Astron. Astrophys. 193, 357–371 (1988)

ASTRONOMY AND ASTROPHYSICS

The Mark III stellar interferometer

M. Shao [1], M.M. Colavita [1], B.E. Hines [2], D.H. Staelin [2], D.J. Hutter [3], K.J. Johnston [3], D. Mozurkewich [3], R.S. Simon [3], J.L. Hershey [4], J.A. Hughes [4], and G.H. Kaplan [4]

[1] Smithsonian Astrophysical Observatory, 60 Garden Street, Cambridge, MA 02138, USA
[2] Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, Cambridge, MA 02139, USA
[3] E.O. Hulburt Center for Space Research, Naval Research Laboratory, Washington, DC 20375, USA
[4] US Naval Observatory, Washington, DC 20390, USA

Received May 25, accepted July 23, 1987

Summary. The Mark III interferometer is an operational long baseline stellar interferometer on Mt. Wilson with four possible baseline configurations from 9 m NE-SW to 20 m N-S. The interferometer was designed to be a highly automated astronomical instrument to measure stellar positions and diameters to a magnitude limit of seven. Initial fringe observations were made in September 1986 with a 12-m N-S baseline. In the following months, semi-automated astrometric and stellar diameter measurements were also made. This paper describes the hardware and software components of the instrument and its operational characteristics.

The interferometer has several novel features. One is the use of optimal estimation and control algorithms (e.g. Kalman filters) in the control loops. Another is the ability to operate both as a closed-loop phased interferometer and eventually as an open-loop or absolute coherent interferometer. High thermal stability and mechanical accuracy should permit the instrument to point blind at an astronomical object and maintain optical path equality to within the limits set by the atmosphere. In this absolute interferometric mode of operation, it should be possible to observe faint astronomical objects that are too dim for phase tracking. In theory, measurements of amplitude, group delay, and closure phase will be possible to 14 mag.

Key words: stellar interferometry – astrometry – stellar diameters – optical array – proper motion

1. Introduction

The Mark III interferometer is the third of a series of stellar interferometers at Mt. Wilson Observatory. The first instrument demonstrated white-light fringe tracking in 1979 (Shao and Staelin, 1980). The second instrument observed fringes in 1982 and was used as a technology test bed for astrometric measurements until 1984 (Shao et al., 1987). Several devices critical to the operation of a stellar interferometer were developed for the Mark II interferometer. One such device was a high speed, ultra high accuracy, laser controlled optical delay line. An operational procedure for rapid switching between stars was also developed. Traditionally, the most difficult task of a stellar interferometer has been finding and tracking the central fringe. In rapid switching, this procedure is totally automated. Thus, operation of the Mark II helped identify the technical problems of making astrometric measurements with a long baseline interferometer.

Like the previous instrument, the Mark III interferometer was also designed to be a test bed for new techniques and technologies relevant to optical interferometry and synthetic optical aperture synthesis. However, the principal purpose of the Mark III interferometer was to initiate a test program of fundamental astrometric measurements to demonstrate a significant improvement in accuracy. Another goal was to initiate a program of accurate stellar diameter measurements. Most important of all, the instrument was designed to be astronomically productive.

Long baseline stellar interferometers have traditionally been considered one of the most technologically challenging astronomical instruments. In the Mark III interferometer, the goal has been to build an instrument that is reliable, easy to operate, and capable of extremely accurate astronomical measurements. In order to achieve these goals, a number of active subsystems were incorporated into the design of the instrument. In implementing these subsystems, the design philosophy involved keeping the mechanical and optical systems as simple as possible and shifting the bulk of the complexity into software.

Although the Mark III was originally designed as a closed loop phased interferometer, operational experience has shown that it has sufficient accuracy and stability to operate as an absolute interferometer. An absolute interferometer is one whose mechanical accuracy and stability in open loop operation, without fringe tracking, is sufficiently high that optical path errors are dominated by atmospheric turbulence. This level of accuracy and stability, usually associated with the primary mirror of a large astronomical telescope, is the level of accuracy exhibited by the Mark III interferometer with the 12-m baseline. Operation with the 20-m baseline is expected to exhibit the same level of accuracy and stability.

The ability to operate as an absolute interferometer makes possible three new modes of interferometric observations for long baseline interferometers. They are photon starved amplitude measurements, photon starved group delay astrometry, and photon starved closure phase measurements. Photon starved amplitude measurements have long been a part of speckle interferometry with single apertures. The importance of photon

Send offprint requests to: M. Shao

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
1988		**00-00-1988 to 00-00-1988**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
The Mark III Stellar Interferometer	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **U. S. Naval Observatory,Washington,DC,20390**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Astronomy and Astrophysics, vol 193, pgs 357-371, 1988

14. ABSTRACT

The Mark III interferometer is an operational long baseline stellar interferometer on Mt. Wilson with four-possible baseline configurations from 9m NE-SW to 20m N-S. The interferometer was designed to be a highly automated astronomical instrument to measure stellar positions and diameters to a magnitude limit of seven. Initial fringe observations were made in September 1986 with a 12-m N-S baseline. In the following months, semi-automated astrometric and stellar diameter measurements were also made. This paper describes the hardware and software components of the instrument and its operational characteristics. The interferometer has several novel features. One is the use of optimal estimation and control algorithms (e.g. Kalman filters) in the control loops. Another is the ability to operate both as a closed-loop phased interferometer and eventually as an open-loop or absolute coherent interferometer. High thermal stability and mechanical accuracy should permit the instrument to point blind at an astronomical object and maintain optical path equality to within the limits set by the atmosphere. In this absolute interferometric mode of operation, it should be possible to observe faint astronomical objects that are too dim for phase tracking. In theory, measurements of amplitude, group delay, and closure phase will be possible to 14 mag.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**15**	

Fig. 1. The Mark III interferometer on Mt. Wilson

starved observations is that they make possible astrometry and aperture synthesis of faint extragalactic objects. These three types of measurements are discussed after the description of the instrument.

2. The instrument

The interferometer is located in California on Mt. Wilson approximately 80 m east of the 100-inch telescope, near the site of the Mark II interferometer. Figure 1 shows a view of the site. Six and ten meters north, south, and east of the central building are the siderostat piers, which are each constructed of steel-reinforced concrete with a 4 × 4 foot base and a height of 8 feet, 4 feet of which is above ground. Siderostats are currently mounted on the innermost piers. Huts provide weather protection for the siderostats. These have hinged roofs which swing back exposing the siderostats to the sky. Currently, there are vacuum pipes between the huts and the central building to eliminate local seeing effects caused by heat in the main building. Initial operation involved only the 12-m baseline connecting the north and south siderostats; however, the east siderostat can be used in place of the north siderostat for a 9-m NE-SW baseline. The fourth window on the central building is for future expansion. In 1988, we plan to operate on the outer set of piers shown in Fig. 1, with 20-m N-S and 16-m NE-SW baselines.

Since the instrument's principal purpose is large angle astrometry, attention was paid to construction practices that affect astrometric accuracy. The central building is air conditioned during the day so that at night there will be a minimum temperature difference between inside and outside. The piers in the huts that hold the siderostats go down to bedrock and are not directly connected to the concrete pads on which the huts are built. The piers are also insulated by two inches of foam. The optics inside the main building that affect astrometric measurements are all located on a super-invar breadboard, except the mirrors in the optical delay line. However, the position of the optical delay line is monitored by a laser interferometer referenced to the same super-invar breadboard. The laser is frequency stabilized with better than 10^{-9} long-term stability.

The instrument can be divided into five subsystems: 1) the star tracker or autoguider, 2) the optical delay line, 3) the stellar fringe acquisition and tracking system, 4) the laser metrology system, and 5) the siderostat pointing and observation sequencing system. A description of the optical layout of the instrument is followed by a more detailed explanation of the five subsystems.

2.1. *Optical layout*

Figure 2 is an optical schematic of the interferometer. Light from the two siderostats is directed toward the central building by fixed mirrors. These fixed mirrors are necessary to keep the polarization vectors from the two interfering beams parallel for observation of stars located over all parts of the sky. After the light enters the main building, the beams are directed by piezoelectrically controlled mirrors towards the vacuum delay lines. After the light reflects from the delay line retroreflectors, the beams are combined at the beamsplitter B. After the beamsplitter, part of the light is directed into optical fibers which feed GaAs phototubes located some distance from the central breadboard, and part of the light is directed toward the star tracker. Other optical components include the laser interferometers, and alignment lasers and telescopes.

2.2. *Star tracker*

The star tracker has been described in detail previously (Clark et al., 1986; Clark, 1986) and we will provide only a brief summary here. The star tracker system is designed to keep the two interfering wavefronts parallel to within a small fraction of an arcsecond. As with all servo systems, the star tracker is composed of an error sensor, a controller, and an actuator. For the Mark III interferometer, two different error sensors were built because a key component (a photon counting camera) for the original design was not delivered before initial observations were made in September 1986. Only the star tracker design that will be permanently attached to the interferometer is described.

An optical schematic of the star tracker is shown in Fig. 3. Light from the star in a 3-inch beam, intercepted by the two siderostats, is combined at a beamsplitter. Light from the south siderostat passes through an annular wedge that deflects the outer portion of the beam by 45″; the inner 2-inch part of the beam is undeflected. After the light is combined at the beamsplitter, an annular mirror at one output of the beamsplitter directs the light in the annulus towards a telescope and imaging detector. The light in the central 2-inch core from both sides of the beamsplitter is fed to photomultipliers that detect the fringes.

When the two components of the light beam that enter the fringe detection phototubes are parallel (as they must be for proper operation of the interferometer), the two components of the annular beam entering the telescope will be tilted by 45″ relative to each other because of the 45″ wedge in the south arm of the interferometer. Thus two star images will appear on the face of the photon counting imaging detector. The optical arrangement of this star tracker is slightly different from that used for some other interferometers. The present design using one telescope for both star trackers eliminates the need to align the two optical systems to each other with subarcsecond accuracy. More important, the relative alignment of two separate optical systems would drift with time and temperature, whereas the star trackers for the Mark III interferometer will maintain alignment indefinitely.

The photon camera used in the star tracker is a commercial detector based on the PAPA camera (Papaliolios et al., 1985) with 256 × 256 pixels. The detector produces a pulse with an x, y address for every detected photon. The digital output of the

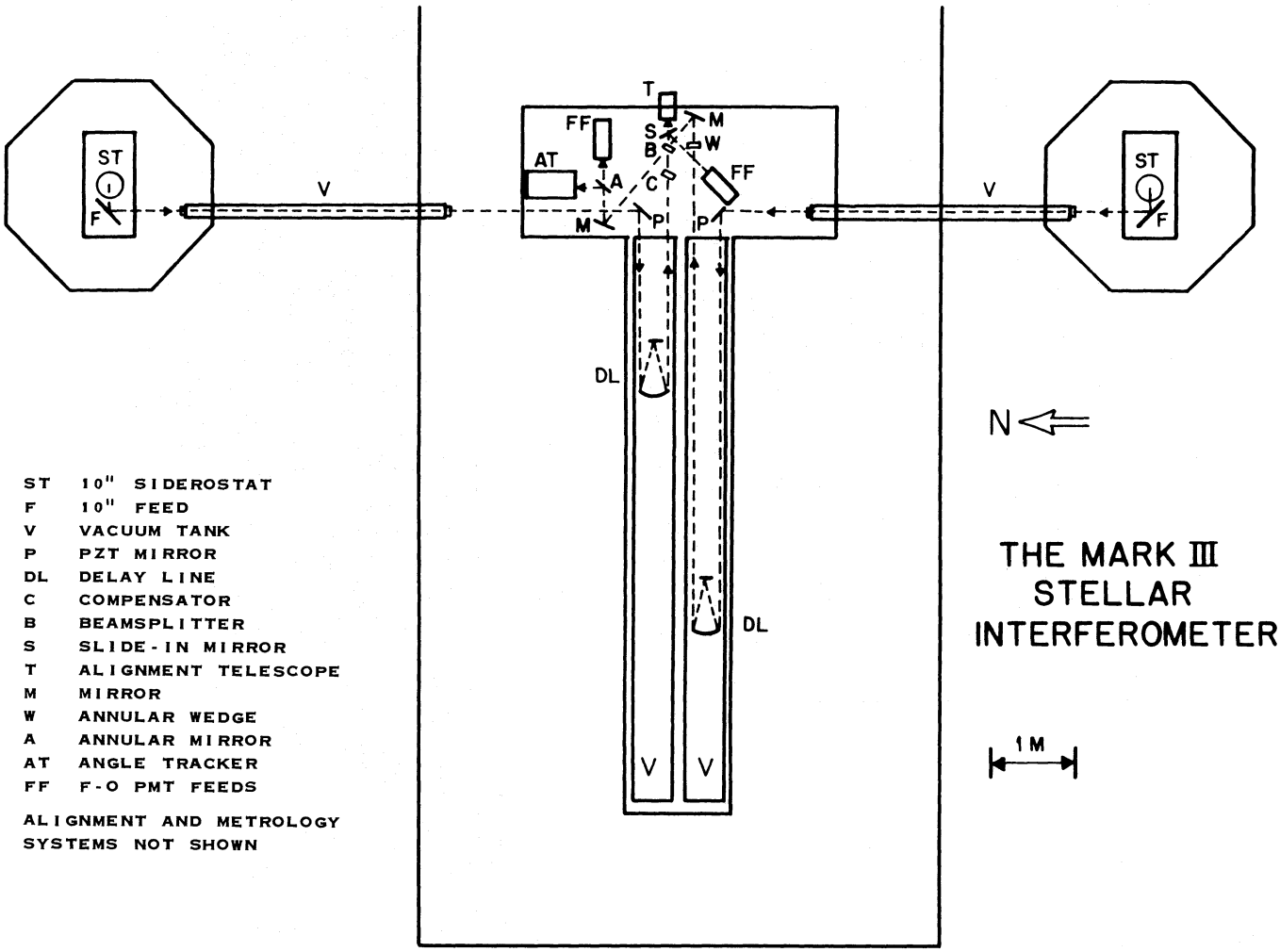

Fig. 2. Optical schematic of the Mark III interferometer

camera is interfaced to a high speed signal processor which locates the centroids of the two stellar images. Two additional optical/mechanical components are shown in Fig. 3. The variable neutral density filter is used to reduce the photon flux to the camera to prevent overloading of the camera electronics when observing bright stars. The setting of the filter is controlled by the star tracker computer. There are also two shutters, one in each arm of the interferometer, which are used for star acquisition. When the siderostats stop slewing, the two stellar images may be sufficiently far from their correct positions that proper identification, i.e., which star came from which siderostat, may not be possible. The shutters are used to remove this ambiguity.

The camera is the error sensor. The actuator is a PZT-driven gymbaled mirror. This 5-inch mirror is driven in two orthogonal directions with a 30″ peak-to-peak motion. The bearings are cross flexure pivots and the mirror mount has its lowest resonance at 200 Hz.

The heart of the star tracker is the control system; it makes extensive use of optimal estimation techniques and only a brief outline is given here. A detailed description is given by Clark et al. (1986). The control problem is separated into a spatial filtering part and a temporal filtering part. The spatial filter takes the photon addresses generated by the photon camera and derives the centroid of the star. The temporal filter takes current and past centroid estimates and generates a set of voltages to drive the piezoelectric actuators.

The image of the star on the camera face is the diffraction pattern of an annular telescope aperture. A cross-correlation centroid estimator was implemented. The estimated x coordinate of the centroid is given by

$$\hat{x} = \iint f(x, y)\, h_x(x, y)\, dx\, dy\,, \tag{1}$$

where f is the detected image and h_x is the correlation function for the x direction. The noise in the detected image is photon noise with a variance given by

$$\sigma_f^2 = f(x, y)\,. \tag{2}$$

Since the theoretical shape of the image is known, there is a closed form solution for the optimal correlation function, i.e., that function $h(x, y)$ which gives a minimum variance for $\hat{x}$ due to photon noise in the measurement of $f(x, y)$. The optimal correlation function is given by

$$h_{\text{opt}}(\alpha, \beta) = \frac{[-\partial f(\alpha, \beta)/\partial x]}{f(\alpha, \beta)}\,. \tag{3}$$

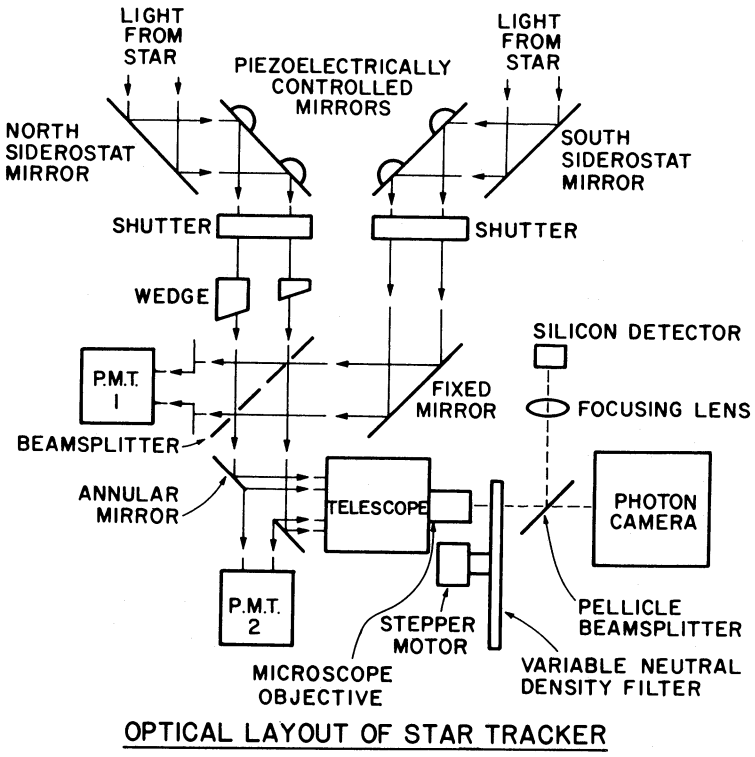

Fig. 3. Optical schematic of star tracker

Table 1

Correlation function	Single photon variance (arcsec)^2
Optimal [Eq. (3)]	0.303
Quad cell [Eq. (4)]	3.09
Linear [Eq. (5)]	6.14

This type of centroid estimator has limited dynamic range but very low variance. An analysis of this estimator for our annular aperture image was made relative to centroid estimators using quad cell and linear correlation functions:

$$\text{Quad cell:} \quad h = \begin{matrix} +1 & x > 0 \\ -1 & x < 0 \end{matrix}, \tag{4}$$

$$\text{linear:} \quad h = x. \tag{5}$$

The results are shown in Table 1. Note that the variance for the optimal correlator estimator is about an order of magnitude better than a quad cell.

The correlation integral [Eq. (1)] is performed by a custom digital signal processor. The processor calculates the integral on a photon by photon basis, rather than a pixel by pixel basis. This is more efficient computationally when the number of detected photons is much less than the number of pixels in the image. Currently, centroid estimates are calculated every 10 ms. The signal processor can operate at photon rates well above the capabilities of the camera. The optical correlation functions, one in x, and one in y for each of the two trackers are downloaded to the correlation processor from the main control computer. The calculation of the optimal correlation function is not done in real time.

The second part of the control system is the temporal filter. While the spatial filter is implemented mostly in digital hardware and microcode, the temporal filter is implemented entirely in software on the star tracker control computer. The temporal filter is a discrete time Kalman filter. The Kalman filter assumes a state variable formulation of the tracking problem and is the linear filter which minimizes the mean square tracking error. This filter takes into account the observation noise (photon noise) by adaptively varying the tracker bandwidth according to the noise variance. The output of the filter is scaled and then directed to a D/A converter and high voltage amplifier to drive the PZTs.

In addition to the software for the Kalman filter, there is software for acquisition, data recording, and for utilities such as moving the shutters and controlling the camera. There is also software for complete automatic operation and for off-line diagnostics and analysis of tracker performance.

2.3. *Optical delay line*

Two optical delay lines are used to equalize the paths from the star to the beamsplitter via the two siderostats 12 m apart. The mechanical and optical arrangement is shown in Fig. 4. The mechanical design is similar to that used in delay lines for high resolution Fourier spectroscopy (Connes, 1975). The optical delay is changed by moving cat's eye retroreflectors along tracks which are housed in two 25-ft vacuum chambers. The peak-to-peak optical path difference of this system is 20 m, and the delay lines can be positioned under laser control to an accuracy of 10 nm, for a dynamic range of 2 10^9 resolution elements in optical path difference. The maximum slew rate is 1 m per second of optical path difference. Twenty seconds is required to go from one end to

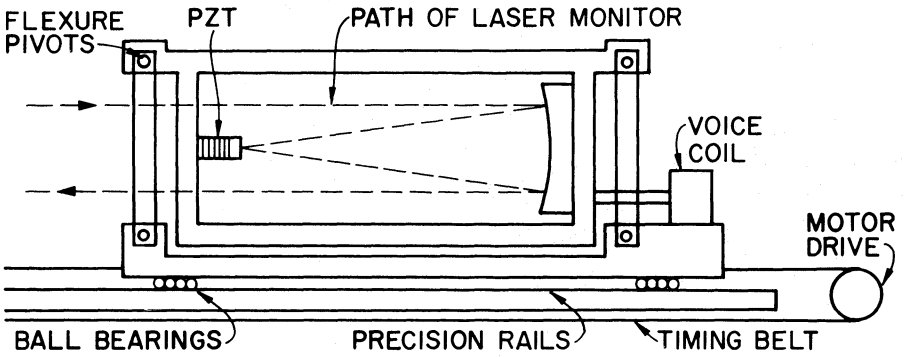

Fig. 4. Three stage (PZT, Voice Coil, Motor) optical delay line

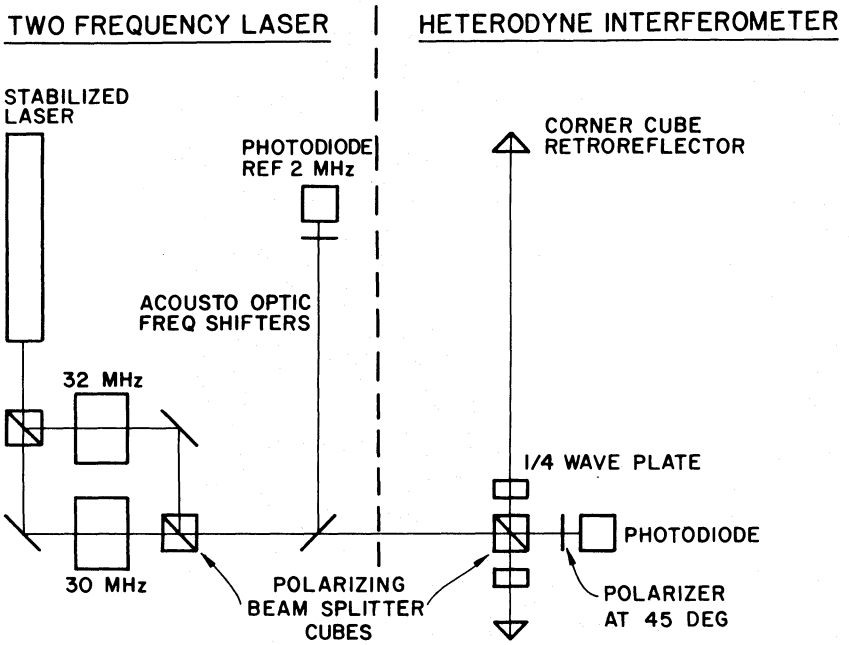

Fig. 5. Heterodyne laser interferometer for delay line

the other. The maximum tracking rate is 2 mm s^{-1}, at which point the rms position accuracy has degraded to 20 nm.

Figure 4 shows the optical path of the laser beam that measures the position of the cat's eye retroreflector. The cat's eye is an 8-inch f/3 parabola with a 1-inch flat mirror at its focus. The 1-inch mirror is mounted on a PZT stack that has a 5-μm peak-to-peak range. The cat's eye is suspended on a flexure bearing over a belt driven translation stage. A voice coil positions the cat's eye over the stage, while a motor at one end of the 25-ft track moves the stage via a timing belt. The voice coil has a 2-mm range and the motor has a 9-μm step size. The three actuators – PZT, voice coil, and motor – have considerable overlap in terms of resolution and dynamic range.

The delay line sensor is a heterodyne laser interferometer (Fig. 5). The light from a stabilized laser is split into two beams by a polarizing beamsplitter. They are then frequency shifted by two acousto-optic bragg cells with a 2-MHz frequency difference. When recombined, the beam has its two orthogonal polarizations at different frequencies with a 2-MHz frequency difference. The laser is a commercial transverse Zeeman stabilized HeNe laser with a one-day stability of 200 kHz at line center, 474 THz. The interferometer is a standard polarization heterodyne interferometer. The phase of the 2-MHz signal from the interferometer is compared to the 2-MHz reference (Fig. 5) and a one wavelength change in the monitored optical path causes a 2π change in the relative phases of the interferometer and reference 2-MHz signals. A custom digital phase comparator circuit counts the fringes with $\lambda/64$ resolution at a maximum slew rate of $2\ 10^6$ fringes per second.

Since the desired optical path difference is given by the difference between the positions of the two delay lines, it is only necessary to actively control one delay line. Thus, vibrations in one delay line are compensated by driving the other in exactly the same way.

The laser interferometers are read at a 1-kHz rate by the fringe tracker computer. The difference between the two interferometers is the measured position, which is subtracted from the desired position to give an error signal. The PZT servo is a simple first-order system, i.e., the PZT drive voltage is just the integral of the error signal. The necessary integral is performed in digital hardware, and the velocity of the PZT mirror is updated every millisecond by the computer. The closed-loop bandwidth of this

servo is approximately 100 Hz, for a time constant of 1 to 2 ms. The small-signal bandwidth of the entire delay line is the bandwidth of this innermost servo.

The voice coil servo is designed to keep the PZT system within its 5-μm dynamic range. The cat's eye flexure suspension has a 1.5-Hz resonance. Since a 5-μm accuracy requires a response significantly faster than 1.5 Hz, the voice coil servo has to be at least a second-order servo. The voice coil servo is implemented entirely in software. The error signal input to the voice coil servo is the distance the PZT mirror is from the center of its range. The output is a current to the voice coil. The closed-loop bandwidth of this servo is approximately 10 Hz.

The motor servo is designed to keep the voice coil in the center of its range. The input to this servo is the position of the voice coil as measured with a lateral-cell transducer. The motor is a microstep-driven stepper motor, interfaced to the computer through a programmable pulse generator. In the tracking mode, it operates as a first-order servo, with a small-signal closed-loop bandwidth of about 1 Hz.

The delay line control system was designed to interact with the rest of the interferometer through a simple interface. The rest of the interferometer does not see three separate actuators and control loops and a laser interferometer. Rather, the delay line subsystem is programmed by setting its position and velocity. At a given time, the delay line is programmed to move to a given position. Thereafter, the optical delay is changed at a rate determined by the programmed velocity. The delay line will move at this rate until the programmed position or velocity is changed.

The composite three-stage servo system has a moderately high bandwidth when the error signal is small (i.e., when delay line position minus target position is less than about 4 μm). However, when switching from one star to the next, the programmed position can suddenly change by 15 m. Since the PZT and voice coil have small dynamic ranges, and in addition have limits on their velocity and acceleration, these servo systems will not behave properly in response to such a large step. Left uncorrected, the resultant oscillations are sufficiently violent to misalign optics. Therefore, in addition to the linear tracking mode of operation described above, the delay line has several other modes of operation. A significant portion of the delay line servo code detects input conditions which would require a change in the servo mode so that the delay line remains well behaved under all conditions.

Table 2 shows the various modes of operation of the three delay-line loops. The modes of the various servos depend on the inputs to the delay line. Rapid changes in the programmed position or velocity is the usual reason for changes in operating mode. Hysteresis in the transition thresholds prevents oscillations caused by changes in the servo parameters between modes. Inputs from the limit switches can also cause changes in the servo modes. For example, two limit switches are used at each end of the track. The first informs the computer that the delay line should start decelerating in order to prevent the delay line from moving beyond the second limit switch, which disables the motors. Other interlocks prevent the laser fringe counter from running at 2 million fringes per second, taking the delay line with it, if the laser becomes misaligned. For a more detailed description of the delay line, the reader is referred to Holm (1986).

Table 2. Inputs: target position, target velocity, current laser positions (2), limit switches (8), voice coil position

Servo	Modes		
PZT	Linear track	Off	
Voice coil	Linear track	Damp	Off
Motor	Linear track	Accel./decel.	Constant vel. Off

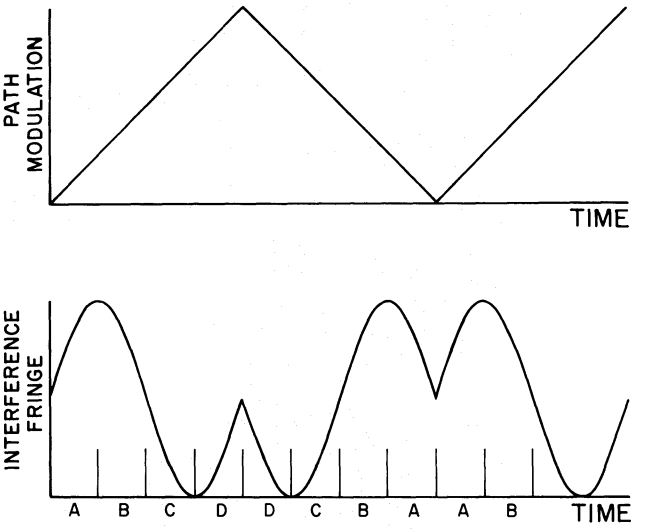

Fig. 6. Stellar fringe modulation/demodulation

2.4. *Stellar fringe detector*

With proper operation of the star tracker and delay line, the two beams combined at the beamsplitter should interfere. This section describes the procedure by which the amplitude and phase of the interference pattern are measured.

In the Mark III interferometer, the interfering wavefronts are parallel. Hence, the two outputs of the beamsplitter are uniform over the aperture and complementary. If the phase difference is 90 degrees, all then light will go to one of the two outputs; if the phase difference is −90 degrees, the light will go to the other output.

The fringe detection algorithm has been described previously (Shao and Staelin, 1977). The basic technique uses path length modulation and synchronous demodulation. The PZT controlled mirror in one of the vacuum delay lines is modulated with a 500-Hz triangle wave. If the path difference in the two arms of the interferometer is within the coherence length of the light (a few wavelengths), the intensity of light at the output of the beamsplitter will be as shown in Fig. 6. During one sweep of the PZT, the detected photons are grouped into four time bins denoted by A, B, C, and D. The amplitude and phase of the fringe are given by

$$\phi = \tan^{-1}\left(\frac{D-B}{C-A}\right) - \frac{\pi}{4}, \tag{6}$$

$$\text{Amp} = \sqrt{(C-A)^2 + (D-B)^2}\,. \tag{7}$$

In the above equations, the $C-A$ and D − B terms are the sine and cosine components of the fringe.

The measured phase is unwrapped so that if the atmosphere moves the fringe position beyond $\pm\pi$, the correct error signal will

be used to track the fringe. The unwrapped phase is used as the error signal in a first-order servo. As mentioned before, the delay line is programmed by position and velocity. The unwrapped phase times a gain term is added to the previous programmed position of the delay line to yield the new programmed position. The programmed velocity is the sidereal rate. The closed-loop bandwidth of this servo is about 20 Hz. A similar fringe tracker was used on the Mark I and Mark II interferometers to track stellar fringes.

The fringe tracking servo, implemented entirely in software, is the simplest part of the stellar fringe detection subsystem. Other software components of the fringe detection subsystem perform functions such as stellar fringe search and initial acquisition, fringe recovery after tracker failure, central fringe identification, and automatic transition between operating modes. In addition to the above software components, other system components that affect the operation of the stellar fringe detector are the PZT stroke servo, and the optics for two color fringe measurements and for accurate visibility measurements.

The instrument was designed to make accurate astrometric measurements and accurate stellar diameter measurements on as many stars as possible. The design of a stellar fringe tracker is complicated by the fact that the requirements of high tracker sensitivity, low systematic astrometric error, and low visibility errors are not entirely consistent.

The stroke of the PZT can be varied by one of the control computers. Since the piezoelectric constant is temperature dependent, the peak-to-peak amplitude of the PZT triangle wave is monitored with a laser interferometer and the drive voltage is adjusted to maintain the proper amplitude to within 0.5 nm. For stellar diameter measurements, the stroke is set at the effective wavelength of the narrow bandpass filter. For astrometry, the stroke is set at the effective wavelength of the red channel.

When the interferometer is set to point at a star, the delay line is programmed to go to the expected position of the fringe plus 60 μm. From past experience, the calculated fringe position is never more than 30 μm in error if the baseline vector used in the calculation is derived from stellar fringe data less than a week old. The delay line is programmed to scan for the fringe at approximately 60 μm per second. The fringe amplitude is calculated every atmospheric coherence time, and if a fringe is detected, the scan is stopped and the servo turned on.

When the tracker loses the fringe, the delay line is moved back 5 μm and a search is restarted. After initial fringe acquisition, the search space is restricted to a 20-μm interval centered about the most recent position where the fringe was found. All of these search parameters are superimposed on the fringe motion due to earth rotation, which is calculated using a baseline vector typically derived from the previous night's data. When the seeing is exceptionally good, i.e., sub-arc-second seeing, the fringe tracker will stay locked on one fringe for many minutes, for up to 10^5 atmospheric coherence times.

If the fringe stays locked for more than one second, a central fringe identification algorithm is enabled. The visibility during that one-second interval is compared with a threshold. If the visibility is below the threshold, it is assumed that the tracker is on the wrong fringe and the tracker will jump one fringe to the right or left. The direction of the tracker jumps is determined by the visibility gradient. Since the tracker has an rms phase error of about 1/6 of a fringe, the visibility gradient can be determined by looking at the distribution of visibility versus phase. Since the central fringe has the highest visibility, the tracker will usually hop in that direction.

The operation of the various systems which comprise the star light fringe detector requires that a number of quantities: phase, amplitude, SNR, visibility gradient, etc. be calculated in real time. Although the system is rather complicated, the resulting tracker is extremely robust. The interferometer can be commanded to point at a star and the tracker will lock onto the fringe within 90 s. Thus, the software complexity was introduced to enable the instrument to operate efficiently under a wide variety of conditions with a minimum of operator intervention.

2.5. *Laser metrology*

During initial operation in 1986, the siderostat laser metrology system was not operational. However, the laser system is important for astrometry and should become operational in the summer of 1987. The basic laser interferometer used to measure changes in instrument geometry was described in the delay line section of this paper. Two such interferometers measure the position of the cat's eye retroreflectors. An additional 12 interferometers will be used to measure the position of the three siderostat mirrors.

The metrology system measures the motion of the pivot point of the siderostat by measuring the distance between four reference corner cubes in an invar plate, which is bonded to the concrete pedestal, and the surface of a zerodur sphere attached to the mirror as shown in Fig. 7. The sphere is made accurate to $\lambda/10$ and the center of the sphere is coincident with the mechanical pivot point to 20 μm. The interferometers have 5 nm resolution; the desired accuracy of the pivot point monitor is 100 nm for one-milliarcsec astrometry with a 20-m baseline. The measurements from these lasers will be used to correct for mechanical errors in the siderostat as well as thermal motion of the siderostat, as described below.

The basic model for the instrument is given by

$$\text{Fringe Pos} = \boldsymbol{S} \cdot \boldsymbol{B} + C, \tag{8}$$

$\boldsymbol{S}$ is the unit vector to the star,
$\boldsymbol{B}$ is the baseline vector,
C is the delay offset.

If the siderostats were perfect, the baseline vector would be the vector that joined the pivot points of the two siderostats. If the delay lines were commanded to move to their zero positions, the delay offset C would be the difference in the optical paths from the beamsplitter to the pivot points of the two siderostats. Motion of the siderostat pivot points affects both $\boldsymbol{B}$ and C. Motion of the internal optics only affects C.

With our 12-m baseline on Mt. Wilson, the rms atmospheric fringe motion was 6 μm under conditions of good seeing. For stellar diameter measurements with the fringe tracker running, neither instrument drifts nor atmospheric motion is important. However, for astrometric measurements, the measured fringe position in Eq. (8) is made relative to the baseline vector and delay offset. Thus, it is necessary that both $\boldsymbol{B}$ and C be constant over a period of time that is sufficiently long for the interferometer to be able to observe the fringe position of a number of stars.

The parameters $\boldsymbol{B}$ and C can change for two reasons: imperfect siderostat bearings and thermal drift. Ideally, the siderostat pivots about a point on the surface of the mirror. In practice, the mirror translates as it rotates in a rather unpredictable way due to irregularities in the bearings and flexure of the mirror cell. The laser metrology system is designed to measure these imperfections. The result, in theory, is that baseline drifts will be limited to thermal drift of the top of the concrete pier to which the reference corner cubes in the laser system are attached. With rapid switching between stars, only baseline drift with a time scale faster than the

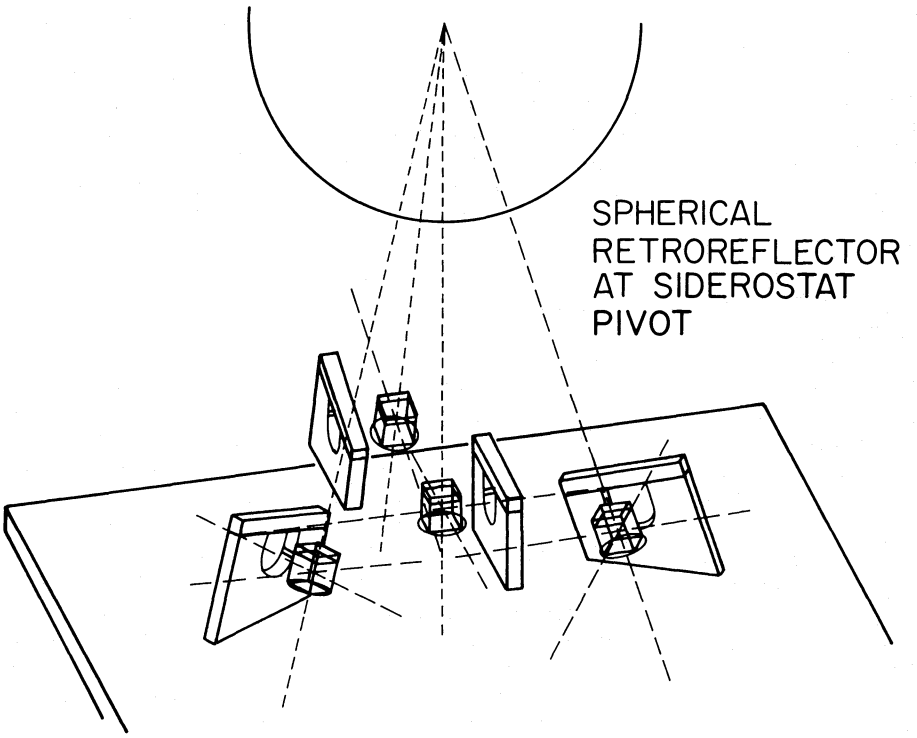

Fig. 7. Laser metrology system for baseline vector monitor

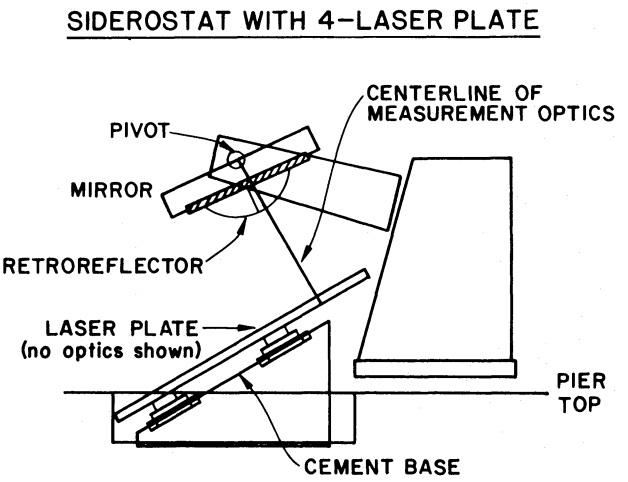

Fig. 8. Arrangement of siderostat and laser plate on Pier

switching cycle results in astrometric errors for large angle relative astrometry.

In order to monitor the position of the pivot point, a minimum of three interferometers is necessary. The fourth interferometer provides the redundancy necessary to solve for the position of the reference corner cubes. The interferometers measure the position of the center of the sphere. Certain parameters, such as the distance between the center of the sphere and the surface of the mirror must be derived from stellar data.

The mechanical accuracy of the instrument without the laser metrology system is approximately 10–15 μm if the baseline was calculated from the previous night's stellar data. The size of this error implies that searching for the fringe would take only a few hundred milliseconds, an insignificant amount of time compared to the 30 to 60 s needed for the siderostats to slew to a new star. The mechanical errors are roughly equivalent to the errors due to atmospheric turbulence, or to the errors in the primary mirror of a high quality large astronomical telescope. With the laser system we hope to improve this accuracy by two orders of magnitude.

While this higher mechanical accuracy is not necessary for stellar diameter measurements, it is crucial for ultra-high accuracy astrometry. This level of accuracy is also necessary for future space-based phase coherent long baseline interferometers.

2.6. Siderostat control

The siderostat control subsystem is in many ways similar to the computer controlled pointing system for a modern astronomical telescope. The siderostat itself is a two axis motorized gymbaled mirror mount (Fig. 8), with the fixed axis tilted 75° from vertical towards the West. The siderostats use conventional precision worm and wheel gears and microstepped stepper motors. The control computer sets programmable pulse generators to control motor velocity. The average step size is 0.″02 to provide smooth tracking, and the maximum slew rate is 1.°8 per second. Typical backlash in the gear train is less than 1″.

The control system performs several tasks. One is open loop or blind pointing, where a geometric model of the siderostat is used to point the siderostat and track the star. A second task is closed loop pointing, where the siderostat is moved to keep the high-speed PZT tilt mirrors in the center of their range. The third task is updating the parameters of the geometric model when a star is

Table 3. Siderostat model parameters

ϕ_F, θ_F	Angular orientation of feed beam (internal alignment)
Az, El (F)	Azimuth and elevation of fixed axis
θ_{axis}	Angle between fixed and movable axis
θ_{mirror}	Angle between movable axis andmirror normal
Z_F, Z_M	Zero point of motor position on the two axes

acquired. In addition, the siderostat control system controls the sequencing of observations during the night.

The geometry of the siderostat described earlier was that of an ideal siderostat. However, the stationary axis may not be exactly 15 degrees from horizontal; it may not be pointed due west, and the two axes may not be exactly perpendicular. Altogether, there are eight parameters in the geometric model of the siderostat. With values of these eight parameters, there is a closed-form relation between the angular positions of the motors and the pointing of the instrument. It is also possible to invert this problem in closed form. By pointing the siderostat at a number (>4) of bright stars whose coordinates are known, and recording the positions of the motors when the stars are centered in the field of view of the interferometer, the eight parameters of the model can be determined in a least squares fashion. In our initial setup, we used the FK4 coordinates of a half dozen stars precessed to the night of observation.

Since the siderostats direct the starlight into the central beam-combining building, the alignment of the interferometer optics inside the building will affect the pointing of the siderostat. The internal alignment of the interferometer is represented by two of the eight parameters describing the siderostat. All of the parameters are listed in Table 3. No attempt was made to model errors in the precision worm and wheel gear or to model flexure. It is most likely that the residual errors are primarily in the gears. Residuals from a least-squares fit for the siderostat parameters are of order 5″. Typical pointing errors for stars which are not used in the least-squares fit are approximately 7″ to 10″.

However, each new star that is observed provides some information on the eight siderostat parameters. Hence, rather than an off-line least-squares solution, a real time algorithm could be implemented to update the eight parameters every time the instrument locked onto a new star. In the Mark III, a Kalman filter is used to update the parameter vector. In operation, the initial values of the parameters at the beginning of the night are the parameters derived from the previous night's observations. These parameters are updated for every new star and, under normal conditions, the initial values would be essentially lost after approximately 6 stars. Thereafter, the parameter estimates would reflect thermal drift of the instrument during the night.

This complex siderostat model is used only for initial acquisitions. After the star tracker locks onto the star, the positions of the PZT tilt mirrors are sent to the siderostat control computer. The siderostat drive rates for the motors are then modified to keep the PZTs in the center of their range. Since the PZT tilt mirrors in effect represent realignment of the internal optics, the corrections to the sideral rate are generated by changing the parameters θ_F, ϕ_F.

The siderostat control computer also calculates the theoretical fringe position and velocity for the optical delay line. These values are sent to the delay line control computer through a shared multiport memory every 2 s. The theoretical fringe position is calculated from a baseline vector derived from the previous night's data. in the future, we plan to update the baseline vector with a Kalman filter in the same way the siderostat parameters are updated. On the first night of interferometric observations, the baseline vector was determined by conventional surveying techniques. The surveying error was approximately 1 mm, or about 20″ of baseline tilt.

The last function of the siderostat control computer is the sequencing of observations. An abbreviated version of the FK4 catalog is kept on-line (In 1987, the FK5 catalog will be kept on-line). A menu driven program enables the operator to list and plot the stars visible at the present time and to easily define the stellar observation sequence. The siderostat control computer also communicates with the other two control computers to enable fully automatic operation once the observing sequence is programmed. The siderostat control computer signals the star tracker computer to initiate a star acquisition sequence. When the star tracker has locked onto a star it signals the fringe tracker computer to initiate a fringe acquisition sequence. When the allotted observing time for a star is over, the siderostat control computer signals the other two computers to disengage their servo systems. For additional information, the reader is referred to Kelly (1986).

3. Operational characteristics

The instrument was built and debugged over a 33 month period. Fringes were observed on the first night of operation, September 12, 1986. Over the following months, observational techniques were refined and automated. Figure 9 is a plot of the observations made on 1986, November 11. The vertical axis is the delay line position where fringes were observed. Each dot represents a 100-s observation. A total of 165 observations of 23 stars were made over a 10.5-h period. Stars over a declination range of 68° and an hour angle range of 8 h were observed. All these observations were of stars 3.0 magnitude or brighter.

The initial measurements in 1986 will be discussed in a series of separate papers. Wide-angle one-color astrometry shows night-to-night repeatability of about 0″.04 (Mozurkewich et al., 1987), and the formal error for seven nights of data is 0″.02 in declination. Significant improvements are anticipated when two-color analysis is applied and when the laser siderostat monitor is installed in the summer of 1987. Preliminary measurements on the atmospheric limitations of two color absolute astrometric measurements indicate that single observations of ~ 100 s should yield accuracies of about 0″.01, a factor of 5 improvement over one-color observations (Colavity et al., 1987). Two-color observations should approach 0″.003 for a single night's observation. Last of all, we have made preliminary stellar diameter measurements of βGem and αCMi. Assuming a uniform disk, we measure 6.15 ± 0.25 mas for αCMi and 8.70 ± 0.15 mas for βGem. The uniform disk diameter for αCMi is not in agreement with prior measurements (Hanbury Brown et al., 1974) in part due to the difference in observation wavelength and uniform disk assumption. A much more careful discussion of accurate stellar diameter measurements is given elsewhere (Shao et al., 1988).

The instrument was never operated near the limiting magnitude of the fringe tracker because the temporary star tracker we were using could not track stars fainter than 3 mag. The tracker using the photon camera should be capable of operation on faint

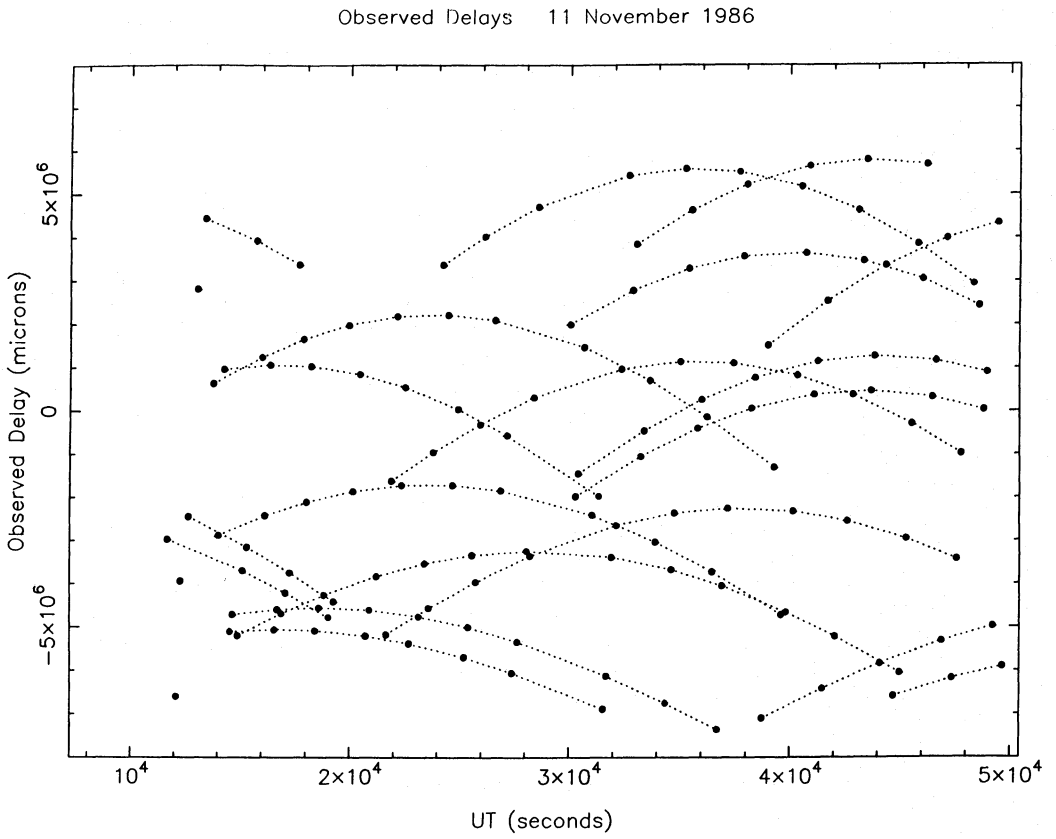

Fig. 9. Position of delay line where fringes were observed, 165 observations of 23 stars

sources, to 12 mag under good seeing conditions when only mechanical errors need to be tracked. When the new star tracker is installed in the spring of 1987, the limiting magnitude will be set by the fringe detection system. We plan to operate the interferometer in two distinct modes that have two very different magnitude limits. The two modes represent two types of interferometers. One is a closed loop phased interferometer. The second is an open loop coherent interferometer or absolute interferometer. We plan to test this second mode of operation in the next two years.

3.1. *Phase-coherent interferometry*

The Mark III was designed to be an active phased interferometer. In this mode of operation, the fringe tracker compensates for all optical path fluctuations, including atmospheric turbulence to $\sim \lambda/6$ rms. Fringe phase measurements in two colors are used for astrometric purposes, while light from a bandpass filter, < 400 Å, is used for stellar diameter measurements. It is also possible to measure the diameter of a star at emission or absorption lines. For diameter measurements in narrow spectral lines, we make use of the phase coherent property of an active interferometer. For narrow spectral lines, the photon flux in the line might be 10^{-3} of the total flux and the number of photons per atmospheric coherence time may be as low as 10^{-2}. However, since the active optics system is phase coherent, the coherent integration time is not limited to the atmospheric coherence time. By choosing a longer coherent integration time, > 1 s, a fringe visibility measurement can be made in 10^4 s with a signal-to-noise ratio of 100.

For astrometry, the fringe phase and the position of the optical delay line is recorded when the fringe tracker is locked on the central fringe. In this case, the astrometric error due to photon noise is negligible and the major errors are instrumental. A detailed discussion of one- and two-color astrometry will be presented in separate papers (Colavita et al., 1987; Mozurkewich et al., 1988). However, the astrometric capability of the instrument affects all other modes of operation of the instrument.

With the known coordinates of FK 5 stars it is possible to derive the baseline vector and delay offset by measuring the fringe position of four stars. This derived baseline is as accurate as the coordinates of the FK 5 stars used. An accurate baseline vector in turn enables the instrument to find the fringe faster and to increase the number of observations from several per night to several hundred per night. This astrometric capability is the key to operation of the instrument as an absolute interferometer, described next.

3.2. *Absolute interferometer*

Within the next year we plan to install the software for real time baseline calibration. In this mode, the interferometer will rapidly switch between 4–6 FK 5 stars to calibrate the baseline, and then point blind at a faint astronomical object for about 10^3 s. This cycle will be repeated to yield the desired integration time on the faint object. Three types of measurements are planned, two of them in the near future. There are significant advantages to phase coherent operation and operation in the absolute mode is reserved for objects too faint for the fringe tracker, 9 mag for 10 cm apertures. All three types of measurements planned for the absolute interferometer are photon starved measurements. Under these conditions, direct phase measurements are not possible and the

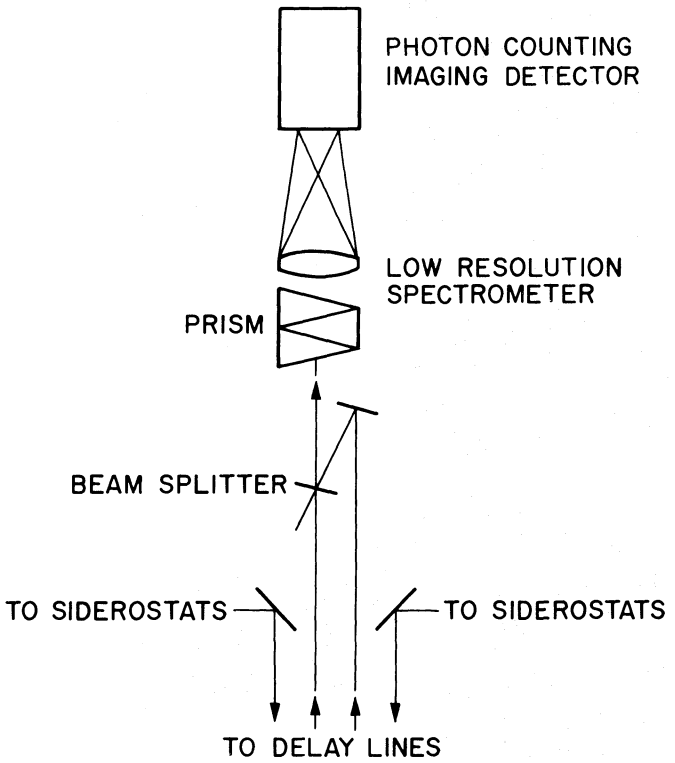

Fig. 10. Dispersed fringe detector schematic

only quantities that can be measured are amplitude, group delay, and closure phase. Photon starved conditions occur when the average number of photons per r_0 aperture, per τ_0 integration time, is less than 1 in the coherent optical bandpass.

In open loop operation, a 12-m interferometer will exhibit rms phase fluctuations of approximately 20 λ at 0.6 μm for 1″ seeing. The fringe detector must be able to accomodate this lack of temporal coherence. On the other hand, to maximize sensitivity, light from all wavelengths should be used for interferometric measurements. The basic fringe detector (Fig. 10) consists of a prism spectrometer with an imaging photon counting detector. The required resolution of the spectrometer is of the order of

$$\Delta\lambda \simeq \lambda \left(\frac{r_0}{L}\right)^{5/6}, \tag{9}$$

where λ is the wavelength of observation, L the length of the baseline, and $r_0 \propto \lambda^{6/5}$ is equal to about 10 cm at 0.6 μm for 1″ seeing. Conservatively, 50 Å channels are assumed in the calculations below. In the following sections we briefly describe the signal processing of the output of the detector for the three different types of photon starved measurements.

3.2.1. Photon-starved amplitude measurements

The estimator for fringe amplitude must be modified to take into account a bias that is important at low photon rates. An unbiased estimator for visibility squared is given by

$$V^2 = \frac{\pi^2}{2}\left[\frac{\langle (A-C)^2 + (B-D)^2 - N\rangle}{\langle N - \mathrm{Dark}\rangle^2}\right] \tag{10}$$

where $N = A + B + C + D$ = total photon flux, Dark = dark count and $\langle\ \rangle$ denotes time average.

The signal-to-noise ratio of the above visibility estimator has been extensively studied both for amplitude interferometry and speckle interferometry. There are two regions of interest, depending on photon rate. The dependence of integration time needed to achieve a given signal-to-noise ratio, given a photon flux N per r_0 aperture, per τ_0 integration time, per spectral channel is given by

$$\text{Integration time} \propto N^{-1},\ N > 1,\ N > \mathrm{Dark},$$
$$\propto N^{-2},\ 1 > N > \mathrm{Dark}. \tag{11}$$

For stellar diameter measurements, the numerator and denominator of the expression in Eq. (10) must be evaluated every 10 ms for each spectral channel, typically 60 channels of 50 Å bandwidth for a 12-m baseline. If collecting apertures greater than r_0 (10 cm) are used, there must be a spectrometer for each r_0 subaperture. In photon starved operation there is no hard magnitude limit as there is in active phase coherent operation. As a practical limit we require a signal-to-noise ratio of 30 in 10^4 s of integration. Assuming an unresolved star, the required photon flux is 1.6 10^{-2} photons in each of 60 spectral channels, per r_0 aperture per 10 ms. This corresponds to a magnitude limit of 14.3 for 10-cm apertures. With 1-m apertures, the limiting magnitude is about 16.8 (see Appendix). Accurate visibility measurements will require much higher signal-to-noise ratios and the magnitude limit will degrade. The optical efficiency of the spectrometer was assumed to be 100% and the quantum efficiency of the imaging photon counter was assumed to be 10%. For these reasons, these estimates should be considered optimistic.

3.2.2. Photon-starved group delay astrometry

The signal from the spectrometer can be processed in a different way to yield astrometric information at faint magnitudes. When the optical path difference in the two arms of the interferometer is different by many wavelengths, say 3 μm, there will be constructive interference at $\lambda = 0.5$ μm and 0.6 μm, and destructive interference in between. More precisely, the spectrometer output will exhibit a fringe pattern due to the interference of light from the two siderostats as given by

$$I(\nu) = I_0(\nu)\,(1 + \sin 2\pi\nu\tau), \tag{12}$$

where $I_0(\nu)$ is the spectrum of the star, ν is optical frequency $(1/\lambda)$ and τ is the optical path difference in the two arms of the interferometer, i.e., the group delay. Equation (12) has a sine rather than a cosine because the Mark III uses a dielectric beamsplitter to combine the beams. The Fourier transform of the "fringed" spectrum $I(\nu)$ will have a delta function at τ. At this delay, the Fourier transform has both an amplitude and phase. In the photon starved mode, where

$$\int I(\nu)\,d\nu < 1\ \text{photon}/r_0^2/10\,\text{ms}, \tag{13}$$

the phase is not measurable.

The basic signal processing algorithm is to average the power spectrum of $I(\nu)$. As in the previous section, each r_0 subaperture must be separately processed if the collecting apertures are large. Since τ changes due to turbulence, a time average of the power spectrum is a gaussian distribution of delta functions. For a 12-m baseline, the standard deviation of this gaussian is ~12 μm for 1″ seeing. A SNR of 30 in this case means that the peak in the gaussian is approximately 30 times the rms fluctuations in the rest of the power spectrum. For 10-cm apertures, the magnitude limit is approximately 14.3 mag. For 1-m apertures, the corresponding limit is 16.8 (see Appendix).

3.2.3. Photon-starved closure phase

Closure phase measurements were first made with radio interferometers (Jennison, 1958), and subsequently suggested for optical interferometers (Rogstad, 1968). Recently, closure phase measurements were made at visible wavelengths by masking the aperture of a large telescope (Baldwin, 1986; Readhead, 1988) at photon rich levels. In closure phase measurements, three interferometers, whose baselines form a closed triangle, measure three complex visibilities corrupted by turbulence. However, the closure phase, the sum of the phases of the three complex visibilities, is independent of the atmosphere. In photon starved measurements, raw phases cannot be measured. Instead, the product of the complex visibilities is averaged, i.e.,

$$\text{triple product} = \langle A_1 e^{i\varphi_1} A_2 e^{i\varphi_2} A_3 e^{i\varphi_3} \rangle, \quad (14)$$

where $A_1 e^{i\varphi_1}$ is the complex visibility from interferometer 1. The importance of averaging the complex visibilities comes from the fact that the product of complex visibilities is not biased in the photon starved region. The closure phase is simply the imaginary part of the log of the triple product (see Appendix). The triple product was first used in high resolution optical imaging for reconstruction of speckle images (Weigelt, 1977), a technique called speckle masking. This technique extends the sensitivity of closure phase measurements more than an order of magnitude beyond conventional photon rich measurements.

For long baseline interferometers, the triple product must be calculated every 10 ms, for each spectral channel, and each r_0 aperture. The signal-to-noise ratio (SNR) of the triple product is qualitatively different from the SNR for amplitude measurements. At high photon fluxes, $N \gg 1$, the SNR of all these quantities varies as $\sqrt{N}$. At low photon fluxes, $N < 1$, the SNR for amplitude measurements varies as N, and that of the triple product varies as $N^{3/2}$ (see Appendix). Again, assuming an SNR of 30 in 10^4 s ($\Delta\phi$ measured to $\lambda/180$), the magnitude limit with 10-cm apertures is approximately 12.5 mag, or a photon flux of 0.08 photons per r_0, per 10 ms, per 50 Å bandwidth. The limiting magnitude for 1-m apertures is 14.2.

Although three siderostats will be in operation in 1987, closure phase measurements are not planned for the immediate future. In order to operate three interferometers simultaneously the second delay line must be converted to high precision operation. Operation as an absolute interferometer will then enable closure phase measurements to be made with a one to two order-of-magnitude increase in sensitivity over active phase coherent operation.

4. Conclusion

In the past seventy years a number of stellar interferometers have been proposed and built. Starting with Michelson, the instruments have gotten progressively more complex and sophisticated (Koechlin et al., 1985; Labeyrie et al., 1986; Davis et al., 1986). The Mark III incorporates many of the ideas proven by these previous instruments. The major differences between this interferometer and others reside primarily in its operational characteristics – it is fully automated, it is easy to operate, and it is mechanically accurate and thermally stable. It is also capable of making literally hundreds of observations in one night; both astrometric and stellar diameter measurements are possible. Last of all, this instrument has the potential to operate as an absolute interferometer which, from an astronomical point of view, holds tremendous promise for observation of faint objects.

To fully exploit the possibilities of absolute interferometry, and to construct images of faint extragalactic objects at milliarcsec resolution, even more complex interferometers with large, multiple apertures will be needed. During the next several years the Mark III should demonstrate much of the necessary technology, as well as produce some of the scientific results that would motivate the construction of a large optical array.

Acknowledgements. The authors wish to acknowledge the important contributions of P. Cheimets, J. Pohlman, and C. Papa in the design and construction of the precise mechanical components of the system as well as the staff of Mt. Wilson Observatory for their cooperation and aid during the construction of the instrument. This project was made possible with support from the following grants and contracts: N 00014-84-C-2082, N 00014-84-C-2137, N 00014-85-C-2212, N 00014-86-C-2105, N 00014-86-C-2114 from the Naval Research Laboratory; from the Smithsonian Institution directly and through its Scholarly Studies Program; AST 79-19553 from the National Science Foundation; NAS-2-50 and NSG-7176 from NASA; ONR N 00014-80-C-0348 from the Office of Naval Research.

Appendix: photon-starved interferometry

The photon-starved signal-to-noise ratio for the three interferometric quantities: amplitude, group delay, and closure phase are discussed in this Appendix. We present limiting magnitudes for these quantities extrapolated from our experience with the Mark III interferometer. Recently, there has been considerable interest in the use of light scattered from the ionosphere from a pulsed laser in order to provide a probe of atmospheric turbulence to operate an active-optics system (Foy and Labeyrie, 1986). The application of this technique is not considered in our calculations.

Estimates of amplitude, group delay, and closure phase each have slightly different magnitude limits. Qualitatively, different effects degrade the sensitivity for the different quantities. In all cases, we assume that the fringe detector is a dispersive spectrometer followed by an imaging photon-counting detector. We also assume that the resolution of the spectrometer is a few times greater than what is needed to maintain high fringe contrast in the presence of the optical path fluctuations caused by atmospheric turbulence. Magnitude limits are calculated assuming an unresolved star, a 12-m baseline, 10-cm collecting apertures, a 10 ms coherent integration time, a 60-channel, 50 Å resolution spectrometer, unit fringe visibility, a total integration time of 10^4 s, and a final signal-to-noise ratio of 30. The detector dark count is neglected, and a total system efficiency of 0.1 is assumed. We ignore the reduction in fringe visibility and hence signal-to-noise ratio due to the use of coherent apertures of 10 cm and integration times of 10 ms. In practice, the seeing would have to be somewhat better than 1″ for these magnitude limits to be achieved. Scaling formulae are given for the case of different seeing conditions.

At high photon fluxes, when the number of photons per r_0 aperture, per atmospheric coherence time τ_0, per spectral channel $\Delta\lambda$, is much greater than one, the signal-to-noise ratio is proportional to the square root of the photon flux for estimates of all three interferometric quantities. We define the quantities N and M as

N = No. of detected photons per frame, i.e., per r_0, τ_0, $\Delta\lambda$,

$$M = \text{No. of frames} = \left(\frac{D}{r_0}\right)^2 \left(\frac{T}{\tau_0}\right) \left(\frac{BW}{\Delta\lambda}\right),$$

where D = diameter of telescope, T = total integration time, and BW = total optical bandwidth of the spectrometer (assumed 0.4 to 0.7 μm). The terms photon-rich and photon-starved apply to the cases $N > 1$ and $N < 1$, respectively.

The signal-to-noise ratio for fringe amplitude measurements has been extensively studied for many types of interferometers. It is usually more convenient to consider estimates of V^2 rather than of V. The signal-to-noise ratio for estimates of V^2 can be written (Tango and Twiss, 1980)

$$SNR(V^2) = \frac{1}{4} M^{1/2} NV^2 \left[1 + \frac{1}{2} NV^2\right]^{-1/2}. \quad \text{(A1)}$$

For the case of a fringe-scanning interferometer with 4 bins per scan, N is replaced by $(8/\pi^2)N$. The qualitative feature of this expression is that at low photon fluxes, $N < 1$, the signal-to-noise ratio is proportional to N, rather than to $N^{1/2}$. Thus, if a star is dimmer by a factor of 10, the number of frames M must increase by a factor of 100 to give the same signal-to-noise ratio.

The noise behavior of the group-delay estimate is similar to the V^2 estimate. The "fringed" spectrum as detected by the spectrometer is

$$I(\nu) = I_0(\nu)(1 + \sin 2\pi\nu\tau'), \quad \text{(A2)}$$

where $I_0(\nu)$ is the spectrum of the star, ν is optical frequency, and τ' is the group delay. The group delay changes due to atmospheric turbulence. Over long time intervals, τ' is a gaussian-distributed random variable with mean τ_1, determined by the position of the star and optical delay line, and variance σ_τ^2, which depends on the seeing and baseline length. The variance is given by

$$\sigma_\tau^2 = 6.88\,(\lambda/2\pi)^2\,(L/r_0)^{5/3}, \quad \text{(A3)}$$

where L is the baseline length and r_0 is the atmospheric coherence diameter.

Each coherence time τ_0, the power spectrum, $P(\tau)$, of $I(\nu)$ is calculated. The power spectrum is then averaged for the length of the observation. There are three important features: a peak at the origin attributable to the Fourier transform of $I_0(\nu)$, a uniform bias due to photon noise, and a gaussian distribution at the mean group delay τ_1 due to atmospheric turbulence. The averaged power spectrum is given by

$$\langle P(\tau)\rangle = \langle |\int e^{2\pi i\nu\tau} I_0(\nu)\,(1 + \sin 2\pi\nu\tau')\,d\nu|^2\rangle, \quad \text{(A4)}$$

where $\langle\ \rangle$ represents an ensemble average over both group delay τ' and photon statistics.

A better estimator is the unbiased, normalized power spectrum $Q(\tau)$, given by

$$Q(\tau) = \frac{\langle P(\tau) - \int I(\nu)\,d\nu\rangle}{\langle \int I(\nu)\,d\nu\rangle^2}, \quad \text{(A5)}$$

where $N_s = \int I(\nu)\,d\nu$ is the total photon flux in a 10 ms interval. Comparison of this expression with the estimator (11) shows that (A5) is just the V^2 estimator evaluated at a particular group delay. Thus, the performance of the group delay estimator can be calculated using (A1).

Given a spectrometer with a bandwidth of 0.4 μm to 0.7 μm, the delay resolution is $\Delta x = (1/0.4 - 1/0.7)^{-1} \sim 1$ μm. Since the group delay fluctuates with a standard deviation of σ_τ, the gaussian distribution of group delay in $Q(\tau)$ will be spread over approximately $\sigma_\tau/\Delta x$ bins, while the amplitude of the gaussian will be reduced by approximately $\sigma_\tau/\Delta x$. Photon-starved statistics apply for $(1/2)\,N_s V^2 < \sigma_\tau/\Delta x$, and using (A1), the signal-to-noise ratio for an estimate of the amplitude of the gaussian using a matched filter is given by

$$SNR(Q(\tau)) = \frac{N_s V^2 M^{1/2}}{7.5}\sqrt{\frac{\Delta x}{\sigma_\tau}}, \quad \text{(A6)}$$

where N_s is the total photon flux in the spectrum. In (A6) sampling losses are neglected, and it is assumed that $\sigma_\tau \gg \Delta x$. Using (9) and (A3) to write (A6) in terms of N, rather than N_s, yields a signal-to-noise ratio for group delay estimates which is nearly identical to the signal-to-noise ratio of a conventional amplitude estimate.

For the case of group-delay, an estimate of the centroid of the group-delay distribution is desired. The standard deviation of such an estimate can be written

$$\sigma_p = \frac{\sqrt{2}\sigma_\tau}{SNR}, \quad \text{(A7)}$$

where the SNR is given by (A6). A Monte Carlo simulation was performed which verified the above analytically derived expressions.

Under photon-starved conditions, the measurement of closure phase is qualitatively different from the measurement of amplitude. The definition of the triple product, Eq. (14), is repeated below:

$$\text{triple product} = \langle A_1 e^{i\varphi_1} A_2 e^{i\varphi_2} A_3 e^{i\varphi_3}\rangle. \quad \text{(A8)}$$

In photon-counting measurements, if zero photons are detected in any one of the three interferometers, the triple product is zero. For a frame to contain some information on the closure phase, at least one photon must be detected in all three interferometers. The Poisson probability function is given by

$$P(m) = \frac{N^m e^{-N}}{m!}, \quad \text{(A9)}$$

where $P(m)$ is the probability of detecting m photons given an average arrival rate of N. For $N \ll 1$ the probability of detecting one photon is N. Thus, the probability of detecting one photon in all three interferometers is N^3, so that the variance for estimates of closure phase should be proportional to N^{-3} under photon starved conditions.

More precisely, the variance of the closure phase calculated via the triple product can be written

$$\sigma_\phi^2 = 3\left(\frac{2}{NV^2}\right) + 6\left(\frac{2}{NV^2}\right)^2 + 4\left(\frac{2}{NV^2}\right)^3 \quad \text{(A10)}$$

for a signal-to-noise ratio $SNR = 1/\sigma_\phi$. In this expression N is the average number of detected photons in each interferometer. Equal visibilities V have been assumed for each of the three interferometers. Thus, at high photon fluxes, the variance is three times the variance in a photon-rich phase measurement. At low photon fluxes, the variance is proportional to N^{-3}, as given by the last term of (A10). For this term, V is the geometric mean of the visibilities in the three interferometers. A Monte Carlo simulation for closure phase measurements was also conducted, which verified the above expression.

The integration time needed to achieve a signal-to-noise ratio of 30 versus photon flux N is given in Fig. 11 for measurements of visibility squared and group delay, and closure phase. For astronomical measurements, a photon flux corresponding to an SNR of 30 on an unresolved star allows a diameter measurement of 4% on a partially-resolved source. An SNR of 30 for group delay corresponds to an astrometric measurement of 0.6 μm in

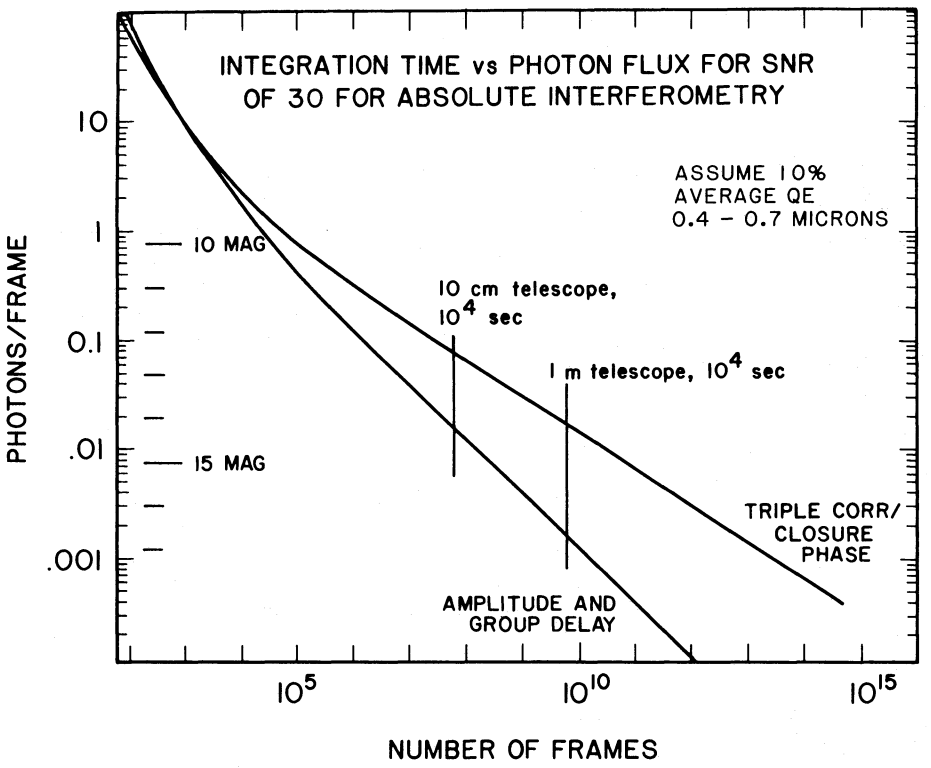

Fig. 11. Integration time vs. magnitude for photon-starved measurements

delay, or 0.″01 in angle. An *SNR* of 30 for the triple product corresponds to an error of $\lambda/180$ in closure phase. The limiting magnitudes for a single-r_0 system are approximately 14.3 and 12.5 for visibility/group delay and closure phase. With 1-m aterpures, the limiting magnitudes are approximately 16.8 and 14.2, respectively.

If r_0 and τ_0 are different from the nominal values assumed, the magnitude limits will change. In general, an instrument will have a fixed collecting aperture. Thus, if r_0 increases, the number of photons, N, per r_0 increases, but the number of frames, M, decreases. Similarly, given a fixed 10^4 s for a measurement, an increase in τ_0 will also increase N and decrease M. If the signal-to-noise ratio were dependent on N and M evenly, the final *SNR* would be independent of seeing. This is the case for photon-rich measurements where the *SNR* is proportional to $\sqrt{NM}$.

For photon-starved measurements, this is not the case. The signal-to-noise ratio for amplitude and delay measurements follow the scaling relation $NM^{1/2}$, while the *SNR* for closure phase follows the relation $N^{3/2}M^{1/2}$. Two additional facts should be considered. As r_0 improves, the rms path length fluctuation between the apertures decreases proportionally. Hence, the number of required spectral channels that must be separately analyzed decreases according to (9). In addition, since the temporal fluctuations of phase are usually modeled by assuming frozen turbulence translating past the instrument at a constant wind speed, τ_0 should increase with r_0. Thus, the scaling of N and M with r_0 is approximately given by

$$N \propto r_0^4, \tag{A11}$$

$$M \propto r_0^{-4}. \tag{A12}$$

The scaling of the magnitude limits for photon-starved measurements with seeing is then given by

$$\text{Amplitude/delay sensittivity} \propto r_0^2, \tag{A13}$$

$$\text{Phase closure sensitivity} \propto r_0^{8/3}. \tag{A14}$$

These results should be compared to the case of a wideband phase tracking system. In this case, only N, not M is important, and since the fringe is tracked, there is no need for the spectrometer. Thus, the sensitivity scales with seeing as r_0^3.

References

Baldwin, J.E., Haniff, C.A., Mackay, C.D., Werner, P.J.: 1986, *Nature* **320**, 595

Connes, P.: 1975, *Appl. Opt.* **14**, 2067

Clark, L.D., Shao, M., Colavita, M.M.: 1986, *Proc. SPIE.* **627**, 838

Clark, L.D.: 1986, M.S. Thesis, Dept. of E.E.C.S., M.I.T., Cambridge, MA

Colavita, M.M., Shao, M., Staelin, D.H.: 1987, *Appl. Opt.* **26**, 4106 and 4113

Davis, J., Tango, W.J.: 1986, *Nature* **323**, 234

Foy, R., Labeyrie, A.: 1986, *Astron. Astrophys.* **152**, L29

Hanbury Brown, R., Davis, J., Allen, L.R.: 1974, *Monthly Notices Roy. Astron. Soc.* **167**, 121

Holm, W.H.: 1986, M.S. Thesis, Dept. of E.E.C.S., M.I.T., Cambridge, MA

Jennison, R.C.: 1958, *Monthly Notices Roy. Astron. Soc.* **118**, 276

Kelly, M.J.: 1986, M.S. Thesis, Dept. of E.E.C.S., M.I.T., Cambridge, MA
Koechlin, L., Rabbia, Y.: 1985, *Astron. Astrophys.* **153**, 91
Labeyrie, A., Schumacher, G., Dugué, M., Thom, C., Bourlon, P., Foy, F., Bonneau, D., Foy, R.: 1986, *Astron. Astrophys.* **162**, 359
Mozurkewich, D., Hutter, D.J., Johnston, K.J., Simon, R.S., Shao, M., Colavita, M.M., Staelin, D.H., Hines, B.E., Hershey, J.L., Hughes, J.A., Kaplan, G.H.: 1988, *Astrophys. J.* (in press)
Papaliolios, C., Nisenson, P., Ebstein, S.: 1985, *Appl. Ppt.* **24**, 287
Readhead, A.C.S., Nakajima, T.S., Pearson, T.J., Neugebauer, G., Oke, J.B., Sargent, W.L.W.: 1988, *Astron. J.* (in press)
Rogstad, D.H.: 1968, *Appl. Opt.* **7**, 585
Shao, M., Staelin, D.H.: 1977, *J. Opt. Soc. Am.* **67**, 81
Shao, M., Staelin, D.H.: 1980, *Appl. Opt.* **19**, 1519
Shao, M., Colavita, M.M., Staelin, D.H., Johnston, K.J., Simon, R.S., Hughes, J.A., Hershey, J.L.: 1987, *Astron. J.* **93**, 1280
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, *Astrophys. J.* **327**, 358
Tango, W.J., Twiss, R.Q.: 1980, in *Progr. Optics*, **XVII**, ed. E. Wolf, North-Holland, Amsterdam, p. 239
Weigelt, G.P.: 1977, *Opt. Comm.* **21**, 55