Studies of Evolved Stars with a Mid-Infrared Interferometer

by

Everett Arthur Lipman

B.A. (University of Chicago) 1991
M.A. (University of California at Berkeley) 1993

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Physics

in the

GRADUATE DIVISION
of the
UNIVERSITY of CALIFORNIA at BERKELEY

Committee in charge:

Professor Charles H. Townes, Co-Chair
Dr. William C. Danchi, Co-Chair
Professor Sumner P. Davis
Professor William J. Welch

Fall 1998

The dissertation of Everett Arthur Lipman is approved:

Wm. H. Townes — Sept. 16, 1998
Co-Chair — Date

William C. Dauchi — Sept. 16 1998
Co-Chair — Date

Sumner P. Davis — Sept. 16, 1998
Date

Wm C Welch — Sept. 17, 1998
Date

University of California at Berkeley

Fall 1998

Studies of Evolved Stars with a Mid-Infrared Interferometer

Contents

Chapter 1

Introduction

The clear night sky is fascinating to behold. At a remote site, unpolluted by artificial lighting, one can see as many as a few thousand stars. These points of light, with their many colors and brightnesses, evoke an undeniable desire for more knowledge of their nature. During the last few hundred years, mankind has been turning ever more powerful optical instruments toward the sky in order to discover where the stars are, what they are made of, how big they are, and ultimately, how they live and die.

This dissertation describes observations of evolved stars which were made with the U.C. Berkeley Infrared Spatial Interferometer (ISI). The ISI is a special combination of two telescopes which is able to obtain information with very high angular resolution in the mid-infrared region of the electromagnetic spectrum. This range of wavelengths, around 10 μm, is where the thermal radiation from room-temperature objects is at its maximum. Evolved stars, those in the late stages of the stellar life cycle, are often surrounded by shells of relatively cool dust, and thus make ideal targets for study with the ISI.

1.1 Limitations of Telescopic Imaging

There are two primary reasons to use optical instruments to study the stars. The first is to gather more light. A typical dark-adapted human pupil is only 5 mm in diameter. By using binoculars with 50 mm objective lenses, one increases the amount of light in the retinal image by a factor of $(50/5)^2 = 100$. For any given detector (the cells of the retina, or a silicon charge-coupled detector, for example), increasing the amount of light

gathered will increase the number of objects available for study, and improve the signal-to-noise ratio on those already visible. Because of this advantage, larger telescopes are preferable, all other factors being equal.

The second reason to use optical instruments is to improve resolution. Because of the wave nature of light, any point in an object being viewed will be mapped, even in a perfect optical instrument, to a diffraction pattern in the image. In the common case of an instrument with a circular entrance pupil, this will be the Airy pattern (pictured in Fig. 1.2). The result of this effect is a fine-scale blurring, since distinct points with small separations in the object will become overlapping disks with faint rings surrounding them in the image.

The light intensity of the central maximum (known as the Airy disk) of the Airy pattern drops to zero at an angular distance of $1.22\lambda/D$ radians from the center,[1] where λ is the wavelength of the light forming the image, and D is the diameter of the instrument's entrance pupil. The diffraction-limited resolution angle of a telescope is usually given as this same angle, $1.22\lambda/D$. This is the Rayleigh criterion, which states that the smallest angle which can separate two resolved points in an object is that angle for which the edge (first intensity null) of the Airy disk due to one point in the image will coincide with the center of the other disk. Although this definition is somewhat arbitrary, it provides a useful measure for comparing optical systems. Using this criterion, the resolution of an eye with a 5 mm pupil, viewing green light,[2] would be about 30 arc seconds ($''$), where $1''$ is 1/3600 of a degree of arc. By comparison, 20/20 vision is defined as the ability to read letters which are $300''$ tall with $60''$ features. At a distance of 17 mm, the approximate focal length of the eye, $30''$ corresponds to 2.5 μm. Notably, the light-sensitive rod and cone cells in the retina have collecting areas of approximately 1 and 4 μm^2, respectively. Thus these cells are just small enough to take advantage of the best possible resolution which the eye can have.

While a few arc seconds might seem to be a very small angle, and not a serious limitation to useful resolution, the enormous distances to the stars result in

[1] The light intensity in the Airy pattern has the functional form $\left(\frac{J_1(\xi)}{\xi}\right)^2$, where J_1 is a Bessel function of first order, and $\xi \equiv (\frac{\pi}{\lambda})D\sin\theta$. The first zero of $J_1(\xi)$ occurs at $\xi = 3.832$, so that $3.832 = (\frac{\pi}{\lambda})D\sin\theta_0$, or $D\sin\theta_0 = 1.22\lambda$. The angular distance from the center of the pattern, θ, is normally extremely small (around 5×10^{-7} rad at optical wavelengths with a 1 m aperture), so $\theta_0 \approx 1.22\lambda/D$. A derivation of this diffraction pattern is given by Lipson and Tannhauser [59].

[2] The eye's sensitivity peaks near the center of the visible spectrum, at a wavelength which corresponds to green light.

their having minuscule angular diameters. Figure 1.1 illustrates the sort of angle which must be resolved in order to see a star as a disk rather than a point of light. Even the

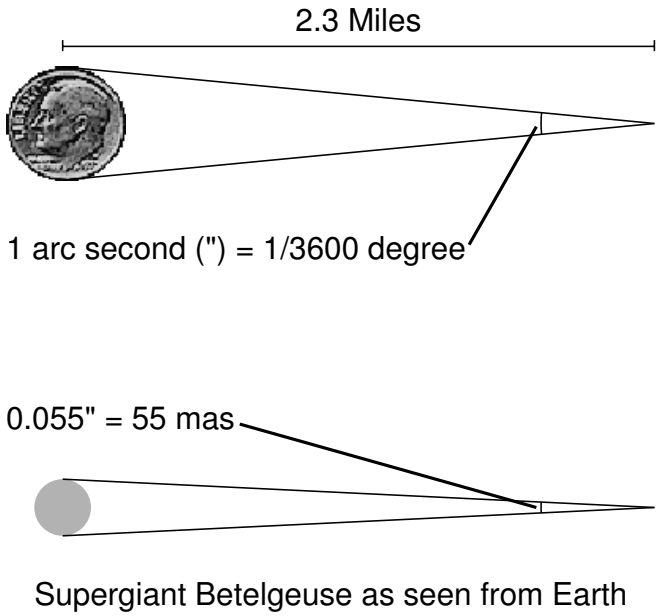


Figure 1.1: Illustration of small angles. As shown in the top part of the figure, $1''$ is the angle subtended by a dime at 2.3 miles. Betelgeuse (α Orionis), a star with one of the largest angular diameters, subtends 55 thousandths of an arc second as seen from Earth.

supergiant Betelgeuse, which is significantly larger than Jupiter's orbit, subtends only 55 mas (thousandths of an arc second) as seen from Earth. Using the Rayleigh criterion, we can quickly compute that in order to resolve, using visible light, two points with the angular separation θ of Betelgeuse's diameter, we would need a telescope with diameter $D \geq 1.22\lambda/\theta = 2.5\,\mathrm{m}$.

Although there are a number of such telescopes, long before we reach this size, another limitation comes into play. For any telescope with a diameter greater than about 25 cm, distortion of the visible-light image by the atmosphere is more of a limitation than diffraction. This situation is illustrated in Fig. 1.2. Recently, adaptive optics systems have been developed which correct for atmospheric fluctuations and allow large telescopes to enjoy nearly diffraction-limited performance.

Although atmospheric turbulence can be avoided by launching a telescope into orbit, or corrected with adaptive optics, diffraction presents a more fundamental

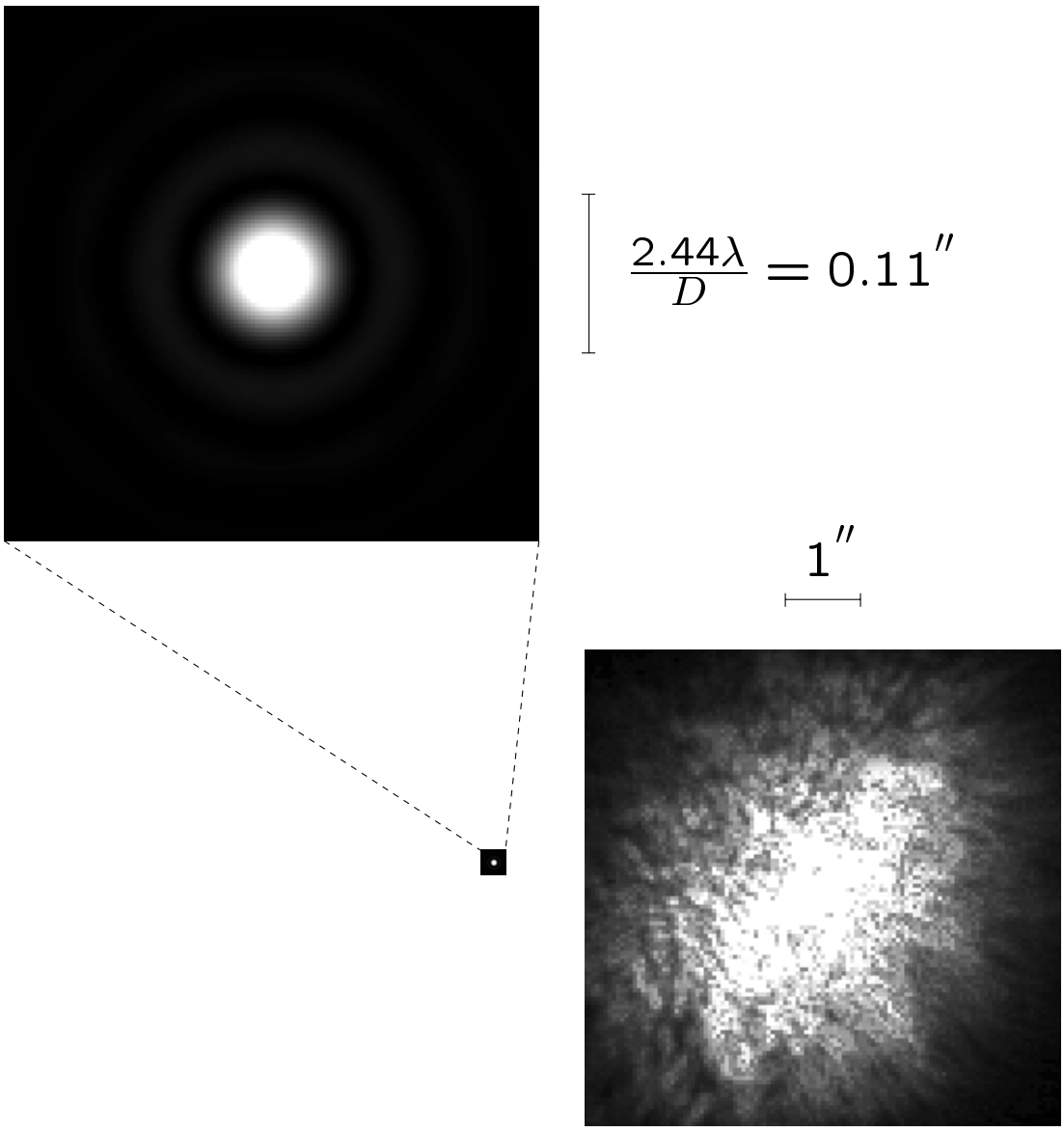


Figure 1.2: Limitations of telescopic imaging. At the top is the diffraction disk which would result from the imaging of a perfect point source in 2.2 μm light by a 10 m telescope. Below, the diffraction disk is reproduced, to scale, next to a real image of a binary star taken at 2.2 μm on the 10 m Keck telescope. For any telescope with an aperture larger than about 25 cm, blurring of the visible-light image due to the atmosphere is more of a limitation than diffraction.

problem. It is not currently possible to build telescopes much larger than 10 m in diameter. Because of this engineering limitation, we must use other methods in order to obtain resolutions higher than $0.01''$.

It was demonstrated late in the nineteenth century that by using optical interference, one can, to some extent, overcome both factors which limit the resolution of ordinary telescopes. This technique of stellar interferometry, described in chapter 2, has been used since the early 1920s to enhance our understanding of the stars. The ISI is one of a number of interferometers presently obtaining information about stars and their surroundings which would not otherwise be available.

1.2 Evolved Stars

As was mentioned earlier, evolved stars, with their shells of relatively cool dust, are ideal targets of observation for the ISI. Such stars are also interesting from a theoretical point of view, since our understanding of the late stages of stellar evolution is still incomplete.

Figure 1.3 illustrates the life cycle of a typical star, such as our own Sun. A complete qualitative description of this life cycle can be found in *The Physical Universe* by Shu [83], and a more technical treatment is given by Hansen and Kawaler in *Stellar Interiors* [35]. In the figure, known as a Hertzprung-Russell diagram, the luminosity of the star (in units of solar luminosity $L_\odot$) is plotted against its effective temperature. The effective temperature is the temperature which a blackbody of the star's size would need to have in order to match the luminosity of the star. When plotted on this diagram, most stars fall within a band known as the main sequence. Position in the main sequence is determined by stellar mass.

Most of the life of a star is spent in the main sequence, as the initial supply of hydrogen is slowly fused into helium at the star's core. When this supply of hydrogen runs out, the core will contract, and fusion will begin in a shell around the helium ash. Helium produced by this "shell burning" accumulates in the core, the additional mass causing further contraction. This contraction will increase the gravity at the surface of the core, and the temperature and density of the shell will rise, enhancing the fusion reaction. Consequently, the star begins to expand, and its surface temperature drops, since the increased luminosity is not sufficient to maintain the larger surface area at the original temperature. These changes can be seen on the diagram as the initial departure

2.2 Theoretical Background

2.2.1 Coherence Length and Starlight

At first, it might seem counterintuitive to use an interferometer to study light from an incoherent, broad-band source such as a star. If one combines the light from two flashlights, the result will still be uniform illumination, with the intensity equal to the sum of the intensities of the two sources. Likewise, sunlight combined by reflection does not normally exhibit interference effects. There are, however, situations where interference effects are created by "white" light. An example of such a situation is the occurrence of brilliant colors in soap bubbles and thin films of oil floating on water. The condition necessary in all systems producing interference fringes from light is that the paths taken from the source to the point of interference by the interfering beams of light must be matched to within a distance l_c, the *coherence length*. For sunlight, $l_c \approx 1\,\mu$m, so in order to see interference, we must have a pathlength difference, such as a round trip across the thickness of a soap bubble, of not much more than $1\,\mu$m.

The light from a flashlight or a star is produced in the form of many photons, each resulting from the transition of a thermally-excited electron. Unlike photons from a laser, these are correlated neither in phase nor in frequency. We can represent a photon with an idealized harmonic pulse as shown in Fig. 2.1. This pulse has the following form:

$$f(t) = \begin{cases} \cos(2\pi\nu_0 t) & -\frac{1}{2}\tau < t < \frac{1}{2}\tau \\ 0 & \text{elsewhere,} \end{cases} \tag{2.1}$$

where τ, the time over which the pulse maintains a continuous phase, is called the *coherence time*. If we were to try to produce an interference pattern with many pulses of this form, for example by passing them through a beam splitter and recombining them on a screen, clearly the two parts of a split pulse would need to arrive at the screen with a relative delay of no more than τ in order for interference to occur. The coherence length l_c is defined as the distance over which a photon will travel during its coherence time, so

$$l_c = c\tau, \tag{2.2}$$

where c is the speed of light.

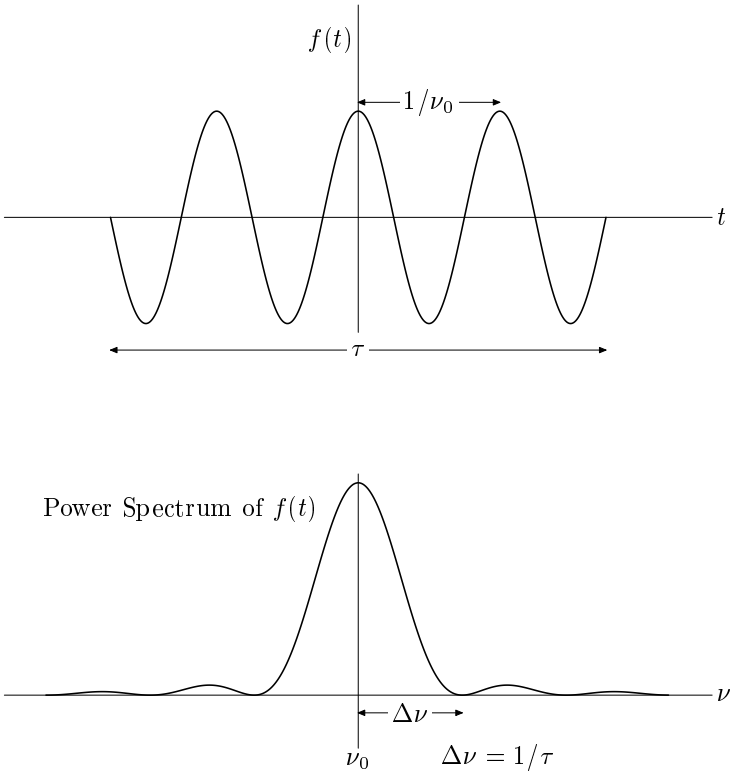


Figure 2.1: Power spectrum of a finite harmonic pulse. Pictured here is a harmonic pulse with duration τ and frequency ν_0. The power spectrum of the pulse has two identical components. One, shown here, is centered about ν_0, and the other is centered about $-\nu_0$. The first maximum in the power spectrum of the pulse has a half-width at its base of $\Delta\nu = 1/\tau$.

We can relate τ, and hence l_c, to the bandwidth of the light by examining the power spectrum of the pulse (Bracewell [17]), which is:

$$g(\nu) = A\,\mathrm{sinc}^2\left[\pi(\nu \pm \nu_0)\tau\right], \tag{2.3}$$

where $\mathrm{sinc}(x) \equiv \sin(x)/x$, and A is a normalization constant, chosen such that the power of the pulse is the same in the time and frequency domains. The half-width $\Delta\nu$ of the central maximum of $g(\nu)$ at its base (Fig. 2.1) is usually chosen to represent the spread of frequencies contained in the pulse (as, for example, in Born & Wolf [14]). From (2.3) we can see that $g(\nu) = 0$ when

$$\Delta\nu = \nu - \nu_0 = 1/\tau. \tag{2.4}$$

Substituting from (2.2), we have

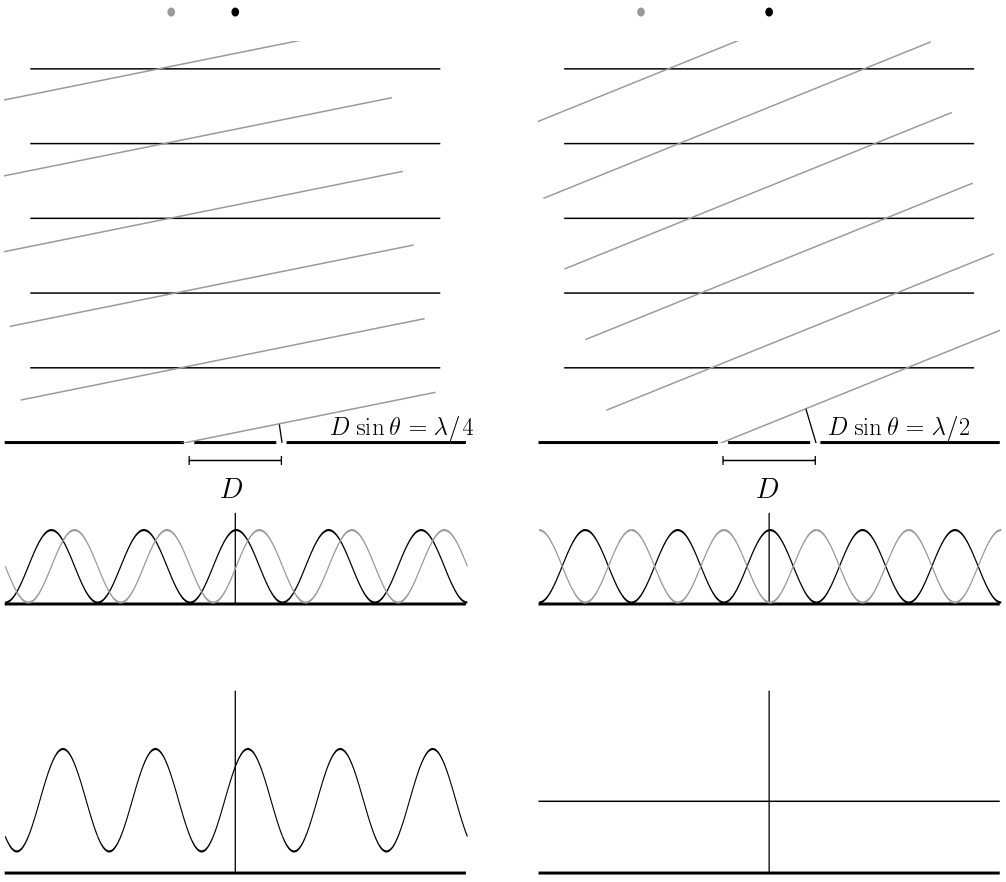
$$l_c = c\tau = c/\Delta\nu. \tag{2.5}$$

Thus there is an inverse relationship between the bandwidth and coherence length of light from a given source. For visible sunlight, which contains wavelengths between 400 and 700 nm, $\Delta\nu \approx 3 \times 10^{14}$ Hz, so $l_c \approx 1\,\mu$m. Typical "monochromatic" light, as from a multimode laser with a bandwidth of 100 MHz, would have a coherence length $l_c = 3$ m.

As will be described in §2.2.2, a simple stellar interferometer consists of two separated telescopes. The light collected by the telescopes is combined to produce fringes, and from the contrast (also known as "fringe visibility") of these fringes, we can recover high resolution information about the brightness distribution of the source on the sky. In order to obtain fringes and make an accurate measurement of the visibility, the interferometer must ensure that the paths taken by a stellar wavefront through the two telescopes to the point of combination are equal to within a fraction of l_c. This is accomplished by placing a variable delay in one of the arms of the interferometer, as pictured in Fig. 2.8. The detection bandwidth of the ISI is about 6 GHz ($l_c \approx 5$ cm), so the delay line, which is a computer-switched bank of cables, must be continually adjusted to maintain the equality of the signal paths to within about 1 cm as the Earth's rotation changes the direction to the star, and hence the geometric delay (Fig. 2.8).

2.2.2 Interference Fringes and Fringe Visibility

The basic principle of stellar interferometry is illustrated in Fig. 2.2. The stellar interferometer is represented as a two-slit experiment, with each slit corresponding to one of the telescopes. The distance between the telescopes, D, is referred to as the *baseline*. The sources pictured are pairs of points, each being smaller than the resolution of the interferometer. They could be thought of as closely spaced binary stars, for example. Although the spacing in the figure has been exaggerated for clarity, it should be understood that *the points comprising each source are separated by an angle θ too small to be resolved by the individual telescopes*. Some typical numbers for the ISI would be $D = 32$ m, $\lambda = 11.15\,\mu$m, and $\theta = 0.036''$. Each ISI telescope has a diameter of 1.65 m, so one can compare θ to the resolution of each telescope, which is given by the Rayleigh criterion (§1.1) as $1.7''$. The distances to all stars other than the Sun are so



$$\text{Fringe Visibility: } V_{\text{M}} \equiv \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}$$

$$\text{Resolution: } \theta_{\min} = \lambda/2D$$

Figure 2.2: Interference fringe visibility. Each of the two point sources, which are separated by an angle θ, produces a fringe pattern when received by the two-slit interferometer. Since the sources are mutually incoherent, the intensities of the individual fringe patterns add to yield the final pattern. The angles and the light wavelength are greatly exaggerated for clarity. See the text for a description of the fringe visibility and resolution.

enormous that the light from any point on a star arrives as a plane wave (excepting for atmospheric distortion, which will be discussed in §2.2.5).

Considering the components of the sources separately for the moment, we can see that the plane wave due to each unresolved point will form an interference pattern in the two-slit experiment. The intensities of the individual patterns are shown directly below the slits in Fig. 2.2. Because these individual patterns are formed by mutually incoherent points, there will be no interference between the two patterns, but rather, *the intensities of the individual fringes will add incoherently to produce the observed pattern*. This is analogous to what happens when two flashlights illuminate the same area. The intensities add, but there is no interference. The observed pattern, which is what would actually be seen on the screen of such a two-slit experiment, is shown below the individual patterns.

Even though the two source components are mutually incoherent, a fringe pattern is still observed in the example on the left-hand side of Fig. 2.2. One can easily see that if the source points were spatially coincident, the individual patterns would have the same phase, and the resulting sum would have full swings between zero and maximum intensity. If we observe sources with components that are separated by an angle θ, the contrast of the observed fringe pattern will decrease until the point shown on the right-hand side of Fig. 2.2, where uniform illumination is observed.

In order to express the degree of fringe contrast, Michelson defined the *fringe visibility*

$$V_{\rm M} \equiv \frac{I_{\rm max} - I_{\rm min}}{I_{\rm max} + I_{\rm min}}, \tag{2.6}$$

where $I_{\rm max}$ and $I_{\rm min}$ are the maximum and minimum intensities in the observed fringe pattern, respectively. When $V_{\rm M}$ drops to zero, the source is said to be resolved. As shown on the right-hand side of Fig. 2.2, $V_{\rm M} = 0$ when $\sin\theta = \lambda/2D$. Using the small-angle approximation,

$$\theta_{\rm min} = \lambda/2D, \tag{2.7}$$

where $\theta_{\rm min}$ is the angle taken to be the resolution limit for the interferometer. Comparing $\theta_{\rm min}$ to the Rayleigh criterion, we can see that the formal resolution of an interferometer is roughly the same as that of a telescope with radius D. Since the largest presently operating infrared telescope has an effective radius of 5 m, an interferometer such as the

ISI has a significant resolution advantage at a given wavelength. With a 32 m baseline, the ISI has a resolution $\theta_{\rm min} = 0.036''$.

An important feature of the observed fringe shown on the left-hand side of Fig. 2.2 is the offset in phase from the center of symmetry of the two-slit experiment. If the source were centered about the axis of symmetry, the fringe would also be centered, with zero phase. In actual practice, the phase of the fringe constantly changes during an observation. The rotation of the Earth moves the source with respect to the interferometer, sweeping the fringe past the center of symmetry. If a photodetector were placed so that it measured the intensity in a small area ($\ll$ one cycle) of the fringe pattern, its output would be an oscillation at the *natural fringe frequency*, which is determined by D, λ, the position of the star, and the rotation rate of the Earth. The natural fringe frequency for the ISI would be about 210 Hz for a 32 m east-west baseline with the star due south on the celestial equator. It turns out that atmospheric fluctuations (§2.2.5), which introduce a random phase difference between the telescopes, make it difficult to measure the absolute phase of the fringe, and with a two-element interferometer, this information is lost. As will be discussed in §2.2.3, lack of phase information requires that some assumptions be made about the symmetry of the source if its brightness distribution is to be reconstructed.

Naturally, it is desirable to measure a wide range of sources with an interferometer, not just those whose angular size happens to coincide with a single combination of wavelength and baseline, as in Fig. 2.2. Recalling that $\theta_{\rm min} = \lambda/2D$, we can see that a source with a given angular size will be resolved if the baseline is increased until this equation is satisfied. In fact, in Michelson's original experiment, this is exactly what was done. A series of measurements was made with the interferometer mirrors at various separations until the fringe visibility dropped to zero, indicating the size of the source (which was the star α Orionis). Measurements at a wide range of baselines are made possible with a fixed interferometer by the technique of *Earth rotation synthesis*. Although the main collecting mirrors of an interferometer are separated by a baseline D, if the direction to the source is not perpendicular to the baseline, the resolution for a given measurement is determined by the *effective baseline* D_e. This is the projected baseline as seen from the direction of the source (Fig. 2.8). By situating the telescopes on an east-west baseline, the observer can make measurements with baselines D_e ranging from D when the source is on the meridian, all the way down to some small fraction

of D as it approaches the horizon. In principle, as the Earth rotates during the course of the night, the interferometer will measure the source's *visibility curve*, which is its fringe visibility $V_{\rm M}$ as a function of effective baseline D_e.

2.2.3 Fringe Visibility in Two Dimensions

If all sources were simple pairs of unresolved points, and all that was desired was the separation of the two points, it would be sufficient to find where the visibility curve for a given source drops to zero. One could then use (2.7) to compute the separation angle. Ultimately, however, the goal of interferometry is to take advantage of high angular resolution to discern structure in source images which appear only as points to a conventional telescope.

In the description that follows, and in the visibility curves shown later, the interferometer baseline will be represented by the *spatial frequency* $u = D/\lambda$, which is the length of the baseline in units of the observation wavelength. As shown in the previous section, $\lambda/2D$ is an angle, which is convenient to express in arcseconds. The spatial frequency will be expressed in inverse arcseconds (arcsec^{-1}), so

$$u \text{ (radians}^{-1}) = \left(\frac{180}{\pi}\right) \cdot 3600 \cdot u \text{ (arcsec}^{-1}).$$

Figure 2.3 illustrates the situation when a real source, with a two-dimensional brightness distribution $B(\xi, \eta)$,[1] is observed with a two-element interferometer. Light from a point p in the source will arrive at each telescope, producing an electric field $\mathbf{E_{1p}}$ at T_1 and $\mathbf{E_{2p}}$ at T_2, where

$$\begin{aligned} \mathbf{E_{1p}} &= E_{1p} e^{i(\omega t + \phi_{1p})} \quad \text{and} \\ \mathbf{E_{2p}} &= E_{2p} e^{i(\omega t + \phi_{2p})}. \end{aligned} \tag{2.8}$$

The phases ϕ_{1p} and ϕ_{2p} depend on the coordinates (ξ, η) of p, as well as the positions of the telescopes, expressed as spatial frequencies u and v. E_{1p} and E_{2p} are the (real) amplitudes after reception by the telescopes. These two fields are combined by the interferometer to produce a fringe pattern with field $\mathbf{E_p} = \mathbf{E_{1p}} + \mathbf{E_{2p}}$. Apart from a

[1] Over a given bandwidth, $B(\xi, \eta)$ measures the power per unit area received at the ground from an element of the source subtending a given solid angle. It could have units of $\frac{\mathrm{W}}{\mathrm{m}^2 \cdot \mathrm{steradian}}$, for example.