Visible Light Fringe Tracking

T. A. ten Brummelaar
CHARA, Georgia State University, Atlanta, GA 30303

I.1. INTRODUCTION

In order to achieve phase closure, or even simple visibility magnitude measurement, the OPLE system must be controlled such that the differential delays between the seven arms of the Array are much smaller than the fringe envelope. While for small baselines it should be possible to use dead-reckoning based on an astrometric model to keep all baselines within the fringe coherence length, this will become very difficult at the larger baselines. Tracking fringes requires generating some kind of error signal that, for each baseline, measures the current position within the fringe envelope and how far away and in which direction the maximum lies. This is a separate task from measuring fringe visibilities, and in the CHARA Array it has been decided that the fringe tracking subsystem will be physically, as well as functionally, separate from the visibility/phase closure subsystems.

The Array visible fringe tracker will make use of Group Delay Tracking (GDT), that is, looking for channel fringes in the dispersed spectra of combined beams and locking onto a pre-determined number of fringes across the spectrum. To our knowledge, this technique is least sensitive to noise and has been used and shown to work on other interferometers, such as the Sydney University Stellar Interferometer (SUSI). It also has the advantage of not requiring any temporal path length modulation, thereby ensuring that it does not restrict either the bandwidth or the choices for beam combination and visibility measurement downstream.

I.2. SEPARATING THE FRINGE TRACKER FROM THE IMAGING SYSTEM

If one were to insist that the fringe tracker utilize the same data as the visibility measurement subsystem, it would unduly complicate the design problem. For example, the fringe tracker must function in real time while this is not necessarily true of the imaging subsystem. By developing the two subsystems independently, they can be optimized for their respective tasks. For example, the fringe tracking subsystem could use a wide optical bandwidth and short sample times in order to lock onto the fringes, while the imaging subsystem could use longer sample times, a smaller optical bandwidth, and higher spectral dispersion. In this way the variation of source structure with wavelength can be studied, and the effect of readout noise in the CCD detector minimized.

Furthermore, the development of independent subsystems can be staggered over time; a fringe tracker can be developed and operational long before imaging is implemented at visible wavelengths. The fringe tracker can be used to produce visibility magnitude measurements, and thus useful science will not be delayed as the more challenging imaging system is brought on line. If the two systems are separate, new techniques for either fringe tracking or imaging, when they become available, could be implemented without disturbing the other subsystem. Tying the two systems together into one package would provide no obvious developmental path.

I.3. BASIC CONCEPTS OF GDT

Group delay tracking has been well described by others, most notably in the theses by Lawson (1993) and Buscher (1988). If two beams are combined such that there are no, or minimal, residual phase difference and tilt, the dispersed spectrum will be channeled with fringes. These fringes will run perpendicular to the direction of dispersion and will be equally spaced in wavenumber. Appendix N contains a discussion of dispersion effects internal to the interferometer and derives an expression for the resulting fringe pattern in each pixel of the detector system, a simplified version of which is

$$I(\nu) = I_s \left[1 + |V(\nu)| \cos\left(2\pi\nu \,\mathrm{OPD} - \phi\right)\right] + I_b \tag{I.1}$$

where ν is the wavenumber ($\frac{1}{\lambda}$), I_s is the intensity of the unmodulated stellar spectrum, OPD is the optical path length difference or delay, $|V(\nu)|$ is the wavenumber-dependent fringe visibility magnitude, ϕ is visibility phase, and I_b is the background intensity.

Bright fringes will appear when the optical path length difference is an integral number of wavelengths, while dark bands will be apparent when the delay causes destructive interference. Thus, if the wavenumber is held constant, fringes can be observed by changing the optical path length difference. Alternatively, fringes can be observed by keeping the optical path length difference constant and changing the wavenumber. These later fringes, each of equal chromatic order, are called channel fringes and the resulting spectrum a channeled spectrum. As the optical path length difference (or delay) is increased, the number of channels across any given bandwidth will also increase. Thus, by counting the number of channel fringes across a fixed bandwidth, one can measure the delay. The number of fringes across the spectrum is given by

$$n = \nu_{\max} \,\mathrm{OPD} - \nu_{\min} \,\mathrm{OPD} \tag{I.2}$$

and therefore

$$\mathrm{OPD} = \frac{n}{\nu_{\max} - \nu_{\min}}. \tag{I.3}$$

The optical path length difference is a linear function of the number of fringes across the spectrum. One effective and simple way of measuring n is to measure the fringe frequency, that is, the number of fringes per unit wavenumber. If the sample pixels are equally spaced in wavenumber, a Fourier transform of the channel spectrum would contain a zero frequency component, corresponding to the total intensity in the frame, along with a peak corresponding to the fringe frequency. Since the detectors will consist of an array of pixels, the discrete or Fast Fourier Transform (FFT) will need to be used, bringing with it a number of associated signal-to-noise issues. These are discussed at length by Lawson (1993), while Appendix J contains an analysis of the performance of the GDT approach and Appendix P contains a discussion of detectors that could be used in the fringe tracker. We plan to use bare, low-noise, CCD chips and use on-chip binning to achieve pixels approximately equally spaced in wavenumber.

I.4. GDT VERSUS PATH LENGTH MODULATION METHODS

Group delay tracking is not the only method available for fringe tracking. Most alternatives, however, involve modulation of the path length of each beam. When only two or three telescopes or siderostats are used in an array, such schemes are advantageous because they

allow the direct determination of the fringe phase. Unfortunately, with seven telescopes to phase these methods are too cumbersome. One such technique, first described by Wyant (1975), that was successfully used in the Mark III interferometer (Shao et al. 1988) uses a sawtooth wave modulation in one arm of a single baseline interferometer. The detected intensity is measured and binned into four temporally equal parts I_A, I_B, I_C, and I_D. If the modulation amplitude is equal to one wavelength, the visibility magnitude and delay can be calculated via (Shao & Staelin 1980)

$$\mathrm{OPD} = \frac{1}{2\pi\nu}\left[\tan^{-1}\left(\frac{I_A - I_C}{I_D - I_B}\right) + \phi\right] \tag{I.4}$$

$$|V(\nu)| \propto \frac{(I_A - I_C)^2 + (I_D - I_B)^2}{I_A + I_B + I_C + I_D}. \tag{I.5}$$

While this method has an inherent 2π ambiguity in the phase measurement, it is possible, assuming the fringe phase does not move by more that π in one sample time, to unwrap the correct phase. As it assumes the modulation introduced is one wave in length, this method is wavelength dependent, although it is possible to perform phase tracking at several wavelengths at once by re-binning and adding dead time to various parts of the modulation.

There are a number of reasons for not choosing a path length modulation method for fringe tracking in the CHARA Array:

- While it is possible to use these systems at multiple wavelengths, they are inherently a single wavelength method. CHARA is intended to be used over a very large bandwidth, and it is unlikely that path length modulation methods will be suitable.
- A large signal to noise ratio is required for these methods to work. For example, numerical simulations by Buscher (1988) and Lawson (1993) indicate that while GDT fails at a signal to noise ratio of 0.7, path length modulation methods require a signal to noise ratio of 3.5. The Mark III fringe tracking subsystem failed when less than ~ 30 photons were detected (Shao & Staelin 1980). Numerical modeling by Lawson (1993) shows that by using GDT with a noiseless detector, active tracking would be possible at ~ 10 photons per frame and passive detection at ~ 3.
- In GDT there is no 2π ambiguity in the fringe location. While it is possible to unwrap these ambiguities in path length modulation schemes, a 2π error in phase measurement can go completely undetected.
- GDT results in a much larger coherence length or fringe envelope size. Looking at channel spectra is probably the oldest method for finding fringes in an interferometer; indeed it was used by Michelson & Pease (1921). Even when the delay is very large, channel fringes can still be observed. The delay must already be very small for path length modulation methods to work.
- Path length modulation methods would not supply visibility magnitude estimates across many wavelength channels, as the proposed GDT system will do.
- As discussed in Section I.7, the use of GDT allows for a graceful decline in imaging information as the number of available photons falls off. A path length modulation method would simply fail at some point, stopping all observation.
- While the implementation of the modulation technique is direct and simple for a single pair of telescopes, when seven telescopes and 21 baselines are involved it is not

immediately apparent how to perform the temporal fringe encoding without losing all fringe information in a maze of path length modulations. Furthermore, it is not clear how one would isolate the imaging subsystem from these modulations. Unless a completely separate interferometer channel and mini-OPLE system were used for each baseline, further reducing the amount of light available for each subsystem, path length modulation methods would become too involved and complex for use in the CHARA Array. The primary advantage of using temporal encoding is the low number of pixels required in the detector.

I.5. OPTICAL LAYOUT

The basic optical layout of the proposed fringe tracking subsystem is shown in Figure I.1 and a hardware tree is given in figure Figure I.2. While all the beams have been shown to be either horizontal or vertical in this diagram, the final design of the fringe tracker will employ reflections less than 90 degrees in order to minimize any potential polarization problems.

After having passed through the beam sampling subsystem, the seven visible light beams enter the fringe tracker subsystem. The fringe tracking optics are based on a modified Mach-Zehnder interferometer, where instead of splitting a single beam and recombining it with itself, we combine two initially parallel beams. This means that there is an offset in each baseline which will have to be compensated in the OPLE. The result of this offset is that the line of equal phase along the seven beams is not perpendicular to the beams but lies along a diagonal. This effect has implications for the alignment procedures used (see Appendix M) and for any optical subsystem downstream (see Appendix K).

Since polarization effects usually result in a loss of correlation, it has been decided to use polarizing beam splitters to divide the light between the fringe tracking and imaging subsystems. Thus each system will receive a single polarization. This makes the specification of the optics used in each system, beam splitters for example, less strict as they are only required to deal with a single polarization. A subset of seven baselines will be sampled in a pairwise fashion within the fringe tracking optics. Each set of combined beams is then spectrally dispersed by a prism and imaged onto the detector system, a bare CCD (refer to Appendix J). Note that seven sets of spectra are imaged onto each detector system, thereby reducing the number of array detectors required. Different wavelength bands can be selected for tracking by adjusting the position of the prisms. A design option here is to put each of the combined beams into multimode fibers. Since the beams are already combined, single-mode fibers wll not be necessary. These fibers can be brought together and all fourteen signals can pass through a single prism and CCD. The drawback of this approach is the light lost due to inefficient coupling to the fibers.

Using the prisms to create a spectrum in the horizontal plane and placing all seven spectra onto one detector will cause some of the image planes to be rotated with respect to the detector surface. Ray tracing has shown that the defocusing caused by this rotation is minimal and should not affect the performance of the device. It would also be possible to use the prisms to create spectra in the vertical plane which, while making the mounts used somewhat more complex, would avoid this problem altogether. Another potential problem with using prisms as a dispersive element is that the resulting spectra are not linear in wavenumber, as required for GDT. This may be overcome by on-chip or software driven binning of the spectra in the CCD detector systems. This binning process will

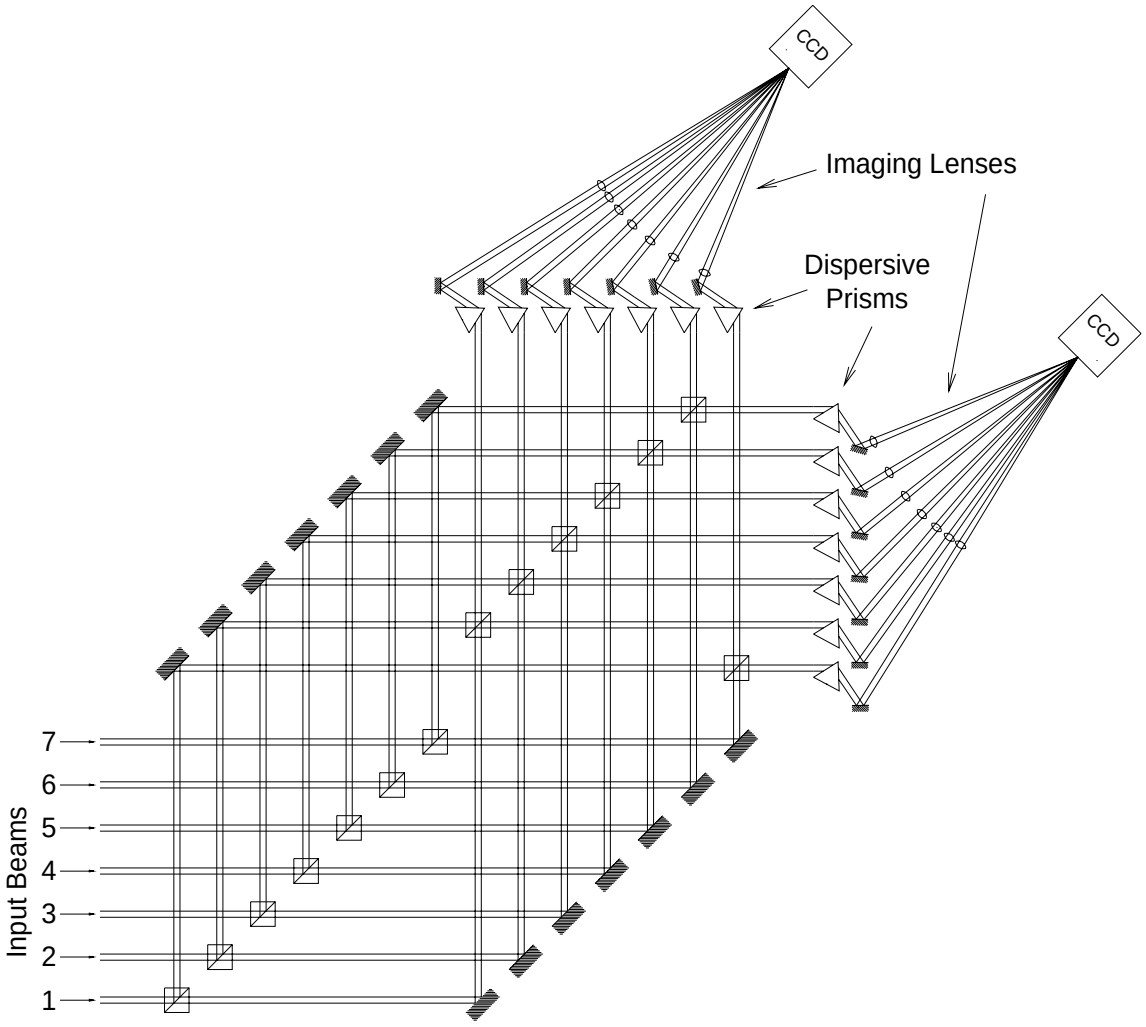

FIGURE I.1. Optical layout for the fringe tracking subsystem. Note that reflections are shown to be 90 degrees. In the real device these angles of reflection will be much less than this in order to minimize polarization problems.

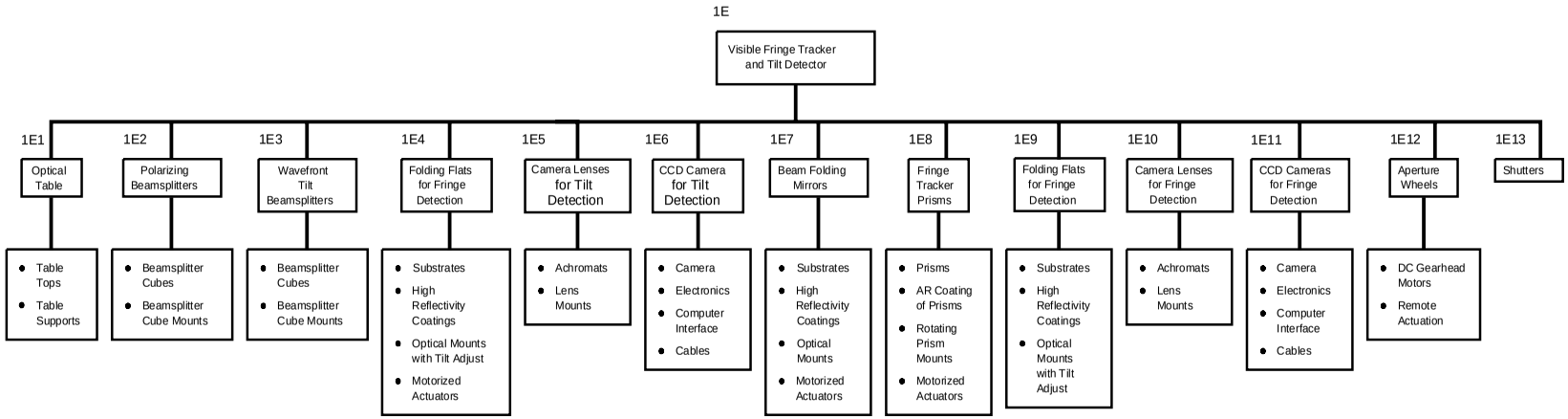

FIGURE I.2. Hardware tree for the fringe tracking subsystem.

be programmable and can therefore be adjusted to suit the current observing conditions (Lawson 1993). For example, if too few photons per pixel are available larger bins could be used. The subsystem performance can in this way be made to degrade gracefully as the system becomes photon starved.

Since it is intended to also use the visible fringe tracking subsystem to make visibility magnitude measurements, a set of APDs will also be included in the optical layout. These will be used instead of the CCD cameras for photon counting at low light levels.

I.6. BASELINE SAMPLING

The GDT scheme requires a pairwise combination of the beams, and unless we use many more beamsplitters (thereby substantially increasing the complexity of the device and further reducing the amount of light allocated to each pair) the maximum number of baselines represented in the fringe tracker will be seven. Fortunately, unlike imaging, the fringe tracker need only have enough data to ensure that each telescope/OPLE combination is correctly phased. Thus it is possible to maintain mutual coherence along all baselines using only a subset of seven baselines, as long as this subset includes each telescope at least once. This baseline sample can be selected by changing the beam switching mirrors in the beam sampler subsystem (refer to Appendix H).

In order to ensure that the maximum visibility, and therefore signal to noise ratio, be present in each fringe tracking channel, the baseline subset chosen for GDT will be those representing the smallest baselines possible that still include all telescopes. By keeping these pairs of beams correctly phased, the remaining baselines will also be in phase.

I.7. FRINGE TRACKING AND VISIBILITY MEASUREMENTS

By separating the fringe tracking and imaging subsystems, as well as having control over the optical bandwidth and on-chip binning, the Array performance can be made to degrade gracefully as the available photon flux reduces with increasing object magnitude. Using a combination of the fringe tracking subsystem and the imaging subsystem, five regimes of fringe measurement are planned: three using the imaging subsystem and two using the fringe tracker.

I.7.1. Visibility Measurements with the Imaging Subsystem

When enough light is available to either actively track or passively measure delay, the imaging subsystem can be used to collect visibility data. The imaging data collection scheme is described in Appendix K. As the magnitude of the target object increases, fewer and fewer photons will be available for both fringe tracking and imaging. Depending on the available photon flux one would choose between the following three methods for data collection:

1. Active Fringe Tracking

   This will be the preferred method for data collection in the CHARA Array. For bright objects and/or objects for which the compact array will be used, the fringe tracking subsystem should be capable of actively measuring the delay every sample and

therefore track the white light fringe position. This means that the visibility/phase closure measurement subsystem downstream will be able to integrate relatively long frames over large enough periods to obtain high SNR visibility magnitude and phase closure data.

2. Semi-Active Fringe Tracking

   For fainter objects, a reduced photon flux will reach the fringe tracker, making it difficult to reliably track the fringes. It will then be necessary to start integrating over many samples in the fringe tracker, reducing the bandwidth of the tracking servo. Thus the frame times may have to be smaller and integration times longer in the imaging subsystem, but both visibility magnitude and phase closure data will be available. Of course, if the tracker looses the fringes, the imaging subsystem would be shut down until the fringes are once again stable.

3. Passive Fringe Tracking

   If the photon flux is further reduced, the fringe tracking subsystem will no longer be able to actively track the fringes. By using short sample times and long integration, an average delay measurement could still be made. While this will do little to aid tracking, such an average delay could be used to renormalize the visibility measurements. The phase closure data would probably be lost, but the imaging subsystem would still be capable of obtaining visibility magnitude data. Short sample times with integration over many samples would be necessary but, since the imaging subsystem is being used, all baselines would be simultaneously available.

I.7.2. Visibility Measurements with the Fringe Tracker

Once the photon flux is so low that it is not possible to use the imaging subsystem at all, the fringe tracking subsystem can become a delay curve measurement device. This method was first described by Tango & Twiss (1980) and was used in the Monteporzio interferometer (Tango 1979), the Sydney University 11.4m Prototype (Davis & Tango 1985), and is currently being used in the Sydney University Stellar Interferometer (Davis et al. 1992).

Using the fringe tracker for visibility measurement means that only seven baselines can be sampled at any given time, but since we are no longer trying to actively track fringes we are no longer required to use the subset representing the smallest baselines. We can therefore observe objects over several nights, each time sampling a different set of baselines. Longer observational runs will be necessary but all baseline data will be available.

Basically, the technique relies on averaging over many small sample times and allowing the atmosphere to move the fringe envelope around the mean tracking position. Following Section I.3, if we consider one pair of narrow band spectral channels (corresponding to one pixel in each of the two sides of a given beam splitter) the intensity of light measured in each pixel will be

$$I_A \propto 1 + |V(\nu)| \cos(2\pi\nu \,\mathrm{OPD} - \phi) \tag{I.6}$$

and

$$I_B \propto 1 - |V(\nu)| \cos(2\pi\nu \,\mathrm{OPD} - \phi) \tag{I.7}$$

where the difference in sign is due to the beamsplitter used to combine the two beams. Thus, when fringes are present, if one channel is bright the other channel will be dark. The

visibility is estimated by integrating the term

$$[|V(\nu)|cos(2\pi\nu\mathrm{OPD} - \phi)]^2 = \frac{(I_A - I_B)^2}{(I_A + I_B)^2} \tag{I.8}$$

over many sample times, each smaller than the time constant of the atmosphere. Since the atmosphere will cause the phase term ϕ to change at random, while still not moving the fringes far compared to the fringe envelope size, this term will average out, thereby yielding a measurement of the current visibility magnitude. A fit of the fringe envelope shape can be performed once this measurement has been made at several delays and an estimate for the value at the white light position obtained.

In order for delay curve measurement to work, very narrow bandwidths must be used, and large integration times are required. The strategy for the CHARA Array is to have many small bandwidth pixels to build up the equivalent of a large bandwidth device. As discussed in Appendix N each channel will have a large fringe envelope size, making GDT and delay curve measurement possible. Note that the array geometry must be precisely known so that the OPLE tracking rates do not cause fringe wander during visibility measurement. Once many bright objects have been tracked by one of the methods outlined above, however, good baseline solutions will be available, reducing the potential errors from this source.

Since this technique for visibility measurement is well understood and relatively easy to implement once the fringe tracker is built, it will probably be the first method used to acquire useful scientific data from the Array. It also represents a low technological risk. There are two potential methods for making use of delay curve measurement techniques, depending on whether there is enough light to close the fringe tracker/OPLE servo loop or not.

1. Active Fringe Envelope Tracking

   In order to get enough photons to measure the delay at low light levels, sample times longer than the atmospheric coherence time t_0 will have to be used. This means that the fringe tracker will only produce an 'average' delay measurement, but one that will be good enough to keep the system close to the fringe envelope center. Thus the visibilities can be measured at, or close to, the envelope maximum with any offset being logged and available for post-processing.

2. Passive Fringe Envelope Tracking

   The final data collection option for the CHARA Array is for when so little light is available that active tracking of the fringe envelope will not be possible. This means that the OPLEs will be required to track an astrometric model of the Array. The integration times in the fringe tracker will be even greater, although not so large that all coherence is lost in the system altogether. In this case, the GDT data is simply recorded rather than used for fringe tracking, and used in post-processing of the data. If nothing else, the GDT data will be an indicator of 'good' and 'bad' data and give a measurement of how far, on average, one is from the fringe envelope maximum. In the best case, the recorded delay information can be used to re-normalize the visibility measurements. In this situation the entire fringe envelope would need to be measured and a curve fitting process used to estimate the visibility maximum.

I.8. REFERENCES

Buscher, D., 1988, "Getting the most out of the COAST," Ph.D. thesis, Cambridge University

Davis, J. & Tango, W. J., 1985, "The Sydney University 11.4m Prototype Stellar Interferometer," *Proc. Astr. Soc. Aust.*, **6**

Davis, J., Tango, W. J., Booth, A. J., Minard, R. A., ten Brummelaar, T. A., & Shobbrook, R. R., 1992, "An update on SUSI", in *High Resolution Imaging by Interferometry II* (E.S.O. Garching)

Lawson, P. R., 1993, "Group Delay Tracking with the Sydney University Stellar Interferometer," Ph.D. thesis, University of Sydney, submitted

Michelson, A. A. & Pease, F. G., 1921, "Measurement of the diameter of α Orionis with the interferometer," *ApJ*, **53**, 249

Shao, M. Colavita, M.M., Hines, B.E., Staelin, D.H., Hutter, D.J., Johnston, K.J., Mozurkewich, D., Simon, R.S., Hershey, J.L., Hughes, J.A., & Kaplan, G.H., 1988, "The Mark III stellar interferometer," *A&A*, **193**, 357

Shao, M. & Staelin, D. H., 1980, " First fringe measurements with a phase tracking stellar interferometer", *Appl.Optics*, **19**, 1519

Tango, W. J., 1979, "The Monteporzio two meter amplitude interferometer," in Proc IAU Colloq. No. 50, *High Angular Resolution Stellar Interferometry*, (Univ of Sydney)

Tango, W. J. & Twiss, R. Q., 1980, "Michelson Stellar Interferometry", *Progress in Optics*, **XVII**

Wyant, J. L., 1975, "Use of an ac heterodyne lateral shear interferometer with real time wavefont correction systems", *Appl.Optics*, **14**