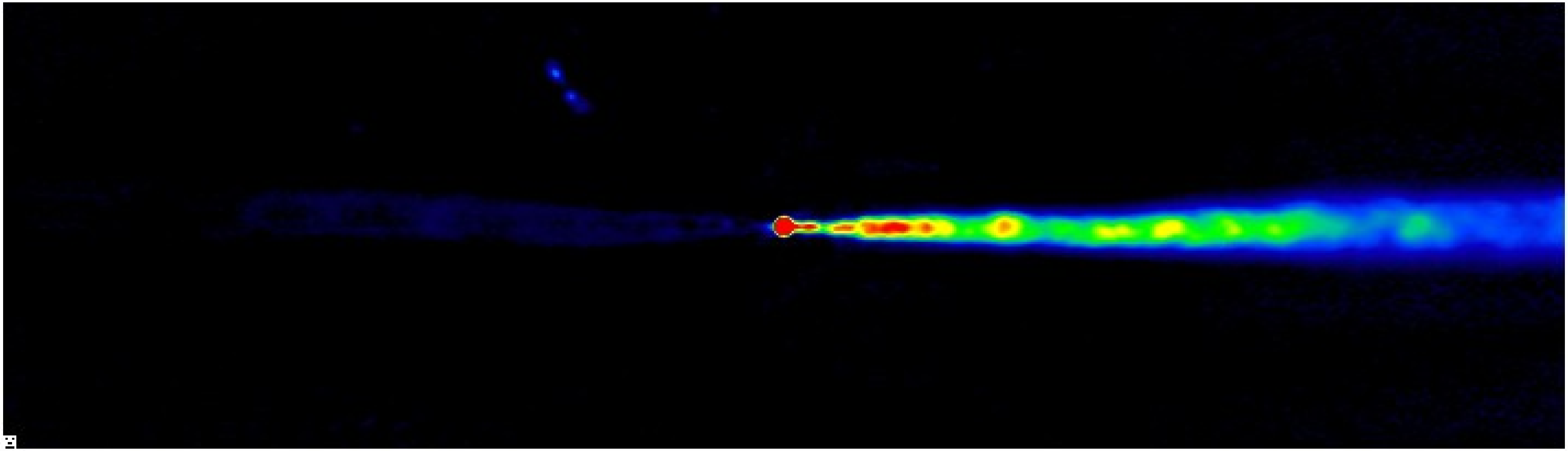
Principles of Radio and Millimetre Interferometry

Robert Laing (ESO)



Outline

- Why radio interferometry?
  - Resolution
  - Astronomical applications: radiation mechanisms
- Fundamentals
  - Young's slit
  - Coherence functions
  - Images and visibilities: the Fourier relation
- How does a radio interferometer work?
  - Follow the signal path
  - Technologies for different frequency ranges
- Current arrays
  - ALMA
  - VLA
  - Very Long Baseline Interferometry

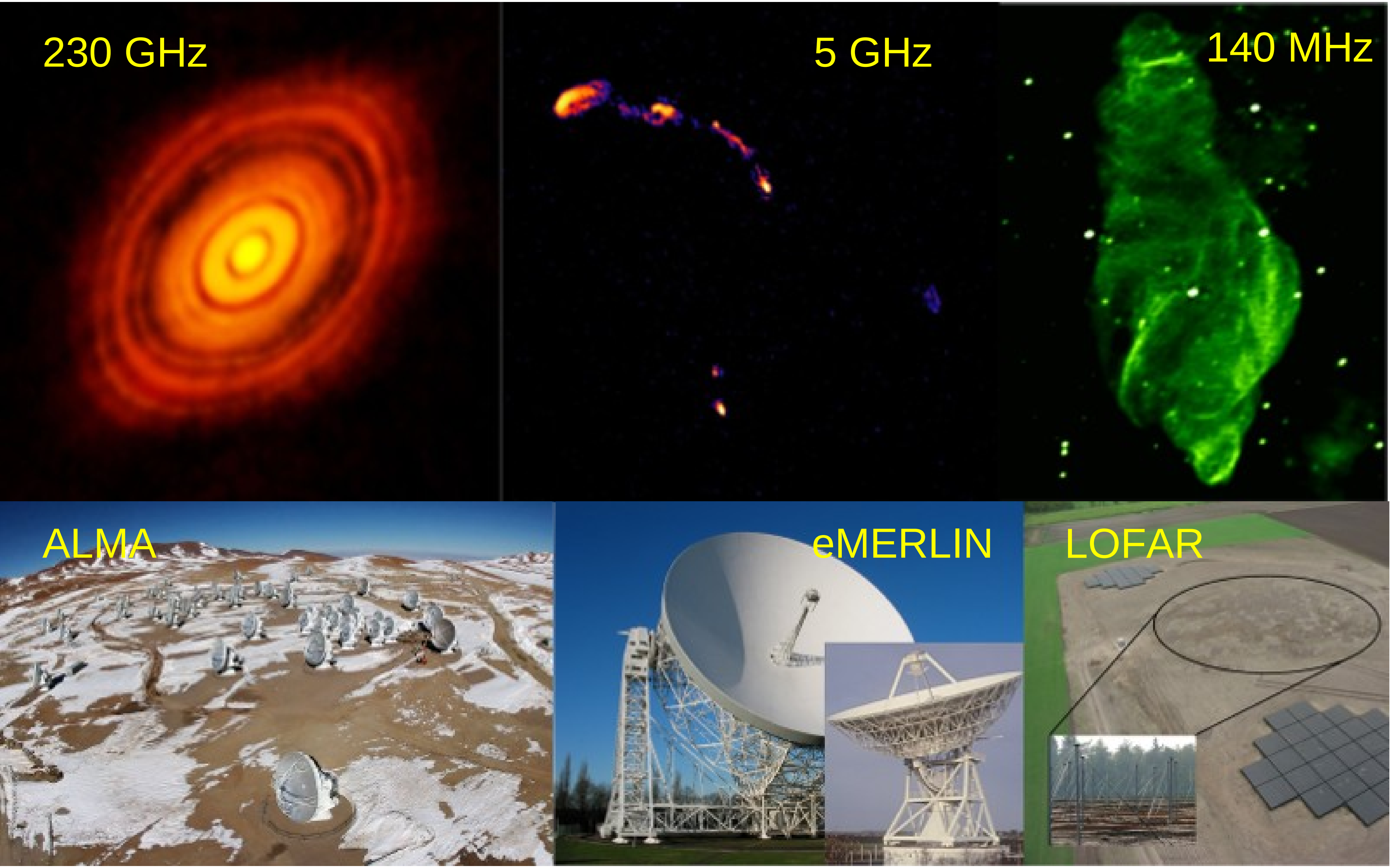
Why interferometry?

- Wavelength/frequency range
  - Limits set by the atmosphere
  - 10 m/30 MHz (ionosphere) to 0.3 mm/1 THz (troposphere)
    - except in exceptional conditions
  - Conventionally “radio” 10 m – 10 mm; “millimetre/sub-millimetre” 10 mm – 0.3 mm (arbitrary)
- Resolution
  - For a single telescope, diameter D, the angular resolution in radians ~λ/D
  - Maximum diameter ~300m (Arecibo), corresponding to 40 arcsec resolution at 6 cm/5 GHz
  - Hopeless by comparison with optical telescopes, Hubble Space Telescope (0.05 arcsec)
- The solution: interferometry
  - Resolution is then ~λ/d, where d is now the separation of the interferometer elements – potentially Earth diameter or larger

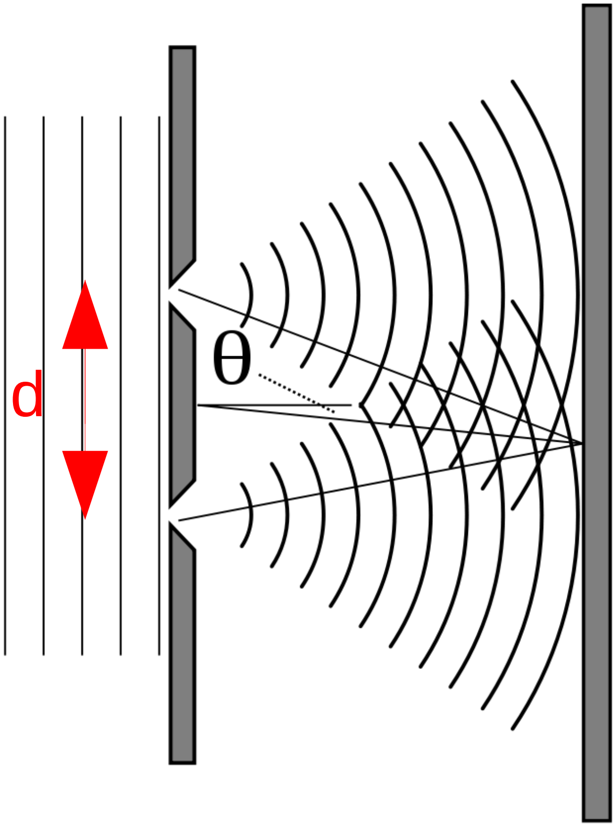
Radiation mechanisms: radio and mm

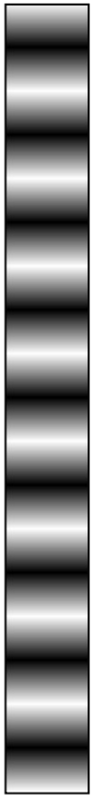
- Continuum mechanisms
  - Synchrotron radiation (relativistic electrons; non-thermal energy distribution).
  - Free-free emission (bremsstrahlung)
  - Thermal emission from dust
- Lines
  - Hyperfine transition of neutral hydrogen (21 cm/1420 MHz)
  - Rotational and vibrational transitions of molecules (CO strongest; many other species including complex organic molecules)
  - Masers (OH, SiO, water, ...)
  - Atomic transitions (e.g. [CII])

From high to low frequencies



Young's Slit Experiment





Angular spacing of fringes = λ/d

Familiar from optics

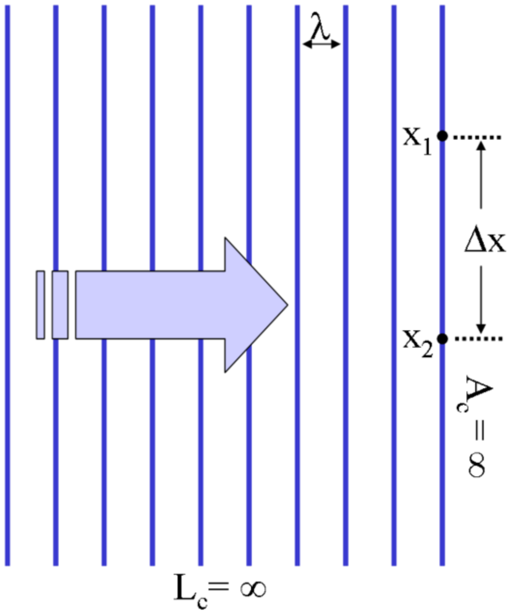
Essentially the way that astronomical interferometers work at optical and infrared wavelengths (e.g. VLTI)

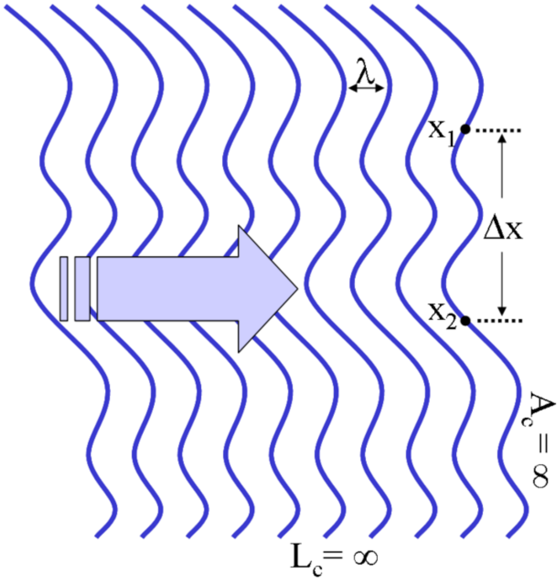
“Direct detection”

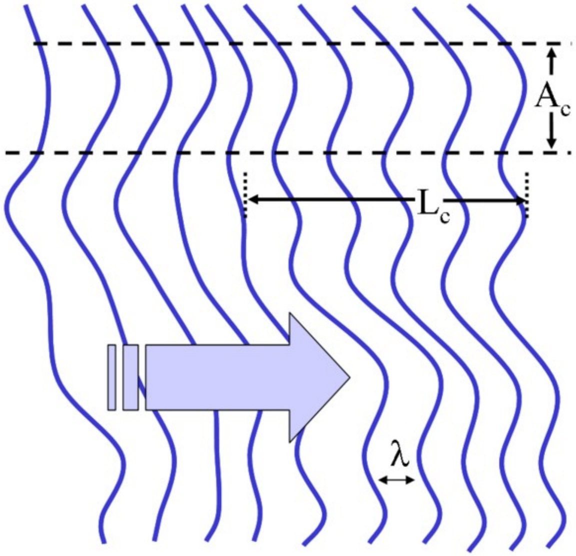
But this is not how radio interferometers work in practice ....

- The two techniques are closely related, and it often helps to think of images as built up of sinusoidal “fringes”
- But radio interferometers collect radiation (“antenna”), turn it into a digital signal (“receiver”) and generate the interference pattern in a special-purpose computer (“correlator”)
- How does this work?
- I find it easiest to start with the concept of the mutual coherence of the radio signal received from the same object at two different places
- No proofs, but I will try to state the simplifying assumptions clearly

Spatial coherence







Coherent plane wave

Coherent, varying profile

Partial coherence

Two waves are spatially coherent if they have a constant phase relation at different points in space.

The ideal interferometer (1)

- Astrophysical source, location **R**, generates a time-varying electric field **E**(**R**,t). EM wave propagates to us at point **r**.
- In frequency components**:** $\mathbf{E}(\mathbf{R},t) = \int \mathbf{E}_\nu(\mathbf{R})\exp(2\pi i\nu t)d\nu$
- Assumptions:
  - Monochromatic radiation
  - Scalar field
  - Source is far away
  - Space is empty

Generalise later

Usually good approximations

- With these assumptions, Huygens' Principle says
  - $E_\nu(\mathbf{r}) = \int E_\nu(\mathbf{R})\{\exp[2\pi i|\mathbf{R}-\mathbf{r}|/c]/|\mathbf{R}-\mathbf{r}|\}\, dA$ (dA is the element of area at distance |**R**|)
- What we can measure is the **correlation** of the field at two different observing locations. This is

$C_\nu(\mathbf{r}_1,\mathbf{r}_2) = \langle E_\nu(\mathbf{r}_1)E^{*}_\nu(\mathbf{r}_2)\rangle$ where <> is a time average

The ideal interferometer (2)

- Assumption: radiation from astronomical objects is not spatially coherent ("random noise")
  - $\langle E_\nu(\mathbf{R}_1)E^*_\nu(\mathbf{R}_2)\rangle = 0$ unless $\mathbf{R}_1 = \mathbf{R}_2$
- Now write $\mathbf{s} = \mathbf{R}/|\mathbf{R}|$ and $I_\nu(\mathbf{s}) = |\mathbf{R}|^2\langle|E_\nu(\mathbf{s})|^2\rangle$ (the observed intensity). Using the approximation of large distance to the source again:
  - $C_\nu(\mathbf{r}_1,\mathbf{r}_2) = \int I_\nu(\mathbf{s}) \exp[-2\pi i\nu\mathbf{s}.(\mathbf{r}_1\mathbf{-r}_2)/c]\, d\Omega$
- $C_\nu(\mathbf{r}_1,\mathbf{r}_2)$, the spatial coherence function, depends only on separation $\mathbf{r}_1\mathbf{-r}_2$, so we can keep one point fixed and move the other around.
- It is a complex function, with real and imaginary parts, or an amplitude and phase.

An interferometer is a device for measuring the spatial coherence function

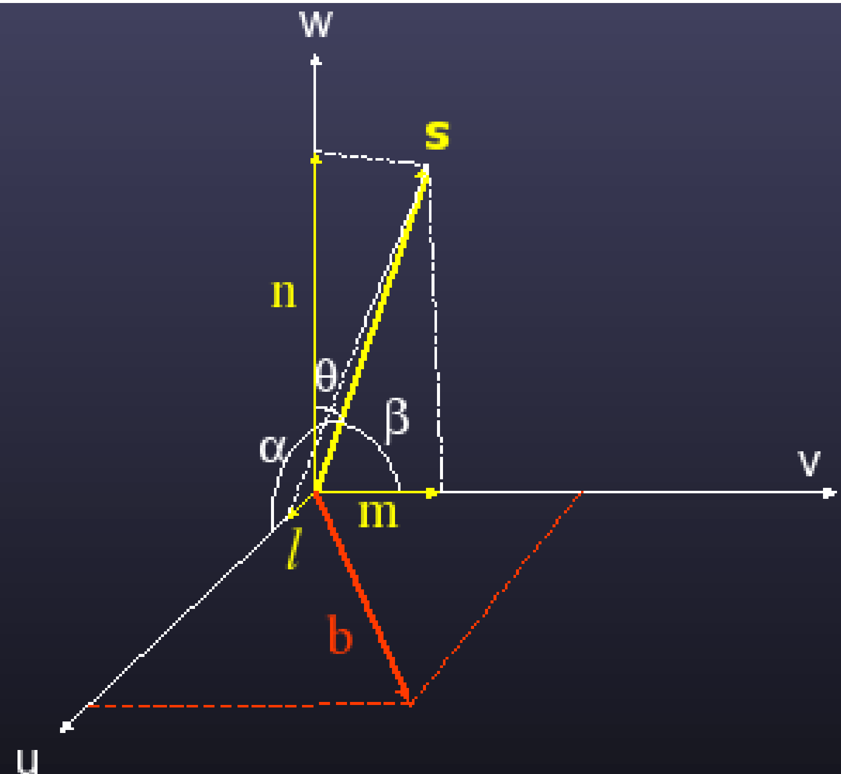
u, v, w and direction cosines

The unit direction vector **s** is defined by its projections on the (u,v,w) axes. These components are called the **Direction Cosines**.

$$l = \cos(\alpha)$$

$$m = \cos(\beta)$$

$$n = \cos(\theta) = \sqrt{1 - l^2 - m^2}$$



The baseline vector **b** is specified by its coordinates (u,v,w) (measured in wavelengths). In this special case,

$$\mathbf{b} = (\lambda \mathrm{u}, \lambda \mathrm{v}, \lambda \mathrm{w}) = (\lambda \mathrm{u}, \lambda \mathrm{v}, 0)$$

The Fourier Relation

- Assumptions
  - Receiving elements have no direction dependence
  - All sources are in a small patch of sky
  - We can measure at all values of $\mathbf{r}_1$-$\mathbf{r}_2$

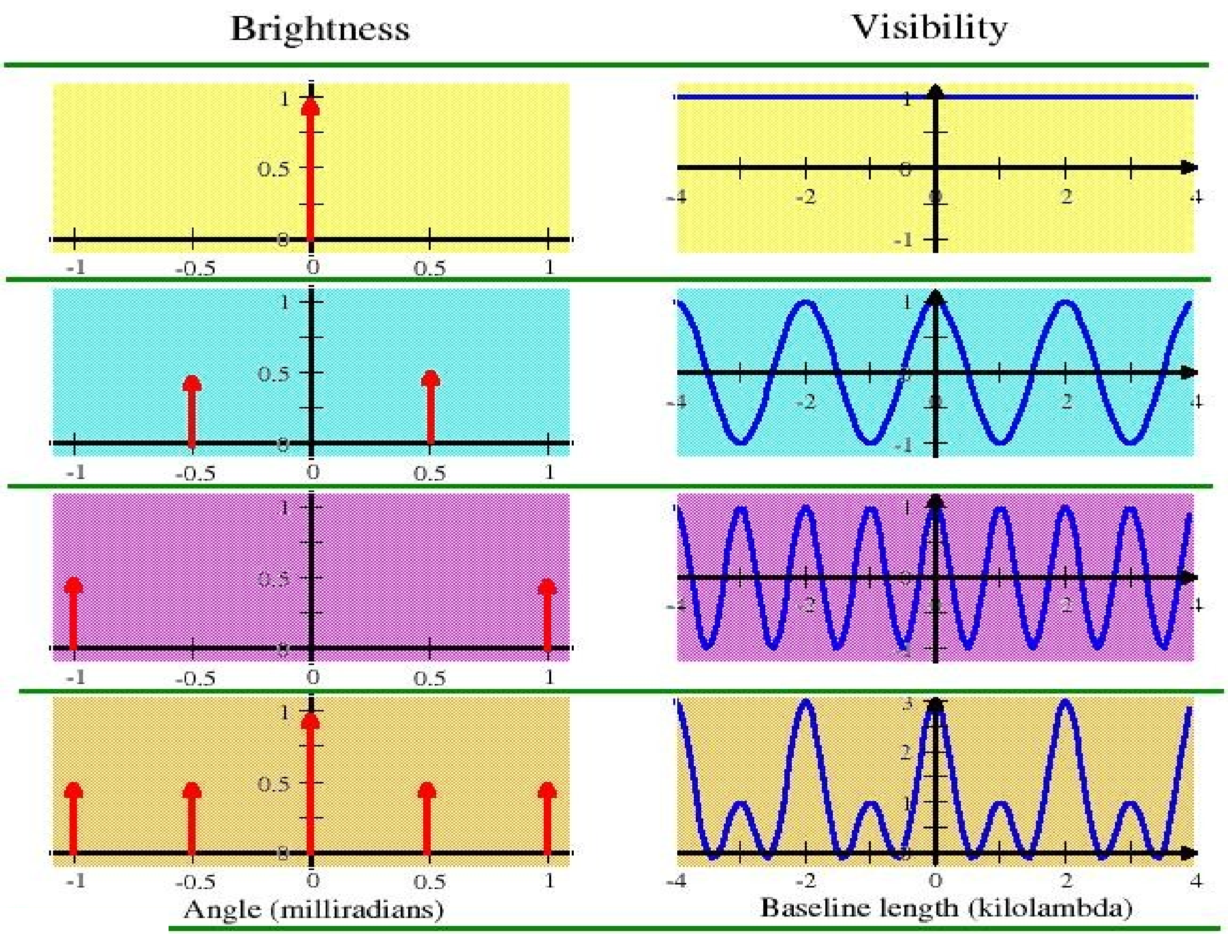
Generalise later

- Pick a special coordinate system such that the phase tracking centre has $s_0 = (0,0,1)$
  - $C(\mathbf{r}_1,\mathbf{r}_2) = \exp(-2\pi i w) V'_\nu(u,v)$
  - $V'_\nu(u,v) = \iint I_\nu(l,m) \exp[-2\pi i(ul+vm)]\, dl\, dm$
- This is a Fourier transform relation between the modified complex visibility V'_ν (the spatial coherence function with separations expressed in wavelengths) and the intensity $I_\nu(l,m)$
- This relation can be inverted to get the intensity distribution, which is what we want:

$$I_\nu(l,m) = \iint V'_\nu(u,v) \exp[2\pi i(ul+vm)]\, du\, dv$$

Interlude: Fourier Transforms

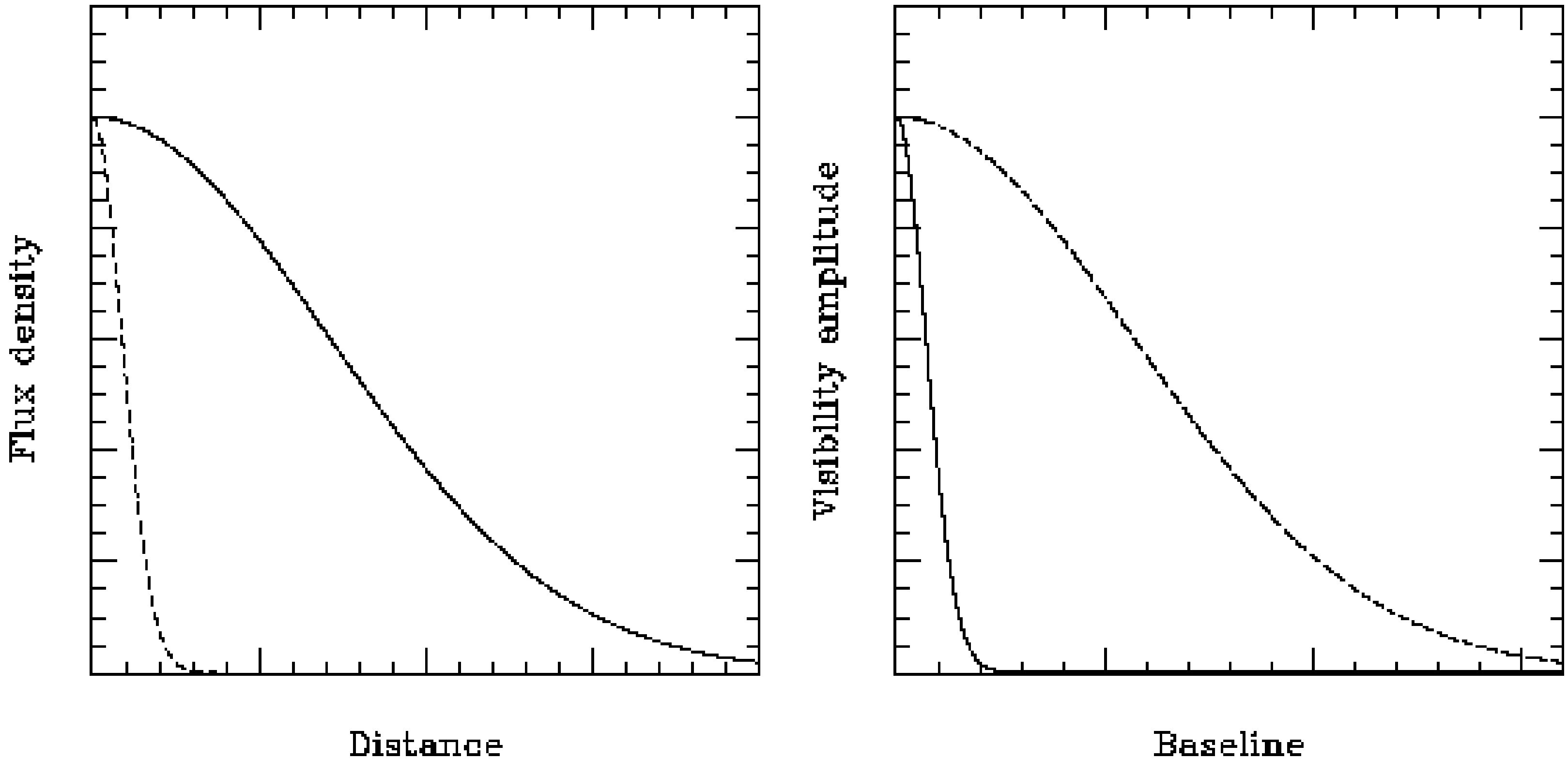
1D Fourier transform pairs



$$F(\nu) = \int_{-\infty}^{\infty} f(t) \exp(-2\pi i \nu) dt$$

$$f(t) = \int_{-\infty}^{\infty} F(s) \exp(2\pi i \nu) d\nu$$

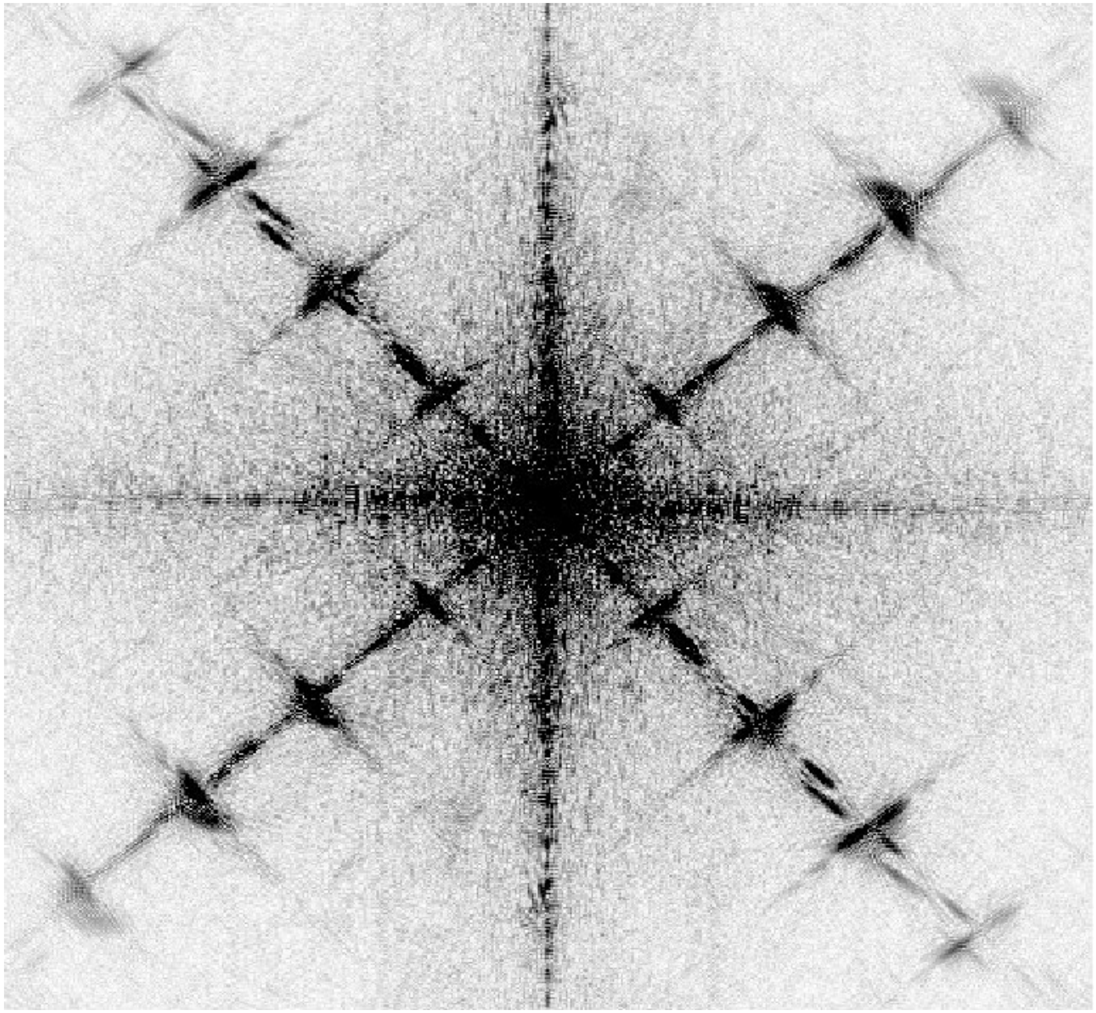
Fourier Transforms (2)



The Fourier transform of a Gaussian function is another Gaussian.

FWHM on sky is inversely proportional to FWHM in spatial frequency: fat objects have thin Fourier transforms and vice versa.

What is this?

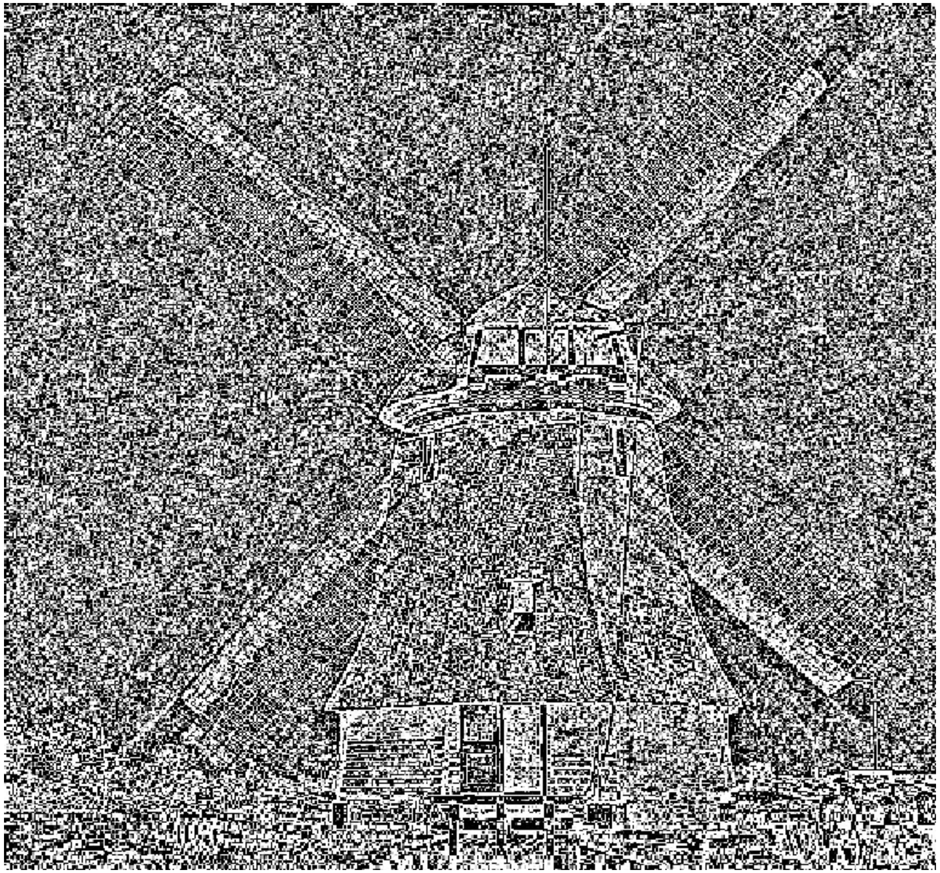


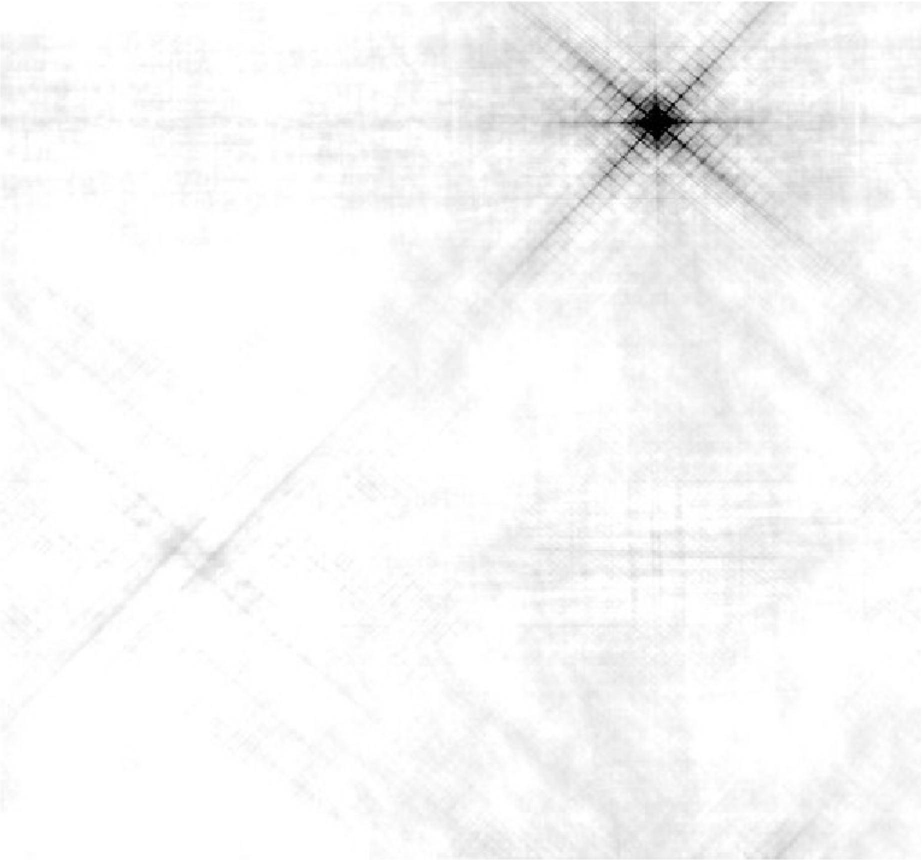
This is the amplitude of the Fourier transform of an image of a well-known object.

Can you:
- Say something about its size, shape and orientation?
- Deduce anything about its fine-scale structure?

The Answer

Phase and Amplitude





Unit amplitude + correct phase

Zero phase + correct amplitude

Reviewing the Assumptions

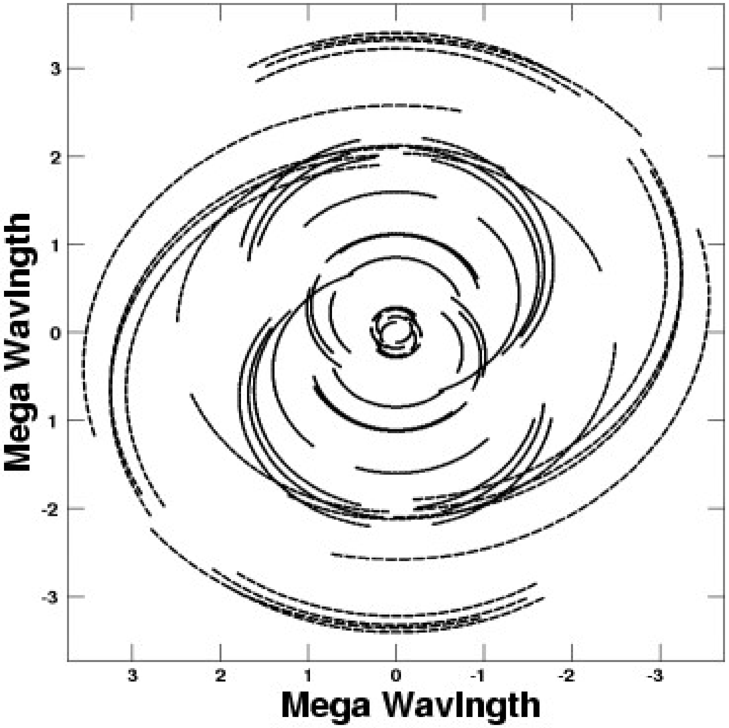
- Monochromatic radiation
  - Can easily divide the band into multiple spectral channels (see later)
- Scalar field
  - Two orthogonal polarization states (crossed linear or RH and LH circular); treat independently
- Distant objects (in the far field)
  - Very good approximation except for some solar-system observations
- Incoherent emission
  - True, except in a few very special cases
- Space is empty and receiving elements have no direction dependence
  - Closely related and **not true** in general
  - More complex algorithms needed (current R&D)
- Small field of view
  - Often not true, but algorithms exist

Sampling

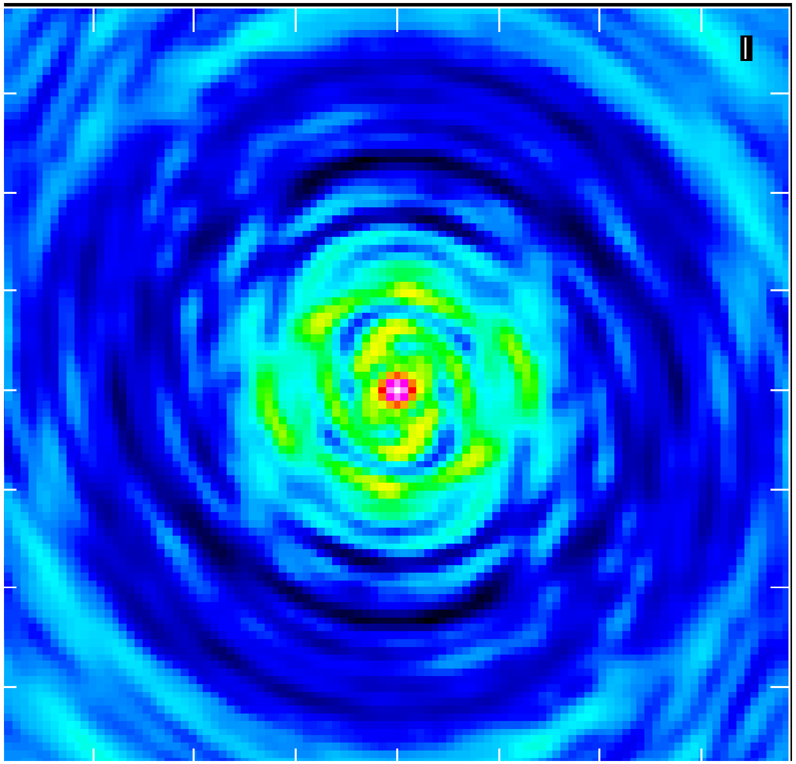
- Remaining assumption: we can measure the coherence function for any spacing. In fact:
  - We have a number of antennas at fixed locations on the Earth (or in orbit around it)
  - The Earth rotates
  - We make many (usually) short integrations over extended periods, sometimes in separate observations
  - So we effectively measure at discrete u, v (and w) positions.
- In 2D, this process can be described by a sampling function S(u,v) which is a delta function where we have taken data and zero elsewhere.
- $I^D_\nu(l,m) = \iint V_\nu(u,v)\, S(u,v) \exp[2\pi i(ul+vm)]\, du\, dv$ is the **dirty image**, which is the Fourier transform of the **sampled** visibility data.
- $I^D_\nu(l,m) = I_\nu(l,m) \otimes B(l,m)$, where the $\otimes$ denotes convolution and

  $B(l,m) = \iint S(u,v) \exp[2\pi i(ul+vm)]\, du\, dv$ is the **dirty beam**

Dirty beam

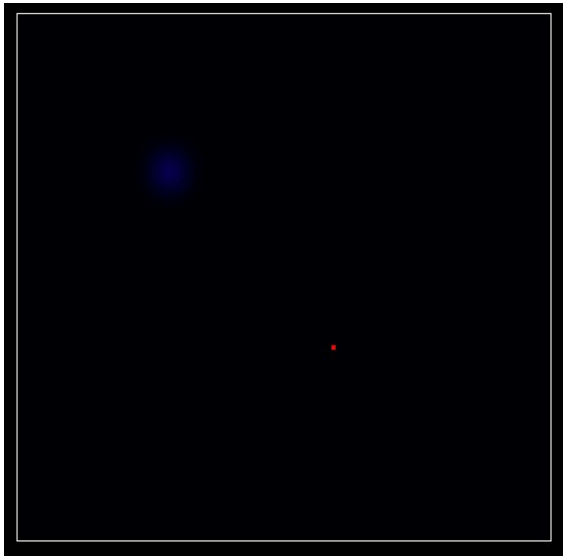


Sampling function

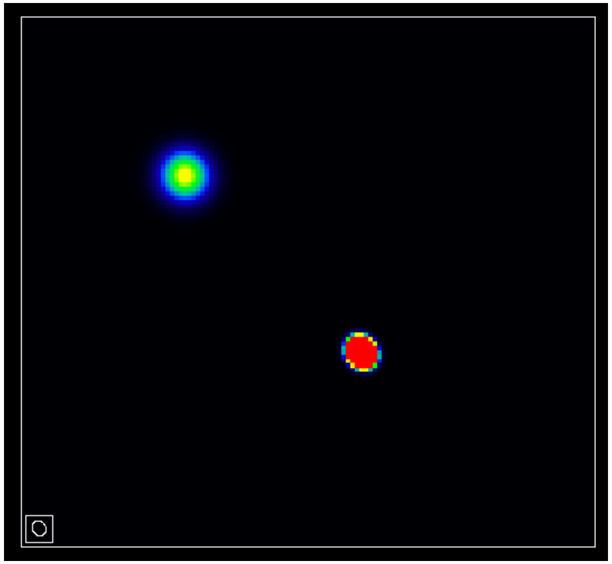


Dirty beam

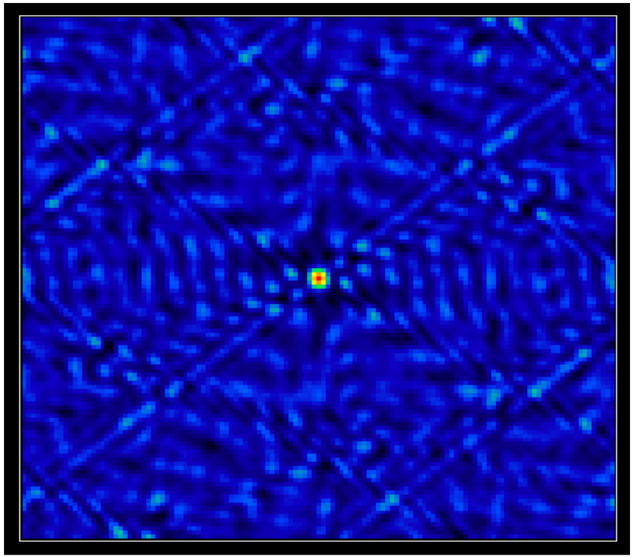
Deconvolution



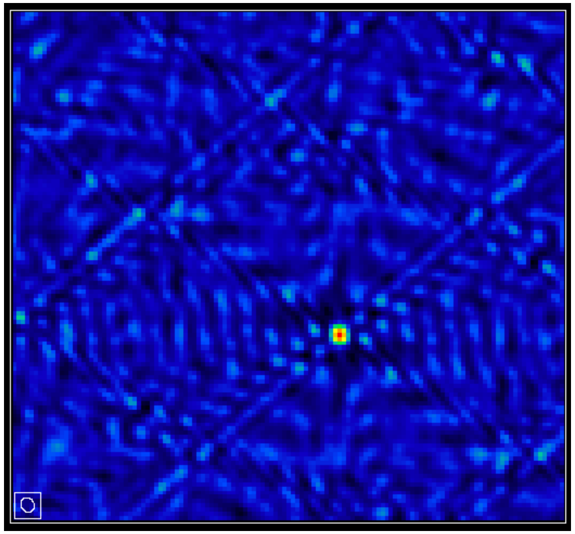
Model



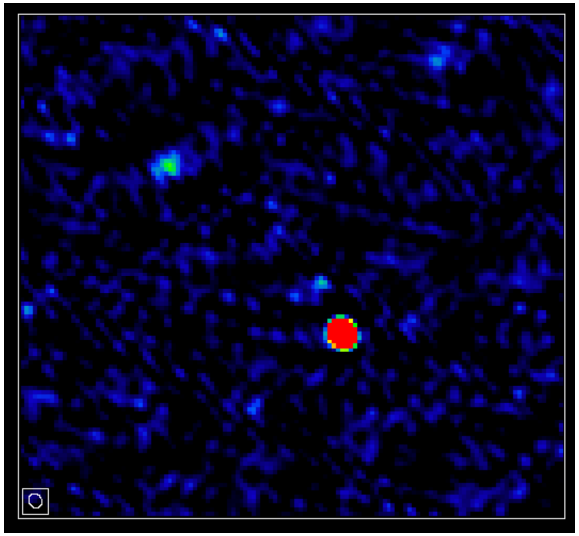
Convolved model



Dirty beam

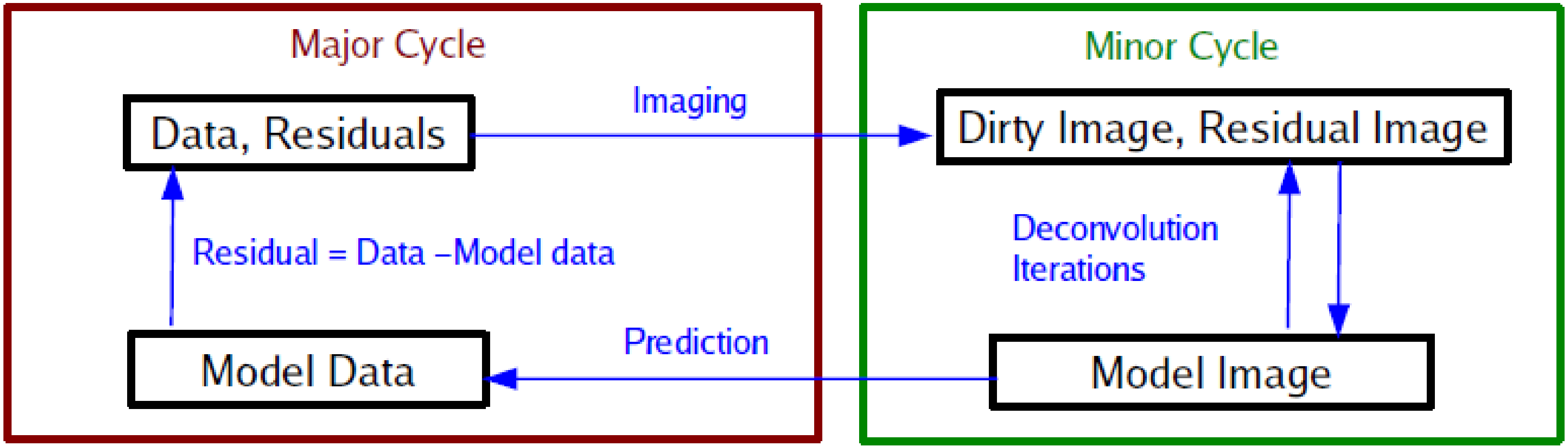


Dirty map



Deconvolved

Imaging Algorithms



Deconvolution algorithms

CLEAN (represent the sky as a set of delta functions)
Multi-scale clean (represent the sky as a set of Gaussians)
Maximum entropy (maximise a measure of smoothness)
....

Resolution and field size

- Some useful parameters:
  - Resolution /rad: $\approx\lambda/d_{max}$
  - Maximum observable scale /rad: $\approx \lambda/d_{min}$
  - Primary beam/rad: $\approx \lambda/D$
- Good coverage of the u-v plane (many antennas, Earth rotation) allows high-quality imaging.
- Some brightness distributions are in principle undetectable:
  - Uniform
  - Sinusoid with Fourier transform in an unsampled part of the u-v plane.
- Sources with all brightness on scales $>\lambda/d_{min}$ are resolved out.
- Sources with all brightess on scales $< \lambda/d_{max}$ look like points,

Key technologies: following the signal path

- Antennas
- Receivers
- Down-conversion
- Measuring the correlations
- Spectral channels
- Calibration

Antennas collect radiation

- Specification, design and cost are frequency-dependent
  - High-frequency: steerable dishes (5 – 100 m diameter)
  - Low-frequency: fixed dipoles, yagis, ...
  - Ruze formula efficiency = $\exp[-(4\pi\sigma/\lambda)^2]$
  - Surface rms error $\sigma < \lambda/20$
  - sub-mm antennas are challenging (surface rms <25 μm for 12m ALMA antennas)

High frequency (ALMA)

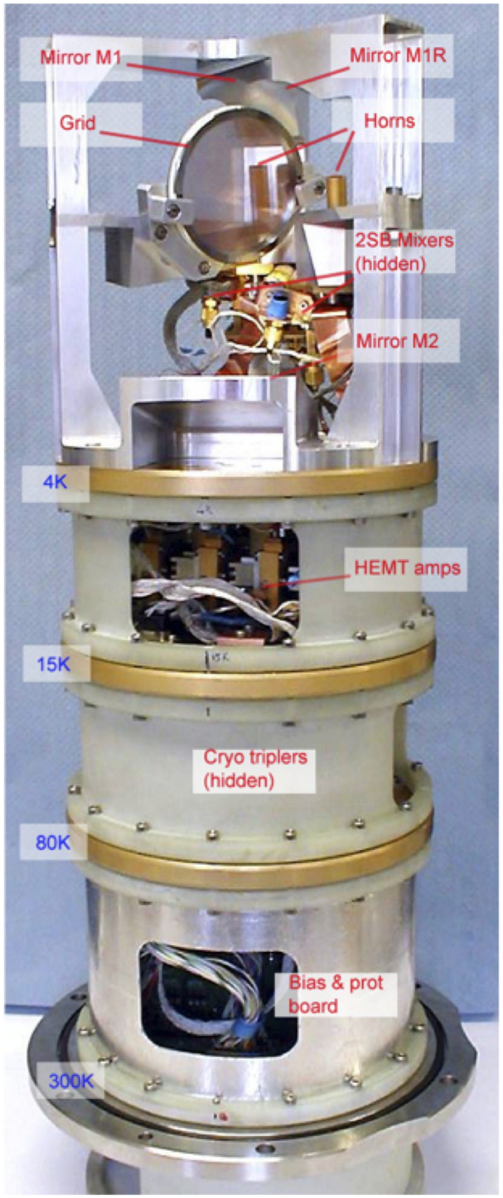
Low frequency (LOFAR)

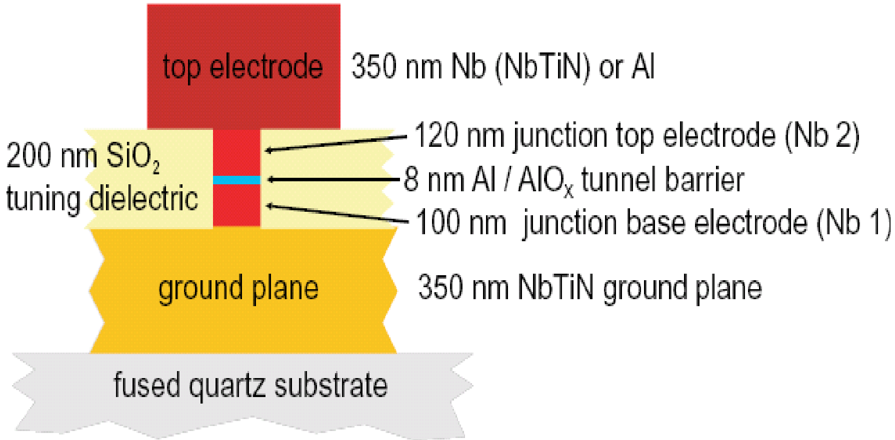
Receivers

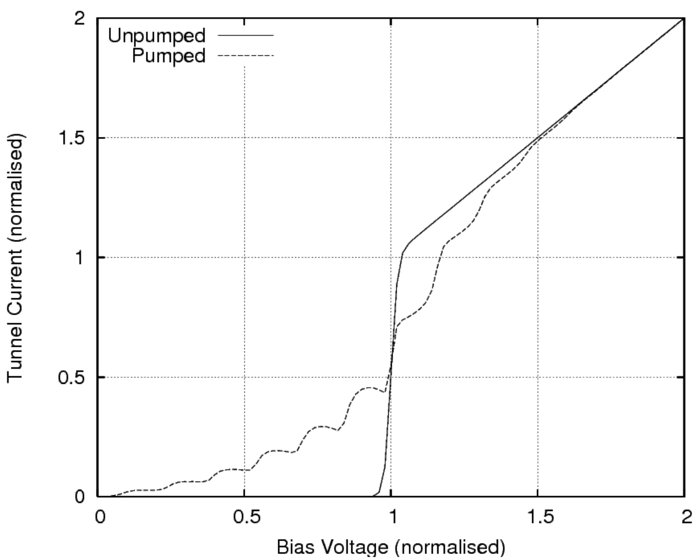
- Purpose is to detect radiation
- Cryogenically cooled for low noise (except at very low frequencies)
- Normally detect two polarization states
  - Separate optically, using wire grid or orthomode transducer
- Optionally, in various combinations:
  - Amplify RF signal (HEMT)
  - Then either:
    - digitize directly (possible up to ~10 GHz) or
    - mix with phase-stable local oscillator signal to make intermediate frequency (IF) → two sidebands (one or both used) → digitize
    - a mixer is a device with a non-linear voltage response
- Digitization typically 3 – 8 bit
- Send to correlator

SIS mixers for mm wavelengths

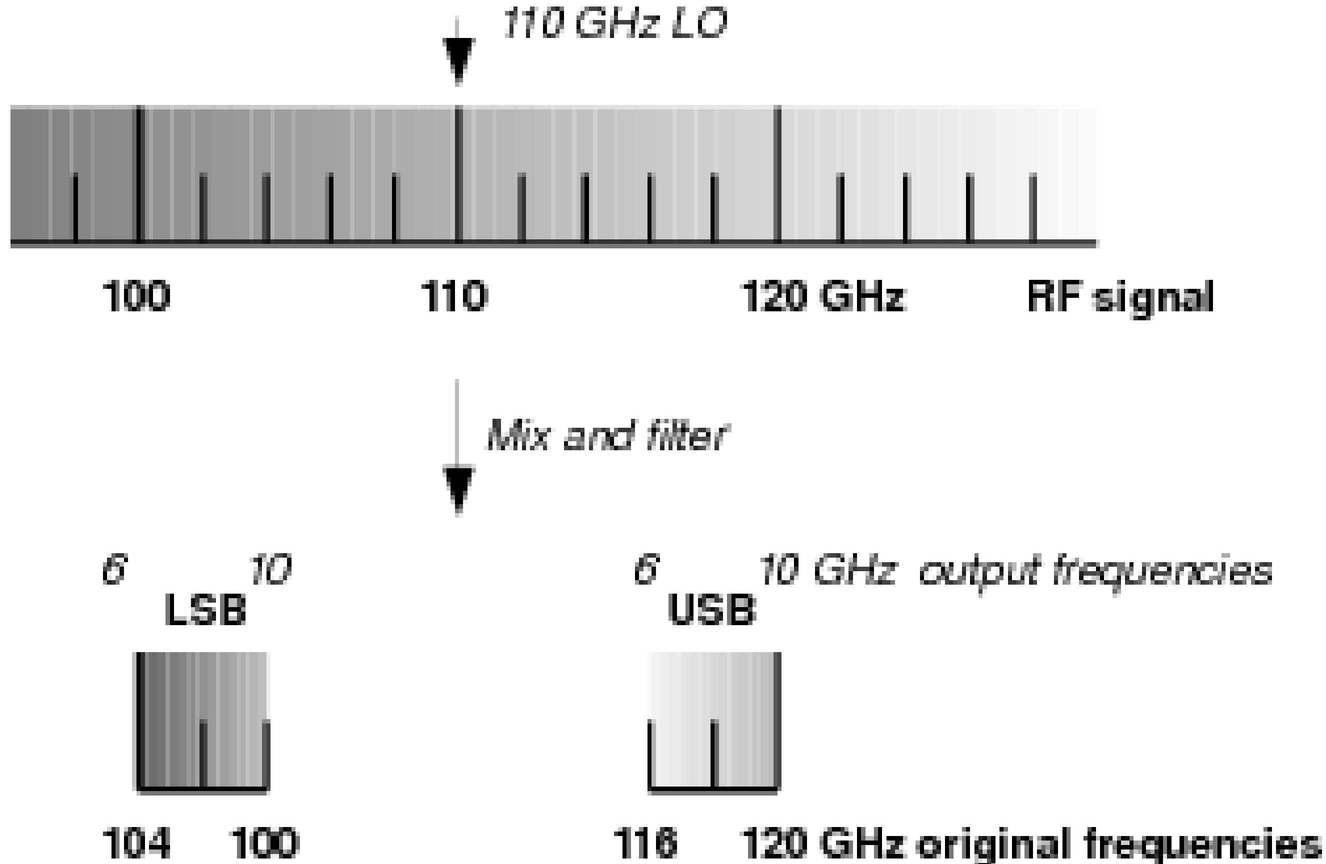
SIS = superconductor – insulator - superconductor







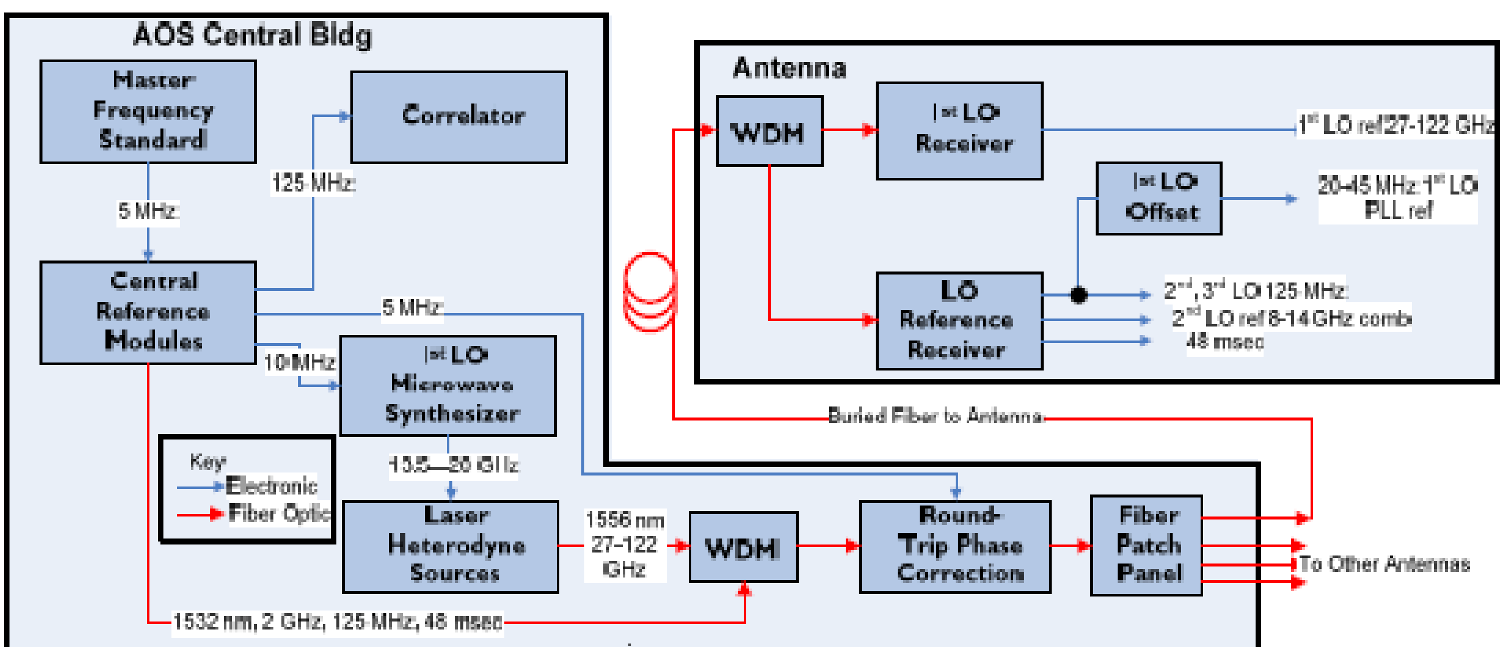
Sidebands



Note inverted order of LSB frequency labels

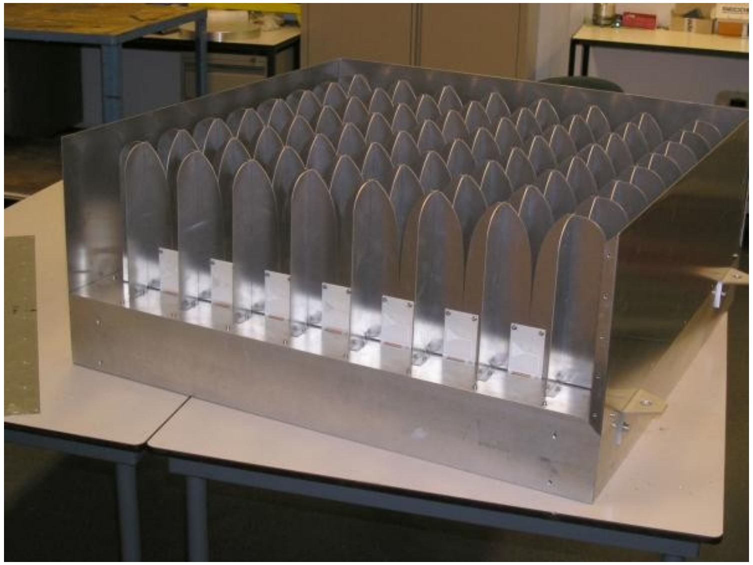
Local Oscillator Distribution

- These days, often done over fibre using optical analogue signal
- Master frequency standard (e.g. H maser)
- Must be phase-stable – round trip measurement
- Slave local oscillators at antennas
  - Multiply input frequency
  - Change frequency within tuning range
  - Apply phase shifts



Spatial multiplexing

- Focal-plane arrays (multiple receivers in the focal plane of a reflector antenna).
- Phased arrays with multiple beams from fixed antenna elements (“aperture arrays") e.g. LOFAR
  - a phased array is an array of antennas from which the signals are combined with appropriate amplitudes and phases to reinforce the response in a given direction and suppress it elsewhere
- Hybrid approach: phased array feeds (= phased arrays in the focal plane of a dish antenna).



Noise

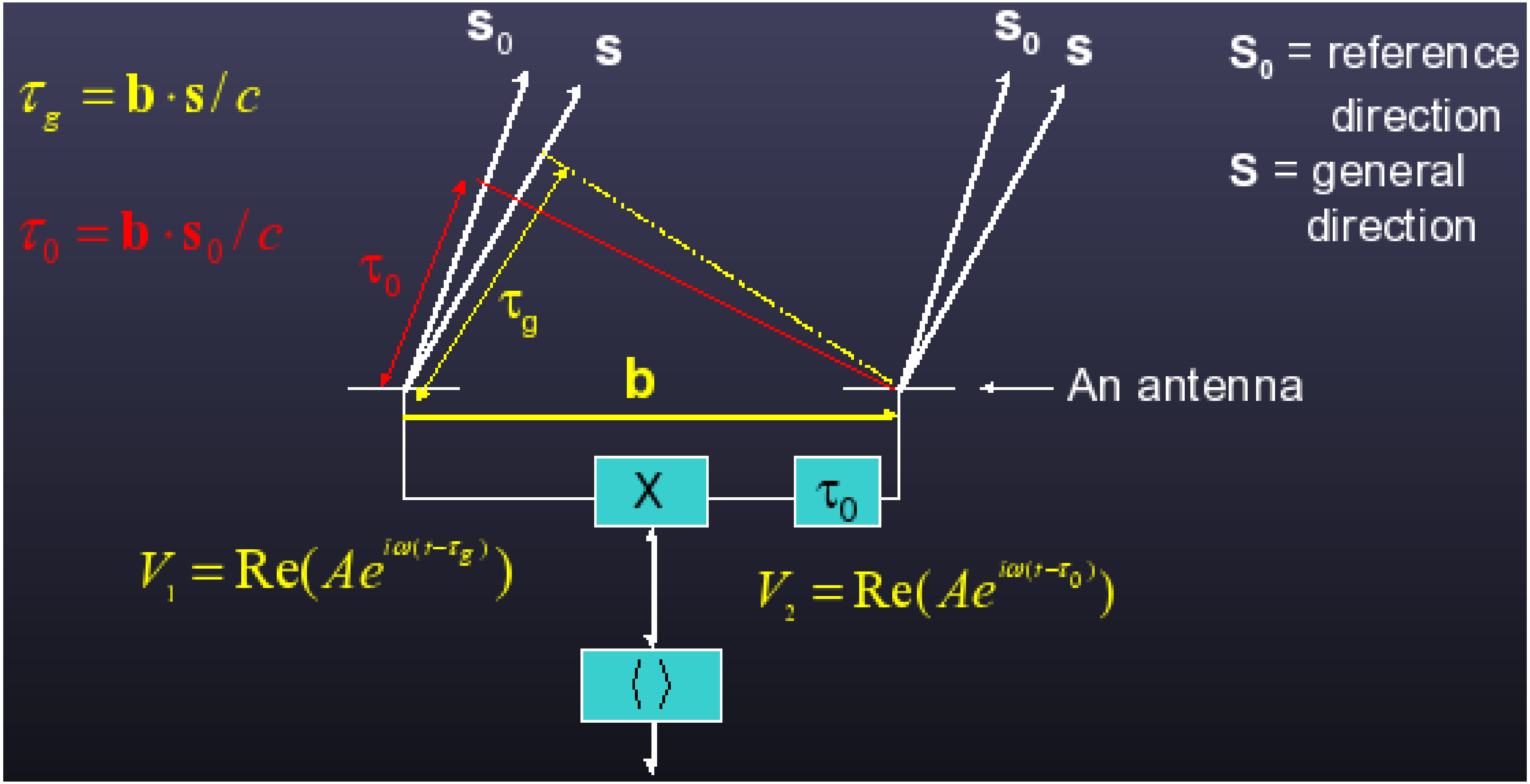
- RMS noise level S_{rms}
  - T_{sys} is the system temperature, A_{eff} is the effective area of the antennas, N_A is the number of antennas, $\Delta\nu$ is the bandwidth, t_{int} is the integration time and k is Boltzmann's constant
- For good sensitivity, you need low T_{sys} (receivers), large A_{eff} (big, accurate antennas), large N_A (many antennas) and, for continuum, large bandwidth $\Delta\nu$.
- Best T_{rec} typically a few x 10K from 1 – 100 GHz, ~100 K at 700 GHz. Atmosphere dominates T_{sys} at high frequencies; foregrounds at low frequencies.

$$S_{rms} = \frac{2kT_{sys}}{A_{eff}\sqrt{N_A(N_A - 1)t_{\mathrm{int}}\Delta\nu}}$$

Delay

- An important quantity in interferometry is the time delay in arrival of a wavefront (or signal) at two different locations, or simply the delay, τ.
  - Constant (“cable”) delay in waveguide or electronics
  - geometrical delay
  - propagation delay through the atmosphere
- Aim to calibrate and remove all of these accurately
- Phase varies linearly with frequency for a constant delay
  - $\Delta\varphi = 2\pi\tau\Delta\nu$

Delay Tracking



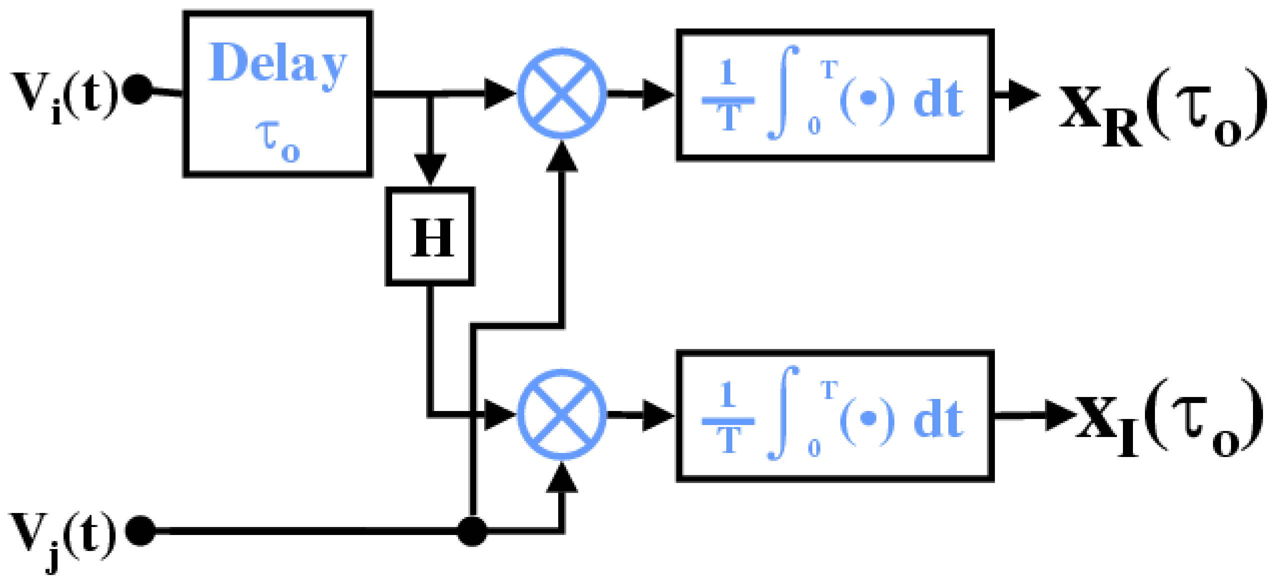
The geometrical delay τ_0 for the delay tracking centre can be calculated accurately from antenna position + Earth rotation model.

Works exactly only for the delay tracking centre. Maximum averaging time is a function of angle from this direction.

What does a correlator do?

Takes digitized signals from individual antennas; calculates complex visibilities for each baseline

Antenna i



Real

Imaginary

Antenna j

- $x_R = x_{ij}(\tau_0)$
- $x_I = x_{ij}(\tau_0 + \Delta\tau)$, **with** $\Delta\tau = 1/\left(4\nu_0\right)$ **(**$\Delta\phi = 90°$**).**

Spectroscopy

- We make multiple channels by correlating with different values of **lag**, τ. This is a delay introduced into the signal from one antenna with respect to another as in the previous slide. For each quasi-monochromatic frequency channel, a lag is equivalent to a phase shift $2\pi\tau\nu$, i.e.

  $V(u,v,\tau) = \int V(u,v,\nu)\exp(2\pi i\tau\nu)\, d\nu$
- This is another Fourier transform relation with complementary variables ν and τ, and can be inverted to extract the desired visibility as a function of frequency.
- In practice, we do this digitally, in finite frequency channels:

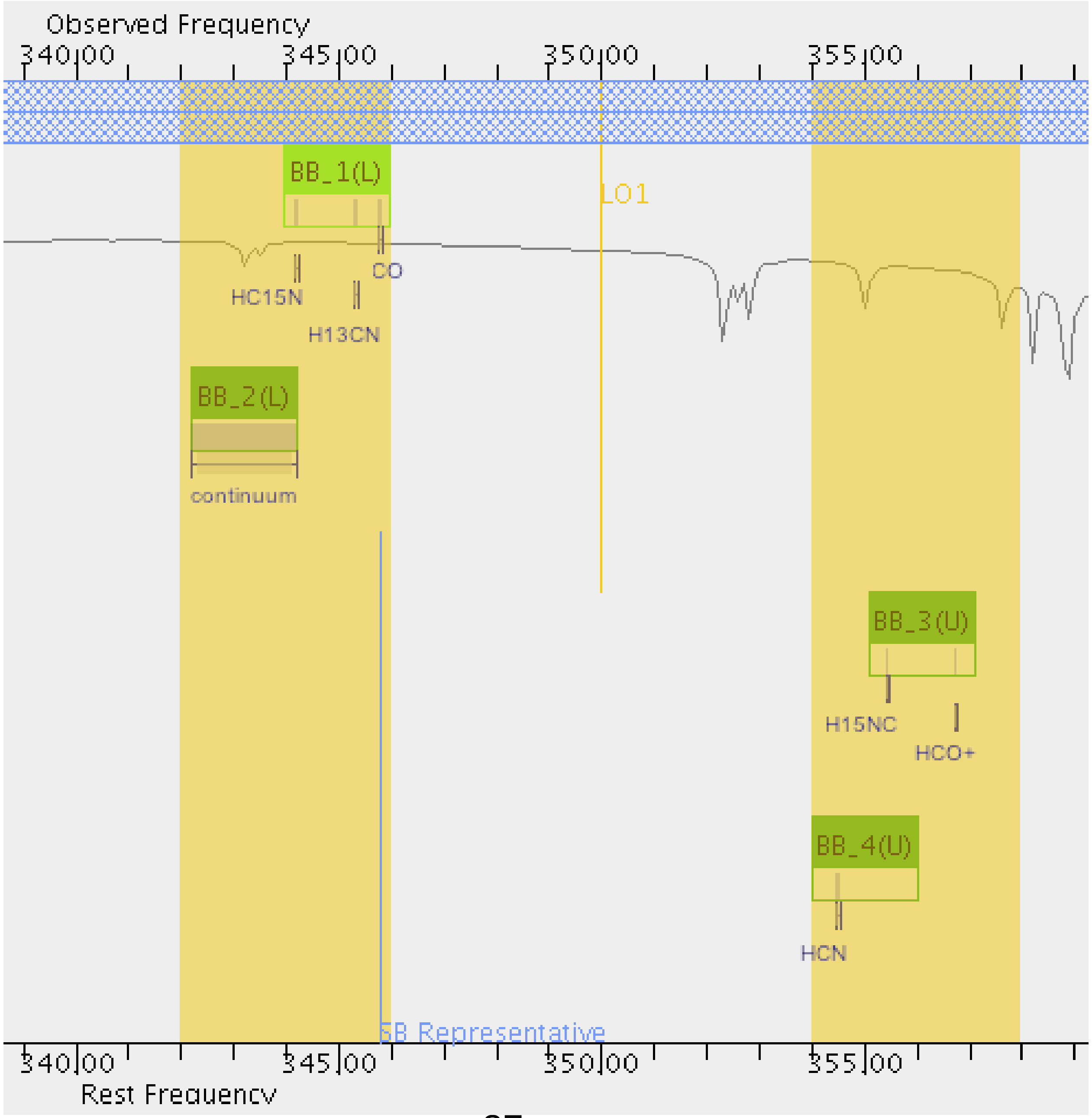
$$V(u,v,j\Delta\nu) = \Sigma_k V(u,v, k\Delta\tau) \exp(-2\pi ijk\Delta\nu\Delta\tau)$$

- Each spectral channel can then be imaged (and deconvolved) individually. The final product is a **data cube**, regularly gridded in two spatial and one spectral coordinate.

Correlation in practice

- Correlators are powerful, but special-purpose computers
  - ALMA correlator 1.6 x 10^{16} operations/s
- Implement using FPGA's or on general-purpose computing hardware + GPU's, depending on application
- Alternative architectures:
  - XF (multiply, then Fourier transform)
  - FX (Fourier transform, then multiply)
  - Hybrid (digital filter bank + XF)
- Most modern correlators are extremely flexible, allowing many combinations of spectral resolution, number of channels and bandwidth.

An example: ALMA spectral setup



Calibration Overview

- What we want from this step of data processing is a set of perfect visibilities $V(u,v,w,\nu)$, on an absolute amplitude scale, measured for exactly known baseline vectors (u,v,w), for a set of frequencies, ν, in full polarization.
- What we have is the output from a correlator, which contains signatures from at least:
  - the Earth's atmosphere
  - antennas and optical components
  - receivers, filters, amplifiers
  - digital electronics
  - leakage between polarization states
- Basic idea:
  - Apply a priori calibrations
  - Measure gains for known calibration sources as functions of time and frequency; interpolate to target
  - Calibrator-target-calibrator cycle can be as short as 10s

Antenna-dependent calibration

- Rationale: most corrupting effects are antenna, not baseline-dependent, e.g.
  - atmosphere
  - antenna optics
  - receiver gain
  - ....
- Therefore associate a complex gain with each antenna. In general, this will be time, frequency, polarization and direction-dependent
- This makes the calibration problem overdetermined, since we have N antennas and N-1 phases, but N(N-1)/2 independent measurements: high precision for a large-N array. Reduce complexity, improve accuracy.

Calibrations (1)

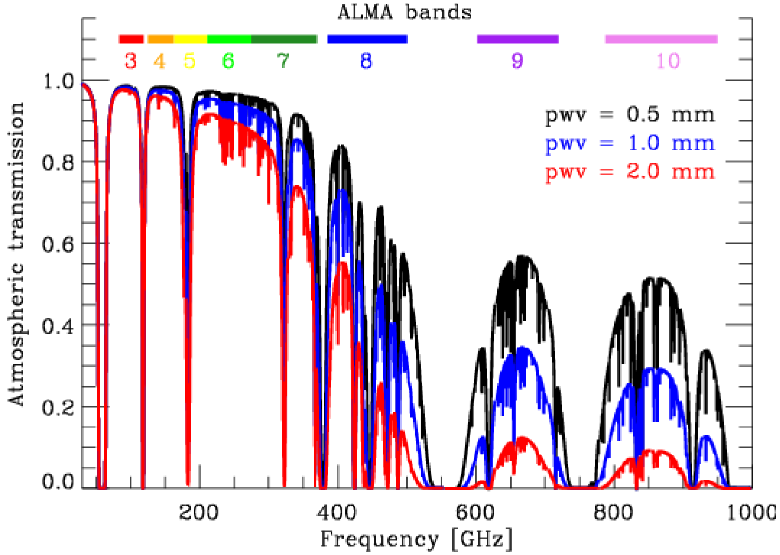
- Array calibrations (these have to be at least approximately right; small errors can be dealt with)
  - Antenna pointing and focus model (absolute or offset)
  - Correlator model (antenna locations, Earth orientation and rate, clock, ...)
  - Instrumental delays
  - Model of the geomagnetic field
- Calibrator models
  - Positions (proper motions, ephemerides, ...)
  - Flux density and polarization (IQUV) as functions of frequency (continuum spectrum, lines if relevant)
  - Images if not point-like
- Previously measured/derived from simulations/.. and applied post hoc
  - Antenna gain curve/voltage pattern as function of elevation, temperature, frequency, .
  - Receiver non-linearity correction, quantization corrections, ...

Calibrations (2)

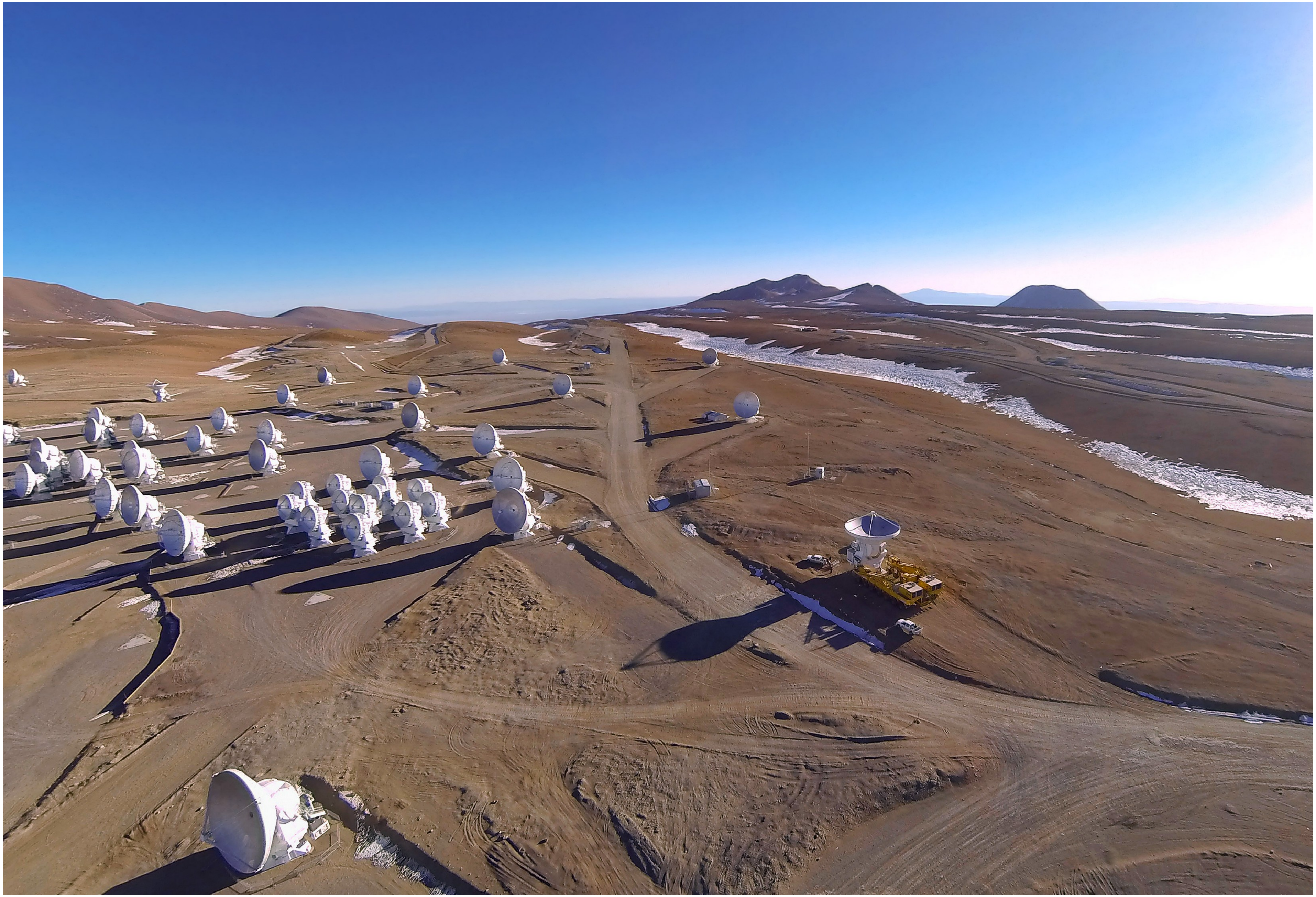
- Total-power monitoring – receiver stability and atmospheric opacity
  - atmosphere + load (mm/sub-mm)
  - continuous, switched noise source (cm/m)
  - Can also be used to monitor phase difference between polarizations
- Auxiliary
  - Water-vapour radiometer (atmospheric path, opacity contribution)
  - Multifrequency satellite (e.g. GPS) signals (electron content)
- On-line calibrations
  - Pointing offsets
  - Focus
  - Delay (from phase slope across band)
- Interpolation from calibrator to target
  - Complex gain (amplitude + phase) - fast
  - Bandpass (gain as a function of frequency) - slow

ALMA

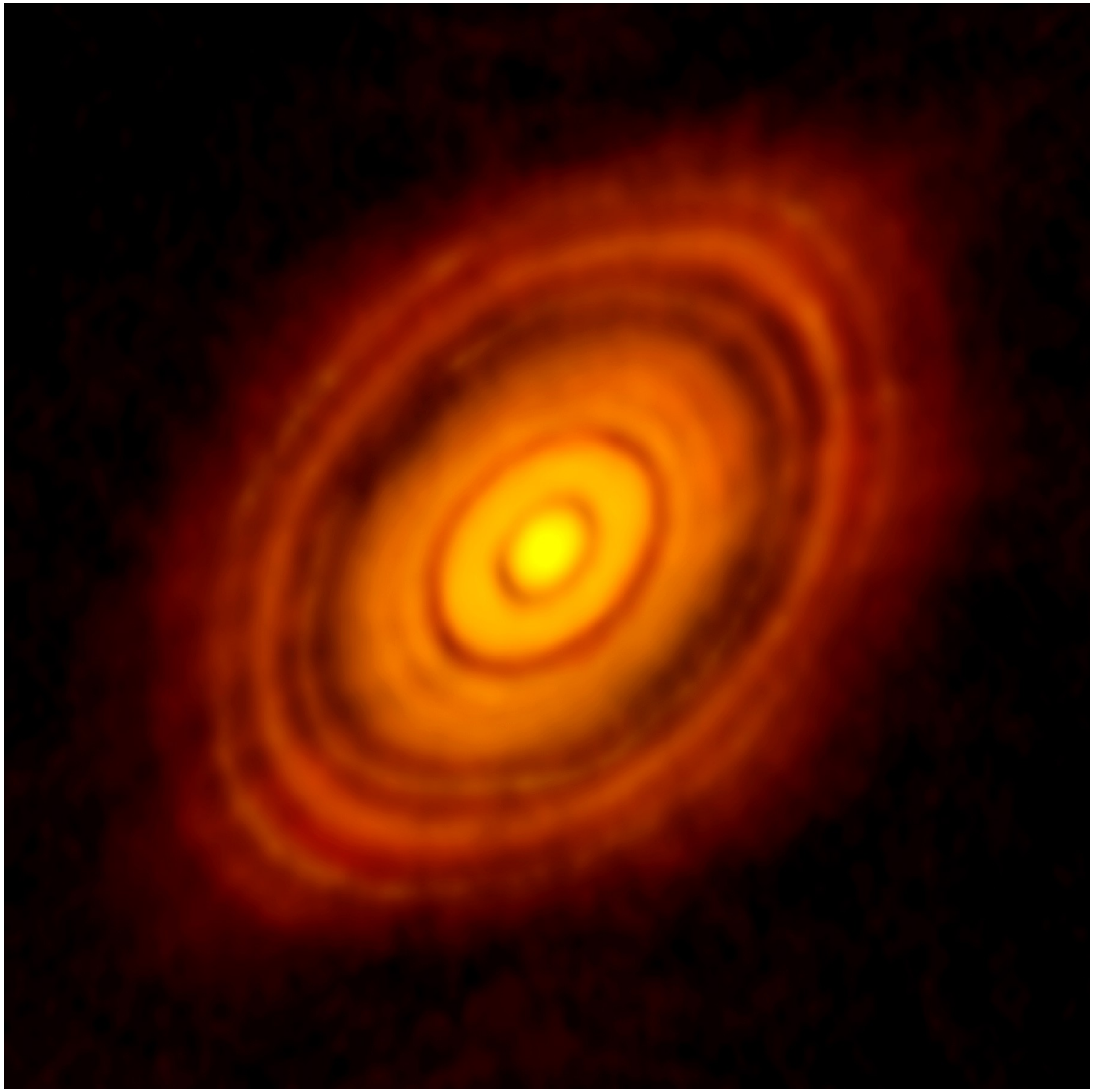
- Atacama Large Millimetre/Submillimetre Array
  - International Project (Europe/North America/East Asia)
  - Optimised for frequency range 35 – 950 GHz
  - 10 receiver bands
  - 54 12m and 12 7m antennas; <25 μm rms surface accuracy; 0.6 arcsec rms offset pointing
  - Reconfigurable; maximum baseline 16 km (resolution 0.005 arcsec at 950 GHz)
  - High, dry site



ALMA



Dust emission from a protoplanetary disk

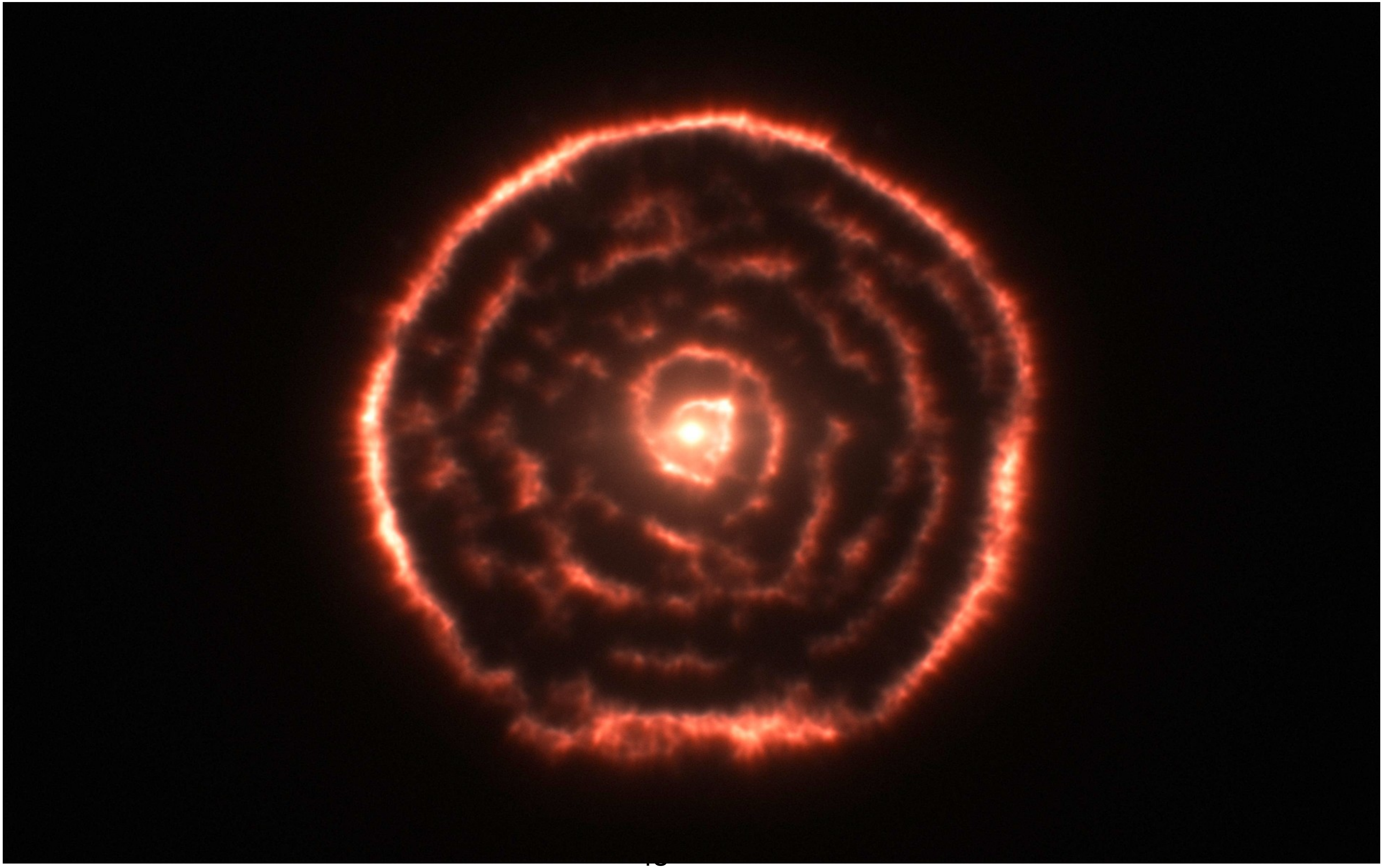


HL Tau

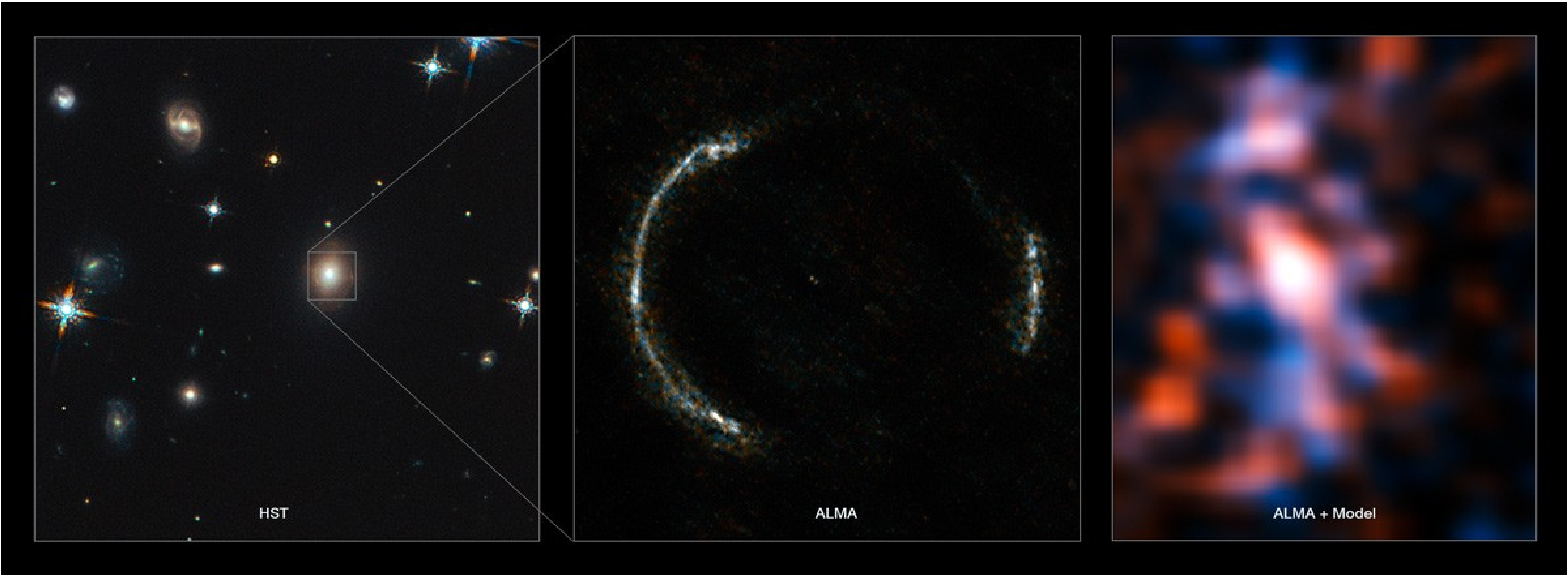
ALMA

1.3 mm

CO emission from an old star

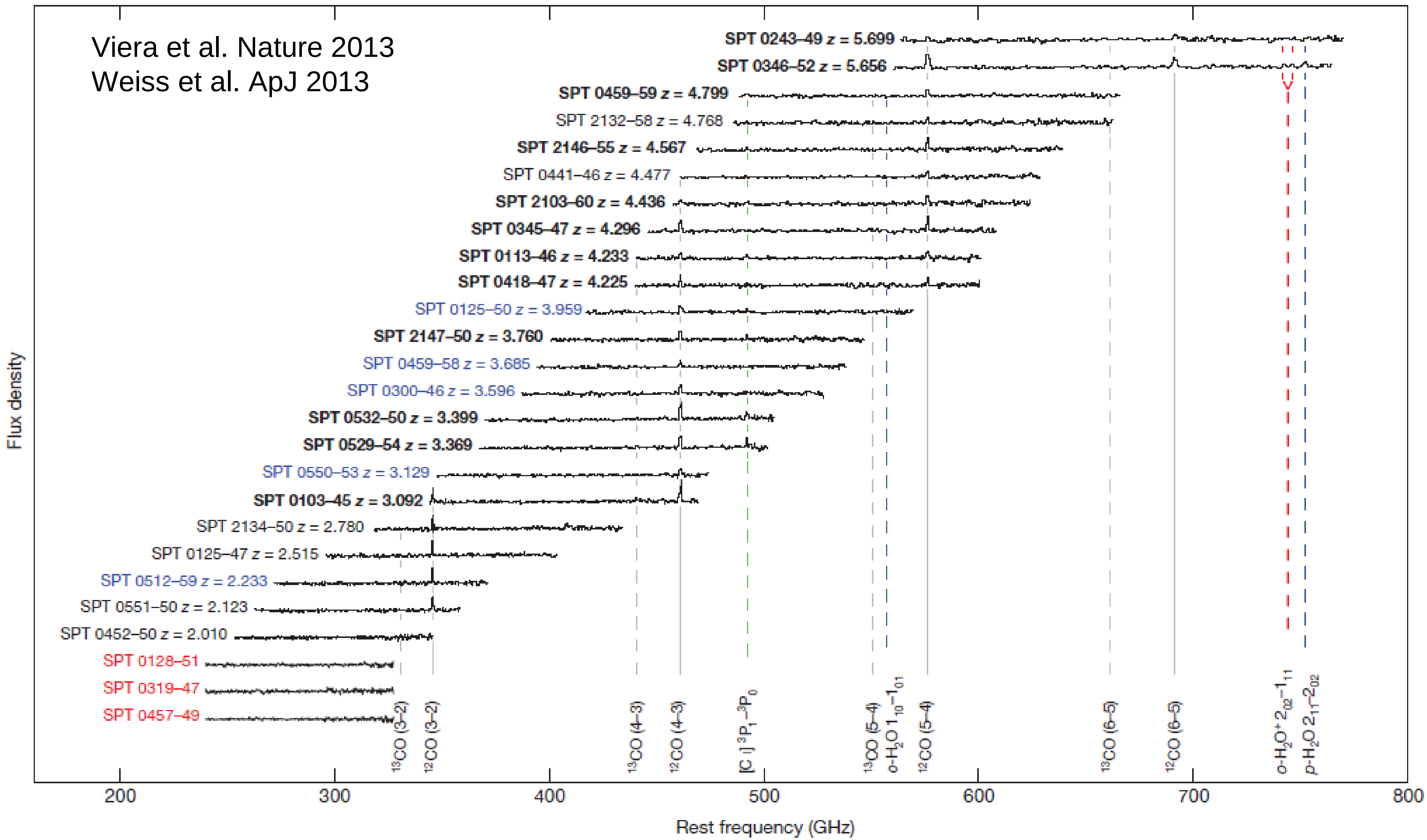


Gravitationally lensed dust emission



Lensing galaxy Lensed image Model

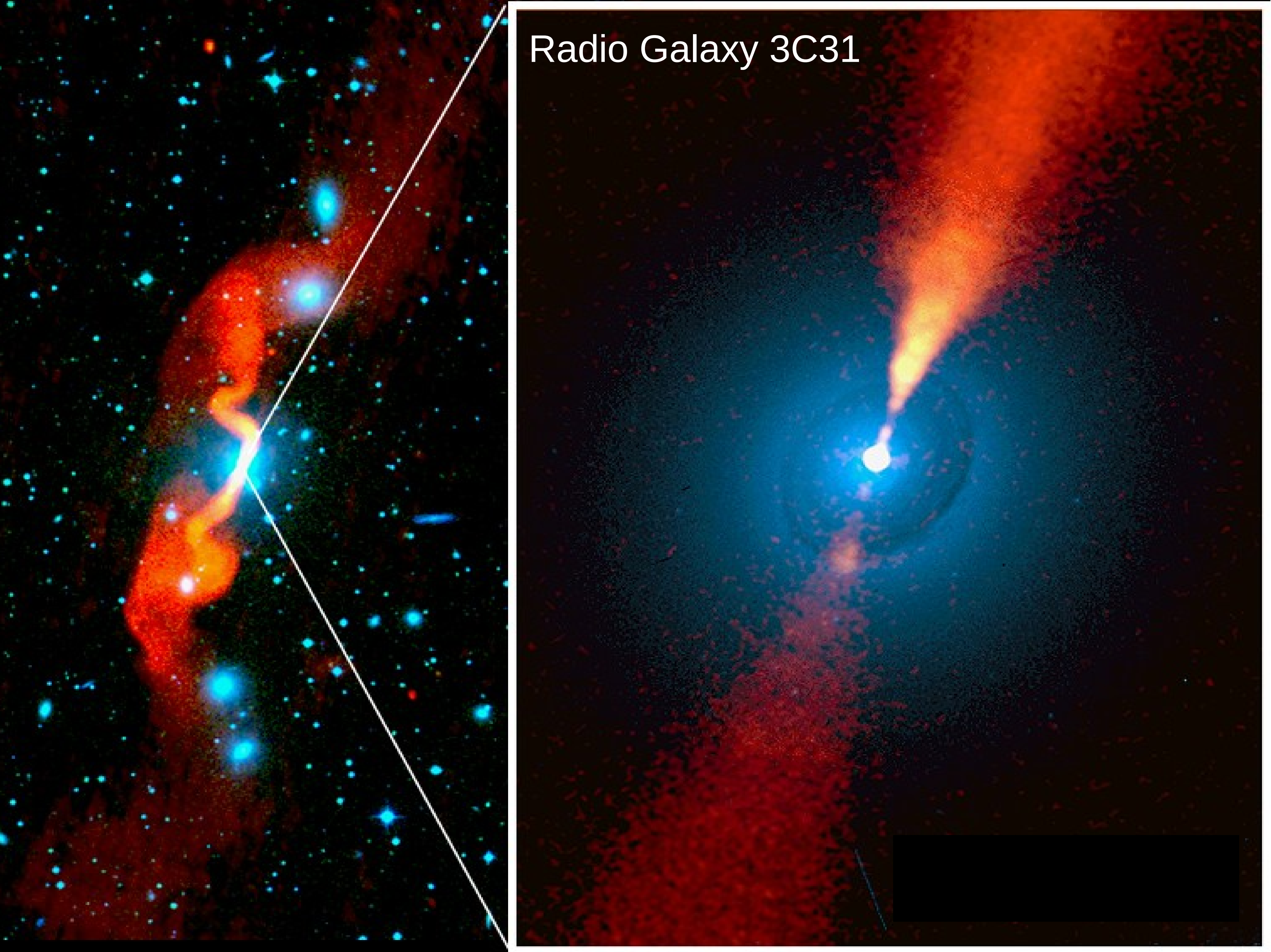
Water in early galaxies



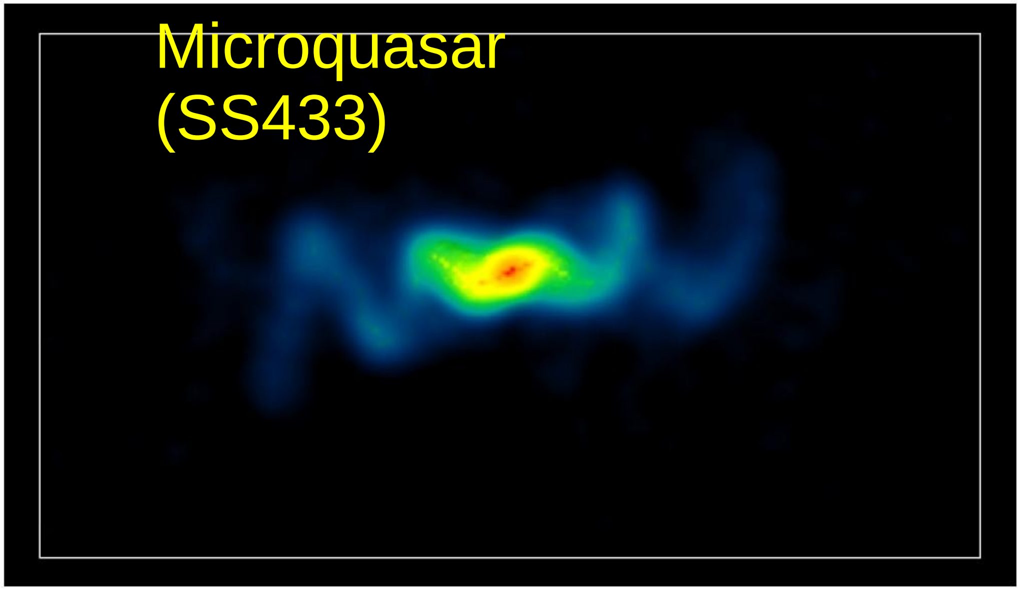
VLA

Karl G Jansky VLA

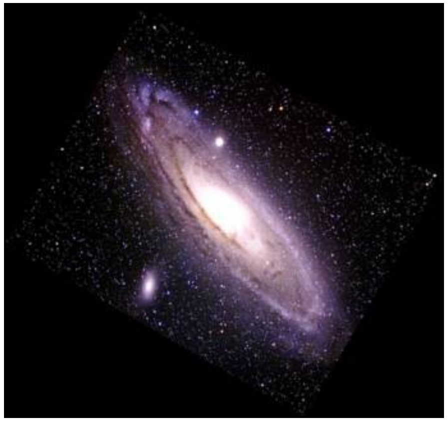
- Original VLA started full operations in 1980
- 27 25m antennas; reconfigurable on railway tracks
- Maximum baseline 35 km
- Major upgrade just completed
  - More sensitive receivers
  - Complete frequency coverage 1 – 50 GHz
  - Wider bandwidth (fibre optic data transport) – up to 8 GHz
  - Much more powerful correlator

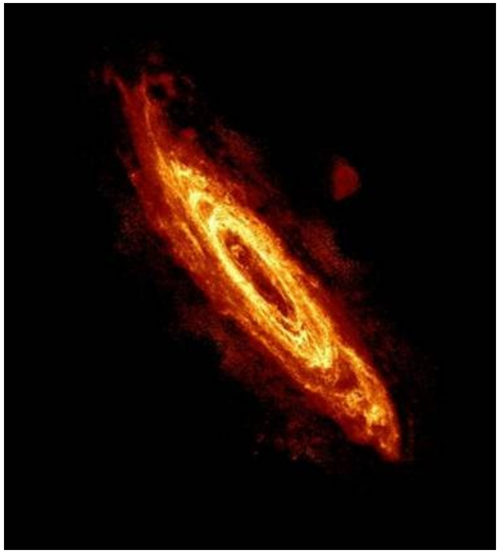

Radio Galaxy 3C31

Galactic synchrotron sources



Neutral Hydrogen





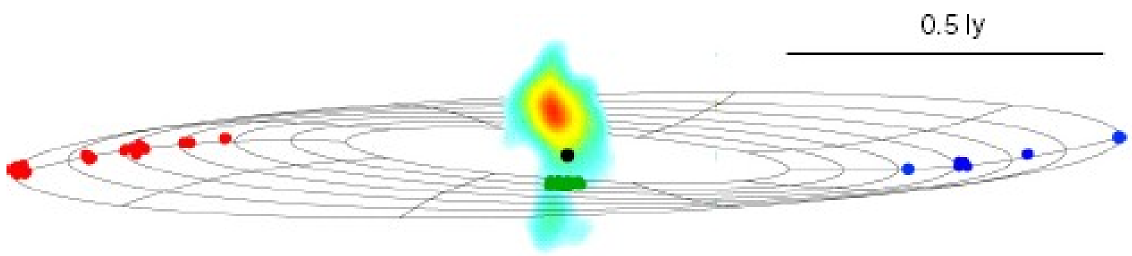
Optical HI

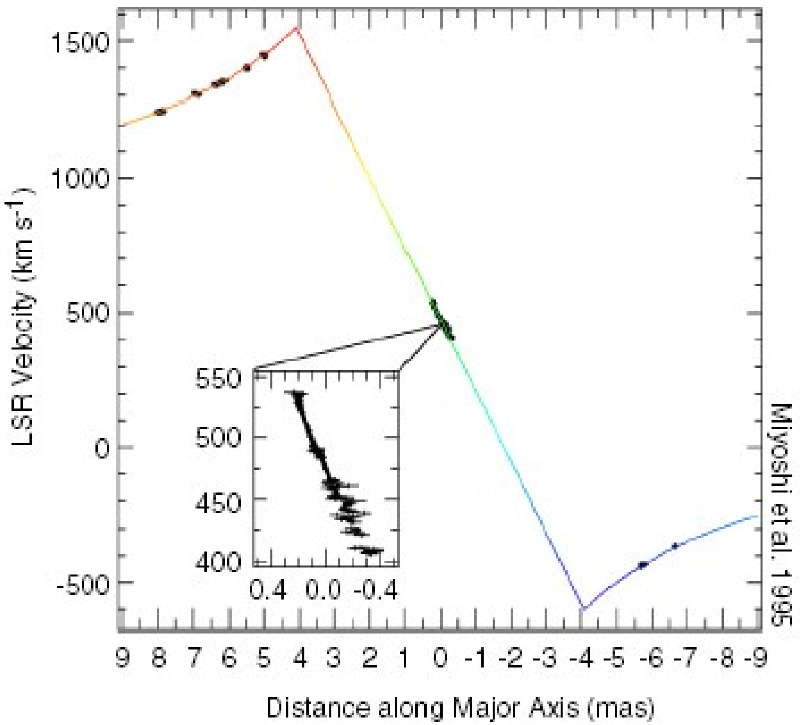
Andromeda galaxy

Very Long Baseline Interferometry

- Antennas not physically connected
- Baselines up to Earth diameter (or even larger for satellites)
  - Radioastron satellite
- Record data from individual antennas; correlate later
  - Requires very precise frequency standard (H maser)
  - Poorly known delays → "fringe search" on calibrator

Measurement of Black Hole Mass in NGC 4258

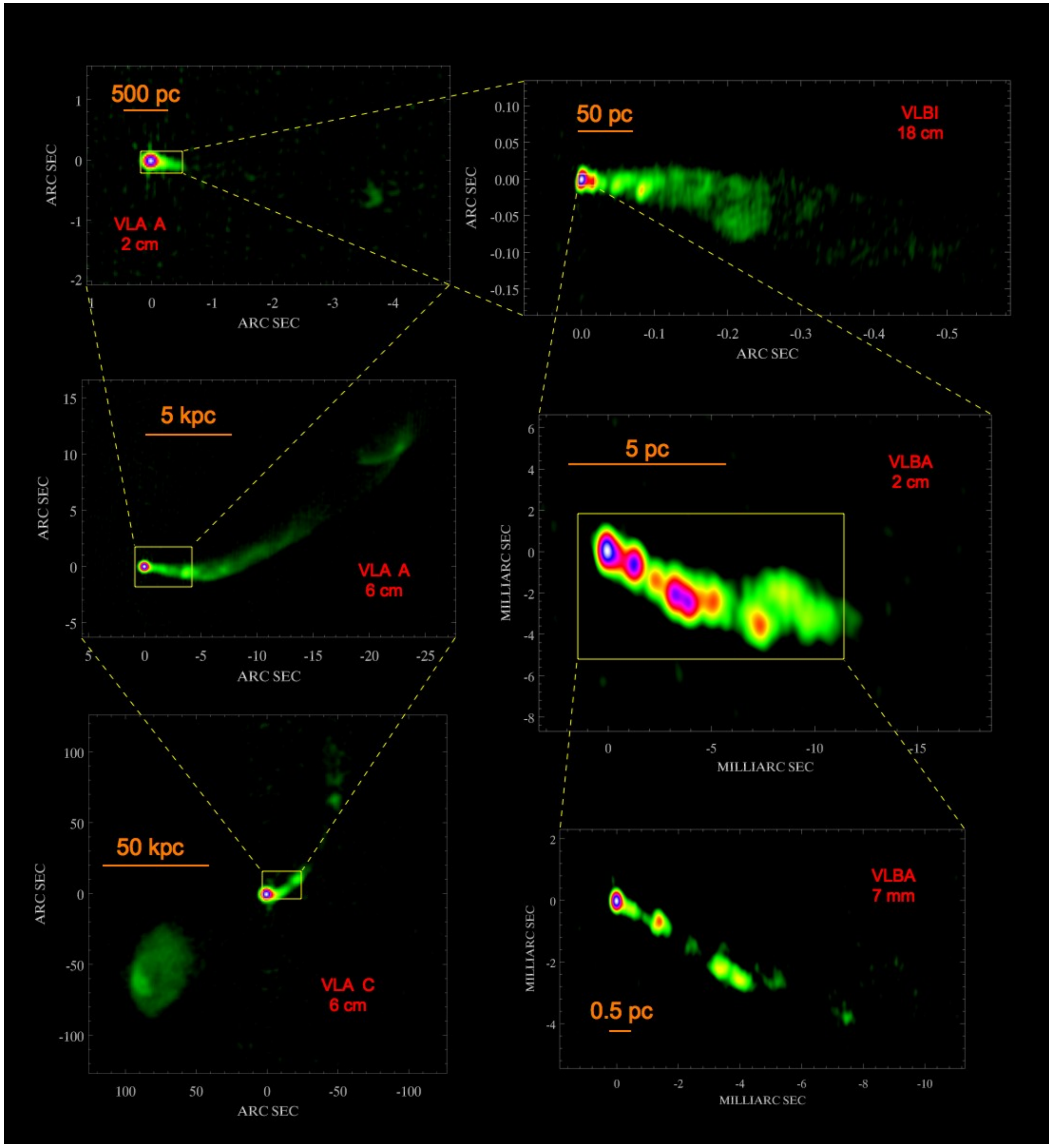


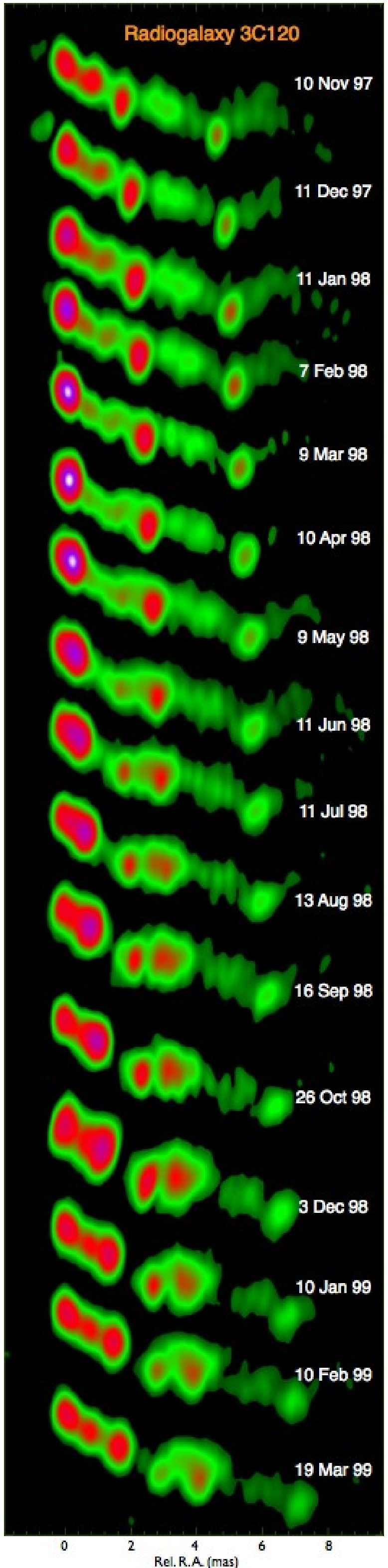


VLBI observations of water masers at 22 GHz

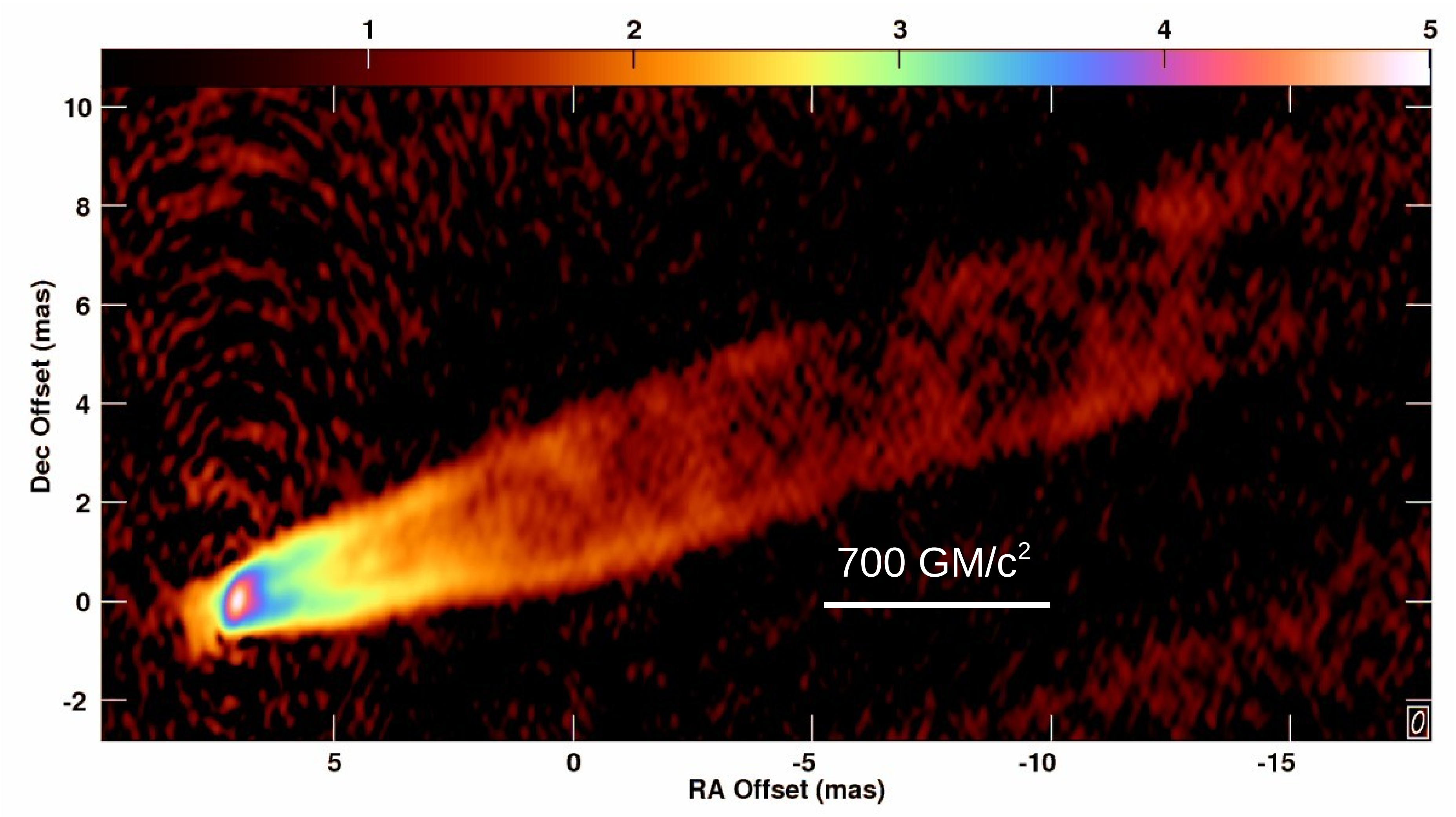
Rotation curve gives mass of central black hole $= 3.5 \times 10^7$ solar masses

Superluminal Motion



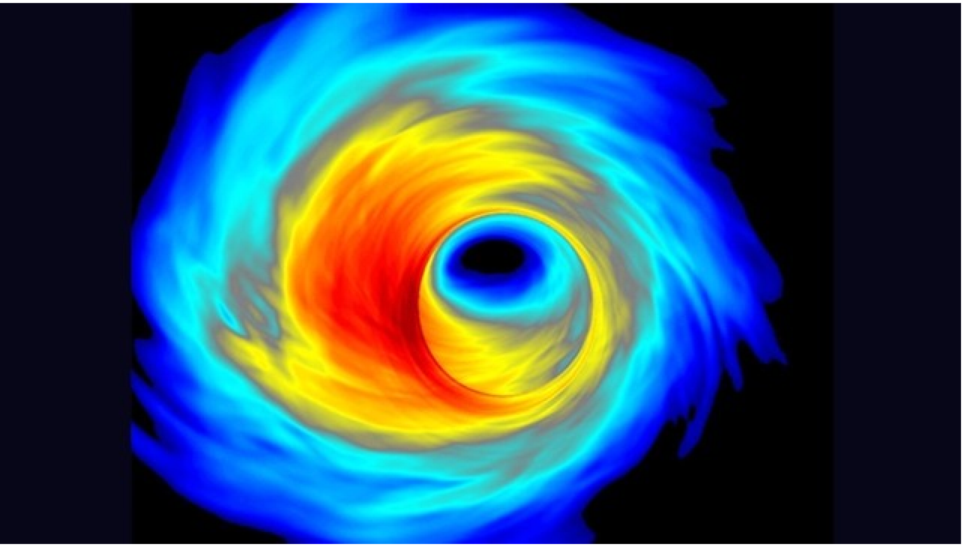


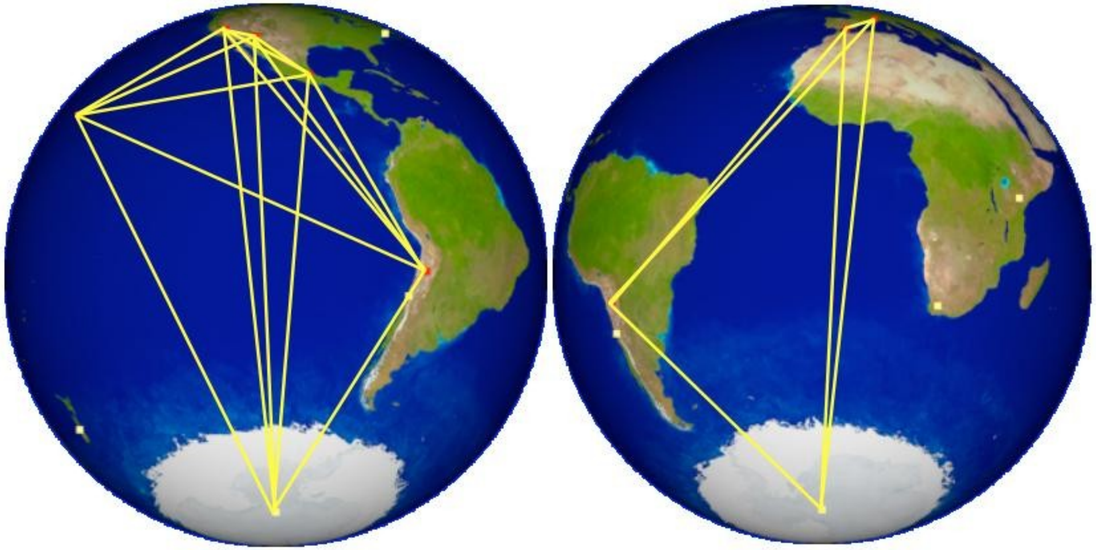
M87: close to the black hole



VLBI image at 7mm wavelength
The jets are collimated on scales <50 times the radius of the black hole event horizon

Imaging a black hole event horizon





To resolve the largest black hole event horizon (at the Galactic Centre) we need a resolution of about 20 µarcsec,
$\lambda/d \approx 10^{-10}$ rad. We can do this with d = 10000 km and λ = 1 mm – now feasible using ALMA as a phased array.

Watch this space.