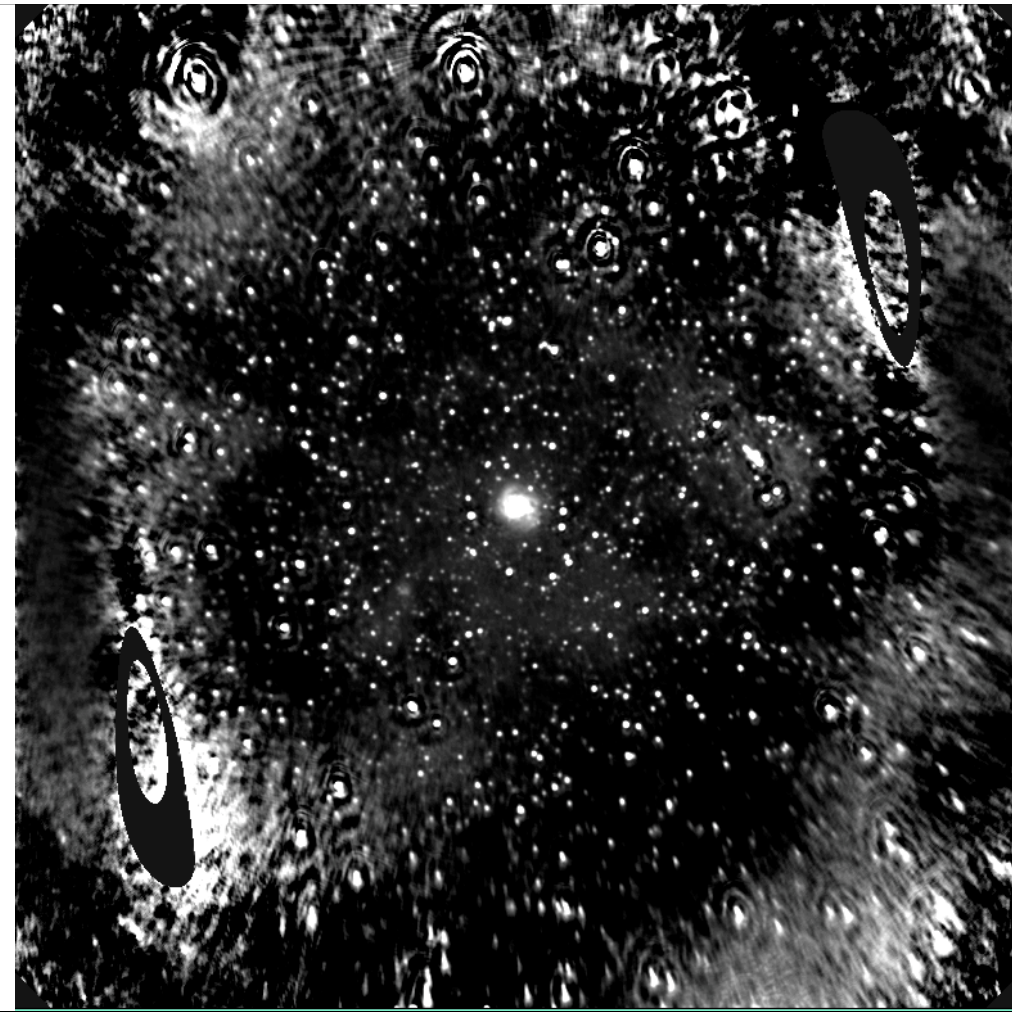
Primary beam effects

$$\mathcal{F}\{\mathcal{V}_{ij}\} = \sqrt{A_i(l,m)A_j(l,m)}I(l,m)$$

- Primary beam acts as a multiplicative factor on the image. Correct by dividing the output of the imaging by the beam model to recover the true intensity.
- This amplifies the noise in the image: dividing by the beam model corrects the intensity but increases the noise at the same time.
- Sensitivity can't be recovered at the nulls (zeros) in the beam (dividing by zero).



This is a LOFAR image of galaxy IC342 and surroundings, with no primary beam correction. Note the ‘fading’, how sources appear dimmer at the edges.



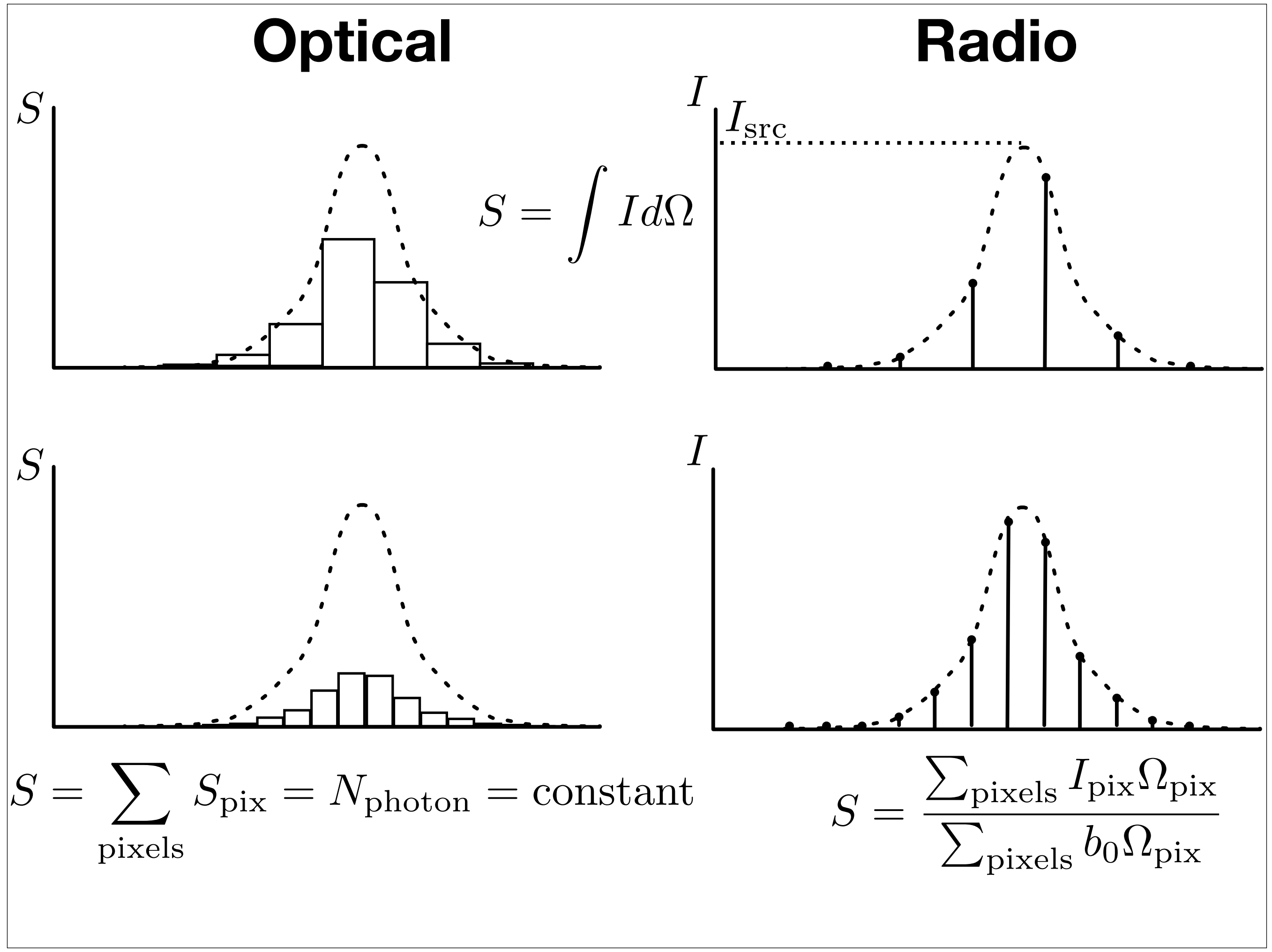
Corresponding image WITH primary beam correction. Note that sources at the edge now look equally bright with the center. Note how the noise is hugely amplified at the edges, especially around the nulls (which are simply blanked in the imaging software).

Imaging

- $I(l,m)_{measured}$ is a continuous quantity, which we can calculate by (inverse) Fourier transforming our visibility measurements

- Since we're producing digital images, this will end up quantized on some pixel grid, like optical images.

- Unlike optical images, radio images are not photon counting! This changes how they react to different pixel sizes, and how intensity and flux density are determined.



Consider an unresolved source broadened by some PSF/synthesized beam (dashed line).

Optical cameras are intrinsically flux density measuring: they integrate over the area of the pixel. Shrinking the pixel changes the signal per-pixel, but not the integrated value.

Radio images are fundamentally different: they report intensity. Increasing the number of pixels increases integral, which is unphysical. Need to normalize by the synthesized beam to get actual flux density (often called ‘integrated flux’ vs ‘peak flux’).

‘Peak flux’, aka intensity, can be read directly off the image. But this is not actual source intensity! It is flux density, if the source is unresolved. If source is resolved, then it’s not well defined.

Integrated flux is found by summing over pixels, and dividing by sum-over-beam (to remove effects of PSF)

Imaging

- The imaging process creates an NxM pixel image from the Fourier transform of the visibilities. The number of pixels, and the size of each pixel, is set by the user.

- The most efficient method is an FFT algorithm, which produces an NxM pixel image from an NxM grid of points in the *u,v* plane.

- Baselines won't fall exactly onto this grid, so there's some **gridding** process which puts the visibilities onto the grid. Visibilities that fall into the same grid point get average together.

Imaging

- NxM and the pixel size are chosen by the user, but generally informed by the observation parameters:
    - Pixel size should be such that there are at least 3-5 pixels across the synthesized beam. The size of the synthesized beam is inversely related to the longest baseline in the observation.
    - Number of pixels should be enough to cover the primary beam*, out to about the first null, and sometimes the first sidelobe if there's a bright source there.

The factor of 3-5 pixels (or more) across the beam is called the *oversampling factor* (because you sample more finely than your resolution). This is important because it means you're sampling the image well in case a source falls halfway between pixels (remember, not photon counting!). In principle, more oversampling is good, except that it increases the number of pixels needed.

*: VLBI is slightly different. Smearing/decorrelation effects tend to dominate and limit the field-of-view to be smaller than the primary beam.