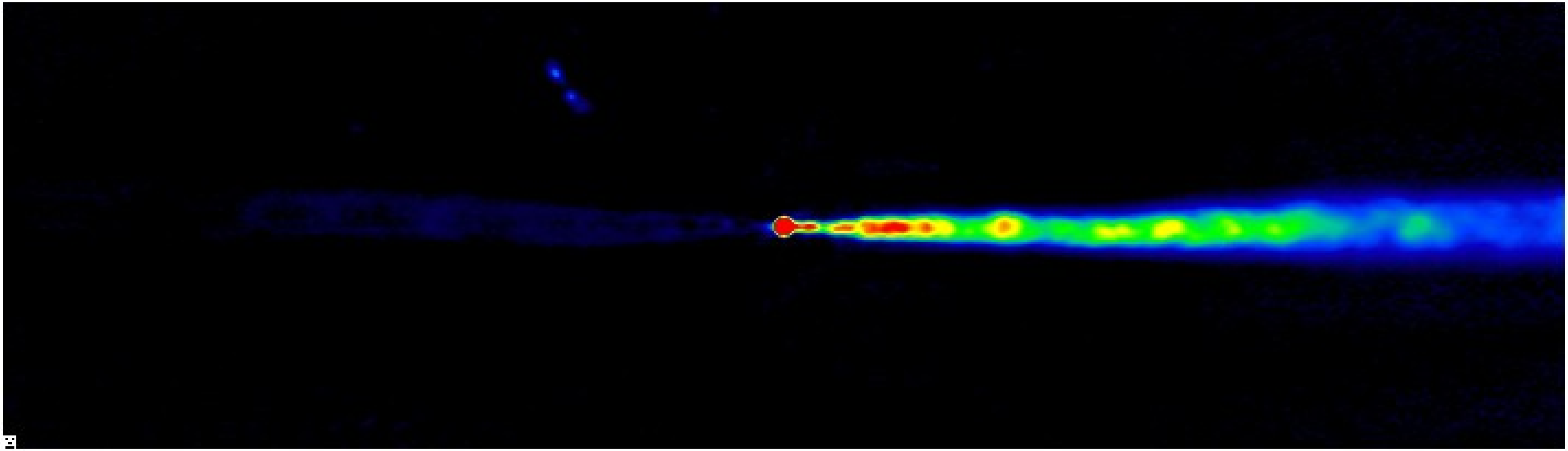
Fundamentals of Radio Interferometry

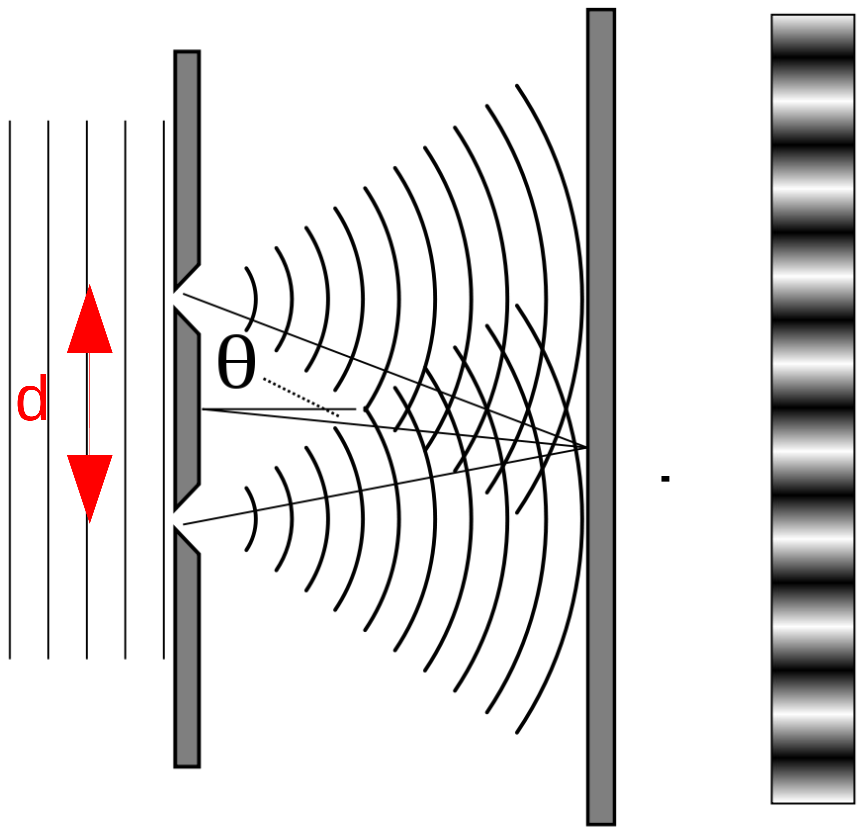
Robert Laing (ESO)



Objectives

- A more formal approach to radio interferometry using coherence functions
  - A complementary way of looking at the technique
  - Simplifying assumptions
- Relaxing the simple assumptions
- How does a radio interferometer work?
  - Follow the signal path
  - Technologies for different frequency ranges

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 and return to them later.

The ideal interferometer (1)

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

  The coefficients $\mathbf{E}_\nu(\mathbf{R})$ are complex vectors (amplitude and phase; two polarizations)
- Simplification 1: radiation is monochromatic

$\mathbf{E}_\nu(\mathbf{r}) = \iiint P_\nu(\mathbf{R},\mathbf{r}) \mathbf{E}_\nu(\mathbf{R})\, dx\, dy\, dz$ where $P_\nu(\mathbf{R},\mathbf{r})$ is the propagator

- Simplification 2: scalar field (ignore polarization for now)
- Simplification 3: sources are all very far away
- This is equivalent to having all sources at a fixed distance – there is no depth information

The ideal interferometer (3)

- 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 $|\mathbf{R}|$)
- 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 and

  ***** means complex conjugation.
- Simplification 5: radiation from astronomical objects is not spatially coherent (“random noise”)
  - $\langle E_\nu(\mathbf{R}_1)E^{*}_\nu(\mathbf{R}_2)\rangle = 0$ unless $\mathbf{R}_1 = \mathbf{R}_2$

The ideal interferometer (4)

- 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

- No requirement to measure the coherence function in real time if you can store the field measurements (VLBI)

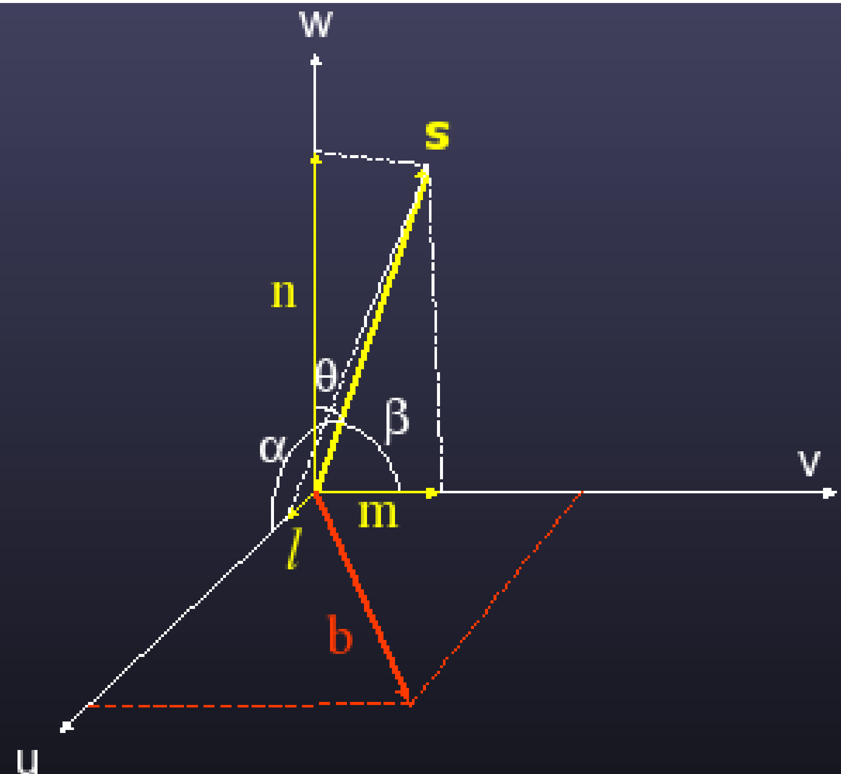
u, v, w and direction cosines

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,

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

The Fourier Relation

- Simplification 6: receiving elements have no direction dependence
- Simplification 7: all sources are in a small patch of sky
- Simplification 8: we can measure at all values of $\mathbf{r}_1$-$\mathbf{r}_2$ and at all times
- 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 $I_\nu(l,m)$

Fourier Inversion

- This relation can be inverted to get the intensity distribution, which is what we want:

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

- This is the fundamental equation of synthesis imaging.
- Interferometrists love to talk about the (u,v) plane. Remember that u, v (and w) are measured in wavelengths.
- The vector (u,v,w) = $(\mathbf{r}_1 - \mathbf{r}_2)/\lambda$ is the **baseline**

Simplification 1

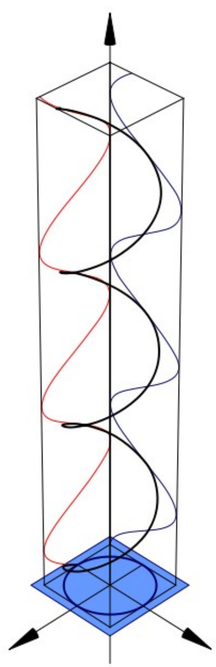
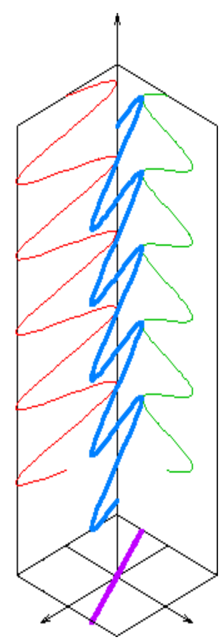
- Radiation is monochromatic
  - We observe wide bands both for spectroscopy (HI, molecular lines) and for sensitive continuum imaging, so we need to get round this restriction.
  - In fact, we can easily divide the band into multiple spectral channels (details later)
  - There are imaging restrictions if the individual channels are too wide for the field size – see imaging lectures.
    - Usable field of view < $(\Delta\nu/\nu_0)(l^2+m_2)^{1/2}$.
    - Not usually an issue for modern correlators

Simplification 2

- Radiation field is a scalar quantity
  - The field is actually a vector and we are interested in both components (i.e. its polarization).
  - This makes no difference to the analysis as long as we measure two states of polarization (e.g. right and left circular, or crossed linear) and account for the coupling between states.
  - Use the **measurement equation** formalism for this (calibration and polarization lectures).

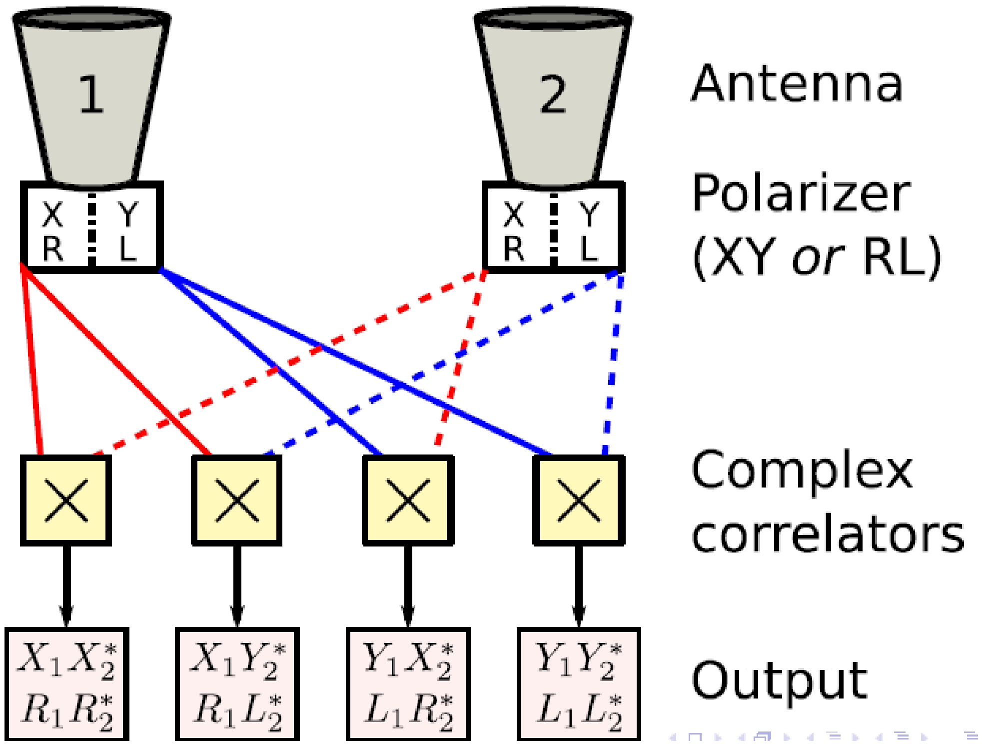
Polarization

- Want to image Stokes parameters:
  - I (total intensity)
  - Q, U (linear)
  - V (circular)
- Resolve into two (nominally orthogonal) polarization states, either right and left circular or crossed linear.

Stokes Parameters and Visibilities

- This assumes the perfect case:
  - no rotation on the sky (not true for most arrays, but can be calculated and corrected)
  - perfect system
- Circular basis
  - $V_I = V_{RR} + V_{LL}$
  - $V_Q = V_{RL} + V_{LR}$
  - $V_U = i(V_{RL} - V_{LR})$
  - $V_V = V_{RR} - V_{LL}$
- Linear basis
  - $V_I = V_{XX} + V_{YY}$
  - $V_Q = V_{XX} - V_{YY}$
  - $V_U = V_{XY} + V_{YX}$
  - $V_V = i(V_{XY} - V_{YX})$



4 coherence functions

$V_{RR} = \langle R_1 R_2^* \rangle$ and so on

Simplifications 3 and 4

- Sources are all a long way away
  - Strictly speaking, in the far field of the interferometer, so that the distance is $>D^2/\lambda$, where D is the interferometer baseline
  - True except in the extreme case of very long baseline observations of solar-system objects
- Radiation is not spatially coherent
  - Generally true, even if the radiation mechanism is itself coherent (masers, pulsars)
  - May become detectable in observations with very high spectral and spatial resolution
  - Coherence can be produced by scattering, since signals form the **same** location in a sources are spatially coherent, but travel by **different** paths through interstellar or interplanetary medium.

Simplifications 5 and 6

- Space between us and the source is empty
- The receiving elements have no direction-dependence
- Closely related and not true in general. Examples:
  - Interstellar or interplanetary scattering
  - Tropospheric and (especially) ionospheric fluctuations which lead to path/phase and amplitude errors, sometimes seriously direction-dependent
  - Ionospheric Faraday rotation, which changes the plane of polarization
  - Antennas are usually highly directional by design
- Standard calibration deals with the case that there is no direction-dependence (i.e each antenna has a single, time-variable complex gain)
- Direction dependence is becoming more important, especially for low frequencies and wide fields.

A special case: primary beam correction

- If the response of the antenna is direction-dependent, then we are measuring

$I_\nu(l,m)\ D_{1\nu}(l,m)D^*_{2\nu}(l,m)$ instead of $I_\nu(l,m)$ (ignore polarization for now)

- An easier case is when the antennas all have the same response

$A_\nu(l,m) = |D_\nu(l,m)|^2$

- In this case, $V'_\nu(u,v) = \iint A_\nu(l,m) I_\nu(l,m) \exp[-2\pi i(ul+vm)]\ dl\ dm$
- We just make the standard Fourier inversion and then divide by the **primary beam** $A_\nu(l,m)$

Simplification 7

- The field of view is small
  - (or antennas are in a single plane)
- Not always true
  - Basic imaging equation becomes:

$$V\nu(u,v,w) = \iint I_\nu(l,m) \{\exp[-2\pi i(ul+vm+(1-l^2-m^2)^{1/2}w)]/(1-l^2-m^2)^{1/2}\}\, dl\, dm$$

  - No longer a 2D Fourier transform, so analysis becomes more complicated (the “w term”)
  - Map individual small fields (“facets”) and combine later
  - w-projection
  - See imaging lectures

Simplification 8

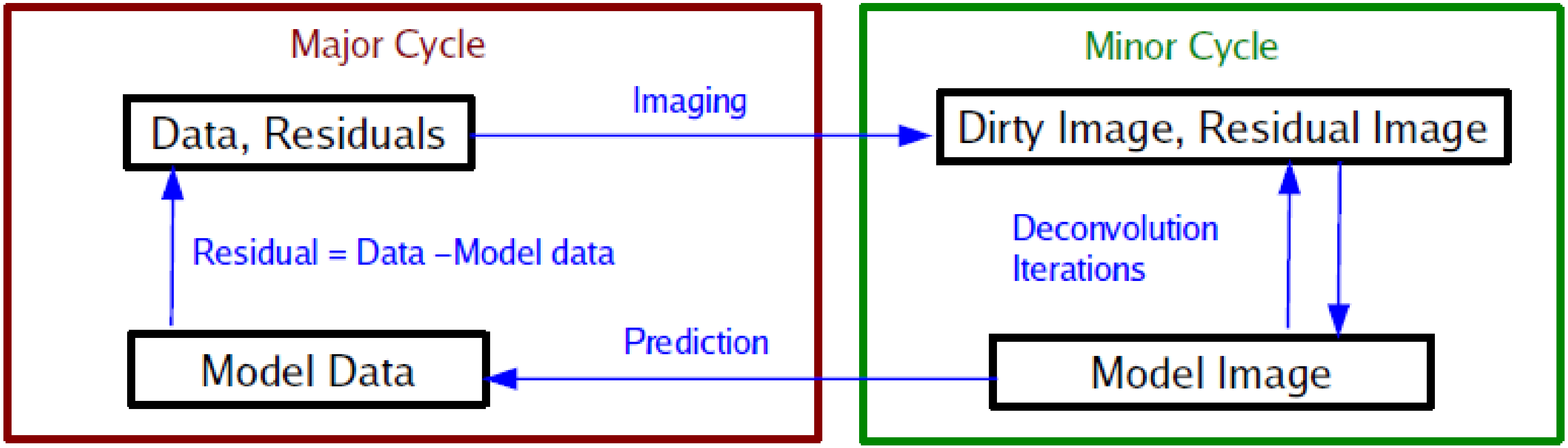
- We can measure the coherence function for any spacing and time.
- In fact:
  - We have a number of antennas at fixed locations on the Earth (or in orbit around it)
  - The Earth rotates
  - We make many (usually) short integrations over extended periods, sometimes in separate observations
  - So effectively we sample at discrete u, v (and w) positions.
  - Implicitly assume that the source does not vary
    - Often a problem when combining observations take over a long time period
  - Also assume that each integration (time average to get the coherence function) is of infinitesimal duration.

Simplification 8 (continued)

- In 2D, this measurement 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**
- The process of getting from $I^D_\nu(l,m)$ to $I_\nu(l,m)$ is **deconvolution** (examples in previous lecture).
- However, perhaps better to pose the problem in a different way: what model brightness distribution $I_\nu(l,m)$ gives the best fit to the measured visibilities and how well is this model constrained?

Modern imaging algorithms usually work more like this



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)
....

Reminder: resolution and field size

d = baseline D = antenna diameter

- 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); offset pointing <0.6 arcsec rms

High frequency (ALMA)

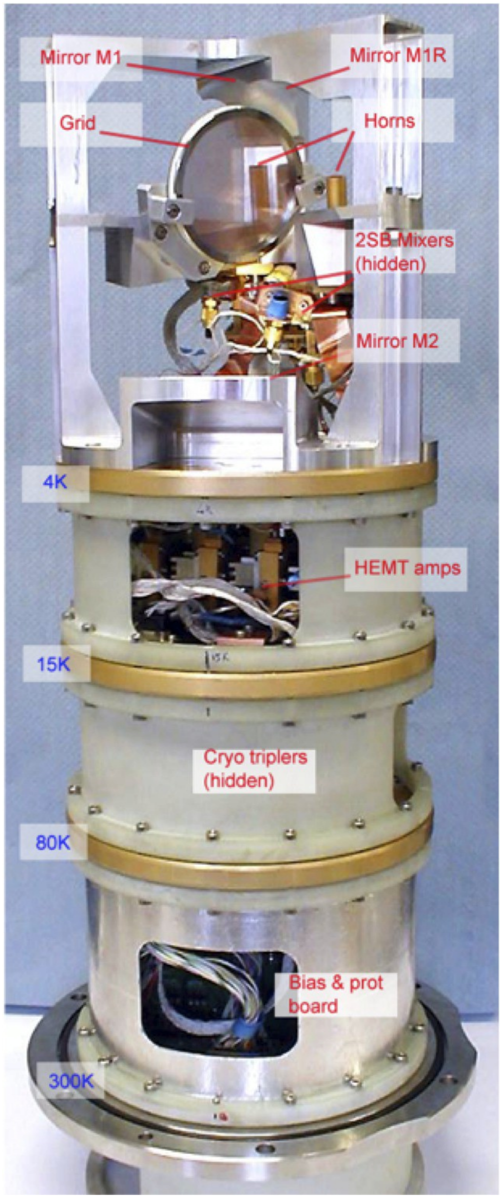
Low frequency (LOFAR)

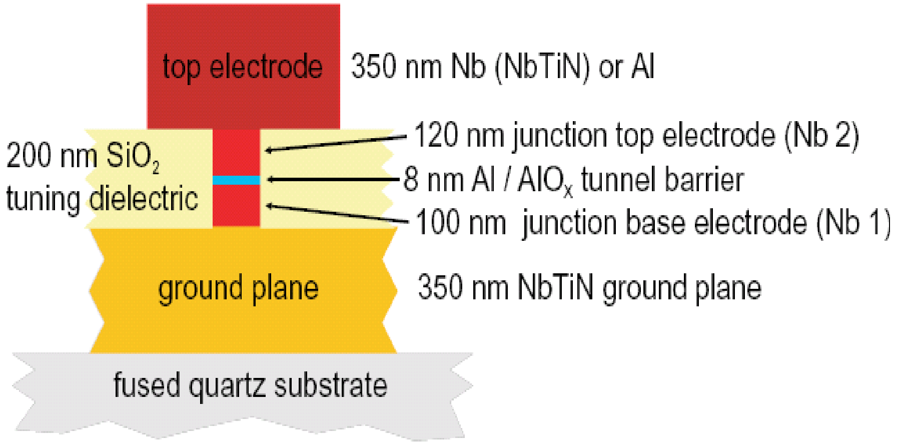
Receivers

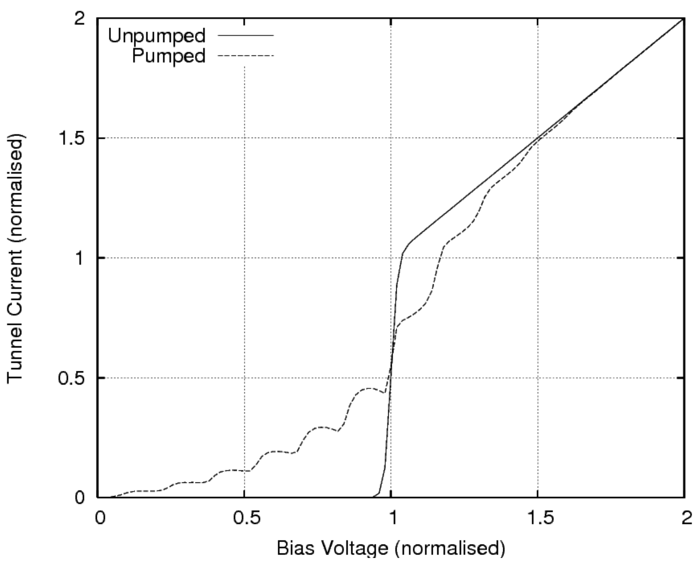
- Purpose is to detect radiation
- Cryogenically cooled for low noise (except at very low frequencies)
- Normally detect two polarization states
  - Separate optically, e.g. using crossed dipoles, wire grid, orthomode transducer, ...
- Optionally, in various combinations:
  - Amplify RF signal (HEMT)
  - Then either:
    - digitize directly (possible up to ~10's 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

Example: SIS mixers for mm wavelengths

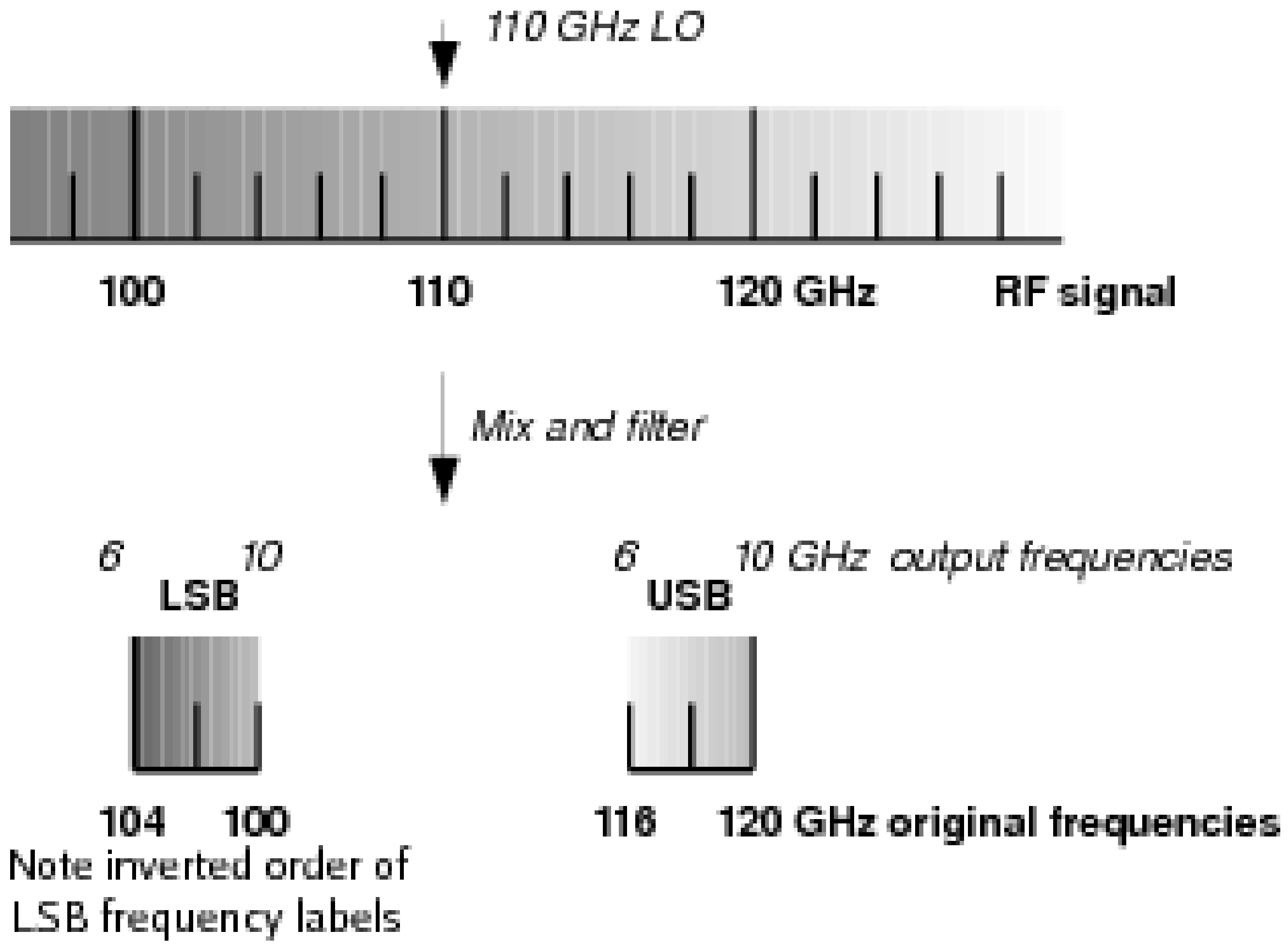
SIS = superconductor – insulator - superconductor







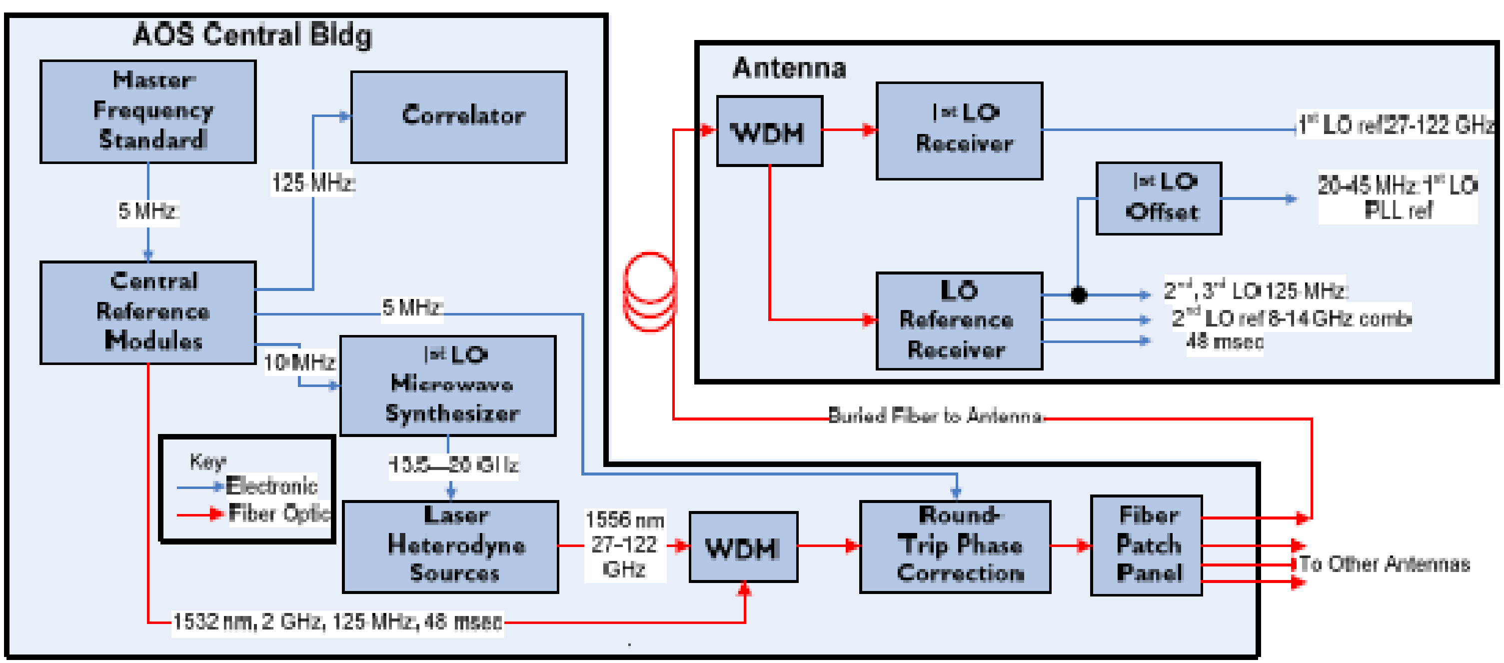
Sidebands



Either separate and keep both sidebands or filter out one.

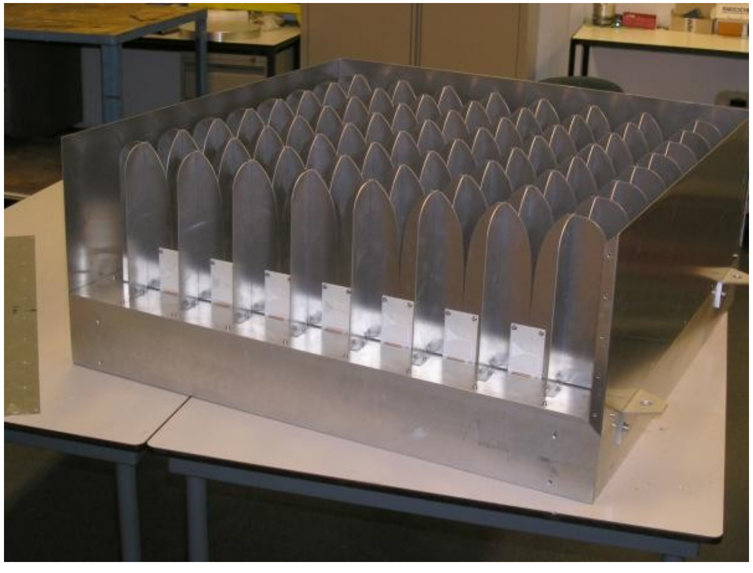
Local Oscillator Distribution

- 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, e.g. APERTIF).



Noise

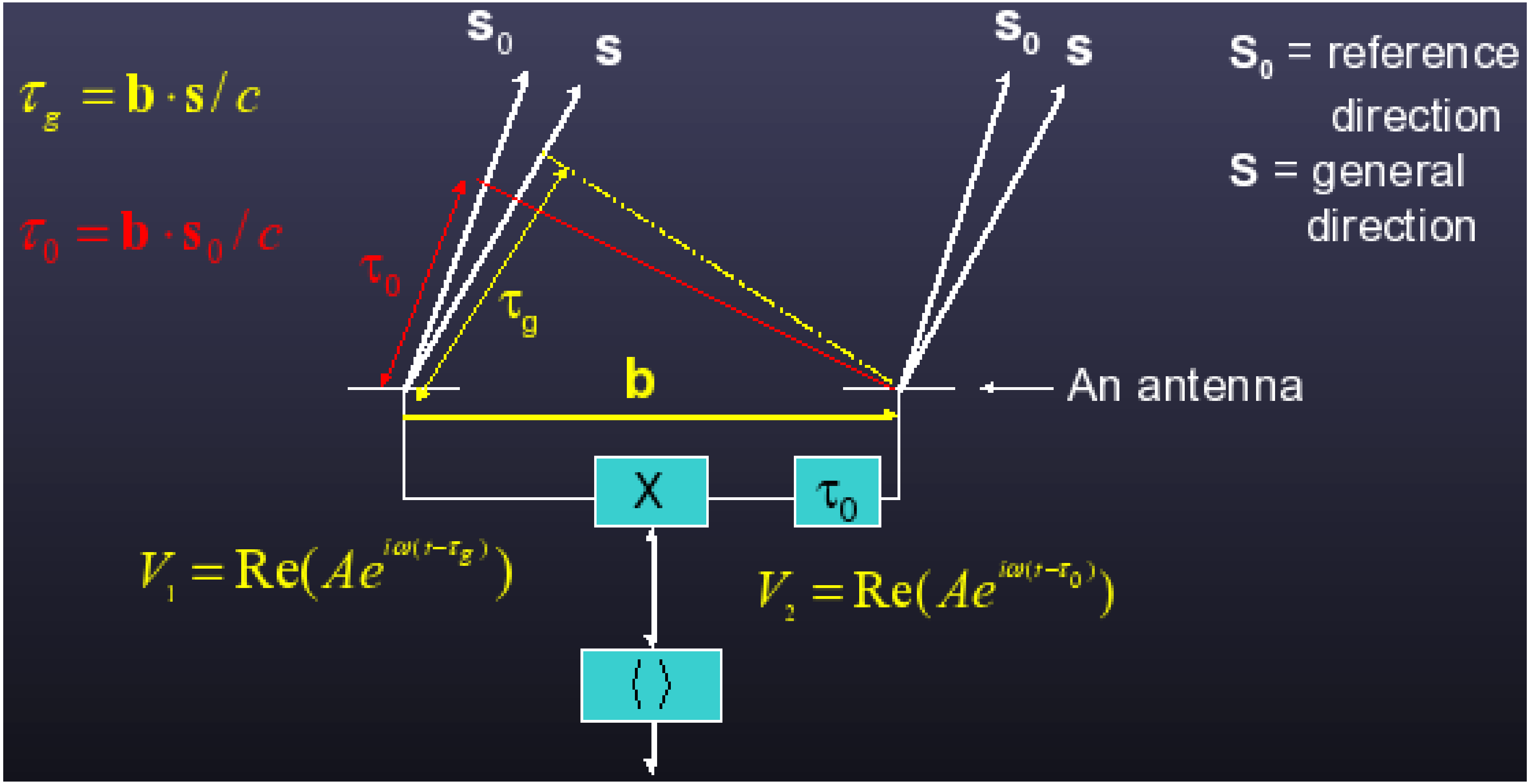
- RMS noise level S_{rms}
  - T_{sys} is the system temperature, A_{eff} is the effective area of the antennas, N_A is the number of antennas, $\Delta\nu$ is the bandwidth, t_{int} is the integration time and k is Boltzmann's constant
- For good sensitivity, you need low T_{sys} (receivers), large A_{eff} (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.
- This directly affects our ability to calculate the coherence function
- Examples:
  - 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
  - $\Delta\varphi = 2\pi\tau\Delta\nu$
  - Characteristic signature

Delay Tracking



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

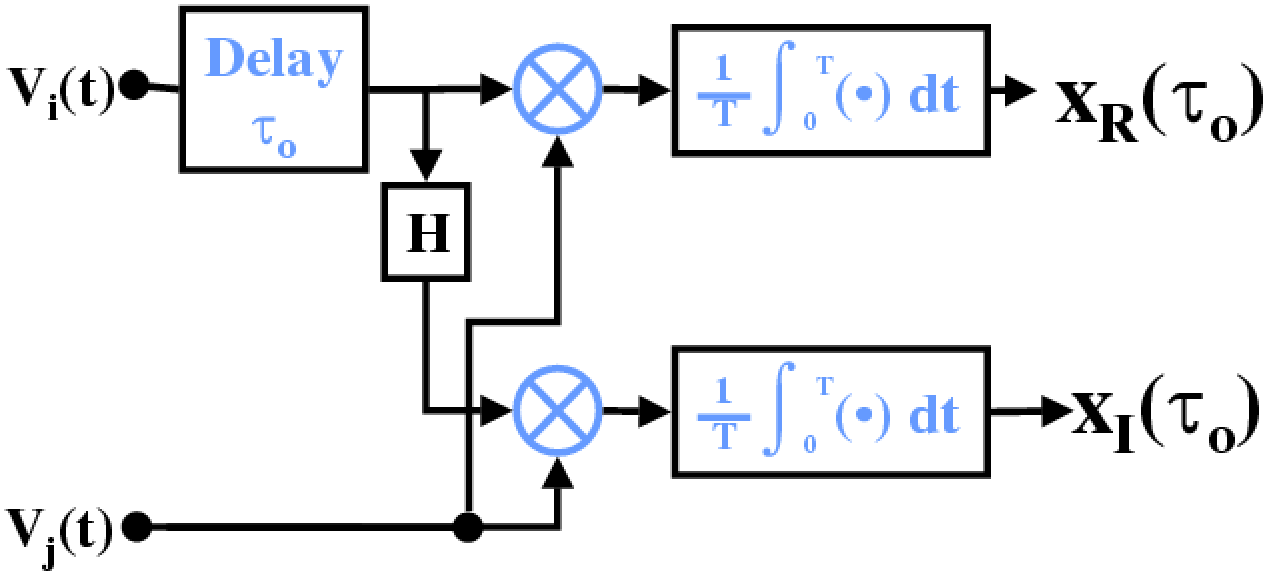
Works exactly only for the delay tracking centre. Maximum averaging time is a function of angle from this direction.

What does a correlator do?

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

Antenna i

Antenna j



Real

Imaginary

- $x_R = x_{ij}(\tau_0)$
- $x_I = x_{ij}(\tau_0 + \Delta\tau)$, **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:

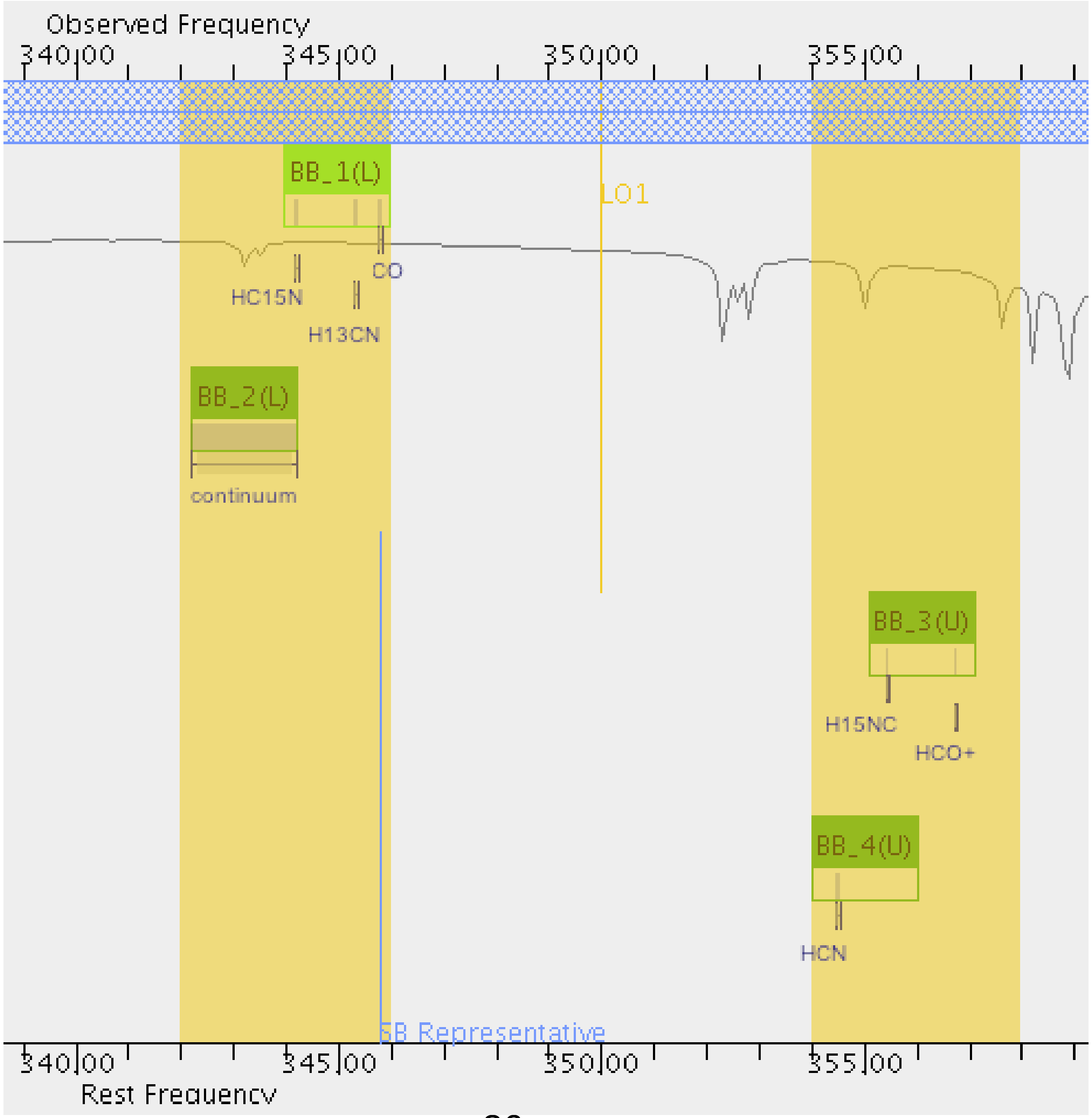
$$V(u,v,j\Delta\nu) = \Sigma_k V(u,v, k\Delta\tau) \exp(-2\pi ijk\Delta\nu\Delta\tau)$$

- 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 $V_\nu(u,v,w)$, 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

A priori calibrations

- Array calibrations (these have to be at least approximately right; small errors can be dealt with). For instance:
  - Antenna pointing and focus model (absolute or offset)
  - Correlator model (antenna locations, Earth orientation and rate, clock, ...)
  - Instrumental delays
  - Model of the geomagnetic field
- 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, ...

Auxiliary Calibrations

- Pre-observation
  - Pointing offsets
  - Focus
  - (approximate) delay
- During observation
  - atmosphere + load (mm/sub-mm)
  - continuous, switched noise source (cm/m)
  - Can also be used to monitor phase difference between polarizations
  - Water-vapour radiometer (atmospheric path, opacity contribution)
  - Multifrequency satellite (e.g. GPS) signals (electron content)
- Calibration source models and measurements

Off-line calibrations are covered in the lecture and tutorial.

The End