# A NEW ALGORITHM FOR THE ABSOLUTE METROLOGY OF OPTICAL SURFACES

**JOHN BENJAMIN MITCHELL**

A thesis submitted in partial fulfilment of the requirements of
the Kingston University for the degree of
Doctor of Philosophy

This research programme was carried out in collaboration with
the National Physical Laboratory, Teddington, Middlesex.

September 1996.

**ABSTRACT**

In this thesis, the problem of the metrology of optical surfaces is examined. The need for accurate metrology is established with regard to optical wavefronts reflecting or refracting at the surfaces of optical components. Optical interferometry is identified as the most useful analytical tool for surface metrology by virtue of its high precision and accuracy. Accordingly the theory of interferometry is briefly presented. The application of the theory to the interferometric instruments is shown together with measuring configurations for the various optical surfaces commonly encountered. An extensive overview of the techniques used to evaluate data from interferometric measurements is given with particular emphasis on precision phase measuring methods.

Most interferometric measurements of surfaces are made relative to a reference surface of high quality. Where the accuracy of the surface to be measured is comparable to, or better than, that of the reference surface, an absolute measurement technique is required in order to give meaningful results. A review is given of the existing methods for the absolute measurement of nominally flat and spherical surfaces and the shortcomings of these methods.

A new algorithm for the absolute testing of flat surfaces is developed, based on relative measurements of pairs from a population of three test flats in a number of positional combinations. The new method has a number of potential advantages over those previously described, particularly since it yields information about the flats over their entire surfaces on a square grid of points. The implementation of the new method on a Zygo Mark IV interferometer is described together with experimental results using both synthesized and actual experimental data. Suggestions for improvements to the method and its implementation are made.

A speculative study of other possible techniques for absolute flatness measurement is presented, including the possible application of the Ritchey-Common test, point diffraction interferometry, phase conjugation and profilometry.

A full and up to date survey of the pertinent literature is given throughout the thesis.

THE FOLLOWING HAS BEEN REDACTED AT THE REQUEST OF THE UNIVERSITY:

CARTOON ON PAGE FOLLOWING ABSTRACT

"Mister Logic" *Viz Comic* 72 June/July 1995

"We must, therefore, recognise that what is designated by the terms "glance", "hand", and, in general, "body" is a system of systems devoted to the inspection of a world and capable of leaping over distances, piercing the perceptual future, and outlining hollows and reliefs, distances and deviations- a meaning in the inconceivable flatness of being".

Maurice Merleau-Ponty, "Indirect language and the voices of silence".

"THE VERY BIG STUPID is a thing which breeds by eating The Future. Have you seen it? It sometimes disguises itself as a good-looking quarterly bottom line, derived by closing the R&D Department".

Frank Zappa, "The Real Frank Zappa Book".

"Heavens!" said Mrs Lambchop.
"Gosh!" said Arthur. "Stanley's flat!"
"As a pancake," said Mr Lambchop. "Darndest thing I've ever seen."
"Let's all have breakfast," Mrs Lambchop said. "Then Stanley and I will go and see Doctor Dan and hear what he has to say."

Jeff Brown, "Flat Stanley".

CONTENTS

# Chapter 1 : Introduction.

## 1.1 The Importance Of Measuring Optical Surfaces.

In any optical system there will be optical components that are designed to have a specific effect on the light travelling through the system. The desired effect may be a change in the direction, shape, polarisation or amplitude of the wave, or a combination of two or more of these.

A conventional optical component consists of some bulk material with certain optical and physical properties bounded by surfaces which define the shape of the component. The optical properties include the refractive index and absorption of the material which may vary with the wavelength of the light and also the direction in the material. Physical properties such as hardness, chemical resistance etc. are important as regards the manufacture and use of the component but do not directly affect its optical performance. The function of the component depends both on the shape (also known as the figure) of its defining surfaces and also, if the light travels through the component, upon the optical properties of the material from which it is made.

If the component is purely reflective (a mirror, for example) its function depends only upon the figure of the surface since the light does not travel through the substrate material. Geometrically, the light obeys the well known rule that the angle of reflection equals the angle of incidence. The direction of an individual ray after reflection is thus dependent upon the orientation of the reflective surface at the point where the ray was incident, as shown in figure 1.1.

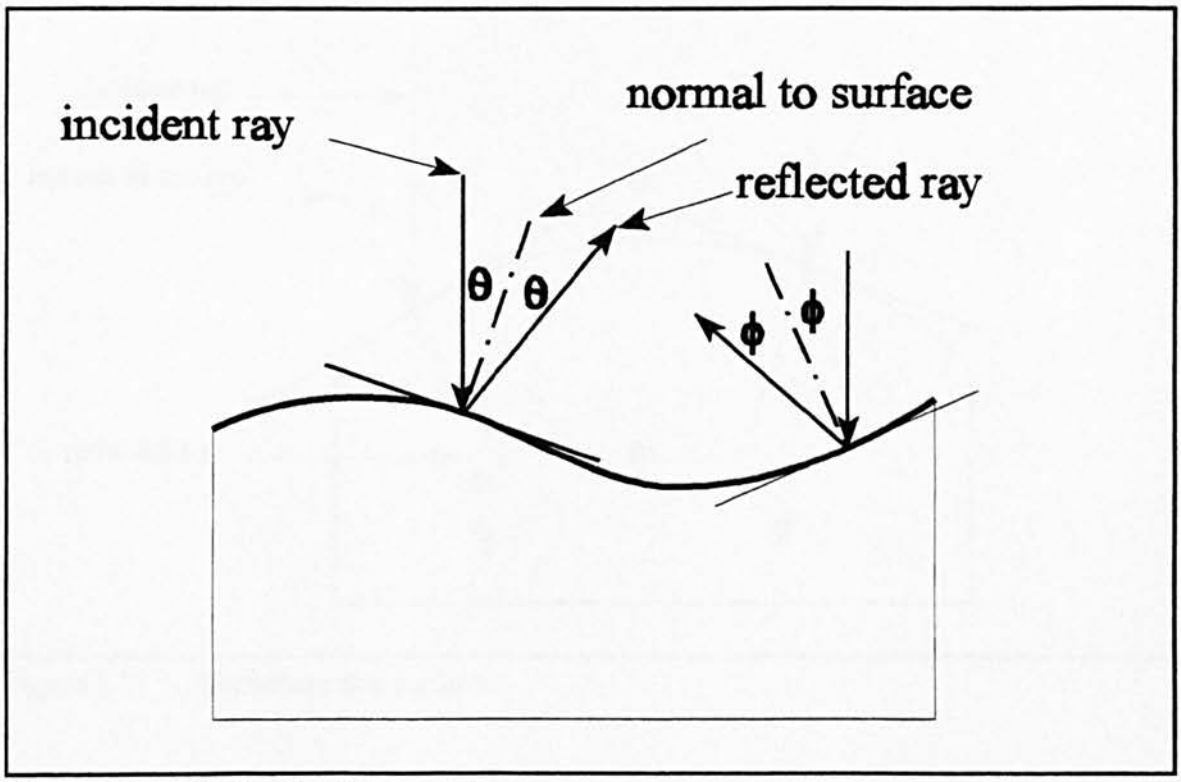


Figure 1.1 Reflection from a surface.

When a component is transmissive, the direction of the transmitted ray depends upon the angle of incidence and also upon the refractive indices of the media on either side of the surface of the component. The relationship between the directions of the incident and transmitted rays is given by Snell's law;

$$n_1 \sin\theta_1 = n_2 \sin\theta_2,$$

where: $n_1$ ,$n_2$ are the refractive indices of the media either side of the surface, and

$\theta_1$, $\theta_2$ are the angles of the incident and refracted rays respectively.

This relationship is illustrated in figure 1.2.

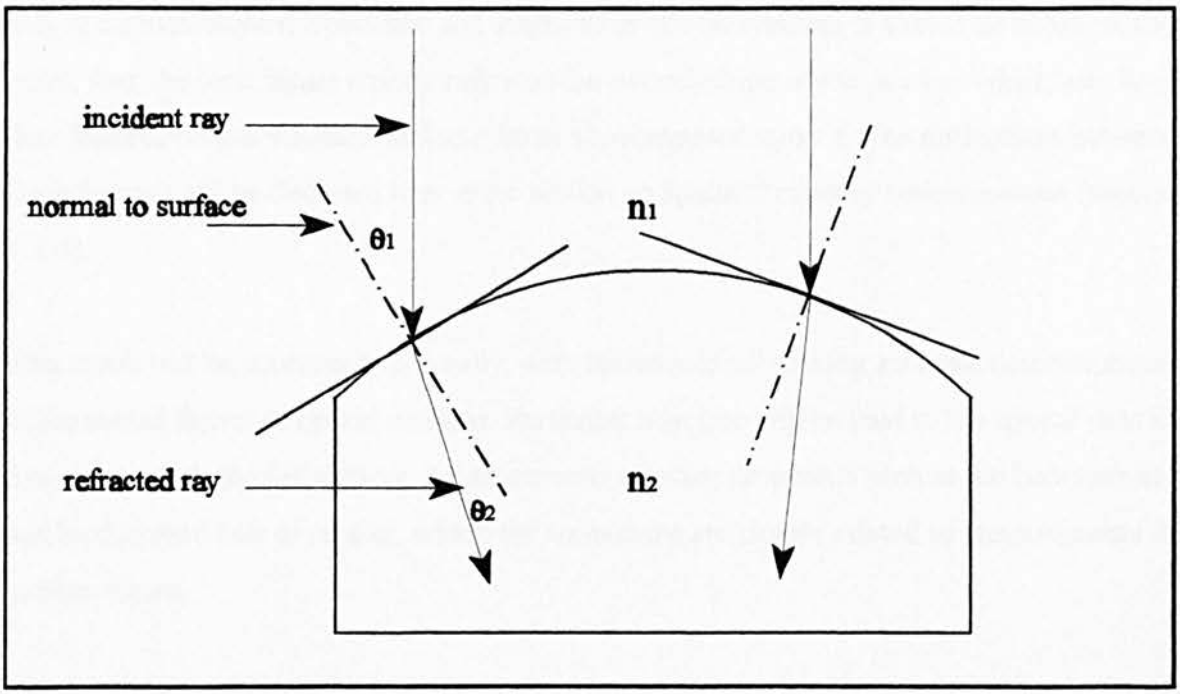

Figure 1.2 Refraction at a surface.

If the refractive indices are known and can be assumed to be constant (homogeneous) throughout the media then the only factors affecting the function of the components are the figures of the surfaces.

To illustrate the importance of the surface figure of optical components, consider an imaging system. For the performance of the system to be diffraction limited (that is, limited by diffraction at the aperture of the system) then the optical components should introduce wavefront aberrations no greater than $\lambda/4$ at the wavelength of the light being imaged. If the imaging system consists of a single optical surface then the maximum allowed deviation from the ideal surface figure is $\lambda/8$ for a reflective component and $\lambda/4(n-1)$ for a refractive component of index n in air. It can thus be seen that the allowed error on the surface of a glass lens ($n \approx 1.5$) is approximately four times higher than that of a mirror. For a system consisting of multiple optical surfaces, the manufacturing tolerances must, of course, be more stringent.

Given that the figure of optical surfaces has such an important effect upon the performance of the optical components of which they are part, their accurate measurement has a vital part to

play in the manufacture, evaluation and diagnosis of optical systems. It should be noted, at this point, that the term figure strictly refers to the overall shape of the surface which may have finer features such as waviness and roughness superimposed upon it. The distinctions between these features will be discussed later in the section on spatial frequency considerations (section 1.2.4).

This thesis will be concerned, primarily, with techniques for making accurate determinations of the overall figures of optical surfaces. Particular attention will be paid to the special case of measuring nominally flat surfaces. Measurements of other properties such as the homogeneity will be discussed only in passing, where the techniques are closely related to measurements of surface figure.

## 1.2 Interferometry.

One of the most sensitive tools for measuring the figures of surfaces is interferometry. The phenomenon of interference of light occurs wherever two or more electromagnetic waves are superposed. If the medium in which the waves meet is linear (which is usually the case unless the wave intensities are very high) the electric field vectors of the waves add algebraically. Consider two waves, $E_1$ and $E_2$ of the same frequency:

$$E_1 = E_{01}\sin(\omega t + \alpha_1)$$

and

$$E_2 = E_{02}\sin(\omega t + \alpha_2)$$

where $\omega$= the angular frequency of the wave,

$\alpha$= -(kx+$\varepsilon$), the phase of the wave at x,

$\varepsilon$= the epochal angle, the phase of the wave at x= 0,

k= the wave number, =$2\pi/\lambda$,

$\lambda$= the wavelength,

t= time,

x= distance from the source.

Where the two waves overlap in space, the resultant is the linear superposition of these waves;

$$E = E_1 + E_2$$

or

$$\begin{aligned} E &= E_{01}(\sin\omega t\cos\alpha_1 + \cos\omega t\sin\alpha_1) \\ &\quad + E_{02}(\sin\omega t\cos\alpha_2 + \cos\omega t\sin\alpha_2) \\ &= (E_{01}\cos\alpha_1 + E_{02}\cos\alpha_2)\sin\omega t \\ &\quad + (E_{01}\sin\alpha_1 + E_{02}\sin\alpha_2)\cos\omega t \end{aligned}$$

Writing;

$$E_0\cos\alpha = E_{01}\cos\alpha_1 + E_{02}\cos\alpha_2$$

and

$$E_0\sin\alpha = E_{01}\sin\alpha_1 + E_{02}\sin\alpha_2$$

where

$$E_0^2 = E_{01}^2 + E_{02}^2 + 2E_{01}E_{02}\cos(\alpha_2 - \alpha_1)$$

and

$$\tan\alpha = \frac{E_{01}\sin\alpha_1 + E_{02}\sin\alpha_2}{E_{01}\cos\alpha_1 + E_{02}\cos\alpha_2}.$$

The resultant wave may now be written;

$$\begin{aligned} E &= E_0(\cos\alpha\sin\omega t + \sin\alpha\cos\omega t) \\ &= E_0\sin(\omega t + \alpha). \end{aligned}$$

The resultant wave is harmonic with the same frequency as the constituent waves but with a different amplitude and phase.

Since the amplitude of a light wave is not directly observable we must consider the resultant intensity of the two waves. The observable intensity of a wave is proportional to the square of its amplitude. The resultant intensity of the interference of two waves is thus proportional to $E_0^2$.

The resultant is the sum of the two intensities of the two constituent waves plus an interference term, $2E_{01}E_{02}\cos(\alpha_2-\alpha_1)$. The important quantity here is the phase difference between the two interfering waves, $\alpha_2-\alpha_1 = \phi$. When $\phi = 0, \pm 2\pi, \pm 4\pi, \ldots$ the resultant is a maximum and

interference is constructive and when $\phi = \pm\pi, \pm 3\pi, \ldots$ the resultant is a minimum and interference is destructive. Note that when the intensities of the two interfering waves are equal, the resultant is zero for destructive interference and 4I for constructive interference.

The interference equation may be written conveniently as;

$$I = I_0(1 + \gamma \cos\phi),$$

where $I_0 = I_1 + I_2$, the DC intensity level,

$\gamma = 2E_{01}E_{02}/I_0$, the fringe visibility.

The significance of these quantities is illustrated in figure 1.3.

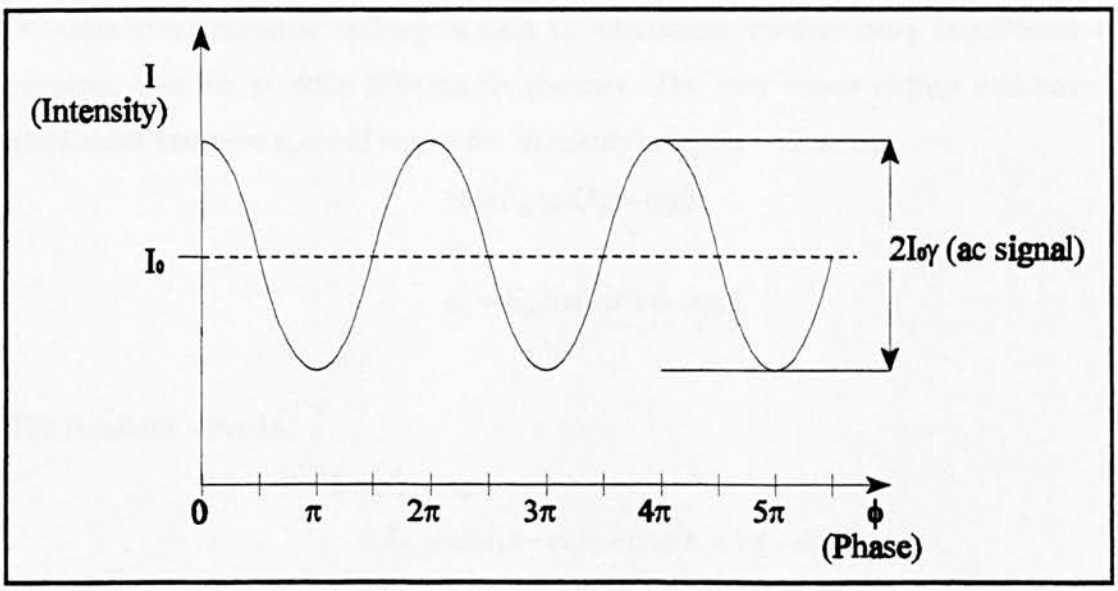


Figure 1.3 Intensity distribution for two-beam interference.

The phase difference, $\phi$, may arise from a difference in the path travelled by the two interfering waves as well as a difference in their epochal angles;

$$\begin{aligned}\phi &= (kx_1 - \varepsilon_1) - (kx_2 - \varepsilon_2) \\ &= \frac{2\pi}{\lambda}(x_1 - x_2) + (\varepsilon_1 - \varepsilon_2)\end{aligned}$$

In interferometry, the two interfering waves generally arise from the same source, so their epochal angles are the same. In this case;

$$\phi = \frac{2\pi}{\lambda}(x_1 - x_2).$$

The quantity $x_1$-$x_2$ is known as the optical path difference (OPD) and it can be seen that constructive interference will occur when the OPD is a whole number of wavelengths, (OPD= 0, $\lambda$, $2\lambda$,...) and destructive interference when the OPD is an odd number of half wavelengths (OPD= $\lambda/2$, $3\lambda/2$, $5\lambda/2$,...).

If the two waves travel through a number of different media with different refractive indices, n, the OPD can be written;

$$OPD = \sum_i x_{1_i} n_i - \sum_j x_{2_j} n_j$$

In some interferometric techniques such as heterodyne interferometry interference occurs between two waves with different frequencies. The two waves (which will have equal amplitudes and zero epochal angles for simplicity) may be written as;

$$E_1 = E_{01} \cos(k_1 x - \omega_1 t)$$

and

$$E_2 = E_{01} \cos(k_2 x + \phi - \omega_2 t).$$

The resultant wave is;

$$\begin{aligned} E &= E_1 + E_2 \\ &= E_{01}\left[\cos(k_1 x - \omega_1 t) + \cos(k_2 x + \phi - \omega_2 t)\right] \\ &= 2E_{01} \cos\tfrac{1}{2}\left[(k_1 + k_2)x + \phi - (\omega_1 + \omega_2)t\right] \\ &\quad \times \cos\tfrac{1}{2}\left[(k_1 - k_2)x + \phi - (\omega_1 - \omega_2)t\right] \end{aligned}$$

This equation may be rewritten as;

$$E = 2E_0 \cos(k_m x + \phi - \omega_m t) \cos(\bar{k}x + \phi - \bar{\omega}t),$$

where;

$$\begin{aligned} \bar{\omega} &= \tfrac{1}{2}(\omega_1 + \omega_2), \text{ the average angular frequency,} \\ \bar{k} &= \tfrac{1}{2}(k_1 + k_2), \text{ the average wave number,} \\ \omega_m &= \tfrac{1}{2}(\omega_1 - \omega_2), \text{ the modulation frequency,} \\ k_m &= \tfrac{1}{2}(k_1 - k_2), \text{ the modulation wave number.} \end{aligned}$$

The resultant wave arising from the interference of two waves with different frequencies is illustrated in figure 1.4.

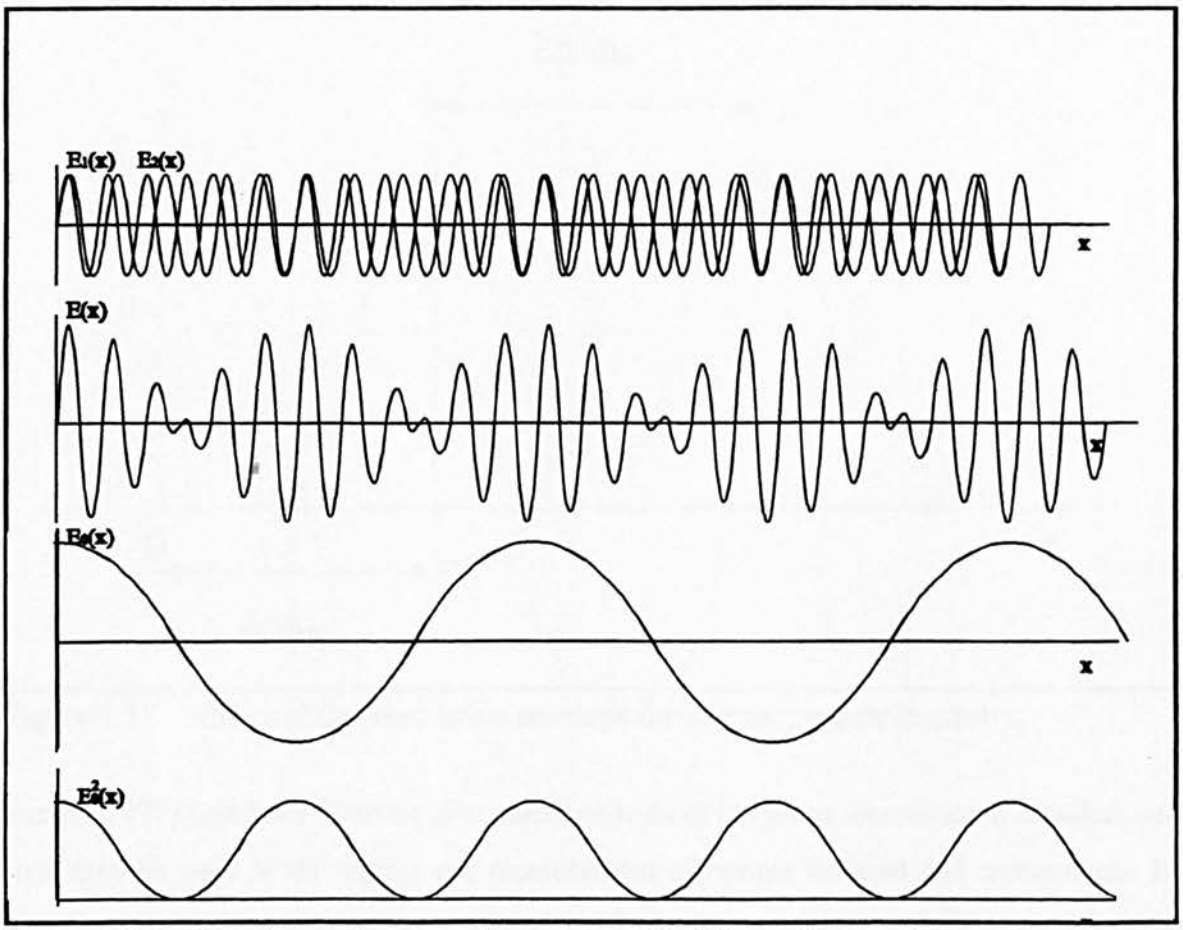


Figure 1.4 Interference between two waves with different frequencies.

When the frequencies of the two waves are similar, the resultant can be regarded as a travelling wave with the average frequency, amplitude modulated by half the difference frequency. Again the observed quantity is the wave intensity, proportional to the square of the amplitude.

$$
\begin{aligned}
E_0^2 &= 4E_{01}^2 \cos^2(k_m x + \tfrac{\phi}{2} - \omega_m t) \\
&= 2E_{01}^2\left[1 + \cos 2(k_m x + \tfrac{\phi}{2} - \omega_m t)\right] \\
&= 2E_{01}^2\left[1 + \cos\{2(k_m x - \omega_m t) + \phi\}\right]
\end{aligned}
$$

So, at some point, where two waves with different frequencies interfere, the resultant intensity is modulated at a frequency $2\omega_m = (\omega_1 - \omega_2)$ known as the beat frequency. The phase difference, $\phi$, between the two interfering waves is replicated in the phase of the modulation envelope and

is thus easily measured as shown in figure 1.5. It is this feature that makes heterodyne interferometry a useful tool for the accurate measurement of distances.

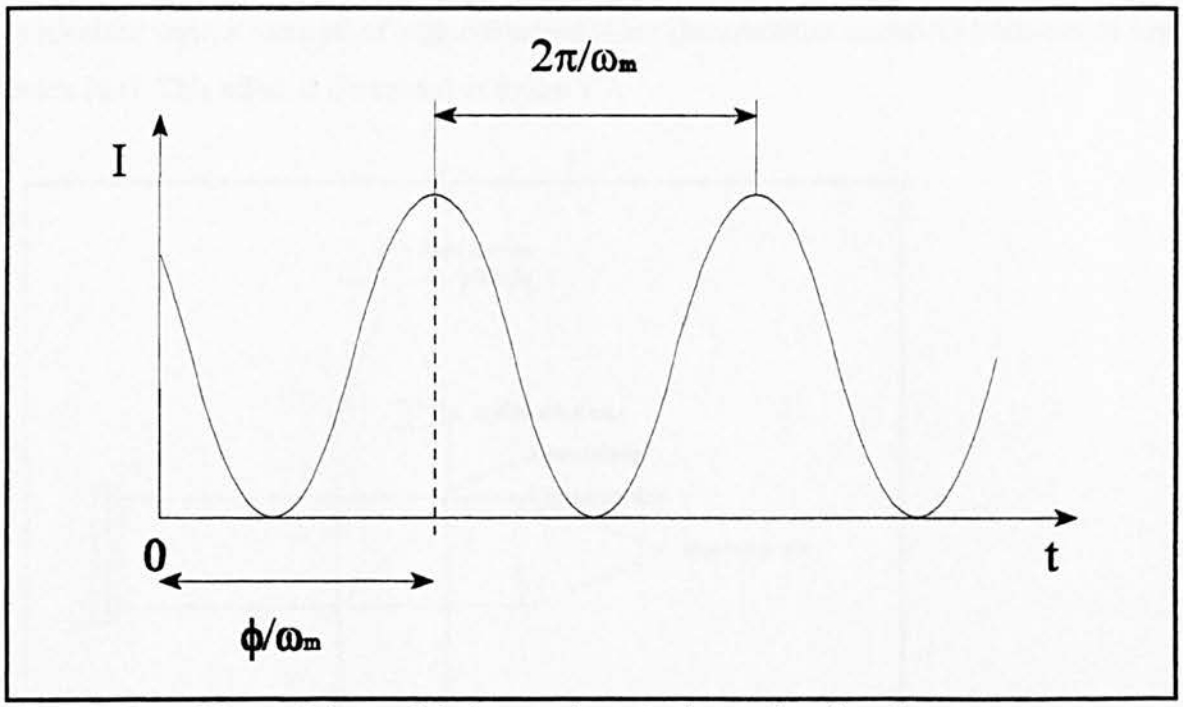


Figure 1.5 Phase of the modulation envelope for heterodyne interferometry.

Briers (1972) and Stahl (1990b) give useful reviews of the many interferometric techniques that may be used in the testing and measurement of optical surfaces and components. In addition, Malacara, Cornejo and Murty, (1975) have compiled a very comprehensive bibliography of optical testing methods, including interferometry, up to the date of publication. For the measurement of surface figure the Twyman-Green and Fizeau interferometer configurations are of the greatest interest.

### 1.2.1 Twyman-Green Interferometer.

The basic Twyman-Green interferometer is shown in figure 1.6. Light from a monochromatic point source is collimated to form a plane wavefront and divided in amplitude by a beamsplitter. The two wavefronts are reflected by plane mirrors M1 and M2 and then recombined by the beamsplitter. Two interference patterns are formed; one directed back to the light source and one towards the observer. The two interference patterns are complementary in order to obey the principle of conservation of energy. Where there is a

bright fringe in one pattern there is a dark one in the other. This occurs because there is a 180° ($\pi$ radians) phase change in the wavefronts upon reflection at the beamsplitter when the light is incident onto a material of high refractive index (beamsplitter material) from one of low index (air). This effect is illustrated in figure 1.7.

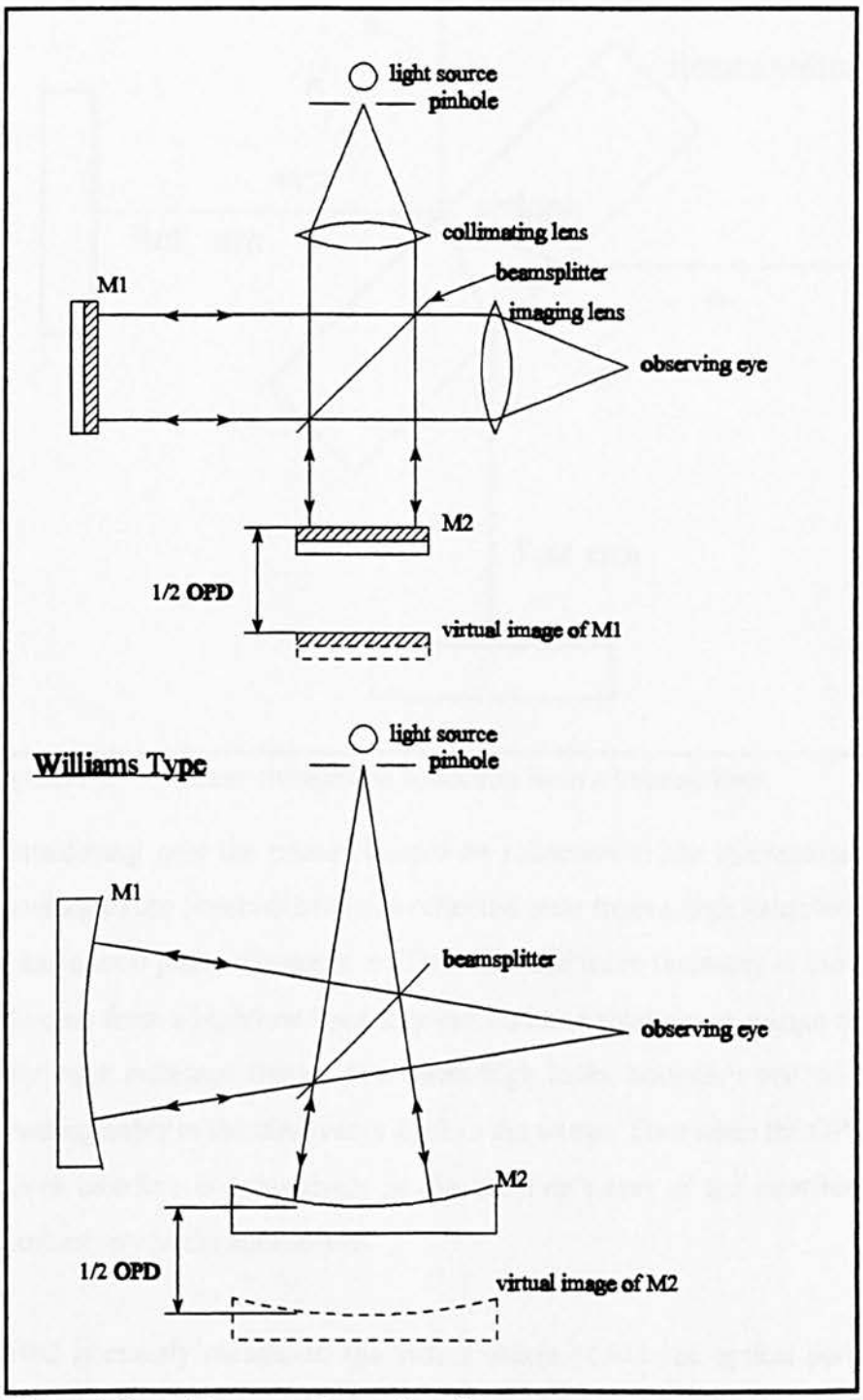


Figure 1.6 Twyman-Green Interferometers.

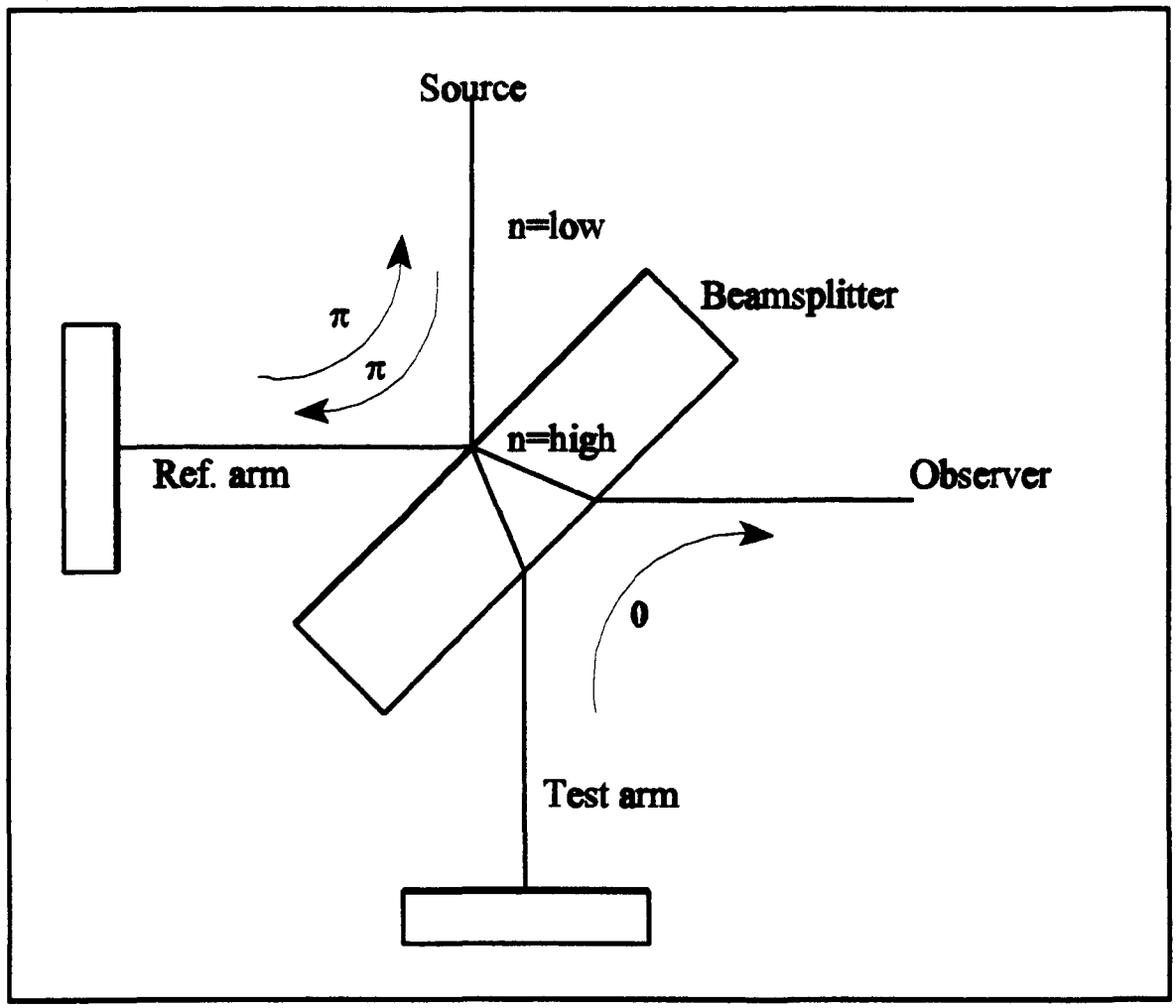


Figure 1.7 Phase changes on reflection from a beamsplitter.

Considering only the phase changes on reflection at the beamsplitter, the reference wave travelling to the observer has been reflected once from a high index/low index boundary and so has a total phase change of $\pi$. The reference wave returning to the source has twice been reflected from a high/low boundary and so has a total phase change of $2\pi$. The test wave is only ever reflected from a low index/high index boundary and so has no phase change travelling either to the observer or back to the source. Thus when the OPD is such that the two waves interfere constructively in the observer's arm of the interferometer, they interfere destructively in the source arm.

If M2 is exactly parallel to the virtual image of M1 the optical path difference (OPD) is constant and there is a single fringe across the field of view. If M1 and M2 are non-parallel the

OPD will vary and there will be a number of fringes across the field of view. Each fringe represents a contour of equal separation between the two mirrors. Since the OPD is twice the separation between the mirrors a change in separation of half a wavelength ($\lambda/2$) corresponds to a change in OPD of a whole wavelength ($\lambda$). Thus the change in separation represented by one interference fringe is $\lambda/2$. Where both mirrors are plane, but with a small wedge angle between them, the fringe pattern will consist of a number of straight and parallel fringes of constant separation depending on the wedge angle. If one mirror (M1, say) is plane and the other distorted, the fringe pattern will represent a contour map of the surface of the distorted mirror. We may then measure the contours or figure of M2 with respect to M1, the reference mirror.

In the general case that neither M1 or M2 are plane, the interference pattern corresponds to the difference in the figure of the two mirrors. This can easily be seen with reference to figure 1.8. If M1 and M2 have the same non-plane figure then their separation is constant and a single fringe will result (or a pattern of equi-spaced straight fringes if there is a wedge angle between them). If the figures of the two surfaces are denoted by $f_{M1}(x,y)$ and $f_{M2}(x,y)$, where x and y are the Cartesian co-ordinates of the surfaces, then the fringe pattern yields the information; $g(x,y)=f_{M1}(x,y)-f_{M2}(x,y)$.

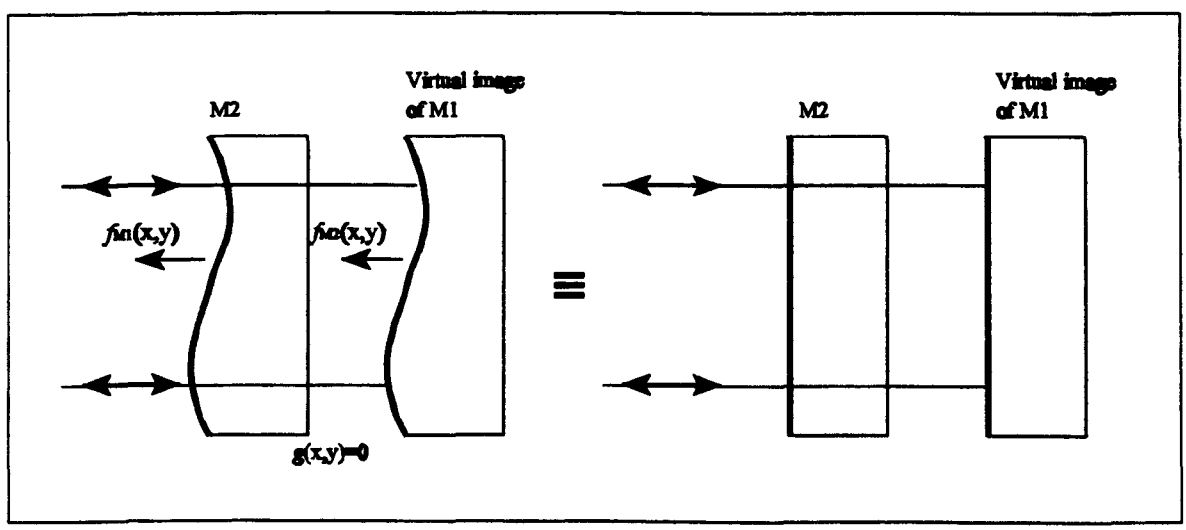


Figure 1.8 OPD in a Twyman-Green Interferometer.

A variation of the Twyman-Green interferometer, sometimes known as a Williams

interferometer, for measuring concave spherical surfaces is also shown in figure 1.6. The two spherical mirrors are placed so that their centres of curvature are at the point source of light. In this case the fringe pattern corresponds to the difference between the figures of the two surfaces with respect to their centres of curvature. Here, the surface figures are denoted by $f_{M1}(\phi,\theta)$ and $f_{M2}(\phi,\theta)$ where $\phi$ and $\theta$ are the spherical co-ordinates of the surfaces and the fringe pattern yields $g(\phi,\theta) = f_{M1}(\phi,\theta)-f_{M2}(\phi,\theta)$.

The use of the Twyman-Green interferometer for performing tests on a wide variety of optical components and surfaces is described by Birch (1979) and Malacara (1992).

### 1.2.2 Fizeau Interferometer.

The basic Fizeau interferometer configuration is shown in figure 1.9.

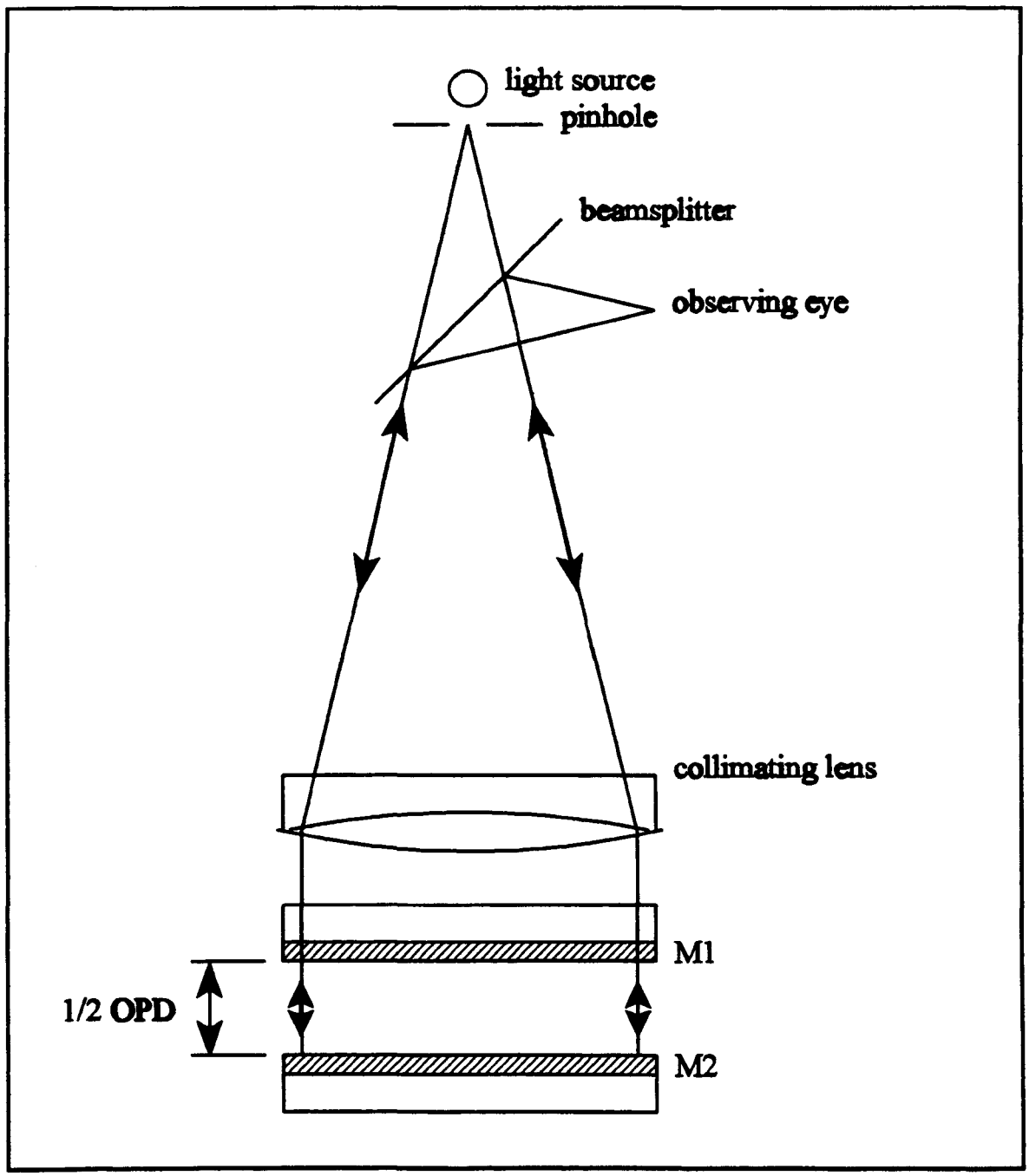


Figure 1.9 Fizeau interferometer.

Light from a monochromatic point source is collimated by the collimating lens. A portion of the light is reflected from the surfaces of M1 and M2 ( this will typically be about 4% for uncoated glass surfaces). The reflected light is redirected by the beamsplitter towards the

observer. In common with the Twyman-Green interferometer, the fringe pattern observed is dependent upon the OPD between the light reflected from M1 and from M2 which is twice the separation between the two surfaces. Once again, in order that energy be conserved, there is a complementary interference pattern to that directed towards the observer. In this case the complementary pattern is transmitted through M2 though it is difficult to observe when the reflectivities of the surfaces are low since it is swamped by the light directly transmitted by both surfaces. As in the case of the Twyman-Green interferometer, if one surface is assumed to be perfectly plane (usually the upper surface, M1), the interference pattern gives the contours of the other surface, M2, and a fringe interval denotes a height interval of $\lambda/2$. For the more general case where neither surface is plane, however, the interferogram yields the sum of the figures of the two surfaces rather than the difference as is the case for the Twyman-green interferometer. When the figures of M1 and M2 are again given by $f_{M1}(x,y)$ and $f_{M2}(x,y)$ respectively then the fringe pattern yields the information $g(x,y)=f_{M1}(x,y)+f_{M2}(-x,y)$. This may easily be seen by reference to figure 1.10.

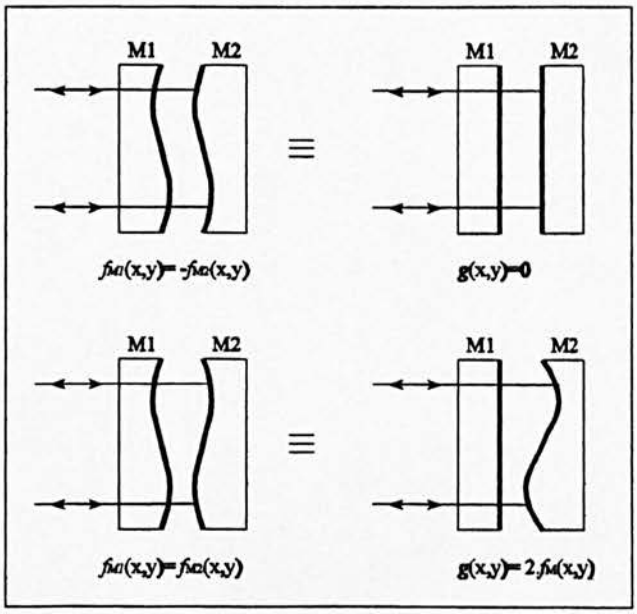


Figure 1.10 OPD in Fizeau interferometer.

The Fizeau interferometer is a very versatile optical measuring instrument which may be configured to test a wide variety of optical components and systems. Harris (1971) devoted his Ph.D. thesis to the design and construction of an "universal" Fizeau interferometer system. A comprehensive discussion of the Fizeau interferometer and its variants is given by Mantravadi (1992a) and an analysis of potential sources of measuring errors is provided by

Huang (1993).

### 1.2.3 Measurement Configurations.

In the descriptions of the Twyman-Green and Fizeau interferometers given above the wavefronts have been assumed to be nominally plane for the measurement of nominally plane surface figures. Either type of interferometer may also be used to measure a number of different surface figures such as spheres and conic sections (conicoids) and general aspheres. These measurements require that the interferometer produce a spherical test wavefront. This may be achieved in a Twyman-Green interferometer by placing a converging or diverging lens in the test arm (a converging lens is most commonly used since it produces both a convergent wavefront before the focus and a diverging wavefront past the focus). In a Fizeau interferometer, the lens may either be placed after a plane reference surface or before a spherical reference surface. The configurations for producing a spherical test wavefront are shown in figure 1.11.

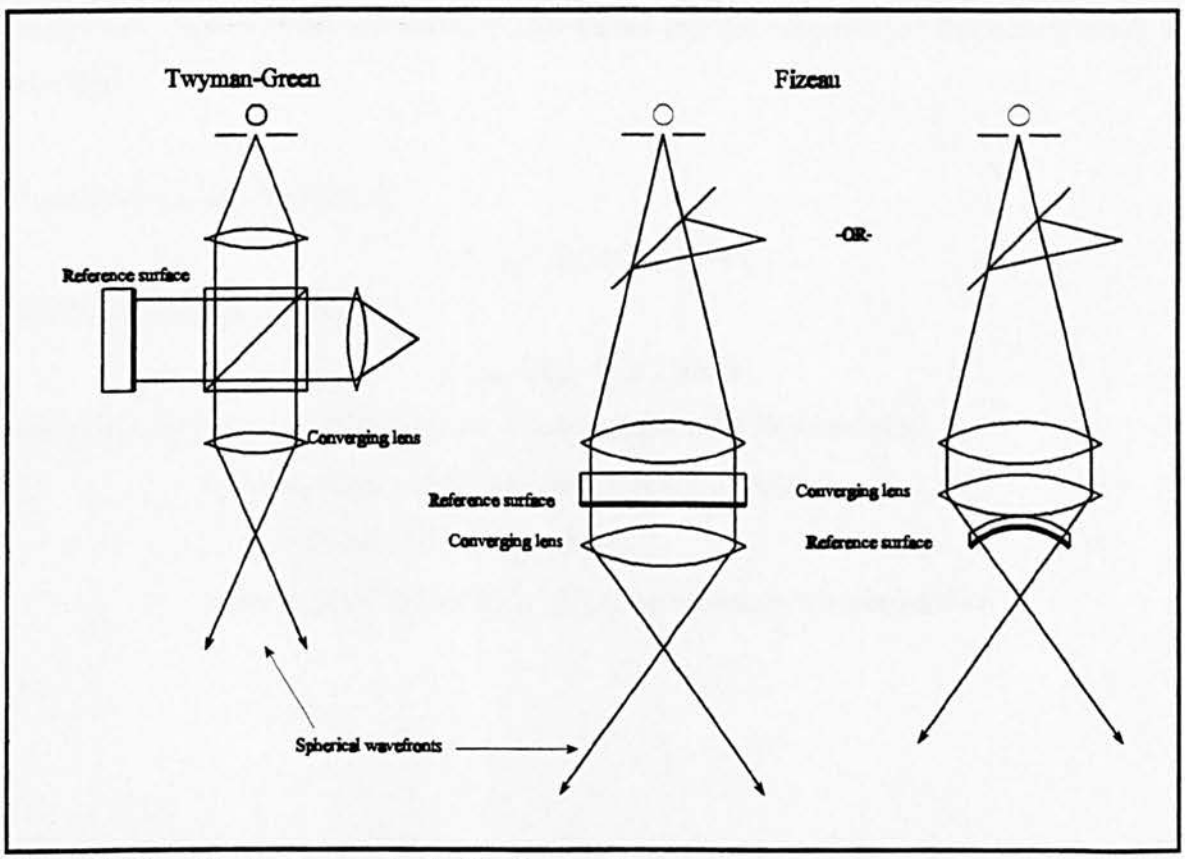


Figure 1.11 Producing a spherical test wavefront.

Figure 1.12 shows test configurations for measuring various non-plane optical surfaces. When testing spherical surfaces, the focus of the converging test wave coincides with the centre of curvature of the spherical test surfaces so that the test wave is normal to the test surface and is reflected back upon itself. Surfaces with conic section figures have two foci and so an auxiliary surface must be included to ensure that the test surface is tested at the appropriate conjugates (Stahl, 1990a). A parabolic surface has one finite real focus and the other at $\pm\infty$. the real focus is placed at the focus of the test wave and the focus at infinity is imaged back onto the real focus by an auxiliary flat mirror. An hyperbolic surface has a real focus, at the focus of the test wave and a virtual focus which is imaged onto the real focus by an auxiliary spherical mirror, centred on the virtual focus. An elliptical surface has two real foci, the first at the focus of the test wave. The second focus is imaged onto the first by an auxiliary spherical mirror centred on the second focus.

It should be noted that, for each of the configurations for testing conic sections, the test wave is reflected twice from the test surface. This means that the sensitivity of the measurement is increased.

For the Fizeau interferometer;

$$g = f_{int} + f_{aux} + 2f_{test}\cos\theta$$

and for the Twyman-Green;

$$g = f_{int} - (f_{aux} + 2f_{test}\cos\theta),$$

where: $f_{int}$ is the figure of the reference surface within the interferometer,

$f_{aux}$ is the figure of the auxiliary reference surface,

$f_{test}$ is the figure of the test surface.

$\theta$ is the angle of incidence of the test wave on the test surface.

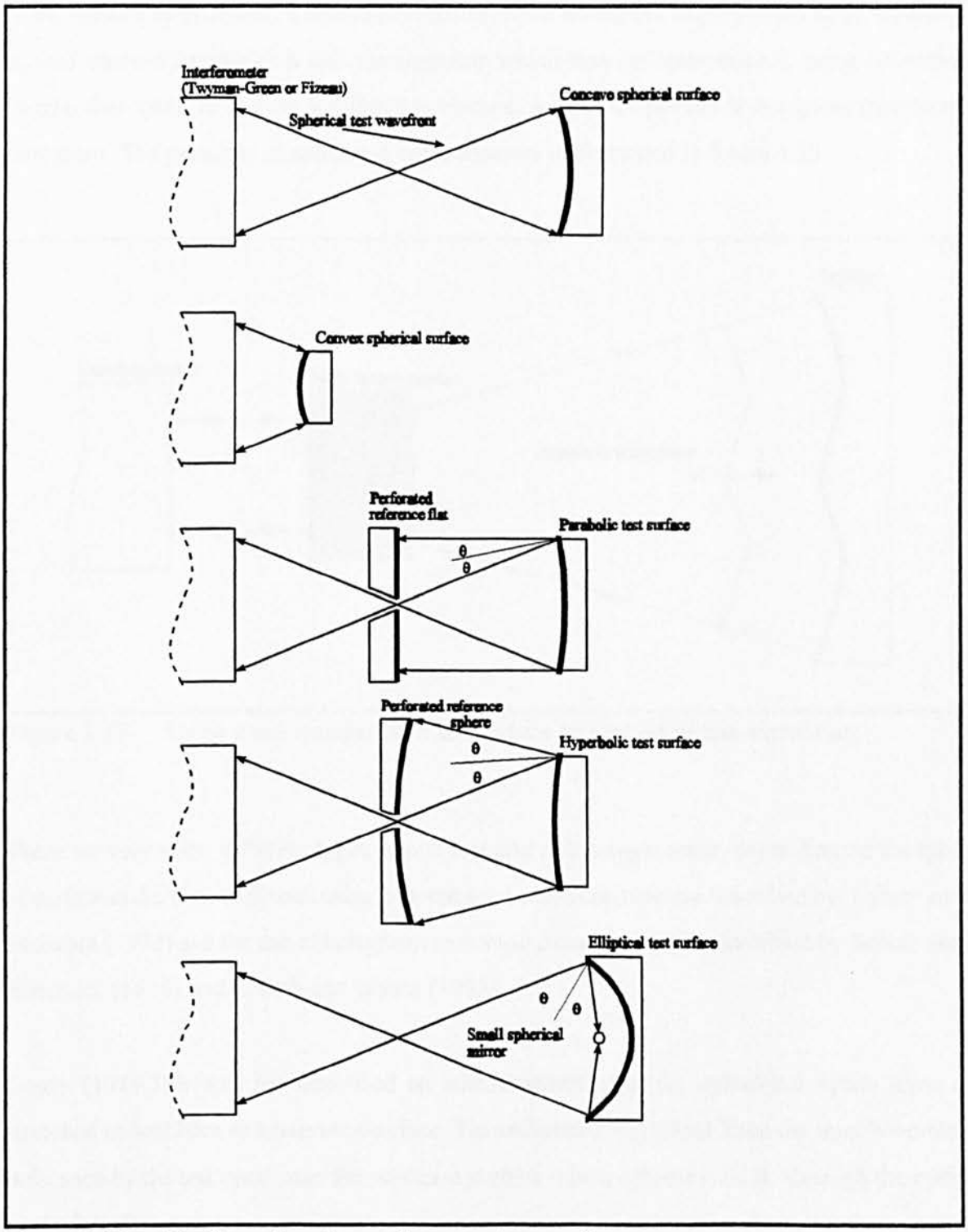


Figure 1.12 Configurations for testing non-plane surfaces.

In order to test the surfaces of general aspheric optical components that do not have simple foci, it is necessary to transform the test wave so that its wavefront matches the nominal shape

of the surface to be tested. The transformation of the wavefront is performed by an auxiliary optical element known as a null compensator which may be conventional, using reflective orrefractive optics, or may be a diffractive element such as a optically or computer generated hologram. The principle of using null compensators is illustrated in figure 1.13.

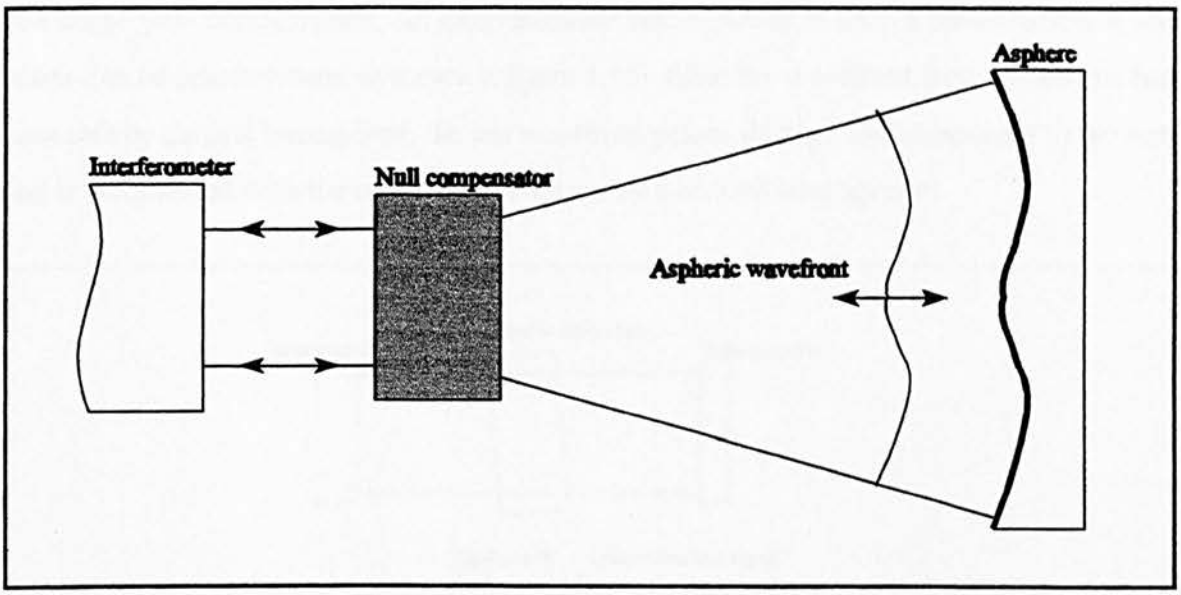


Figure 1.13 Using a null compensator to produce an aspherical test wavefront.

There are very many different types of null test and null compensator, depending on the type of surface under test. Null tests using conventional compensators are described by Offner and Malacara (1992) and the use of holograms as compensator elements is described by Schulz and Schwider (1976) and Creath and Wyant (1992).

Geary (1989,1987a,b) has described an interferometric test for cylindrical optics using a stretched optical fibre as a reference surface. The collimated wavefront from the interferometer is focused by the test optic onto the reference surface which reflects it back through the optic to the interferometer.

As well as measuring the figures of optical surfaces, the Fizeau and Twyman-Green interferometers are well suited to measuring the properties of many optical components and systems in transmission. Test configurations for some typical measurements are shown in

figure 1.14. When using a Twyman-Green or Fizeau interferometer to test a component in transmission, the test wavefront passes twice through the test object. Where the wavefront aberrations introduced by the component are large the returning wavefront may take a quite different path through the component than the outgoing wavefront which will make meaningful analysis of the interferogram difficult. In such cases it is preferable to measure the component in a single pass configuration. An interferometer that is suited to such a measurement is the Mach-Zender interferometer as shown in figure 1.15. After the wavefront from the source has been split by the first beamsplitter, the test wavefront passes through the component under test and is recombined with the reference wavefront by a second beamsplitter.

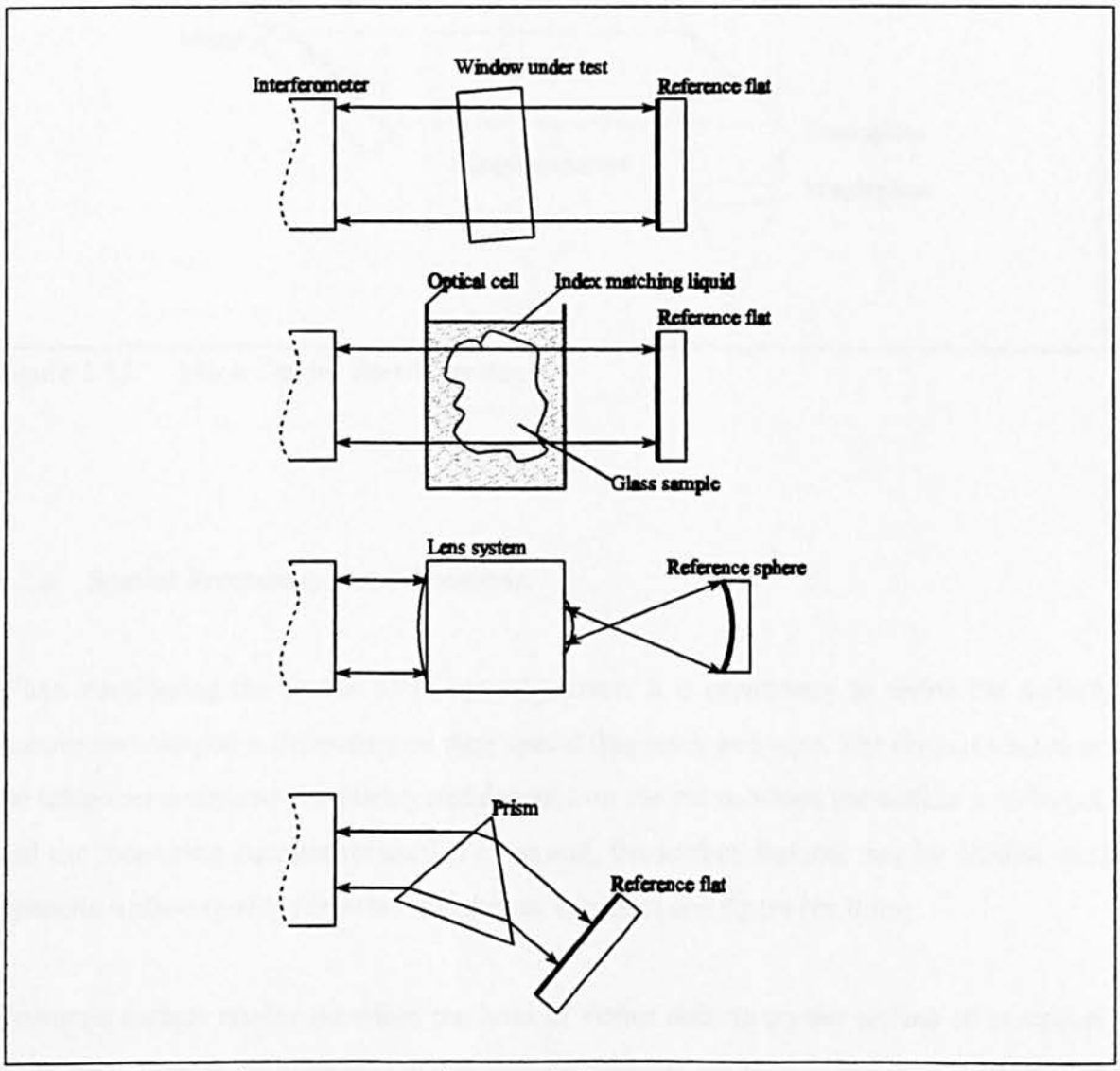


Figure 1.14 Testing components in transmission.

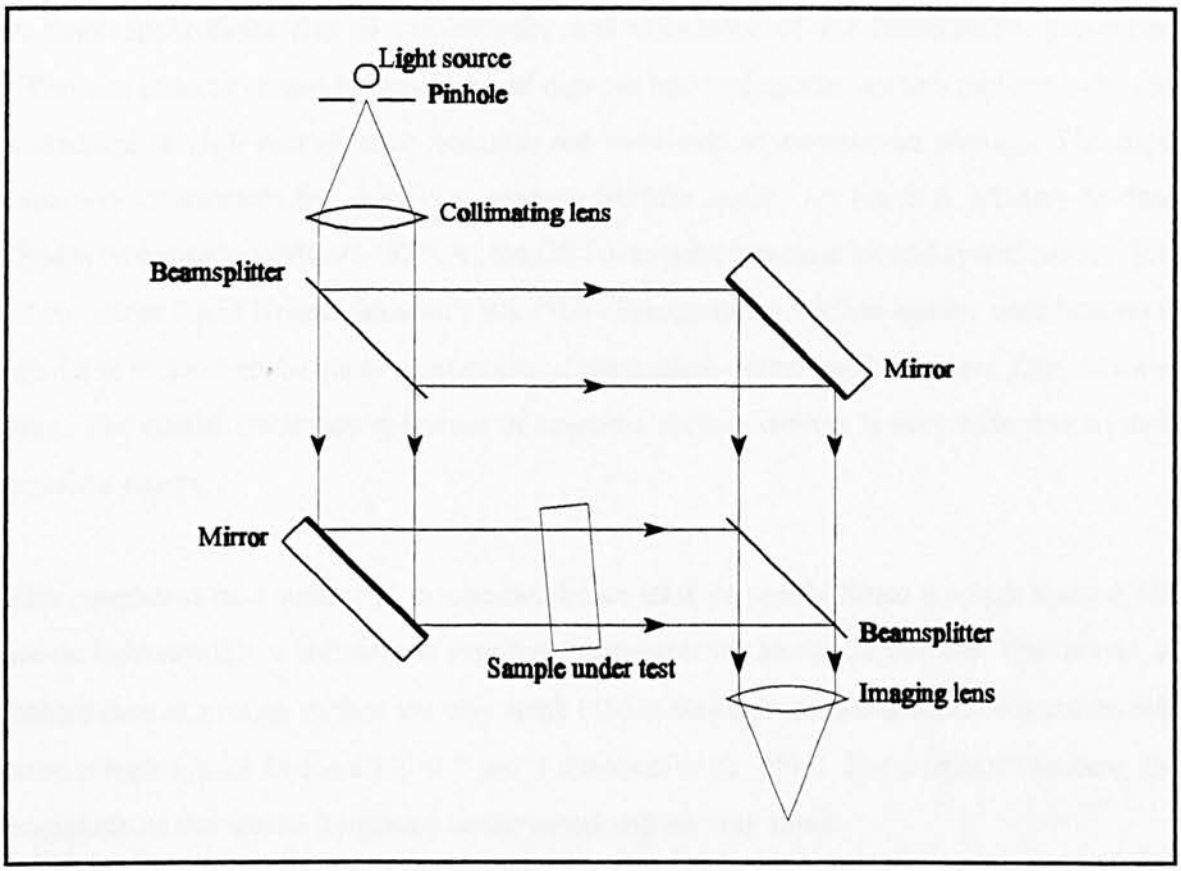


Figure 1.15 Mach-Zender interferometer.

### 1.2.4 Spatial Frequency Considerations.

When considering the profile of an optical surface, it is convenient to divide the surface features into categories depending on their spatial frequency and type. The divisions between the categories is somewhat arbitrary and depends on the use to which the surface is to be put and the measuring instruments used. To this end, the surface features may be divided into; cosmetic surface quality (defects), roughness, waviness and figure (or form).

Cosmetic surface quality describes the level of visible defects on the surface of an optical component. Defects are localized features such as scratches and digs (small pits or craters) on the polished optical surface. These features are important because they can have a serious

adverse effect on performance due to light scattering. Scattering can be particularly important in laser applications due to the intensity and coherence of the illumination. Unwanted diffraction patterns caused by scratches and digs can lead to degraded system performance, and scattering of high energy laser radiation can even lead to component damage. The most common conventions for specifying cosmetic surface quality are the U.S. Military Surface Quality Specification, MIL-0-13830A , the DIN (Deutsche Industrie Norm) specification, DIN 3140, sheet 7 and British Standard, BS 4301. The cosmetic surface quality specification is arrived at subjectively by visual comparison of the surface defects with standard scratches and digs. The spatial frequency spectrum of cosmetic surface defects is very wide due to their transient nature.

The roughness of a surface is determined by its state of polish. Since a rough surface will scatter light strongly, it is clearly an important parameter of the optical surface. The individual feature sizes of a rough surface are very small (10μm down to atomic dimensions) and so will have a high spatial frequency (>0.1 $\mu m^{-1}$) (Bennett et al, 1991). For a smooth surface, the amplitude of the spatial frequency components will be very small.

Surface waviness is a feature commonly found on optical surfaces that have been formed by diamond turning on a lathe. The features arise as a result of the groove formed in the substrate (frequently a metal such as copper) by inelastic deformation of the material as the diamond tool tip removes material to form the required surface figure. The spatial frequencies of the waviness features is usually well defined due to the regularity of the grooves left by the turning process and typically lie in the range 0.1-100 $mm^{-1}$.

The surface feature with the lowest spatial frequencies is the figure or form of the surface. The surface figure describes the overall shape of the surface (spherical, flat etc.) onto which the other surface features are superimposed. The lower limit of the spatial frequency spectrum due to the surface figure is due to the dimensions of the surface itself. The upper limit is generally bound by the spatial resolution of the instrument used to measure the surface form.

The different features of a general optical surface and their spatial frequency spectra are shown

in figure 1.16.

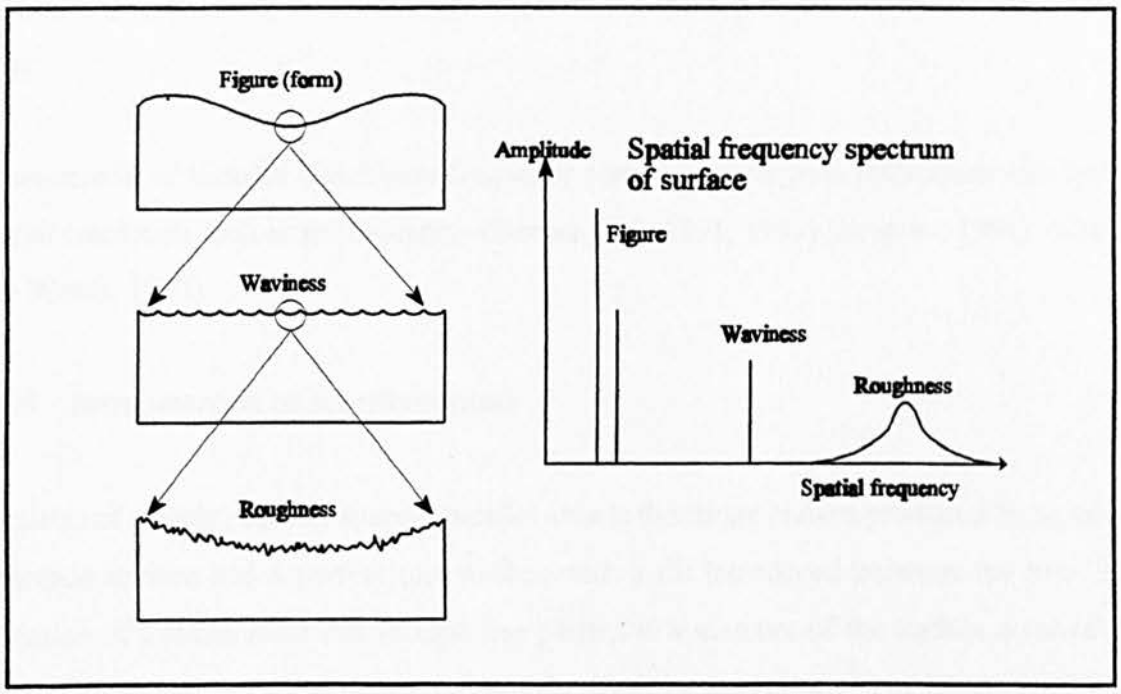


Figure 1.16 Spatial frequency spectrum of a general optical surface.

This thesis is concerned mainly with measurements of surface figure whose highest spatial frequency component is limited by the spatial resolution of the imaging device in the measuring instrument. In the case of the interferometers considered, the spatial resolution is determined by the number of pixels (picture elements) in the electronic camera which images the surface under test. Where the image diameter covers (say) 200 camera pixels the minimum resolved spatial wavelength will be 100$^{th}$ the diameter of the image. This is due to the Nyquist sampling theorem which states that for a signal to be resolved, the sampling frequency must be at least twice the maximum frequency component of the signal.

To determine the spatial frequency spectrum of a surface, it would normally be necessary to determine the surface figure and then to transform this into the spatial frequency domain by a Fourier transform. Hariharan (1996) describes a method for directly determining the spatial frequency spectrum of a nominally flat surface. The method involves making two interferograms of a surface compared to a reference flat, with the surface translated a distance,

x, between the first and second interferograms. Taking the Fourier transform of the difference between the two interferograms reveals the spatial frequency spectrum of the test surface. The lower limit of the spatial frequency range which can be determined by this technique is set by the lateral displacement, x, and the upper limit by the spacing of the elements in the detector array.

Measurement of features with higher frequency components requires techniques with higher spatial resolution such as profilometry (Bennet et al, 1991, 1993) (Bristow, 1991) (Creath and Wyant, 1990).

### 1.2.5 Interpretation of Interferograms.

A pattern of straight, equally spaced, parallel lines is the fringe pattern produced by a perfect reference surface and a perfect test surface with a tilt introduced between the two. The deviation of a fringe from this straight line pattern is a measure of the surface error of the element under test. A review of interferogram analysis methods is given by Malacara (1990).

Test surfaces whose surface accuracy is $\lambda/10$ or worse can be evaluated by visual observation of the fringe pattern produced by the interferometer. Since one fringe separation corresponds to $\lambda/2$, an error of 0.2 fringes corresponds to an error of $\lambda/10$. This can easily be estimated by eye. If the required accuracy is better than $\lambda/10$ then the fringe pattern can be photographed and accurately measured by manual or semi-automatic means. The best accuracy is obtained by capturing the fringe pattern with an electronic camera and performing automatic analysis as described in the section on phase measuring techniques (section 1.2.6).

A typical interferogram might appear as shown in figure 1.17.

Superimposed upon the fringe pattern is a grid corresponding to perfect straight fringes with the average spacing of the interference fringes. Remembering that one fringe spacing corresponds to a wavefront error of $\lambda/2$, the wavefront error at a point on an interference fringe is the deviation of that point from the average fringe grid divided by twice the fringe

spacing. Wavefront error= A/B wavelengths. This is the method adopted by the American Society for Testing and Materials for the manual interpretation of interferograms (Glassman, 1978, Bissinger, 1978). Determination of whether the wavefront error is positive or negative depends on knowledge of the order of the fringes. The fringe order can be determined by finding the thin end of the wedge between the reference and test surfaces since the fringe order decreases towards the thin end (the zero order fringe occurs where the OPD is zero). The direction of the wedge is easily found by slightly changing the tilt between the two surfaces. If the fringe density increases then the wedge angle is increasing.

Two disadvantages of manual methods of interferogram interpretation are, (i) that quantitative analysis is time consuming and tedious and, (ii) that information may only be derived at the fringe positions. These problems may, to some extent, be alleviated by semi-automatic analysis methods using computer programs (Swantner, 1985). The centres of the interference fringes are digitized (either manually or by computer processing of a CCD camera image) and passed to a computer program for analysis. An example of a commercial program of this type is APEX interferogram analysis software (Telic Optics, 576 Boston Rd.E., Marlborough, Mass. USA). The problem of the sparsity of data points is addressed by interpolation or fitting of polynomial functions to the existing data points. A family of polynomial functions well suited to the representation of interferogram data are the Zernike polynomials.

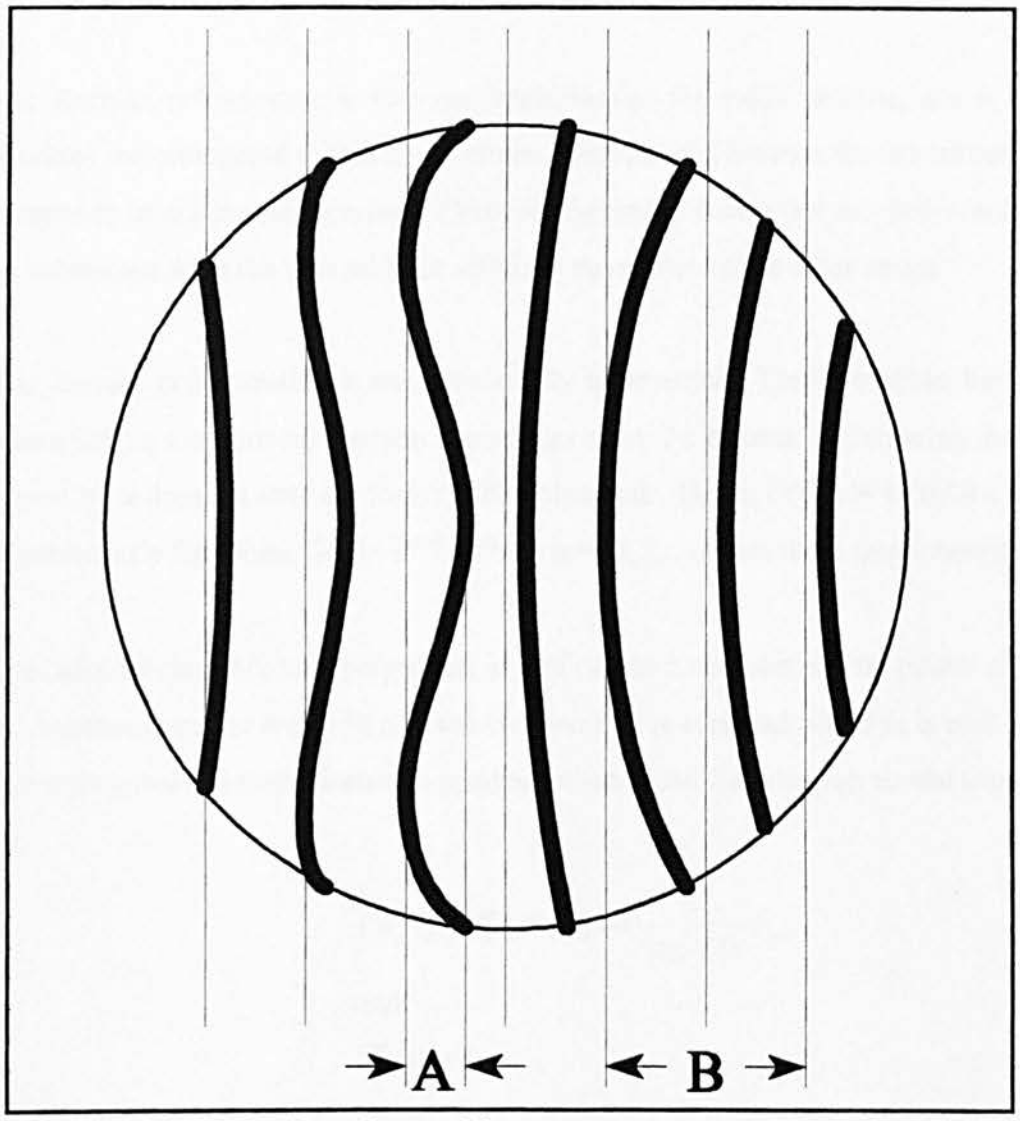


Figure 1.17 Typical interference fringe pattern.

#### 1.2.5.1 Zernike Polynomials.

It is often convenient to express wavefront data in a polynomial form as a means of interpolating between a sparse array of data points or of dissecting the data into physically meaningful terms. A number of properties of the Zernike polynomials make them particularly well suited to this task (Wyant and Creath, 1992) .

The individual terms of the Zernike polynomials correspond to many of the aberrations commonly found in optical tests (defocus, astigmatism, coma etc.).

The Zernike polynomials, in two real variables ($\rho$, the radial variable, and $\theta$, the angular variable) are orthogonal within a unit circle. This is useful because the wavefronts under test commonly have a circular aperture. Their orthogonality means that any polynomial term may be subtracted from the data without affecting the values of the other terms.

The Zernike polynomials are also rotationally symmetrical. They are of the form $R(\rho)G(\theta)$ where $G(\theta)$ is a continuous function that repeats every $2\pi$ radians and rotating the coordinate system by $\alpha$ does not alter the form of the polynomial. Hence $G(\theta,\alpha)= G(\theta)G(\alpha)$. The set of trigonometric functions, $G(\theta)= e^{\pm im\theta}$, where m=0,1,2,... meets these requirements.

The radial function $R(\rho)$ is a polynomial in $\rho$ of degree n and contains no power of $\rho$ less than m. Another property is that $R(\rho)$ must be even if m is even and odd if m is odd.
The orthogonal and normalization properties of the radial Zernike polynomial terms are given by;

$$\int_0^1 R_n^m(\rho)R_{n'}^m(\rho)\rho d\rho = \frac{1}{2(n+1)}\delta_{nn'}$$

and

$$R_n^m(1) = 1$$

The radial polynomial may conveniently be factored into;

$$R_{2n-m}^m(\rho) = Q_n^m(\rho)\rho^m,$$

where $Q^m_n(\rho)$ is a polynomial of order 2(n-m) and can be written;

$$Q_n^m(\rho) = \sum_{s=0}^{n-m}(-1)^s \frac{(2n-m-s)!}{s!(n-s)!(n-m-s)!}\rho^{2(n-m-s)}.$$

In practice the radial polynomials are combined with sine and cosines rather than the complex exponential form of the angular term. The final Zernike polynomial series, representing the

wavefront, W, can be written as

$$W = \sum_{1}^{\infty}\left[A_n Q_n^0(\rho) + \sum_{m=1}^{n} Q_n^m(\rho)\rho^m(B_{nm}\cos m\theta + C_{nm}\sin m\theta)\right]$$

where $A_n$, $B_{nm}$, $C_{nm}$ are individual polynomial coefficients. The table below lists the first few Zernike polynomials together with the common optical aberration which they represent (where appropriate). Notice that each term contains the appropriate amount of each lower order term to make it orthogonal to those terms.

The first 16 Zernike Polynomials.

| n | m | No. | Polynomial | Aberration |
|---|---|---|---|---|
| 0 | 0 | 0 | 1 | Piston |
| 1 | 1 | 1 | $\rho\cos\theta$ | x-tilt |
| | | 2 | $\rho\sin\theta$ | y-tilt |
| | 0 | 3 | $2\rho^2-1$ | Defocus |
| 2 | 2 | 4 | $\rho^2\cos2\theta$ | 0° astigmatism and defocus |
| | | 5 | $\rho^2\sin2\theta$ | 45° astigmatism and defocus |
| | 1 | 6 | $(3\rho^2-2)\rho\cos\theta$ | Coma and x-tilt |
| | | 7 | $(3\rho^2-2)\rho\sin\theta$ | Coma and y-tilt |
| | 0 | 8 | $6\rho^4-6\rho^2+1$ | Third order spherical and defocus |
| 3 | 3 | 9 | $\rho^3\cos3\theta$ | |
| | | 10 | $\rho^3\sin3\theta$ | |
| | 2 | 11 | $(4\rho^2-3)\rho^2\cos2\theta$ | |
| | | 12 | $(4\rho^2-3)\rho^2\sin2\theta$ | |
| | 1 | 13 | $(10\rho^4-12\rho^2+3)\rho\cos\theta$ | |
| | | 14 | $(10\rho^4-12\rho^2+3)\rho\sin\theta$ | |
| | 0 | 15 | $20\rho^6-30\rho^4+12\rho^2-1$ | |

The fitting of Zernike polynomials to a discrete set of wavefront data points is usually accomplished by first fitting an x-y polynomial to the measured data. The x-y polynomials are

then transformed to a linear series of Zernike polynomials by means of a linear transformation using matrix multiplication. This method has the disadvantage that errors are introduced by the intermediate x-y polynomial stage because of rounding errors and because the x-y polynomials are orthogonal on a unit square and not a circle. Malacara et al (1990) describe a method of overcoming these drawbacks by performing a least squares fitting using orthogonal polynomials over a discrete set of data points in a unit circle which tend towards the Zernike polynomials as the number of data points becomes large.

### 1.2.6 Phase Measuring Techniques.

There are two main categories of automatic phase measurement techniques. The phase may either be determined electronically or analytically. The electronic methods are mostly heterodyne techniques. Many of the first techniques were developed before the advent of area detector arrays and desktop computers. Because most electronic methods measure only one point at a time they must scan to examine an area, their main application today is in distance measuring interferometry.

Analytical techniques make use of area detector arrays and easily available computer processing to extract, essentially in real time, phase information from an entire interferogram at once. Analytical techniques include phase shifting and spatial carrier techniques.

#### 1.2.6.1 Heterodyne Techniques.

The term heterodyne refers to the mixing of two different frequencies to produce a beat signal at the difference frequency. As the name implies, in a heterodyne interferometer, the light waves in the reference and test beams have different frequencies. The difference in frequencies is very slight compared to the frequency of visible light (~$5 \times 10^{14}$ Hz) and may be of the order of a few MHz, depending on the technique used.

The beat (difference) frequency produced by the interference between the reference and test beams is compared to a reference sinusoidal signal which may be produced optically or

electronically. The time delay between the crossing of the zero phase points of the test and reference sinusoidal signals is a measure of the phase. Every time the test signal passes through another zero in the same direction, another fringe is counted. If the beam is interrupted as the detector is scanned across the interferogram, the fringe count is corrupted and the measurement needs to begin again.

Today, heterodyne techniques are used mainly in distance measuring interferometers. An example is shown in figure 1.18 though there are many different implementations of this type of instrument.

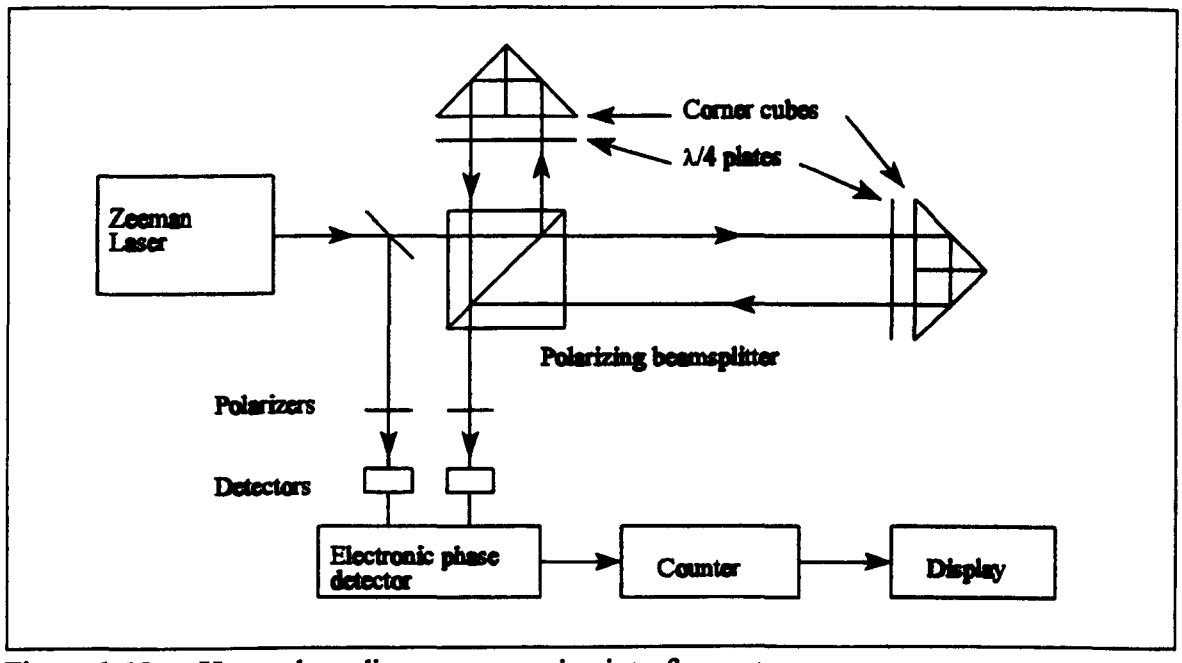


Figure 1.18 Heterodyne distance measuring interferometer.

The light source is a Zeeman laser which produces two orthogonally polarised outputs with frequencies a few MHz apart. A small fraction of both outputs is split off before the polarization beamsplitter to provide the reference signal. A polarizer at 45° is used to combine the beams for the reference signal. The reference signal is sinusoidal in time at the beat frequency between the two beams.

The reference and test beams are split by a polarizing beam splitter (PBS). Corner cubes are used to reflect the beams back into the PBS. Quarter wave plates in each beam rotate the polarizations by 90° to the direction they entered. The test and reference beams are

recombined by a polarizer at 45° and detected by a second detector.

The phase, modulo $2\pi$, is found by electronically detecting the phase difference between the signals from the two detectors. As the corner cube in the test arm is moved in and out relative to the rest of the interferometer, the electronics count the number of fringes which pass by. This enables the distance over which the test beam corner cube is moved to be measured. Using sophisticated electronics, distances may easily be measured to $\lambda/1000$ or better.

The very high sensitivity of heterodyne techniques have led to its use as an optical profiling tool (eg. Sommargren, 1981). Profilometers, whatever their measurement technique, are usually scanning instruments and commonly scan just a line on an object rather than a whole area. This means that the need to scan the detector is not a disadvantage when using heterodyne techniques for profilometry.

Some authors (Massie, 1980, Barnes, 1987) have reported heterodyne interferometers for measuring areas.

Massie's interferometer is shown, simplified, in figure 1.19.

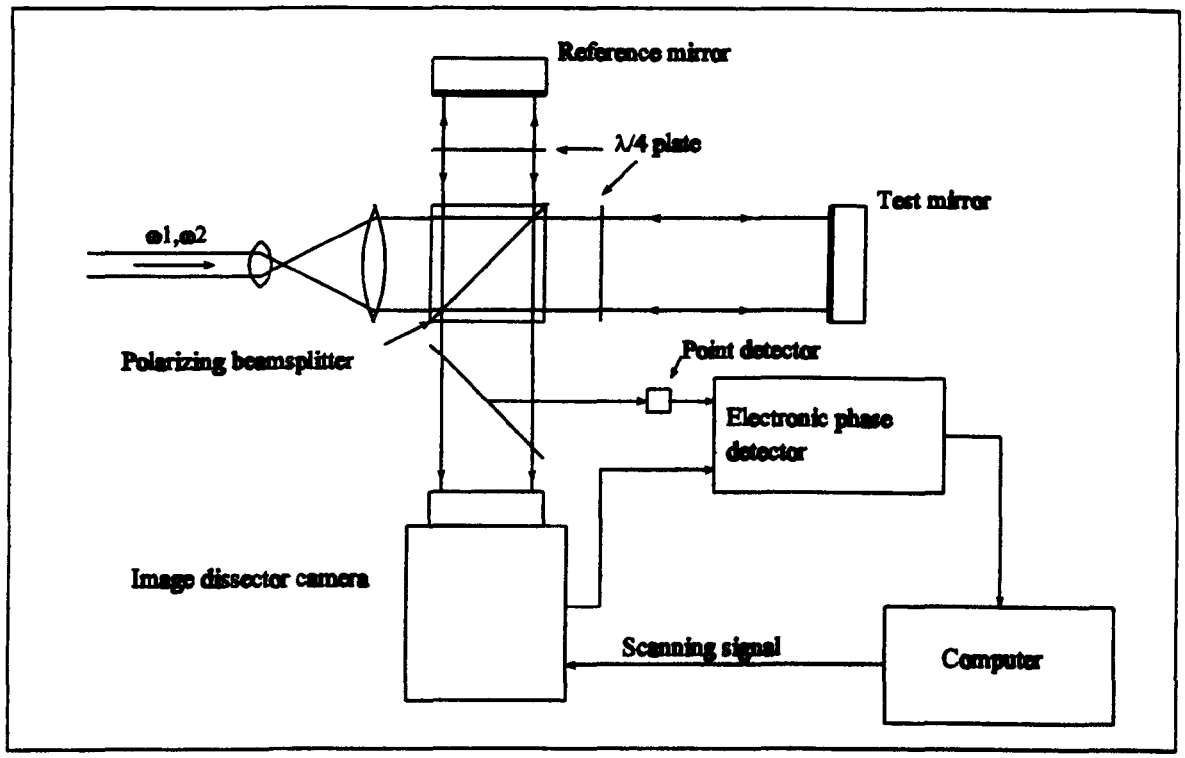

Figure 1.19 Massie's heterodyne interferometer.

The light source consists of two optical frequencies, $\omega_1$ and $\omega_2$ with orthogonal polarizations. The two frequencies are generated from a single frequency emitted by a Krypton ion laser in an auxiliary Mach-Zender interferometer (not shown for clarity). Each arm of the Mach-Zender interferometer includes an acousto-optic Bragg cell to shift the frequency in that arm. One Bragg cell is driven at 42 MHz and the other at 43 MHz and so the recombined beams have a frequency difference of 1 MHz. One arm of the Mach-Zender interferometer contains a $\lambda/2$ waveplate to rotate the polarization in that arm by 90° so that the two output frequencies are orthogonally polarized.

The main interferometer is a polarizing Twyman-Green type with $\omega_1$ in one arm and $\omega_2$ in the other. A reference signal for the electronic phase detector is sampled from a point in the image plane while the fringe pattern, modulated at the difference frequency, 1 MHz, is detected by an image dissector camera. In an image dissector camera, the image is formed on a photo-cathode which emits electrons when incident photons are absorbed. The electrons are focused by field coils and accelerated towards an anode where they are detected. The point on the photo-cathode from which the electrons are detected is scanned by controlling the field coil

currents. The phase difference between the reference and camera signals is electronically detected, digitized and stored on a computer which also generates the scanning control signals. The scan rate is limited only by the settling time of the electronic phase detector and is quoted as 50μs per point for this system which is very much faster than could be achieved by mechanical scanning and could rival standard video rates with more sophisticated electronics. The quoted phase accuracy is $\lambda/70$ which could also, no doubt, be improved.

The principle of Barnes' heterodyne Fizeau interferometer is shown in figure 1.20.

The laser light is diffracted by a rotating radial diffraction grating. Spatial filter, SF1, allows only the +1 and -1 diffracted orders to pass and these have a frequency difference of 2f, where f is the line passing frequency of the rotating grating. The separation of the two diffracted orders is greatly exaggerated in the figure for clarity. The two diffracted orders are collimated by a collimating lens and reflected from the reference and test surfaces. The two surfaces are tilted slightly so that only the +1 order from one surface and the -1 order from the other surface pass through spatial filter SF2. The light from the test and reference surfaces interferes in the focal plane of the imaging lens to produce an interference pattern modulated at frequency 2f. A stationary detector, D1, in the image plane produces a phase reference signal and detector D2 is scanned across the interference pattern. The relative phase of the signals from D1 and D2 is measured by an electronic phase detector to give a phase map of the interference pattern as D2 is scanned.

A resolution of $\lambda/200$ is claimed for the interferometer but the long measurement time necessitated by the mechanical scanning of the detector means that the interferometer accuracy is limited by slow drifts in the orientation of the test and reference surfaces and other components.

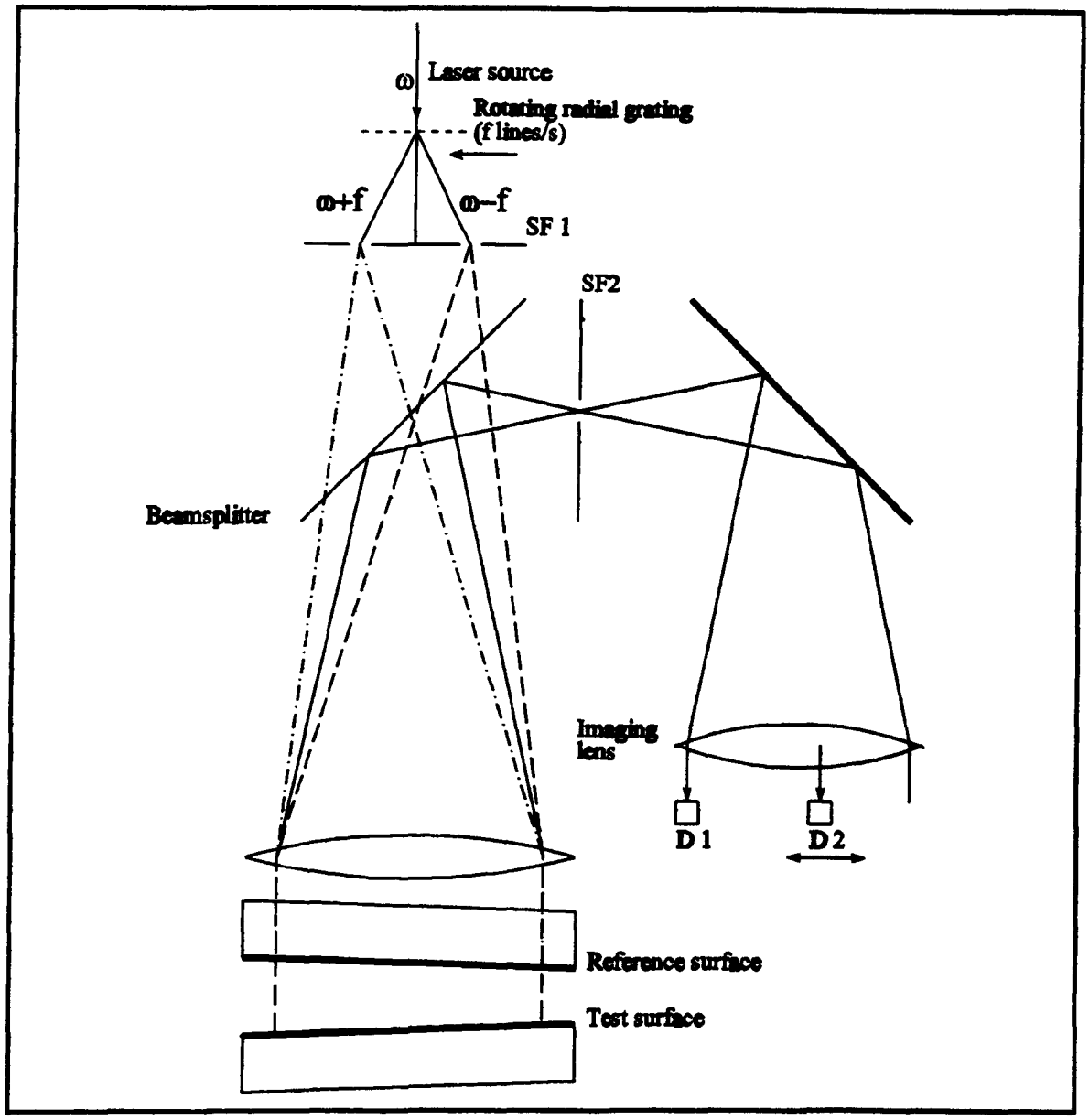


Figure 1.20 Barnes' heterodyne interferometer.

### 1.2.6.2 Phase Shifting (Quasi-Heterodyne) Techniques.

The development of area image detector arrays (principally charge-coupled devices (CCDs)) has meant that whole fringe pattern images may be electronically captured, digitized and processed by computer. The finite integration (exposure) time of CCDs and the fact that the data is read out in a serial fashion (one pixel at a time) means that true heterodyne techniques may not be used.

In phase shifting techniques, several images (frames) of the interference pattern are acquired by the CCD camera and stored by a computer. The phase difference between the reference and test beams in the interferometer is shifted by a known amount between the frames. The phase shift may be accomplished by a number of methods (Wyant and Creath, 1985, Crane 1969) but the most common is to vary the OPD by translating either the reference or test surface along the optical axis by means of a piezo-electric transducer.

As the phase shifter is moved, the phase at each point in the interferogram changes giving the appearance that the fringes are moving across the interferogram. Because of this, the techniques are sometimes called fringe scanning or fringe shifting techniques.

The intensity at each pixel is given by:

$$I(x,y) = I_0(x,y)\{1+\gamma_0(x,y)\cos[\phi(x,y)+\alpha(t)]\}$$

where: I(x,y) is the intensity at a single detector point,

$I_0(x,y)$ is the average intensity,

$\gamma_0(x,y)$ is the fringe visibility,

$\phi(x,y)$ is the phase of the wavefront being measured,

$\alpha(t)$ is the phase shift as a function of time, t.

Since the detector has to integrate for some finite time, the detected intensity at a single point can be written as the integral of the instantaneous intensity over the integration time, $\Delta$. The average phase shift for the $i^{th}$ frame of data is $\alpha_i$.

$$I_i(x,y) = \frac{1}{\Delta}\int_{\alpha_i-\Delta/2}^{\alpha_i+\Delta/2} I_0(x,y)\{1+\gamma_0(x,y)\cos[\phi(x,y)+\alpha(t)]\}d\alpha(t)$$

After integrating over $\alpha(t)$, the intensity of the detected signal becomes

$$I_i(x,y) = I_0(x,y)\{1+\gamma_0(x,y)\mathrm{sinc}\left(\tfrac{\Delta}{2\pi}\right)\cos[\phi(x,y)+\alpha_i]\}$$

where $\mathrm{sinc}(\Delta/2) = \sin(\Delta/2)/(\Delta/2)$.

There are two basic methods of varying the phase between frames. The phase may be varied continuously while frames are acquired or may be varied in a step-wise fashion with the phase constant during the acquisition period. The first method is potentially faster but requires more sophistication from the electronics.

Mathematically, the only difference between the phase ramping and phase stepping techniques is a reduction in the detected fringe modulation for the ramping method. This is due to the sinc term. When the phase is stepped ($\Delta=0$), the sinc term has a value of 1. When the phase is ramped ($\Delta \leq \alpha$) and $\alpha = \pi/2$ (90°) then the sinc term will be between 0.9 and 1. $\Delta$ will generally be less than $\alpha$ because of the finite time it takes to read out the CCD array between frames.

Once a number of frames of data have been acquired the phase is determined computationally by one of a number of techniques (Malacara, 1992). Some of these will now be described.

#### 1.2.6.2.1 Three-Frame Technique.

Three frames of intensity data are the minimum required to calculate the wavefront phase since there are three unknowns, $I_0$, $\gamma_0$ and $\phi$ in the interferogram intensity equation.

Using phase shifts of $\alpha_i = \pi/4$, $3\pi/4$ and $5\pi/4$, the intensity distributions of the interferograms may be expressed as:

$$I_1 = I_0\left[1+\gamma\cos\left(\phi+\frac{\pi}{4}\right)\right] = I_0\left[1+\frac{\sqrt{2}}{2}\gamma(\cos\phi-\sin\phi)\right]$$

$$I_2 = I_0\left[1+\gamma\cos\left(\phi+\frac{3\pi}{4}\right)\right] = I_0\left[1+\frac{\sqrt{2}}{2}\gamma(-\cos\phi-\sin\phi)\right]$$

$$I_3 = I_0\left[1+\gamma\cos\left(\phi+\frac{5\pi}{4}\right)\right] = I_0\left[1+\frac{\sqrt{2}}{2}\gamma(-\cos\phi+\sin\phi)\right]$$

Note that the

$$I_3 - I_2 = \sqrt{2} I_o \gamma \sin\phi$$
$$I_1 - I_2 = \sqrt{2} I_o \gamma \cos\phi$$

(x,y) dependencies are still implied and $\gamma$ is $\gamma_0$ multiplied by the constant sinc term.

The phase at each detector point is then simply;

$$\phi = \tan^{-1}\left(\frac{I_3 - I_2}{I_1 - I_2}\right)$$

The fringe visibility can be calculated using;

$$\gamma = \frac{\sqrt{(I_3 - I_2) + (I_1 - I_2)}}{\sqrt{2} I_0}$$

It is easy to check the signal modulation ($2I_0\gamma$) and to set a threshold on it of about 5-10%. If the modulation is less than this value at any given data point, that point is flagged as "bad". Bad points are usually caused by noisy pixels and can be due to scratches, dust, scattered light etc.

When a general phase shift, is used, the three intensity measurements become:

$$I_1 = I_0[1 + \gamma\cos(\phi - \alpha)]$$
$$I_2 = I_0[1 + \gamma\cos\phi]$$
$$I_3 = I_0[1 + \gamma\cos(\phi + \alpha)]$$

The phase can be calculated using;

$$\phi = \tan^{-1}\left[\left(\frac{1 - \cos\alpha}{\sin\alpha}\right)\frac{I_1 - I_3}{2I_2 - I_1 - I_3}\right]$$

Phase shifts of $\alpha = 2\pi/3$ (120°) are also commonly used, in which case the phase is given by;

$$\phi = \tan^{-1}\left(\sqrt{3}\frac{I_1 - I_3}{2I_2 - I_1 - I_3}\right)$$

#### 1.2.6.2.2 Synchronous Detection.

An early technique for phase measurement utilized methods of communication theory to perform synchronous detection. To synchronously detect a noisy signal, it is correlated (or multiplied ) with sinusoidal and cosinusoidal signals of the same frequency and averaged over many periods of oscillation. The method of synchronous detection as applied by Bruning (1974) to phase measurement involves N measurements that are equally spaced over one modulation period with phase shifts such that;

$$\alpha_i = \frac{i2\pi}{N}, \text{ with } i = 1,\ldots,N$$

the phase can be calculated from;

$$\phi(x,y) = \tan^{-1}\left[\frac{\sum I_i(x,y)\sin\alpha_i}{\sum I_i(x,y)\cos\alpha_i}\right]$$

Note that N can be any number of frames. The more frames of data the less error. This technique does not take up a large amount of memory for a large number of frames, because only running sums of the intensities multiplied by the sine and cosine of the phase shift need to be kept.

#### 1.2.6.2.3 Carré (1966) Technique.

This technique uses four frames of intensity data and assumes that the phase shift is not known, but that it is constant from frame to frame. In this case, the measured intensity data frames can be written as:

$$I_1 = I_0\left[1+\gamma\cos\left(\phi-\frac{3\alpha}{2}\right)\right]$$
$$I_2 = I_0\left[1+\gamma\cos\left(\phi-\frac{\alpha}{2}\right)\right]$$
$$I_3 = I_0\left[1+\gamma\cos\left(\phi+\frac{\alpha}{2}\right)\right]$$
$$I_4 = I_0\left[1+\gamma\cos\left(\phi+\frac{3\alpha}{2}\right)\right]$$

where the phase shift is assumed to be linear. From these equations, the phase shift can be calculated using;

$$\alpha = 2\tan^{-1}\left[\sqrt{\frac{3(I_2-I_3)-(I_1-I_4)}{(I_2-I_3)+(I_1-I_4)}}\right]$$

and the phase at each point is;

$$\phi = \tan^{-1}\left\{\tan\left(\frac{\alpha}{2}\right)\left[\frac{(I_1-I_4)+(I_2-I_3)}{(I_2+I_3)-(I_1+I_4)}\right]\right\}$$

These two equations combine to yield;

$$\phi = \tan^{-1}\left\{\frac{\sqrt{[(I_1-I_4)+(I_2-I_3)][3(I_2-I_3)-(I_1-I_4)}}{(I_2+I_3)-(I_1+I_4)}\right\}$$

For this technique, the fringe visibility is given by;

$$\gamma = \frac{1}{2I_0}\sqrt{\frac{[(I_1-I_4)+(I_2-I_3)]^2+[(I_2+I_3)-(I_1+I_4)]^2}{2}}$$

where it assumed that $\alpha$ is near $\pi/2$.

An obvious advantage of this technique is that the phase shift does not have to be calibrated. It is, however, computationally intensive.

### .2.6.2.4 Four Frame Technique.

This technique is the same as that for synchronous detection with N=4 and leads to an easy-to-calculate equation. The four frames of intensity data are given by;

$$I_1 = I_0[1+\gamma\cos\phi]$$

$$I_2 = I_0[1+\gamma\cos\left(\phi+\frac{\pi}{2}\right)] = I_0[1-\gamma\sin\phi]$$

$$I_3 = I_0[1+\gamma\cos(\phi+\pi)] = I_0[1-\gamma\cos\phi]$$

$$I_4 = I_0[1+\gamma\cos\left(\phi+\frac{3\pi}{2}\right)] = I_0[1+\gamma\sin\phi]$$

The phase at each point is;

$$\phi = \tan^{-1}\left(\frac{I_4 - I_2}{I_1 - I_3}\right),$$

and the fringe visibility is;

$$\gamma = \frac{\sqrt{(I_4 - I_2)^2 + (I_1 - I_3)^2}}{2I_0}$$

### 1.2.6.2.5 Five Frame Technique.

The most popular technique used in commercial interferometers today was developed by Hariharan et al (1987). It utilizes 5 frames of intensity data, and is popular because it is quite insensitive to common systematic errors which are present in phase measuring interferometers and data can still be taken rapidly because the number of frames is relatively small.

This algorithm is designed to reduce the possibility of having the numerator and denominator tend to zero, and thereby reduce the uncertainty in the calculation. This algorithm uses five frames of intensity data with relative phase shifts of $\alpha=\pi/2$.

$$I_1 = I_0[1+\gamma\cos(\phi-\pi)] = I_0[1-\gamma\cos\phi]$$
$$I_2 = I_0\left[1+\gamma\cos\left(\phi-\frac{\pi}{2}\right)\right] = I_0[1+\gamma\sin\phi]$$
$$I_3 = I_0[1+\gamma\cos\phi]$$
$$I_4 = I_0\left[1+\gamma\cos\left(\phi+\frac{\pi}{2}\right)\right] = I_0[1-\gamma\sin\phi]$$
$$I_5 = I_0[1+\gamma\cos(\phi+\pi)] = I_0[1-\gamma\cos\phi]$$

The phase calculated from this set of intensities is given by;

$$\phi = \tan^{-1}\left[\frac{2(I_2-I_4)}{2I_3-I_5-I_1}\right]$$

This is a very simple calculation and has a large tolerance to miscalibration of phase shift. For this technique, the fringe visibility is given by;

$$\gamma = \frac{\sqrt{[2(I_2-I_4)^2]+(2I_3-I_5-I_1)^2}}{4I_0}$$

#### 1.2.6.2.6 "2+1" Frame Technique.

When measuring large optical components such as telescope mirrors, not all of the components will fit onto a single vibration isolation system. Because of this, there is likely to be a lot of vibration and air turbulence causing the interference fringes to move around. In this case, the data need to be acquired as fast as possible to freeze the interference fringes in time. Since the data need to be taken fast, the number of data frames needs to be minimized. The "2+1" frame technique is aimed at solving these problems.

Three frames of data are required by this technique. They are written as;

$$I_1 = I_0[1+\gamma\cos\phi]$$
$$I_1 = I_0\left[1+\gamma\cos\left(\phi-\frac{\pi}{2}\right)\right] = I_0[1+\gamma\sin\phi]$$
$$I_3 = \frac{1}{2}\{I_0[1+\gamma\cos\phi]\} + \frac{1}{2}\{I_0[1+\gamma\cos(\phi+\pi)]\} = I_0$$

The first two frames of data are taken very quickly with a $\pi/2$ phase shift between them. The third frame is the DC intensity (the average of two frames with a phase shift of $\pi$ between them) which can be acquired at any time.

There are only two frames of data which can be affected by vibration and air turbulence. If these two frames are taken on either side of the interline transfer in a standard CCD camera, 1ms exposures can be taken as quickly as 1μs apart. This will freeze most vibrations and air turbulence that may affect the measurement. These three frames of data can then be combined to calculate the wavefront phase from;

$$\phi = \tan^{-1}\left(\frac{I_2 - I_3}{I_1 - I_3}\right)$$

The fringe visibility for this technique is given by;

$$\gamma = \frac{\sqrt{(I_2 - I_3)^2 + (I_1 - I_3)^2}}{I_0}$$

#### 1.2.6.2.7 Scanning Phase Shift Technique.

Another technique developed for use in the presence of vibration and air turbulence utilizes a large number of data frames. This technique was originally developed (Vikhagen, 1990) for TV holography and looking at large structures that could not be isolated; however it can be useful for other applications.

For this technique, many frames of data with random phase shifts are collected.

$$I_i = I_0[1+\gamma\cos(\phi+\alpha_i)].$$

$\alpha_i$ will be a random value between 0 and $2\pi$. Every time a new data frame is recorded, the maximum and minimum intensity value at each detector point is determined. When the number of frames becomes large, these values will approach the maximum and minimum fringe intensities, $I_{max}$ and $I_{min}$. This means that these values can be used to determine the DC intensity, $I_0$, and the fringe visibility, $\gamma$, at each data point, leaving the phase, $\phi$, as the only remaining unknown. The phase can be calculated using;

$$\phi = \cos^{-1}\left(\frac{I_i - I_0}{\gamma I_0}\right)$$

where $I_i$ is the intensity data frame with a phase shift of $\alpha_i$, and $I_0$ and $\gamma$ are calculated from;

$$I_0 = \frac{I_{max} + I_{min}}{2},$$

and

$$\gamma = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} = \frac{I_{max} - I_{min}}{2I_0}$$

Once a large number of frames has been recorded, the phase can be calculated as each new intensity frame is recorded. This technique will work as long as vibration and air turbulence do not cause the image of the object to move more than a fraction of a pixel from frame to frame. This algorithm may easily be programmed in an array processor to calculate the phase modulo $2\pi$ at video rates.

#### 1.2.6.2.8 Error Sources in Phase Measuring Techniques.

There are many sources of error in interferometric measurement. Some sources of error are common to all interferometric techniques. Phase measuring techniques have some extra potential sources of error which can propagate through the data processing algorithm to give erroneous results (Grievenkamp and Bruning, 1992). Reviews of the various sources of error in phase shifting interferometry are given by Cochran (1992), van Wingerden et al (1991) and Stahl and Tome (1988).

All phase shifting algorithms rely on shifting the relative phases of the test and reference beams

in a known manner. In some cases, such as Carre's algorithm (section 1.2.6.2.3), the actual value of the phase shift is unimportant so long as it is the same between each of the frames of data. Generally, the dependence of the accuracy of a phase shifting algorithm on the accuracy of the phase shift varies between algorithms. Inaccuracies in the phase shift may arise from a number of sources. The piezo electric devices commonly used for phase shifting are generally non-linear in that their change in dimensions does not relate linearly to the applied voltage and in addition, they also exhibit hysteresis. The effect of phase shifter non-linearity may be minimized by calibration of the piezo electric devices and by compensation for their errors in the control electronics or in the processing of the acquired data. The extent to which phase shifter non-linearity affects the various different phase shifting algorithms has been described by a number of authors (Cheng and Wyant, 1985, Ai and Wyant, 1987, Huang, 1992).

In the descriptions given above of the various phase measuring algorithms, the measured quantities from which the phase has been determined are the intensities of the fringes. In order to perform the algorithm, the intensities of the fringes must be converted into digital values for processing by computer. Two sources of potential error are thus detector non-linearity and analogue to digital converter errors. In general, silicon photodetectors such as CCDs are quite linear over a large range of incident intensities. At very low intensities a signal present in the absence of any light called the dark current can be significant. At high intensities, the detector can begin to saturate resulting in non-linearity. Very high contrast fringes with a wide range of intensity values can thus suffer from non-linear detection and the accuracy of the phase measurement can suffer. Again, the extent to which detector non-linearities affect the accuracy of the phase measurement varies depending on the algorithm used as has been shown by Kinnstaetter et al (1988). The analogue to digital (A/D) converters used to convert the CCD data into digital values are easily made very linear but a potential source of errors arises from their quantization of the data. An A/D divides the continuously variable input signal into a number of discreet steps. Video A/D converters commonly convert the camera data into 8 or 10 bit words meaning that the data are divided into 256 ($2^8$) or 1024 ($2^{10}$) steps. Brophy (1990) has derived the error in the computed phase due to quantization of the data for most of the common phase shifting algorithms. For most purposes, sufficient accuracy is achieved by 8 bit A/D converters and errors of less than 0.0001 waves can be achieved by the use of 10 bits.

The long coherence length associated with the lasers usually used as the light sources in phase measuring interferometry can give rise to the presence of coherent noise which can affect the accuracy of the measurement. Coherent noise is caused by spurious reflections and scattering of light from surfaces in the interferometer other than those which are being measured and results in unwanted interference fringes and speckle being superimposed on the desired interference pattern. Coherent noise can be minimized by careful interferometer design (see sections 3.1 and 6.4.2). The effect of spurious reflections has been described by Ai and Wyant (1988) together with means by which it may be minimized for some of the phase shifting algorithms.

Effects due to instabilities in the interferometer caused by instabilities of the laser source, vibration and air turbulence can affect all interferometric measurements. The effect on phase measuring methods can, however, be particularly severe since they rely on the variations in the fringe pattern being due only to the phase shifter and not to other causes. These effects may be minimised by use of stabilized laser sources, careful vibration isolation and thermal control.

Imperfections in the optics of the interferometer other than those in the reference surface can have an effect on the accuracy of phase measurements due to propagation errors. If the light reflected from the test and reference surfaces does not accurately retrace its path through the optical system then the errors in the optics will be apparent in the interference pattern due to wavefront shear (see section 5.3). These effects may be minimized by careful nulling of the interference fringes so that tilts between the interfering wavefronts are a minimum. An analysis of propagation errors is given by Huang (1992,1993) and Joswicki (1991).

#### 1.2.6.3 Spatial Techniques.

##### 1.2.6.3.1 Fourier Transform Technique.

The Fourier transform technique (Roddier and Roddier, 1987) is a way to extract phase from a single interferogram. It is commonly used in non-destructive testing and stellar interferometry where it is difficult to get more than a single interferogram. A flowchart of the basic technique

is shown in figure 1.21.

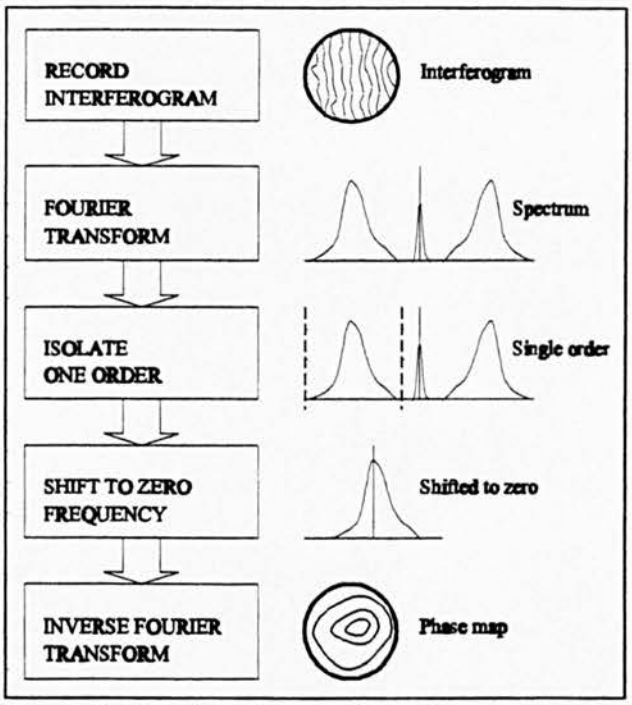


Figure 1.21 Fourier transform phase measuring technique.

A single interferogram is recorded, with sufficient tilt fringes that none of the fringes are closed and that their order increases monotonically across the interferogram. The recorded interferogram intensity distribution is Fourier transformed and one order of the frequency spectrum (usually the +1 or -1 order) is isolated and shifted to zero frequency. After an inverse Fourier transform, the result is the phase. The choice of which order to isolate depends on a priori knowledge of the direction in which the order of the wedge fringes increases (in other words, which is the thick end of the wedge). If the wrong order is chosen, the phase map will have the wrong sign (for example, a convex wavefront will be interpreted as a concave wavefront.

This can be illustrated mathematically, by rewriting the interference equation as;

$$I = I_0(x,y) + c(x,y)e^{i2\pi f_0 x} + c^*(x,y)e^{-2\pi f_0 x},$$

where

$$c(x,y) = I_0(x,y)\gamma(x,y)e^{i\phi(x,y)},$$

and the * indicates a complex conjugate. The term c(x,y) contains the phase information we wish to extract. After performing a one dimensional Fourier transform, we have

$$I(\xi,y) = I_0(\xi,y) + \bar{c}(\xi - f_0, y) + \overline{c^*}(\xi + f_0, y),$$

where $\xi$ is the spatial frequency in the x direction, $f_0$ is the spatial frequency of the tilt fringes and the bars indicate Fourier transforms. The next step is to filter out and isolate the second term, and then perform the inverse Fourier transform to yield c(x,y). The wavefront modulo $2\pi$ is then given by

$$\phi(x,y) = \tan^{-1}\left(\frac{\mathrm{Im}[c(x,y)]}{\mathrm{Re}[c(x,y)]}\right).$$

#### 1.2.6.3.2 Spatial Synchronous Detection.

Another spatial technique is known as spatial synchronous detection (Womak, 1984) (Dorband et al, 1990) (Kuchel, 1990) (Freischlad et al, 1990b) . In this technique, the interferogram is multiplied by a reference pattern in the spatial domain. Again we rewrite the interference equation including a spatial carrier frequency;

$$I(x,y) = I_0(x,y)\{1 + \gamma(x,y)\cos[\phi(x,y) + 2\pi f_0 x]\}.$$

$f_0$ is the spatial carrier frequency created by adding many tilt fringes. A cosine reference pattern can be written as;

$$R_{\cos}(x,y) = \cos(2\pi f_0 x).$$

This pattern can either be generated analytically in a computer or electronically. These two patterns are multiplied together to yield;

$$I(x,y)R_{\cos}(x,y) = I_0(x,y)\left(\cos(2\pi f_0 x) + \gamma(x,y)\{\cos[\phi(x,y) + 4\pi f_0 x] + \cos[\phi(x,y)]\}\right)$$

The last term is filtered from the rest of the terms to yield a cosine function at the fundamental frequency.

$$C(x,y) = I_0(x,y)\gamma(x,y)\cos[\phi(x,y)].$$

Similarly, a second function is created by multiplying the interferogram by a sine reference pattern

$$R_{\sin}(x,y) = \sin(2\pi f_0 x).$$

This yields;

$$S(x,y) = I_0(x,y)\gamma(x,y)\sin[\phi(x,y)].$$

The phase is then calculated from;

$$\phi(x,y) = \tan^{-1}\left(\frac{S(x,y)}{C(x,y)}\right).$$

The biggest problem associated with all spatial phase measuring techniques is the need to introduce a large amount of tilt between the reference and test waves in the interferometer. The large tilt means that the two wavefronts travel quite different paths in the interferometer which can introduce errors due to aberrations in the interferometer optics (Huang, 1993).

#### 1.2.6.4 Phase Unwrapping.

All of the phase measurement algorithms described result in a phase calculation modulo $2\pi$ due to the periodic nature of the trigonometric functions employed. The results of the computation thus contain discontinuities where the phase values jump from $2\pi$ to zero. These discontinuities are easily removed by using the knowledge that the surface being measured is continuous and the assumption that its derivative is also continuous (that is, that the surface has no steps). The results of the computation are made continuous simply by adding $2\pi$ to the phase values at every discontinuity until the resultant phase map is a continuous function.

## 1.3 Absolute Interferometric Measurements.

Using phase measuring interferometry it is possible to determine the wavefront due to the interference of light reflected from two surfaces to a precision of $\lambda/500$ or better. When the measurement is made to determine the figure of one of the surfaces, however, it must be

remembered that the wavefront is due to the sum of the figures of the reference and test surfaces (Fizeau interferometer) or their difference (Twyman-Green interferometer).
The measurement is thus a relative measurement and the uncertainty in the result cannot be less than the uncertainty in the figure of the reference surface. Commercially produced reference surfaces typically have a figure error certified to be within $\lambda/20$ to$\lambda/10$. The method used to arrive at the figure accuracy is usually not specified and a contour map of the part is very seldom supplied, the quoted accuracy usually being a peak-to-valley (P-V) value.

For many applications the accuracy of a relative measurement is sufficient but there are a growing number of cases where this is not so. As an example, optical surfaces for use at very short wavelengths (soft X-ray telescopes or ultraviolet lasers) will usually be tested at visible wavelengths because of the availability of convenient visible laser light sources and detector arrays. While a surface accuracy of $\lambda/10$ at the test wavelength may be quite sufficient for an equivalent component operating at visible wavelengths, this will translate to an accuracy of several wavelengths at the design wavelength. For critical applications, this uncertainty in the surface figure will be quite unacceptable.

Clearly then there is a need to be able to make measurements that exceed the accuracy of the reference surface (Elssner et al, 1992). Such measurements, that are independent of the errors in the reference surface, are known as absolute measurements (as distinct from relative measurements). Various methods have been developed for the absolute testing of optical surfaces. The absolute measurement techniques are generally dependent on the nominal surface figure of the component under test (flat, spherical etc.).

### 1.3.1 Flat Surfaces.

Techniques for the absolute interferometric testing of flat surfaces can conceptually be divided into two groups. The first involves the use of a reference surface whose shape can be confidently predicted because it is defined by physical laws and not by manufacturing effort. Such a surface is the undisturbed level surface of a liquid which naturally forms a near perfect

flat (actually a sphere of a radius equal to that of the earth). The second group of techniques use analytical methods to deduce the figures of the surfaces in the interferometer. These techniques generally involve manipulation of the data acquired from several interferograms comparing pairs of flats in different orientations.

#### 1.3.1.1 Liquid Surface Interferometry.

The surface of a liquid will, in the absence of external forces other than gravity, tend to conform to the shape of a gravitational equi-potential of the gravitational field to which it is subject. In the case of a liquid at the Earth's surface this corresponds, for practical purposes, to a sphere with a radius of 6371 km. The existence of such an almost perfect, naturally occurring, flat surface has led to its use, by a number of workers, as the reference surface in the interferometric examination of flats. Historically, the first use of a liquid flat for forming interference was by Lord Rayleigh (1893), though his use of water was far from ideal as will be seen. For some authors ( Dukhopel,1971, Gorshkov, 1986, Ketelsen and Anderson, 1988) the principal attraction of liquid flats is the ease with which they may be made with large apertures for testing large diameter flats. Chen et al (1993) have developed a large aperture phase shifting interferometer for flatness measurement which includes an optional liquid flat for calibration purposes though they provide little detail of this aspect of their instrument.

The National Physical Laboratory (NPL) (Teddington, Middlesex, U.K.) have a liquid surface interferometer which is used for calibration of their 300mm diameter master flats (Dew, 1966). A simplified view of their interferometer is shown in figure 1.22 as described by Debenham and Dew (1980) .

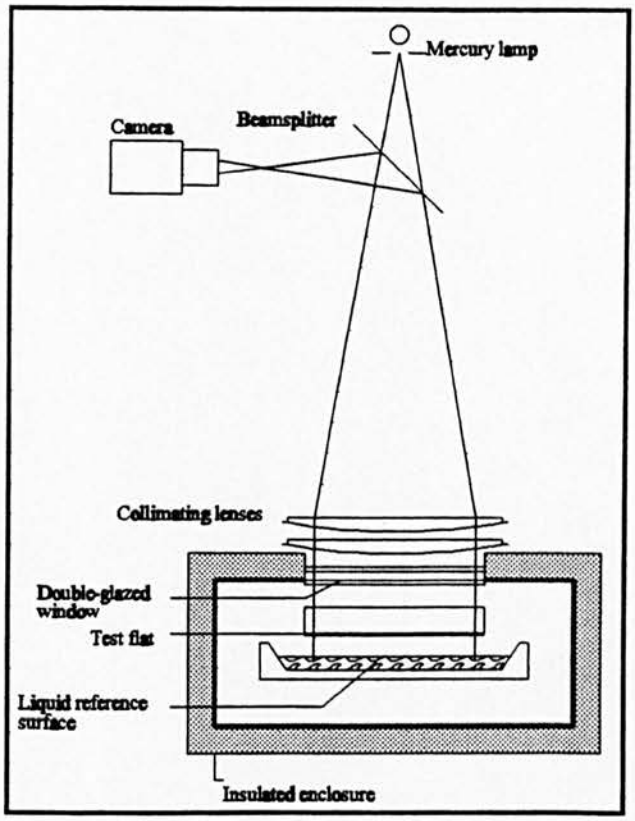


Figure 1.22 NPL's liquid surface Fizeau interferometer.

The interferometer is of the Fizeau type with a low pressure mercury lamp source filtered to isolate the green line ($\lambda$=546.07 nm). The fringe patterns are recorded on photographic plates for analysis. The liquid surface is approximately 3mm thick to damp vibrations and has a larger diameter than the test flat so that edge effects due to surface tension are negligible over the aperture of the interferometer. The liquid and test flats are enclosed within a heavily insulated box with a double- glazed window for optical access. This is necessary since the liquids used have low thermal conductivities and high thermal coefficients of expansion. To further reduce thermal effects, the laboratory in which the interferometer is situated is controlled to ±0.05°C. To minimise vibrational disturbances, the floor of the laboratory is isolated from the main building, being supported on rubber blocks and a deep bed of sand. Additionally, to reduce environmental noise, measurements are performed during "silent hours" when there is little activity in the vicinity of the laboratory. Recently the mercury source has been replaced by a He-Ne laser ($\lambda$= 632.8 nm) and the photographic camera by a CCD camera to allow automatic

wavefront analysis by a Fourier transform or spatial synchronous technique.

The choice of liquid is very important for obtaining accurate results from liquid surface interferometry (Dew, 1966). The liquid must acquire and maintain a stable flat surface and be immune to as many environmental factors as possible. Factors that may influence the figure of the liquid surface include vibration, absorption of atmospheric water, evaporation, electrostatic charges and thermal effects. To counter these effects, the ideal liquid would have high viscosity, be hydrophobic, have low vapour pressure and resistivity and have high thermal conductivity and low thermal expansion. Mercury has many of these features (Bunnagel, 1966) but a high reflectivity resulting in a low fringe contrast for uncoated test surfaces. Being opaque it is also unsuitable for configurations where it is necessary for the test surface to be below the liquid surface. The liquid of choice is usually a silicone vacuum pump oil. This has the desirable properties of high viscosity, low vapour pressure and hydrophobia. Thermal effects are minimised by stringent control of temperature and electrostatic charges are allowed to leak away by increasing the humidity of the atmosphere in the vicinity of the surface. The high viscosity of the oil necessitates a relaxation time of at least 24 hours for the surface figure to settle after the liquid has been poured, making the measurement time rather long.

One disadvantage of liquid surface interferometry is that the test surfaces may only be measured in a horizontal configuration. This can be a problem when the eventual orientation of the surface in use is not horizontal since sag due to gravity can alter the shape of the surface. This problem is especially severe for large flats where the diameter to thickness ratio is large. Debenham and Dew (1980) have addressed this problem by measuring the flats with the test surface both facing downwards and upwards. The true surface figure, free from gravitational sag, is then the mean of the two measurements . The upward facing measurement is made by submerging the flat in the oil. The calculation of the mean surface figure must be weighted to take account of the bouyancy of the flat when submerged in the oil. The reference wavefront is reflected from the surface of the oil and the test wavefront from the oil/glass boundary. One problem with this technique is that the fringe contrast is low due to the reflectivity from the oil/air boundary being much higher than the reflectivity from the oil/glass boundary.

### 1.3.1.2 Three-Flat Test.

In the absence of a natural perfect flat to act as the reference surface in a flatness test it is necessary to deduce the figure of the flats in the test by analytical means. The simplest of the analytical tests, and that upon which most other flatness tests are based, is the so called "three-flat" test (Emerson, 1952, Primak, 1967, Polster, 1968, Schulz and Schwider, 1976, Schwider, 1991). As the name implies, the three-flat test uses three nominally flat surfaces. The three surfaces, which shall be labelled A, B and C, are tested in pairs, AB, CB and AC, in a Fizeau interferometer to yield the sums of the figures of each pair of flats. The three positional combinations are shown in figure 1.23.

Analysis of the three interferograms yields the functions:

$$g_{AB}(x,y) = f_A(x,y) + f_B(-x,y),$$
$$g_{CB}(x,y) = f_C(x,y) + f_B(-x,y),$$
$$g_{AC}(x,y) = f_A(x,y) + f_C(-x,y).$$

Note the reversal of the coordinate system of the second flat in each combination. Because, in a Fizeau interferometer, the two surfaces face each other, the second surface must be flipped over (in this case, about the y-axis) and the direction of the x-axis is reversed.

This system of three equations has four unknowns: $f_A(x,y)$, $f_B(-x,y)$, $f_C(x,y)$ and $f_C(-x,y)$ and so is soluble only when x= -x= 0 and reduces to three unknowns.

When x= 0, the solutions to the equations are:

$$f_A(0,y) = \frac{g_{AB} + g_{AC} - g_{CB}}{2},$$
$$f_B(0,y) = \frac{g_{AB} + g_{CB} - g_{AC}}{2},$$
$$f_C(0,y) = \frac{g_{CB} + g_{AC} - g_{AB}}{2}.$$

It is thus only possible to deduce the absolute profile of the surfaces under test along a single

diameter using three positional combinations of three flats. Determination of the profiles of other diameters is only possible by using additional positional combinations.

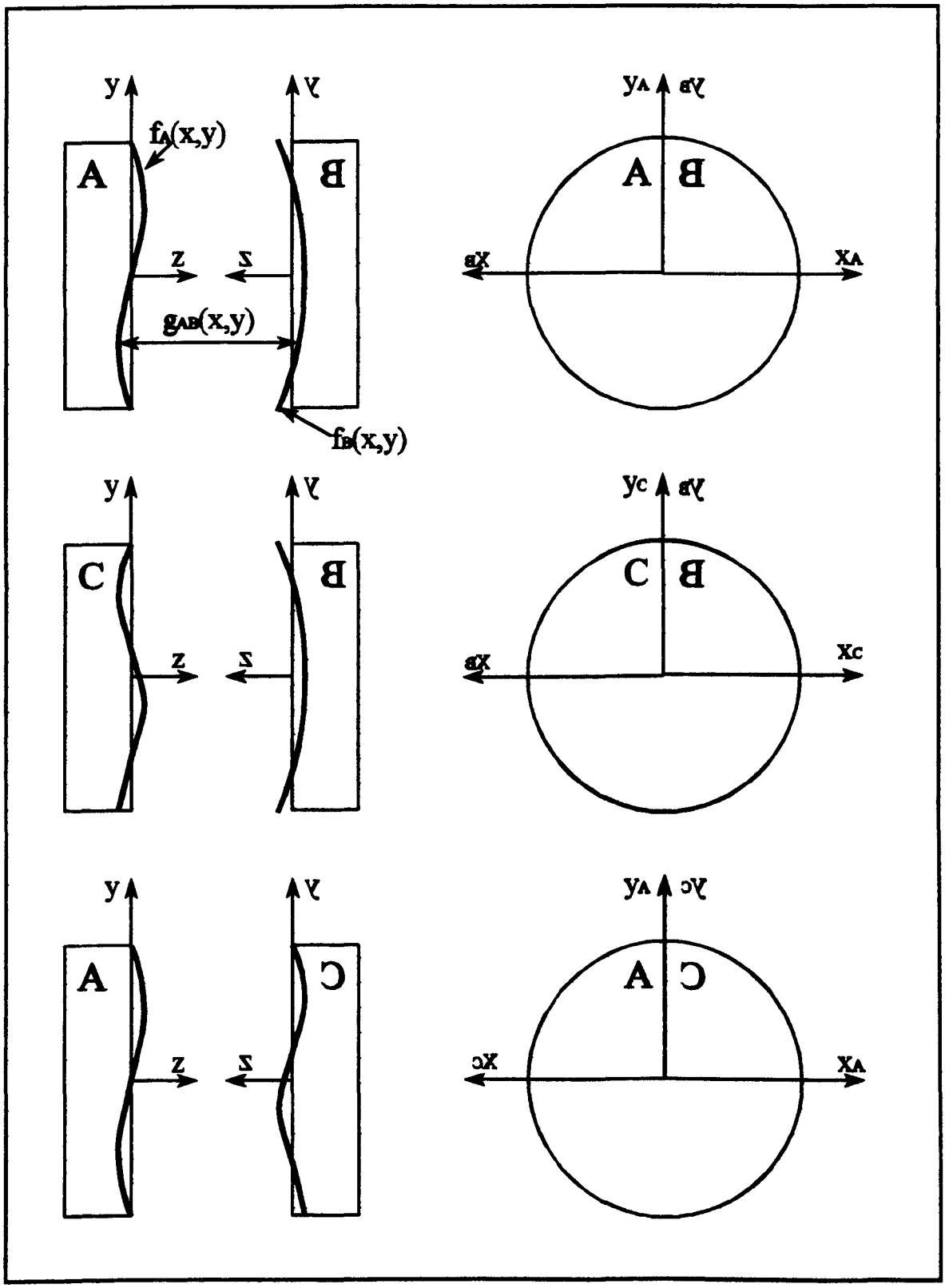


Figure 1.23 Positional combinations for the three-flat test.

It is important to note that the success of this method is due to the interferograms yielding information about the sum of the figure of the surfaces under test in a Fizeau interferometer. Comparing the same three pairs of flats in a Twyman-Green interferometer would yield the following functions:

$$g_{AB}(x,y) = f_A(x,y) - f_B(x,y),$$
$$g_{CB}(x,y) = f_C(x,y) - f_B(x,y),$$
$$g_{AC}(x,y) = f_A(x,y) - f_C(x,y).$$

Though this system of three equations has only three unknowns it has no unique solutions and is thus insoluble since the three equations are not independent.

#### 1.3.1.3 Rotational Methods.

##### 1.3.1.3.1 Schulz's Method.

Extension of the basic three flat technique to obtain the absolute profiles of additional diameters may easily be achieved by recording an additional interferogram for each extra diameter. This approach is, however, very tedious for a large number of diameters. It is, in fact, possible to derive the profiles of an arbitrarily large number of diameters from only four interferograms (Schulz, 1967, Schulz and Schwider, 1967, Schulz et al, 1971, Schulz and Schwider, 1976, Stahl, 1991). The first three interferograms are the same as the basic three flat test to determine a single diameter. In the fourth interfrogram, one of the flats is rotated about the optical axis by an angle $\phi = 2\pi M/N$ in order to determine the contours along N diameters. M and N are natural numbers, prime to each other. To facilitate the reduction of the interferogram data the surfaces of the flats are defined as shown in figure 1.24 for M=1 and N=5.

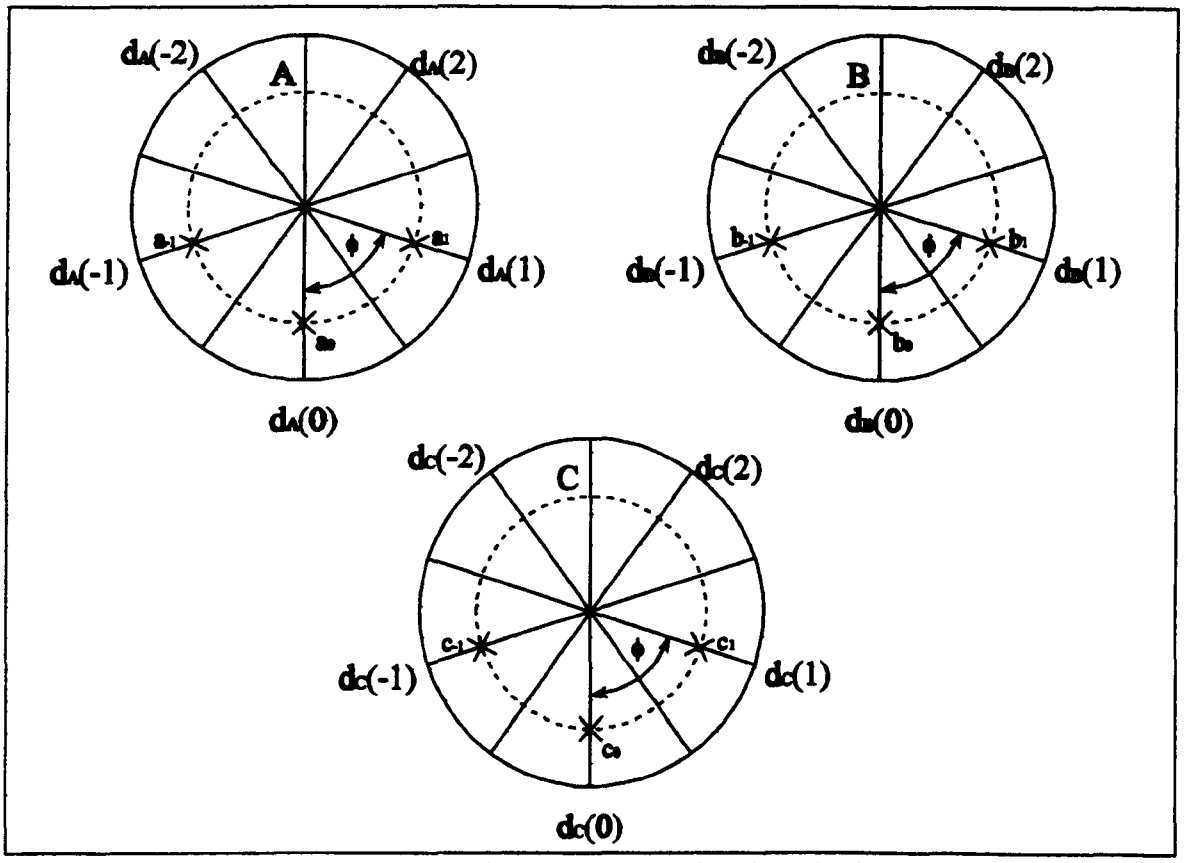


Figure 1.24 Definition of surfaces for Schulz's rotational method.

The five diameters whose profiles are to be determined are labelled $d_A(0,\pm1,\pm2)$ for flat A and similarly for flats B and C. Also points $a_0$ and $a_{\pm1}$ are defined on $d_A(0, \pm1)$ which are equidistant from the centre of the flat and similarly for flats B and C. These points define the nominal plane from which the deviations of each flat will be measured. These points are defined by;

$$f_{a_0} = f_{a_{-1}} = f_{a_1} = 0,$$
$$f_{b_0} = f_{b_{-1}} = f_{b_1} = 0,$$
$$f_{c_0} = f_{c_{-1}} = f_{c_1} = 0.$$

The seemingly arbitrary choice of three points to define the nominal plane of each flat only defines their orientation in space and has no effect on their figure.

The four positional combination of the three flats are shown in figure 1.25.

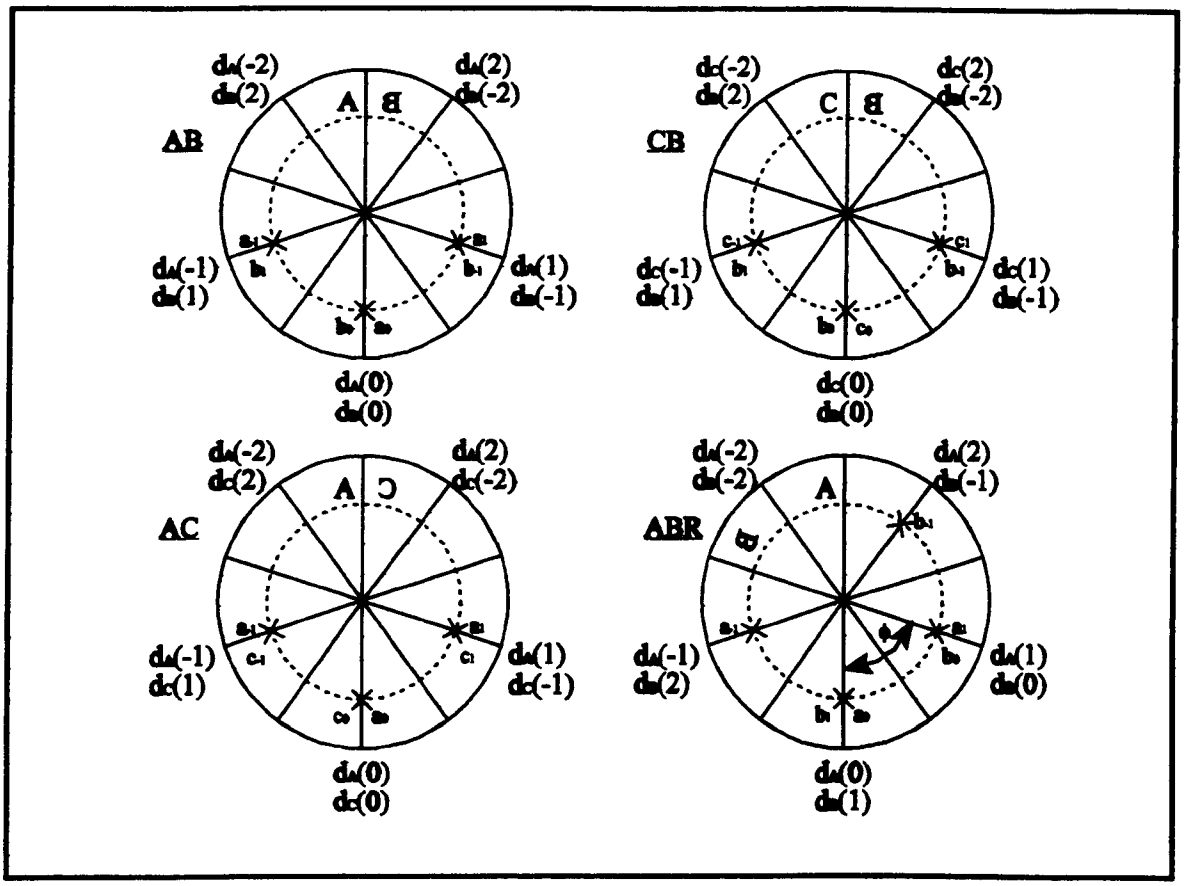


Figure 1.25 Positional combinations for Schulz's rotational method.

The measured data for the four $G_{AB}$, $G_{CB}$, $G_{AC}$ and $G_{ABR}$ will not generally be consistent with the given definition of the nominal plane of each flat. This is because the tilt between the flats during the measurement cannot be precisely known prior to the measurement. Considering the combination AB, it can be seen from figure 1.25 that the points $a_{-1}$, $a_0$, $a_1$ on flat A are coincident with the points $b_1$, $b_0$, $b_{-1}$, respectively, on flat B. Given that the value of the surface functions of the flats is zero at all the points, the value of $G_{AB}$ should be zero at these points also. In order to make the measured data consistent it is necessary to subtract a linear term corresponding to a plane, P, defined by the measured values $G_{AB}(a_{-1})$, $G_{AB}(a_0)$ and $G_{AB}(a_1)$. Note that the coordinate system for the measured data is the same as that of the upper flat in the measured combination. The corrected data are;

$$g_{AB} = G_{AB} - P(G_{AB}(a_{-1}), G_{AB}(a_0), G_{AB}(a_1))$$

similarly, for combinations CB and AC;

$$g_{CB} = G_{CB} - P(G_{CB}(c_{-1}), G_{CB}(c_0), G_{CB}(c_1))$$
$$g_{AC} = G_{AC} - P(G_{AC}(a_{-1}), G_{AC}(a_0), G_{AC}(a_1))$$

The data sets $g_{AB}$, $g_{CB}$ and $g_{AC}$ are now consistent with the definitions of the nominal planes for flats A, B and C.

The basic three flat test method may now be used to determine the absolute contours of the three flats along the diameters $d_A$, $d_B$ and $d_C$. The value of the surface figure function is now known at the central point (coordinates (0,0)) of each flat and this fact may now be used to make the measured data $G_{ABR}$ consistent with the definitions of the nominal planes.

It can be seen from figure 1.25 that, in combination ABR, the points $a_0$ and $b_1$ are coincident and the points $a_1$ and $b_0$ are coincident. In addition the central points of each flat are coincident. The measured data at the centre will in general not be equal to the sum of the surface figure functions of A and B at the centre and so $G_{ABR}$ may be corrected thus;

$$g_{ABR} = G_{ABR} - P\left(G_{ABR}(a_0), G_{ABR}(a_1), \left[G_{ABR}(0,0) - \{f_A(0,0) + f_B(0,0)\}\right]\right)$$

The known profile of the diameter, $d_A(0)$ on flat A may now be subtracted from $g_{ABR}$ to give the profile of diameter $d_B(1)$ on flat B. Similarly the profile of diameter $d_B(0)$ may be subtracted from $g_{ABR}$ to give the profile of diameter $d_A(1)$ on flat A.

$$f_B[d_B(1)] = g_{ABR}[d_A(0)] - f_A[d_A(0)],$$
$$f_A[d_A(1)] = g_{ABR}[d_A(1)] - f_B[d_B(0)]$$

The profiles of these two new diameters may now be subtracted from $g_{AB}$ to give the profiles of two further diameters;

$$f_B[d_B(-1)] = g_{AB}[d_A(1)] - f_A[d_A(1)],$$
$$f_A[d_A(-1)] = g_{AB}[d_A(-1)] - f_B[d_B(1)].$$

The profiles of these new diameters may then be subtracted from $g_{ABR}$ to give;

$$f_B[d_B(2)] = g_{ABR}[d_A(-1)] - f_A[d_A(-1)],$$
$$f_A[d_A(2)] = g_{ABR}[d_A(2)] - f_B[d_B(-1)],$$

which may be subtracted from $g_{AB}$ to give;

$$f_B[d_B(-2)] = g_{AB}[d_A(2)] - f_A[d_A(2)],$$
$$f_A[d_A(-2)] = g_{AB}[d_A(-2)] - f_B[d_B(2)].$$

The absolute profiles of five diameters on flats A and B have now been found for the case where N= 5.

Note that for a rotational angle $\phi = 2\pi \times Q$ where Q is an irrational number (one that cannot be expressed as a ratio of integers, M/N) the number of diameters that may be evaluated is theoretically infinite.

Since its original conception the rotation method of Schulz has undergone several improvements which have been made possible by the emergence, in the intervening period, of computer aided phase measuring interferometry.

Schulz and Grzanna (1992) redefined the nominal planes from which the surface figure functions would be derived. Rather than define these planes by three discrete points, they were defined as the best fit planes to the surfaces being measured in the least squares sense. The least squares best fit plane to the measured data for any combination of flats, with the best fit tilt terms removed, equals the sum of the best fit planes of the two surfaces comprising that combination. This approach reduces the possible error introduced into the analysis by a tilt correction based upon only three points since the least squares fit to the data takes account of every data point.

Schulz (1993) modified the method by introducing a fifth positional combination with an angle of rotation, $K\phi$, equal to an integral multiple of the rotation angle in the fourth, $\phi$. This has the effect of reducing the maximum number of steps required to determine the profile of each diameter, hence reducing the accumulation of measuring errors by propagation through the

analysis.

Elssner et al (1994) review the developments to date in Schulz's rotation method and describe the experimental conditions necessary to obtain the best measurement accuracy with the technique.

A problem with the rotation method is that, in general, points on diameters other than those parallel with the x and y axes will not coincide with the measured grid of points. This is because the analysis of data works on a radial grid of data and the measured data will be on a rectangular or square grid dictated by the geometry of the CCD camera used in the interferometer. Interpolation of data is therefore necessary in order to convert the rectilinear array of measured data into a radial array for processing. In order for the results to be in a useable form after the analysis, it is likely that they will have to be converted back into a rectangular array by further interpolation. This data interpolation is a likely source of error in the analysis.

A further refinement of Schulz's method to avoid the problem of interpolating data has been described by Grzanna (Grzanna and Schulz, 1990, Grzanna, 1994). In this method, the rotation angle is 90° . The x and y axes of the flats are thus rotated onto each other and interpolation of data is not necessary. A further positional combination is used with one flat translated along the x or y axis with respect to the other in order to facilitate the determination of the absolute deviations from flatness on a square grid of points. The five positional combinations used in this method are similar to those used in the method independently developed by this author and to be described in chapter two, though the details of the data analysis are quite different.

#### 1.3.1.3.2 Ai's Method.

A function in a cartesian coordinates system can be expressed as the sum of four functions with different classes of symmetry. The four classes are; even symmetry in both x and y, even in x and odd in y, odd in x and even in y, odd in both x and y.

$$F(x,y) = F_{ee} + F_{eo} + F_{oe} + F_{oo},$$

where;

$$F_{ee} = \frac{[F(x,y) + F(-x,y) + F(x,-y) + F(-x,-y)]}{4},$$

$$F_{eo} = \frac{[F(x,y) + F(-x,y) - F(x,-y) - F(-x,-y)]}{4},$$

$$F_{oe} = \frac{[F(x,y) - F(-x,y) + F(x,-y) - F(-x,-y)]}{4},$$

$$F_{oo} = \frac{[F(x,y) - F(-x,y) - F(x,-y) + F(-x,-y)]}{4}.$$

Ai and Wyant (1992a) describe a method for the absolute testing of flats where the figures of each flat are decomposed into these four classes of symmetry which are determined separately.

For convenience, two operators, $[\ ]^x$ and $[\ ]^\theta$, are defined, describing a flipping of the function in x and a rotation of the function by θ respectively.

$$[F(x,y)]^x = F(-x,y),$$

$$[F(x,y)]^\theta = F(x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta).$$

The measured data for the test are obtained from eight positional combinations of three flats where, in some combinations, one flat is rotated by 180°, 90° or 45°. The measured functions are;

$$g_1 = f_A + f_B^x, \qquad g_5 = f_A^{180} + f_C^x,$$

$$g_2 = f_A^{180} + f_B^x, \qquad g_6 = f_B + f_C^x,$$

$$g_3 = f_A^{90} + f_B^x, \qquad g_7 = f_B^{90} + f_C^x,$$

$$g_4 = f_A^{45} + f_B^x, \qquad g_8 = f_B^{45} + f_C^x.$$

$g_1$, $g_2$, and $g_5$ may then be written;

$$g_1 = f_{[A]ee} + f_{[A]oo} + f_{[A]oe} + f_{[A]eo} + f_{[B]ee} - f_{[B]oo} - f_{[B]oe} + f_{[B]oo},$$

$$g_2 = f_{[A]ee} + f_{[A]oo} - f_{[A]oe} - f_{[A]eo} + f_{[B]ee} - f_{[B]oo} - f_{[B]oe} + f_{[B]oo},$$

$$g_5 = f_{[A]ee} + f_{[A]oo} - f_{[A]oe} - f_{[A]eo} + f_{[C]ee} - f_{[C]oo} - f_{[C]oe} + f_{[C]oo}.$$

The odd-even and even odd parts of the surface functions of the flats may thus be obtained from;

$$f_{[A]oe} + f_{[A]eo} = \frac{g_1 - g_2}{2},$$

$$f_{[B]oe} + f_{[B]eo} = \left[\frac{g_1 - [g_1]^{180}}{2} - f_{[A]oe} + f_{[A]eo}\right]^x,$$

$$f_{[C]oe} + f_{[C]eo} = \left[\frac{g_5 - [g_5]^{180}}{2} - [f_{[A]oe} + f_{[A]eo}]^{180}\right]^x.$$

The even-even parts of the surface functions may also be found from;

$$f_{[A]ee} = \frac{\overline{g_1} + \overline{g_5} - \overline{g_6} + [\overline{g_1} + \overline{g_5} - \overline{g_6}]^x}{4},$$

$$f_{[B]ee} = \frac{\overline{g_1} + [\overline{g_1}]^x - 2f_{[A]ee}}{2},$$

$$f_{[C]ee} = \frac{\overline{g_5} + [\overline{g_5}]^x - 2f_{[A]ee}}{2},$$

where;

$$\overline{g_1} = \frac{g_1 + [g_1]^{180}}{2},$$

$$\overline{g_5} = \frac{g_5 + [g_5]^{180}}{2},$$

$$\overline{g_6} = \frac{g_6 + [g_6]^{180}}{2}.$$

All of the even-even, odd-even and even odd terms are now solved, leaving only the odd-odd terms which are not soluble from the data sets employed so far. An odd-odd function in a cartesian coordinate system is an odd function in a polar coordinate system with the same origin and has a period of 180°. Thus an odd-odd function, $F_{oo}(x,y)$ may be expressed as a Fourier sine series;

$$F_{oo}(x,y) = \sum_{m=1} h_{2m} \sin(2m\theta),$$

and

$$[F_{oo}(x,y)]^{90} = -\sum_{m=odd} h_{2m} \sin(2m\theta) + \sum_{m=even} h_{2m} \sin(2m\theta),$$

where $x^2+y^2=$ constant and $h_{2m}$ are the Fourier coefficients. To emphasise the periodicity of $F_{oo}$ a subscript $2\theta$ is added.

$$F_{oo,2\theta} = F_{oo,2odd\theta} + F_{oo,2even\theta},$$

$$[F_{oo,2\theta}]^{90} = -F_{oo,2odd\theta} + F_{oo,2even\theta},$$

where

$$F_{oo,2even\theta} = \sum_{m=even} h_{2m} \sin(2m\theta) = \sum_{m=1} h_{4m} \sin(4m\theta) = F_{oo,4\theta},$$

$$F_{oo,2odd\theta} = \sum_{m=odd} h_{2m} \sin(2m\theta),$$

$F_{oo,2even\theta}$ has a fundamental frequency of 4 and can be further divided into two groups;

$$F_{oo,4\theta} = F_{oo,4odd\theta} + F_{oo,4even\theta},$$

$$[F_{oo,4\theta}]^{45} = -F_{oo,4odd\theta} + F_{oo,4even\theta}.$$

An odd-odd function can thus be expressed as the sum of mod(mθ) terms where m= 2, 4, 8, 16,...;

$$F_{00,2\theta} = F_{00,2odd\theta} + F_{00,4odd\theta} + F_{00,8odd\theta} + F_{00,16odd\theta} + \ldots$$

Note that each term represents a broad spectrum of Fourier sine terms. For example, $F_{00,2odd\theta}$ includes components of sin(2θ), sin(6θ), sin(10θ), sin(14θ)... and $F_{00,4odd\theta}$ includes components of sin(4θ), sin(12θ), sin(16θ), sin(20θ)... A smooth surface may then be well represented by the first two terms.

The odd-odd components of the surface figure function may now be determined;

$$f_{[A]oo,2odd\theta} = \frac{g_1' - g_3'}{2},$$

$$f_{[B]oo,2odd\theta} = \frac{g_6' - g_7'}{2},$$

$$f_{[C]oo,2odd\theta} = \frac{[g_7']^{-90} - g_6'}{2},$$

where $g_1'$, $g_3'$, $g_6'$ and $g_7'$ are $g_1$, $g_3$, $g_6$ and $g_7$ with the even-even, odd-even and even-odd components subtracted;

$$g_1' = f_{[A]oo,2\theta} - f_{[B]oo,2\theta},$$

$$g_3' = [f_{[A]oo,2\theta}]^{90} - f_{[B]oo,2\theta},$$

$$g_6' = f_{[B]oo,2\theta} - f_{[C]oo,2\theta},$$

$$g_7' = [f_{[B]oo,2\theta}]^{90} - f_{[C]oo,2\theta}.$$

The 4odd$\theta$ terms are;

$$f_{[A]oo,4odd} = \frac{g_1'' - g_4''}{2},$$

$$f_{[B]oo,4odd} = \frac{g_6'' - g_8''}{2},$$

$$f_{[C]oo,4odd} = \frac{[g_8'']^{-45} - g_6''}{2},$$

The terms, g", are the measured data with the previously determined terms removed;

$$g_1'' = f_{[A]oo,4\theta} - f_{[B]oo,4\theta}$$

$$g_4'' = [f_{[A]oo,4\theta}]^{45} - f_{[B]oo,4\theta}$$

$$g_6'' = f_{[B]oo,4\theta} - f_{[C]oo,4\theta}$$

$$g_8'' = [f_{[A]oo,4\theta}]^{45} - f_{[C]oo,4\theta}$$

The figures of the three surfaces may now be written;

$$f_A = f_{[A]ee} + f_{[A]oe} + f_{[A]eo} + f_{[A]oo,2odd\theta} + f_{[A]oo,4odd\theta}$$
$$+\text{higher order terms,}$$
$$f_B = f_{[B]ee} + f_{[B]oe} + f_{[B]eo} + f_{[B]oo,2odd\theta} + f_{[B]oo,4odd\theta}$$
$$+\text{higher order terms,}$$
$$f_C = f_{[C]ee} + f_{[C]oe} + f_{[C]eo} + f_{[C]oo,2odd\theta} + f_{[C]oo,4odd\theta}$$
$$+\text{higher order terms,}$$

The higher order terms may be found from further positional combinations with different rotational angles. For instance the 8odd$\theta$ terms may be found from combinations with a 22.5° rotation.

Realization that some of the eight positional combinations required for the above method are not linearly independent has resulted in refinements of the method and reduction of the number of combinations required. Ai and Wyant (1992b) reduce the number of combinations to six and Shao et al (1992) reduce the number to the minimum of four. The four combinations described by Shao et al are the same as for Schulz's method with a rotation angle of 45°. This allows determination of the odd-odd functions up to the 4odd$\theta$ terms. Using, instead, an angle of 22.5° or further subdivision would permit determination up to the 8odd$\theta$ terms or higher.

#### 1.3.1.3.3 Rotational Methods Using Polynomials (Fritz's Method).

The orthogonal and rotationally invariant properties of the Zernike polynomials (see section 1.2.5.1) make them an attractive proposition for the analysis of wavefront data where rotations of wavefronts occur. Fritz (1983, 1984) has described a method using these properties for the absolute measurement of flatness.

Four positional combinations of three flats are used, with a rotation of one flat by an angle, $\phi$, about the optical axis, in the same way as for the rotational method according to Schulz (figure 1.25). The figures of the three flats are each described by the sum of a series of Zernike polynomials;

$$f_A = \sum_1^{\infty}\left[A_{[A]n}Q_n^0(\rho) + \sum_{m=1}^{n} Q_n^m(\rho)\rho^m(B_{[A]nm}\cos m\theta + C_{[A]nm}\sin m\theta)\right],$$

$$f_B = \sum_1^{\infty}\left[A_{[B]n}Q_n^0(\rho) + \sum_{m=1}^{n} Q_n^m(\rho)\rho^m(B_{[B]nm}\cos m\theta + C_{[B]nm}\sin m\theta)\right],$$

$$f_C = \sum_1^{\infty}\left[A_{[C]n}Q_n^0(\rho) + \sum_{m=1}^{n} Q_n^m(\rho)\rho^m(B_{[C]nm}\cos m\theta + C_{[C]nm}\sin m\theta)\right].$$

Because of the rotationally invariant form of the Zernike polynomials, the figure of flat B rotated through angle, $\phi$, can be written as;

$$f_B = \sum_1^{\infty}\left[A_{[B]n}Q_n^0(\rho) + \sum_{m=1}^{n} Q_n^m(\rho)\rho^m(B_{[B]nm}\cos m(\theta-\phi) + C_{[B]nm}\sin m(\theta-\phi)\right],$$

$$= \sum_1^{\infty}\left[A_{[B]n}Q_n^0(\rho) + \sum_{m=1}^{n} Q_n^m(\rho)\rho^m(\overline{B_{[B]nm}}\cos m\theta + \overline{C_{[B]nm}}\sin m\theta\right]$$

where;

$$\overline{B_{[B]nm}} = B_{[B]nm}\cos m\phi - C_{[B]nm}\sin m\phi$$

$$\overline{C_{[B]nm}} = B_{[B]nm}\sin m\phi + C_{[B]nm}\cos m\phi$$

These four equations can be combined to give the Zernike polynomials for the four test combinations. Equating these with the polynomial terms fitted to the measured data for the four combinations allows the polynomial coefficients for the three surfaces to be determined. The absolute contours of the three flats are thus determined up to the order of the fit.

Since the number of polynomial terms fitted to the data must be finite, the spatial resolution of the analysis is limited by the highest order terms. For example a 37 term Zernike polynomial fit (a common maximum for commercial software packages for interferogram analysis) has a maximum angular term of sin5θ and will not describe any surfaces features with angular spatial frequency components of greater than 5 cycles/rev..

However for sufficiently smooth surfaces (and surfaces sufficiently accurate to warrant absolute testing are likely to be very smooth) the method can yield very accurate results. Fritz (1984) quotes an accuracy of λ/100 when using this method for the certification of 250 mm

(10 inch) diameter flats.

### 1.3.2 Spherical Surfaces.

The absolute testing of spherical surfaces is much more straightforward than the absolute testing of flats. The same techniques as described for flatness testing may be adapted for spherical testing. The fact that the centre of curvature of spheres is accessible, however, provides us with the means of deriving the absolute contours of spheres over their whole surfaces from as few as three separate measurements (Truax, 1988, Creath and Wyant, 1990). A report by this author entitled "Absolute interferometric testing of microspheres." was commissioned by the National Physical Laboratory (NPL). The report describes in detail the various different possibilities for the absolute testing of very small radius spheres, though the techniques described are applicable to general spherical surfaces. A copy of the report is included in this thesis as Appendix A.

### 1.3.3 Other Surfaces.

In section 1.2.3 configurations for testing conicoid surfaces were described where the surfaces were compared with spherical and flat reference surfaces. Clearly, if the absolute contours of the spherical and flat surfaces were known by the absolute measurement techniques described above or to be described then the tests of the conicoid surfaces can, themselves, be made absolute.

An interesting possibility for the absolute measurement of conicoids is presented by the properties of a rotating liquid surface. In a rotating fluid adding the vectors of the centripetal and gravitational accelerations gives a surface that has the shape of a paraboloid. Using a reflecing liquid one therefore gets a reflecting paraboloid that could be used as a reference surface in an interferometric test. The focal length of the mirror, L is related to the acceleration due to gravity, g, and the angular velocity of the liquid, $\omega$ by;

$$L = \frac{g}{2\omega^2}.$$

To date, most interest in rotating liquid mirrors lies in their potential as very large telescope primary mirrors for astronomy or LIDAR. A review paper by E. F. Borra, "Liquid mirrors: a review." is available from the Word Wide Web at; http://astrosun.phy.ulaval.ca/lmt/lmt.home.html. One group of workers are investigating the possibility of using rotating liquid mirrors as reference surfaces at Centre Spatial de Liege (CSL). Though there are, as yet, no publications in the literature on this aspect of liquid mirror research, the group at CSL have a Web page under construction at; http://astrosun.phy.ulaval.ca/lmt/lmt-csl.html.

Note: Since first submission of this thesis, the following paper has appeared in the literature:

Ninane, N.M., Jamar, C. A., "Parabolic liquid mirrors in optical shop testing." *Applied Optics* **35** (31) 6131-6139 (1996).

## Chapter 2: Description Of New Method For Flatness Testing.

The method for absolute flatness testing to be described is based upon the basic three-flat test described in chapter 1 (section 1.3.1.2). The test is extended in order to overcome the limitation that it only yields the surface contour along a single diameter for each flat.

The extension involves taking one of the combinations involved in the three-flat test and modifying it by rotating one flat through 90°. This rotation of 90° is a special case to which the objections raised in the section on rotational methods (section 1.3.1.3) do not apply. A square (cartesian) coordinate system rotated through 90° is equivalent to the same coordinate system but with the axes interchanged. There is thus no requirement for interpolation in order to operate on pairs of points in rotated and non-rotated coordinate systems.

An additional modification of the basic combination is made by translating one of the flats laterally along one axis. Again, this causes no problem with interpolation of data if the translation is by an integral number of units along the axis where 1 unit is the spacing between adjacent measured points (commonly the spacing between CCD camera pixels).

### 2.1 Definition Of Surfaces.

The flatness measuring method requires measurements to be made on a number of different combinations of pairs from three flats. In order that data from these measurements may be meaningfully processed it is necessary that the coordinate systems for each flat be defined. For three flats denoted by the labels A, B, and C the coordinate systems are defined as shown in figure 2.1.

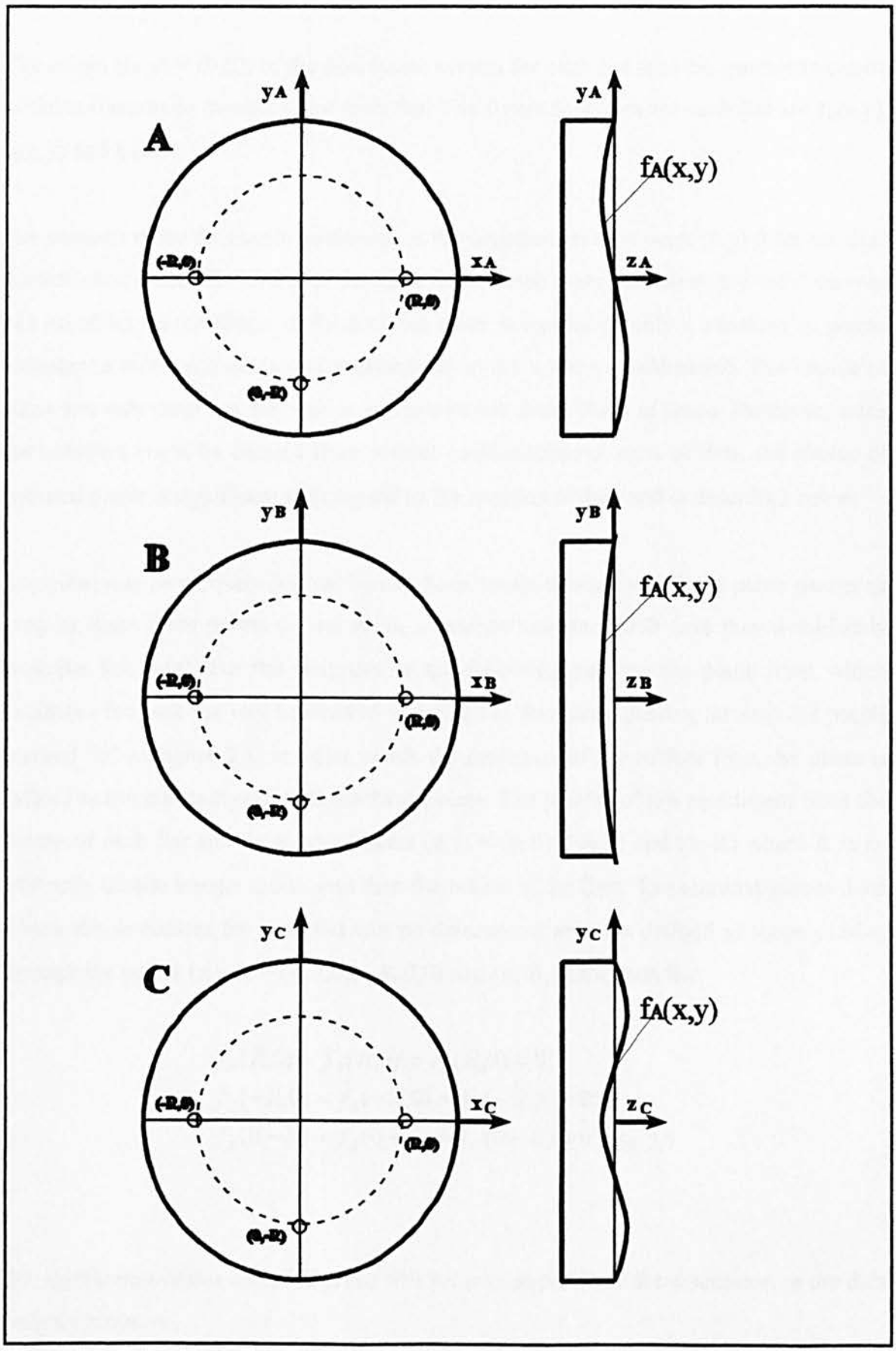


Figure 2.1. Definition of surfaces used in the flatness test.

The origin ((x,y) = (0,0)) of the coordinate system for each flat is at the geometric centre of the surface to be measured for each flat. The figure functions for each flat are $f_A(x,y)$, $f_B(x,y)$ and $f_C(x,y)$.

The contours of the flats are to be derived as the deviations at each point (x,y) from an ideal mathematical plane. The choice of the plane from which these deviations are to be derived has no effect on the shape of the surfaces since it represents only a variation in piston (constant z term) and tilt (terms proportional to the x and y coordinates). The choice of plane thus only describes the position and orientation of the plane in space. However, since the contours are to be derived from several combinations of pairs of flats, the choice of nominal plane is significant with regard to the analysis of data and is described below.

Any plane may be uniquely defined by any three points through which the plane passes so long as these three points do not lie in a straight line (in which case they would only describe that line). For the purposes of the following analysis the plane from which deviations for each flat will be derived is defined as that plane passing through the points marked "o" in figure 2.1. In other words the deviation of the surface from the plane is defined as being zero at each of these three points. The points "o" are equidistant from the centre of each flat and lie at coordinates (x,y) = (R,0), (-R,0) and (0,-R) where R is an arbitrarily chosen integer radius less than the radius of the flats. The nominal planes from which the deviations for each flat will be determined are thus defined as those passing through the points (x,y,z) = (R,0,0), (-R,0,0) and (0,-R,0) for each flat;

$$\begin{aligned} f_A(R,0) &= f_B(R,0) = f_C(R,0) = 0, \\ f_A(-R,0) &= f_B(-R,0) = f_C(-R,0) = 0, \\ f_A(0,-R) &= f_B(0,-R) = f_C(0,-R) = 0. \end{aligned} \quad \text{Eq.2.1}$$

The significance of this choice of plane will become apparent as the discussion of the data analysis proceeds.

## 2.2 Raw Data To Be Analysed.

The positional combinations in which the three flats are to be measured are shown in figure 2.2. The combinations are labelled AB, CB, AC, ABR and ABS. In each case, the first letter of the label refers to the "upper" flat or flat nearest to the collimating lens of the Fizeau interferometer. It can be seen that the coordinate systems for the upper flats are as shown in figure 2.1. In the first three combinations AB, CB, and AC the direction of the x-axis has been reversed for the "lower" flat by virtue of it having been inverted. In the fourth combination, ABR, the lower flat, B, has been rotated by 90° anti-clockwise about the origin as seen from flat A. Hence the coordinate system for flat B has also been rotated. In the fifth combination the lower flat, B, has had its origin translated by a distance, s, along the x-axis in the positive direction with respect to flat A. The y-axes of A and B have thus been displaced by distance, s.

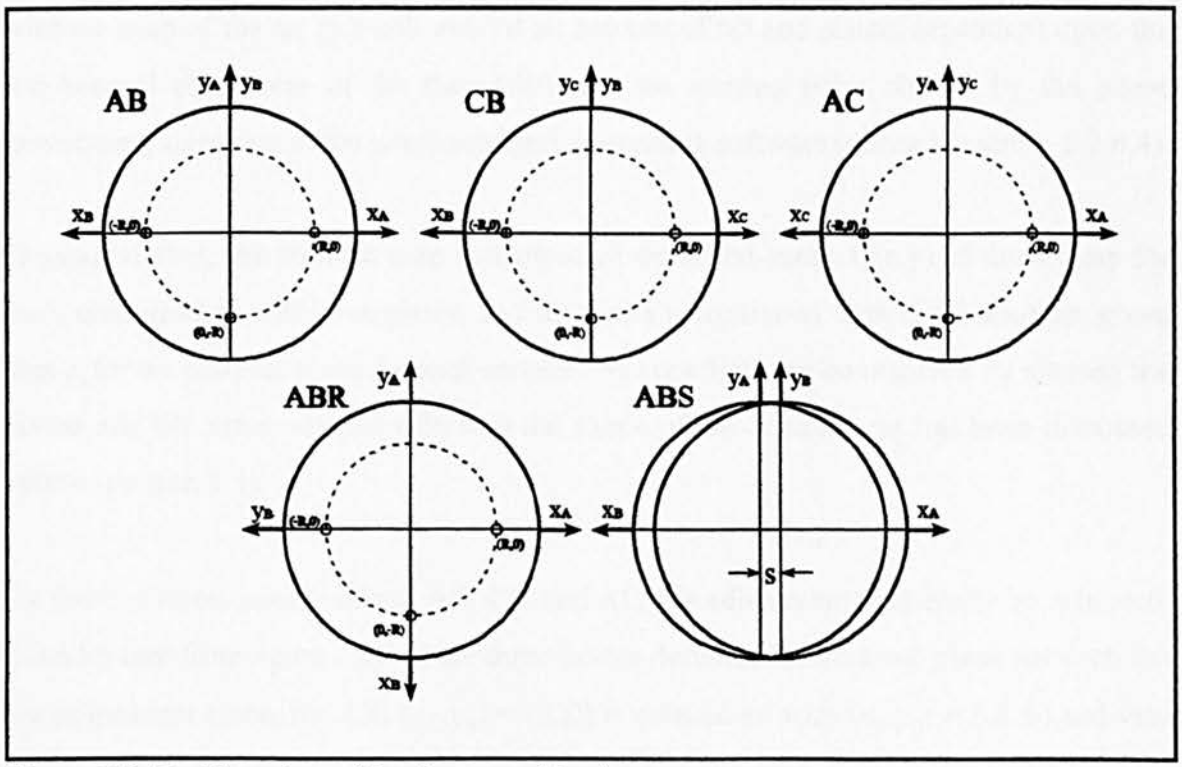


Figure 2.2. Positional combinations used in the test.

In each of the combinations described above, the interferometer measures the sum of the deviations from the nominal planes for the two flats in the combination, plus an unknown

piston and tilt associated with the mechanical positioning of the flats during the measurement. It is these arrays of data, from the interferometric measurements of each pair of flats, that form the raw data from which the contours of each flat will be derived. These arrays of data shall henceforth be known as the "uncorrected deviation-sums"; $G_{AB}(x,y)$, $G_{CB}(x,y)$, $G_{AC}(x,y)$, $G_{ABR}(x,y)$ and $G_{ABS}(x,y)$.

## 2.3 Analysis Of Data.

### 2.3.1 Normalisation Of Initial Data.

When setting up a pair of flats to form a cavity in a Fizeau interferometer, it is generally not possible to know precisely the amount of piston and tilt between the flats prior to the measurement. Therefore, when the analysis of the fringe patterns is performed, the resulting contour map of the air gap will exhibit an amount of tilt and piston dependent upon the mechanical placement of the flats (tilt) and the starting point chosen by the phase unwrapping algorithm in the interferometer's processing software (piston) (section 1.2.6.4).

In general then, the contour map (uncorrected deviation-sum), G(x,y) of the air gap for each combination will have piston and tilt terms inconsistent with the definition, given above, for the nominal plane for each surface. This conflict may be resolved by altering the piston and tilt terms without affecting the shape of the contours, as has been discussed before (section 2.1).

For the first three combinations, AB, CB and AC this adjustment may easily be achieved. It can be seen from figure 2.2 that the three points defining the nominal plane for each flat are coincident since, for AB, $(x_A,y_A) = (R,0)$ is coincident with $(x_B,y_B) = (-R,0)$ and vice versa, and points $(x_A,y_A) = (x_B,y_B) = (0,-R)$ are coincident. This is similarly true for combinations CB and AC. Since the deviation from the nominal plane at each of the three points "o" for each flat is defined as being zero, it follows that the deviation-sum at these three points must also be equal to zero. To correct the piston and tilt terms it is thus merely necessary to calculate the planes, P, which pass through the measured, uncorrected,

deviation-sums at the three points "o" for each combination and then subtract that plane from the uncorrected deviation-sum for each combination, AB, CB and AC. The results are the corrected deviation-sum for the pairs of flats, which will now simply be called the "deviation-sums"; $g_{AB}(x,y)$, $g_{CB}(x,y)$ and $g_{CB}(x,y)$ for AB, CB and AC respectively;

$$g_{AB}(x,y) = G_{AB}(x,y) - P_{[G_{AB}(R,0),G_{AB}(-R,0),G_{AB}(0,-R),]}(x,y),$$
$$g_{CB}(x,y) = G_{CB}(x,y) - P_{[G_{CB}(R,0),G_{CB}(-R,0),G_{CB}(0,-R),]}(x,y),$$
$$g_{AC}(x,y) = G_{AC}(x,y) - P_{[G_{AC}(R,0),G_{AC}(-R,0),G_{AC}(0,-R),]}(x,y). \quad \text{Eq.2.2}$$

The equation of a plane, z=P, passing through points $(x_1,y_1,z_1)$, $(x_2,y_2,z_2)$ and $(x_3,y_3,z_3)$, as a function of (x,y)is given by;

$$P_{[(x_1,y_1,z_1),(x_2,y_2,z_2),(x_3,y_3,z_3)]}(x,y) = \frac{I(x_1 - x) + J(y_1 - y) + Kz_1}{K}$$

where

$$I = (y_2 - y_1)(z_3 - z_1) - (z_2 - z_1)(y_3 - y_1),$$
$$J = (z_2 - z_1)(x_3 - x_1) - (x_2 - x_1)(z_3 - z_1),$$
$$K = (x_2 - x_1)(y_3 - y_1) - (y_2 - y_1)(x_3 - x_1). \quad \text{Eq2.3}$$

Thus far it is possible to correct the tilts only for the first three combinations. Further analysis, however, will lead to the correction for combinations ABR and ABS.

### 2.3.2 Initial Three-Flat Test.

Referring to section 1.3.1.2, it can be seen that the combinations AB, CB and AC are exactly those required to perform the standard three-flat test for absolute flatness measurement. It can be seen from figure 2.2 that it is only along the y-axis for each flat that the contour may be solved since flat C has been inverted about the y-axis between combinations CB and AC. The three-flat test is applied to the deviation-sum data for AB, AC and CB to yield the absolute contours for flats A and B along their y-axes;

$$f_A(0,y) = \frac{g_{AB}(0,y) + g_{AC}(0,y) - g_{CB}(0,y)}{2},$$
$$f_B(0,y) = \frac{g_{AB}(0,y) + g_{CB}(0,y) - g_{AC}(0,y)}{2},$$
$$f_C(0,y) = \frac{g_{CB}(0,y) + g_{AC}(0,y) - g_{AB}(0,y)}{2}. \quad \text{Eq.2.4}$$

Flat C will play no further part in the analysis and the following analysis will work towards finding the whole-surface contours of flats A and B only.

### 2.3.3 Extension To Orthogonal Diameters.

Given that the absolute contours of flats A and B are now known along the y-axis of each flat, it follows that the deviations from the nominal planes for A and B are now known at the origins $(x_A,y_A) = (x_B,y_B) = (0,0)$. These may be denoted by $f_A(0,0)$ and $f_B(0,0)$.

The correction of piston and tilt for combination ABR may now be performed. Referring to figure 2.2.1, it can be seen that points $(x_A,y_A) = (R,0)$ and $(x_B,y_B) = (0,-R)$ are now coincident and thus the deviation-sum here should equal zero. The same is true for points $(x_A,y_A) = (0,-R)$ and $(x_B,y_B) = (R,0)$. Additionally, the origins of A and B being coincident, the deviation-sum here should be $f_A(0,0)+f_B(0,0)$. The piston and tilt correction required for ABR is therefore that the plane defined by the difference between the measured, uncorrected, combination sum, $G_{ABR}(x,y)$ and the known deviations at these three points be subtracted from the measured, uncorrected, deviation-sum. The result is the corrected deviation-sum which will now simply be called the deviation sum of ABR, $g_{ABR}(x,y)$;

$$g_{ABR}(x,y) = G_{ABR}(x,y) - P_{[G_{ABR}(R,0),G_{ABR}(0,-R),\{G_{ABR}(0,0)-[f_A(0,0)+f_B(0,0)]\}]}(x,y). \quad \text{Eq.2.5}$$

It is now possible to find the contours of the x-axes of A and B. Subtracting the known contour of A's y-axis from the ABR deviation sum data yields the contour of B's x-axis and vice versa;

$$f_A(x,0) = g_{ABR}(x,0) - f_B(0,-y),$$
$$f_B(x,0) = g_{ABR}(0,-y) - f_A(0,-y). \quad \text{Eq.2.6}$$

### 2.3.4 Extending the analysis across the flats.

Once the profiles of flats A and B have been found along the x and y axes it is possible to proceed to find the profiles of an array of equally spaced chords parallel to the y-axis. First, however, it is necessary to correct the piston and tilt in the deviation-sum array for combination ABS. The correction of piston and the component of tilt proportional to the x coordinate (x-tilt) is straightforward and will be treated first. The correction of the component of tilt proportional to the y coordinate (y-tilt) is more difficult.

#### 2.3.4.1 Correction Of Piston and X-Tilt In ABS Data.

Referring to figure 2.2 it can be seen that the x-axes of A and B in the ABS combination are superimposed but translated by distance, s. This is shown in side view in figure 2.3.

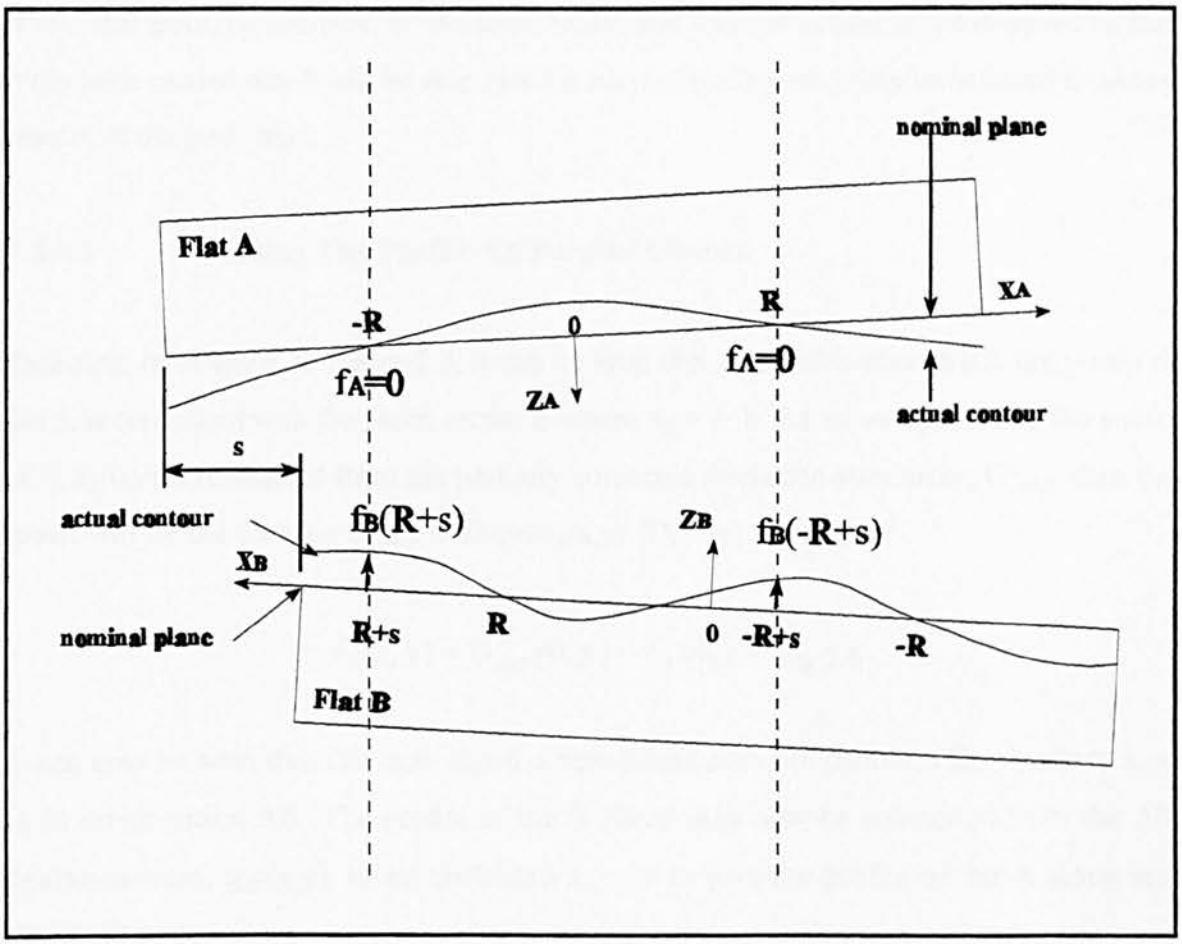


Figure 2.3. Section along x-axes for combination ABS.

For the purposes of this discussion, consider two points on A; $(x_A,y_A) = (R,0)$ and $(x_A,y_A)= (-R,0)$ where, by definition $f_A = 0$. The coincident points on B are $(x_B,y_B) = (-R+s,0)$ and $(x_B,y_B)= (R+s,0)$ respectively. At these points the deviation sum should be equal to $f_B(-R+s,0)$ and $f_B(R+s,0)$ respectively. The x-tilt and piston may thus be corrected in ABS by finding a plane defined by the difference between the measured, uncorrected error-sum at these two points and the required values. There are, of course, an infinite number of such planes since two points can only define a straight line and any straight line lies in an infinite number of different planes. It is convenient, therefore, just to set the y component of the plane to be removed equal to zero. The ABS array thus corrected for piston and x-tilt will be referred to as the "partially corrected deviation-sum", $G^P_{ABS}(x,y)$.

$$G^P_{ABS}(x,y) = G_{ABS}(x,y) - P_{[G_{ABS}(-R+s,0),G_{ABS}(R+s,0)]}(x,y). \quad \text{Eq.2.7}$$

It is not possible, at present, to find the correction required for the y-tilt in the ABS data.

It will, therefore, be assumed, for the time being, that this correction is not required or that it has been carried out. It will be seen that the correction required may be inferred from the results of the next step.

### 2.3.4.2 Finding The Profiles Of Parallel Chords.

Referring, once again, to figure 2.2, it can be seen that, in combination ABS, the y-axis of flat A is coincident with the chord on flat B where $x_B = s$. If the known profile of the y-axis of A, $f_A(0,y)$ is subtracted from the partially corrected deviation-sum array, $G^P_{ABS}$, then the result will be the contour of the B chord $f_B(s,y)$. That is;

$$f_B(s,y) = G^P_{ABS}(0,y) - f_A(0,y). \quad \text{Eq.2.8}$$

It can now be seen that this new chord is coincident with the chord on flat A where $x_A = -s$ in combination AB. The profile of the B chord may now be subtracted from the AB deviation-sum, $g_{AB}(x,y)$, along the chord $x_A = -s$ to give the profile of flat A along that chord. That is;

$$f_A(-s,y) = g_{AB}(-s,y) - f_B(s,y). \quad \text{Eq.2.9}$$

This new chord may now be subtracted from $G^P_{ABS}(x,y)$ to give the profile of the chord where $x_B = 2s$. That is;

$$f_B(2s,y) = G^P_{ABS}(-s,y) - f_A(-s,y). \quad \text{Eq.2.10}$$

This process may be continued, subtracting the profile of each new chord from $G^P_{ABS}$ or $g_{AB}$ to give the profile of a further chord until the edge of the flat is reached. The process will then have found the profiles of chords on flat A where $x_A = -s, -2s, -3s....$ and chords on flat B where $x_B = s, 2s, 3s....$

A similar process may be carried out using the known profile of flat B along its y-axis. Subtracting this profile from $G^p_{ABS}(x,y)$ gives the profile of the chord on A where $x_A = s$. That is:

$$f_A(s,y) = G^p_{ABS}(s,y) - f_B(0,y). \quad \text{Eq.2.11}$$

Then, subtracting this new chord from $g_{AB}(x,y)$ gives the profile of the chord on B where $x_B = -s$. That is:

$$f_B(-s,y) = g_{AB}(s,y) - f_A(s,y). \quad \text{Eq.2.12}$$

And so on.

The process will then have found the profiles of chords on flat A where $x_A = s, 2s, 3s....$ and chords on flat B where $x_B = -s, -2s, -3s....$

This procedure for finding the profiles of the chords is illustrated in figure 2.4 For the sake of simplicity and clarity a constant term, c, has been added to the values of $G^p_{ABS}$ and $g_{AB}$ in order to separate the surfaces of A and B.

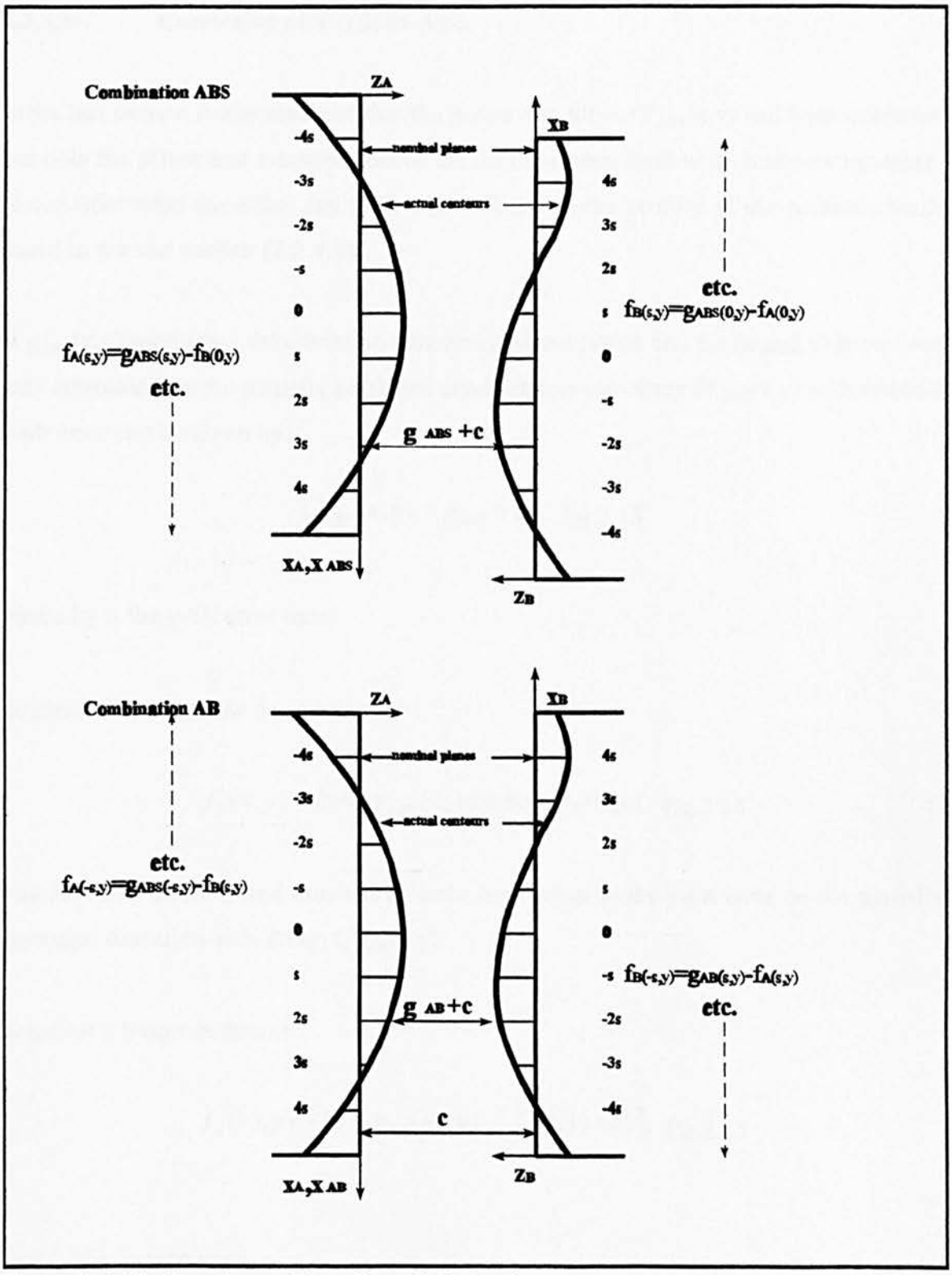


Figure 2.4. Finding the profiles of parallel chords.

### 2.3.4.3 Correction of Y-Tilt In ABS.

In the last section it was assumed that the piston and tilt in $G^p_{ABS}(x,y)$ had been corrected but only the piston and x-component of the tilt have been dealt with. It is now necessary to consider what the effect any y-tilt error will be on the profiles of the parallel chords found in the last section (2.3.4.2).

If $g_{ABS}(x,y)$ represents the deviation-sum array where piston and tilt (x <u>and</u> y) have been fully corrected then the partially corrected combination-sum array $G^p_{ABS}(x,y)$ with residual y-tilt error can be given by:

$$G^p_{ABS}(x,y) = g_{ABS} + ky, \quad \text{Eq.2.13}$$

where ky is the y-tilt error term.

Equation 2.8 may now be rewritten;

$$f_B(s,y) + ky = g_{ABS}(0,y) + ky - f_B(0,y). \quad \text{Eq.2.14}$$

The first new chord found thus has an error term equal to the y-tilt error on the partially corrected deviation-sum array, $G^p_{ABS}(x,y)$.

Equation 2.9 now becomes:

$$f_A(-s,y) - ky = g_{AB}(-s,y) - [f_B(s,y) + ky], \quad \text{Eq.2.15}$$

and equation 2.10 becomes:

$$f_B(2s,y) + 2ky = g_{ABS}(-s,y) + ky - [f_A(-s,y) - ky]. \quad \text{Eq.2.16}$$

The next B chord has thus gained an extra ky error term.

In general each chord on the B flat, $f_B(ns,y)$, has an error term equal to nky and each chord on the A flat, $f_A(ns,y)$, has an error term equal to -nky.

In order to find the value of the y-tilt term, k, the array of B-chords is inverted and rotated by 90° and then added to the array of A-chords in the same way as the flats A and B are combined in combination ABR. This is illustrated in figure 2.5. Naturally, the chords may only be added at the points where they cross, giving a square array of values spaced a distance, s, apart. Let this array of values be denoted by $G'_{ABR}(x,y)$.

Generally $G'_{ABR}(x,y)$ is given by;

$$\begin{aligned} G'_{ABR}(x,y) &= [f_A(p,q)+kpq]+[f_B(-p,-q)+kpq] \\ &= f_A(p,q)+f_B(-p,-q)+2kpq, \end{aligned} \qquad \text{Eq.2.17}$$

where x = ps, y = qs and p,q = 0, 1, -1, 2, -2, ....

For the corrected deviation-sum array for combination ABR, $g_{ABR}(x,y)$, the values are given by:

$$g_{ABR} = f_A(x,y)+f_B(-x,-y). \qquad \text{Eq.2.18}$$

If this array is sampled on a square grid of spacing, s, and the array, $G'_{ABR}(p,q)$, is subtracted from it, the result reveals the value of k.

$$\begin{aligned} G'_{ABR}(p,q)-g_{ABR}(p,q) &= 2kpq, \\ k &= \frac{G'_{ABR}(p,q)-g_{ABR}(p,q)}{2pq}. \end{aligned} \qquad \text{Eq.2.19}$$

Thus the value of the y-tilt error in $G^p_{ABS}(x,y)$ has been determined and may now be removed to give the corrected deviation-sum array, $g_{ABS}(x,y)$ for combination ABS.

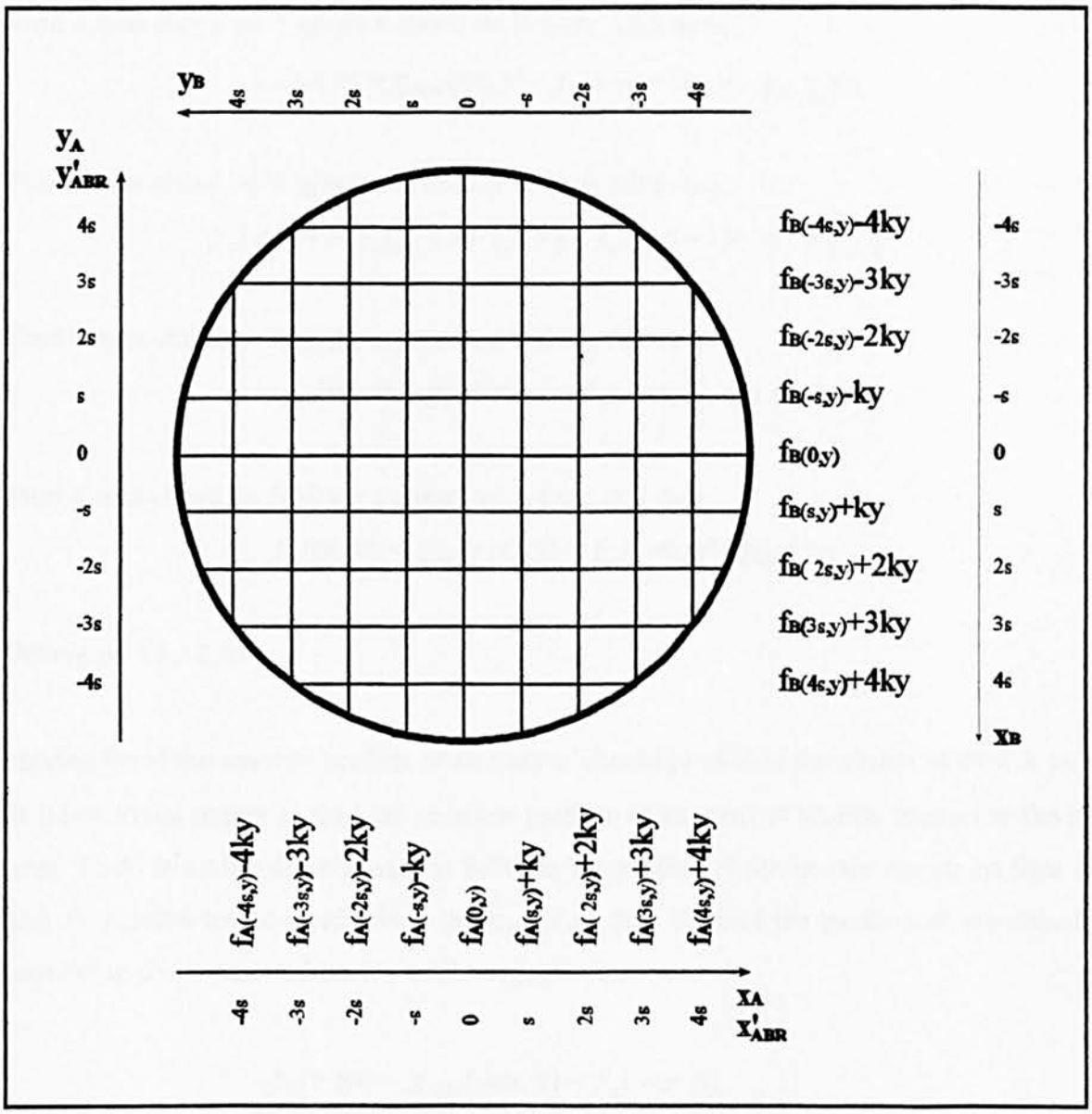


Figure 2.5. Addition of uncorrected chords to find k, the y-tilt error term.

The absolute contours of the surfaces may now be found with the correct spatial orientation between the parallel chords in the array. This may be achieved either by individually removing the error term, nky from each chord or by using the corrected error-sum array for combination ABS, $g_{ABS}(x,y)$ and re-performing the analysis from equation 2.8 onwards.

The equations for finding the profiles of the chords may be written generally as:

Find a new chord on A given a chord on B from ABS data;

$$f_A(ns,y) = g_{ABS}(ns,y) - f_B((-n+1)s,y). \quad \text{Eq.2.20}$$

Find a new chord on B given a chord on A from ABS data;

$$f_B(ns,y) = g_{ABS}(-(n-1)s,y) - f_A(-(n-1)s,y). \quad \text{Eq.2.21}$$

Find a new chord on A given a chord on B from AB data;

$$f_A(ns,y) = g_{AB}(ns,y) - f_B(-ns,y). \quad \text{Eq.2.22}$$

Find a new chord on B given a chord on A from AB data;

$$f_B(ns,y) = g_{AB}(-ns,y) - f_A(-ns,y). \quad \text{Eq.2.23}$$

Where n= ±1,±2,±3...

Having found the absolute profiles of an array of chords parallel to the y-axes of flats A and B it is a trivial matter to find the absolute profiles of an array of chords parallel to the x-axes. To do this it is only necessary to subtract the profiles of the known chords on flats A and B, parallel to the y-axis from the $g_{ABR}(x,y)$ data to yield the profiles of the chords parallel to the x-axis on flats B and A respectively;

$$f_B(x,ns) = g_{ABR}(-ns,y) - f_A(-ns,y),$$
$$f_A(x,ns) = g_{ABR}(ns,y) - f_B(-ns,y). \quad \text{Eq.2.24}$$

The analysis described yields the absolute contours of the flats at all points on a square grid of chords spaced by distance s in the x and y directions. The analysis therefore yields more information than the test described by Grzanna (1990, 1994) (see section 1.3.1.3.1) which finds the values of the surface figure functions only on a square grid of points from a similar set of positional combinations. The array of points found in Grzanna's method are included in the results found by the method described above, being the points of intersection on the array of chords.

## 2.4 Small Scale Manual Demonstration of the Algorithm.

In order to demonstrate the validity of the above algorithm and to illustrate the steps in the analysis it is helpful to perform a manual simulation using small arrays of data. A random array of data is used to simulate the surface function of each flat. The use of a random array ensures that the results of the algorithm are not the accidental result of an inappropriately chosen figure function. The simulated figure functions for the three flats are shown below. For the purposes of this simulation R =2 and so the appropriate points defining the nominal planes for the flats have been set to zero (Equation 2.1)

FLAT A-$f_A(x,y)$

| y↓ x→ | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|---|---|---|
| 4 | 63 | 76 | 80 | 33 | 81 | 0 | 12 | 5 | 68 |
| 3 | 93 | 18 | 31 | 93 | 69 | 94 | 32 | 1 | 6 |
| 2 | 17 | 13 | 97 | 85 | 22 | 79 | 33 | 62 | 39 |
| 1 | 64 | 60 | 85 | 43 | 68 | 44 | 82 | 42 | 47 |
| 0 | 80 | 74 | 0 | 17 | 43 | 82 | 0 | 2 | 92 |
| -1. | 86 | 97 | 68 | 45 | 46 | 25 | 62 | 81 | 93 |
| -2. | 17 | 7 | 42 | 49 | 0 | 3 | 41 | 30 | 30 |
| -3. | 81 | 16 | 52 | 95 | 37 | 51 | 34 | 61 | 78 |
| -4. | 94 | 17 | 43 | 95 | 39 | 46 | 20 | 40 | 4 |

FLAT B-$f_B(x,y)$

| y↓ x→ | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|---|---|---|
| 4 | 41 | 14 | 3 | 96 | 68 | 70 | 81 | 59 | 57 |
| 3 | 99 | 61 | 44 | 27 | 85 | 21 | 77 | 80 | 42 |
| 2 | 73 | 89 | 85 | 55 | 78 | 47 | 82 | 72 | 41 |
| 1 | 98 | 2 | 91 | 45 | 75 | 88 | 93 | 19 | 44 |
| 0 | 11 | 73 | 0 | 85 | 47 | 91 | 0 | 98 | 68 |
| -1. | 80 | 49 | 79 | 72 | 93 | 97 | 44 | 17 | 4 |
| -2. | 61 | 30 | 45 | 63 | 0 | 81 | 44 | 77 | 34 |
| -3. | 10 | 8 | 16 | 37 | 97 | 94 | 43 | 24 | 99 |
| -4. | 24 | 77 | 24 | 52 | 20 | 36 | 52 | 11 | 35 |

FLAT C-$f_C(x,y)$

| y↓ x→ | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|---|---|---|
| 4 | 61 | 30 | 55 | 73 | 27 | 49 | 42 | 36 | 666 |
| 3 | 4 | 26 | 14 | 17 | 26 | 84 | 43 | 2 | 22 |
| 2 | 65 | 81 | 94 | 95 | 36 | 71 | 70 | 26 | 3 |
| 1 | 87 | 65 | 31 | 97 | 7 | 77 | 20 | 77 | 94 |
| 0 | 47 | 31 | 0 | 89 | 48 | 22 | 0 | 84 | 98 |
| -1. | 56 | 39 | 95 | 1 | 52 | 55 | 53 | 6 | 93 |
| -2. | 55 | 7 | 6 | 99 | 0 | 17 | 46 | 14 | 27 |
| -3. | 17 | 76 | 66 | 55 | 0 | 31 | 71 | 83 | 76 |
| -4. | 69 | 95 | 12 | 21 | 12 | 70 | 60 | 82 | 77 |

To demonstrate the normalisation of initial data (section 2.3.1) the data for the AB positional combination has had an x and y tilt added.

Uncorrected data for combination AB, $G_{AB}(x,y)$

| y↓ x→ | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|---|---|---|
| 4 | 128 | 145 | 173 | 117 | 165 | 114 | 35 | 41 | 133 |
| 3 | 142 | 107 | 119 | 127 | 169 | 138 | 95 | 83 | 128 |
| 2 | 64 | 93 | 189 | 144 | 114 | 150 | 136 | 171 | 134 |
| 1 | 113 | 86 | 187 | 142 | 156 | 104 | 190 | 63 | 166 |
| 0 | 152 | 178 | 8 | 118 | 102 | 181 | 16 | 93 | 123 |
| -1. | 93 | 119 | 119 | 151 | 150 | 110 | 156 | 147 | 192 |
| -2. | 53 | 88 | 92 | 138 | 10 | 78 | 100 | 76 | 109 |
| -3. | 181 | 43 | 100 | 196 | 142 | 99 | 63 | 84 | 105 |
| -4. | 129 | 30 | 99 | 137 | 67 | 106 | 56 | 131 | 44 |

According to equation 2.3, the equation of the plane to be removed from $G_{AB}$ (x,y) is;

$$P(x,y) = 2x + y + 12.$$

Subtracting this plane from $G_{AB}$(x,y) gives $g_{AB}$(x,y), the corrected deviation sum for combination AB. As expected, $g_{AB}$(x,y) equals the sum of the surface figure functions for flats A and B;

$$g_{AB}(x,y) = f_A(x,y) + f_B(-x,y).$$

It will be assumed that this correction has also been performed for the measured data, $G_{CB}$(x,y) and $G_{AC}$(x,y) to give $g_{CB}$(x,y) and $g_{AC}$(x,y) respectively. Only the data for the y-axes of these two combinations will be shown since these are the only points used in the analysis.

Corrected data for combination AB, $g_{AB}$(x,y)

| y↓ x→ | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|---|---|---|
| 4 | 120 | 135 | 161 | 103 | 149 | 96 | 15 | 19 | 109 |
| 3 | 135 | 98 | 108 | 114 | 154 | 121 | 76 | 62 | 105 |
| 2 | 58 | 85 | 179 | 132 | 100 | 134 | 118 | 151 | 112 |
| 1 | 108 | 79 | 178 | 131 | 143 | 89 | 173 | 44 | 145 |
| 0 | 148 | 172 | 0 | 108 | 90 | 167 | 0 | 75 | 103 |
| -1. | 90 | 114 | 112 | 142 | 139 | 97 | 141 | 130 | 173 |
| -2. | 51 | 84 | 86 | 130 | 0 | 66 | 86 | 60 | 91 |
| -3. | 180 | 40 | 95 | 189 | 134 | 88 | 50 | 69 | 88 |
| -4. | 129 | 28 | 95 | 131 | 59 | 98 | 44 | 117 | 28 |

Corrected data for combinations CB and AC, $g_{CB}(0,y)$ and $g_{AC}(0,y)$

| y | | $g_{CB}(0,y)$ | | $g_{AC}(0,y)$ |
|---|---|---|---|---|
| 4 | | 95 | | 108 |
| 3 | | 111 | | 95 |
| 2 | | 114 | | 58 |
| 1 | | 82 | | 75 |
| 0 | | 95 | | 91 |
| -1 | | 145 | | 98 |
| -2 | | 0 | | 0 |
| -3 | | 97 | | 37 |
| -4 | | 32 | | 51 |

Equations 2.4 may now be applied to the data to yield the contours of each flat along the y-axis.

| y | $g_{AB}(0,y)$ | $g_{CB}(0,y)$ | $g_{AC}(0,y)$ | $f_A(0,y)=(g_{AB}+g_{AC}-g_{CB})/2$ | $f_B(0,y)=(g_{AB}+g_{AC}-g_{CB})/2$ | $f_C(0,y)=(g_{AB}+g_{AC}-g_{CB})/2$ |
|---|---|---|---|---|---|---|
| 4 | 149 | 95 | 108 | 81 | 68 | 27 |
| 3 | 154 | 111 | 95 | 69 | 85 | 26 |
| 2 | 100 | 114 | 58 | 22 | 78 | 36 |
| 1 | 143 | 82 | 75 | 68 | 75 | 7 |
| 0 | 90 | 95 | 91 | 43 | 47 | 48 |
| -1 | 139 | 145 | 98 | 46 | 93 | 52 |
| -2 | 0 | 0 | 0 | 0 | 0 | 0 |
| -3 | 134 | 97 | 37 | 37 | 97 | 0 |
| -4 | 59 | 32 | 51 | 39 | 20 | 12 |

As expected, these results agree with the initial values entered for the surface figure functions of the three flats.

It is now possible to consider the measured data for combination ABR, $G_{ABR}(x,y)$ and find the profiles of the x-axes of A and B. Following the discussion in section 2.3.3 the value of the corrected data at the origin should equal the sum of the two surface functions for A and B at the origin;

$$g_{ABR}(0,0) = f_A(0,0) + f_B(0,0).$$

From equation 2.3, the plane which must be subtracted from $G_{ABR}(x,y)$ is;

$$P(x,y) = 3x - 2y + 20.$$

Uncorrected data for combination ABR, $G_{ABR}(x,y)$

| y↓ x→ | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|---|---|---|
| 4 | 104 | 178 | 159 | 140 | 104 | 95 | 91 | 36 | 116 |
| 3 | 109 | 84 | 128 | 106 | 156 | 160 | 82 | 32 | 109 |
| 2 | 24 | 64 | 192 | 189 | 38 | 177 | 100 | 103 | 91 |
| 1 | 166 | 96 | 152 | 113 | 171 | 137 | 169 | 106 | 129 |
| 0 | 156 | 170 | 92 | 109 | 110 | 198 | 26 | 128 | 144 |
| -1. | 166 | 131 | 131 | 152 | 149 | 147 | 171 | 206 | 163 |
| -2. | 110 | 99 | 142 | 163 | 24 | 74 | 105 | 106 | 118 |
| -3. | 154 | 113 | 144 | 137 | 161 | 97 | 143 | 110 | 127 |
| -4. | 167 | 78 | 106 | 164 | 135 | 81 | 88 | 176 | 79 |

Corrected data for combination ABR, $g_{ABR}(x,y)$

| y↓ x→ | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|---|---|---|
| 4 | 104 | 175 | 153 | 131 | 92 | 80 | 73 | 15 | 92 |
| 3 | 107 | 79 | 120 | 95 | 142 | 143 | 62 | 9 | 83 |
| 2 | 20 | 57 | 182 | 176 | 22 | 158 | 78 | 78 | 63 |
| 1 | 160 | 87 | 140 | 88 | 153 | 116 | 145 | 79 | 99 |
| 0 | 148 | 159 | 78 | 92 | 90 | 175 | 0 | 99 | 112 |
| -1. | 156 | 118 | 115 | 133 | 137 | 122 | 143 | 175 | 129 |
| -2. | 98 | 84 | 124 | 142 | 0 | 47 | 85 | 73 | 82 |
| -3. | 140 | 96 | 124 | 114 | 135 | 68 | 111 | 85 | 89 |
| -4. | 151 | 59 | 84 | 139 | 107 | 50 | 54 | 139 | 39 |

Using equations 2.6 the profiles of flats A and B along the x-axes are given as follows.

| x/y | $g_{ABR}(x,0)$ | $f_B(0,-y)$ | $f_A(-x,0)=$ $g_{ABR}(-x,0)-f_B(0,-y)$ | $g_{ABR}(0,-y)$ | $f_A(0,-y)$ | $f_B(x,0)=$ $g_{ABR}(0,-y)-f_A(0,-y)$ |
|---|---|---|---|---|---|---|
| 4 | 112 | 20 | 92 | 107 | 39 | 68 |
| 3 | 99 | 97 | 2 | 135 | 37 | 98 |
| 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 175 | 93 | 82 | 137 | 46 | 91 |
| 0 | 90 | 47 | 43 | 90 | 43 | 47 |
| -1 | 92 | 75 | 17 | 153 | 68 | 85 |
| -2 | 78 | 78 | 0 | 22 | 22 | 0 |
| -3 | 159 | 85 | 74 | 142 | 69 | 73 |
| -4 | 148 | 68 | 80 | 92 | 81 | 11 |

It is now possible to consider the simulated measured data for combination ABS, $G_{ABS}(x,y)$. The simulated data is shown below. Note that the left hand column is empty, corresponding to the region where the two flats do not overlap due to the lateral shift of flat B.

Uncorrected data for combination ABS, $G_{ABS}(x,y)$

| y↓ x→ | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|---|---|---|
| 4 | | 154 | 157 | 129 | 163 | 77 | 114 | 11 | 82 |
| 3 | | 84 | 132 | 188 | 105 | 191 | 68 | 51 | 70 |
| 2 | | 81 | 193 | 188 | 87 | 172 | 100 | 156 | 134 |
| 1 | | 134 | 131 | 160 | 177 | 137 | 142 | 145 | 58 |
| 0 | | 175 | 128 | 44 | 158 | 150 | 103 | 17 | 177 |
| -1. | | 137 | 118 | 129 | 170 | 142 | 155 | 178 | 157 |
| -2. | | 80 | 155 | 126 | 111 | 30 | 128 | 96 | 78 |
| -3 | | 157 | 115 | 174 | 164 | 178 | 98 | 101 | 107 |
| -4. | | 97 | 96 | 186 | 111 | 99 | 102 | 91 | 105 |

From the discussion in section 2.3.4., it is only possible, at this point to determine the necessary corrections for piston and x-tilt in the data. From equation 2.7, the equation of the plane that must be subtracted from the data to perform these corrections is;

$$P(x,y) = -3x + 24.$$

The partially corrected data, $G^{p}_{ABS}(x,y)$ are shown below.

Partially corrected data for combination ABS, $G^{p}_{ABS}(x,y)$.

| y↓ x→ | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|---|---|---|
| 4 | | 121 | 127 | 102 | 139 | 56 | 96 | -4 | 70 |
| 3 | | 51 | 102 | 161 | 81 | 170 | 50 | 36 | 58 |
| 2 | | 48 | 163 | 161 | 63 | 151 | 82 | 141 | 122 |
| 1 | | 101 | 101 | 133 | 153 | 116 | 124 | 130 | 46 |
| 0 | | 142 | 98 | 17 | 134 | 129 | 85 | 2 | 165 |
| -1. | | 104 | 88 | 102 | 146 | 121 | 137 | 163 | 145 |
| -2. | | 47 | 125 | 99 | 87 | 9 | 110 | 81 | 66 |
| -3. | | 124 | 85 | 147 | 140 | 157 | 80 | 86 | 95 |
| -4. | | 64 | 66 | 159 | 87 | 78 | 84 | 76 | 93 |

Using this data, and the data set $g_{AB}(x,y)$, the equations 2.8 to 2.12 may be applied to find the profiles of a parallel array of chords on each flat with the y-tilt error terms included.

Following the discussion in section 2.3.4.3, the array of chords on flat A is added to the flipped and rotated array of chords on flat B to give the square array of values, $G'_{ABR}(x,y)$. The value of the y-tilt error term is found by subtracting the corrected data for the ABR positional combination, $g_{ABR}(x,y)$ from $G'_{ABR}(x,y)$ (equation 2.19).

Profile of parallel chords on Flat A before y-tilt correction.

| y↓ x→ | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|---|---|---|
| 4 | 111 | 112 | 104 | 45 | 81 | -12 | -12 | -31 | 20 |
| 3 | 129 | 45 | 49 | 102 | 69 | 85 | 14 | -26 | -30 |
| 2 | 41 | 31 | 109 | 91 | 22 | 73 | 21 | 44 | 15 |
| 1 | 76 | 69 | 91 | 46 | 68 | 31 | 66 | 23 | 25 |
| 0 | 80 | 74 | 0 | 17 | 43 | 82 | 0 | 2 | 92 |
| -1 | 64 | 78 | 52 | 42 | 46 | 28 | 68 | 90 | 105 |
| -2. | -7 | -11 | 30 | 43 | 0 | 9 | 53 | 48 | 54 |
| -3. | 45 | -11 | 34 | 86 | 37 | 60 | 52 | 88 | 114 |
| -4. | 46 | -19 | 19 | 83 | 39 | 58 | 44 | 76 | 52 |

Profile of parallel chords on Flat B before y-tilt correction.

| y↓ x→ | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|---|---|---|
| 4 | 89 | 50 | 27 | 108 | 68 | 58 | 57 | 23 | 9 |
| 3 | 135 | 88 | 62 | 36 | 85 | 12 | 59 | 53 | 6 |
| 2 | 97 | 107 | 97 | 61 | 78 | 41 | 70 | 54 | 17 |
| 1 | 120 | 21 | 107 | 58 | 75 | 85 | 87 | 10 | 32 |
| 0 | 11 | 73 | 0 | 85 | 47 | 91 | 0 | 98 | 68 |
| -1. | 68 | 40 | 73 | 69 | 93 | 100 | 60 | 36 | 26 |
| -2. | 37 | 12 | 33 | 57 | 0 | 87 | 56 | 95 | 58 |
| -3. | -26 | -19 | -2 | 28 | 97 | 103 | 61 | 51 | 135 |
| -4 | -24 | 41 | 0 | 40 | 20 | 48 | 76 | 47 | 83 |

Sum of A chords and flipped and rotated B chords with y-tilt error, $G'_{ABR}(x,y)$.

| y↓ x→ | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|---|---|---|
| 4 | 200 | 247 | 201 | 155 | 92 | 56 | 25 | -57 | -4 |
| 3 | 179 | 133 | 156 | 113 | 142 | 125 | 26 | -45 | 11 |
| 2 | 68 | 93 | 206 | 188 | 22 | 146 | 54 | 42 | 15 |
| 1 | 184 | 105 | 152 | 94 | 153 | 110 | 133 | 61 | 75 |
| 0 | 148 | 159 | 78 | 92 | 90 | 175 | 0 | 99 | 112 |
| -1. | 132 | 100 | 103 | 127 | 137 | 128 | 155 | 193 | 153 |
| -2. | 50 | 48 | 100 | 130 | 0 | 59 | 109 | 109 | 130 |
| -3. | 68 | 42 | 88 | 96 | 135 | 86 | 147 | 139 | 161 |
| -4 | 55 | -13 | 36 | 115 | 107 | 74 | 102 | 211 | 135 |

$G'_{ABR}(x,y)-g_{ABR}(x,y)$

| y↓ x→ | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|---|---|---|
| 4 | 96 | 72 | 48 | 24 | 0 | -24 | -48 | -72 | -96 |
| 3 | 72 | 54 | 36 | 18 | 0 | -18 | -36 | -54 | -72 |
| 2 | 48 | 36 | 24 | 12 | 0 | -12 | -24 | -36 | -48 |
| 1 | 24 | 18 | 12 | 6 | 0 | -6 | -12 | -18 | -24 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| -1. | -24 | -18 | -12 | -6 | 0 | 6 | 12 | 18 | 24 |
| -2. | -48 | -36 | -24 | -12 | 0 | 12 | 24 | 36 | 48 |
| -3. | -72 | -54 | -36 | -18 | 0 | 18 | 36 | 54 | 72 |
| -4 | -96 | -72 | -48 | -24 | 0 | 24 | 48 | 72 | 96 |

It can easily be seen from the values of $G'_{ABR}(x,y)-g_{ABR}(x,y)$ and from equation 2.19 that the y-tilt error term, k=-3. Therefore the plane that must be subtracted from the partially corrected data for combination ABS, $G^{P}_{ABS}(x,y)$ is P(x,y)=-3y. This gives the fully corrected data for combination ABS, $g_{ABS}(x,y)$.

Corrected data for combination ABS, $g_{ABS}(x,y)$

| y↓ x→ | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|---|---|---|
| 4 | | 133 | 139 | 114 | 151 | 68 | 108 | 8 | 82 |
| 3 | | 60 | 111 | 170 | 90 | 179 | 59 | 45 | 67 |
| 2 | | 54 | 169 | 167 | 69 | 157 | 88 | 147 | 128 |
| 1 | | 104 | 104 | 136 | 156 | 119 | 127 | 133 | 49 |
| 0 | | 142 | 98 | 17 | 134 | 129 | 85 | 2 | 165 |
| -1. | | 101 | 85 | 99 | 143 | 118 | 134 | 160 | 142 |
| -2. | | 41 | 119 | 93 | 81 | 3 | 104 | 75 | 60 |
| -3. | | 115 | 76 | 138 | 131 | 148 | 71 | 77 | 86 |
| -4. | | 52 | 54 | 147 | 75 | 66 | 72 | 64 | 81 |

Equations 2.8 to 2.12 may now be applied to this data to derive the profiles of the array of parallel chords with the correct spatial relationship to each other. Notice that the results of this step are identical to the synthesised data for flats A and B which were input at the start of the analysis. This verifies the validity of this flatness measuring algorithm.

Profiles of parallel chords on Flat A derived from corrected data.

| y↓ x→ | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|---|---|---|
| 4 | 63 | 76 | 80 | 33 | 81 | 0 | 12 | 5 | 68 |
| 3 | 93 | 18 | 31 | 93 | 69 | 94 | 32 | 1 | 6 |
| 2 | 17 | 13 | 97 | 85 | 22 | 79 | 33 | 62 | 39 |
| 1 | 64 | 60 | 85 | 43 | 68 | 44 | 82 | 42 | 47 |
| 0 | 80 | 74 | 0 | 17 | 43 | 82 | 0 | 2 | 92 |
| -1. | 86 | 97 | 68 | 45 | 46 | 25 | 62 | 81 | 93 |
| -2. | 17 | 7 | 42 | 49 | 0 | 3 | 41 | 30 | 30 |
| -3. | 81 | 16 | 52 | 95 | 37 | 51 | 34 | 61 | 78 |
| -4. | 94 | 17 | 43 | 95 | 39 | 46 | 20 | 40 | 4 |

Profiles of parallel chords on flat B derived from corrected data.

| y↓ x→ | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|---|---|---|
| 4 | 41 | 14 | 3 | 96 | 68 | 70 | 81 | 59 | 57 |
| 3 | 99 | 61 | 44 | 27 | 85 | 21 | 77 | 80 | 42 |
| 2 | 73 | 89 | 85 | 55 | 78 | 47 | 82 | 72 | 41 |
| 1 | 98 | 2 | 91 | 45 | 75 | 88 | 93 | 19 | 44 |
| 0 | 11 | 73 | 0 | 85 | 47 | 91 | 0 | 98 | 68 |
| -1. | 80 | 49 | 79 | 72 | 93 | 97 | 44 | 17 | 4 |
| -2. | 61 | 30 | 45 | 63 | 0 | 81 | 44 | 77 | 34 |
| -3. | 10 | 8 | 16 | 37 | 97 | 94 | 43 | 24 | 99 |
| -4. | 24 | 77 | 24 | 52 | 20 | 36 | 52 | 11 | 35 |

## 2.5 Extending the Spatial Resolution of the Test Using Two Different Lateral Shifts.

It has been shown that the method described is capable of finding the absolute contours of a pair of flats along a square array of parallel chords. In principle the method may be used to find the absolute contours at every point on the measured array of points by making the lateral shift in positional combination ABS equal to one pixel. In this case, however, the number of steps required to find the contours towards the edges of the flats would be large where the pixel spacing is small compared to the radius of the flats. The number of steps required to reach the $n^{th}$ chord is 2n.

Where a large number of steps is required to find the absolute contour at a datapoint the effect of the accumulation of random errors in the measured datapoints will become large. The effect of experimental errors will be considered in greater detail in the next chapter. It will be shown below that the number of steps required to determine the absolute contours of the flats at every datapoint may be greatly reduced by using two positional combinations, ABS1 and ABS2 with lateral shifts where the two shifts are by different distances, $s_1$ and $s_2$.

Suppose that, initially, the method described above is followed using only one of the shifted positional combinations, ABS1. This will yield the profiles of flats A and B along a parallel array of chords spaced a distance $s_1$ apart. These profiles may be written as $f_A(ns_1,y)$ and $f_B(ns_1,y)$ respectively. Suppose now that a similar proceedure is undertaken using the data from positional combination ABS2. Instead, however, of using the known profiles of flats A and B along their vertical diameters as a starting point for finding the profiles of new chords, the profiles $f_A(ns_1,y)$ and $f_B(ns_1,y)$ found from the analysis of the ABS1 data are used.

Equations 2.20 to 2.23 may now be rewritten thus:

Find a new chord on A given a chord on B from ABS2 data;

$$f_A(ns_1 + ms_2, y) = g_{ABS}(ns_1 + ms_2, y) - f_B\big((-n+1)s_1 + (-m+1)s_2, y\big). \quad \text{Eq.2.25}$$

Find a new chord on B given a chord on A from ABS2 data;

$$f_B(ns_1 + ms_2, y) = g_{ABS}(-(n-1)s_1 - (m-1)s_2, y) - f_A(-(n-1)s_1 - (m-1)s_2, y). \quad \text{Eq. 2.26}$$

Find a new chord on A given a chord on B from AB data;

$$f_A(ns_1 + ms_2, y) = g_{AB}(ns_1 + ms_2, y) - f_B(-ns_1 - ms_2, y). \quad \text{Eq.2.27}$$

Find a new chord on B given a chord on A from AB data;

$$f_B(ns_1 + ms_2, y) = g_{AB}(-ns_1 - ms_2, y) - f_A(-ns_1 - ms_2, y). \quad \text{Eq.2.28}$$

Where n= ±1, ±2, ±3..., m= ±1, ±2, ±3...

If the two lateral shifts, $s_1$ and $s_2$ are mutually prime integer numbers of pixels (having no common integer factors) then any chord may be found from 2(n+m) steps. If $s_1$ and $s_2$ are chosen such that, as well as being mutually prime, they are of the order of the square root of the radius of the flats then the maximum number of steps to find the profile of any chord will be minimized.

## Chapter 3: Implementation of the New Flatness Testing Algorithm on the Zygo Mark IV Interferometer.

The algorithm for the absolute testing of flat surfaces described in chapter 2 was implemented on a Zygo Mark IV phase measuring Fizeau interferometer (Zygo Corporation, Laurel Brook Road, P.O. Box 448, Middlefield, CT 06455-0448, U.S.A.) at the National Physical Laboratory (NPL, Teddington, Middlesex, U.K. TW11 0LW).

This Chapter describes the development of software for the implementation of the testing algorithm as well as hardware to facilitate the necessary measurements. The debugging and testing of the software using synthesized data are described. Limitations of the software, dictated by the interferometer's internal computer architecture, are identified as an obstacle to the full practical implementation of the algorithm using real experimental data. Suggestions as to how these problems may be overcome and the algorithm improved further are discussed together with an analysis of the propagation of experimental errors and how this may be minimized.

### 3.1 The Zygo Mark IV Interferometer.

The Zygo Mark IV interferometer is a phase measuring Fizeau instrument with a Helium-Neon laser source (Domenicali and Hunter, 1980). A simplified view of the internal features of the interferometer is shown in figure 3.1. The 2 mW Helium-Neon laser light source is spatially filtered using a microscope objective/ pinhole combination and converted to a plano wavefront by a collimating lens to give a measurement aperture of 100 mm (4 inches). The plane wavefront illuminates the interference cavity formed by the two optical flats. The position of the first flat in the cavity is modulated by a piezo-electric phase shifter to allow phase measurement as described in section 1.2.6.2. The light reflected by the interference cavity is directed along one of two paths. In alignment mode the reflected light falls on a screen placed at the back focus of the collimating lens and the images of the pinhole formed there due to the two surfaces in the interference cavity are imaged onto the CCD TV camera. This allows the elements in the interference cavity to be aligned by adjusting their tilts such that the two point

images are central and coincident in the camera's field of view. In measurement mode the fringe pattern formed by the interference cavity is imaged onto a rotating diffuser screen which reduces the coherence of the light. The reduction in the coherence of the light greatly reduces the problems that would otherwise arise owing to the formation of spurious interference fringes due to stray reflections in succeeding optical elements. The image of the interference pattern is transferred to the CCD device in the camera via a zoom lens with a 6:1 ratio. This allows the interference patterns due to test elements smaller than the maximum interferometer aperture of 100 mm (4 inches) to fill the field of view.

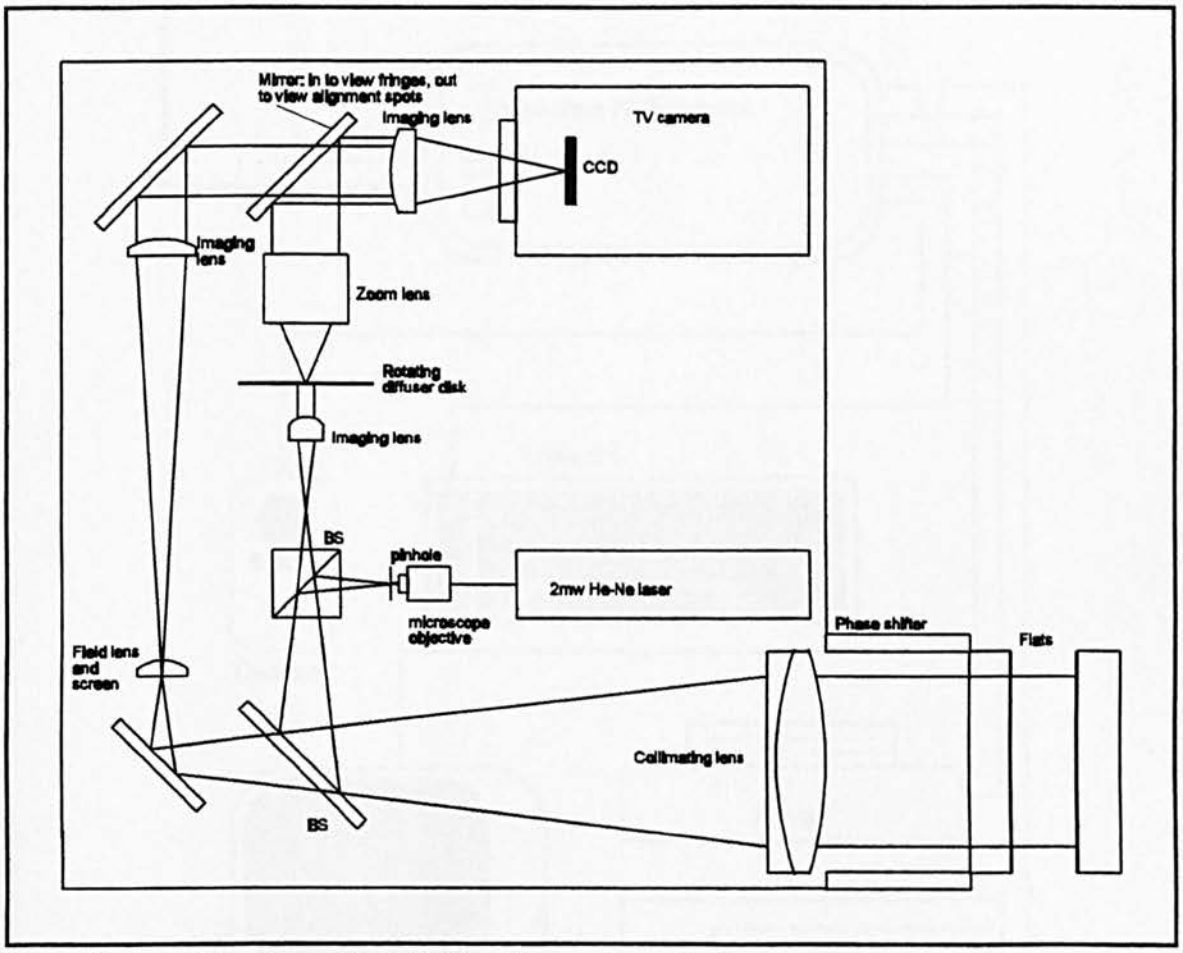


Figure 3.1 The Zygo Mark IV interferometer optical system.

Data acquisition and control of the Zygo Mark IV interferometer is accomplished by a proprietary computer system. A block diagram of the integrated Mark IV interferometer system is shown in figure 3.2.

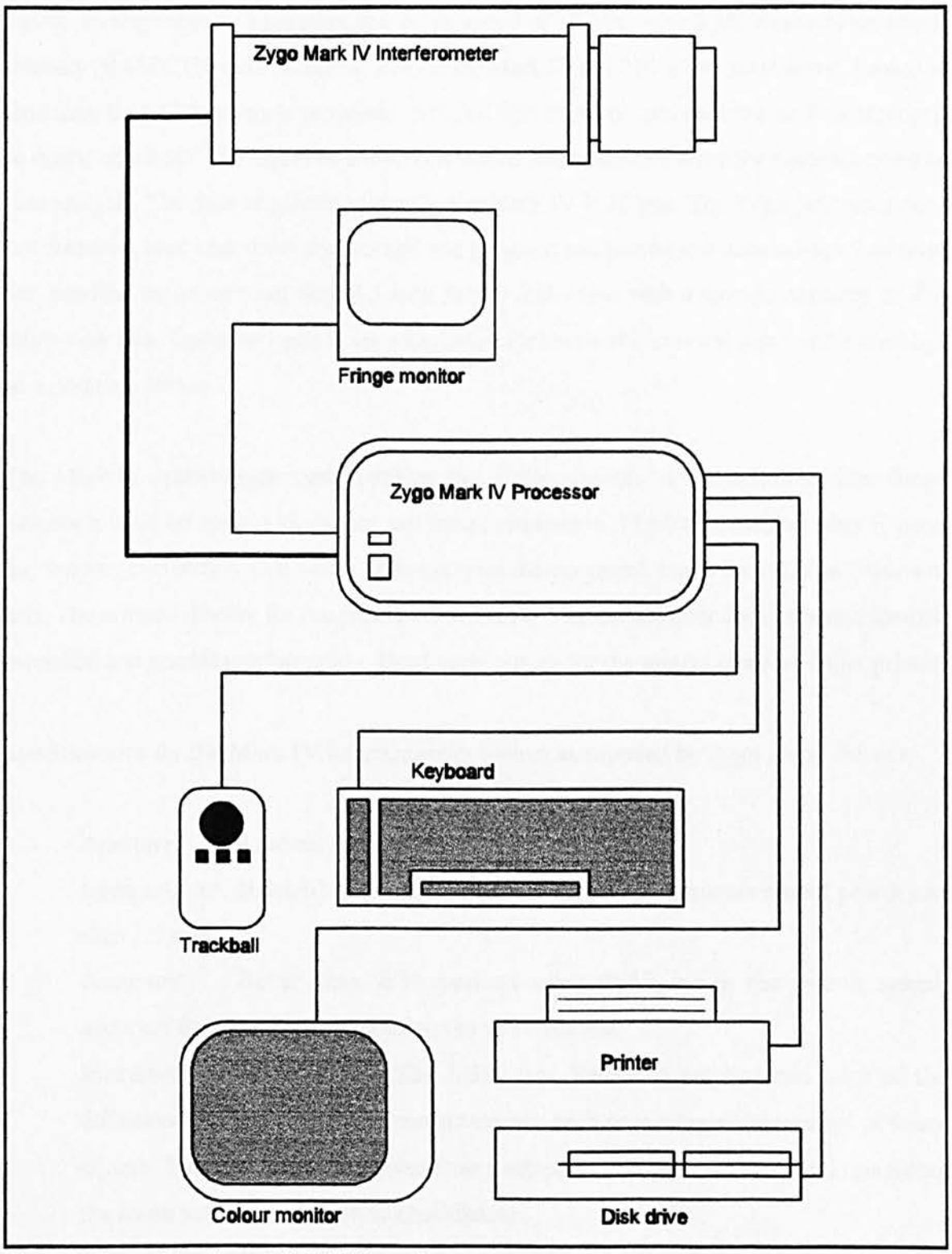


Figure 3.2 The Zygo Mark IV interferometer system.

The control, acquisition and analysis functions of the interferometer system are handled by the Zygo Mark IV processor and peripheral devices. The processor is based on the Motorola

68000 microprocessor operating at a clock speed of 10 Mhz with 2 Mbytes random access memory (RAM). The data "window" used in the Mark IV is a 210 x 280 pixel array. Analogue data from the CCD camera is processed to digital light intensity data utilizing an 8 bit analogue to digital converter. The digital information is loaded into a memory array for manipulation and data analysis. The data acquisition time for the Mark IV is 100ms. The Zygo processor does not feature a hard disk drive for storage and program and permanent data storage functions are handled by an external dual 3.5 inch floppy disk drive with a storage capacity of 720 Kbytes per disk. Operator input is via a keyboard for text and numerical input and a trackball as a pointing device.

The Mark IV interferometer system utilizes two display devices. A monochrome video fringe monitor is used for system alignment and fringe acquisition. The fringe monitor also displays any cursors currently active which define current measurement windows or act as alignment aids. The primary display for the processor system is a colour monitor for displaying textual, numerical and graphical information. Hard copy output for the system is via an inkjet printer.

Specifications for the Mark IV interferometer system as supplied by Zygo are as follows:

- Aperture: 4 inches (100mm)
- Light source: Helium-Neon laser operating at 632.8 nm. Maximum output power less than 1.5 mW.
- Accuracy: Better than $\lambda/50$ peak-to-valley (P-V), being the overall system accuracy for absolute testing using the three-flat test.
- Instrument precision: better than $\lambda/500$ rms, being the residual rms error of the difference of two consecutive measurements, each consisting of an average of 4 sets of data. The specification is derived from a sample of 100 measurements and represents the mean value plus $2\sigma$ (98 % confidence).
- Repeatability of P-V and rms: better than $\lambda/100$ and $\lambda/1000$ respectively, being for a population of 100 measurements, where each sample consists of an average of 10 sets of data. The specifications are for the $2\sigma$ (98 % confidence) repeatability of the data.
- Spatial resolution: 210 x 280 pixels.

- Computational resolution: 1/512 fringe.

### 3.1.1 Holding The Test Flats.

The Zygo interferometer system features a standardized bayonet type system for the mounting of reference optics and accessories. Reference optics, be they flat or spherical, are held in an aluminium cell which mates with a bayonet receptacle on the front of the phase shifter on the interferometer. The bayonet mount ensures stable and repeatable positioning of the reference optics and the receptacle features x and y tilt adjustment controls to align the surface to the interferometer. For holding the test surfaces relative to the interferometer, a variety of mounts with different numbers of degrees of freedom are available. For flatness testing, the required degrees of freedom are x and y-tilt and translation in the x and y directions. The mounts are equipped with the standard bayonet receptacle to accommodate surfaces mounted in bayonet equipped cells. For unmounted surfaces, there is a spring-loaded self-centring three jaw chuck which mates with the bayonet receptacle on the adjustable mount.

For the absolute flatness testing method to be implemented, three flats, A, B and C are used in various positional combinations. In one positional combination, CB, flat C is mounted in the interferometer. In all of the other positional combinations, AB, AC, ABR, ABS1 and ABS2, flat A is mounted in the interferometer. Flats A and C are thus fitted in standard bayonet fitting cells. The third flat, B, is always mounted externally to the interferometer and requires to be rotated through an angle of 90° for combination ABR and translated in the x direction by a known amount in combinations ABS1 and ABS2. To facilitate this, a special cell was constructed for flat B, an otherwise unmounted flat.

The cell for mounting flat B is shown in figure 3.3. The cell is machined from aluminium in the form of a ring. The optical flat is held in place by a flange at the rear of the cell and by three nylon screws equi-spaced around the edge. The cell is designed to be held by the self-centring three jaw chuck in the test surface mount. The cylindrical jaws of the chuck locate in three of the v-section grooves machined around the periphery of the cell. The v-section grooves are located every 30° around the periphery of the cell which allows the cell to be accurately

rotated through 90° for positional combination ABR. In order to further facilitate the positioning of flat B, a ring shaped accessory (figure 3.4) bearing fiducial pointers is attached to the front of the cell. The ring is attached by screws into tapped holes in the front of the cell. The screws pass through arc shaped slots machined into the ring which allow the angular position of the ring to be adjusted. This adjustment is necessary since the angular position of the jaws of the three jaw chuck depend on the diameter of the object held and is thus not predetermined.

In use, the fiducial pointers are used to align flat B with cursors on the interferometer's fringe monitor in order to ensure the correct lateral translations in positional combinations ABS1 and ABS2.

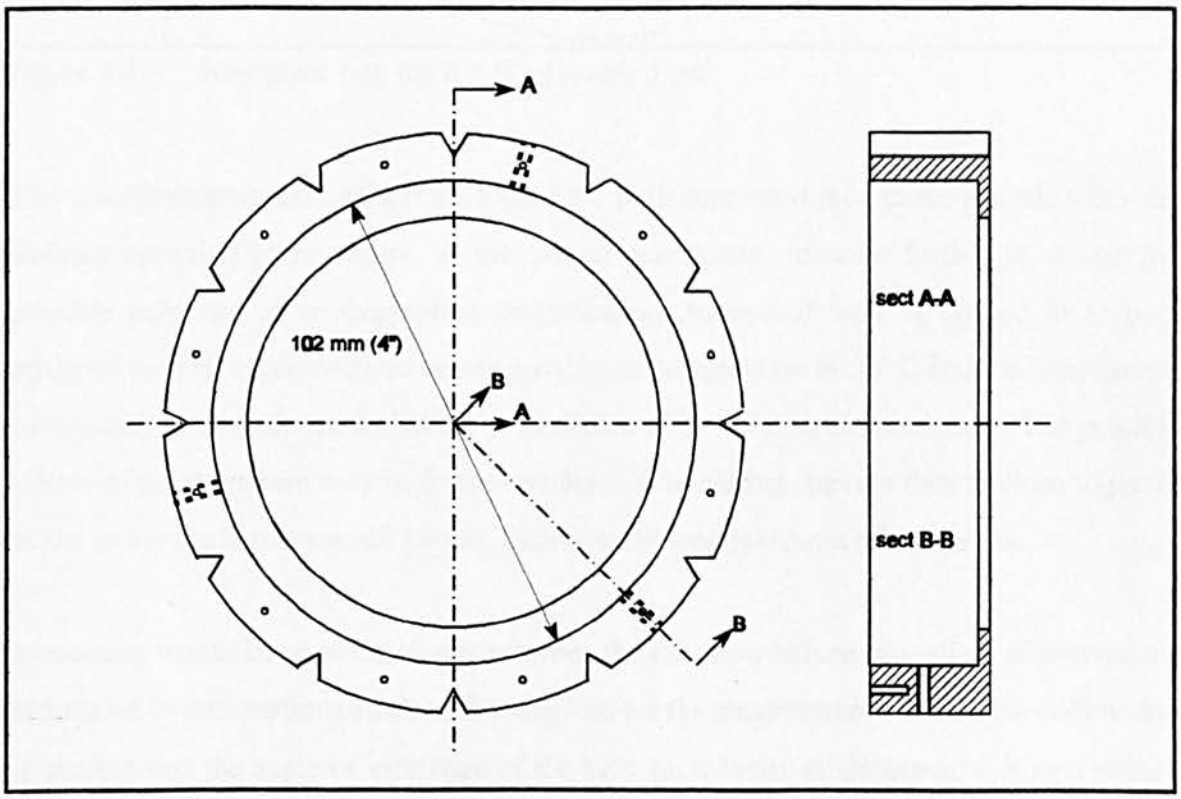


Figure 3.3 Cell for mounting flat B.

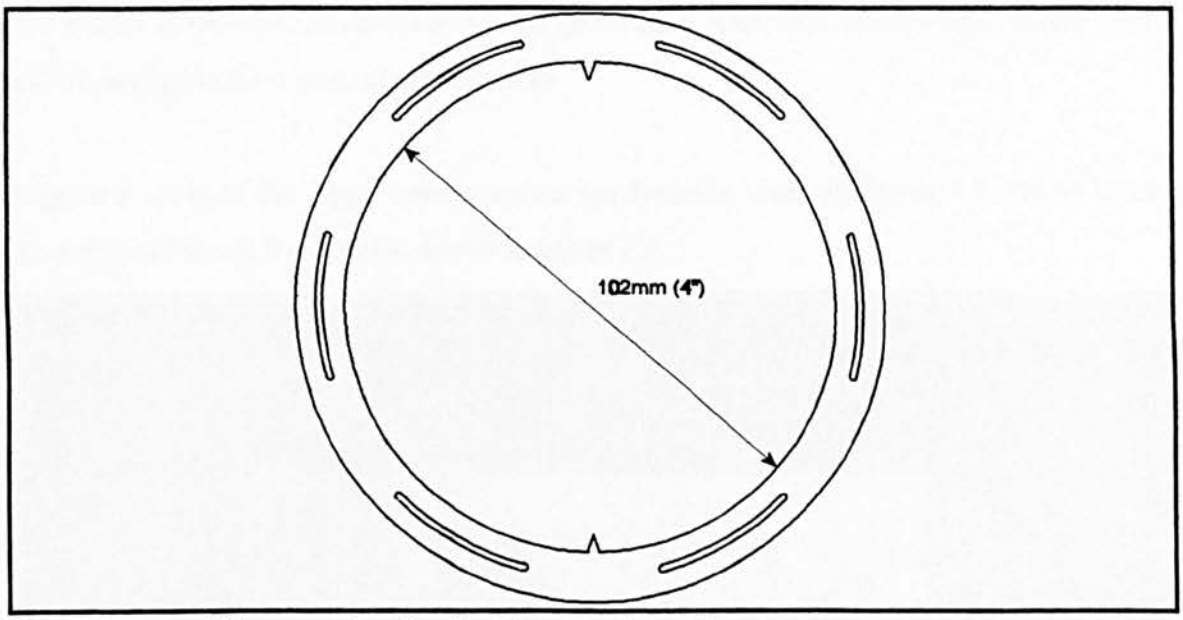


Figure 3.4 Alignment ring for flat B's mounting cell

The interferometer and test surface mount are both supported on a pneumatically vibration isolated optical table to ensure stable mutual positioning. In order further to reduce the possible influence of environmental disturbances, the optical table is housed in a room equipped with air conditioning to ensure a stable air temperature of 20°C and the interference cavity may be surrounded by baffles to minimize effects due to air movement. The possible effects of air movement may be further minimized by placing the two flats as close together as the mounting hardware will permit, thus reducing the thickness of the air gap.

Minimizing the thickness of the air gap between the flats also reduces the effect of aberrations introduced by the interferometer's collimating lens on the measurement. Where the collimation is perfect and the angle of incidence of the light on a cavity of thickness, t, is everywhere normal, the OPD between test and reference wavefronts is 2t. Where the collimation is imperfect and the illumination angle varies from normal by $\delta\theta$, the OPD is $2t/\cos\delta\theta$. The OPD thus varies over the field of view by $2t(1-1/\cos\delta\theta)$ even when t is constant, thus introducing an error into the measurement. Since the magnitude of this error is proportional to t, it is desirable to keep t as small as is practicable. Reducing t also minimizes the effect of aberrations introduced into the test wavefront by errors in the back surface of, or inhomogeneities in, the reference flat by the same argument. When t is zero, aberrations introduced by the collimator

or substrate of the reference flat only cause a geometrical distortion of the fringe pattern which will be negligible for a well adjusted system.

A general view of the Zygo interferometer hardware is given in figure 3.5. Detail of the mounting hardware for flat B is shown in figure 3.6.

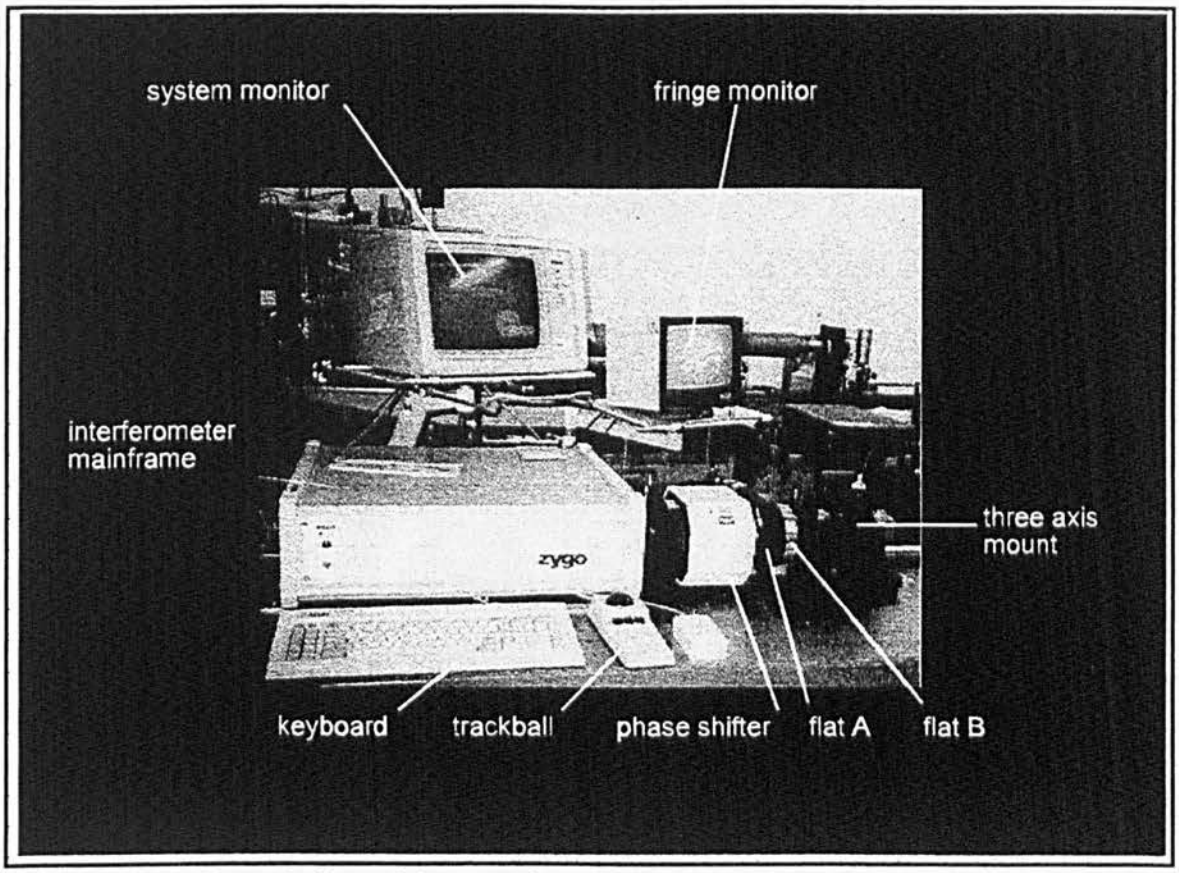


Figure 3.5 General view of the Zygo interferometer hardware.

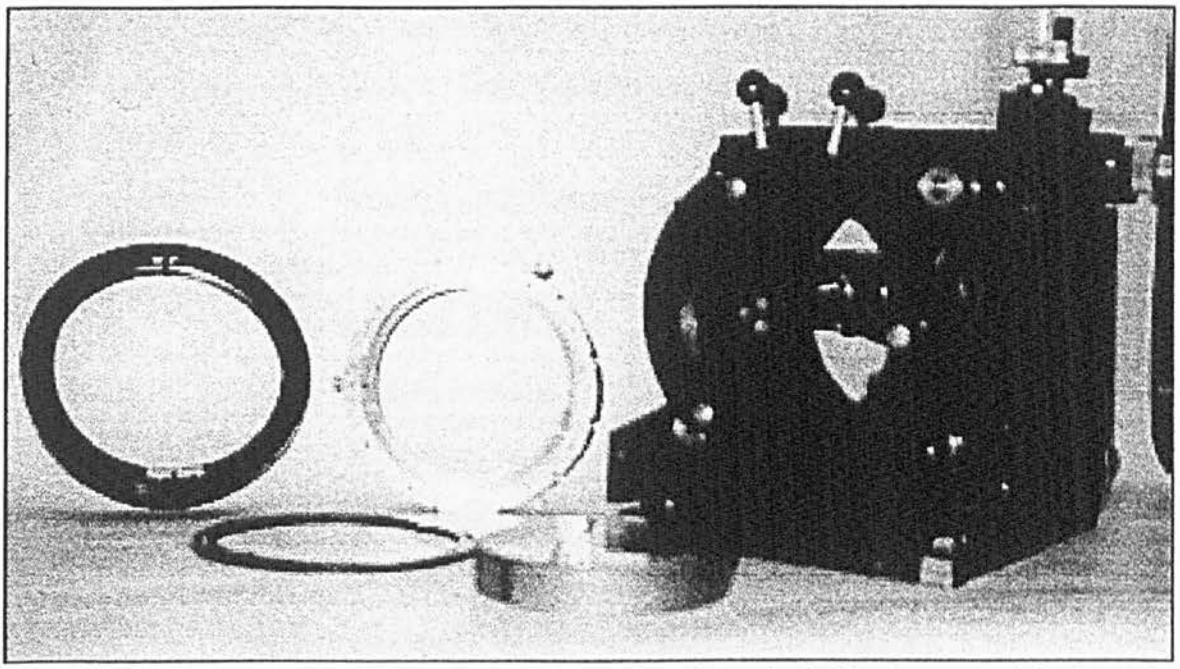

Figure 3.6 Detail of flat mounting hardware showing, clockwise from top left: flat A in Zygo bayonet mount, mounting cell for flat B, Zygo three axis mount, flat B unmounted and alignment ring for flat B.

## 3.2 The Zygo Interferometry Programming Language (ZIPL).

The Zygo Mark IV interferometer system is supplied with standard software for performing common relative measurement functions. Where the standard user interface software does not fulfill the requirements for the analysis of a specialist measurement technique, a programming language is supplied to allow the user to write customized processing software. The proprietary programming language is called the Zygo Interferometry Programming Language (ZIPL).

ZIPL is an interpreted computer language very much like BASIC. Most of the common BASIC commands (gosub, for..to..next etc.) are common to ZIPL. In addition ZIPL includes many commands that are specific to the control of, and analysis of data in, a Zygo interferometer system. A brief description of those specialized ZIPL commands used in the flatness analysis software are described in the table below. More detailed descriptions of these commands and those not described may be found in the ZIPL user manual. One useful feature of most interpreted BASIC languages that is missing from ZIPL is a renumbering facility. This

means that it is important, when writing ZIPL code, to leave plenty of unused line numbers between lines of code to allow for later additions and corrections.

Functions of Specialized ZIPL Commands

| Command | Function |
|---|---|
| acquire | Acquire light level data from interferometer. |
| average | Average multiple datasets. |
| clr | Remove user defined variables from memory. |
| connect | Convert unconnected phases to connected phases. |
| convert | Convert light level data to phase data. |
| coordinates | Specify pixel coordinate system (squared or rectangular). |
| cursor | Set up cursors on fringe monitor. |
| fmat | Sets fields for display of numerical information. |
| initialize | Initialize a disk or volume U. |
| invert | Invert the data. Used for Mark IV interferometer. The processor unit may also be used with the Maxim profiler which does not use this command. |
| isometric | Display data in isometric plot format. |
| mask | Restrict area of operation. |
| multiply | Multiplies current dataset by a factor. |
| printer is | Selects the print device. 1=monitor, 2=printer. |
| scroll | Put the graphics terminal into scroll mode. |
| square | Convert the data into squared format. The interferometer has a rectangular CCD array. This command makes the data look like it came from a square array. |
| subtract | Subtract dataset in user volume from current data. |
| units | Specify units for parameter input or analysis output. |
| window | Specify data area of interest. |
| zgen | Generate a surface from the ZGEN system variables. |

| | |
|---|---|
| zremove | Subtract a surface defined by the ZGEN system variables from the current data set. |

In addition to featuring special commands for interferometry, ZIPL defines a number of System Variables which are distinguished by being written in capital letters. The table below describes the significance of those ZIPL System Variable used in the flatness analysis software.

Significance of System Variables

| System Variable | Significance |
|---|---|
| CURSOR[n][m] | The coordinate extents of cursor n. |
| DATA | The name of the array into which data sets are loaded in order for them to be operated upon. |
| DATAPOINT[x][y] | The value of an element of DATA at coordinates (x,y). |
| ERROR | Query error number. |
| ERROR$ | Query text description of error. |
| TIMER | Set millisecond timer. Query current value. |
| ZGEN[n] | Access value of $n^{th}$ Zernike coefficient.ZGEN[36],[37],[38] are x and y centre coordinates of the function and radius of the function respectively. |

The storage of data in ZIPL is in one of four areas known as "volumes". These volumes are named A, B, M and U. The A and B storage volumes are the two floppy disk drives, each with a capacity of 720 Kbytes. There is no hard disk drive. Volumes M and U are the "main" and "user" volumes respectively and occupy RAM of which there is a total of 2 MBytes. It is unclear how much of the 2 MBytes is available to volumes M and U since the operating system and program code occupy some of this space. Obviously, to maximize program execution speed it is desirable that the data being processed is stored in RAM but when several large datasets are in use there is insufficient RAM available so the floppy disk volumes, A and B, must be used to store some data during program execution as will be seen in the description of the flatness testing software below. A feature of ZIPL code is that elements of measured datasets may only be accessed when that dataset is loaded into the system variable array,

DATA. This is a major constraint which dominates the structure of the code, as will be seen.

### 3.3 Development of ZIPL code to implement the flatness testing algorithm.

The overall structure of the program code closely parallels the development of the flatness testing algorithm as described in chapter 2. The complete ZIPL program listing is included in this thesis as Appendix B. Important sections of the code are described in some detail below where this helps to demonstrate the functions being performed.

The program is structured from a group of subroutines. The main program body is contained in line numbers 10 to 999.

```
10    rem REFERENCE FLAT CALIBRATION PROGRAM
20    rem BY JOHN MITCHELL (KINGSTON UNIVERSITY)
25    on error gosub @_aftermath
30    gosub @_init
40    gosub @_cursors
50    gosub @_acq_cb
60    gosub @_acq_ac
70    gosub @_acq_abr
80    gosub @_acq_ab
90    gosub @_acq_abs1
100   gosub @_acq_abs2
110   gosub @_make_blanks
120   gosub @_do_fff
130   gosub @_do_schulz
140   gosub @_adj_abs
150   gosub @_solve
900   ? ERROR: ? ERROR$
999   end
```

Lines 30 and 40 call subroutines to initialize constants and cursors etc. for later use by the program. Lines 50 to 110 call subroutines which handle the measurement and acquisition of the various datasets from the different positional combinations used in the absolute flatness testing algorithm. Lines 120 to 150 call subroutines which perform the major sections of the data processing to derive the absolute contours of the flats from the measured datasets. Subroutine @_do_fff performs the basic three flat test to find the profile of the first diameters on each flat. @_do schulz finds the profiles of the diameters orthogonal to those already found. @_adj_abs corrects datasets ABS1 and ABS2 for piston and tilt. Finally, @_solve finds the solution for the contours of the flats over their entire surfaces.

In order to facilitate easy understanding and reading of the program code, the first line number and a brief description of the function of each subroutine is given in the table below.

Position and function of the program subroutines

| Subroutine name | Line number | Function |
|---|---|---|
| @_acq_ab | 6000 | Acquisition of AB dataset. |
| @_acq_abr | 5000 | Acquisition of ABR dataset. |
| @_acq_abs1 | 7000 | Acquisition of ABS1 dataset. |
| @_acq_abs2 | 8000 | Acquisition of ABS2 dataset. |
| @_acq_ac | 4000 | Acquisition of AC dataset. |
| @_acq_cb | 3000 | Acquisition of CB dataset. |
| @_acq_data | 10000 | Acquisition of datasets from interferometer. |
| @_adj_abr | 18000 | Correction of piston and tilt in ABR dataset . |
| @_adj_abs | 23000 | Correction of piston and tilt for ABS1 and ABS2 datasets. |
| @_adj_data1 | 11000 | Correction of piston and tilt for AB, AC and CB datasets. |
| @_adj-xtilt | 24000 | Correction of piston and x-tilt for ABS1 and ABS2 datasets. |
| @_aftermath | 32000 | Provides diagnostic data in the event of a fatal program error. |

| | | |
|---|---|---|
| @_chord_print | 37000 | Prints the data for a horizontal chord for diagnostic purposes. |
| @_cursors | 2000 | Sets up cursors to define the extent of acquired data and fringe monitor cursors to aid positioning of the flats. |
| @_delay | 22000 | Delays program execution before data acquisition to allow settling of fringe pattern. |
| @_diagnost | 31000 | Provides diagnostic data at various stages of program execution. |
| @_do_fff | 15000 | Performs standard three flat test on AB, AC and CB datasets to find the profiles of the vertical diameters. |
| @_do_schulz | 17000 | Performs Schulz's method to find profiles of horizontal diameters. |
| @_find_plane | 12000 | Determines the coefficients of the plane passing through three points. |
| @_find_ytilt | 25000 | Determines and removes the y-tilt error in the ABS1 and ABS2 datasets. |
| @_init | 1000 | Sets the value of constants used during program execution and allows the operator to choose whether to acquire new data, use old data or synthesize data for test purposes. |
| @_init_temp | 16000 | Initializes two arrays for the temporary storage of data. |
| @_make_blanks | 14000 | Creates blank datasets in which to write the results of processing. |
| @_minmax | 21000 | Determines the extent and centres of the datasets from cursor data. |
| @_new | 20000 | Prepares for the acquisition of new data. |
| @_old | 19000 | Prepares for the use of old data. |
| @_remove_plane | 13000 | Removes a plane from the current dataset given the coefficients of that plane. |
| @_results_disk | 35000 | Prepares a floppy disk for storage of the final results of the program. |

| | | |
|---|---|---|
| @_solve | 30000 | Determines the contours of the whole of flats A and B once all data has been corrected. |
| @_step_chords1 | 26000 | Determines the profiles of additional chords on A and B starting from a known chord on flat A. |
| @_step_chords2 | 27000 | Determines the profiles of additional chords on A and B starting from a known chord on flat B. |
| @_sub_temp1a | 28000 | Routine used by @_step_chords1 (performs equation 2.26). |
| @_sub_temp1b | 28500 | Routine used by @_step_chords2 (performs equation 2.25). |
| @_sub_temp2a | 29000 | Routine used by @_step_chords1 (performs equation 2.27). |
| @_sub_temp2b | 29500 | Routine used by @_step_chords2 (performs equation 2.28). |
| @_synth | 34000 | Synthesizes datasets for the purpose of testing the operation of the program. |
| @_temp_diag | 33000 | Prints the contents of arrays temp1 and temp2 for diagnostic purposes. |
| @_zgen_clear | 36000 | Sets the values of Zernike coefficients ZGEN[0] to ZGEN[35] equal to zero. |

A flow chart showing the structure of the first section of the program, dealing with initialization and data acquisition is shown in figure 3.7.

First, constants for later use by the program are set. Constants, s1 and s2, are the pixel shifts for positional combinations ABS1 and ABS2, respectively. Constant, bad= 99999.0, is the value given by the processor to a measured pixel when the phase of the wavefront at that pixel cannot be correctly determined. This can be either because of an optical defect at that point such as a speck of dust on one of the surfaces, or because that point is outside the aperture of the interference cavity formed by the two surfaces. At frequent points during the program, it is necessary to check whether a point is bad since otherwise a bad point may then be converted

by a mathematical operation to an erroneous value that would not be recognized as such by the processor. The constant, multfact, is used to increase the dynamic range of the calculations in the processing. The significance of this constant will be described later.

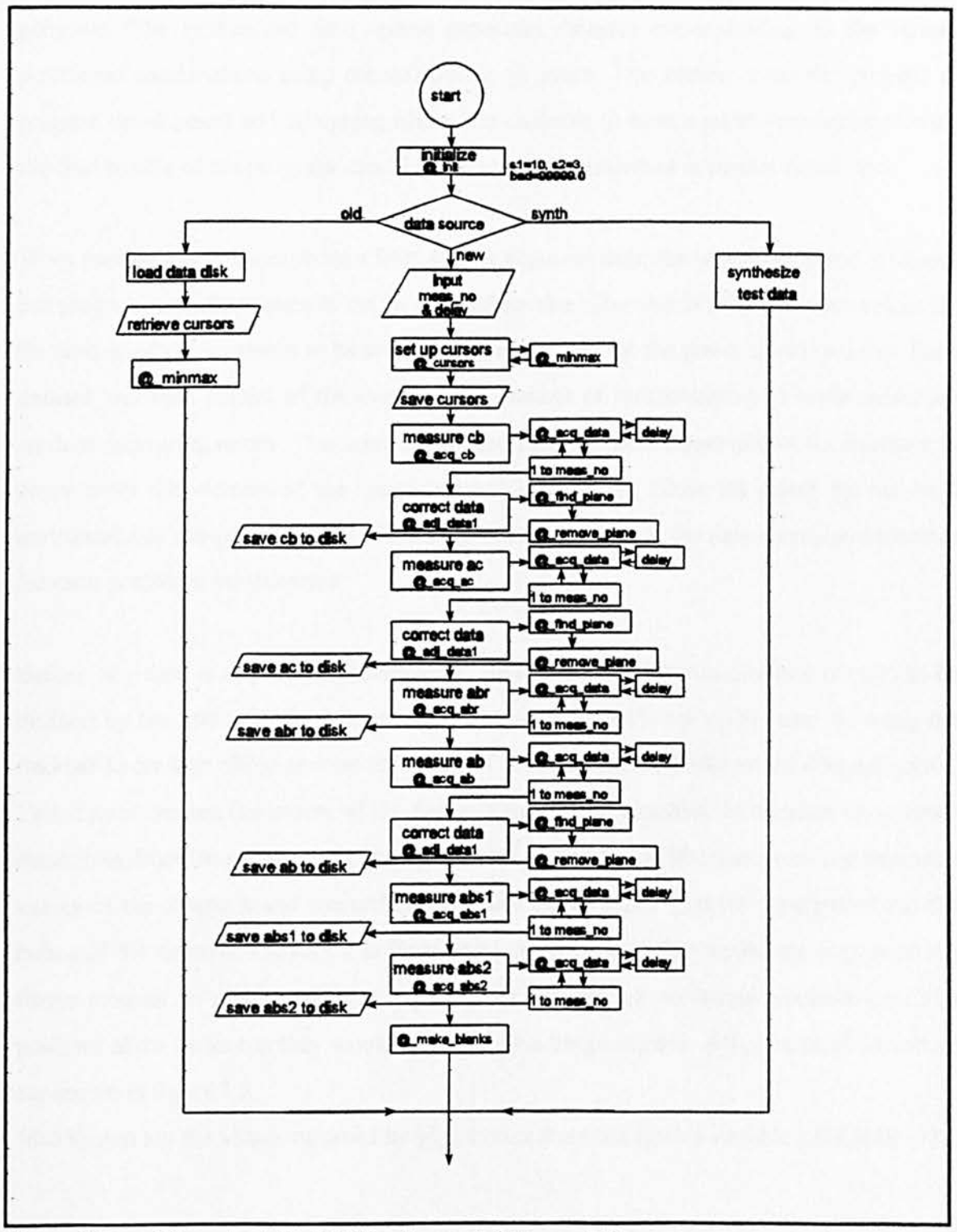


Figure 3.7 Flow chart for initialization and data acquisition.

At this point the user is asked to decide upon the source of the data to be processed by the program. If the user chooses to process old data or synthesized data, the section of the program devoted to the acquisition of new data is skipped. If old data is chosen, then the user must supply a floppy disk containing measured data acquired during a prevoius run of the program. The synthesized data option generates datasets corresponding to the various positional combinations using the subroutine @_synth. This option is for the purpose of program development and debugging where it is desirable to have a priori knowledge of what the final results of the program should be. This will be described in further detail later.

When making a new measurement from freshly acquired data, the new data option is chosen and program execution passes to the @_new subroutine. The user is asked to input values for the number of measurements to be averaged, meas_no, and for the pre-acquisition delay. Each dataset will then consist of the average of a number of measurements in order to reduce random measuring errors. The delay introduced before measurement allows the operator to retire from the vicinity of the interferometer in order to allow his effect on the local environment to dissipate. Control of the program then passes to the data acquisition routines for each positional combination.

Before any data is acquired, subroutine @_cursors allows the measurement cursors to be defined by the user. Cursor 1 is an elliptical cursor which is set by the user by using the trackball to draw an ellipse around the image of the interference cavity on the fringe monitor. This cursor defines the extent of the data which is to be acquired. Subroutine @_minmax determines, from the system variables for cursor 1, CURSOR [1]the maximum and minimum values of the dataset x and y coordinates as well as the x and y centre coordinates and the radius of the datasets. Cursors 2 to 9 are point cursors that, when turned on, appear on the fringe monitor to provide alignment guides for the various positional combinations. The positions of the cursors as they would appear on the fringe monitor if they were all turned on are shown in figure 3.8.
Also shown are the values returned by @_minmax from the system variable CURSOR [1].

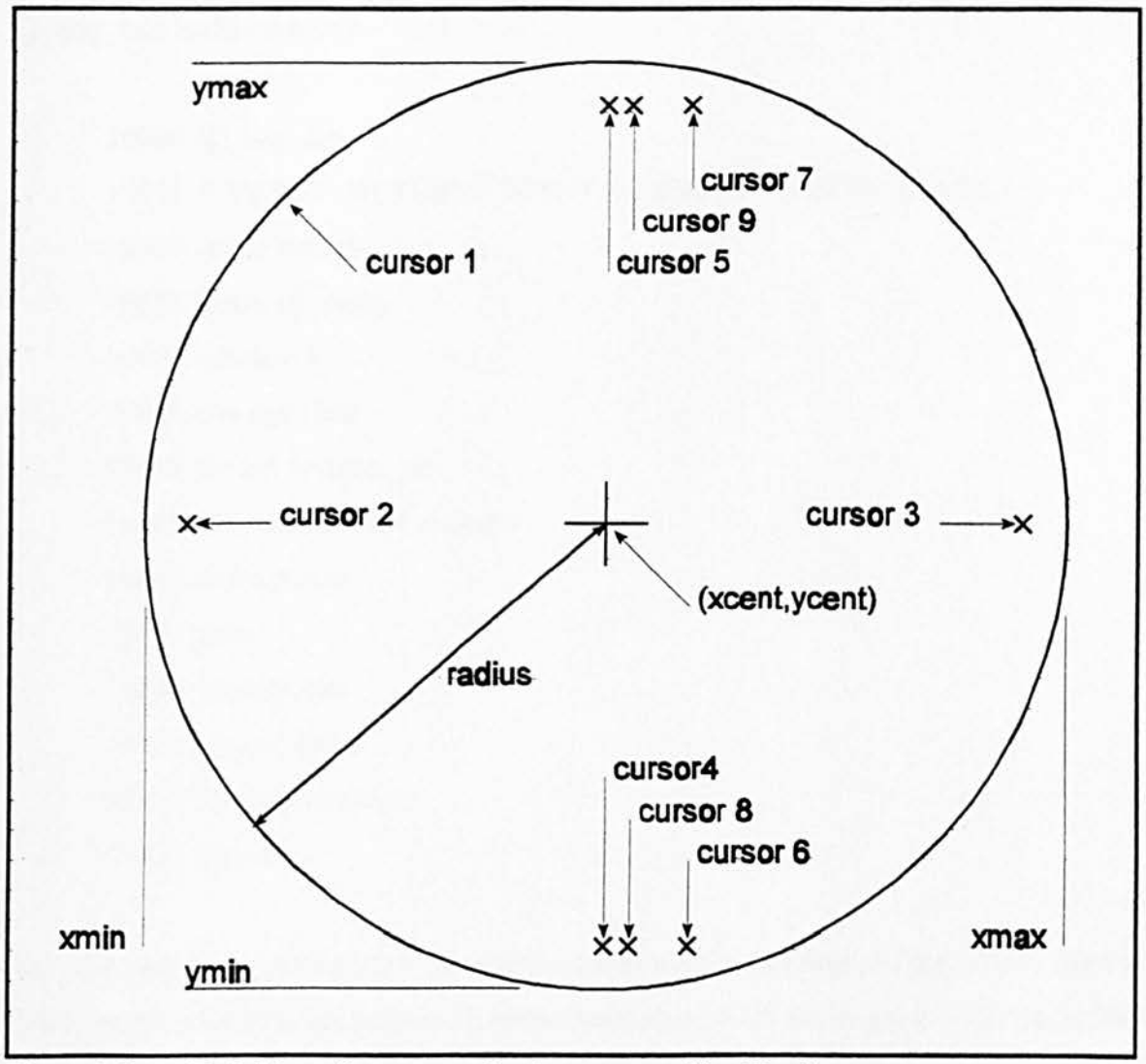


Figure 3.8 Definition of cursors and values returned by @_minmax.

The order in which the various positional combinations is measured is chosen to minimize handling of the three flats. For example, after measuring the first combination, CB, the flat mounted in the interferometer is always flat A which need not be touched from that point onwards.

The routines, @_acq_cb to @_acq_abs2, which handle the acquisition of the datasets for the various positional combinations each prompt the user to position the appropriate flats in the interferometer and in the test flat mount with respect to the appropriate alignment cursors which are turned on for the positional combination which is to be measured. The subroutine, @_acq_data, is then called which handles the actual data acquisition. The subroutine,

@_acq_data is shown below.

```
10000 @_acq_data:
10010 ? "PRESS <RETURN> WHEN READY TO ACQUIRE DATA."
10020  input ready$
10025 gosub @_delay
10030 window 1
10040 average clear
10050 for i=1 to meas_no
10060 acquire : convert: connect
10070 average sum
10080 next i
10090 average calc
10100 square: invert
10105 multiply multfact
10110 return
```

Once the user has confirmed that the interferometer is correctly aligned (line 10020) there is a delay set by subroutine @_delay to allow for environmental effects to settle. The data is then acquired meas_no times in order to minimize random errors in a loop (lines 10050 to 10080). The command, "acquire", acquires the fringe pattern intensity data from the interferometer's CCD camera. The command "convert" then converts the intensity data to phase data using a five frame phase measuring algorithm (section 1.2.6.2.5). "Connect" then performs a phase unwrapping algorithm to remove phase discontinuities from the phase data (section 1.2.6.4). After the measurement loop, the measurements are averaged (line 10070). The CCD camera in the Mark IV interferometer has a rectangular pixel array but a square grid of data is more useful. The command "square" converts the rectangular array of measured data into a square array. The processor in the Mark IV interferometer system is capable of performing control and analysis functions for a variety of Zygo interferometric instruments (for example, the Zygo Maxim interferometric profiler). Depending on the type of interferometer (the Maxim profiler is a Mirau-type interferometer), the phase data calculated by the acquire/ convert/ connect

commands may be the negative of their true values. This is the case with the Zygo Mark IV Fizeau interferometer and so the command "invert" inverts the calculated dataset to give the correct values.

Once the phase data have been acquired the datasets for positional combinations AB, CB and AC are adjusted for piston and tilt as described in section 2.3.1. This is accomplished by subroutine @_adj_data1 which picks out the three points at which the nominal plane for the positional combination is defined. Subroutine @_find_plane determines the coefficients of the plane to be removed and subroutine @_remove _plane subtracts that plane from the current dataset. In the first implementation of @_remove _plane, the plane coefficients were used to calculate the z coordinate values of the plane which were then subtracted from the dataset a point at a time. This process was found to be exceedingly slow, taking some 20 minutes to execute for each dataset. Subroutine @_remove_plane was rewritten so that the plane to be removed was described in terms of the first three Zernike polynomials via the system variables ZGEN [0] , ZGEN [1] and ZGEN [2]. System variables ZGEN [36], ZGEN [37] and ZGEN [38] define the centre and radius of the surface defined by the Zernike polynomials. The command zremove 3 subtracts the plane so described from the current dataset. The execution of the rewritten subroutine was considerably faster, taking only seconds to execute. It appears that the ZIPL language is not optimized for operations that interrogate the DATA system variable on a point wise basis via the DATAPOINT system variable. This feature has a major impact on the processing speed of later portions of the program where there is no choice but to operate on the data on a pointwise basis.

```
13000 @_remove_plane: ? "REMOVING PLANE"
13010 rem REMOVES A PLANE FROM THE DATASET GIVEN THE PLANE
13020 rem COEFFICIENTS I,J,K AND THE POINT z0,y0,z0
13030 ZGEN [0]=(I*x0+J*y0)/K +z0: ? ZGEN [0]
13040 ZGEN [1]=-1*radius*I/K: ? ZGEN [1]
13050 ZGEN [2]=-1*radius*J/K: ? ZGEN [2]
13060 ZGEN [36]= xcent
13070 ZGEN [37]= ycent
13080 ZGEN [38]= radius
```

```
13090 zremove 3
13900 return
```

After the data for each positional combination have been acquired and, where appropriate, adjusted for piston and tilt, they are stored on floppy disk drive B. There is insufficient RAM to hold this many datasets in memory at once.

The next stage in the data processing is to find the absolute profile of the vertical diameters of flats A and B by the three flat method as described in section 2.3.2. This section of processing is accomplished by subroutine @_do_fff. A flow chart showing the sequence of processing is shown in figure 3.9.

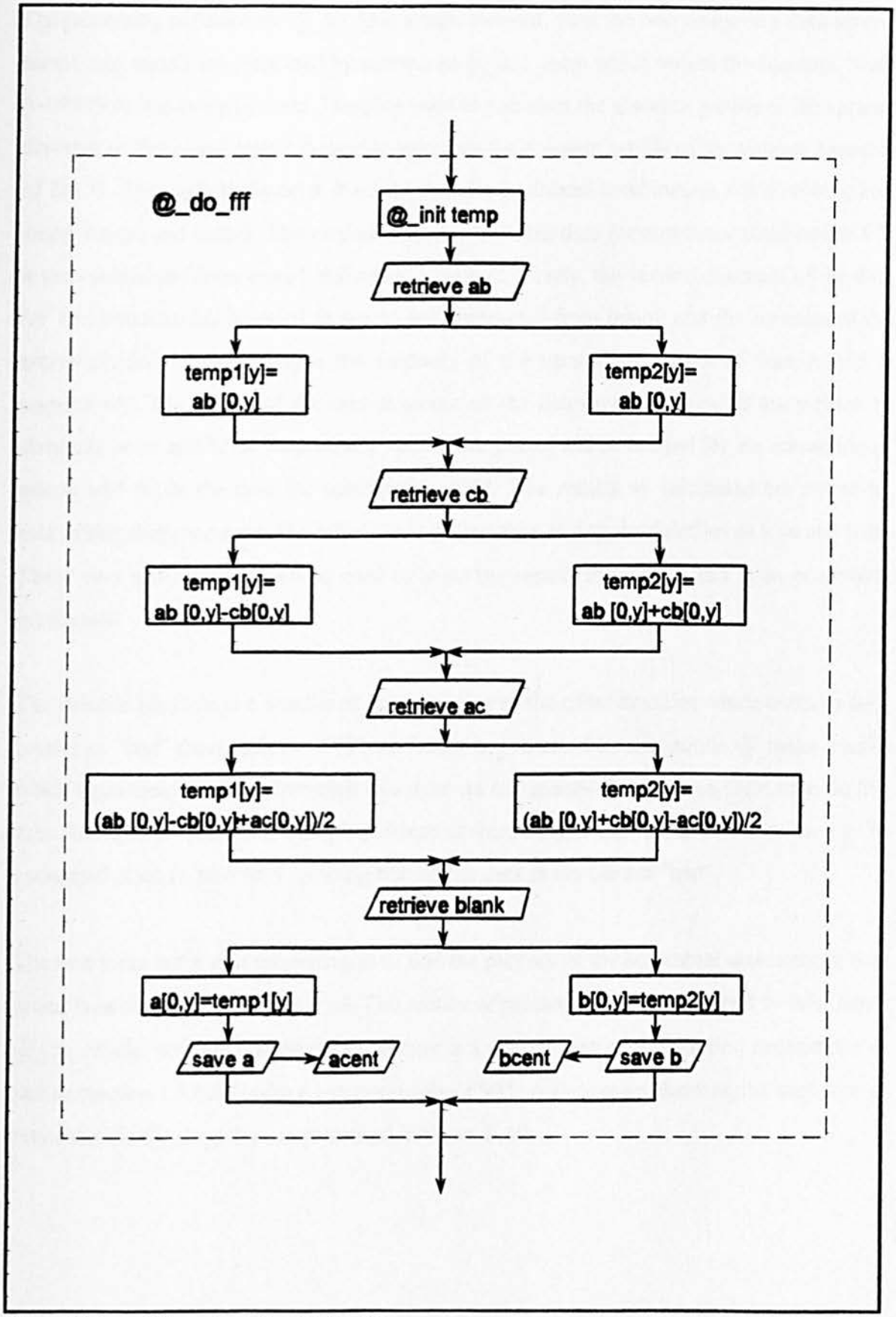


Figure 3.9 Flow chart for @_do_fff.

The processing sequence for @_do_fff is straightforward. First the two temporary data arrays, temp1 and temp2 are initialized by subroutine @_init_temp which writes the constant "bad" (= 99999.0) into every position. Temp1 is used to calculate the absolute profile of the vertical diameter of flat A and temp2 is used to calculate the absolute profile of the vertical diameter of flat B. The vertical diameter from the data for positional combination AB is written into arrays temp1 and temp2. The vertical diameter from the data for positional combination CB is then subtracted from temp1 and added to temp2. Finally, the vertical diameter of the data for combination AC is added to temp1 and subtracted from temp2 and the contents of the arrays divided by two to give the contours of the vertical diameters of flats A and B respectively. The values of the central points on the diameters of A and B are written to constants acent and bcent respectively since these points will be needed for the correction of piston and tilt in the data for combination ABR. The results so calculated are stored by overwriting them into a datafile called blank.da and then saving the datafiles as a.da and b.da. These two new datafiles will be used to store the results for flats A and B as processing progresses.

The datafile blank.da is a datafile of the same size as the other datafiles which contains only invalid or "bad" datapoints (= 99999.0). Blank.da is created by subroutine @_make_blanks which again uses Zernike polynomials as a short cut to generate the data in a short time. In this case, the system variable ZGEN[38], which corresponds to the radius of the surface to be generated is set to zero thus ensuring that all the data in the file are "bad".

The next stage in the data processing is to find the profiles of the horizontal diameters of flats A and B as discussed in section 2.3.3. This section of processing is accomplished by subroutine @_do_schulz, so named since the procedure is a special case of the rotation method due to Schulz (section 1.3.1.3.1) with a rotation angle of 90°. A flow chart showing the sequence of processing in @_do_schulz is presented in figure 3.10.

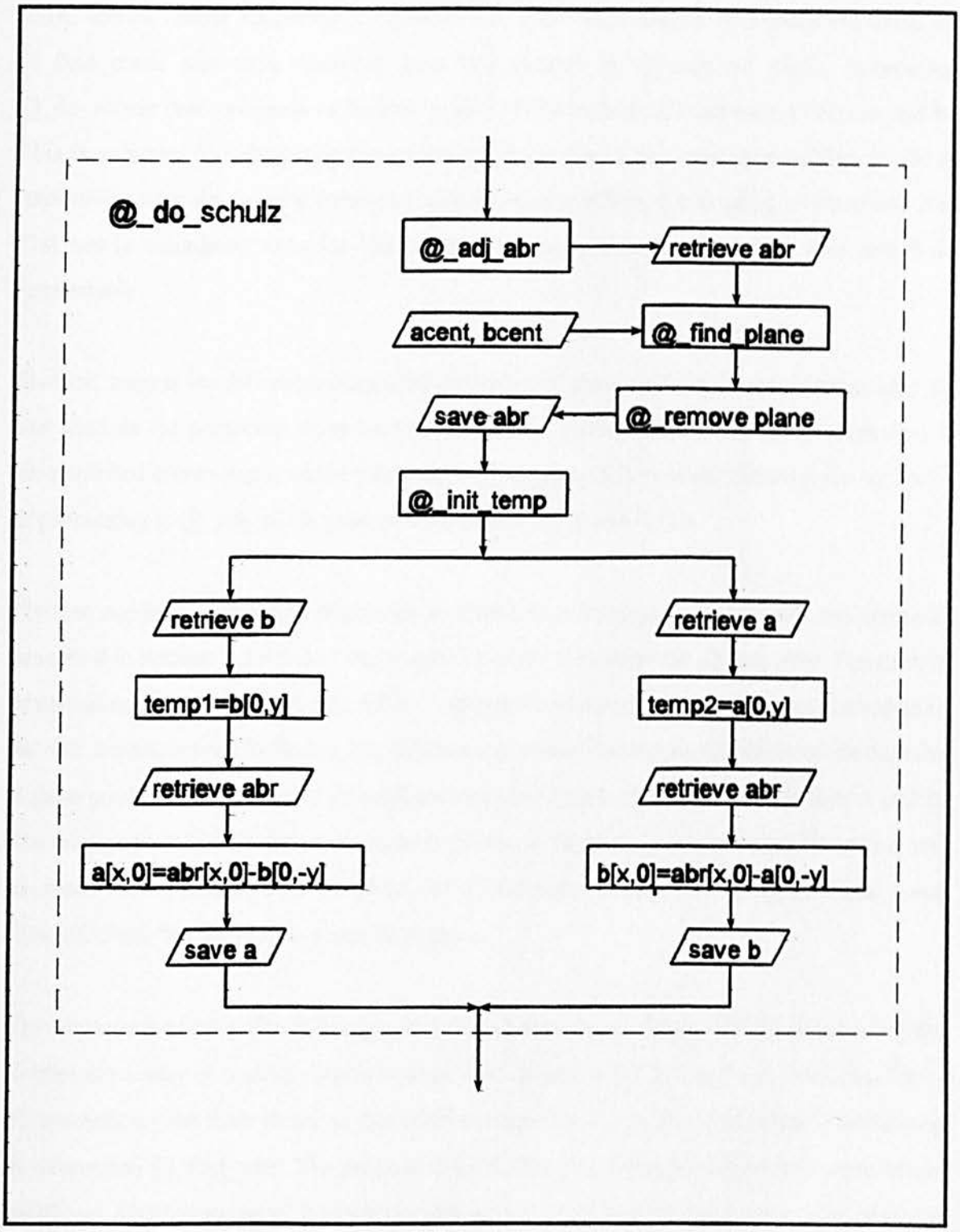


Figure 3.10 Flow chart for @_do_schulz.

The first stage of the processing in @_do_schulz is the correction of piston and tilt for positional combination ABR. This correction is accomplished by subroutine @_adj_abr which finds the three points through which the plane to be removed from the dataset for combination

ABR, abr.da, passes according to equation 2.5. The coefficients for this plane are found by @_find_plane and then removed from the dataset by @_remove plane. Subroutine @_do_schulz then proceeds to find the profile of the horizontal diameters of flats A and B. This is achieved by subtracting the known profile of the vertical diameters of flats B and A respectively from abr.da using temporary arrays temp1 and temp2 according to equations 2.6. The newly calculated data for flats A and B are written into datafiles a.da and b.da respectively.

The next stage in the data processing is the correction of piston and tilt for the datasets abs1.da and abs2.da for positional combinations ABS1 and ABS2 respectively. This correction is accomplished in two stages within subroutine @_adj_abs. A flow chart showing the sequence of processing in @_adj_abs is presented in figures 3.11a and 3.11b.

The first step in the correction of abs1.da and abs2.da is the adjustment of x-tilt and piston as described in section 2.3.4.1 and this is accomplished by subroutine @_adj_xtilt. For each of positional combinations ABS1 and ABS2, two points to define the plane to be subtracted from the data are determined by finding the difference between the measured values of the datasets at these points and the expected values from the known horizontal diameters of flats A and B. The third point needed to define the plane is chosen so that the y component of that plane will be zero. The coefficients of the planes to be subtracted are found by @_find_plane and removed from the data by @_remove_plane.

The corrections of the y-tilts in datasets abs1.da and abs2.da are determined by first finding the profiles of an array of parallel chords as described in section 2.3.4.2 and then deducing the y-tilt correction term from these, as described in section 2.3.4.3. This procedure is performed by subroutine @_find_ytilt. The procedure for finding the y-tilt correction for combinations ABS1 and ABS2 is the same. Accordingly, datasets abs1.da and abs2.da are, in turn, renamed abs.da and processed in an identical manner.

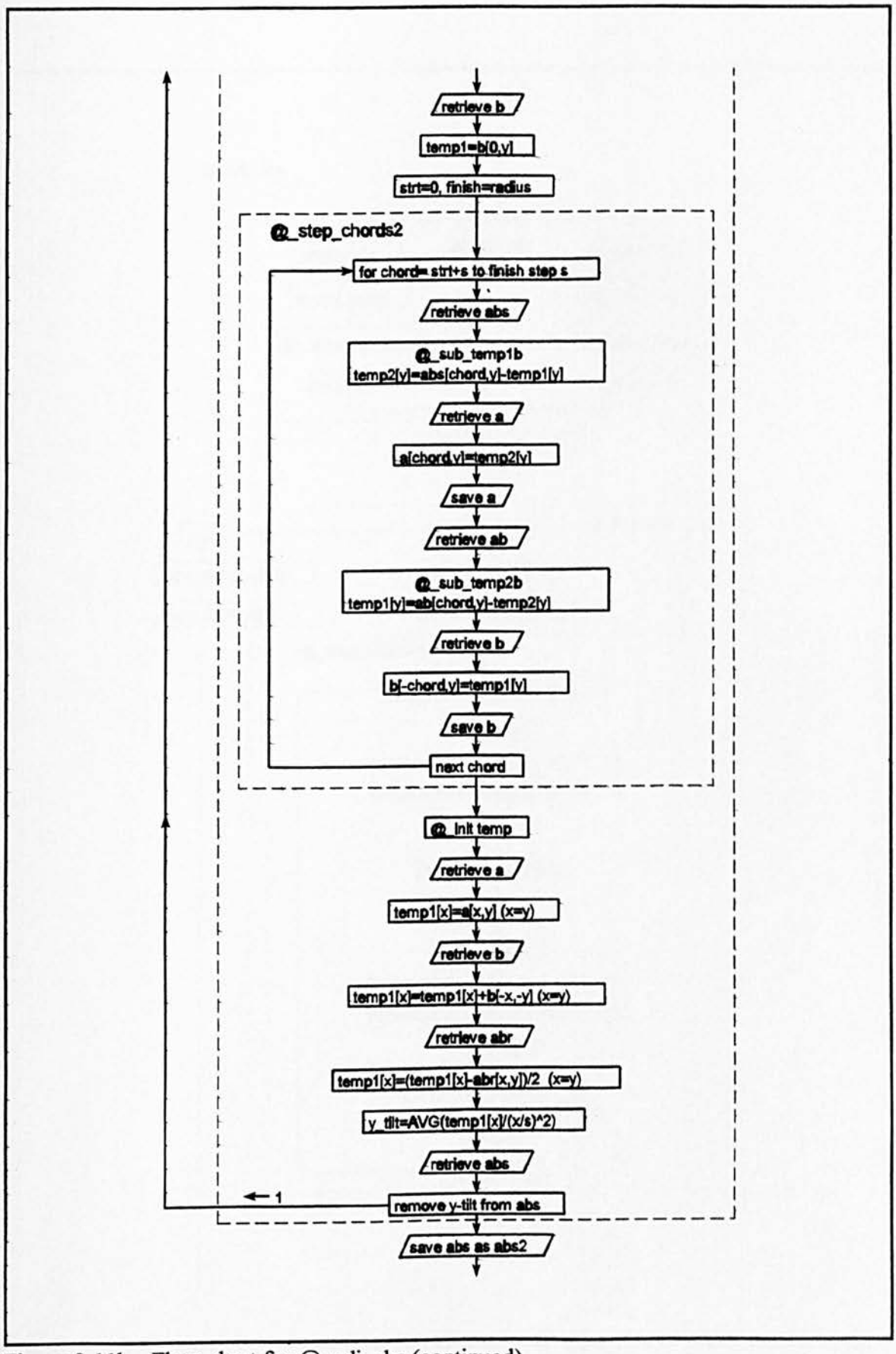


Figure 3.11b Flow chart for @_adj_abs (continued).

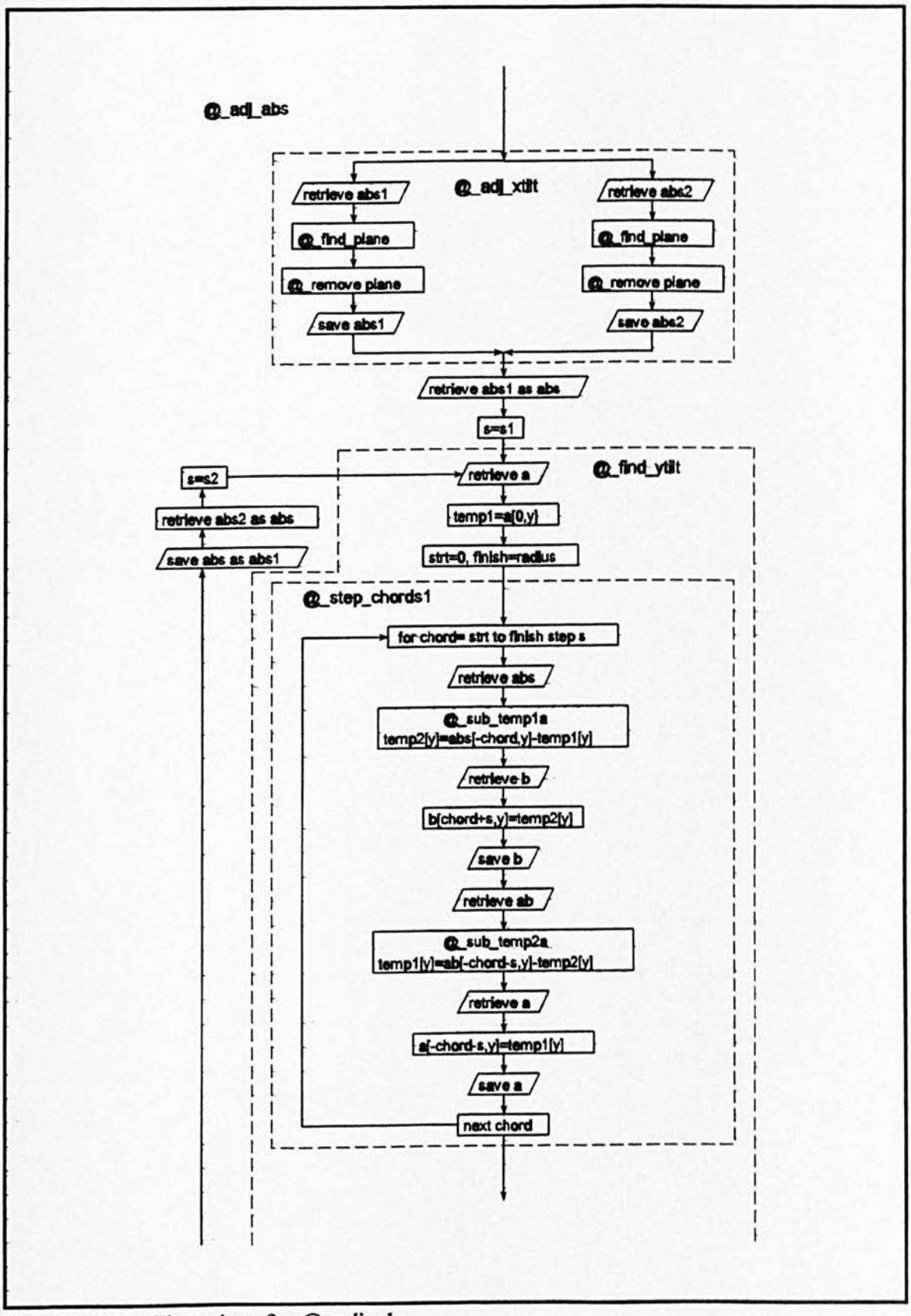


Figure 3.11a Flow chart for @_adj_abs.

The profiles of parallel chords are found by subroutines @_step_chords1 and @_step_chords2 which use, as a starting point, the previously known profile of a chord on flat A and flat B respectively. In this case, the previously known chords are the vertical diameters of the flats. Starting with a known chord on flat A, @_step_chords1 calls subroutine @_sub_temp1a which finds the profile of a new chord on flat B from the data for positional combination ABS (equation 2.26). The profile of this new chord on B is then used by subroutine @_sub_temp2a to find the profile of a new chord on flat A from the data for positional combination AB (Equation 2.27). @_sub_temp1a and @_sub_temp2a are performed, in turn, within a loop until the edge of the data is reached. Starting with a known chord on flat B, @_step_chords2 operates in a similar manner to find the profiles of new chords using subroutines @_sub_temp1b (Equation 2.25) and @_sub_temp2b (Equation 2.26).

Having found the profiles of a parallel array of chords on flats A and B, the y-tilt correction required for the ABS dataset is found according to the discussion in section 2.3.4.3 and equation 2.19. A mean value for the y-tilt correction term is found by solving equation 2.19 at each point where p=q. The y-tilt terms are then removed from the ABS datasets using the ZIPL "zremove" command.

Having corrected piston and tilt in the ABS1 and ABS2 datasets, the final stage in the data processing is to find the contours of the whole surface of flats A and B using both of the shifted datasets as described in section 2.5. This procedure is performed by subroutine @_solve for which a flow chart is shown in figure 3.12.

Initially, a set of parallel chords, spaced by distance s2, is found starting from the known profiles of the vertical diameters of flats A and B using subroutines @_step_chords1 and @_step_chords2 from the data for positional combination ABS2. These chords are then used, in turn, to find the entire set of chords covering flats A and B from the data for positional combination ABS2, again using subroutines @_step_chords1 and @_step_chords2. Once all the chords have been found, the complete data for flats A and B are stored on a floppy disk.

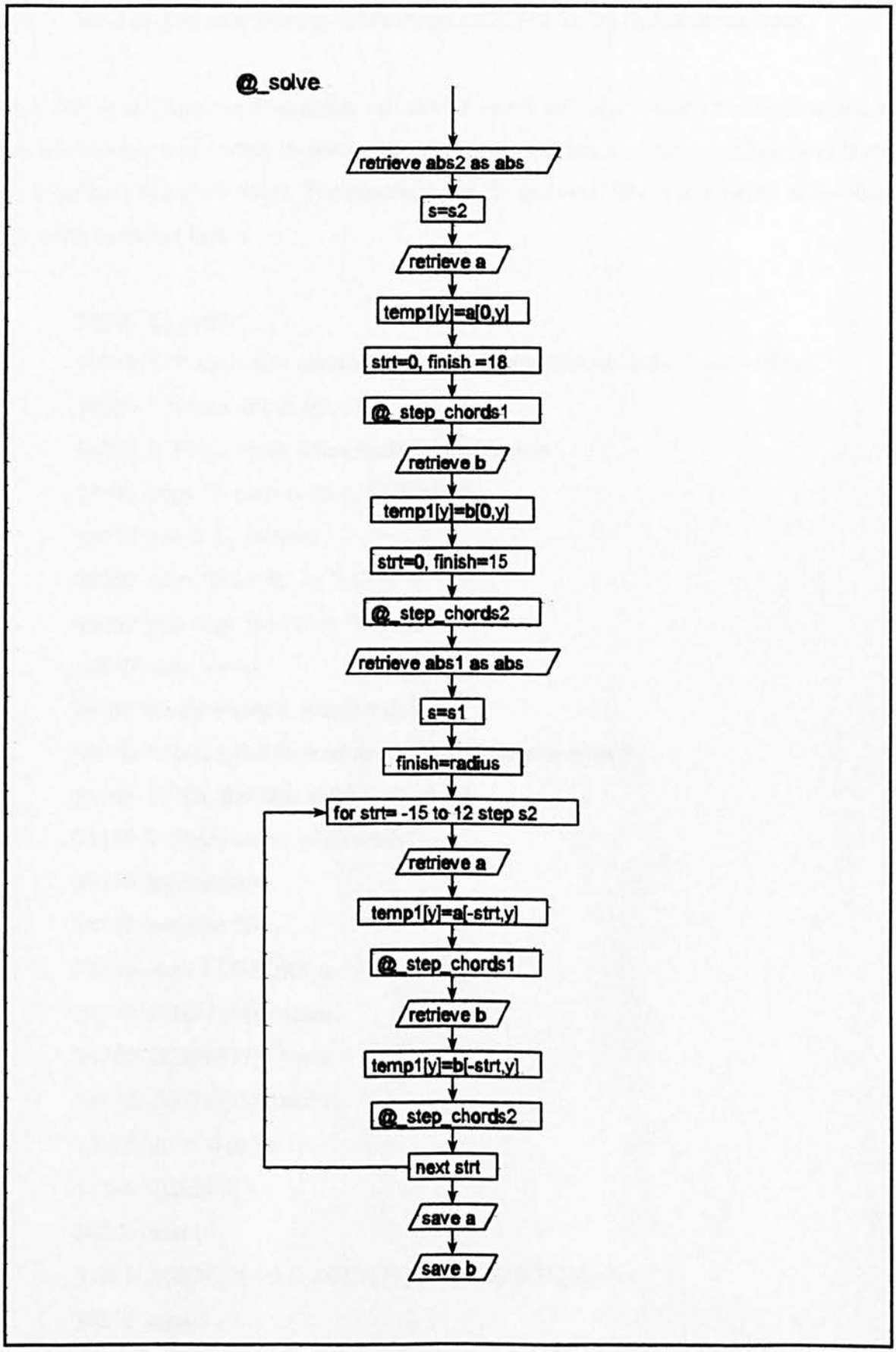


Figure 3.12 Flow chart for @_solve.

## 3.4 De-Bugging and Testing of the Program Code Using Synthesized Data.

In order to facilitate the de-bugging and testing of the absolute flatness testing program, a section of code was written to synthesize data for the various positional combinations from three surfaces of known shape. The generation of the synthetic data is handled by subroutine @_synth as shown below.

```
34000 @_synth:
34010 ? "Place a disk containing a datafile named ab.da and corresponding "
34020 ? "cursor file in drive B"
34025 ? "Press return when ready": input ready$
34030 copy "B:cursors.cu to CURSORS
34040 gosub @_minmax
34050 copy "B:ab.da" to DATA
34060 rem copy DATA to "U:A.da"
34070 units waves
34080 window calc 1: window data 1
34090 ? "place disk to receive synthesized data in drive B"
34100 ? "NB. this disk will be initialized
34110 ? "Press return when ready"
34115 input ready$
34120 initialise "B"
34130 copy CURSORS to "B:cursors.cu"
34150 ZGEN [36]= xcent
34160 ZGEN [37]= ycent
34170 ZGEN [38]= radius
34180 for i= 0 to 35
34190 ZGEN [i]= 0
34200 next i
34201 ZGEN [0]= 0.1: ZGEN [1]= 0.05: ZGEN [2]= 0.1
34202 zgen 5
```

```
34203 copy DATA to "U:A.da"
34210 rem AB
34215 units waves: gosub @_zgen_clear
34220 ZGEN [4]= 0.1
34230 zgen 5
34240 invert
34250 copy DATA to "U:X.da"
34260 copy "U:A.da" to DATA
34270 subtract "U:X.da"
34274 multiply multfact
34275 gosub @_adj_data1
34280 copy DATA to "B:ab.da"
34285 ? "AB": rows= 19: cols= 19: spacing= 10: gosub @_diagnost
34290 rem CB
34295 gosub @_zgen_clear
34296 units waves
34300 ZGEN [3]= 0.1
34320 zgen 5
34330 subtract "U:X.da"
34335 multiply multfact
34340 gosub @_adj_data1
34350 copy DATA to "B:cb.da"
34355 ? "CB": gosub @_diagnost
34360 rem AC
34365 gosub @_zgen_clear
34370 delete "U:X.da": delete "U:X.at"
34375 units waves
34380 ZGEN [3]= 0.1
34400 zgen 5
34410 invert
34420 copy DATA to "U:X.da"
```

```
34430 copy "U:A.da" to DATA
34440 subtract "U:X.da"
34445 multiply multfact
34450 gosub @_adj_data1
34460 copy DATA to "B:ac.da"
34465 ? "AC": gosub @_diagnost
34470 rem ABR
34475 gosub @_zgen_clear
34480 delete "U:X.da": delete "U:X.at"
34490 units waves
34500 ZGEN [4]= -0.1
34510 zgen 5
34520 invert
34530 copy DATA to "U:X.da"
34540 copy "U:A.da" to DATA
34550 subtract "U:X.da"
34555 multipy multfact
34560 copy DATA to "B:abr.da"
34565 ? "ABR": gosub @_diagnost
34570 rem ABS1
34580 delete "U:X.da": delete "U:X.at"
34590  ZGEN [36]= xcent+s1
34595 units waves
34600 ZGEN [4]= 0.1
34620 zgen 5
34630 invert
34640 copy DATA to "U:X.da"
34650 copy "U:A.da" to DATA
34660 subtract "U:X.da"
34665 multiply multfact
34670 copy DATA to "B:abs1.da"
```

```
34675 ? "ABS1": gosub @_diagnost
34680 rem ABS2
34690 delete "U:X.da": delete "U:X.at"
34700 ZGEN [36]= xcent+s2
34705 units waves
34710 ZGEN [4]= 0.1
34730 zgen 5
34740 invert
34750 copy DATA to "U:X.da"
34760 copy "U:A.da" to DATA
34770 subtract "U:X.da"
34775 multiply multfact
34780 copy DATA to "B:abs2.da"
34785 ? "ABS2": gosub @_diagnost
34790 for i= 0 to 38
34800 ZGEN [i]= 0
34810 next i
34820 initialize "U"
34830 gosub @_make_blanks
34999 return
```

Subroutine @_synth, as shown, generates each synthetic dataset from three synthetic surfaces corresponding to flats A, B and C which are each defined in terms of the Zernike polynomials via the ZIPL ZGEN system variables. The extent of the surfaces are defined in terms of the CURSOR system variable recovered from a previously saved cursor file from an actual surface measurement.

For the case of @_synth as shown, flat A is defined as a perfectly flat surface characterized by 0.1 waves of piston (ZGEN [0]= 0.1), 0.05 waves of x-tilt (ZGEN [1]= 0.05) and 0.1 waves of y-tilt (ZGEN [2]= 0.1). Flat B is defined as having 0.1 waves of 0° astigmatism (ZGEN [4]= 0.1). Flat C is defined as a spherical surface with 0.1 waves of defocus error (ZGEN [3]=

0.1).

To generate the synthetic datasets for each positional combination, the second surface for that combination is generated in the DATA system variable, inverted and then stored in the user volume of memory as "U:X.da". The first surface for the combination is then generated in DATA and "U:X.da" is subtracted from it. It is necessary to use the invert then subtract sequence since ZIPL does not have a command for adding two datasets. For the rotated positional combination, ABR, flat B is described by -0.1 waves of 0° astigmatism since rotating a purely astigmatic surface by 90° is equivalent to inverting it. For the shifted positional combinations, ABS1 and ABS2, the centre of the dataset describing flat B is shifted by adding s1 and s2 respectively to ZGEN [36] which defines the central x coordinate of the dataset.

Before being stored the datasets for each positional combination are multiplied by factor multfact. This factor was introduced to extend the dynamic range of the calculations as will be explained later. For the time being, factor multfact should be considered to be absent or equal to one.

As an aid to diagnosing bugs in the program, three subroutines, @_temp_diag, @_diagnost and @_chord_print, were written to provide diagnostic output at various points during processing. When called, @_temp_diag outputs the contents of the temporary data arrays, temp1 and temp2. When called, @_diagnost outputs a sample of the current contents of DATA on a square grid, the spacing of which is defined before the routine is called. When called, @_chord_print prints the values of a selected horizontal chord from DATA.

In order to illustrate the use of the diagnostic output of these sections of code, the diagnosis of the last major bug to be found in the program is described below.

The program was run using the synthesized data option to construct simulated datasets as described above. The program run time is very long, approximately five hours, considering its complexity. This is probably due to the slow access to elements of the DATA system variable

as described above. This has a dramatic effect on the speed at which the program runs since the algorithm requires the data to be processed on a point-wise basis in a number of long program loops. The long run time means that it is impractical to stop the program during its execution to check on the progress of the data processing and it is necessary to rely on the output of the diagnostic routines at pre-chosen key points in the data processing scheme.

Three dimensional plots of the final results of the program for flats A and B are shown in figures 3.13 and 3.14. The plots were obtained by loading the result datasets a.da and b.da into DATA and using the ZIPL "isometric" command to produce the plot.

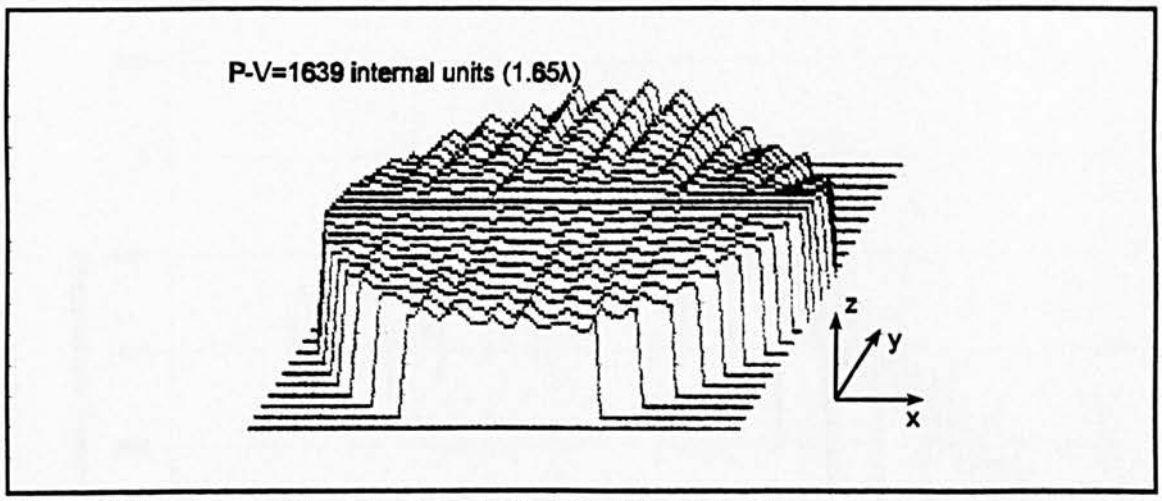


Figure 3.13 Isometric plot of results for flat A (with bug).

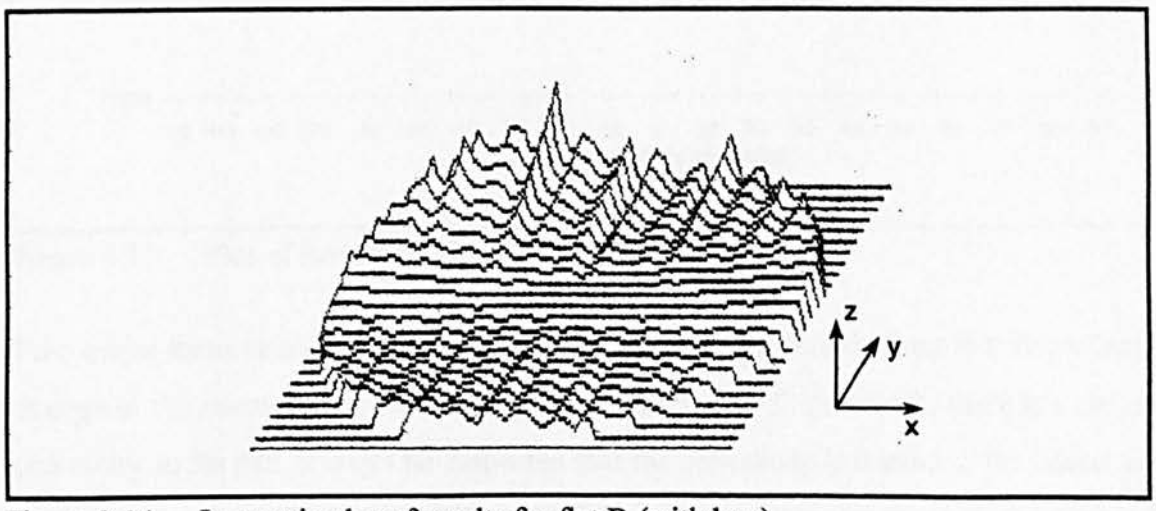


Figure 3.14 Isometric plot of results for flat B (with bug).

It can easily be seen that the results are far from the expected plane (flat A) and astigmatic (flat B) surfaces figures. It can also be seen that the predominant feature of the results is a waviness that increases away from the horizontal diameter (y=0). The profile of the horizontal diameter of flat A appears to be straight as expected and the profile of the horizontal diameter of flat B can be seen to be slightly curved which may be consistent with the desired astigmatic figure. Figure 3.15 shows a plot of the data for flat A along a horizontal chord (y=60). The data was output by subroutine @_chord_print and the y-axis units are ZIPL "internal units". One internal unit equals 1/512 fringe (1/1024 λ).

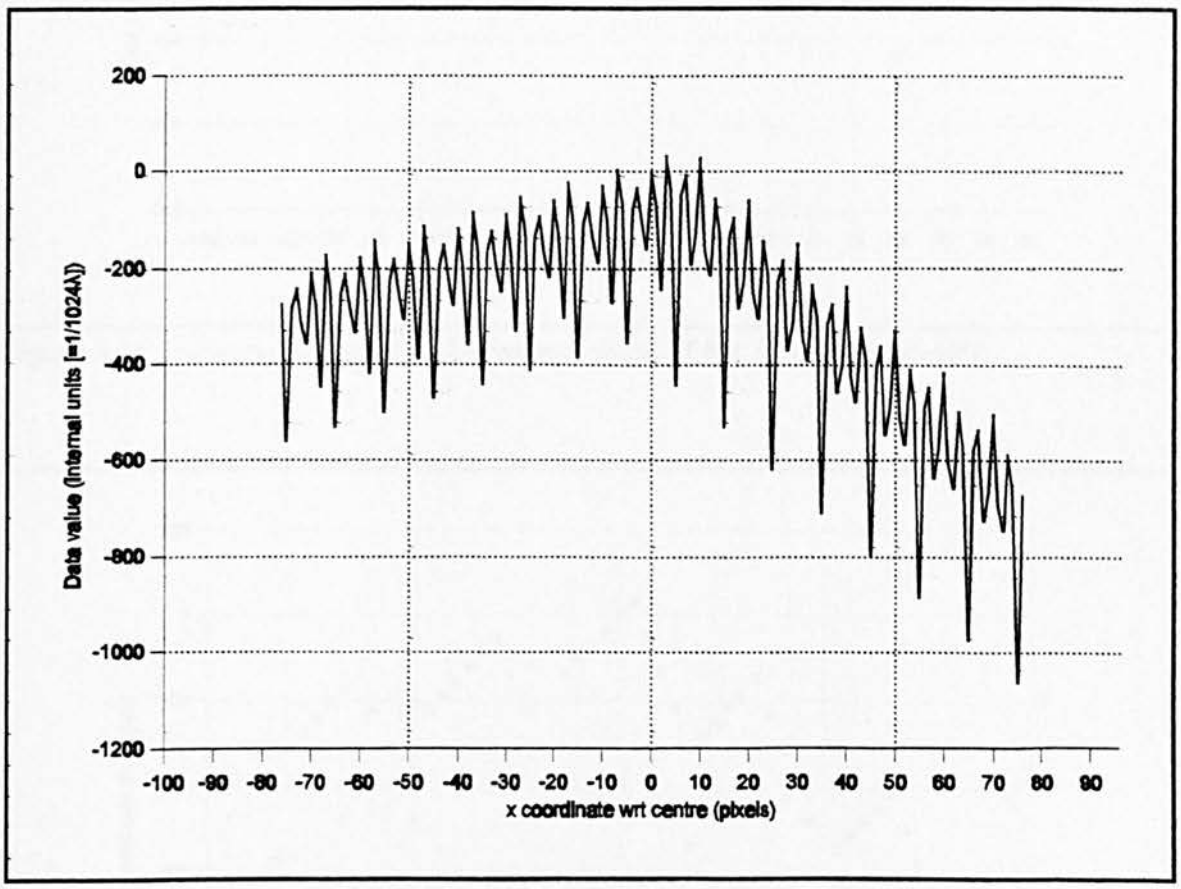


Figure 3.15 Plot of horizontal chord of flat A with bug (y=60).

Two major features are immediately obvious from the plot. Firstly, there is a discontinuous change in the overall slope of the line in the vicinity of x=0. Secondly, there is a complex periodicity to the plot. It might be suspected that the periodicity is related to the lateral shifts in the positional combinations ABS1 and ABS2 and indeed this becomes obvious when the data are re-plotted using every 10$^{th}$ point (s1= 10) and every 3$^{rd}$ point (s2= 3) in figures 3.16

and 3.17 respectively.

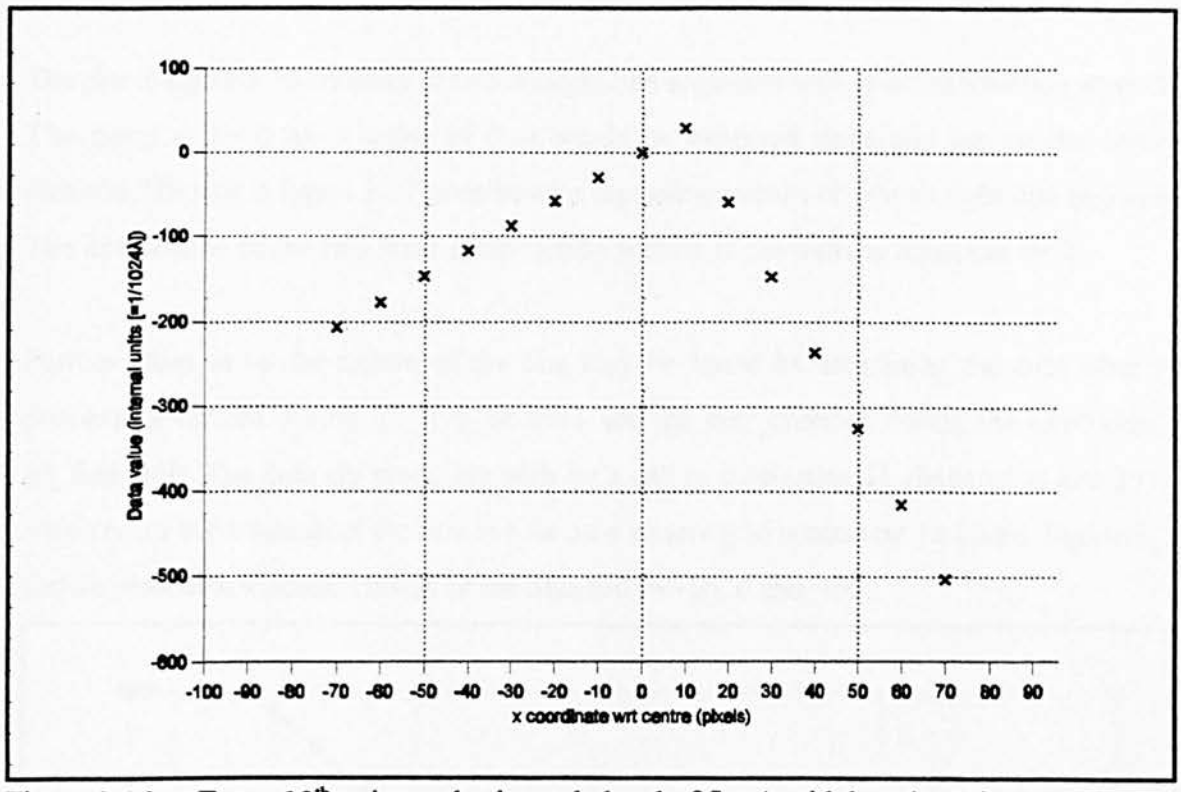


Figure 3.16 Every 10th point on horizontal chord of flat A with bug (y=60).

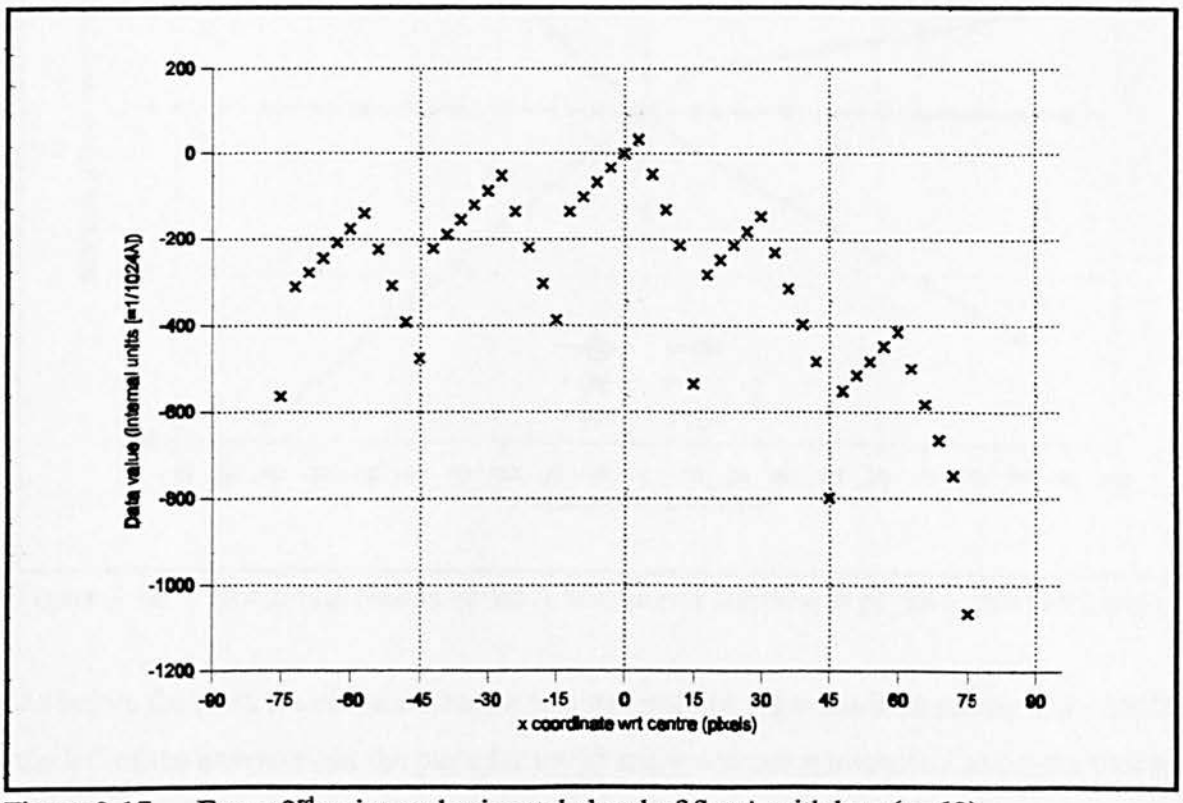


Figure 3.17 Every 3rd point on horizontal chord of flat A with bug (y=60).

The plot in figure 3.16 consists of two straight line segments with their intersection at x= 10. The point at x= 0 has a value of 0 as would be expected since this lies on the vertical diameter.The plot in figure 3.17 consists of a repeating pattern of two straight line segments. The intersection of the two lines in the middle section of the pattern occurs at x= 3.

Further clues as to the nature of the bug may be found by examining the data after the processing carried out by @_step_chords1 and @_step_chords2 during the execution of @_find_ytilt. The data are made available by a call to subroutine @_diagnost at line 25132 which prints out a sample of the data in a.da on a square grid spaced by 10 pixels. Figure 3.18 shows plots of horizontal chords of the data for y= -50, 0 and +50.

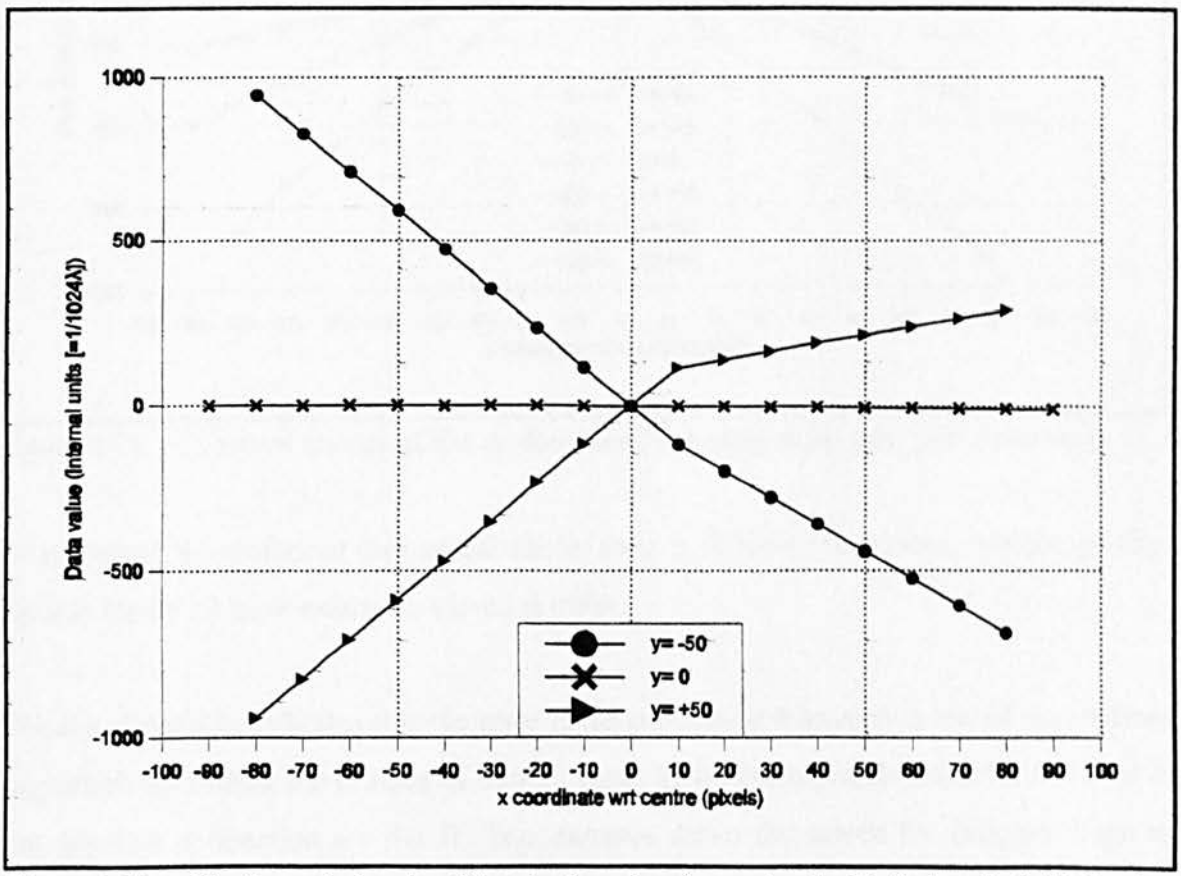


Figure 3.18 Horizontal chords of flat A after chord stepping in @_find_ytilt (with bug).

As before the plots are characterized by two straight line segments intersecting at x= 10. To the left of the intersections the plots for y= 50 and y=-50 are symmetrical about the plot for

y= 0 as would be expected for the straight vertical chords for a flat surface. To the right of the line intersections, however, the symmetry is no longer present indicating that the profile of the vertical chords for x> 10 are no longer straight. This can be seen by plotting the profiles of a few vertical chords from the same data as shown in figure 3.19.

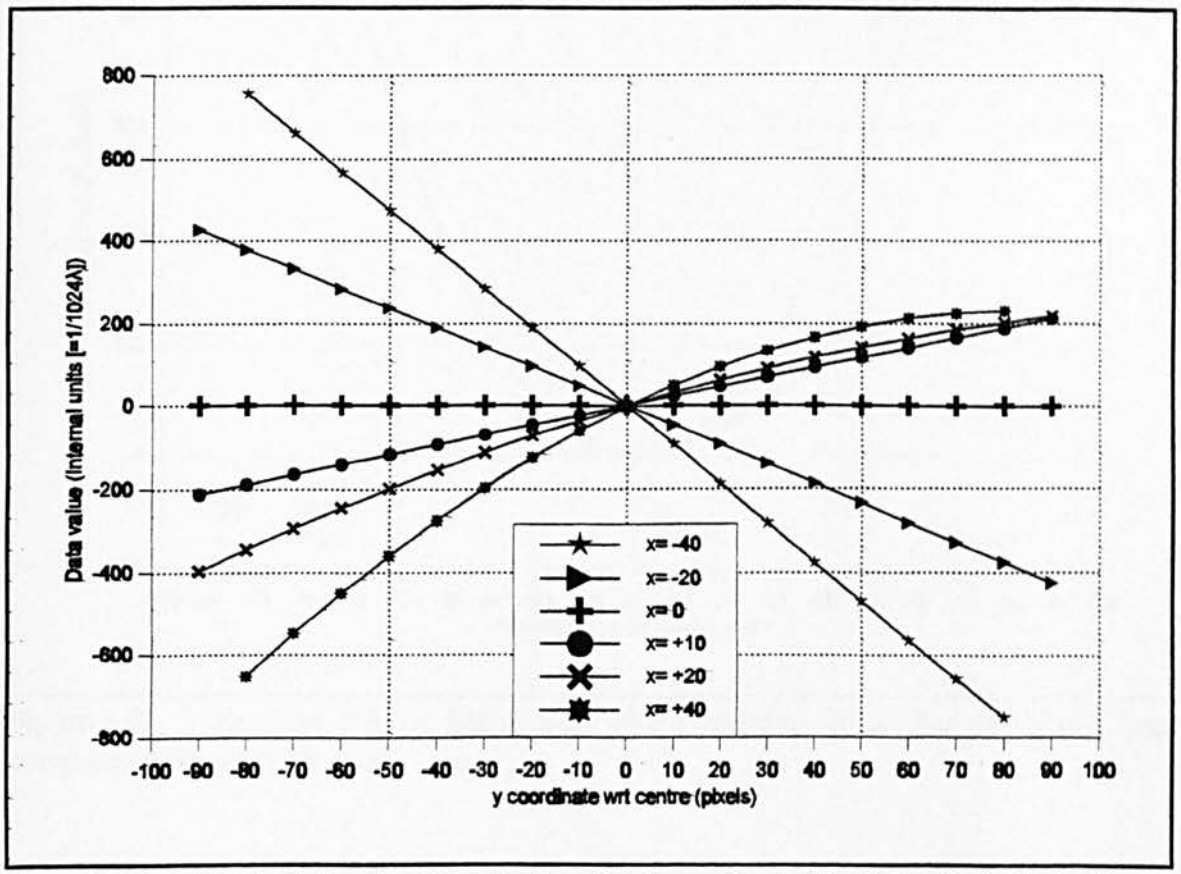


Figure 3.19 Vertical chords of flat A after chord stepping in @_find_ytilt (with bug).

As suspected, the profiles of the vertical chords for x ≤ 10 have the correct, straight, profile. Chords for x> 10 have incorrect, curved profiles.

The above evidence indicates that the error in the processing is located in one of the routines responsible for finding the profiles of chords, either in the positive x direction for flat A or in the negative x direction for flat B. This narrows down the search for program bugs to subroutines @_step_chords2, @_sub_temp1b or @_sub_temp2b. Further evidence for this may be obtained by plotting the data for the diagonal diameter of flat A after processing by @_find_ytilt as shown in figure 3.20. It is this data that is used by @_find_ytilt to calculate the y-tilt error in the ABS datasets and is expected to follow a quadratic form as discussed in

section 2.3.4.3 (equation 2.19).

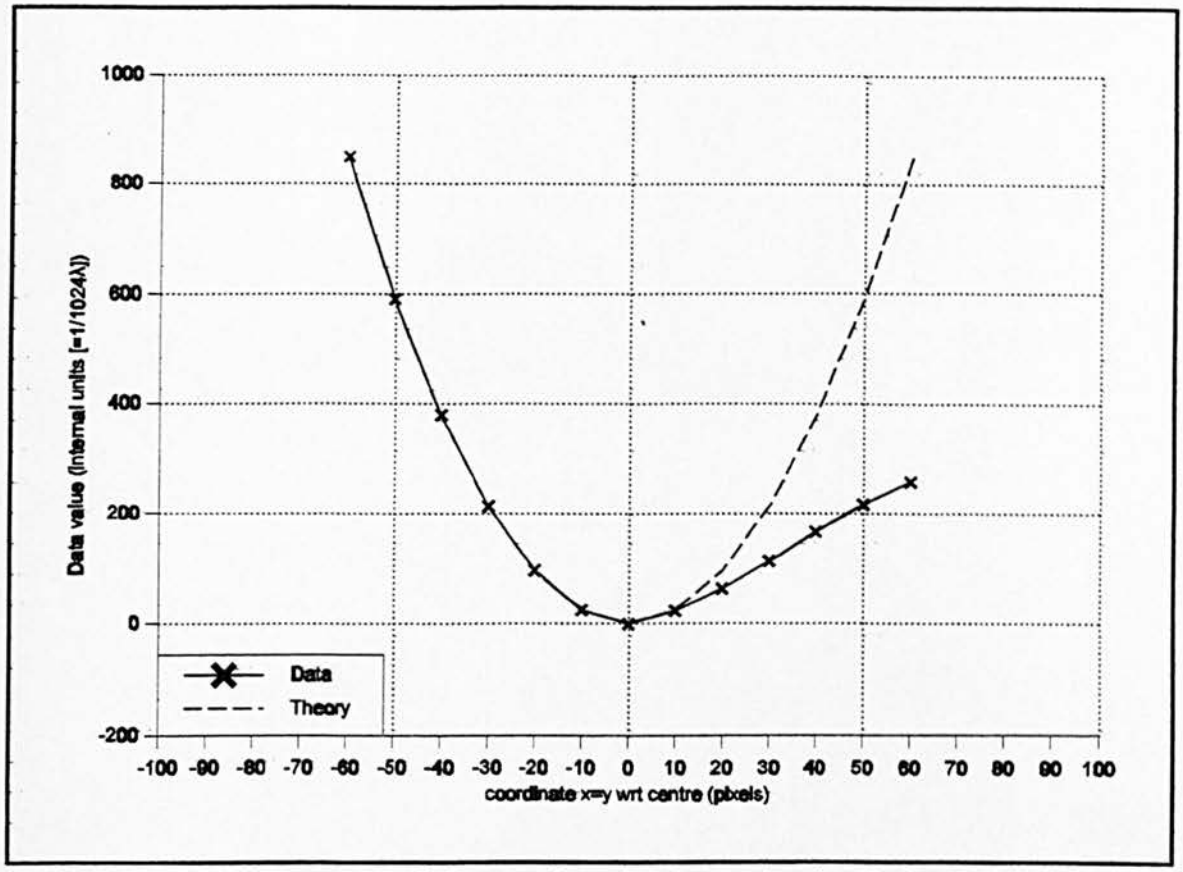


Figure 3.20 Diameter x=y of flat A after chord stepping in @_find_ytilt (with bug) compared with expected result.

It can be seen that the data follows the expected quadratic form for $x \leq 10$. For $x > 10$ the data deviates from the expected form. Hence one obvious effect of the bug in the program is that the calculation of the y-tilt adjustment will be in error.

Further clues as to the location of the bug in the program may be obtained by examination of the data output for flat B. Figure 3.21 shows the profiles of horizontal chords from the data for flat B after the chord stepping procedures in @_find_ytilt.

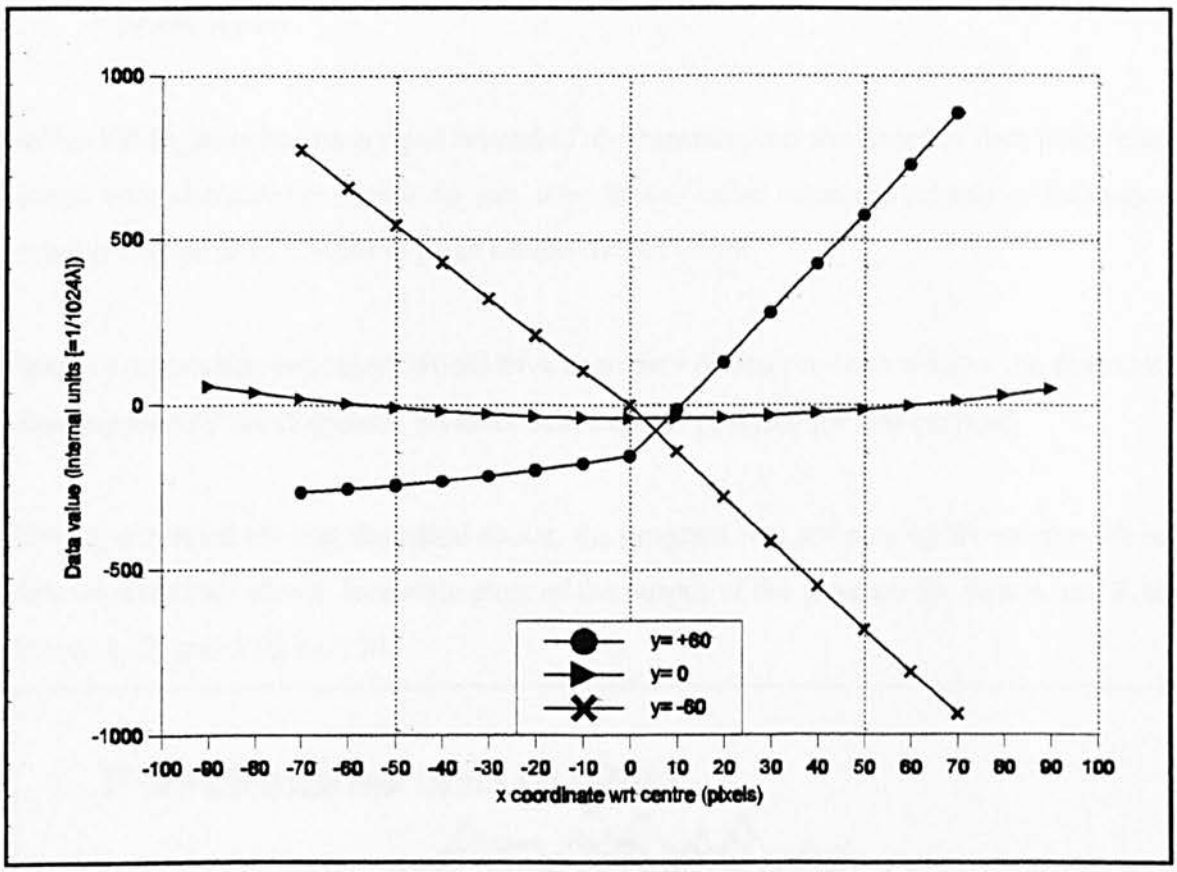


Figure 3.21 Horizontal chords of flat B after chord stepping in @_find_ytilt (with bug).

Once again the plots are characterized by a discontinuity though for flat B it occurs at x=0. This fact, together with the fact that the discontinuity in the data for flat A does not occur until x= 10, indicates that the probable location of the bug is in the routine that finds the profile of the first new chord on flat B in B's negative x direction. This subroutine is @_sub_temp2b and indeed, close inspection of this section of code revealed the bug shown below.

```
29500 @_sub_temp2b:
29510 for dpy= ymin to ymax
29520 temp1 [dpy]= bad
29525 dpoint= DATAPOINT [xcent+chord][ycent]
29530 if dpoint= bad then goto 29560
29540 if temp2 [dpy]= bad then goto 20560
29550 temp1 [dpy]= dpoint- temp2[dpy]
29560 next dpy
```

29900 return

In line 29525, *ycent* had been typed instead of *dpy* meaning that the chord of data from which temp2 was subtracted each time @_sub_temp2b was called consisted entirely of data values equal to the value of the centre point on the correct chord.

Such a program bug and others would have been very difficult to spot without the diagnostic data supplied by the diagnostic routines built into the program for that purpose.

Having corrected the bug described above, the program was rerun using the same synthetic data as described above. Isometric plots of the output of the program for flats A and B are shown in figures 3.22 to 3.24.

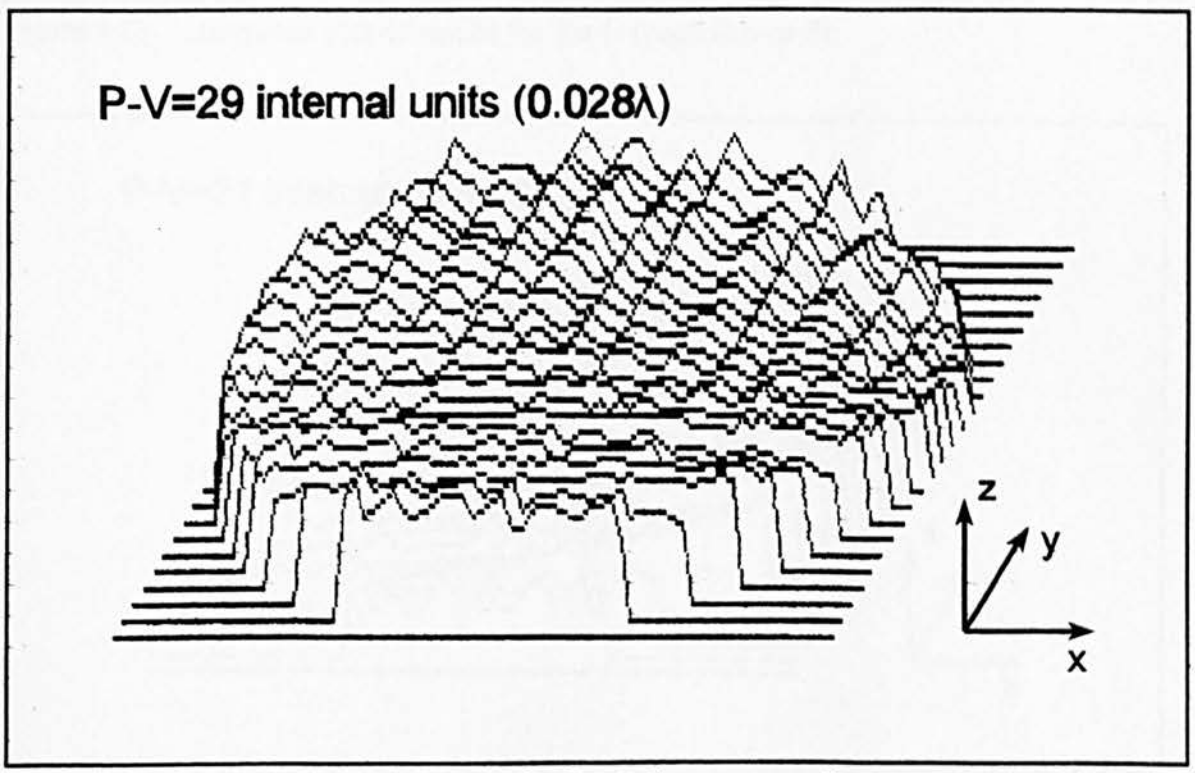


Figure 3.22 Isometric plot of results for flat A (bug corrected).

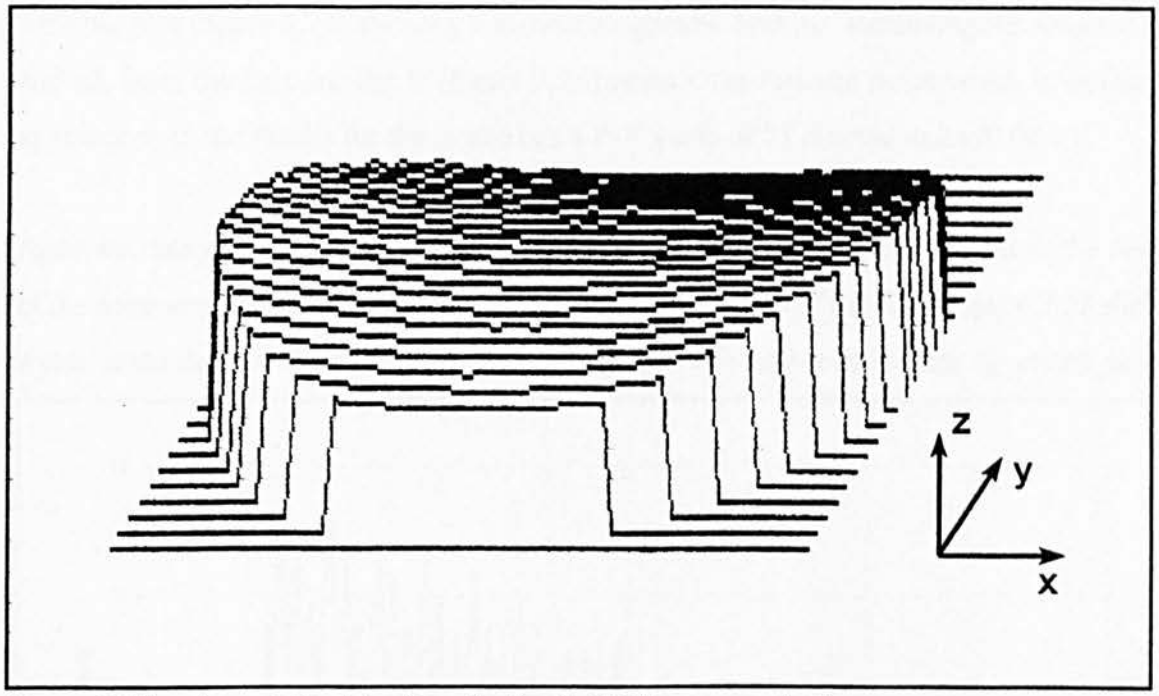


Figure 3.23 Isometric plot of results for flat B (bug corrected)

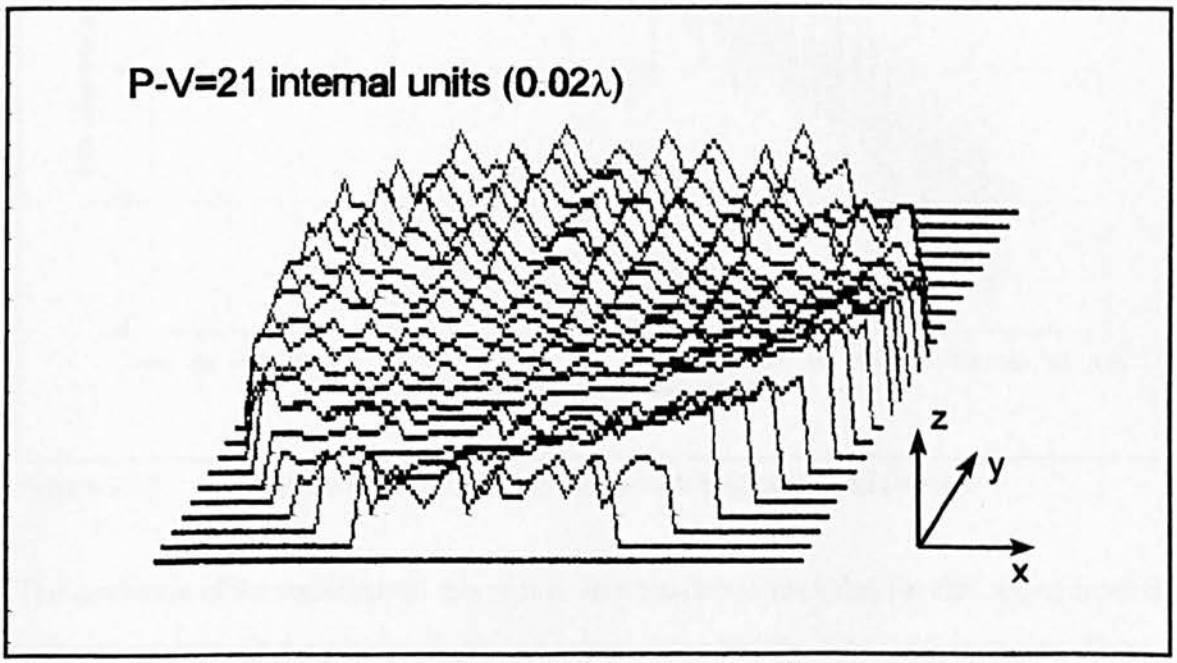


Figure 3.24 Isometric plot of results for flat B with astigmatism removed (bug corrected).

At first glance, the results for flat A (figure 3.22) appear no better than those previously obtained but this is due to the difference vertical scaling between the plots. The P-V value for this data is only 29 internal units (0.028 λ) whereas the P-V value for the previous results for flat A was 1693 internal units (1.65 λ). The results for flat B appear more promising, however,

with the plot (figure 3.23) showing a distinct astigmatic contour. Removing the astigmatism and tilt from the data for flat B (figure 3.24) reveals the residual noise which is similar in appearance to the results for flat A and has a P-V value of 21 internal units (0.02 λ).

Again the data output by the diagnostic routines may be used to provide clues as to the origin of the noise which is superimposed upon the correct contours of the flats. Figure 3.25 shows a plot of the data for a horizontal chord of flat A (y= 60) output by routine @_chord_print.

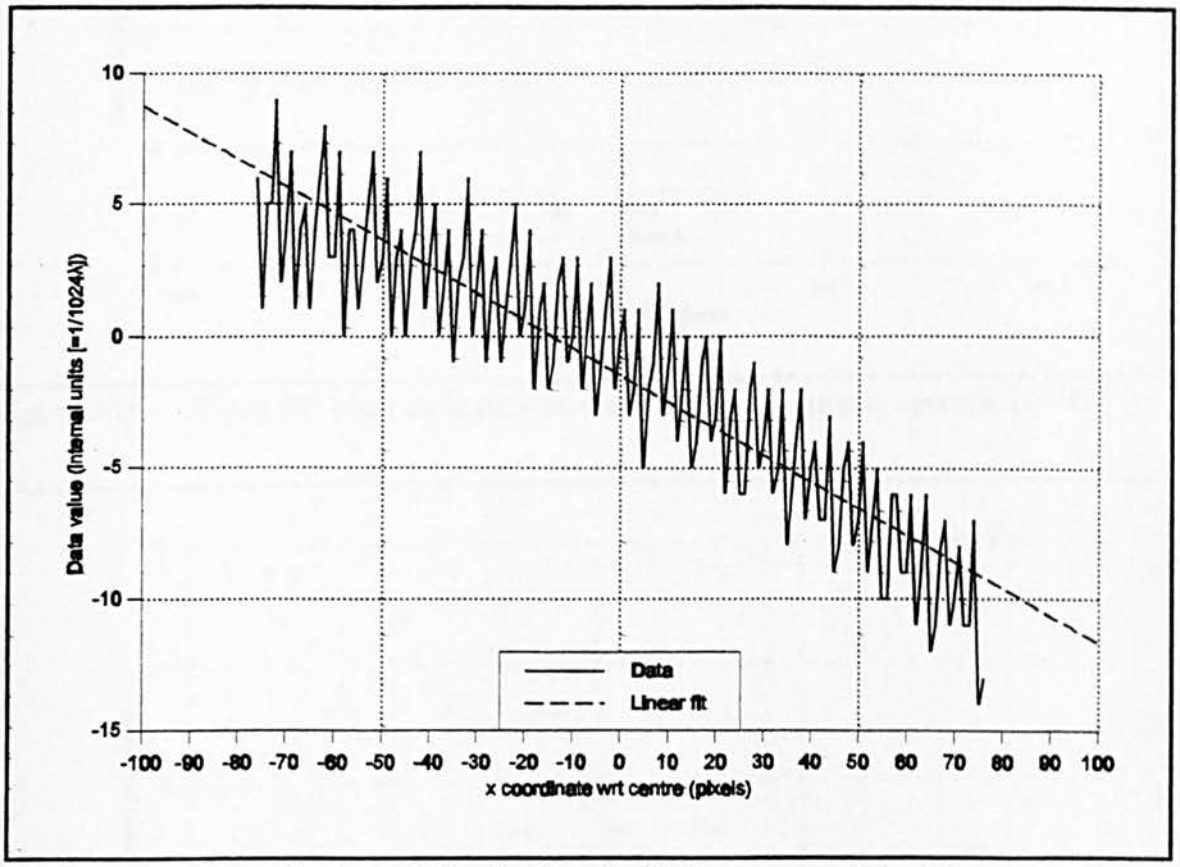


Figure 3.25 Plot of horizontal chord of flat A with bug corrected (y=60).

The amplitude of the variations in this plot is very much less than that for the output from the bugged program and the previously obvious slope discontinuity is no longer present. There is still an overall slope error (the ideal output being a straight horizontal line with data value equal to zero) and there is an apparently periodic oscillation about the mean slope. The periodic nature of the line is again related to the two lateral shifts of 10 and 3 pixels as can be seen from figures 3.26 and 3.27 respectively.

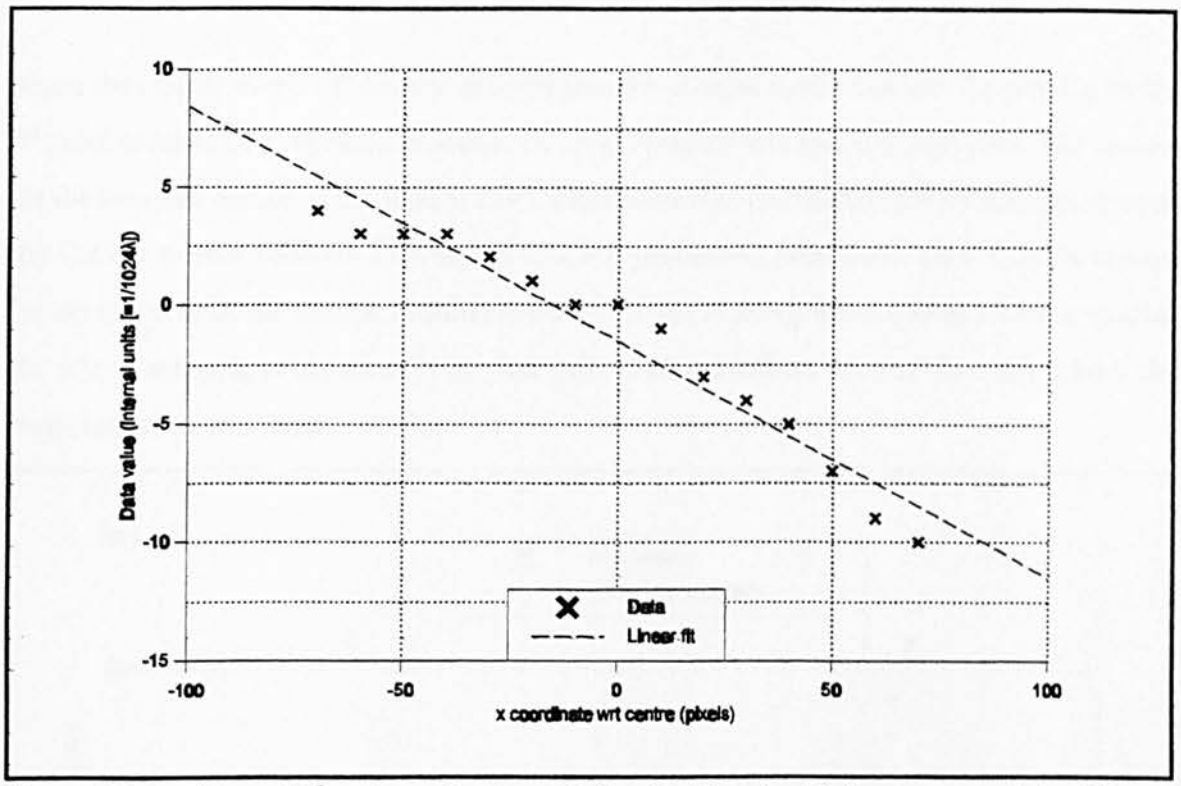


Figure 3.26 Every 10th point on horizontal chord of flat A with bug corrected (y=60).

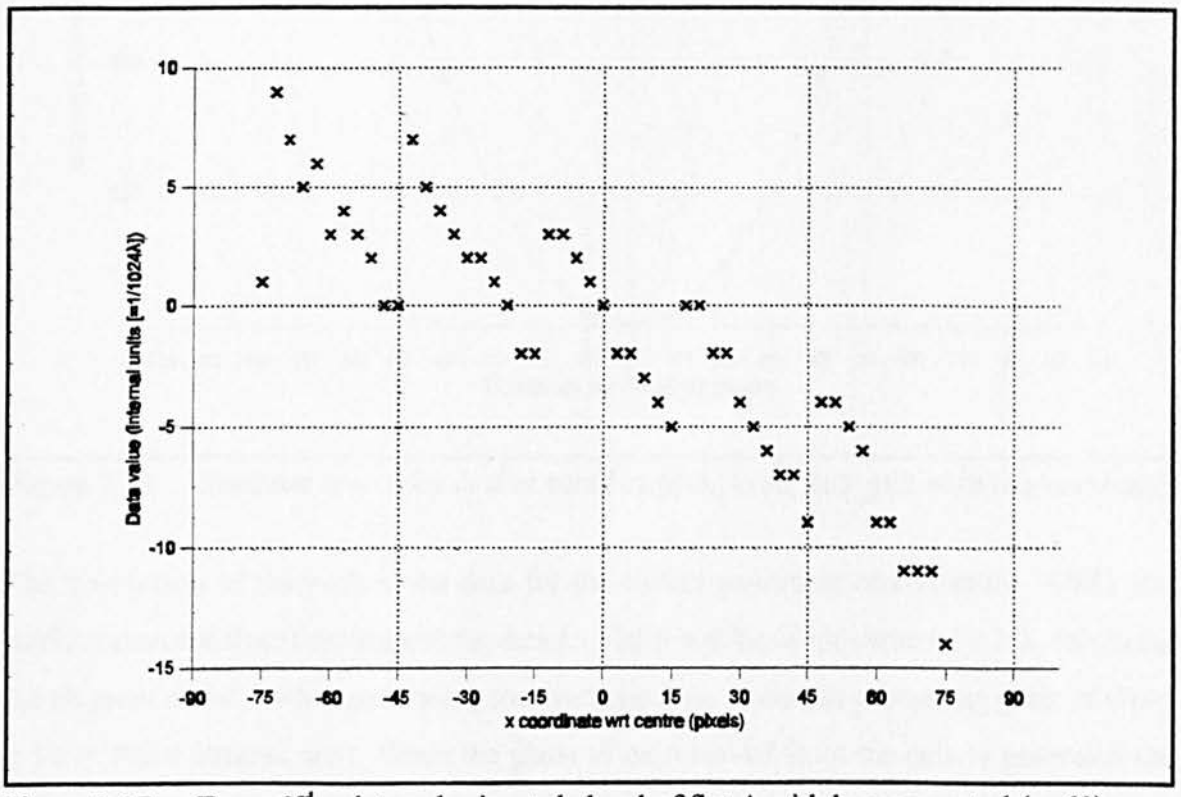


Figure 3.27 Every 3rd point on horizontal chord of flat A with bug corrected (y=60).

Again the plot for every 10th point is an approximately straight single line and the plot for every 3rd point consists of a repeated sequence of approximately straight line segments. The reason for the step-like nature of the lines is ZIPL's use of integer calculation for its manipulation of the DATA system variable. This aspect of ZIPL processing provides a clue as to the reason for the tilt error in the results. Plotting the data for flat A along the diagonal where x=y after the chord stepping procedures in @_find_ytilt (figure 3.28) shows that the data follow the expected quadratic form.

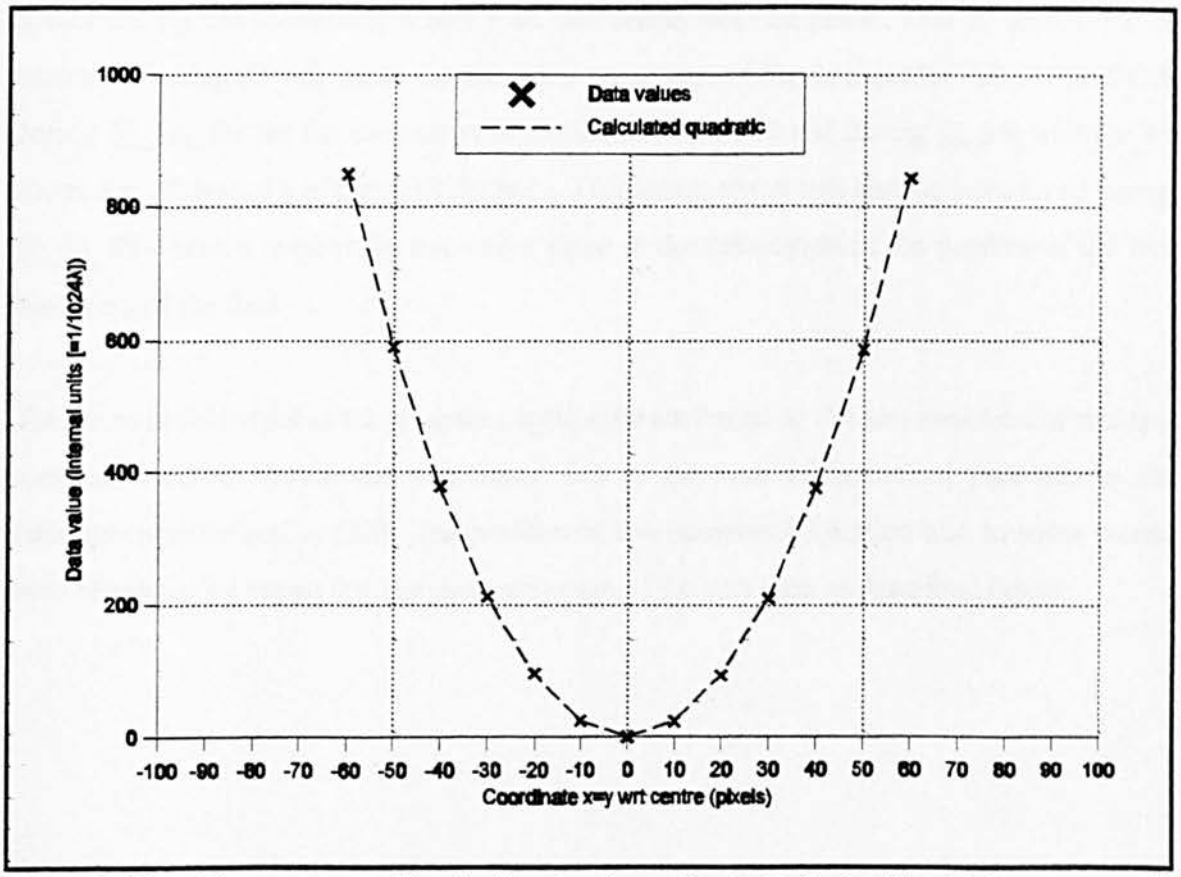


Figure 3.28 Diameter x=y of flat A after chord stepping in @_find_ytilt with bug corrected.

The correction of the y-tilt in the data for the shifted positional combinations, ABS1 and ABS2, calculated from this data and the data for Flat B will be of the order of 0.1 λ, this being the tilt given to flat A when generating the synthetic data. From the processing point of view, 0.1λ = 102.4 internal units. When the plane to be removed from the data is generated the zremove command will truncate the argument supplied to it to an integer value thus introducing a significant truncation error. The linear fit to the data shown in figure 3.26

indicates that the slope of the data at y= 60 for every 10th point on flat A is about -10 internal units per 100 pixels. From the discussion in section 2.3.4.3 and equation 2.19, this indicates a residual y-tilt error in the dataset for ABS1 of approximately 0.6 internal units over the radius of the dataset. This value is below the resolution available for generation of a correction plane to be removed from the data. In addition, the values of the plane at any point will have been rounded (or truncated) to integer values, so introducing further errors. The introduction of truncation errors due to ZIPL's use of integer arithmetic is not solely confined to the removal of y-tilt errors in the ABS datasets. The same problem will introduce errors at all points during the processing where a tilt correction must be made. That is, errors will be introduced during @_adj_data1 for the initial correction of the datasets for AB, AC and CB, during @_adj_abr for the correction of the dataset for ABR and during @_adj_xtilt for the correction of the x-tilt in the ABS datasets. Truncation errors will also be introduced during @_do_fff where a division by two takes place in the calculation of the profiles of the first diameters of the flats.

The errors in the output of the program may thus be attributed to the accumulation of multiple truncation errors during the processing due to the lack of numerical precision in the calculations performed by ZIPL. The problem of low numerical precision has, to some extent, been alleviated by expanding the dynamic range of the raw data as described below.

### 3.4.1 Increasing the Numerical Precision of the Calculations.

As described above, the Zygo Mark IV processor performs arithmetic functions on the DATA system variable using internal units (= 1/1024 $\lambda$) and integer arithmetic. The value of a data point may take any integer value between -32767 and +32767. Any attempt to write a DATA value outside of this range causes the program to crash with an "overflow" error. Writing a non-integer value to DATA causes the value to be rounded to the nearest integer value. When processing datasets corresponding to nearly flat surfaces with little tilt, the range of data values covers only a small range of the allowed values about zero and so the errors caused by truncation of the data are significant in comparison to the data. This is particularly so for the algorithm described above since the errors introduced at any step in the processing will propagate through to the next processing step.

The effect of truncation errors on the processing were reduced by the introduction of the multiplication factor, "multfact" (line 1050). It should again be stressed that this factor had not been introduced to the program at the time the above results were generated. The introduction of multfact increases the numerical precision by multiplying the data values of the raw data at the time of acquisition (line 10105 in @_acq_data) or at the time of synthesis (in @_synth) by multfact.

Multiplication of the data does not increase the precision of the raw data but increases the dynamic range available to the processing. For example, a multiplication factor, multfact= 10 would increase the numerical precision to the equivalent of having one decimal point of precision. The choice of value for multfact is a compromise between increasing the numerical precision and avoiding the possibility of a data overflow. Even so, during the processing for @_find ytilt the data values can become very large due to the uncorrected y-tilt error and so the data are tested for potential overflow before being written into DATA and ignored if they fall outside the allowed range. This only has the effect of potentially reducing the number of data points available for the calculation of the y-tilt error. The value of multfact used for the results given below is 50.

When the processing of the data is complete, the dynamic range of the data is re-compressed by dividing the data by multfact. Figures 3.29 to 3.31 show isometric plots of the results of the processing using the same synthetic data as above with dynamic range expanded 50 times.

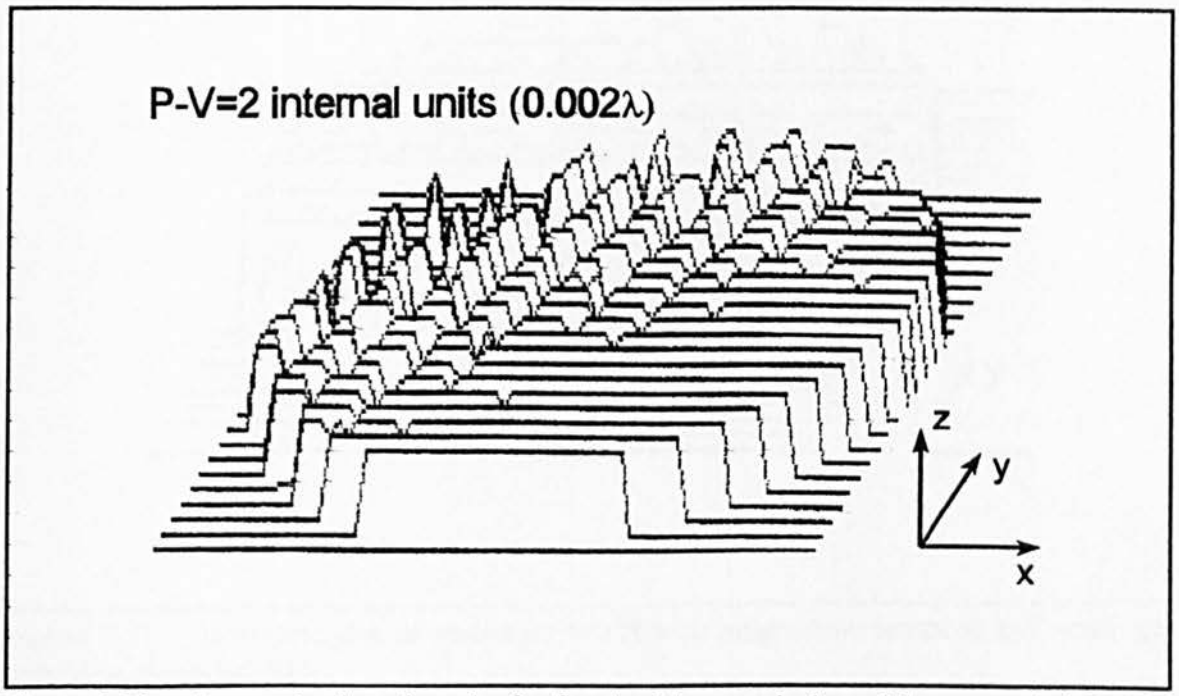


Figure 3.29 Isometric plot of results for flat A with numerical precision expanded 50X

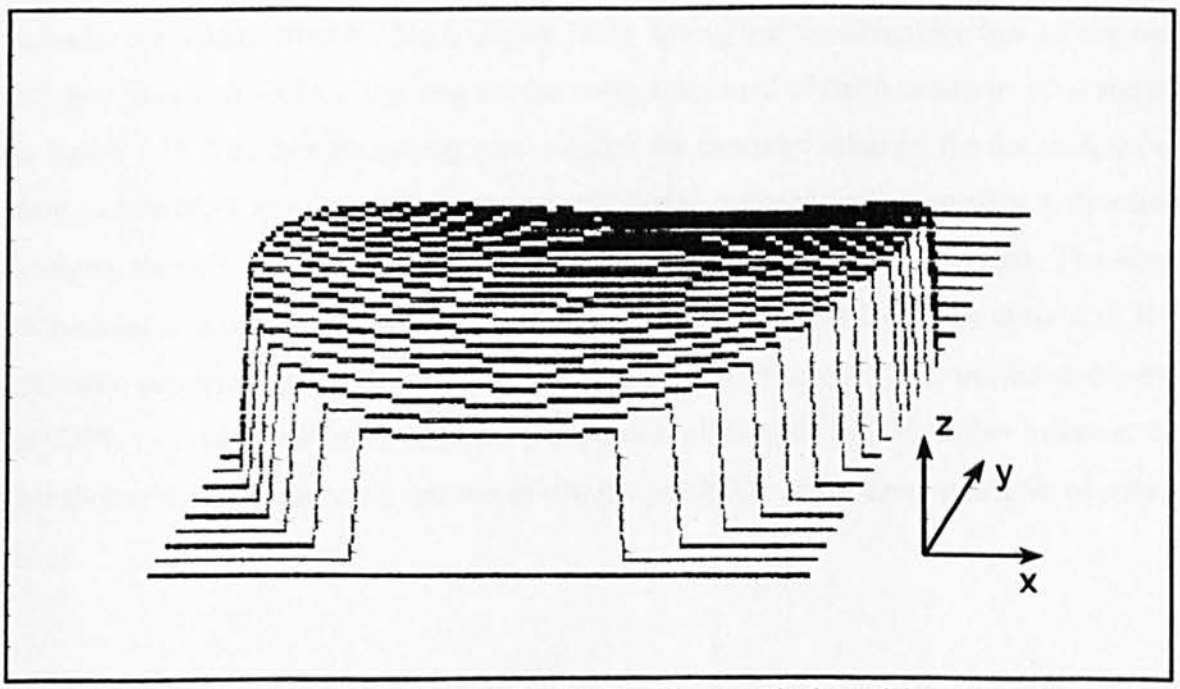


Figure 3.30 Isometric plot of results for flat B with numerical precision expanded 50X.

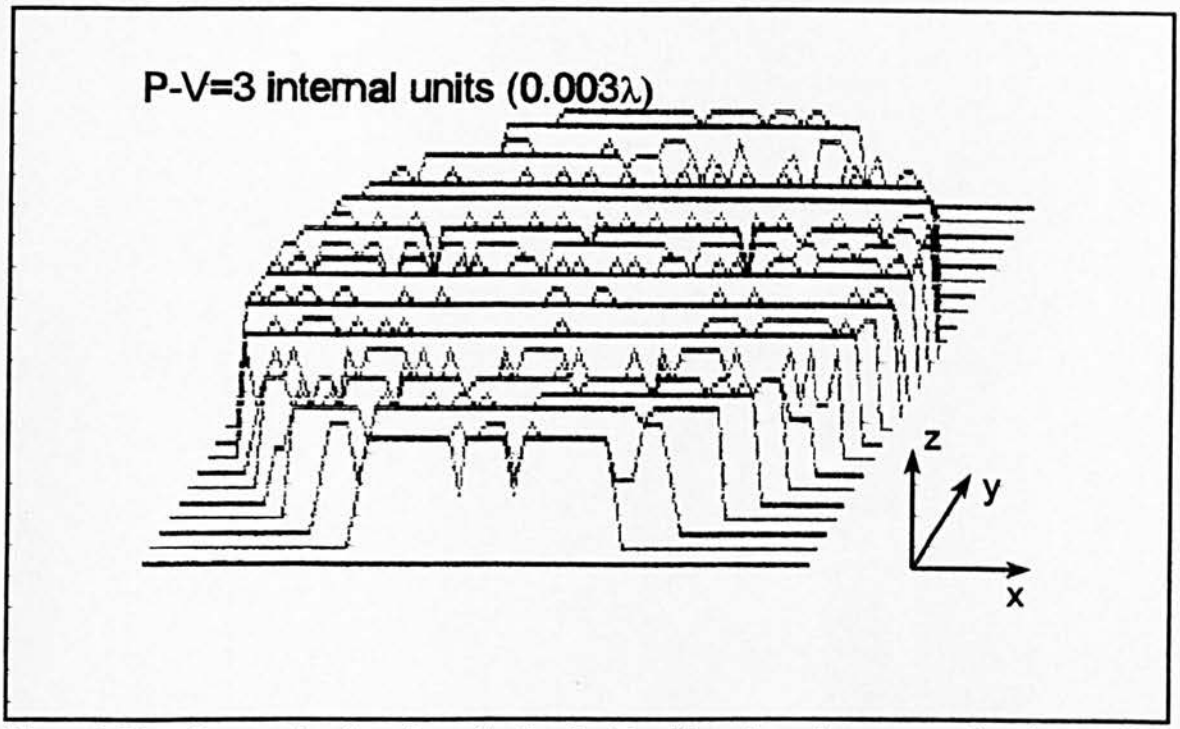


Figure 3.31 Isometric plot of results for flat B with astigmatism removed and numerical precision expanded 50 X.

The P-V value for the data for flat A (figure 3.29) is 2 internal units (0.002λ) and the P-V value for the residual error for flat B (figure 3.31), having had the astigmatic figure removed, is 3 internal units (0.003λ). The data for the horizontal chord of flat A where y= 60 is shown in figure 3.32. The data are mostly zero which is the expected value for the flat surface but data values of -1 internal units become increasingly common in the negative x direction implying the very small residual tilt of the data shown by the linear fit to the data. The slope of the linear fit is of the order of 0.5 internal units (0.0005 λ) over the aperture of the flat. The increased numerical precision has thus reduced the slope errors due to the integer arithmetic of ZIPL to a very low level. Increasing the value of multfact would further increase the precision of the calculations but that would risk the possibility of the data causing an overflow error.

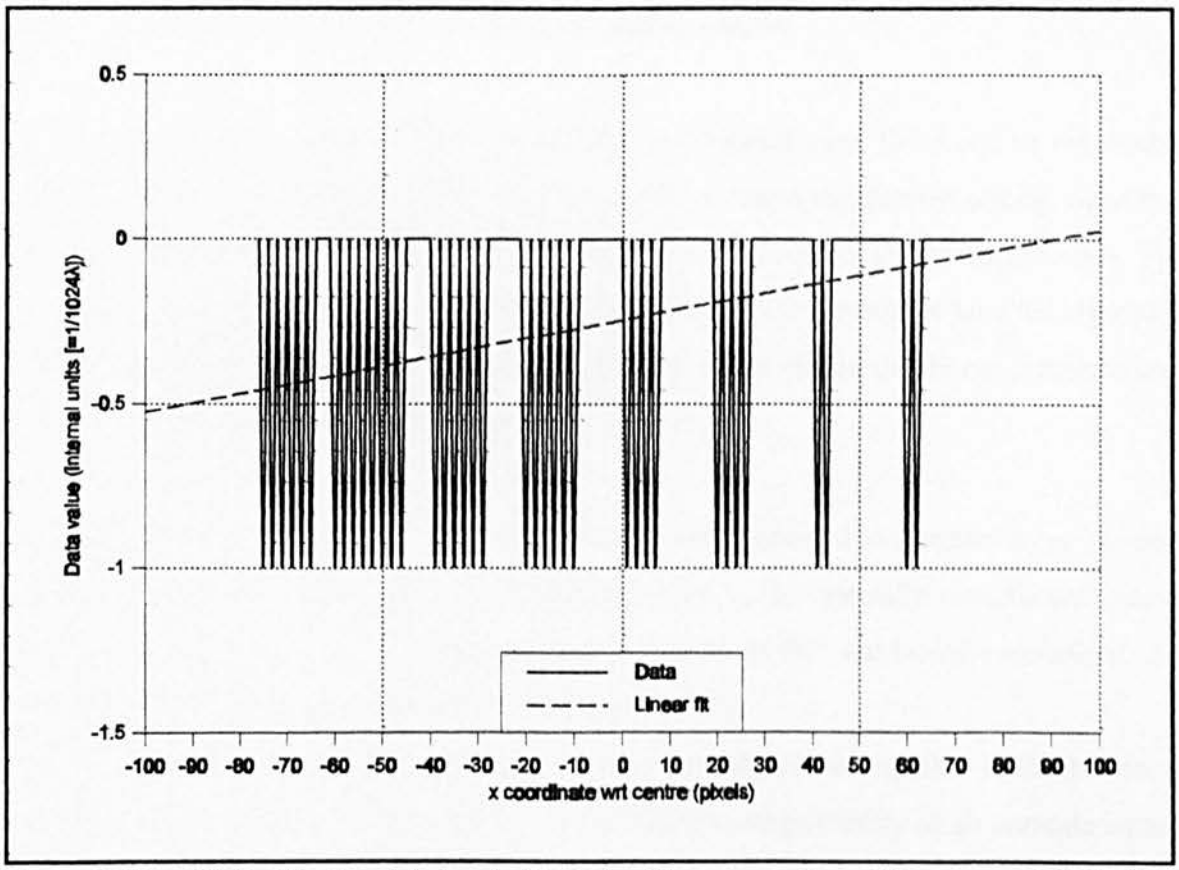


Figure 3.32 Plot of horizontal chord of flat A with numerical precision expanded 50 X.

## 3.5 Results of Experiments Using Real Measurements.

Having performed the above experiments using synthesized data, there can be reasonable confidence that the software is free from bugs and performs the flatness testing algorithm described in chapter two within the limits imposed by the Zygo Mark IV system software. The next stage in the development of the flatness testing method is to investigate how the algorithm performs using data obtained from actual measurements of the various positional combinations. This was carried out as described below.

Of the three flats used in the test, two (flats A and C) were mounted in standard Zygo bayonet mounting cells. The third flat (flat B) was mounted in the specially constructed mount described in section 3.1.1 to facilitate accurate rotation by 90° and lateral translations. All three flats have a nominal diameter of 100 mm (4 inches).
The two flats for each positional combination were placed as close together as the mounting hardware would allow (about 10 mm) in order to minimize the possibility of air currents within the interferometer cavity disturbing the fringe pattern. Close spacing of the two surfaces also reduces the magnitude of measurement errors due to aberrations in the collimator as discussed in section 3.1.1. The cavity was adjusted to give a nulled interference pattern (one fringe over the field of view). This was necessary in order to ensure that the multiplication of the wave front data would not cause an overflow error. The allowed DATA values of -32767 to 32766 correspond to a maximum fringe density of 64 fringes (32 waves) across the field of view. This corresponds to a maximum number of fringes before multiplication by multfact of 64÷multfact fringes. For multfact= 50, the maximum number of fringes is therefore 1.28. A larger value of multfact would make it very difficult to set up a sufficiently nulled cavity to avoid an overflow error. The cavity was then surrounded by acrylic baffles to exclude ambient laboratory air currents.

The variable meas_no was set to 10 so that each dataset would be the average of ten measurements. The variable delay was set to 300 to give a five minute delay between instructing the interferometer to proceed with a measurement and the measurement taking place. This was to allow the operator to leave the laboratory while the measurement took place

in order to minimize any possible environmental effects on the interferometer due to the proximity of the operator.

Plots of the results of the processing of the acquired interference patterns for flats A and B are shown in figures 3.33 and 3.34.

The plots show the, now familiar, ridged appearence of data where the correction of y-tilt has been in error. The P-V value for flat A is 119 internal units (0.116 $\lambda$) and for flat B, 125 internal units (0.122 $\lambda$). A plot of the data for the horizontal chord of A where y= 60 is shown in figure 3.35. The reason for the y-tilt error is likely to be due to the sensitivity of the algorithm to experimental errors accumulating during the chord stepping process as will be discussed in section 3.5.1.

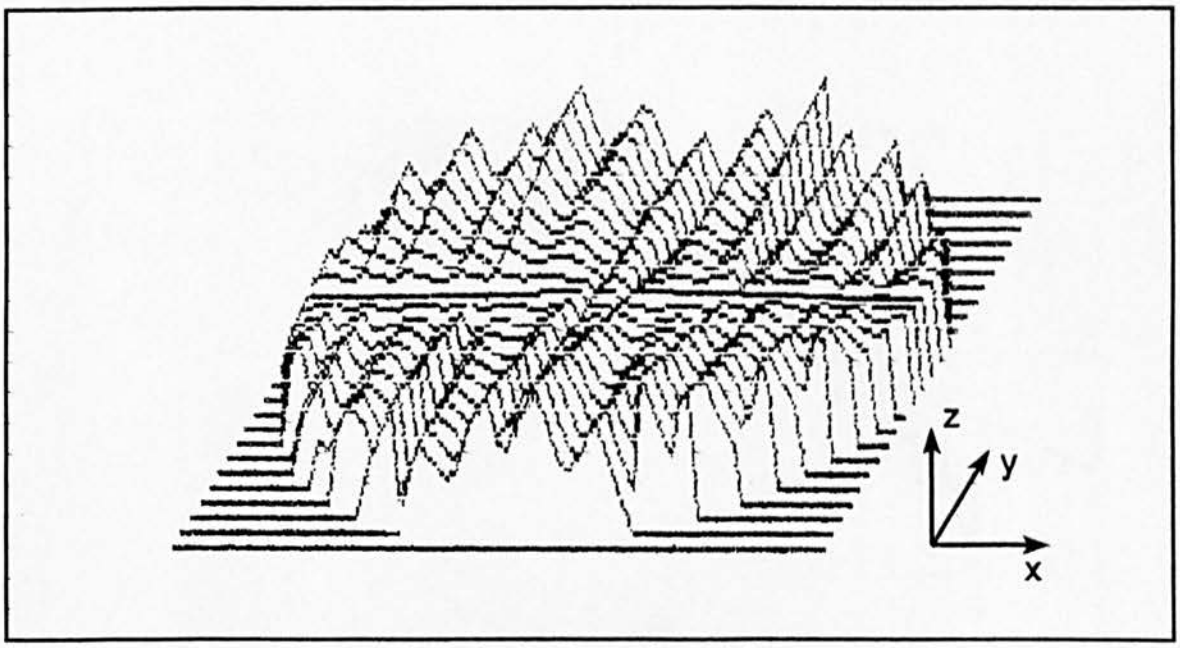


Figure 3.33 Isometric plot of results for flat A derived from measured data.

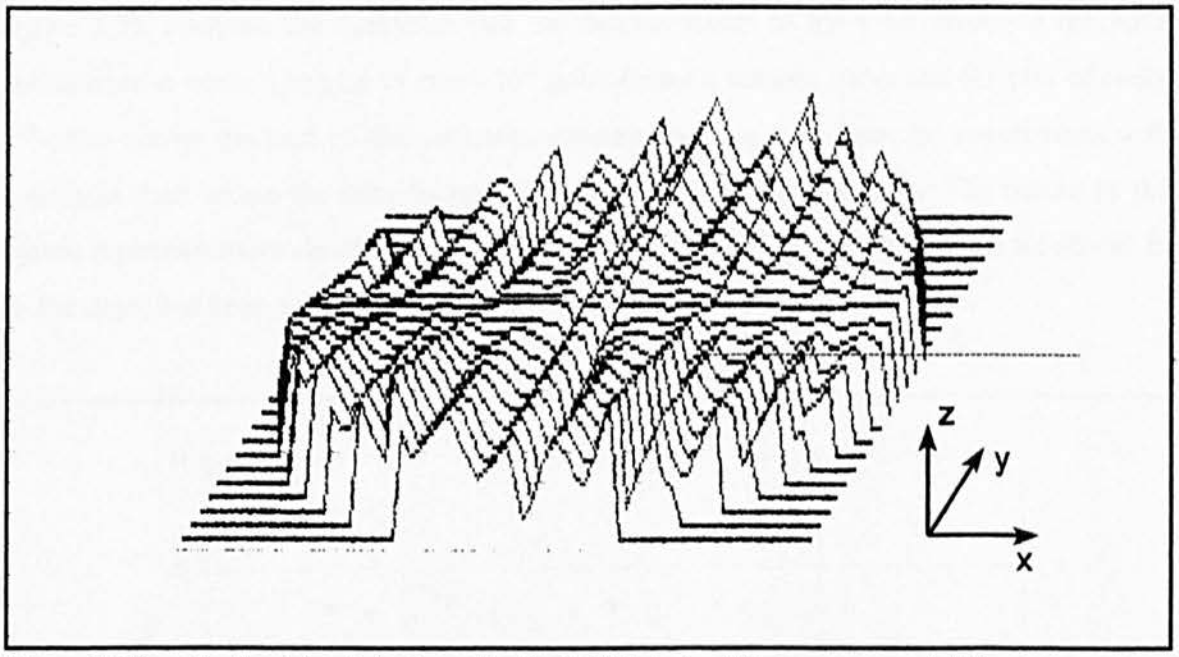


Figure 3.34 Isolmetric plot of results for flat B derived from measured data.

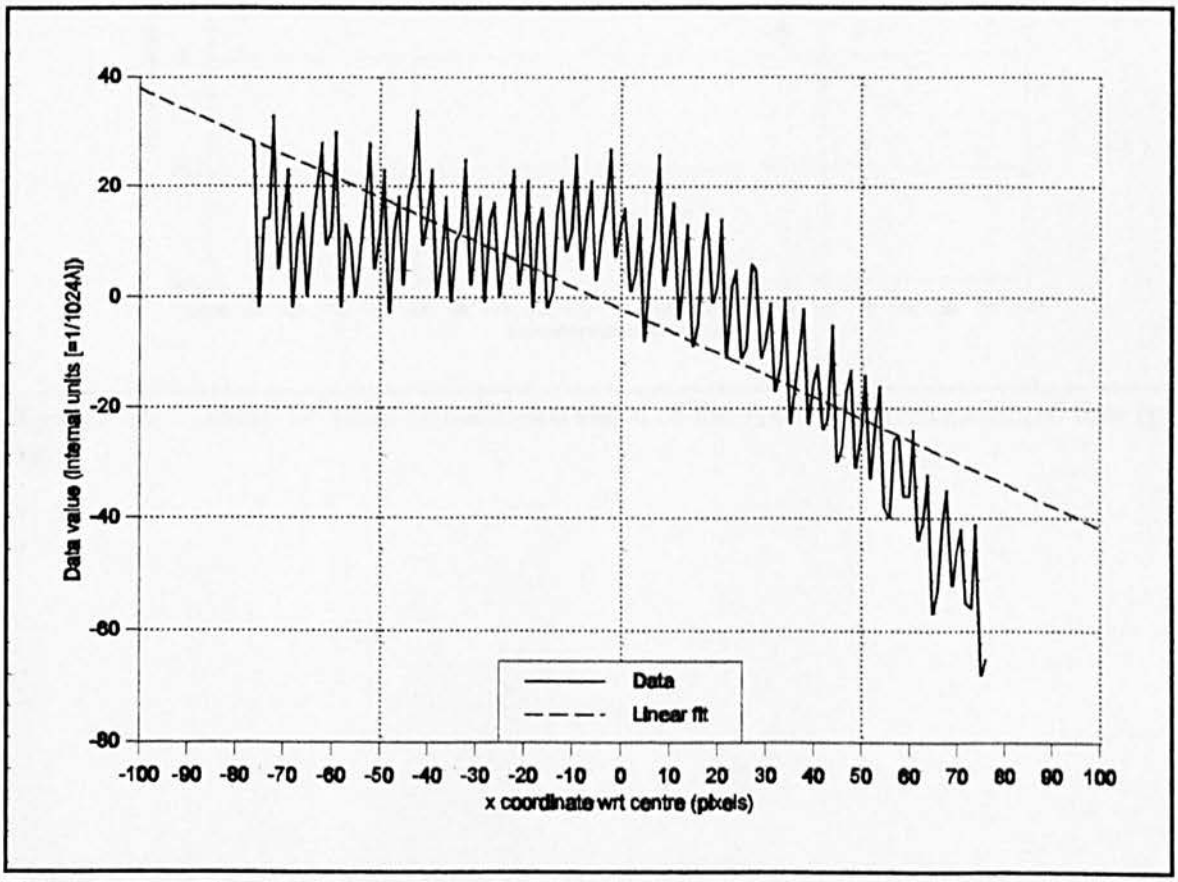


Figure 3.35 Plot of horizontal chord of flat A derived from measured data (y= 60).

Plotting every 10th point (figure 3.36) and every 3rd point (figure 3.37) of the data shown in

figure 3.35 confirms the suspicion that the determination of the y-tilt errors in the ABS datasets are in error. The plot of every 10th point forms a smooth curve and the plot of every 3rd point shows the kind of discontinuous repeated pattern seen from the experiments with synthetic data before the introduction of enhanced numerical precision. The nature of the curves is perhaps more clearly shown in figures 3.38 and 3.39 where the overall tilt (linear fit to the data) has been subtracted.

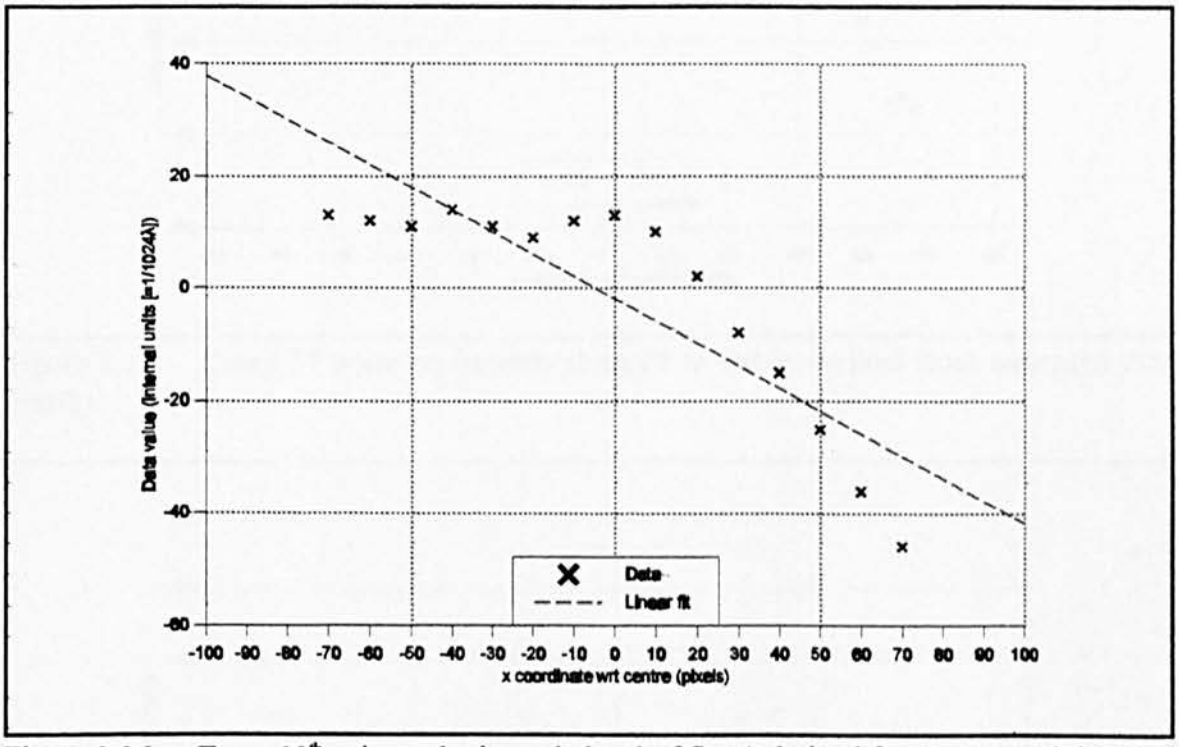


Figure 3.36 Every 10th point on horizontal chord of flat A derived from measured data (y= 60).

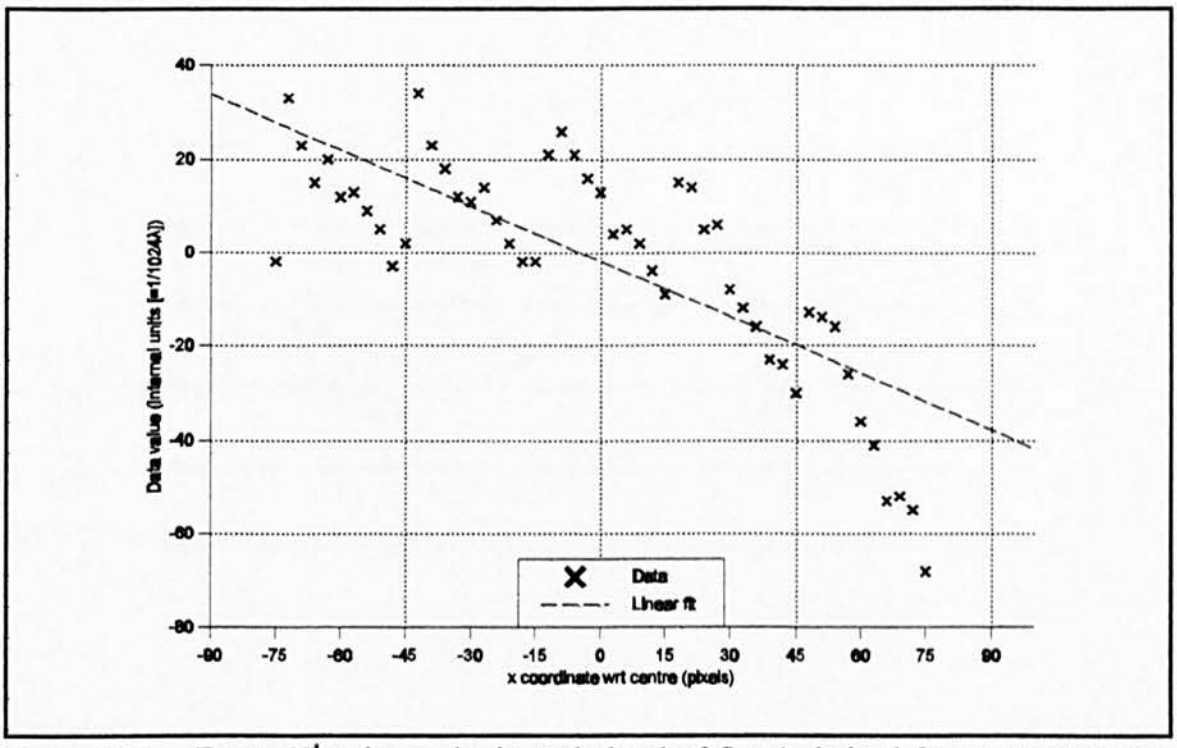


Figure 3.37 Every 3rd point on horizontal chord of flat A derived from measured data (y=60)

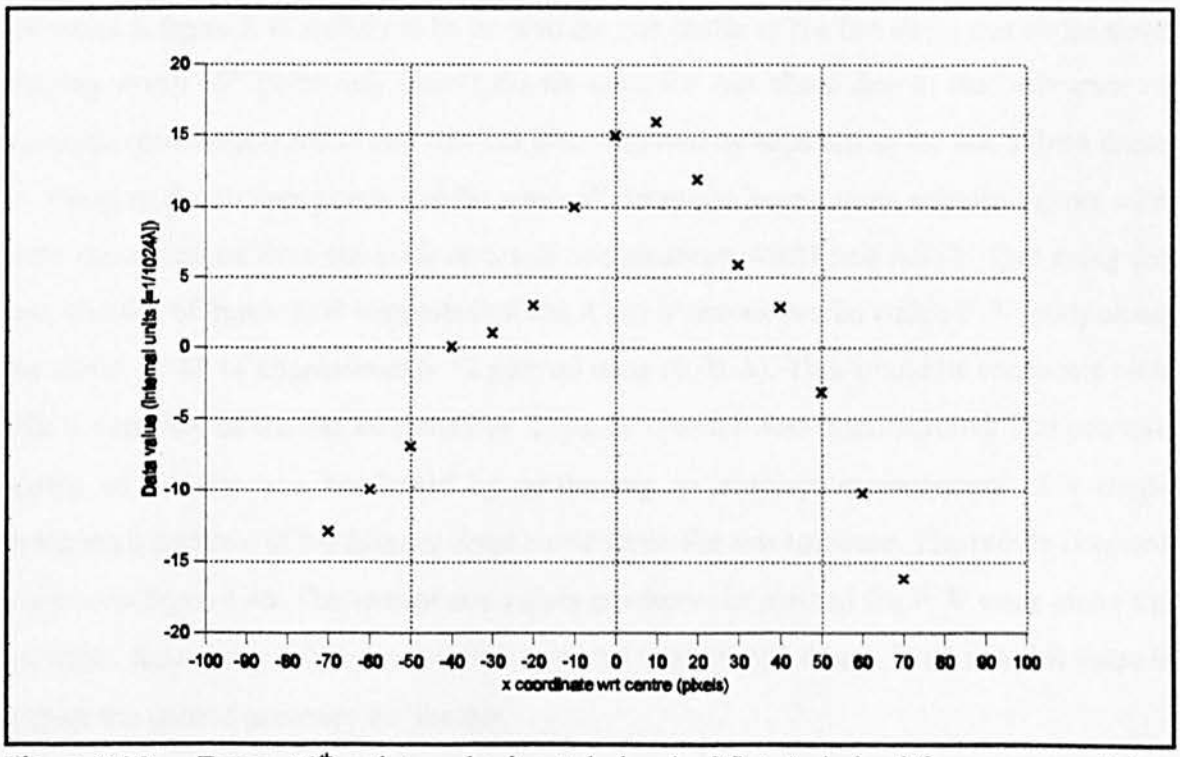


Figure 3.38 Every 10th point on horizontal chord of flat A derived from measured data minus linear fit (y=60).

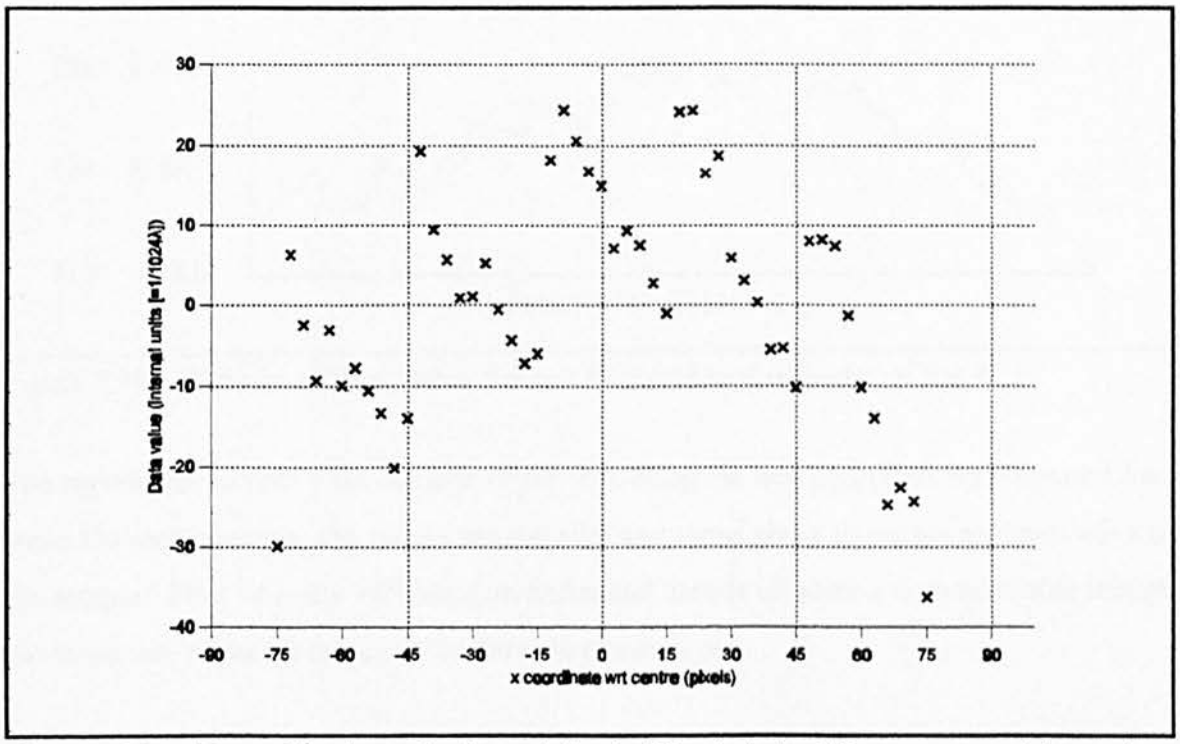


Figure 3.39 Every 3rd point on horizontal chord of flat A derived from measured data minus linear fit (y= 60).

The curve in figure 3.38 is likely to be close to the true profile of the flat along that chord since plotting every 10th point only shows the tilt error for that chord due to the y-tilt error of positional combination ABS1 and this has been removed by subtracting the line of best linear fit. The plots for all data points and for every 3rd datapoint have a more complex nature with error contributions from the y-tilt errors of combinations ABS1 and ABS2. That being the case, the plot of figure 3.38 suggests that flat A has a convex profile with a P_V value along the chord y= 60 of approximately 32 internal units (0.03 $\lambda$). This would be consistent with 1/20 $\lambda$ accuracy of the flat as quoted by Zygo, by whom it was manufactured. The concave profile of the flat was confirmed by performing an absolute measurement of a single (horizontal) diameter of flat A using Zygo's own three-flat test software. The profile obtained is shown in figure 3.40. The vertical axis values are shown in μm and the P_V value along the horizontal diameter is 0.042 μm (=0.066 $\lambda$, =1/15 $\lambda$ at $\lambda$= 632.8 nm). Note that this value is outside the quoted accuracy for the flat.

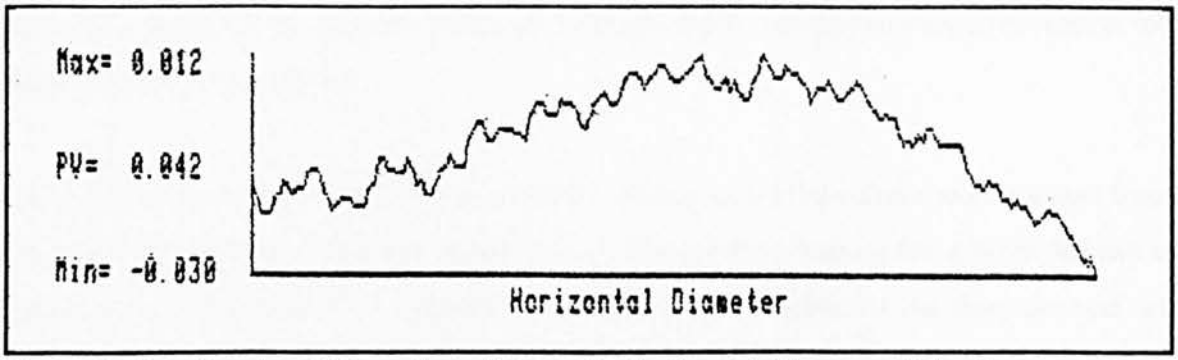


Figure 3.40 Results of Zygo three-flat test for horizontal diameter of flat A.

The experiment to derive the contour of the flats using the new algorithm was repeated four times. On each occasion, the results were similar and those given above are representative of the sample. Plots of every 10th point on horizontal chords all show a convex profile though the amplitude of the tilt errors in the data vary in each case.

Having obtained error free results using synthesized data, the source of errors when using measured data must derive from experimental errors. Sources of experimental errors are discussed in the next section.

### 3.5.1 Sources of Experimental Error.

Sources of error that may affect the results of the flatness measurement technique can be classified into two categories. Firstly there are errors introduced by the interferometer itself (not including reference surface errors). Secondly there are errors introduced by the experimental technique including environmental factors.

#### 3.5.1.1 The Effect of Interferometer Errors.

Interferometer errors will affect the accuracy of the measurements of the OPD associated with the interferometer cavity. Interferometer errors may result from a variety of sources such as phase shifter nonlinearity or camera nonlinearity as has been discussed in section 1.2.6.2.8. Since the interferometer used is a commercial instrument, the details of the individual factors which affect its measurement accuracy are not known. For example, no information is available

about any feedback mechanisms that may be employed to compensate for hysteresis in the piezo-electric phase shifters.

It is possible, however, to estimate the overall accuracy of any wavefront measurement from the manufacturer's specifications (section 3.1). The system accuracy for a three-flat test is specified as better than $\lambda/50$. Assuming the three measurements for the three-flat test are independent and applying probability theory (Boas ,1983, chapter 16) then for equation 2.4;

$$\begin{aligned}\mathrm{var}(f_A) &= \mathrm{var}\left(\frac{g_{AB}+g_{AC}-g_{CB}}{2}\right)\\ &= \left(\tfrac{1}{2}\right)^2 \mathrm{var}(g_{AB}) + \left(\tfrac{1}{2}\right)^2 \mathrm{var}(g_{AC}) + \left(-\tfrac{1}{2}\right)^2 \mathrm{var}(g_{CB})\\ &= \tfrac{1}{4}\left[\mathrm{var}(g_{AB}) + \mathrm{var}(g_{AC}) + \mathrm{var}(g_{CB})\right].\end{aligned}$$

$$\begin{aligned}\sigma_f &= \sqrt{\tfrac{3}{4}\sigma_g^2}\\ &= \sqrt{\tfrac{3}{4}}\sigma_g.\end{aligned}$$

$$\sigma_g = \sqrt{\tfrac{4}{3}}\sigma_f,$$

assuming that the individual measurement errors are statistically distributed about some mean and that the probable errors for each positional combination are the same, $\sigma_g$. The probable error for any measurement is thus $\sqrt{(4/3)} \times 1/50\ \lambda = 0.023\lambda = 23$ internal units.

Disregarding, for the moment, errors in the results of the calculations due to errors in tilt correction the errors due to the interferometer will propagate through the calculation as follows. The calculations to find the value of new points involve only simple subtractions of the values of known points from the dataset of one of the positional combinations. From the discussion in section 2.5, the number of steps required to find the value of a new point is 2(n+m) where, for the algorithm as implemented here, n is the number of 10 pixel jumps in x to reach the point from the $3m^{th}$ chord from the centre. For a flat with an approximate radius of 100 pixels, the maximum value of 2(n+m) is approximately 30. The maximum expected error in the result is then $\sqrt{30}\ \sigma_g = 125$ internal units ($0.12\ \lambda$).

The effect of interferometer errors on the determination of tilt errors will now be considered. For the correction of the tilt of the datasets for positional combinations AB, CB and AC, the tilt is determined from the values of three points on a 60 pixel radius circle using equation 2.3. The error in x and y-tilt determinations is therefore given by;

$$\begin{aligned}\sigma_{x-tilt} &= \sigma_{y-tilt} \\ &= \sqrt{\tfrac{2}{120^2}}\sigma_g \\ &= 0.012\sigma_g \\ &= 0.28 \text{ internal units/pixel.}\end{aligned}$$

For the tilt correction of the dataset for ABR, the errors in the determination of the x and y-tilt are given by;

$$\begin{aligned}\sigma_{x-tilt} &= \sigma_{y-tilt} \\ &= \sqrt{\tfrac{2}{60^2}}\sigma_g \\ &= 0.024\sigma_g \\ &= 0.54 \text{ internal units/pixel.}\end{aligned}$$

For the determination of the x-tilt correction in the data for the ABS positional combinations, two points are used, 120 pixels apart. The error is given by;

$$\begin{aligned}\sigma_{x-tilt} &= \sqrt{\tfrac{2}{120^2}}\sigma_g \\ &= 0.012\sigma_g \\ &= 0.28 \text{ internal units/pixel.}\end{aligned}$$

For the determination of the y-tilt correction in the data for the ABS positional combinations the values of the chords determined for A along a 45° diagonal are added to the corresponding values of the chords on B, rotated. The difference is found between the sum and the corresponding value from the ABR dataset (section 2.3.4.3) The average of these differences, weighted by $1/n^2$, where n is the index of the chords is the y-tilt correction;

$$y-tilt = \frac{1}{sm}\sum_{n=1}^{m}\frac{1}{n^2}\left(A_{(ns,ns)} - B_{(-ns,-ns)} - ABR_{(ns,ns)}\right)$$

The error on the values of each chord is $\sqrt{(2n)}\sigma_g$ as discussed above and so the error on the

y-tilt determination is given by;

$$\sigma_{y-tilt} = \sqrt{\frac{1}{m}\sum_{n=1}^{m}\left[\left(\frac{1}{sn^2}\right)^2 2\sqrt{2n}\sigma_g + \sigma_{g.}\right]}$$

Because of the $1/n^4$ term, a good approximation of the error may be obtained from the first (n=1) term of the summation;

$$\sigma_{y-tilt} = \frac{1}{s^2}\left(2\sqrt{2}+1\right)\sigma_g$$
$$= 0.9 \text{ internal units / pixel for ABS1}$$
$$= 10 \text{ internal units / pixel for ABS2}$$

The magnitudes of these probable errors are sufficient to explain the y-tilt errors encountered when analysing data from measurements made with the interferometer. In fact the errors seen in the results are much smaller than this (about 0.07 internal units per pixel x-tilt error) suggesting that the manufacturer specifications on which the estimates are based may be rather conservative. The section on recommended further work (section 4.2.2) will suggest means by which the effect of interferometer errors may be reduced.

It is assumed that the imaging system in the interferometer is free from geometrical distortions. If this is not the case then errors may be introduced due to loss of registration between the coordinate system of the flats and the coordinate system of the data as the chords are stepped across the aperture. The geometrical distortion would have to be greater than the pixel spacing for this to have a large effect on the results and it is hoped that the assumption of low distortion is justified. An investigation to confirm this optimism will be suggested in section 4.2.4.

#### 3.5.1.2 Errors Due to Experimental Technique.

Aspects of experimental technique that may introduce errors into the measurements may be classified into two categories. Firstly, perturbations of the interferometer cavity due to environmental disturbances would affect the measurement accuracy. Secondly, misalignments of the test flats forming the interferometer cavity could affect the correct interpretation of the data during the execution of the processing algorithm.

Perturbations of the interferometer cavity during the acquisition of the data may arise from vibration, instabilities of the mounting hardware, air currents or thermal effects. Vibration effects are minimized by mounting the interferometer and mounting hardware on a pneumatically isolated optical table and by the absence of the operator at the moment the data are acquired. Mounting hardware instabilities should be minimal since the mounts are rigid and stably supported on the same platform as the interferometer. Air currents are minimized by the shielding of the interferometer cavity and the absence of the operator during the acquisition of data, as has been described above. Thermal effects are minimized by the temperature controlled environment of the laboratory and by minimal handling of the hardware between data acquisitions. Any thermal effects due to handling were minimized by the wearing of cotton gloves and leaving an interval of 15 minutes between handling and initializing the acquisition process. It is believed that the precautions taken to minimize environmental disturbances should be such as to allow the interferometer's accuracy to attain its full potential.

For a test flat diameter of 100mm, the field of view of the interferometer covers a diameter of approximately 200 pixels on the CCD camera. The separation between pixels thus corresponds to approximately 0.5 mm. This implies that, to achieve correct registration between the surfaces of the flats and their coordinate systems assumed by the processing algorithm, the flats should be positioned to an accuracy of better than 0.5 mm in x and y. For flats A and C, it is assumed that the bayonet cells in which they are mounted and the registration of the mounting hardware are sufficient to ensure adequate positioning accuracy. This may not be the case in practice. For the positioning of flat B, the accuracy relies on alignment of the fiducial marks on its mounting cell with cursor marks on the fringe monitor. It is unlikely that this method can

be relied upon to position the flat to an accuracy of better than 0.5mm. Techniques by which the registration accuracy of the flats may be improved will be suggested in the section on recommended further work (section 4.2.3).

# Chapter 4: Conclusions and Suggestions for Further Work.

## 4.1 Conclusions.

In Chapter 1 the need for accurate metrology of optical surfaces was identified. Since this metrology is usually accomplished by the comparison of the surface with a reference surface, the need for absolute measurement methods was established. A review of the current state of the art of absolute measurement techniques, with the emphasis on the measurement of flat surfaces, was presented. The need for an improved flatness measuring algorithm suited to the characteristics of the interferometric measuring instruments and, in particular, their detector geometry was identified.

Such an algorithm was developed in Chapter 2. The algorithm was based on and developed from the basic three-flat absolute measurement technique described in Chapter 1. Extra positional combinations of the three flats were introduced involving a 90° rotation of one flat about the optical axis and a lateral translation perpendicular to the optical axis. Reduction of the data from measurement of these positional combinations in a Fizeau-type interferometer resulted in the derivation of the surface figure of each flat on a square grid of points. Proof of the validity of the algorithm was provided by a small scale manual demonstration of the algorithm using arrays of randomly generated numbers as simulated surface figure functions. The principle of reducing the propagation of experimental errors by introducing extra positional combinations with different lateral shifts was introduced.

In Chapter 3 the implementation of the algorithm developed in Chapter 2 was described. Some detailed attention was given to an example of the diagnostic procedures used in de-bugging the software since this accounted for a good deal of the effort that went into program development. During the development of the implementation of the algorithm, it became clear that the Zygo Mark IV computer hardware and Zygo's proprietary computer language, ZIPL, have serious shortcomings with regard to their capabilities for the processing required. These shortcomings forced a rather unnatural program structure and resulted in an extremely long run time. More serious, however, was the lack of numeric precision due to the integer

arithmetic which forced the somewhat clumsy "fix" described in section 3.4.1. With the benefit of hindsight, it would have been advantageous to have used the Zygo hardware for data acquisition only with the data reduction functions handled by an auxiliary computer. Indeed, this is recommended below as a prerequisite to further development of the flatness measurement algorithm.

Once the program had been debugged and the problem of the lack of numerical precision had been overcome, results obtained from synthesized data confirmed the validity of the algorithm. The results obtained from real measurements of pairs of flats suffered from large errors. Possible sources of these errors were identified as the accumulation of interferometer errors through the data processing and also as the result of errors in the mechanical positioning of the flats. Means by which these errors may be reduced are recommended below with other suggestions for further work.

Ultimately, even if the propagation of errors through the data processing and errors due to experimental technique can be reduced to negligible proportions, the accuracy of the result will be limited by the basic accuracy of the interferometer. The basic accuracy of the Zygo Mark IV is quoted by the manufacturers as being $\lambda/50$ which sets a limit on the ultimate accuracy attainable by any algorithm. To achieve more accurate results it would be necessary to obtain (or construct !) an interferometer with significantly higher basic accuracy.

## 4.2 Suggestions for Further Work.

### 4.2.1 Transfer of Data Processing to PC.

A large proportion of the effort that has gone into this research project has been involved with the development and de-bugging of the software to perform the flatness testing algorithm. The difficulties inherent in this process have been due largely to the nature of the computer hardware and Zygo's ZIPL programming language. Given that the program as it now stands takes some five hours to process the data, it is felt that further program development using the Zygo processor and ZIPL would be unwieldy and lead to an even more unacceptably long run

time.

It is suggested that the role of the Zygo processor and ZIPL be limited to the functions of data acquisition. The raw data thus acquired by the Zygo system would then be transferred via the processor's serial (RS232 standard) communication link to an auxiliary (probably an IBM PC compatible) computer for processing. The processing algorithm may be implemented in any of the many programming languages available (BASIC, C, etc.).

The advantages to using a PC to perform the processing functions of the algorithm would be manyfold:

- Perhaps the most important benefit would be the availability of floating point arithmetic to eliminate the problem of lack of numerical precision due to ZIPL's integer arithmetic.
- A PC would benefit from increased memory and hard disk storage which would remove the processing bottleneck associated with the need to transfer data to and from the slow floppy disk media in the ZIPL program.
- The necessity to load data to the DATA system variable in ZIPL would be removed. This would simplify the program structure since each dataset could be immediately available in its own array or file. Since the main reason for the slow execution of the ZIPL program is believed to be the slow access to DATA on a point-wise basis, the execution of the algorithm should be vastly speeded up.
- Since program development and subsequent data processing could take place on any PC remote from the interferometer, program development and many experiments would not be subject to the availability of the Zygo system.
- One possible approach to data processing on a PC would be to program a spreadsheet package (Microsoft Excel or Borland Quatro Pro, for example) to execute the algorithm. This would have the advantage that it would provide a built-in, user friendly means to examine the data, both numerically and graphically, during and after processing.
- A PC software package could process data originating from Fizeau interferometers

other than the Zygo Mark IV. It would simply be necessary to convert the data files produced by each interferometer to a common format for which the PC software is designed.

### 4.2.2 Improvements to Data Processing Algorithm.

The flatness measuring algorithm described in preceding chapters has been shown to yield correct results for synthetic data where no measurement errors exist. For data resulting from real measurements of flats where random errors are inevitable, however, the algorithm, in its present form, has been shown to be vulnerable to the accumulation of errors to the extent that the results are unacceptably inaccurate. The following improvements and modifications to the algorithm are suggested in order to reduce the propagation of errors during data reduction. It is suggested that these improvements are implemented after the transfer of processing functions from the Zygo processor to a PC to facilitate easier program development and de-bugging.

The correction of tilt in the raw data for positional combinations AB, CB, AC and ABR as the algorithm currently stands, depends upon finding the equation of a plane passing through three points. The error in the coefficients of the plane is thus highly dependent on the errors in the height of the three points. It is suggested that, rather than define the nominal plane of each flat in terms of three points, it be defined as the best fit plane to the surface of the flat, in a least squares sense. The coefficients of the best fit plane would then be a function of every point on the surface and would thus have a much lower error since the RMS error of a random distribution is very much less than the P-V error for a large number of points. The correction of the piston and tilt of the datasets for combinations AB, CB, AC and ABR would then involve finding the best fit plane to the measured data and subtracting that plane from the data. This is valid since the best fit to the sum of two sets of data (the individual surfaces) is equal to the sum of the best fits to each set of data. This approach can only work for the positional combinations where every point on each flat is coincident with a point on the other flat in the combination. For the shifted combinations ABS1 and ABS2 some points on each flat do not contribute to the sum of the surface figures represented by the measured wavefront and so the

best fit plane to the data will not be a true measure of the required tilt correction. The tilt correction for these datasets must therefore be carried out in the way described in Chapter 2. The accuracy of the tilt correction may, however, be increased by considering more data points. For the correction of piston and x-tilt as described in section 2.3.4.1, only two points along the already determined horizontal diameters of flats A and B were considered. By considering every point at which the diameters of flats A and B are coincident in the ABS1 and ABS2 datasets the piston and x-tilt may be found more accurately by a least squares fit.

In the ZIPL implementation of the algorithm, the y-tilt in the datasets for ABS1 and ABS2 was found by deriving the coefficient of the quadratic surface defined by equation 2.19. To derive the coefficient, only those points on the surface where x=y were considered. The determination of the quadratic coefficient and therefore the y-tilt would be more accurate by considering every point on the surface and finding the best fit quadratic surface, in the least squares sense.

The extent to which the measurement errors accumulate at a point depends on the number of steps required to derive the value of the surface figure function of the flat at that point. The algorithm, as it stands, reduces the maximum number of steps required to find the surface figure function by using two different lateral shifts for combinations ABS1 and ABS2 (section 2.5). The maximum number of steps may be further reduced by adding a third laterally shifted positional combination (or more). If all three lateral shifts were mutually prime integer numbers of pixels then the number of steps would be reduced by an extension of the argument put forward in section 2.5. An alternative approach would be to make the third lateral shift the product of (or integer multiple thereof) the original two shifts. The profiles of chords found using this large shift would have small errors since they would reach the edges of the flats in a small number of steps. Some of the chords found from the smaller shifts would be coincident with those found from the large shift and so their errors could be compensated and controlled.

It is suggested that the efficacy of these enhancements be tested experimentally by introducing random errors to simulated datasets from which the results are to be derived. This would not be easy to achieve in a ZIPL program since there is no random number generating function

available. Most, if not all, programming languages for the PC have a random number generating function.

### 4.2.3 Improvements to Experimental Procedure.

As discussed in section 3.5.1.2, the most likely sources of errors due to experimental technique arise from errors in the positioning of the flats in the interferometer cavity. It is suggested that, to ensure proper alignment, the flats be fitted with a removable cross wire graticule. When aligning the flats, the exact position of the cross wire would be monitored by observing the phase values returned by the phase measuring process. Where the cross wire obscures the interferometer cavity, "bad" values will be returned and the position of the flat will be adjusted so that these values fall along the desired x and y axes of the pixel array. When mounted in the reference flat position in the interferometer, a flat is adjustable only for x and y-tilt. In order to ensure, therefore, that the cross wire positions are consistent for the flats mounted in the reference position (flats A and C), means must be provided either for x-y translation of the flat in this position or for x-y translation of the cross wire on the flat. Additionally, rotation of the cross wire must be allowed for to ensure that the axes of the pixel array and the cross wire are parallel. For the flats mounted in the test position, the necessary degrees of freedom are provided by the mounting hardware. The cross wires must, of course, be removed when data measurements are being made and so a means must be provided whereby they may be repeatably re-positioned. This may be achieved by means of a kinematic type mount (for a discussion of kinematic mounts, see section 6.2.2).

### 4.2.4 Miscellaneous Suggestions.

In section 3.5.1.1, it was mentioned that an assumption upon which the flatness measuring algorithm relies is that there be negligible geometrical distortion in the imaging system of the interferometer. It is suggested that an experiment be carried out to confirm (or otherwise) the validity of this assumption. This might be achieved simply by placing a calibrated artifact (such as a graticule with a grid or checkerboard pattern) in the interferometer cavity. The function mapping the coordinate system of the flats to the coordinate system of the interferometer's

CCD array would then be determined by examining the position of the artifact's features in the acquired data array.

The algorithm, as it stands, is designed to find the absolute contours of the flats at every point on the measured array. For a 100 mm diameter flat covering an approximately 200 pixel diameter area on the CCD array this corresponds to a square array of approximately 0.5 mm pitch. According to the Nyquist sampling theorem, therefore, the surface figure functions can be determined up to a maximum spatial frequency of 1 cycle per mm. It is likely that smoothly polished optical flats will have surface figure functions with maximum spatial frequency components much less than 1 cycle per mm (see section 1.2.4). If this is the case then an array of data values for every CCD pixel contains a large amount of data that is redundant in an accurate description of the surface figure. It is suggested that the spatial frequency spectrum of the surfaces of the flats to be measured be investigated. This could be achieved by measuring the spatial frequency spectrum of the wavefront produced by a pair of flats forming an interferometer cavity. This would be valid since it is highly unlikely that the highest spatial frequencies of the two flats would cancel each other. Cancellation of lower spatial frequencies is more likely but this is not the section of the spectrum of interest in this case. It is irrelevant from which flat the high frequency components originate since it is the maximum spatial frequency component within the population of flats to be measured which is of interest. The spatial frequency spectrum of the data would be derived by performing a Fourier transform. If this experiment revealed that the maximum spatial frequency component of the population of flats to be measured was much less than that requiring measurement at every pixel then the algorithm could easily be modified, by a different choice of lateral shifts, to calculate the contour on a sparser grid of points. This would result in a lower accumulation of errors since the maximum number of steps to calculate the surface figure function at any point on the grid would be reduced. The values of the surface figure function at intermediate pixel positions would be found by interpolation with negligible loss of accuracy.

## Chapter 5: Discussion of Other Approaches to the Problem of Absolute Flatness Measurement.

In previous chapters the discussion of the measurement of flatness has considered that the absolute contours must be derived from relative measurements made by conventional (Fizeau) interferometers. It has been shown that the absolute contours may be derived from a number of relative measurements between pairs from a population of several flats. The derivation of the contours requires a good deal of data reduction which has been shown, inevitably, to lead to the accumulation of experimental errors, though these may be minimized. In this chapter the possibility of a more direct approach to the measurement of flatness will be considered. To this end, several techniques have been identified which may, individually or in combination, lend themselves to flatness measurement. Techniques to be considered include; the Ritchey-Common test, common-path interferometry, shearing interferometry, photo-refractive crystals for phase conjugation and wavefront storage, and profilometry. Of these, profilometry is an odd one out since it generally involves examining a surface on a point by point basis. The other techniques lend themselves to a whole-surface examination. In order to simplify the structure of the discussion, the separate techniques will first be described in isolation. Any promising avenues consisting of a combination of techniques will be discussed afterwards.

### 5.1 The Ritchey-Common Test.

Given that the normals to a nominally flat surface are nominally parallel it seems natural that an optical test of that surface should be conducted in parallel or collimated light whose wavefront is also flat. The problem with this approach lies in the difficulty of generating an error free collimated beam of light (one with a perfectly flat wavefront). In general the best that can be done is to collimate a diverging spherical wavefront with a lens or concave mirror which is likely to introduce aberrations into the beam, however small. In the Fizeau interferometer, the aberrations in the collimated wave are largely compensated for if the interference cavity is nulled and the test and reference surfaces are close together. Under these conditions the test and reference wavefront nominally travel the same path and so the aberrations due to the collimating optics are cancelled so long as they are small. Unfortunately, as has been shown

in previous chapters, the aberrations due to the reference surface are difficult to determine.

In contrast to a collimated wavefront, a perfect diverging spherical wave is easy to generate by a spatial filter. A laser beam focused onto an aperture (pinhole) whose diameter is equal to or smaller than the Airy spot size will diverge after the aperture, in the far field, as a perfect spherical wave having had any aberrated components filtered out in the focal plane by the pinhole. One test for flat surfaces utilizing a divergent spherical wavefront is the Ritchey-Common test (Ritchey,1904, Shu,1983). The basic arrangement of the Ritchey-Common test is shown in figure 5.1.

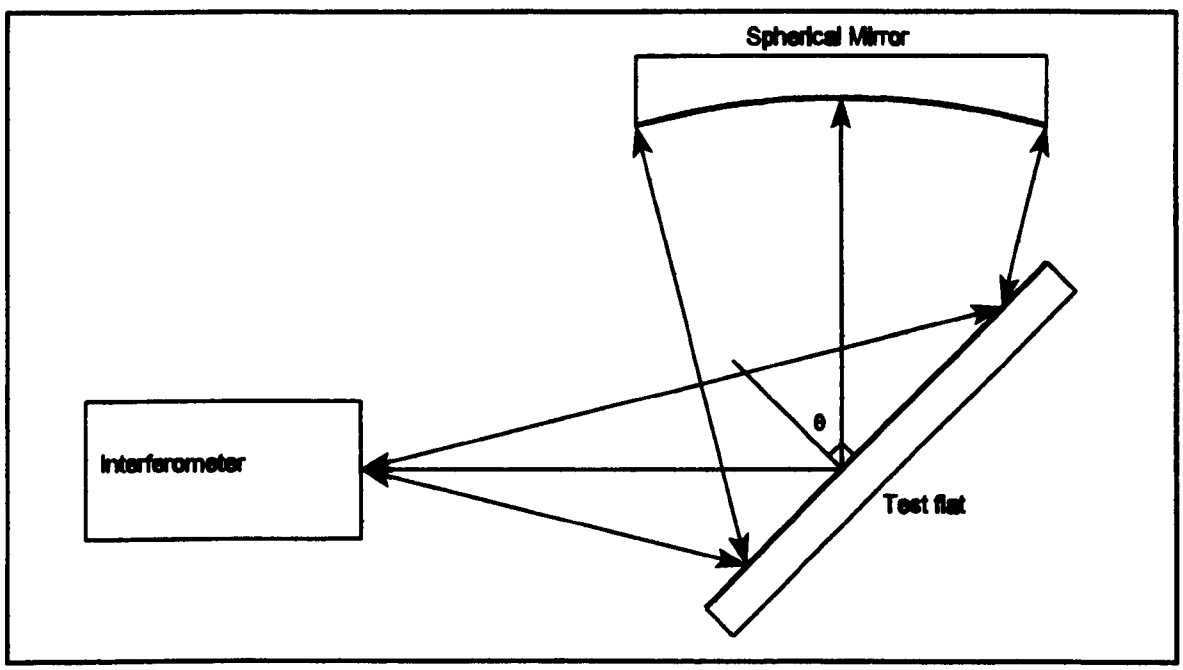


Figure 5.1 Ritchey-Common test.

A diverging spherical wave is generated by the interferometer and reflected from the test flat at some angle of incidence, θ. A spherical mirror is placed with its centre of curvature at the image of the point source and reflects the test wave back to the interferometer where the wavefront is measured. The measured wavefront contains information about the interferometer errors, the spherical mirror errors and the test flat errors. If the errors due to the interferometer and the spherical mirror are previously determined then the test flat errors may be determined except for those terms contributing to tilt and focus. The errors due to the interferometer and spherical mirror may be found relatively straightforwardly where the interferometer is of the Fizeau or Twyman-Green type by the absolute sphericity test proposed by Jensen (1973).

Jensen's sphericity test is fully described in Appendix A (section 6.3.2.1).

Since tilt and focusing are adjustment parameters for the test setup, they cannot be determined by measurement. It is possible, however to determine these "free" parameters unambiguously by computation provided that the results of the measurements taken at two different angles of incidence are combined as described by Kuchel (1986). The test configurations used for the measurements are shown, slightly simplified, in figure 5.2.

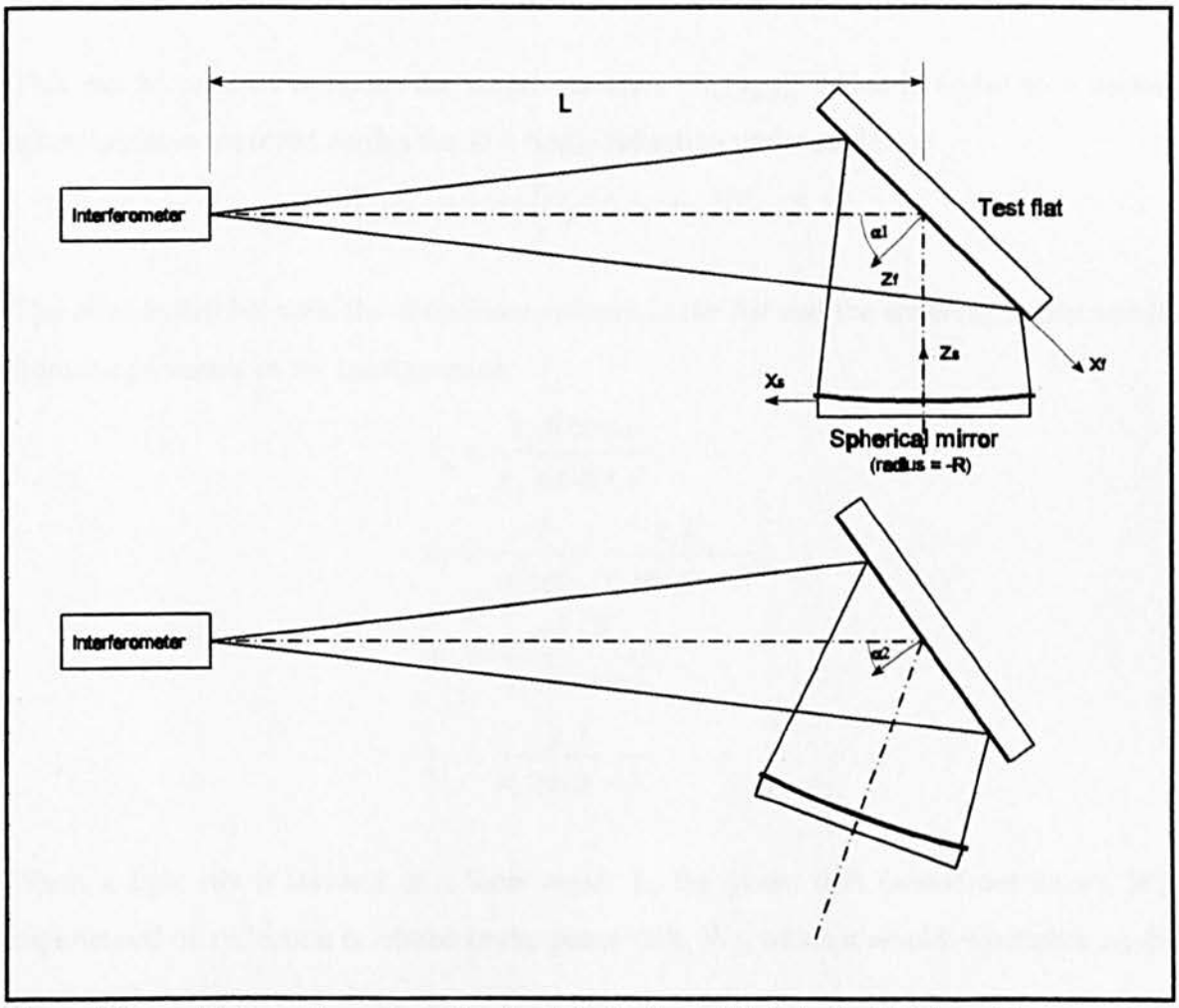


Figure 5.2 Test configurations for absolute testing of flats in the Ritchey- Common test.

The vertex distance, L, from the point source to the test flat is measured along the optical axis. Two coordinate systems are defined; $x_f$, $y_f$, $z_f$ in the vertex of the test flat and $x_s$, $y_s$, $z_s$ in the vertex of the spherical mirror which has a radius of curvature of -R. Wavefront measurements are made at two different angles of incidence, $\alpha_1$ and $\alpha_2$. It is assumed that the wavefront errors due to the auxiliary components (the interferometer and the spherical mirror) have been

found previously by means of Jensen's test or otherwise.

The reference coordinate system for the measurements in the Ritchey-Common configuration is $x_s$, $y_s$, $z_s$ in the vertex of the spherical mirror. The two configurations shown are used to measure the wavefronts $U_{\alpha 1}$ $(x_s,y_s)$ and $U_{\alpha 2}$ $(x_s,y_s)$. The two wavefronts include the wavefront error, H $(x_s,y_s)$ and the wavefront error caused by the test flat.

$$U_\alpha(x_s,y_s) = H(x_s,y_s) + 2W_\alpha(x_s,y_s).$$

This can be used to compute the wavefront error $W_\alpha(x_s,y_s)$ which is added to a perfect spherical wave mirrored on this flat in a single reflection under angle $\alpha$;

$$W_\alpha(x_s,y_s) = \tfrac{1}{2}\left[U_\alpha(x_s,y_s) - H(x_s,y_s)\right].$$

The relationship between the coordinate systems in the flat and the spherical mirror results from the geometry of the configuration;

$$x_s = \frac{x_f R\cos\alpha}{x_f \sin\alpha + L},$$

$$x_f = \frac{-1}{\cos\alpha} \times \frac{x_s L}{x_s \tan\alpha - R},$$

$$y_s = \frac{-y_f R}{x_f \sin\alpha + L},$$

$$y_f = \frac{y_s L}{x_s \tan\alpha - R}.$$

When a light ray is incident at a local angle, $\beta$, the phase shift (wavefront error), $W_\beta$, experienced on reflection is related to the phase shift, $W_\perp$, which it would experience on the same surface at normal incidence;

$$W_{\perp} = \frac{W_{\beta}}{\cos\beta},$$

$$W_{\perp}(x_f, y_f) = \frac{W_{\alpha}(x_f, y_f)}{\cos\left[\beta(x_f, y_f)\right]},$$

where

$$\beta = \arctan\left[\frac{\left(x_f + L\sin\alpha\right)^2 + y_f^2}{L\cos\alpha}\right]^{\frac{1}{2}}.$$

The wavefront error, $W_{\alpha}(x_s,y_s)$, may thus be unambiguously converted into a test flat inherent wavefront error, $W_{\perp}(x_f,y_f)$. This, however does not unambiguously describe the flat surface since those components of $W_{\alpha}(x_s,y_s)$ describing piston, tilt and focus are dependent upon the test setup and cannot be determined by measurement. Varying defocussing components, $d\times(x_s^2+y_s^2)$, in $W_{\alpha}(x_s,y_s)$ lead to different results for $W_{\perp}(x_f,y_f)$. These free parameters may, however, be determined by computation from the results of two measurements at different angles of incidence. A non-linear optimization routine is used for the determination of the focusing terms $d_1$ and $d_2$ in$W_{\alpha 1}(x_s,y_s)$ and $W_{\alpha 2}(x_s,y_s)$ such that the wavefront errors $W_{\perp 1}(x_f,y_f)$ and $W_{\perp 2}(x_f,y_f)$ show the best possible correspondence.

$$W_{\alpha 1}(x_s, y_s) = R_{\alpha 1}(x_s, y_s) + d_1 \times (x_s^{\ 2} + y_s^{\ 2})$$
$$W_{\perp 1}(x_f, y_f) = R_{\perp 1}(x_f, y_f) + e_1 x_f + f_1 y_f + g_1$$

$$W_{\alpha 2}(x_s, y_s) = R_{\alpha 2}(x_s, y_s) + d_2 \times (x_s^{\ 2} + y_s^{\ 2})$$
$$W_{\perp 2}(x_f, y_f) = R_{\perp 2}(x_f, y_f) + e_2 x_f + f_2 y_f + g_2$$

$$\Phi(d_1, d_2) = \sum_{i,j}\left[R_{\perp 1}(x_{fi}, y_{fj}) - R_{\perp 2}(x_{fi}, y_{fj})\right] \rightarrow \text{minimum}.$$

Where d, e, f and g are focus, x-tilt, y-tilt and piston terms respectively. $R_{\alpha 1}$ and $R_{\alpha 2}$ are the wavefront errors obtained by measurement from which the piston, tilt and power terms are eliminated. For the computation of the merit function, $\Phi$, the tilt and piston included in $W_{\perp}$ are not taken into account. Since the shape of the test flat is the same in the two measurements, the focus terms are optimized until the difference between the functions $R_{\perp 1}$ and $R_{\perp 2}$, $\Phi$ is a

minimum.

Kuchel reports that, using angles of incidence of 38° and 52°, several flats of diameter 356 mm were measured with an RMS uncertainty of less than 1 nm.

The Ritchey-Common test for measuring flatness is of particular interest since several of the techniques to be considered below show promise for the absolute measurement, at the centre of curvature, of spherical mirrors or spherical wavefronts.

## 5.2 The Point Diffraction Interferometer.

An interesting form of interferometer which relies on the fact that a perfect spherical wave is produced on diffraction from a small circular aperture is the Point Diffraction Interferometer (PDI). The PDI was first described by Linnik (1933) and rediscovered by Smartt (Smartt and Steel, 1975). An English language translation of Linnik's original paper appears (without figures) in a paper by Speer et al (1979) describing the use of a PDI for measuring grazing incidence X-ray optics. The principle of the PDI is shown in figure 5.3. A converging wave from the system under test comes to a focus on a neutral density (ND)filter which has a small pinhole aperture. Light falling on the pinhole is diffracted to form a diverging spherical reference wavefront. The rest of the incident wavefront passes attenuated, but otherwise unchanged through the filter where it interferes with the spherical reference wave to give a fringe pattern directly representative of the deviations of the incident wavefront from a truly spherical wave. Tilt fringes may be introduced by displacing the pinhole laterally from the centre of the focused beam as shown in figure 5.3. Similarly, de-focus fringes may be introduced by displacing the pinhole axially from the focus of the incident wave. A simple point diffraction interferometer is manufactured and marketed by Ealing Electro-Optics Inc (89 Doug Brown Way, Holliston, MA 01746, USA) consisting of an ND filter with pinhole, a three axis positioning mount and ground glass viewing screen. The optimum size of the pinhole is about the size of the Airy disk that would be produced by the incident wavefront if it were free from aberrations. The optimum attenuation of the ND filter is chosen so that the test and reference waves are of equal intensity, so maximizing the fringe visibility. Smartt and Steel

(1975) recommend transmittances for the ND filter of between 0.005 and 0.05 with 0.01 being a useful general purpose value.

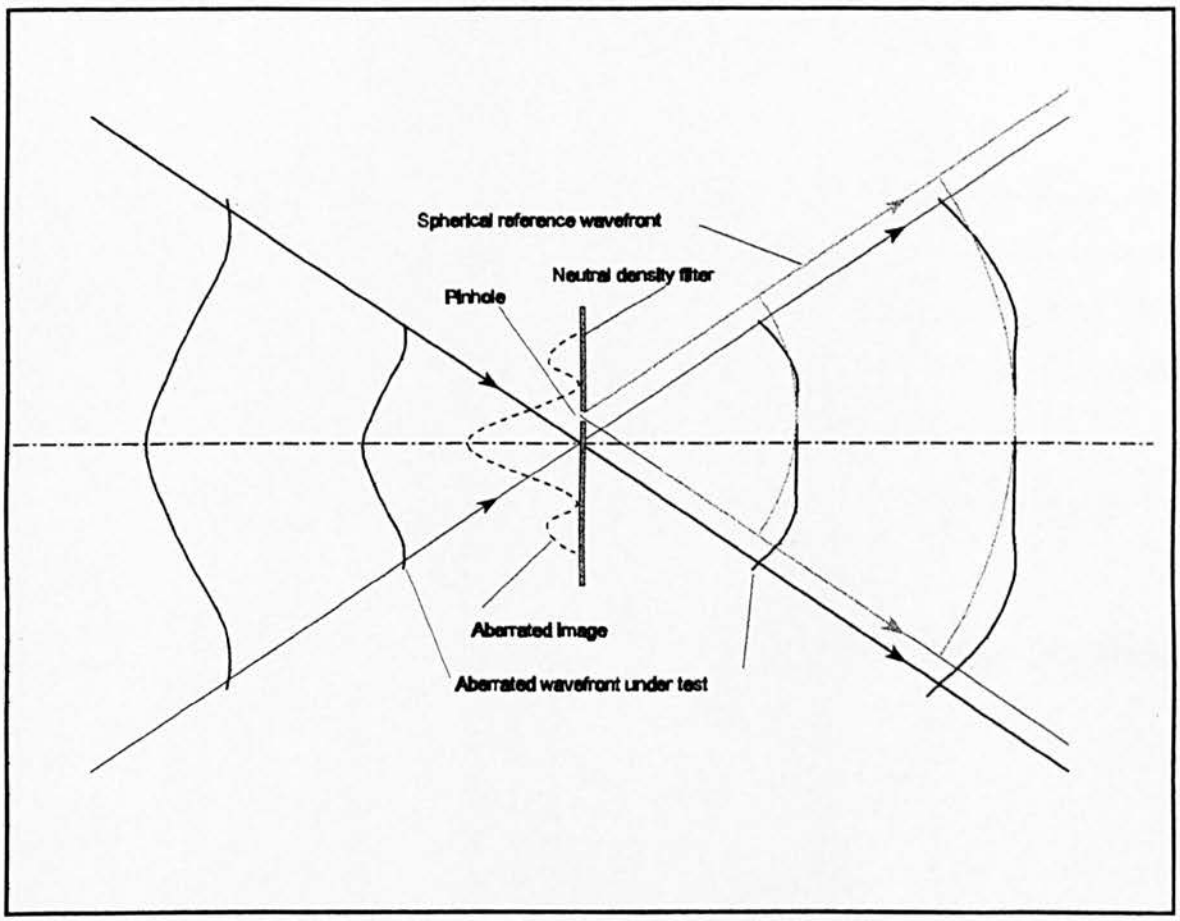


Figure 5.3 Principle of the point diffraction interferometer (PDI).

The operation of the point diffraction interferometer is similar to that of the phase contrast interferometer (Zernike, 1943). In the phase contrast interferometer the phase of the reference wave is shifted relative to the aberrated wave. In the point diffraction interferometer, the amplitude of the two waves are altered.

The complex amplitude of the incident wave may be written as;

$$A(x,y) = \exp\left[ i \frac{2\pi}{\lambda} W(x,y) \right].$$

where W(x,y) is the aberration function of the wavefront. This may be re-written as;

$$A(x,y) = \exp\left[i\frac{2\pi}{\lambda}\overline{W}(x,y)\right] + \left\{\exp\left[i\frac{2\pi}{\lambda}W(x,y)\right] - \exp\left[i\frac{2\pi}{\lambda}\overline{W}(x,y)\right]\right\}$$

where $\overline{W}(x,y)$ represents an ideal spherical wavefront. The complex amplitude is thus expressed as a perfect wavefront (first term) plus aberrations (second term). Since the two terms are differently distributed in the image plane, they may be separately modified. This is the technique used in the phase contrast and point diffraction interferometers. The complex amplitude of the wave, having passed through the interferometer is given by;

$$A'(x,y) = \alpha\exp(i\delta)\exp\left[i\frac{2\pi}{\lambda}\overline{W}(x,y)\right] + \exp\left[i\frac{2\pi}{\lambda}W(x,y)\right] - \exp\left[i\frac{2\pi}{\lambda}\overline{W}(x,y)\right]$$

where $\alpha$ is the amplitude transmittance and $\delta$ the phase difference introduced to the reference wave. The irradiance of the fringe pattern is then given by;

$$I(x,y) = \alpha^2 + 4\sin^2\left\{\frac{\pi}{\lambda}\left[\overline{W}(x,y) - W(x,y)\right]\right\} - 4\sin\left\{\frac{\pi}{\lambda}\left[\overline{W}(x,y) - W(x,y)\right]\right\} \times\alpha\sin\left\{\frac{\pi}{\lambda}\left[\overline{W}(x,y) - W(x,y)\right] + \delta\right\}.$$

When $\alpha = 1$ and $\delta = 0$, the trivial case of unity irradiance over the entire field results.

For the point diffraction interferometer, the irradiance function may be obtained by setting $\delta = 0$;

$$I(x,y) = \alpha^2 + (1-\alpha)4\sin^2\left\{\frac{\pi}{\lambda}\left[\overline{W}(x,y) - W(x,y)\right]\right\}.$$

For a basic PDI consisting of a pinhole in an absorbing film the phase shift, $\delta$, is constant (conveniently regard as zero as above). This means that the phase shifting methods of

evaluating the wavefront (section 1.2.6.2) may not be applied. Spatial methods of wavefront analysis (section 1.2.6.3) may be applied if large numbers of tilt fringes are introduced into the interference pattern as described above.

In order to apply phase shifting techniques to the PDI, a number of workers (Wu et al, 1984, Kadono et al, 1987, Mercer and Creath, 1996) have developed versions where the phase of the reference wave may be shifted relative to the test wave. Wu et al and Kadono et al describe a polarization technique for performing the phase shifting. The technique of Kadono et al is shown in figure 5.4.

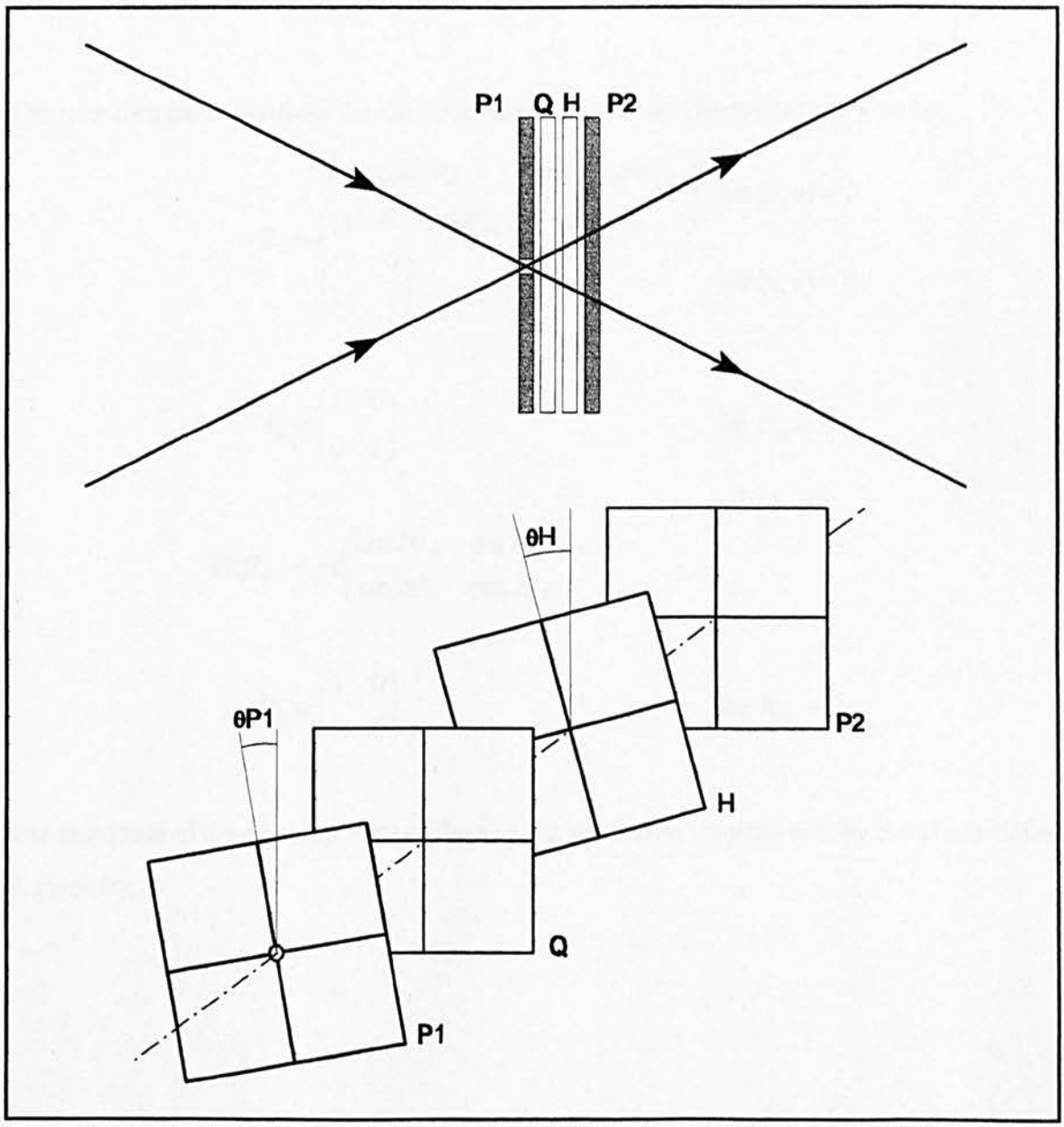


Figure 5.4 Polarization technique for phase shifting PDI (Kadono et al, 1987).

The neutral density filter with pinhole of the basic PDI is replaced by a linear polarizing film, P1, oriented at an angle $\theta_{P1}$ to the vertical, with a small pinhole. The incident wavefront is plane polarized at an angle, $\phi$, to the vertical. The dc portion of the incident wave, $W_{dc}$ (the reference wave) passes through the pinhole and the aberrated portion, $W_{ac}$ (the object wave) passes through the polarizer. Both waves then pass through a quarter wave plate, Q, with its fast axis vertical, a half wave plate with its fast axis at an angle $\theta_H$ to the vertical and then a polarizer oriented vertically. The amplitude of the wave in front of P1 is given, using the Jones vector as;

$$A_0(x,y) = [W_{dc}(x,y) + W_{ac}(x,y)]\begin{pmatrix}\cos\phi \\ \sin\phi\end{pmatrix}.$$

The transformation matrices for the components of the interferometer are give by;

$$\mathbf{P}_1 = \begin{cases} \begin{pmatrix} \cos^2\theta_{P1} & \sin\theta_{P1}\cos\theta_{P1} \\ \sin\theta_{P1}\cos\theta_{P1} & \sin^2\theta_{P1} \end{pmatrix} & \text{for } (x,y) \neq 0 \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \text{for } (x,y) = 0 \end{cases}$$

$$\mathbf{Q} = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \qquad \text{for } \theta_Q = 0$$

$$\mathbf{H}(\theta_H) = -i\begin{pmatrix} \cos 2\theta_H & \sin 2\theta_H \\ \sin 2\theta_H & \cos 2\theta_H \end{pmatrix}$$

$$\mathbf{P}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{for } \theta_{P2} = 0$$

The amplitude of the dc component of the incident wave after transmission by the phase shifter is given by;

$$W'_{dc} = \mathbf{P}_2\mathbf{H}(\theta_H)\mathbf{Q}W_{dc}\begin{pmatrix}\cos\phi\\ \sin\phi\end{pmatrix}$$

$$= W_{dc}\begin{pmatrix}\sin\phi\sin 2\theta_H - i\cos\phi\cos 2\theta_H\\ 0\end{pmatrix}$$

$$= W_{dc}\begin{pmatrix}f(\theta_H,\phi)\exp[i\psi_{dc}(\theta_H)]\\ 0\end{pmatrix}$$

where

$$\psi_{dc}(\theta_H) = \tan^{-1}(\tan\phi\tan 2\theta_H) - \tfrac{\pi}{2},$$

$$f(\theta_H,\phi) = \sqrt{(\sin^2\phi\sin^2 2\theta_H + \cos^2\phi\cos^2\theta_H)}.$$

Similarly, the amplitude of the ac component of the transmitted wave is given by;

$$W'_{ac} = \mathbf{P}_1\mathbf{P}_2\mathbf{H}(\theta_H)\mathbf{Q}\mathbf{P}_1W_{ac}\begin{pmatrix}\cos\phi\\ \sin\phi\end{pmatrix}$$

$$= \cos(\theta_{P1}-\phi)W_{ac}\mathbf{P}_2\mathbf{H}(\theta_H)\mathbf{Q}\begin{pmatrix}\cos\phi\\ \sin\phi\end{pmatrix}$$

$$= \cos(\theta_{P1}-\phi)W_{ac}\begin{pmatrix}\sin\theta_{P1}\sin 2\theta_H - i\cos\theta_{P1}\cos 2\theta_H\\ 0\end{pmatrix}$$

$$= W_{ac}\begin{pmatrix}\cos(\theta_{P1}-\phi)f(\theta_H,\phi)\exp[i\psi_{ac}(\theta_H)]\\ 0\end{pmatrix}$$

where

$$\psi_{ac} = \tan^{-1}(\tan\theta_{p1}\tan 2\theta_H) - \tfrac{\pi}{2}.$$

When the angle $\theta_{P1}$ of linear polarizer P1 is chosen to be $\theta_{P1} = -\phi$, the relative phase difference between the ac and dc components is given by;

$$\Delta\psi = \psi_{ac}(\theta_H) - \psi_{dc}(\theta_H)$$
$$= -2\tan^{-1}(\tan\phi\tan 2\theta_H).$$

Thus the relative phases of the reference and object waves may be varied by rotating half wave plate, H and a phase shifting algorithm (section 1.2.6.2) may be employed to determine the object wavefront, $W_{ac}$.

Mercer and Creath (1996) describe a phase shifting point diffraction interferometer utilizing a film of liquid crystal material as the phase shifting element. The principle of the liquid crystal point diffraction interferometer is illustrated in figure 5.5.

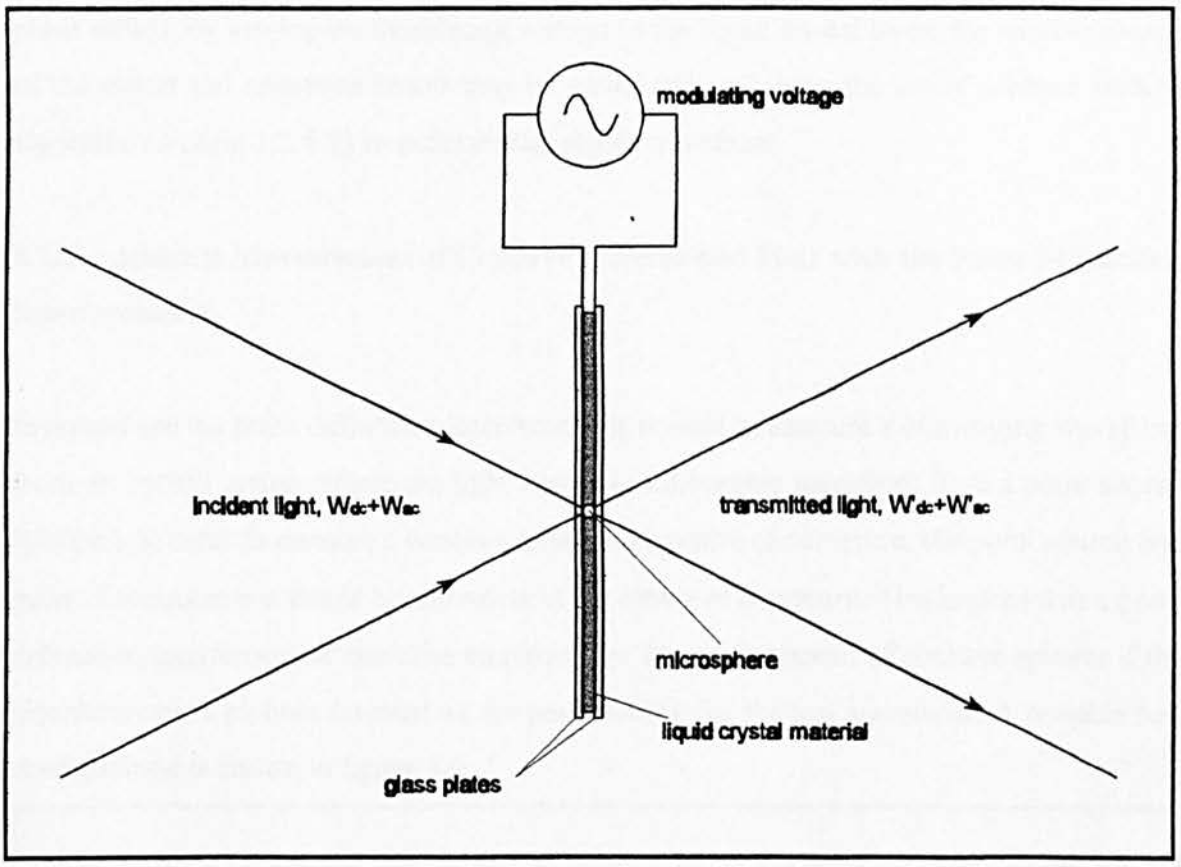


Figure 5.5 Mercer and Creath's liquid crystal point diffraction interferometer.

A 9 μm thick film of nematic liquid crystals is sandwiched between two thin glass plates with transparent conductive coatings on their inner surfaces. A transparent plastic microsphere of 9 μm diameter is embedded in the liquid crystal layer and acts as the diffracting aperture for the reference wave. The object beam is phase shifted by modulation of the voltage across the liquid crystals, which alters their refractive index. The glass plates are prepared so that the birefringent liquid crystals are homogeneously aligned with their directors (long axes) oriented vertically, parallel to the plates. This configuration allows phase modulation of vertically polarized light travelling through the layer. Horizontally polarized light is not phase shifted. The uniaxial liquid crystal layer has a refractive index equal to the extraordinary refractive index for light polarized parallel to the aligned directors. As the amplitude of the applied electric field increases, the molecules rotate and the refractive index of the layer shifts towards

the ordinary index of refraction. The refractive index equals the ordinary refractive index when the molecules are aligned perpendicular to the plates (parallel to the electric field). The light diffracted by the embedded microsphere does not pass through the liquid layer and so is not phase shifted. By varying the modulating voltage to the liquid crystal layer, the relative phases of the object and reference beams may be varied, thus allowing the use of a phase shifting algorithm (section 1.2.6.2) to evaluate the object wavefront.

### 5.2.1 Absolute Measurement of Concave Spheres and Flats with the Point Diffraction Interferometer.

In normal use the point diffraction interferometer is used to measure a converging wavefront from an optical system where the light source is a diverging wavefront from a point source (pinhole). In order to measure a concave sphere at its centre of curvature, the point source and point of measurement should be coincident at the centre of curvature. This implies that a point diffraction interferometer could be employed for the measurement of concave spheres if the interferometer's pinhole doubled as the point source for the test wavefront. A possible test configuration is shown in figure 5.6.

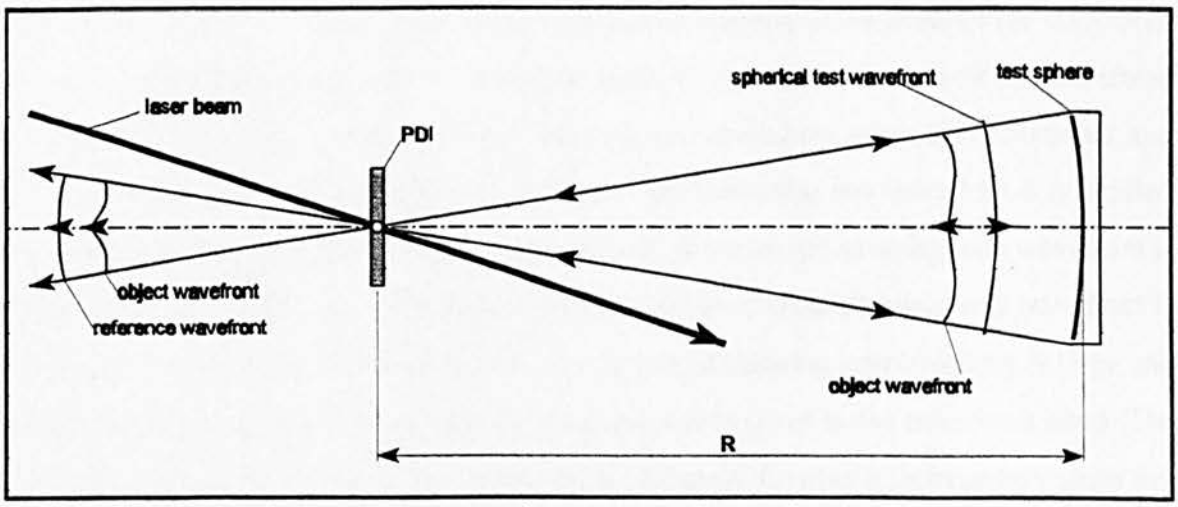


Figure 5.6 Absolute measurement of a concave sphere using the PDI.

A laser beam is incident on the pinhole of the PDI and is diffracted to form a diverging spherical test wavefront. The laser beam is oriented in such a way that the undiffracted portion escapes from the optical system in order to avoid spurious light interfering with the desired wavefronts. The test sphere, positioned with its centre of curvature coincident with the

pinhole, reflects the test wavefront to give the aberrated object wavefront. The object wavefront passes through the PDI and a portion of it is diffracted at the pinhole to form the spherical reference wavefront. The interference pattern formed by the object and reference waves contains information only about the aberrations introduced by the test sphere and so gives an absolute measurement of the sphere contours.

In order to measure flat surfaces, the arrangement for the absolute testing of spheres may be modified to encompass the Ritchey-Common test (section 5.1) whereby the path of the test beam is folded by the flat.

## 5.3 Shearing Interferometry.

In shearing interferometry, there is no separate reference surface or reference beam with which the object wavefront is compared. Instead the object beam is split and one half displaced (or sheared) in some fashion. The two beams are then recombined to form an interference pattern. The object beam is thus compared with a sheared version of itself. The wavefront shear may take one of a number of forms, the most common of which is a lateral shear (Briers, 1972, Mantravadi, 1992b) where the object beam is displaced laterally in the plane of the wavefront (if the wavefront is nominally plane) or rotated about its centre of curvature (if the wavefront is nominally spherical). Other forms of shearing interferometry are radial, rotational and reversal shear (Malacara, 1992, Briers, 1972). In radial shearing one wavefront is magnified with respect to the other before they are recombined. In rotational shearing, one wavefront is rotated about the optical axis with respect to the other. In reversal shearing, one wavefront is reversed about a diameter of the wavefront. The subject of shearing interferometry is large and a full discussion is beyond the scope of this thesis and is given in the references cited. The discussion here will be restricted to the basic theory of lateral shearing interferometry since this has possible applications for flatness measurement in combination with other techniques.

The principle of lateral shearing interferometers is shown for collimated and converging light in figure 5.7

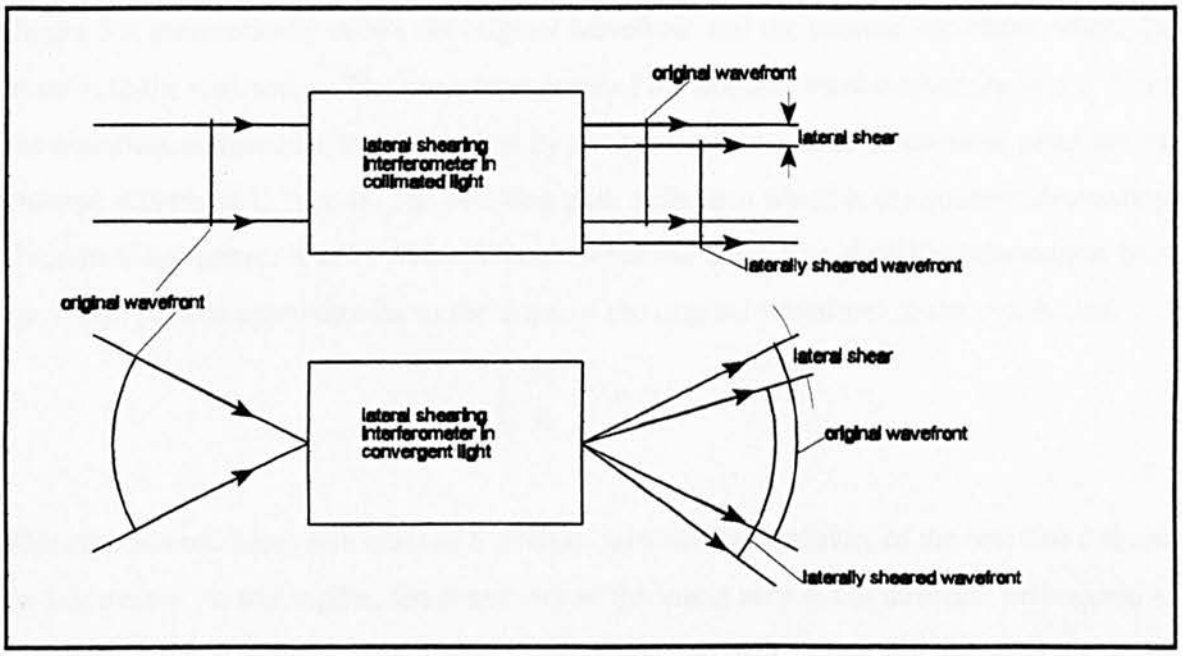


Figure 5.7 Schematic illustration of lateral shearing in collimated and converging light.

Very many optical arrangements exist for producing lateral shear. Two simple arrangements for collimated and convergent light, based on the Michelson interferometer are shown in figure 5.8.

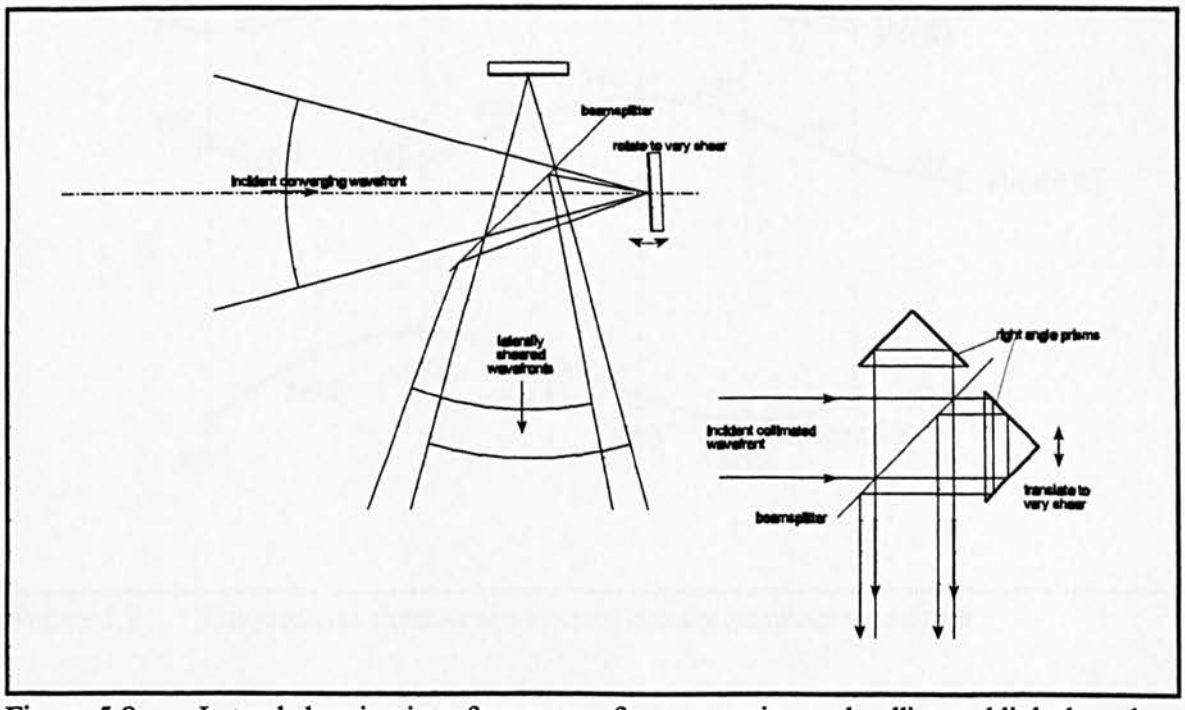


Figure 5.8 Lateral shearing interferometers for converging and collimated light based on the Michelson interferometer.

Figure 5.9 schematically shows the original wavefront and the sheared wavefront where the shear is in the x-direction. The wavefront error of the original wave is given by W(x). When the wavefront is sheared in the x direction by an amount S, the error at the same point for the sheared wavefront is W(x-s). The resulting path difference which is the quantity determined from the fringe pattern is $\Delta W = W(x) - W(x-s)$. When the shear, S, is small the information from the fringe pattern approximates to the slope of the original wavefront in the x-direction;

$$\left(\frac{\delta W}{\delta x}\right) S = n\lambda.$$

This equation becomes more exact as S tends to zero but the sensitivity of the test also reduces as S decreases. As this implies, the sensitivity of the test is zero in the direction orthogonal to the shear direction and to find the slope in both x and y directions two interferograms must be measured with shears in the x and y directions.

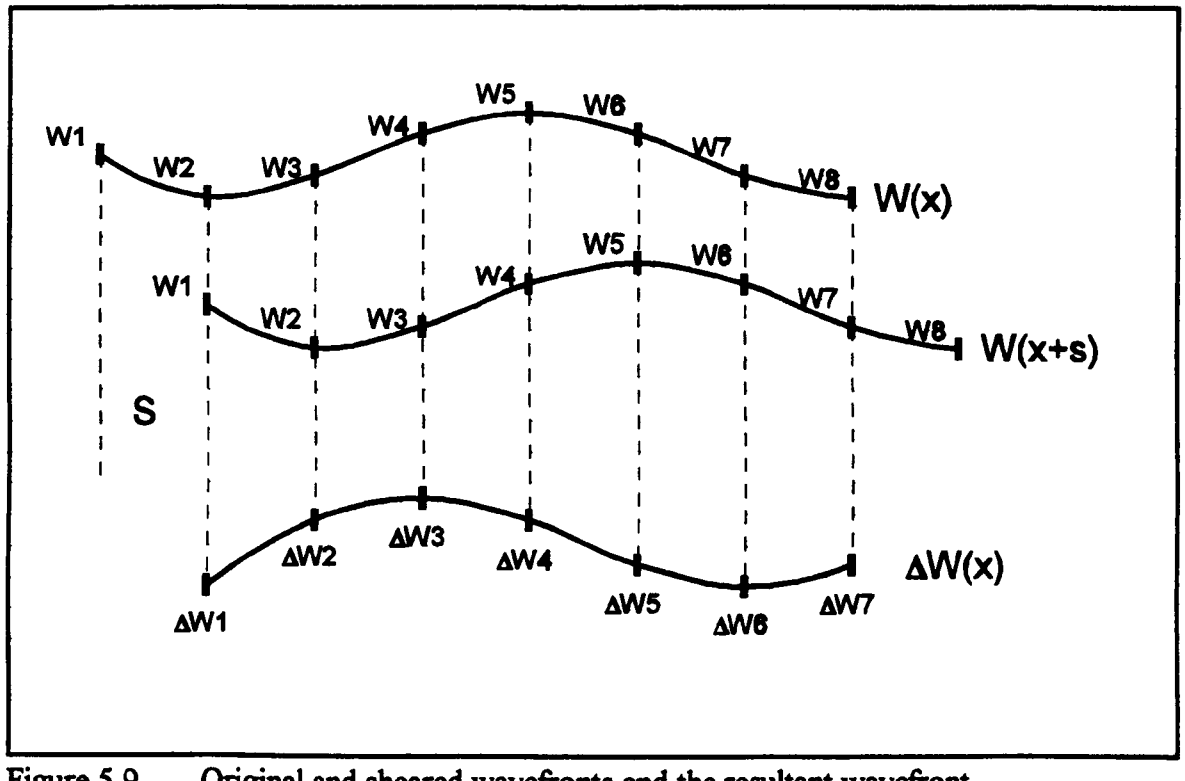


Figure 5.9 Original and sheared wavefronts and the resultant wavefront.

Since lateral shearing interferometry does not measure the shape of the wavefront directly, the data obtained from the interferograms must be processed to recover the original wavefront

shape. Where S is small and the data approximate the wavefront slope, the wavefront shape is obtained by integrating the slope. Various methods have been described for evaluating lateral shearing interferograms. Rimmer (1974) describes a method whereby the wavefront may be determined on a rectangular grid from the data obtained from two orthogonal shears. Fischer and Stahl (1993) describe a method whereby the interferometric slope data are fit to the derivatives of the Zernike polynomials (section 1.2.5.1) in order to derive a representation of the original wavefront in terms of the Zernike polynomials.

The use of shearing interferometry to measure flat surfaces may be achieved by shearing the wavefront reflected from the surface. There are, however, difficulties in making the test absolute because the original and sheared wavefronts must, by definition, follow different paths through whatever optical system is employed to produce the shear. Since the interferometer optics will inevitably contain unknown residual aberrations the two wavefronts will be aberrated differently and the test will not be absolute. The absolute calibration of a shearing interferometer is likely to be at least as difficult a task as the calibration of a conventional interferometer and this, together with the fact that the wavefront is not directly determined, makes the shearing interferometer an unlikely candidate for absolute measurement. It will be seen, however, in section 5.4.2, that shearing may be possible without auxiliary optics by the use of photorefractive non-linear optical devices.

## 5.4 Photorefractive Non-Linear Optics.

Photorefractive non-linear optical media are materials in which the refractive index depends on the local electric field and hence the local amplitude of light in the material (Feinberg, 1988). The effect is non-linear since its magnitude varies with the magnitude of the field. Most materials are linear at normal light intensities and hence show no non-linear characteristics at intensities less than those produced by focused high power lasers. A few materials, however, exhibit marked photorefractive effects at the intensities produced by low power light such as that produced by the helium-neon laser. Single crystal Barium Titanate ($BaTiO_3$) is the most efficient of these materials for use with a He-Ne laser. Other photorefractive media requiring the higher intensities available from an Argon ion laser are Lithium Niobate ($LiNbO_3$), Bismuth

Silicon Oxide (BSO) and Strontium Barium Niobate (SBN).

If two or more light waves overlap within a photorefractive medium an interference pattern results. The local refractive index of the medium is modulated by the intensity variations of the interference pattern and so a phase grating is written into the material which mimics the interference pattern. Diffraction by the phase grating will couple light from each incident light wave to the others (a process known as mixing). This mixing effect may be used to generate the phase conjugate of any incident wavefront (Yeh, Chapter 6,1993, Boyd et al, 1987).

The principle of four wave mixing to produce phase conjugation is shown in figure 5.10. Two beams are incident on, and interfere inside, a photorefractive medium. Beam 1 is called the object beam which contains spatial information and beam two is called the reference beam and is a plane wave.

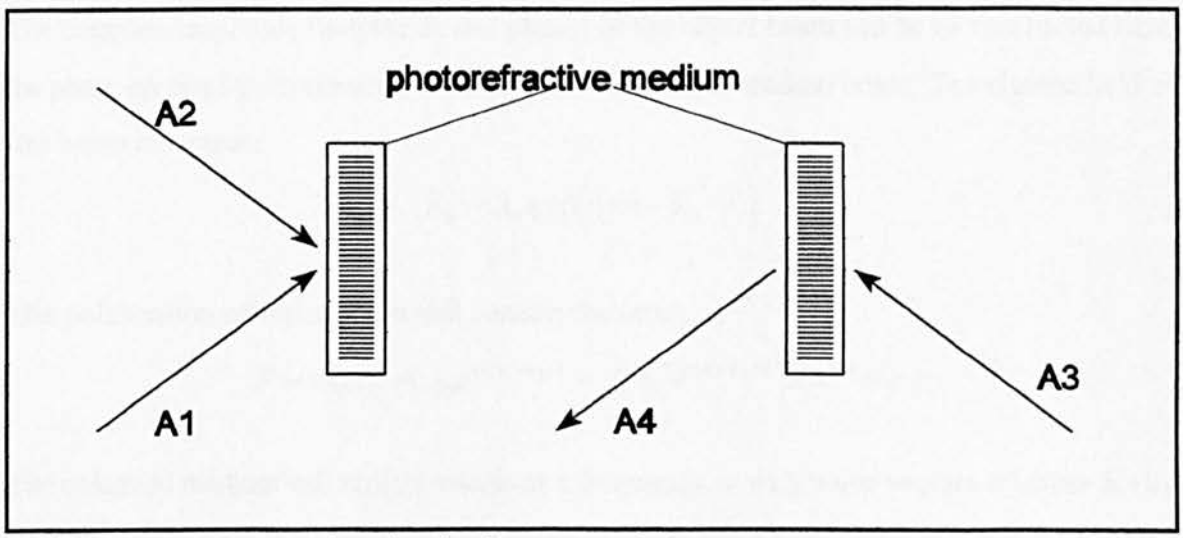


Figure 5.10 Phase conjugation by four wave mixing.

The electric fields of the beams are written;

$$E_1 = A_1 \exp\left[i\left(\omega t - \mathbf{k}_1 \cdot \mathbf{r}\right)\right]$$
$$E_2 = A_2 \exp\left[i\left(\omega t - \mathbf{k}_2 \cdot \mathbf{r}\right)\right]$$

where $\mathbf{k}_1$ and $\mathbf{k}_2$ are wave vectors.

As a result of the interference between the two beams, the intensity can be written;

$$I = |A_1|^2 + |A_2|^2 + A_1^* A_2 e^{-i\mathbf{K}\cdot\mathbf{r}} + A_1 A_2^* e^{i\mathbf{K}\cdot\mathbf{r}}$$

where $\mathbf{K} = \mathbf{k}_2 - \mathbf{k}_1$.

In the photorefractive medium the intensity variation results in a variation of refractive index. Assuming the change in the index, $\Delta n$, to be proportional to the intensity, then $\Delta n$ may be written;

$$\Delta n = n_0 |A_1|^2 + n_0 |A_2|^2 + \left\{ n_0 A_1 A_2^* e^{i\mathbf{K}\cdot\mathbf{r}} + \text{c.c.} \right\}$$

where c.c. denotes the complex conjugate of the preceding term. The wavefront information of the object beam is thus recorded in the medium in terms of $\Delta n$, a phase grating. The process is exactly analogous to the recording of an object wave in holography.

The complex amplitude (amplitude and phase) of the object beam can be reconstructed from the phase grating by illuminating the medium with a third, readout beam. The electric field of this beam is written;

$$E_3 = A_3 \exp\left[ i(\omega t - \mathbf{k}_3 \cdot \mathbf{r}) \right].$$

The polarization of the medium will contain the term;

$$P = n n_0 \varepsilon_0 \left[ A_1^* A_2 e^{-i(\mathbf{K}+\mathbf{k}_3)\cdot\mathbf{r}} + A_1 A_2^* e^{i(\mathbf{K}-\mathbf{k}_3)\cdot\mathbf{r}} \right] A_3 e^{i\omega t} + c.c.$$

The polarized medium will radiate waves at a frequency $\omega$ with wave vectors of either $\mathbf{K}+\mathbf{k}_3$ or $\mathbf{K}-\mathbf{k}_3$. If the readout beam is counter-propagating relative to the reference beam ($\mathbf{k}_3 = -\mathbf{k}_2$) then $\mathbf{K}+\mathbf{k}_3 = -\mathbf{k}_1$ and the first term of the polarization will radiate efficiently. The electric field of the radiated beam is written;

$$E_4 = A_4 \exp\left[ i(\omega t - \mathbf{k}_4 \cdot \mathbf{r}) \right]$$

where;

$$A_4 = c n_0 (A_2 A_3) A_1^*$$
$$\mathbf{k}_4 = -\mathbf{k}_1$$

$E_4$ is exactly the phase conjugate of the object wave, $E_1$, its complex amplitude, $A_4$, is proportional to the complex conjugate of the amplitude of the object wave, $A_1^*$ and its wave vector is equal and opposite to that of the object wave.

The practical result of this is that when an object beam illuminates a photorefractive medium pumped by two counter-propagating beams, the phase conjugate of the object beam is generated. The shape of the phase conjugate wave is exactly the same as that of the object wave but propagates in the opposite direction. The phase conjugate wave will exactly retrace the path of the object wave and any reciprocal transformations, including aberrations, that were performed on the object wave will be cancelled in the phase conjugate wave. A photorefractive crystal used to generate the phase conjugate of an object wave is called a phase conjugate mirror (PCM).

Feinberg (1982) discovered that phase conjugation could be achieved in $BaTiO_3$ without the use of externally applied pump (reference and readout) beams. It was discovered that when the crystal was illuminated, at an appropriate angle, by the object beam, the pump beams were spontaneously generated within the crystal by scattering from defects and total internal reflection from the crystal faces. In the words of Feinberg (1988) the effect is "caused by dirt". A possible mechanism for this effect is as follows. When the object beam enters the photorefractive crystal, a certain proportion of the light is scattered from internal or surface features. Some of the scattered light will propagate at the correct angle to be totally internally reflected from the crystal faces and phase gratings will build up where the various beams overlap. Those gratings with the correct orientation to fulfill the requirements for efficient four wave mixing will reinforce the intensity of the beams which formed them by coupling light from the incident radiation. These gratings will then tend to build up at the expense of the less efficient gratings until counter-propagating pump beams are formed within the crystal and the phase conjugate of the object beam is generated. A typical arrangement of beams within a self-pumped phase conjugator (SPPC) is shown in figure 5.11.

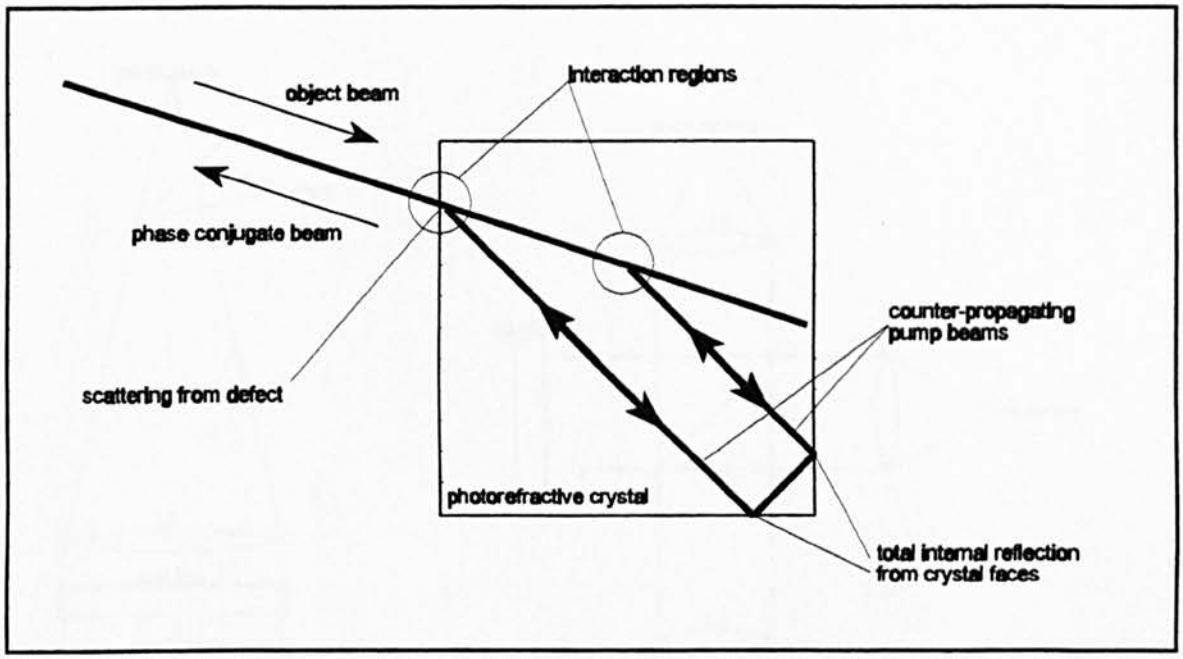


Figure 5.11 Self-pumped phase conjugator.

The fidelity of the phase conjugate beam from a SPPC is greater than that from an externally pumped crystal since the self-generated pump beams are automatically exact conjugates of each other.

### 5.4.1 Interferometers Using Phase Conjugate Mirrors.

Conventional interferometers employ combinations of lenses, mirrors and beamsplitters. In practice none of these elements is ever perfect and they introduce aberrations to the wavefronts that may add errors to the determination of the wavefront to be measured. The property of phase conjugate mirrors (PCMs) that cancel the aberrations added to an incident object beam in the phase conjugate reconstruction make it interesting to investigate the properties of interferometers employing PCMs in place of conventional mirrors (Yeh, chapter 9, 1993). The obvious point at which to start such an investigation is to replace the reference mirror in a conventional Fizeau (Gauthier et al, 1989) or Twyman-Green (Feinberg, 1983) interferometer with a PCM as shown in figure 5.12.

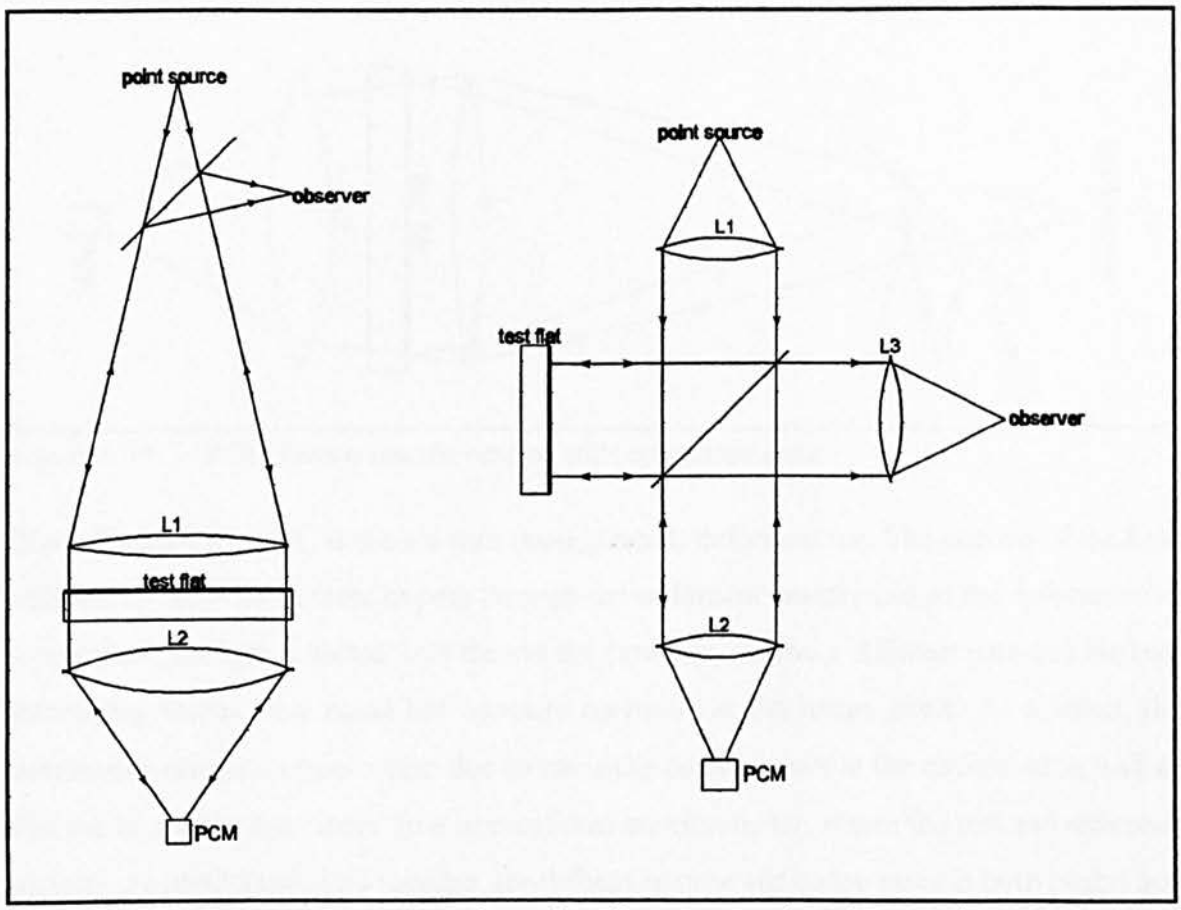


Figure 5.12 Fizeau and Twyman-Green interferometers using phase conjugate mirrors.

The quality of the lenses, L2, used to focus the light onto the PCMs is unimportant since the aberrations introduced will be cancelled on reflection. It might be imagined, at first sight, that such arrangements would yield an absolute measurement of the figure of the test flats since the PCMs apparently provide a "perfect" reference wavefront but this assumes that the collimation provided by lenses L1 is perfect. Consider the case where the collimation is not perfect shown in figure 5.13 for a Fizeau interferometer and also in figure 6.12 for a Twyman-Green interferometer.

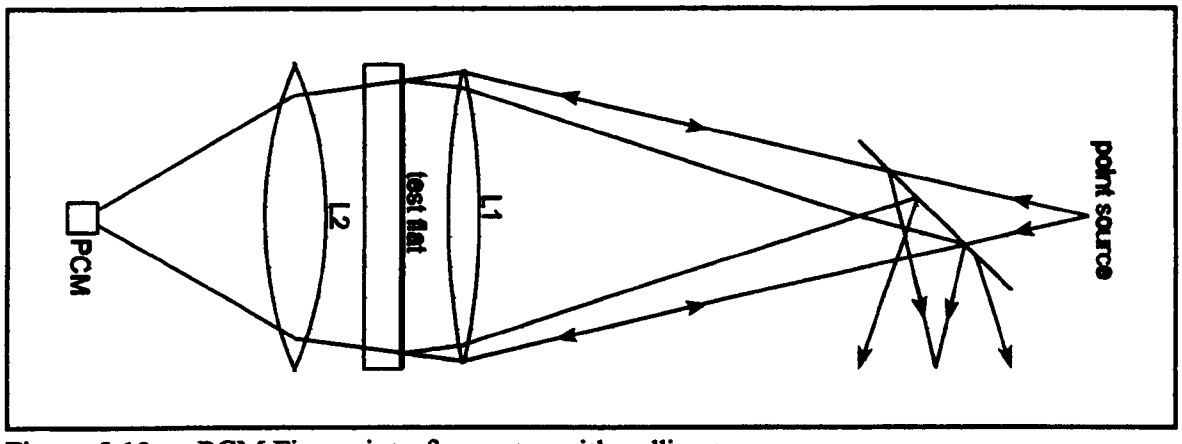

Figure 5.13 PCM Fizeau interferometer with collimator error.

The collimating lens, L1, is shown with (exaggerated) defocus error. The portion of the light reflected by the PCM retraces its path through the collimator exactly and so the defocus error is cancelled. The light reflected from the test flat, however, follows a different path and the two interfering beams have equal but opposite curvature at the image plane. As a result, the interference pattern contains a term due to twice the defocus error in the collimator as well as that due to the test flat errors. In a conventional interferometer, where the test and reference surfaces are parallel and close together, the defocus error would be the same in both beams and so would cancel in the interference pattern. The same argument applies to collimator aberrations other than defocus. Howes (1986) uses this sensitivity of the PCM interferometer to collimator errors as a test for collimation accuracy by using a Twyman-Green interferometer with PCM and assuming a high quality flat surface. It is impossible simultaneously to test for collimation accuracy and for the accuracy of the flat.

Some authors (Simova et al, 1993, Sasaki et al, 1993, Wang et al, 1994) have exploited the aberration doubling effect of interferometers with PCMs to increase the sensitivity of conventional interferometers by a factor of two. The arrangement described by Simova et al is shown in figure 5.14.

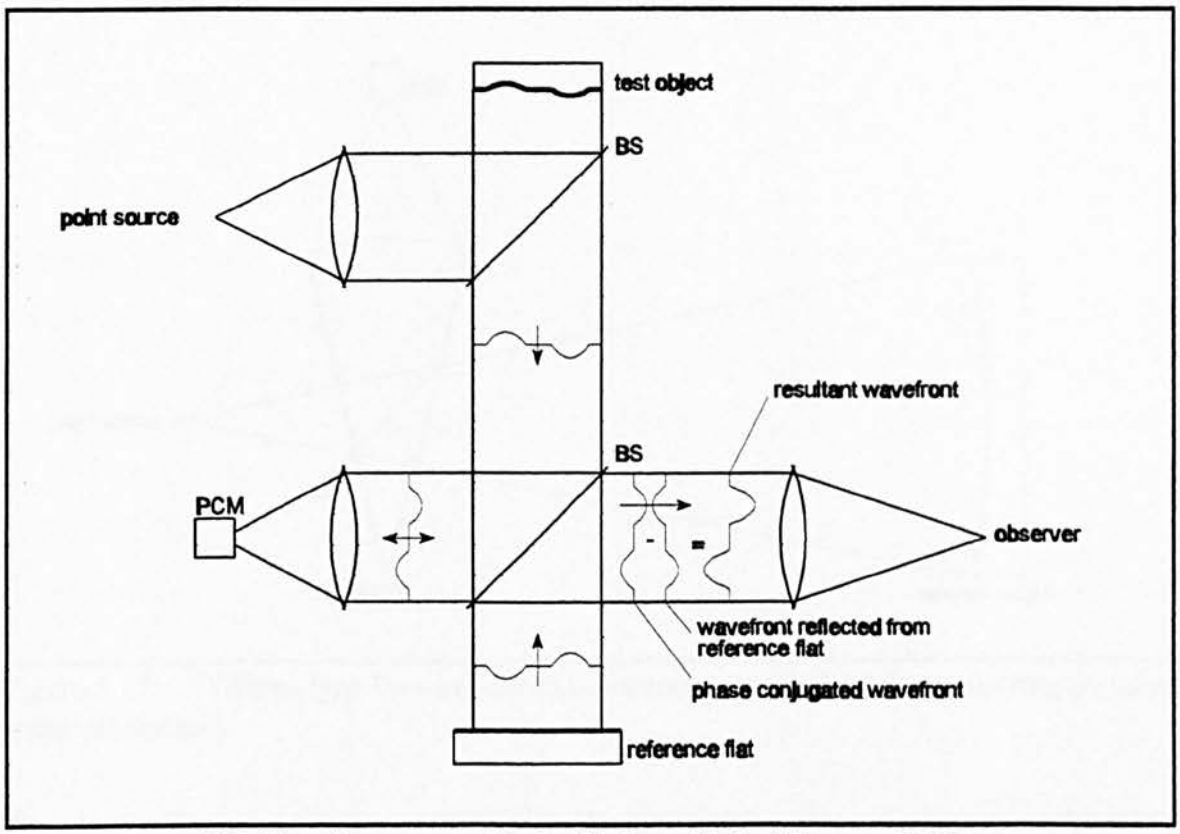


Figure 5.14 PCM interferometer with doubled sensitivity (Simova, 1993).

Light reflected from the test object is used as the input to a phase conjugate interferometer of the Twyman-Green type. The wavefronts reflected from the PCM and the reference mirrors contain equal but opposite aberrations due to the test object. The sensitivity to the aberrations caused by the test object is thus double that of a conventional Twyman-Green interferometer. The interferometer is still, of course, sensitive to the aberrations due to the reference flat and the collimating lens. The arrangements described by Sasaki et al and Wang et al are similar but employ a Fizeau PCM interferometer in place of the Twyman-Green.

To test a concave spherical surface, it is possible to dispense with the need for a collimating (or other beam shaping) lens by the use of a Twyman-Green interferometer of the Williams type (section 1.2.1, figure 1.6). Shukla et al (1990a,b) describe a phase-conjugate interferometer for testing concave spherical surfaces as shown in figure 5.15.

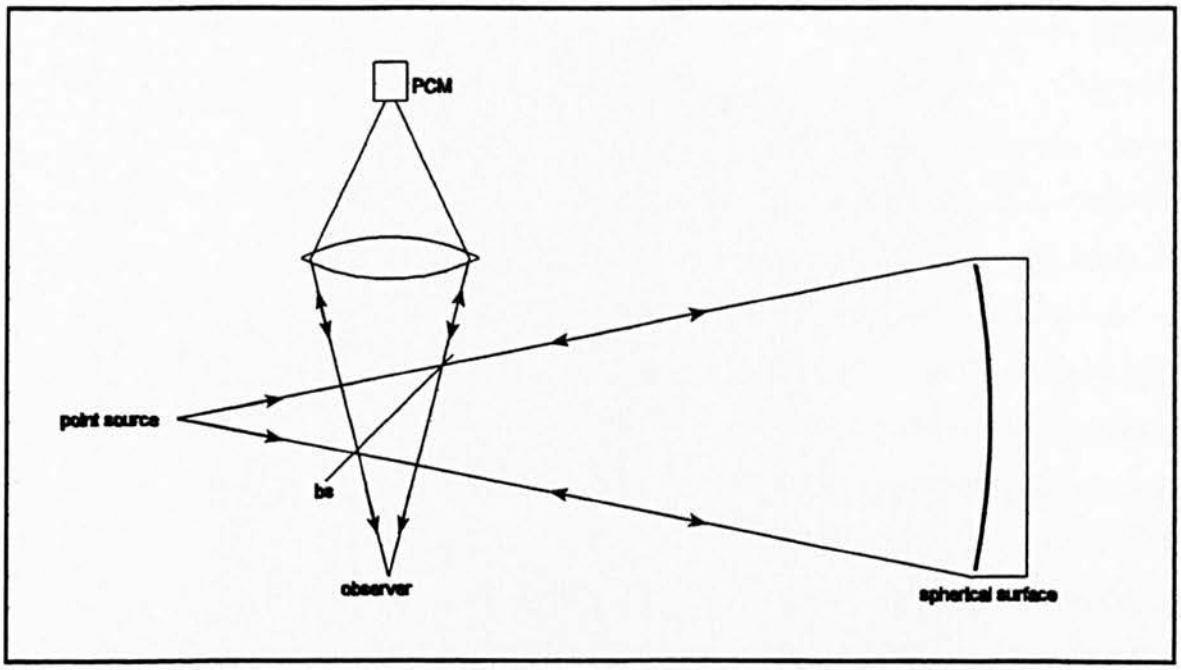


Figure 5.15 Williams type Twyman-Green interferometer with PCM for measuring concave spherical surfaces.

The interferometer would appear to offer an absolute test for concave spherical surfaces but unfortunately it is sensitive to wavefront aberrations due to the beamsplitter. If the beamsplitter aberrations could be made insignificantly small or be calibrated then the interferometer would offer the possibility of absolute flatness testing by combination with the Ritchey-Common test (section 5.1).

### 5.4.2 Photorefractive Phase Conjugators for Wavefront Storage and Shearing Interferometry.

A feature of photorefractive non-linear materials that was not mentioned in the treatment given above (section 5.4) is their slow response time. For example, at power levels of a few mW from a He-Ne laser, the phase gratings in Barium Titanate can take tens of seconds or even minutes to build up. The response time is inversely related to intensity and speeds up as the power level increases. For applications requiring the processing of rapidly changing optical signals this can be a considerable disadvantage. The effect, however, can be put to use to store and retain wavefront information for some time after that wavefront has changed or ceased to exist. For example, the phase conjugate of a wave illuminating an externally pumped

photorefractive crystal will continue to be produced for some time after the original wave has been blocked by a shutter. This effect may be used to produce an interference pattern between a wavefront and that same wavefront recorded at some earlier time. The change in the wavefront between the two recordings may be due to some dynamic effect such as fluid flow or mechanical loading of some component or may be introduced by displacing (or shearing) an optical component. Several authors have described interferometers utilizing this effect. The interferometer described by Chang et al (1988) utilizes two separate photorefractive BSO crystals to record the wavefront at different points in time which are then interfered together. Yang et al (1991), Yang and Siahmakoun (1993) and Sun et al (1996) describe interferometers where the two wavefronts are sequentially recorded in the same Barium Titanate crystal. The interferometer described by Sun et al improves the fidelity of the phase conjugate waves by generating the pump beams by self-pumped phase conjugation of a single incident external pump beam. This interferometer is illustrated in figure 5.16.

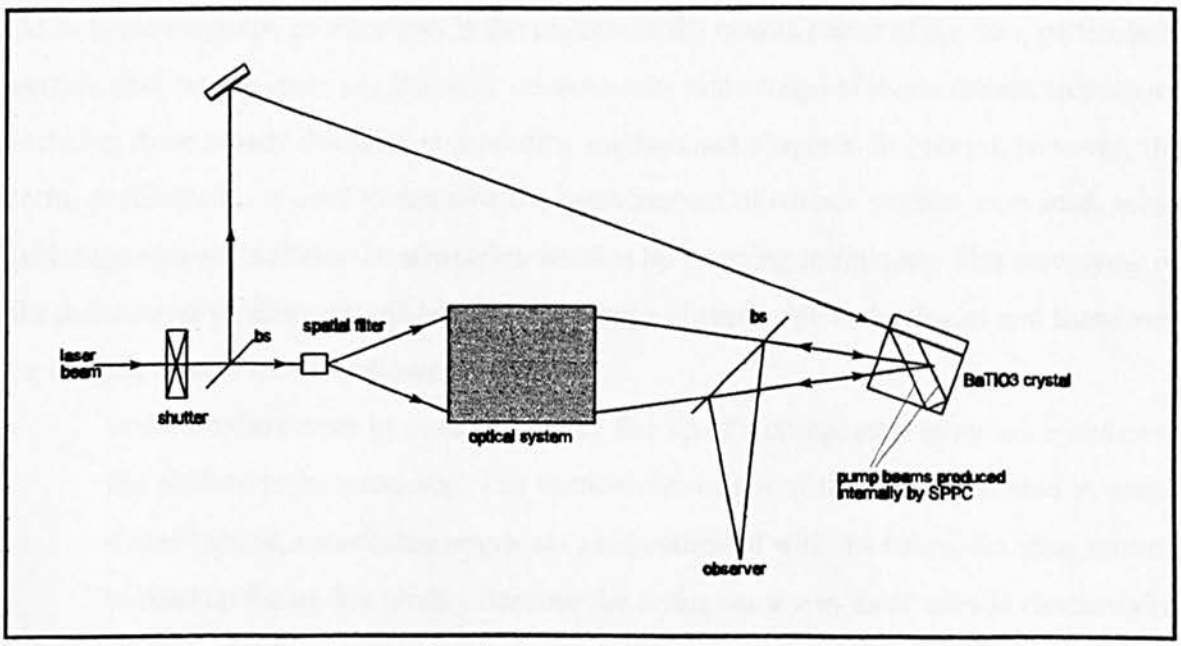


Figure 5.16 Phase conjugate shearing interferometer (Sun et al, 1996)

The object wave is incident upon the photorefractive crystal and the phase gratings build up to produce the phase conjugate wave. The phase conjugate wave is directed to the observer but, at this point, there are no interference fringes. Once the conjugate wave has built up, the shutter is closed and no light falls on the crystal though the phase gratings will persist for some time. The component in the optical system which is to be measured is then displaced by a

known distance and the shutter re-opened. The new object wave is then incident on the crystal and a new set of phase gratings begin to build up within it. For some time, both the old and new sets of phase gratings exist within the crystal and the phase conjugates of both old and new object waveforms are produced and are directed to the observation plane where they interfere. The interference pattern so produced is equivalent to a shearing interferogram but the shear is produced by a mechanical movement and not by an optical shear. The interferogram thus contains information only about the optical component. This arrangement could be used to test flats by using an object wave reflected from the flat surface and shifting the flat laterally between closing and re-opening the shutter to produce a lateral shearing interferogram.

## 5.5 Profilometry.

As its name suggests, profilometry is the science of the measurement of profiles, particularly surface profiles. As such, profilometry covers a very wide range of measurement techniques including those already described in preceding sections and chapters. In general, however, the term, profilometry, is used to describe the measurement of surface profiles over small areas with high spatial resolution or measuring profiles by scanning techniques. This narrowing of the definition of profilometry still leaves a large range of applicable technologies and these may be roughly categorized as follows;

- Stylus profilers work by tracing the very fine tip of a compliantly mounted stylus over the surface to be measured. The vertical movement of the stylus is sensed by some means (optical, capacitance sensor etc.) and combined with the lateral scanning motion to build up the surface profile. Because the stylus has a very small area in contact with the sample the contact pressure can be very high resulting in permanent surface damage, particularly of soft materials.
- Scanning probe microscopes (SPMs) are a family of instruments which profile a surface by moving a fine tip in close proximity to the surface. The first instrument of this type was the scanning tunnelling microscope (STM). The STM works by moving a conductive tip towards a conductive surface with a voltage between the two until a tunnelling current is detected. Tunnelling is a quantum mechanical effect whereby an

electrical current can flow across an insulating gap if the gap is sufficiently narrow (less than 1nm). The magnitude of the current is highly dependent on the thickness of the gap. The profile of the surface is determined by monitoring a feedback loop which keeps the current constant by controlling the height of the probe as it is scanned across the surface. STMs have vertical resolutions of Angstroms and lateral ranges of tens of μm. The other common form of SPM is the atomic force microscope (AFM). The AFM measures the attractive atomic force which occurs between the probe tip and the surface when their separation is of the order of 1 nm by detecting the deflection of a cantilever beam on which the probe is mounted. The sample need not be electrically conducting. The resolution and range of the AFM is similar to that of the STM. Bennett et al (1993) present the results of a study comparing the results obtained with an AFM to those obtained with a conventional stylus profiler on a variety of super smooth surfaces.

- Optical profilers are a large family of instruments which use light to determine the profiles of surfaces. They may be divided into two categories depending on whether they employ interferometric or non-interferometric techniques. They may also be sub-divided into three further categories of those which measure surface height directly, those which measure surface slope or those which measure surface curvature.

### 5.5.1 Non-Interferometric Optical Height Profilers.

The most common form of non-interferometric optical height profiler is the focus sensor. This type of instrument works by scanning a focused laser spot across a surface and adjusting the height of the objective lens to maintain focus as the surface height varies. The focus is determined by monitoring the shape and position of the focus of the reflected light which returns through the objective lens. The height resolution of a focus sensor depends on the depth of field, and thus the focal length of the objective lens but is not as great as that of interferometric instruments. Instruments based on focus sensors are described by Breitmeier and Ahlers (1987) and Lou et al (1984).

### 5.5.2 Interferometric Optical Height Profilers.

Interferometric optical height profilers can be divided further into two categories; those which measure profile simultaneously over a small area and those which build up a profile by scanning a measurement point over a surface.

Height profilers which measure simultaneously over an area are basically miniaturized conventional interferometers built into optical microscopes as shown in figure 5.17.

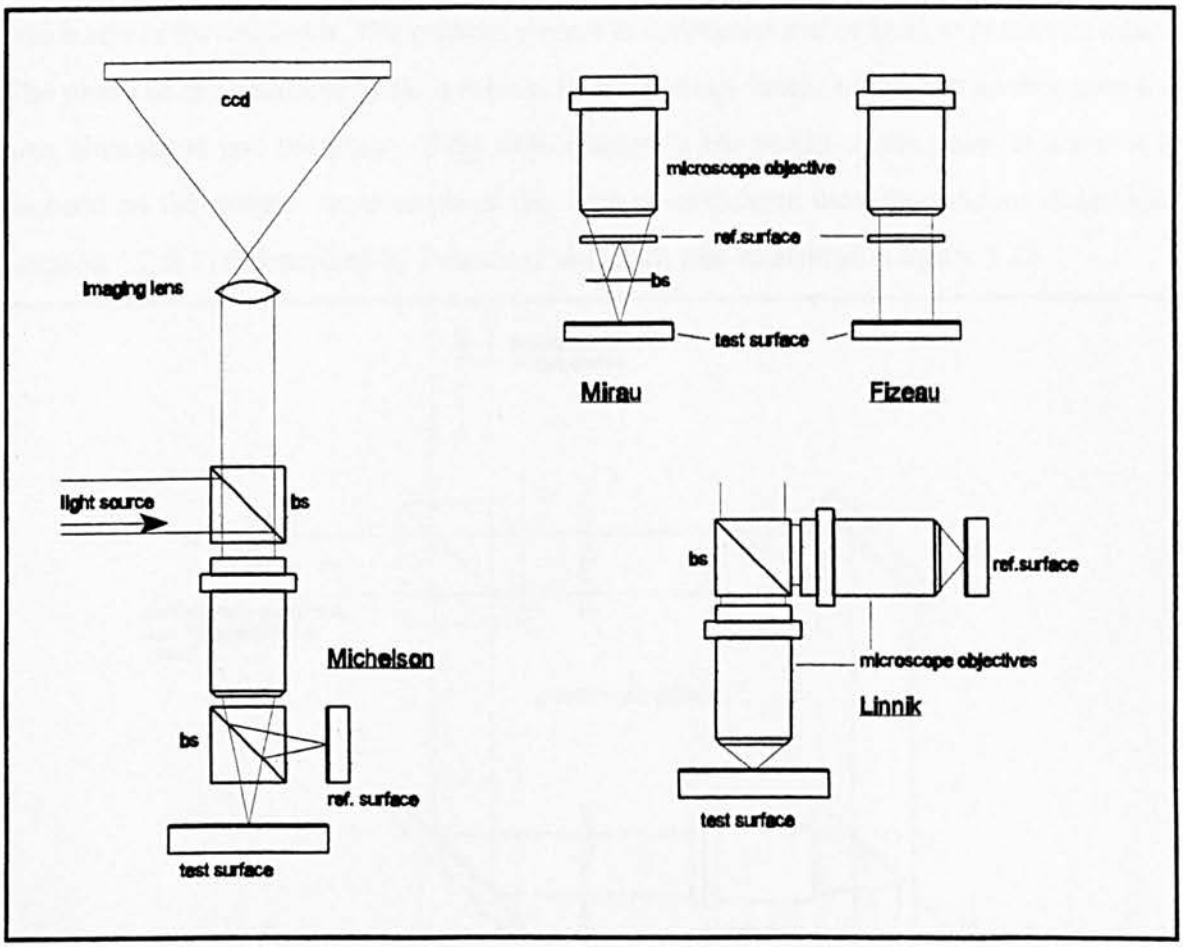


Figure 5.17 Interferometric microscopes for height profiling.

Each of the interferometer types requires a reference surface and so the measurements obtained are relative. In order to allow phase measurement of the interference pattern imaged onto the CCD camera, the reference surfaces are mounted on piezo-electric transducers. Commercial instruments of this type are the Zygo (Middlefield, CT, USA) Maxim and Wyko (Tuscon, AZ, USA) TOPO profilers. The Wyko instrument utilizes a white light source and Michelson,

Mirau or Linnik objectives, depending on the required magnification. The Fizeau objective cannot be used with a white light source because of the unequal path length between test and reference beams. The Zygo instrument utilizes a laser source and Fizeau objectives (Biegen, 1988).

Another type of interferometric height profiler is the concentric beam interferometer which does not require a separate reference surface. The concentric beam interferometer employs a type of radial shear called exploded shear where the diameter of one beam is reduced to a point which acts as the test beam. The reference beam is collimated and coaxial with the test beam. The phase of the reference beam is related to the average height of the test surface over the area illuminated and the phase of the object beam to the height of the point at which it is focused on the sample. An example of this type of instrument using heterodyne techniques (section 1.2.6.1) is described by Pantzer et al (1986) and illustrated in figure 5.18.

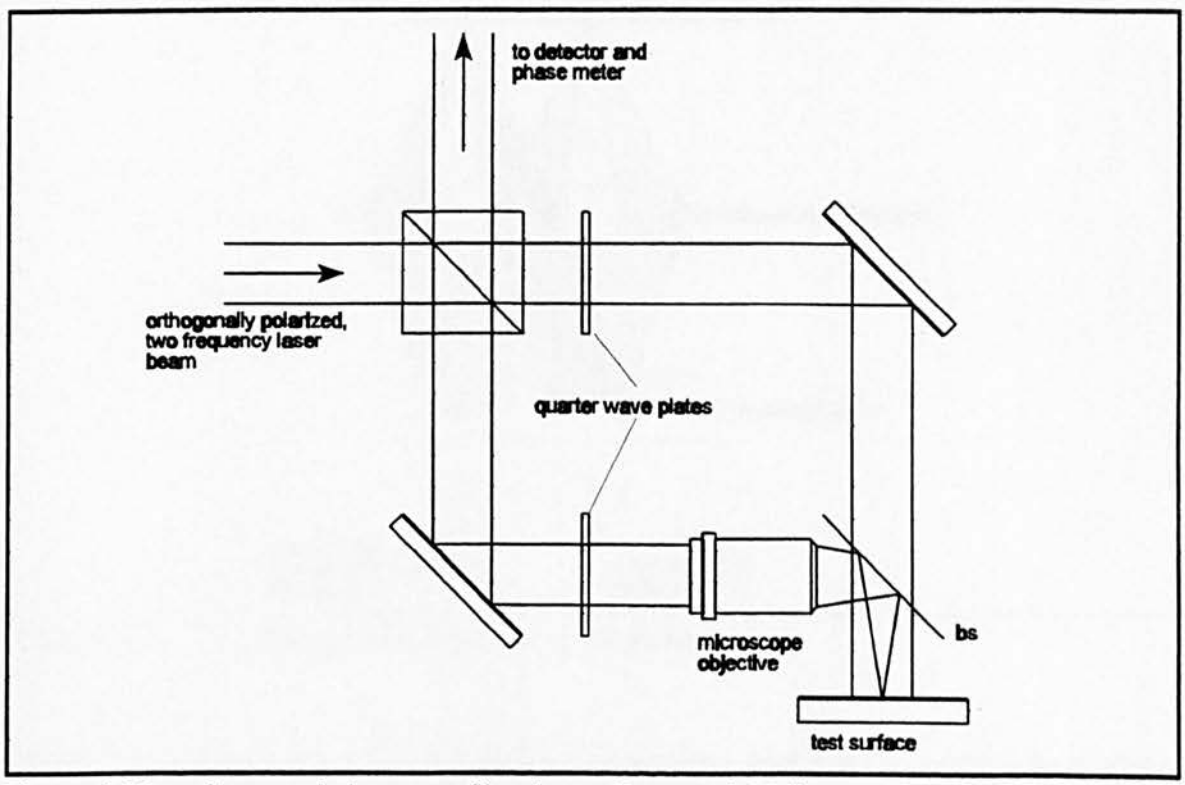


Figure 5.18 Concentric beam profiler due to Pantzer (1986).

Another method for producing concentric reference and test beams is described by Downs et al (1989) using a birefringent lens. A birefringent lens has different focal lengths for orthogonal polarization states and so it is possible to produce both a collimated reference beam and a

focused test beam using the same lens.

A problem with the concentric beam type profiler is that it is insensitive to surface features with low spatial frequency because of the averaging of the phase of the reference beam. A profiling instrument described by Sommargren (1981) overcomes this problem by focusing two beams onto the surface to be profiled (figure 5.19). The reference beam spot is focused onto a point on the surface which is on an axis about which the surface is rotated. The test beam spot traces out a circular path as the surface is rotated and the profile is measured along that path. The instrument uses heterodyne phase measurement and a Wollaston prism is used to separate and recombine the two orthogonally polarized laser frequencies.

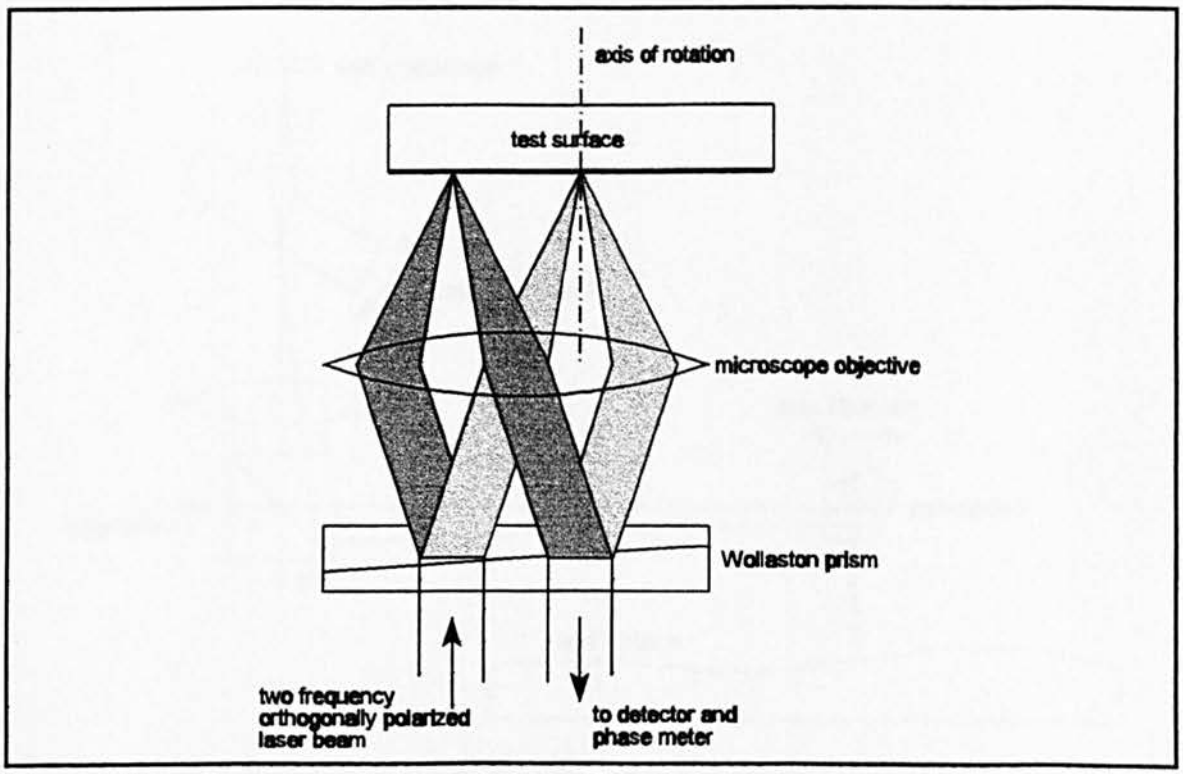


Figure 5.19 Sommargren's heterodyne profilometer.

Whilst the interferometric height measuring systems described directly measure the profile of the test surface, they are prone to errors, over long scan lengths, resulting from errors in the mechanical slides employed to perform the scanning. Ultimately, the accuracy which can be achieved over a linear scan is limited by the straightness of the scanning motion and so the techniques are best suited to the measurement of the high spatial frequencies of a profile rather

than the overall shape.

### 5.5.3 Non-Interferometric Slope-Measuring Profilers.

An approach to profiling surfaces which avoids the need for a reference surface is to measure the surface slope or gradient. The surface profile may then be determined by numerically integrating the slope data. A non-interferometric method of measuring surface slope is to monitor the angle of a beam of light reflected from a surface as the beam is scanned across the surface. This is the approach employed by Virdee (1993, 1995) who uses an auto-collimating technique as shown in figure 5.20.

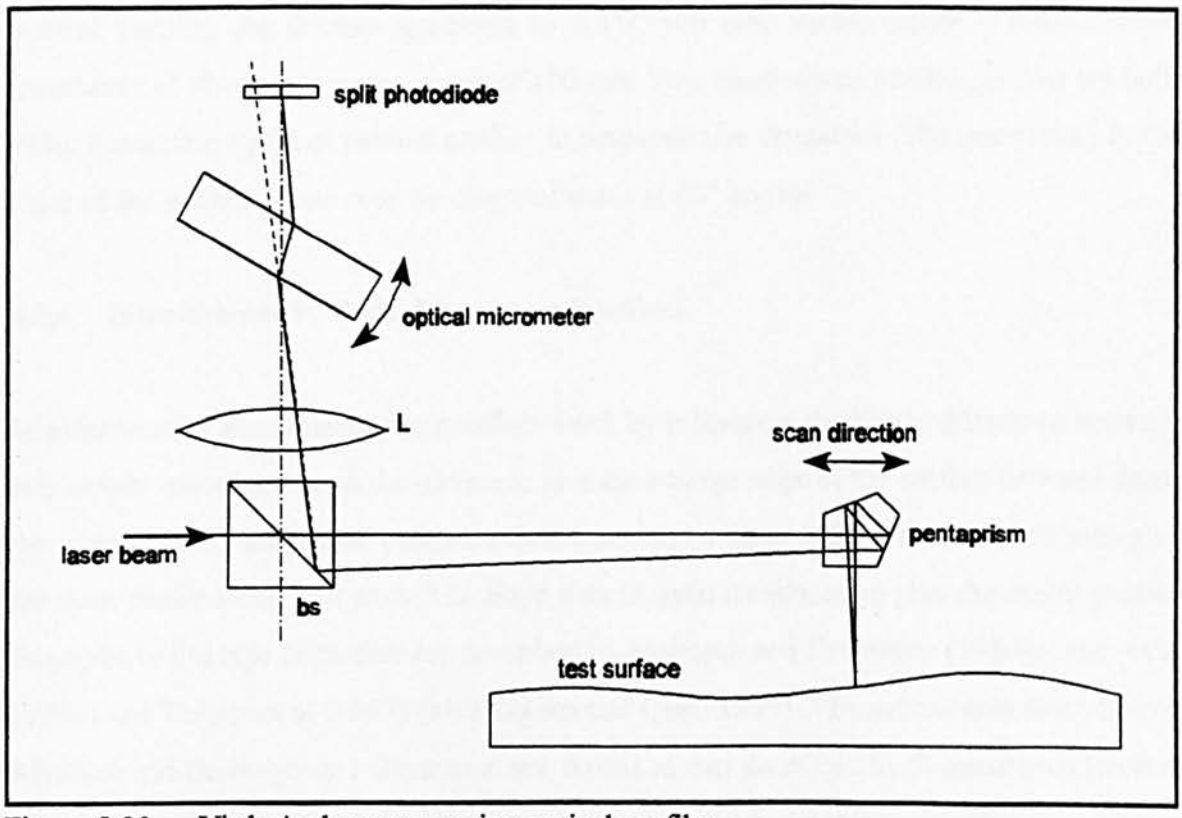


Figure 5.20 Virdee's slope-measuring optical profiler.

The laser beam is scanned across the surface to be measured by a pentaprism. The rest of the instrument is fixed in relation to the test surface. A pentaprism has the property that it reflects the incident light through precisely 90° regardless of its orientation. The slope of the test surface at which the laser beam is incident causes the reflected beam to return to the instrument at a small angle to the incident beam, as shown. The reflected beam is focused onto a position

sensitive split photodiode via an "optical micrometer" consisting of a tiltable parallel sided glass plate. In the absence of the optical micrometer, the focused spot would generally fall on the split photodiode off centre giving a non-nulled signal. The signal is nulled by a feedback loop which tilts the optical micrometer to centre the focused spot. The angular deviation of the reflected beam and thus the slope of the surface are determined by the tilt of the optical micrometer necessary to give a nulled signal. The sensitivity of this auto-collimating system is such that it has a resolution of 0.01 arc seconds. The line profiles of the surface are then obtained by integrating the slope values with respect to distance. The major limitation of the accuracy of this technique is due to thermal gradients giving rise to refractive index variations in the air which cause the laser beams to deviate from straightness. By careful environmental control keeping the thermal gradients to 0.1°C/300 mm Virdee claims a measurement uncertainty of ±2 nm over a scan length of 100 mm. Two dimensional surface profiles are built up by measuring a grid of parallel profiles in perpendicular directions. The uncertainty in the twist of the surface is removed by diagonal scans at 45° angles.

### 5.5.4 Interferometric Slope-Measuring Profilers.

Interferometric slope-measuring profilers work by measuring the height difference between two closely spaced points on the surface to give the average slope of the surface between those two points. The measurement points are then scanned in a linear path across the surface to give the slope profile along that path. The slope data is again integrated to give the height profile. Examples of this type of profiler are described by Makosch and Drollinger (1984), Omar et al (1990) and Takacs et al (1987) (also Takacs and Qian, 1989). The instruments described by Makosch and Drollinger and Omar et al are similar to that described by Sommargren (section 5.5.2) except that the two measurement points are scanned in a linear path rather than a circle in order to measure slope rather than height. The instrument described by Takacs et al is shown, slightly simplified, in figure 5.21.

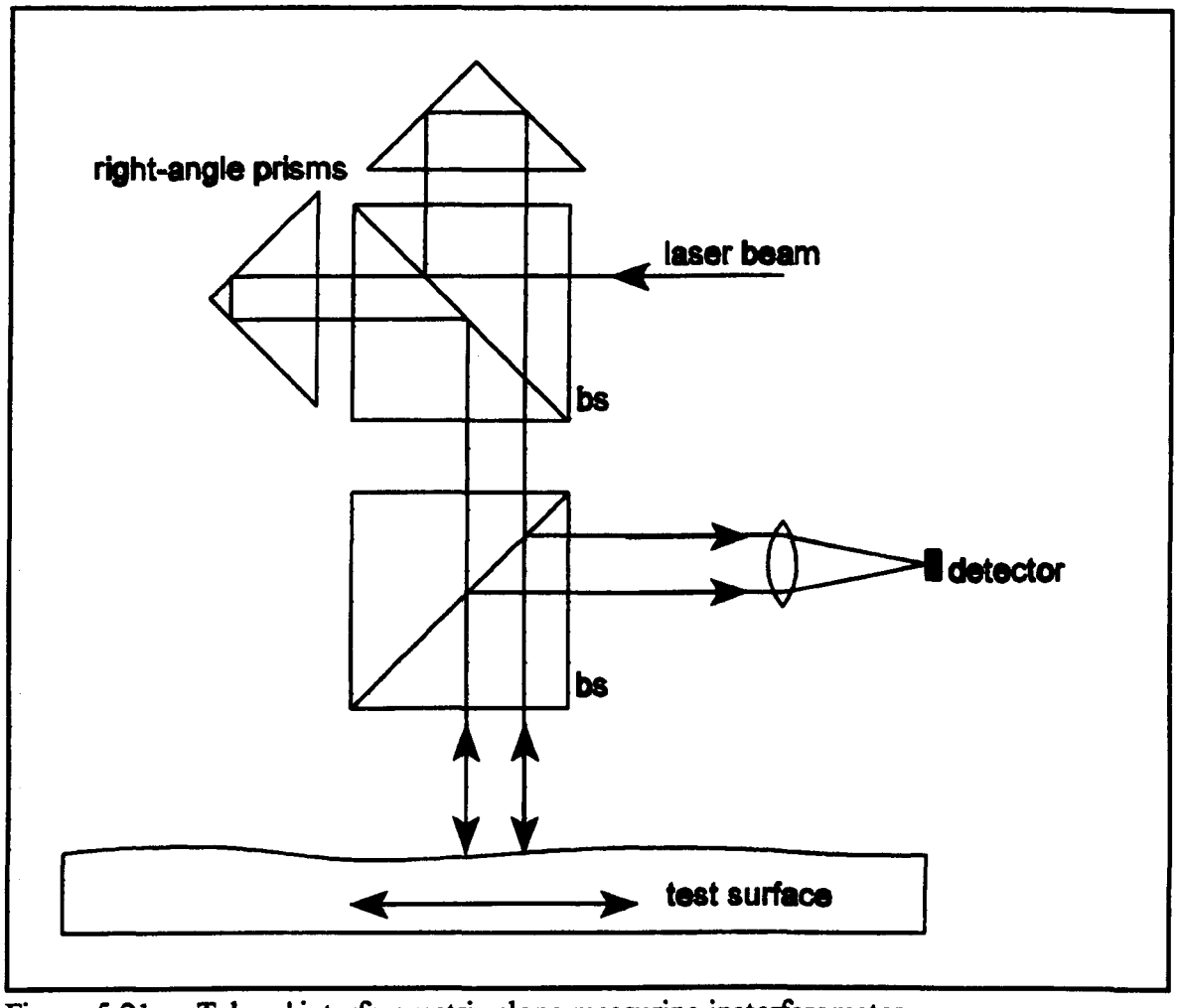


Figure 5.21 Takacs' interferometric slope measuring ineterferometer.

As the measuring head is scanned over the test surface the OPD between the two beams varies as the height difference between the two measurement spots changes. The varying OPD is detected from the interference fringe pattern focused on the detector. The whole measuring head including the detector is mounted on a precision air bearing slide to scan the test surface. The interferometer is thus not immune to variations in tilt of the measuring head due to slide inaccuracies which will show up as errors in the determination of the surface slope and thus its profile.

### 5.5.5 Non-Interferometric Curvature-Measuring Profilers.

It can be seen that profiling instruments are sensitive to relative movements between the measuring instrument and the test surface and thus to the straightness of the mechanism used to perform the scanning. A property of the shape of a surface which is insensitive to the position from which it is measured is its second derivative or curvature. An instrument which measures the curvature of a surface will be insensitive to variations in the piston and tilt of the measuring position. The height profile of a surface may be obtained by twice integrating the curvature data. A non-interferometric instrument of this type is described by Glenn (1990a,b) and manufactured by Bauer Associates (Wellesley , MA, USA). The measurement scheme employed by Glenn is similar to the auto-collimating technique used by Virdee (section )5.5.3 except that the difference in the angles of two, initially parallel, beams reflected from closely spaced spots on the surface is measured in order to give the derivative of the slope and thus the second derivative of the height.

### 5.5.6 Interferometric Curvature-Measuring Profilers.

This author is unaware of any published descriptions of curvature measuring interferometers in the literature but presents here an untested concept for such an instrument. Two versions of the proposed instrument are shown in figures 5.22 and 5.23 which use the same basic principle. In one version (figure 5.22) the measured quantity is the polarisation state of the light and in the other (figure 5.23) the measured quantity is the phase of an optical heterodyne signal. The principle will be described assuming the polarisation version of the profiler with reference to figure 5.22.

Light from a polarised laser source is split into three parallel beams by multiple beamsplitter, BS. The first beam forms the reference probe and is reflected by the polarising beamsplitter, PBS. The second beam forms the test probe and is also reflected by PBS. The third beam passes from the system, having had its plane of polarisation rotated 90° by half-wave retarder plate, $\lambda/2$.

The test beam is focused onto the sample by objective lens, L2, reflected and then transmitted by PBS, having had its polarisation rotated 90° by double passage through quarter-wave retarder plate, $\lambda/4$-1. The test beam then returns to the sample, having been reflected by mirror, M, via lens, L1 and is again reflected. The test beam is now reflected by PBS, having had its plane of polarisation rotated a further 90° by $\lambda/4$-1 and returns to BS.

The propagation of the reference beam proceeds in a similar fashion except that it is focused onto the sample at two points symmetrically about the test beam. This is achieved by its off-axis passage through L1 returning it along a parallel but displaced path for the second pass to the sample surface. The reference beam is brought to two foci either side of the test beam by having its direction of propagation deviated by the wedge prisms. The spacing between test and reference spots is determined by the wedge deviation angle, $\phi$ and the focal length of L2, f; spacing=f $\tan\phi$.

After its second pass, the reference beam is reflected by PBS to BS via retarder $\lambda/2$ which rotates its polarisation by 90°. The test and reference beams recombine in BS and then pass through quarter-wave plate $\lambda/4$-2 to give a linearly polarised beam whose polarisation vector depends on the relative phase of the test and reference beams. The relative phase of the beams depends on their optical path difference (OPD). The OPD will vary as the sample is scanned beneath the probe beams depending on the surface topography. The changing state of the polarisation, which may be measured electronically as shown, or otherwise, is thus related to the profile of the scanned surface. Since the reference spots are symmetrically placed either side of the test spot, the interferometer is self-compensating for tilts of the sample and would thus not require a precision translation stage or a controlled laboratory environment to operate effectively.

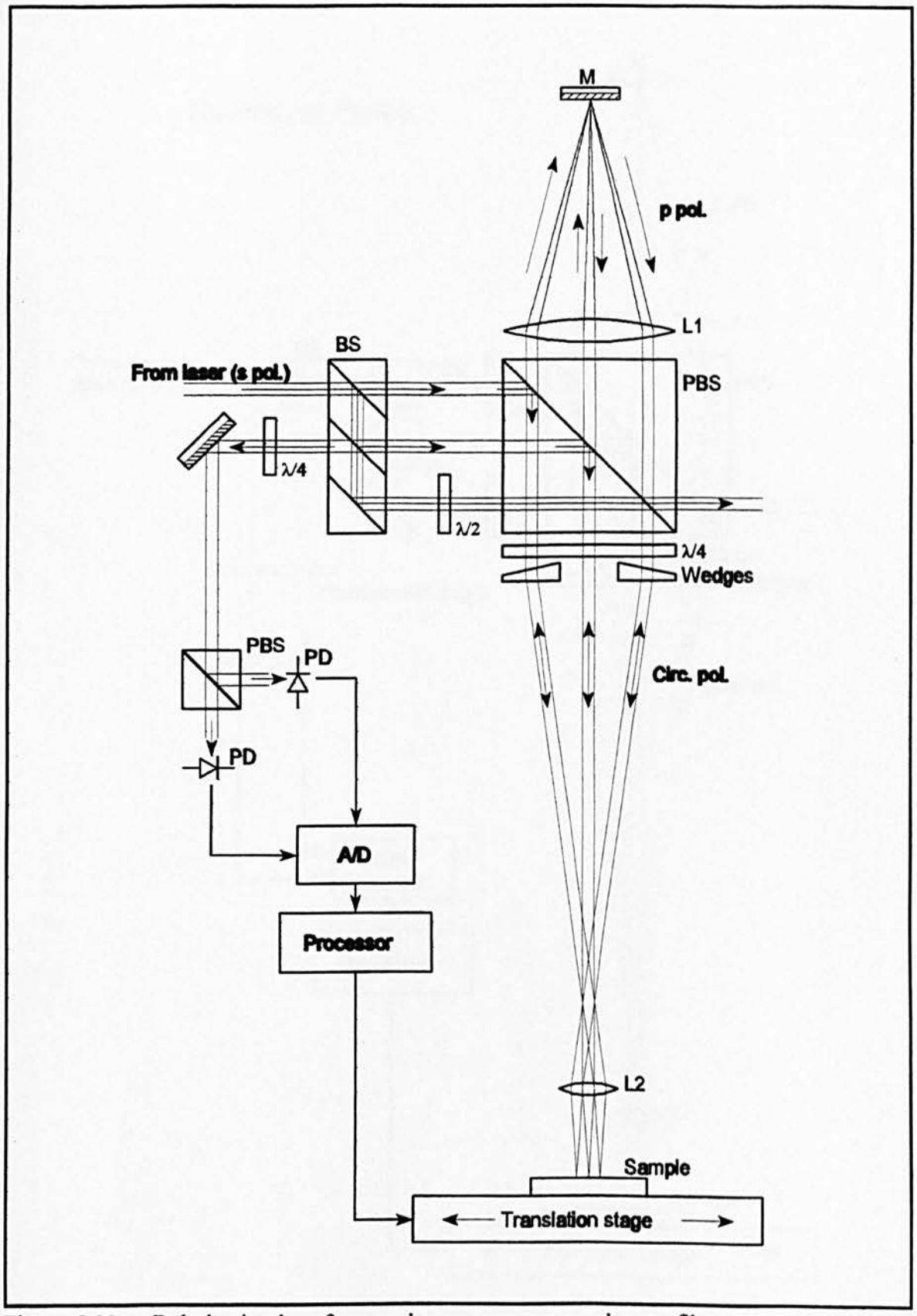


Figure 5.22 Polarization interferometric curvature-measuring profiler.

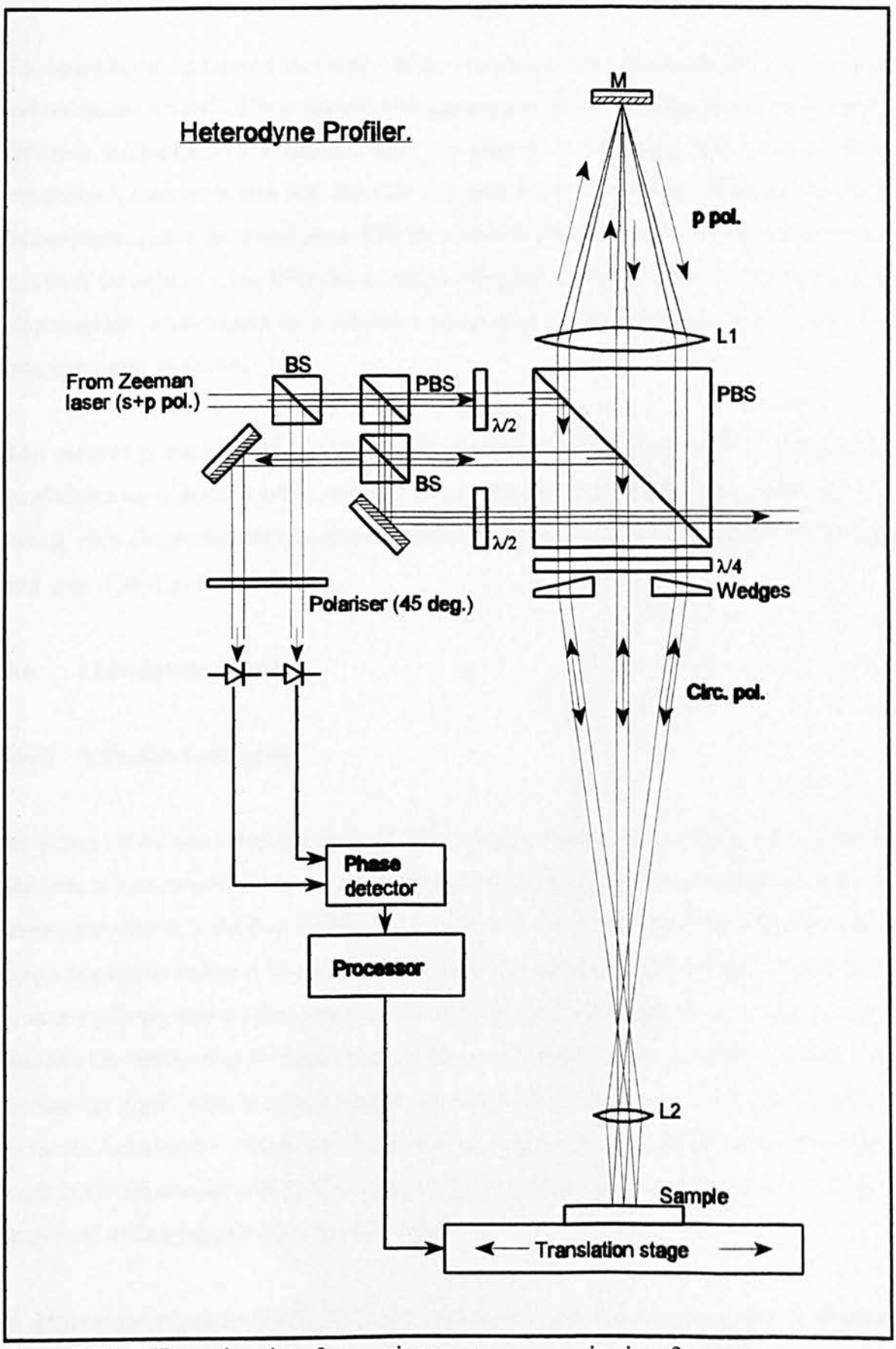


Figure 5.23 Heterodyne interferometric curvature-measuring interferometer.

The operation of the heterodyne version of the instrument is similar except that the reference and test beams are derived from the two orthogonally polarised oscillation modes with slightly different frequencies of a Zeeman laser (or other two frequency laser source). When recombined, the two beams beat together to produce a signal at their difference frequency whose phase mimics the optical phase difference between the two beams. This signal frequency will be of the order of a few MHz (for a Zeeman laser) and so its phase may easily be measured electronically with respect to a reference signal derived from the laser beam before the interferometer, as shown.

This concept is derived from the design of an interferometer to measure thermal expansion coefficients on which this author worked at the GEC research labs (Wembley, Middx. UK.) during 1984. Descriptions of related interferometer designs are given by Bennett (1977), Okaji and Imai (1984) and Birch (1987).

## 5.6 Miscellaneous ideas.

### 5.6.1 A Perfect Collimator?

In section 1.3.3 it was noted that the surface of a liquid rotating about a vertical axis assumes the form of a paraboloid. If a sufficiently high quality paraboloid surface could be achieved by these means then it could form the basis of a collimator for the interferometric testing of flats by placing a point source at its principal focus. Rotating liquid mirrors have been used as large diameter primary mirrors for telescopes where their reflectivity must be high. Consequently, mercury has been used as the liquid. As was discussed in section 1.3.1.1 mercury is not an ideal choice for liquid surface interferometry because of its low viscosity. For optical testing purposes, however, the reflectivity of such a mirror need not be high and so a more ideal liquid such as the silicone oil used for flat liquid surface interferometry could be used, resulting in improved surface figure over a mercury mirror.

A disadvantage of using a rotating liquid mirror as a collimator is that the axis of the collimated

beam would necessarily be vertically orientated. The flat being tested would then have to be positioned in a horizontal plane and thus subject to gravitational sag. This is one of the major disadvantages of conventional liquid surface interferometry.

### 5.6.2 An Absolute Reference Plane for Height Measuring Profilometry.

The central (zero order)fringe of the interference pattern formed by two coherent point sources defines a plane in three dimensions. The other fringes in the pattern are hyperbolic in cross section. If a means could be devised to reference the position of the measuring head of a height measuring profilometer to the central maximum of the zero order fringe of such an interference pattern then the problem of the slide errors in the scanning mechanism of the profilometer would be solved. One way in which this might be achieved is to devise a sensor which samples the interference pattern from the two point sources at a point in space and uses a feedback mechanism to lock the sensor position to the fringe centre. This could be achieved with high precision by using a heterodyne technique and two point sources with slightly different optical frequencies (see section 1.2.6.1). In order to totally define the position of the measuring head, it would have to be equipped with three such sensors to control the tip and tilt as well as the height. The problem of locking the sensors to the zero order fringe could be solved by using two or more wavelengths. In this way the fringes formed by each wavelength would only coincide exactly for the zero order.

## Appendix A: Absolute Interferometric Testing of Microspheres.

**(Report commissioned by NPL, March 1996)**

### 6.1 Introduction.

Recent advances in areas such as fibre and integrated optics and optical computing have led to the emergence of a new class of very small scale optical components. Some examples are:
optical microspheres employed as lenses to couple light between optical fibres and integrated optical devices;
microlens arrays used to facilitate free-space communication between devices in optical processing and computing applications.
The surface figure of these micro-optical devices has a profound effect on their performance and so must be tightly controlled. In order to facilitate the accurate manufacture of small spherical surfaces, techniques must be developed to perform accurate measurements of their surface figure.

In a previous report, Stevens (1994) describes experimental investigations of interferometric techniques for measuring the surfaces of microspheres made at NPL. Non-absolute measurements were made using a variety of interferometers of the Fizeau and Twyman-Green type with both low coherence mercury sources and high coherence laser sources. Some of the practical problems associated with the small radius of curvature of the test surfaces were identified.

This report develops the work undertaken at NPL. Particular emphasis is placed on making *absolute* measurements of the surface figures of microspheres.

### 6.2 Choice Of Type Of Interferometer.

Interferometers for the measurement of surface contours may be broadly classified into two general classes, shearing and non-shearing interferometers. Shearing interferometers are

considered unsuitable for the given task because of the difficulty of interpreting the interference fringes and the accumulation of errors that can occur as a result of integrating the slope data that the fringes yield in order to derive the surface contours (Stevens,1994, Malacara,1992, Saunders,1961). Of the non-shearing interferometers, those most suitable for measuring surface contours are those which analyse the wavefront reflected from the surface under test (as distinct from the transmitted wavefront). Of these, the two basic interferometer types are the Fizeau and Twyman-Green interferometers.

The Fizeau and Twyman-Green interferometers will each be considered and their relative merits for the measurement of microspheres will be examined.

### 6.2.1 The Fizeau Interferometer.

In a Fizeau interferometer, interference fringes are generated by a low finesse cavity formed by reference and test surfaces where the surfaces are illuminated at nominally normal incidence. In its most common form, for the measurement of plane surfaces, the cavity is formed by plane reference and test surfaces which are nominally parallel, though there may be a small wedge to introduce tilt fringes (Malacara, 1992). The fringe pattern formed by a Fizeau interferometer yields information about the distance between conjugate points on the reference and test surfaces. If the reference surface is perfect, with no deviations from the ideal surface figure, then the fringe pattern will relate directly to the figure of the test surface. If, however, the reference surface is imperfect then the fringe pattern will relate to the sum of the figures of both surfaces.

The Fizeau interferometer is relatively insensitive to environmental disturbances since, for the most part, the reference and test wavefronts follow a common path.

For the measurement of spherical surfaces, the cavity consists of concentric reference and test surfaces. This arrangement is shown in figure 6.1. The cavity spacing is modulated by the piezo-electric transducer, PZT, in order to phase step the interference fringe under computer control. The fringes are acquired by the CCD camera and analysed by the computer using one

of the phase measuring algorithms developed for wavefront analysis (Malacara,1992; Creath,1992). Commercial software packages exist for this purpose. NPL possesses Wyko's DOS/RTI software which handles PZT control, acquisition of CCD camera data and data analysis.

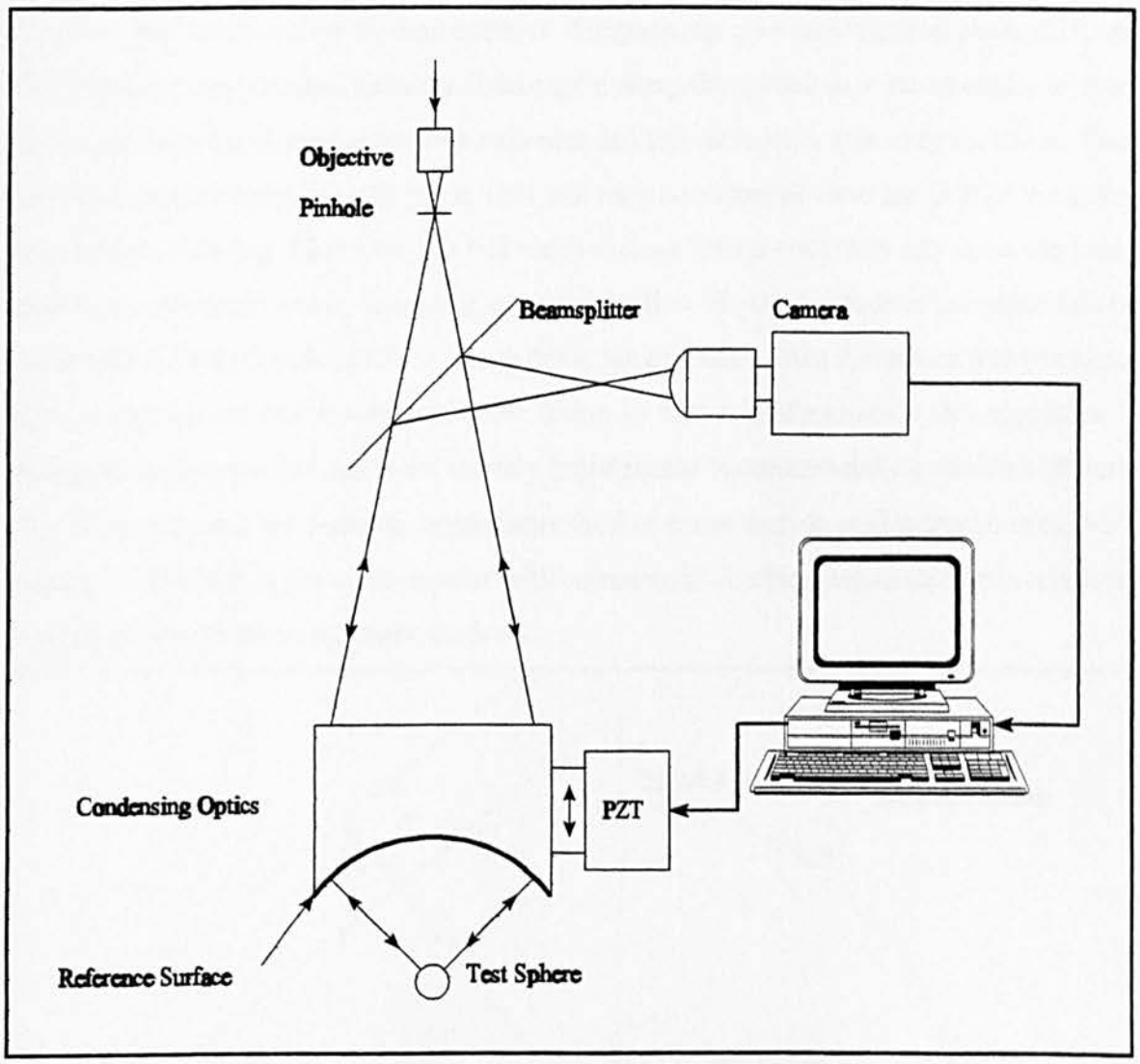


Figure 6.1 Fizeau interferometer for measuring spheres.

NPL possesses a commercial phase measuring Fizeau interferometer (Zygo Mk IV) which has a variety of reference spheres of different numerical apertures. For reasons which remain unclear, the Zygo interferometer seems unable to focus its CCD camera onto very small spheres which results in spurious fringe patterns making analysis of surface shape difficult or impossible.

A significant drawback to the use of a Fizeau interferometer for the measurement of small spheres is the requirement for a reference surface of high numerical aperture. Not only is the manufacture of such a surface of high quality difficult, but a high numerical aperture can result in errors in the phase measurement algorithm. By reference to figure 6.2, it can be seen how the phase shift varies with numerical aperture. Suppose, for a certain required phase shift, the PZT translates the reference surface a distance, $\delta x$, along the optical axis. At an angle, $\alpha$, from the optical axis, the separation between reference and test surfaces is now only $\delta x \cos \alpha$. Thus for high numerical apertures, the phase shift will vary considerably over the field of view. For most phase measuring algorithms, this will result in large errors since they rely upon the phase shift being a particular value, depending on the algorithm. However, there is one algorithm by Carré (Carré,1966; Creath,1992) for which this is not a problem since it assumes that the phase shift is not known but is constant from frame to frame. Unfortunately this algorithm is computationally intensive and is not usually implemented in commercially available software. The Wyko software, for example, implements the five frame technique (Hariharan et al,1987; Creath,1992) which is the most popular with commercial interferometers since it is relatively insensitive to common systematic errors.

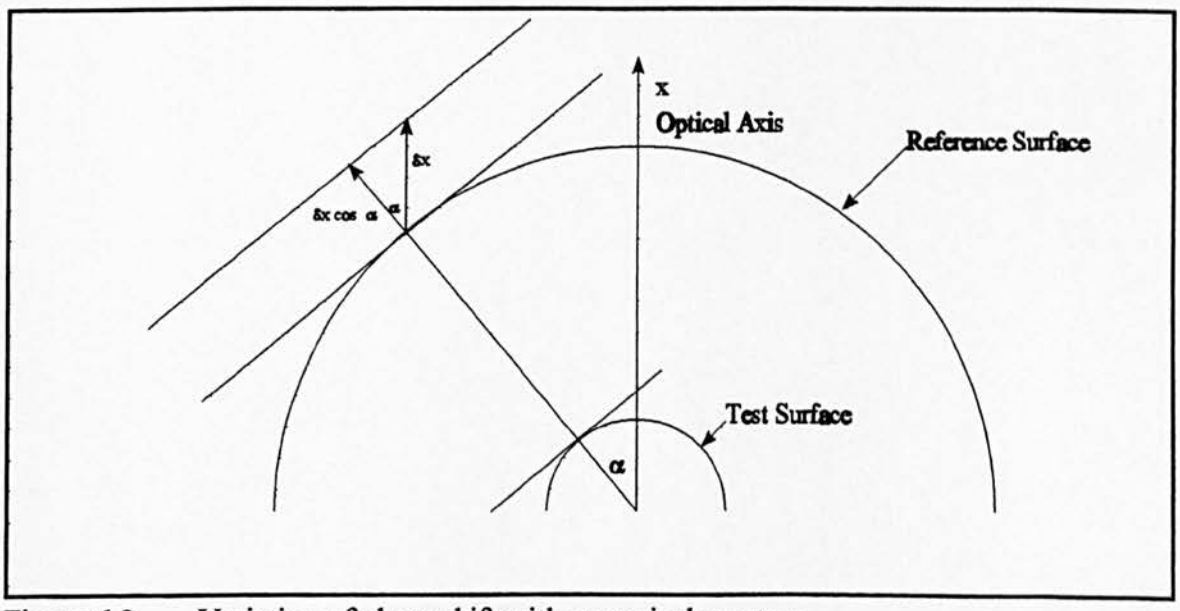


Figure 6.2 Variation of phase shift with numerical aperture.

In order to avoid the variation of phase shift with numerical aperture, the component translated by the PZT (the reference surface) must be used in collimated light and so must be plane. This

may be achieved by transforming the plane wavefront into a spherical wavefront between the plane reference surface and the test sphere as shown in figure 6.3. In this arrangement, however, the reference surface is not compared directly with the test sphere but with the combination of test sphere and condensing optics. There are absolute measurement techniques which enable the errors in both the reference surface and the condensing optics to be isolated and so removed and these will be described later.

It should be noted that the problem of the variation of phase shift with numerical aperture may also be avoided if the phase shift is achieved by means other than a mechanical translation of the reference or test surface. The optical path length between the two surfaces may be altered by changing the refractive index of the medium between them. This may be accomplished by varying the pressure where the medium is gaseous (usually air). Alternatively, the optical path length may be kept constant and the phase difference between the wavefronts reflected from the two surfaces varied by modulating the wavelength of the light source. This would require a tunable source such as a semiconductor diode laser.

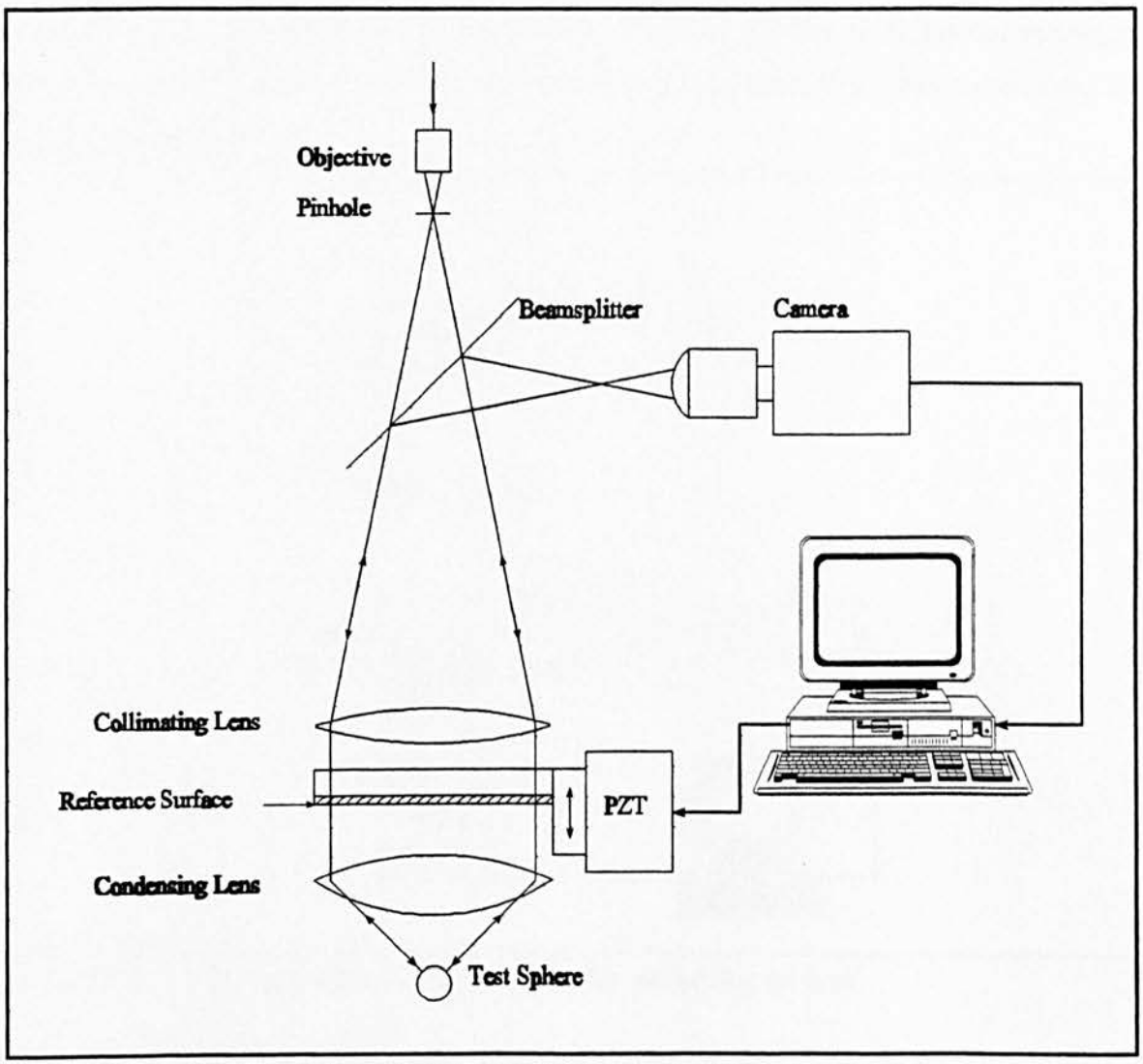


Figure 6.3 Modified Fizeau interferometer for measuring microspheres.

### 6.2.2 The Twyman-Green Interferometer.

In a Twyman-Green interferometer, light is divided by a beamsplitter and travels separately to the test and reference surfaces. The light reflected from these surfaces is then recombined by the beamsplitter to form an interference pattern (figure 6.4). When the reference and test arms are of the same length the virtual image of the reference surface is in the same position as the test surface (rather than opposing it as in the case of the Fizeau interferometer). In this case, then, the fringe pattern represents the difference between the figures of the two surfaces. Again, since the converging lens which transforms the collimated light into a spherical wavefront is effectively between the reference and test surfaces, the test wavefront is due to

both the converging optics and the test surface. The same absolute measurement techniques mentioned above may be used to isolate errors arising from the reference surface and condensing optics.

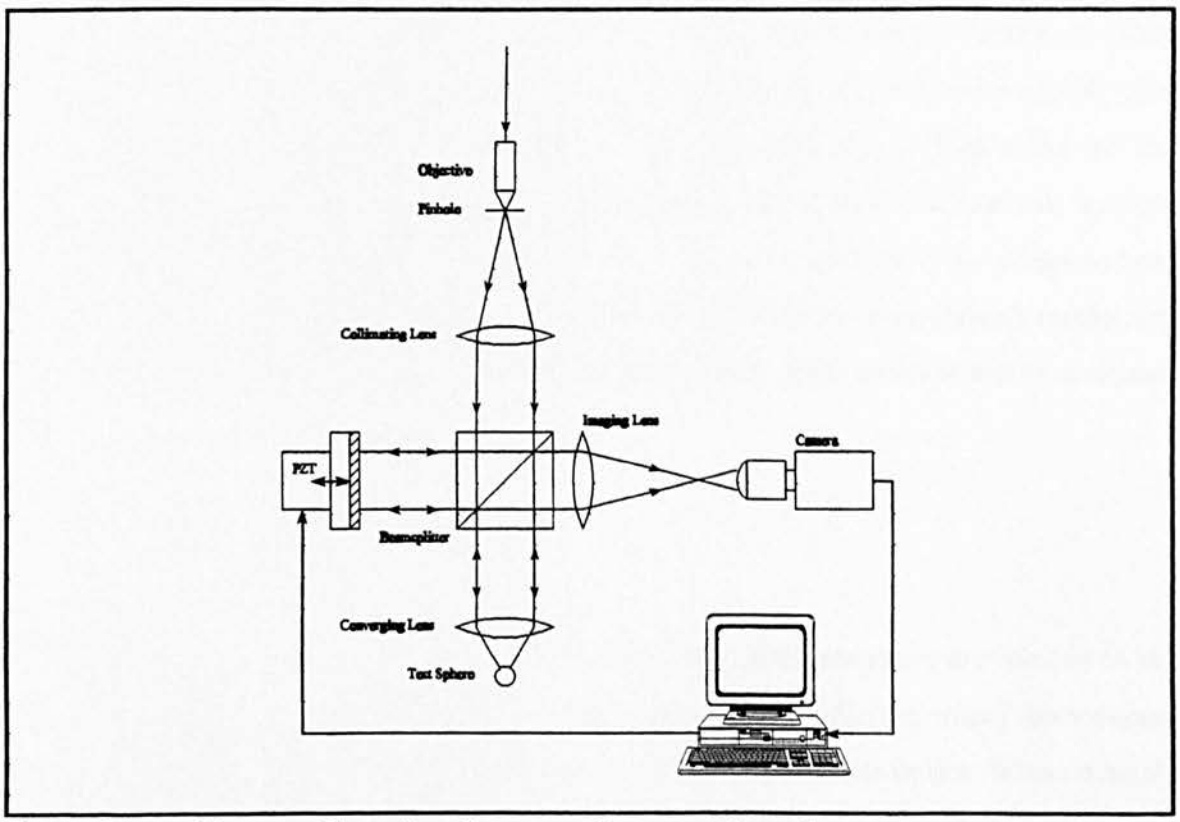


Figure 6.4 Twyman-Green interferometer for measuring spheres.

Unlike the Fizeau interferometer, the test and reference arms follow separate paths and so the Twyman-Green interferometer is more susceptible to environmental disturbances. However, since an instrument designed to measure small spheres would be of compact dimensions, it would be straightforward to provide vibrational isolation and shielding from other disturbances such as air currents. Having separate reference and test arms does have some advantages. The reference surface is not required to transmit any light which greatly simplifies its mounting on the PZT and also its substrate need not be of high optical quality. The Twyman-Green arrangement also allows for greater flexibility in the positioning of, and access to, the various optical components.

A particular problem encountered by both Fizeau and Twyman-Green interferometers when using highly coherent light sources (lasers) is that of coherent noise. The high degree of spatial

and temporal coherence of lasers means that any scattered light from dust or optical imperfections, or any spurious reflections from other than the test or reference surfaces, that pass through the exit pupil of the system will cause speckle noise or spurious fringes across the field of view. The obvious solution to this problem is to use a less coherent light source such as a filtered low pressure mercury lamp or a laser passed through a moving ground glass screen (Schwider and Falkenstorfer,1995). With a less coherent source, so long as the test and reference paths are matched to within the coherence length, fringes will be formed but spurious light will be outside the coherence length and so will not contribute to the fringe pattern. Unfortunately a highly coherent source is necessary for the absolute measurement methods to be described. Other practical methods of reducing the coherent noise problem will be discussed later.

## 6.3 Absolute Measurement Methods.

Measurements of surfaces by Fizeau or Twyman-Green interferometers are relative to the reference surface and any other beam shaping optics. This is sufficient where the accuracy required of the measurement is less than the uncertainty in the reference optics. Where a highly accurate measurement must be made, however, an absolute measurement method must be used to separate the interferometer and test surface errors.

Such methods may be broadly classified into two categories. First, there are methods which make relative measurements between pairs of surfaces from a population of three nominally spherical surfaces. The data so obtained are used to deduce the absolute figures of the spheres. Such methods are directly comparable to the well known three-flat technique (Malacara,1992) and its derivatives for the absolute measurement of flat surfaces and will, accordingly, be referred to as "three-sphere" techniques. Second, there are methods which make use of the fact that the centre of curvature of a converging spherical wavefront is accessible (the so called "cat's-eye" position) in order to separate interferometer and test surface errors. These methods will be referred to as "cat's-eye" techniques.

### 6.3.1 Three-Sphere Techniques.

Let us consider three spherical surfaces, A, B, C, as shown in figure 6.5. The deviations, for each surface, from an ideal sphere are $f_A(\phi,\theta)$, $f_B(\phi,\theta)$,$f_C(\phi,\theta)$, respectively, where $\phi$ and $\theta$ are the angular co-ordinates with respect to their centres of curvature.

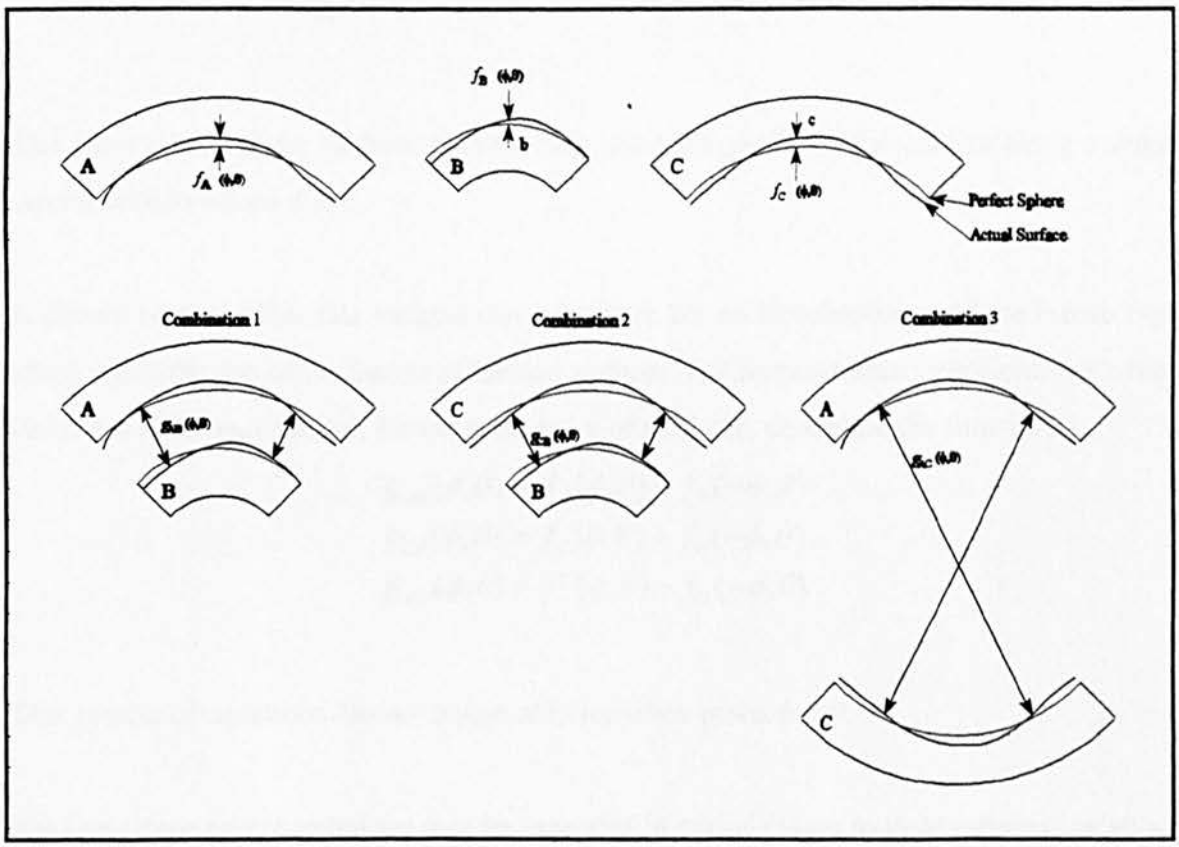


Figure 6.5 Three sphere absolute test.

These three spheres are measured in pairs, AB, CB, AC as shown in figure 6.5 in a Fizeau interferometer where the upper surface acts as the reference surface and the lower as the test surface. The measurements determine the functions respectively, which are defined as follows:

$$g_{AB}(\phi,\theta) = f_A(\phi,\theta) + f_B(-\phi,\theta)$$
$$g_{CB}(\phi,\theta) = f_C(\phi,\theta) + f_B(-\phi,\theta)$$
$$g_{AC}(\phi,\theta) = f_A(\phi,\theta) + f_C(-\phi,\theta)$$

zsince the Fizeau interferometer yields the sum of the deviations of each surface from the ideal. We now have a system of three equations with four unknowns: . The system has solutions only

where $\phi = 0$ and reduces to three unknowns. The solutions are:

$$f_A(0,\theta) = \frac{g_{AB}(0,\theta) + g_{AC}(0,\theta) - g_{CB}(0,\theta)}{2}$$
$$f_B(0,\theta) = \frac{g_{AB}(0,\theta) + g_{CB}(0,\theta) - g_{AC}(0,\theta)}{2}$$
$$f_C(0,\theta) = \frac{g_{CB}(0,\theta) + g_{AC}(0,\theta) - g_{AB}(0,\theta)}{2}$$

The basic three-sphere method can thus only yield the profile of the spheres along a single central section where $\phi = 0$.

It should be noted that this method can only work for an interferometer of the Fizeau type which yields the *sum* of the figures of the two surfaces. A Twyman-Green interferometer which yields the *difference* would, for the same pairs of surfaces, determine the functions:

$$g_{AB}(\phi,\theta) = f_A(\phi,\theta) - f_B(-\phi,\theta)$$
$$g_{CB}(\phi,\theta) = f_C(\phi,\theta) - f_B(-\phi,\theta)$$
$$g_{AC}(\phi,\theta) = f_A(\phi,\theta) - f_C(-\phi,\theta)$$

This system of equations has no unique solution even when $\phi = 0$.

The basic three-sphere technique may be extended in various ways to yield information about further central sections or the whole of the spherical surfaces (Schulz and Shwider,1976; Gubin and Sharonov, 1990). Each of these methods involves a fourth positional combination of two of the three spheres where one sphere is rotated about the optical axis through a suitably chosen angle. The method of Schulz and Schwider involves stepping the profile of the central section originally determined by the basic three-sphere test around the optical axis to yield the profiles of a number of further central sections determined by the chosen rotation angle. This method is prone to an accumulation of experimental errors as the number of new profiles increases. Gubin and Sharonov describe a method similar to that described by Fritz (Fritz,1984) for flat surfaces. The surfaces and the four interferograms are approximated by a system of polynomials that are orthogonal on a circular part of a sphere (Zernike polynomials) and the polynomials are manipulated to give a polynomial description of the

absolute figure of the surfaces. The method is computationally intensive and its accuracy is highly dependent upon the accuracy to which the rotation of the sphere can be achieved.

All three-sphere techniques rely upon the direct comparison of pairs of spherical surfaces in a Fizeau interferometer. This means that where the surfaces have high numerical aperture, the problem of the variation of phase shift over the field of view (see section 2.1) will always be a serious limitation on the accuracy of these methods.

### 6.3.2 Cat's-Eye Techniques.

These techniques may be used in either Fizeau or Twyman-Green interferometers which have an external and accessible focus known as the cat's-eye position. Placing a reflective surface at this position produces an interferogram which contains information only about the errors in the interferometer. Manipulation of this information with that obtained when the test surface is included may be used to determine separately the errors due to the test surface and the interferometer.

#### 6.3.2.1 Three Position Cat's-Eye Technique.

This technique was first introduced by Jensen (Jensen,1973) and later developed by others (Bruning et al,1974; Truax,1988; Elssner et al,1989; Creath and Wyant,1990). It involves three separate measurements which are combined to determine the surface minus the errors due to the interferometer and the reference surface. It works equally well in a Fizeau or Twyman-Green interferometer.

The three measurements are shown in figure 6.6.

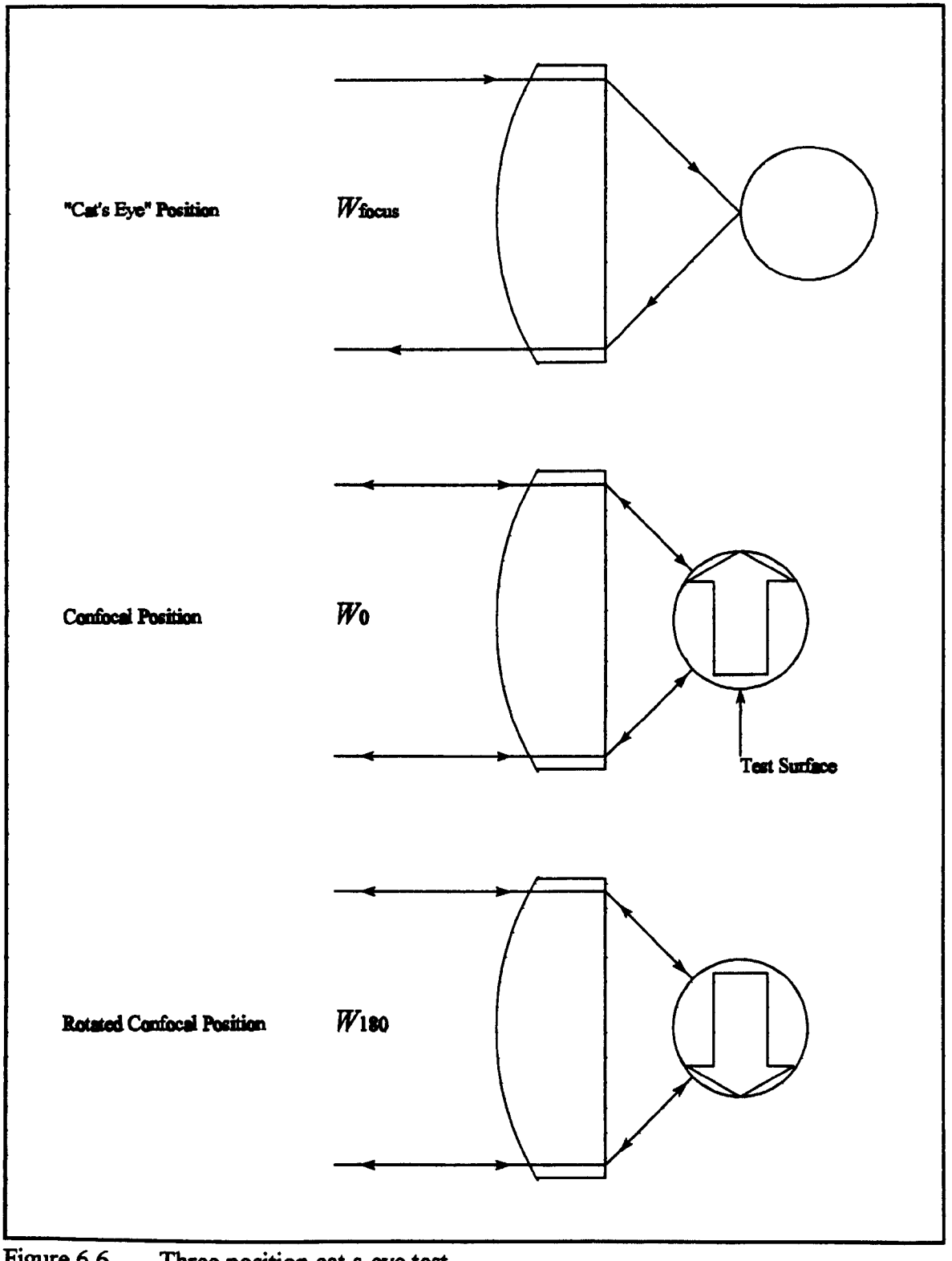


Figure 6.6 Three position cat,s-eye test.

The first measurement is with the test surface at the focus of the converging lens (also known as the cat's-eye position).

The second measurement is with the test surface positioned such that its centre of curvature is at the focus of the converging lens (also known as the confocal position).

The third measurement is taken after rotating the test surface 180° about the optical axis (the rotated confocal position).

Assuming a single measurement is equal to the difference in the wavefronts due to the test and reference arms of the interferometer,

$$W_{meas} = W_{testarm} - W_{refarm}$$

For measurement #1,

$$W_{testarm} = \frac{W_{conv} + \overline{W}_{conv}}{2}$$

Where:

$W_{conv}$ = the aberrations due to a double pass of the optics in the test arm or the interferometer (this is usually just the converging lens, but can include other optics),

$\overline{W}_{conv} = W_{conv}$ rotated by 180°.

This measurement can be written as

$$W_{focus} = \frac{W_{conv} + \overline{W}_{conv}}{2} - W_{refarm}$$

This equation works for either a Twyman-Green or a Fizeau interferometer.

For measurement #2 the test arm wavefront is given as

$$W_{testarm} = W_{conv} + W_{testsurf}$$

where $W_{testsurf}$ = effects due only to the test surface.

Thus, measurement #2 can be written as

$$W_{0°} = W_{testsurf} + W_{conv} - W_{refarm}$$

For measurement #3, the test surface is rotated by 180°. With the test surface contributions given by $\overline{W}_{testsurf}$, this measurement can be written as

$$W_{180°} = \overline{W}_{testsurf} + W_{conv} - W_{refarm}$$

The three measurements can then be used to solve for the test surface using

$$W_{testsurf} = \frac{W_{0°} + \overline{W}_{180°} - W_{focus} - \overline{W}_{focus}}{2}$$

Which is simply calculated with additions, subtractions and 180° rotations of the three measurements. (In the computer, the data are rotated about their centres, not the centre of the CCD array).

The aberrations in the interferometer and errors due to the reference surface can be obtained by calculating

$$W_{conv} - W_{refarm} = \frac{W_{0°} - \overline{W}_{180°} + W_{focus} - \overline{W}_{focus}}{2}$$

This reference wavefront can then be subtracted from measurements of subsequent test spheres as long as the radii of curvature are similar. If there is a large difference in radii of curvature, a new reference wavefront must be generated.

This is the method implemented in Wyko's DOS/RTI software.

#### 6.3.2.2 Alignment Of Test Sphere And Interferometer.

The critical part of this test is the alignment of the test sphere relative to the interferometer. Incorrect location of the optical axis causes a wavefront shear upon rotation of the data which introduces errors that are proportional to the errors present in the individual wavefronts

(Truax, 1988). The optical axis is defined by the first measurement in the cat's-eye position with the fringes nulled. The detector in the interferometer should be centred on the optical axis.

Next, the test surface needs to be aligned relative to the optical axis in order to rotate the test surface by 180° without altering the interference pattern. Figure 6.7 illustrates the possible misalignments of the sphere and the axis of rotation with respect to the optical axis. The necessary alignments are:

the vertex of the sphere must lie on the optical axis,

a) the axis of rotation defined by the rotation stage must coincide with the optical axis,

b) the centre of curvature of the test surface must lie at the focus of the converging lens.

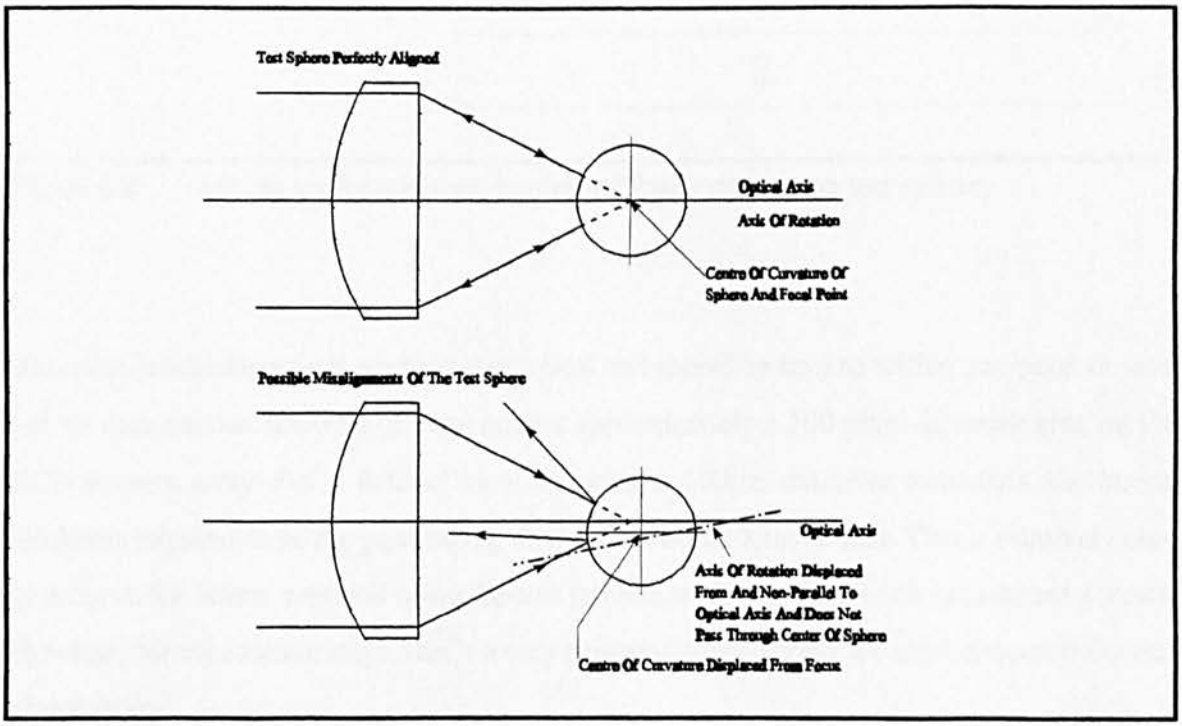


Figure 6.7 Alignment of test sphere.

In order to achieve this, a mount with eight degrees of freedom as shown in figure 6.8 is required. The test sphere is mounted on a small x-y stage in order to align the centre of curvature of the sphere with the axis of the rotation stage, $\theta$. The axis of the rotation stage is made parallel to the optical axis by tip/tilt stage, $\alpha$-$\beta$. The centre of curvature of the sphere is brought into coincidence with the focus of the converging lens by the X-Y-Z translation stages.

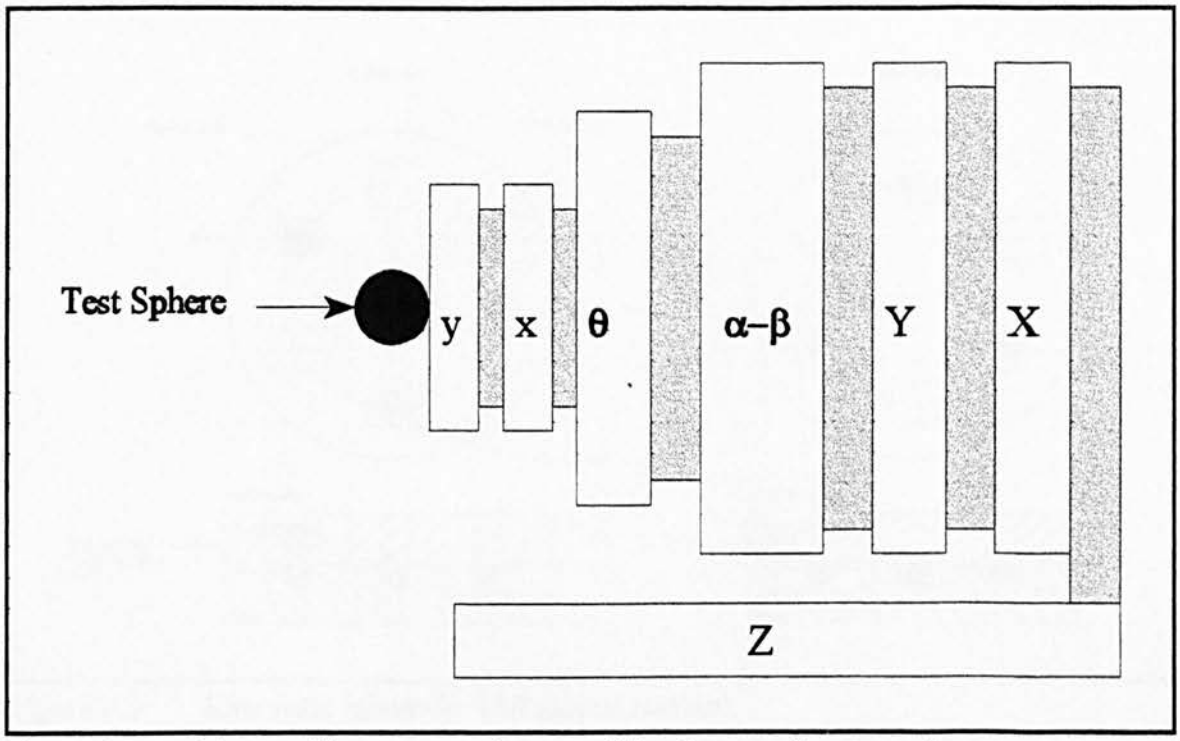

Figure 6.8 Mount with necessary degrees of freedom to align test sphere.

Generally, lateral alignments are the most critical and should be held to within one pixel or less. Let us assume that the field of view covers approximately a 100 pixel diameter area on the CCD camera array. For a field of view covering a 100μm diameter microlens, the lateral resolution required from the positioning devices would be 1μm or less. This is relatively easy to achieve for lateral motions using flexure mounts with very fine pitch adjustment screws. However, for the rotation stage, this is a very stringent specification for axial concentricity and repeatability.

Since the rotation mount does not require continuous adjustment, but needs only two positions, 180° apart, it may be possible to achieve the specification using a two-position kinematic mount as shown in figure 6.9.

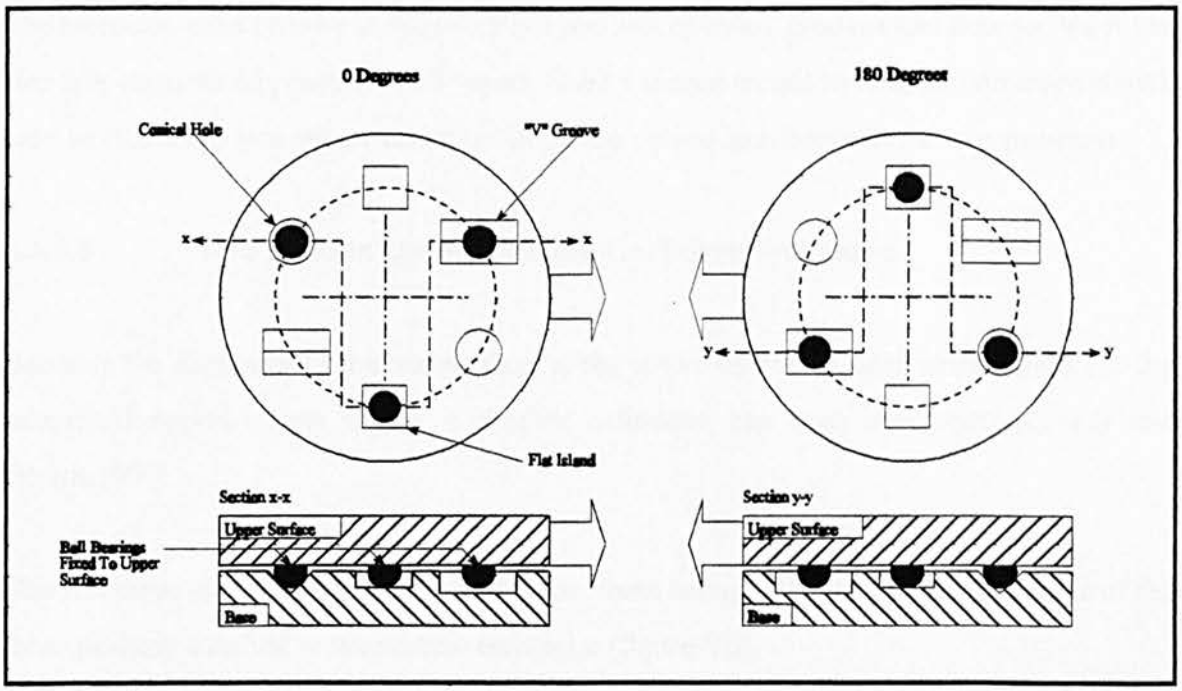


Figure 6.9 Kinematic mount for 180 degree rotation.

Any optical mount's position can be defined uniquely in terms of six independent co-ordinates; three translations and three rotations with respect to some arbitrary fixed co-ordinate system. A mount is said to be kinematic when the number of degrees of freedom (axes of free motion) and the number of physical constraints applied to the mount total six. This is equivalent to saying that any physical constraints applied are independent (non-redundant). A kinematic mount therefore has six independent constraints. The most common type of kinematic mount is the "cone, groove, and flat" mount illustrated in figure 6.9. Consider the optic as being attached to the co-ordinate system of the three balls in the figure and its corresponding mount having the cone, the groove and the flat. If the optic is first seated in the cone, three degrees of freedom (x, y, z translations) are eliminated without redundancy.

At this point the optic can still rotate freely about all axes. Next, the second ball is seated in the groove which is aligned toward the cone. This constrains two more degrees of freedom, pitch and yaw. Finally, there is only one degree of freedom left to constrain, which is roll. This is accomplished by seating the third sphere on the flat. Six non-redundant constraints make this a kinematic mount.

The kinematic mount shown in figure 6.9 has two sets of cones, grooves and flats so that it has two fully constrained positions 180° apart. Such a mount would have to be constructed with care so that there was not a translation along the optical axis between the two positions.

#### 6.3.2.3 Two Position Quasi-Absolute Cat's-Eye Technique.

Because the alignment of the test surface, in the above technique, gets more difficult as the numerical aperture gets larger, a simpler technique has been developed (Creath and Wyant, 1990).

This technique only requires two measurements, these being the first two measurements of the three position absolute measurement technique (figure 6.6).

The first measurement is with the test surface at the focus of the converging lens:

$$W_{focus} = \frac{W_{conv} + \overline{W}_{conv}}{2} - W_{refarm}$$

The second measurement is with the test surface positioned such that its centre of curvature is at the focus of the converging lens:

$$W_{0°} = W_{testsurf} + W_{conv} - W_{refarm}$$

When the first measurement is subtracted from the second, the result is the wavefront due to the surface plus an error term due to the converging lens:

$$W_{0°} - W_{focus} = W_{testsurf} + \frac{W_{conv} - \overline{W}_{conv}}{2}$$

For aberrations with even symmetry, such as defocus, spherical and astigmatism, the error term is zero since these aberrations cancel out ($W_{conv} = \overline{W}_{conv}$). The difference of the two measurements then becomes

$$W_{0°} - W_{focus} = W_{testsurf}$$

For aberrations with odd symmetry, such as coma, $W_{conv} = -\overline{W}_{conv}$, the difference between the two measurements will be

$$W_{0°} - W_{focus} = W_{testsurf} + W_{conv}$$

Most spherical surfaces do not have coma in them, and because a misalignment of the spherical test surface would not introduce coma into the measurement, it can be assumed that any coma in the measurement should be due to the interferometer. As long as the coma is assumed to be in the interferometer and not in the test surface, it can be subtracted from the measurement to yield the test surface independent of the interferometer.

For higher order aberrations, those with even symmetry will cancel while those with odd symmetry will not cancel and should be subtracted from the measurement result as long as they are not in the test surface.

The odd aberrations in the result may be found by rotating the data set by 180° and subtracting it from the data set before rotation:

$$W_{odd} = (W_{0°} - W_{focus}) - \overline{(W_{0°} - W_{focus})}$$

This technique is much simpler to implement than the three position cat's-eye technique but makes assumptions about the test surface that render it only quasi-absolute. The assumptions made may be valid for spherical surfaces manufactured by conventional means. There are, however, probably no grounds for such assumptions for surfaces formed by other means, such as the reflow of islands of photo-resist to form spherical microlenses.

A very important feature of both cat's-eye methods is the inversion of the wavefront in the cat's-eye position. As can be seen from figure 6.6, in the cat's-eye position, the incident wavefront does not retrace its path upon reflection from the test surface but is inverted. This means that the light source must be temporally and spatially coherent. If a partially coherent source were used in order to reduce coherent noise problems then fringes would not be formed in the cat's eye position unless the wavefront in the reference arm were inverted also (for instance, by using a cat's-eye reflector in place of the plane reference surface). This would

make the absolute measurement techniques unworkable since the contribution of the reference arm, , to the fringe pattern would then be different for the cat's-eye and confocal positions.

## 6.4 Proposal for the Absolute Measurement of Microspheres.

Of the absolute measurement techniques examined above, the method of choice is the three position cat's-eye technique (3.2.2). It is a truly absolute technique, requiring only simple mathematical manipulation of experimental data and does not suffer from limitations on accuracy due to variation of phase shift at high numerical aperture. Moreover, the software to perform the data manipulation exists as an integral part of the Wyko DOS/RTI phase measuring interferometer software which NPL already possesses.

The Twyman-Green is the preferred configuration of the interferometer to perform the measurements. The flexibility in the positioning of optical components that the Twyman-Green affords is particularly advantageous with regard to the measurement of small radius spheres. Because the surface of the spheres will be very close to the focus of the converging lens, the position of the image planes of the surface will be highly sensitive to sphere radius. It will require considerable manoeuvrability in the imaging optics in order to cope with a range of sphere diameters. Indeed, this requirement may, in some measure, explain the apparent inability of the Zygo Fizeau interferometer to properly image the surfaces of small spheres (see section 2.1).

### 6.4.1 Features of the Proposed Twyman-Green Interferometer

A schematic diagram of the proposed Twyman-Green interferometer is shown in figure 6.10. The design has some similarities to that proposed by Shwider and Falkenstorfer (1995) for measuring microspheres except that it utilises highly coherent light in order to permit absolute testing.

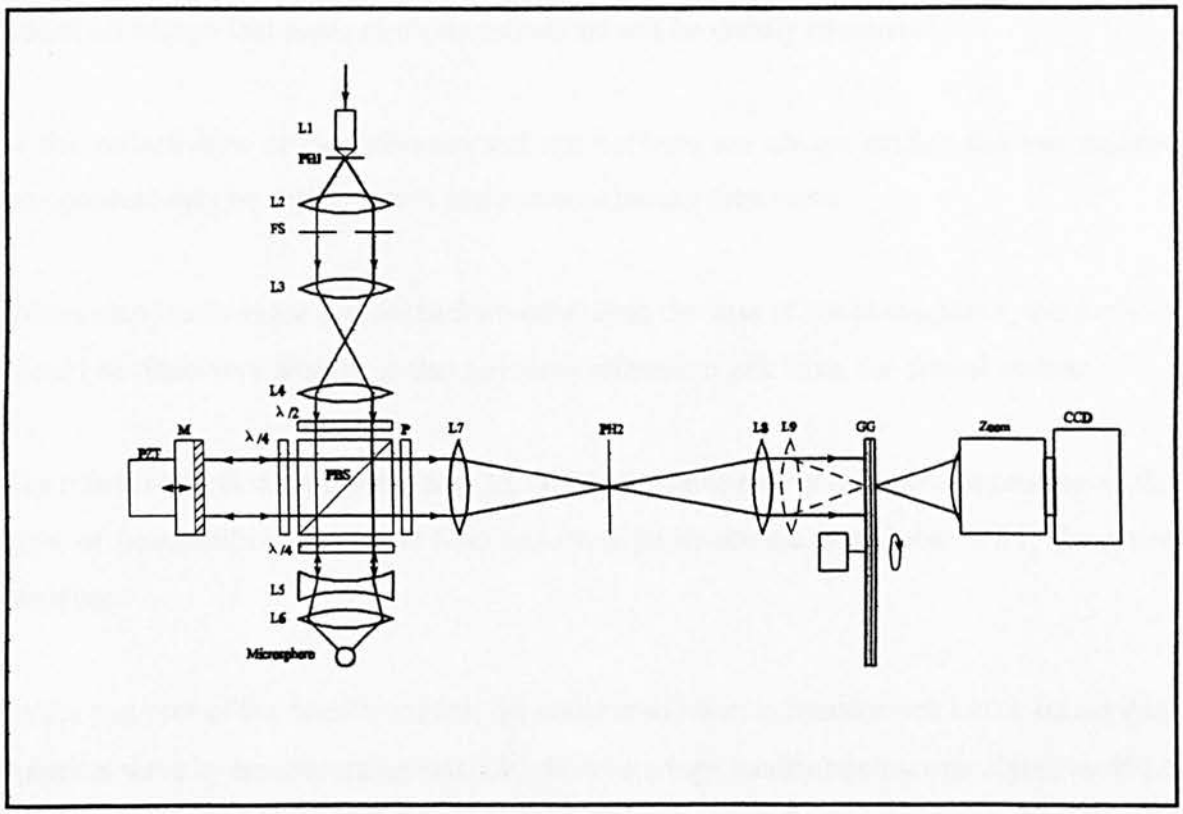


Figure 6.10 Schematic of proposed Twyman-Green interferometer for absolute measurement of spheres.

The light source is a helium-neon laser ($\lambda$ = 632.8 nm), expanded by a spatial filter (objective lens, L1 and pinhole PH1) and collimated by lens L2.

An adjustable field-stop aperture, FS, is imaged onto the test surface by the 4*f* system of lenses, L3 and L4. The field-stop may prove useful, for example, to avoid illuminating the substrate or mount, where the numerical aperture of the interferometer exceeds that of the spherical surface to be measured, thus minimising spurious reflections. If this facility is not required, then FS, L3 and L4 may be omitted.

The beamsplitter is of the polarising variety, consisting of a polarising beamsplitter cube, PBS, half- and quarter-wave retardation plates, $\lambda/2$ and $\lambda/4$, and analyser, P. The quarter-wave plates rotate the plane of polarisation by 90° on double-pass and the half-wave plate is rotated to adjust the intensities in either arm of the interferometer to give high contrast interference fringes. The orthogonally polarised reference and test beams are made to interfere, when they are recombined by the beamsplitter, by analyser, P. Using a polarising beamsplitter has the

added advantage that some spurious reflections will be greatly attenuated.

If the reflectivities of the reference and test surfaces are always similar, the beamsplitter components may be replaced by a conventional beamsplitter cube.

Where plane surfaces are present in the beams, as in the case of the beamsplitter, the surfaces should be tilted very slightly so that any stray reflections exit from the optical system.

The reference surface is a plane mirror, M, which may be coated or uncoated depending on the type of beamsplitter used. It is mounted on a piezo-electric transducer, PZT, for phase stepping.

In the test arm of the interferometer, the collimated beam is transformed into a converging spherical wave by the converging lens, L6, which is a high quality microscope objective. If L6 is not corrected for infinite conjugates, then the negative lens, L5, must be included to simulate the back conjugate (16 cm for standard microscope tube length).

Lenses, L2 to L6 should be chosen for minimum aberration since any misalignment of the test surface introduces errors proportional to the interferometer aberrations (see section 3.2.2). This may, in itself be a good reason for omitting L2 and L3. The most critical component is L6 since it operates at a high numerical aperture and should be of the highest possible quality.

In the exit arm of the interferometer, a magnified real image of the test surface is formed on the rotating ground glass screen, GG, by lenses L7 and L8 . The interference fringe pattern is observed by a CCD camera equipped with a zoom lens.

An additional lens, L9, may be inserted into the beam in order to form a real image of the source pinhole on the ground glass screen. This is to facilitate alignment of the interferometer by superimposing images of the source pinhole reflected from the reference and test surfaces.

An intermediate real image of the source pinhole exists at the focus of lens L7. A pinhole

aperture may be placed at this focus to reduce the problem of spurious reflections and coherent noise as will be described in the next section.

### 6.4.2 Reduction of Coherent Noise.

Since the interferometer must use highly coherent light, there is likely to be a problem of coherent noise due to spurious reflections and scattered light. The coherent noise manifests itself as unwanted high frequency spurious fringes and speckle noise superimposed upon the interference pattern due to the reference and test surfaces. Since the required fringe patterns will be nulled (as near as is possible, a single fringe across the field of view), they will be well separated from the noise in terms of spatial frequency.

Schwider et al (1986) propose removing the noise by a posteriori digital spatial filtering of the fringe pattern by convolution of the data with a suitable function. A similar effect could be achieved in the Wyko DOS/RTI software by representing the fringe patterns in terms of the Zernike polynomials and then setting the high order terms equal to zero.

A simpler and more direct method of removing the coherent noise would be by direct optical spatial filtering. Since the fringe pattern is at the front focus of lens L8, its Fourier transform is located at the back focus, which is also the front focus of lens L7. It is then a simple matter to remove the high spatial frequencies from the fringe pattern by placing a suitable diameter circular aperture (pinhole, PH2) at the rear focus of L8. In effect, L7, L8 and PH2 are a simple $4f$ spatial filter. Geometrically speaking, only light reflected from the reference and test surfaces and forming a real image of the source pinhole at the front focus of L7 will pass through the pinhole PH2. The diameter of the pinhole should be just a little larger than the diameter of the point-spread function of the image of the source pinhole.

Once the fringe pattern has been imaged by L8, the coherence of the light is destroyed by the rotating ground glass screen. This will prevent any further coherent noise being generated by spurious reflections or scattering within the CCD camera or its lens. This feature is common in commercial phase measuring interferometers such as those manufactured by Wyko and Zygo

(see, for instance, Domenicali and Hunter (1980)).

### 6.4.3 Mapping the Test Result to the Spherical Surface.

Having performed the absolute measurement procedure, the result is a two dimensional array of deviations from a perfect spherical surface mapped onto the flat surface of the detector array. In order for the result to be meaningful and useful, it is necessary to know the geometrical relationship between conjugate points on the sphere and the detector array. In other words, the geometrical distortions of the imaging optics must be determined.

Given that a condition for accurate absolute measurement is that the optics be of a very high quality, it may be sufficient to model the geometrical distortions in the imaging optics. This would simply be achieved by entering the prescriptions and positions of the optical elements into an optical ray tracing program.

A preferable approach, however, would be to make direct measurements of the distortions of the imaging optics. One method, used by Gates (1960), is to place an engraved steel ball in the object space of the interferometer and measure the positions of the images of the engraved marks in the image space. This technique would, however, be difficult to implement for very small spheres. An alternative approach may be to illuminate a suitable graticule in the image plane of the interferometer (that normally occupied by the CCD array) and measure its image in the object space of the interferometer (that normally occupied by the test sphere) using a travelling microscope.

Whichever method is used, it will have to be repeated for each radius of sphere tested because of the sensitivity of the image plane position to sphere radius (see section 4).

For most microlenses, the numerical aperture of the interferometer will be sufficient to cover the surface of the lens in a single measurement. For microspheres, however, a single measurement will cover only a sub-aperture of the entire spherical surface. To find the figure of the whole surface of a sphere, multiple sub-apertures must be measured and linked together.

Theoretical approaches to this problem have been reported in the literature (Williams and Kwon,1987; Day et al,1987). The technique involves modelling the full spherical surface in terms of a set of spherical harmonics. The full aperture spherical harmonic coefficients are then estimated from the individual sub aperture Zernike coefficients. At the time of writing, however, these authors had not yet verified their approach in practice.

One of the practical difficulties of measuring the full aperture of a sphere by sub aperture testing is that of controlling the change in orientation of the sphere between measurements. A suggestion made by Shinn (1996) is to adapt a method used in the gem faceting trade. The gem to be worked is waxed onto a "stick" that has an indexing pin. When all the possible facets have been worked with this grip, the gem is transferred to another "stick" with the aid of a fixture, thus retaining the orientation. When making interferometric measurements, the contamination of the test surface by wax would be undesirable and so a gentler method of gripping the spheres, for example by a low vacuum at the end of a small diameter tube, would be more appropriate.

## 6.5 Optical Phase Conjugation.

Phase conjugation is a non-linear optical effect whereby an optical wavefront incident upon the phase conjugating medium is reflected in such a way that its spatially dependant characteristics are reversed. The reflected wavefront thus retraces its incident path exactly, so cancelling any aberrations introduced into the incident beam by optical elements in its path. The effect is therefore also known as wavefront reversal.

Phase conjugation may be produced by a number of different non linear optical interactions such as three wave and four wave mixing, stimulated Brillouin scattering, stimulated Raman scattering and photon echoes. For interferometry the most useful method of producing phase conjugation is by degenerate four wave mixing in photorefractive materials (Yeh,1993). A photorefractive material is one where the local refractive index of the material varies with the local electric field strength, and thus with the local optical intensity.

Figure 6.11 shows, diagrammatically, the four wave mixing process. Consider four waves overlapping in an interaction region within the photorefractive medium. The incident wave, $W_{inc}$ interacts with the write beam, $W_{write}$ to create a spatial interference pattern in the region where they overlap. The resultant optical intensity distribution "writes" a volume phase diffraction grating into the material by the photorefractive effect. This process is analogous to volume holography where the object wave is $W_{inc}$ and the reference wave is $W_{write}$. The read beam, $W_{read}$, is conjugate to $W_{write}$ and interacts with the diffraction grating to produce the phase conjugate of the incident wave, $W_{conj}$. Note that the read and write beams can be interchanged and that interaction between any combination of $W_{inc}$ or $W_{conj}$ with $W_{write}$ or $W_{read}$ would produce the same diffraction grating. The four wave mixing is thus said to be degenerate.

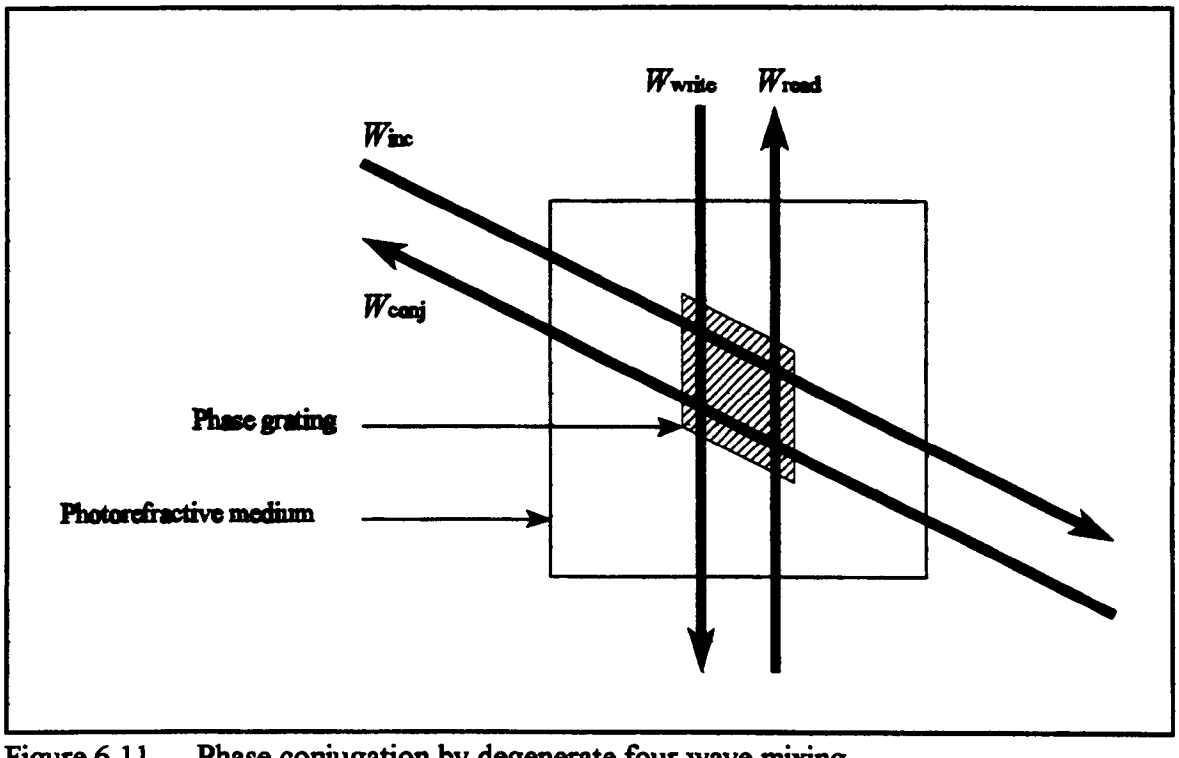


Figure 6.11 Phase conjugation by degenerate four wave mixing.

A phase conjugating device used to produce the conjugate of an incident wave is commonly known as a phase conjugate reflector or phase conjugate mirror (PCM). The read and write beams may be externally applied to the photorefractive medium or may be internally generated by total internal reflection of light scattered from the incident beam. This is known as self-

pumped phase conjugation (Feinberg,1982). A common photorefractive material for interferometry is single crystal Barium Titanate ($BaTiO_3$) which will exhibit photorefractive effects with the relatively low intensities available from helium-neon lasers. A potentially useful feature of externally pumped PCMs is that, for sufficiently intense read and write beams, the reflectivity can be more than 100%.

### 6.5.1 Applications of Phase Conjugation to Absolute Interferometry.

Because of their ability to cancel aberrations in the incident wavefront, phase conjugate mirrors (PCMs) would seem to show great potential for developing interferometer configurations where the test surface is the only source of aberration in the observed fringe pattern. Such an interferometer would be intrinsically absolute since it would be free from error sources due to reference surfaces, beam shaping optics etc..

Several authors have described interferometers using PCMs (for example, Fainman et al,(1981); Ikeda et al,(1982); Gauthier et al,(1989); Shukla et al, (1990); Wang et al, (1994)). Each of these authors has, in a variety of different configurations, generated the reference wavefront by reflection from a PCM rather than from a conventional reference surface. None of the interferometers, however, are intrinsically absolute since optical imperfections that were, in the conventional interferometer, self compensating, have become significant. This appears to be a generally applicable feature of phase conjugate interferometers and is illustrated for the case of a simple Twyman-Green interferometer in figure 6.12.

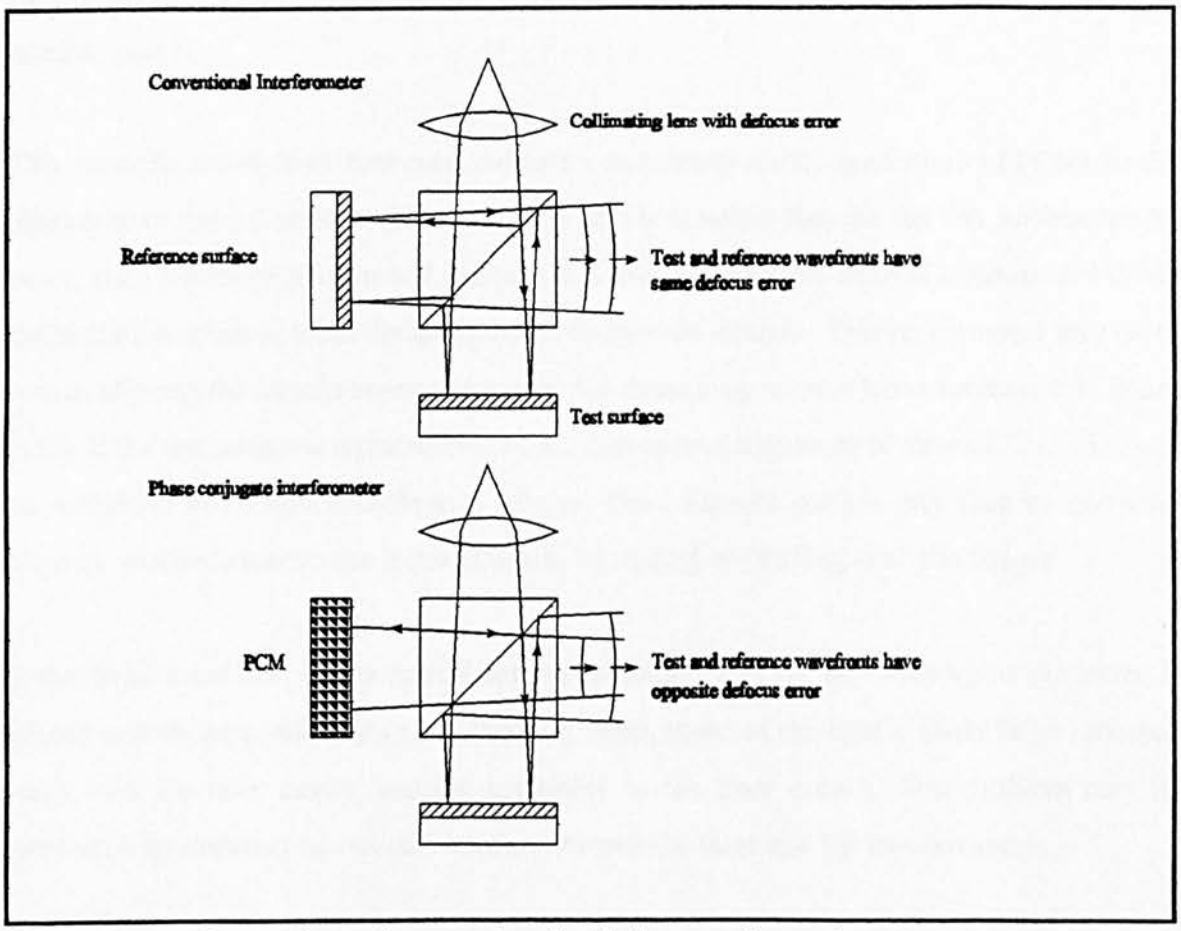


Figure 6.12 Comparison of conventional and phase conjugate interferometers.

Let us assume that the collimating lens in the interferometer has some aberrations (illustrated, in this example, by simple de-focus). In the case of the conventional interferometer the de-focus error is common to both the wavefronts reflected from the test surface and from the reference surface. When the two wavefronts recombine, the defocus error cancels and the fringe pattern is due only to the difference between the reference and test surfaces. For the phase conjugate interferometer, the wavefront reflected from the test surface still contains the de-focus error but this has been eliminated from the reflected reference wave. The fringe pattern is now due to the test surface errors and the collimator errors. In the conventional interferometer the accuracy is limited by the reference surface; in the phase conjugate interferometer the accuracy is limited by the collimator. Neither configuration can be said to have any particular advantage over the other. The principle illustrated by the above example seems to apply wherever PCMs are employed to compensate for the errors in one part of an

interferometer. Whatever advantage is gained is negated by the exaggeration of errors in another part.

The example above does, however, indicate a potentially useful application of PCMs to the alignment of optical systems (Howes,1986). If it is assumed that the flat test surface has no error, then the fringe pattern will indicate how well the incident beam is collimated. Perfect collimation is achieved when the interference fringes are straight. This arrangement may be of use in aligning the interferometer proposed for measuring microspheres (section 4.1, figure 6.10). If the test surface is replaced by a PCM, then correct alignment of lenses L2, L3, L4 will be indicated by straight interference fringes. The reference surface may then be correctly aligned, perpendicular to the incident beam, by nulling or "fluffing out" the fringes.

It should be noted that, in any optical system containing a PCM, or where a plane reflector is placed exactly perpendicular to a collimated beam, some of the light is likely to be reflected back into the laser cavity, causing instability in the laser output. This problem may be prevented by inserting an optical isolator between the laser and the interferometer.

In an interferometer for measuring convex spherical surfaces, the interferometer must produce a converging spherical wavefront to match the nominal shape of the test surface. The optics used to transform the wavefront from the light source into the converging wavefront are likely to have by far the largest aberrations of the whole optical system and so will be the limiting factor on the intrinsic accuracy of the interferometer. It is easy to produce a diverging spherical wavefront free from aberration by using a spatial filter. Reversing the spherical wavefront by using a PCM will produce a converging spherical wavefront free from aberration.

A simple interferometer using this principle is illustrated in figure 6.13. The diverging spherical wavefront is produced by the spatial filter consisting of the objective and pinhole. The wavefront is reversed by the PCM and may be phase modulated by an electric field applied to the photo-refractive crystal in order to allow phase measuring interferometry. The beamsplitter reflects a portion of the converging spherical wave to the test surface where it is reflected to form the test wavefront. The beamsplitter also reflects a portion of the diverging spherical

wavefront to form the reference wavefront. Assuming the beamsplitter to be perfect, the interference pattern due to the test and reference wavefronts will contain information only about the test wavefront. Unfortunately since the beamsplitter cannot be assumed to be perfect, the interferometer is not intrinsically absolute. For some applications, however, where the complicated nature of absolute measurement procedures are unjustified, this arrangement may be of utility since, for a high quality beamsplitter, the aberrations would be much less than for a high numerical aperture converging lens.

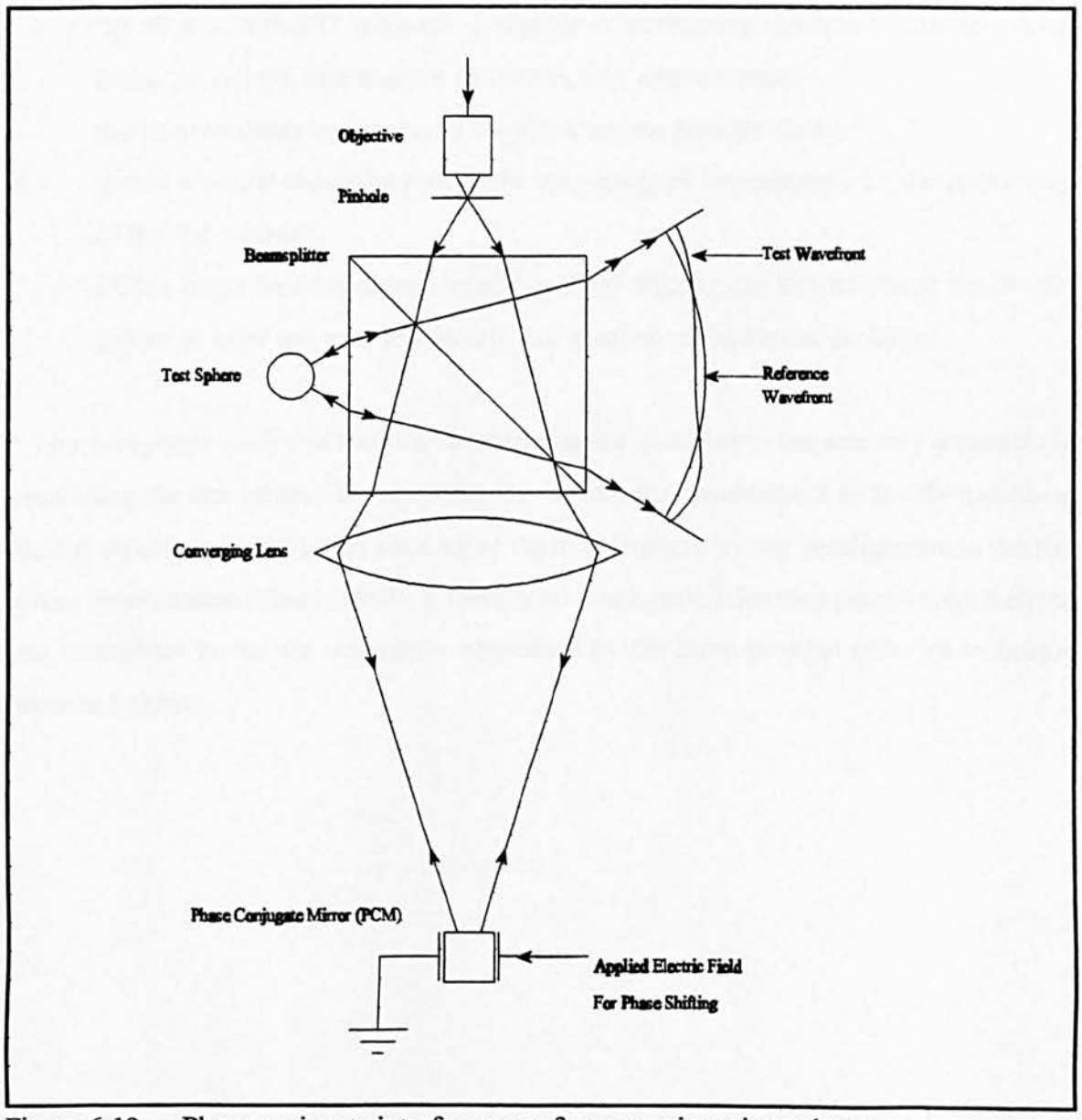


Figure 6.13 Phase conjugate interferometer for measuring microspheres.

## 6.6 Summary.

From the above discussion, recommendations and conclusions are as follows:
the three position cat's-eye absolute measurement technique is that best suited to testing high numerical aperture spherical surfaces;

- a laser Twyman-Green phase measuring interferometer should be constructed with a plane reference surface in one arm and the spherical test surface in the other;
- the Wyko DOS/RTI software is capable of performing the interferometer control functions and the data analysis of the absolute measurement;
- the interferometer optics should be of the highest possible quality*;
- special attention should be paid to the very stringent requirements for the positioning of the test surface*;
- PCMs might be considered a useful tool for aligning the interferometer but do not appear to offer any new solutions to the absolute measurement problem.

* The acceptable quality of the interferometer optics is related to the accuracy attainable in positioning the test sphere. The error introduced into the measurement by the aberrations in the test wavefront is due to the shearing of those aberrations by any misalignment in the test sphere upon rotation (Truax, 1988). If there is no misalignment (perfect positioning) then the test wavefront errors are completely eliminated by the three position cat's eye technique described above.

**Appendix B: Listing of the ZIPL Implementation of the Flatness Testing Algorithm.**

```
10      rem REFERENCE FLAT CALIBRATION PROGRAM
20      rem BY JOHN MITCHELL (KINGSTON UNIVERSITY)
25      on error gosub @_aftermath
30      gosub @_init
40      gosub @_cursors
50      gosub @_acq_cb
60      gosub @_acq_ac
70      gosub @_acq_abr
80      gosub @_acq_ab
90      gosub @_acq_abs1
100     gosub @_acq_abs2
110     gosub @_make_blanks
120     gosub @_do_fff
130     gosub @_do_schulz
140     gosub @_adj_abs
150     gosub @_solve
900     ? ERROR: ? ERROR$
999     end
1000    @_init: clr
1010    rem SETS UP CONSTANTS ETC. FOR LATER USE BY PROGRAM
1020    plotter is 1 : scroll : printer is 1
1030    dim temp1[289] : dim temp2[289]
1040    bad=99999.0
1050    multfact=50
1070    s1= 10: s2= 3: rem pixel shifts for abs1 and abs2
1080    on error goto 900
1090    initialize "M": initialize "U"
1500    ? "DO YOU WISH TO USE OLD DATA <O> OR ACQUIRE NEW DATA
<N>"
```

```
1505  ? "OR MAKE SYNTHETIC DATA <S>?"
1510  input old_new$
1520  if old_new$= "O" or old_new$= "o" then gosub @_old: goto120
1530  if old_new$= "N" or old_new$= "n" then gosub @_new: goto1900
1535  if old_new$= "S" or old_new$= "s" then gosub @_synth: goto 120
1540  ? "INVALID INPUT- TRY AGAIN": goto 1500
1900  return
2000  @_cursors:
2010  rem SETS UP THE CURSORS NEEDED LATER IN THE PROGRAM
2020  coordinates squared
2030  ? "SET UP AN ELLIPTICAL CURSOR AROUND THE IMAGE ON THE
FRINGE"
2040  ? "MONITOR USING THE TRACKBALL"
2050  cursor 1 ellipse
2060  gosub @_minmax
2120  rem HORIZONTAL FIDUCIAL CURSORS FOR abr
2130  cursor 2 point xmin+5,  ycent
2140  cursor 3 point xmax-5, ycent
2150  rem VERTICAL FIDUCIAL CURSORS FOR ab
2160  cursor 4 point xcent, ymin+5
2170  cursor 5 point xcent, ymax-5
2180  rem VERTICAL FIDUCIAL CURSORS FOR abs1
2190  cursor 6 point xcent+s1, ymin+5
2200  cursor 7 point xcent+s1, ymax-5
2210  rem VERTICAL FIDUCIAL CURSORS FOR abs2
2220  cursor 8 point xcent+s2, ymin+5
2230  cursor 9 point xcent+s2, ymax-5
2240  cursor all off
2250  cursor 1 on
2260  copy CURSORS to "B:cursors.cu"
2900  return
```

```
3000  @_acq_cb:
3010  ? "PLACE FLAT C IN INTERFEROMETER AND FLAT B IN MOUNT."
3020  ? "ADJUST FOR A SUITABLE FRINGE PATTERN"
3030  ?
3040  gosub @_acq_data
3050  gosub @_adj_data1
3060  copy DATA to "B:cb.da"
3070  return
4000  @_acq_ac:
4010  ? "PLACE FLAT A I N INTERFEROMETER AND FLAT C IN MOUNT."
4020  ? "ADJUST FOR SUITABLE FRINGE PATTERN ."
4030  ?
4040  gosub @_acq_data
4050  gosub @_adj_data1
4060  copy DATA to "B:ac.da"
4070  return
5000  @_acq_abr:
5010  cursor 2 on: cursor 3 on
5020  ? "PLACE FLAT A IN INTERFEROMETER AND FLAT B IN MOUNT"
5030  ? "SUCH THAT THE ALIGNMENT FIDUCIAL IS ALIGNED WITH    THE"
5040  ? "HORIZONTAL CURSOR MARKS ON THE FRINGE MONITOR."
5050  ? "ADJUST FOR SUITABLE FRINGE PATTERN."
5060  ?
5070  gosub @_acq_data
5080  copy DATA to "B:abr.da"
5090  cursor 2 off: cursor 3 off
5100  return
6000  @_acq_ab:
6010  cursor 4 on: cursor 5 on
6020  gosub @_acq_abv
6030  gosub @_adj_data1
```

```
6040  copy DATA to "B:ab.da"
6050  cursor 4 off: cursor 5 off
6060  return
7000  @_acq_abs1:
7010  cursor 6 on: cursor 7 on
7020  gosub @_acq_abv
7030  copy DATA to "B:abs1.da"
7040  cursor 6 off: cursor 7 off
7050  return
8000  @_acq_abs2:
8010  cursor 8 on: cursor 9 on
8020  gosub @_acq_abv
8030  copy DATA to "B:abs2.da"
8040  cursor 8 off: cursor 9 off
8050  return
9000  @_acq_abv:
9010  ? "PLACE FLAT A IN INTERFEROMETER AND FLAT B IN MOUNT"
9020  ? "SUCH THAT THE ALIGNMENT FIDUCIAL IS ALIGNED WITH"
9030  ? "THE VERTICAL CURSOR MARKS ON THE FRINGE MONITOR."
9040  ? "ADJUST FOR A SUITABLE FRINGE PATTERN."
9050  ?
9060  gosub @_acq_data
9070  return
10000 @_acq_data:
10010 ? "PRESS <RETURN> WHEN READY TO ACQUIRE DATA."
10020  input ready$
10025 gosub @_delay
10030 window 1
10040 average clear
10050 for i=1 to meas_no
10060 acquire : convert: connect
```

```
10070 average sum
10080 next i
10090 average calc
10100 square: invert
10105 multiply multfact
10110 return
11000 @_adj_data1:
11010 ? "ADJUSTING TILT OF DATA"
11020 x0=0: y0=-60: z0=DATAPOINT [xcent][ycent-60]
11030 x1=60: y1=0: z1=DATAPOINT [xcent+60][ycent]
11040 x2=-60: y2=0: z2=DATAPOINT [xcent-60][ycent]
11050 gosub @_find_plane
11060 gosub @_remove_plane
11070 return
12000 @_find_plane:
12010 ? "FINDING COEFFICIENTS OF PLANE"
12020 I=((y1-y0)*(z2-z0)-(z1-z0)*(y2-y0))
12030 J=((z1-z0)*(x2-x0)-(x1-x0)*(z2-z0))
12040 K=((x1-x0)*(y2-y0)-(y1-y0)*(x2-x0))
12050 return
13000 @_remove_plane: ? "REMOVING PLANE"
13010 rem REMOVES A PLANE FROM THE DATASET GIVEN THE PLANE
13020 rem COEFFICIENTS I,J,K AND THE POINT z0,y0,z0
13030 ZGEN [0]=(I*x0+J*y0)/K +z0: ? ZGEN [0]
13040 ZGEN [1]=-1*radius*I/K: ? ZGEN [1]
13050 ZGEN [2]=-1*radius*J/K: ? ZGEN [2]
13060 ZGEN [36]= xcent
13070 ZGEN [37]= ycent
13080 ZGEN [38]= radius
13090 zremove 3
13900 return
```

```
14000 @_make_blanks:
14010 rem CREATE BLANK DATAFILES IN WHICH TO WRITE a AND b    DATA
AS IT  IS CALCULATED
14020 rem DATAFILES WILL INITIALLY CONTAIN "BAD DATA" ie.
14030 rem ALL DATAPOINTS= 99999.0
14040 ZGEN [0]=1
14050 ZGEN [36]= xcent
14060 ZGEN [37]= ycent
14070 ZGEN [38]= 0
14075 zgen 1
14080 mask 1
14090 copy DATA to "B:blank.da"
14110 return
15000 @_do_fff: units internal
15005 gosub @_results_disk
15010 rem PERFORMS STANDARD THREE FLAT TEST TO FIND 1st
      DIAMETER   (VERTICAL)
15020 rem FIND DIAMETER OF FLATS A AND B
15030 gosub @_init_temp
15040 copy "B:ab.da" to DATA
15050 for dpy= ymin to ymax
15060 temp1 [dpy]= DATAPOINT [xcent][dpy]
15065 temp2 [dpy]= DATAPOINT [xcent][dpy]
15070 next dpy
15075 cent= ycent: gosub @_temp_diag
15080 copy "B:cb.da" to DATA
15090 for dpy= ymin to ymax
15100 if temp1 [dpy]= bad then goto 15130
15110 if DATAPOINT [xcent][dpy]= bad then goto 15130
15120 temp1 [dpy]= temp1 [dpy]-DATAPOINT [xcent][dpy]
15125 temp2 [dpy]= temp1 [dpy]+DATAPOINT [xcent][dpy]
```

```
15130 next dpy
15135 gosub @_temp_diag
15140 copy "B:ac.da" to DATA
15150 for dpy= ymin to ymax
15160 if temp1 [dpy]= bad then goto 15190
15170 if DATAPOINT [xcent][dpy]= bad then goto 15190
15180 temp1 [dpy]= (temp1 [dpy]+DATAPOINT [xcent][dpy])/2
15185 temp2 [dpy]= (temp2 [dpy]-DATAPOINT [xcent][dpy])/2
15190 next dpy
15195 gosub @_temp_diag
15200 acent= temp1[ycent]
15210 copy "B:blank.da" to DATA
15220 for dpy= ymin to ymax
15230 DATAPOINT [xcent][dpy]= temp1[dpy]
15240 next dpy
15245 ? "vertical diameter of A"
15246 rows= 19: cols= 1: spacing= 10
15247 gosub @_diagnost
15250 copy DATA to "M:a.da"
15440 bcent= temp2 [ycent]
15450 copy "B:blank.da" to DATA
15460 for dpy= ymin to ymax
15470 DATAPOINT [xcent][dpy]= temp2 [dpy]
15480 next dpy
15485 ? "vertical diameter of B"
15486 gosub @_diagnost
15490 copy DATA to "M:b.da"
15500 return
16000 @_init_temp:
16010 rem INITIALISES TWO ARRAYS FOR TEMPORARY STORAGE OF  DATA
16020 for i= 0 to 288
```

```
16030 temp1 [i]= bad
16040 temp2 [i]= bad
16050 next i
16060 return
17000 @_do_schulz:
17005 gosub @_minmax
17010 rem FINDS HORIZONTAL DIAMETERS OF FLATS A AND B BY
METHOD
17020 rem ACCORDING TO SCHULZ
17030 gosub @_adj_abr
17040 rem FIND HORIZONTAL DIAMETER OF FLAT A
17050 gosub @_init_temp
17060 copy "M:b.da" to DATA
17070 for dpy= ymin to ymax
17080 temp1 [dpy]= DATAPOINT[xcent][dpy]
17090 next dpy
17100 copy "U:abr.da" to DATA
17200 for dpx= xmin to xmax
17210 if temp1 [ycent+xcent-dpx]= bad then goto 17240
17220 if DATAPOINT [dpx][ycent]= bad then goto 17240
17230 temp2 [dpx]= DATAPOINT [dpx][ycent]-temp1 [ycent+xcent-dpx]
17240 next dpx
17250 copy "M:a.da" to DATA
17260 for dpx= xmin to xmax
17270 DATAPOINT [dpx][ycent]= temp2 [dpx]
17280 next dpx
17281 ? "horizontal diameter of A"
17282 rows= 1: cols= 19: spacing = 10
17283 gosub @_diagnost
17285 delete "M:a.da": delete "M:a.at"
17290 copy DATA to "M:a.da"
```

```
17300 rem FIND HORIZONTAL DIAMETER OF FLAT B
17310 gosub @_init_temp
17320 copy "M:a.da" to DATA
17330 for dpy= ymin to ymax
17340 temp1 [dpy]= DATAPOINT [xcent][dpy]
17350 next dpy
17360 copy "U:abr.da" to DATA
17370 for dpy= ymin to ymax
17380 if temp1 [dpy]= bad then goto 17410
17390 if DATAPOINT [xcent][dpy]= bad then goto 17410
17400 temp2 [dpy]= DATAPOINT [xcent][dpy]-temp1 [dpy]
17410 next dpy
17420 copy "M:b.da" to DATA
17430 for dpx= xmin to xmax
17440 DATAPOINT [dpx][ycent]= temp2 [ycent+xcent-dpx]
17450 next dpx
17451 ? "horizontal diameter of B"
17452 gosub @_diagnost
17455 delete "M:b.da": delete "M:b.at"
17460 copy DATA to "M:b.da"
17470 return
18000 @_adj_abr:
18010 ? "ADJUSTS TILT OF abr DATASET"
18020 copy "B:abr.da" to DATA
18030 x0= 0: y0= 0: z0=(DATAPOINT [xcent] [ycent])-acent-bcent
18040 x1= 60: y1= 0: z1= DATAPOINT [xcent+60][ycent]
18050 x2= 0: y2= -60: z2= DATAPOINT [xcent][ycent-60]
18060 gosub @_find_plane
18070 gosub @_remove _plane
18074 delete "B:abr.da": delete "B:abr.at"
18075 ? "abr after adj_abr"
```

```
18076 gosub @_diagnost
18080 copy DATA to "B:abr.da": copy DATA to "U:abr.da"
18090 return
19000 @_old:
19010 ? "PLACE DISK CONTAINING DATA PREVIOUSLY ACQUIRED IN
      DRIVE B":?
19020 ? "PRESS <RETURN> WHEN READY"
19030 input ready$
19040 copy "B:cursors.cu" to CURSORS
19050 gosub @_minmax
19900 return
20000 @_new:
20010 ? "*************************************************************"
20020 ? "IMPORTANT- PLACE A BLANK DISK OR DISK CONTAINING NO
      IMPORTANT DATA"
20030 ? "IN DRIVE B. THIS DISK WILL BE INITIALISED, IE. WIPED CLEAN
      BEFORE"
20040 ? "PROCESSING BEGINS
20050 ? "*************************************************************"
20060 ? "PRESS <RETURN> WHEN READY"
20070 input ready$
20080 initialise "B"
20090 ? "INPUT NUMBER OF MEASUREMENTS TO BE AVERAGED"
20100 input meas_no
20110 ? "ENTER DELAY BEFORE ACQUISITION (SECONDS)"
20120 input delay
20900 return
21000 @_minmax:
21010 xmin= CURSOR [1][1]
21020 ymin= CURSOR [1][2]
21030 xmax= CURSOR [1][3]
```

```
21040 ymax= CURSOR [1][4]
21050 xcent= int((xmin+xmax)/2)
21060 ycent= int((ymin+ymax)/2)
21070 radius= int((xmax-xmin)/2)
21080 if int((ymax-ymin)/2) > radius then radius= int((ymax-ymin)/2)
21900 return
22000 @_delay:
22010 TIMER= 0
22020 if TIMER > delay*1000 then goto 22900
22030 goto 22020
22900 return
23000 @_adj_abs:
23005 ? "adjusting abs"
23010 gosub @_adj_xtilt
23015 copy "B:ab.da" to "U:ab.da": copy "B:ab.at" to "U:ab.at"
23020 copy "B:abs1.da" to "M:abs.da"
23025 copy "B:abs1.at" to "M:abs.at"
23026 ? "find y-tilt"
23030 s= s1: rows= 19: cols= 19: spacing = 10
23031 ? "abs1"
23040 gosub @_find_ytilt
23050 delete "B:abs1.da": delete "B:abs1.at"
23060 copy DATA to "B:abs1.da"
23070 delete "M:abs.da": delete "M:abs.at"
23080 copy "B:abs2.da" to "M:abs.da"
23085 copy "B:abs2.at" to "M:abs.at"
23090 s=s2: rows=19: cols= 19: spacing = 3
23091 ? "abs2"
23100 gosub @_find_ytilt
23110 delete "B:abs2.da": delete "B:abs2.at"
23120 copy DATA to "B:abs2.da"
```

```
23900 return
24000 @_adj_xtilt:
24001 ? "adj x tilt"
24010 copy "M:b.da" to DATA
24020 P1= DATAPOINT [xcent+60+s1][ycent]: ? "P1=";P1
24025 P3= DATAPOINT [xcent+60+s2][ycent]: ? "P3=";P3
24050 P2= DATAPOINT [xcent-60+s1][ycent]: ? "P2=";P2
24060 P4= DATAPOINT [xcent-60+s2][ycent]: ? "P4=";P4
24070 copy "B:abs1.da" to DATA
24080 x0= -60: y0= 0: z0= DATAPOINT [xcent-60][ycent]-P1
24090 x1= 60: y1= 0: z1= DATAPOINT [xcent+60][ycent]-P2
24100 x2= x0: y2= 10: z2=  z0
24105 ? "abs1"
24110 gosub @_find_plane
24120 gosub @_remove_plane
24125 rows= 1: cols= 19: spacing = 10
24126 ? "Horizontal diameter of abs1 after adj_xtilt"
24130 delete "B:abs1.da": delete "B:abs1.at"
24140 copy DATA to "B:abs1.da"
24150 copy "B:abs2.da" to DATA
24160 x0= -60: y0= 0: z0= DATAPOINT [xcent-60][ycent]-P3
24170 x1= 60: y1= 0: z1= DATAPOINT [xcent+60][ycent]-P4
24180 x2= x0: y2= 10: z2= z0
24185 ? "abs2"
24190 gosub @_find_plane
24200 gosub @_remove_plane
24201 ? "Horizontal diameter of abs2 after adj_xtilt"
24205 gosub @_diagnost
24210 delete "B:abs2.da": delete "B:abs2.at"
24220 copy DATA to "B:abs2.da"
24900 return
```

```
25000 @_find_ytilt:
25010 copy "M:a.da" to DATA
25020 for dpy= ymin to ymax
25030 temp1[dpy]= DATAPOINT [xcent][dpy]
25040 next dpy
25050 strt= 0: finish= radius
25055 gosub @_step_chords1
25060 copy "M:b.da" to DATA
25070 for dpy= ymin to ymax
25080 temp1 [dp]= DATAPOINT[xcent][dpy]
25090 next dpy
25100 strt= 0: finish= radius
25110 gosub @_step_chords2
25115 ? "B after stepchords2"
25116 gosub @_diagnost
25120 gosub @_init_temp
25130 copy "M:a.da" to DATA
25131 ? "A after stepchords2"
25132 gosub @_diagnost
25140 for dp= s to radius step s
25150 temp1 [xcent+dp]=  DATAPOINT [xcent+dp][ycent+dp]
25160 temp1 [xcent-dp]= DATAPOINT [xcent-dp][ycent-dp]
25170 next dp
25180 copy "M:b.da" to DATA
25190 for dp= s to radius step s
25200 if DATAPOINT [xcent-dp][ycent-dp]= bad then temp1 [xcent+dp]= bad
25205 if temp1 [xcent+dp]= bad then goto 25220
25210 temp1 [xcent+dp]= temp1 [xcent+dp]+ DATAPOINT [xcent-dp][ycent-dp]
25220 if DATAPOINT [xcent+dp][ycent+dp]= bad then temp1 [xcent-dp]= bad
25225 if temp1 [xcent-dp]= bad then goto 25240
25230 temp1 [xcent-dp]= temp1 [xcent-dp]+DATAPOINT [xcent+dp][ycent+dp]
```

```
25240 next dp
25260 copy "U:abr.da" to DATA
25270 for dp= s to radius step s
25280 if DATAPOINT [xcent+dp][ycent+dp]= bad then temp1[xcent+dp]= bad
25285 if temp1 [xcent+dp]= bad then goto 25300
25290 temp1 [xcent+dp]= (temp1 [xcent+dp]- DATAPOINT
[xcent+dp][ycent+dp])/2
25300 if DATAPOINT [xcent-dp][ycent-dp]= bad then temp1[xcent-dp]= bad
25305 if temp1 [xcent-dp]= bad then goto 25320
25310 temp1 [xcent-dp]= (temp1 [xcent-dp]-DATAPOINT [xcent-dp][ycent-dp])/2
25320 next dp
25330 n= 0: tot= 0
25340 for dp= s to radius step s
25350 if temp1 [xcent+dp]= bad then goto 25380
25360 n= n+1
25370 tot= tot+(temp1 [xcent+dp]*s*s/(dp*dp))
25380 if temp1 [xcent-dp]= bad then goto 25410
25390 n=n+1
25400 tot= tot+(temp1 [xcent-dp]*s*s/(dp*dp))
25410 next dp
25420 y_tilt= tot*radius/(s*n)
25430 ZGEN [0]= 0
25440 ZGEN [1]= 0
25450 ZGEN [2]= y_tilt
25460 ZGEN [36]= xcent
25470 ZGEN [37]= ycent
25480 ZGEN [38]= radius
25490 copy "M:abs.da" to DATA
25500 zremove 3
25900 return
26000 @_step_chords1: fmat 0,0,4,0: printer is 2: ? "step_chords1"
```

```
26010 for chord= strt to finish step s
26015 if xcent+chord+s > xmax then goto 26100
26020 copy "M:abs.da" to DATA
26030 gosub @_sub_temp1a
26040 copy "M:b.da" to DATA
26045 if xcent+chord+s > xmax then goto 26080
26050 for dpy= ymin to ymax
26055 if xcent+chord+s > xmax then goto 26070
26056 if temp2 [dpy] > 32766 then goto 26070
26057 if temp2 [dpy] < -32767 then goto 26070
26060 DATAPOINT [xcent+chord+s][dpy]= temp2 [dpy]
26070 next dpy
26080 delete "M:b.da": delete "M:b.at"
26090 copy DATA to "M:b.da"
26100 copy "U:ab.da" to DATA
26110 gosub @_sub_temp2a
26115 if xcent-chord-s < xmin then goto 26180
26120 copy "M:a.da" to DATA
26130 for dpy= ymin to ymax
26135 if xcent-chord-s < xmin then goto 26150
26136 if temp1 [dpy] > 32766 then goto 26150
26137 if temp1 [dpy] < -32767 then goto 26150
26140 DATAPOINT [xcent-chord-s][dpy]= temp1 [dpy]
26150 next dpy
26160 delete "M:a.da": delete "M:a.at"
26170 copy DATA to "M:a.da"
26180 next chord
26900 return
27000 @_step_chords2: ? "step_chords2"
27010 for chord= strt+s to finish step s
27020 copy "M:abs.da" to DATA
```

```
27030 gosub @_sub_temp1b
27040 copy "M:a.da" to DATA
27050 for dpy= ymin to ymax
27051 if temp2 [dpy] > 32766 then goto 27070
27052 if temp2 [dpy] < -32767 then goto 27070
27060 DATAPOINT [xcent+chord][dpy]= temp2 [dpy]
27070 next dpy
27080 delete "M:a.da": delete "M:a.at"
27090 copy DATA to "M:a.da"
27100 copy "U:ab.da" to DATA
27110 gosub @_sub_temp2b
27120 copy "M:b.da" to DATA
27130 for dpy= ymin to ymax
27131 if temp1 [dpy] > 32766 then goto 27150
27132 if temp1 [dpy] < -32767 then goto 27150
27140 DATAPOINT [xcent-chord][dpy]= temp1 [dpy]
27150 next dpy
27160 delete "M:b.da": delete "M:b.at"
27170 copy DATA to "M:b.da"
27180 next chord
27900 return
28000 @_sub temp1a:
28010 for dpy= ymin to ymax
28020 temp2 [dpy]= bad
28025 dpoint= DATAPOINT [xcent-chord][dpy]
28030 if dpoint= bad then goto 28060
28040 if temp1 [dpy]= bad then goto 28060
28050 temp2 [dpy]= dpoint-temp1 [dpy]
28060 next dpy
28499 return
28500 @_sub_temp1b:
```

```
28510 for dpy= ymin to ymax
28520 temp2 [dpy]= bad
28525 dpoint= DATAPOINT [xcent+chord][dpy]
28530 if dpoint= bad then goto 28560
28540 if temp1 [dpy]= bad then goto 28560
28550 temp2 [dpy]= dpoint-temp1 [dpy]
28560 next dpy
28900 return
29000 @_sub_temp2a:
29005 if xcent-chord-s < xmin then goto 29499
29010 for dpy =ymin to ymax
29020 temp1 [dpy]= bad
29024 if xcent-chord-s < xmin then goto 29060
29025 dpoint= DATAPOINT [xcent-chord-s][dpy]
29030 if dpoint= bad then goto 29060
29040 if temp2 [dpy]= bad then goto 29060
29050 temp1 [dpy]= dpoint-temp2 [dpy]
29060 next dpy
29499 return
29500 @_sub_temp2b:
29510 for dpy= ymin to ymax
29520 temp1 [dpy]= bad
29525 dpoint= DATAPOINT [xcent+chord][dpy]
29530 if dpoint= bad then goto 29560
29540 if temp2 [dpy]= bad then goto 20560
29550 temp1 [dpy]= dpoint- temp2[dpy]
29560 next dpy
29900 return
30000 @_solve:
30005 ? "solve"
30006 ? "find 3-spaced starting chords"
```

```
30010 delete "M:abs.da": delete "M:abs.at"
30020 copy "B:abs2.da" to "M:abs.da"
30030 copy "B:abs2.at" to "M:abs.at"
30040 s= s2
30050 copy "M:a.da" to DATA
30060 for dpy= ymin to ymax
30070 temp1 [dpy]= DATAPOINT [xcent][dpy]
30080 next dpy
30090 strt= 0: finish= 18
30095 rows= 19: cols= 11: spacing= 3
30100 gosub @_step_chords1
30105 ? "A after stepchords1"
30106 gosub @_diagnost
30110 copy "M:b.da" to DATA
30120 for dpy= ymin to ymax
30130 temp1 [dpy]= DATAPOINT [xcent][dpy]
30140 next dpy
30150 strt= 0: finish= 15
30160 gosub @_step_chords2
30165 ? "B after step_chords2"
30166 gosub @_diagnost
30170 delete "M:abs.da": delete "M:abs.at"
30180 copy "B:abs1.da" to "M:abs.da"
30190 copy "B:abs1.at" to "M:abs.at"
30200 s=s1
30205 ? "step on 10s"
30210 for strt= -15 to 12 step s2
30215 ? "strt="; strt
30220 copy "M:a.da" to DATA
30230 for dpy= ymin to ymax
30240 temp1 [dpy]= DATAPOINT [xcent-strt][dpy]
```

```
30250 next dpy
30260 finish= radius
30270 gosub @_step_chords1
30280 copy "M:b.da" to DATA
30290 for dpy= ymin to ymax
30300 temp1 [dpy]= DATAPOINT [xcent-strt][dpy]
30310 next dpy
30320 finish= radius
30330 gosub @_step_chords2
30340 next strt
30350 copy "M:a.da" to DATA: multiply (1/multfact)
30360 copy DATA to "A:a.da"
30370 copy "M:b.da" to DATA: multiply (1/multfact)
30380 copy DATA to "A:b.da"
30900 return
31000 @_diagnost:
31010 fmat 0,0,11,0
31020 units internal
31030 firstrow= ycent-(int(rows/2)*spacing)
31040 firstcol= xcent-(int(cols/2)*spacing)
31050 printer is 2
31060 ? "first row of ";rows;"is ";firstrow
31070 ? "first column of ";cols;"is ";firstcol
31080 ? "spacing is ";spacing;"pixels"
31085 ?:?
31090 for dpy= firstrow to firstrow+((rows-1)*spacing) step spacing
31095 ind= 0
31100 for dpx= firstcol to firstcol+((cols-1)*spacing) step spacing
31110 ? DATAPOINT [dpx][dpy];"   ";
31115 ind= ind+1
31116 if ind= 6 then ?: ind=0
```

```
31120 next dpx
31130 ?
31140 next dpy
31145 ?:?
31999 return
32000 @_aftermath:
32010 ? ERROR: ? ERROR$
32020 ? "Aftermath after fatal error"
32030 ? "contents of A datafile"
32040 copy "M:a.da" to DATA
32050 rows= 19: cols= 19: spacing= 10
32060 gosub @_diagnost
32070  spacing= 3
32080 gosub @_diagnost
32090 ? "contents of B datafile"
32100 copy "M:b.da" to DATA
32110 spacing = 10
32120 gosub @_diagnost
32130 spacing= 3
32140 gosub @_diagnost
32210 end
32999 return
33000 @_temp_diag:
33001 units internal
33005 printer is 2
33010 ? "temp1                    temp2"
33020 fmat 0,0,20,0
33030 ? "center pixel is ";cent
33050 for dp= cent-90 to cent+90 step 10
33060 ? temp1 [dp];
33070 ? temp2 [dp]
```

```
33080 next dp
33999 return
34000 @_synth:
34010 ? "Place a disk containing a datafile named ab.da and corresponding "
34020 ? "cursor file in drive B"
34025 ? "Press return when ready": input ready$
34030 copy "B:cursors.cu to CURSORS
34040 gosub @_minmax
34050 copy "B:ab.da" to DATA
34060 rem copy DATA to "U:A.da"
34070 units waves
34080 window calc 1: window data 1
34090 ? "place disk to receive synthesised data in drive B"
34100 ? "NB. this disk will be initialised
34110 ? "Press return when ready"
34115 input ready$
34120 initialise "B"
34130 copy CURSORS to "B:cursors.cu"
34150 ZGEN [36]= xcent
34160 ZGEN [37]= ycent
34170 ZGEN [38]= radius
34180 for i= 0 to 35
34190 ZGEN [i]= 0
34200 next i
34201 ZGEN [0]= 0.1: ZGEN [1]= 0.05: ZGEN [2]= 0.1
34202 zgen 5
34203 copy DATA to "U:A.da"
34210 rem AB
34215 units waves: gosub @_zgen_clear
34220 ZGEN [4]= 0.1
34230 zgen 5
```

```
34240 invert
34250 copy DATA to "U:X.da"
34260 copy "U:A.da" to DATA
34270 subtract "U:X.da"
34274 multiply multfact
34275 gosub @_adj_data1
34280 copy DATA to "B:ab.da"
34285 ? "AB": rows= 19: cols= 19: spacing= 10: gosub @_diagnost
34290 rem CB
34295 gosub @_zgen_clear
34296 units waves
34300 ZGEN [3]= 0.1
34320 zgen 5
34330 subtract "U:X.da"
34335 multiply multfact
34340 gosub @_adj_data1
34350 copy DATA to "B:cb.da"
34355 ? "CB": gosub @_diagnost
34360 rem AC
34365 gosub @_zgen_clear
34370 delete "U:X.da": delete "U:X.at"
34375 units waves
34380 ZGEN [3]= 0.1
34400 zgen 5
34410 invert
34420 copy DATA to "U:X.da"
34430 copy "U:A.da" to DATA
34440 subtract "U:X.da"
34445 multiply multfact
34450 gosub @_adj_data1
34460 copy DATA to "B:ac.da"
```

```
34465 ? "AC": gosub @_diagnost
34470 rem ABR
34475 gosub @_zgen_clear
34480 delete "U:X.da": delete "U:X.at"
34490 units waves
34500 ZGEN [4]= -0.1
34510 zgen 5
34520 invert
34530 copy DATA to "U:X.da"
34540 copy "U:A.da" to DATA
34550 subtract "U:X.da"
34555 multipy multfact
34560 copy DATA to "B:abr.da"
34565 ? "ABR": gosub @_diagnost
34570 rem ABS1
34580 delete "U:X.da": delete "U:X.at"
34590  ZGEN [36]= xcent+s1
34595 units waves
34600 ZGEN [4]= 0.1
34620 zgen 5
34630 invert
34640 copy DATA to "U:X.da"
34650 copy "U:A.da" to DATA
34660 subtract "U:X.da"
34665 multiply multfact
34670 copy DATA to "B:abs1.da"
34675 ? "ABS1": gosub @_diagnost
34680 rem ABS2
34690 delete "U:X.da": delete "U:X.at"
34700 ZGEN [36]= xcent+s2
34705 units waves
```

```
34710 ZGEN [4]= 0.1
34730 zgen 5
34740 invert
34750 copy DATA to "U:X.da"
34760 copy "U:A.da" to DATA
34770 subtract "U:X.da"
34775 multiply multfact
34780 copy DATA to "B:abs2.da"
34785 ? "ABS2": gosub @_diagnost
34790 for i= 0 to 38
34800 ZGEN [i]= 0
34810 next i
34820 initialize "U"
34830 gosub @_make_blanks
34999 return
35000 @_results_disk:
35010 ? "Please replace program disk in drive A"
35020 ? "with a disk to record the results"
35030 ? "This disk will be initialised !"
35040 ? "Press return when ready"
35050 input ready$
35060 ? "Are you sure you have removed the program disk? <Y>"
35070 input ready$
35080 if ready$= "Y" or ready$= "y" then initialize "A": goto 35999
35090 goto 35010
35999 return
36000 @_zgen_clear:
36010 for i= 0 to 35
36020 ZGEN [i]= 0
36030 next i
36999 return
```

```
37000 @_chord_print:
37010 fmat 0,0,6,0
37020 units internal
37030 printer is 1
37040 ? "Prints out entire contents of selected horizontal chord from DATA"
37050 ? "Enter which chord (in relation to center) you wish to print out"
37060 input ch
37070 printer is 2
37074 ind=0
37080 for dpx= xmin to xmax
37100 ? DATAPOINT [dpx][ycent+ch]; "  ";
37110 ind= ind+1
37120 if ind= 10 then ?: ind= 0
37130 next dpx
37135 ?:?:?:?
37140 printer is 1
37150 return
```

## REFERENCES


Ai, C., Wyant, J. C.,"Effect of piezoelectric transducer nonlinearity on phase shift interferometry." *Applied Optics* **26** (6) 1112-1116 (1987)

Ai, C., Wyant, J. C., "Effect of spurious reflection on phase shift interferometry." *Applied Optics* **27** (14) 3039-3045 (1988).

Ai, C., Wyant, J. C., "Absolute testing of flats decomposed to even and odd functions.", *Proc. SPIE* **1756** 73-83 (1992a).

Ai, C., Wyant, J. C.,"Modified three flat method using even and odd functions.",*OSA Tech. Digest* **24** 71-74 (1992b).

Barnes, T. H., "Heterodyne Fizeau interferometer for testing flat surfaces." *Applied Optics* **26** (14) 2804-2809 (1987).

Bennet, J. M., Elings, V., Kjoller, K., "Recent developments in profiling optical surfaces." *Applied Optics* **32** (19) 3442-3447 (1993).

Bennett, S. J. "An absolute interferometric dilatometer." *J. Phys. E: Sci. Instrum.* **10** 525-530 (1977).

Betend-Bon, J. P., Wosinski, L., Breidne, M., Robertsson, L.,"Fiber optic interferometer for testing conic section surfaces.", *Applied Optics* **30** (13) 1715-1722 (1991).

Biegen, J. F., "Interferometric surface profiler." *U.S. Patent* 4732483 (1988).

Birch, K. G., "Interferometric examination of lenses and other components." *Proc. SPIE* **163** 112-120 (1979).

Birch, K. P., "An automatic absolute interferometric dilatometer." *J.Phys. E: Sci Instrum.* **20** 1387-1392 (1987).

Bissinger, H. D., "How to interpret an interferogram." *ASTM STP* **666** 11-28 (1978).

Boas, M. L. , *Mathematical Methods in the Physical Sciences* (2nd Edition) J. Wiley & Sons (1983).

Boyd, R.W., Habashy, T.M., Jacobs, A. A., Mandel, L., Nieto-Vesperinas, M., Tompkin, W.R., Wolf, E., "Nature of the interference pattern produced on reflection at a phase-conjugate mirror." *Optics Letters* **12** (1) 42-44 (1987).

Breitmeier, U., Ahlers, R.-J., "Dynamically focusing electro-optical sensor system for microprofilometry," *Proc. SPIE.* **802** 170-173 (1987).

Briers, J. B., "Interferometric testing of optical systems: a review." *Optics and Laser Technology* 28-41 February (1972).

Bristow, T. C., "Surface metrology testing- frequency footprints." *Optics and Photonics News* 18-20 May (1991).

Brophy, C. P., "Effect of intensity error correlation on the computed phase of phase shifting interferometry." *J. Opt. Soc. Am.* **7** 537 (1990).

Bruce, C. F., Sharples, F. P., "Relative flatness measurement of uncoated optical flats." *Applied Optics* **14** (12) 3082-3085 (1975).

Bruning, J. H., Herriott, D. R., Gallagher, J. E., Rosenfeld, D. P., White, A. D., Brangaccio, D. J., " Digital wavefront measuring interferometer for testing optical surfaces and lenses", *Applied Optics* **13** (11) 2693-2703 (1974)

Bunnagel, R., "Herstellung und verwendung eines gegen erschutterungen unenpfindlichen quecksilberspiegels als ebenheitsnormal." *Optica Acta* **3** (2) 81-84 (1966).

Carré, P. "Installation et utilization du comparateur photoelectrique et interferentiel du Bureau International Des Poids Et Mesures", *Metrologia* **2** (1), 13-23 (1966).

Chen, J., Song, D., Zhu, R., Chen, D., Wang, Q., Chen, L., Zhang, Y., "Large aperture high accuracy phase shifting digital flat interferometer." *Proc. SPIE* **2003** 367-375 (1993).

Chang, S., Isono,M., Sato, T., "Dynamic interferometry using two phase conjugate waves." *Applied Optics* **27** (22) 4735-4738 (1988).

Cheng, Y., Wyant, J. C., "Phase shifter calibration in phase shifting interferometry." *Applied Optics* **24** (18) 3049-3052 (1985).

Cochran, E. R., "Limitations of phase-measuring interferometry for surface characterisation and testing." *Proc. SPIE* **1776** 151-157 (1992).

Crane, R, "Interference phase measurement." *Applied Optics* **8** 538-542 (1969).

Creath, K., *Advanced Interferometric Optical Testing*, SC39 Short Course Notes, SPIE'S International Symposium on Optical Science and Engineering, San Diego (1992).

Creath, K, Wyant, J. C., "Absolute measurement of spherical surfaces" *Proc. SPIE* **1332** 2-7 (1990).

Creath, K, Wyant, J.C., "Holographic and speckle tests." in Malacara (1992).

Day, R. D., Beery, T. A., Lawrence, G. N., "Sphericity Of Full Spheres Using Sub Aperture Optical Testing Techniques" *Proc. SPIE* **661**, 334-341 (1986).

Debenham, M., Dew, G. D., "Dtermining the gravitational deformation of an optical flat by liquid surface interferometry." *Precision Engineering* **2** 93-97 (1980).

Dew, G.D., "The measurement of optical flatness." *J. Sci. Instrum.* **43** 409-415 (1966).

Domenicali, P. L., Hunter, G. C., "Optical interferometer system with CCTV camera for measuring a wide range of aperture sizes", *U.S. Patent* 4201473 (1980).

Dorband, B., Wiedmann, W., Wegmann, W., Kubler, W., Freischlad, K. R., "Software concept for the new Zeiss interferometer." *Proc. SPIE* **1332** 664-672 (1990).

Downs, M. J., McGiven, W. H., Ferguson, H. J., "Measurement of the profiles of supersmooth surfaces using optical interferometry." *Proc. SPIE.* **1009** 14-17 (1989).

Dukhopel, I. I., "Interference methods and instruments for inspecting optical flats." *Sov. J. Opt. Technol.* **38** (9) 570-578 (1971).

Elssner, K. E., Burow, R., Schulz, G., Vogel, A., "Absolute interferometric surface testing: a stringent need." *Proc. SPIE* **1781** (1992).

Elssner, K. E., Vogel, A., Grzanna, J., Schulz, G., "Establising a flatness standard." *Applied Optics* **33** (13) 2437-2446 (1994).

Elssner, K.E., Burow, R., Grzanna, J., Spolaczyk, R., "Absolute sphericity measurement", *Applied Optics* **28** (21) 4649-4661 (1989).

Emerson, W. B., "Dterminationb of the planeness and bending of optical flats.", *J. Res. Nat. Bur. Stand.* **49** (4) 241-247 (1952).

Evans, C. J., Estler, W. T., Parks, R. E., "Absolute calibration of optical flats using a commercial phase measuring interferometer." *OSA Tech. Digest* **24** 75-77 (1992).

Fainman, Y., Lenz, E., Shamir, J., "Contouring by phase conjugation", *Applied Optics* **20** (1), 158-163 (1981).

Feinberg, J., "Self-pumped continuous wave phase conjugator using internal reflection", *Optics Letters* 7 486-488 (1982).

Feinberg, J., "Interferometer with self-pumped phase-conjugating mirror." *Optics Letters* **8** (11) 569-571 (1983).

Feinberg, J., "Photorefractive non-linear optics." *Physics Today* 46-52 (October 1988).

Fischer, D. J., Stahl, H. P., "Mechanism for surface fitting of interferometric slope data." *Proc SPIE* **2003** 224-232 (1993).

Freischlad, K., Kuchel, M., Wiedmenn, W., Kaiser, W., Mayer, M., "High precision interferometric testing of spherical mirrors with long radius of curvature." *Proc. SPIE* **1332** 8-17 (1990a).

Freischlad, K., Kuchel, M., Schuster, K. H., Wegman, U., Kaiser, W., "Real-time wavefront measurement with lambda/10 fringe spacing for the optical shop." *Proc. SPIE* **1332** 18-24 (1990b).

Fritz, B. S., "Absolute calibration of an optical flat", *Proc.SPIE* **433** 123-130 (1983).

Fritz, B. S., "Absolute calibration of an optical flat", *Optical Engineering* **23** (4) 379-383 (1984).

Gates, J. W., "Sphericity interferometer", *Proc Symposium on Interferometry*, NPL **11** 283-284, pub. HMSO (1960).

Gauthier, D. J., Boyd, R.W., Jungsquist, R. K., Lisson, J. B., Voci, L. L., "Phase conjugate

Fizeau interferometer", *Optics Letters* **14** (6), 323-325 (1989).

Geary, J. M., Parker, L. J., "New test for cylindrical optics." *Optical Engineering* **26** (8) 813-820 (1987).

Geary, J. M., "Testing cylindrical lenses." *Optical Engineering* **26** (12) 1219-1224 (1987).

Geary, J. M., Maeda, R., "Interferometry on grazing incidence optics." *Optical Engineering* **26** (12) 1225-1228 (1987).

Geary, J. M., "Data analysis in fiber optic testing of cylindrical optics." *Optical Engineering* **28** (3) 212-216 (1989).

Glassman, A. T., "A manual method for interpretation of nominally plane wavefronts." *ASTM STP* **666** 3-10 (1978).

Glenn, P., "Angstrom level profilometry for sub-millimeter to meter scale surface errors." *Proc. SPIE.* **1333** (1990a).

Glenn, P., "Robust, Angstrom level circularity profilometry." Proc. SPIE. **1333** (1990b).

Gorshkov, V. A., "Interferometer for testing the shape of plane surfaces of large optical elements." *Sov. J. Opt. Technol.* **53** (3) 138-140 (1986).

Greivenkamp, J. E., Bruning, J. H., "Phase shifting interferometry." in Malacara (1992).

Grzanna, J., Schulz, G., "Absolute testing of flatness standards at square grid points." *Optics Communications* 77 (2,3) 107-112 (1990).

Grzanna, J., "Absolute testing of optical flats at points on a square grid: error propagation." *Applied Optics* **33** (28) 6654-6661 (1994).

Gubin, V. B., Sharonov, V. N., "Absolute calibration of spherical surfaces", *Soviet Journal of Optical Technology* **57** (9) 554-555 (1990).

Hariharan, P., Oreb, B. F., Eiju, T., "Digital Phase-Shifting Interferometry: A Simple Error-Compensating Phase Calculation Algorithm", *Applied Optics* **26** (3), 2504-2505 (1987).

Hariharan, P., "Optical flat surfaces: direct measurements of the spatial frequency spectrum of surface errors." *Optical Engineering* **35**, 484-485 (1996).

Harris, J. S., "The universal Fizeau interferometer." Ph.D. Thesis, University of Reading, U.K. (1971).

Howes, W. L., "Lens collimation and testing using a Twyman-Green interferometer with a self pumped phase conjugating mirror", *Applied Optics* **25**, 473-474 (1986).

Huang, C., "Evaluation of phase shift, misalignment, and propagation errors in Fizeau interferometry." *OSA Tec. Digest* **24** 87-89 (1992).

Huang, C., "Propagation errors in precision Fizeau interferometry." *Applied Optics* **32** (34) 7016-7021 (1993).

Ikeda, O., Suzuki, T., Sato, T., "High Accuracy Surface Profile Measuring System Using A BSO Phase Conjugating Mirror", *Applied Optics* **21** (24), 4468-4472 (1982).

Jensen, A. E., "Absolute calibration method for laser Twyman-Green wavefront testing interferometers", (abstract only), *Journal of the Optical Society of America* **63** 1313 (1973).

Jozwicki, R., "Influence of aberrations of Fizeau interferometer elements on measurement errors." *Applied Optics* **30** (22) 3126-3132.

Kadono, H., Takai, N., Asakura, T., "New common-path phase shifting interferometer using

a polarization technique." *Applied Optics* **26** (5) 898-904 (1987).

Ketelsen, D. A., Anderson, D. S., "Optical testing with large liquid flats." *Proc SPIE* **966** 365-371 (1988).

Kinnstaetter, K., Lohmann, A. W., Schwider, J., Streibl, N., "Accuracy of phase shifting interferometry." *Applied Optics* **27** (24) 5082-5089 (1988).

Kuchel, M., "Absolute measurement of flat mirrors in the Ritchey-Common Test." *OSA Technical Digest. Workshop on Optical Fabrication and Testing, University of Arizona, Ed. R. E. Parks.* 114-119 (1986).

Kuchel, M., "The new Zeiss interferometer." *Proc.SPIE* **1332** 655-663 (1990).

Linnik, W., "Simple interferometer to test optical systems." *Comptes Radnus de l'Academie des Sciences d l'U.R.SS.* **1** 208 (1933).

Lou, D. Y., Martinez, A., Stanton, D., "Surface profile measurement with a dual beam optical system." *Applied Optics* **23** (5) 746-751 (1984).

Makosch G., Drollinger, B., "Surface profile measurement with a scanning differential ac interferometer." *Applied Optics* **23** (24) 4544-4553 (1984).

Malacara, D., Cornejo, A., Murty, M. V. R. K., "Bibliography of various optical testing methods." *Applied Optics* **14** (5) 1065-1080 (1975).

Malacara, D., "A review of interferogram analysis methods." *Proc SPIE* **1332** 678-689 (1990a).

Malacara, D., Carpio-Valadez, J. M., Sanchez-Mondragon, J. J., "Wavefront fitting with discrete orthogonal polynomials in a unit radius circle." *Optical Engineering* **29** (6) 672-675

(1990b)

Malacara, D. (Ed.), *Optical Shop Testing* (2nd Edition), J. Wiley & Sons (1992).

Mantravadi, M.V., "Newton, Fizeau, and Haidinger interferometers." in Malacara (1992a).

Mantravadi, M. V., "Lateral Shearing Interferometers." in Malacara (1992b).

Marioge, J. P., Bonino, B., Mullot, M., "Standard of flatness: its application to Fabry-Perot interferometers." *Applied Optics* **14** (9) 2283-2285 (1975).

Massie, N. A., "Real-time digital heteriodyne interferometry: a system." *Applied Optics* **19** (1) 154-160 (1980).

Mercer, C. R., Creath, K., "Liquid-crystal point-diffraction interferometer for wave-front measurements." *Applied Optics* **35** (10) 1633-1642 (1996).

Offner, A., Malacara, D., "Null tests using compensators." in Malacara (1992).

Okaji, M., Imai, H., "A practical measurement system for the accurate dtermination of linear thermal expansion coefficients." *J. Phys. E: Sci. Instrum.* **17** 669-673 (1984).

Omar, B. A., Holloway, A. J., Emmony, D. C., "Differential phase quadrature surface profiling interferometer." *Applied Optics* **29** (31) 4715-4719 (1990).

Pantzer, D., Politch, J., Ek, L., "Heterodyne profiling instrument for the Angstrom region." *Applied Optics* **25** (22) 4168-4172 (1986).

Parks, R. E., "Removal of test optic errors." *Proc SPIE* **153** 56-63 (1978).

Polster H. D., "The determination of the absolute contours of optical flats. II." *Applied Optics*

7 (5) 977-978 (1968).

Primak, W., "The determination of the absolute contours of optical flats." *Applied Optics* **6** (11) 1917-1923 (1967).

Primak, W., "Optical flatness standard."*Optical Engineering* **23** (6) 806-815 (1984).

Primak, W., "Optical flatness standard II: reduction of interferograms."*Proc SPIE* **954** 375-381 (1988).

Primak, W., "Optical flatness standard: comment."*Optical Engineering* **28** (8) 934 (1989).

Rayleigh, Lord, "Interference bands and their application." *Nature* **48** 212-214 (1893).

Rimmer, M. P., "Method for evaluating lateral shearing interferograms." *Applied Optics* **13** (3) 623-629 (1974).

Ritchey, G.W., "On the modern reflecting telescope and the making and testing of optical mirrors". *Smithson. Contrib. Knowl.* **34** 3 (1904).

Roddier, C., Roddier, F., "Interferogram analysis using Fourier transform technique." *Applied Optics* **26** (9) 1668-1673 (1987).

Sasaki, O., Wang, X., Takebayashi, Y., Suzuki, T., "Sinusoidal phase-modulating Fizeau interferometer using phase conjugate wave." *Proc. SPIE.* **2003** 12-20 (1993).

Saunders, J. B., "Precise topography of optical surfaces." *J. Res. Nat. Bur. Stand.* **47** (3) 148-155 (1951).

Saunders, J. B., "Measurement of wavefronts without a reference standard I: The Wavefront Shearing Interferometer", *J. Res. Nat. Bur. Stand.* **65B**, 239 (1961).

Schulz, G., "Ein interferenzverfahren zur absoluten ebenheitsprufung langs beliebiger." *Optica Acta* **14** 375-388 (1967).

Schulz, G., Schwider, J.,"Precise measurement of planeness." *Applied Optics* **6** (6) 1077-1084 (1967).

Schulz, G., Schwider, J., Hiller, C., Kicker, B., "Establishing an optical flatness standard." *Applied Optics* **10** (4) 929-934 (1971).

Schulz, G., Schwider, J., "Interferometric testing of smooth surfaces", *Progress In Optics* **XIII** Ed. Wolf, E.. North Holland (1976).

Schulz, G., Schwider, J., "Comments on the paper "Optical flatness standard"." *Applied Optics* **26** (6) 559-560 (1987).

Schulz, G., Grzanna, J., "Absolute flatness testin by the rotation method with optimal measuring error compensation." *Applied Optics* **31** (19) 3767-3780 (1992).

Schulz, G., "Absolute flatness testing by an extended rotation method using two angles of rotation." *Applied Optics* **32** (7) 1055-1059 (1993).

Schwider, J., Elssner, K. E., Grzanna, J., Spolaczyk, R., "Results and error sources in absolute sphericity measurement", *IMEKO Symposium On Laser Applications In Precision Measurement*, Budapest, 93-103 (1986)

Schwider, J., "Fizeau- and Michelson-type interferograms and their relation to the absolute testing of optical surfaces." *Optik* **89** (9) 113-117 (1992).

Schwider, J., Falkenstorfer, O., "Twyman-Green interferometer for testing microspheres", *Optical Engineering* **34** (10), 2972-2975 (1995).

Shao, L., Parks, R. E., Ai, C., "Absolute testing of flats using four data sets." *Proc. SPIE* **1776** 94-97 (1992).

Shinn, A. L. "Re: Holding small balls ?" Private communication (1996).

Shu K. L. "Ray-trace analysis and data reduction methods for the Ritchey-Common test." *Applied Optics* **22** 1879 (1983).

Shukla, R. P., Dokhanian, M., Venkateswarlu, P., George, M.C., "Phase conjugate Twyman-Green interferometer for testing spherical surfaces and lenses and for measuring refractive indices of liquids or solid transparent materials", *Optics Communications* **78** (5,6), 407-415 (1990a).

Shukla, R. P., Dokhanian, M., Venkateswarlu, P., George, M.C., "Phase conjugate Twyman-Green interferometer for testing conicoidal surfaces." *Proc. SPIE.* **1332** 274-286 (1990b).

Simova, E., Zhivkova, S., Stoilov, G., Miteva, M., "Highly sensitive interferometric system for optical surface profiling." *Optics and Lasers in Engineering* **18** 53-59 (1993).

Slomba, A. F., Hull-Allen, C. G., Takacs, P. Z., Evans, C. J., Bennett, J. M., "Flatness intercomparison measurements made on an optical flat." *OSA tech. digest* **24** 1-3 (1992).

Smartt, R. N., Steel, W. H., "Theory and application of point diffraction interferometers." *Japan. J. appl. Phys.* **14** 351-356 (1975).

Sommargren, G.E., "Optical heterodyne profilometry." *Applied Optics* **20** (4) 610-618 (1981).

Speer, R. J., Chrisp, D., Turner, D., Mrowka, S., Tregidgo, "Grazing incidence interferometry: the use of the Linnik interferometer for testing image forming reflection systems." *Applied Optics* **18** (12) 2003-2012 (1979).

Stahl, H. P., Tome J. A., "Phase-measuring interferometry: performance characterization and calibration." *Proc. SPIE* **954** 78-87 (1988).

Stahl, H. P., "Aspheric surface testing techniques." *Proc. SPIE* **1332** 66-76 (1990a).

Stahl, H. P., "Review of phase measuring interferometry." *Proc. SPIE* **1332** 704-719 (1990b).

Stahl, H. P., "Absolute interferometric measurement: flats and windows." *Lasers and Optronics* 27 (November 1991).

Stevens, R. F., "Interferometry for the measurement of form of near-spherical surfaces", *NPL Report MOM(E)14* (1994).

Sun, C.-C., Tsou, R.-H., Shen, W., Chang, H.-H., Chang, J.-Y., Chang, M.-W., "Shearing interferometer with a Kitty self-pumped phase conjugator." *Applied Optics* **35** (11) 1815-1819 (1996).

Swantner, W., "Wavefront analysis on a personal computer." *Proc. SPIE.* **540** 85-88 (1985.

Takacs, P.Z., Qian, S.-N., Colbert, J., "Design of a long trace surface profiler." *Proc. SPIE.* **749** 59-64 (1987).

Takacs, P.Z., Qian, S.-N., "Surface profiling interferometer." *U.S. Patent 4884697 (1989).*

Truax, B. E. "Absolute interferometric testing of spherical surfaces", *Proc. SPIE* **966** 130-137 (1988).

Van Wingerden, J., Frankena, H. J., Smorenburg, C., "Linear approximation for measurement erors in phase shifting interferometry." *Applied Optics* **30** (19) 2718-2729 (1991).

Vikhaven, E., "Nondestructive testing by use of TV holography and deformation phase

gradient calculation." *Applied Optics* **29** (1) 137-144 (1990).

Virdee, M. S., "A new approach for establishing a reference plane for absolute mjeasurement of shape and flatness with nanometre precision." *NPL Report MOM 114* (1993).

Virdee, M. S., "Non-contacting straightness measurement to nanometre accuracy." *Int. J. Mach. Tools Manufact.* **35** (2) 157-164 (1995).

Wang, Q., Chen, J., Zhu, R., Chen, L., Zhang, Y., "A new technique for testing large optical flat." *Proc SPIE* **2003** 389-397 (1993).

Wang, X., Sasaki, O., Takebayashi, Y., Suzuki, T., Maruyama, T., "Sinusoidal phase modulating Fizeau interferometer using a self- pumped phase conjugator for surface profile measurements", *Optical Engineering* **33** (8), 2670-2674 (1994).

Williams, R. A., Kwon, O. Y., "Three dimensional interferometric testing of optical spheres", *Journal of the Optical Society of America, A.* **4** (10) 1855-1860 (1987).

Womak, K.,"Interferometric phase measurement using spatial synchronous detection." *Optical Engineering* **23** 391-395 (1984).

Wu, S., Xu, C., Wang, Z., "New application of point diffraction interferometer- polarization fringe scanning PDI." *ICO-13 Conference Digest- Optics in Modern Science and Technology* 458-459 (1984).

Wyant, J. C., Creath, K., "Recent advances in interferometric optical testing." *Laser Focus/ Electro-optics* 118-132 (1985).

Wyant, J.C., Creath, K., "Basic wavefront aberration theory for optical metrology." *Applied Optics and Optical Engineering* **XI.** ed. Shannon, R. R., Wyant, J. C. Academic Press. 1-53 (1992).

Yang, G., Siahmakoun, A., Khorana, B.M., "Tunable self-reference phase conjugate interferometer." *Applied Optics* **30** (19) 2714-2717 (1991).

Yang, G., Siahmakoun, A., "Time-delay Fizeau phase conjugate interferometer." *Applied Optics* **32** (19) 1578-1582 (1993).

Yeh, P. *Introduction To Photorefractive Non Linear Optics*, J. Wiley & Sons (1993).

Zernike, F., "Diffraction theory of knife-edge test and its improved form, the phase contrast." *Mon. Not. R. Astron. Soc.* **94** 371 (1934).

**Acknowledgements.**

There are many people without whose help and encouragement this research project would have been far more difficult, less enjoyable or impossible. To remember and name every one would be a task as formidable as any I have described in the pages of this thesis. I would like to extend my sincere thanks to the following individuals for their particular support:

At Kingston University;

Professor David Briers, my supervisor, for putting this project my way, understanding and patience.

Martin Abbott and Paul Haycock, of the Faculty of Science workshop, for their skilled craftsmanship.

The postgraduate research community for making life at Kingston a stimulating and fun experience.

At the National Physical Laboratory;

Dr Keith Birch, for originating this project and many useful discussions.

Dave Putland, for tolerating my intrusions into his working environment.

Dr Richard Stevens and Professor Mike Hutley for extending my horizons beyond flatness.

My Mother, Father and Sister without whom I would probably never have finished this.

Gaenor too!