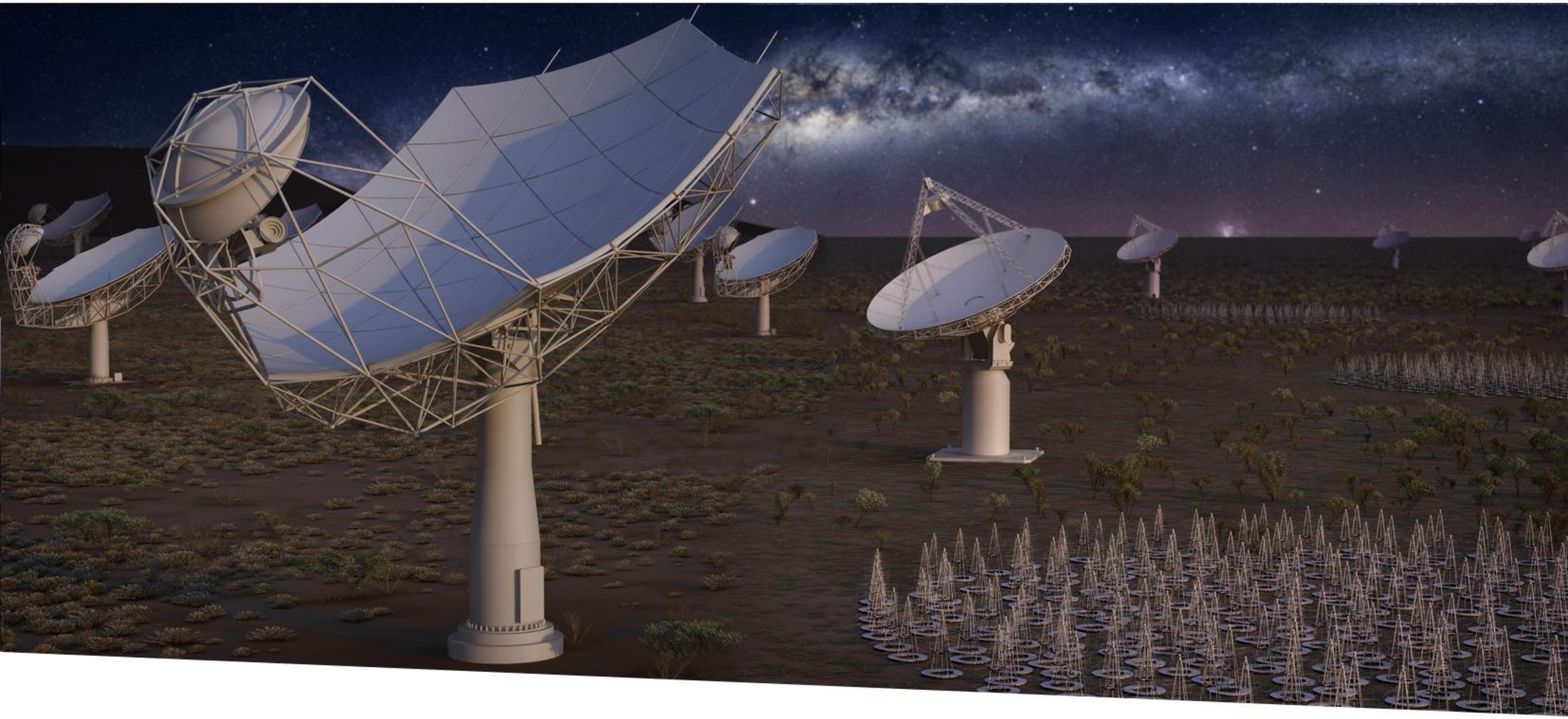
Introduction to interferometry



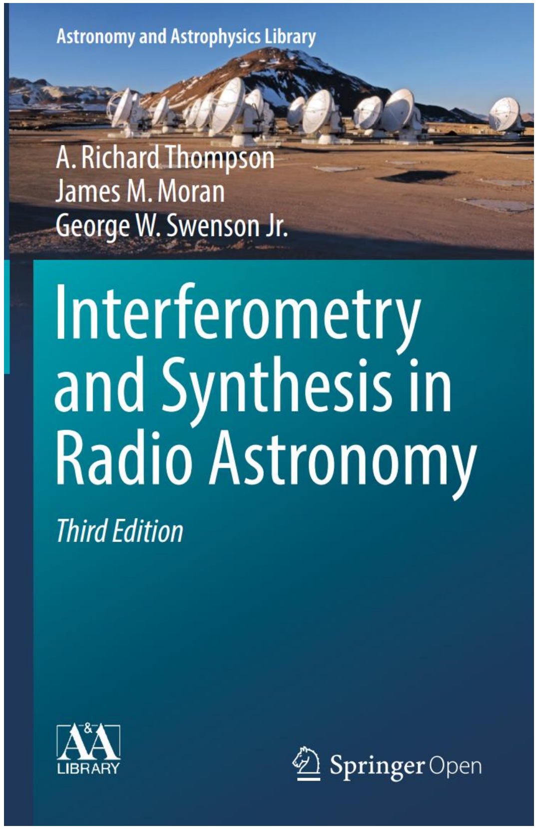
SQUARE KILOMETRE ARRAY

Exploring the Universe with the world's largest radio telescope

Daniel Hayden

27 May 2019

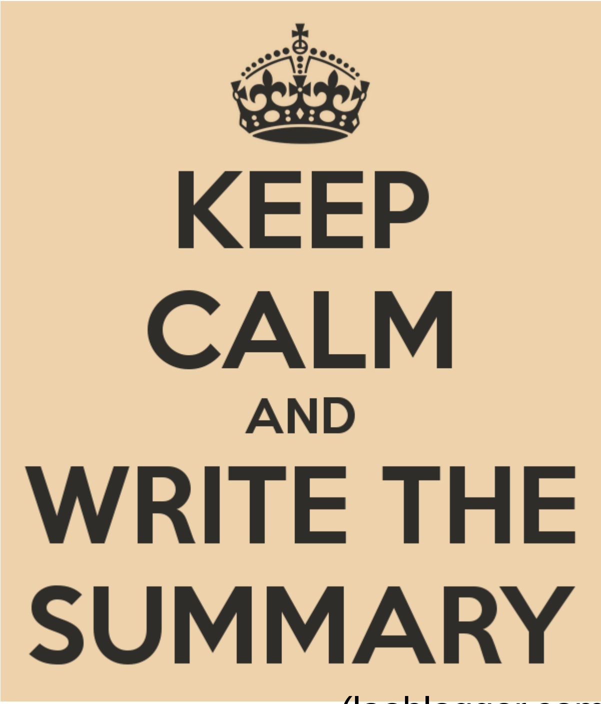
Acknowledgements



- This presentation is partly based on chapter 1 of Interferometry and Synthesis in Radio Astronomy, Third Edition, by Thompson et. al
- Some slides are also adapted from:
  - Perley (Green Bank Interferometry school 2015)
  - Shafi (AVN 2017)

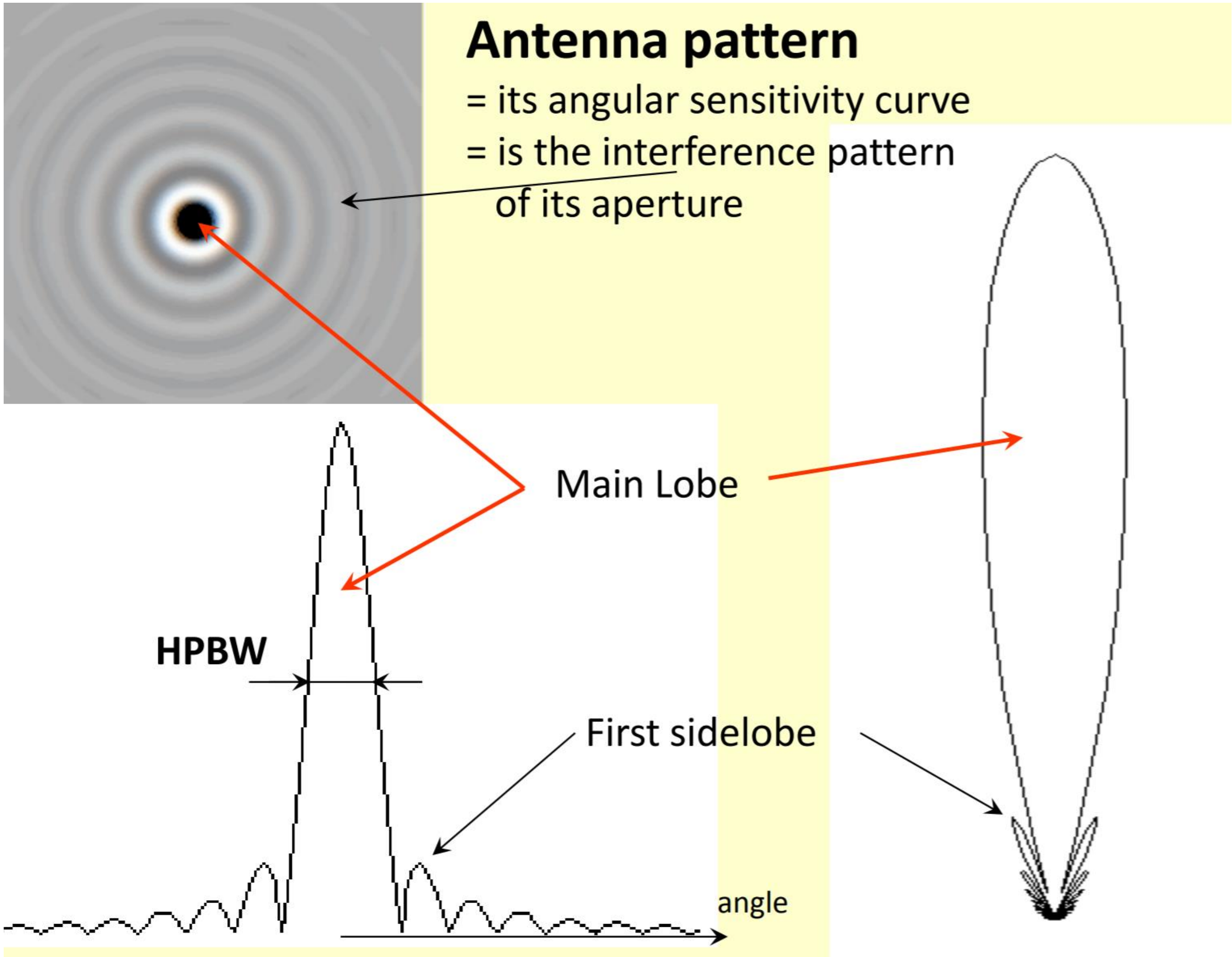
Scope

- This lecture will mostly focus on the <u>two element interferometer</u>.
- It will explain how it works, why it is needed, and look at a few historical developments tracing its use.
- I'll occasionally stop for **questions** along the way



(laoblogger.com)

Angular resolution - recap



(Slide taken from presentation by J.Koppen)

Why we need good angular resolution in radio astronomy

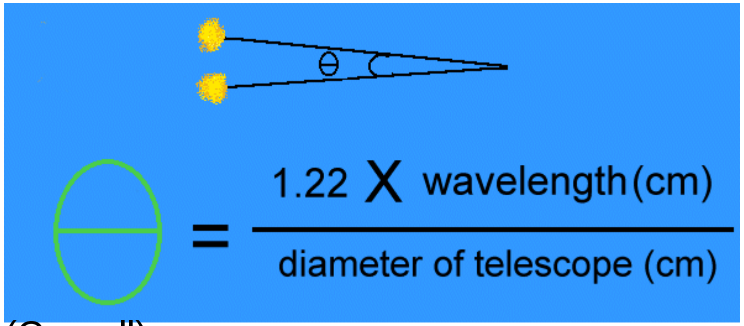
- Let's take optical telescopes: resolution achievable from ground using conventional techniques ~0.5 arcseconds.
- Remember 1 degree = 3600 arcseconds. 1 arcsecond = angular size of a small coin at ~ 2 miles.
- In radio astronomy it is very important to:
  - measure the positions of radio sources with enough accuracy to identify them with objects detected in the optical and other parts of the EM spectrum.
  - measure intensity, polarisation and frequency spectrum with similar angular resolution in both radio and optical.

So can we get the same kind of resolution with single radio antennas as with single optical antennas? No. But Why?

Southern African Large Telescope
Ang. Res < 0.6 arcsec

Single radio dishes become too large!

- Practical limitations on dish size limit the resolution to a few tens of arcseconds.
- E.g. 100m antenna at 20cm wavelength – resolution ~7 arcmin.
- But Why? Angular resolution depends on wavelength and dish diameter. For larger wavelength the dish diameter needs to be larger to get the same resolution.



(Cornell)

- If we want 1 arcsecond resolution, we would need an aperture ~400 times larger, or 40km in size!


America's most powerful radio telescope IS . . .

ZAPPED!

. . . by hostile space aliens!

BEFORE

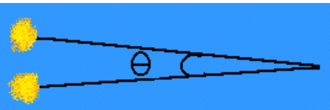
AFTER

(NRAO)

Instead of 1 dish.. many small dishes

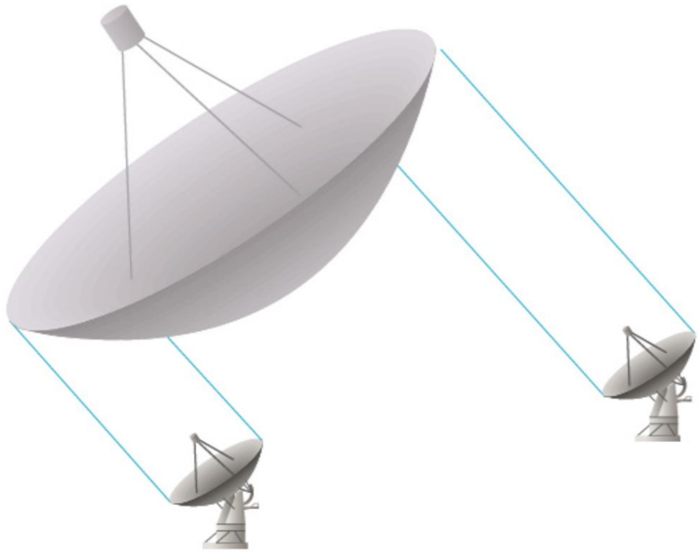
But which of the two configurations below gives higher sensitivity?

- It turns out that one can get the resolution of a 40km diameter dish by instead using 2 dishes 40 km apart! This is an interferometer.
- This distance between the dishes is called the baseline length.
- The angular resolution of an interferometer is:



$$\Theta = \frac{1.22 \times \text{wavelength (cm)}}{\text{Baseline length (cm)}}$$

$$\theta \sim \lambda / B$$

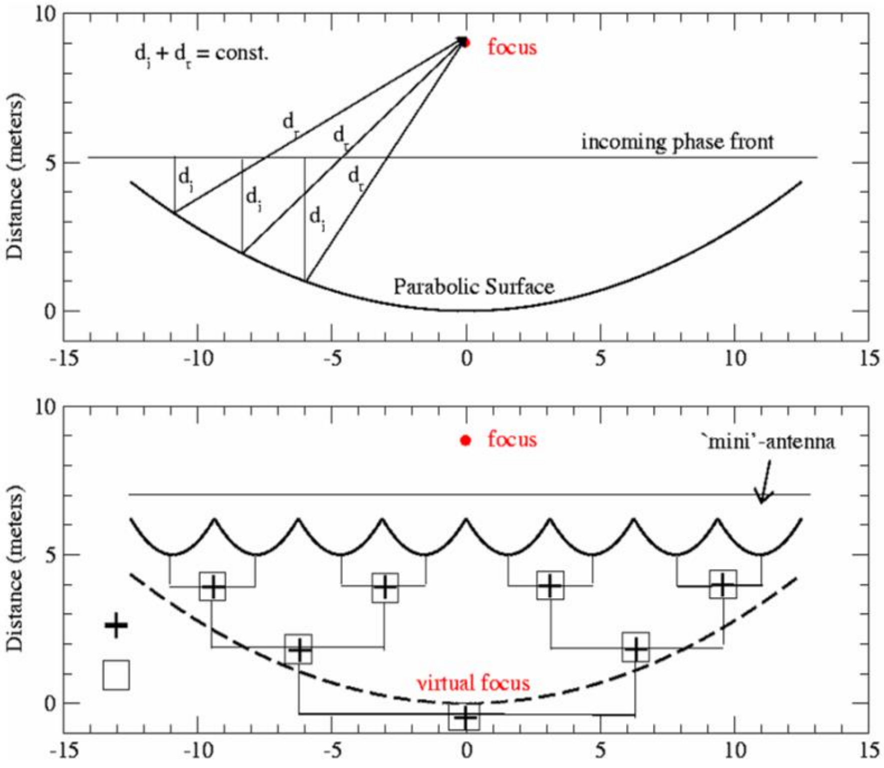


Copyright © Addison Wesley

Most radio interferometers have angular resolutions between 0.1 and 10 arcseconds.

Why? does this work?

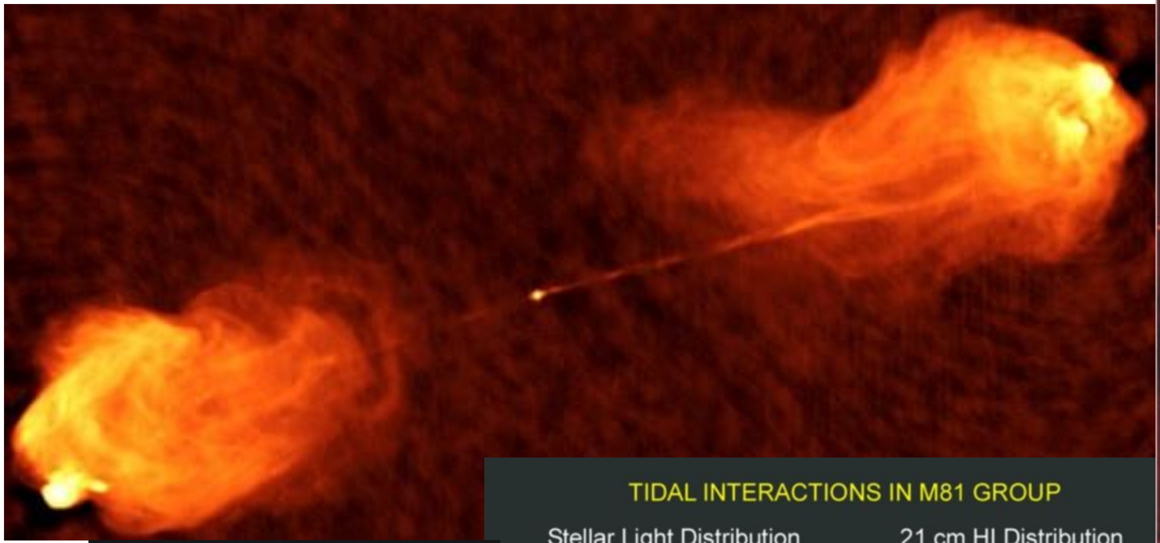
- A single parabolic dish is actually an interferometer.
- An incoming plane wave is coherently summed at the focus.
- **But instead** one could replace the single dish with a number of smaller dishes, whose individual signals are summed at a ‘virtual focus’.
- The ‘virtual focus’ sum is equivalent to the single dish focus.

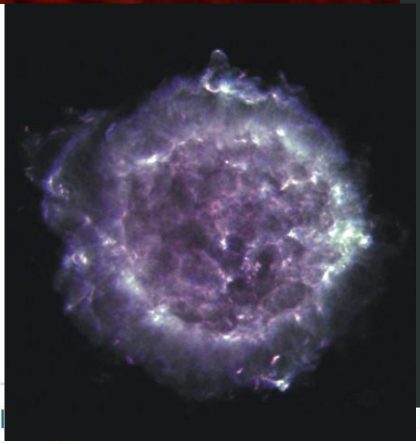


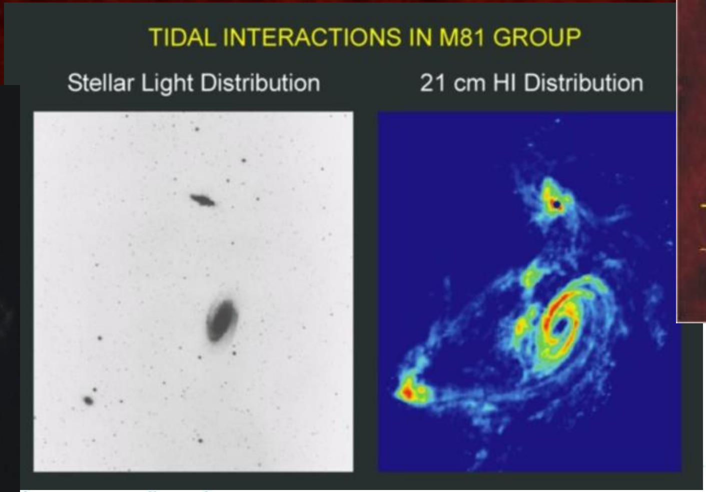
(Perley, 2015)

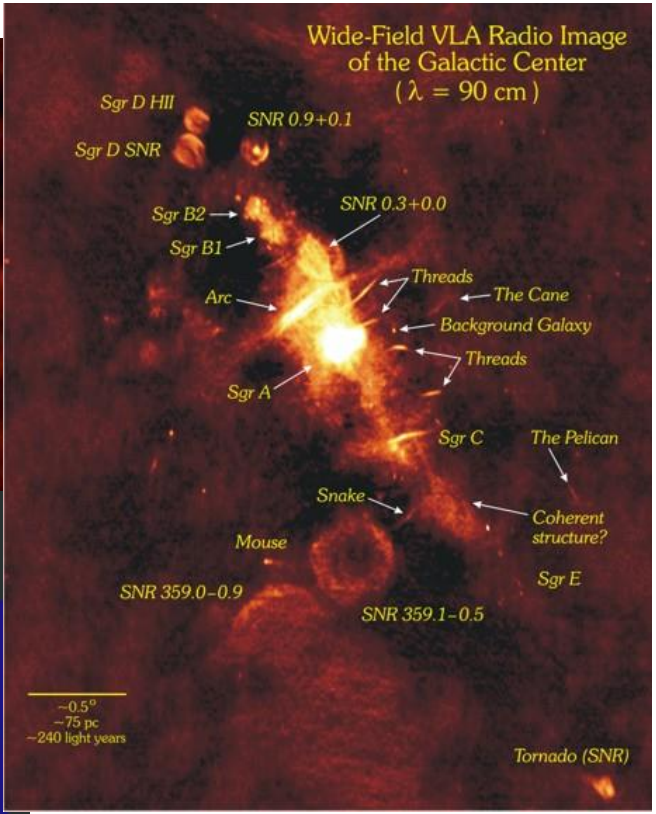
Some examples of interferometric imaging

- Some examples of radio emission from celestial sources







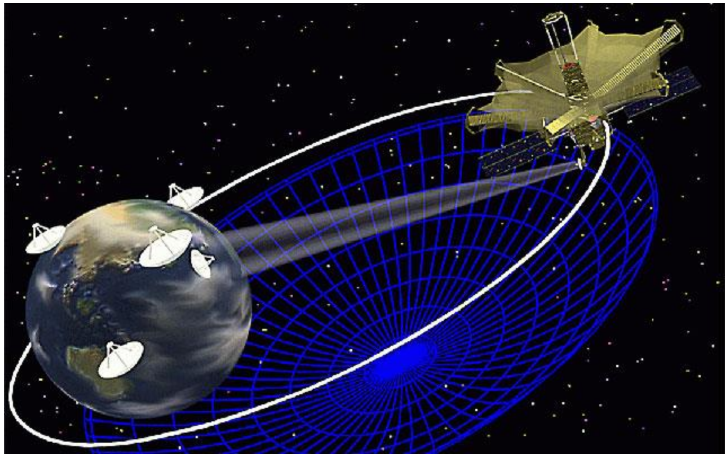


(NRAO)

So can we get angular resolution MUCH higher than 1 arcsec?

- Yes. By simply increasing the distance between antennas. This can be done to 1000s of kilometres to give angular resolutions of milli-arcseconds – more on this in a later talk.

- **Some applications** of higher angular resolution are:
  - Detailed imaging of radio emission from celestial sources.
  - Establishing distance scale of the universe, and testing general relativity (astrometry).
  - Measuring shifts of earth's axis relative to its surface – allows study of earth dynamics (astrometry)
  - Relative motions of tectonic plates by determining the precise distance between antennas.



(NASA)

Space Very Long Baseline Interferometry

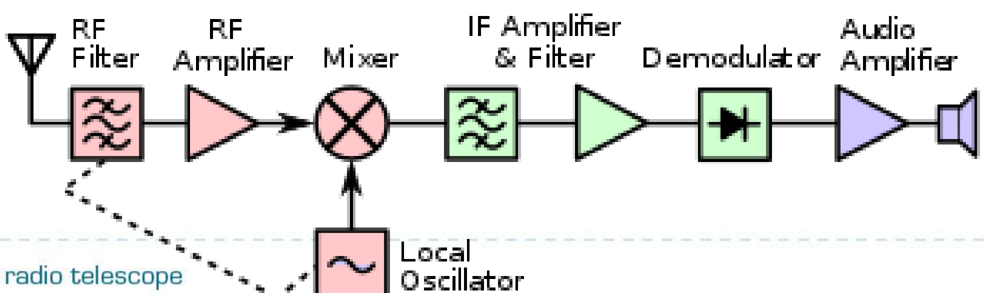
Other contributing factors to attaining high resolution in radio astronomy

- RF signals can be processed easily with high precision. Why? Because the heterodyne principle lets you convert RF signals to a IF baseband by mixing them with a signal from a local oscillator (LO).

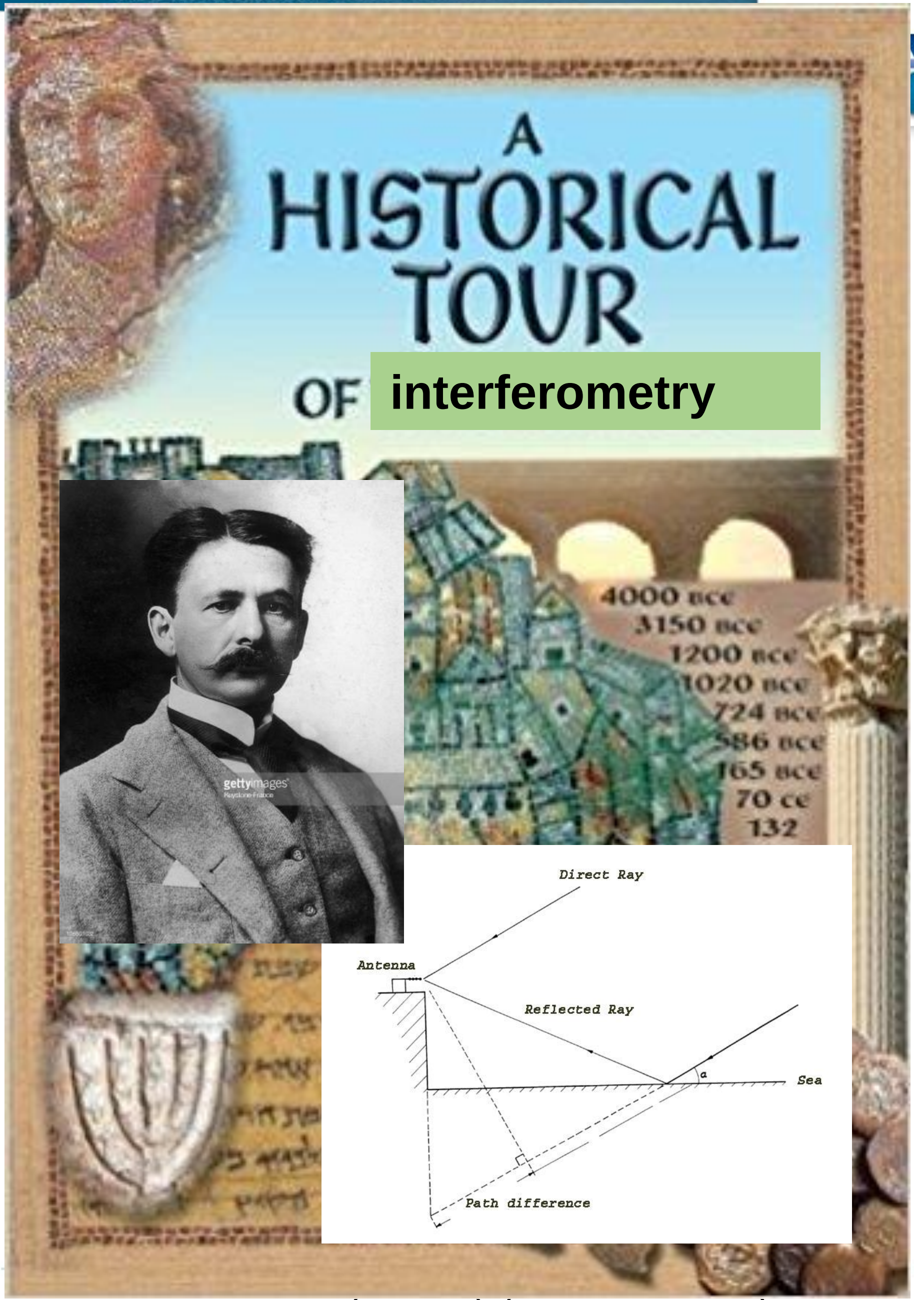
  Why? convert from RF to IF?

  – To convert several different frequencies to a single common frequency for processing. Amplifiers and filters can be tuned to a fixed frequency.
  – Lower frequency transistors have higher gains so fewer stages are needed.
  – It's easier to make sharply selective filters at lower frequencies.

- Another advantage of the radio domain is that phase variations induced by the Earth's atmosphere are less severe than at shorter wavelengths.



Let’s now take a historical tour to understand the two-element interferometer!

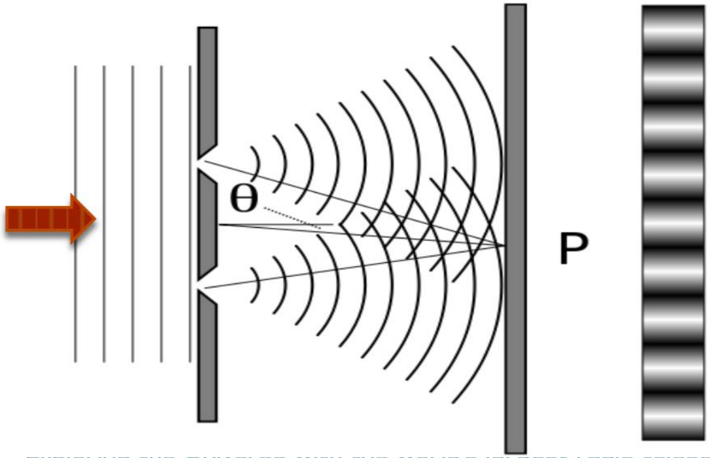


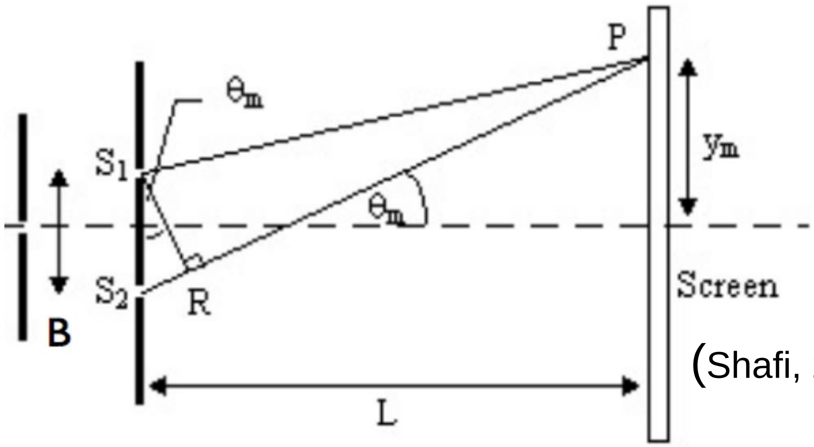
(CSIRO) (B. Ratzer 2010)

1801 – Young's double slit experiment

(Veritaseum)

- Demonstrates the wavelike nature of light
- Waves arriving in phase interfere constructively and waves arriving out of phase interfere destructively.
- B is simply the baseline length described earlier.
- If $B \ll L$ the path difference is $B \sin\theta$.
- **Constructive interference occurs when** $B \sin\theta = m\lambda$ (the path difference is an integer multiple of the wavelength).





(Shafi, 2017)

Young’s double split experiment

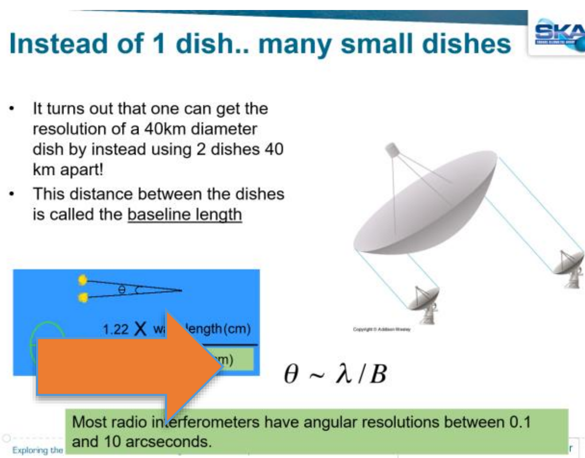
- Similarly, destructive interference occurs when $B \sin\theta = (m + 1/2)\lambda$
- Let’s call the points of constructive interference ‘fringes’.
- Re-arranging the expression for constructive interference gives:

$$\sin\theta = m\lambda/B$$

- For $y \ll L$ one can use the small angle approximation $\theta \sim \sin\theta$
- The angular separation between ‘fringes’ is therefore:

$$\theta \sim \lambda/B$$

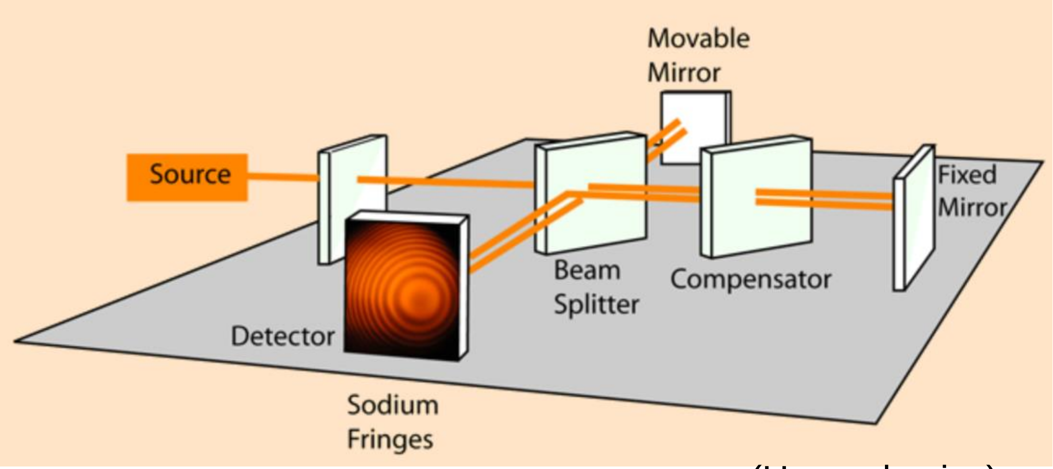
- **Remember this equation?** It’s the same equation for the angular resolution of an interferometer that we saw earlier



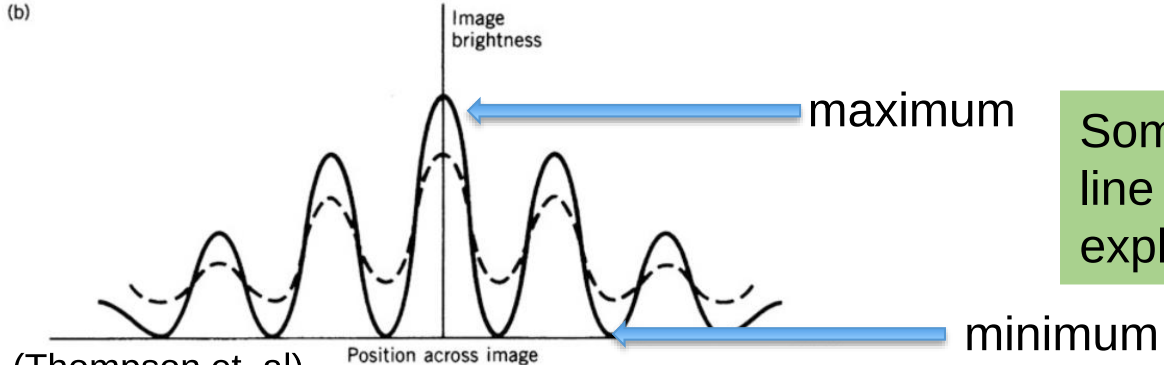
Moving on – Michelson interferometer 1890-1920

Early optical interferometers were used to measure the diameters of some near, large stars such as Arcturus and Betelgeuse.

Beams of light from a star fall upon two apertures and are combined in a telescope. This creates an interference pattern, or fringes, like in the double slit experiment.



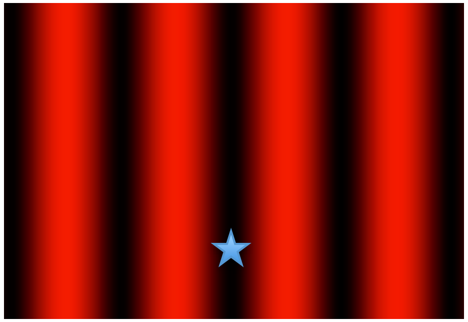
(Hyperphysics)



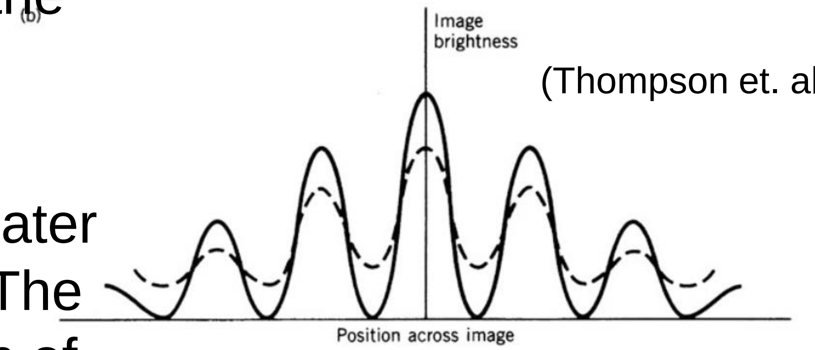
(Thompson et. al)

Sometimes you see the dotted line not the solid line. What explains this?

Michelson interferometer 1890-1920

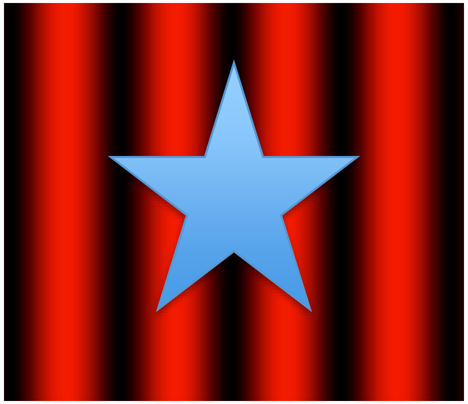


- **Scenario 1**: angular width of star is small compared to fringe spacing. A single fringe pattern shows on the screen. This gives the solid line in the figure.



(Thompson et. al)

- **Scenario 2:** width of the star is greater than the spacing between fringes. The resultant image is the superposition of fringe patterns from a series of points across the star. Maxima and minima do not coincide so the fringe amplitude is attenuated. This gives the dotted line in the figure.



1890-1920 Michelson interferometer

- Michelson defined the fringe visibility

$$\mathcal{V}_M = \frac{\text{intensity of maxima - intensity of minima}}{\text{intensity of maxima + intensity of minima}}$$

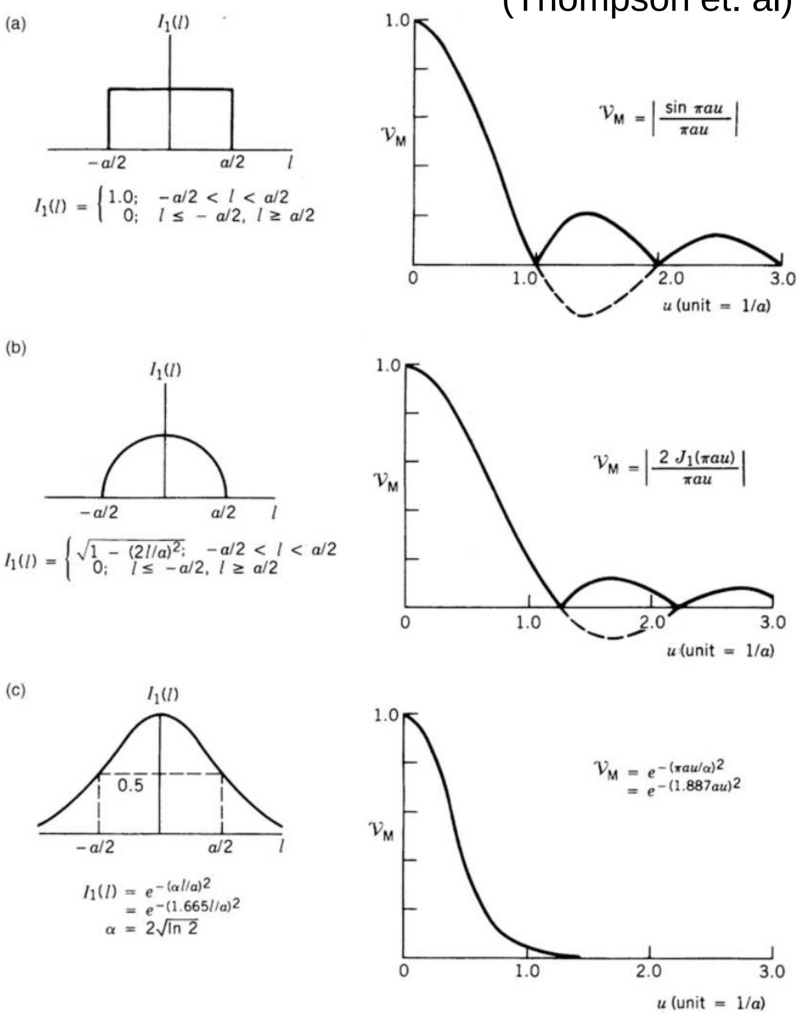
- Note that this quantity is normalised to unity when the intensity at the minima is zero i.e. when the width of the star is small compared to the fringe width.

- If the fringe visibility is measurably less than 1, the star is said to be **resolved** by the interferometer.

1890-1910 Michelson interferometer

- The right column shows the fringe visibility as a function of the spacing of the receiving apertures (baseline) in units of the wavelength.
- The left column shows the one-dimensional intensity profiles $I(l)$ for 3 intensity models of a star (uniform rectangular source, uniform circular source, circular Gaussian distribution).
- The two quantities are related, and the fringe visibility can be measured. Can the intensity profile somehow be inferred from this?
- Yes! See green box below.

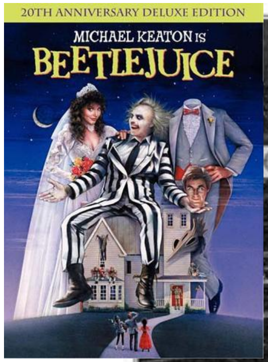
(Thompson et. al)

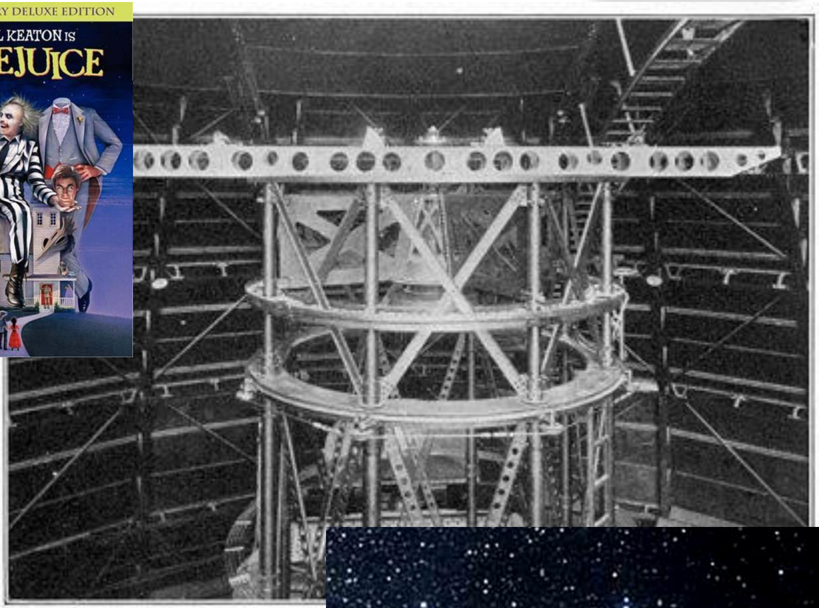


The fringe visibility is proportional to the modulus of the Fourier transform of $I(l)$ with respect to the spacing of the apertures measured in wavelengths. More on this…later.

Michelson interferometer 1890-1910

- Michelson and Pease used the circular model to understand their observations.
- **How did they determine stellar diameter?**
- By adjusting the aperture spacing until they found the first minimum in the fringe visibility function.
- Betelgeuse's diameter was found to be 380 million kilometres – <u>300 times larger than the sun.</u>



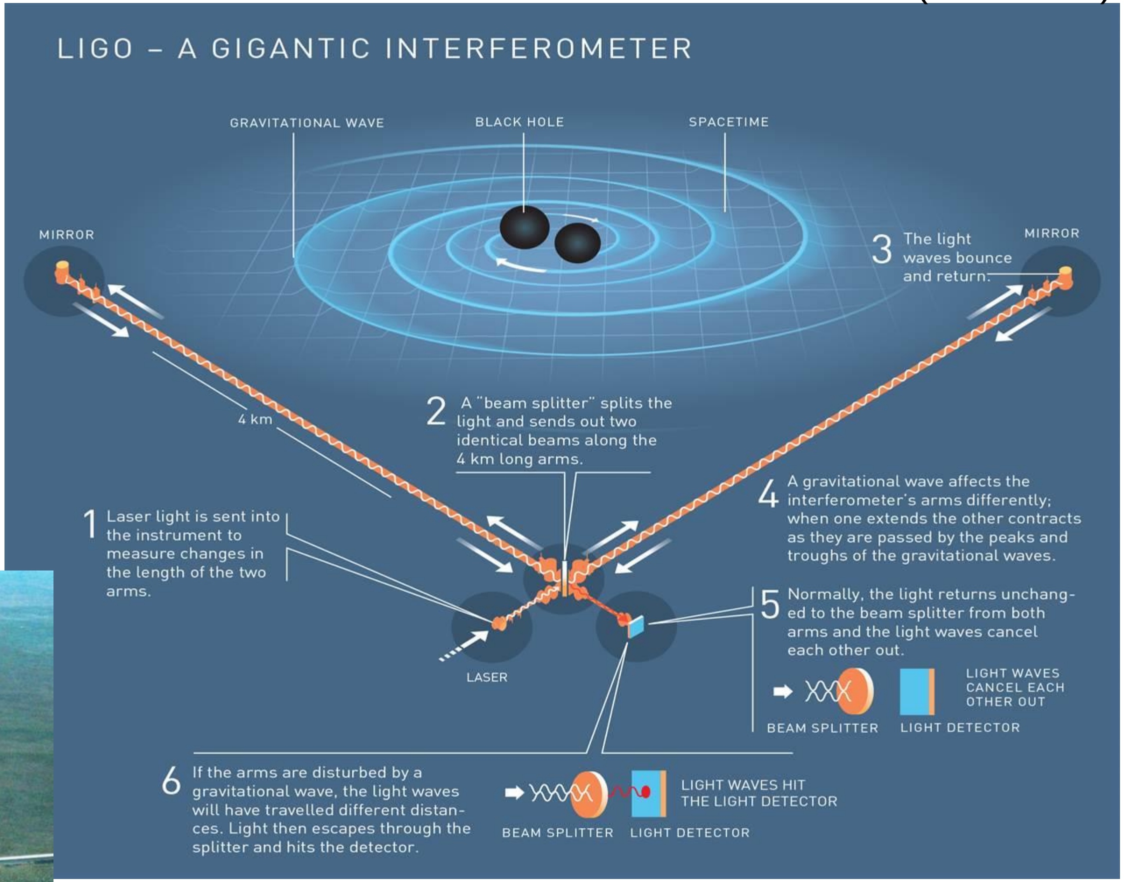


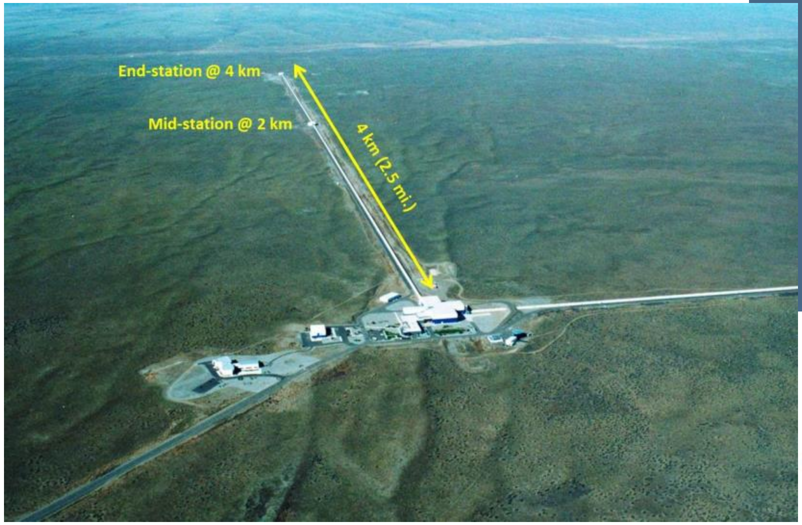
Due to atmospheric turbulence, the fringe pattern moved erratically over timescales of 10-100ms. The fringe visibility had to be estimated by eye.

LIGO

- Remember LIGO's detection of gravitational waves in 2015 from two black holes merging?
- The LIGO interferometer is fundamentally a Michelson interferometer. This diagram explains how it works.

(NBCNews)



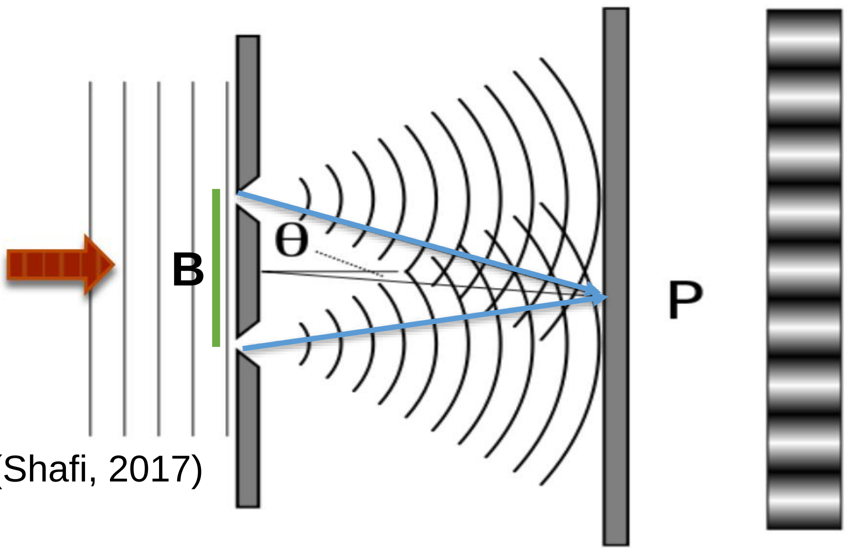


(Caltech)

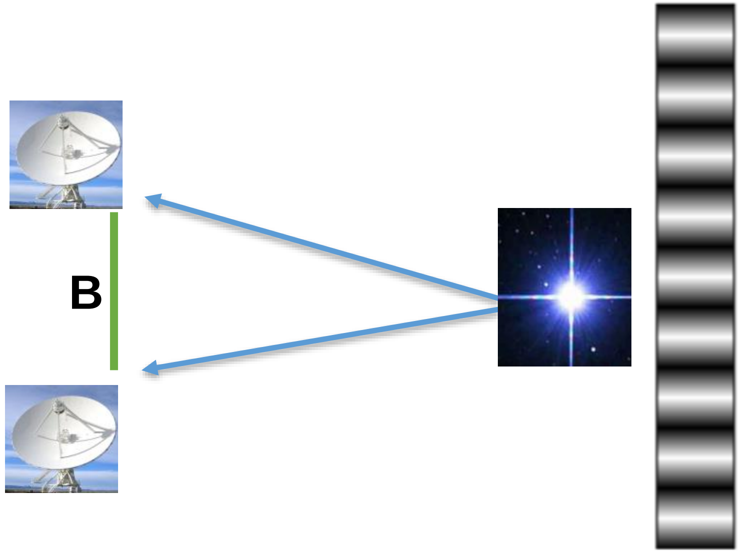
SKA1-Mid, with MeerKAT, will also contribute to this field by measuring how distances to pulsars change!

From optical to radio interferometer

- **Optical** – the fringe pattern can be seen on a screen. It is therefore a function of position.
- **Radio** – There is no screen which is a function of position. Instead, the radio wave from a source induces a voltage at two antennas which is summed together. Fringes occur as a function of time as the radio source changes position in the sky.



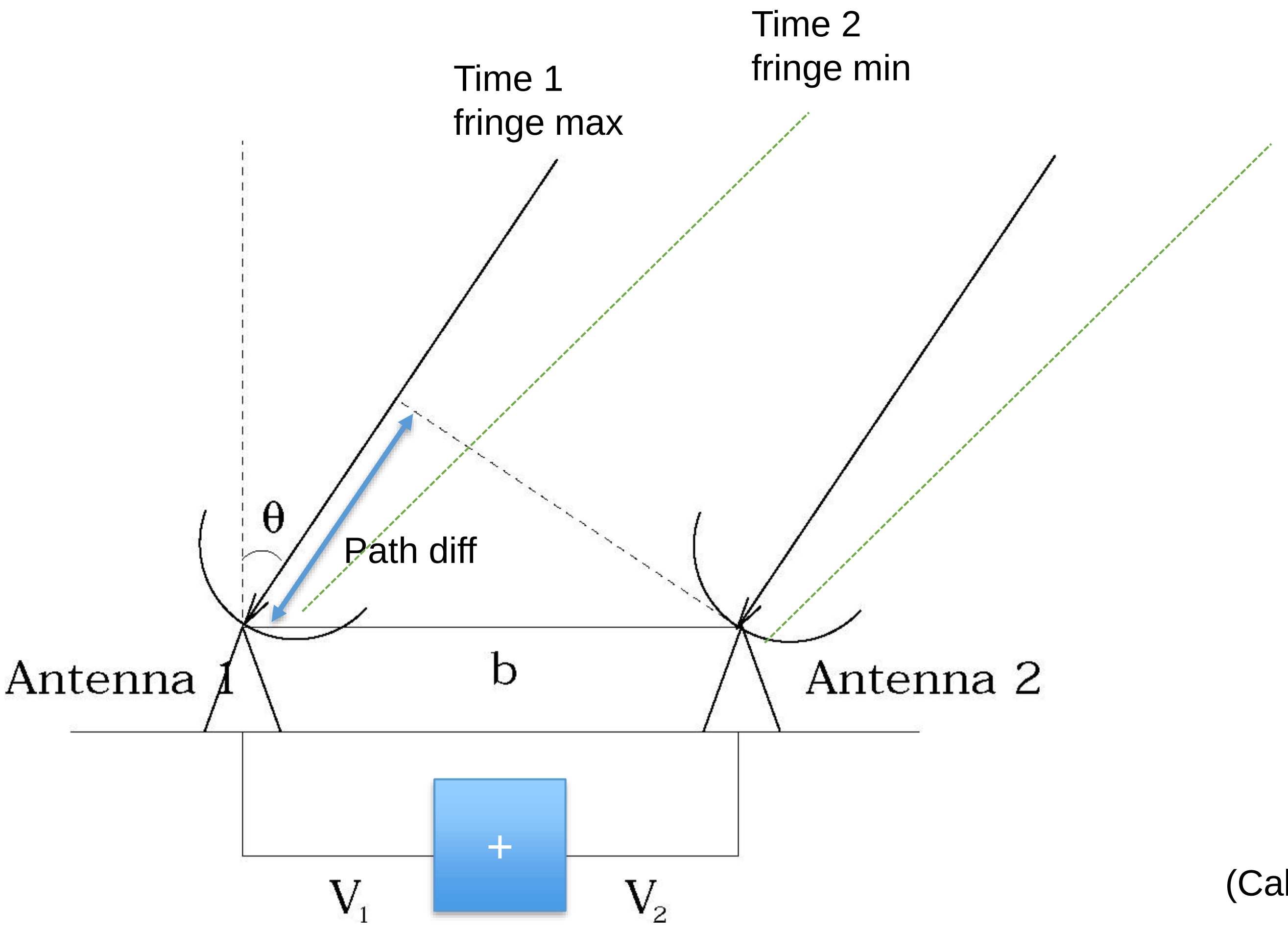
(Shafi, 2017)



Optical – fringe pattern in space

Radio – fringe pattern in time

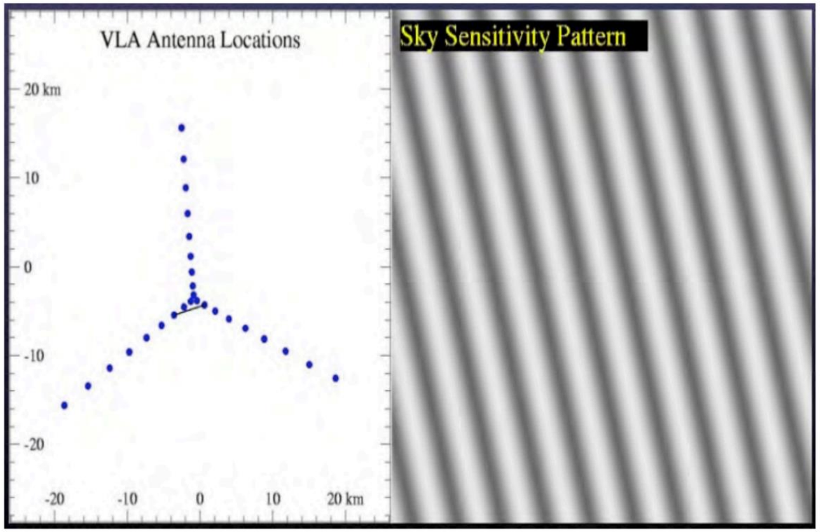
Two element radio interferometer



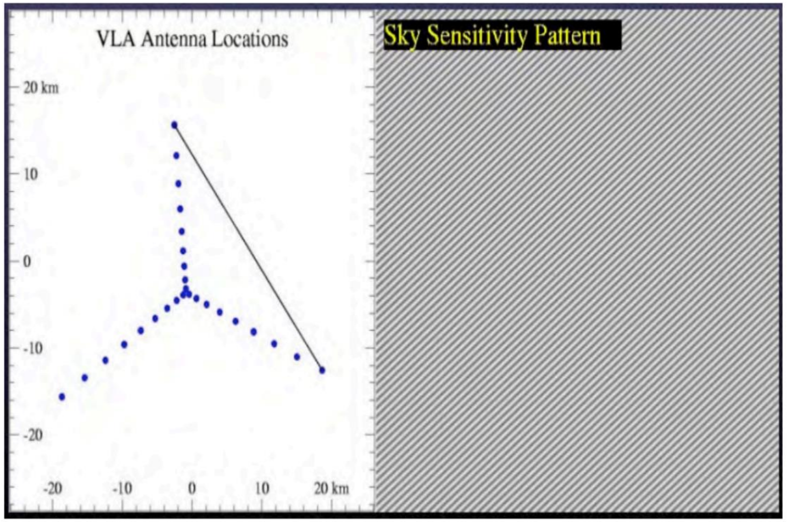
(Caltech)

Two element radio interferometer

- The fringe spacing decreases (i.e. angular resolution improves) with increasing baseline.
- Therefore long baselines are very sensitive to changes in source position. Short baselines are less sensitive.



Short VLA baseline



Long VLA baseline

(Nebiya, 2017)

1946 Ryle and Vonberg

Ryle at work

(Alchetron.com)

- Radio interferometer to investigate cosmic radio emission.
- Used dipole arrays at 175 MHz, with a baseline that was variable between 17 and 240m.

- **Simplifying assumptions:**
  - Receiver is sensitive to a narrow band of frequencies so in a simplified analysis we can consider monochromatic signals at centre frequency v_0.
  - The source is very far so we can assume incoming waves are parallel.

- Signal voltage from antenna 2 (see two slides previous) is $V\sin(2\pi v_0 t)$

- The longer path length to antenna 1 introduces a time delay $\tau = (B/c)\sin\theta$ where B is the baseline length, θ is the angular position of the source and c is the speed of light.

- The signal voltage from antenna 1 is therefore $V\sin[2\pi v_0(t-\tau)]$

1946 Ryle and Vonberg

- The detector generates a response proportional to the squared sum of the 2 signal voltages:

$$\{V\sin(2\pi v_0 t) + V\sin[2\pi v_0(t-\tau)]\}^2$$

- The output of the detector is averaged in time, which removes terms in t. Also, a lowpass filter removes the higher frequency harmonic terms. The detector output, in terms of the power P_0 generated by either antenna alone, is:

$$P = P_0[1+\cos(2\pi v_0\tau)]$$

- Substituting for the time delay τ we get:

$$P = P_0\left[1+\cos\left(\frac{2\pi v_0 B\sin\theta}{c}\right)\right]$$

- Thus, as the source moves across the sky, P varies between 0 and $2P_0$ i.e. you get fringes!

1946 Ryle and Vonberg

- The phase of the fringes ∅ is given by the argument of the cos function:

$$\emptyset = \frac{2\pi v_0 B \sin\theta}{c}$$

For what values of θ does this give a maximum in P? We get the same result as in the double slit experiment, that the angular spacing between fringes is $\theta \sim \lambda/B$

- The change in phase as a function of source angle is:

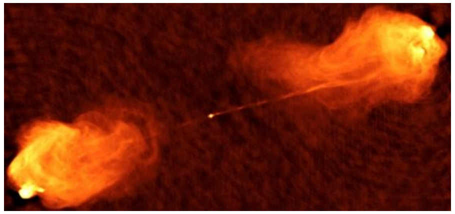
$$\frac{d\emptyset}{d\theta} = \frac{2\pi v_0 B \cos\theta}{c}$$

- The "fringe frequency" is then:

$$\frac{d\emptyset}{dt} = \frac{2\pi v_0 B \cos\theta}{c}\frac{d\theta}{dt} = \frac{2\pi\omega v_0 B \cos\theta}{c}$$

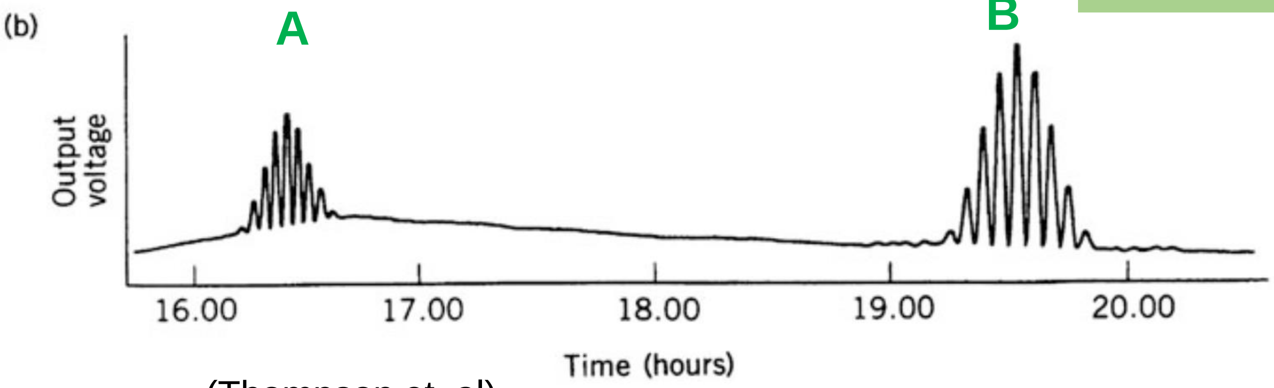
1946 Ryle and Vonberg

Their results, see below. Note the response is modulated by the beam pattern of the antennas.



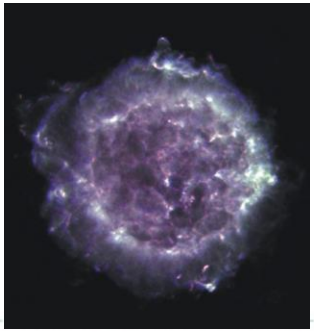
A – Cygnus A

Using an interferometer rather than single dish results in an increase in precision of determining the time of transit of the source.



(Thompson et. al)

B – Cassopeia A



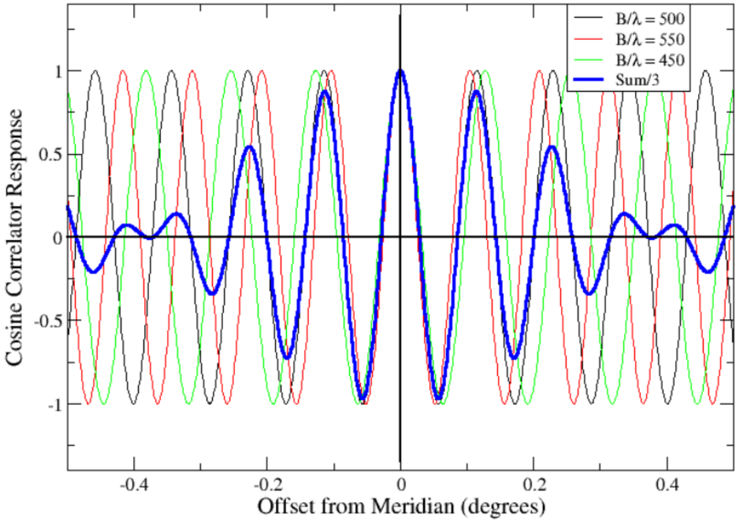
(NRAO et. al)

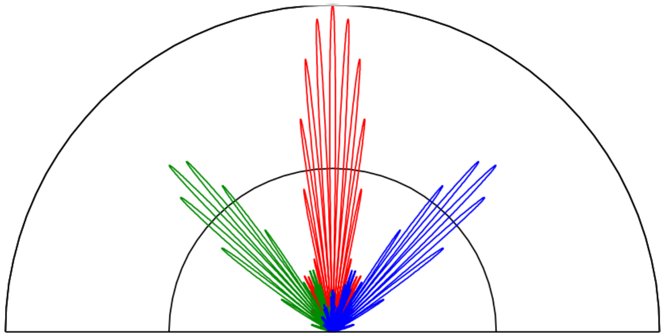
A quick note on delay tracking

- We assumed earlier that the received signals were monochromatic.

- In reality, the effect of having multiple frequency components is an <u>attenuation of the fringe amplitude</u> from sources far from the meridional plane (zenith).

- To overcome this, one can shift the entire ‘fringe packet’ to the position of interest by adding a time delay to the antenna closest to
the source.



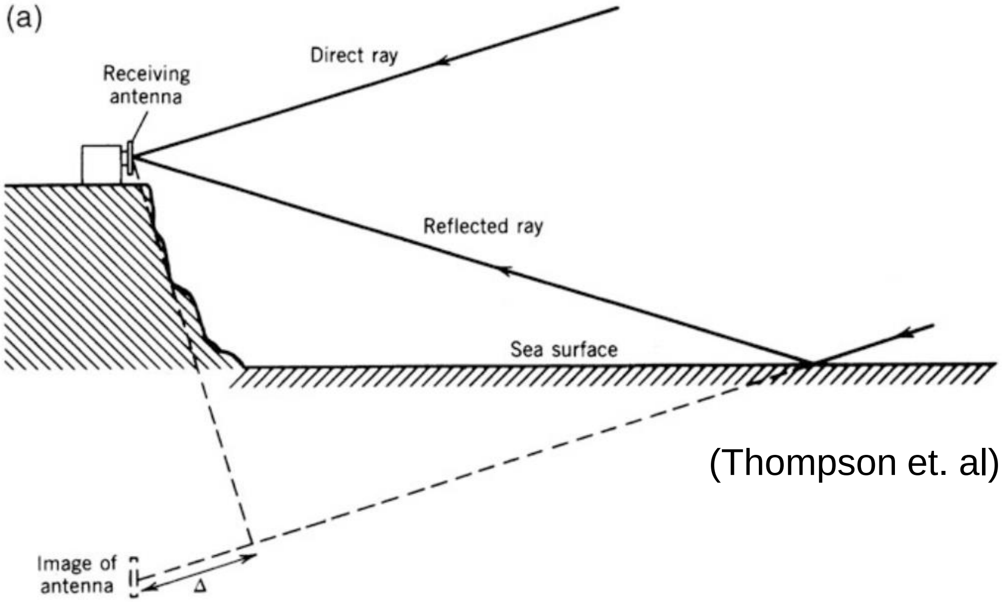


(Perley, 2015)

How can one build a ‘two element’ radio interferometer using only one antenna?

~1947 Sea interferometer

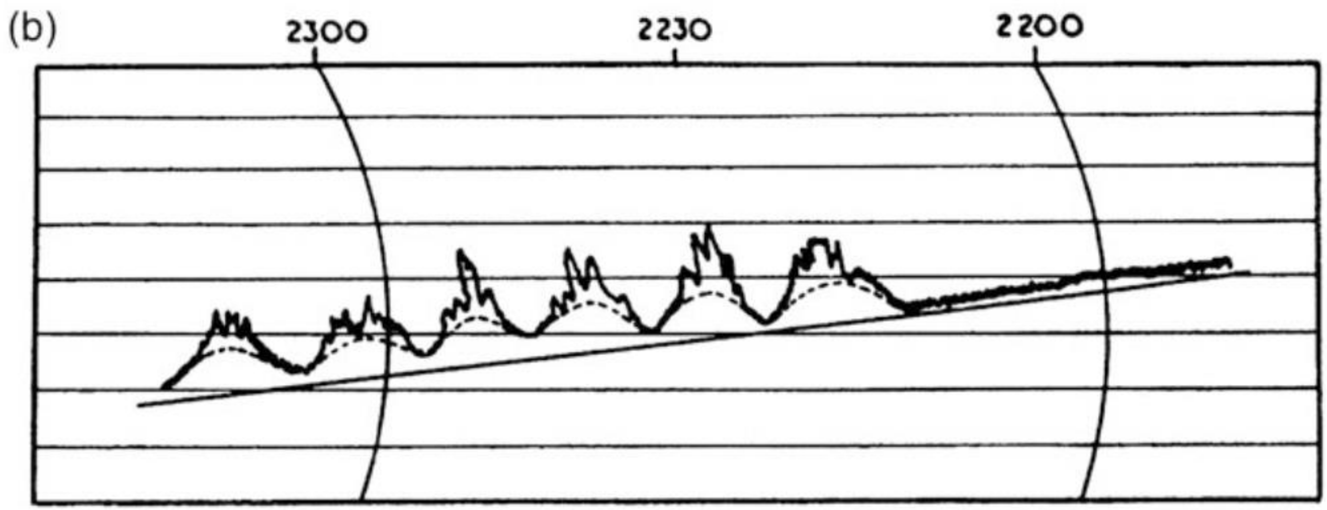
- Using a horizon pointing antenna installed for radar during WW2.
- And using the reflective ability of the sea!
- Received radiation from sources rising over the horizon.
- Optimal frequency ~200 MHz because of ionospheric effects at lower frequencies and sea roughness at higher frequencies.



(Thompson et. al)

The fringe pattern is similar to what would be obtained with the actual antenna and one at the position of its image in the sea

~1947 Sea interferometer



(Thompson et. al)

- Record of Cygnus A at 100 MHz. First evidence of existence of discrete non-solar radio source!
- Why? is there both a steady component and a fluctuating component?
- What are some disadvantages of this set-up?
  - Difficulty of varying the length of the baseline (this is set by the cliff height).
  - Long atmospheric paths.
  - Roughness of sea surface.

(whitecatcoaching.com)

- But there is a problem with the previous two setups (Ryle and Vonberg, Sea interferometer).

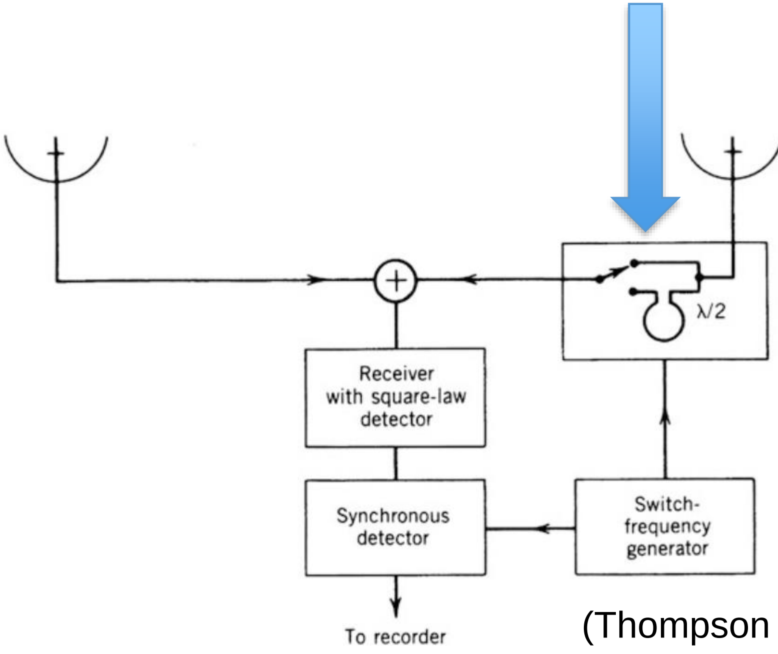
- In addition to the signal from the source, the output of the receiver contains other noise components:
  - Galactic background
  - Thermal noise from ground in sidelobes
  - Noise generated by amplifiers
- Component from the source is usually << total noise power.
- Thus one has to remove a large offset.
- But the drifts in this offset degrade the detectability of weak sources and accuracy of measurement of fringes.

In the 1950s, offset drifts sometimes caused the chart recorder pen to go off scale!

1952 Phase switching interferometer

- The introduction of phase switching removed the unwanted components of the receiver output. **How?**
- If V_1 and V_2 are the signal voltages from the two antennas, output of the adding interferometer is proportional to $(V_1 + V_2)^2$.
- In the phase switching system, the phase of one of the signals is periodically reversed.
- The output of the detector alternates between $(V_1 + V_2)^2$ and $(V_1 - V_2)^2$
- A detector takes the difference between these two terms which is proportional to $V_1 V_2$

The phase is reversed by switching to use a $\lambda/2$ longer length of wire

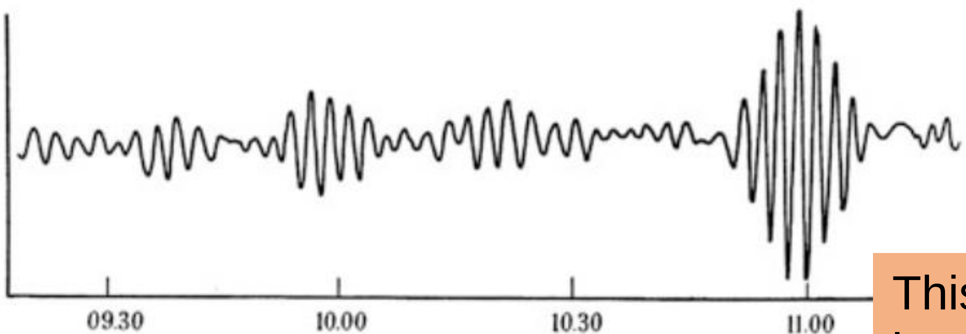


(Thompson et. al)

1952 Phase switching interferometer

- When V_1 is multiplied with V_2 rather than added and squared, the constant term in the output P (from earlier) disappears.
- Thus the output is the time average of the product of the signal voltages i.e. proportional to the cross-correlation between the signals.
- The circuitry that multiples and averages the signals in a modern interferometer (though by a different mechanism) is known as a correlator.

The removal of the constant term reduces sensitivity to instrumental gain variation. This allows installation of amplifiers at antennas to overcome attenuation in transmission lines.



(Thompson et. al)

This significant advance allowed for using longer antenna spacings and larger arrays.

1950s Optical identification and calibration sources

- Interferometer observations gave evidence of numerous discrete sources.
- But <u>how to identify the optical counterpart</u> of these sources?
- This required accurate measurement of radio positions. How to do this?
- Determine the time of transit of the central fringe using an E-W baseline.
- The declination is related to the fringe frequency, as derived earlier.

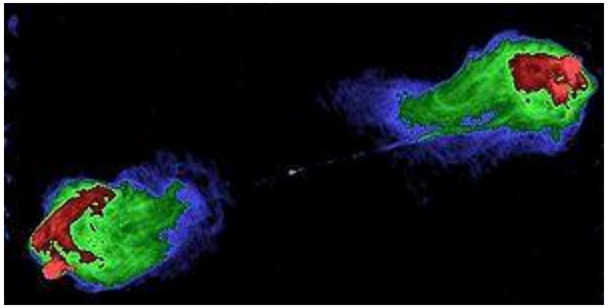
Why? is this difficult to do?

- Position measurement is only as accurate as knowledge of the fringe pattern. This is determined by the relative locations of the electrical centres of antennas.
- Any inequality in the electrical path lengths from the antennas to where they are combined introduces a phase term which offsets the fringe pattern.

1950s Optical identification and calibration sources

Results

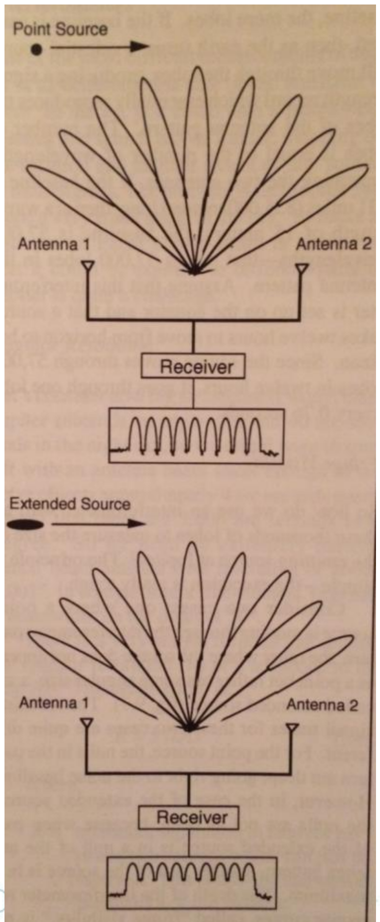
- Cygnus A was proved to be a distant galaxy.
- Cassiopeia A was proved to be a supernova remnant.



But you get something extra

Once you have identified radio sources with optical objects, for which you know the positions, you can calibrate for the unknowns in the baseline parameters and fringe phases! This removed the need for absolute calibration of the antennas and receiving system.

Experiment How to use two-element interferometer to measure the angular size of emitting source at Jupiter?



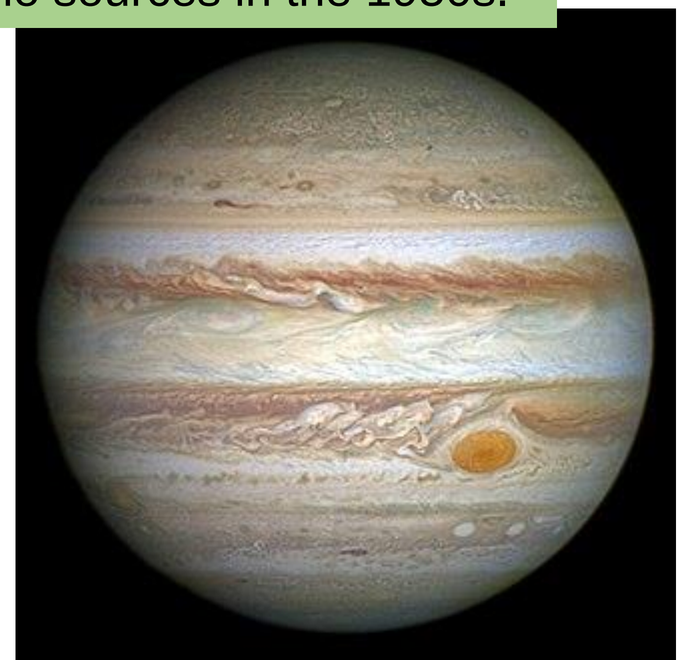
Consider two cases:

1. a point source is moving through the fringe pattern
2. an extended source is moving through the fringe pattern.

In 1. the nulls are deep, going right to the noise baseline.

In 2. the nulls are not as deep, because when part of the extended source is in the null of the fringe pattern, another part of the source is in a maximum.

This was the kind of experiment done for various radio sources in the 1950s.



As with the Michelson interferometer, the ‘fringe visibility’ is related to the angular size of the source.

(NASA)

Jupiter experiment - continued

- So how do you measure the size of the emitting source? **Do you remember** how Michelson and Pease determined the diameter of Betelgeuse?

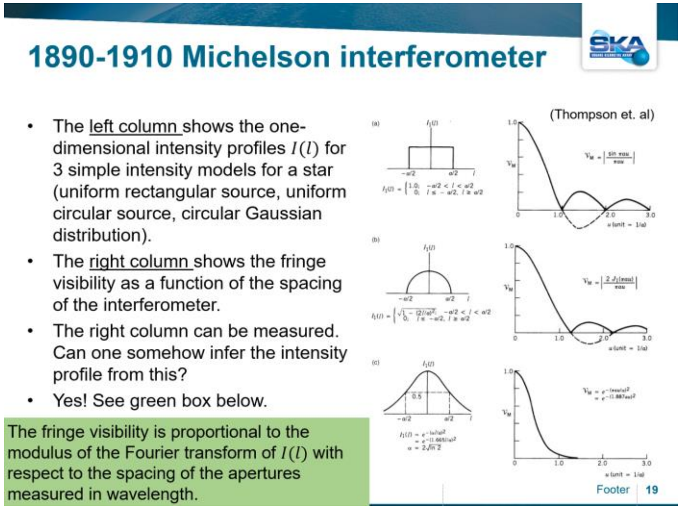
- Start with the two interferometers fairly close together and keep increasing the baseline until the fringe visibility indicates that the source has been resolved.

- Why? is this experiment challenging in a way that the Michelson experiment isn't?

- Because unlike the Michelson interferometer, where the aperture spacing can be adjusted in the instrument, this experiment requires actually physically moving the antennas. **Good luck!**

But what about imaging?

- Okay. So two-element interferometers can give us angular sizes of stars and planets, positions of sources, etc. But what about imaging? How do interferometers give us images such as this?
- The clue is this slide from earlier..



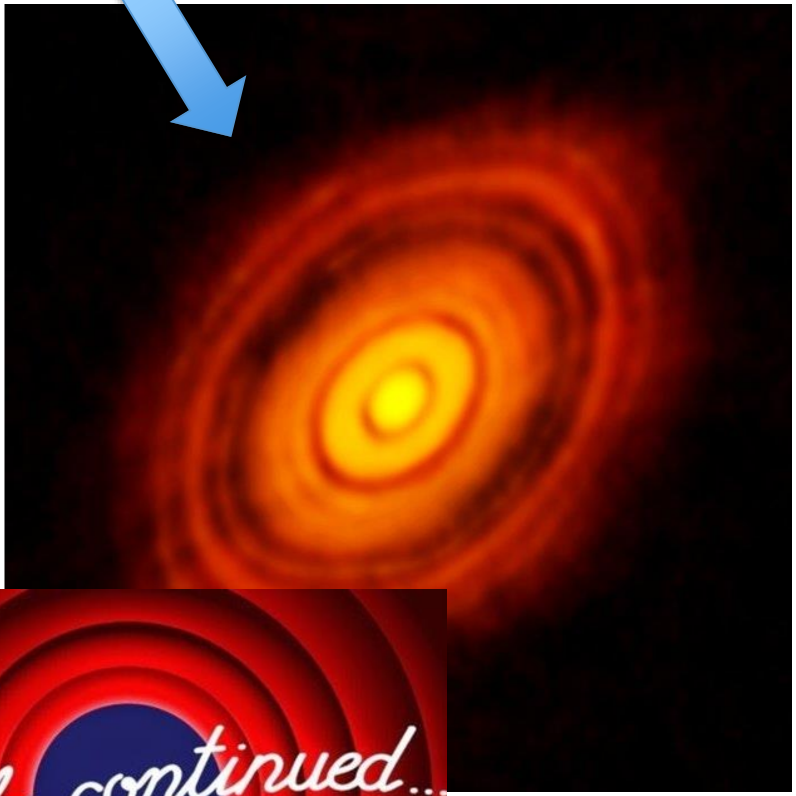
1890-1910 Michelson interferometer

- The left column shows the one-dimensional intensity profiles $I(l)$ for 3 simple intensity models for a star (uniform rectangular source, uniform circular source, circular Gaussian distribution).
- The right column shows the fringe visibility as a function of the spacing of the interferometer.
- The right column can be measured. Can one somehow infer the intensity profile from this?
- Yes! See green box below.

The fringe visibility is proportional to the modulus of the Fourier transform of $I(l)$ with respect to the spacing of the apertures measured in wavelength.

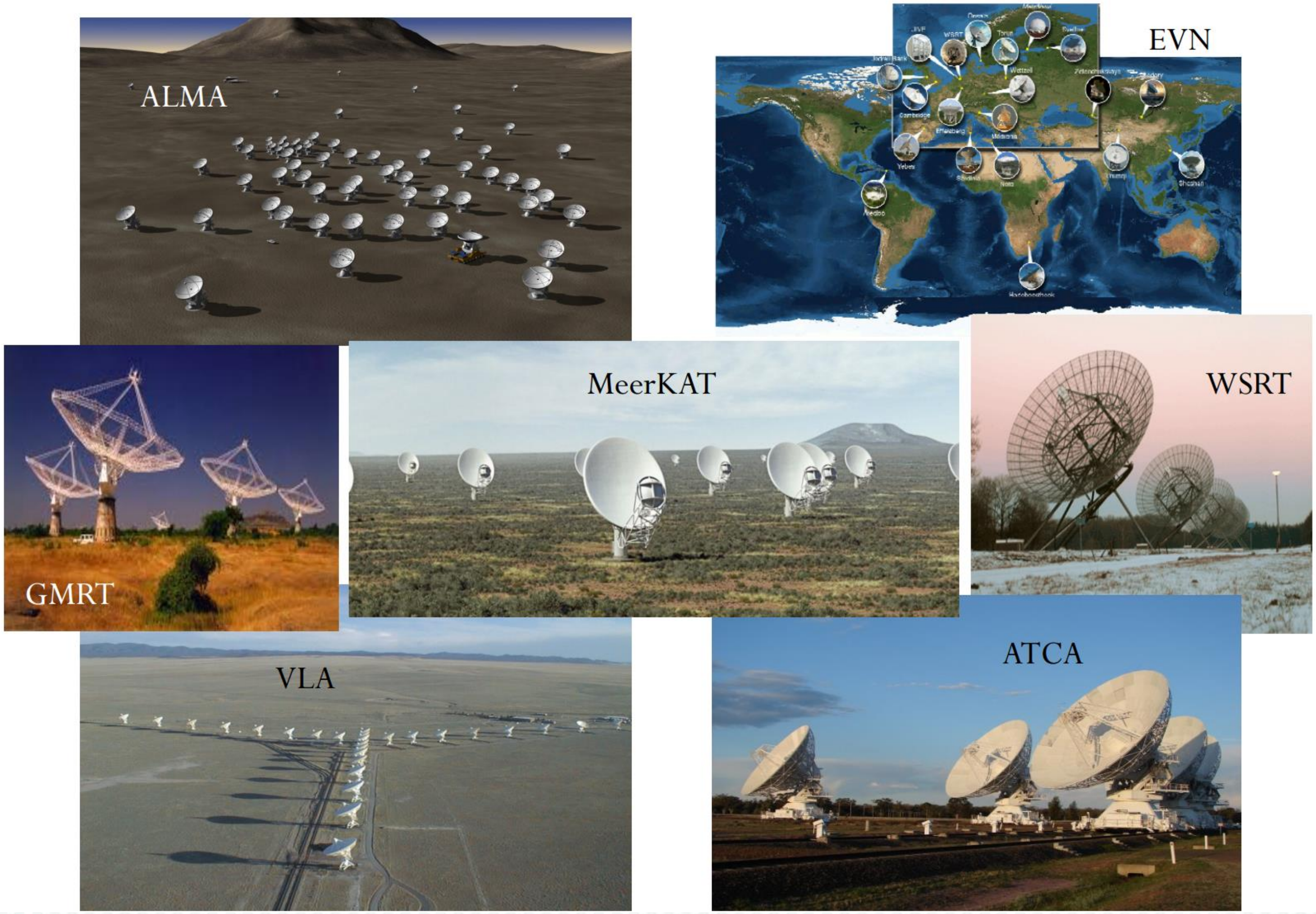
(Thompson et. al)

Footer 19



Interferometers around the world

(Slide taken from Shafi 2017)



Thank you!