

Chapter 8

Phase and Group Delay Estimation

PETER R. LAWSON

JET PROPULSION LABORATORY
PASADENA, CALIFORNIA

In this Chapter I would like to describe the methods that are used for phase and group-delay estimation, and to outline the expected performance of each technique. The methods of phase estimation that are described here have been applied (with variations) at the Mark III interferometer (Shao *et al.*, 1988), the Palomar Testbed Interferometer (PTI) and the Navy Prototype Optical Interferometer (NPOI). Likewise, the methods of group-delay tracking to be described here, although different in each case, have been applied at PTI (Colavita *et al.*, 1999), NPOI (Benson *et al.*, 1998), the Sydney University Stellar Interferometer (Davis *et al.*, 1995; Lawson, 1995), and the Grand Interféromètre à 2 Télescopes (Koechlin *et al.*, 1996).

There are other methods of fringe detection and measurement, which do not rely on phase or group delay estimation as described in this chapter. These other approaches are methods of *coherence envelope* tracking and detect the location of the fringe packet by sweeping or scanning the delay line back and forth through the entire fringe envelope, with a throw several times larger than the coherence length. Such approaches have been used successfully at SUSI* (Davis *et al.*, 1999), COAST (Baldwin *et al.*, 1994), and IOTA (Traub, 1998). The advantage of envelope tracking is that it is straightforward to implement. The sweep is generally made much larger than the coherence envelope and therefore the envelope need only be roughly centered in the sweep. If the baseline solution is accurate and the path variations introduced by the atmosphere are small, corrections to the tracking position

*The method used at SUSI, described by Davis *et al.* (1999), does not use a fast sweep but instead uses steps through the coherence envelope and a method of fringe measurement first described by Tango and Twiss (1980).

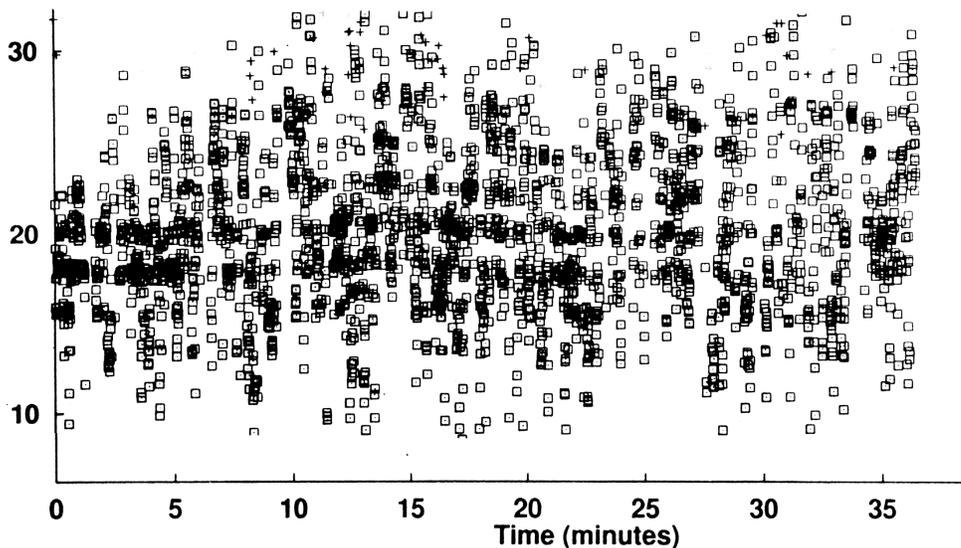


Figure 8.1: Group delay as a function of time observed with PTI in conditions of poor seeing. Several parallel fringes are visible in the plot, illustrating the behavior of the phase-tracking algorithm as it makes mistakes in identifying the central fringe. This information can be used to correctly unwrap the phase during post-processing of the data.

may be made infrequently—perhaps every few minutes. However, the fringes are only sampled for a fraction of the observation time, and the process may be labour intensive, time consuming, and inefficient.

In the following I will describe methods of phase and group delay estimation that allow the fringes to be observed on or near the peak of the coherence envelope. Methods of envelope tracking will not be considered further.

8.1 Motivation

All visibility measurements must be made in such a way that calibration of the data is later possible by observing unresolved reference sources. The visibility loss experienced by fringes observed in source and calibrator should be the same in each case, and fringes should therefore be measured at the same fixed position on the coherence envelope. Without some form of servo control the coherence envelope would move during an observation, because of errors in the astrometric model and random path variations ($\sim 10 \mu\text{m}$ rms per meter of baseline) introduced by the atmosphere. If the observations are to be made always at the same place on the coherence envelope (e.g. at the peak) some form of phase or group-delay tracking must be used.

Astrometric interferometers estimate the relative angular position of celestial objects through accurate measurements of the delay required to obtain fringes on each source. All astrometric interferometers currently use a combination of phase and group delay estimation to provide accurate delay measurements. Phase estimation is used to provide the highest resolution, and group-delay estimation is used so that the phase is unwrapped with reference to the peak of the coherence envelope, as illustrated in Figure 8.1.

Imaging interferometers that estimate closure phase must be capable of simultaneous phase measurements over at least three baselines. In practice this is accomplished by imposing a different spatial or temporal modulation to each baseline and then estimating phase through Fourier techniques or phase-tracking algorithms. Although phase tracking is not always necessary for imaging, as shown for example by the COAST interferometer (Baldwin *et al.*, 1994), phase-estimation methods are nonetheless used to estimate closure phase.

8.2 Phase and Group Delay

The optical path difference between the combined wavefronts in an interferometer can be expressed in terms of the indices of refraction of the different media n_i and the path lengths in each arm of the interferometer that the light traverses, x_{1i} and x_{2i} :

$$x(\kappa) = \sum_{k=0}^K (x_{1i} - x_{2i}) n_i(\kappa) \quad (8.1)$$

where $\kappa = 1/\lambda$ is the spectroscopic wavenumber at a wavelength λ .[†] If we now make the change of variables $x_i = x_{1i} - x_{2i}$, and assume that we have a vacuum delay x_0 and K dispersive media, we have that the phase of the fringes can be given by

$$2\pi\kappa x(\kappa) = 2\pi\kappa \left[x_0 + \sum_{k=1}^K x_i n_i(\kappa) \right] \quad (8.2)$$

In general the optical path-difference x is wavelength dependent, because the light that travels to the beam combiner from each arm of the interferometer may have passed through different paths in vacuum, and different dispersive pathlengths in air and glass: at shorter wavelengths the index of refraction is higher, light travels slower through the media, and the optical path-difference is larger. If we consider an interferometer that only admits a restricted bandwidth of light, there is light of a particular wavelength that arrives “first” and one that arrives “last,” and we can speak of the ensemble of waves as traveling as a

[†]Here I distinguish between the *spectroscopic* wavenumber, $\kappa = 1/\lambda$, and the wavenumber that is customarily used in optics, $k = 2\pi/\lambda$. This distinction is suggested by Born and Wolf (1980) and will be used throughout this chapter.

group or packet whose mean progress is described by a group velocity. The group delay is proportional to the rate-of-change of phase as a function of wavenumber, evaluated at the center of the band.

$$\text{group delay } (\kappa_0) = \left. \frac{d}{d\kappa} \kappa x(\kappa) \right|_{\kappa_0} \quad (8.3)$$

$$= x_0 + \sum_{k=1}^K \left. \frac{d}{d\kappa} \kappa x_i n_i(\kappa) \right|_{\kappa_0} \quad (8.4)$$

If we have simply a vacuum path-difference, $x(\kappa) = x_0$, the group delay is independent of wavelength, and the fringe phase is a linear function of wavenumber.

A phase-tracking algorithm would seek the location of a position of constant phase, where the fringe visibility is highest. A group-delay tracking algorithm would seek a location of constant group delay—where the number of fringes across the bandwidth is maintained constant.

A long-baseline stellar interferometer must typically compensate for a *vacuum* delay of several tens of meters and random variations of air path of several tens of microns rms. In principle, all ground-based interferometers should use vacuum delay lines, but many do not either because of funding restrictions or the perception that longitudinal dispersion is less of a problem at infrared wavelengths. Examples of fringes distorted by longitudinal dispersion have been presented by Lawson (1997) and include an illustration of the wavelength dependence of group delay. If the dispersion is uncompensated then fringes will appear to have a reduced visibility and fringes in different wavelength bands will arrive delayed one from the other, making it impossible to record fringes in two or more bands simultaneously. The effects of dispersion have been studied by numerous authors (Lacasse and Traub, 1988; Tango, 1990; ten Brummelaar, 1995; du Foresto *et al.*, 1995; Lawson and Davis, 1996; Léveque *et al.*, 1996; Davis *et al.*, 1998; Daigne and Lestrade, 1999) and dispersion compensators have been implemented, mostly on an experimental basis, at the I2T, SUSI, and PTI. Dispersion compensators will be used with the GI2T/REGAIN interferometer, the Keck Interferometer, the VLTI, and the CHARA Array—all of which use air delay lines.

It is interesting to note that if dispersion is present the fringe phase at the peak of the coherence envelope (at zero group delay) will not necessarily be zero. As the delay line introduces an increasing dispersive air path, the brightest fringe in the coherence envelope (the central fringe) will move away from the position of zero group delay. When it has moved more than half a wavelength, the fringe that followed it will be closer to the peak of the envelope and will become the new central fringe. The position of zero-phase and zero-group-delay will appear to move with respect to each other in a sawtooth pattern as a function of delay.

8.3 Model of the Fringe

The intensity of a fringe pattern can be expressed (cf. Chapter 7, Born and Wolf 1980) as follows:

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2} |\gamma_{12}| \cos(\phi_{12} - \varphi), \quad (8.5)$$

where

$$\varphi = \frac{2\pi}{\lambda}(s_2 - s_1). \quad (8.6)$$

λ is the wavelength of the light, I_1 and I_2 are the intensities of the light in each arm of the interferometer, and γ_{12} is the complex degree of coherence with modulus $|\gamma_{12}|$ and argument $\phi_{12} - \varphi$, where φ arises from the path difference, and ϕ_{12} contains information about the source. The parameters s_1 and s_2 are two optical pathlengths. The contrast, or visibility, of the fringes is the ratio of the fringe amplitude to the total background illumination,

$$\mathcal{V} = \frac{2\sqrt{I_1 I_2} |\gamma_{12}|}{I_1 + I_2}. \quad (8.7)$$

If we make a change of variables introducing the spectroscopic wavenumber $\kappa = 1/\lambda$, letting

$$I_s = 2\sqrt{I_1 I_2}, \quad I_b = I_1 + I_2 - I_s, \quad (8.8)$$

and

$$x = (s_2 - s_1), \quad (8.9)$$

then we have[‡]

$$I(\kappa, x) = I_s \left[1 + |\gamma_{12}| \cos(2\pi\kappa x - \phi_{12}) \right] + I_b. \quad (8.10)$$

The substitution of $x = s_2 - s_1$ in this case is to indicate that the phase offset is entirely piston phase, with no tilt component. Under conditions where $I_1 = I_2$ then the visibility of the fringes is the modulus of the complex degree of coherence,

$$\mathcal{V} = |\gamma_{12}|. \quad (8.11)$$

A source that is *non* quasi-monochromatic may still be treated as such if it is observed with an instrumental bandwidth that is sufficiently small. Bright fringes will occur wherever the path difference x is an integer multiple of 2π at most wavelengths. This will be so at *all* wavelengths only when x is zero and when the dispersion is the same in each arm of the interferometer. The reduction in the fringe visibility at increasing values of x is described by the coherence envelope.

[‡]Although not explicitly indicated here, a phase offset should be introduced to distinguish between fringes that are produced by light traveling single or double-pass through the beam-splitter. The offset phase is $\pi/2$ if the beam-splitter is used in the normal way: two beams of starlight enter the beam-splitter from opposite sides and are combined. However, the phase offset is 0 if the source of light is an artificial star that shines out through the beam-splitter, sending two beams out which are then autocollimated and returned. This is discussed at greater length by Traub in Chapter 3 (Section 3.2.3).

8.3.1 Coherence Envelope

If we observe fringes using a finite bandwidth (ie. not quasi-monochromatic), the recorded intensity is the integral of $I(\kappa, x)$ over wavenumber, weighted by a filter function $W(\kappa)$ that describes the bandpass.

$$I(\bar{\kappa}, x) = \int_{-\infty}^{\infty} W(\kappa - \bar{\kappa}) I(\kappa, x) d\kappa, \quad (8.12)$$

where $\bar{\kappa}$ is the center of the passband.

The filter function includes both the shape of the bandpass and the frequency response of the detector; it has values that are large within the bandwidth, and near zero outside. The result of this averaging is to reduce the sensitivity of the interferometer to fringes of large delay: when the bandwidth $\Delta\kappa$ partially spans a fringe (in the wavenumber domain κ) then the visibility appears to be reduced. This is simple to illustrate.

Let us introduce a change of variables, such that $\kappa' = \kappa - \bar{\kappa}$, and perform the integration in Equation 8.12 with respect to κ' . If we insert Equation 8.10 into Equation 8.12, ignore the background I_b for now, and rearrange the terms we have

$$\begin{aligned} I(\bar{\kappa}, x) = I_s [& 1 + |\gamma| \cos(2\pi\bar{\kappa}x - \phi) \int_{-\infty}^{\infty} W(\kappa') \cos(2\pi\kappa'x) d\kappa' \\ & - |\gamma| \sin(2\pi\bar{\kappa}x - \phi) \int_{-\infty}^{\infty} W(\kappa') \sin(2\pi\kappa'x) d\kappa'] \end{aligned} \quad (8.13)$$

where the subscripts have been dropped from γ_{12} and ϕ_{12} . Now if we define $\Omega(x)$ as the Fourier transform of $W(\kappa)$, then we have:

$$\Omega(x) = |\Omega(x)| e^{j\phi_\Omega} = \int_{-\infty}^{\infty} W(\kappa) e^{j2\pi\kappa x} d\kappa, \quad (8.14)$$

and therefore

$$|\Omega(x)| \cos \phi_\Omega = \int_{-\infty}^{\infty} W(\kappa) \cos(2\pi\kappa x) d\kappa, \quad (8.15)$$

$$|\Omega(x)| \sin \phi_\Omega = \int_{-\infty}^{\infty} W(\kappa) \sin(2\pi\kappa x) d\kappa. \quad (8.16)$$

Equation 8.12 may therefore be written in the form

$$I(\bar{\kappa}, x) = I_s \left[1 + |\gamma_x| \cos(2\pi\bar{\kappa}x - \phi + \phi_\Omega) \right] \quad (8.17)$$

where the apparent visibility $|\gamma_x|$ is the product of the true visibility and the modulus of the Fourier transform of the filter function, evaluated at the current delay:

$$|\gamma_x| = |\gamma| |\Omega(x)|. \quad (8.18)$$

The transfer function $\Omega(x)$ describes the *coherence envelope*. If $W(\kappa)$ is symmetric then $\Omega(x)$ is real valued, $\phi_\Omega = 0$, and only at zero delay, where the envelope is at its peak, is the true visibility observed.

A Rectangular Bandpass

If a detector has a rectangular bandpass then its coherence envelope would resemble a sinc function:

$$W(\kappa) = \begin{cases} 0, & |\kappa| > \Delta\kappa/2 \\ 1, & |\kappa| < \Delta\kappa/2 \end{cases} \quad \text{and} \quad \Omega(x) = |\Delta\kappa| \frac{\sin \pi x \Delta\kappa}{\pi x \Delta\kappa}. \quad (8.19)$$

If a bandwidth of $\Delta\lambda$ is used at a wavelength λ , the same interval expressed in wavenumber is as follows:

$$\Delta\kappa = \frac{1}{(\lambda - \Delta\lambda/2)} - \frac{1}{(\lambda + \Delta\lambda/2)}, \quad \text{therefore} \quad \Delta\kappa = \frac{\Delta\lambda}{\lambda^2 - (\Delta\lambda/2)^2}. \quad (8.20)$$

If we assume that the fractional bandwidth is very small we can ignore the second term in the denominator.

$$\Delta\kappa \simeq \frac{\Delta\lambda}{\lambda^2}. \quad (8.21)$$

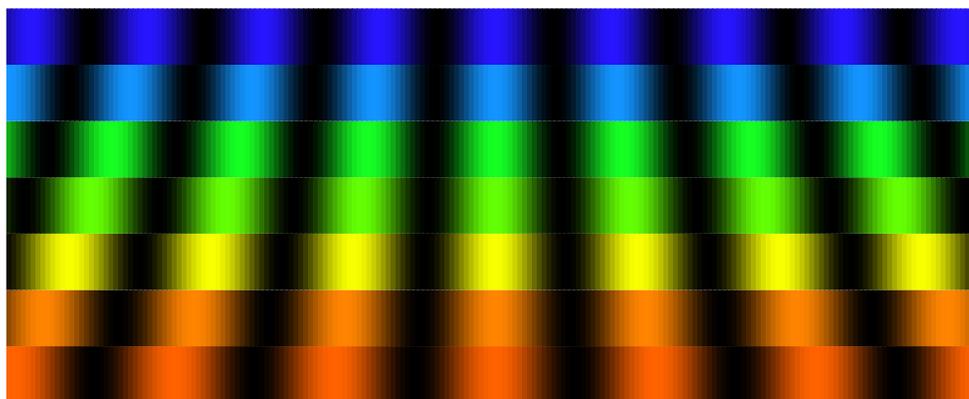
The sinc function is characterized by the location of its first zero crossing, where $x = 1/\Delta\kappa$. This distance can be thought of as the coherence length of the starlight under observation.

The Coherence Envelope and the Color of Fringes

Figure 8.2 illustrates in a more intuitive way how the coherence envelope would become narrower when fringes are observed over an increasingly large bandwidth, which encompasses many different colors (wavelengths). Fringes from a Young's double-slit experiment are shown in Figure 8.2(a), as seen through filters at different wavelengths. Figure 8.2(b) shows the corresponding fringe pattern if all the colors are viewed simultaneously. Note that as you increase the path-difference and move further away from the central fringe, either left or right from the center of the page, the fringes become less and less distinct because the fringes at each color cease to add constructively. The larger the range of colors, the sooner the fringes disappear as the pathlength is increased.

8.3.2 Channeled Spectrum

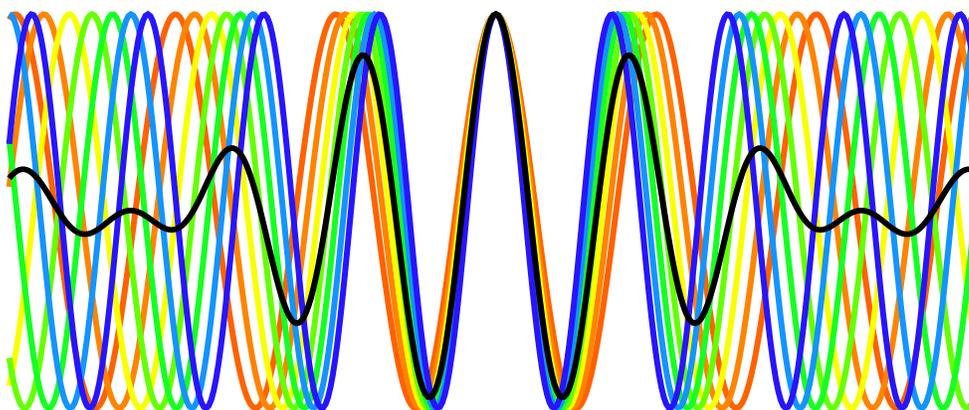
Fizeau and Foucault (1845) were the first to point out that fringes are still observable in the spectrum of interfered light, even when the path differences were so large that the white-light fringes have completely vanished. Suppose what you observed was a pattern similar to the one illustrated in Figure 8.2(b) and you somehow separated the different



(a)



(b)



(c)

Figure 8.2: Narrow band and white-light fringes: (a) Fringes from a Young's double-slit experiment, shown for seven different wavelengths or colors of light. (b) The corresponding "white-light" fringe after summing together the different colored fringes. (c) A cross section of the fringes at each color and the corresponding white-light fringe. Note that the color-wavelength relationship is not to scale. [After A.A. Michelson, *Light Waves and Their Uses* (University of Chicago Press: 1902), Plate II.]

colors to obtain the fringe patterns of Figure 8.2(a). You can now see fringes at large path differences (left and right in the figure) where none were previously visible. Which is to say, if we disperse the light in a spectrometer so that a bandwidth $\Delta\kappa$ is separated into M smaller bands of width $\Delta\kappa/M$, the coherence length for fringe detection changes from $1/\Delta\kappa$ to $M/\Delta\kappa$; it becomes M times larger. Fringes observed in dispersed light have been termed Edser-Butler fringes, or fringes of equal chromatic order, and produce a spectrum that is *channeled* with fringes (Steele, 1987).

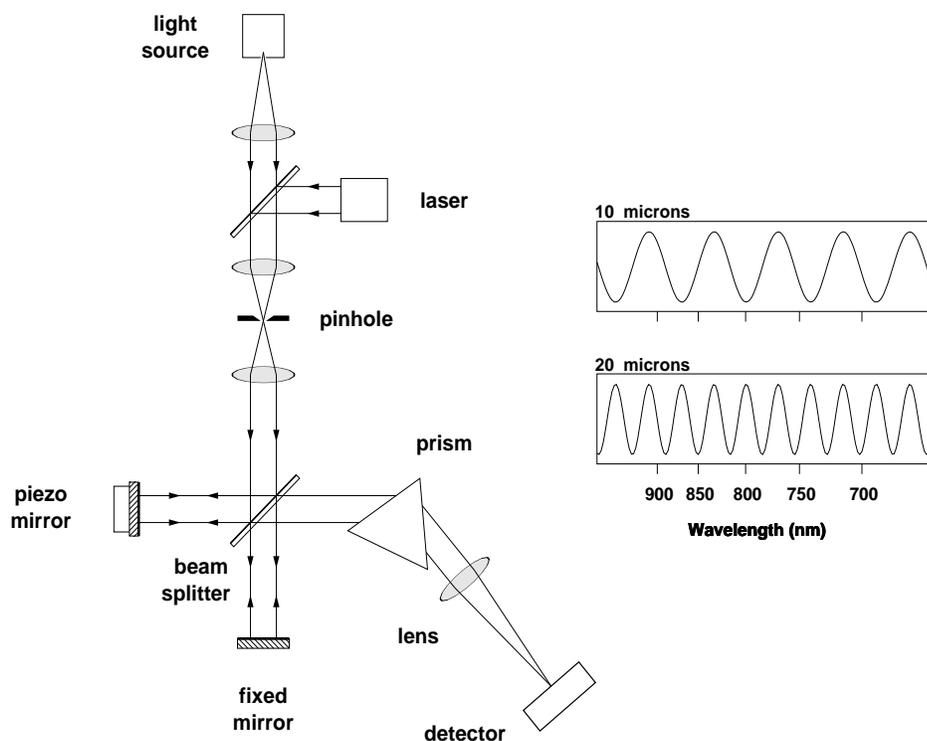


Figure 8.3: Layout of a Michelson interferometer used to produce channeled spectra. Simulations of channeled spectra are shown for path-differences of 10 and 20 μm , using a prism of BK7 glass.

We can re-write Equation 8.17 as a channeled spectrum (here arbitrarily setting ϕ_Ω to zero) in the following way:

$$f_k(x) = I_k \left[1 + |\gamma_k| \cos(2\pi\bar{\kappa}_k x - \phi_k) \right], \quad (8.22)$$

where k is an index number that counts the pixels across the spectrometer; such that at pixel k the wavelength and wavenumber are λ_k and κ_k respectively, I_k is the intensity of the stellar spectrum, $|\gamma_k|$ is the fringe visibility amplitude, and ϕ_k is the fringe phase.

As can be seen in Equation 8.22 the number of fringes in the spectrum across a given bandwidth is directly proportional to the optical path-difference. If p fringes are counted

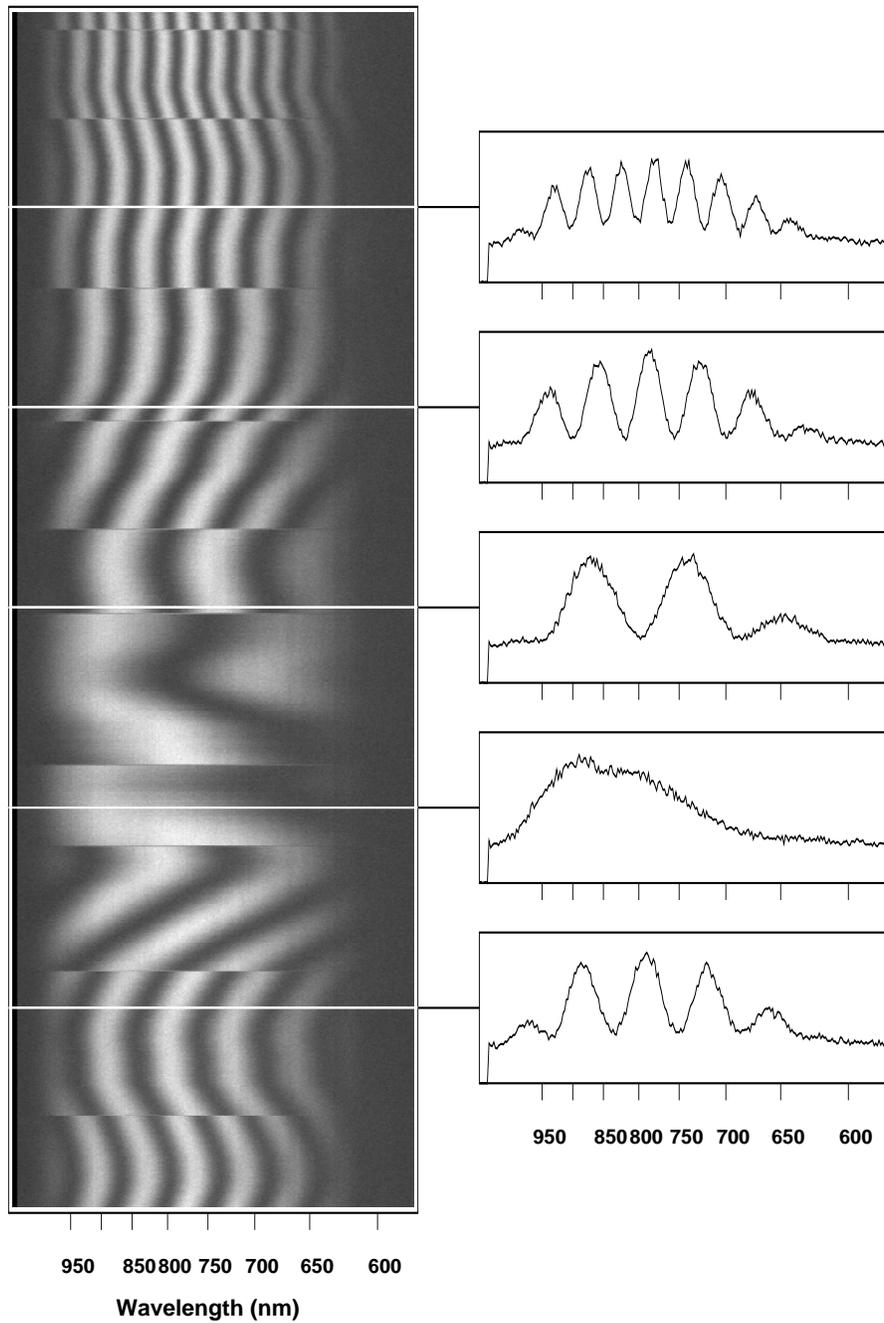


Figure 8.4: Channeled spectra produced at COAST using an artificial star. Various examples are shown as the delay line is stepped from one side of path-equality to the other.

between wavelengths λ_{\min} and λ_{\max} , then we have

$$p = \frac{1}{2\pi} \left[\frac{2\pi x}{\lambda_{\min}} - \frac{2\pi x}{\lambda_{\max}} \right], \quad (8.23)$$

and therefore

$$x = \frac{p}{\Delta\kappa}, \quad (8.24)$$

where

$$\Delta\kappa = \left[\frac{1}{\lambda_{\min}} - \frac{1}{\lambda_{\max}} \right]. \quad (8.25)$$

The optical path-difference can therefore be determined simply by estimating the number or frequency of fringes in the channeled spectrum.[§] Figure 8.3 shows a laboratory Michelson interferometer with a prism on its output. Simulated channeled spectra, as would be seen on the detector, are illustrated. Figure 8.4 shows a time sequence of actual channeled spectra recorded with a CCD camera as the path-difference is stepped from one side of zero path-difference to the other.

Numerous applications for channeled spectra have been described in the literature, including their use for analyzing spectroscopic measurements (Edser and Butler, 1898), for measuring absolute phase shift and dispersion (Sanderman, 1971), and for the analysis of thin films (Feldman, 1984).

8.4 Methods of Tracking Fringe Phase

Let us now return to a description of interference fringes, and look at the approaches that have been used to measure fringe phase with stellar interferometers.

All methods of phase-measurement interferometry involve a modulation of the fringes. Stellar interferometers that measure phase typically use a temporal modulation of the optical path-difference to sample the fringe. The modulation frequency is chosen fast enough so that (for the ground-based instruments) atmospheric path-fluctuations are effectively frozen.

We can rewrite the equation of the fringes, Equation 8.10, using the trigonometric identity

$$\cos(a - b) = \cos(a) \cos(b) + \sin(a) \sin(b), \quad (8.26)$$

and express the fringe pattern as follows:

$$I(\kappa) = \frac{1}{\tau} [N + X \cos(2\pi\kappa x) + Y \sin(2\pi\kappa x)], \quad (8.27)$$

where $\kappa = 1/\lambda$, τ is the measurement period, and

$$N = \tau I_s, \quad (8.28)$$

$$X = \tau I_s |\gamma_{12}| \cos(\phi_{12} - \phi_0), \quad (8.29)$$

$$Y = \tau I_s |\gamma_{12}| \sin(\phi_{12} - \phi_0), \quad (8.30)$$

[§]In this description, the number of fringes in the channeled spectrum only determines the distance from a zero group-delay, but not whether the offset is positive or negative. At least two measurements of channeled fringes at different delays are required to remove the ambiguity in sign.

where we have assumed the background I_b is identically zero. It can now be seen that with at least three measurements at different *known* positions of x we can solve for the three fringe parameters N , X , and Y , and that with many more measurements we can improve the accuracy of the estimate by applying a least-squares algorithm. The numerous different approaches to phase measurement interferometry are described in detail by Creath (1988). In the following it will be assumed that if the total measurement period is τ , and that if M measurements are made of the fringe, they are each made over a time τ/M .

8.4.1 Phase Tracking with Four Quarter-Wavelength Steps

For the purpose of illustration, let us assume the pathlengths in the interferometer are stepped in quarter-wavelength steps. The fringe is given, as above, by

$$I(\kappa, n) = \frac{1}{\tau} [N + X \cos(2\pi\kappa x_n) + Y \sin(2\pi\kappa x_n)], \quad (8.31)$$

and the pathlength modulation is

$$x_n = \frac{n}{4} \frac{1}{\kappa_0}, \quad n = 0, 1, 2, 3 \quad (8.32)$$

where λ_0 is the wavelength of the fringe measurement, ie: $\kappa_0 = 1/\lambda_0$. The four measurements, each integrated for a time $\tau/4$, can be written

$$A = (N + X) / 4, \quad (8.33)$$

$$B = (N + Y) / 4, \quad (8.34)$$

$$C = (N - X) / 4, \quad (8.35)$$

$$D = (N - Y) / 4, \quad (8.36)$$

and these can be reduced to the following three equations

$$A + B + C + D = N, \quad (8.37)$$

$$A - C = X/2, \quad (8.38)$$

$$B - D = Y/2. \quad (8.39)$$

We now arrive at a simple expression for the phase,

$$\phi_{12} - \phi_0 = \tan^{-1} \left(\frac{B - D}{A - C} \right). \quad (8.40)$$

The way that the A , B , C , and D counts are combined to produce phasors depends both on the form of the equation describing the fringe and the initial phase of the sweep—it is therefore not surprising that there are a confusing number of correct but somewhat different descriptions of the four-bin algorithm.

8.4.2 Phase Tracking with a Linear Pathlength Sweep

No stellar interferometer actually uses a phase-*stepping* approach. In practice the pathlength is not stepped, but is varied as a triangle or sawtooth wave with a peak-to-peak amplitude of one wavelength, $\lambda_0 = 1/\kappa_0$, or some integer number of wavelengths. Here we will consider only the case of a single-wavelength sweep. The following derivation is also described in a slightly different way by Wyant (1975, Equation 15ff) and by Colavita (1985, Chapter 4).

During the course of one cycle of duration τ , the photon counts A , B , C , and D are each recorded over a time interval of $\tau/4$. The measurement comprises separate integrations which are each an average across part of the total sweep,

$$\frac{1}{\Delta x} \int_{x_n}^{x_{n+1}} I(\kappa, n) dx \quad \text{where} \quad \Delta x = x_{n+1} - x_n. \quad (8.41)$$

If we have four separate integrations each across a quarter wavelength, then we set $\Delta x = 1/(4\kappa_0)$. We will also set the boundaries of the integration so that the total sweep is one wavelength, symmetric about a zero path-difference:

$$x_n = \frac{n}{4} \frac{1}{\kappa_0} - \frac{3}{4} \frac{1}{\kappa_0}, \quad n = 1, 2, 3, 4, 5. \quad (8.42)$$

It is straightforward to perform the integrations described by Equations 8.41, 8.42, and 8.27. We have that A is bounded by (x_1, x_2) , B is bounded by (x_2, x_3) , C is bounded by (x_3, x_4) , and D is bounded by (x_4, x_5) . We also have that

$$\int_{x_n}^{x_{n+1}} \cos(2\pi\kappa x) dx = \frac{1}{2\pi\kappa} [\sin(2\pi\kappa x_{n+1}) - \sin(2\pi\kappa x_n)], \quad (8.43)$$

and

$$\int_{x_n}^{x_{n+1}} \sin(2\pi\kappa x) dx = \frac{1}{2\pi\kappa} [\cos(2\pi\kappa x_n) - \cos(2\pi\kappa x_{n+1})]. \quad (8.44)$$

We can now write the A , B , C , and D as follows:

$$A = \frac{N}{4} + X \frac{1}{2\pi} [\sin(-\pi/2) - \sin(-\pi)] + Y \frac{1}{2\pi} [\cos(-\pi) - \cos(-\pi/2)], \quad (8.45)$$

$$B = \frac{N}{4} + X \frac{1}{2\pi} [\sin(0) - \sin(-\pi/2)] + Y \frac{1}{2\pi} [\cos(-\pi/2) - \cos(0)], \quad (8.46)$$

$$C = \frac{N}{4} + X \frac{1}{2\pi} [\sin(\pi/2) - \sin(0)] + Y \frac{1}{2\pi} [\cos(0) - \cos(\pi/2)], \quad (8.47)$$

$$D = \frac{N}{4} + X \frac{1}{2\pi} [\sin(\pi) - \sin(\pi/2)] + Y \frac{1}{2\pi} [\cos(\pi/2) - \cos(\pi)], \quad (8.48)$$

which may be simplified to

$$A = \frac{N}{4} + \frac{1}{2\pi} (-X - Y), \quad (8.49)$$

$$B = \frac{N}{4} + \frac{1}{2\pi} (X - Y), \quad (8.50)$$

$$C = \frac{N}{4} + \frac{1}{2\pi}(X + Y), \quad (8.51)$$

$$D = \frac{N}{4} + \frac{1}{2\pi}(-X + Y). \quad (8.52)$$

$$(8.53)$$

We can now write

$$\frac{B - D}{A - C} = \frac{Y - X}{X + Y}. \quad (8.54)$$

Dividing through by X and using the relationship $Y/X = \tan(\phi_{12} - \phi_0)$, we have

$$\frac{B - D}{A - C} = \frac{\tan(\phi_{12} - \phi_0) - 1}{1 + \tan(\phi_{12} - \phi_0)}. \quad (8.55)$$

Using the trigonometric relationship

$$\tan(a - b) = \frac{\tan a - \tan b}{1 + \tan a \tan b}, \quad (8.56)$$

we have finally that

$$\tan^{-1}\left(\frac{B - D}{A - C}\right) = \phi_{12} - \phi_0 - \frac{\pi}{4}, \quad (8.57)$$

$$\phi_{12} - \phi_0 = \tan^{-1}\left(\frac{B - D}{A - C}\right) + \frac{\pi}{4}. \quad (8.58)$$

The phase shift of $\pi/4$ corresponds to a lag of half an integration bin, as would be expected.

Although the minimum number of bins that one could use would be three, algorithms with more than four bins are also in use. For example, the Navy Prototype Optical Interferometer uses an eight-bin algorithm where the real and imaginary components of the phase are calculated as follows (Benson, 1998):

$$\begin{aligned} \cos(\phi) &\propto (A - E) + 0.5\sqrt{2}(B - D - F + H), \\ \sin(\phi) &\propto (C - G) + 0.5\sqrt{2}(B + D - F - H), \end{aligned}$$

with the integration bins now extending from A through H . The phase offset in this case would be $\pi/8$.

Cassaing *et al.* (2000) have suggested that a servo that only seeks to find the zero-phase position, and does not share data with the science instrument, need only estimate the sine phasor, as that would suffice for a zero-seeking servo.

8.4.3 Simultaneous Phase Measurements at Several Wavelengths

Astrometric interferometers often make phase measurements at several different wavelengths simultaneously. The actual modulation is chosen to be one wavelength at the longest wavelength where data is measured. For data at shorter wavelengths, the data acquisition is halted briefly when the modulation exceeds one wavelength, and then resumed when it returns. The boundaries of the A , B , C , and D bins are re-defined according to the timing appropriate at each wavelength.

8.4.4 Triangle vs Sawtooth Modulations

A sawtooth waveform is preferable when the detector being used to record the fringe has appreciable readout noise.

A triangle wave produced by a moving piezo will have non-linearities that are different in its up and down-strokes. These non-linearities will produce a bias in the fringe phase measurement—a phase measured using an up-stroke will have a different bias than a phase measured using a down-stroke. The problem arises that all phase estimates should have the *same* bias, so that phase-difference measurements (used for the astrometry) are unbiased.

At the Mark III interferometer, each phase estimate was made as an average between a measurement made on the up-stroke and a measurement made on the down-stroke. All phase estimates therefore had the same bias. In practice several cycles of up and down-strokes were averaged if the atmospheric conditions permitted. Photomultiplier tubes were used that photon-counted without incurring read noise.

At the Palomar Testbed Interferometer, the detector that is used is a NICMOS III detector. It has very high read noise, and even with multiple non-destructive reads, the read noise is about 12 electrons rms. Each phase estimate should use the absolute minimum number of reads necessary, so that the highest signal-to-noise is achieved in each cycle. A sawtooth waveform is therefore used. It has only an up-stroke, and so every phase measurement has the same bias—and only four reads (one measurement of A , B , C , D) are used rather than eight reads per phase estimate.

8.5 Methods of Tracking Group-Delay

Group-delay tracking has a very long history of use in stellar interferometry. Michelson and Pease (1921) applied this technique to acquire fringes by eye with the 20-ft interferometer using a direct-view prism. Labeyrie (1975) used an identical approach when he demonstrated that fringes could be acquired with two separated telescopes. The I2T interferometer acquired fringes this way up until about 1984, and the GI2T also routinely used a direct-view prism up until about 1995.

The idea of applying this technique using photon counting detectors was no doubt obvious to Labeyrie and was also suggested by Tango and Twiss (1980). The probable limitations of group-delay tracking have been described by numerous authors since then. These include simulations for the IOTA interferometer performed by Nisenson and Traub (1987) and Traub *et al.* (1990); simulations for the COAST interferometer performed by Buscher (1988); signal-to-noise predictions for “photon-starved” operation with the Mark III interferometer (Shao *et al.*, 1988); simulations for SUSI considering photon noise only (Lawson, 1995); and further simulations taking into account detector read noise (ten Brummelaar, 1997) and visibility fluctuations (Lawson *et al.*, 1999). Simulations have also shown that

with the use of *a priori* knowledge and Bayesian analysis methods that performance limits could be further extended (Meisner, 1996; Padilla *et al.*, 1998; Morel and Koechlin, 1998).

Channeled spectra were recorded by Kim (1989) using a PAPA camera at the Mark III interferometer, although without implementing a servo loop. Similar observations were carried out by the author (Lawson, 1994) at the Sydney University Stellar Interferometer, along with observations of atmospheric path fluctuations seen in channeled spectra (Davis *et al.*, 1995). In 1994 the GI2T began to automate a low-bandwidth servo for pathlength control using dispersed fringes (Koechlin *et al.*, 1996). When the Palomar Testbed Interferometer was commissioned in 1995, it automated a method of group-delay tracking using phasor measurements (Colavita *et al.*, 1999), as did the NPOI at about the same time (see for example Benson *et al.* 1998 and Hummel 2000). Future applications of group-delay tracking with space-borne interferometers have been reviewed by Shao and Colavita (1992).

Approaches to Group-Delay Estimation

The group delay can be measured if the combined beams from an interferometer are dispersed in a spectrometer. The detected spectrum of the star will be channeled with fringes whose number is proportional to the optical path-difference.

Approaches to group-delay tracking can be broadly classified according to the form of the measurement and the type of data processing that is used. The measurement will be either of a single channeled spectrum or of fringe phasors.*

1. *Channeled spectrum*: single snap-shot of the stellar spectrum channeled with fringes, but no pathlength modulation (e.g. Lawson 1995, with delay estimation from a transform of real-valued data).
2. *Multi-wavelength phasor measurements*: fringe phasors recorded at multiple wavelengths, with pathlength modulation and methods of phase measurement interferometry (e.g. Colavita *et al.* 1999; Armstrong *et al.* 1998, with delay estimation from a transform of a series of complex numbers).

Approaches to Data Reduction

The approaches to data processing for group-delay tracking could include a cross-correlation with an optimal filter, a least-squares modeling of the data, or a method of power spectrum analysis. Each of these approaches would use a simplified model of the fringe, based on a small set of free parameters, such as fringe frequency, amplitude, phase, and some

*This distinction is perhaps artificial. What I have called the *fringe phasor* approach, would be an n -bin method of phase measurement applied at multiple wavelengths, for which n snap-shots of channeled spectra are required. Although the sign of the delay cannot be determined from a single channeled spectrum, with two or more channeled spectra recorded with a known shift in delay, as is the case with measuring the fringe phasors, the sign is straightforward to determine.

assumptions—including the relationship between detector pixels and wavelength on the spectrometer. The utility of a particular technique depends on how well its assumptions model the data. For instance, the fast Fourier transform (FFT) assumes that the data represent a series of harmonically related sinusoids, but if there exists one sinusoid that does not coincide with any of the harmonics, then the corresponding power spectrum will be poorly reconstructed. Likewise, if the data were sampled at irregular intervals then the sampling will also bias the group delay.

Advances in spectrum analysis have come from deriving power spectra from more accurate assumptions, as is illustrated in the review paper by Kay and Marple Jr. (1981). Spectacular improvements are possible if the proper model is chosen and a strong signal is present. Unfortunately, all of these methods will fail when they are asked to derive power spectra from processes that deviate from their model. This can occur in some cases simply by adding observation noise to the data: at low signal-to-noise levels the resolution is often no better than an FFT approach, and consequently many of the methods are ill adapted for real-time processing. While it may be possible to determine the parameters that describe the power spectrum, one must then recalculate the power spectrum numerous times to locate the fringe peak, performing lengthy calculations. This has meant that only the relatively prosaic FFT and several simple variations of least-squares methods have been used in stellar interferometry. These will now be described.

8.5.1 Channeled Spectrum: Fast Fourier Transform

The most straightforward method is to use, despite its limitations, is the fast Fourier transform. If we choose the FFT to process the data, we can cast the problem in terms of an estimate of visibility using the Discrete Fourier Transform (DFT) with the same model for the fringes used by Walkup and Goodman (1973).

If we can assume that wavenumber is mapped linearly onto the detector then an FFT could be used. If the mapping is non-linear then a DFT would be used with the actual wavenumbers corresponding to each sample in the spectrometer. Let us look at how an FFT would be implemented.

Linear Mapping

If wavenumber κ is mapped linearly onto the detector coordinates ξ ,

$$\kappa = c_0 \xi \tag{8.59}$$

then it is straightforward to describe the sampled and transformed data, where we ignore DFT artifacts in the following discussion. If the detector has M pixels of width $\Delta\xi$, and wavelengths from λ_{\min} to λ_{\max} mapped onto it, then we have

$$\Delta\xi = \frac{1}{Mc_0} \left[\frac{1}{\lambda_{\min}} - \frac{1}{\lambda_{\max}} \right], \quad \text{or} \quad \Delta\xi = \frac{\Delta\kappa}{Mc_0}, \tag{8.60}$$

where $\Delta\kappa$ is the corresponding interval in wavenumber between λ_{\min} and λ_{\max} .

$$\Delta\kappa = \frac{1}{\lambda_{\min}} - \frac{1}{\lambda_{\max}}. \quad (8.61)$$

The samples therefore lie at intervals of wavenumber given by

$$\kappa_m = \kappa_{\min} + mc_0\Delta\xi, \quad m = 0, 1, \dots, M. \quad (8.62)$$

The transform determines the spatial frequency of the fringes detected across the array, that is to say p fringes per M pixels. We have therefore $x = p/\Delta\kappa$,

$$x_p = \frac{p}{Mc_0\Delta\xi}, \quad p = 0, 1, \dots, M/2. \quad (8.63)$$

and p is an index of spatial frequency. Using the expressions for κ_m and x_p we have

$$x\kappa = \frac{mp}{M} + \frac{p\kappa_{\min}}{\Delta\kappa}. \quad (8.64)$$

If we can describe the fringes as in Equation 8.10:

$$I(m) = I_s [1 + |\gamma| \cos(2\pi\kappa x + \phi_\gamma)] + I_b, \quad (8.65)$$

then inserting Equation 8.64 into 8.65 yields

$$I(m) = I_s \left[1 + |\gamma| \cos \left(\frac{2\pi mp_0}{M} + \phi \right) \right] + I_b, \quad (8.66)$$

where $I(m)$ is the average intensity at the m^{th} pixel of the detector,

$$\phi = \frac{2\pi p\kappa_{\min}}{\Delta\kappa} + \phi_\gamma. \quad (8.67)$$

The average total number of photons in each frame of data can be written

$$M_t = M(I_s + I_b), \quad (8.68)$$

where I_s and I_b are the average stellar spectrum and background per pixel.

The Discrete Fourier Transform of the detected channeled spectrum would be

$$\mathcal{I}(p) = \sum_{m=0}^{M-1} I(m) \exp \left[j \frac{2\pi pm}{M} \right], \quad (8.69)$$

whose real and imaginary parts are

$$\text{Re}[\mathcal{I}(p)] = \sum_{m=0}^{M-1} I(m) \cos \left(\frac{2\pi pm}{M} \right) = \begin{cases} M(I_s + I_b) & p = 0 \\ M(I_s|\gamma|/2) \cos \phi & p = p_0 \\ 0 & \text{otherwise} \end{cases} \quad (8.70)$$

and

$$\text{Im}[\mathcal{I}(p)] = \sum_{m=0}^{M-1} I(m) \sin \left(\frac{2\pi pm}{M} \right) = \begin{cases} M(I_s|\gamma|/2) \sin \phi & p = p_0 \\ 0 & \text{otherwise} \end{cases} \quad (8.71)$$

where the factor M arises in performing the sum of the DFT, and the fringe amplitude contains a factor of $1/2$ because the DFT calculates both positive and negative frequency components. These are complex conjugates, this being the transform of real-valued data, and no information would be lost if we discarded the negative frequency half.

The features at $\pm p_0$ are not truly delta functions, but sinc functions whose nulls lie at the locations of the other samples in the spatial frequency domain. For example, if the fringe frequency was in fact somewhere part-way between the frequencies sampled by the FFT, the convolution of the sinc function with the sampling would be more obvious.

Power Spectrum and Periodogram

It would be normal in most approaches of group-delay tracking to form a power spectrum from the complex transform described by Equations 8.70 and 8.71 and afterwards integrate the power spectra to improve the signal-to-noise ratio. From Equations 8.66 and 8.68 we have that the amplitude of the power spectrum would be

$$|\mathcal{I}(p)| = \begin{cases} M_t & p = 0 \\ M(I_s|\gamma|/2) & p = \pm p_0 \\ 0 & \text{otherwise} \end{cases}, \quad (8.72)$$

Examples of fringe signals detected in channeled spectra are shown in Figure 8.5. These are from internal fringes formed with SUSI used in autocollimation. Note the peak at the fringe frequency (located between 30 and 70 cycles) and the large peak at the zero frequency.

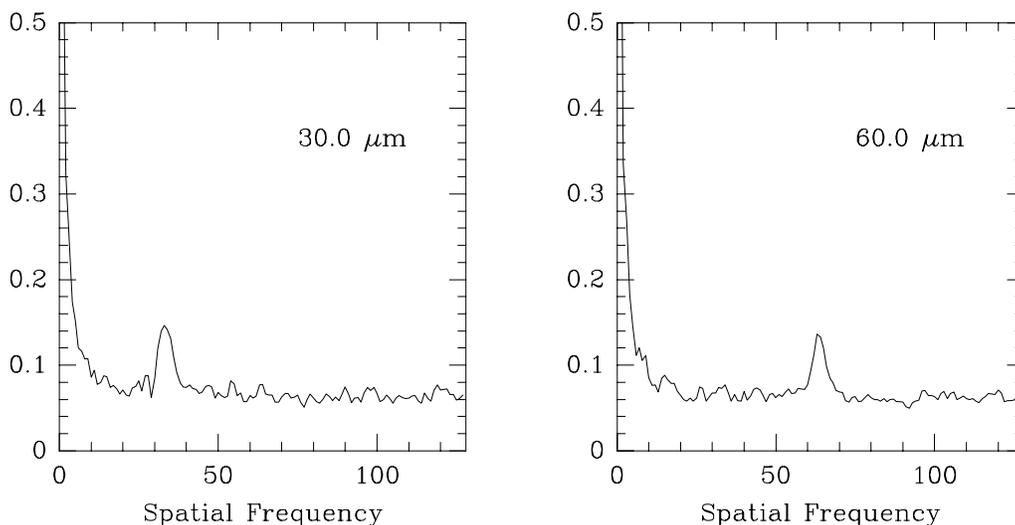


Figure 8.5: Examples of the fringe signal from an FFT processing of channeled spectra. Note the large signal at a spatial frequency of 0. The distances indicated are relative positions of a piezo actuator and are not with respect to the location of zero path-difference. [From Figure 10.7 of Lawson (1994).]

Our ability to locate the spatial frequency of the peak is somewhat better than the “resolution” of the power spectrum. The resolution in the power spectrum is the smallest separation in spatial frequency for which *two* unresolved line-features can be distinguished as separate.[†] If the spectrum only has a single feature (the fringe) the problem reduces to finding the best fit to the data of a known transfer function (i.e. a sinc function in the case of a square bandpass). The peak can be located by padding the data with its mean value before applying the DFT to produce samples of the power spectrum at shorter intervals. The peak can be approximately located this way (it isn’t practical to infinitely pad the data) and can be further determined by applying a three-point parabolic interpolation around those samples nearest the peak.

An advantage of using power spectrum analysis is that it allows incoherent integration of the fringe signal. It is the poor resolution in delay, when compared with phase-tracking methods, that makes this approach attractive. Small changes in delay may be unresolved in the power spectrum, making it possible to integrate numerous noisy power spectra and to thereby improve the sensitivity. Group-delay tracking with the FFT has been seen as particularly suited to low-light-level conditions in which methods of phase tracking would fail.[‡]

The methods of phase measurement require a modulation, and consequently the fringes are very slightly blurred in each sample—the fringes move by $\lambda/4$ per sample, reducing the fringe visibility by $\sim 10\%$. If we are processing a single channeled spectrum as described here without modulating the delay line, the sensitivity for fringe detection is slightly better in comparison.

A significant drawback of this approach is that without some subtle changes to the data processing, it is difficult to track fringes at zero delay. Because in each cycle we only measure a single channeled spectrum we cannot determine the sign of the delay, making this method a poor candidate for implementation as a zero-seeking servo. Furthermore, to avoid the fringe signal being buried in the low spatial frequency profile of the stellar spectrum, the zero-frequency signal must be subtracted in each frame. This is somewhat complicated because of variations in the intensity of the stellar spectrum due to scintillation.

Non-Linear Mapping

A problem that is common to all implementations of group-delay tracking is that the mapping from wavenumber to pixel number will most likely not be linear. Most spectrometers

[†]The resolution is inversely proportional to the total bandwidth detected by the array, and the DFT produces estimates at intervals in delay of $\Delta x = 1/\Delta\kappa$ up until a cut-off of $x = \pm M/(2\Delta\kappa)$, corresponding to a distance of half the coherence length.

[‡]This comparison is only valid if we assume that the sensitivity and noise characteristics (dark current and read noise for example) are the same for the detectors used in each approach.

use either a prism or a grating, neither of which have dispersions that are constant in κ .[§] Consequently the distance between fringes will change throughout the detected spectrum, and the fringes will be partly stretched or compressed. This “chirp” means that although the number of fringes would be the same, the associated frequency is more difficult to identify. The Fourier transform yields a fringe *frequency*, not a number-of-fringes. The transform of the fringes is therefore not a delta function and the peak is broadened with its height reduced. For instance, if the spacing between fringes doubles from one edge of the detector to the other, then the peak would be spread between these two frequencies. This effect is more severe the more fringes are present: at larger path differences the peak becomes progressively broader and lower in height. It becomes more difficult to detect the peak in the presence of noise, and the broadening means that the peak is less well defined. Careful calibration of the wavelength scale of the detector along with the use of a discrete Fourier transform is required to overcome these losses. The effects of longitudinal dispersion within the interferometer must also be understood and accounted for, otherwise the sensitivity of tracking will be degraded, sometimes in unexpected ways (Lawson and Davis, 1996).

8.5.2 Channeled Spectrum: Lomb-Scargle Periodogram

One of the disadvantages of the FFT approach is that the periodogram is not normalized in a way that allows thresholding against noise. It is therefore difficult to judge the significance of a peak in the power spectrum relative to the noise. A better method would be a power spectrum derived from a least-squares fit to the data. This would then give us some measure of the residuals and goodness of the fit.

The Lomb-Scargle periodogram, discussed by Press *et al.* (1992), is such an approach. The advantage of this method is that it normalizes the periodogram (power spectrum) so that it is possible to ignore noise peaks below a set threshold. Although it appears to be a computer intensive approach, W.J. Tango at the University of Sydney has implemented it for real-time fringe tracking with a CCD detector.

Because the Lomb-Scargle approach requires that the data sets be real-valued only, I have not yet seen how it could be adapted to process fringe phasors.

8.5.3 Channeled Spectrum: Least-Squares Fit

Traub *et al.* (1990) describe a method of group-delay tracking which uses a cross-correlation of the data with model functions. It is assumed that a family of functions exist which will closely fit the data providing certain parameters are adjusted. It follows that if these parameters are chosen correctly then it will minimize the least-squared difference between the data and the model. If we were to consider the delay by itself then we would perform

[§]An exception to this is a 60° prism of BK7 glass, which is closely linear over the wavelength range of 600–1000 nm.

the minimization by taking the partial derivative with respect to the delay, x , of the mean squared difference, equating it to zero, and solving for the delay. If the data are represented by the set g_k and the model is $f_k(x)$ then we have

$$\frac{\partial}{\partial x} \left[\sum_{k=1}^K [g_k - f_k(x)]^2 \right] = 0, \quad (8.73)$$

which can be written in full as

$$\frac{\partial}{\partial x} \left[\sum_{k=1}^K [g_k^2 - 2g_k f_k(x) + f_k^2(x)] \right] = 0. \quad (8.74)$$

In this equation only the cross term is of interest. The sum of the g_k^2 terms is a constant and contributes nothing to the minimization. Furthermore, if the model $f_k(x)$ was normalized correctly then the sum of the $f_k^2(x)$ terms would be independent of x , and therefore would also be a constant. We can now express the minimization of the mean square difference as

$$\frac{\partial}{\partial x} \left[\sum_{k=1}^K g_k f_k(x) \right] = 0, \quad (8.75)$$

where the function $f_k(x)$ maximizes the sum of the cross terms. The quantity in brackets is simply the cross-correlation between the model and the data, calculated at zero lag. Traub *et al.* (1990) presented simulations of pathlength motions with peak-to-valley excursions of $1.3 \mu\text{m}$ over 1 second with $|\gamma| = 1.0$. He concluded that delay tracking should be possible at count rates as low as 10 photons per coherence time, with a position uncertainty of $\sim 0.2\lambda$.

8.5.4 Multi-Wavelength Phasor Measurements: Fast or Direct Fourier Transform

Multiple-wavelength phase measurements are complicated somewhat because you can never modulate a *pathlength* to produce the same phase shift at all wavelengths. However, if you can control the timing of your detector then it is possible to bin the data separately at each wavelength and at the same time ignore data at wavelengths where the phase introduced by the phase-shifter (piezo) has already changed by 1λ or more. This is the approach that was used at the Mark III interferometer and which is currently used at PTI and NPOI.

We have then that the four bins A , B , C , and D that characterize ϕ as a function of wavenumber κ are recorded in each cycle of modulation. Let us then assume that the phase ϕ arises from a vacuum path-difference x , such that

$$\phi(\kappa) = 2\pi\kappa x, \quad (8.76)$$

If we measure the quantities A , B , C , D , at M wavenumbers we can calculate

$$h_c(\kappa_m) = A(\kappa_m) - C(\kappa_m) \quad m = 0, \dots, M - 1 \quad (8.77)$$

$$h_s(\kappa_m) = B(\kappa_m) - D(\kappa_m) \quad m = 0, \dots, M - 1 \quad (8.78)$$

so that we now have

$$h_c(\kappa_m) \propto \cos(2\pi\kappa_m x), \quad (8.79)$$

$$h_s(\kappa_m) \propto \sin(2\pi\kappa_m x). \quad (8.80)$$

We can now define the complex number series

$$h(\kappa_m) = h_c(\kappa_m) + jh_s(\kappa_m), \quad m = 0, \dots, M - 1 \quad (8.81)$$

and take its discrete Fourier transform

$$H(x) = \sum_{m=0}^{M-1} h(\kappa_m) \exp(j2\pi\kappa_m x). \quad (8.82)$$

The value of x that locates the peak in the power spectrum $|H(x)|^2$ corresponds to the group delay of the fringes.

The advantage of this method is that the modulation allows the sign of the delay to be unambiguously determined, and because phasors are processed there is no zero-frequency term in the power spectrum: it is straightforward to track at zero group delay. This allows the tracking to be implemented as a zero-seeking servo, and because the mean tracking position can be zero the astrometric error introduced by incorrectly scaling spatial frequencies to delays is of less consequence.

The disadvantage of this brute-force approach is that you cannot weigh the data to distinguish between good and bad estimates of the sine and cosine of the phase. This ability to weight the data is important if we know beforehand that certain pixels in our array are noisier than others. If for example we know that all the sine and cosine measurements estimates are noisy, we would like to have some figure-of-merit to allow us to judge the usefulness of the derived delay estimate.[†]

8.5.5 Multi-Wavelength Phasor Measurements: “Optimal” Estimator

With the conventional approach to group-delay estimation, the delay is inferred from the *spatial frequency* of the fringes in a channeled spectrum. It is assumed that the source is so faint and the atmosphere so unstable that coherent integration is limited to time-scales less than $\sim 2t_0$, and thus *incoherent* integration (the integration of power spectra) is used. The phase information in the complex Fourier transform of the fringes is simply thrown away, because it is assumed to be so corrupted by noise that it is unrecoverable.

At high light levels where a sufficient signal-to-noise is achievable in a time less than t_0 , the phase of the channeled spectra can indeed be extracted. It is then possible to formulate a group-delay estimate using this phase and thereby greatly improve the resolution of the estimates (cf. Equations 8.70 and 8.71 yield the phase ϕ at the fringe frequency, p_0).

[†]Dave Mozurkewich has pointed out that a simple but perhaps heavy-handed way of weighting the data is to ignore data points that are suspect.

This approach is a two-step procedure: one must first correctly identify the spatial frequency of fringes in the channeled spectra (normal group-delay estimate); and secondly extract the phase of those fringes and interpret it in terms of a delay. This approach is *optimal* in the sense that it provides a group-delay estimate with the variance of a phase estimator. This will only work well at high light levels and under circumstances where the dispersion is well understood. Lawson *et al.* (2000) have described an implementation using phasors for use at PTI. The approach has also been independently considered by Mozurkewich, Hummel, and Benson for use at the NPOI (Mozurkewich, 2000). Although the method is not yet in routine use at either interferometer—in part because of the difficulty in modeling the changing atmospheric dispersion—it may ultimately allow noise in group-delay estimates to be greatly reduced.

8.6 Variance of Phase and Group Delay Estimates

The derivations that follow have been previously described by Lawson *et al.* (2000).

8.6.1 Variance of Phase

The expected signal-to-noise ratio (SNR) and rms phase error σ_ϕ for a four-bin phase estimate has been derived by Wyant (1975):

$$\text{SNR} = \frac{2}{\pi} \sqrt{NV^2}, \quad \sigma_\phi = \frac{\pi}{2} \frac{1}{\sqrt{NV^2}}, \quad (8.83)$$

where V is the fringe visibility, N is the number of photons per frame, and the effects of background and detector read noise have been ignored.

8.6.2 Variance of Group Delay

Variance of Phase-Slope Derivation

The group delay is proportional to the slope of the phase as a function of wavenumber. The group delay is defined as

$$\frac{1}{2\pi} \frac{\partial \phi}{\partial \kappa}, \quad (8.84)$$

where $\kappa = 1/\lambda$. If we have several noisy samples of phase at independent wavenumbers, the variance of the group delay is proportional to the variance of the slope of a straight line fit to that data.

If there are M data points and the i^{th} data point has for its coordinate x_i and a variance of σ_i^2 , then the variance of the slope σ_b^2 of a least-squares fit of a line to that set of data is given as (Press *et al.* 1992, Section 15.2 “Fitting data to a straight line,” Equations 15.2.4 to 15.2.9):

$$\sigma_b^2 = \frac{S}{\Delta}, \quad (8.85)$$

where

$$S = \sum_{i=1}^M \frac{1}{\sigma_i^2}, \quad \Delta = S S_{xx} - (S_x)^2, \quad (8.86)$$

and

$$S_x = \sum_{i=1}^M \frac{x_i}{\sigma_i^2}, \quad S_{xx} = \sum_{i=1}^M \frac{x_i^2}{\sigma_i^2}. \quad (8.87)$$

The group delay is $(1/2\pi)$ times the slope of phase with respect to wavenumber, $\kappa = 1/\lambda$. The variance of the group delay is therefore proportional to the variance of the phase—as measured at M data points across a band $\Delta\kappa$. If the variance of the central-fringe phase is σ_ϕ^2 and we assume that the light from the broadband channel is divided equally amongst M pixels, the variance of the phase in each pixel will be $M\sigma_\phi^2$. We can therefore write that

$$S = \frac{1}{\sigma_\phi^2}, \quad (8.88)$$

$$S_x = \frac{1}{\sigma_\phi^2} \left[\frac{1}{M} \sum_{i=1}^M \kappa_i \right] \rightarrow \frac{1}{\sigma_\phi^2} \frac{1}{\Delta\kappa} \int_{\bar{\kappa}-\Delta\kappa/2}^{\bar{\kappa}+\Delta\kappa/2} \kappa d\kappa = \frac{1}{\sigma_\phi^2} \bar{\kappa}, \quad (8.89)$$

$$S_{xx} = \frac{1}{\sigma_\phi^2} \left[\frac{1}{M} \sum_{i=1}^M \kappa_i^2 \right] \rightarrow \frac{1}{\sigma_\phi^2} \frac{1}{\Delta\kappa} \int_{\bar{\kappa}-\Delta\kappa/2}^{\bar{\kappa}+\Delta\kappa/2} \kappa^2 d\kappa = \frac{1}{\sigma_\phi^2} \left[\bar{\kappa}^2 + \frac{\Delta\kappa^2}{12} \right]. \quad (8.90)$$

We have therefore that

$$\Delta = \frac{1}{\sigma_\phi^4} \frac{\Delta\kappa^2}{12}, \quad (8.91)$$

and it follows from Equation 8.85 that the variance of the slope of phase with respect to wavenumber is

$$\sigma_b^2 = 12 \frac{\sigma_\phi^2}{\Delta\kappa^2}, \quad (8.92)$$

independent of the number of pixels M . The rms variations in group delay can therefore be written

$$\sigma_{gd} = \frac{\sqrt{12}}{2\pi} \frac{\sigma_\phi}{\Delta\kappa}. \quad (8.93)$$

Matched Filter Derivation

One can also derive the signal-to-noise ratio for the amplitude group-delay estimator from a simple matched-filter argument. Assume the input data are the phases as a function of wavenumber

$$\phi = \phi_0 + (\kappa - \kappa_0) \frac{\partial\phi}{\partial\kappa}. \quad (8.94)$$

The orthonormal basis functions over the bandwidth $\Delta\kappa$ are respectively,

$$\frac{1}{\sqrt{\Delta\kappa}} \quad \text{and} \quad \sqrt{12} \frac{(\kappa - \kappa_0)}{(\Delta\kappa)^{3/2}}. \quad (8.95)$$

Thus

$$(\Delta\kappa)^{1/2}\phi_0 \quad \text{and} \quad \frac{1}{\sqrt{12}}(\Delta\kappa)^{3/2}\frac{\partial\phi}{\partial\kappa} \quad (8.96)$$

are each estimated with the same error.

8.6.3 Phase and Group Delay Variance Compared

From Equation 8.93 we now have that the ratio of standard deviations of the phase and group delay estimates is

$$\frac{\sigma_{\phi d}}{\sigma_{gd}} = \left[\frac{1}{2\pi} \frac{\sigma_{\phi}}{\kappa} \right] \left[\frac{\sqrt{12}}{2\pi} \frac{\sigma_{\phi}}{\Delta\kappa} \right]^{-1} = \frac{1}{\sqrt{12}} \frac{\Delta\kappa}{\kappa}, \quad (8.97)$$

where $\sigma_{\phi d}$ is the rms path fluctuation corresponding to phase variations σ_{ϕ} .

As an example of the difference between phase and group-delay variations, with PTI and an observation bandwidth of 2.0–2.4 μm and a mean observing wavelength of $\lambda = 2.2 \mu\text{m}$, we have $1/\Delta\kappa = 12 \mu\text{m}$, and can conclude that delay estimates derived from phase estimates will have rms variations 19 times smaller those derived from group-delay estimates.

8.7 Conclusion

The methods of phase and group delay estimation are routinely used in modern stellar interferometers to locate fringes and maintain the observations at a fixed location on the coherence envelope. Other methods of fringe measurement, in particular coherence envelope tracking, may also be used to estimate fringe parameters but are generally less efficient and more labour intensive.

In this Chapter we have reviewed the various methods of phase and group delay estimation that are currently being used in stellar interferometers, with emphasis on the techniques used at SUSI, PTI, and NPOI.

Acknowledgments

This work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. The author would like to thank Mark Colavita for discussions leading to the derivations of Section 8.6.2.

References

- J.T. Armstrong, D. Mozurkewich, L.J. Rickard, D.J. Hutter, J.A. Benson, P.F. Bowers, N.M. Elias II, C.A. Hummel, K.J. Johnston, D.F. Buscher, J.H. Clark III, L. Ha, L.-C. Ling, N.M. White, and R.S. Simon, “The Navy Prototype Optical Interferometer (NPOI),” *Astrophys. J.* **496**, 550–571 (1998).
- J.E. Baldwin, R.C. Boysen, G.C. Cox, C.A. Haniff, J. Rogers, P.J. Warner, D.M.A. Wilson, and C.D. Mackay, “Design and performance of COAST,” in *Amplitude and Intensity Spatial Interferometry II*, J.B. Breckinridge, ed., Proc. SPIE **2200**, 118–128 (1994).
- J.A. Benson (1998), private communication.
- J.A. Benson, D. Mozurkewich, and S.M. Jefferies, “Active optical fringe tracking at the NPOI,” in *Astronomical Interferometry*, R.D. Reasenberg, ed., Proc. SPIE **3350**, 493–496 (1998).
- G.D. Bergland, “A guided tour of the fast Fourier transform,” *IEEE Spectrum* **6**, 41–52 (1969).
- M. Born and E. Wolf, *Principles of Optics*, 6 edn. (Oxford, UK: Pergamon Press, 1980), Section 10.3.1.
- T.A. ten Brummelaar, “Differential path considerations in optical stellar interferometry,” *Appl. Opt.* **34**, 2214–2219 (1995).
- T.A. ten Brummelaar, “Correlation measurement and group delay tracking in optical stellar interferometry with a noisy detector,” *Mon. Not. R. Astron. Soc.* **285**, 135–150 (1997).
- D.F. Buscher, *Getting the most out of COAST*, Ph.D. thesis, University of Cambridge (1988).
- F. Cassaing, B. Fleury, C. Coudrain, P.-Y. Madec, E. Di Folco, A. Glindemann, and S. Lévêque, “Optimized fringe tracker for the VLTI/PRIMA instrument,” in *Interferometry in Optical Astronomy*, P.J. Léna and A. Quirrenbach, eds. Proc. SPIE **4006**, 152–163 (2000).
- M.M. Colavita, *Atmospheric limitations of a two-color astrometric interferometer*, Ph.D. thesis, Massachusetts Institute of Technology (1985).
- M.M. Colavita, J.K. Wallace, B.E. Hines, Y. Gursel, F. Malbet, D.L. Palmer, X.P. Pan, M. Shao, J.W. Yu, A.F. Boden, P.J. Dumont, J. Gubler, C.D. Koresko, S.R. Kulkarni, B.F. Lane, D.W. Mobley, and G.T. van Belle, “The Palomar Testbed Interferometer,” *Astrophys. J.* **510**, 505–521 (1999).
- K. Creath, “Phase-measurement interferometry techniques,” *Prog. Opt.* **26**, 349–393 (1988).
- G. Daigne and J.F. Lestrade, “Astrometric optical interferometry with non-evacuated delay lines,” *Astron. Astrophys. Supp. Ser.* **138**, 355–363 (1999).
- J. Davis, P.R. Lawson, A.J. Booth, W.J. Tango, and E.D. Thorvaldson, “Atmospheric path variations for baselines up to 80 m measured with the Sydney University Stellar Interferometer,” *Mon. Not. R. Astron. Soc.* **273**, L53–L58 (1995).
- J. Davis, W.J. Tango, A.J. Booth, E.D. Thorvaldson, and J. Giovannis, “The Sydney University Stellar Interferometer – II. commissioning observations and results,” *Mon. Not. R. Astron. Soc.* **303**, 783–791 (1999).

- J. Davis, W.J. Tango, and E.D. Thorvaldson, “Dispersion in stellar interferometry: simultaneous optimization for delay tracking and visibility measurements,” *Appl. Opt.* **37**, 5132–5136 (1998).
- E. Edser and C.P. Butler, “A simple method of reducing prismatic spectra,” *Phil. Mag.* **46**, 207–216 (1898).
- A. Feldman, “Determination of fringe order in the channel spectra of thin films,” *Appl. Opt.* **23**, 1193–1196 (1984).
- H. Fizeau and L. Foucault, “Sur le phénomène des interférences entre deux rayons de lumière dans le cas de grandes différence de marche,” *C. R. Acad. Sci.* **30**, 146–159 (1845).
- V. Coudé du Foresto, G. Perrin, and M. Boccas, “Minimization of fibre dispersion effects in double Fourier stellar interferometers,” *Astron. Astrophys.* **293**, 278–286 (1995).
- C.A. Hummel, “Practice of interferometry at NPOI,” in *Interferometry in Optical Astronomy*, P.J. Léna and A. Quirrenbach, eds. *Proc. SPIE* **4006**, 459–464 (2000).
- S.M. Kay and S.L. Marple Jr., “Spectrum analysis—a modern perspective,” *Proc. IEEE* **69**, 1380–1418 (1981).
- E.J. Kim, *Dispersed Fringe Group Delay Astrometry using the Mark III Stellar Interferometer*, M.S.E.E, Mass. Inst. Technology (1989).
- L. Koechlin, P.R. Lawson, D. Mourard, A. Blazit, D. Bonneau, F. Morand, Ph. Stee, I. Tallon-Bosc, and F. Vakili, “Dispersed fringe tracking with the multi- r_0 apertures of the Grand Interféromètre à 2 Télescopes,” *Appl. Opt.* **35**, 3002–3009 (1996).
- A. Labeyrie, “Interference fringes obtained on Vega with two optical telescopes,” *Astrophys. J.* **196**, L71–L74 (1975).
- M.G. Lacasse and W.A. Traub, “Glass compensation for an air filled delay line,” in *High-Resolution Imaging by Interferometry*, F. Merkle, ed., ESO Conf. Workshop, **29**, 959–970 (Garching, Germany: European Southern Observatory, 1988).
- P.R. Lawson, *Group Delay Tracking with the Sydney University Stellar Interferometer*, Ph.D. thesis, University of Sydney (1994).
- P.R. Lawson, “Group-delay tracking in optical stellar interferometry with the fast Fourier transform,” *J. Opt. Soc. Am. A* **12**, 366–374 (1995).
- P.R. Lawson, “Dispersion compensation and fringe tracking,” in *High Angular Resolution in Astrophysics*, A.-M. Lagrange, P. Léna, and D. Mourard, eds. *NATO ASI*, **3350**, 349–353 (Dordrecht: Kluwer Academic, 1997).
- P.R. Lawson and J. Davis, “Dispersion compensation in stellar interferometry,” *Appl. Opt.* **35**, 612–620 (1996).
- P.R. Lawson, T.R. Scott, and C.A. Haniff, “Group-delay tracking and visibility fluctuations in long-baseline stellar interferometry,” *Mon. Not. R. Astron. Soc.* **304**, 218–224 (1999).
- P.R. Lawson, M.M. Colavita, P.J. Dumont, and B.F. Lane, “Least-squares estimation and group delay in astrometric interferometers,” in *Interferometry in Optical Astronomy*, P.J. Léna and A. Quirrenbach, eds. *Proc. SPIE* **4006**, 397–406 (2000).

- S. Léveque, B. Koehler, and O. von der Luhe, “Longitudinal dispersion compensation for the very large telescope interferometer,” *Astrophys. Space. Sci.* **239**, 305–314 (1996).
- J.A. Meisner, “Atmospheric delay tracking in a long-baseline optical stellar interferometer,” *Opt. Eng.* **35**, 1927–1935 (1996).
- A.A. Michelson and F.G. Pease, “Measurement of the diameter of α Orionis with the interferometer,” *Astrophys. J.* **53**, 249–259 (1921).
- S. Morel and L. Koechlin, “Fringe tracking using a priori information on the optical path difference drift,” in *Astronomical Interferometry*, R.D. Reasenberg, ed., Proc. SPIE **3350**, 1057–1064 (1998).
- D. Mozurkewich (2000), private communication.
- P. Nisenson and W. Traub, “Magnitude limit of the group delay tracking method for long baseline interferometry,” in *Interferometric Imaging in Astronomy*, J.W. Goad, ed., 129–133 (Tucson, AZ: National Optical Astronomy Observatories, 1987).
- C.E. Padilla, V.I. Karlov, L.K. Matson, K. Soosaar, and T. ten Brummelaar, “High-performance fringe tracking algorithms utilizing statistical models at atmospheric turbulence,” in *Astronomical Interferometry*, R.D. Reasenberg, ed., Proc. SPIE **3350**, 1045–1056 (1998).
- W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes in C* (Cambridge, UK: Cambridge University Press, 1992).
- R.J. Sanderman, “Use of channeled spectra to measure absolute phase shift and dispersion in two beam interferometry,” *Appl. Opt.* **10**, 1087–1091 (1971).
- M. Shao and M.M. Colavita, “Long-baseline optical and infrared stellar interferometry,” *Ann. Rev. Astron. Astrophys.* **30**, 457–498 (1992).
- M. Shao, M.M. Colavita, B.E. Hines, D.H. Staelin, D.J. Hutter, K.J. Johnston, D. Mozurkewich, R.S. Simon, J.L. Hershey, J.A. Hughes, and G.H. Kaplan, “The Mark III stellar interferometer,” *Astron. Astrophys.* **193**, 357–371 (1988).
- W.H. Steele, *Interferometry* (Cambridge, UK: Cambridge University Press, 1987).
- W.J. Tango, “Dispersion in stellar interferometry,” *Appl. Opt.* **29**, 516–521 (1990).
- W.J. Tango and R.Q. Twiss, “Michelson stellar interferometry,” *Prog. Opt.* **17**, 239–277 (1980).
- W.A. Traub, “Recent results from the IOTA interferometer,” in *Astronomical Interferometry*, R.D. Reasenberg, ed., Proc. SPIE **3350**, 848–855 (1998).
- W.A. Traub, M.G. Lacasse, and N.P. Carlton, “Spectral dispersion and fringe detection in IOTA,” in *Amplitude and Intensity Spatial Interferometry*, J.B. Breckinridge, ed., Proc. SPIE **1237**, 145–152 (1990).
- J.F. Walkup and J.W. Goodman, “Limitations of fringe parameter estimation at low light levels,” *J. Opt. Soc. Am.* **63**, 399–407 (1973).
- J. Wyant, “Use of an ac heterodyne lateral shear interferometer with real-time wavefront correction systems,” *Appl. Opt.* **14**, 2622–2626 (1975).

