

PRACTICAL CONSIDERATIONS IN GROUND-BASED OPTICAL/INFRARED INTERFEROMETRY

Chris Haniff¹

Abstract. In this second lecture I review some of the practical considerations associated with developing and implementing ground-based interferometry at optical and near-infrared wavelengths. Familiarity with these logistical concerns is relevant not only for understanding how arrays such as the VLTI actually operate but also for identifying the most important questions that need to be addressed when planning a campaign of interferometric observations. I outline some of the strategies that have been adopted to address these issues and point to some areas ripe for future development.

1 Introduction

In the first lecture of this school I outlined the physics underlying the use of interferometry as a high-angular resolution tool, and sketched out in broad terms the methods used to produce high resolution images of astronomical sources. In today's class I want to focus on the first stage of this process and explore the experimental difficulties associated with making measurements of the spatial coherence function from the ground. Two major challenges make this a less than straightforward task: (i) the need for opto-mechanical precision at the level of an optical wavelength, i.e. on micron scales and (ii) the need to cope with the presence of a turbulent and rapidly fluctuating atmosphere. How these two problems can be overcome will form a substantial component of our discussion.

As in my first lecture, the space available here does not allow a comprehensive treatment of many of the issues and problems raised (interested readers are again directed towards the excellent series of lectures in Lawson (2000) for more information). The solutions to the myriad challenges facing interferometrists have

As ever, I would like to thank the many colleagues who have contributed to the material presented here and helped further the practice of optical/infrared interferometry over the years.

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

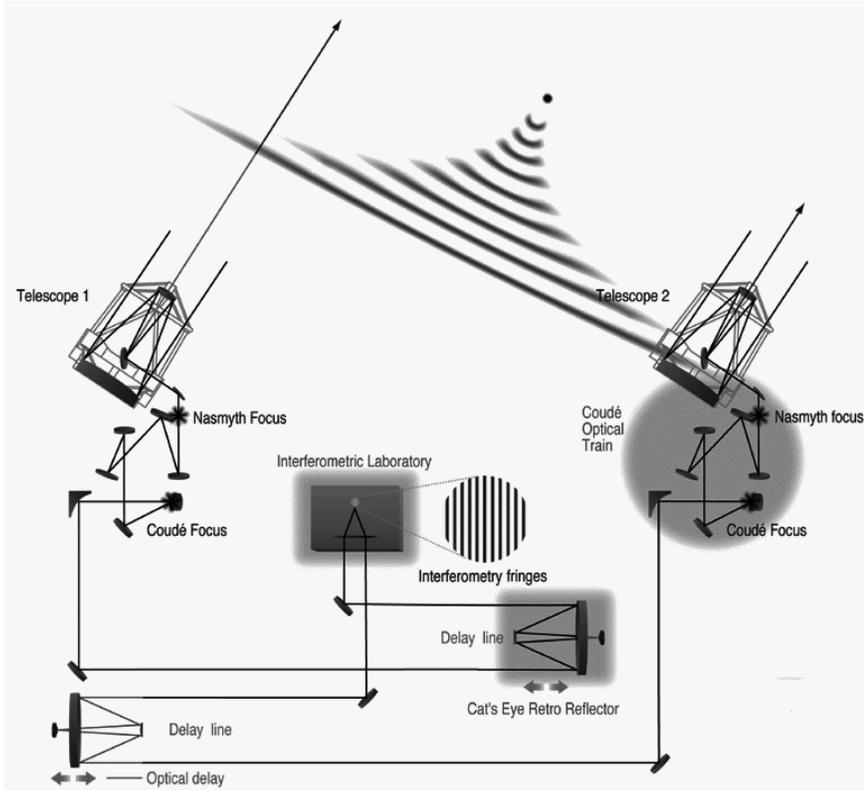


Fig. 1. A schematic view of the principle of operation of the VLTI. The source is assumed to be at infinity so that the incident wavefronts are plane. Note that any given wavefront arrives at the two telescopes at different times. See main text for further details.

required huge investments of effort, and I apologize in advance for dealing with them so cursorily. I hope my readers will recognize the elegance and subtlety of many of the solutions that we now take advantage of, and take heart in the knowledge that while ground-based interferometry is undoubtedly difficult it has not yet been able to foil the ingenuity of dedicated interferometric scientists working at many different laboratories worldwide!

2 Tasks to be tackled

A useful starting point from which to begin our survey of the implementational difficulties associated with ground-based interferometry is a list of the tasks required to measure the visibility function of an astronomical source. These are depicted schematically in Fig. 1 which shows the light from a distant point source arriving at two collectors of an interferometric array, in this case the VLTI. As

the wavefronts travel to their final destination, the interferometric hardware has to perform, as a minimum, the following roles in sequence:

- Sample the radiation from the source at two locations r_1 and r_2 .
- Deliver the light from each collector to a central laboratory.
- Equalize the optical paths for the separate light beams.
- Mix the light to form fringes.
- Detect and record the interference fringes.
- Extract the amplitude and phase of the spatial coherence function.

Fortunately, these functions are sufficiently distinct that we can treat each of them quasi-independently. However, there *are* two themes that underly the chosen solutions to all of these design questions. These are the need to maintain good sensitivity and the need to be able to cope with the atmospheric perturbations experienced at ground-based observatory sites. Not surprisingly, the overall scientific productivity of an array is a strong function of its sensitivity limit, and as we will see later, this is most strongly constrained from the ground by the presence of the atmosphere.

3 Sampling the incident electric field

The first two steps on our interferometric journey involve sampling the incoming radiation field and delivering these samples to a distant laboratory. This is conventionally performed using afocal telescopes which produce collimated beams of light at their outputs. As described in the first lecture, the optimum location and number of samples of the radiation measured will be determined by a series of rules-of-thumb related to the type of information required. Some of the more practical questions that an observer will need to ask herself include:

- What interferometer stations should I use, and what range of hour angles should I observe at given the co-ordinates of the source?
- What will the maximum time allowed for building up the coverage of the uv plane be?
- Will any of the interferometer stations shadow each other at the times chosen for observations?
- Will the zenith distance be too large at any of the times of observation, i.e. either lead to poor seeing at low elevations or exceed the capability of the delay lines?
- Will the source be too resolved to permit monitoring of the atmospheric fluctuations during the measurements (see later)?

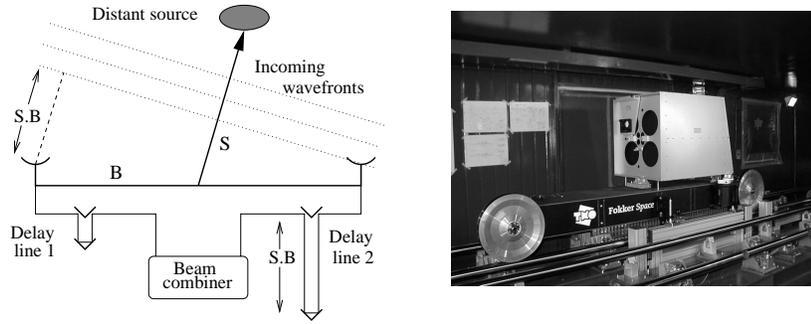


Fig. 2. Optical path equalization in a two-element interferometer. The origin of the optical path difference needing correction is the differential arrival time of the incoming wavefronts at the two collectors. The magnitude of the delay is equal to the projection of the baseline vector B in the direction towards the source S . The right hand panel shows one of the delay line carriages for the VLTI: two of its wheels and the precision rails on which it travels are clearly visible. It has a length of 2.5 metres and a mass of 250 kg.

Once sampled, the radiation from the source is usually delivered to a central optical laboratory via free-space propagation in air-filled or evacuated pipes. Provided the beams are sufficiently wide that diffraction losses are small (see, e.g., Horton *et al.* 2001), and the beampaths are matched so as not to introduce any differential polarization (see, e.g., Traub 1988) or field rotation between the beams, this is an efficient method of beam transport. More recently there has been interest expressed in using single-mode optical fibres as beam transport devices (Perrin *et al.* 2000). Whether these can compete in terms of overall throughput, optical bandpass and ease of use will be explored over the next few years.

4 Equalizing the optical paths

Once the interferometric beams have been delivered to a beam-combining laboratory it is usual to correct them for their differential arrival times at the two primary collectors. This compensation for the geometric delay (see Fig. 2) is, in most cases, accomplished using an optical “trombone” in each arm of the interferometer. These are movable carriages that can be driven along precision tracks and which carry a retro-reflecting optical assembly. By controlling the locations of the carriages appropriately, delays can be introduced into each arm of the interferometer so as to match the optical paths in each arm continuously as a source is tracked across the sky.

For truly monochromatic observations, this OPD matching need not actually be performed — interference fringes will be visible even for large path mismatches. But for any finite bandpass, $\Delta\lambda$, there will be a characteristic optical path dif-

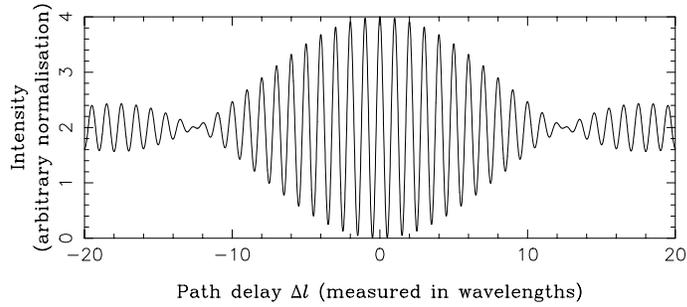


Fig. 3. The two-beam pupil plane fringe pattern expected for a polychromatic unresolved source. The fringes are shown as a function of optical path difference between the interfering beams. A top-hat spectral bandpass with $\Delta\lambda/\lambda = 0.08$ ($l_{\text{coh}} = 12.5\lambda$) causes the fringes to be modulated with a sinc function “coherence envelope” with its first null at an OPD of l_{coh} . Note that the apparent fringe visibility is only equal to the expected value of unity close to the centre of the coherence envelope.

ference (OPD), measured by the coherence length $l_{\text{coh}} \sim \lambda^2/\Delta\lambda$, at which the observed fringe contrast will have decreased to zero (see Fig. 3 for an example of this). For useful science measurements to be secured the geometric delay must thus be corrected well enough that any residual OPD is small compared to l_{coh} . The VLTI delay line carts allow for the insertion of up to 120 m of extra optical path and can be moved at velocities of up to 0.5 ms^{-1} with a position jitter of no more than 14 nm rms. Note that this stability is much better than the coherence length requirement for any bandpass that might be used at near or mid-infrared wavelengths with the initial suite of instruments at the VLTI, and is actually related to the ability of the carriages to provide sub-wavelength stabilization of the OPD over long periods of time.

There are two additional aspects of delay correction that are worth mentioning here. The first is most clearly explained with reference to Fig. 2. It is clear from this diagram that the use of a moveable carriage permits correction of the geometric delay for only one direction in the sky at a time¹. Thus any OPD matching will degrade as a function of off-axis field angle. This gives rise to the usual field-of-view limitation for a pupil-plane interferometer with maximum baseline B of:

$$\theta_{\text{max}} < [\lambda/B] \times [\lambda/\Delta\lambda] , \quad (4.1)$$

i.e. that the field of view can usually be no greater than the product of the spatial and spectral resolutions.

¹In actual fact, this will depend on whether the beam combination is performed homothetically, i.e. maintaining the relative pupil separations and sizes as seen from the source. No such beam combiners have been designed for separated-element arrays yet.

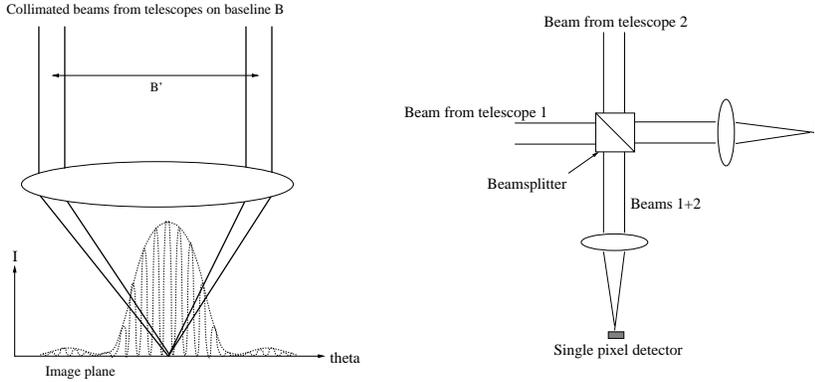


Fig. 4. Schematic diagrams of the optical set-ups required for two-beam image plane (left) and pupil plane (right) beam combination.

Second, for delay lines where the geometric delay is compensated for in a medium with a wavelength-dependent refractive index, such as air in the case of the VLTI, then even for a single pointing direction the delay can only be compensated for at a single wavelength. This so called “longitudinal dispersion” (see, e.g. Tango 1990; Lawson & Davis 1996) is, generally speaking, only a small problem at long wavelengths. However, for long baselines and at large zenith angles with broad bandpasses it must be corrected for with additional optical elements. To get a feel for the magnitude of this effect at the VLTI we can consider measurements of a source 50° from the zenith with a 100 m baseline. Over the extent of the near-infrared K band, from $2.0 - 2.5 \mu\text{m}$, the dispersion-induced differential delay will be $\sim 10 \mu\text{m}$ and hence comparable to the coherence length of the light. From a practical point of view this implies that keeping the visibility losses due to this effect below 10% will need a spectral resolution, $R \geq 5$, or a value of $R > 12$ for measurements made in the J band at $1.25 \mu\text{m}$.

5 Combining the optical beams

The correlation of the signals in an optical/infrared interferometer is usually referred to as “beam combination”. The essential feature of this process is the formation of a set of interference fringes whose contrast and location encode the amplitude and phase of the spatial coherence function (Haniff, these proceedings). In practice there are numerous ways in which such fringe patterns can be realised, but often the two simplest schemes are employed.

The first of these is so-called “image plane” combination (see left panel of Fig. 4). Here, collimated beams from each interferometric collector are brought together with a lens or mirror to a common focus. The resulting focal plane image will be crossed by fringes in an identical manner to that seen in Young’s double-

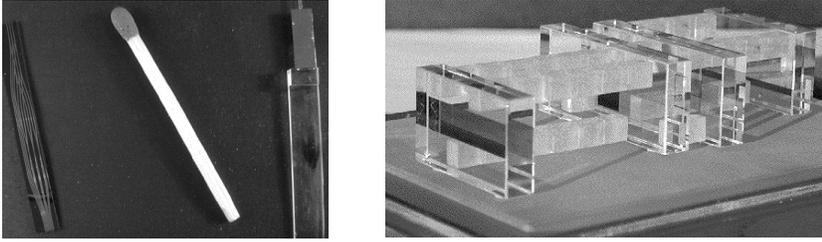


Fig. 5. Examples of two and three-way integrated optics beam combiners (left) and a four-way bulk optics pupil-plane beam combiner (right). The integrated optics combiners are very small indeed, whereas the bulk optics combiner is ~ 10 cm long. Both are physically very stable devices.

slit experiment. The angular size of the focal plane image will be governed by the diameters of the incoming beams, and will thus set the minimum detector array size. The fringe period, and hence the maximum detector pixel size, will depend on the physical separation of the incoming beams at the focussing lens/mirror. This last point is particularly important because it underlies the method by which multi-way beam combination can be realised: if many beams are being correlated, then the input pupil configuration must be chosen so that every pair of beams corresponds to a uniquely identifiable vector separation and hence spatial fringe period on the detector. In other words, the fringe encoding relies upon using a non-redundant input pupil. Often this can be arranged with a 1-dimensional configuration of beams which then allows the other spatial dimension of an array detector to be used for wavelength dispersion (see, for example, the discussion of the optical layout of AMBER (Petrov, these proceedings)).

The other favoured method for optical correlation goes by the name of “pupil plane” beam combination (see right panel of Fig. 4). In this scheme the beams to be correlated are mixed at a beamsplitter (usually a 50:50 mirror, but sometimes an equivalent fibre or integrated optics coupler) and the two complementary outputs are focused onto separate single pixel detectors. The intensity at each detector is then measured as a function of the optical path difference between the input beams. This is equivalent to the optical arrangement of a laboratory Michelson interferometer, and, as in that experiment, the interference fringes are visualized by deliberately modulating the optical path in one of the input beams. In multi-beam pupil plane combiners (see, e.g. Mozurkewich 1994) non-redundant modulation of the optical paths in each beam is used to give rise to a modulated intensity output that is separable into different temporal frequency components, each of which is associated with a single pair of input beams. The amplitude and phase of each of these temporal fringe patterns is what the interferometrist has to measure.

Both of these types of beam combination will be represented in the initial complement of facility instruments at the VLTI. In the near infrared, AMBER will use an image plane beam-combiner (Petrov, these proceedings) while in the

mid-infrared, MIDI will combine its beams using a 50 : 50 beamsplitter (Perrin, these proceedings). These choices highlight an important point: while the two approaches of spatial and temporal encoding are formally equivalent, the choice as to which to employ will always depend on practical considerations such as the availability of low-noise array detectors, access to linear and fast path modulators, and the choice of suitable optical components. Currently, working beam combiners have been implemented using bulk optics, integrated optics (Berger *et al.* 2001), and fibre-optics (Coudé du Foresto & Ridgway 1992) components. Which of these will be favoured in the future is likely to depend on which can provide the best stability and throughput, the best broad-band performance, and the best functionality for large numbers of interferometric collectors.

6 Coping with atmospheric fluctuations

Thus far our discussion has implicitly ignored the most important challenge for ground-based interferometers, i.e. how to overcome the Earth's atmosphere. The fluctuations this introduces into the wavefronts from distant sources are the most important factors limiting the exploitation of interferometric methods at optical and infrared wavelengths, and so an understanding of these perturbations is of particular importance for us today. A full treatment of the spatio-temporal fluctuations produced by the atmosphere is beyond the scope of this lecture, but interested readers are encouraged to consult Roddier (1981) for an excellent overview of the subject. The question I want to concentrate on here is more pragmatic: exactly how do these perturbations affect the ability of interferometers to secure high quality and astronomically interesting data, and how can they be overcome?

6.1 Spatial fluctuations

The spatial fluctuations introduced by the atmosphere are conventionally characterised by a spatial scale, Fried's parameter r_o . This is roughly equal to the diameter of the circular aperture over which the root mean-square wavefront perturbation is 1 radian. Fried's parameter scales as $\lambda^{6/5}$, and at the best astronomical sites takes values of order 15 cm at 500 nm. Broadly speaking telescope diameters, D , smaller or greater than r_o will give instantaneous images that are either diffraction limited ($D < r_o$) or highly distorted and speckled ($D > r_o$).

The impact of these spatial wavefront fluctuations on interferometric measurements is summarized in the left-hand panel of Fig. 6. This shows the expected root-mean-square visibility amplitude for observations of an unresolved source ($V_{\text{intrinsic}} = 1.0$) with telescopes of different sizes, and with and without the use of a fast autoguiding system (Buscher 1988). The major effects of the wavefront corrugations are a rapid drop in V_{rms} and a rapid increase in the fluctuations of V as the telescope size increases. Together these lead to a loss in sensitivity — in the photon limited case the signal-to-noise ratio for fringe parameter measurements scales as V — and an increased difficulty in calibration. The situation is

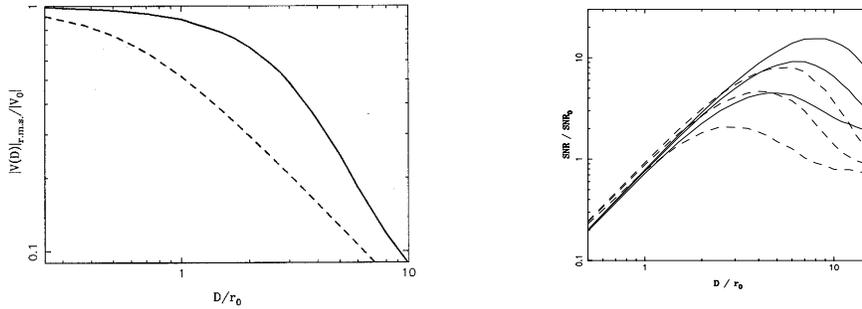


Fig. 6. The effect of spatial wavefront corrugations on interferometric measurements. The left hand panel shows the rms visibility amplitude expected for an unresolved source observed with telescopes of different diameters D . Results for uncorrected (dashed line) and tip-tilt corrected (solid line) optical systems are given. The right hand panel shows the signal-to-noise ratio for amplitude squared measurements for telescopes with different levels of modal adaptive correction. Results are shown for perfect 2-, 5- and 9-Zernike mode correction, with (solid line) and without (dashed line) the use of an optical fibre spatial filter. Higher order corrections give enhanced signal-to-noise.

improved with the use of a fast autoguider, but this is only effective for aperture sizes of approximately $\leq 3r_o$. Beyond this, higher order adaptive correction becomes necessary to limit the precipitous drop in V_{rms} as the aperture size is increased.

Two work-arounds for this problem have been proposed:

- The use of an adaptive optics (AO) system to correct higher order components of the wavefront perturbations beyond tip and tilt. At the VLTI this will be required for efficient use of the Auxiliary Telescopes for wavelengths shorter than $1.65 \mu\text{m}$ and for all wavelengths shorter than $5 \mu\text{m}$ for the UTs (see table 1). The principal difficulty of this approach is the need to find a suitable guide star for the AO system. This must not only be bright enough to drive the wavefront sensor but also be close enough to the science target to

$\lambda/\mu\text{m}$	1.25	1.65	2.2	3.5	5.0
ATs	4.0	2.9	2.0	1.6	0.8
UTs	17.8	12.7	9.0	5.2	3.4

Table 1. The ratio of D/r_o for the auxiliary (AT) and unit telescopes (UT) of the VLTI at different near and mid-infrared wavelengths. A median value of 15 cm for r_o at $0.5 \mu\text{m}$ has been assumed. Values of $D/r_o > 3$ imply the need for some form of moderate order adaptive optics correction to allow the larger telescope aperture to be used effectively for interferometry.

ensure that it experiences very similar wavefront perturbations to the light from it. We can quantify these constraints for the VLTI by considering the 60-element MACAO AO systems for the UTs. These are expected to deliver a Strehl ratio of 0.6 at $2.2\ \mu\text{m}$ for bright on-axis reference stars ($m_v < 10$) but a Strehl ratio of only 0.1 when $m_v = 16.5$. Furthermore, an off-axis offset of $\sim 40''$ between the source and reference star will lead to a similar reduction in their AO performance in the K band.

- Spatially filtering the light arriving from the interferometric collectors. This passive approach (see, e.g., Shaklan & Roddier 1988) exchanges the reduced and fluctuating visibility signal provided by a non-adaptively corrected array for a signal where the fringe visibility remains constant but where the overall optical throughput varies as a function of time. Fig. 7 outlines the basis of this technique. Both pinholes and pieces of single-mode optical fibre have been proposed as suitable spatial filters (for a critical comparison of the two the reader is referred to Keen *et al.* 2001) though to date most interferometric implementations have used fibre components only.

A useful strategy is to take advantage of both of these approaches. We can quantify how beneficial this might be by considering the signal-to-noise ratio (S/N) for low-light-level fringe amplitude measurements. In this regime the observable of interest is the square of the visibility amplitude, the S/N of which is plotted in the right-hand panel of Fig. 6 (Buscher & Shaklan 1994). The curves show how this S/N scales with telescope size for different levels of AO correction, both with (solid lines) and without (dashed lines) single-mode fibre spatial filtering. Several important results are evident:

- It is always beneficial to use a spatial filter unless the interferometric collector size is smaller than or comparable to r_o .
- Higher order adaptive correction is more beneficial than lower order correction.
- For any given order of correction there is an optimum telescope size, beyond which the S/N will *decrease*.
- For perfectly corrected wavefronts the $S/N \propto D$

To summarize then, reductions in fringe visibility and signal-to-noise ratio arising from spatial perturbations in the incoming wavefronts can be ameliorated by using a combination of active (AO) and passive (spatial-filtering) methods. The latter can be used for any type of observation, whilst the success of the former will rely upon the source being bright enough to act as a wavefront reference or the presence of a close bright natural guide-star. In this sense, the correct way of viewing adaptive correction is not primarily as a route to enhanced limiting sensitivity, but rather as a means of improving the sky coverage of interferometers and allowing observations of relatively bright targets to be undertaken more rapidly.

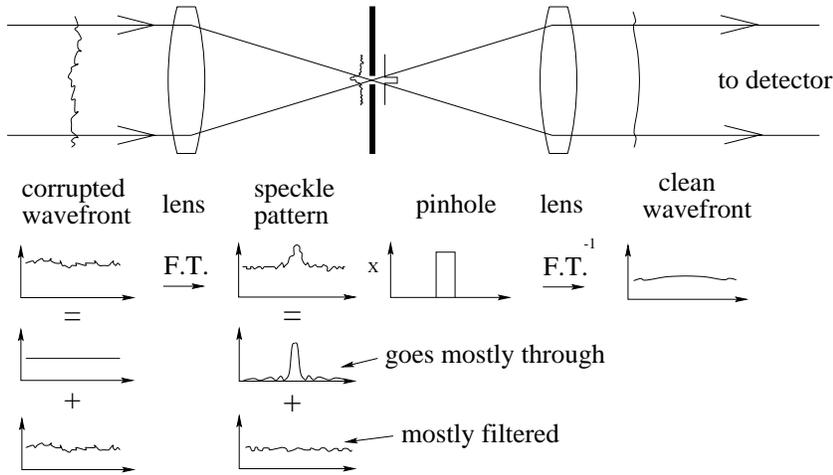


Fig. 7. A schematic illustration of how spatial filtering with a pinhole works. Three key ideas are represented: (i) the decomposition of the incident wavefront into perfect and disturbed components (ii) the origin of the focal plane fields as the Fourier transform of the fields in the aperture (iii) the action of the pinhole as a filter for the focal plane field distribution. The action of a piece of single-mode optical fibre is similar though with a transfer function described by the electric field profile within the fibre core.

6.2 Temporal fluctuations

The temporal fluctuations in the atmosphere are characterised in a similar fashion to the spatial fluctuations, i.e. we can define a coherence time, t_o , which measures the time over which the rms variation of the wavefront phase at a fixed point is one radian. The precise relationship between the spatial scale r_o and the coherence time is complex. One scenario pictures the temporal fluctuations arising from the wind-driven motion of a “frozen” layer of turbulence past the interferometer, while another has some type of in-situ “boiling” of the wavefronts taking place (see, e.g. St.-Jacques & Baldwin 2000 for a discussion of this point). For a frozen screen, where the wavefront evolution time is assumed to be much greater than the time for the screen to blow past the array, the coherence time can be written as $t_o = 0.314r_o/v$, where v is a characteristic wind velocity. Note that in this picture large scale spatial perturbations will be associated with timescales longer than the coherence time. Values of t_o typical of modern observatory sites are between 2 and 20 ms at 500 nm and scale with wavelength in the same manner as r_o . For the VLTI, recent data from Paranal give a median value of approximately 20 ms at $2.2\mu\text{m}$, or equivalently 3 ms at $0.5\mu\text{m}$ (Glindemann, priv. comm.). The small value of the coherence time, as compared to the exposure times use for conventional astronomical measurements, is the fundamental problem that ground-

based interferometrists have to grapple with.

From an interferometric perspective, the question we should be asking now is: how will these temporal fluctuations affect the measurement of the spatial coherence function? The answer to this is quite simple: since the fluctuations arise from changes in the optical paths along the lines of sight from the interferometric collectors to the source, they will alter the position of the white-light fringe. This has three quite distinct implications:

- Short-term fluctuations will move the interferometric fringes past the detectors and so any fringe measurements must be made on timescales shorter than t_o to limit this blurring.
- Longer timescale fluctuations will lead to large offsets in the position of the white-light fringe. Assuming Kolmogorov turbulence, it is straightforward to show that the rms optical path difference, σ_{opd} , for a baseline of length B will be equal to $0.417\lambda(B/r_o)^{5/6}$ (Davis *et al.* 1995). Hence, on all but the very shortest interferometer baselines the white-light fringe motion will likely exceed the coherence length of the radiation being observed. This implies a need for slow dynamic tracking of the white-light fringe motion.
- Unless the white-light fringe motion can be monitored absolutely, measurements of the phase of the interferometer fringes will no longer characterise the phase of the spatial coherence function.

The last of these points deserves special mention. The mathematical basis for interferometric imaging is the Fourier transform relationship between the coherence function and the sky brightness distribution. If we are unable to measure the phase of the coherence function then simple Fourier inversion becomes impossible! Fortunately, there are routes that circumvent this impasse. The first option is simply to track the atmospheric excursions dynamically at the sub-wavelength level. The measured fringe phase will then correspond to the phase of the coherence function. An alternative, though less direct approach, is to measure a phase-type quantity that is independent of the atmospheric fluctuations, for example a relative phase or a closure phase. This can then be used to infer information about the source structure. These two different approaches are outlined below.

6.2.1 Tracking the atmospheric fluctuations

The direct measurement, and subsequent correction of, the optical path fluctuations in an astronomical interferometer is generically referred to as fringe tracking. Under this general descriptor there are at least three levels of tracking that have either been employed or proposed for current arrays. The first of these simply ensures that fringe measurements are made close to the coherence envelope of the fringe pattern. This will not in itself guarantee that fringes are observable, but it will mean the region of delay space to be searched is manageably small. A second level of correction involves keeping any residual OPD a small fraction of the width of

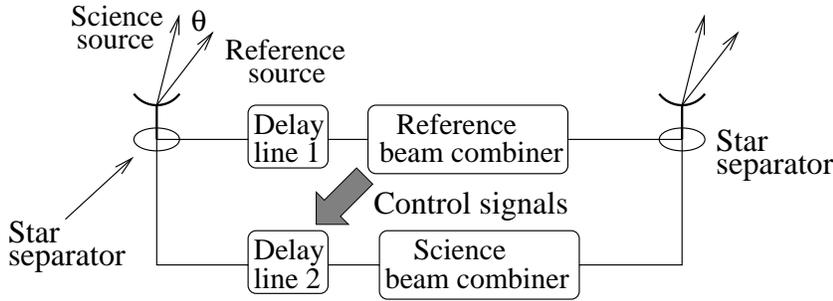


Fig. 8. A schematic layout of a dual-beam phase-tracking interferometer. The unit collectors simultaneously observe two sources, separated by a small angle θ . The fringes from the reference source are monitored in real-time and corrections fed to a separate delay-line which matches the optical paths for the beams from the science target. A high-precision internal metrology system, not shown in the diagram, is used to tie the optical paths of the source and reference signals together.

the coherence envelope. This is sometimes referred to as “coherencing” or “envelope tracking” in the literature. The most ambitious scheme is when the motion of the white-light fringe is tracked and corrected for in real-time at the sub-fringe spacing level. This is usually referred to as “phase tracking” and is exemplified at the VLTI by the PRIMA instrument (Delplancke *et al.* 2000). Only the last of these strategies allows for direct Fourier inversion of the visibility measurements, and so the first two approaches still require the measurement of observables that are resistant to residual fluctuations.

Nowadays a favoured approach to coherencing is to use group-delay tracking (see, e.g. Lawson 1995). This basically involves observing the interferometric fringes in dispersed light, mapping them onto a 2-dimensional space of wavenumber vs fringe phase, and examining the 2-dimensional “spatial frequency” of the resulting fringes. The value of this spatial frequency then measures the group delay, i.e. the location of the centre of the coherence envelope. The details of this method can be found elsewhere (Lawson 2000) but for our purposes there are two main points to note. First, the resolution of this method for measuring the group delay is related to the optical bandpass being dispersed. More precisely, the resolution in delay will be equal to the coherence length of the total bandpass, $\lambda^2/\Delta\lambda$. So, for example, for a tracker using the whole of the infrared K-band, the resolution will be $\sim 12\ \mu\text{m}$. Second, since it will take some time for the atmosphere to move the coherence envelope by this amount, this method allows for incoherent integration of the delay signal for many coherence times. For example for the 20% bandpass of the K-band an integration time of order $30t_o$ would be suitable.

This method can be contrasted with phase tracking, where the phase of the white-light fringe is measured in a time short compared to the coherence time,

and the fringe motion corrected for instantaneously. This requires fast enough sampling and a high enough instantaneous signal-to-noise ratio so that the change in fringe phase between successive exposures is $\leq \pi$ radians. Failure to satisfy these demands will produce errors in “unwrapping” the measured phase perturbations and lead to tracking failure. In comparison to group-delay tracking, however, this method does allow sub-wavelength stabilization of the interference fringes, and so, in principle, the coherent integration time can be increased without limit. The downside of this approach is the short allowable exposure time ($t_{\text{exp}} \leq t_o/2$) and hence its poorer limiting sensitivity (a factor of ~ 2.5 magnitudes) as compared to group-delay tracking.

Interestingly, there is a close relationship between our earlier discussion of adaptive optics and our review of the pros and cons of phase tracking. Since both are active real-time processes, both rely on the presence of a suitable reference source to provide the feedback signal to drive the correcting device. This signal can come from the science target itself, perhaps using light in a broadband channel adjacent to the science bandpass, or from a nearby reference star. And, if the latter approach is taken, this reference needs to be close (typically within a few tens of arcseconds at $2.2 \mu\text{m}$) to the science target. The major difference between adaptive optics and phase-tracking becomes apparent only when an off-axis reference source is being used. In AO systems the off axis beam is usually easy to deal with, because the telescope/instrument field-of-view is usually much greater than the off-axis offset angle; this is not the case for a long-baseline interferometer. Here, the light from the reference target must be separated from that of the science object at the interferometric collectors and then be delivered to a separate beam combiner in the optical laboratory using an additional differential delay line (see Fig. 8). Moreover, the optical paths of the target and reference sources through the interferometer must be monitored with high precision so as to allow correction for any internal differential path fluctuations. The cost and complexity of this extra hardware, together with an additional requirement that the reference source not be resolved by the interferometer baseline (hence precluding the use of laser guide stars) means that the implementation of off-axis phase tracking, usually referred to as “dual-feed” interferometry, is a non-trivial exercise.

It seems likely then that in any given situation observers will have to choose between a number of possible options for coping with the temporal fluctuations introduced by the atmosphere. For the brightest targets on-source phase tracking should be possible, while for fainter objects either self-referenced envelope tracking or off-axis envelope or phase tracking may be available. The presence, or not, of a suitable off-axis reference source will be a critical factor.

6.2.2 Measuring “good” observables

Whether or not any type of fringe tracking is taking place, ground-based interferometers almost always exploit an additional strategy when information about the phase of the coherence function is required. This involves measuring quantities that are related to the coherence function phase, but which themselves are not

perturbed by the atmosphere.

Consider, for example, making simultaneous measurements of the visibility function of a source on a given baseline at two similar wavelengths. In the absence of a phase tracker, the instantaneous fringe phase at each wavelength will be equal to the true visibility phase plus an unknown phase error associated with the atmospherically-induced OPD. However, because the atmospheric optical path fluctuations are to first order achromatic — the refractive index of air varies only slowly with wavelength — the errors at one wavelength will be related those at any other. Hence, the phase error measured at one wavelength can be used to correct for the phase error at another. This method is most useful when the source is known to be unresolved at one wavelength (the reference wavelength), so that its visibility phase is known to be zero there. In this case, the difference in the measured fringe phases becomes a proxy for the true visibility phase at the science wavelength and the effects of the atmosphere can be removed. This “differential phase” technique (Stee, these proceedings) can in fact be used whatever the source structure is. All that is required is that the visibility phase be known at some reference wavelength. However, the success of this method will be limited by the precision with which the dispersive effects of the atmosphere can be modelled.

Similar “resistance” to the atmospheric fluctuations can be achieved by measuring visibility phases simultaneously on multiple interferometer baselines, where the baselines in question are formed by traversing a closed path between at least three telescopes (see Fig. 9). Under these conditions, the sum of the observed visibility phases round the loop, the so-called “closure phase”, is also a robust observable. The atmospheric errors on the individual visibility phases cancel out in this linear combination and the resulting quantity is immune to atmospheric perturbations and solely dependent on the source morphology. It is worth mentioning that the number of independent closure triangles, N_c , that can be measured with an N -element interferometer is equal to $(N - 1)(N - 2)/2$ so that the fraction of the visibility phase information retained in the closure phases ($F = 1 - \frac{2}{N}$) rises rapidly as N increases. The increasing priority given to the number of array elements in modern interferometer designs, in large part, stems from this result. Exactly how closure-phases are interpreted and used for interferometric science is covered in more detail in the contributions from Monnier and Buscher in these proceedings.

Before concluding, it will be useful to compare and contrast the differential and closure phase methods. Both are self-referenced techniques, both rely upon simultaneous measurements of different perturbed fringe phases, and both provide a signal that can be coherently integrated over many integration times. On the other hand, differential phase methods are most useful when the source visibility function is known at some wavelength, whereas the closure phase technique is wholly independent of the source morphology. Closure methods thus provide a powerful model-independent way of eliminating certain classes of perturbations, so-called antenna-dependent complex gain errors. The temporal fluctuations of the atmosphere give rise to just such errors but other types of error can, and do, exist and can only be overcome using other methods.

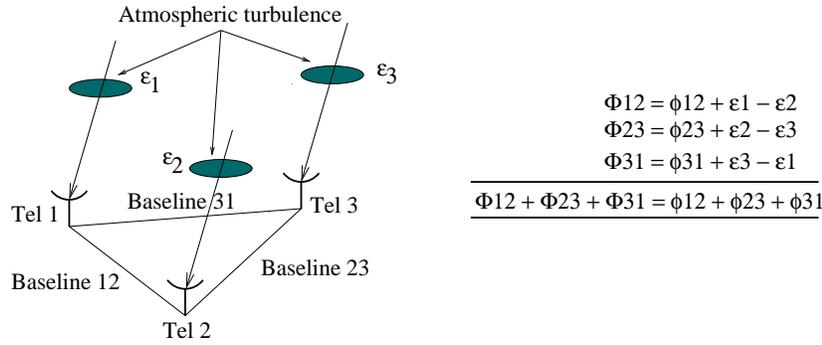


Fig. 9. The basic principle of closure phase measurement as applied to a simple 3-element interferometer. Each measured visibility phase, Φ_{ij} , is equal to the true coherence function phase, ϕ_{ij} , to which have been added phase errors, ϵ_i and ϵ_j , associated with the unknown optical paths above telescopes i and j . All of these terms cancel out when the measured visibility phases are summed “round the loop”. This produces the closure phase, a phase-type quantity that only depends on the source coherence function.

7 Sensitivity

In the previous lecture we introduced the idea of quantifying the sensitivity of an interferometric array by asking two parallel questions of a given source:

- Is it bright enough to provide a signal with which to stabilize the array against temporal, and possibly spatial, atmospheric wavefront perturbations?
- Is it bright enough that the integrated signal-to-noise ratio on the visibility amplitudes and differential/closure/visibility phases is high enough after some moderate integration time (perhaps 5 minutes or so) to be useful.

The first of these constraints is the one I want to concentrate on here since if fringe stabilization has failed it is unlikely that any useful faint source data will be secured. Fringe tracking will thus be a necessary, but not sufficient, condition for useful science. In the low-signal limit, the situation that we will actually be interested in analysing is the limiting performance of envelope tracking — the faintest sources will clearly be too faint to drive any phase-tracker. The detailed derivation of the expression for the signal-to-noise ratio for envelope tracking is beyond the scope of this introductory discussion, but its behaviour will be of some interest to us. Broadly speaking it will take the form of a power-spectrum signal-to-noise ratio:

$$(S/N) \propto \frac{[VN]^2}{\sqrt{[(N + N_{\text{dark}})^2 + 2(N + N_{\text{dark}})N^2V^2 + 2(N_{\text{pix}})^2(\sigma_{\text{read}})^4]}}, \quad (7.1)$$

where V is the apparent fringe visibility, ranging between 0 and 1, N is the total number of photon counts detected from the source in the integration time, N_{dark} is the number of dark or background counts detected in the integration time, N_{pix} is the number of detector pixels over which the fringe signal is spread out, and σ_{read} is the readout noise per pixel. For photon-limited detection ($N_{\text{dark}} = \sigma_{\text{read}} = 0$) the signal-to-noise ratio will become:

$$(S/N) \sim \frac{[VN]^2}{\sqrt{[N^2 + 2N^3V^2]}} \quad (7.2)$$

$$\sim [V^2N]^\alpha, \quad (7.3)$$

where α is in the range 0.5 – 1.0. For useful tracking we will clearly require $(S/N) > 1$.

The expressions presented above deserve very careful attention. For those who are interested in either designing interferometers or using them for faint source science, there are at least four points to note:

- The use of array detectors with high readout noise brings with it a significant signal-to-noise penalty.
- If an array detector with finite readout noise is used, then the fringe signal should be detected with as small a number of pixels as possible.
- Losses in visibility will have a more important effect on the signal-to-noise ratio as compared to the same fractional loss in photon rate.
- Sources that are significantly resolved ($V \ll 1$) will be much more difficult to observe than point-like sources.

This last point is perhaps best demonstrated through an numerical example. Imagine an astronomer is observing a faint Cepheid variable with a two-element interferometer equipped with a photon-limited fringe sensor and is using light from the source to track the motion of the coherence envelope. She is interested in measuring the diameter of the star, and so chooses to observe it twice, once with a baseline where the star is only moderately resolved ($V_{\text{source}} = 0.75$) and again on another baseline where the visibility has dropped to 0.15. If we assume that envelope tracking on the short baseline can be performed with a S/N of say 3, then she will find that on the longer baseline successful tracking will not be possible. More precisely, the signal-to-noise ratio will be lower by a factor of ~ 6 which corresponds to an effective reduction in the source brightness by a factor of 25 (3.5 magnitudes). If the source is even more resolved on the longer baseline, say $V_{\text{source}} = 0.05$, the loss in S/N will be a factor of ~ 40 , i.e. an effective decrease in the source flux of 6 magnitudes! So, in this case, a Cepheid that is observable on a short baseline will be between 3 and 6 magnitudes too faint to be observed on exactly those baselines that can put good limits on its diameter.

It is this that is the real challenge for ground-based interferometry: how to observe faint and *resolved* targets. Those that are unresolved will unlikely be interesting — the role for an interferometer is after all to resolve their structure —

while those that are resolved may not be bright enough to stabilise the array. Perhaps we can console ourselves with the fact that the faintest sources will probably be too small to resolve with today's generation of interferometers anyhow: a 0.5 milliarcsecond diameter blackbody with a temperature of 2500 K will have an apparent magnitude in the near-infrared of between 7 and 8. This is ~ 5 magnitudes brighter than the group-delay tracking limit for an interferometer employing 2-m-class collectors, and so sources with surface brightnesses 10^2 times lower should still be observable.

8 Calibration strategies

The final aspect of optical/infrared interferometry that I wish to touch on briefly in this introduction is the role of calibration. We have already heard that the modulation and phase of the fringes detected at the end of the long optical train depicted in Fig. 1 should encode the amplitude and phase of the spatial coherence function of the source being observed. And in principle, the reliability of these data should be governed by their signal-to-noise ratios, i.e. depend on factors such as the exposure time, source magnitude, collecting area and optical throughput of the array. In practice, however, most interferometric studies are usually limited by systematic errors rather than the intrinsic S/N of the data. Uncertainties in the interferometric response, i.e. the effective complex transfer function of the experiment, are the underlying problem. These can arise from numerous effects, but in the main are traceable to changes in the local seeing conditions which feed through to all of the real-time control systems. In a very real sense then, the art of optical/infrared interferometry is to design both an array and an observing strategy that delivers reliable interferometric data in the presence of fluctuating observing conditions.

Among the techniques that have been used in the past several are worth commenting on:

- Measurement of sources that have known coherence functions. This usually means observing unresolved sources close in both time and space to the science target. The difficulties associated with finding suitable calibrator stars, which themselves should be of comparable brightness to the science target, are dealt with elsewhere in these proceedings (Boden 2002). Suffice to say they are considerable!
- Measurement of quantities that are less easily disturbed by systematic errors. The best example of this is the use of phase, rather than amplitude, data. Atmospheric fluctuations will generally only ever reduce the measured visibility amplitude of a source, whereas they will not bias any phase measurement. Differential or closure-phase techniques are thus much more robust against atmospherically induced systematics, though they can be compromised by instrumental biases, e.g. baseline dependent correlator errors.
- Careful design of instruments so as to make them only weakly sensitive to

changes in the seeing. This is exemplified by spatial filtering which, as we have seen, trades fluctuations in fringe visibility for fluctuations in throughput, these being much easier to treat in the subsequent data analysis.

I think the lesson to take home here is thus relatively simple: don't forget the need to calibrate your interferometric measurements, as these may ultimately limit what science you can do. But then again, capitalize on the techniques that others have developed and always know how small you *need* your error bars to be.

9 Concluding remarks

Once again this has been a long lecture, and today it has mainly focused on the large number of potential complications associated with securing scientifically valuable interferometric measurements. What I hope you have picked up is a measure of the complexity of an interferometric array, some feeling for the issues that delimit its possibilities, and an appreciation of the ingenuity of the engineers and physicists who have been responsible for putting them together. However, the lesson to be learnt is not that interferometry is difficult, but rather that all of these problems are indeed soluble.

The reality is that optical/infrared interferometry with arrays like the VLTI is a valuable and moreover, relatively unexplored tool for astrophysics. With 80 papers in the refereed literature in the past 2 years alone, there is little doubt that ground-based interferometry has reached some level of success and is delivering science. The position you are in today, on the verge of the era of modern facility arrays, is thus a unique one. You may choose to turn away from this opportunity, but I hope that the remainder of this winter school will encourage you to press ahead. There is no question that the road ahead will be difficult, from both the experimental and theoretical point of view, but I cannot believe the journey will be anything but intellectually challenging, scientifically rewarding and ultimately valuable for contemporary astrophysics.

References

- Berger, J.P., Haguenaer, P., Kern, P., Perraut, K., Malbet, F., Schanen, I., Severi, M., Millan-Gabet, R. & Traub, W. 2001 *A*, 376, L31
- Buscher, D.F. 1988, *MNRAS*, 235, 1203
- Buscher, D.F. & Shaklan, S.B. 1994 *Proc. SPIE*, 2201, 980
- Coudé du Foresto, V. & Ridgway, S.T. 1992, in *“High Resolution Imaging by Interferometry II”*, ed. J.M. Beckers & F. Merkle (European Southern Observatory: Garching, Germany), p. 731
- Davis, J., Lawson, P.R., Booth, A.J., Tango, W.J. & Thorvaldson, E.D. 1995 *MNRAS*, 273, L53
- Delplancke, F., Leveque, S., Kervella, P., Glindemann, A. & d’Arcio, L. 2000 *Proc. SPIE*, 4006, 365
- Horton, A.J., Buscher, D.F. & Haniff, C.A. 2001 *MNRAS*, 327, 217

- Keen, J.W., Buscher, D.F. & Warner, P.J. 2001 MNRAS, 326, 1381
- Lawson, P.R. 1995, JOSA-A, 12, 366
- Lawson, P.R. 2000, editor, "*Principles of Long Baseline Stellar Interferometry*", JPL Publication 00-009 (Pasadena: NASA)
- Lawson, P.R. & Davis, J. 1996 Appl. Opt., 35, 612
- Mozurkewich, D. 1994 Proc. SPIE, 2200, 76
- Perrin, G., Lai, O., Lena, P.J. & Coude du Foresto, V. 2001 Proc. SPIE, 4006, 708
- Roddier, F. 1981 Prog. Optics 19, 281
- St.-Jacques, D. & Baldwin, J.E. 2000 Proc. SPIE, 4006, 951
- Shaklan, S.B. & Roddier, F. 1988 Appl. Opt., 27, 2334
- Tango, W.J. 1990 Appl. Opt., 29, 516
- Traub, W.A. 1988, in "*High Resolution Imaging by Interferometry*", ed. F. Merkle (European Southern Observatory: Garching, Germany), p. 1029