

AN INTRODUCTION TO INTERFEROMETRY

Chris Haniff¹

Abstract. In this paper I provide a brief introduction to astronomical interferometry at optical and infrared wavelengths. Two key concepts, central to understanding the basis and practice of interferometry are introduced: image formation with conventional telescopes, in particular the Fourier decomposition of images, and the nature and utility of measurements of the coherence function or mutual intensity. Thereafter I focus on optical/infrared interferometry, outlining how measurements of the coherence function are made at these wavelengths, how they can be used to interpret a source's structure, and what the principles of interferometric imaging tell us about the limitations expected for the current generation of arrays such as the VLTI.

1 Introduction

The opening to the astronomical community of facility optical/infrared interferometers such as the VLTI is a recent development. Moreover, interferometry at these wavelengths is a sufficiently unusual technique that it demands a broadening of the understanding of a “typical” astronomer to its possibilities and limitations. The goal of teaching the essentials of interferometric imaging to an audience of non-experts is thus both important and timely. However, to attempt to review the field in a single presentation would be a bold and ambitious challenge, and my aim here is somewhat more modest. What I want to do instead is to introduce the basic “language” of interferometry to those who may ultimately wish to exploit arrays like the VLTI for their favourite astrophysics. Many will be unfamiliar with this language so I have assumed no prior knowledge of interferometry at all. What follows then will not be a comprehensive and detailed treatment of optical/infrared interferometry, but rather a brief guide to the territory, with pointers to areas that

I would like to thank the numerous colleagues whose have provided figures with which to illustrate this manuscript, and whose ideas have help clarify mine over the years.

¹ University of Cambridge, Astrophysics Group, Cavendish Laboratory, Madingley Road, Cambridge, CB3 0HE, UK; e-mail: cah@mrao.cam.ac.uk

differ from their radio counterparts. I hope it will give most of my audience the confidence to ask the right questions when assessing the possible merit of optical/infrared interferometry for their own research. A useful reference for further details on many of the topics mentioned here is Lawson (2000).

Before we begin I must make one small confession: I am not an interferometric “salesman”. The historical fact that conventional optical/infrared telescopes continue to provide adequate angular resolution and sensitivity for most astronomical programmes implies that ground-based interferometric arrays should at present rightly be considered “niche” instruments, i.e. as telescopes with a limited remit. Nevertheless, they provide unique and valuable scientific information on the sources that they can observe. The role that they can, and will, play though is not in doubt - an increase of two orders of magnitude in any capability — in this case angular resolution — is guaranteed to lead to new astrophysical insights. The goal for today’s generation of astronomers, and my audience here today, will be to identify and realise these prizes.

2 Image formation with conventional telescopes

An understanding of imaging with conventional telescopes is one of the most useful starting points for investigating what we can realistically expect from the current generation of interferometric arrays. We begin then with the fundamental equation for incoherent imaging (see, e.g., Goodman 1996):

$$I(l, m) = \iint P(l - l', m - m') O(l', m') dl' dm' , \quad (2.1)$$

where $O(l, m)$ is the true source brightness distribution, $I(l, m)$ is the observed brightness distribution, and $P(l, m)$ is the response to a point source, or point-spread function, of whatever imaging system is being used. All of these are functions of l and m which are angular co-ordinates on the sky. This convolutional relationship, which more precisely describes the behaviour of all linear space-invariant (isoplanatic) systems can be written more concisely by taking the Fourier transform of the previous equation to give:

$$\tilde{I}(u, v) = \tilde{P}(u, v) \times \tilde{O}(u, v) \quad (2.2)$$

$$= T(u, v) \times \tilde{O}(u, v) . \quad (2.3)$$

In this equation, functions with a tilde refer to the Fourier transforms of their real space counterparts, and u and v are spatial frequencies with dimensions of radians⁻¹. The importance of this description is that the essential features of the imaging system, whatever form it takes, are encapsulated in a complex multiplicative transfer function, $T(u, v)$. Thus, the amplitude and phase of $T(u, v)$ quantify which spatial frequency components of the object are represented in the image, and how faithful that reproduction is.

An important question we can ask here is: what determines the form of $T(u, v)$? In general, the transfer function is obtained from the normalized auto-correlation

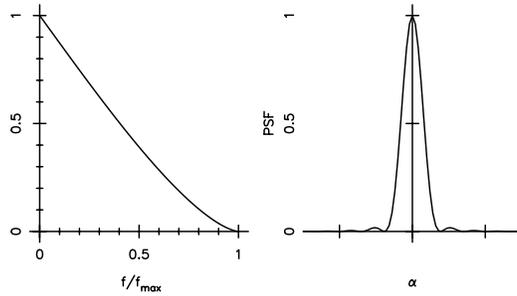


Fig. 1. The transfer function (left) and normalized point-spread function (right) of a perfect telescope with an unobscured circular aperture of diameter D . The maximum frequency transmitted by the aperture is $f_{\max} = D/\lambda$. Note the enhanced weighting of the lower spatial frequencies in $T(f)$ and the relatively poor transmission of frequencies close to f_{\max} . The first null of the point-spread function, an Airy pattern, occurs at an angular co-ordinate $\alpha = 1.22\lambda/D$.

of the transmission function of the telescope pupil, so that:

$$T(u, v) = \frac{\iint P^*(x, y)P(x + u, y + v) dx dy}{\iint |P(x, y)|^2 dx dy}, \quad (2.4)$$

where x and y refer to co-ordinates in the telescope pupil and the integration is taken over all possible values of x and y . A number of key features of this formalism are worth noting carefully. First, in the absence of aberrations, $P(x, y)$ will be equal to 1 where the aperture is transparent and will be 0 otherwise. The presence of aberrations will lead to changes in the phase of $P(x, y)$ and a reduction in the value of the complex valued transfer function. Second, for each spatial frequency, u , there will exist a physical baseline, B , in the pupil, of length λu . Finally, for an unaberrated circularly symmetric aperture, the transfer function can be written as a function of a single co-ordinate such that $T(u, v) = T(f)$, with $f^2 = u^2 + v^2$. Fig. 1 shows a simple example of this behaviour in practice. The left hand panel shows the transfer function for a circular aperture of diameter D , while the right hand panel shows the resulting normalized point-spread function. In this case this is none other than the familiar Airy function with its first null at an angular co-ordinate $\alpha = 1.22\lambda/D$, and a full width at half maximum of $0.9\lambda/D$.

The story thus far may appear relatively straightforward so it is worthwhile reiterating the essential points we should draw from the arguments presented. There are five key ideas to digest:

- The decomposition of an image into a series of spatially separated compact point source responses (point-spread functions), and the formal equivalence of this to a superposition of non-localized co-sinusoidal functions with appropriate amplitudes and phases.

- The description of an image in terms of these co-sinusoidal basis functions each of which has a particular angular period on the sky. These are nothing more than the Fourier components of the image.
- The action of an imaging system as a linear filter for the true spatial Fourier spectrum of the source. Note that in this scheme even images from, for example, the Hubble Space Telescope are filtered versions of the true source brightness distributions.
- The association of each Fourier component (or spatial frequency) with a distinct physical baseline within the aperture collecting light from the source.
- The form of the point-spread function arising from the relative sampling, and hence weighting given to, the different spatial frequencies (and hence baselines) measured by the pupil of the imaging system.

As we will see later, these ideas are so close to the heart of interferometric imaging that it really makes little sense to distinguish imaging with interferometers from the imaging we may already be used to when using conventional telescopes.

3 Coherence functions

The next step in our journey takes us into somewhat more unfamiliar territory but, as we shall see, provides the link between interferometric measurements and the description of imaging we have just reviewed. Fuller treatments of what we will discuss can be found in many other sources (see, e.g., Born & Wolf 1999) and so here the aim will simply be to identify the essential physical bases for imaging with interferometers.

Let us begin by considering the spatio-temporal correlations of the electric field arising from a distant astronomical source (see Fig. 2). We imagine measuring the electric field at two locations r_1 and r_2 and at times t_1 and t_2 where these

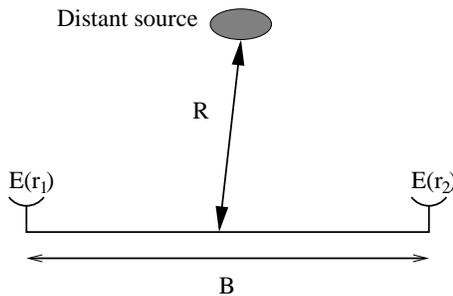


Fig. 2. A schematic diagram of the set-up needed to measure the coherence function of the radiation from a distant source. The electric fields, $E(r_1)$ and $E(r_2)$, from the source are measured at two locations separated by a distance B , with $R \gg B$.

fields result from the summation of different contributions from all elements of the source whose radiation is collected by the antennae. We define the spatio-temporal coherence function (sometimes referred to as the mutual intensity) of the radiation as:

$$V(r_1, t_1, r_2, t_2) = \langle E(r_1, t_1) \times E^*(r_2, t_2) \rangle , \quad (3.1)$$

where the asterisk denotes complex conjugation and the angle brackets indicate an average over a time long compared to the period of oscillation of the electric field. Two forms of this time-averaged complex cross-product will be of interest: the first when $r_1 = r_2$, the so-called temporal coherence function, and the second when $t_1 = t_2$, which is called the spatial coherence function.

3.1 The temporal coherence function

When measurements are made at the same location, but at different times, the spatio-temporal coherence function for an incoherent astronomical source takes a particularly simple form such that:

$$\langle E(r_1, t_1) \times E^*(r_2, t_2) \rangle = V(t_1 - t_2) = V(\tau) . \quad (3.2)$$

Note that the argument of this coherence function depends only on the difference between the times at which measurements of the electric field are made. Physically what the temporal coherence function quantifies is the extent to which the fields along a given wave train are correlated: it is the same quantity measured by a laboratory Michelson interferometer where light from a single source is split, propagated along two unequal paths, and then recombined and the resulting intensity examined.

The usefulness of the temporal coherence function arises from an important result in statistical optics, which we will not prove here (see, e.g. Born & Wolf 1999), the Wiener-Khintchine theorem. This states that the normalized value of $V(\tau)$ is equal to the normalized Fourier transform of the spectral energy distribution, $B(\omega)$, of the source:

$$V(\tau) = \frac{\int B(\omega) e^{-i\omega\tau} d\omega}{\int B(\omega) d\omega} . \quad (3.3)$$

Referring to the elementary properties of Fourier transforms we can see that a broad spectral energy distribution will lead to a coherence function that decays rapidly with τ , since τ and ω are reciprocal co-ordinates. More quantitatively, we can define a coherence time $\tau_{\text{coh}} \sim 1/\Delta\nu$, with $\Delta\nu = \Delta\omega/2\pi$, such that the value of the temporal coherence function for $\tau \sim \tau_{\text{coh}}$ will be close to zero. From the point of view of experimental physics the value of the Wiener-Khintchine theorem arises from the uniqueness of the Fourier transform relationship between $V(\tau)$ and $B(\omega)$: measurements of $V(\tau)$ in principle allow unambiguous recovery of the source spectrum via an inverse Fourier transform. This is spectroscopy without the use of a dispersing element!

3.2 The spatial coherence function

The treatment above is followed closely when we consider the second special case of $V(r_1, t_1, r_2, t_2)$, i.e. when $t_1 = t_2$. In this case, for the spatially incoherent sources that astronomers usually look at, the spatio-temporal coherence function can be written as:

$$\langle E(r_1, t_1) \times E^*(r_2, t_2) \rangle = V(r_1 - r_2) = V(\vec{\rho}) , \quad (3.4)$$

where the time dependence has now vanished and again we are left with a function of a single argument. This is now a spatial co-ordinate, $\vec{\rho}$, equal to the vector separation between the two locations at which the electric field is measured. What $V(\vec{\rho})$ measures then is the extent to which the fields at different points in a plane perpendicular to the incoming wavevector are correlated. In the context of a laboratory analogue, the spatial coherence function is what a Young's slit set-up with two circular "slits" measures: the apertures define the two regions of the incoming wavefront to be examined, and the intensity at an on-axis focal point provides a measure of $V(\vec{\rho})$.

The use of the spatial coherence function for imaging derives from another elegant piece of physics, the van Cittert-Zernike theorem. This states that, for spatially incoherent sources in the far field, the normalized value of the spatial coherence function $V(\vec{\rho})$ is equal to the normalized Fourier transform of the sky brightness distribution, $I(\vec{\alpha})$:

$$V(\vec{\rho}) = \frac{\int I(\vec{\alpha}) e^{-i(\vec{\alpha} \cdot \vec{\rho}) 2\pi/\lambda} d\vec{\alpha}}{\int I(\vec{\alpha}) d\vec{\alpha}} , \quad (3.5)$$

or in slightly different notation:

$$V(u, v) = \frac{\iint I(l, m) e^{-i2\pi(ul+vm)} dl dm}{\iint I(l, m) dl dm} . \quad (3.6)$$

Here u and v are the components of the baseline $\vec{\rho}$ measured in wavelengths and projected onto a plane perpendicular to the incident wavevector, and l and m are angular co-ordinates on the sky (see Fig. 3).

The conclusion here is as remarkable as for the temporal coherence function discussed above. Measurements of the spatial coherence, or "visibility" function — we will use these two phrases interchangeably hereafter — will be related to the sky brightness simply through a Fourier transform. If a suitable number of measurements of $V(u, v)$ can be secured then one can recover the sky brightness distribution without using imaging optics at all!

3.3 Putting it all together

From the point of view of a didactic presentation it is important to note the mathematical equivalence of the spatial coherence function, $V(\vec{\rho}, \tau = 0)$ and the Fourier decomposition of an arbitrary source that we referred to earlier. While the former

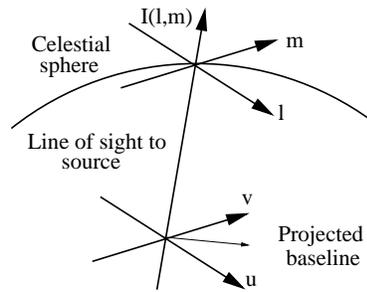


Fig. 3. The (u, v) and (l, m) co-ordinate systems that are conventionally used in interferometry. The physical baseline $\vec{\rho}$ is projected onto a plane perpendicular to the line of sight to the source to obtain u and v . The source brightness distribution is described by the function $I(l, m)$ which exists on the tangent plane to the celestial sphere at the point at which the observer's line of sight intersects it.

results from a particular limiting case of the van Cittert-Zernike theorem appropriate for most astrophysical sources, the latter is merely a convenient method for characterising the content of an image. However, the fact that both involve the Fourier transform of the sky brightness distribution means that we have finally arrived at a powerful means of understanding how interferometry can be exploited for imaging. This can be summarized as follows:

- Any source in the sky can be described as a superposition of co-sinusoids each of which corresponds to a given spatial frequency or angular scale on the sky.
- Measurements of the spatial coherence function are nothing more than measurements of the amplitudes and phases of these Fourier components.
- Interferometers are devices that measure the spatial coherence function.
- Hence, an interferometer with two telescopes with a projected separation, B , will measure the value of the Fourier transform of the source brightness distribution at a spatial frequency $u = B/\lambda$.

So, the practice of interferometry at optical/infrared wavelengths involves nothing more than measuring $V(u, v)$ well enough so that the image resulting from inverse Fourier transforming the coherence function is an adequate representation of what is in the sky. Exactly what “adequate” implies will be explained in a following section. This process of measuring the coherence function and inverse Fourier transforming it is, of course, what conventional telescopes do all the time, albeit in a manner that is completely hidden from the user.

4 Measuring the visibility function at optical/infrared wavelengths

Thus far we have said nothing about how, in practice, one can get access to field quantities like $E(r_1, t)$ and $E(r_2, t)$ so as to compute their time averaged product $\langle E(r_1) \times E^*(r_2) \rangle$. This is often a source of much confusion and so I shall spend a little time outlining how this is performed here.

The simplest way to approach this question is to consider a Young's slit experiment (see Fig. 4). In the focal plane of the set-up a detector will register an intensity given by the modulus squared of the total electric field $E_1 + E_2$. Here E_1 and E_2 refer to the contributions arising from each of the sub-apertures in the screen that intercepts the incoming wavefronts. At the centre of the fringe pattern these contributions will have taken the same time to travel from the screen to the focal plane, and so if we consider only this central part of the fringe pattern we can ignore any temporal coherence effects. The time averaged intensity can then be written as:

$$\langle (E_1 + E_2) \times (E_1 + E_2)^* \rangle = \langle |E_1|^2 \rangle + \langle |E_2|^2 \rangle + \langle E_1 E_2^* \rangle + \langle E_1^* E_2 \rangle \quad (4.1)$$

$$= \langle |E_1|^2 \rangle + \langle |E_2|^2 \rangle + \langle 2|E_1||E_2|\cos(\phi) \rangle, \quad (4.2)$$

where ϕ is the phase difference between E_1 and E_2 .

The intensity thus comprises two parts: a constant term given by the sum of the intensities from each sub-aperture, and an oscillatory part which bears a very close resemblance to the spatial coherence function $\langle E_1 E_2^* \rangle$. The features of the observed fringe pattern that will be of interest are its modulation and its phase. The modulation, or visibility, V , is defined as:

$$V = \frac{[I_{\max} - I_{\min}]}{[I_{\max} + I_{\min}]}, \quad (4.3)$$

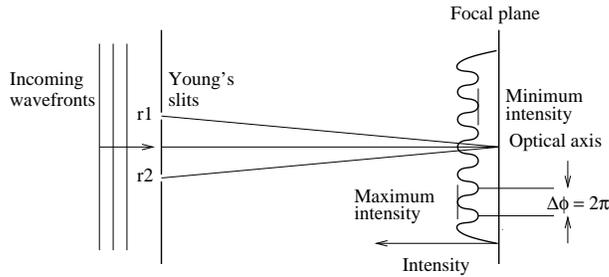


Fig. 4. A schematic diagram of a Young's slit experiment. Light enters from the left, is intercepted by a screen with two small apertures at locations r_1 and r_2 and is subsequently focussed onto a distant screen. The maximum and minimum intensities in the resulting fringe pattern are shown, as is the definition of a phase shift of 2π , i.e. a shift in the fringe pattern of one wavelength.

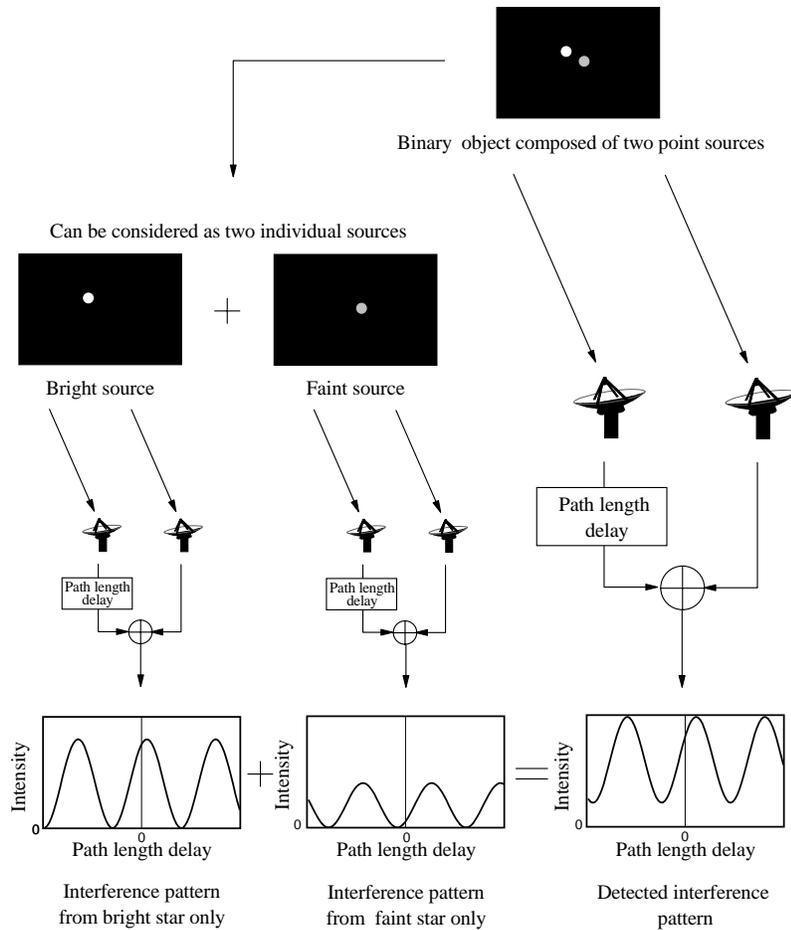


Fig. 5. The relationship between a particular source brightness distribution and the fringes observed in a two-element interferometer. In this case the object is a binary star comprising two unequal and unresolved components. Each component produces its own interference pattern with an amplitude dependent on its brightness, and a phase, ϕ , related to its position in the sky. Note that since both components are unresolved each produces a fringe pattern with a visibility $V = 1$. The overall fringe pattern observed is the superposition of the intensity patterns from each component of the source. The information about the relative brightness and location of the two components is contained in the visibility and phase of this resulting fringe pattern.

where I_{\max} and I_{\min} refer to the maximum and minimum intensities in the fringe pattern, while the phase of the fringe pattern simply refers to the location of the white-light fringe relative to some reference point, usually chosen to be the optical

axis. From an interferometric point of view these two observables are critical because they encode the amplitude and phase of the spatial coherence function respectively. So, the answer to our initial question is simple: to measure the spatial coherence function we need to interfere the fields from two points on the wavefront incident from the source to create a fringe pattern, and we then need to measure the modulation depth and spatial location of the white-light fringe. The way in which the source brightness distribution is actually encoded in the properties of the observed fringe pattern, i.e. how the van Cittert-Zernike theorem comes into play, is explained pictorially in Fig. 5.

5 Science with interferometers

The framework for interferometric imaging we have described above can be summarized as follows:

- The Fourier transform of the source brightness distribution is the spatial coherence or visibility function $V(u, v) = V(B_x/\lambda, B_y/\lambda)$, with B_x and B_y the components along two orthogonal directions of the projected interferometer baseline.
- Measurements of $V(u, v)$ are made with many different interferometer baselines so as to secure as complete information on $V(u, v)$ as is possible.
- An inverse Fourier transform of these data is performed producing a, hopefully faithful, representation of the sky.

Whilst this picture is in principle correct, it begs a number of important questions. For example, what does a typical visibility function look like, and how complete does it have to be measured to be useful? And perhaps more importantly: how will the sampling of $V(u, v)$ impact on the type of images that can be made with an interferometer? The following subsections address these issues in turn.

5.1 Simple visibility functions

The properties of the visibility functions of three simple source types that might be observed with the VLTI are displayed pictorially in Fig. 6. Let us first consider an unresolved source of intensity A_1 located at an angle l_1 relative to the pointing direction of the interferometer. If we orient our co-ordinates so that the interferometer baseline is parallel to the vector separation between the pointing centre and the off-axis source, the visibility function will be a function of a single co-ordinate u , such that:

$$V(u) = \frac{\int A_1 \delta(l - l_1) e^{-i2\pi ul} dl}{\int A_1 \delta(l - l_1) dl} \quad (5.1)$$

$$= e^{-i2\pi ul_1} \quad (5.2)$$

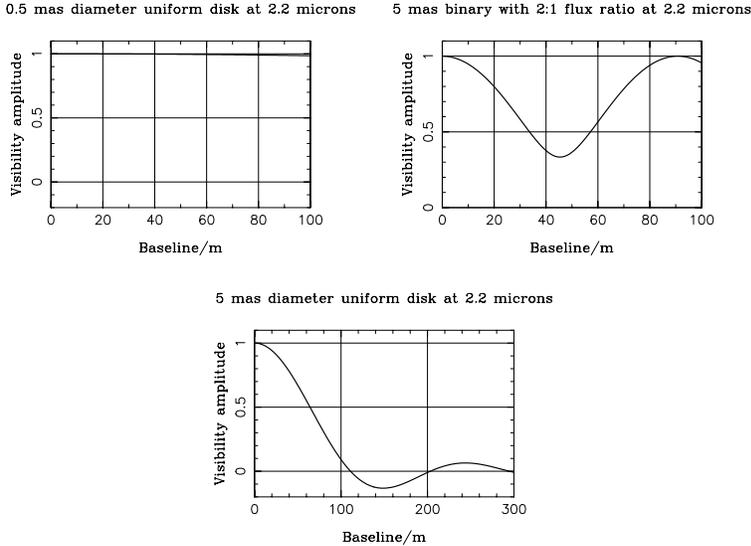


Fig. 6. The predicted normalized visibility amplitude for three types of simple source. In each case a wavelength of $2.2 \mu\text{m}$ has been assumed and the source size has been chosen to match the baselines accessible with the VLTI. For the 5 mas resolved disk (lower panel), which is assumed to be located at the centre of the field of view, amplitudes that are negative correspond to baselines where the phase of the visibility function has flipped from 0 to π radians.

The visibility amplitude will thus be unity for all baseline lengths, while the visibility phase will vary linearly with u ($= B/\lambda$). In practice few sources will be truly unresolved so instead the first panel of Fig. 6 shows the visibility amplitude for a 0.5 milli-arcsecond diameter disk for baselines in the range 0–100 m. This size of disk roughly corresponds to the apparent angular size of an M5 dwarf at 6 parsecs and for this small a source the visibility amplitude is virtually indistinguishable from that of a delta-function. Sources such as these will give interferometric fringes of high contrast and thus be easy to observe, but will be too small to be good targets for imaging studies.

More interesting behaviour is shown by a binary star, here assumed to comprise two point sources of intensity A_1 and A_2 at angles $l_1 = 0$ and l_2 relative to the pointing centre. As before, a suitable alignment of co-ordinate axes allows the visibility function to be written as a function of a single spatial frequency so that we have:

$$V(u) = \frac{\int [A_1 \delta(l) + A_2 \delta(l - l_2)] e^{-i2\pi ul} dl}{\int [A_1 \delta(l) + A_2 \delta(l - l_2)] dl} \quad (5.3)$$

$$= \frac{[A_1 + A_2 e^{-i2\pi u l_2}]}{[A_1 + A_2]} . \quad (5.4)$$

This coherence function is plotted for a 5 milli-arcsecond separation binary, with a flux ratio of 2:1, in the upper right panel of Fig. 6. In this case both the visibility amplitude and phase oscillate with the baseline length B , and baselines from 0 to λ/l_2 meters are required to recognize the binary nature of the source. A further point to note is that as the brightness ratio of the two components increases — for example one of the stars may be a very low mass companion — the modulation of the visibility function will become smaller and smaller. Thus while the fringe contrast itself may well be high, its periodic modulation may be difficult to discern unless very high signal-to-noise ratio measurements can be secured.

The final panel of Fig. 6 shows an example of the visibility function of a fully resolved target, in this instance a centrally-located uniform disk with diameter, θ , equal to 5 milli-arcseconds, for example a M1 Ia supergiant at 1 kpc. The rotational symmetry of the source means that the visibility function can be written in terms of Bessel functions so that:

$$V(u_r) \propto \int_0^{\theta/2} \rho J_0(2\pi\theta u_r) d\rho \quad (5.5)$$

$$= \frac{2J_1(\pi\theta u_r)}{(\pi\theta u_r)} , \quad (5.6)$$

where ρ is the radial circular polar co-ordinate, u_r is its corresponding conjugate spatial frequency, and J_0 and J_1 are the zeroth and first-order Bessel functions of the first kind. Two features of this visibility function deserve note. First, baselines from 0 to beyond λ/θ meters are required to identify the source as a resolved disk (see the plot of $V(u_r)$). And second, the visibility amplitude falls very rapidly as the baseline increases. This is a crucial new fact: since information on scales smaller than the disk diameter will correspond to spatial frequencies $u > 1/\theta$, obtaining such information will require the measurement of fringes with very low contrast. This is a significant observational challenge and implies that the study of resolved stellar disks will be much more difficult than the previous examples we have considered above.

An example of the visibility function expected for a more complex astrophysical source is displayed in Fig. 7. This shows the un-normalized visibility function for an object comprising 6 elliptical Gaussian components with a total flux of 4 units. To allow comparison with the 1-dimensional examples given above, the visibility function has been projected onto a direction parallel to the longest dimension of the source. While complicated in detail, certain features of the visibility function are quite straightforward to explain. The most important points to note are (i) the rapid fall in V from a value of 4 at zero baseline to a value of approximately 1 at a baseline of $\sim 5 \times 10^6 \lambda$ and (ii) the approximate constancy of the visibility amplitude at longer baselines. What can we infer from this? First, that most of the flux from the source is resolved on angular scales of order $\frac{1}{5 \times 10^6}$ radians, and second, that the source contains an unresolved component contributing approximately 25% of

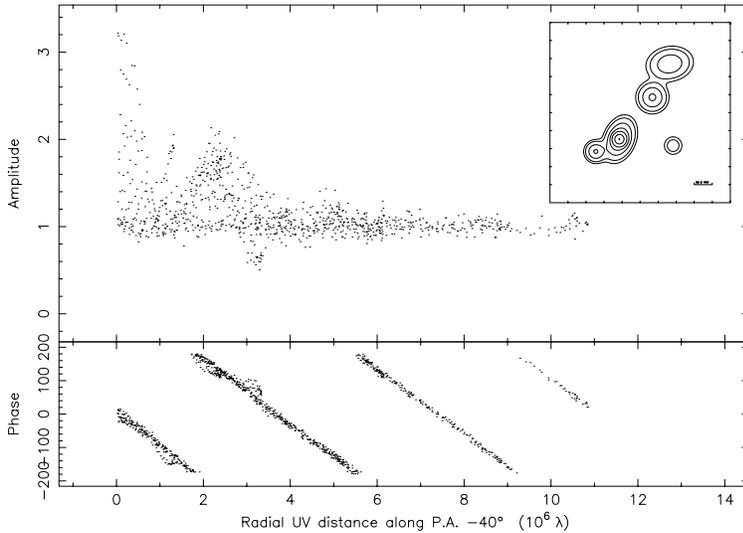


Fig. 7. Samples of the predicted visibility amplitude (arbitrary units) and phase (in degrees) for the multi-component resolved model show in the top right insert. The x axis (baseline length) is measured in mega-lambda, while the visibility has a zero-baseline value of 4 (i.e. is just off the top of the plot). For ease of presentation the 2-dimensional visibility function has been projected along the longer dimension of the source at PA -40° . North is up and East to the left. The map of the source is contoured at 1, 2, 5, 10, 20, 50, and 99% of the peak flux and shows a small scale bar of length 30 milli-arcseconds.

the total flux. In fact, the linear variation of visibility phase with baseline length indicates that this compact feature is offset from the pointing centre, with the rate-of-change of phase identifying the magnitude of this offset.

There is perhaps one final comment to make here. The source discussed in the last example is one with structure that *can* be resolved, and hence is interesting to image, but which has a visibility function that stays notably high at long baselines by virtue of its compact core (c.f. the rather different behaviour of a uniform disk visibility function). As we shall see in the next lecture, the signal-to-noise ratio for fringe parameter measurement is a strong function of the visibility amplitude and if V is too low a high signal-to-noise ratio on a given measurement can be very difficult to achieve. Hence, this type of source, with resolvable structure but with some unresolved flux too, typifies the class of complex objects that is likely to be best suited for detailed interferometric study with the optical/infrared arrays available today.

5.2 Image recovery

The examples above have demonstrated that many features (e.g. separations and flux ratios) of both simple and more complex brightness distributions can be inferred directly from measurements of $V(u, v)$. However, the most exciting challenge for optical/infrared interferometry will be to image complex astrophysical phenomena that have hitherto remained unobservable. What progress in this area can we realistically expect in the near to mid-term?

We begin by reminding ourselves of the fundamental relationship between the visibility function and the normalized sky brightness distribution:

$$I_{\text{norm}}(l, m) = \iint V(u, v) e^{+i2\pi(ul+vm)} du dv . \quad (5.7)$$

In practice, however, what our measurements give us is a sampled version of $V(u, v)$, and so the image we can recover will be the so-called “dirty map”:

$$I_{\text{dirty}}(l, m) = \iint S(u, v) \times V(u, v) e^{+i2\pi(ul+vm)} du dv \quad (5.8)$$

$$= B_{\text{dirty}}(l, m) * I_{\text{norm}}(l, m) , \quad (5.9)$$

where $S(u, v)$ is the sampling function describing which measurements of $V(u, v)$ have been secured, and $B_{\text{dirty}}(l, m)$ is the Fourier transform of the sampling distribution, otherwise known as the “dirty beam”. The dirty-beam is nothing more than the point source response of the interferometer. It is usually far less attractive than an Airy pattern, for example it will often have strong and numerous

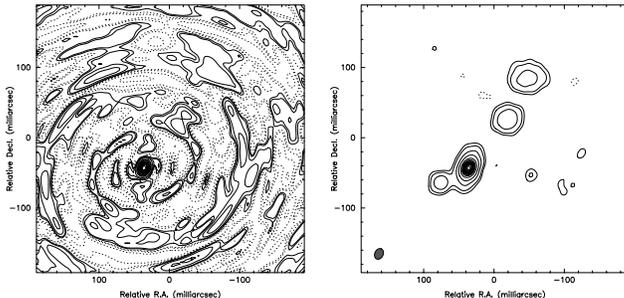


Fig. 8. The dirty (left) and deconvolved (right) image for a simulated observation of the test source presented in Fig. 7. In both panels contours are plotted at -10, -5, -2, -1, 1, 2, 5, 10, 20, 30, 40, 50, 60, 70, 80, and 90% of the peak flux. Negative contours are shown dashed. Despite the fact that the source is barely visible in the dirty image, it is straightforward to deconvolve the map and correct for the interferometer point-spread function. The residual noise features in the right hand panel reflect the imperfect sampling of the uv plane in this simulation. The small grey ellipse in the bottom left hand corner of the deconvolved map shows the size of the core of the point-spread function, and hence the effective resolution in the restored image.

sidelobes, but it is completely determined by the known sampling of the uv plane. So, the lesson to learn here is that despite the unusual form of the interferometric point-spread function its behaviour will generally be very well understood and it can be accounted for very straightforwardly. This process of correcting an interferometric image for the uv plane sampling is known as deconvolution and can be performed using many schemes such as CLEAN, MEM, and WIPE (see, e.g., Cornwell *et al.* 1999; Lannes *et al.* 1994). Fig. 8 shows an example of this type of deconvolution in practice with “before” and “after” images for a simulated observation of the model source whose visibility function is displayed in Fig. 7. It is interesting to note how difficult the source is to discern in the dirty map, yet how successful the deconvolution is.

For those who are unfamiliar with interferometric imaging it is fortunate that much of the useful experience gleaned by radio astronomers over the years, and which will still apply to imaging with optical/infrared arrays, can be encapsulated in a small number of useful “rules-of-thumb”. These include the following:

- The number of visibility data measured should be greater than or at least equal to the number of filled pixels required in the recovered image. Note that it is the number of pixels that will have flux in that is relevant here, not the total number of pixels in the image. So for an observation comprising snapshot measurements with R reconfigurations of an N -telescope array this implies: $\frac{N(N-1)}{2} \times R \geq \text{number of filled pixels}$.
- The distribution of samples of the visibility function should be as uniform as possible so as to make deconvolution straightforward. In essence this is simply saying that large gaps in the uv plane coverage will be difficult to deal with, or equivalently that a faithful representation of the source will not be possible if measurements of large ranges of spatial frequencies are missing from the interferometric dataset.
- The range of angular scales displayed in the interferometric image will be governed by the range of projected interferometer baselines, i.e. B_{\max}/B_{\min} . For interferometers with small numbers of collectors, and where the time available to image the source permits few reconfigurations of the array, this ratio may be very small indeed, e.g. < 5 .
- The overall extent of the source being observed will determine how finely $V(u, v)$ should be sampled. More precisely, for a source of size θ_{\max} radians sampling very much finer than $\Delta u \sim 1/\theta_{\max}$ will be unnecessary.

To see what these rules mean in practice let us consider an example of the uv coverage expected for a possible future phase for the VLTI, using eight telescopes (see Fig. 9). This corresponds to a six hour observation where all 28 possible baselines are measured every ten minutes. The total number of visibility data is thus $28 \times 36 = 1008$. This observing strategy gives good uniformity in the uv coverage, though the temporal sampling of $V(u, v)$ is perhaps a little too frequent for the size of the target, and more importantly the range of angular scales sampled

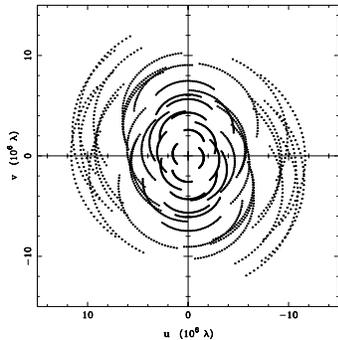


Fig. 9. The uv coverage expected for an eight-element array at Paranal for a source at declination -65° . The points plotted indicate the spatial frequencies sampled by the array at ten minute intervals over a six hour observation. Note that even with this dense coverage of the uv plane the ratio of maximum to minimum baseline length is only ~ 9 .

is only a factor of ~ 9 . Exactly how these factors impact on the interferometric imaging process can be seen in the map shown in the right hand panel of Fig. 8. This was reconstructed from simulated data assuming the same uv coverage as in Fig. 9 and with only small errors in the amplitude and phase of the measured visibility function ($\Delta V/V \simeq 0.05, \Delta\phi \simeq 10^\circ$). Overall the reconstructed image is of high quality and is very similar to the model source used to generate the test data (compare the upper right panels of Figs. 7 and 8). The interferometric map shows emission with a range of angular scales from 1 to ~ 10 times the interferometric resolution and extending over an area of approximately 15×15 resolution elements. This would clearly have been a successful observation, but note how many samples of $V(u, v)$ were required and the high signal-to-noise ratio of each datum. Maps of this quality come at an expensive price!

Before leaving this section on imaging there are two further topics we need to touch on. These are the field-of-view and the image quality that we should expect from interferometry.

The first of these often features prominently in discussions of interferometric imaging, but usually only when optical/infrared astronomers participate! I think the reason for this is a historical accident which needs explaining. It is now commonplace for conventional optical and near-infrared telescopes to have high-resolution imaging cameras with fields of view measured in minutes of arc, i.e. hundreds of times the size of the available angular resolution. Astronomers used to this type of instrumentation may naturally expect optical/infrared interferometers to have similarly large fields of view. However, as has been outlined above, this would require measurements of the coherence function on baselines with lengths ranging from some minimum value to hundreds of times that value. The practical difficulties in sampling the uv plane over two orders of magnitude in spatial frequency are certainly non-trivial, and so while specialised optical con-

figurations for achieving this do exist in principle, a range in measured spatial frequencies of a factor of ten is more realistic (see, e.g. Fig. 9). So, in the near to mid-term fields-of-view of up to 10×10 resolution elements is what we should be expecting. It is perhaps some consolation to know that even mature radio arrays like the 27-element VLA spend most of their time making maps with linear extents of perhaps $40\times$ the interferometric resolution. So this tale does have a positive side to it: all the evidence from radio-astronomy shows that surprisingly small interferometric fields of view can still lead to excellent science!

My final comments concern the quality of the maps that an interferometer can provide. This is difficult to answer in a fully quantitative sense because the fidelity of an interferometric image will depend on whether or not any particularly strong spatial frequency components in the source have been left unmeasured. This will depend on source morphology itself and so perhaps the only useful statement we can make is that the uv plane coverage should be as complete as possible so as to cope with any possible source structure! However, we can make a more quantitative statement about the levels of random noise expected in an interferometric map. This measure of quality is conventionally characterized by the “dynamic range”. This is the ratio of the maximum intensity in the map to the intensity of the weakest believable feature. If we assume that our measurements sample the visibility function pretty uniformly then, broadly speaking, the dynamic range will be given by

$$\text{Dynamic range} \sim \sqrt{N_{\text{data}}} \times (S/N)_{\text{per-datum}} \quad (5.10)$$

$$\sim \frac{\sqrt{N_{\text{data}}}}{\sqrt{[(\Delta V/V)^2 + (\Delta\phi)^2]}}, \quad (5.11)$$

where S/N is the signal-to-noise ratio, $\Delta V/V$ is the typical fractional visibility amplitude error per datum, $\Delta\phi$ is the typical phase error in radians per datum, and N_{data} is the total number of visibility measurements available.

Thus, for an 8-element array and with visibility and phase errors of a few percent this equation implies a dynamic range of approximately 500:1 or around 7 magnitudes. The fact that the imaging simulation of Fig. 8 just fails to achieve this is actually because the uv plane coverage is not quite uniform enough. Real observations, where slightly more care is taken in minimizing any large gaps in the sampling of the visibility function, should thus be able to detect very faint structures close to brighter cores if this level of random and systematic error can be achieved.

6 Sensitivity

Any discussion of a new experimental tool for astrophysics would not be complete without some mention of its sensitivity and thereby the range of source types for which it could usefully be exploited. However, the complexity involved in deploying optical/infrared interferometers means that the actual limiting sensitivity

achievable is a very strong function of the details of the particular implementation employed. To discuss, for example, the sensitivity of the VLTI would thus be out of place in an introductory lecture like this. What I think is useful though is to talk a little about what sensitivity actually means in an optical/infrared interferometric context.

Since all ground-based interferometers will have to overcome the phase perturbations introduced by the atmosphere a convenient and more useful re-casting of the question “How faint can I go?” is actually the following: “How bright must my source be such that I can secure useful interferometric data?” The answer to this question comes in two quite separate parts:

First, the source must be bright enough to provide a sufficiently strong signal to allow real-time correction of any phase perturbations in the incoming wavefronts. These may come from the atmosphere or may be internal to the interferometer.

Second, the source must be bright enough to allow a reasonable signal-to-noise ratio for the fringe parameters to be built up over some convenient total integration time. Because the rotation of the Earth will sweep the physical baselines through the uv plane at the sidereal rate, this time will generally be measured in minutes and not hours.

Once these two criteria have been satisfied, the faintest structures detectable will be governed by the total number and quality of the visibility data as outlined above in our discussion of dynamic range. In the case of imaging, the flux of these weak features will be measured relative to the brightest feature in the map.

This is quite clearly a rather different approach to quantifying sensitivity than is normally used by optical/infrared astronomers. However, it highlights an important similarity between two contemporary methods for achieving high angular resolution in optical/infrared astronomy: adaptive optics and interferometry. Both of these overcome the seeing limit through real-time correction of the atmospheric corrections. Both require a reference signal with which to sense the atmosphere and unless such a signal exists neither will be effective no matter what other steps are taken. We will return to the topic of sensitivity in more detail in the next lecture where we will see to what extent the “source” we have been referring to need actually be the science target of interest. But it will perhaps come as no surprise to find that the limiting “sensitivity” of optical/infrared interferometry will not be dissimilar to that of natural guide star adaptive optics.

7 Current examples and expectations for the near to mid-term

This has been a long lecture and so I want to finish with something that will be easy to digest. Many of you will have come to this school to learn what the VLTI can do for you. To set your expectations at the right level I want to present three examples of real interferometric data — I stress these are not simulations — that typify what optical/infrared interferometers can realistically provide.

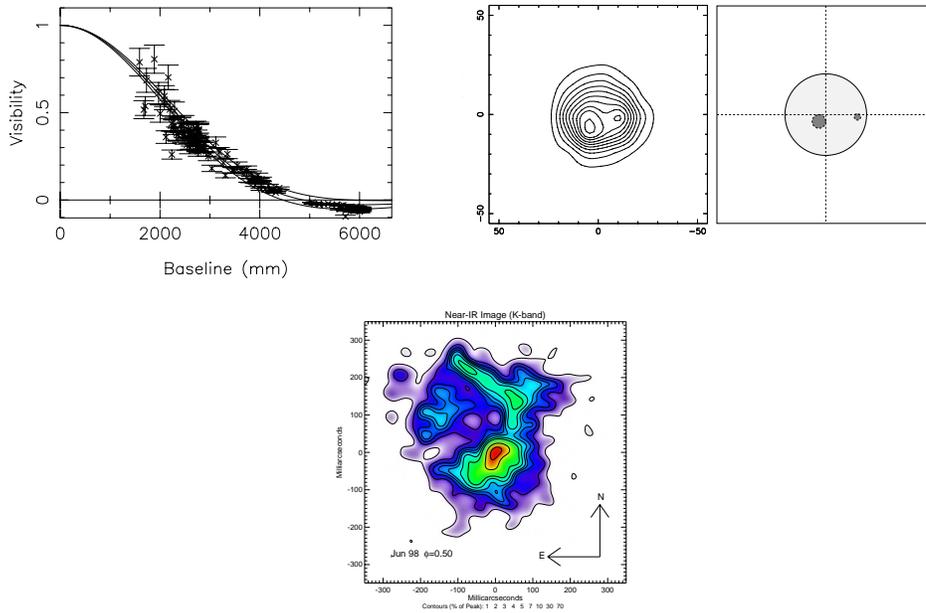


Fig. 10. Three examples of interferometric data obtained with ground-based optical/infrared interferometers. The top left panel shows the visibility function of the M2 Iab supergiant α Ori measured with a 3-element array at 830 nm. To its right are an interferometric image and sketch of the surface structure of the M5 Ib supergiant α Her measured with a 5-element array at 633 nm. The lower panel shows an interferometric map of the K-band emission surrounding the carbon star IRC+10216 obtained with a 21-element interferometer. See text for references. Note the increasing complexity of the analysis that becomes possible as the number of array elements is increased.

The first of these is shown in the upper left panel of Fig. 10, and is the visibility function of α Ori, a M2 Iab supergiant star, measured with a 3-element array at 830 nm (Burns *et al.* 1998). The data extend beyond the first null of the visibility function and so can be used to investigate both the angular size and limb-darkening profile of the star.

With more array elements, and hence better Fourier plane coverage, model-independent imaging becomes possible. The upper right hand panel of Fig. 10 shows an interferometric image of the surface of the M5 Ib supergiant α Her (Tuthill *et al.* 1997) together with a schematic decomposition of the image into three components: a uniform disk, and two unequally bright spots. The image, which was obtained at 633 nm using a conventional telescope converted into a 5-element interferometer through the use of an aperture mask (see, e.g., Haniff *et al.* 1987 for an explanation of this technique), clearly resolves the 50 milli-arcsecond diameter stellar disk and contains perhaps 3×3 filled pixels. It is a rudimentary

image, but shows structure that would have been impossible to discern using any other direct imaging method.

Finally, with large (≥ 8) numbers of array elements, imaging of complex source structures becomes feasible. The lower panel of Fig. 10 shows an interferometric map of the carbon star IRC+10216 in the near-infrared K-band obtained using a 21-element aperture-masked telescope (Tuthill *et al.* 2000). This image has a dynamic range of $\sim 100:1$, an angular resolution of better than 50 milli-arcseconds, and contains of order 10×10 filled pixels. More importantly, it is equivalent in quality to any comparable radio interferometric map made under phase-unstable conditions. It quite clearly demonstrates what optical/infrared interferometry can deliver if dense and high quality sampling of the coherence function is available.

In conclusion then, what type of interferometric science can we expect from large ground-based arrays in the near to mid-term? I think the answers to this question are fairly clear:

- 2 telescope interferometers will allow simple parametric model fitting.
- 5 telescope arrays will allow rudimentary, but nevertheless unambiguous and model-independent, imaging.
- Arrays with large numbers of telescopes (≥ 8) will provide model-independent imaging of complex astrophysical phenomena.

The practical problems we have to overcome to realise these goals will be the subject of our next lecture.

References

- Born, M. & Wolf, E. 1999, *Principles of Optics*, (Cambridge UK: Cambridge University Press)
- Burns, D., Baldwin, J.E., Boysen, R.C., Haniff, C.A., Lawson, P.R., Mackay, C.D., Rogers, J., Scott, T.R., Warner, P.J., Wilson, D.M.A. & Young, J.S. 1998, MNRAS, 297, 462
- Cornwell, T.J., Braun, R. & Briggs D.S. 1999, ASP Conference series, 180, 151
- Goodman, J.W. 1996, *Introduction to Fourier Optics* (McGraw-Hill: New York)
- Haniff, C.A., Mackay, C.D., Titterton, D.J., Sivia, D., Baldwin, J.E. & Warner, P.J. 1987, Nat, 328, 694
- Lannes, A., Anterrieu, E. & Bouyoucef, K. 1994, J. Mod. Opt., 41, 1537
- Lawson, P.R. 2000, editor, *Principles of Long Baseline Stellar Interferometry*, JPL Publication 00-009 (Pasadena: NASA)
- Tuthill, P.G., Haniff, C.A., & Baldwin, J.E. 1997, MNRAS, 285, 529
- Tuthill, P.G., Monnier, J.D., Danchi, W.C., & Lopez, B. 2000, ApJ, 543, 284