VLTI Summer School 2021

Introduction to interferometry

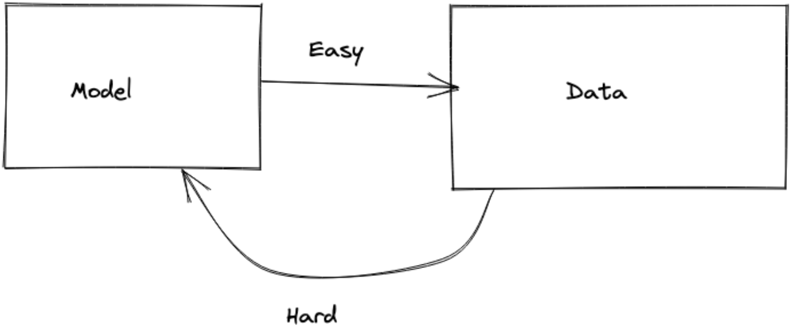
David Buscher

2021-06-07

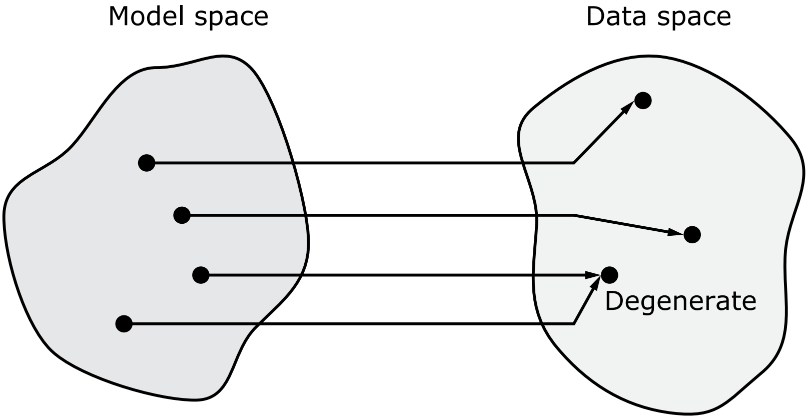
- How to learn interferometry
- The phase-coherent interferometer
  - Making fringes
  - From fringes to images
- The atmosphere and quantum noise
- Interferometry in practice

The need for interferometry

Science consists of solving inverse problems



Degeneracies mean that the inverse is not unique

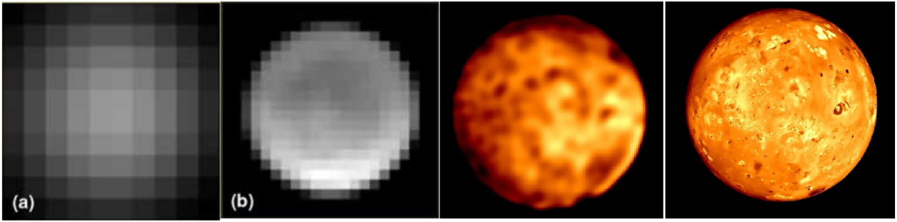


We need to take measurements such that different models predict different data

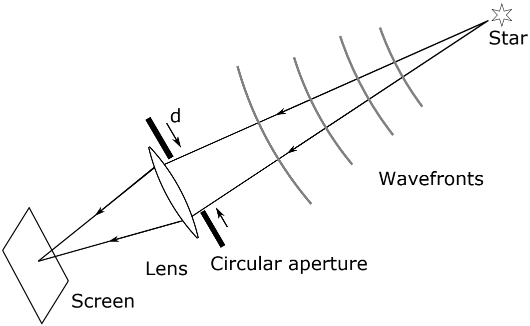
Bayes' theorem formalises this:

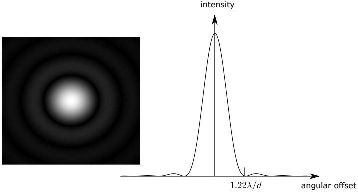
$$P(\text{model}|\text{data}) \propto P(\text{data}|\text{model})P(\text{model})$$

Angular resolution is the sensitivity to structure on small spatial scales

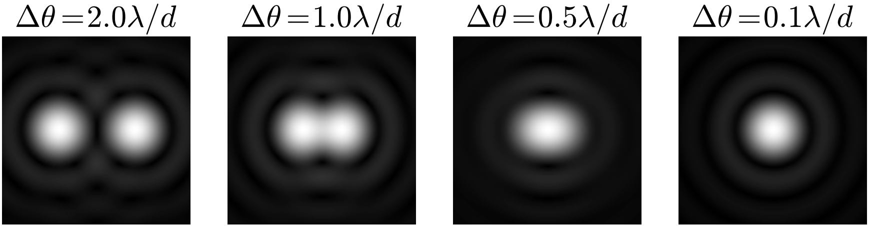


The resolution of a conventional telescope is limited by diffraction





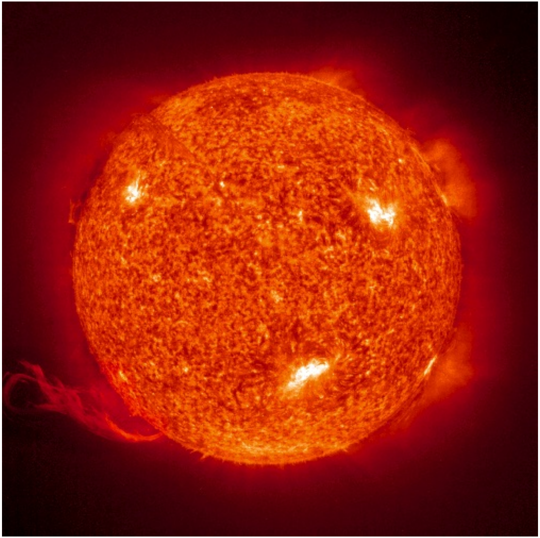
Below a given separation, two stars are indistinguishable from one brighter star



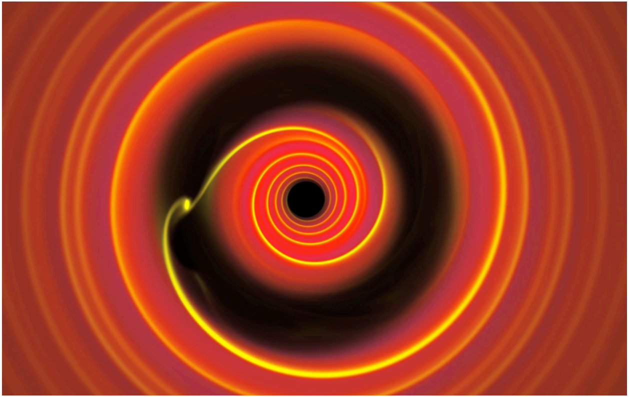
An ELT is not big enough

$d = 39\,\mathrm{m}, \lambda = 2.2\ \mu\mathrm{m} \rightarrow 1.22\lambda/d = 14$ milliarcsec

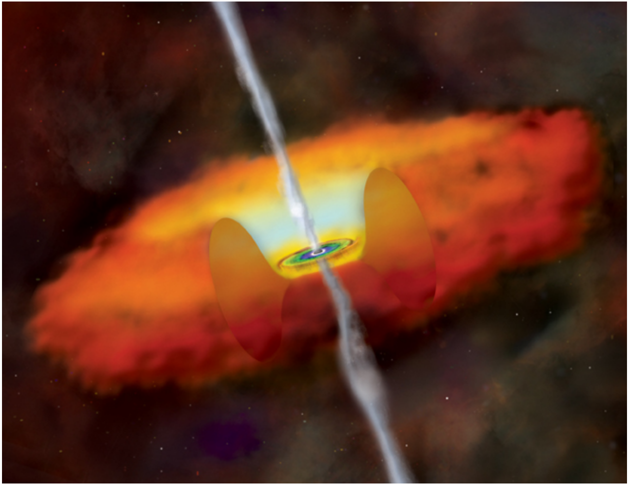
The Sun at 10 pc is 0.9 mas



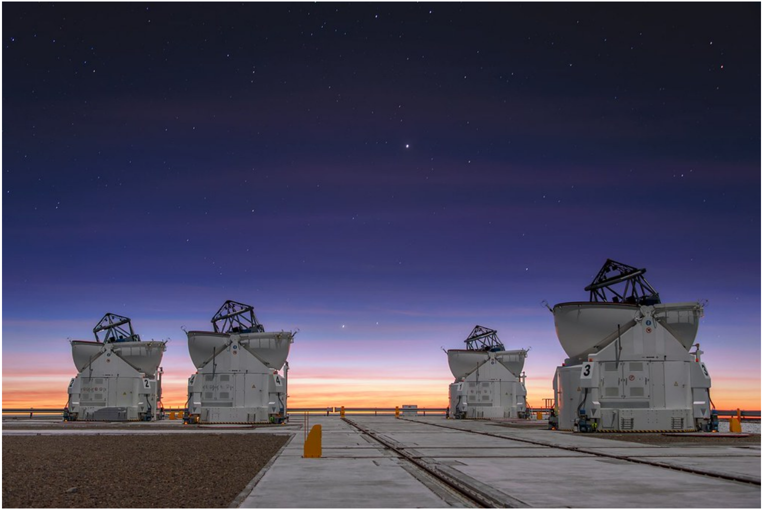
The Earth forming at 150 pc is 6 mas



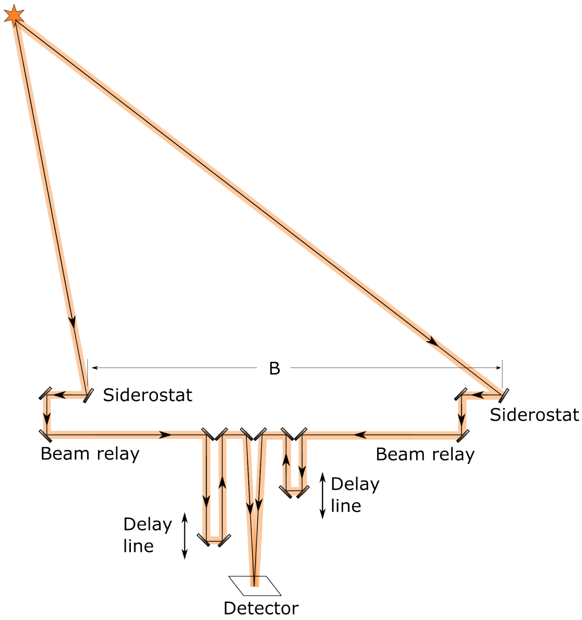
AGN dust tori are a few mas

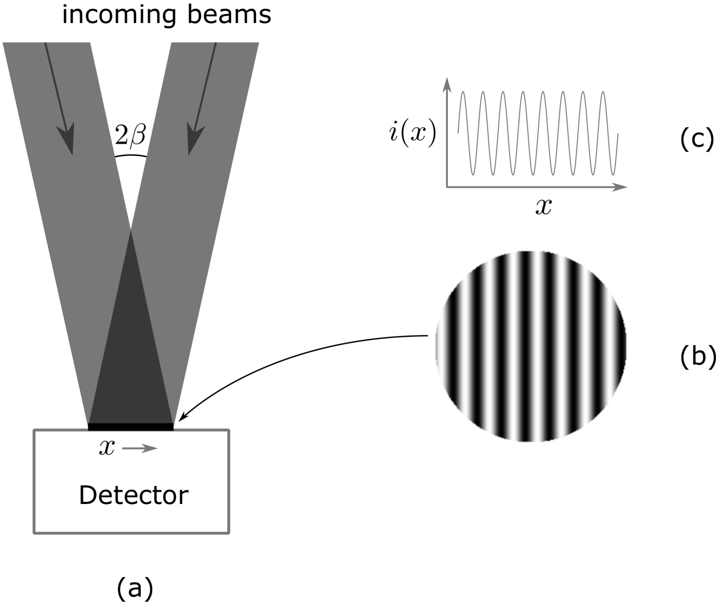


We need measurements which are sensitive to structure on (sub-)mas scales



Making fringes


B
Siderostat
Siderostat
Beam relay
Beam relay
Delay line
Delay line
Detector


incoming beams
2β
$i(x)$
x
(c)
(b)
x
Detector
(a)

We can derive the interferometric measurement equation in three steps

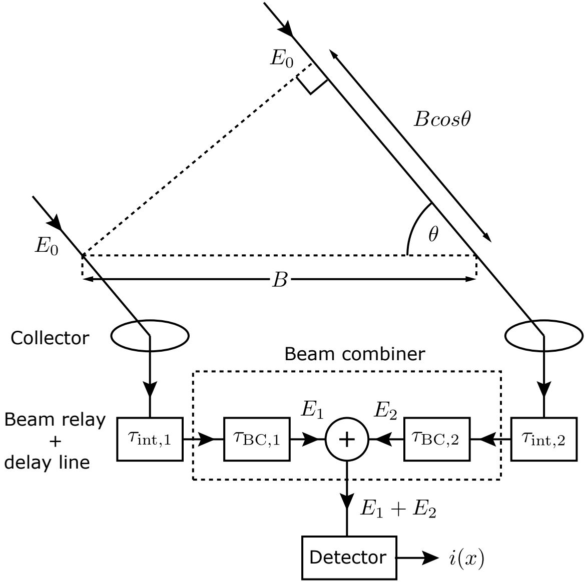
The fringe pattern for 1 star

The fringe pattern for 2 stars

The fringe pattern for an arbitrary object

One star

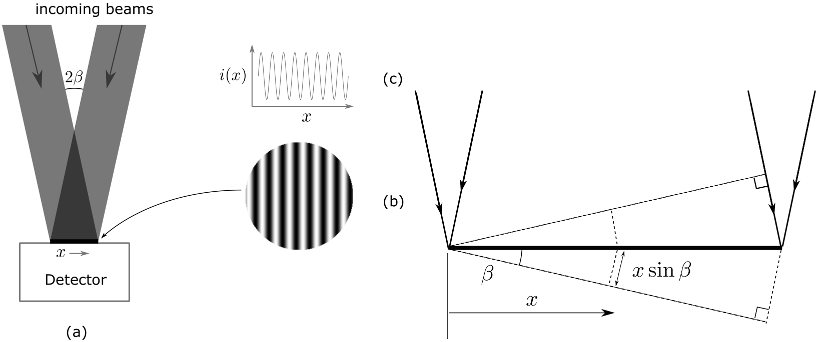
An interferometer adds time-delayed light beams



In the optical we normally talk about Optical Path Difference (OPD)

$$\tau_{\mathrm{ext},2} - \tau_{\mathrm{ext},1} = \mathrm{OPD}_{\mathrm{ext}}/c = B\cos\theta/c$$

A beam combiner allows us to sample a range of time delays



We represent light waves in terms of complex coefficients

$$
\begin{aligned}
E_0 &\propto \mathrm{Re}\left[\Psi_0 e^{-2\pi \mathrm{i}\nu t}\right] \\
E_1(x) &\propto \mathrm{Re}\left[\Psi_1(x) e^{-2\pi \mathrm{i}\nu t}\right] \\
E_2(x) &\propto \mathrm{Re}\left[\Psi_2(x) e^{-2\pi \mathrm{i}\nu t}\right] \\
\Psi_1(x) &= \Psi_0 e^{2\pi \mathrm{i}\nu[\tau_{\mathrm{ext},1}+\tau_{\mathrm{int},1}+\tau_{\mathrm{BC},1}(x)]} \\
\Psi_2(x) &= \Psi_0 e^{2\pi \mathrm{i}\nu[\tau_{\mathrm{ext},2}+\tau_{\mathrm{int},2}+\tau_{\mathrm{BC},2}(x)]}
\end{aligned}
$$

The interference term depends on the phase difference of the beams

$$
\begin{aligned}
i(x) =& \epsilon_0 \left\langle (E_1(x) + E_2(x))^2 \right\rangle \\
&= \left\langle |\Psi_1(x)|^2 + |\Psi_2(x)|^2 \right\rangle \\
&\quad + 2\mathrm{Re}\left[\langle \Psi_1(x) \Psi_2^*(x) \rangle \right]
\end{aligned}
$$

The phase difference can be written in terms of delay differences

$$\Psi_1(x)\Psi_2^*(x) = |\Psi_0|^2 \, e^{2\pi \mathrm{i}\nu[\tau_{12}+\tau_{\mathrm{BC},12}(x)]}$$

where

$$\tau_{12} = (\tau_{\mathrm{ext},1} - \tau_{\mathrm{ext},2}) + (\tau_{\mathrm{int},1} - \tau_{\mathrm{int},2})$$
$$\tau_{\mathrm{BC},12}(x) = \tau_{\mathrm{BC},1}(x) - \tau_{\mathrm{BC},2}(x).$$

The phase of the sinusoidal fringe pattern depends on τ_{12}

$$i(x) = 2F_0 \left(1 + \mathrm{Re}\left[e^{\mathrm{i}\phi_{12}} e^{2\pi \mathrm{i} s x}\right]\right)$$

where

$$\begin{aligned} F_0 &= |\Psi_0|^2 \\ \phi_{12} &= 2\pi\nu\tau_{12} \\ s &= \nu\tau_{BC,12}(x)/x \end{aligned}$$

The phase of the fringes changes if the star moves

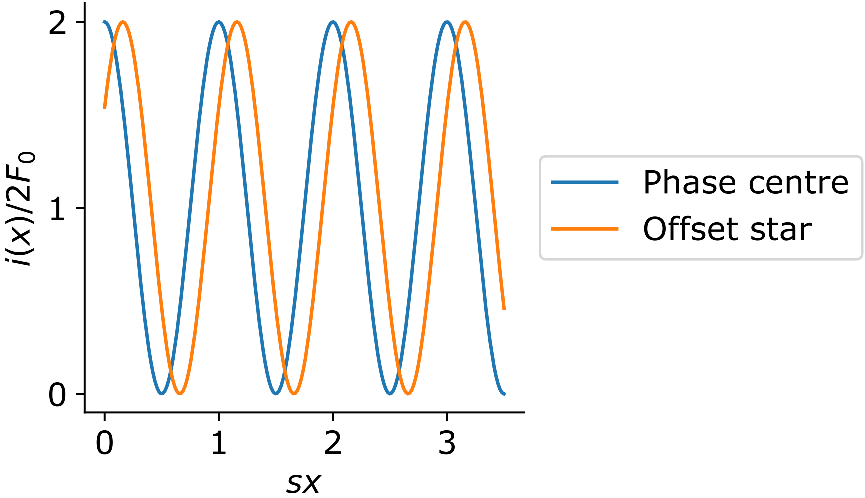
A star at direction θ_0 is at the phase centre if $\tau_{12} = 0$.

For a star offset by $\Delta\theta$ from the phase centre:

$$
\begin{aligned}
c\tau_{12} &= B\cos(\theta_0 + \Delta\theta) - B\cos\theta_0 \\
&\approx -\Delta\theta B\sin\theta_0 \\
\Rightarrow \phi_{12} &= -2\pi u\Delta\theta
\end{aligned}
$$

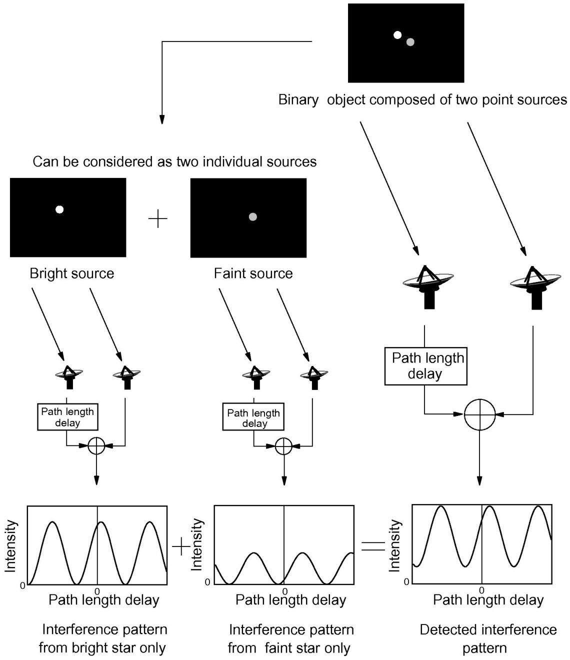
where

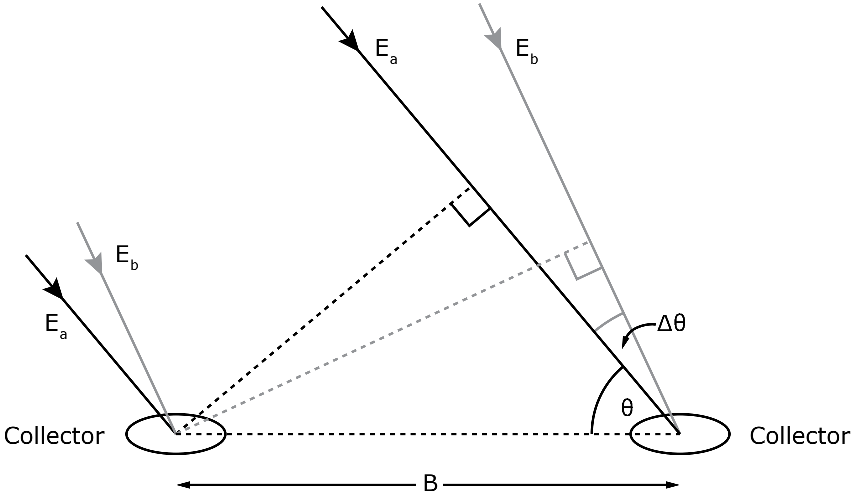
$$u = B\sin\theta_0/\lambda.$$



2
1
0
$i(x)/2F_0$
0
1
2
3
sx
Phase centre
Offset star

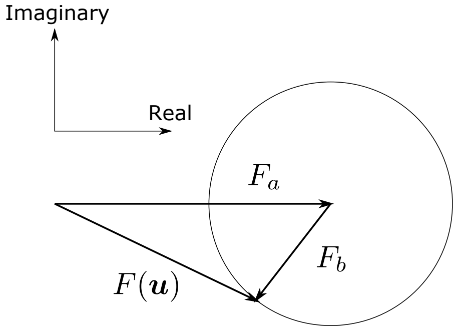
Two stars


Binary object composed of two point sources
Can be considered as two individual sources
Bright source
Faint source
Path length delay
Path length delay
Path length delay
Intensity
0
0
Path length delay
Intensity
0
0
Path length delay
Intensity
0
0
Path length delay
Interference pattern from bright star only
Interference pattern from faint star only
Detected interference pattern


E_a
E_b
E_b
E_a
Δθ
θ
Collector
Collector
B

The fringe pattern depends on the angular offset and not just the intensities

$$
\begin{aligned}
i(x) &= F_a\left(1 + \mathrm{Re}\left[e^{2\pi \mathrm{i} s x}\right]\right) + F_b\left(1 + \mathrm{Re}\left[e^{2\pi \mathrm{i}(\Delta\theta u + s x)}\right]\right) \\
&= F_a + F_b + \mathrm{Re}\left[\left(F_a + F_b e^{2\pi \mathrm{i}\Delta\theta u}\right) e^{2\pi \mathrm{i} s x}\right]
\end{aligned}
$$



We are sensitive to source structure on small angular scales

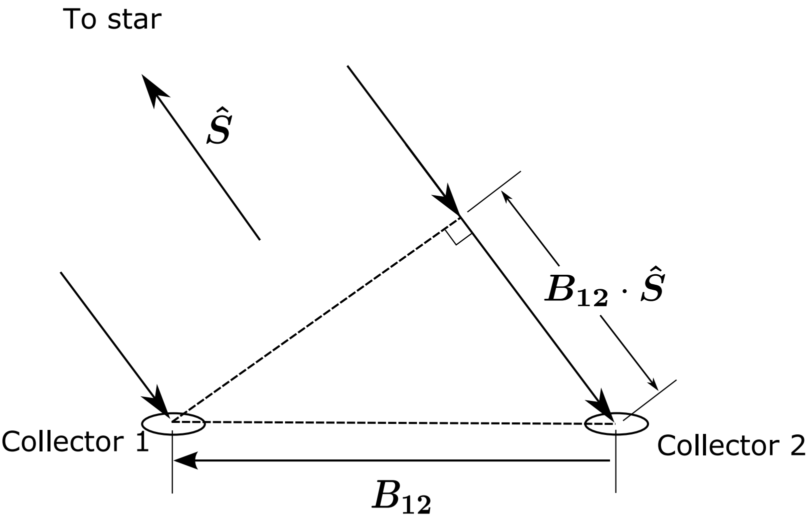
Noticeable effect when $\Delta\theta \gtrsim 1/u \approx \lambda/B$

This is the same as the resolution of a telescope of diameter $\sim B$.

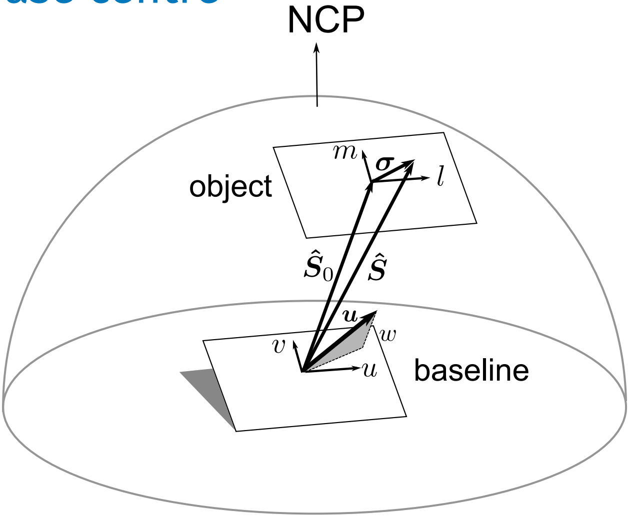
If $B = 100\,\mathrm{m}$ and $\lambda = 500\,\mathrm{nm}$ then $\lambda/B \approx 1\,\mathrm{mas}$

Arbitrary 2-D object

Vector formulation



We use axes based on the tangent plane at the phase centre



For small fields of view, only u and v matter

$$
\begin{aligned}
\tau_{12} &= \mathbf{B}_{12} \cdot \hat{\mathbf{S}}/c - \mathbf{B}_{12} \cdot \hat{\mathbf{S}}_0/c, \\
&= \mathbf{B}_{12} \cdot \boldsymbol{\sigma}/c \quad \text{where } \boldsymbol{\sigma} = \hat{\mathbf{S}} - \hat{\mathbf{S}}_0 \\
\Rightarrow \phi_{12} &= 2\pi \mathbf{u} \cdot \boldsymbol{\sigma} \quad \text{where } \mathbf{u} = \mathbf{B}_{12}/\lambda \\
\mathbf{u} \cdot \boldsymbol{\sigma} &= ul + vm + nw \\
n &\approx \frac{1}{2}(l^2 + m^2) \ll |\boldsymbol{\sigma}| \quad \text{if } l, m \ll 1 \\
\Rightarrow \mathbf{u} \cdot \boldsymbol{\sigma} &\approx ul + vm
\end{aligned}
$$

Hereafter write $\mathbf{u} = (u, v)$ and $\boldsymbol{\sigma} = (l, m)$

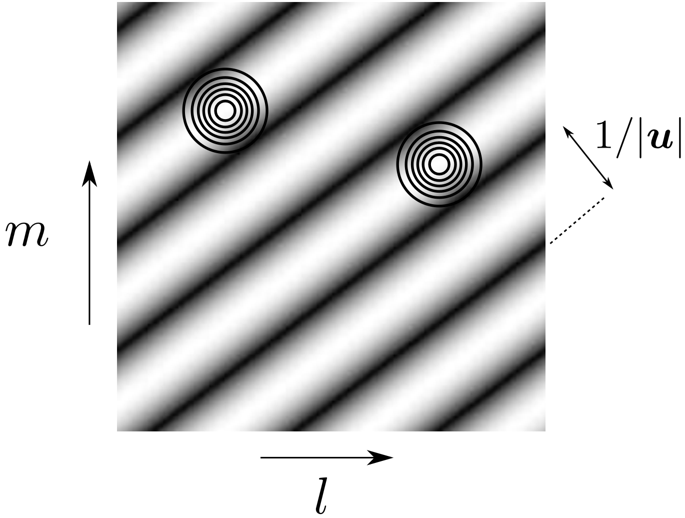
We sum the fringe patterns from all sources

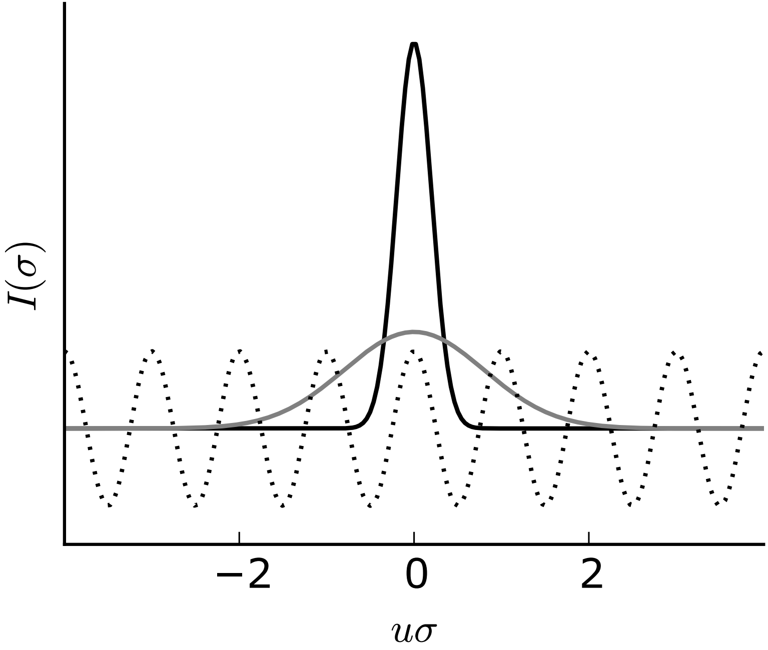
$$
\begin{aligned}
i(x) &= \iint_{-\infty}^{\infty} I(\boldsymbol{\sigma}) \left(1 + \mathrm{Re}\left\{e^{-2\pi \mathrm{i} \boldsymbol{\sigma}\cdot\boldsymbol{u}} e^{2\pi \mathrm{i} s x}\right\}\right) \, dl \, dm \\
&= F(0) + \mathrm{Re}\left[F(\boldsymbol{u}) e^{2\pi \mathrm{i} s x}\right]
\end{aligned}
$$

where

$$
\begin{aligned}
F(\boldsymbol{u}) &= \iint_{-\infty}^{\infty} I(\boldsymbol{\sigma}) e^{-2\pi \mathrm{i} \boldsymbol{\sigma}\cdot\boldsymbol{u}} \, dl \, dm \\
\Rightarrow F(0) &= \iint_{-\infty}^{\infty} I(\boldsymbol{\sigma}) \, dl \, dm
\end{aligned}
$$

$F(u)$ weights the object $I(\sigma)$ with a sinusoidal "mask" and sums the result





$I(\sigma)$
−2
0
2
$u\sigma$

Nomenclature

$F(\boldsymbol{u})$: coherent flux, correlated flux

$V(\boldsymbol{u}) \equiv F(\boldsymbol{u})/F(0)$: object (complex) visibility

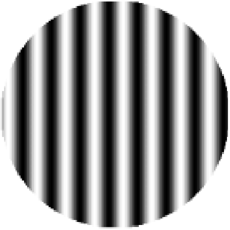
V_{12}: fringe (complex) visibility

$|V(\boldsymbol{u})|$: object visibility (modulus)

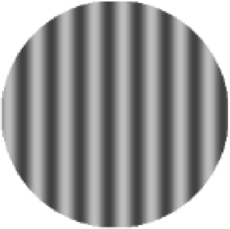
$|V_{12}|$: fringe contrast, fringe visibility (modulus)

Michelson fringe visibility

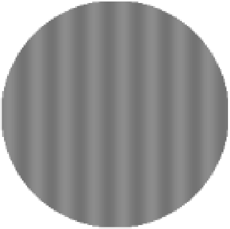
$|V| = 1.0$



$|V| = 0.5$



$|V| = 0.1$



Coherence and fringes

Mutual intensity of two beams Ψ_1 and Ψ_2:

$$M_{12} = \langle \Psi_1 \Psi_2^* \rangle$$

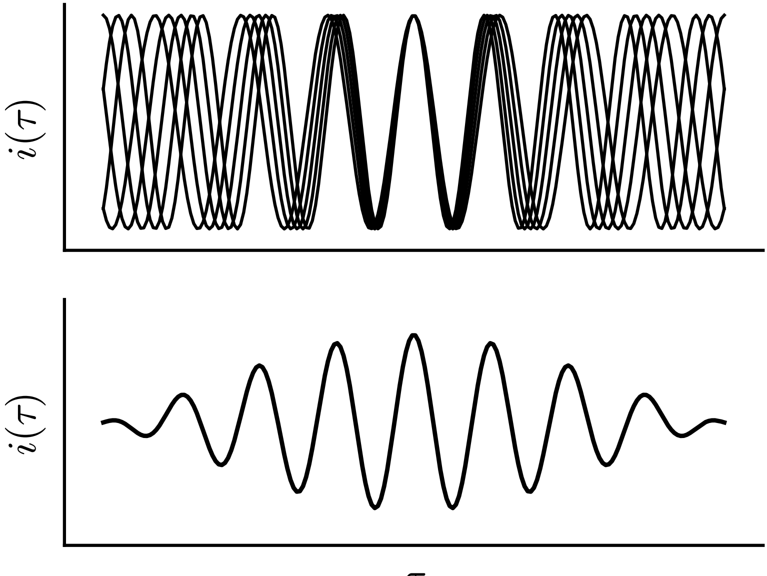
Degree of coherence:

$$C_{12} = \frac{M_{12}}{\sqrt{M_{11} M_{22}}}.$$

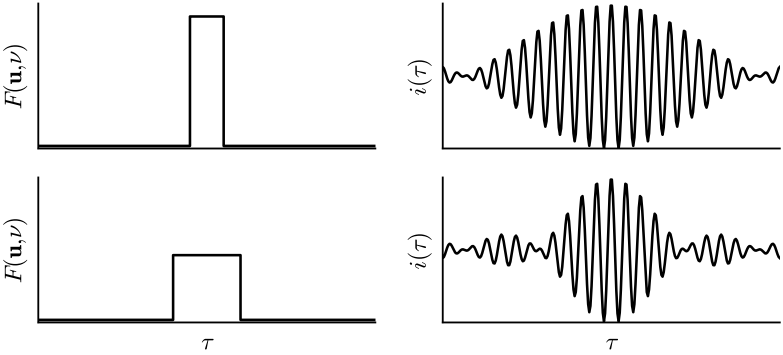
Interference measures cross-correlation:

$$i(x) = \left\langle |\Psi_1(x)|^2 + |\Psi_2(x)|^2 \right\rangle + 2\mathrm{Re}\left[\langle \Psi_1(x) \Psi_2^*(x) \rangle\right]$$

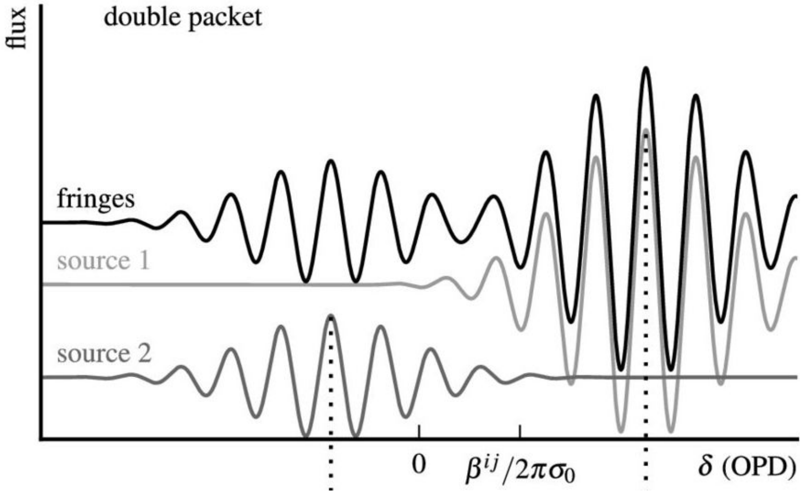
Polychromatic interferometry



The coherence length varies inversely with the bandwidth



A wide bandwidth can limit the field of view



From fringes to images

Fourier transforms

Can compose any function f from sine waves:

$$f(\mathbf{x}) = \iint_{-\infty}^{\infty} g(\mathbf{s}) e^{2\pi i \mathbf{s}\cdot\mathbf{x}} \, ds_x \, ds_y$$

where coefficients are

$$g(\mathbf{s}) = \iint_{-\infty}^{\infty} f(\mathbf{x}) e^{-2\pi i \mathbf{s}\cdot\mathbf{x}} \, dx \, dy.$$

Fourier transform $g(\mathbf{s}) = \mathcal{F}[f(\mathbf{x})]$

Inverse Fourier transform: $f(\mathbf{x}) = \mathcal{F}^{-1}[g(\mathbf{s})]$

A fringe measurement measures a single Fourier component of the image

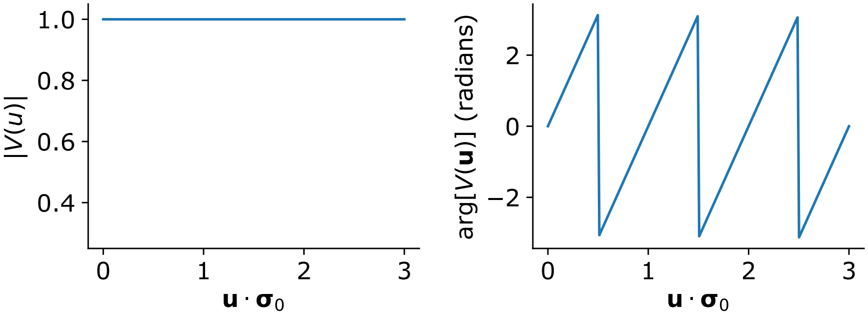
$$F(\boldsymbol{u}) = \iint_{-\infty}^{\infty} I(\boldsymbol{\sigma}) e^{-2\pi i \boldsymbol{\sigma} \cdot \boldsymbol{u}} \, dl \, dm$$

is equivalent to $F(\boldsymbol{u}) = \mathcal{F}\left[I(\boldsymbol{\sigma})\right]$

If we measure $F(\boldsymbol{u}) \forall \boldsymbol{u}$ we can in principle invert this.

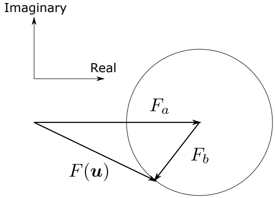
Better to first develop your Bayesian intuition: what models give what data?

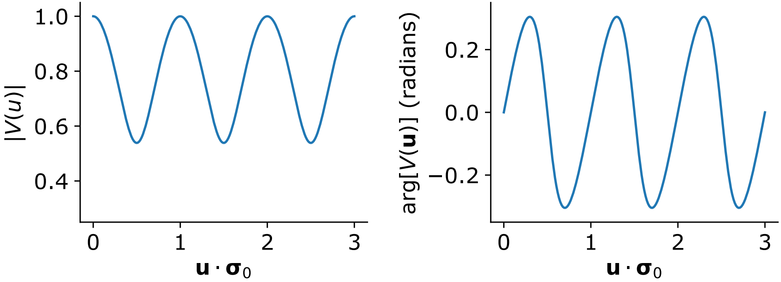
Point source



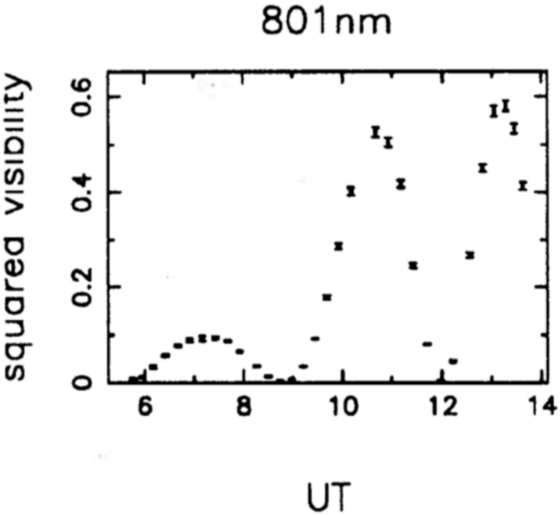
$$
\begin{aligned}
I(\boldsymbol{\sigma}) &\propto \delta(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0) \\
\Rightarrow F(\boldsymbol{u}) &\propto e^{2\pi \mathrm{i} \boldsymbol{u}\cdot\boldsymbol{\sigma}_0}
\end{aligned}
$$

Binary star system

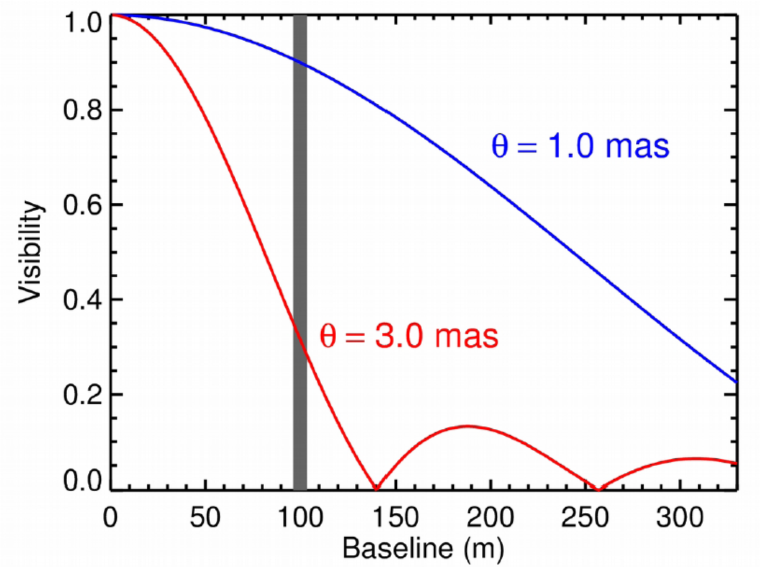




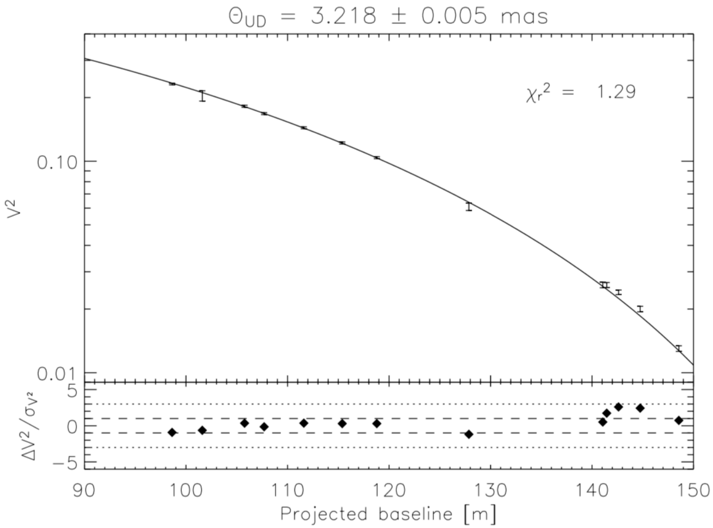
The binary star Capella (Hummel+ 1994)



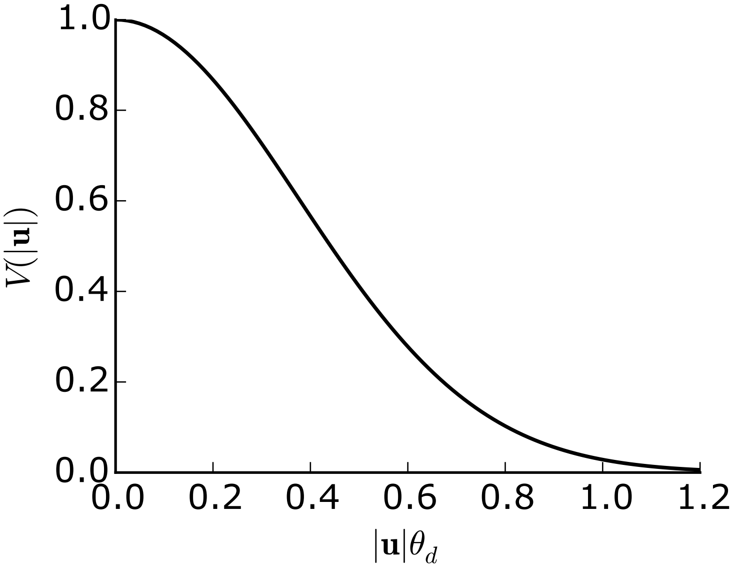
Uniform disc



Vega (Absil+ 2006)



Gaussian disc



Objects offset from the phase centre

Convolution theorem: $\mathcal{F}\{f * g\} = FG$ where $F(\mathbf{s}) = \mathcal{F}\{f(\mathbf{x})\}$ and $G(\mathbf{s}) = \mathcal{F}\{g(\mathbf{x})\}$, where

$$(f_1 * f_2)(\mathbf{x}) \equiv \iint_{\text{All space}} f_1(\mathbf{x}')f_2(\mathbf{x} - \mathbf{x}')\, dA$$

$$\begin{aligned} f(\mathbf{x} - \mathbf{x}_0) &= f(\mathbf{x}) * \delta(\mathbf{x} - \mathbf{x}_0) \\ \Rightarrow \mathcal{F}\{f(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0)\} &= F e^{2\pi i \mathbf{u}\cdot\boldsymbol{\sigma}_0} \end{aligned}$$

Rules of thumb

$V(0) = 1$ and $|V(\boldsymbol{u})| \leq 1$

Significant deviations from $|V| = 1$ ("resolved") when $B \gtrsim \lambda/\theta$

Sharp-edged structures show "ringing" sidelobes

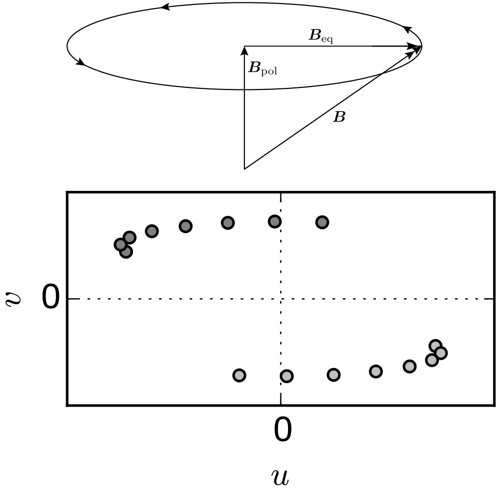
Symmetric objects have real visibility functions

For all objects $V(-\boldsymbol{u}) = V^*(\boldsymbol{u})$ (Hermitian symmetry)

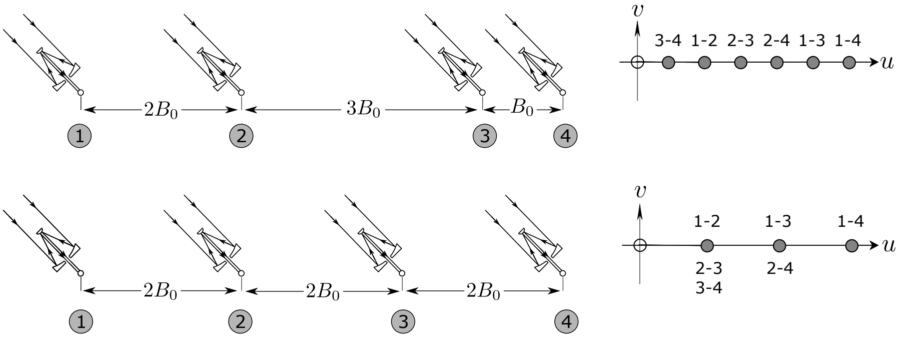
How to sample the Fourier plane

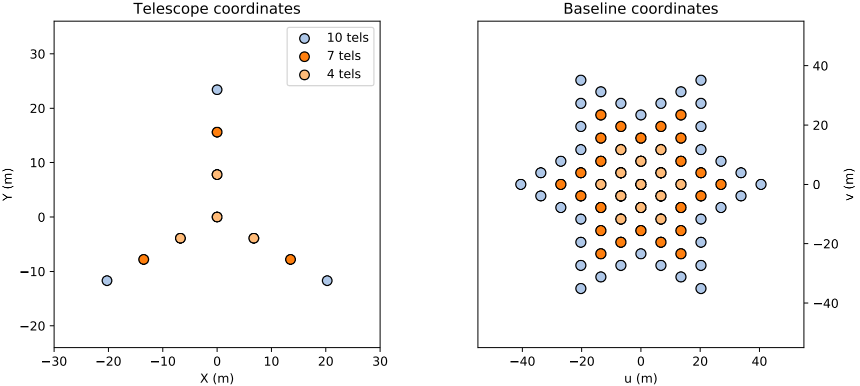
We can change the baseline by moving the telescopes

Earth rotation is carbon-neutral telescope transportation



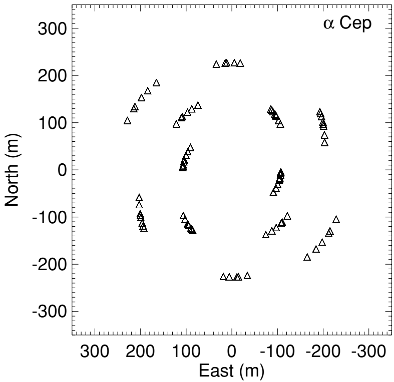
Multiple telescopes sample many (u, v) points

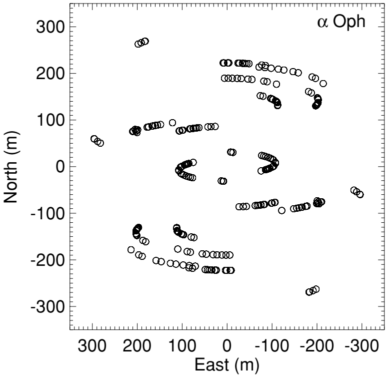




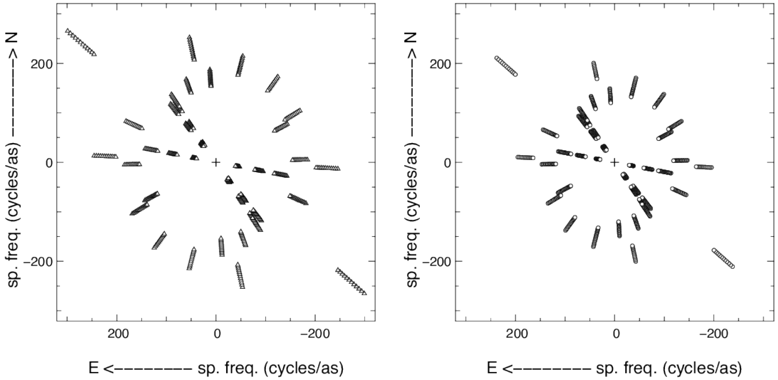
$$N_{\mathrm{bas}} = N_{\mathrm{tel}}(N_{\mathrm{tel}} - 1)/2$$

We typically use multiple telescopes in tandem with Earth rotation





Observing in multiple spectral channels gives additional coverage



How does finite sampling affect the image?

Fourier inversion leads to a "dirty image"

The data sampled at locations $\{\boldsymbol{u}_k\}$ is

$$\hat{F}(\boldsymbol{u}) = F(\boldsymbol{u}) \sum_k \delta(\boldsymbol{u} - \boldsymbol{u}_k)$$

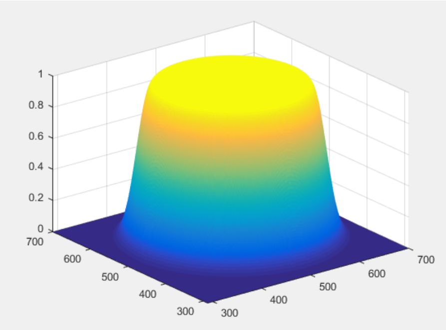
The "synthesised image"/"dirty image" is

$$\begin{aligned} \hat{I}(\sigma) &= \mathcal{F}^{-1}\left[\hat{F}(\boldsymbol{u})\right] \\ &= I(\boldsymbol{\sigma}) * b(\boldsymbol{\sigma}) \end{aligned}$$

where $b(\boldsymbol{\sigma})$ is the "dirty beam"

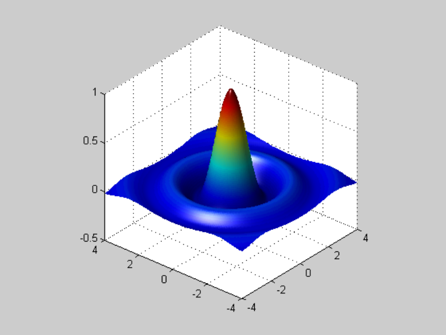
$$b(\boldsymbol{\sigma}) = \mathcal{F}^{-1}\left[\sum_k \delta(\boldsymbol{u} - \boldsymbol{u}_k)\right].$$

Angular resolution depends on the maximum baseline



(u, v) coverage

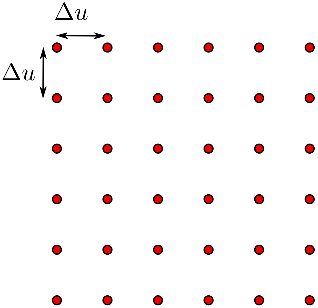
$|u| < u_{\mathrm{max}}$



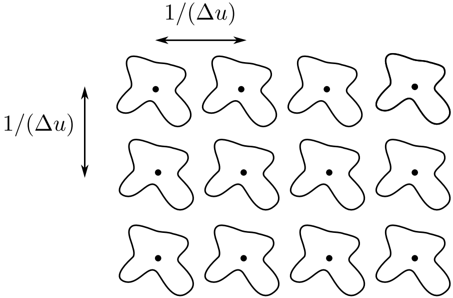
dirty beam

FWHM $\sim 1/u_{\mathrm{max}}$

Field of view depends on the density of sampling

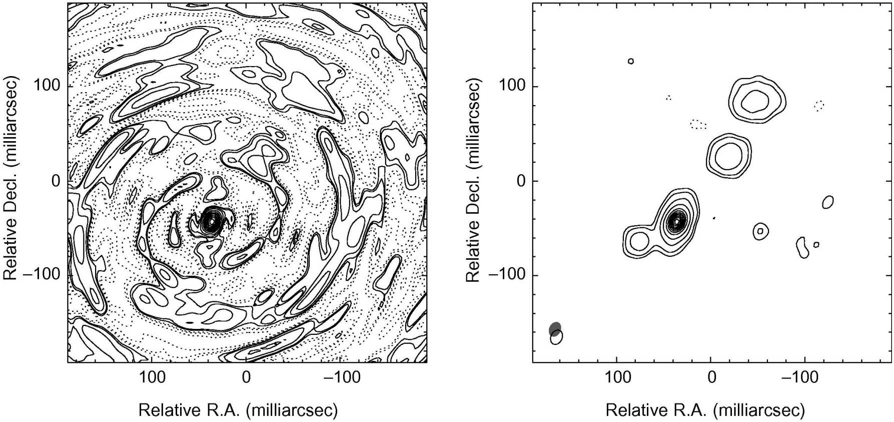


(u, v) coverage



dirty image

Deconvolution can ameliorate imperfections in the sampling



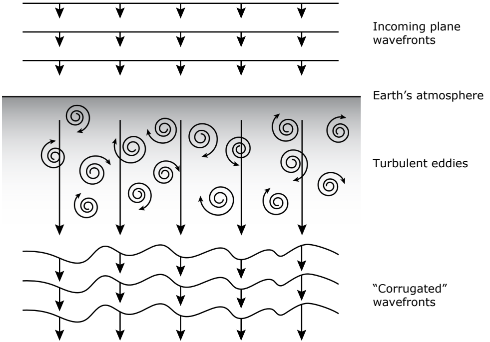
Must still meet minimum sampling criteria to get an adequate image

Longest baselines $u_{\mathrm{max}} \gtrsim 1/\theta_{\mathrm{min}}$

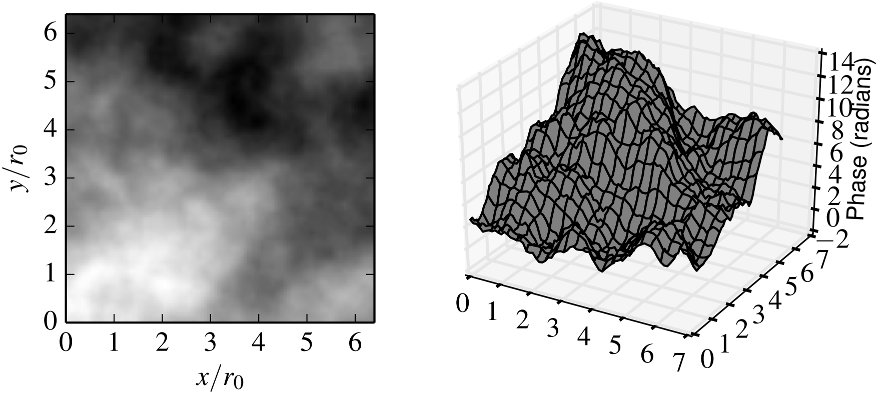
Largest sampling "holes" $\Delta u \lesssim 1/\theta_{\mathrm{max}}$

Atmospheric seeing

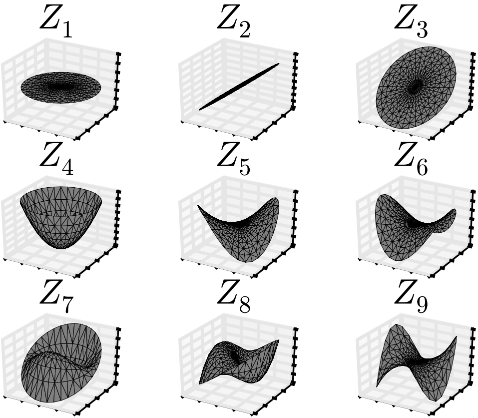
Turbulent mixing of *existing* gradients of refractive index causes random wavefront perturbations



The spatial structure is fractal



The corrugations are dominated by low-order spatial modes

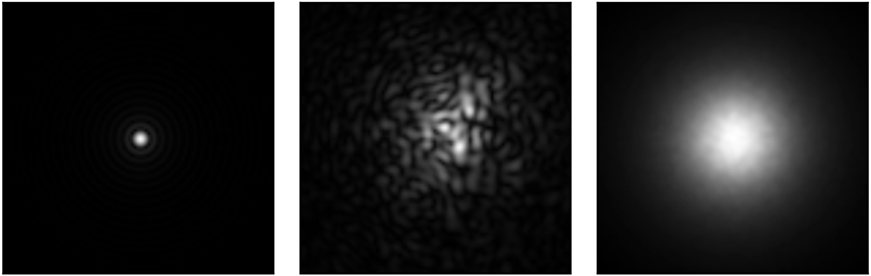


We can define a characteristic spatial scale

$$D_\phi(\mathbf{r}, \mathbf{r}') \equiv \left\langle \left|\phi(\mathbf{r}' + \mathbf{r}) - \phi(\mathbf{r}')\right|^2 \right\rangle$$

$$D_\phi(\mathbf{r}) = 6.88(r/r_0)^{5/3}$$

Images are blurred on scales $\sim \lambda / r_0$



$$\lambda / r_0 \approx 1\,\mathrm{arcsec} \text{ for } \lambda = 500\,\mathrm{nm},\ r_0 = 10\,\mathrm{cm}$$

Temporal seeing can be modelled from assuming “frozen turbulence”

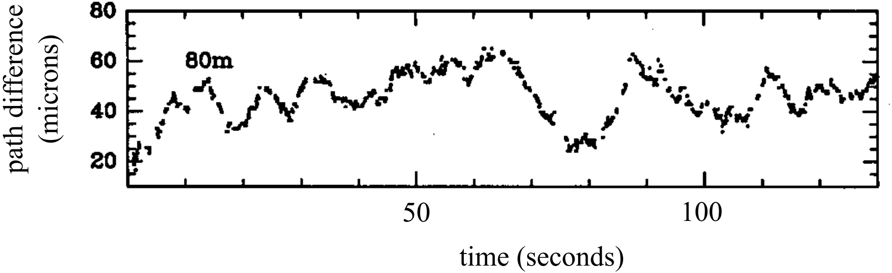
$$D_\phi(t) \equiv \left\langle \left|\phi(\mathbf{r}, t' + t) - \phi(\mathbf{r}, t')\right|^2 \right\rangle = (t/t_0)^{5/3}$$

$$t_0 = 0.314 r_0/v \sim \text{milliseconds}$$

`https://share.streamlit.io/dbuscher/megascreen/tests/demos/streamlit_movie.py`

First-order effects on interferometers

The differential piston between telescopes is hundreds of wavelengths

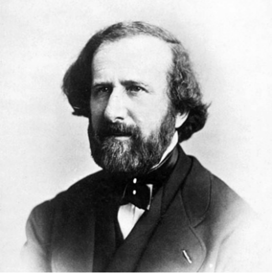


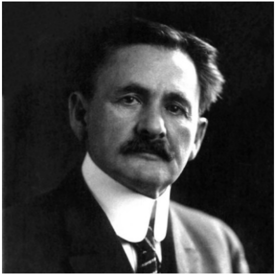
Visibility phase is “meaningless”

Visibility modulus is the only good observable

We need the phase to make images

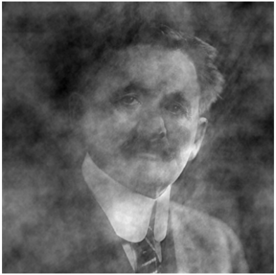
$i_1(x, y)$





$i_2(x, y)$

$i_{12}(x, y)$



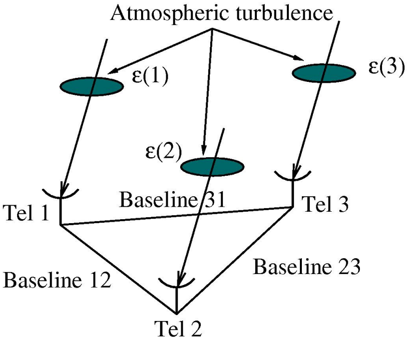
$i_{21}(x, y)$

Closure phase

$\frac{1}{2}N(N-1)$ object visibility phases ϕ_{ij}

$N-1$ unknown phase perturbations ϵ_j

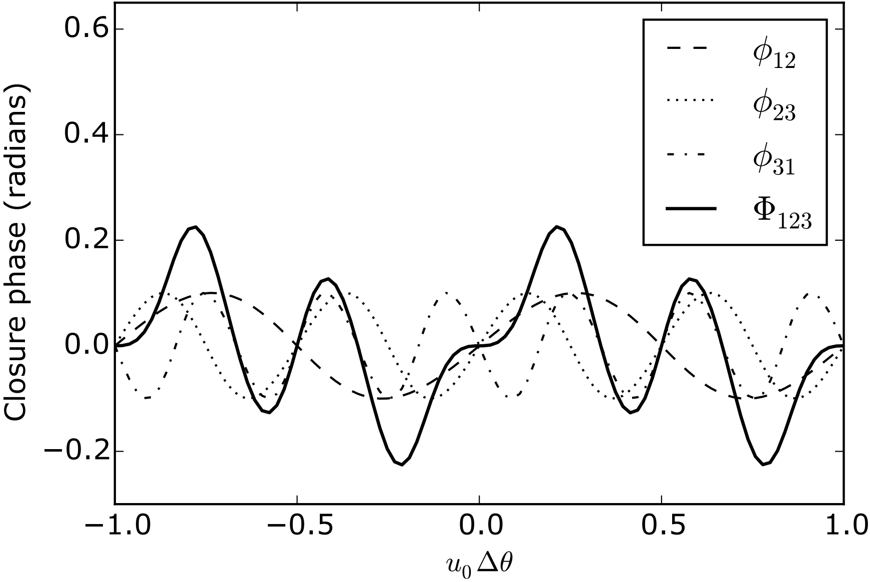
Can solve for $\frac{1}{2}(N-1)(N-2)$ perturbation
-independent terms - “closure phases”



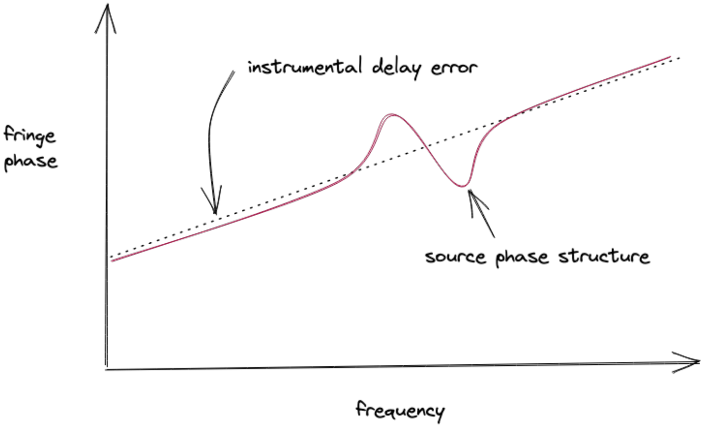
$$\Phi(12) = \phi(12) + \varepsilon(1) - \varepsilon(2)$$
$$\Phi(23) = \phi(23) + \varepsilon(2) - \varepsilon(3)$$
$$\Phi(31) = \phi(31) + \varepsilon(3) - \varepsilon(1)$$

$$\Phi(12) + \Phi(23) + \Phi(31) = \phi(12) + \phi(23) + \phi(31)$$

Binary star example

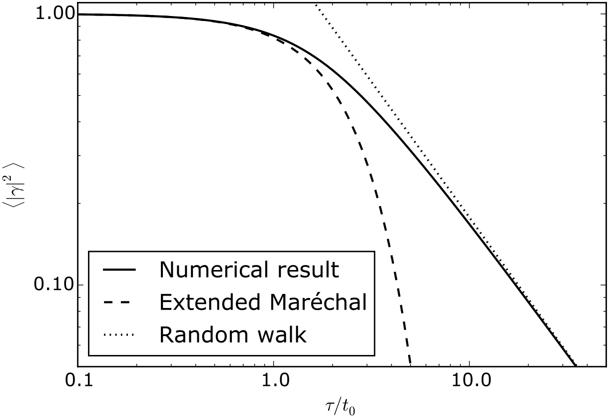


(Spectral) differential phase can retrieve wavelength-dependent structure

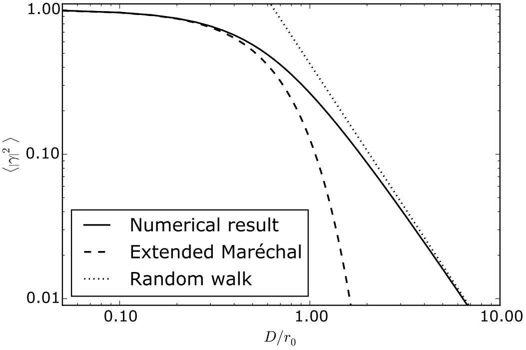


Higher-order effects

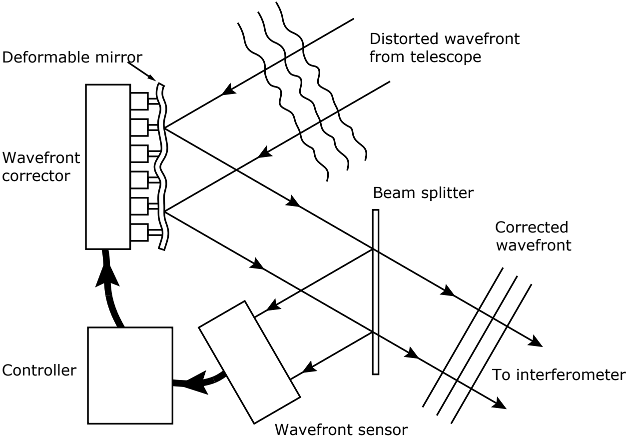
Fringes smear out if the exposure time is too long



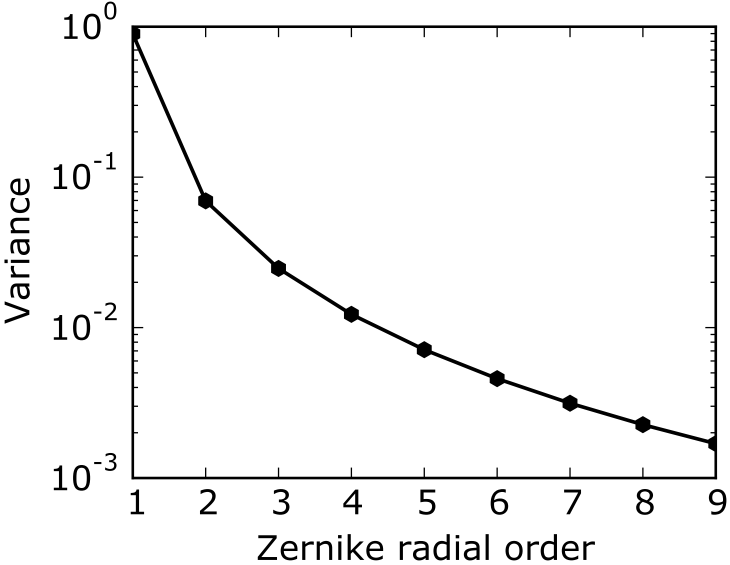
The fringes distort if the aperture is too large



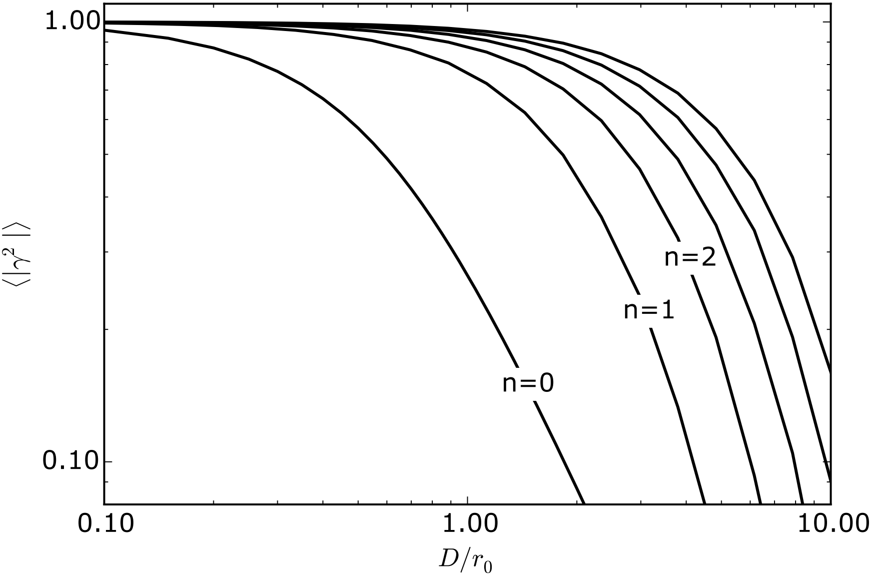
Adaptive optics can correct the spatial fluctuations



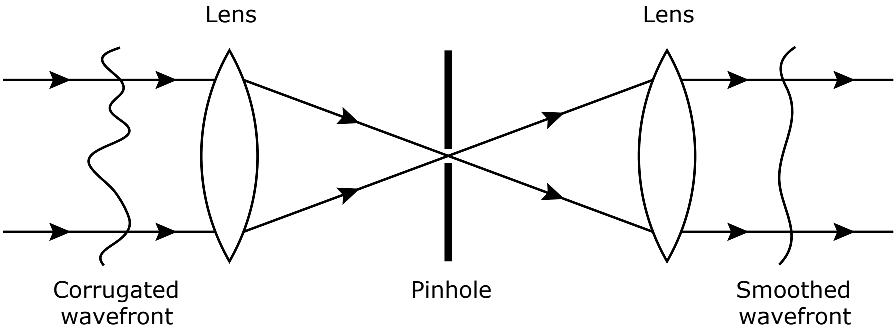
Tip/tilt correction is 90% of the battle



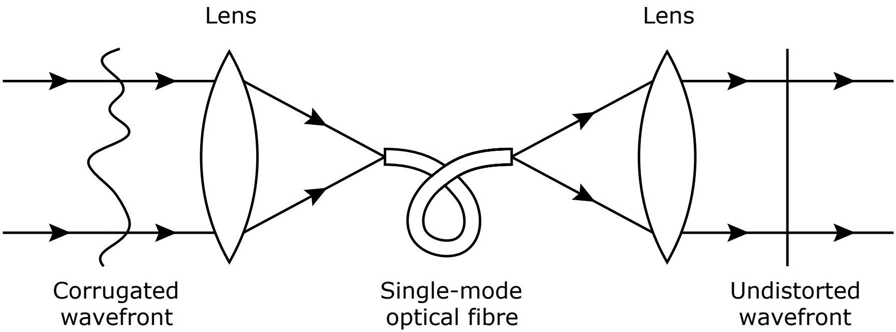
AO allows you to use larger apertures



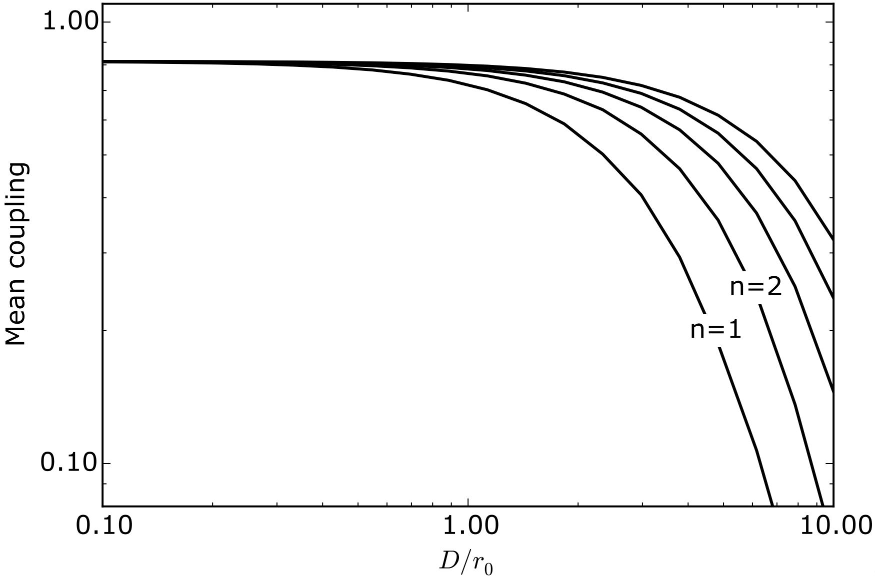
Spatial filtering is “passive AO”



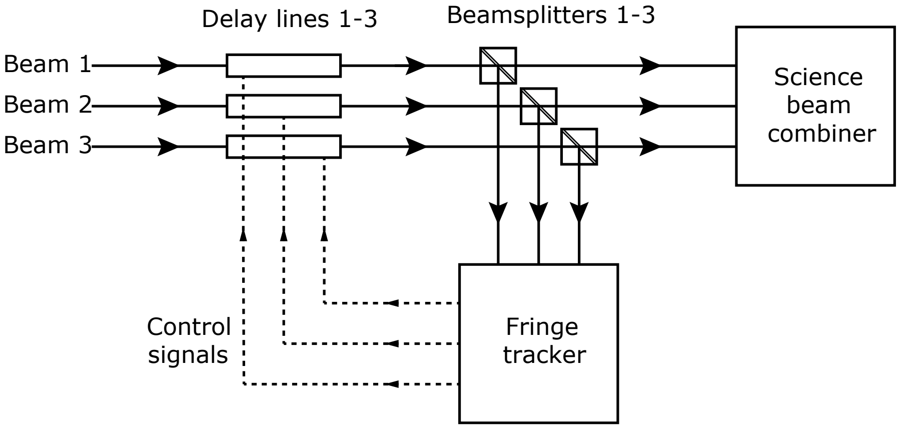
Single-mode fibres are perfect spatial filters



AO increases light coupling



Fringe tracking is “piston AO”



Cophasing and coherencing

Ideally, we want to track the piston to $\ll \lambda$ to "freeze" the fringe phase and have long integrations (cophasing or coherent integration).

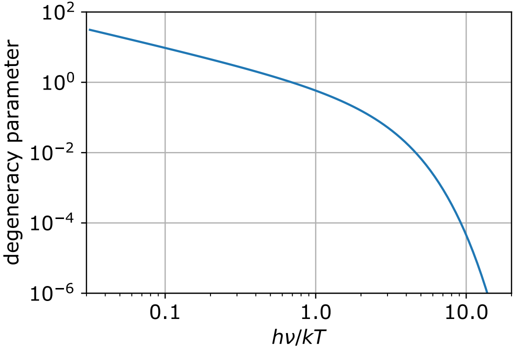
At low light levels, the fringe tracker fails

Can still use a "group delay" tracker for "coherencing" — keep the fringe "envelope" centred.

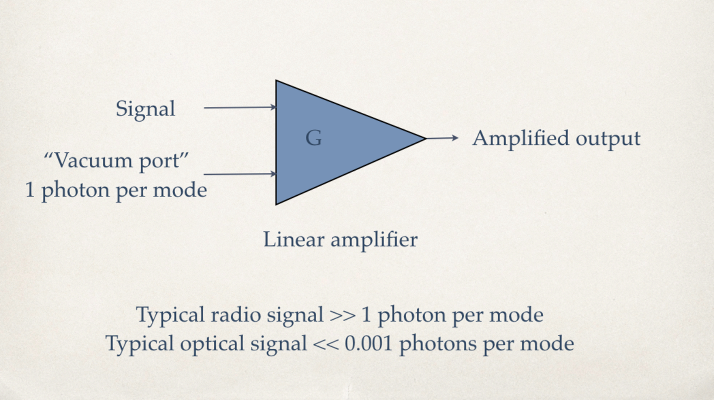
Measurement noise and data reduction

The degeneracy parameter of optical radiation sources is tiny

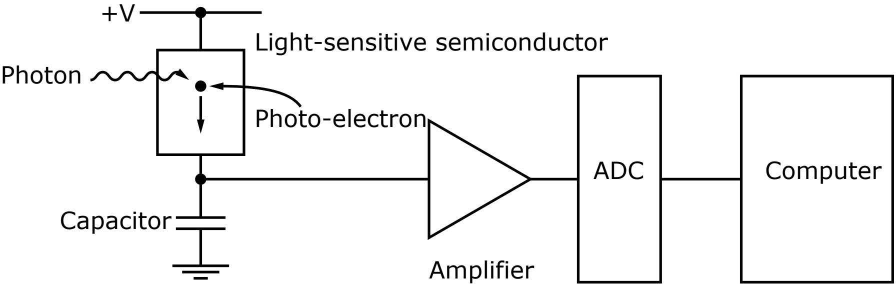
Radio $h\nu/kT \lesssim 0.5$
Optical $h\nu/kT \gtrsim 5$



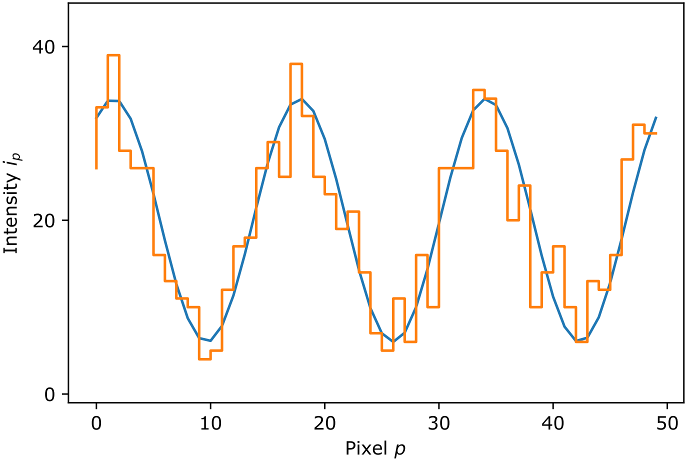
Coherent optical amplifiers are useless so we can't split light between many baselines



We need to detect un-amplified fringe intensities



The detected fringe pattern is discrete and noisy



We can model the detected interferogram in terms of the noiseless fringe pattern Λ

$$\Lambda_p = \frac{1}{N_{\mathrm{pix}}} \left(\overline{N}_{\mathrm{phot}} + \mathrm{Re}\left\{ F_{ij} e^{2\pi \mathrm{i} s_{ij} p} \right\} \right)$$

$$i_p = \Lambda_p + n_p$$

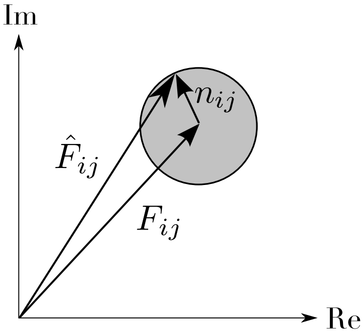
The noise on the fringe parameters is the sum of the noise on each pixel

Using a Discrete Fourier Transform (DFT) to extract F_{ij}

$$\hat{F}_{ij} = 2 \sum_{p=0}^{N_{\mathrm{pix}}-1} i_p e^{-2\pi \mathrm{i} s_{ij} p}$$

$$\hat{F}_{ij} = F_{ij} + n_{ij}$$

$$n_{ij} = 2 \sum_{p=0}^{N_{\mathrm{pix}}-1} n_p e^{-2\pi \mathrm{i} s_{ij} p}$$



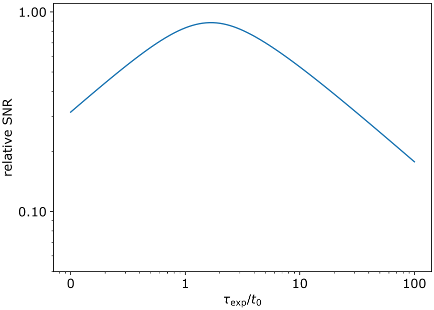
Low fringe visibility is more of a problem than low flux

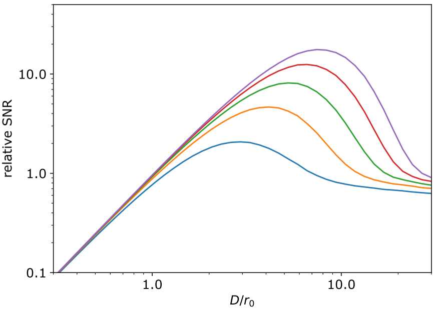
$$\mathrm{SNR}(\hat{F}_{ij}) = \frac{\left|\left\langle \hat{F}_{ij} \right\rangle\right|}{\sigma_{ij}}$$

In the photon-noise-dominated regime:

$$\mathrm{SNR}(\hat{F}_{ij}) \approx \frac{1}{2}|V_{ij}|\sqrt{\overline{N}_{\mathrm{phot}}},$$

There is a maximum SNR we can reach in a single exposure





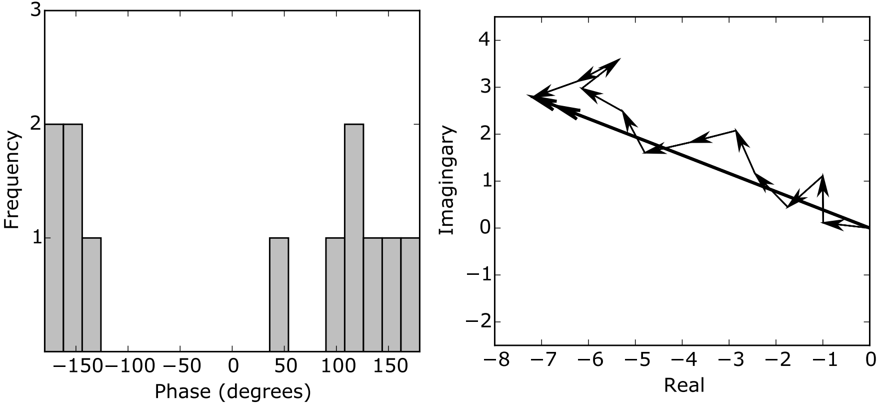
We need to average the results from many exposures

Can have many 1000s of interferograms in a few-minute observation

If we have a reliable source of phase data, can do “coherent integration”, effectively increasing the exposure time

Then do “incoherent integration” — average power spectrum ($|F|^2$) and bispectrum.

Don’t average the closure phase!

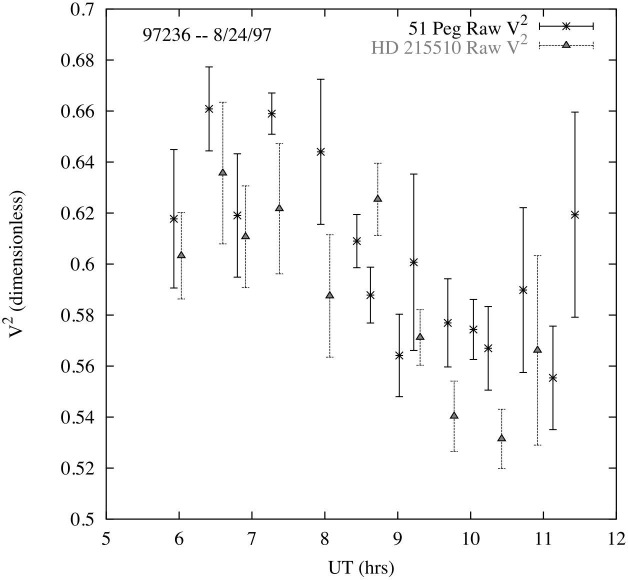


The bispectrum (triple product)

$$T_{pqr} = T(\boldsymbol{u}_{pq}, \boldsymbol{u}_{qr}) = F(\boldsymbol{u}_{pq})F(\boldsymbol{u}_{qr})F(-\boldsymbol{u}_{pq} - \boldsymbol{u}_{qr})$$

$$\arg(T_{pqr}) = \phi_{pq} + \phi_{qr} + \phi_{rp}$$

The fringe visibility depends on the seeing



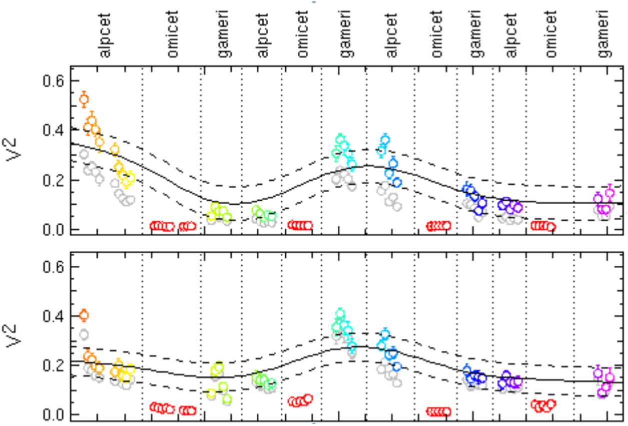
We use measurements of stars with known $|V(\boldsymbol{u}_{ij})|$ to calibrate the transfer function

$$\left\langle \left| \hat{V}_{ij} \right|^2 \right\rangle = \left\langle \left| \gamma_{ij} \right|^2 \right\rangle |V(\boldsymbol{u}_{ij})|^2$$

Assume that the transfer function $\left\langle \left| \gamma_{ij} \right|^2 \right\rangle$ is stable between observations of the target and calibrator stars.

At mid-IR wavelengths, we can calibrate the coherent flux F_{ij} rather than the visibility V_{ij}

It is best to bracket the target with calibrators



Interferometers in practice

Interferometric facilities

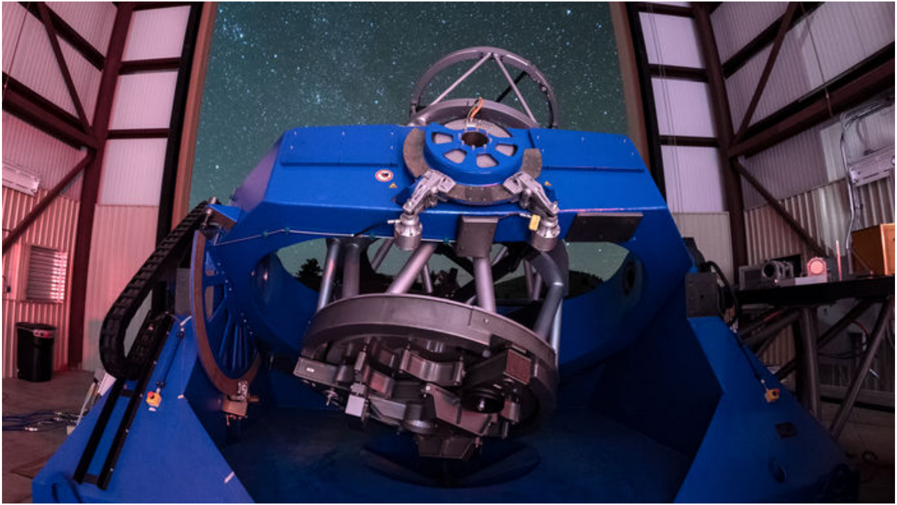
NPOI: 6×12 cm collectors, max baseline 450 m

CHARA: 6×1 m telescopes, max baseline 330 m

VLTI: 4×8 m UTs + 4×1.8 m ATs, max baseline 200 m

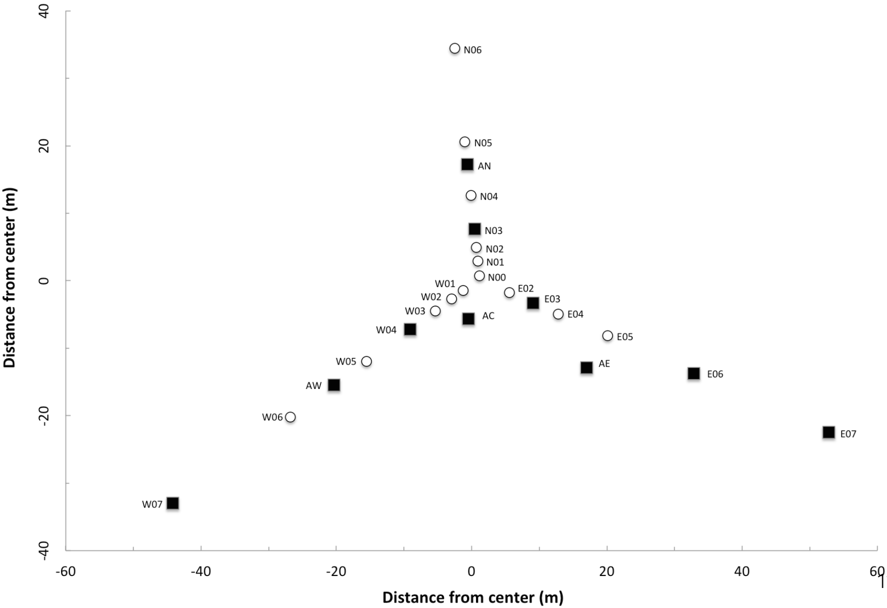
MROI: 10×1.4 m telescopes, max baseline 350 m

Collectors

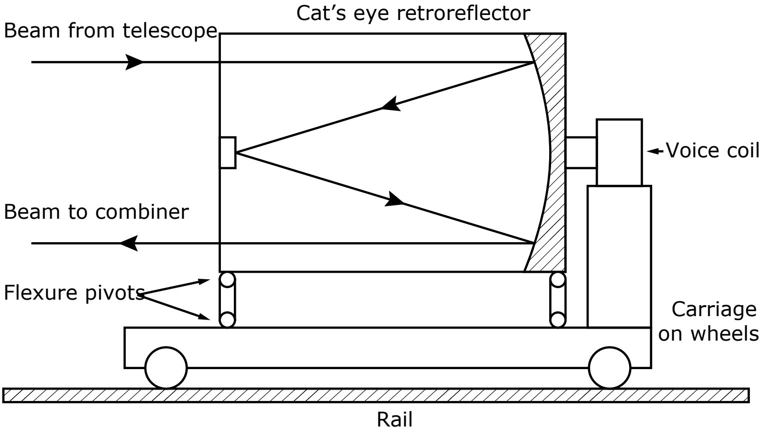


Beam relay

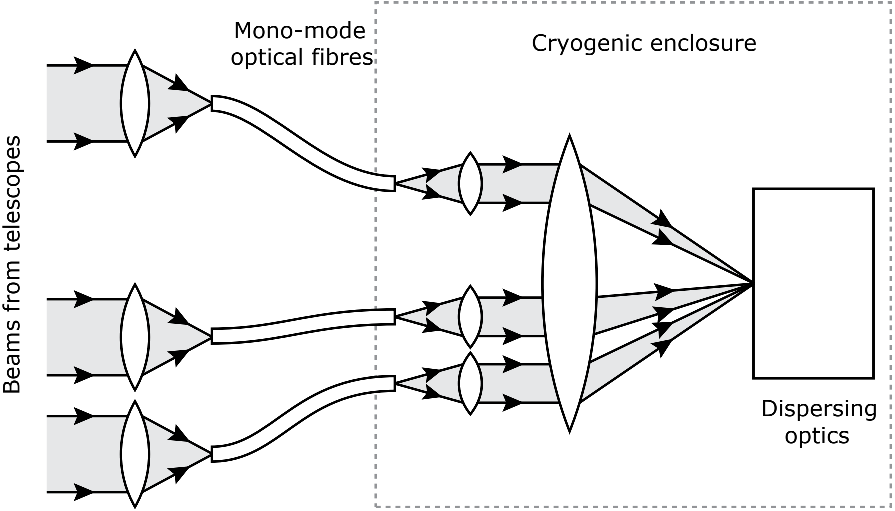
Array layout



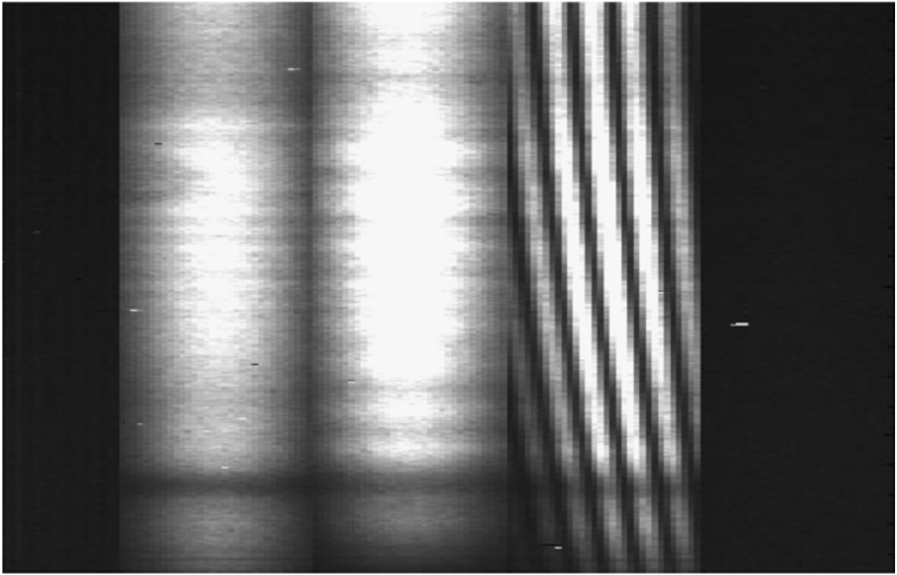
Delay lines


Cat’s eye retroreflector
Beam from telescope
Voice coil
Beam to combiner
Flexure pivots
Carriage
on wheels
Rail

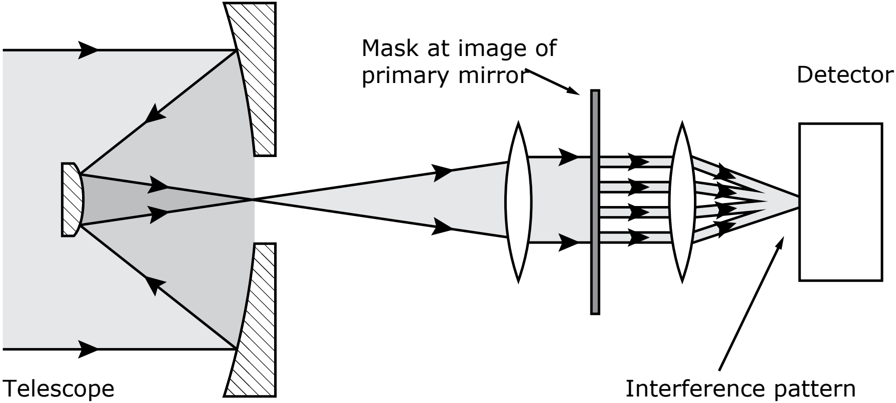
Beam combiners



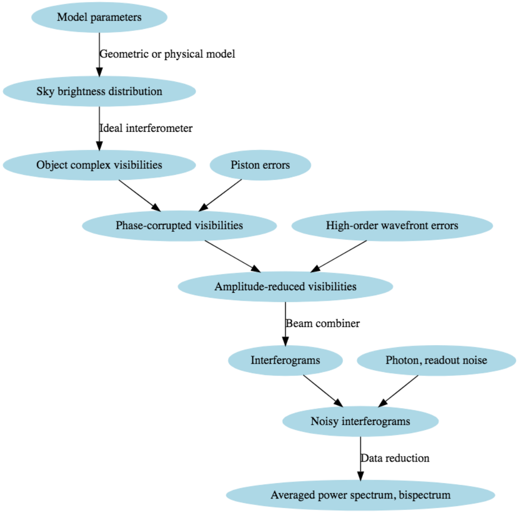
Dispersed fringes



Aperture masking



We have built a “forward model” of an interferometric measurement



Now we just need to solve our inverse problem

1. Model our targets (YSOs etc)
2. Model our observations (ASPRO)
3. Observation and data reduction
4. Model-fitting & image reconstruction