Optical interferometry – a gentle introduction to the theory

Chris Haniff

Astrophysics Group, Cavendish Laboratory,
Madingley Road, Cambridge,
CB3 0HE, UK

Motivation

- A – Uninterested: I’m here for the holiday.
- B – Might be interested: I’m sceptical: prove it to me!
- C – Possibly interested: I need to learn more.
- D – Interested: I want to work to understand this.

What we will cover

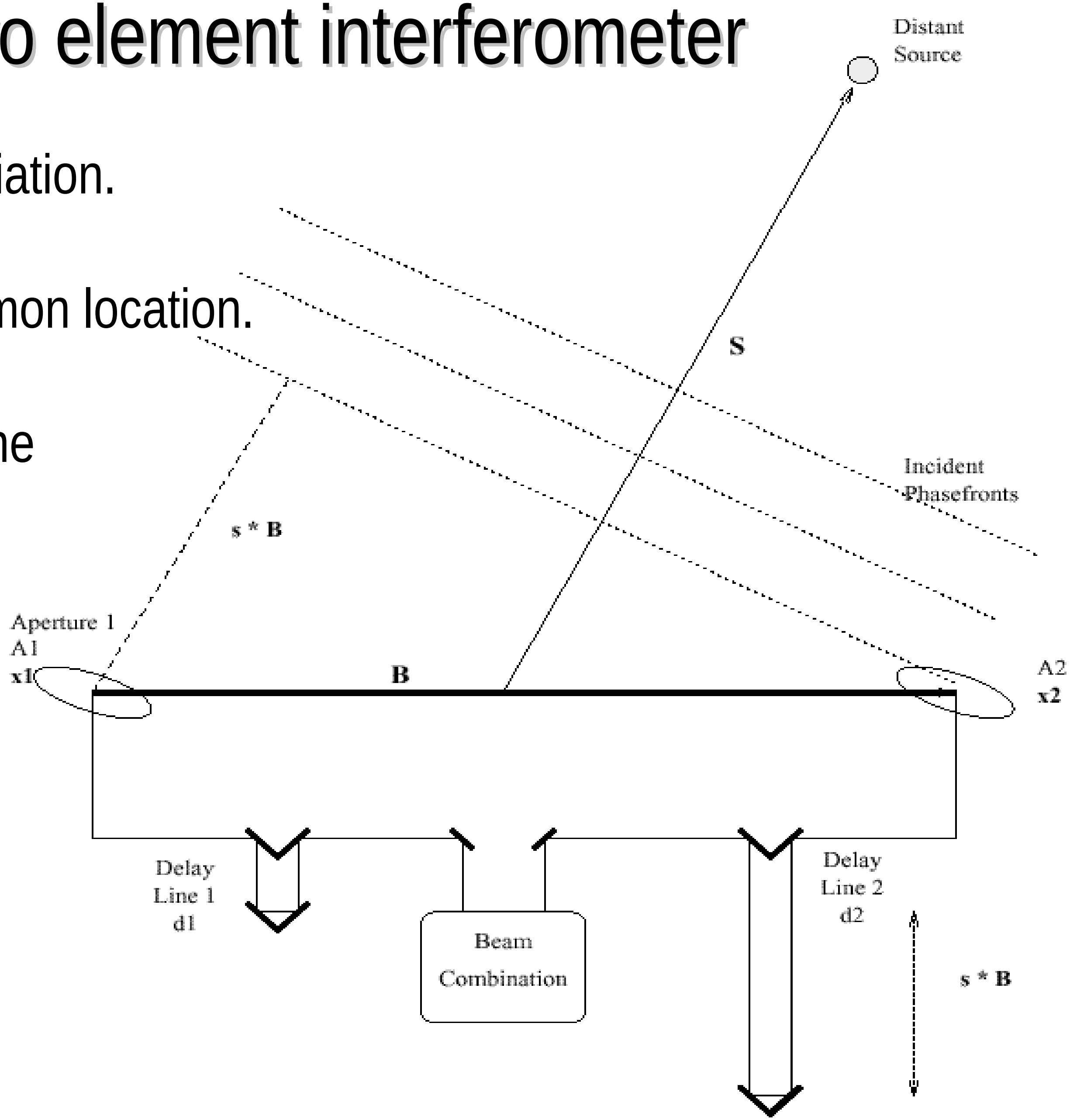
- What no one ever tells you (or admits to).
- What an interferometer does.
- The output of an interferometer.
- How to describe interference fringes.
- What interferometers tell you about sources (qualitative).
- What interferometers tell you about sources (quantitative).
- Visibility functions.
- Imaging with interferometers.

Preamble

- Learning interferometry is like learning to surf:
  - You have to want to do it.
  - You start in the shallows.
  - Having an expensive surf-board doesn't help.
  - You don't have to know how to make surf-boards.
  - Knowing how to surf won't help you escape a charging tiger.
- This is a school:
  - I will assume you know nothing - you should assume the same.
  - Don't guess - physics is not intuitive.
  - Ask questions - last year those who didn't went away confused.
  - If you don't understand ask.
- I am not trying to sell you a surf board:
  - Interferometry is a niche technique - it's not the solution to every astronomical problem.

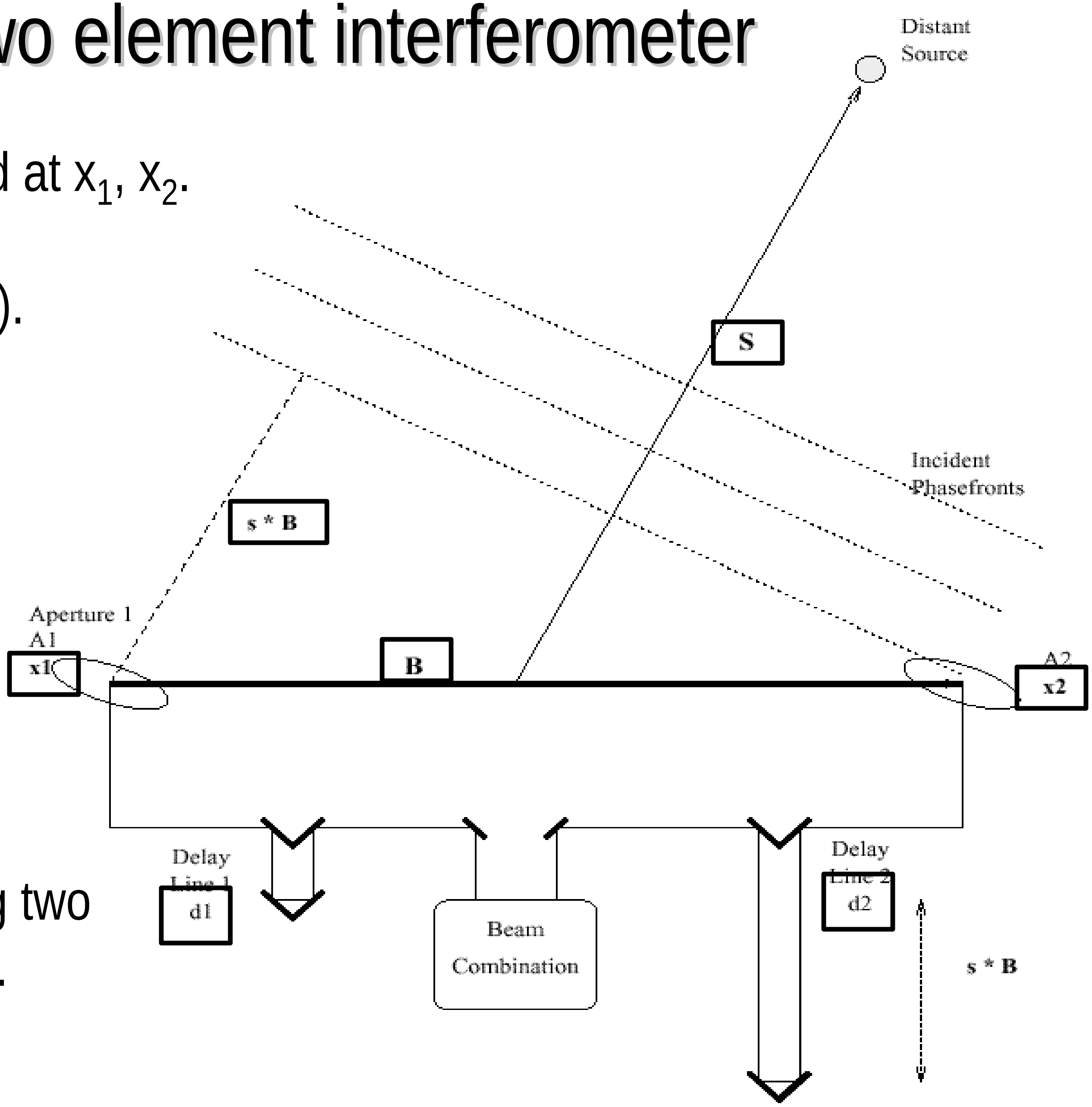
A two element interferometer

- Sampling of the radiation.
- Transport to a common location.
- Compensation for the geometric delay.
- Combination of the beams.
- Detection of the resulting output.



A two element interferometer

- Telescopes located at x_1, x_2.
- Baseline B = (x_1-x_2).
- Pointing direction is S.
- Geometric delay is s.B, where s = S/|S|.
- Optical paths along two arms are d_1 and d_2.



Key Ideas 1

- Functions of an interferometer:
  - Sampling.
  - Optical path matching.
  - Combination of electric fields.
  - Detection.

- Nomenclature:
  - Baseline.
  - Pointing direction.
  - Geometric delay.

The output of a 2-element interferometer (i)

- At combination the E fields from the two apertures can be described as:
  - $\psi_1 = A \exp(ik[s.B + d_1]) \exp(-i\omega t)$ and $\psi_2 = A \exp(ik[d_2]) \exp(-i\omega t)$

- So, summing these at the detector we get:

$$\Psi = \psi_1 + \psi_2 = A \left[\exp(ik[s.B + d_1]) + \exp(ik[d_2])\right] \exp(-i\omega t)$$

- And hence the time averaged intensity, $\langle \Psi\Psi^* \rangle$, will be given by:

$$\langle \Psi\Psi^* \rangle \propto \langle [\exp(ik[s.B + d_1]) + \exp(ik[d_2])] \times [\exp(-ik[s.B + d_1]) + \exp(-ik[d_2])] \rangle$$
$$\propto 2 + 2\cos(k[s.B + d_1 - d_2])$$
$$\propto 2 + 2\cos(kD)$$

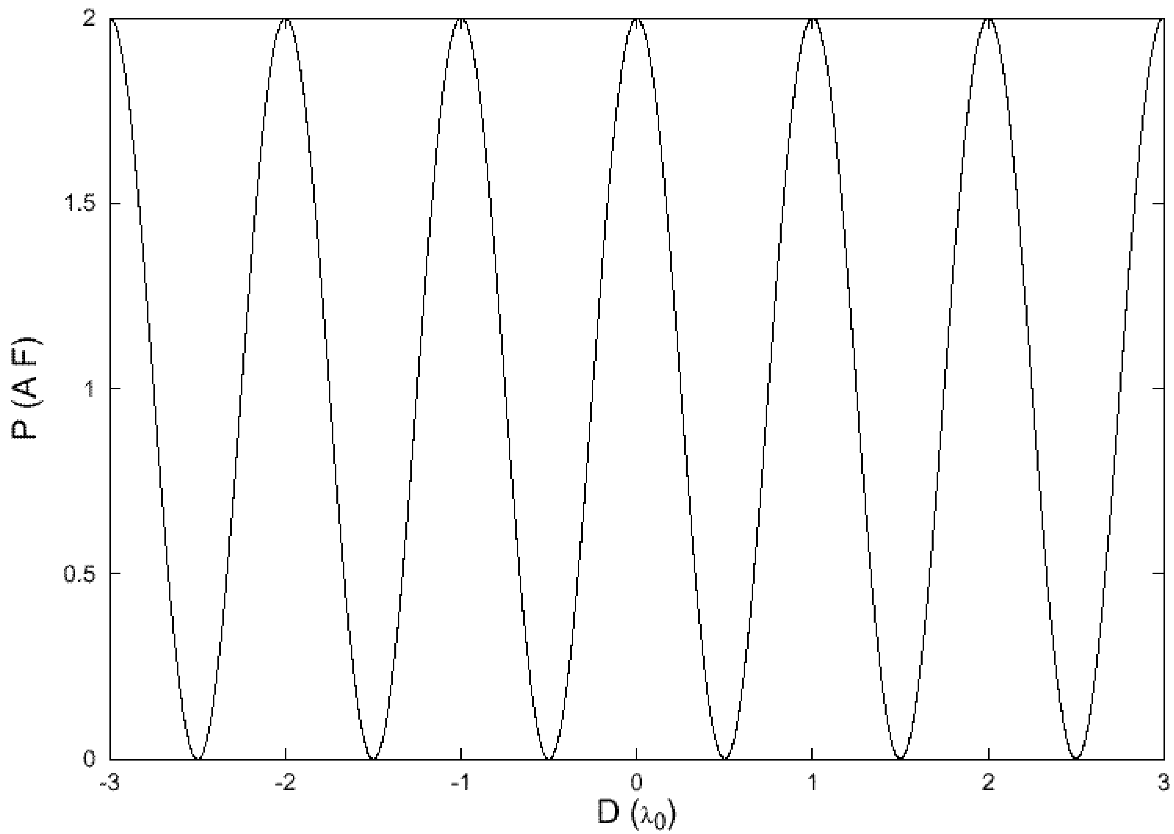
Note, here $D = [s.B + d_1 - d_2]$.

This is a function of the path lengths, d_1 and d_2, the pointing direction and the baseline.

The output of a 2-element interferometer (ii)

Detected power, $P = \langle \Psi\Psi^* \rangle \propto 2 + 2 \cos (k [s.B + d_1 - d_2])$

$\propto 2 + 2 \cos (kD)$, where $D = [s.B + d_1 - d_2]$



– The output varies co-sinusiodally with D.

– Adjacent fringe peaks are separated by $\Delta d_{1\ or\ 2} = \lambda$ or $\Delta s = \lambda/B$.

Key ideas 2

- The output of the interferometer is a time averaged intensity.
- It has a cosinusoidal variation - these are the “interference fringes”.
- The cosinusoidal variation is a function of k.D, which in turn can depend on many things:
  - The wavevector, $k = 2\pi/\lambda$.
  - The baseline, B.
  - The pointing directions, s.
  - The optical path difference between the two arms of the interferometer, d_1-d_2.

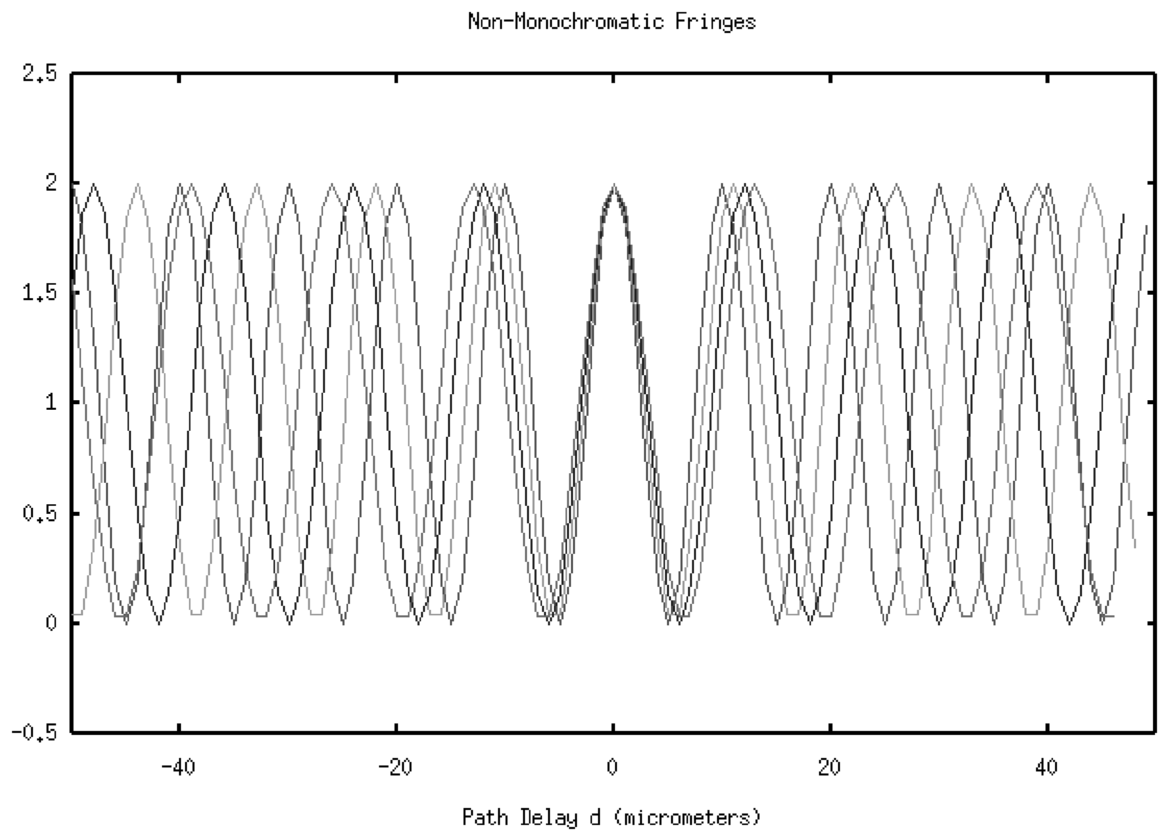
- Note that if you adjust things correctly, the output is fixed. This is what most interferometers actually aim to do.

Extension to polychromatic light

- We can integrate the previous result over a range of wavelengths:
  - E.g for a uniform bandpass of $\lambda_0 \pm \Delta\lambda/2$ (i.e. $\nu_0 \pm \Delta\nu/2$) we obtain

$$P \propto \int_{l_0-\Delta l/2}^{l_0+\Delta l/2} [2+2\cos(kD)]\, d\, l$$

$$= \int_{l_0-\Delta l/2}^{l_0+\Delta l/2} 2\,[1+\cos(2pD/\,l\,c)]\, d\, l$$



Extension to polychromatic light

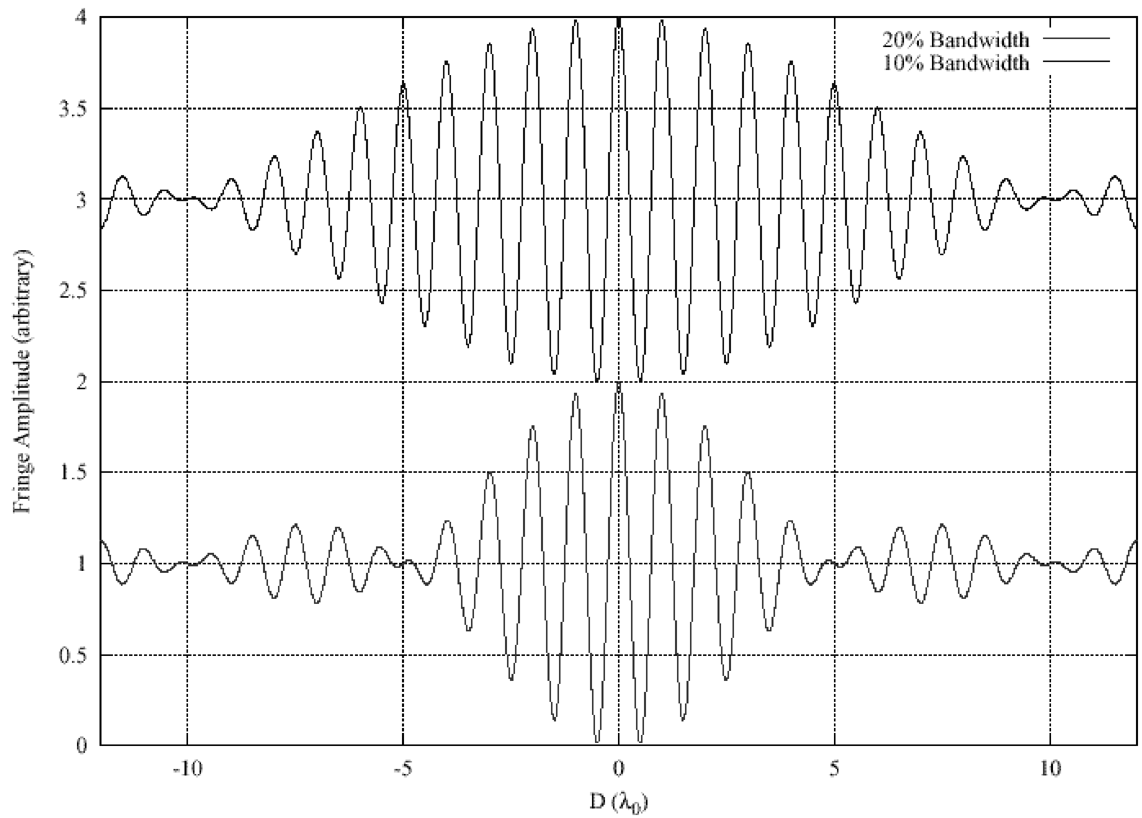
- We can integrate the previous result over a range of wavelengths:
  - E.g for a uniform bandpass of $\lambda_0 \pm \Delta\lambda/2$ (i.e. $\nu_0 \pm \Delta\nu/2$) we obtain:

$$P \propto \int_{l_0-\Delta l/2}^{l_0+\Delta l/2} [2+2\cos(kD)]\, dl$$

$$= \int_{l_0-\Delta l/2}^{l_0+\Delta l/2} 2[1+\cos(2pD/lc)]\, dl$$

$$= \Delta l \left[1 + \frac{\sin pD\Delta l / l^2{}_0}{pD\Delta l / l^2{}_0} \cos k_0 D \right]$$

$$= \Delta l \left[1 + \frac{\sin pD/\Lambda_{coh}}{pD/\Lambda_{coh}} \cos k_0 D \right]$$



So, the fringes are modulated with an envelope with a characteristic width equal to the coherence length, $\Lambda_{coh} = \lambda^2{}_0/\Delta\lambda$.

Key ideas 3

- The response for a polychromatic source is given by integrating the intensity response for each color.
- This alters the interferometric response and can lead to “removal” of the fringe modulation completely:
  - The correct response is only achieved when k [s.B + d_1 - d_2]) = 0.
  - This is the so called white-light condition.

This is the primary motivation for matching the optical paths in an interferometer and correcting for the geometric delay.

- The narrower the range of wavelengths detected, the smaller is the effect of this modulation:
  - This is usually quantified via the coherence length, $\Lambda_{coh} = \lambda^2_0/\Delta\lambda$.
  - But narrower bandpasses mean less light!

Fringe parameters of interest

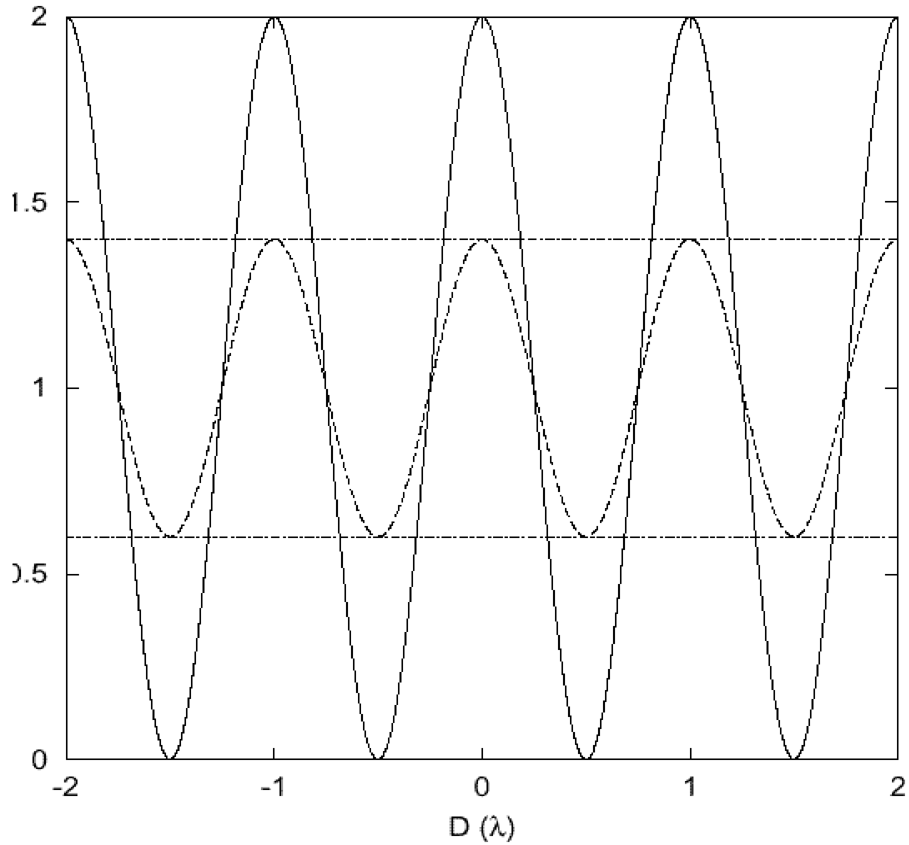
- From an interferometric point of view the key features of any interference fringes are their modulation and their location with respect to some reference point.

- In particular we can identify:

  - The fringe visibility:

$$V = \frac{[I_{max} - I_{min}]}{[I_{max} + I_{min}]}$$

  - The fringe phase:
    - The location of the white-light fringe as measured from some reference (radians).

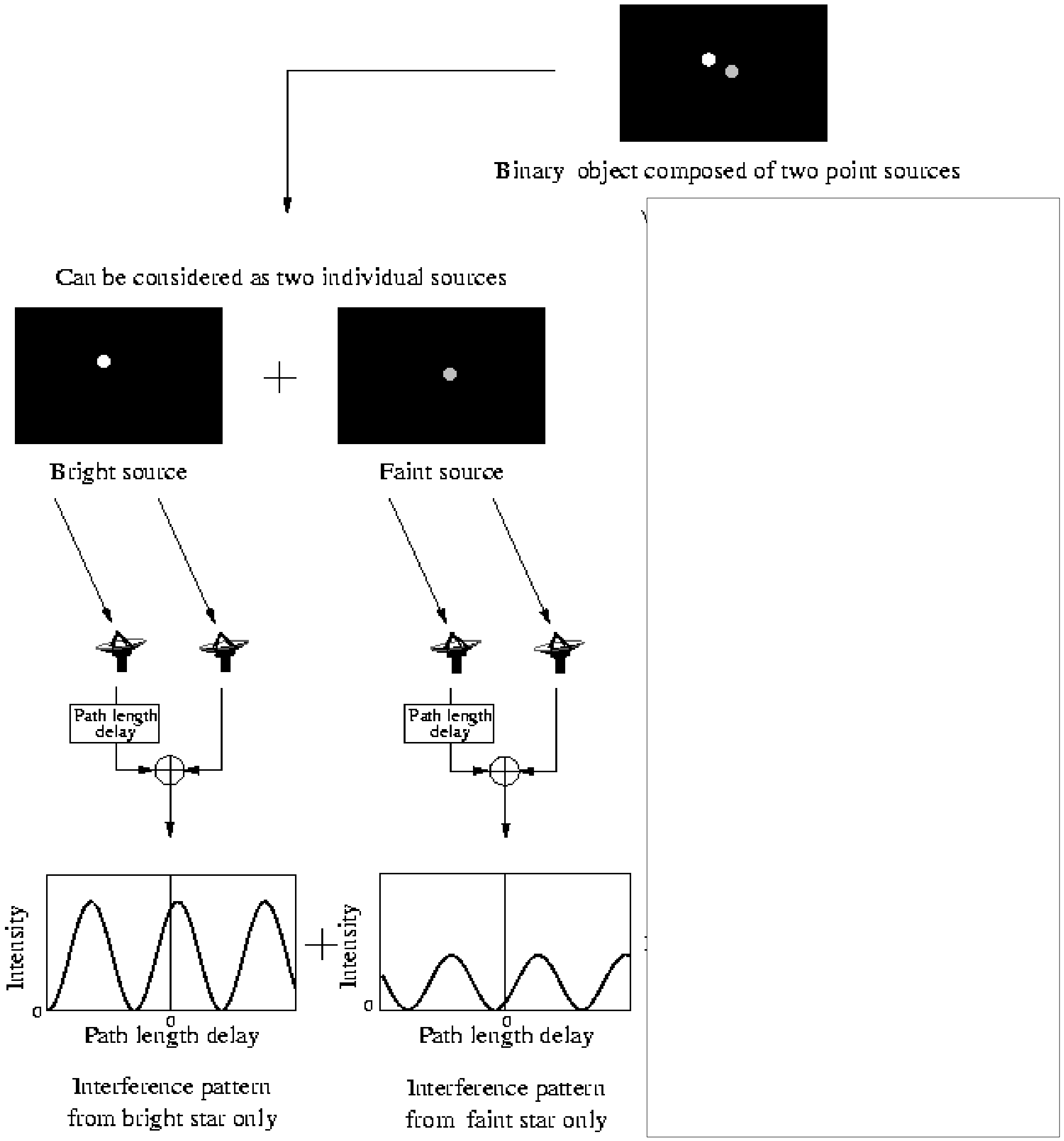


Key ideas 4

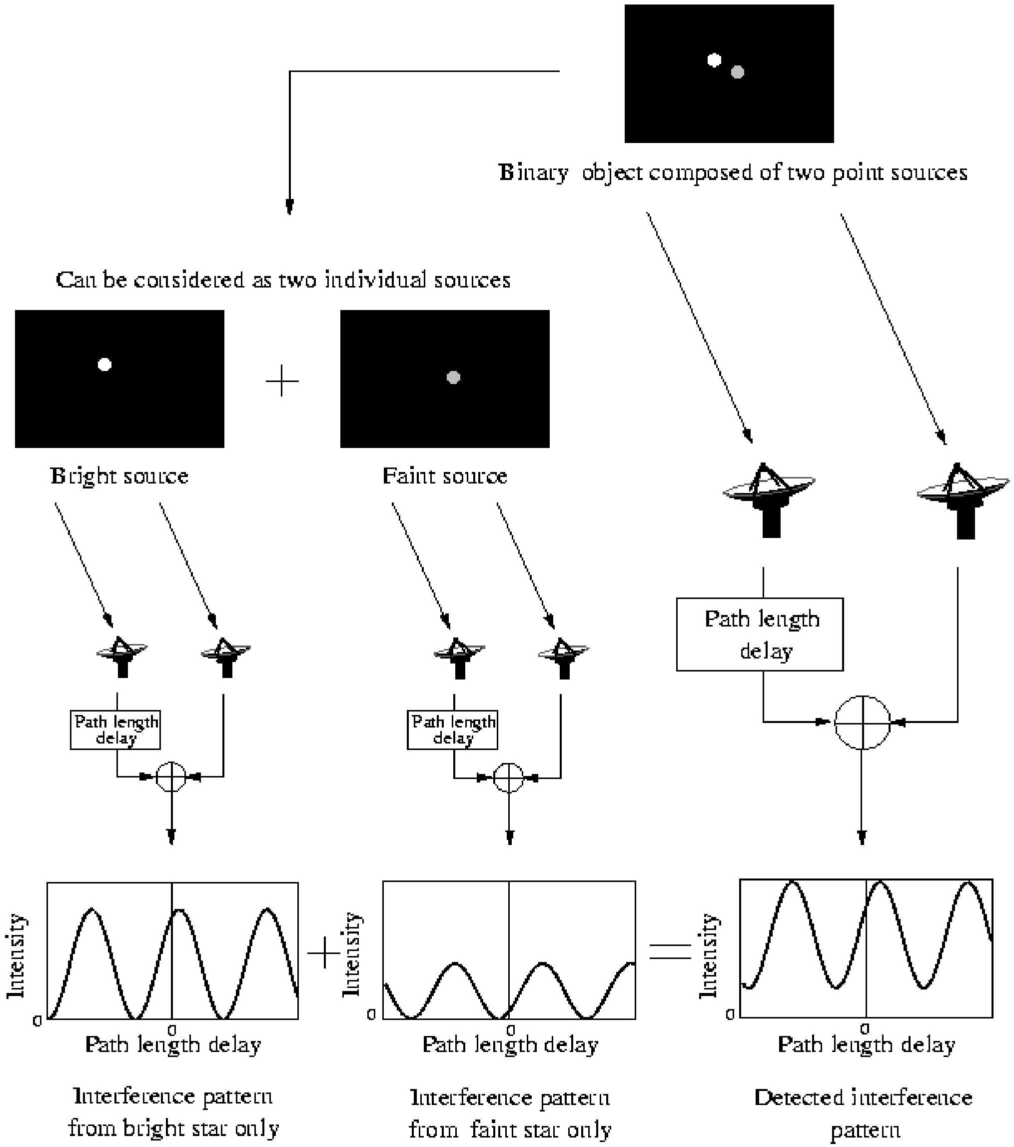
- The parameters of interference fringes that we are usually interested in are:
  - The fringe contrast (<u>excluding any finite bandwidth effects</u>).
  - The fringe phase.
- We are usually not interested in:
  - The fringe period.

- The question you should all be asking now is:
  - Why is it that <u>these</u> are the parameters of interest?
  - And what do they tell us?

Heuristic operation of an interferometer



Heuristic operation of an interferometer



- The resulting fringe pattern has a modulation depth that is reduced with respect to that from each source individually.
- The positions of the sources are encoded in the fringe phase.

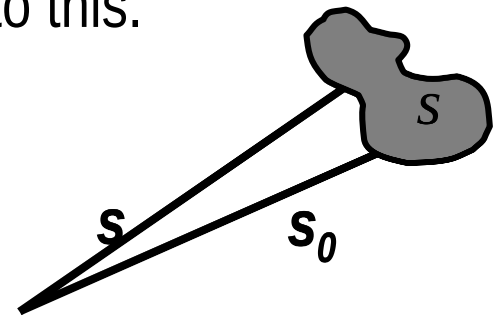
Key ideas 5

- A general source can be described as a superposition of point sources.
- Each of these produces its own interference pattern.
- The superposition of these is what is actually measured.
  - Technically this is know as the “spatially-incoherent source” approximation.

The modulation and phase of the resulting fringe pattern encode the source structure (albeit in an apparently complicated way).

Response to a distributed source

- Consider looking at an incoherent source whose brightness on the sky is described by I(s). This can be written as $I(s_0+\Delta s)$, where s_0 is the pointing direction, and Δs is a vector perpendicular to this.



- The detected power will be given by:

$$P(s_0, B) \propto \int I(s)\left[1+\cos kD\right] d\Omega$$

$$\propto \int I(s)\left[1+\cos k(s.B+d_1-d_2)\right] d\Omega$$

$$\propto \int I(s)\left[1+\cos k([s_0+\Delta s].B+d_1-d_2)\right] d\Omega$$

$$\propto \int I(s)\left[1+\cos k(s_0.B+\Delta s.B+d_1-d_2)\right] d\Omega$$

$$\propto \int I(\Delta s)\left[1+\cos k(\Delta s.B)\right] d\Omega'$$

The van Cittert-Zernike theorem (i)

- Consider now adding a small phase delay, δ, to one arm of the interferometer. The detected power will become:

$$P(s_0, B, d) \propto \int I(\Delta s)\left[1+\cos k(\Delta s.B+d)\right] d\Omega'$$

$$\propto \int I(\Delta s)\, d\Omega' + \cos kd \cdot \int I(\Delta s) \cos k(\Delta s.B)\, d\Omega'$$

$$- \sin kd \cdot \int I(\Delta s) \sin k(\Delta s.B)\, d\Omega'$$

- We now define something called the complex visibility V(*k*,*B*):

$$V(k, B) = \int I(\Delta s) \exp[-ik\Delta s.B]\, d\Omega'$$

so that we can write our interferometer output as:

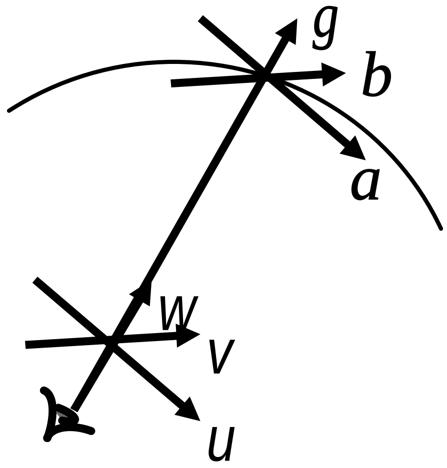
$$P(s_0, B, d) \propto \int I(\Delta s)\ \ d\Omega' + \cos kd\ \mathrm{Re}[V] + \sin kd\ \mathrm{Im}[V]$$

$$P(s_0, B, d)\ =\ I_{total} + \mathrm{Re}\left[V \exp[-ikd\,]\right]$$

What is this complex visibility thing?

- Lets assume s_0 = (0,0,1) and Δs is $\approx$ (α,β,0), with α and β small angles measured in radians.

$$V(k,B) = \int I(\Delta s)\exp[-ik\Delta s.B]\, d\Omega'$$
$$= \int I(a\,,\,b\,)\exp[-ik(aB_x + bB_y)]\, da\ db$$
$$= \int I(a\,,\,b\,)\exp[-i2p\,(au + bv)]\, da\ db$$



- Here, *u* (= B_x/λ) and *v* (= B_y/λ) are the projections of the baseline onto a plane perpendicular to the pointing direction.
  - These are usually referred to as spatial frequencies and have units of rad^{-1}.

So, the complex visibility is the Fourier Transform of the source brightness distribution.

The van Cittert Zernike theorem (ii)

- We can put this all together as follows:
- Our interferometer measures $$P(s_0, B, d) = I_{total} + \mathrm{Re}\left[V \exp[-ikd]\right]$$
- So, if we make measurements with, say, two value of $\delta = 0$ and $\lambda/4$, this recovers the real and imaginary parts of the complex visibility.
- And, since the complex visibility is nothing more than the Fourier transform of the brightness distribution, we have our final result:

The output of an interferometer measures the Fourier transform of the source brightness distribution.

This is the van Cittert-Zernike theorem.

Key ideas 6

- The complex visibility is also known as the “spatial coherence” function.
- Since the FT is a linear transform, if we know the complex visibility we can recover the source brightness distribution.

- Since the visibility function is complex, it has an amplitude and a phase.
- The amplitude and phase of the interference fringes we spoke of earlier, are actually the amplitude and phase of the complex visibility.
- To measure these quantities we have to adjust D.

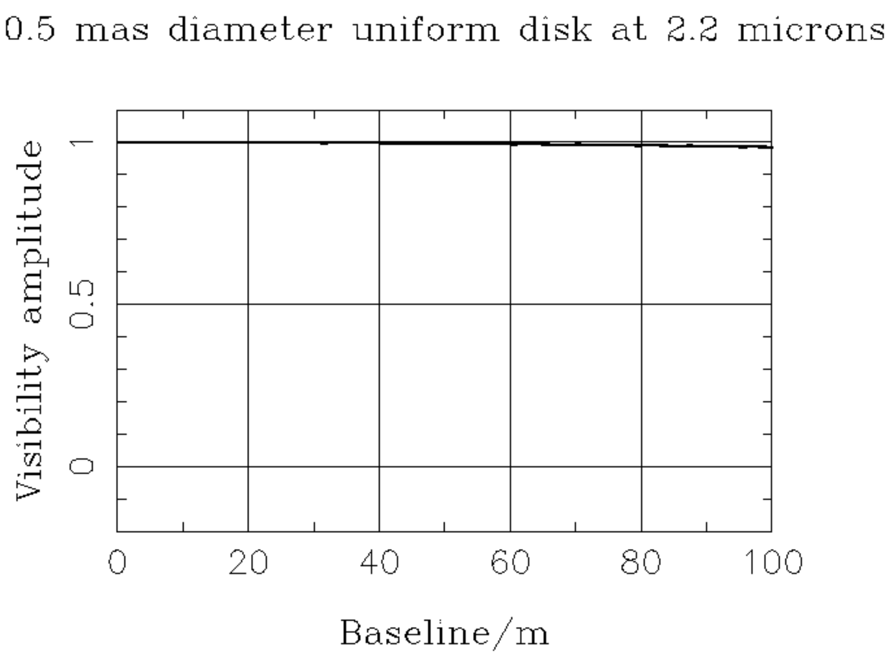
- A measurement from a single interferometer baseline gives a measurement of one value of the FT of the source brightness distribution.
- Long interferometer baselines measure small structures on the sky, and short baselines, large structures.

Visibility functions of simple sources (i)

$$V(u) = \int I(\alpha)\, e^{-i2\pi(u\alpha)}\, d\alpha / \int I(\alpha)\, d\alpha$$

- Point source of strength A_1 and located at angle α_1 relative to the optical axis.

$$V(u) = \int A_1\delta(\alpha-\alpha_1)\, e^{-i2\pi(u\alpha)}\, d\alpha \,/ \int A_1\delta(\alpha-\alpha_1)d\alpha$$
$$= e^{-i2\pi(u\alpha_1)}$$



- The visibility amplitude is unity ∀ *u.*
- The visibility phase varies linearly with *u* (= B/λ).
- Since |V| is close to unity, the interference
  fringes have high contrast.

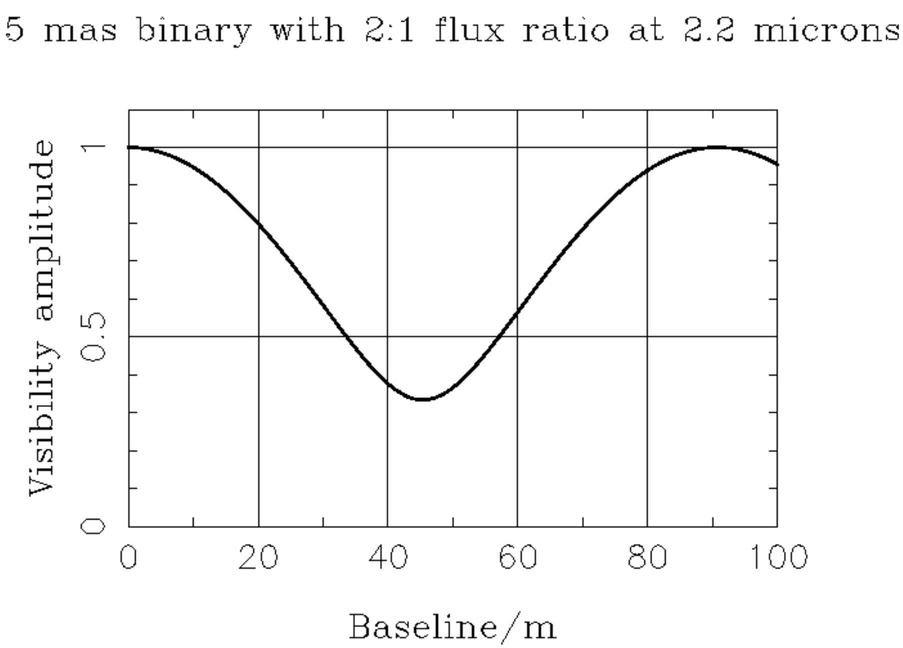
Visibility functions (ii)

$$V(u) = \int I(\alpha)\, e^{-i2\pi(u\alpha)}\, d\alpha / \int I(\alpha)\, d\alpha$$

- A double source comprising point sources of strength A_1 and A_2 located at angles 0 and α_2 relative to the optical axis.

$$V(u) = \int [A_1\delta(\alpha) + A_2\delta(\alpha-\alpha_2)]\, e^{-i2\pi(u\alpha)}\, d\alpha \,/ \int [A_1\delta(\alpha) + A_2\delta(\alpha-\alpha_2)]\, d\alpha$$

$$= [A_1 + A_2 e^{-i2\pi(u\alpha_2)}] / [A_1 + A_2]$$



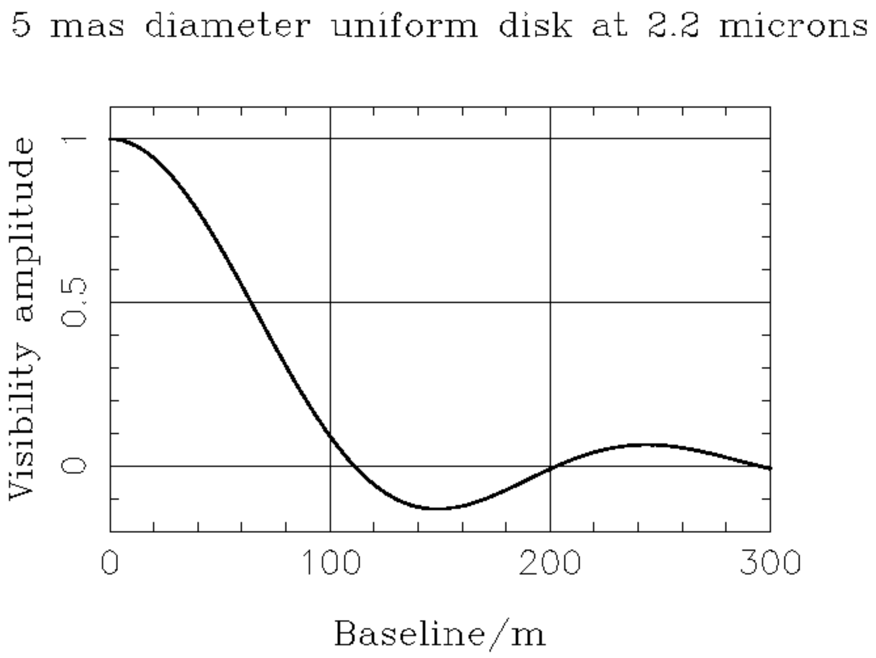
- The visibility amplitude and phase oscillate as functions of *u*.
- To identify this as a binary, baselines from $0 \rightarrow \lambda/\alpha_2$ are required.
- The modulation of the visibility function tells us the separation and brightness ratio of the components.

Visibility functions (iii)

$$V(u) = \int I(\alpha)\, e^{-i2\pi(u\alpha)}\, d\alpha / \int I(\alpha)\, d\alpha$$

- A uniform on-axis disc source of diameter θ.

$$V(u_r) \propto \int^{\theta/2} \rho\, J_0(2\pi\rho u_r)\, d\rho$$
$$= 2J_1(\pi\theta u_r) \,/\, (\pi\theta u_r)$$



- To identify this as a disc requires baselines from $0 \rightarrow \lambda/\theta$ at least.
- The visibility amplitude falls rapidly as u_r increases.
- Information on scales smaller than the disc diameter correspond to values of u_r where $V << 1$, where the interference fringes have very low contrast.

Key ideas 7

- Unresolved, sources have visibility functions that remain high, giving produce high contrast fringes for all baseline lengths.
- Resolved sources have visibility functions that fall to low values at long baselines, giving fringes with very low contrast.

  => Fringe parameters for resolved sources will be difficult to measure.

- Imaging with many resolution elements generally needs measurements where the fringe contrast is both high and low (to pick out large scale and small scale features respectively).

- To usefully constrain a source, the visibility function must be measured adequately. Measurements on a single, or small number of, baselines are normally not enough for unambiguous image recovery.

Introduction to interferometric imaging

- The visibility function, *V(u, v)* is the Fourier transform of the source brightness distribution:

$$V(u,v) = \int I(a\,,\,b\,)\exp[-i2p\,(au + bv)]\,da\;db$$

- So the idea is to measure V for as many values of u and v as possible & perform an inverse FT:

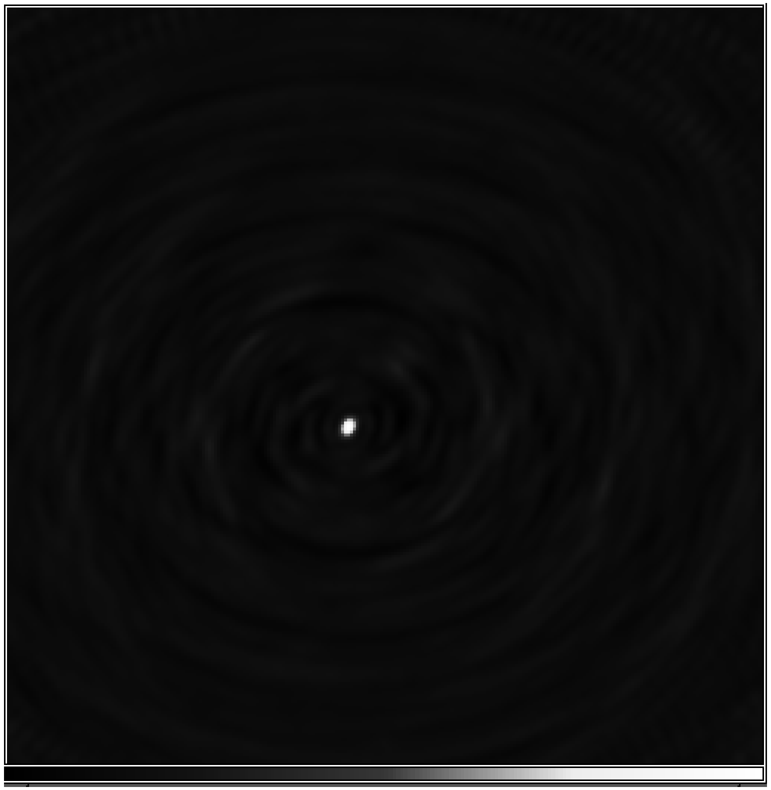
$$\iint V(u,v)\exp[+i2p\,(ua\;+vb\,)]\,du\;dv = I_{norm}(a\,,\,b\,)$$

- But since what we measure is a sampled version of *V(u, v)*, what we actually recover is the so-called “dirty map”:

$$\iint S(u,v)\,V(u,v)\exp[+i2p\,(ua\;+vb\,)]\,du\;dv = I_{dirty}(a\,,\,b\,)$$
$$= I_{norm}(a\,,\,b\,) * B_{dirty}(a\,,\,b\,)$$

B_{dirty}(l,m) is the Fourier transform of the sampling distribution, and is known as the dirty-beam.

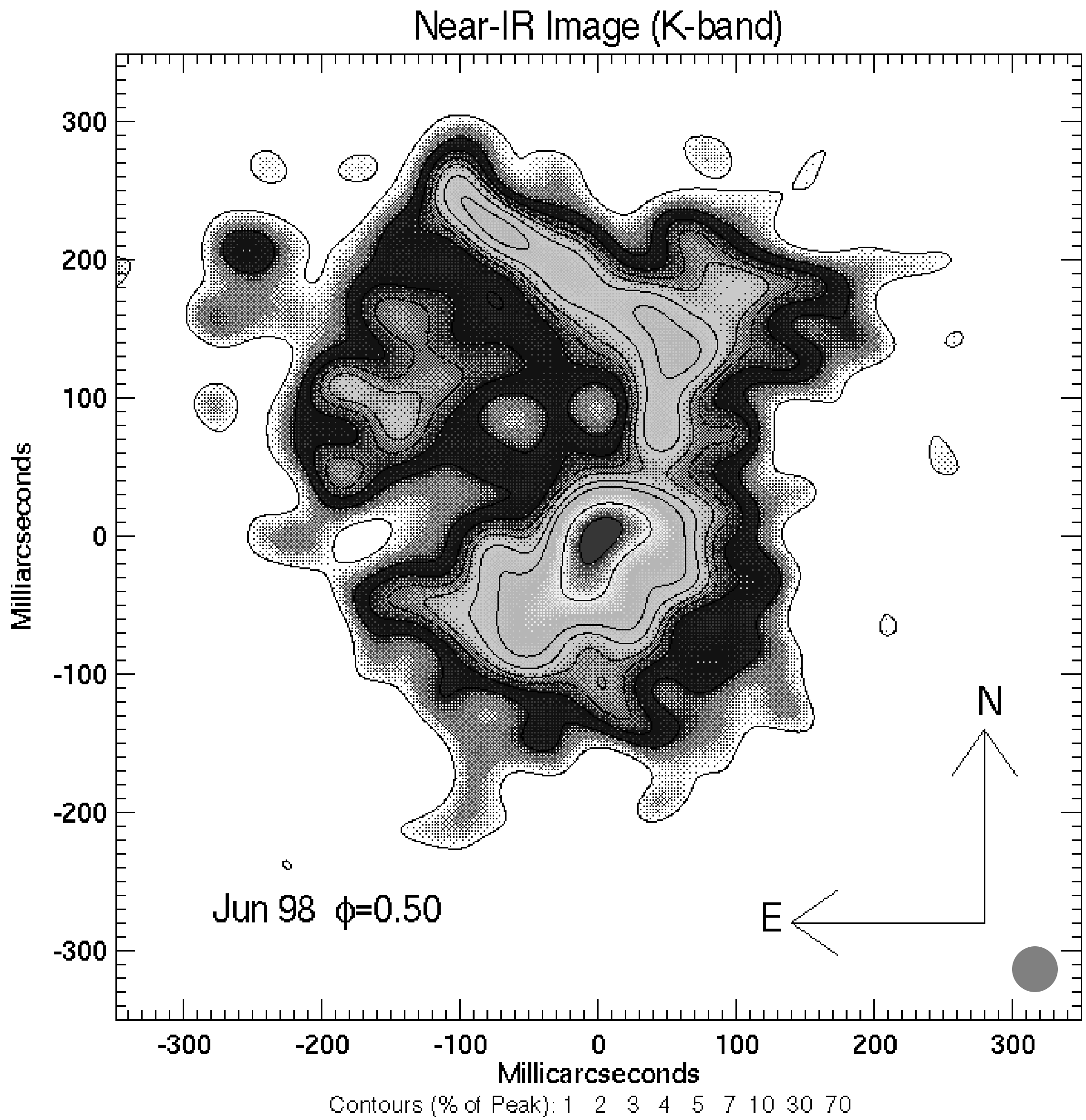
Dirty (and corrected) interferometric images



- Raw interferometric maps generally look awful - but correcting for dirty beam (known as deconvolution- CLEAN, MEM, WIPE) is straightforward.

A real astronomical example

K-band image of IRC+10216. Image courtesy of Peter Tuthill and John Monnier.



Key ideas 8

- Imaging with an interferometer => measuring the visibility function for a wide range of baselines.
- It also => measuring its amplitude and phase.

- The map you get will ONLY contain information corresponding to the baselines you measured.
  - This applies to conventional imaging as well

- There is no such thing as the “correct” image.

Summary

- Interferometers are machines to make fringes.
- The fringe modulation and phase tell you what you are looking at.
- More precisely, these measure the amplitude and phase of the FT of the source brightness distribution.
- A measurement with a given interferometer measures one value of the FT of the source brightness distribution.
- Multiple baselines are obligatory to build up an image.
- Once many visibility measurements are made, an inverse FT delivers a representation of the source that may (or may not) be useful!