A novel heterodyne interferometer for scanning optical microscopy

by

M. J. Offside

Thesis submitted for
The Degree of Doctor of Philosophy
of the
University of London

Department of Electronic and Electrical Engineering
University College London

November 1990

Abstract

A phase sensitive scanning optical microscope is described which can measure surface height changes of about 1 Å. The system is based on a heterodyne version of the Michelson interferometer, and has been designed to reject phase noise caused by vibration in the optics and the sample. A specially constructed objective lens is used to direct two laser beams onto the object surface. The first beam forms a tightly focused spot to probe the sample structure and the second remains collimated, acting as a large area on sample reference beam. In the simplest implementation, the objective may be fabricated by drilling a hole in a lens singlet.

The configuration allows the relative areas illuminated by the two beams to be varied both arbitrarily and independently, thus guaranteeing an accurate absolute phase measurement. This is an important advantage over existing techniques, in which the range of suitable samples is restricted by the limited size of the on sample reference beam. The two beams reflected from the sample are interfered with a third frequency shifted beam, so forming two heterodyne Michelson interferometers in parallel. The light from each interferometer is detected separately, resulting in two AC signals. The phase of these signals are then compared to provide the object surface phase structure. Path length fluctuations due to microphonics are common to both interferometers and are cancelled by this comparison.

Results from a bench top version of the system are presented which demonstrate the principle of the technique and a detailed study is made of the factors limiting the sensitivity of the phase measurement. The conclusions of this study have been applied to the design of a prototype microscope and this has been used to record micrographs of a number of representative samples. In addition the particular imaging characteristics of the system are discussed using a combination of geometrical optics and a transfer function approach.

Acknowledgements

I would like to express my gratitude to Dr. M.Somekh, my supervisor, for his help and guidance over the past three years. His influence will undoubtedly have a lasting effect on my approach to research in the future. Without his enviable ability to obtain financial backing, much of the work described in this volume would not have been possible.

I also wish to thank Dr. C.W.See, who has made significant contributions to the idea behind the microscope. During the first eighteen months of my PhD, he passed on a great deal of invaluable information gained during his own work in optical microscopy. I am also indebted to Prof. C.W.Pitt for his help during my final year, after Dr. Somekh joined the University of Nottingham.

In addition, I would like to thank Dr. R.K.Appel and Dr. S.Valera, with whom I have had many productive technical discussions and Eric Ngau, John Lyle and Freddie Chin for their help in the development of the software for prototype of the microscope.

I would like to acknowledge the Science and Engineering Research Council and Unilever Plc. for their financial support. In particular, I would like to thank my CASE supervisor, Dr. R.Miller of Unilever Research, for making available the extensive facilities at Port Sunlight Laboratory.

To my mother and father,
Sarnia and Lainya

and to Punam

Contents

List of figures

CHAPTER 1

INTRODUCTION

1. Introduction

Since its invention nearly four hundred years ago, the optical microscope has been instrumental in the development of such fields as biology, medicine, metallurgy and crystallography. During most of this period, workers have endeavoured to improve the quality of the optics, since lens aberrations seriously limited the imaging performance of the early full field microscopes. Without exception, these systems were intensity instruments, in that they were only capable of revealing the reflection or absorption properties of the material.[1] For many applications, such as imaging transparent biological tissue or viewing objects with very fine surface topography, this proved a major drawback. These structures primarily affect the phase of the illuminating beam, imposing little or no amplitude modulation on it. Since the human eye (or any optical detector for that matter) is insensitive to this phase modulation, these so called weak phase objects remain largely invisible. A technique was therefore required to convert the phase modulation into an intensity modulation.

A number of instruments were developed with this idea in mind, involving modification of the diffracted spectrum of the light beam. Some notable examples[2] are the central dark ground microscope, in which a small stop is used to block the central order beam; the Schlieren method in which all spectral components one side of the central order are removed and the oblique dark ground method where the central order is excluded also. The most powerful of such techniques is probably Zernike's phase contrast method,[3] in which the central order is phase shifted by $\pi/2$ with respect to the diffracted spectrum allowing

[1] In monochromatic light intensity modulation in the image results from absorption by the object, whereas for non-monochromatic light selective spectral absorption may alter its colour.

[2] A.H.Bennett, H.Jupnik, H.Osterberg and O.W.Richards, "*Phase Microscopy*", J.Wiley, New York (1951).

[3] F.Zernike, "Phase contrast, a new method for the microscopic observation of transparent objects" Physica, **9**, pp 686-698 and pp 974-986, (1942).

interference to take place. This method has the advantage that the resultant intensity variations bear a linear relationship with the original phase changes produced by the object. However, satisfactory phase contrast is only possible with small objects, and in addition the contours may be confused by the halo effect.

These limitations probably contributed to an increased interest in what may be termed conventional interference microscopy, whereby the light from the object is generally interfered with a reference beam of constant wavefront. Any spatial phase modulation imposed on the object beam then becomes visible as spatial intensity modulation on an interferogram. Such techniques date back to the last decade of the nineteenth century, but appear to have been slow in gaining wide acceptance, until Zernike's method highlighted the advantages of phase microscopy.

By the early 1960's about one hundred different microscopes had been developed which incorporated some form of interferometer. The principles underlying these configurations vary greatly and attempts to classify them are inevitably cumbersome.[4] In view of this diversity, it is difficult to find a concise, general definition for the interference microscope. In essence, the incident light is split into two or more coherent beams, some or all of which come into contact with the object. These beams are then subsequently recombined and interference takes place. The methods of dividing the incident beam include the use of transparent surfaces, birefringent elements, diffraction components or diaphragms. Recombination of the beams may achieved by the same component which does the division for a reflection system, or by using a second similar component for transmission systems.

Some notable examples include the Linnik system, which is based on the well known Michelson interferometer and the arrangements of Mirau

[4] W.Krug, J.Rienitz, and G.Schulz, "*Contributions to Interference Microscopy*", Hilger & Watts, London (1964).

and Dyson in which the objective has its own built in reference mirror and semi-transparent layer, both at 90^0 to the optical axis. Nomarski's differential interference contrast microscope is probably the most famous example of a system employing a birefringent beam divider, in this case a Wollaston prism. The relative merits of these and many other instruments are considered in two volumes published in the early 1960's.[4,5] Since then, a number of important advances have been made in optical microscopy, of which the introduction of the laser as the source of illumination and the use of scanning are probably the most significant.

The scanning optical microscope may be traced as far back as 1951 to the "flying spot" system of Young and Roberts.[6,7] Although the method used to perform the scanning (using a cathode ray tube) has been virtually abandoned in microscopy, some of the potential advantages of scanning were suggested. The first laser scanning interference microscopes began to appear in the late 60's shortly after the commercial availability of the laser. The most sensitive of these were the heterodyne techniques, in which the phase information is encoded on a carrier frequency. This, together with use of scanning made them particularly amenable to subsequent electronic signal processing, allowing data to be acquired in serial form. As a result these are among the first optical microscopes capable of recording truly quantitative information. Although such systems were initially developed to measure the amplitude and phase of surface acoustic waves,[8,9,10]

---

[5] M.Francon, "*Progress in microscopy*", Pergamon Press, London (1961).

[6] J.Z.Young and F.Roberts, "A flying-spot microscope", Nature, **167**, pp. 231 (1951).

[7] F.Roberts and J.Z.Young, "The flying-spot microscope", Proc. IEE, **99**, IIIA, pp. 747-757 (1952).

[8] R.L.Whitman, L.J.Laub and W.J.Bates, "Acoustic surface displacement measurements on a wedge-shaped transducer using an optical probe technique", IEEE Trans., **SU-15**, pp. 186-189 (1968).

[9] R.L.Whitman and A.Korpel, "Probing of acoustic surface perturbations by coherent light", Appl. Optics, **8**, pp. 1567-1576 (1969).

[10] R.M.De La Rue, R.F.Humphryes, I.M.Mason and E.A.Ash, "Acoustic-surface-wave amplitude and phase measurements using laser probes", Proc. IEE, **119** (2), pp. 117-126 (1972).

their application to general profilometry and microscopy soon followed.[11,12,13] However, they no longer achieved the same sensitivity due to the deleterious effect of microphonics. These are small vibrations in the optical components, which introduce path length fluctuations into the interferometer. This is not such a severe problem when measuring surface acoustic waves, since they typically have frequencies in the MHz range, whereas the spectrum of microphonics tends to be well below 100 kHz. Furthermore, coherent detection methods are usually employed to extract the signal at the acoustic frequency, and so fluctuations which fall inside the measurement bandwidth are negligible.

One way to overcome the problem of microphonics in a general purpose microscope is to use a common path interferometer. The interferometer is designed so that as far as possible, the two beams follow similar paths in the optical system. Microphonics then affect both beams equally and the phase fluctuations will be largely eliminated from the detected signal. Clearly, such common path systems require that both interfering beams illuminate the sample. The various systems that achieve common path operation can be divided into two main categories.

The first are the differential systems, so called because they only respond to changes in surface phase structure as the sample is scanned. Typically, two beams are focused to closely spaced adjacent points on the sample surface and are recombined at a subsequent interference plane. The paths followed by each beam are therefore almost identical and thus the interferometer is extremely common path.

In many applications however, the true profile is required rather than

---

[11] L.J.Laub, "Apparatus and methods for scanning phase profilometry", U.S. Patent, Nº 3,796,495 (1974).

[12] G.E.Sommargren and B.J.Thompson, "Linear Phase Microscopy", Appl. Optics, **12** (9), pp. 2130-2138 (1973).

[13] T.Sawatari, "Optical heterodyne scanning microscope", Appl. Optics, **12** (11), pp. 2768-2772 (1973).

its differential. This is the domain of the second type of system, which may be termed absolute phase measuring interferometers. The design of common path absolute interferometers is not as straightforward as the differential systems. One of the beams to be interfered is the focused optical probe as before, whereas the second must maintain a constant phase as the sample is scanned, in order to provide a reference beam. This is the important difference between the two methods, since the beams in the differential systems are indistinguishable in terms of a probe and a reference beam.

One approach by Sommargren[14] involved modifying the scan pattern of what is essentially a differential system of the type discussed above. The sample is circularly scanned with one beam focused on the centre of rotation. This beam therefore remains constant while the other interrogates the sample in a circular fashion. Although the system gives the same sensitivity as the differential systems, the nature of the scanning is rather inconvenient and requires accurate alignment of the reference beam at the centre of the scan.

Another way to keep the phase of the reference beam fixed is to somehow make it illuminate a large area of the sample. Effectively, the reference beam averages the surface phase structure over this large area and thus its phase remains constant as the sample is scanned. This procedure does not impose any constraints on the scanning, allowing the more conventional raster scan or line trace to be employed. Existing systems which use this method, such as those by Downs[15] and Huang[16] have been designed for measurement of surface roughness of very smooth surfaces. For samples exhibiting a greater variation in phase structure, a larger reference area on the sample is required to provide a constant

---

[14] G.E.Sommargren, "Optical heterodyne profilometry", Appl. Optics, **20** (4), pp. 610-618 (1981).

[15] M.J.Downs, W.H.McGivern and H.J.Ferguson, "Optical system for measuring the profiles of super smooth surfaces", Precision Eng., **7** (4), pp. 211-215 (1985).

[16] C.C.Huang, "Optical heterodyne profilometer", Optical Engineering, **23** (4), pp. 365-370 (1981).

reference phase. The size of the reference area in both of the above systems is limited, since any increase relative to the area illuminated by the probe beam is accompanied by a reduction in interference efficiency.

The system which is the subject of this thesis has been designed to overcome this limitation. This is achieved by having two *interferometers* which follow similar paths, instead of two interfering beams which follow similar paths. Each interferometer has excellent interference efficiency and give signals which if treated separately suffer from microphonics. However, the phase fluctuations are common to both interferometers, and can be removed by subtracting the phase of the two output signals. One of the interferometers has a collimated sample illuminating beam, to provide a constant reference signal, whereas the other has a tightly focused beam to probe the sample surface. Both interferometers interact with the same optical components, including the sample and are therefore affected almost identically by microphonics. In addition, the interferometers are heterodyne, so the phase subtraction is conveniently performed in the coherent detection process. Strictly then, the system is not a common path interferometer, since the effects of microphonics are removed electronically rather than optically. However, as demonstrated in subsequent chapters, the system does give common path performance, without the associated limitations of existing absolute phase measuring systems.

The arrangement of this thesis is as follows: in the next chapter a review is made of the current state of interference microscopy. The instruments are classified into full field and scanning systems, and these are further subdivided into homodyne and heterodyne techniques. This review demonstrates a clear need for a system capable of making accurate absolute phase measurements. A basic requirement of such a system is that it be insensitive to microphonics.

Chapter 3 describes the optical configuration of the system and theory

describing the insensitivity to microphonics is presented. Experimental results from the first embodiment of the system demonstrate its important attributes, namely the insensitivity to vibration and the quality of the absolute phase measurement. Two appendices following the chapter provide useful practical information concerning the alignment of the optics and optimisation of the signals.

The various limitations which affect the sensitivity of the phase measurement are examined in Chapter 4. More emphasis is given to those factors which are specific to this particular system, with a view to improving its performance.

Chapter 5 is devoted to a prototype version of the microscope. Two dimensional scanning is employed in this system and the images so acquired are compared with those obtained by a number of other techniques.

Chapter 6 is concerned with the image formation theory of the microscope. The two interferometers comprising the system are treated separately. A straightforward ray theory approach is shown to be valid for the reference interferometer and a more complex diffraction theory is applied to the probe interferometer. The results from the two methods are then combined to determine the overall response of the system to various representative sample features.

The final chapter provides an overall assessment of the microscope and includes a number of suggestions for improvements. In addition, a new technique incorporating electronic beam scanning is proposed and its merits and limitations compared to existing systems are discussed.

CHAPTER 2

REVIEW OF INTERFEROMETRIC MICROMETROLOGY

2. Review of interferometric optical micrometrology

The field of optical micrometrology is too diverse to be covered adequately in one chapter. For this reason the scope of this review has been restricted to techniques based on interferometry. These systems have repeatedly demonstrated the greater sensitivity of interferometry compared to other methods. There are, however, a number elegant optical techniques which do not make use of interferometry and these are summarised briefly below.

One important group of instruments rely solely on deviations of the illuminating beam by the sample according to ray optics. Consequently, these arrangements are often described as geometrical systems. Most commonly, the systems detect the focusing error of the objective lens as they scan the surface under test, a method originally employed in the optical pickups for digital audio disks. The various techniques for focus detection include the astigmatic method,[17] the Foucault method[18] and the critical angle method.[19] Since such methods are insensitive to the phase changes imposed by the sample, they do not present any ambiguity in interpreting the results. This is a problem encountered in systems which do measure phase, since the phase change could be due to topography or reflectivity variations. In addition, phase techniques have a limited dynamic range caused by phase wrap around. Again, this is not a problem with geometrical techniques, and so the the measurement range is potentially far greater. The main disadvantage of such instruments is the sensitivity to microphonics. For this reason, these systems do not give the sensitivities that common path interferometric systems provide.

---

[17] D.Y.Lou, A.Martinez and D.Stanton, "Surface profile measurement with a dual-beam optical system", Appl. Optics **23** (5), pp. 746-751 (1984).

[18] G.Bouwhuis and J.J.M.Braat, "Video disk player optics", Appl. Optics **17** (13), pp. 1993-2000 (1978).

[19] T.Kohno, N.Ozawa, K.Miyamoto and T.Musha, "High precision optical surface sensor", Appl. Optics **27** (1), pp. 103-108 (1988).

The confocal microscope[20] is another instrument with some extremely useful characteristics. Basically, the sample is illuminated with a focused beam from a spatially coherent source, and the reflected light is collected by a point detector.[21] This arrangement strongly rejects light that illuminates out of focus sample structure, and so the system has very good depth discrimination. This optical sectioning property has proved invaluable in biological applications. Interferometers also demonstrate good depth discrimination, and in fact, an interferometric system[22] has recently been demonstrated which can image through a scattering medium in a similar manner to the confocal system.

The interferometric techniques considered in this review have been divided into two categories. Section 2.1 deals with scanning systems and section 2.2 discusses full field instruments. These sections have each been subdivided into homodyne and heterodyne, depending on whether a relative optical frequency shift has been introduced between the interfering beams.

2.1 Scanning techniques

There are many advantages to be gained in microscopy by scanning the sample and building up the image point by point. The two most common methods are: (i) mechanically scanning the object under the focused beam, and (ii) optical scanning using galvanometer mirrors, acousto-optic beam deflectors, rotating multifaceted mirrors or by mechanically scanning the objective lens. Object scanning is easier to implement and has the advantage that off axis aberrations in the lenses are not a

[20] See for example: T.Wilson and C.R.Sheppard, "*Theory and practice of scanning optical microscopy*", Academis Press, London, 1984.

[21] In practice the detector is finite and a criterion for detector size to give true confocal operation has been determined by: T.Wilson and A.R.Carlini, "Size of the detector in confocal imaging sytems", Optics Lett., **12** (4), pp.227-229 (1987).

[22] M.Toida, M.Kondo, T.Ichimura and H.Inaba, "Experimental verification of image detection in highly scattering media using antenna properties of optical heterodyne microscope scheme", Electron. Lett., **26** (11), pp. 700-702 (1990).

serious problem, since all measurement points are affected equally. The major drawback is the slow scan speed unless the scan area is small. Optical scanning does offer faster image acquisition times, but requires optics of the quality used in full field systems or better. The use of scanning in some forms of microscopy is essential, because the nature of the contrast mechanism requires a point by point measurement. This is the case with techniques such as photothermal, optical beam induced current (OBIC), and scanning tunneling microscopy. In optical microscopy, scanning provides freedom to perform sensitive AC detection and signal processing schemes. This is particularly important in interferometric systems, where the optical phase may be imposed on a sinusoidal carrier frequency by the method of heterodyning. In addition, various modulations may be imposed on the optical beams such that signals extracted at different frequencies can have different optical transfer functions[23], or carry additional information as in the combined differential intensity and phase system discussed in Section 2.1.2. At present such systems have no full field counterparts, due to the complexity of performing the detection and signal processing for each measurement point in parallel.

2.1.1 Scanning techniques - homodyne

One notable instrument that may be classified under this heading is the polarisation interferometer of Downs et al.,[24,25] which is illustrated in Fig. 2.1. Referring to the diagram, the light from the laser is expanded and passed through a half wave plate, polariser and birefringent lens. The half wave plate controls the total light intensity entering the interferometer through the polariser. The birefringent lens is used in conjunction with a conventional microscope objective to provide a high power objective with different focal lengths for the two orthogonal

[23] M.G.Somekh, "Image formation in scanned heterodyne microscope systems", J. Microscopy, to be published (1990).

[24] M.J.Downs, "Surface profile interferometer", United Kingdom patent Nº 2109545B, filed 14th October 1982.

[25] M.J.Downs, W.H.McGivern and H.J.Ferguson, "Optical system for measuring the profiles of super-smooth surfaces", Precision Eng., **7** (4), pp. 211-215 (1985).

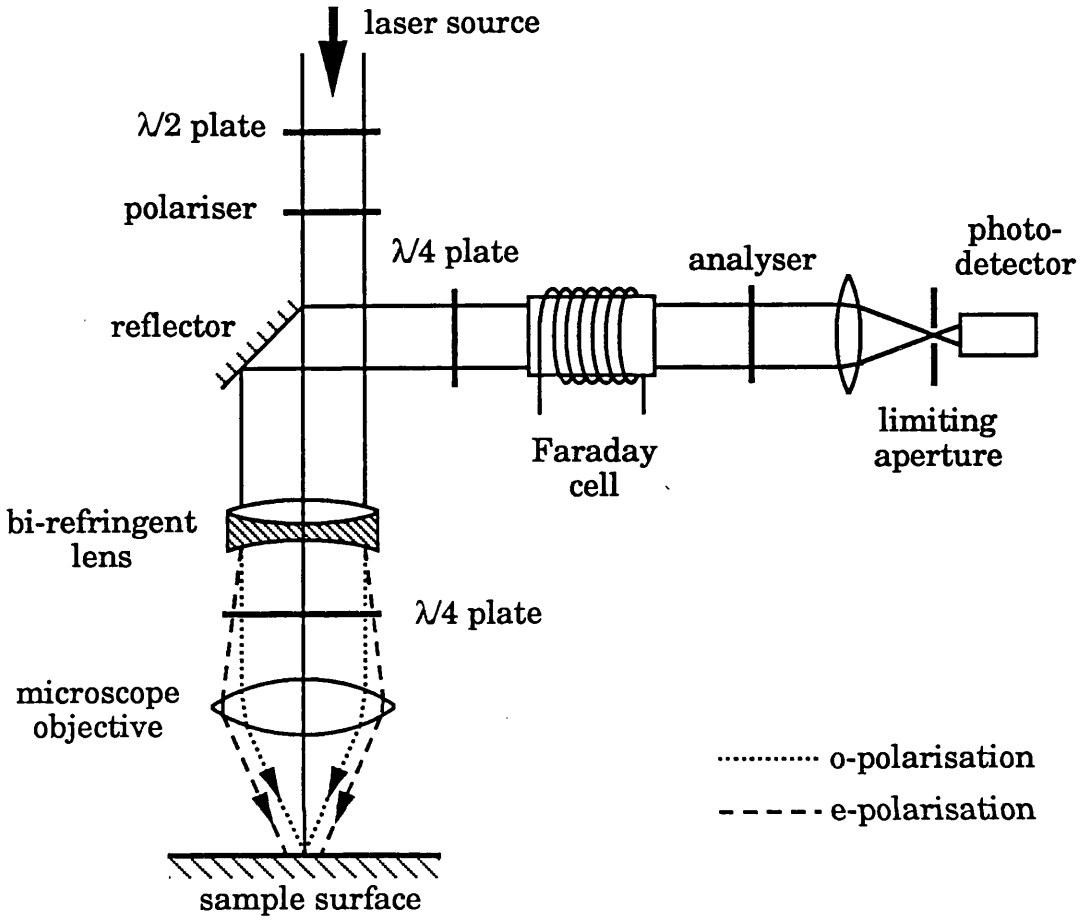


Figure 2.1. The Downs polarisation interferometer.

polarisations of the beam. The polarisation of shorter focal length is focused onto the sample so that the other is out of focus and thus illuminates a larger area of the surface. The idea is that this beam will average the surface structure over the illuminated region in order to provide an on sample reference beam. This objective may be thought of as an axial version of the Nomarski differential interference objective referred to in Chapter 1. This uses a birefringent element such as a Wollaston prism together with the objective to focus the orthogonal polarisations to adjacent points on the sample. A quarter wave plate placed between the birefringent lens and the microscope objective exchanges the polarisation directions of the two components after reflection. In this way, the birefringent lens approximately corrects for the mismatch in the two wavefronts, in order to improve the interference efficiency.

The detection scheme is somewhat unusual in microscopy, in that it is based on polarimetry.[26] The beam returning through the birefringent lens is first directed to a quarter wave plate orientated at 45° to the orthogonal components of the beam. The beam is elliptically polarised at this point due to the phase difference imposed on the orthogonal components by the sample. The quarter wave plate therefore produces a linearly polarised beam whose azimuth is dependent on this phase difference. The polarimeter is then used to detect variations in the plane of vibration of the linearly polarised beam as the sample is scanned, thus providing a measure of the surface phase structure.

The interferometer is almost entirely common path,[27] and thus gives good rejection of phase errors caused by both piston type vibration in the optics and thermal effects. The sensitivity of the system is 0.03 Å measured in a 200 Hz bandwidth. In addition, the instrument is

[26] R.J.King and K.W.Raine, "Polarimetry applied to alignment and angle measurement", Optical Engineering, **20** (1), pp. 039-043 (1981).

[27] Such systems can never be 100% common path since they would be insensitive to the sample structure as well as microphonics.

relatively insensitive to sample tilt, since both beams are affected almost equally.

The system has been designed primarily for measuring the profiles of super-smooth surfaces, such as laser gyro mirrors. For samples which exhibit a greater variation in surface topography however, the area illuminated by the reference beam needs to be larger to maintain a truly constant reference phase as the sample is scanned. This is particularly important when measuring step heights, where, if the reference area is insufficient, the measured heights will be systematically smaller than the true values. Typically, the ratio of the probe and reference spot diameters on the sample is about 1:10, giving an area difference of 100.

Problems arise if this ratio is increased significantly, because the wavefront of the returning out of focus reference beam would no longer match that of the focused probe beam, giving a reduction in interference efficiency. The correction of the wavefront mismatch by exchanging the polarisation direction of the beams before they return back through the birefringent lens gives two slightly divergent beams. If the focal lengths presented by the birefringent lens are f and $f+\Delta f$, then the effective source distances U_p and U_r for the returning probe and reference beams are given by

$$U_p = -\left(\frac{f^2}{\Delta f} + f\right) \tag{2.1}$$

$$U_r = -\left(\frac{f^2}{\Delta f} - f\right) \tag{2.2}$$

where the negative signs indicate that the beams are divergent. If Δf is small compared to f then the first term in each expression dominates and therefore the two wavefronts will be well matched. As Δf is increased in order to make the reference area larger, the second terms become more significant and this results in a mismatch in the wavefronts.

However, it may be possible to correct for this mismatch by introducing a second birefringent lens of appropriate strength. The increase in the size of the reference area would inevitably make the system less common path, but this would probably be a worthwhile trade off to make the measurement more quantitative for a broader range of samples. As with all polarisation interferometers, an ambiguity exists in the interpretation of results from birefringent surfaces.

The second scanning homodyne system of Adachi et al.[28] is interesting, because it manages to reject phase noise caused by vibration, despite the fact that it employs an interferometer which is not common path. Fig. 2.2 shows the optical configuration. The light source is an interference filtered high pressure mercury lamp. This is directed onto the sample through a Mirau interferometer as introduced in Chapter 1. The light reflected by the sample proceeds to the detector head which is positioned at the image plane. Here it interferes with the light reflected from the reference surface of the Mirau interferometer. The detector head contains four separate detectors, A, B, C, D. The light from the particular measurement point on the sample passes through a pinhole placed at the image plane and the interference signal is detected by photodetector D. The lateral resolution of the instrument is therefore determined largely by the size of this pinhole. The remaining detectors surround D and have large areas, so that their interference signals carry spatially averaged information corresponding to large areas of the sample. These therefore act as reference signals.

The phase of the interference term from each of the four detectors is extracted by a method known as phase shifting. Since this method is more usually applied to full field homodyne systems, the essentials of such techniques are discussed in Section 2.2.1. The phase output obtained from the three reference detectors are averaged and then subtracted from the phase of the signal detector D. In this way the

---

[28] M.Adachi, H.Miki, Y.Nakai and I.Kawaguchi, "Optical precision profilometer using the differential method", Optics Letters, 12 (10), pp.792-794 (1987).

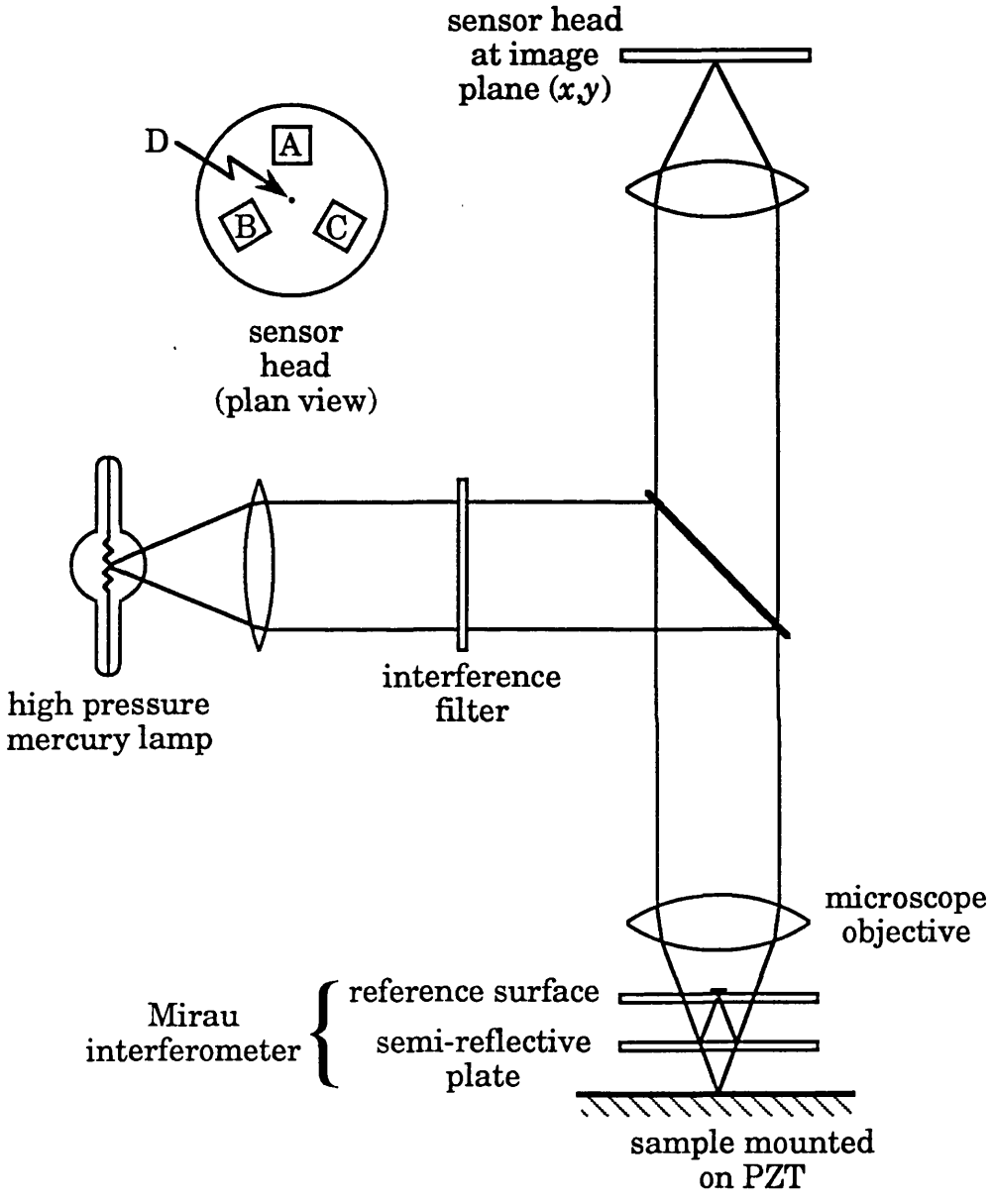


Figure 2.2. The homodyne optical interferometer of Adachi et al.

phase noise caused by piston type vibration of the optics is removed from the signal, thus improving the sensitivity. The sample is then scanned relative to the beam and a profile or image is built up. Since the reference detectors average the light from a large area of the sample, the phase from each remains constant as the sample is scanned. Three reference detectors are used to provide an accurate means of correcting for sample tilt.

The system does not have the high signal to noise ratio associated with most scanning systems, since the probe signal is obtained from a pinhole sized region of a full field interferogram. It differs from the Downs interferometer in that the phase noise caused by vibration is present in the detected signal and is then removed electronically. The system does not reject phase noise caused by air turbulence however, since this requires the beams to overlap spatially. The stability of the phase output corresponds to a surface height of 0.5 Å, indicating that this is not a severe limitation.

2.1.2 Scanning techniques - heterodyne

Optical heterodyne interferometry is a technique whereby the two beams in the interferometer have a relative frequency difference. Methods of introducing the frequency shift include acousto- or electro-optic modulators, or the use of a Zeeman split laser. Interference between two such beams generates a moving fringe pattern, which, when detected results in an AC electrical signal at the difference in the optical frequency. This signal acts as a carrier for the optical phase information. A heterodyne microscope may therefore be thought of as a transducer which converts optical phase to electrical phase. The heterodyne interferometers discussed in this section have been selected because they demonstrate at least some degree of insensitivity to microphonics and/or turbulence in the propagating medium. For most applications these tend to be the factors limiting the measurement.

The systems in this category are presented in chronological order, and the first of these is the optical heterodyne profilometer of Sommargren.[29] The optical configuration of this instrument is shown in Fig. 2.3. A Zeeman split HeNe laser with frequency stabilisation is used as the light source, and emits two collinear, orthogonally linearly polarised beams with a frequency difference of ω_Z, the Zeeman frequency. Part of this beam is tapped off with a beamsplitter, and passes to a polariser at 45^0 to each polarisation, permitting them to interfere at the photodetector. This provides an optically generated reference signal at the Zeeman frequency.

The remainder of the original beams are transmitted to a Nomarski objective, comprising a Wollaston prism and a conventional microscope objective. This focuses the orthogonally polarised beams to two distinct points on the object surface. One of these beams is aligned on the axis of rotation of a rotary table. The reflected beams are recombined by the Wollaston prism and pass to a second polariser and photodetector arrangement. The detected signal is proportional to the intensity I, which is given by

$$I = I_+ + I_- + 2\sqrt{I_+ I_-}\cos(\omega_Z t + \delta) \qquad (2.3)$$

Where I_+ and I_- are the intensities of the frequency upshifted and downshifted Zeeman components respectively. The last term is due to interference and contains the phase information δ, which is dependent on the phase difference between the two points illuminated on the sample. It is extracted using the optically generated reference signal by coherent detection at the Zeeman frequency. By scanning the sample in a circular fashion, a profile is obtained along a circle of radius equal to the separation of the two focused spots. The rms stability and reproducibility of the phase measurement are both quoted as <1 Å in terms of surface topography.

---

[29] G.E.Sommargren, "Optical heterodyne profilometry", Appl. Optics, **20** (4), pp. 610-618 (1981).

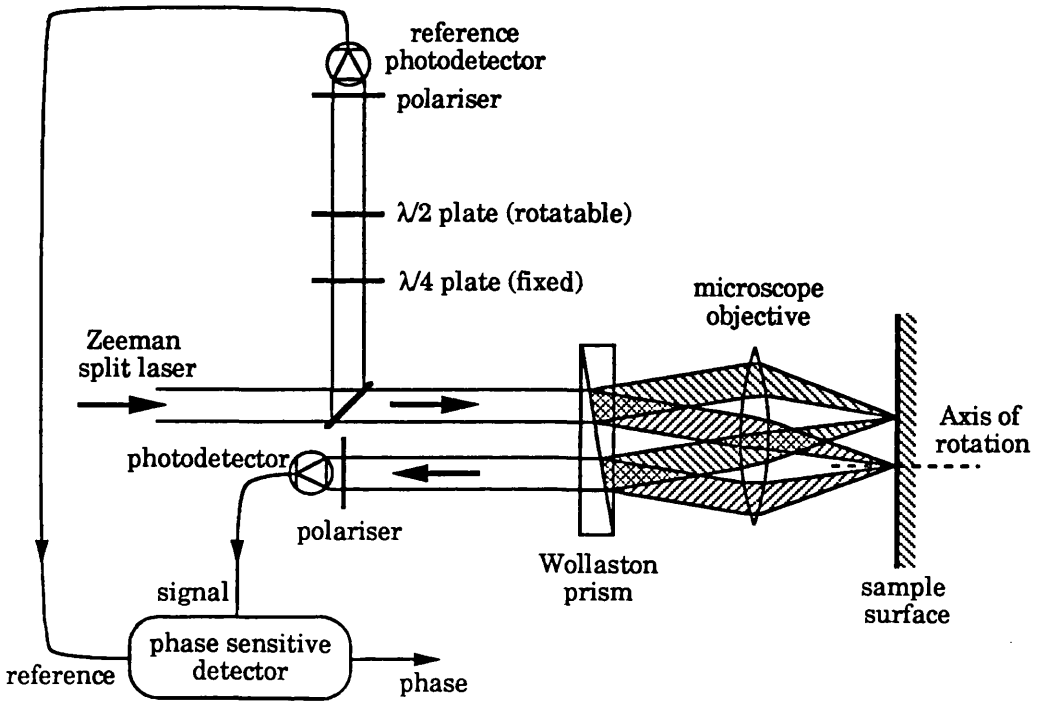


Figure 2.3. The Sommargren heterodyne interferometer.

As for the Downs interferometer, this instrument has been developed for the measurement of surface roughness of super-smooth surfaces, and so the circular scan scheme is acceptable for such an application. For general phase imaging however, a line scan or raster type scan is normally required. This system would provide a differential phase measurement if the sample were scanned in such a manner. An absolute phase measurement is therefore only obtained if one of the focused beams is aligned accurately on the axis of rotation, so that it maintains a constant reference phase.

The second technique does not impose any constraints on the scanning, and uses an acousto-optic device to provide the heterodyne frequency. A diagram of the system, designed by Huang[30] is illustrated in Fig. 2.4. The light from a linearly polarised HeNe laser is directed to an acousto-optic modulator (in this case a Bragg cell) The zeroth order or straight-through beam passes through a beam expander and fills the aperture of the objective lens. This beam is therefore focused to a small spot and forms the object probe beam. The first order diffracted beam from the Bragg cell is shifted in frequency by an amount equal to the frequency of its electrical drive signal ω_B. It is transmitted by a polarising beamsplitter (PBS), passes through a beam reducer and is then combined telecentrically with the object probe beam so that it too is focused onto the sample. Since this beam has a such a small diameter compared to the probe beam, the effective numerical aperture is smaller, so that it focuses to a larger spot than the probe beam, thus acting as a large area, on sample reference beam. The ratio of the areas illuminated by these two beams is therefore equal to the ratio of their diameters before the objective lens. The collimated laser beams returning from the objective lens have their polarisation direction shifted by 90^0 after the double passage through the quarter wave plate and are partially reflected and partially transmitted by the beamsplitter. The reflected beams pass through the beam reducer and are reflected at

[30] C.C.Huang, "Optical heterodyne profilometer", Optical Engineering, **23** (4), pp. 365-370 (1981).

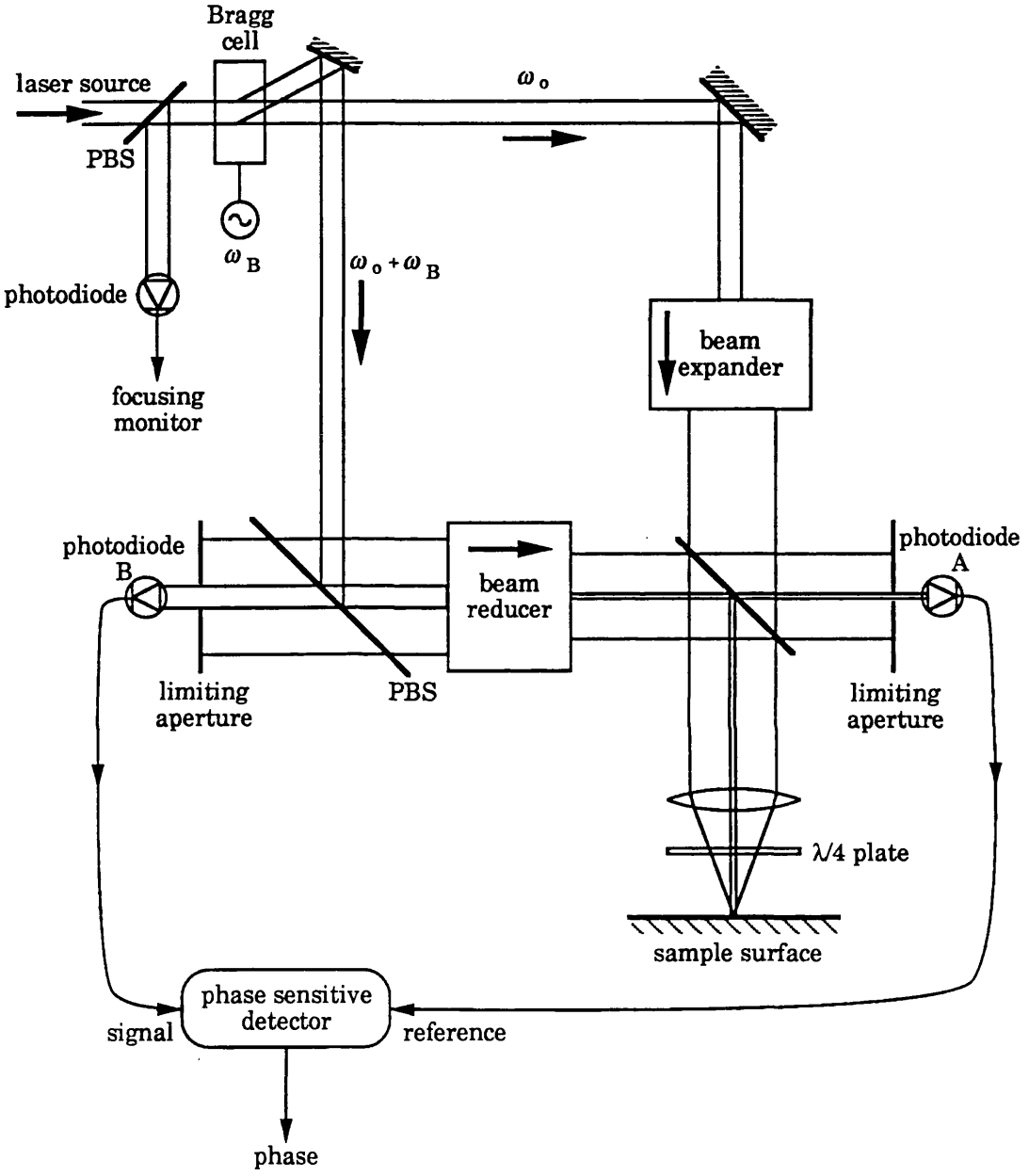


Figure 2.4. The Huang heterodyne interferometer.

the polarising beamsplitter to photodiode B where they interfere to provide the probe signal. This signal is of the same form as Eqn. (2.3), but now the interference term is at the Bragg frequency ω_B. The transmitted beams pass all the way back through the Bragg cell to a focusing monitor. This allows the position of best focus to be located by maximising the detected intensity. Only part of the outward beams reach the sample, the rest is directed by the non-polarising beamsplitter to photodiode A. The interference between these two beams provides an optically generated reference signal and is used to extract the phase in the same manner as the Sommargren system.

The rejection of phase errors due to piston type vibration and thermal drift is achieved by the combination of two features. Firstly, the optically generated reference signal will give a certain degree of phase noise which is common to the probe signal wherever the beams have followed the same path. There are however, regions where the beams reaching photodiodes A and B have not followed similar paths. This occurs from the non polarising beamsplitter onwards, i.e. the beams that interfere to form the reference signal propagate to photodiode A and the beams that interfere to form the probe signal propagate to the sample and back to photodiode B. Throughout these regions, the two beams that interfere at each photodiode have followed the same path so any phase noise introduced is removed when they interfere. This combination of electronic common mode rejection and common optical paths gives near perfect elimination of the optical phase error due to piston type vibration.[31] Thermal drift and air turbulence may at first sight appear to be a problem since the expanded and reduced beams do not completely overlap. However, interference will only occur where the collimated beams overlap spatially, and so by definition interference can only occur between common path beams. This explains why the system has such high sensitivity, quoted as 0.1 Å for the rms stability.

---

[31] Due to the difference in illumination angles of the two sample beams by the obliquity effect, piston of the sample will not be entirely rejected.

The mismatch in the beams also leads to a significant limitation of the technique, since there will also be an accompanying loss of interference signal. The ratio of the beam diameters is normally 1:25, and so only 1/625th of the cross section of the expanded beam contributes to the interference signal.[32] Since the remainder of the beam would only produce extra shot noise at the detector, it is shown spatially filtered in the diagram. This imposes an unwelcome constraint on the ratio of the areas illuminated by the probe and reference beams on the sample. As discussed previously with the Downs interferometer, the reference area must be large to achieve a true absolute phase response. In addition, the loss of the phase information carried by this light will inevitably degrade the imaging performance.

The third system presented in this section, was developed by Jungerman et al.[33] and uses a Bragg cell to introduce the heterodyne frequency as well as to scan the optical beam. The optical arrangement is depicted in Fig. 2.5. The Bragg cell divides the incident beam equally into an undiffracted optical beam of frequency f_0 and a Bragg diffracted beam upshifted to a frequency f_0+f_B, where f_B is the frequency of the electrical drive to the Bragg cell. The objective lens is positioned such that its back focal plane is located at the centre of the Bragg cell. This focuses the two beams onto the sample and converts their angular deviation produced by the Bragg cell to a lateral separation. After reflection from the sample, the beams retrace their paths to the Bragg cell. Part of the beam which was diffracted on its first passage through the Bragg cell passes back through undeflected, with no further frequency shift and reaches the signal detector. In addition, part of the previously undiffracted beam is down shifted in frequency to f_0-f_B and also reaches this detector.

The two beams at f_0+f_B and f_0-f_B interfere at the detector to produce a

---

[32] In practice the reduction in interference would not be quite this bad since we are dealing with Gaussian beams.

[33] R.L.Jungerman, P.C.D.Hobbs and G.S.Kino, "Phase sensitive scanning optical microscope", Appl. Phys. Lett., **45** (8), pp. 846-848 (1984).

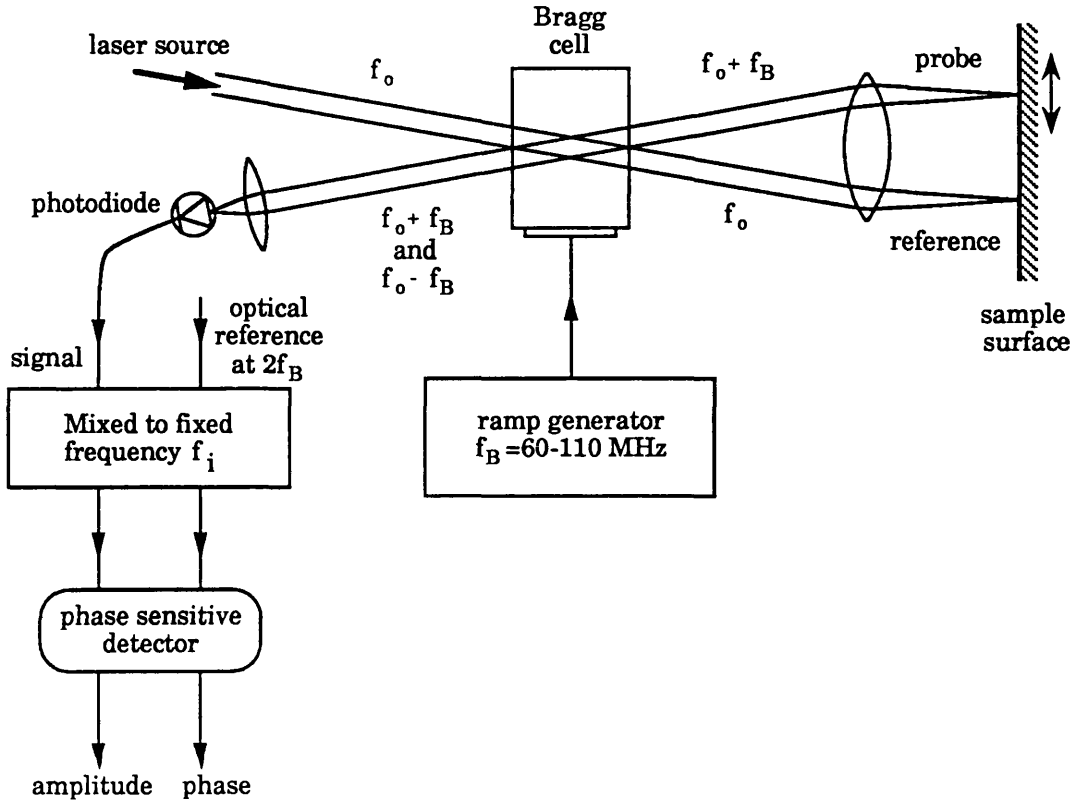


Figure 2.5. The heterodyne interferometer of Jungerman et al.

heterodyne signal at $2f_B$, whose phase is dependent on the path difference between the two beams impinging on the sample. Scanning is achieved by varying the Bragg cell drive frequency f_B. This changes the angular deflection on the first order diffracted beam, so that it scans the sample linearly relative to the fixed zeroth order beam. The detected signal is therefore given by Eqn. (2.3) with the interference signal now at $2f_B$. Extraction of the phase from the interference signal using phase sensitive detection requires a reference signal which is also at $2f_B$. Since this frequency is ramped to provide the scanning, the reference cannot be derived by doubling the Bragg cell drive due to the finite acoustic delay (this will cause a slight difference in the optical and electrical frequencies). For this reason an optically generated reference is used. Part of the incident beam is split off out of the plane of the other beams and directed into the Bragg cell in a plane perpendicular to the first. This produces a second set of spots which illuminate another part of the sample. If this area is chosen to be relatively smooth, the interference signal detected from this second interferometer can be used as a reference for the first. In order to perform the phase sensitive detection at a fixed frequency, the reference signal is first mixed with a local oscillator at frequency f_i say. The lower sideband at $2f_B$-f_i is then mixed with the signal to provide a fixed frequency signal at f_i.

The sensitivity of an initial implementation of the system is about 100 Å which is at least an order of magnitude outside what would normally be expected for this type of technique. The reason for this is not clear, but may be associated with non-ideal performance of the signal processing scheme. The requirement that part of the sample be flat to provide a reference surface may restrict the possible applications of the system. The authors suggest that the beams of the reference interferometer could illuminate larger area of the surface as a means of overcoming the problem. However, this would complicate the optics and make the system less common path. The main advantage of the instrument therefore is the high scan rate that may be obtained with the Bragg cell, though the sensitivity needs to be improved before it can rival the other

techniques presented in this chapter.

The final system in this section has been chosen as as example of scanning heterodyne technique which records differential phase. Although there are various differential systems available,[34] this one is unique in that it also measures the differential intensity information simultaneous with the differential phase. The optical configuration of this system, described by See et al.[35,36] is illustrated in Fig. 2.6. The light from the laser is directed into a Bragg cell which is driven by two signals of frequency ω_1 and ω_2. The Bragg generates two first order diffracted beams corresponding to these signals and they propagate with an angular separation. As in the previous technique of Jungerman et al., these beams are focused to adjacent points on the sample by the objective lens. The reflected beams are recombined by the Bragg cell and subsequently interfere at the photodetector. The detected interference signal at frequency $(\omega_1-\omega_2)$ is therefore given by Eqn. (2.3), but now the phase δ represents a differential measurement, since the sample is scanned relative to the two probe beams of fixed separation.

The differential intensity measurement depends on the difference in reflected light from the the two probe areas. It is obtained by amplitude modulating the two Bragg cell drive signals in phase quadrature at frequency ω_a say. This results in a several frequency components being detected. The differential intensity response (I_{DI}) is at frequency $2\omega_a$ and is given by

---

[34] R.K.Appel, "Differential scanning optical profilometer for simultaneous measurement of reflectivity and phase variation", PhD thesis, Chpt. 2, University of London, (1990).

[35] C.W.See, R.K.Appel and M.G.Somekh, "Scanning differential optical profilometer for simultaneous measurement of amplitude and phase variation", Appl. Phys. Lett., **53** (1), pp. 10-12 (1988).

[36] R.K.Appel, M.G.Somekh and C.W.See, "A scanning differential intensity and phase system for optical metrology", Proc. SPIE, **1164** , *Surface Characterization and Testing II*, pp.250-261 (1989).

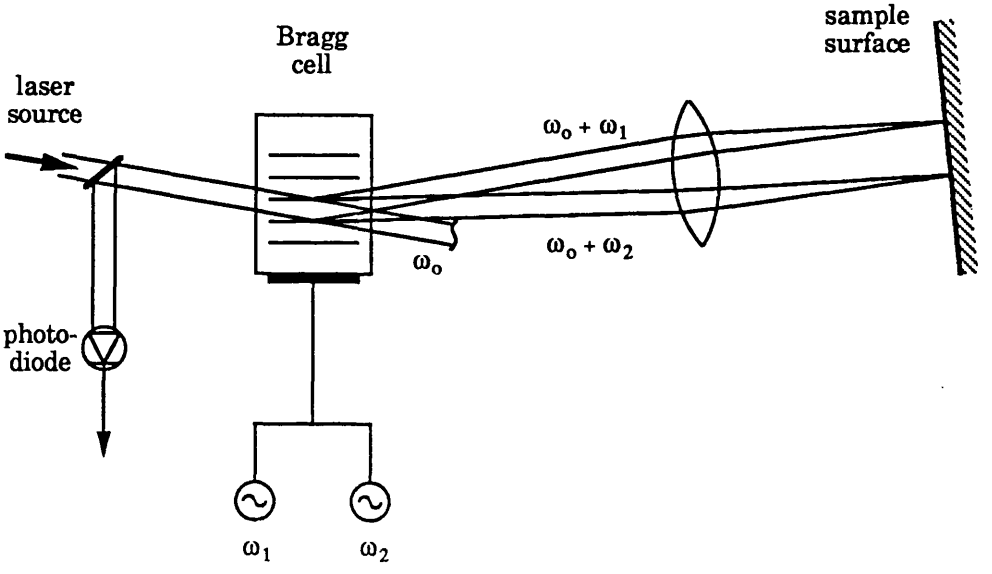


Figure 2.6. The differential intensity and phase interferometer of See et al.

$$I_{DI} \propto (R_1 - R_2) \cos(2\omega_a t) \tag{2.4}$$

where R_1 and R_2 are the intensity reflection coefficients of the probed areas. This measurement therefore records changes in reflectivity in the direction of the scan.

The system is highly common path and completely insensitive to piston type vibration. The only region where air turbulence and thermal drift cause problems is between the Bragg cell and the sample, where the beams do not overlap spatially. The system is also relatively insensitive to sample tilt. Its main advantage however is that the simultaneous recording of differential intensity information allows the ambiguity associated with the phase measurement to be more easily resolved. The other systems presented in this chapter assume that the observed phase changes are entirely due to topography, which may not always be the case. Compositional variations across the sample will also affect the phase, and in this system they will be highlighted by the differential intensity measurement. It should be noted however, that the phase measurement is also differential, and for some applications it may be desirable to measure the true profile i.e. absolute phase.

2.2 Full field techniques

Although the use of scanning provides many advantages as detailed in the previous section, for imaging applications, it tends to be rather slow. This is especially true when the object is mechanically scanned, as opposed to optical scanning using galvanometer mirrors for example. For this reason, full field systems have developed in parallel with the scanning methods, as they offer shorter acquisition times, limited only by the detection arrangement.

2.2.1 Full field techniques - homodyne

The first full field system considered is that of Wyant et al.,[37] depicted in Fig. 2.7. As in the scanning homodyne system of Adachi, described in Section 2.2.1, the system uses a two beam Mirau interferometer. A tungsten-halogen lamp, spectrally filtered to 40 nm bandwidth provides an extended source for full field illumination. The main adaptation of this system from the original Mirau interferometer is the use of modern phase stepping methods to extract quantitative phase information. A CCD array is used to detect each interferogram which is digitised for subsequent processing. The intensity at each pixel (x,y) is given by the usual expression for a homodyne interferometer

$$I = I_s + I_r + 2\sqrt{I_s I_r}\cos\delta(x,y) \tag{2.5}$$

Where I_s and I_r are the intensities of the sample and reference beams respectively which photomix at the pixel (x,y), and $\delta(x,y)$ is the phase difference between these two beams. Clearly this expression is a special case of Eqn. (2.3), in which the frequency difference is zero, so that the interference term is at DC. Thus the phase information is distorted by the cosine term and is also mixed with non-interference terms. In this form, variations in the intensities I_s and I_r will confuse the measurement. These variations may be caused by non-uniformities across the light source, or by reflectivity variations over the sample and reference surface. The phase $\delta(x,y)$ is determined independent of these parameters by the "integrated-bucket" technique.[38] The reference surface is moved at constant velocity by a piezoelectric transducer, thereby changing the phase of the reference beam at a constant rate $\phi(t)$. The intensity at each pixel is integrated while $\phi(t)$ varies from 0 to $\pi/2$, $\pi/2$ to π and π to $3\pi/2$. From these three measurements, a

---

[37] J.C.Wyant, C.L.Koliopoulos, B.Bhushan and D.Basila, "Development of a three-dimensional noncontact digital optical profiler", J. Tribology, **108**, pp. 1-8, (1986).

[38] B.Bhushan,J.C.Wyant and C.L.Koliopoulos, "Measurement of surface topography of magnetic tapes by Mirau interferometry", Appl. Optics **24** (10), pp. 1489-1497 (1985).

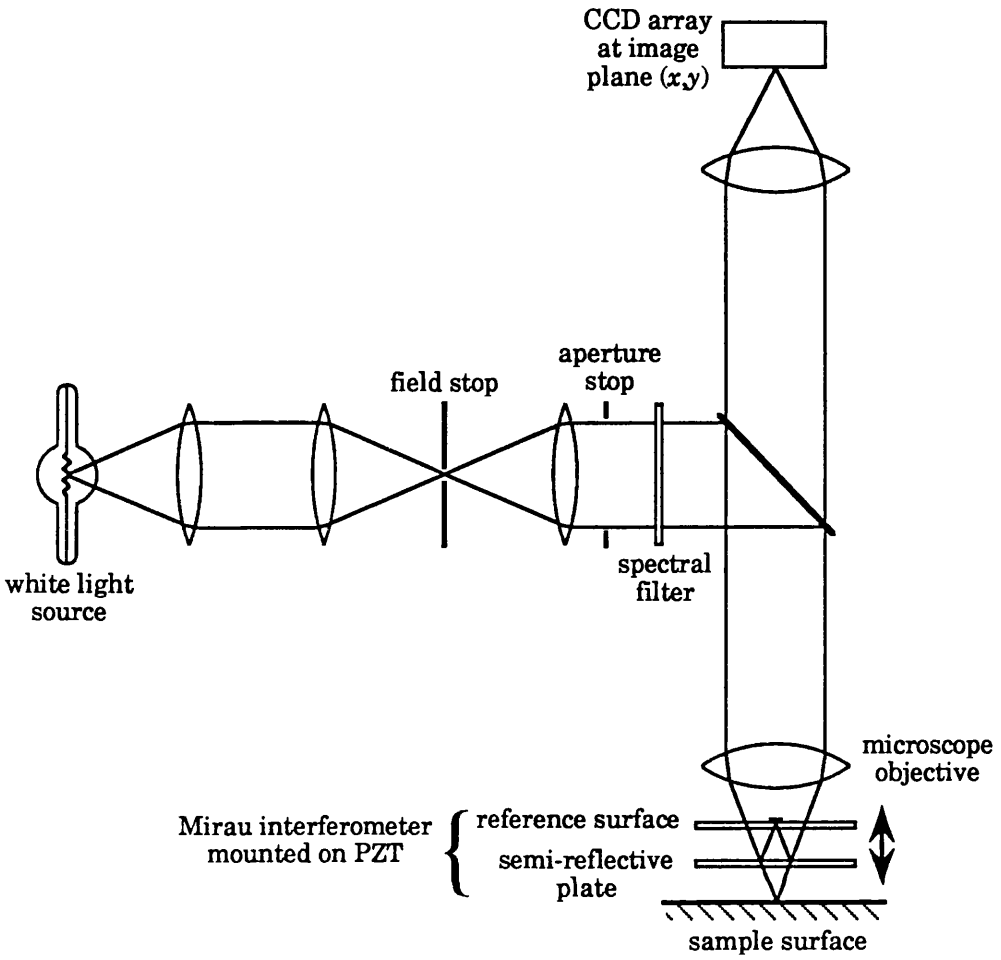


Figure 2.7. The "Wyko" full field phase shifting interferometer.

simple algorithm is used to determine $\delta(x,y)$. In a similar approach known as phase stepping,[39] a series of measurements are taken in which the phase is stepped in discrete multiples of $\pi/2$. As in the previous case, an appropriate algorithm is then used to extract the value $\delta(x,y)$. The integrated-bucket technique has the advantage that moving the reference at a constant velocity causes less vibration than phase stepping. It is interesting to note that a ramped phase change is equivalent to an optical frequency shift (or alternatively, the moving reference effectively introduces a Doppler shift in frequency), and so the integrated bucket technique essentially simulates heterodyne operation. In fact all phase shifting techniques may be thought of as digital heterodyne methods.[40]

Ideally, the reference surface should be perfectly flat and smooth. Any phase variations over the reference surface are measured using a super smooth mirror in place of the sample. This data is stored and subtracted from subsequent measurements by the software. The only significant drawback of the system is that the Mirau interferometer is not common path. The minimum resolvable height surface change for the instrument is quoted as 1 Å in the product literature.[41] Although this this value is comparable to that obtained with common path interferometers, more stringent environmental conditions are required to reduce vibration and thermal effects (and possibly a smaller measurement bandwidth as well, though this is not specified).

---

[39] P.Hariharan, B.F.Oreb and T.Eiju, "Digital phase-shifting interferometry: a simple error-compensating phase calculation algorithm", Appl. Optics **26** (13), pp. 2504-2505 (1987).

[40] K.Freischlad and C.L.Koliopoulos, "Fourier description of digital phase-measuring interferometry", J.Opt. Soc. Am. A, **7** (4), pp.542-551 (1990).

[41] Sales brochure for "TOPO-2D" (1988), Wyko corporation, Tucson, Arizona, U.S.A.

An example of a full field homodyne technique employing a common path interferometer is that of Iwata et al.,[42] illustrated in Fig. 2.8. Light from a laser diode is first passed through a Soleil compensator,[43] which is used to introduce a known phase difference between the two orthogonal polarisations of the beam. The collimated light emerging from the phase shifter enters a double focus lens, which acts as the beam divider for the interferometer. This lens contains a biconcave element (shown shaded) composed of birefringent material and thus it imposes different wavefront curvatures on the two orthogonal components of the beam. It is constructed so that the ordinary ray will pass through undeviated. The extra-ordinary ray focuses at the back focal plane of a conventional microscope objective, and therefore illuminates the sample as a collimated beam. This beam is the object beam of the interferometer shown as a solid line. The ordinary ray is focused onto the sample and the reflected light becomes the reference wave, denoted by the broken line. A quarter wave plate placed between the double focus lens and the objective exchanges the direction of polarisation of the two beams after the double passage. Thus the returning object beam passes back through the double focus lens without deflection, whereas the reference wave is deflected by it. The remaining lenses in the system serve to match the two wavefronts for optimum interference. A polariser oriented at 45° to these two orthogonally polarised beams is also necessary for them to interfere. As for the Wyko system, the interferogram is recorded by a CCD array and the phase extracted using a phase stepping method. In this case, the phase steps are introduced with the Soleil compensator.

Since this interferometer is common path, it is potentially able to achieve greater sensitivity than the Wyko system. The reproducibility obtained from an initial version of the Iwata interferometer is quoted as 200 Å

---

[42] K.Iwata and T.Nishikawa, "Profile measurement with a phase-shifting common-path polarisation interferometer", Proc. SPIE, **1162** -43, *Laser interferometry-Quantitative analysis of interferograms*, (1989).

[43] See for example: E.Hecht, *Optics*, 2nd Ed., p.305, Addison-Wesley (1987).

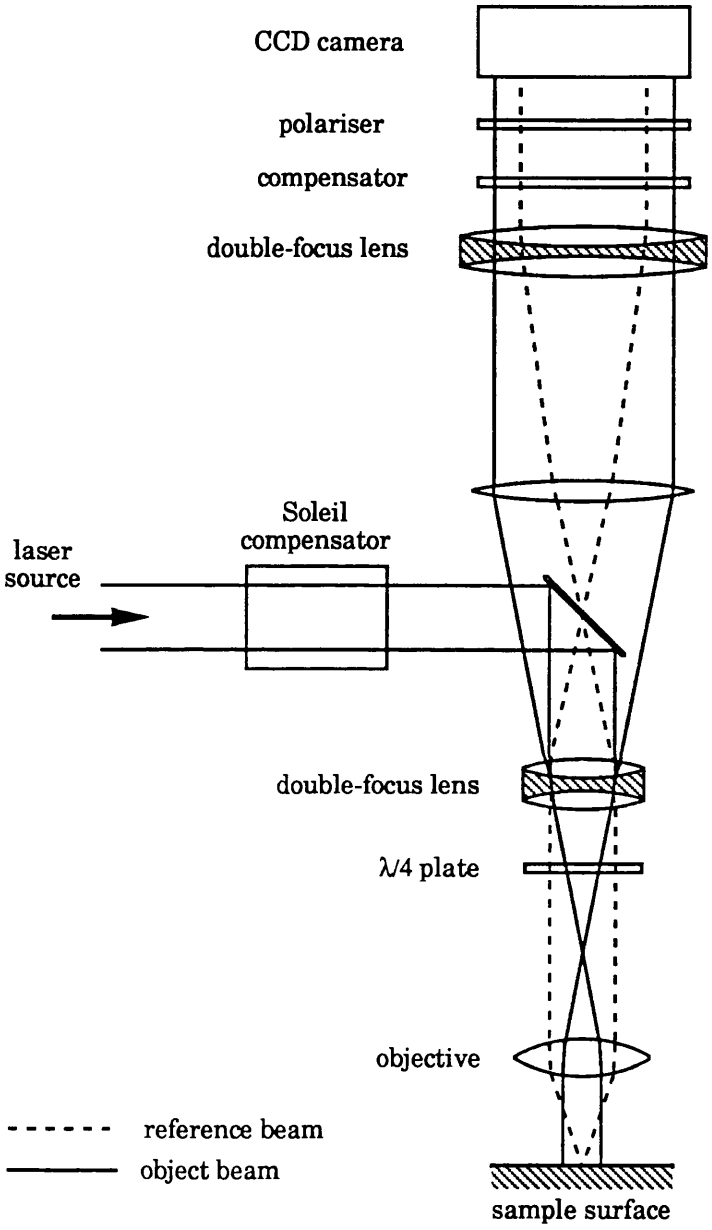


Figure 2.8. The common path phase shifting interferometer of Iwata et al.

however, and this rather poor result is attributed to back reflections in the optical system by the designers. Another possible source of problems is the quarter wave plate, since the the two polarisation components do not pass through it at the same angle. Since the retardance introduced by such a plate is dependent on the thickness presented to each beam, it cannot act as a true quarter wave plate for both components.

The arrangement could be converted to a scanning common path interferometer if the focussed beam is used as the sample probe and the collimated beam is used as the large area reference. In this mode there would no longer be any need for the quarter waveplate, and the last two lenses used to match the wavefronts and to image the interferogram onto the CCD array. The double focus lens would then act as both the beam divider and recombiner for the interferometer in much the same way as the Downs interferometer described earlier.

2.2.2 Full field techniques - heterodyne

The only system to date which can be classified under this heading is that of Massie et al.[44] The reason for such scarcity is probably due to the difficulty in performing full field heterodyne detection. Fig. 2.9 shows the optical configuration. The laser source is split into two beams which are shifted in optical frequency by ω_1 and ω_2, and then recombined collinearly such that the two frequencies can each be identified with an orthogonal polarisation. The combined beam is then directed into a polarisation sensitive Twyman-Green interferometer[45] (i.e. it incorporates a polarising beamsplitter). As a result one frequency component passes into the reference arm and the other passes into the test arm. A quarter wave plate placed in each arm rotates the

---

[44] N.A.Massie, R.D.Nelson and S.Holly, "High -performance real time heterodyne interferometry", Appl. Optics 18 (11), pp. 1797-1803 (1979).

[45] This is a slight modification to the Michelson interferometer to allow testing of optical components. See for example: E.Hecht, *Optics*, 2nd Ed., p.385, Addison-Wesley (1987).

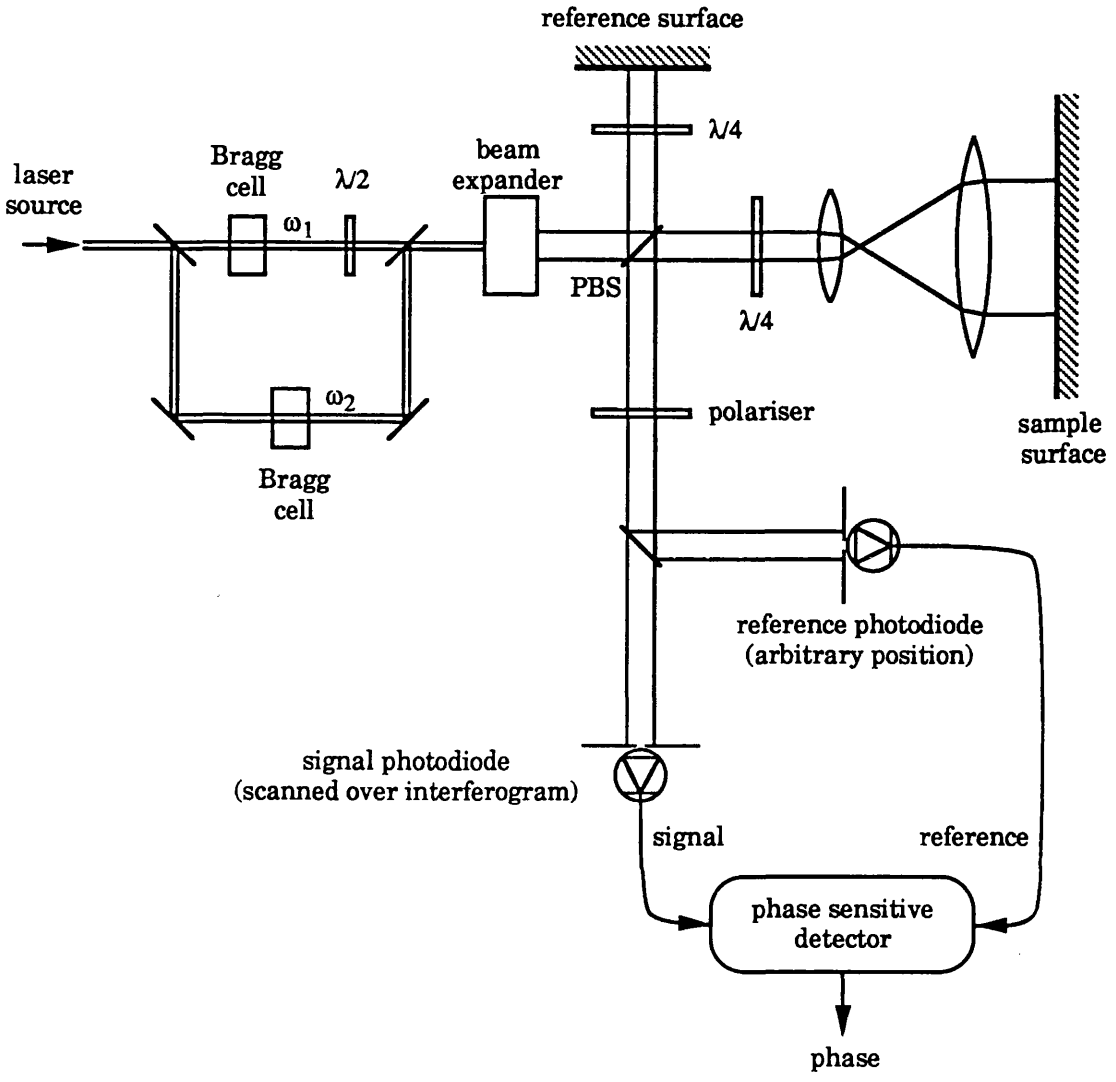


Figure 2.9. The full field heterodyne interferometer of Massie et al.

polarisation of each beam by 90^0 after the second passage, so that the returning beams are both directed to the polariser. As in the Iwata interferometer such a polariser at 45^0 to the two orthogonal polarisations allows them to interfere. The intensity at each point (x,y) at the detection plane is given by

$$I = I_s + I_r + 2\sqrt{I_s I_r} \cos\{(\omega_1 - \omega_2)t + \delta(x,y)\} \qquad (2.6)$$

The system is full field in that the whole interferogram is directly measurable at all times. The phase $\delta(x,y)$ is extracted using phase sensitive detection. In the embodiment described however, detection was achieved by scanning a point detector over the interferogram, since a practical full field phase sensitive detector has yet to be realised.

An interesting feature of the interferometer is the way the reference signal is derived from an arbitrary point in the interferogram. In this way phase noise caused by piston type vibration in the optics is common to both signals and is removed by the detection scheme. This is essential for high accuracy since the interferometer is not common path. A similar approach, taken a stage further,[46] has been used to remove vibration effects in the system presented in this thesis. Results from an initial breadboard version of the Massie arrangement have demonstrated $\lambda/100$ phase accuracy, and the designers claim that a well engineered instrument would probably achieve even greater sensitivity.

Despite a full field image being formed by the system, only a fraction of the total power is monitored at any one time, thereby reducing the signal to noise ratio compared to scanning techniques, where all the power is focused on each point in turn. Such a reduction could only be justified by taking advantage of the full field property to achieve faster image acquisition relative to scanning systems.

---

[46] In the system described in this thesis, instead of deriving the reference signal from part of the same interferometer, as is the case with the Massie system, a second interferometer is introduced, the sole purpose of which is to provide a reference signal with phase noise common to the probe interferometer.

2.3 Summary

A number of interferometers have been presented which record quantitative phase information. The most sensitive of these are those which exhibit insensitivity to phase noise caused by microphonics. There are two methods by which this is achieved. The most common way is to design the interferometer so that the interfering beams are equally affected by the microphonics. In this way the phase noise is removed optically when the beams interfere, so that the phase of the detected signal is noise free. The second way is to remove the phase noise in the post detection stage, by subtraction, using a reference signal which carries the same noise. This is the method used in the systems of Massie and Adachi, where the reference signal is derived from a portion of the interferogram. The instrument designed by Huang uses a combination of both methods to remove the phase noise.

It could be argued that removing the noise optically is the best approach because the electronic method may not give perfect common mode rejection. However, the electronic method is far less restrictive on the design of the interferometer, since it does not require that it be common path. This is important, since it gives freedom to design an arrangement which provides a true absolute phase measurement and which has optimum interference efficiency. The most sensitive techniques are those of Downs and Huang, but they only record absolute phase on samples which are extremely smooth.

The system described in this thesis has been designed to reject the phase noise whilst maintaining a quantitative absolute phase measurement for a broader range of samples. The idea of removing the noise electronically has been extended to the use of two parallel interferometers, whose signals are subtracted. This has lead to the development of a system which is simple in both concept and construction. This is an important consideration if the system is to progress beyond the development stage to that of a widely used, rugged laboratory tool.

CHAPTER 3

THE ABSOLUTE PHASE SCANNING HETERODYNE INTERFEROMETER

3. The absolute phase scanning heterodyne interferometer

The heterodyne interferometer described in this thesis is a development of the AC version of the Michelson interferometer. The main criterion which governed the design of the system was that of achieving insensitivity to the limiting effects of vibration in the optics, whilst also maintaining a true absolute phase measurement. The general principle of the microscope is discussed and theory describing the extraction of the signal is formulated. Later sections discuss the main optical components of the system, placing more emphasis on the design of the specialised objective. Finally, practical details concerning the alignment of the optics are given in appendices at the end of the chapter.

3.1 System configuration

A schematic diagram of the optical system is shown in Fig. 3.1. Two laser beams interact with the sample, and these are subsequently interfered with a third frequency shifted beam, so forming two heterodyne Michelson interferometers. The phase of each AC signal from these interferometers is compared in order to separate the sample phase information from the phase noise caused by microphonics.

After beamsplitter BS, the transmitted beam passes to the so-called 'element A', the nature of which is discussed in detail in Section 3.3.5. In its simplest implementation, it consists of a lens with a centrally drilled hole and serves to direct two beams onto the sample. The first beam (the sample probe) forms a tightly focussed spot on the sample. The second (the sample reference) passes through the hole and remains collimated thus illuminating a relatively large area of the surface. Each beam passes back through 'element A' to separate photodetectors D1 and D2.

The light from the laser which is reflected at beamsplitter BS passes through a Bragg cell to a reference mirror. The double passage through

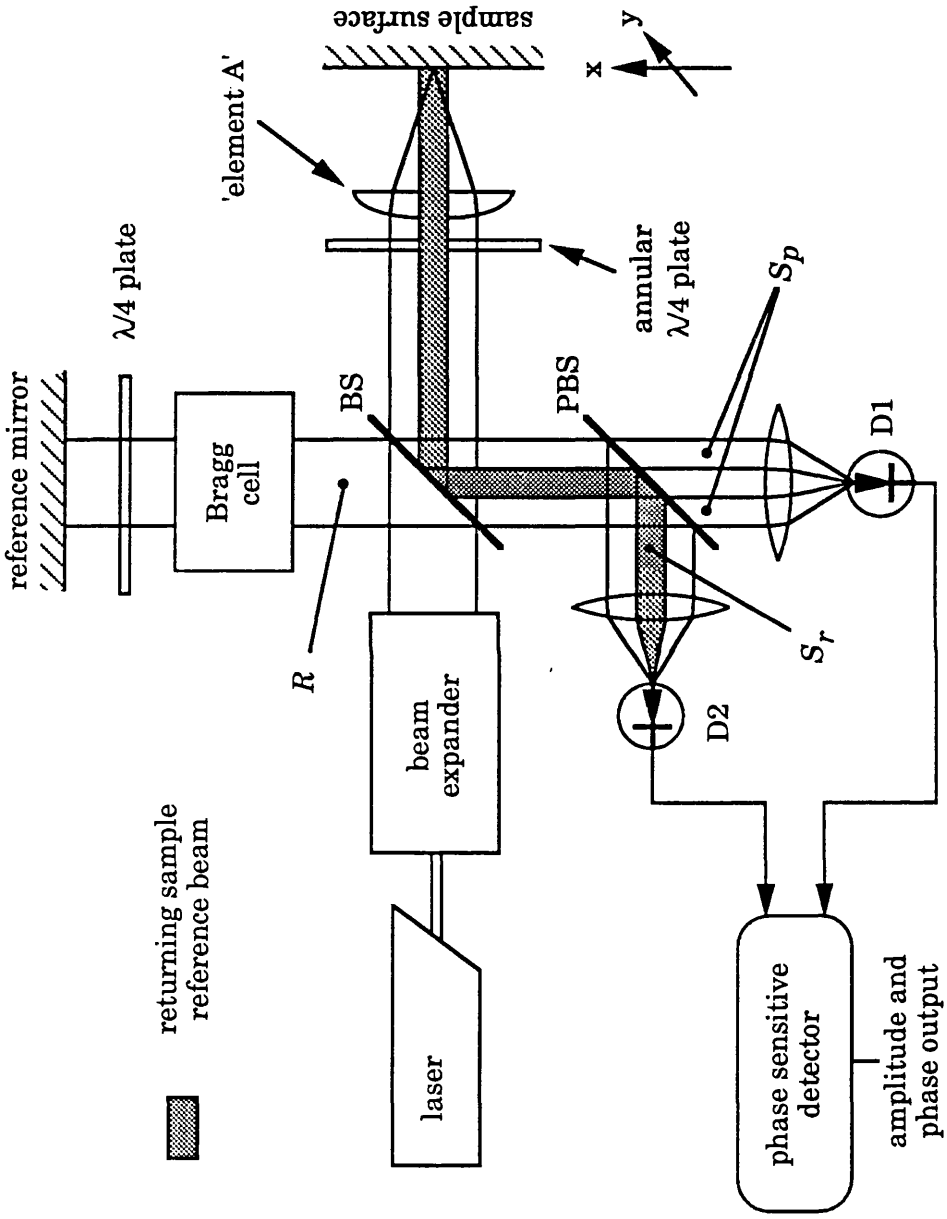


Figure 3.1 Optical configuration of the microscope

the Bragg cell imposes an upshift in frequency of $2\omega_B$ to the returning beam, where ω_B is the Bragg frequency. This beam [the common reference (R)] interferes with both the sample probe (S_p) and the sample reference (S_r). The resulting interferometers are termed the probe and reference interferometers respectively, for reasons which will become apparent later. The common reference beam is required because the sample probe and sample reference beams cannot be interfered directly. The reason is that the two collimated beams returning from the sample do not overlap spatially. This highlights a difficulty encountered when using a large area on sample reference, in that beams which are spatially distinct at one plane (i.e. the sample surface), will not generally interfere efficiently at a subsequent detection plane.[47] The use of two parallel interferometers avoids this problem, so there is no longer any serious limitation on the area illuminated by the reference beam, thus ensuring an absolute phase measurement. The use of the common reference beam in this way has been termed indirect interference. It has proved to be a useful method of overcoming a number of constraints imposed by conventional direct interference. Further applications of the technique are discussed in Chapter 7.

The isolation of the signals from each interferometer may be achieved by placing an annular quarter wave plate in the sample arm, matched to the diameter of the hole drilled in the objective. The separation is then provided by the polarising beamsplitter PBS. In addition, by adjusting the rotation angle of a quarter wave plate placed in the reference arm of the interferometer, the amount of optical power from the common reference beam directed to each detector may be varied. This is a useful way of equalising the amplitudes of the two detected signals in order to minimise the effect of shot noise. Alternatively, the quarter wave plates may be removed and PBS replaced with an ordinary beamsplitter, if appropriate spatial filters are placed in front of each detector lens. For

[47] A.Korpel and R.L.Whitman, "Visualization of a coherent light field by heterodyning with a scanning laser beam", Appl. Optics, **8** (8), pp. 1577-1580, Appendix A (1969).

the reference interferometer, this takes the form of a circular aperture matched to the hole drilled in the objective, whereas for the probe interferometer its inverse is required (i.e. an opaque spot deposited on a glass plate). This latter approach is the safest method when examining birefringent samples and is generally favoured due to its simplicity.

3.2 Theory of signal extraction

The three beams described above may be represented by

$$S_p = A_p \exp j(\omega_0 t + \phi_p + \theta_1) \tag{3.1}$$

$$S_r = A_r \exp j(\omega_0 t + \phi_r + \theta_1) \tag{3.2}$$

$$R = B \exp j\{(\omega_0 + 2\omega_B t) + \phi_c + \theta_2\} \tag{3.3}$$

Where A and B are the appropriate amplitudes, ϕ is the phase of each beam, θ_1 and θ_2 are phase errors due to microphonics and ω_o is the optical frequency. Subscripts p, r and c refer to the sample probe, sample reference and common reference beams respectively. The output current I_p from photodetector D1, results from the interference between S_p and R

$$I_p = \xi G(S_p + R)(S_p + R)^* \tag{3.4}$$

$$\frac{I_p}{\xi G} = A_p^2 + B_p^2 + 2A_p B_p \cos[2\omega_B t + (\phi_c - \phi_p) + (\theta_2 - \theta_1)] \tag{3.5}$$

$$\xi = \frac{\eta e}{h\nu} \tag{3.6}$$

Where * denotes the complex conjugate quantity. B_p refers to the part of the common reference beam which interferes with the sample probe beam. η is the quantum efficiency of the photodetector, e is the electronic charge, h is Planck's constant, ν is the optical frequency and G is an internal gain factor. Similarly for photodetector D2 the

output current I_r results from the interference between S_r and R

$$I_r = \xi G(S_r + R)(S_r + R)^* \tag{3.7}$$

$$\frac{I_r}{\xi G} = A_r^2 + B_r^2 + 2A_r B_r \cos[2\omega_B t + (\phi_c - \phi_r) + (\theta_2 - \theta_1)] \tag{3.8}$$

B_r refers to the part of the common reference beam which interferes with sample reference beam. The phase of these signals can be compared using a coherent detection method such as a vector voltmeter. The resultant phase is $(\phi_p - \phi_r)$, since the phase errors due to microphonics are common to both signals and are therefore removed in the comparison. Finally, since the sample reference beam illuminates a relatively large area of the surface, ϕ_r remains constant as surface is scanned and so the resultant phase is directly related to the sample phase structure.

3.3 The optical arrangement

3.3.1 The laser source

A standard 5 mW linearly polarised HeNe laser was used in the first implementation. This had a $1/e^2$ beam waist of 0.8 mm and a longitudinal mode mode spacing of 435 MHz.

3.3.2 The optical isolator

This was a Faraday rotator and provided greater than 40 dB attenuation of the light returning to the laser cavity. The importance of such isolation is discussed in Chapter 4. There are other methods of preventing light returning into the laser cavity, and these are considered as follows.

The first approach involves introducing a slight misalignment into the system so that the returning beam misses the output aperture of the

laser head. This may be enhanced by placing a spatial filter assembly in front of the laser, such that the misaligned beam misses the pinhole. This problem with this technique is the difficulty in providing sufficient misalignment to achieve isolation, whilst also maintaining adequate alignment of the interferometer.

Another method, favoured in the interferometer of Downs,described in the previous chapter, removes the path for light back into the laser cavity, by placing a 45° mirror over half of the aperture of the objective, in order to collect the light reflected from the sample. This is a convenient, inexpensive solution, but has the disadvantage of halving the numerical aperture of the objective in one direction, thus degrading the imaging performance. In addition, back reflections from optical surfaces preceding the mirror could still pose problems.

Alternatively, an isolator may be constructed from a polariser and a quarter wave plate. The light from the laser first passes through the polariser and then emerges circularly polarised from the quarter wave plate. Any light returning through the quarter wave plate strikes the polariser at 90° to its pass direction and is extinguished. This method is limited by the quality of the quarter wave plate, which typically has $\lambda/500$ retardation tolerance,[48] and normally provides about 20-30 dB isolation.

3.3.3 The beam expanders

Two beam expanders are used in the system. The first is required to expand the sample probe beam to fill the objective aperture. The second is used as a telescope in order to reduce the common reference beam diameter back down to the restricted aperture of the Bragg cell. At first sight it would seem that one beam expander, placed in the sample arm after the beamsplitter would suffice. However, the detected beams would then have a maximum cross section diameter of 0.8 mm, resulting in serious diffraction problems and making spatial separation of the

[48] Melles Griot product catalogue, *Optics Guide 4*, chpt. 15, p. 28 (1988).

interferometers difficult.

Identical x10 beam expanders were used, each possessing better than $\lambda/4$ wavefront distortion. In a mechanically scanned system, such high quality optics are not essential, since aberrations affect all measurement points equally. However, they are recommended, in order to maintain an acceptable level of interference efficiency. The diameter of the beam emerging from the beam expander was 8 mm.

3.3.4 The Bragg cell

The Bragg cell in the system is used solely as an optical frequency shifter, in order to perform heterodyning. Consequently, no amplitude modulation is imposed on the incident beam. In order to separate the frequency upshifted first order beam from the zeroth order, sufficient propagation distance must be allowed for. The Bragg cell used in the system had a centre frequency of 40 MHz which gives an angle of 7.6 mrad between the zeroth and first order beams at 632.8 nm optical wavelength. Therefore a safe spatial separation of 2 mm centre to centre between beams when they return to Bragg cell requires the reference mirror to be placed a distance of about 132 mm from the centre of the Bragg cell.

3.3.5 The microscope objective

The objective for the system must direct two beams onto the sample, the tightly focussed sample probe and the collimated sample reference. The simplest way to achieve this is to use a lens singlet with a centrally drilled hole. Alternatively, a planoconvex lens may be used if the central region of the spherical surface is polished flat, parallel to the planar surface. These objectives will be termed the drilled and polished lenses respectively.

In the initial implementation, a drilled planoconvex lens was used.

This had a focal length of 15 mm and a 2 mm diameter hole was drilled in its centre. This was achieved by immersing the lens in water bath with abrasive powder, and applying a tunable ultrasonic drill. In order to avoid damaging the optical surface of the lens, it was first dipped in hot wax. This protective layer was then removed after drilling using acetone. The laser beam diameter entering the lens was 8 mm, thus defining the effective numerical aperture (NA) as about 0.27.

The use of a singlet lens is adequate for low power objectives, but becomes impractical for higher numerical apertures. Spherical aberration becomes increasingly significant in singlet lenses at higher powers, and the effects of this are considered in detail in Chapter 4. For this reason, a multi-element objective is required which is corrected for such geometrical aberrations. This is the approach used in conventional refracting microscope objectives, but clearly, drilling a hole through such an arrangement is not feasible. In section 3.6, two methods are described which have been successfully employed to simulate the effect of this objective at higher powers.

3.3.6 The detector assembly

The detection arrangement used in the first implementation was based on the simple spatial filter approach discussed in Section 3.1. This was favoured for its simplicity and reliability, since the method of separating the interferometers according to polarisation gave a significant degree of crosstalk, due to corruption of the polarisation in the optics. It was found in practice that complete spatial separation of the interferometers was facilitated by the existence of an annular gap between the returning sample probe and sample reference beams, when viewed in cross section. The origin of this gap is discussed in Section 4.3.

The photodetectors were built to a design which had already been developed within the laboratory, and incorporated a standard silicon photodiode with an active area of 1 mm^2. These detectors were AC

coupled to remove the non-interference signals and had a bandwidth ranging from 2 KHz to over 100 MHz (for the -10 dB points).

3.4 The detection electronics and data acquisition

The signals from the photodetectors were fed into separate low noise amplifiers, each with about 50 dB gain. The amplified signal from the reference detector was then passed through a bandpass filter, centred on 80 MHz (i.e. twice the Bragg cell drive frequency). This filter had a pass bandwidth of 8 MHz (for the -3 dB points) and was used primarily to remove the 40 MHz signal due to electrical breakthrough from Bragg cell driver. This was then fed into the reference input of a Hewlett-Packard 8508A vector voltmeter. The amplified signal from the probe interferometer was fed directly into the signal input of the vector voltmeter. Although the amplitude and phase information could have been derived from the vector voltmeter, instead it was used to generate two 20 KHz frequencies, essentially beaten down versions of the original 80 MHz signals, in order to allow a lock-in amplifier to be employed. This gave more control over the measurement bandwidth, and also allowed a phase offset to be introduced in the final output. This proved useful in situations where the phase output happened to be near the wrap around point.

The amplitude and phase output from the lock-in amplifier were then digitised and stored using an IBM personal computer and later a Macintosh SE. Object scanning was achieved with a translation stage, actuated using a DC motor and drive belt to turn the micrometer lead screw. The position sensing consisted of a Schaevitz linear variable differential transformer (LVDT) attached to the translation stage, the output of which was recorded with the amplitude and phase data.

3.5 Results from the first implementation

In this section, preliminary results demonstrating the system performance are presented. These measurements were obtained using the 0.27 NA drilled lens singlet, described in Section 3.3.5.

The first measurement demonstrates the phase noise rejection capabilities of the system, by comparing the stability of the phase output with that of a basic (i.e. non-common path) heterodyne interferometer. Two measurements of the phase of the probe signal were recorded simultaneously, using different reference signals. The first was obtained using an electrically derived reference of arbitrary phase, generated by frequency doubling the Bragg cell drive signal. The second was recorded using the optically derived reference from the reference interferometer. The results for a measurement bandwidth of 33 Hz are shown in Figs. 3.2(a) and 3.2(b) respectively. These stabilities correspond to an rms variation in topography of 41.2 and 2.8 Å respectively, and thus the optical reference gives better than an order of magnitude improvement in sensitivity.

In order to establish that the system can record quantitative phase information, line scans were taken of a silicon wafer in which several tracks were plasma etched. This sample had previously been measured using a mechanical talystep, and the heterodyne differential interferometer of See et al. discussed in Chapter 2. The nature of the sample is shown schematically in Fig. 3.3, and consists of a series of tracks of increasing depth. The height of the steps A-H according to the three instruments are shown in Table 3.1 and demonstrate good agreement. Much of the discrepancy between readings represent true variations, since the profiles were measured at slightly different positions on the sample. An important observation is that the average of the measurements from each of the stylus probe and the absolute phase system agree to with 0.3%. This demonstrates that the sample reference beam averages the phase structure over a sufficiently large area to

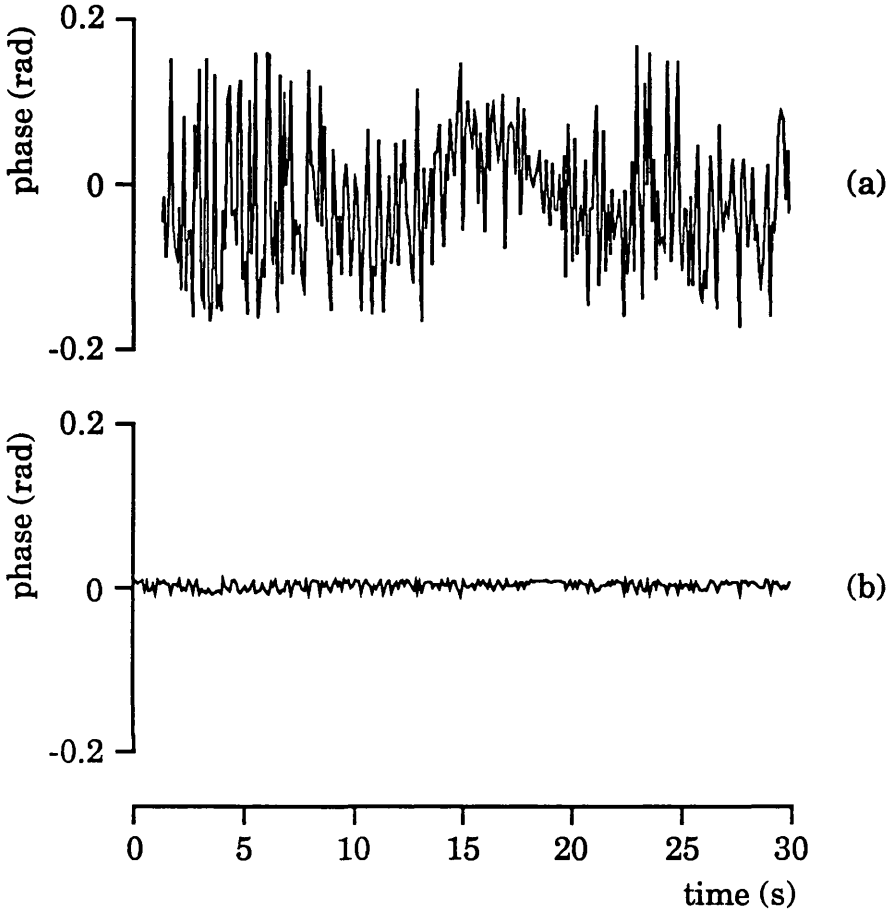


Figure 3.2. Stability of phase output (a) using an electronically derived reference signal and (b) using the optically generated reference signal.

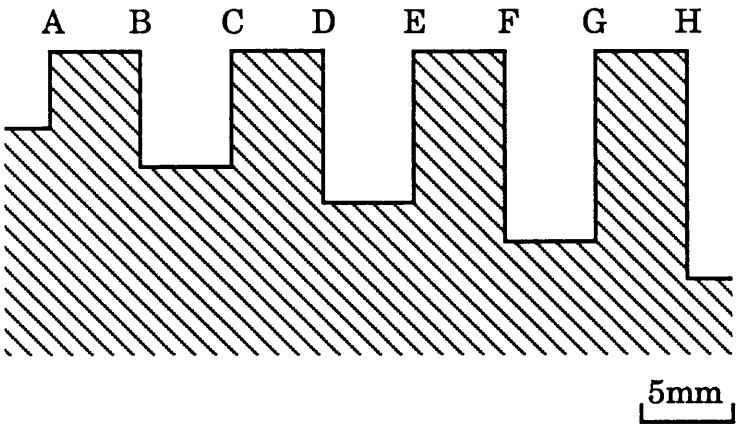


Figure 3.3. Schematic profile of the etched silicon wafer test sample, showing the steps A-H, measured with the system.

Step	A	B	C	D	E	F	G	H
Mechanical stylus (Å)	180	310	315	430	440	530	540	700
Differential phase system (Å)	163	326	302	424	490	579	609	713
Absolute phase system (Å)	166	302	309	457	432	575	552	661

Table 3.1. Comparison of step heights measured with a stylus probe, a differential optical phase profilometer and the absolute phase system.

provide a true absolute phase measurement. If this was not the case, the phase output of the reference interferometer would change significantly over the step, resulting in systematically smaller measured step heights.

The quality of the absolute phase measurement is also illustrated by the line scan shown in Fig. 3.4. The sample is a periodic structure, composed of chrome-gold tracks deposited on an optical flat. It was fabricated by sputtering chrome onto the surface of the optical flat which was covered with a mask consisting of a series of lines, 32 μm wide, with a 32 μm spacing. The chrome bars deposited on the surface in this way were then coated with a thin layer of chrome, followed by a thin layer of gold (chrome was used first, since this adheres well to the glass surface). The resulting sample therefore had a track width of 32 μm and a track height of about 800 Å. The apparent ringing at each edge is fact true topography, generated when the mask used in the fabrication of the sample was lifted away. This was confirmed by mechanical stylus measurements, and later using a confocal microscope and a homodyne differential interference contrast microscope. The results from the latter two systems are presented in Chapter 5.

3.6 Methods to improve the lateral resolution

In this section, two specialised objectives are described which achieve a high numerical aperture for the sample probe beam, whilst maintaining a large area collimated sample reference beam. This cannot be achieved with a lens singlet, due to the limiting effect of spherical aberration. The first two subsections deal with each approach, and results from the methods are compared in Section 3.6.3. The latter section also includes an overall assessment of the objectives, including the drilled/polished singlet.

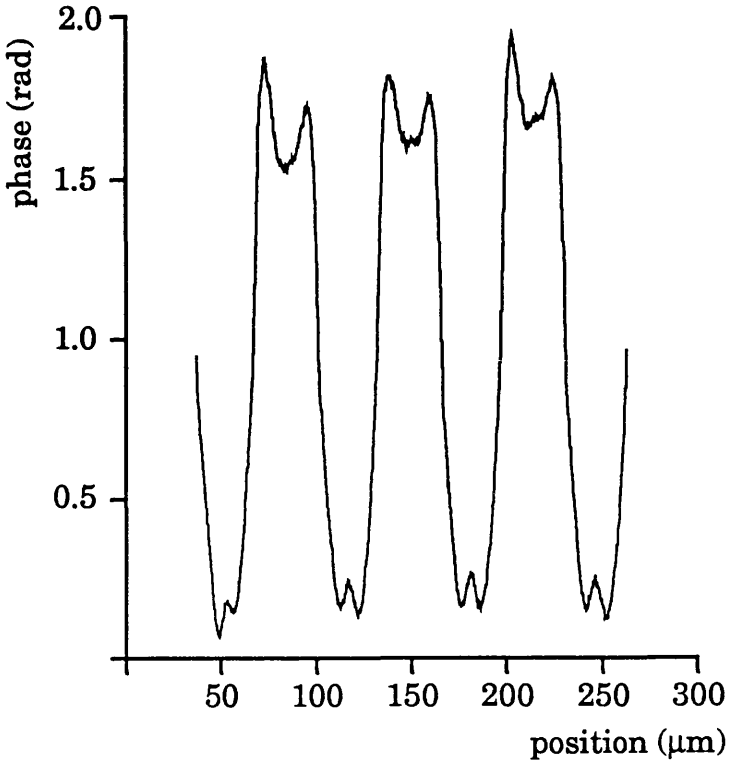


Figure 3.4. Line scan of the chrome-gold periodic track sample, taken using the 0.27 NA drilled planoconvex lens.

3.6.1 The reflecting objective

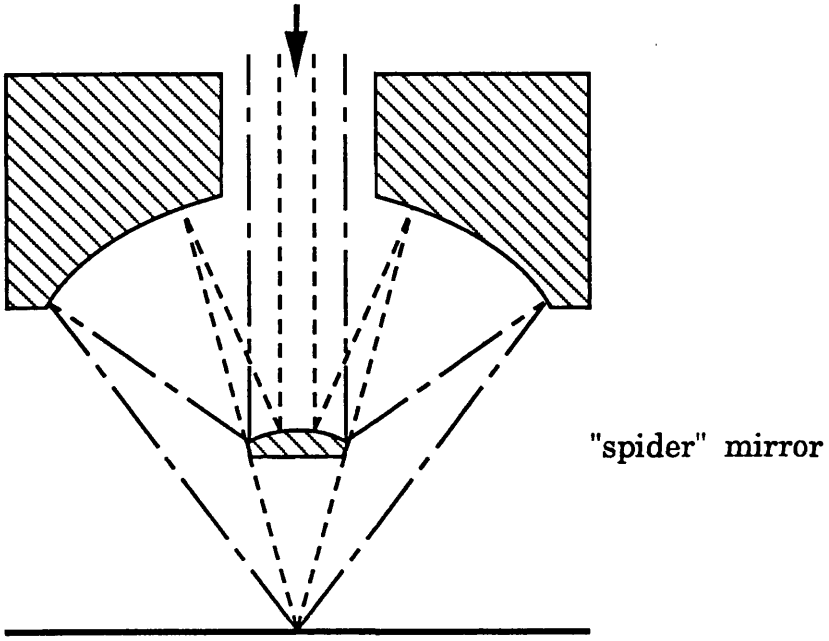
The relatively large size of reflecting objectives and the simplicity of their design makes drilling a central hole for the sample reference beam a feasible proposition. These objectives also have the added advantage of a long working distance and are free from spherical aberration. A schematic of the objective is shown in Fig. 3.5. The central region of the smaller convex mirror or "spider" forms an obscuration to the incident illuminating beam. This region may therefore be drilled out to provide the collimated sample reference beam, without affecting the power in the focussed probe beam.

Such an objective was implemented by modifying an Ealing Electro-optics standard 0.65 NA reflecting objective. This objective has a working distance of about 3.5 mm and the central obscuration of 1.9 mm diameter was drilled out, defining the size of the collimated reference beam. The diffraction limited sample probe spot for this objective is 1.2 μm diameter, which gives a ratio between the areas illuminated by the two sample beams of 2.6×10^6.

One disadvantage of this objective is the obstruction caused by the three spider arms which support the convex mirror. Since the objective is used in reflection, the three arm pattern gives the effective obstruction of a six arm spider support. The optimum pattern would be one with four arms at at 90° to each other. This arrangement gives the same obstruction after reflection due to its diagonal symmetry.

3.6.2 The projecting objective

In this method,the image of either of the low power lens singlets discussed in Section 3.3.5 is projected onto a high numerical aperture conventional refracting microscope objective.



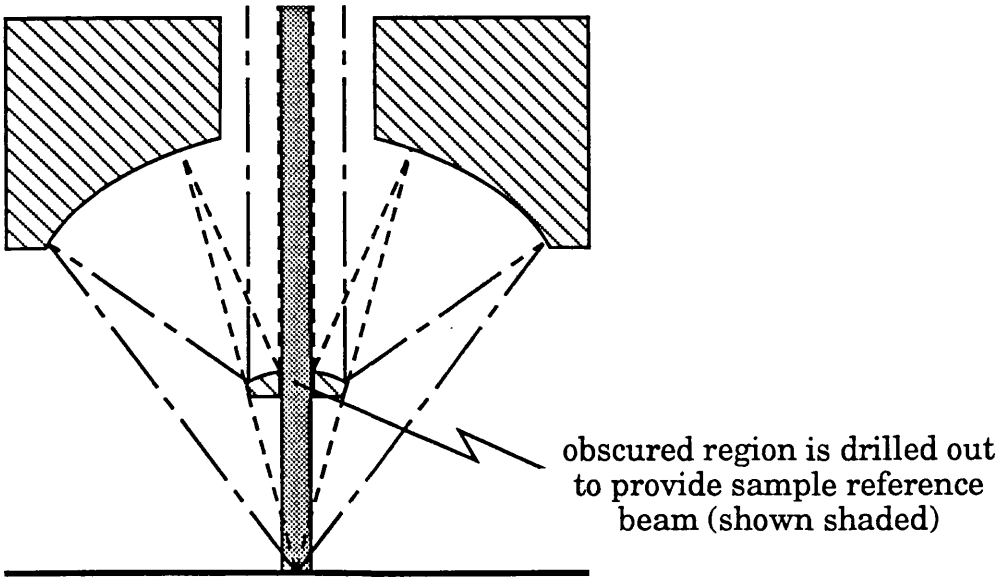


Figure 3.5. Modification to a standard reflecting objective to provide the collimated sample reference beam.

The ray diagram for this composite objective is shown in Fig. 3.6. The focal length of the intermediate lens (f_m), which matches the aperture of a given low power lens (radius a_l) to that of the high power objective (radius a_h) may be expressed as

$$f_m = \frac{a_h f_l}{a_l} \tag{3.9}$$

Where f_l is the focal length of the low power lens. The radius of the collimated reference beam on the sample (a_s) is given by

$$a_s = \frac{\varepsilon a_l f_h}{f_m} \tag{3.10}$$

Where the radius of the hole has been expressed as a fraction of the lens aperture, i.e. εa_l. The value ε has been termed the annulus parameter and is consistent with notation used in subsequent chapters.

The alignment of such multi-element lens arrangements is usually quite difficult, but in this case the procedure is almost trivial. First, the separation of drilled/polished singlet and the matching lens is adjusted until the probe beam emerges collimated from these two components. The final high power lens is then added and adjusted until the reference beam emerges collimated from the complete system. Once aligned in this way, the entire assembly may be treated as a single unit which may be interchanged with the lower power objectives discussed above.

In order to demonstrate the method, the image of a 60 mm focal length lens with a 2 mm diameter hole was projected onto a 0.65 NA refracting microscope objective. The apertures of these components were matched using a 40 mm focal length intermediate lens. From Eqn. (3.9), the aperture of the beam entering the microscope objective is therefore just under 2.7 mm radius (using a value of 4 mm for a_l, the radius of the beam emerging from the beam expander). The aperture of the

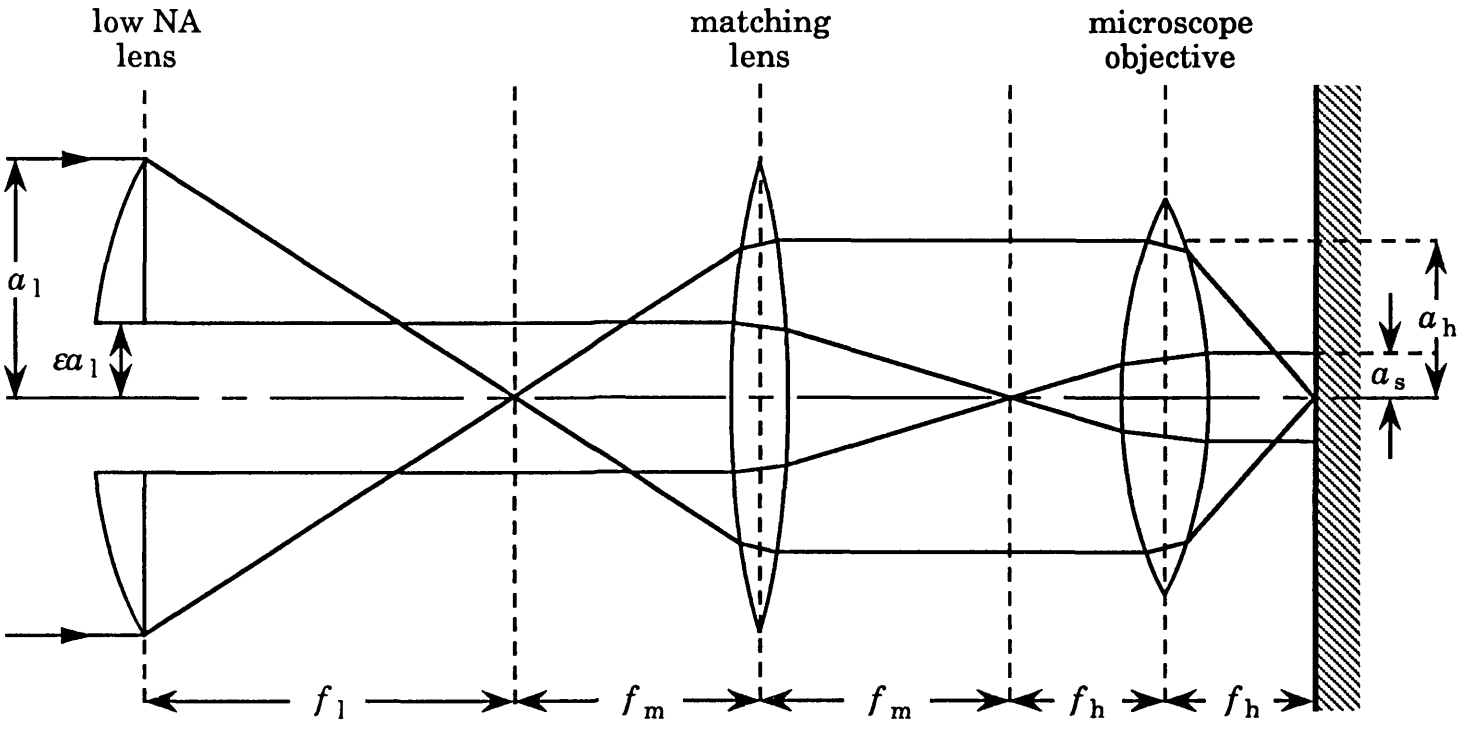


Figure 3.6. Ray diagram of the projecting objective

microscope objective was 2.6 mm radius, and thus nearly all the optical power is accommodated by the arrangement. The low power lens has an effective NA of 0.067 for the 8 mm diameter input beam, and this is therefore increased by almost a factor of ten by projecting it onto the microscope objective.

The radius of the collimated beam on the sample is given by Eqn. (3.10) as 0.1 mm, using the values f_h=4 mm and ε=0.25. A disadvantage of the technique therefore, is the reduction in the area illuminated by the sample reference beam, which is 100 times smaller than that provided by the drilled 0.65 NA reflecting objective. Associated with this reduction however, is a proportionate reduction in the angular deviation of the sample reference beam caused by sample tilt. For the drilled singlet and reflecting objectives, a sample tilt of α results in a deviation in the sample reference beam of 2α. For the projecting objective, this is reduced by an amount equal to the ratio f_h/f_m, i.e. the same reduction as for the diameter of the beam, given in Eqn. (3.10). Thus for the 0.65 NA projecting objective described above, the angular deviation of the sample reference beam due to tilt is reduced by a factor of 10. This is a considerable advantage for sloped samples, and the implications are discussed further in Section 4.4.

3.6.3 Results and assessment of the approaches

Figures 3.7(a) and 3.7(b) show line scans recorded with the 0.65 NA drilled reflecting objective and 0.65 NA projecting objective respectively. These line scans were taken on the same sample measured with the 0.27 NA drilled singlet, shown in Fig. 3.4. Although the three traces were taken over different regions of the sample, they do allow a comparison of the performance to be made. The increase in resolution provided by the both the reflecting and projecting objectives is evident from these scans. Another point worth noting is the difference in the step heights indicated by the three methods. Part of the variation is undoubtedly

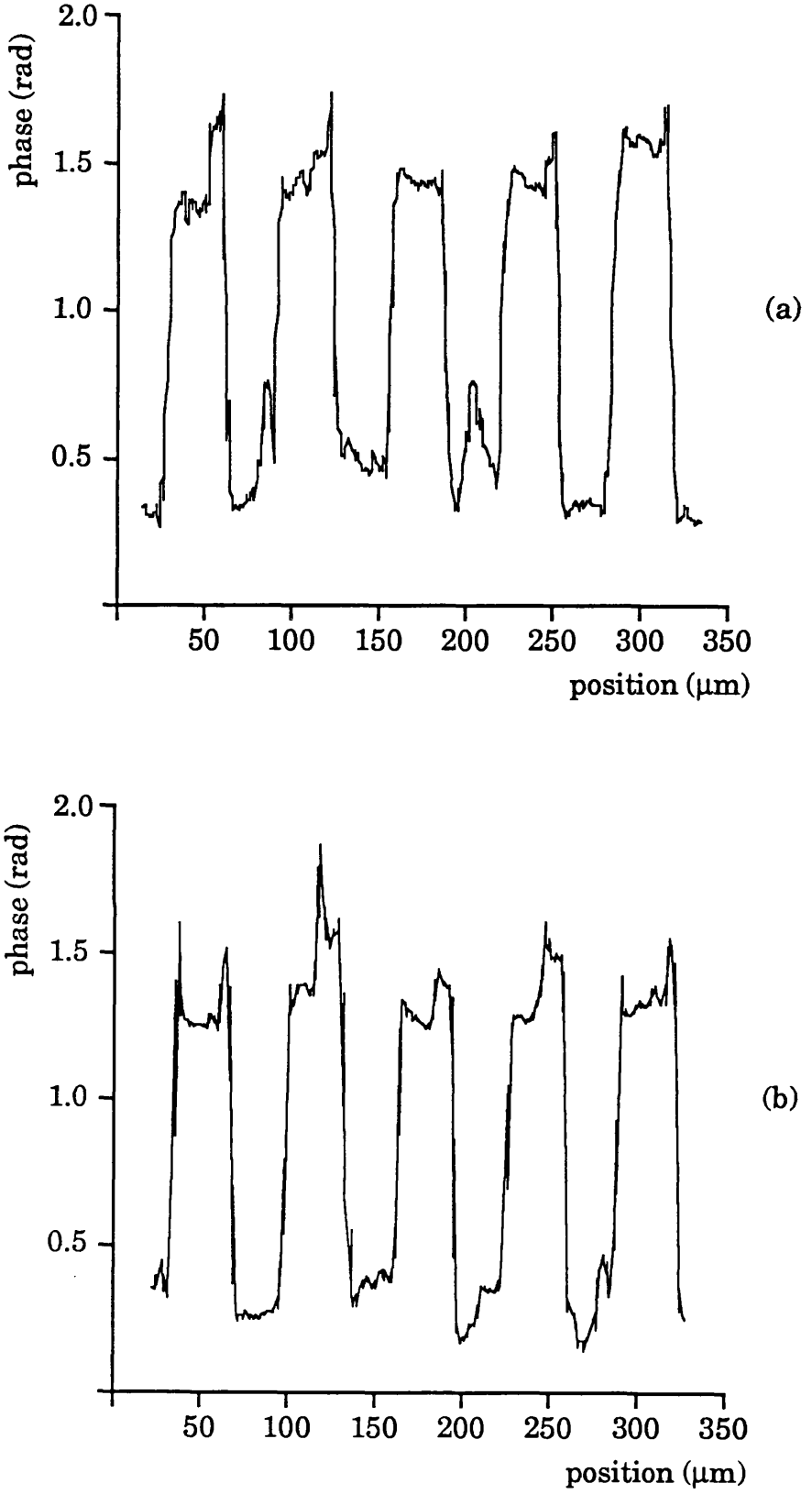


Figure 3.7. Line scan of the chrome-gold periodic track sample, taken using (a) the 0.65 NA drilled reflecting objective, and (b) the 0.65 NA projecting objective.

genuine, but a consistent discrepancy has been observed between the phase response using the drilled singlet, compared to those recorded with the two higher power objectives. This is attributed to the increasingly oblique illumination angle at higher numerical apertures, which reduces the phase change of each ray by a $\cos\theta$ factor. This effect is considered in more detail in Section 4.2.1.

Although the two methods achieve similar resolution, 15 dB greater signal to noise ratio was obtained for the probe interferometer signal when using the projecting objective. This is partly due to the obstruction to the probe beam caused by the arms supporting the convex mirror in the reflecting objective, as well as a slight mismatch in the beam diameter and its aperture. These factors cannot account for all of the reduction in signal however, and this implies that the reflecting objective does have aberrations, despite the manufacturers claims to the contrary. The projecting objective is virtually free of spherical aberration, since the drilled singlet and matching lens have such low NA's. The high power is provided by the conventional microscope objective, which is well corrected for geometrical aberrations.

For applications which do not require high lateral resolution, the drilled or polished lens singlet provides the most straightforward and inexpensive solution. The main disadvantage of this approach is the reduction in interference efficiency caused by spherical aberration, which becomes significant for numerical apertures greater than 0.1 (this is demonstrated in Section 4.3). This may be overcome by modifying achromatic doublets or the "best-form" laser lenses[49] that are now available. An advantage of operating the system with a low power objective (up to about 0.25 NA) is that the system gives better rejection of phase errors caused by microphonics of the sample. For higher power objectives, the rays of the probe beam beam illuminate the sample at a more oblique angle, and consequently suffer a slightly different phase error than the sample reference beam. This effect is

[49] See for example *Ealing Electroptics Product Guide*, p. K-24 (1990).

considered in detail in Section 4.2.1.

The reflecting objective is best suited to the intermediate magnification range, since it is presently only available up to 0.65 NA. It is more compact than the projecting objective and incorporates into the interferometer as easily as the lens singlet. The major drawback is its cost which is far in excess of the other methods.

The projecting objective is essential for the highest NA's and also provides a cheaper alternative to the reflecting objective in the intermediate range. One disadvantage is its long length in the axial direction, making it difficult to interchange it with the lower power objectives. However, if this interferometer is developed into a commercial instrument, it should be possible to design and build a compact form of this objective, with diffraction limited performance. As stated earlier, the projecting optics reduces the diameter of the sample reference beam, and this would tend to favour the reflecting objective. The reference area could be increased by drilling a larger hole in the lens singlet, but this would reduce the power in the sample probe beam, especially since the beam has a Gaussian profile. However, since the projecting objective gives better interference efficiency, there is at least some room to increase the size of the hole and still obtain the signal to noise ratio achieved by the reflecting objective. The most significant advantage of the projecting objective is the improved tolerance it gives to sample tilt, by reducing the angular deviation of sample reference beam. This effect alone justifies its use when the most accurate quantitative phase information is required. As stated earlier, this is discussed in more detail in Section 4.4.

3.7 Summary

The principle of the system has been discussed, and results from a bench top implementation have been presented. These results demonstrate the important attribute of the technique, namely the

quantitative absolute phase response. The stability of the phase corresponds to 2.8 Å for the minimum detectable height variation, measured in a 33 Hz bandwidth. This was achieved without a sophisticated vibration isolation system, which usually accompanies such an instrument. For all the measurements presented in this chapter, and indeed throughout the thesis, the isolation consisted of merely an optical table supported by four inflated car tyre inner tubes. Two methods to improve the lateral resolution of the system have been demonstrated. The results indicate that even the highest numerical apertures may be obtained for the sample probe beam by the projection method, whilst also providing a collimated sample reference beam.

A3.1 Aligning the optical system

1. Start with all components in their rough positions and minimise propagation distances.
2. Mark these positions and then remove all except the laser and scan stage with a mounted sample.
3. Add the beam expander and align so that expanded beam is coaxial with the input beam.

Sample arm:

4. Adjust the sample tilt until the light retraces its path back into the laser head.
5. Add the chosen objective (this may take one of three forms as discussed in Sections 3.3.5 and 3.6) and beamsplitter BS, and align so that the light still retraces its path. Adjust until the sample is in focus by viewing the reflected beam after BS. This beam will have an annular cross section and is collimated when the sample is in focus.
6. If the sample is mechanically scanned, the scan direction must be perpendicular to the optic axis. This requires a second tilt stage beneath the scanner which we will call the scan tilt. Scan the sample back and forth under the illuminating beam and view the annular beam after BS on a screen. If the scan direction is tilted the annulus will grow or contract, as a result of sample defocus. Adjust the scan tilt appropriately to reduce this problem. This will inevitably misalign the sample, which has to be readjusted using the sample tilt stage, until the light retraces into the laser head. Try scanning the sample again to check if the beam stays in focus. If it is not, repeat the above procedure (i.e. scan tilt with compensating sample tilt adjustments) iteratively until tilt of the scan direction is removed. The scan tilt stage should then be locked, since further adjustment is unnecessary, even if the sample is changed.
7. Put the optical isolator in place to prevent light re-entering the laser cavity, taking care to adjust its orientation, in order to maximise the power in the forward direction.

Reference arm:

8. Place the beam reducer in the reference arm just after BS and align so that the reduced beam is coaxial with the input beam.
9. Add the Bragg cell and reference mirror, taking care to allow sufficient propagation distance between them for the zeroth and first order diffracted beams (which propagate at a slight angle to each other) to be separated spatially.
10. Adjust the angle of the Bragg cell until the power diffracted into the first order is maximised.
11. Adjust the tilt of the reference mirror until the the first order beam retraces its path to the Bragg cell. This beam propagates to the detectors with a shift in optical frequency equal to twice the Bragg cell frequency, and is termed the common reference beam.

Detector assembly:

12. Add the second beamsplitter and the spatial filters to each resulting detector arm. The beamsplitter may or may not be polarising, depending on the method of isolating the interferometers, as discussed in Section 3.1. One of the spatial filters is simply an aperture matched to the hole in the objective lens, to isolate the reference interferometer. The second is the negative version of this (i.e. an opaque spot deposited on a glass slide) and isolates the probe interferometer. These are essential if the second beamsplitter is non-polarising, but are only required for added safety if the interferometers are separated according to polarisation.
13. Add the detectors and the collector lenses and align until the two AC signals are located.
14. Maximise these signals by making small adjustments to the optics. Do not attempt to adjust all components as this may result in irretrievable misalignment. The following are recommended:
Bragg angle, reference mirror tilt, sample tilt, focus, detector assembly.

A3.2 Eliminating crosstalk

The importance of eliminating crosstalk between the two interferometers has been demonstrated in Section 4.3. The following procedure is suggested as a rigorous method of removing crosstalk.

(i) Crosstalk in the probe interferometer

Defocus the sample and examine the probe signal on a spectrum analyser or oscilloscope. Any crosstalk from the reference interferometer will not change with defocus and will produce a small residual signal when the true probe signal is entirely lost. This may be removed by adjusting the lateral position of the appropriate spatial filter (i.e. the spot). If this adjustment proves ineffectual and assuming the signal has been attributed crosstalk (as opposed to electrical breakthrough from the Bragg cell drive) the diameter of the spot must be increased. When the crosstalk has been removed refocus the sample until the probe signal is maximised.

(ii) Crosstalk in the reference interferometer

The reference interferometer is more susceptible to sample tilt than the probe interferometer and this provides a useful means of distinguishing any probe signal present with the reference signal. Tilt the sample slightly and refocus by maximising the probe signal. Continue this procedure in small steps to avoid severe misalignment, until the reference interferometer is entirely lost, leaving the crosstalk from the probe interferometer. If this residual signal is from the probe interferometer, it will be sensitive to defocus. As before, adjustments may now be made to the lateral position of the appropriate spatial filter (i.e. the aperture) to eliminate the crosstalk. If this does not help, the diameter of the aperture must be reduced. Once this has been achieved, regain the reference signal by adjusting the sample tilt and refocus the sample to maximise the probe signal.

CHAPTER 4

FACTORS AFFECTING INSTRUMENT PERFORMANCE

4. Factors affecting the performance of the instrument

This chapter is primarily concerned with factors limiting the sensitivity of the phase measurement, since at present it is much worse than the fundamental limit normally set by electrical noise. Section 4.1 determines this limit when the measurement is limited by shot noise, whereas Section 4.2 considers the practical limit set by microphonics and thermodynamic fluctuations. The next two sections are devoted to the absolute accuracy of the measurement rather than its stability. This is affected by crosstalk as discussed in Section 4.3 and by sample tilt considered in Section 4.4. Finally, Section 4.5 consists of a brief discussion of some effects that have been observed in experiments, which have been attributed to the laser source.

4.1 Electrical noise

In this section the effect of electrical noise on the two interferometer signals is examined. An expression is derived for the minimum resolvable sample phase structure, assuming the system is shot noise limited. This is then compared to the value obtained for a basic heterodyne interferometer.

In the presence of noise, the expressions for the probe and reference signals from Chapter 3 become

$$\frac{I_p}{\gamma G_p} = C_p + E_p \cos(2\omega_B t + \psi_p) + i_p(t) \tag{4.1}$$

$$\frac{I_r}{\gamma G_r} = C_r + E_r \cos(2\omega_B t + \psi_r) + i_r(t) \tag{4.2}$$

where

$$C_n = A_n^2 + B_n^2 \tag{4.3}$$

$$E_n = 2A_nB_n \tag{4.4}$$

$$\gamma = \frac{\eta e}{h\nu} \tag{4.5}$$

$$\psi_n = (\phi_c - \phi_n) + (\theta_2 - \theta_1) \tag{4.6}$$

$$\begin{aligned} i_p(t) &= x_p \cos(2\omega_B t + \psi_r) + y_p \sin(2\omega_B t + \psi_r); \\ i_r(t) &= x_r \cos(2\omega_B t + \psi_p) + y_r \sin(2\omega_B t + \psi_p) \end{aligned} \tag{4.7}$$

and the subscript n refers to either the probe (p) or reference (r). G_n is an internal gain factor, associated with each photodetector; $i_n(t)$ represents the band-limited rms noise current in each signal. $x_n(t)$ and $y_n(t)$ are the amplitudes of the in phase and quadrature noise components respectively. The noise has been resolved in this manner for mathematical convenience. Clearly, $x_n(t)$ and $y_n(t)$ have identical statistical properties, so the choice is arbitrary.

The phase difference ϕ between the two signals is extracted using the method of coherent detection.[50] Normally, this involves the non linear mixing of a given signal with a local oscillator signal of the same frequency. The mixer is followed with a baseband filter which rejects the sum frequency, thus isolating the lower sideband at DC which contains the phase information.

In our case, the local oscillator signal (which has negligible noise in a conventional coherent detection arrangement) is replaced by the signal from the reference interferometer and thus the associated noise must be considered. The signal after the mixer is proportional to the product of Eqns. (4.1) and (4.2) :

$$\begin{aligned} \frac{I_pI_r}{\alpha} &= E_pE_r \cos(2\omega_B t + \psi_p)\cos(2\omega_B t + \psi_r) \\ &\quad + E_p i_r \cos(2\omega_B t + \psi_p) + E_r i_p \cos(2\omega_B t + \psi_r) \end{aligned} \tag{4.8}$$

where α is the proportionality constant and the product of the two noise

[50] J.Brown and E.V.D.Glazier, "*Telecommunications*", Chapman &Hall(1964).

terms has been neglected. Expanding Eqn. (4.8) using (4.7) gives

$$\begin{aligned}\frac{I_p I_r}{\alpha} &= \tfrac{1}{2}E_p E_r \{\cos(4\omega_B t + \psi_p + \psi_r) + \cos(\psi_p - \psi_r)\} \\ &+ \tfrac{1}{2}E_p \{x_r [\cos(4\omega_B t + 2\psi_p) + 1] + y_r \sin(4\omega_B t + 2\psi_p)\} \\ &+ \tfrac{1}{2}E_r \{x_p [\cos(4\omega_B t + 2\psi_r) + 1] + y_p \sin(4\omega_B t + 2\psi_r)\} \end{aligned} \quad (4.9)$$

Using Eqn. (4.6) the signal after the baseband filter may be written as

$$\frac{(I_p I_r)_{LP}}{\alpha} = \tfrac{1}{2}E_p E_r \cos(\phi_p - \phi_r) + \tfrac{1}{2}E_p x_r + \tfrac{1}{2}E_r x_p \quad (4.10)$$

The first term is the required DC output and the last two are the noise terms. In a practical coherent detector, such as a lock-in amplifier, part of the reference signal is phase shifted by $\pi/2$ and applied, together with part of the probe signal, to a second mixer/baseband filter stage. This provides a second signal given by

$$\frac{(I_p I_r)'_{LP}}{\alpha} = \tfrac{1}{2}E_p E_r \sin(\phi_p - \phi_r) + \tfrac{1}{2}E_p y_r + \tfrac{1}{2}E_r y_p \quad (4.11)$$

The two signals given by Eqns. (4.10) and (4.11) are then divided to give $\tan(\phi_p\text{-}\phi_r)$ and this removes the dependence on E_p and E_r. It is assumed here that this procedure does not significantly degrade the signal to noise ratio. Normally, the phase differences to be measured are small and the approximation $\tan(\phi_p\text{-}\phi_r) \approx \phi_p\text{-}\phi_r$ may be made. Considering the noise terms in more detail, we may express the total mean square noise current as

$$x_n^2 + y_n^2 = 2x_n^2 = 2eI_o B + \frac{4k_B T_R BF}{R_L} \quad (4.12)$$

where the first term represents shot noise and the second term Johnson (thermal) noise increased by the preamplifier noise. I_o is the time average of the current flowing in the photodiode; B is the measurement bandwidth; F is the noise figure of the photodiode preamplifier; k_B is

the Boltzmann constant; R_L is the photodiode noise resistor and T_R is the load-resistor-noise temperature. Since the signals are detected at frequencies in the MHz range, $1/f$ noise is negligible.

From the above equation it is clear that the Johnson noise may be minimised by choosing R_L to be large (200 Ω say), and by using suitable low noise preamplifiers. Under these conditions, if the laser power is above 100 μW, the Johnson noise may be neglected. In the absence of further limitations such as phase noise due to microphonics, the phase measurement is said to be shot noise limited. The signal to noise ratio is then given by

$$S/N = \frac{\left[\frac{1}{2}\gamma^2 E_p E_r G_p G_r (\phi_p - \phi_r)\right]^2}{eB\left[(\frac{1}{2}\gamma E_p G_p)^2 \gamma C_r G_r^2 + (\frac{1}{2}\gamma E_r G_r)^2 \gamma C_p G_p^2\right]} \tag{4.13}$$

Using the relation $(\phi_p-\phi_r)=2\,\delta\theta$ where $\delta\theta$ represents the sample phase structure, the above equation becomes

$$S/N = \frac{\gamma(4\delta\theta A_p B_p A_r B_r)^2}{eB\left[(A_p B_p)^2 (A_r^2 + B_r^2) + (A_r B_r)^2 (A_p^2 + B_p^2)\right]} \tag{4.14}$$

For optimum performance, the total laser power available should be divided equally between each of the interfering beams and therefore

$$A_p^2 = B_p^2 = A_r^2 = B_r^2 = P \tag{4.15}$$

The final expression for the signal to noise ratio then reduces to

$$S/N = \frac{\eta P (2\delta\theta)^2}{h\nu B} \tag{4.16}$$

The minimum detectable phase change is determined for a signal to noise ratio of unity, and therefore

$$\delta\theta_{min} = \frac{1}{2}\sqrt{\frac{h\nu B}{\eta P}} \qquad (4.17)$$

This is a factor of root two worse than the sensitivity of a basic heterodyne interferometer, assuming the same amount of power is focussed onto the sample surface. However, in order to achieve this condition, twice as much total laser power is required to cater for the additional reference interferometer. Normally, the total laser power available is fixed and in this instance, the common path system is a factor of two worse than the basic heterodyne interferometer, in terms of shot noise limited performance. The justification for incorporating a second interferometer is that in practice, the basic heterodyne interferometer is not shot noise limited. This is due to its greater sensitivity to phase errors caused by microphonics.

A typical value for the quantum efficiency of the photodiode is about 0.73 for a laser wavelength of 632.8 nm. The power in each of the four interfering beams is normally about 0.5 mW. Thus in a typical measurement bandwidth of 1 kHz, the phase resolution evaluated from Eqn. (4.17) is about 4.64×10^{-7} rad, equivalent to 2.34×10^{-4} Å in terms of topography.

4.2 Noise generated in the optical system

In this section an investigation is made into the ability of the system to reject phase noise normally encountered in practical laboratory set up. In essence, it forms a study of the degree to which the system is common path. No system of this type can be perfectly common path, since it would be insensitive to the sample phase structure as well as the phase noise. Section 4.2.1 is concerned with microphonics in the optical components of the system, whereas section 4.2.2 deals with the effect of air turbulence causing refractive index variations in the propagation paths.

4.2.1 Microphonics

Throughout most of the optical system piston type microphonics (i.e. movement in the direction of beam propagation) do not cause a problem, since both interferometers suffer identical phase distortion. There are however, a number of elements for which vibration introduces phase errors which may not be common to both interferometers. The relative importance of each case is considered as follows:

(i) Vibration of each detector lens and photodiode.

Vibration of the components forming each detector assembly does not pose a problem despite the fact that the interferometers follow different paths at this point. The reason is that each interferometer is common path in these regions, so that the noise is removed optically. In any event, the detector assembly could be made common to both interferometers by replacing the final beamsplitter and the two detector arrangements with a single specially constructed bi-cell detector. This detector would have a central active region for the reference interferometer, surrounded by a separate annular active region for the probe interferometer. Care would have to be taken however, to ensure that crosstalk between the interferometers is eliminated, and this is considered in Section 4.3.

(ii) Vibration of the sample.

This is an important problem to consider if the sample is mechanically scanned. The system has been designed to give a relatively high degree of immunity to piston type movement of the sample. However, due to the obliquity effect of the rays focussed onto the sample, the probe interferometer will suffer a slightly different phase error than the reference interferometer, which illuminates the sample normally. If the sample is displaced axially a distance δx, the reference

interferometer will register a phase change θ_r given by

$$\theta_r = 2k\ \delta x \tag{4.18}$$

Due to the obliquity effect each plane wave component of the sample probe beam will suffer a slightly smaller phase change θ_p, which may be calculated by reference to Fig. 4.1. The phase of the in focus ray is compared to that observed when the sample is displaced a distance δx from the focal plane. The resultant phase change, corresponding to the path length AB in Fig. 4.1(b), is given by

$$\theta_p = 2k\ \delta x \cos\alpha \tag{4.19}$$

where α is the illumination angle of the plane wave component. In order to obtain the overall phase, the weighted average over all plane wave components of the sample probe beam must be calculated. The desired value is well known in micro-interferometry as the NA factor, so named because it is a function of the numerical aperture of the objective. In its usual context, the NA factor (F_{NA}) is a coefficient which multiplies the measured phase to correct for the oblique illumination angle. It is therefore experimentally defined as

$$F_{NA} = \frac{\phi_{act}}{\phi_{meas}} \tag{4.20}$$

where ϕ_{meas} is the measured phase and ϕ_{act} is the actual value. Since oblique illumination will make the measured phase variation systematically smaller, the NA factor is clearly always greater than unity. For the probe interferometer ϕ_{act} and ϕ_{meas} are given by

$$\phi_{act} = 2k\ \delta x = \theta_r \tag{4.21}$$

$$\phi_{meas} = (\theta_p)_{AV} = 2k\ \delta x\ (\cos\alpha)_{AV} \tag{4.22}$$

where $(\ )_{AV}$ indicates that a weighted average is performed over all plane

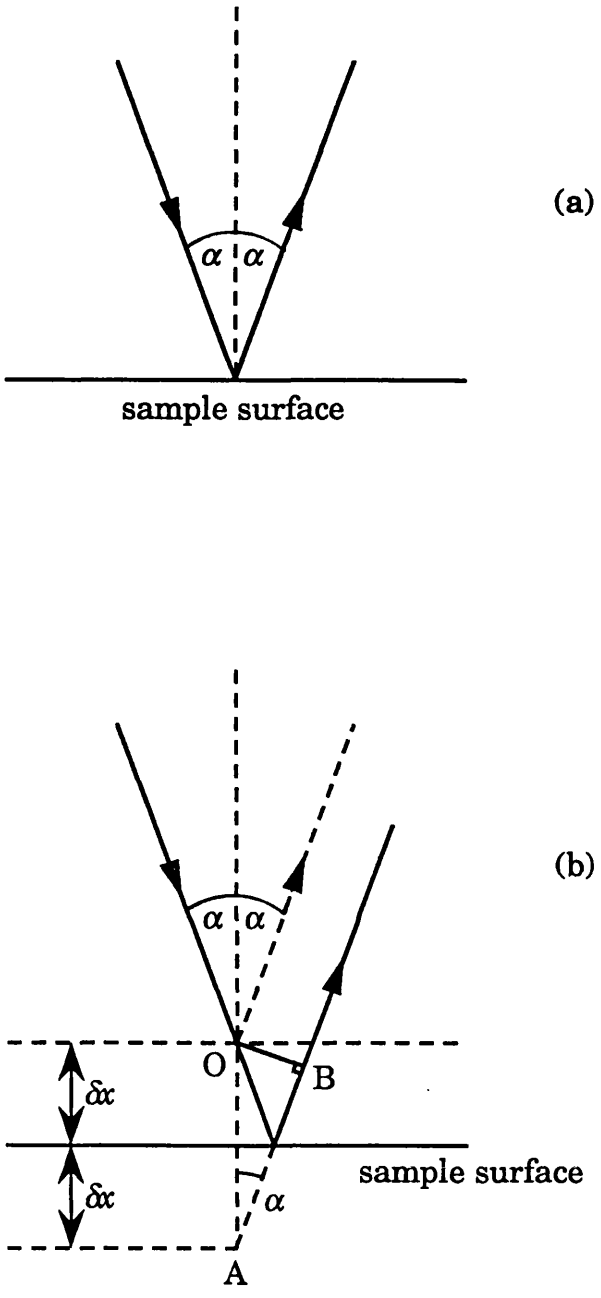


Figure 4.1. Geometry for calculation of the effect of sample vibration on the sample probe beam. (a) In focus position, (b) Sample displaced axially a distance δx.

wave components of the beam. Thus it follows that

$$F_{NA} = \frac{1}{(\cos \alpha)_{AV}} \tag{4.23}$$

The value $(\cos \alpha)_{AV}$ (and thus F_{NA}) is derived for an annular lens in Appendix A4.1. The phase error ϕ_E introduced by a lens displacement δx is therefore

$$\phi_E = \theta_r - (\theta_p)_{AV} \tag{4.24}$$

which leads to the result

$$\phi_E = 2k\ \delta x \left[1 - \frac{1}{F_{NA}}\right] \tag{4.25}$$

The fraction (μ) of the vibrational amplitude that the system is capable of rejecting is therefore given by

$$\mu = \frac{\phi_{act} - \phi_E}{\phi_{act}} = \frac{1}{F_{NA}} \tag{4.26}$$

The NA factor calculated in Appendix A4.1 is plotted as a function of numerical aperture in Fig. 4.2. The annulus parameter ε, defined as the ratio between the lens hole radius to that of the full aperture, was equal to 0.25 for this plot. For a 0.1 NA objective, F_{NA} is equal to 1.003, giving a value for μ of 0.997 (i.e. 99.7% of the vibration of this objective is rejected). For a 0.5 NA objective, F_{NA} is 1.083 and μ is then equal to 0.923. The simple ray theory approach used to calculate F_{NA} in Fig. 4.2 breaks down at high numerical apertures, as indicated by the asymptotic rise to infinity as the numerical aperture tends to unity. The values for F_{NA} can be taken as valid up to numerical apertures of about 0.5. This assertion is based on experimentally determined values for F_{NA} by Creath[51].

[51] K.Creath, "Calibration of numerical aperture effects in interferometric microscope objectives", Appl. Optics, **28**, (15), pp. 3333-3338 (1989).

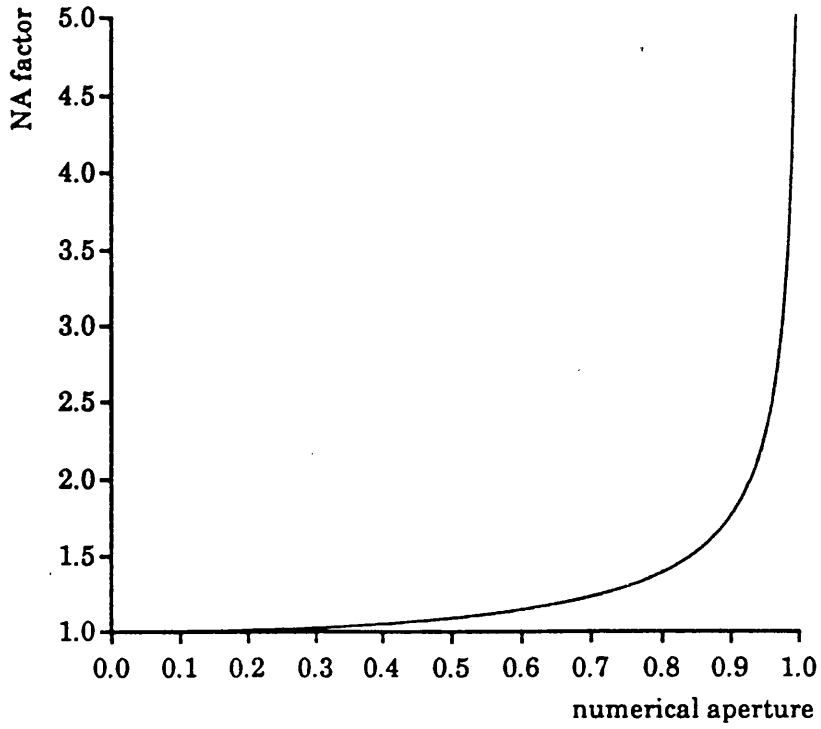


Figure 4.2. NA factor for an annular lens with ε=0.25 as a function of numerical aperture.

For numerical apertures above 0.5, the NA factor may be determined by comparing measurements of standard step heights with their true values according to Eqn. (4.20). As a guide to the value of F_{NA} at high NAs we use the value obtained for a conventional objective measured by Creath. The value quoted for a 0.95 NA objective is 1.225 which gives a value for μ of 0.816.

The dependence of the vibration rejection on numerical aperture is therefore quite significant and for quantitative measurements of step heights it is probably better to sacrifice lateral resolution and use a low NA objective. This assumes that the height difference could be taken between two points some distance from the edge and that the quality of the edge is not important to the user.

(iii) Vibration of the objective lens.

The nature of the objective lens is such that piston type vibration only affects the sample probe beam. The sample reference beam passes through the hole in the lens and is therefore unaffected if the lens moves axially. If the lens is displaced axially a distance δx towards the sample, the average optical path length of the sample probe beam between the lens and the sample will decrease by $2\delta x/F_{NA}$. The path length on the other side of the lens will, however, increase by $2\delta x$. The net path length change ΔOPL experienced by the sample probe beam is therefore given by

$$\Delta OPD = 2\,\delta x \left[1 - \frac{1}{F_{NA}}\right] \tag{4.27}$$

Since the phase of the sample reference beam remains unchanged, the the above path length change produces a measured phase error identical to that given by Eqn. (4.25). The system is thus equally immune from microphonics arising from vibration of the sample and the lens. It

is therefore equally important to suppress vibration in both the objective lens and the sample, though the latter effect will undoubtedly dominate as long as mechanical scanning is employed. An objective has been proposed which overcomes the problem of vibration in both components, by producing a sample reference beam which illuminates at the same angle as the average illumination angle of the sample probe beam. This is discussed in detail Chapter 7.

The ideas discussed in this section lead to a simple condition which must be met for a system to be able to reject the effects of piston. There must be at least two beams propagating through all components. As far as possible, the configuration should be arranged so that both beams are affected equally by the vibration. If the two beams are from the same interferometer, the phase noise is removed optically by the common path effect. If the beams are from separate interferometers, the phase noise is detected and then removed electronically by subtraction.

This section has been devoted to the effects of piston type vibration of components, since this has a direct effect on the phase of the beams. The optics may also vibrate in a direction perpendicular to beam propagation or tilt relative to the optic axis. Lateral movement will not affect the phase of the beams, and tilt of most components will only cause a lateral shift of the beam. A possible problem caused by such movement could be significant tilting motion of the sample, introduced by imperfect mechanical scanning. This is discussed in Section 4.4.

4.2.2 Air turbulence

This section deals with the effect of refractive index variations in the propagation path on system sensitivity. This is important, since the probe and reference interferometers do not spatially overlap. Therefore, phase errors due to turbulence will not be entirely common to both interferometer signals. Two cases are considered in this section. Firstly, an investigation is made of the fundamental limit set by

thermodynamic fluctuations. It is important to note the distinction between this effect and non equilibrium temperature changes, which are the subject of the second part of this section. The two phenomena are therefore referred to as fluctuations and turbulence respectively, for the sake of clarity.

Assuming that the propagating medium (in this case air) is in equilibrium, there will always be fundamental fluctuations in the thermodynamic quantities that specify its state. The approach used to quantify this effect is based on the work by Glenn,[52] which considers the effect of thermodynamic fluctuations in optical fibre based interferometers. The state of the propagating medium is normally described in terms of two quantities such as volume and temperature or entropy and pressure. Considering statistical fluctuations in the volume and temperature, the resulting change in refractive index of the medium is given by

$$\Delta n = \left(\frac{\partial n}{\partial V}\right)_T \Delta V + \left(\frac{\partial n}{\partial T}\right)_V \Delta T \tag{4.28}$$

The mean square fluctuation is

$$\overline{\Delta n^2} = \left(\frac{\partial n}{\partial V}\right)_T^2 \overline{\Delta V^2} + \left(\frac{\partial n}{\partial T}\right)_V^2 \overline{\Delta T^2} + \left(\frac{\partial n}{\partial V}\right)_T \left(\frac{\partial n}{\partial T}\right)_V \overline{\Delta V \Delta T} \tag{4.29}$$

The quantities $\overline{\Delta V^2}$, $\overline{\Delta T^2}$ and $\overline{\Delta V \Delta N}$ are calculated in standard texts on thermodynamics[53] and may be expressed as

$$\overline{\Delta V^2} = -k_B T \left(\frac{\partial V}{\partial P}\right)_T \tag{4.30}$$

$$\overline{\Delta T^2} = -\frac{k_B T^2}{m C_v} \tag{4.31}$$

[52] W.H.Glenn,"Noise in interferometric optical systems: an optical Nyquist theorem", *I*EEE J. Quant. Elect., **25**, (6), pp. 1218-1224 (1989).

[53] L.D.Landau, E.M.Lifshitz, "*Statistical Physics*", Part 1, 3rd Ed., p. 338, Pergamon Press (1980).

$$\overline{\Delta T \Delta V} = 0 \tag{4.32}$$

Where k_B is the Boltzmann constant, m is the total mass of the medium, C_v is its specific heat at constant volume and P is the pressure. Eqn. (4.32) demonstrates that the volume and temperature fluctuations are uncorrelated. Since the volume V is fixed, determined by the volume occupied by the optical beams, it is more relevant in this instance to consider fluctuations of the number of particles within the volume. Dividing both sides of Eqn. (4.30) by N^2, where N is the average number of particles, we obtain

$$\overline{\left(\Delta \frac{V}{N}\right)^2} = -\frac{k_B T}{N^2}\left(\frac{\partial V}{\partial P}\right)_T \tag{4.33}$$

Since V is constant, we may write

$$\Delta \frac{V}{N} = V \Delta \frac{1}{N} = -\frac{V}{N^2}\Delta N \tag{4.34}$$

Substituting this into Eqn. (4.33) gives

$$\overline{\Delta N^2} = -\frac{k_B T N^2}{V^2}\left(\frac{\partial V}{\partial P}\right)_T \tag{4.35}$$

Making the assumption that the medium obeys the ideal gas law

$$PV = N k_B T \tag{4.36}$$

Eqn. (4.35) reduces to

$$\overline{\Delta N^2} = N \tag{4.37}$$

and thus Eqn. (4.29) may now be reformulated as

$$\overline{\Delta n^2} = \left(\frac{\partial n}{\partial N}\right)_T^2 N + \left(\frac{\partial n}{\partial T}\right)_N^2 \frac{k_B T^2}{m C_v} \tag{4.38}$$

In order to evaluate the partial derivatives in the above expression, we make use of the approximate relation for the refractive index of a gas[54]

$$n \approx 1 + \frac{N}{V} \cdot \frac{e^2}{2\pi m_e (f_0 - f)} \tag{4.39}$$

Where e and m_e are the electronic charge and mass, f_0 is their natural frequency of oscillation, and f is the optical frequency. Differentiating with respect to N we obtain

$$\left(\frac{\partial n}{\partial N}\right)_T = \frac{(n-1)}{N} \tag{4.40}$$

Substituting Eqn. (4.36) for an ideal gas into Eqn. (4.39) gives

$$n \approx 1 + \frac{P}{k_B T} \cdot \frac{e^2}{2\pi m_e (f_0 - f)} \tag{4.41}$$

Differentiating with respect to T and using Eqns. (4.36) and (4.39), we obtain

$$\left(\frac{\partial n}{\partial T}\right)_N = \frac{-(n-1)}{T} \tag{4.42}$$

Substituting Eqns. (4.40) and (4.42) into (4.38), the mean square fluctuation in the refractive index becomes

$$\overline{\Delta n^2} = (n-1)^2 \left[\frac{1}{N} + \frac{k_B}{m C_v}\right] \tag{4.43}$$

In order to evaluate this expression, the average number and total mass of the molecules with the volume V of the laser beams must be determined. According to Avogadro's Law we may write

---

[54] N.H.Frank, "*Introduction to Electricity and Optics*", 1st Ed., p. 305, McGraw-Hill (1940).

$$N = V \cdot \frac{N_A}{V_A} \tag{4.44}$$

Where N_A is the Avogadro's Number and V_A is the volume occupied by 1 mole of ideal gas at standard temperature and pressure (STP). In addition, the total mass m of these molecules is given by

$$m = \frac{V}{V_A} m_m \tag{4.45}$$

Where m_m is the molar mass in kilograms. Using these relations, the rms fluctuation in the refractive index may be written as

$$n_{rms} = \overline{(\Delta n^2)}^{\frac{1}{2}} = (n-1)\left(\frac{V_A}{V}\right)^{\frac{1}{2}}\left[\frac{1}{N_A} + \frac{k_B}{m_m C_v}\right]^{\frac{1}{2}} \tag{4.46}$$

Using the numerical values in Table 4.1 (at the end of this section), the above expression yields values of $1.2\text{x}10^{-14}$ and $4.5\text{x}10^{-14}$ for the fluctuation in the refractive index of the volumes occupied probe and reference signals respectively. Since the beams of the probe interferometer surround those of the reference interferometer, there will probably be some correlation between the phase errors suffered by each. Therefore, a certain proportion of the phase error due to thermodynamic fluctuations will be removed in the final phase output. For the purpose of obtaining an order of magnitude value for the limit to sensitivity, we take the worst case where the phase errors are totally uncorrelated. The refractive index fluctuation may be related to the measured phase fluctuation according to

$$\phi_{rms} = \frac{2\pi L}{\lambda} n_{rms} \tag{4.47}$$

where L is the total distance propagated. The refractive index fluctuations therefore correspond to topographical variation of about $2.3\text{x}10^{-4}$ Å. This value cannot be directly compared with the limit set by

electrical noise since it is does not include a bandwidth dependence. In fact, this value represents the fluctuations summed over the entire frequency spectrum and therefore the noise falling within the measurement bandwidth will be significantly below that of electrical noise as calculated in Section 4.1. For this reason, an attempt has not been made to calculate the spectral distribution of these fluctuations. In summary then, such fluctuations will only be a fundamental limitation when using laser powers well above the current milliwatt range, i.e. when shot noise ceases to be the limiting factor.

Non-equilibrium temperature changes represent a much more significant effect than the fundamental fluctuations discussed above. The nature of thermal turbulence depends on the environmental conditions, which, unless actively controlled tend to be highly variable. Hodera[55] has investigated the effect of air turbulence on laser propagation through the atmosphere. Although the turbulence encountered in the laboratory is less severe, the general principles described therein still apply. The approach normally used is to model the turbulent medium in terms of both a spatial and a temporal correlation function of the refractive index variation. The effect the turbulence has on the microscope is dependent on the correlation distance relative to the beam diameter. Thus, in order to give common path performance, the outer beam diameter should be well within the correlation distance, in order that both interferometers are affected equally by the turbulence. The value of the integration time (i.e. the measurement bandwidth) relative to correlation time is not as critical, since both interferometers are affected equally by temporal variations in refractive index.

As an order of magnitude calculation, we may relate a temperature change δT to the resultant change in the refractive index of the air δn by the finite differential version of Eqn. (4.42)

---

[55] H.Hodera, "Laser wave propagation through the atmosphere", Proc. IEEE (**54**), (3), pp.368-375 (1989).

$$\delta n = \frac{-(n-1)}{T}\delta T \tag{4.48}$$

From this and Eqn. (4.47), the phase error per unit change in temperature i.e. $\delta\phi_e/\delta T$ evaluates to 9.29 rad/K, using Table 4.1 and a value of room temperature of 295 K. This result does seem rather large, demonstrating that thermal effects are significant, but it should be noted that it refers to the overall difference in the average temperature of the regions occupied by each interferometer. This is not an easy measurement to perform, especially since the stability of the phase output for the first implementation (2.8 Å in terms of topography) implies that the temperature difference varies by no more than 0.6 mK. For this reason, the best way to investigate the effect of turbulence is probably an optical method. Such a direct measurement also has the advantage that it accounts for the constituents of the air, including CO_2 content and humidity.

The experiment was devised to provide a thorough analysis of the various sources of phase noise, including noise due to imperfect signal processing and microphonics, as well as noise due to thermal turbulence. The approach taken was to compare stability measurements from the instrument after making various modifications to the configuration. These are detailed as follows:

(a) The usual configuration with the drilled reflecting objective.

(b) The same as (a) but with the drilled lens singlet.

(c) The same as (a) but with no objective. This removes the effect of non common mode noise from sample vibrations due to the obliquity effect, as discussed in Section 4.2.1. Thus, by comparing the result with those of (a) and (b) the significance of sample microphonics may be investigated.

(d) The same as (c) but with the spatial filters also removed, so that both detectors view the same light field. By comparing the results with those of (c), the noise contribution due to the lack of spatial overlap (i.e. thermal effects) may be determined.

(e) Using the same signal (from the reference detector) as both the reference and the probe signal. This reveals information concerning the common mode rejection capabilities of the signal processing arrangement.

Obviously, only the configurations of (a) and (b) are useful instruments, with the remaining arrangements chosen simply to isolate the various noise sources. With configuration (d) the result is effectively that for a 100% common path interferometer, which as such would be completely insensitive to surface phase structure. Stability measurements were recorded for each of the above arrangements by monitoring the phase over a period of 30 s. The experiment was performed using the prototype microscope which forms the basis of Chapter 5. The results in terms of an rms variation in topography are shown in Table 4.2, for measurement bandwidths (BW) of 33 Hz and 1 KHz. These bandwidths are compatible with those used for line trace measurements and image acquisition respectively.

A comparison of the values of (a) through to (c) demonstrates conclusively that sample microphonics is currently the limiting factor. The progressive improvement in performance from (a) to (c) is a consequence of the obliquity effect. However, the difference between measurements (b) and (c) (especially for the 1 kHz bandwidth) seems large, considering that the vibration rejection factor μ defined by Eqn. 4.24 is 0.98 for the 0.27 NA lens singlet. The reason is probably the increase in signal to noise ratio obtained for (c) of about 20 dB, due to the absence of the aberrating effect of the objective. The difference between measurements (c) and (d) indicates that thermal effects would become significant if the obliquity problem could be overcome. A method for achieving this is proposed in Section 7.3.1. Measurement (e) represents the best sensitivity that could be achieved with the present coherent detection scheme. There is a significant discrepancy between the values of (d) and (e), bearing in mind that for (d), both detectors are illuminated by the same light field. This indicates that a significant non common

mode noise term originates from the use of separate detectors and amplifiers. The noise may be introduced the components themselves, but could also be caused by each detection arrangement having different transfer functions. The latter effect would cause the optically generated noise terms in each signal to differ slightly, thus degrading the performance of the subtraction process.

Symbol	Quantity	Value
n	Refractive index of air at 630 nm	1.0002760
m_m	Mass of 1 mole of air	2.8964×10^{-2} Kg
$C_v=C_p$	Specific heat at constant volume	240.1 $JK^{-1}Kg^{-1}$
k_B	Boltzmann constant	1.3807×10^{-23} JK^{-1}
V_A	Volume of 1 mole of ideal gas at STP	2.24136×10^{-2} m^3
N_A	Avogadro's Number	6.02217×10^{23}
V_r	Volume of reference interferometer[56]	3.1×10^{-6} m^3
V_p	Volume of probe interferometer	1.3×10^{-6} m^3

Table 4.1. Data for evaluating the thermodynamic fluctuation of the refractive index of air.[57,58]

Configuration	rms stability (Å)	
	33 Hz BW	1 kHz BW
(a) reflecting objective	4.0	8.6
(b) singlet objective	1.4	5.3
(c) objective removed	0.8	2.2
(d) spatial filters also removed	0.6	1.7
(e) reference signal also used as probe signal	0.3	1.4

Table 4.2. Results of an experiment to determine the relative significance of the various sources of phase noise in the interferometer.

[56] The volumes occupied by the probe and reference interferometers were calculated assuming that the total path length is 1 m and that the laser beam fills an 8 mm diameter aperture objective lens, with a 2 mm diameter central hole.

[57] R.C Weast,"*Handbook of Chemistry and Physics*", CRC Press (1989).

[58] C.Nordling, J.Osterman, "*Physics Handbook*", Chartwell-Bratt (1982).

4.3 Crosstalk between the interferometers

In this section the effect of imperfect isolation of the two output signals from the microscope is considered. This may occur if light from the probe interferometer reaches the reference detector, or if light from the reference interferometer reaches the probe detector. The first section considers the effect this has on the phase output. Section 4.3.2 is primarily concerned with what has been termed the "gap effect", which aids spatial separation of the interferometers. Further experimental observations relevant to the discussion of crosstalk are presented in the latter section, and these have led to some unexpected general conclusions as to the nature of heterodyne interference.

4.3.1 The effect of crosstalk on the phase measurement

In Chapter 3, the probe and reference signals (I_p and I_r respectively) were represented by

$$\frac{I_p}{\gamma G} = A_p^2 + B_p^2 + 2A_pB_p \cos\left[2\omega_B t + (\phi_c - \phi_p) + (\theta_2 - \theta_1)\right] \tag{4.49}$$

$$\frac{I_r}{\gamma G} = A_r^2 + B_r^2 + 2A_rB_r \cos\left[2\omega_B t + (\phi_c - \phi_r) + (\theta_2 - \theta_1)\right] \tag{4.50}$$

Where the DC terms may be ignored if the detectors are AC coupled. In the presence of crosstalk, the distorted signals may be written as

$$I_p' = I_p + \xi_r I_r \tag{4.51}$$

$$I_r' = I_r + \xi_p I_p \tag{4.52}$$

where ξ_p and ξ_r are the fractions of the two misplaced interference terms reaching each detector. If a 50:50 beamsplitter together with the appropriate spatial filters are used to isolate the two interferometers, ξ_p

and ξ_r may take values between zero and unity. These extremes correspond to perfect isolation and no isolation respectively. The pure signals I_p and I_r are represented in a phasor diagram in Fig. 4.3(a), and exhibit the true phase difference ϕ, where $\phi=\phi_p-\phi_r$. Fig. 4.3(b) shows how the phase is distorted to ϕ' due to crosstalk. In order to calculate ϕ' we must first find the resultant phase of I_p' and I_r'. Considering Fig. 4.4(a), for I_p' we may write

$$\tan\phi_p' = \frac{\xi_r I_r \sin\phi}{(I_p + \xi_r I_r \cos\phi)} \tag{4.53}$$

Where ϕ_p' is the phase error introduced into the probe signal. Similarly, from Fig. 4.4(b), the phase error in the reference signal (ϕ_r') is given by

$$\tan\phi_r' = \frac{\xi_p I_p \sin\phi}{(I_r + \xi_p I_p \cos\phi)} \tag{4.54}$$

The total phase error ϕ_e is thus

$$\phi_e = \phi_p' + \phi_r' \tag{4.55}$$

and the measured phase difference is reduced to ϕ', where

$$\phi' = \phi - \phi_e \tag{4.56}$$

In order to minimise shot noise I_p and I_r are made equal (see Section 4.1) Eqn. (4.55) then becomes

$$\phi_e = \tan^{-1}\left[\frac{\xi_p \sin\phi}{(1+\xi_p\cos\phi)}\right] + \tan^{-1}\left[\frac{\xi_r \sin\phi}{(1+\xi_r\cos\phi)}\right] \tag{4.57}$$

If I_p and I_r are equal, Eqns. (4.53) and (4.54) vary in the same manner with ξ_p and ξ_r respectively. This is illustrated graphically in Fig. 4.5, where the phase error has been normalised to the true phase ϕ and

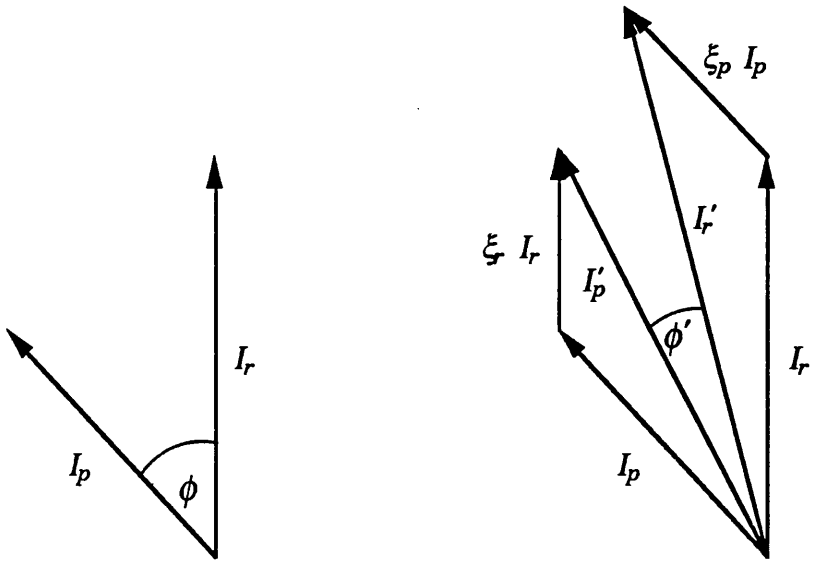


Figure 4.3. Phasor diagram of the probe and reference signals; (a) with perfect isolation, and (b) in the presence of crosstalk.

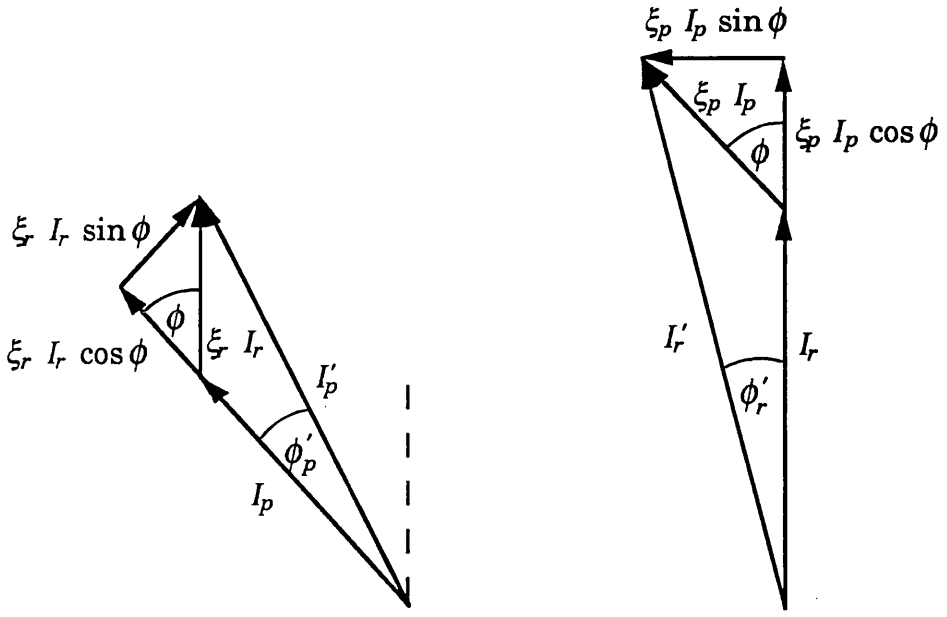


Figure 4.4. Phase error introduced by crosstalk; (a) in the probe interferometer, and (b) in the reference interferometer.

plotted against ξ_p or ξ_r. As either ξ_p or ξ_r tend to unity the resulting phase error tends to $\phi/2$. Therefore, if there is zero isolation in both detectors (i.e. spatial filters removed), the total phase error is ϕ and from Eqn. (4.56) the measured phase difference reduces to zero. This is as expected since the detectors will produce identical output signals in this case, giving no sensitivity to sample phase structure.

For a given degree of crosstalk, the phase error is also dependent on the true phase difference ϕ between the two interferometer signals. This illustrated in Fig. 4.6, where the phase error in either the probe or reference signals is plotted as a function of ϕ. The maximum phase error occurs when $\phi=\pi/2$, and for the probe interferometer it is given by

$$(\phi_p')_{max} = \tan^{-1}\xi_p \tag{4.58}$$

Normally the crosstalk is small (i.e. $\xi_p<<1$) and thus

$$(\phi_p')_{max} \approx \xi_p \tag{4.59}$$

A similar expression may be written for $(\phi_r')_{max}$ for the reference signal. Therefore, a small amount of crosstalk in both interferometer signals, results in an overall maximum phase error of

$$(\phi_e)_{max} \approx \xi_p + \xi_r \tag{4.60}$$

It is this final expression which is most useful in determining the maximum degree of crosstalk that can be tolerated, in a practical system.

Crosstalk in either the probe or the reference detector will cause any quantitative measurements to be systematically lower than the true value. In an imaging application, the presence of crosstalk results in a reduction in contrast. It is therefore imperative that the degree of

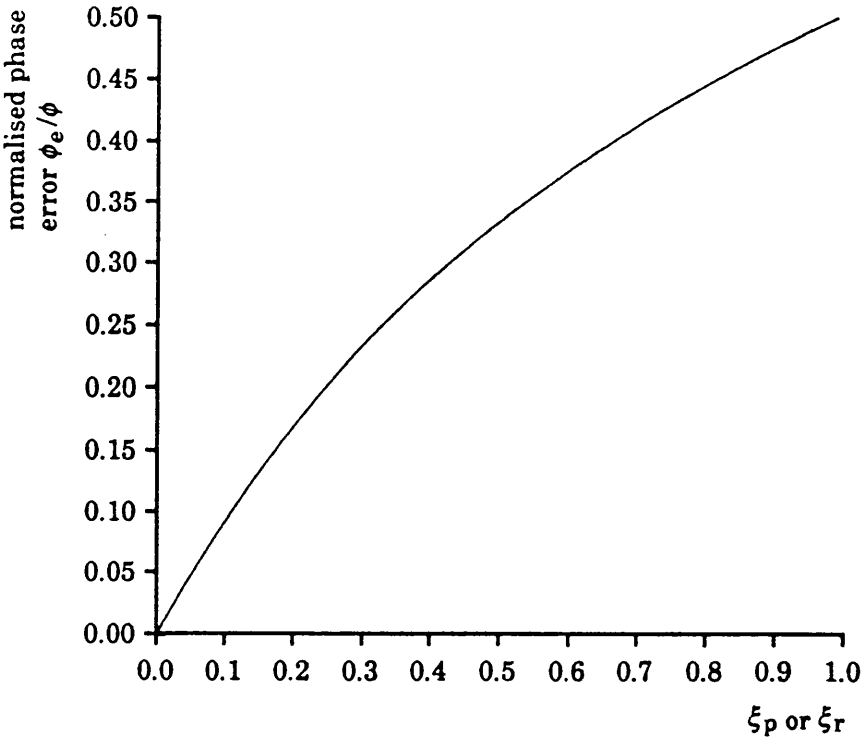


Figure 4.5. Phase error due to crosstalk in the probe or reference signal normalised to ϕ as a function of ξ_p or ξ_r.

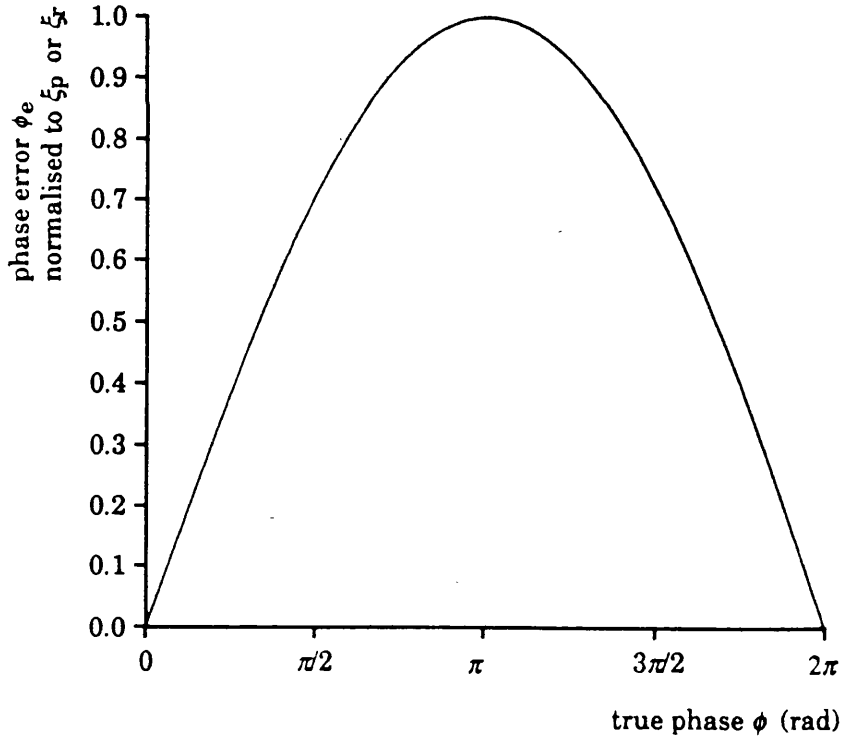


Figure 4.6. Phase error due to crosstalk in the probe or reference signal, normalised by dividing by ξ_p or ξ_r (assuming ξ_p or $\xi_r << 1$), plotted as a function of the true phase difference ϕ.

crosstalk is kept to a minimum. It may be largely eliminated by placing a guard ring (i.e. an annular field stop) around the hole in the lens, to ensure spatial separation of the interferometers. However, this reduces the power in the probe beam and is thus offset against a reduction in signal to noise ratio.

4.3.2 Observations and analysis concerning the degree of crosstalk

In practice, it was found that a guard ring was not necessary, since the returning sample probe and reference beams were separated by an annular gap when viewed in cross section. The origin of this gap has been investigated in detail for a lens singlet, using ray tracing, and has been attributed to spherical aberration. The effect of spherical aberration of a lens is best illustrated by plotting its caustic surface for collimated incident rays. This is a graph of lateral distance of the ray from the optic axis against its focal distance. The caustic for the drilled singlet used in the first implementation is shown in Fig. 4.7, and illustrates how the rays focus to progressively shorter distances at the extremities of the lens aperture. The dashed line represents the radius of the central hole.

In order to trace the rays to the sample and back to the detector plane, a position of best focus for the sample is required. Experimentally, the position of best focus is located by maximising the detected interference signal. At this position, the loss of interference from out of focus rays is minimised. The best focus was taken as the average of the paraxial focus and the minimum focal length for the ray entering at the extreme of the aperture. Since the range of focal lengths of the probe beam rays is given by the caustic between h_a and h_b, this average will tend to bias the rays entering the lens closer to the optic axis. This what is required, since the Gaussian beam profile and the slower variation of the caustic near the optic axis will tend to favour the more central rays. In this instance, the position of the so called circle of least confusion is not a

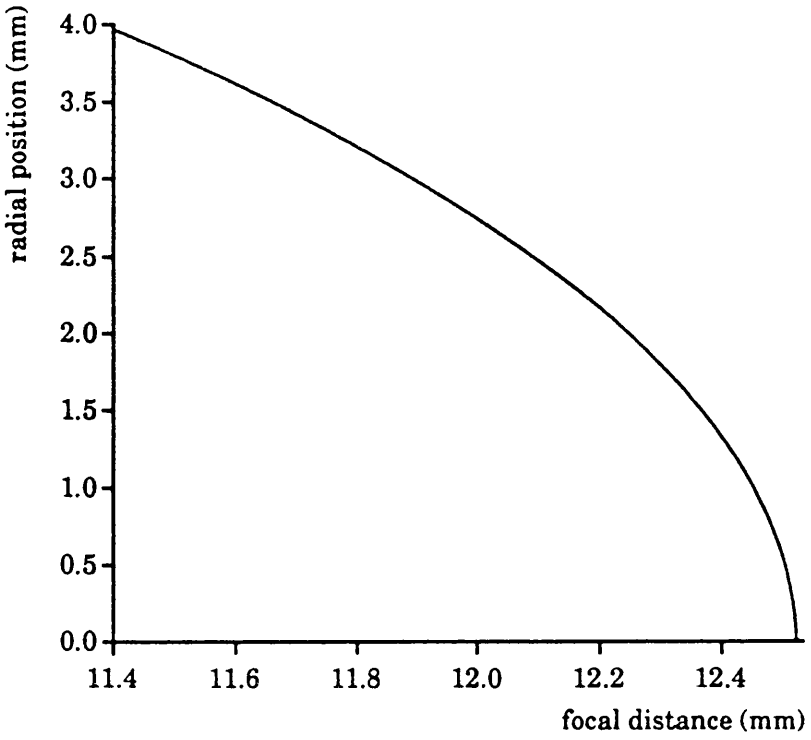


Figure 4.7. Caustic surface for the spherically aberrated lens used in the first implementation.

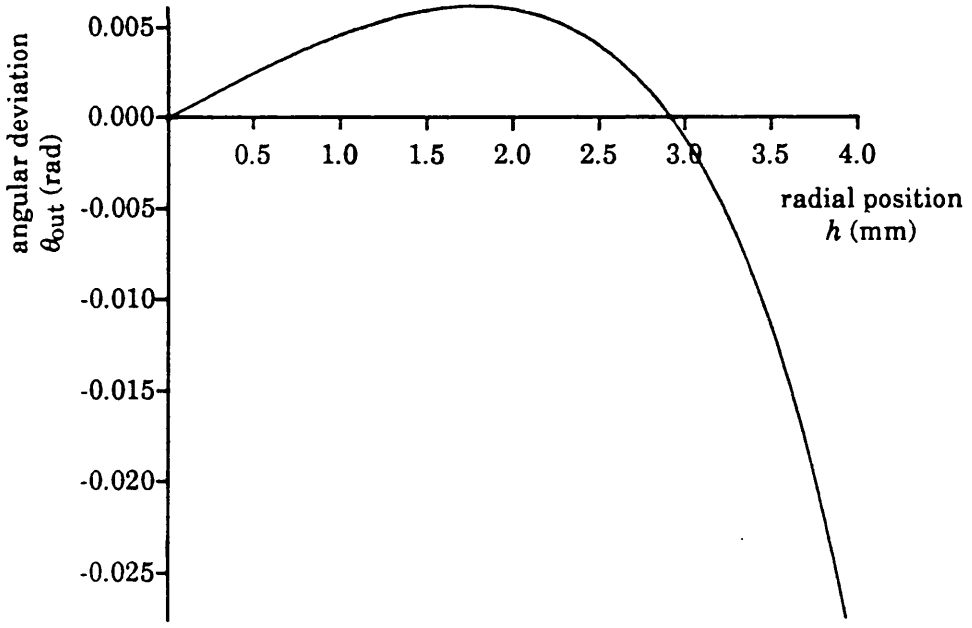


Figure 4.8. Angular deviation of the probe beam rays propagating to the detector as a function of the distance that these rays entered the lens from the optic axis.

good estimate of the position of best focus, since this would tend to bias the outermost rays. Obviously, the value taken for the best focus is approximate, nonetheless it still provides sufficient accuracy to preserve the pertinent information concerning the formation of the annular gap.

The results of the ray trace are revealed in Fig. 4.8, where the angle to the optic axis with which the rays propagate to the detector plane (θ_{out}) is plotted against the distance h that these rays entered the lens from the optic axis. As in Fig. 4.7, the dashed line represents the radius of the central hole in the lens. There are two incident rays for which θ_{out} is zero, corresponding to a collimated returning ray. These are for the central ray at h=0, which as expected remains undeviated and for the ray which focuses to the average of the focal lengths used as the best focal position for the sample. This latter ray which enters the lens at h_{av} say, is in fact, the only one which is exactly in focus at the sample. Those rays entering the lens above h_{av} return convergent to the detector plane. The gradient of the graph becomes progressively steeper in this region as h increases, since the outer rays suffer greater aberration. Below h_{av}, the returning rays are divergent, and the gradient reduces, as the spherical aberration decreases. This gradient eventually passes through a stationary point and θ_{out} reduces to zero for the on axis ray.

Since the lens has a central hole, only the region $h>h_a$[59] is relevant for the probe beam, but considering all rays helps to explain the form of this graph. The divergent rays for $h<h_{av}$ are responsible for the observed gap between the sample probe beam and the sample reference which passes through the hole. These effects are illustrated more clearly on the ray diagram of Fig. 4.9, which shows the arrangement in transmission for clarity. This diagram also illustrates a less significant effect contributing to the gap, whereby the innermost rays strike the drilled surface of the hole after reflection from the sample, and are

[59] In fact the lower limit on h is slightly greater than h_a, since light entering the lens near the hole is refracted into the drilled surface. Thus, even in the absence of spherical aberration there will be a residual gap effect.

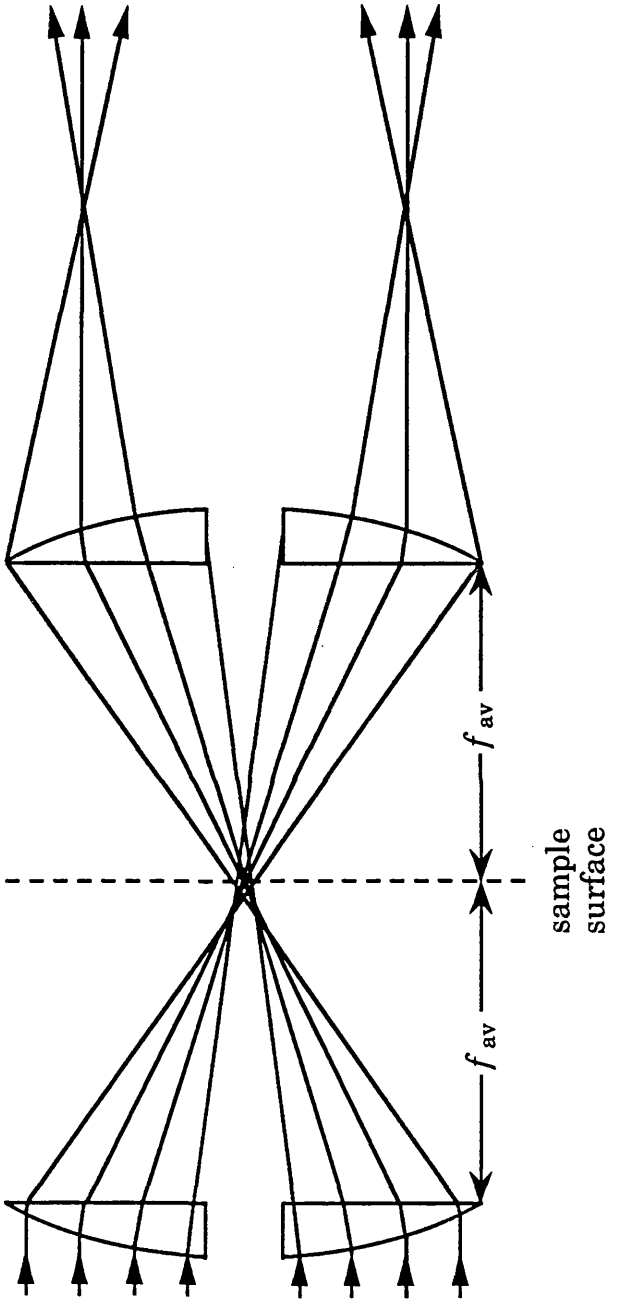


Figure 4.9. Ray diagram for a spherically aberrated objective lens, showing how the gap forms between the sample probe and sample reference beams.

therefore scattered. The envelope of all probe beam rays propagating to the detector plane, calculated using ray tracing, is shown in Fig. 4.10(a). This is a radial plot, and is thus circularly symmetric about the optic axis i.e. h=0. The dotted line indicates the radius of the drilled hole.

In order to compare the theoretical plot of Fig. 4.10(a) with that observed experimentally, profiles of the returning sample beams were measured at various axial positions. These were obtained by scanning an optical power meter across the beams which had been masked with a 100 μm pinhole. Typical plots recorded in this way are illustrated in Fig. 4.11, which shows the profiles measured at a position 100 mm from the 0.27 NA drilled objective lens. Fig. 4.11(a) is the profile across both the sample probe and sample reference beams. Fig. 4.11(b) was taken with the reference tilted away, though care was taken to refocus the sample probe beam. Using the latter profile and more recorded at various positions, the envelope of rays propagating to the detector plane was plotted, producing the result shown in Fig. 4.10(b). Qualitatively, the main features of this plot, agree well with the theoretical envelope of Fig. 4.10(a), i.e. the thinning of the annular probe beam to a minimum width, followed by a gradual expansion. However, the absolute agreement is not as good as might have been expected. It is evident from the experimental plot, that diffraction effects are significant, since this envelope clearly cannot be composed of a number of rays propagating at different angles.

The ray theory approach used to obtain Fig. 4.10(a) is therefore too simplistic to describe the situation completely. A more rigorous theory would have to account for the Gaussian beam profile, and also explicitly determine the best focal position, by modelling the interference with the common reference beam. The ray theory approach is however, sufficiently accurate to demonstrate that the gap effect is due to spherical aberration. This is supported by observations with the projecting objective, where thinning of the annular probe beam beam

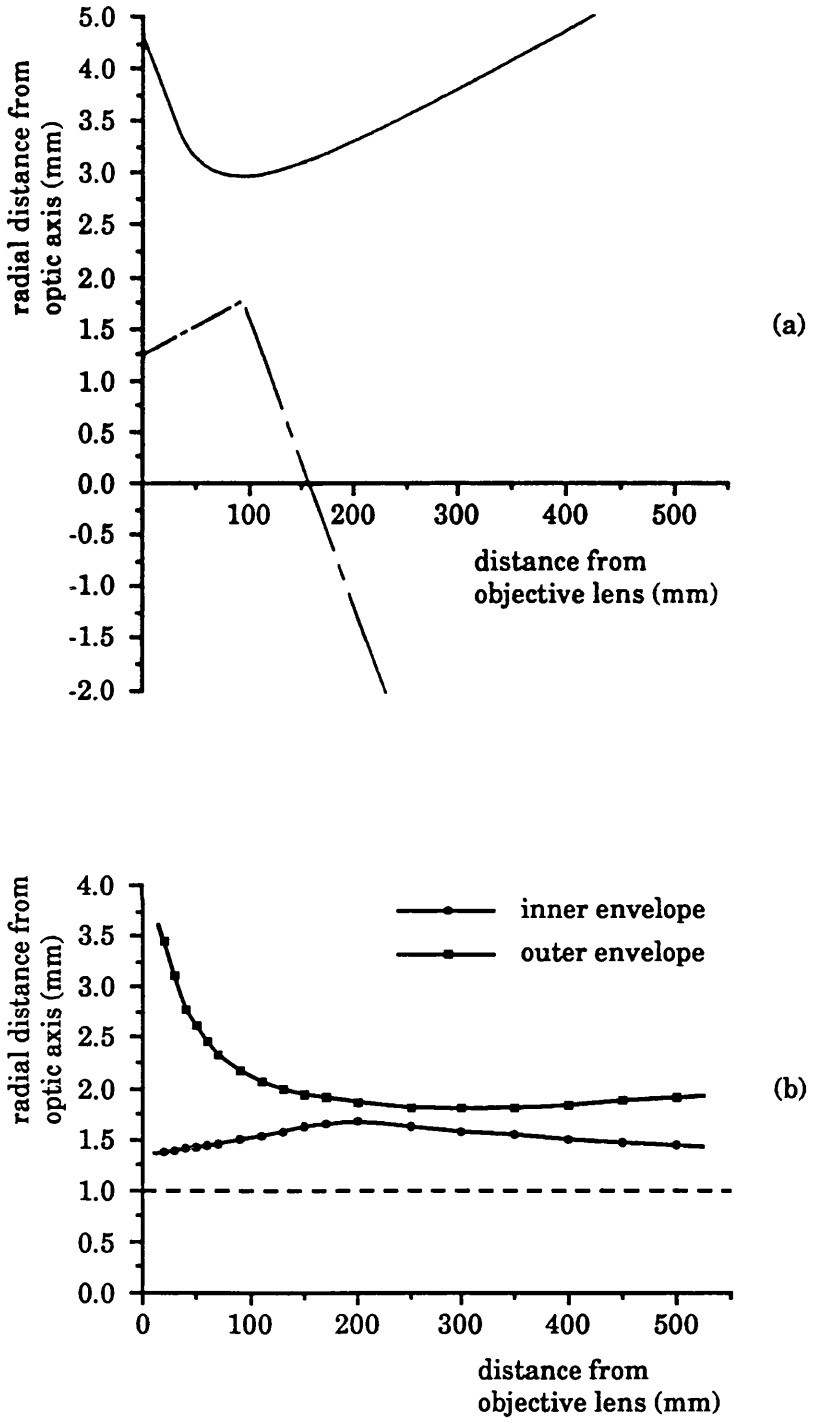


Figure 4.10. Envelope of all probe beam rays propagating from the objective to the detector plane, for the 0.27 NA lens used in the first implementation, (a) according to the ray trace (b) observed experimentally.

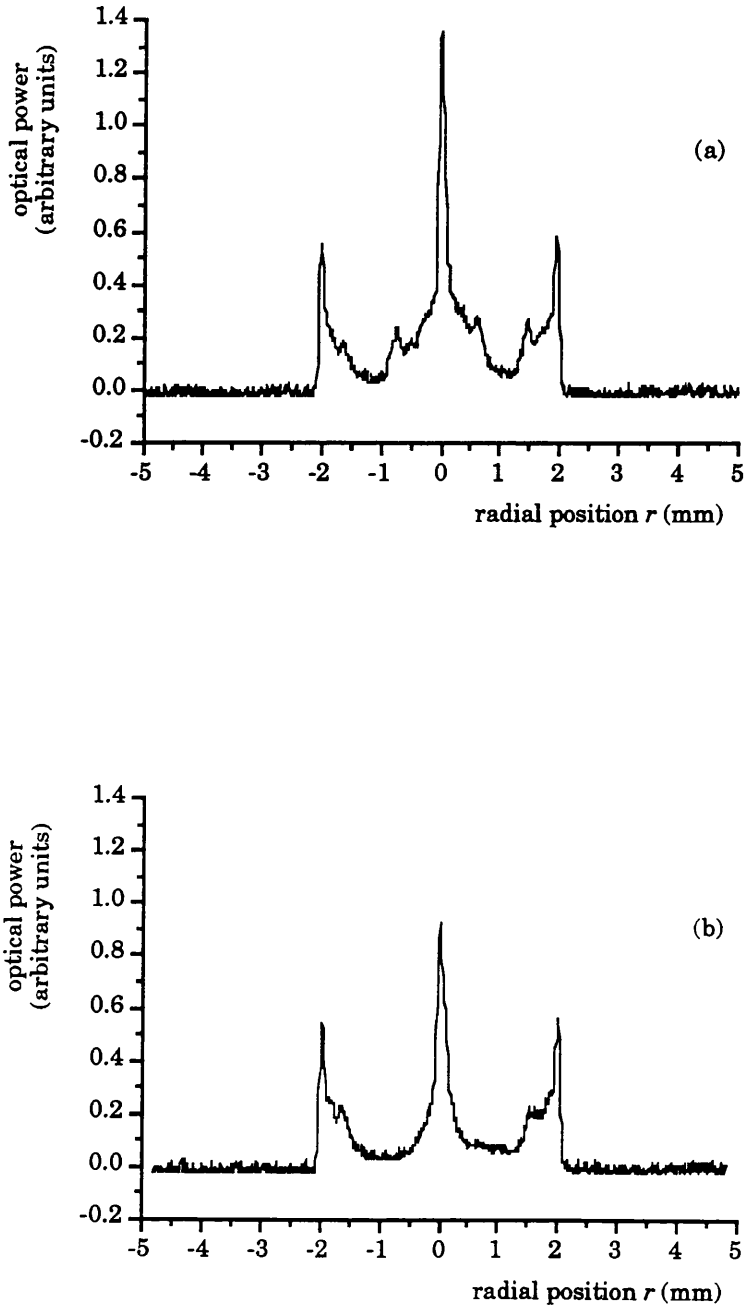


Figure 4.11. Profiles of the beams after propagating 100 mm from the lens to the detector plane, for (a) both the sample probe and sample reference beams, (b) the probe beam only (i.e. sample reference tilted away).

was found to be absent, due to the reduced aberration in these optics. For this reason there seemed to be no justification for pursuing the rigorous analysis, and attention was therefore turned to other interesting features of the beam profiles shown in Fig. 4.11.

Referring to Fig. 4.11(b) there are three sharp peaks in the profile. The outer two are the result of light concentrating at the perimeter of the probe beam, as it thins due to spherical aberration. The central peak results from light diffracted from the annulus into the central region. Thus there is a significant level of probe beam light present in the region occupied by the sample reference beam, as illustrated by Fig. 4.11(a). The above profiles are somewhat deceptive, since the optical power should be integrated with respect to the radius r. Fig. 4.12 shows the corresponding radial power plots of the profiles, obtained by multiplying them by $2\pi r\delta r$ where δr is the sampling interval. By summing the area under these plots, the total probe power diffracting into the region of the sample reference was determined. This enabled the ratio of this stray optical power to the true sample reference power to be calculated, and was found to be 0.73, assuming the spatial filters to be positioned 100 mm from the lens, where the profiles were measured. Obviously, the degree of diffraction increases with propagation distance, and after 300 mm this value increases to 1.49, i.e. there is more stray probe beam light than sample reference light in the central region.

The above situation appears disastrous in view of the crosstalk analysis presented earlier and one might expect a correspondingly large value for the crosstalk parameter ξ_p. In practice however, the crosstalk in the detected heterodyne reference signal was found to be buried in the residual signal due to electrical breakthrough from the Bragg cell driver, even when the spatial filters are placed 300 mm from the objective. This was determined according to the procedure outlined in Appendix A3.2 of the previous chapter, and indicates that the rejection of stray light from the probe beam by the heterodyning process is about 40 dB.

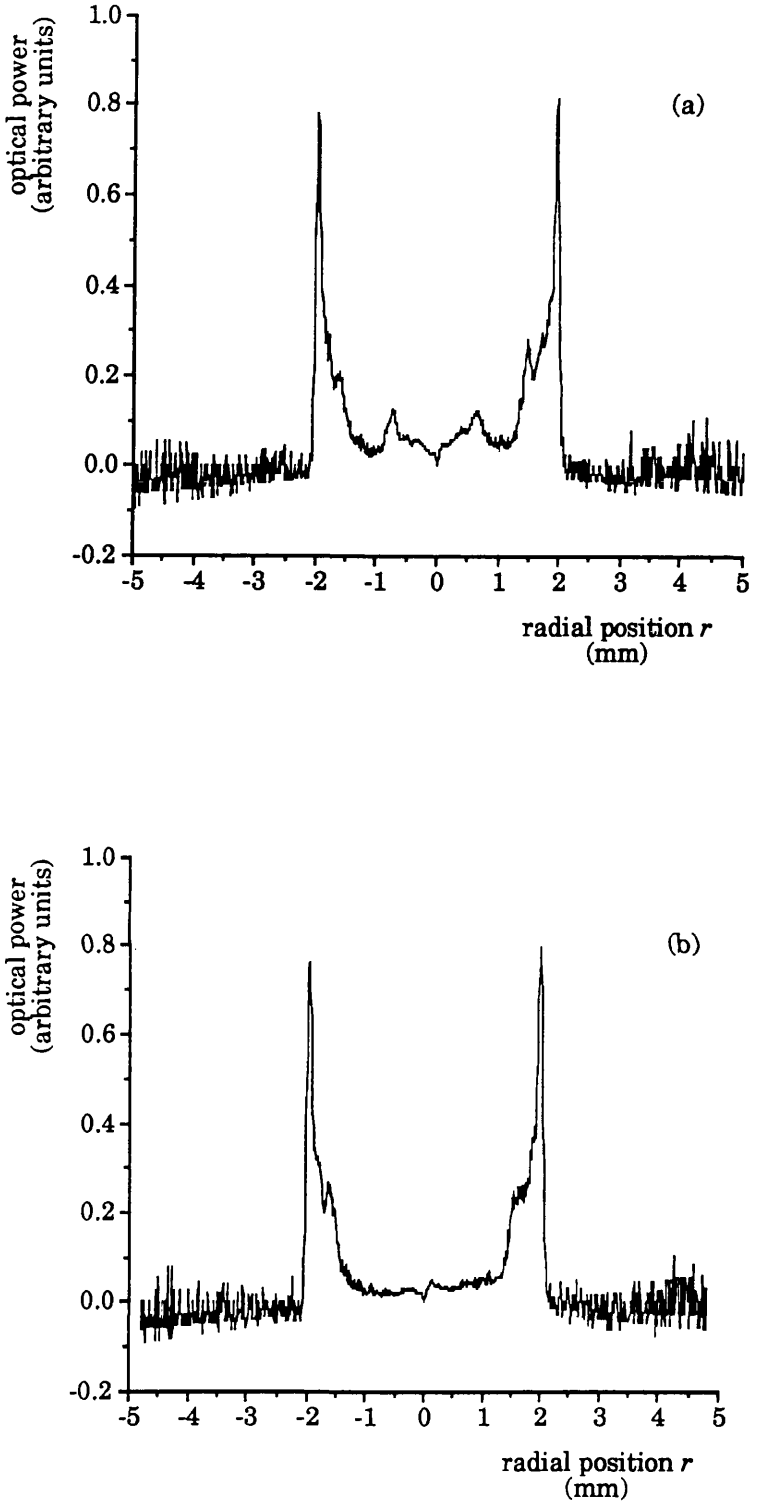


Figure 4.12. Radial power plots of the beams after propagating 100 mm to the detector plane, for (a) both the sample probe and sample reference beams, (b) the probe beam only.

It would appear that the wavefront presented by the diffracted light does not match that of the frequency shifted common reference beam. As a result, these beams do not interfere efficiently, and therefore do not produce a significant signal at the heterodyne frequency. The result of interfering plane wavefronts which have a relative inclination is considered in next section and illustrates the antenna properties of heterodyne photomixing.[60] In a similar fashion, light scattered or diffracted by the object from the collimated sample reference beam into the annular probe beam is also rejected. Furthermore, distorted wavefronts due to aberrations do not contribute to the interference signal. This is demonstrated by the drilled lens singlet, which gave 15 dB less signal to noise ratio than the projecting objective, due to spherical aberration. It is the outermost rays entering the objective which are most aberrated, and these are therefore rejected, thus reducing the effective numerical aperture. This reduction is offset by the fact that the objective is then more likely to be diffraction limited, giving an overall improvement in resolution.

4.3.3 Summary

The experimental results demonstrate that light scattered or diffracted between the interferometers does not result in a significant degree of crosstalk. Consequently, the crosstalk parameters ξ_p and ξ_r should refer solely to the misplaced heterodyne signals and not merely stray light. There are two possible sources of heterodyne crosstalk. The first is electrical crosstalk in the post detection signal processing. This is easily investigated by obstructing the light reaching each detector in turn, and was not found to be a problem for the detectors and amplifiers used in the experiments. The second possibility is optical crosstalk, caused by inadequate spatial filtering of the two interferometers (or in the case of separation according to polarisation, it may result from birefringence in the optics or the sample).

---

[60] A.E.Siegman, "The antenna properties of heterodyne optical receivers", Appl. Optics, **5**, (10), pp. 1588-1594 (1966).

Optical crosstalk can be eliminated by following the procedure outlined in Appendix A3.2. If a drilled lens singlet is used as the objective, spherical aberration produces a gap between the probe and sample reference beams, and this aids isolation of the interferometers by the spatial filtering method. However, the aberration also degrades both resolution and signal to noise ratio, and should naturally be avoided as far as possible. Both the ray tracing theory and experimental results demonstrate the formation of a maximum gap at a certain position along the path to the detector plane. Clearly, this would therefore seem to be the optimum position to place the spatial filters. However diffraction of light from the annular probe beam into the central region must also be avoided by minimising the propagation distance, since this wastes optical power. In cases where the gap is less apparent, such as for the reflecting and projecting objectives, optical crosstalk can be avoided by adjusting the size of the spatial filters, and again, this is detailed in Appendix A3.2

4.4 Tilt of the sample

This section considers the effect of sample tilt on system performance. In the conventional scanning heterodyne interference microscope, tilt is not a serious limitation, since the objective lens translates the tilt of the reflected beam to a lateral displacement at the interference plane. Although there is is a slight loss of signal amplitude due to the resulting overlap mismatch of the interfering beams, there is no error in the phase measurement. This because the interfering wavefronts will still be well matched.

For the common path system, the above discussion applies to the probe interferometer, which does use a lens to focus the illuminating beam onto the sample. The problem arises with the reference interferometer, since its sample illuminating beam passes through the hole in the lens and remains collimated. This results in a reflected beam which remains tilted with respect to the optic axis at the interference plane.

The following discussion considers the effect this has on the phase measurement, in order to calculate the maximum slope that can be tolerated.

In order to simplify the calculation to one dimension, we shall assume the beams are rectangular in cross section. If the sample is tilted from the normal by an angle α in the x direction, as in Fig. 4.13, the sample reference beam from Chapter 3 is modified to

$$S_r = A_r \exp j\{\omega_0 t + \phi_r + \theta_1 + (p + qx)\} \tag{4.61}$$

where

$$p = kl\left(\frac{1}{\cos 2\alpha} - 1\right) \tag{4.61a}$$

$$q = -k\tan\alpha\left(\frac{1}{\cos 2\alpha} + 1\right) \tag{4.61b}$$

This beam is interfered with the common reference beam given by

$$R = B\exp j\{(\omega_0 + 2\omega_B)t + \phi_c + \theta_2\} \tag{4.62}$$

The current density $J(x)$ generated at the photodiode is therefore

$$J(x) = \frac{\eta eG}{2Dhf}(S_r + R)(S_r + R)^* \tag{4.63}$$

Where D is the half width of the rectangular beam in the x direction. This may be written as

$$J(x)\frac{2Dhf}{\eta eG} = A_r^2 + B_r^2 + 2A_r B_r \cos\{\Phi + (p + qx)\} \tag{4.64}$$

where

$$\Phi = 2\omega_B t + (\phi_c - \phi_r) + (\theta_2 - \theta_1) \tag{4.65}$$

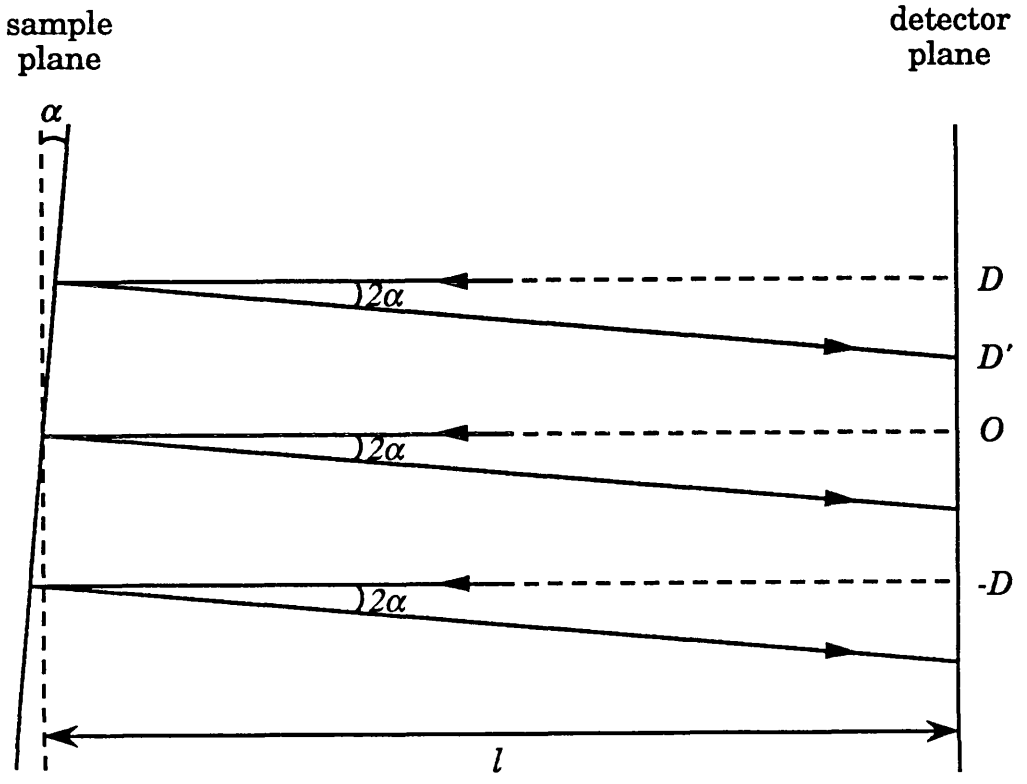


Figure 4.13. Geometry for calculating the effect of sample tilt.

$J(x)$ must be integrated over the whole detector to give the total current. The AC component of the signal $(I_r)_{ac}$ may therefore be written as

$$(I_r)_{ac} = 2A_r B_r \left(\frac{\eta eG}{hf}\right)\frac{1}{2D}\int_{-D}^{D'} \cos\{\Phi + (p+qx)\}\,dx \tag{4.66}$$

Where the upper limit to integration is reduced from D to D' by the lateral misalignment of the sample reference beam at the detector plane and is given by

$$D' = D - \tan 2\alpha(l - D\tan\alpha) \tag{4.67}$$

Where l is the mean distance from the sample to the detector. The integral evaluates to

$$(I_r)_{ac} = 2A_r B_r \left(\frac{\eta eG}{hf}\right)\frac{\sin\{q(D'+D)/2\}}{qD}\cos\left\{\Phi + p + \frac{q}{2}(D'-D)\right\} \tag{4.68}$$

The fraction immediately before the cosine term is the attenuation of the signal ρ_T caused by a reduction in interference efficiency. The phase error θ_e is given directly as the terms added to the true value Φ. Assuming the tilt is small enough so that $\tan 2\alpha \approx 2\alpha$ these parameters reduce to

$$\rho_T \approx \frac{\sin\{2k\alpha(D-\alpha l)\}}{2k\alpha D} \tag{4.69}$$

$$\theta_e \approx 2k\alpha^2(l - D\alpha) \approx 2kl\alpha^2 \tag{4.70}$$

If the diameter of the sample reference beam is 2 mm and l is 40 cm, a tilt angle of 10^{-4} rad gives a phase error of 79 mrad which is equivalent to 40 Å in terms of topography. The corresponding attenuation of the signal is a factor of about 0.48. Evidently, tilt is the most serious

limitation to system performance.

The reason for this large error in the reference phase, accompanied by a severe loss in signal is the fact that the sample reference beam is collimated at the sample surface. The deviation of the beam from the normal is therefore not translated to a lateral displacement by the objective lens as for the sample probe beam. A constant tilt is not a problem to the phase measurement as long as the degree of attenuation is acceptable, since it results in a constant phase offset. The problem arises with warped samples giving a variation in tilt.

The projecting objective does help to alleviate this difficulty somewhat, by reducing the deviation of the returning sample reference beam, as discussed in Section 3.6.2. Since this objective also reduces the diameter of the sample reference beam in proportion to the deviation, the attenuation due to to tilt is the same as for a comparable reflecting objective or lens singlet, with the same area illuminated by the sample reference beam. This is because the product αD which dominates in Eqn. (4.69) is constant. However, to a good approximation Eqn.(4.70), which gives the associated phase error, is independent of the beam diameter. As a result, the projecting objective gives a smaller phase error than a comparable reflecting objective or singlet for the same degree of tilt. This improvement is helped by the fact that α is squared in Eqn. (4.70).

As an example, the 0.65 NA projecting objective presented in Chapter 3 reduces the deviation by a factor of 10, as calculated in Section 3.6.2. Consequently, the phase error is reduced by a factor of 100. Thus the phase error calculated above (applicable for the drilled reflecting objective or singlet) is reduced to 0.79 mrad, or 0.4 Å. Clearly then, the projecting objective is the best objective to use for quantitative measurements in the presence of surface slope.

Even for the projecting objective, attenuation of the signal due to slope is

a severe limitation which needs to be overcome. A method has been described in the literature[61] of achieving heterodyne detection with less critical angular alignment, but it does not seem possible to incorporate this into the optical configuration and maintain the high degree of common path performance. The approach would involve focusing the sample reference beam onto the detector, where it would interfere with an unfocused common reference beam. The idea is that since the wavefronts of the Airy disk of the focused beam are plane and perpendicular to the optic axis,[62] they will match the wavefront of the common reference beam even in the presence of sample tilt. Another limitation of the approach is the reduction in signal to noise ratio, since most of the the power in the common reference beam is wasted. Alternatively, a feedback mechanism using a split detector has been proposed as a possible solution to this problem, by actively realigning the system as the sample is scanned. This is discussed in Chapter 7.

4.5 Laser instabilities

Two sources of noise due the laser have been observed experimentally,and either is capable of limiting the sensitivity of the phase measurement. If the heterodyne signal is viewed on a spectrum analyser, a number of sidebands are normally present about the fundamental at twice the Bragg cell drive frequency. These sidebands are usually (but not always) symmetrical about the fundamental and progressively decrease in amplitude away from it. They move cyclically in frequency, in and out of the fundamental, growing in amplitude as they near it, and decaying as they move away. At one stage in the cycle, the sidebands disappear totally into the fundamental, and not surprisingly, this is when the phase output is found to be most stable.

---

[61] W.S.Read and D.L.Fried, "Optical heterodyning with noncritical angular alignment", Proc. IEEE, **51**, p.1787 (1963).

[62] M.Born and E.Wolf, *Principles of Optics*, Pergamon Press, 6th Ed., pp. 445-449 (1980).

The cause of this effect is not entirely understood, but is almost certainly the result of thermal changes in the laser head. The evidence for this lies in variation of the period of the cycle from the time the laser is switched on. For the laser used for the measurements (detailed in Section 3.3.1), the period was found to increase from a few seconds initially, to about 20 minutes after a half hour warm up time. At the extreme of the cycle, the sideband spacing was about 130 kHz. The system was originally constructed with a 10 mW HeNe laser, but this had to be changed, since the above effect was so severe that phase lock could only be achieved at the stable stage in the cycle. This is probably due to the greater heat generated by this laser, but may also be related to its increased cavity length which supports more longitudinal modes.

The effect was found to be totally absent when using a frequency stabilised laser, but these have the disadvantage of low output powers, typically 0.5 mW. Although the laser used in the experiments did exhibit the effect, it did not seem to be limiting the phase measurement (i.e. there was no observable variation in the phase stability as the cycle progressed, as found with the 10 mW laser), and therefore, it was favoured over the stabilised laser due to its greater output power.

The second effect is caused by light passing back into the laser head, resulting in cavity detuning. The light returning to the laser from the external reflector may be thought of as forming a secondary cavity, whose length is fluctuating, due to microphonics. This is a possible explanation of the observed noise on the spectrum analyser, which is highly erratic, in contrast to the more regular thermal effect described above. Since the system is so sensitive to sample tilt, the degree of alignment required means that as much light is directed into the cavity as is detected. Under such conditions, the interference signals are too unstable for the detection electronics to cope and the system is completely unusable. The method employed to eliminate the problem was to incorporate a Faraday isolator, which typically provides about 40 dB isolation. This proved sufficient to remove the observed instability.

4.6 Summary

The fundamental limit to the phase measurement has been investigated theoretically in terms of two processes, namely electrical noise and thermodynamic fluctuations. It has been demonstrated that electrical noise is the dominant factor, although even this limit is several orders of magnitude better than the sensitivity achieved in experiments (2.6 Å for the first implementation). The small value obtained for the noise caused by thermodynamic fluctuations demonstrates that the lack of spatial overlap of the two interferometers is not a fundamental limitation.

The practical limit to the phase measurement is set by microphonics and temperature variations. Only piston like microphonics causing a relative movement between the sample and the objective affect the phase. This is a result of the obliquity effect, and thus vibration rejection is better at low numerical apertures. The results of the stability experiment of Section 4.2.2 demonstrate that sample microphonics represent the current limit to sensitivity. These measurements were recorded with the scan stage static, and are likely to be greater for fast mechanical object scanning. This justifies the development of an objective with greater immunity to vibration, such as that proposed in Section 7.3.1. The experimental results also indicate that noise due to the lack of spatial overlap of the interferometers i.e. thermal turbulence, would become significant if the present limitation were to be relaxed by such an objective. Consequently, it is doubtful whether the system can ultimately achieve the sensitivities attained by systems with more spatial overlap, such as those of Downs and Huang described in Chapter 2, without operating in temperature controlled environment. The advantage offered by the technique however is the larger area illuminated by the sample reference beam which provides a true absolute phase response for a greater range of samples.

The analysis concerning the NA factor in Section 4.2.1 also highlights the need to account for the obliquity effect when relating the measured

phase to a corresponding height variation. The obliquity effect is not specific to this particular system and it is normal practice to account for it by calibrating each objective using standard samples according to Eqn. 4.22.

The work on crosstalk and sample tilt has revealed some important general characteristics of heterodyne interference. This has led to some important advantages as well a major limitation of the technique. Firstly, light diffracted or scattered between the two samples beams does not interfere efficiently with the frequency shifted common reference beam, and thus crosstalk is avoided in the detected heterodyne signals. In addition, the effect of lens aberrations are reduced, and this is a strong argument for not placing a similar objective lens in the reference arm to match the wavefronts. This would improve the interference signal, but would probably degrade the resolution.[63] The limitation referred to is slope of the sample, which deviates the collimated sample reference beam. The projecting objective is the best objective to use for such samples since it reduces the deviation, resulting in a dramatic reduction in the phase error.

[63] There is a trade off here between rejection of unwanted aberrated wavefronts (from rays traversing the outer regions of the objective) and the associated loss of the higher object spatial frequency information that these carry.

Appendix A4.1 Calculation of the NA factor for an annular lens

The NA factor may be calculated by summing the obliquity effect of the focussed sample probe beam rays over all possible angles. Referring to Fig. A4.1 the average value of cos θ may be written as

$$(\cos\theta)_{AV} = \frac{\int_{\alpha_1}^{\alpha_2} 2\pi r dr \cos\theta}{\int_{\alpha_1}^{\alpha_2} 2\pi r dr} \tag{A4.1}$$

Where α_1 and α_2 are minimum and maximum propagation angles respectively. Clearly, α_2 defines the numerical aperture of the lens, whereas α_1 is determined by the radius of the central hole. The term $2\pi r dr$ is the weighting factor for all rays illuminating the sample at an angle θ, and the integral in the denominator is the appropriate normalising factor. In order to write the integrals solely in terms of θ, the following substitutions must be made

$$r = f\tan\theta \tag{A4.2}$$

$$dr = f\mathrm{cosec}^2\theta\, d\theta \tag{A4.3}$$

Eqn. (A4.1) then becomes

$$(\cos\theta)_{AV} = \frac{\int_{\alpha_1}^{\alpha_2} \frac{\sin\theta\, d\theta}{\cos^2\theta}}{\int_{\alpha_1}^{\alpha_2} \frac{\sin\theta\, d\theta}{\cos^3\theta}} \tag{A4.4}$$

Performing the integration[64] yields

$$(\cos\theta)_{AV} = \frac{\left[\frac{1}{\cos\theta}\right]_{\alpha_1}^{\alpha_2}}{\left[\frac{1}{2\cos^2\theta}\right]_{\alpha_1}^{\alpha_2}} \tag{A4.5}$$

---

[64] I.S.Gradshteyn and I.M.Ryzhik, *Table of Integrals Series and Products*, Academic Press, 4th Ed., p. 137 (1965)

and therefore

$$F_{NA} = \frac{1}{(\cos\theta)_{AV}} = \frac{(\cos\alpha_1 + \cos\alpha_2)}{2\cos\alpha_1 \cos\alpha_2} \qquad \text{(A4.6)}$$

This expression is plotted graphically in Fig. 4.2.

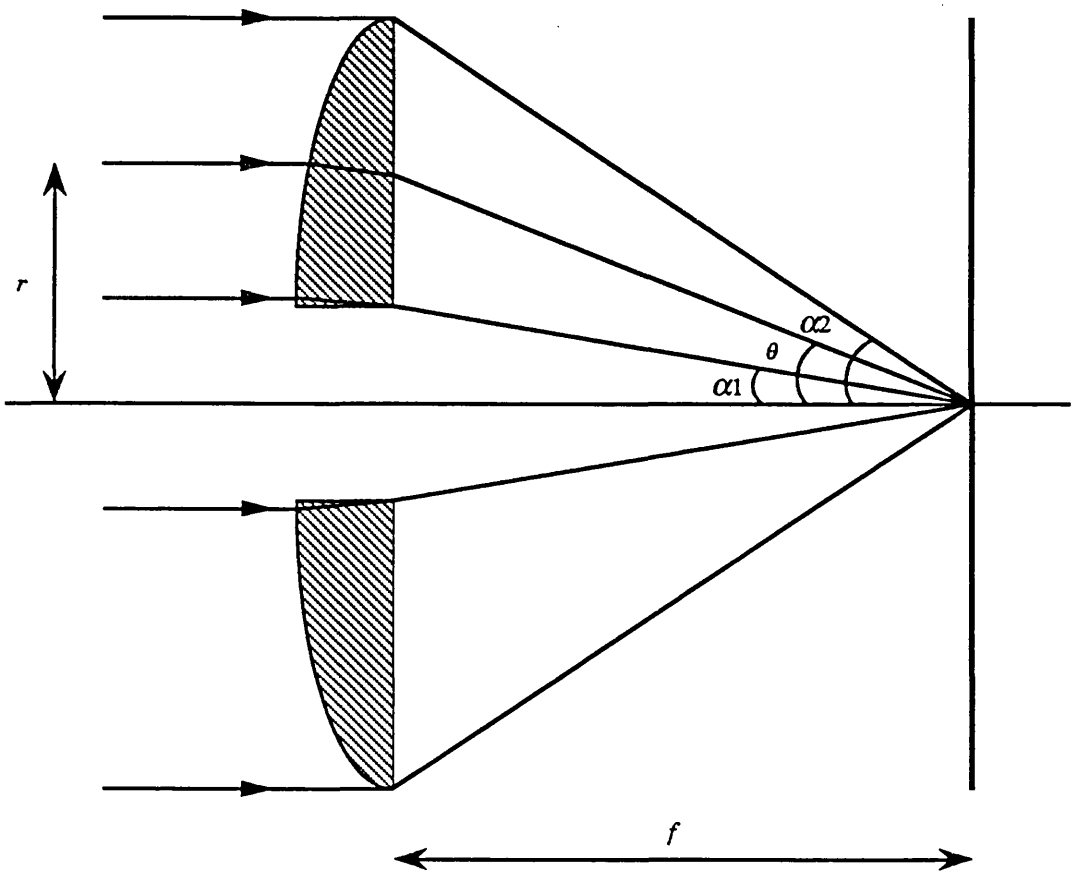


Figure A4.1. Geometry for the calculation of the NA factor of an annular lens.

CHAPTER 5

THE PROTOTYPE VERSION OF THE MICROSCOPE

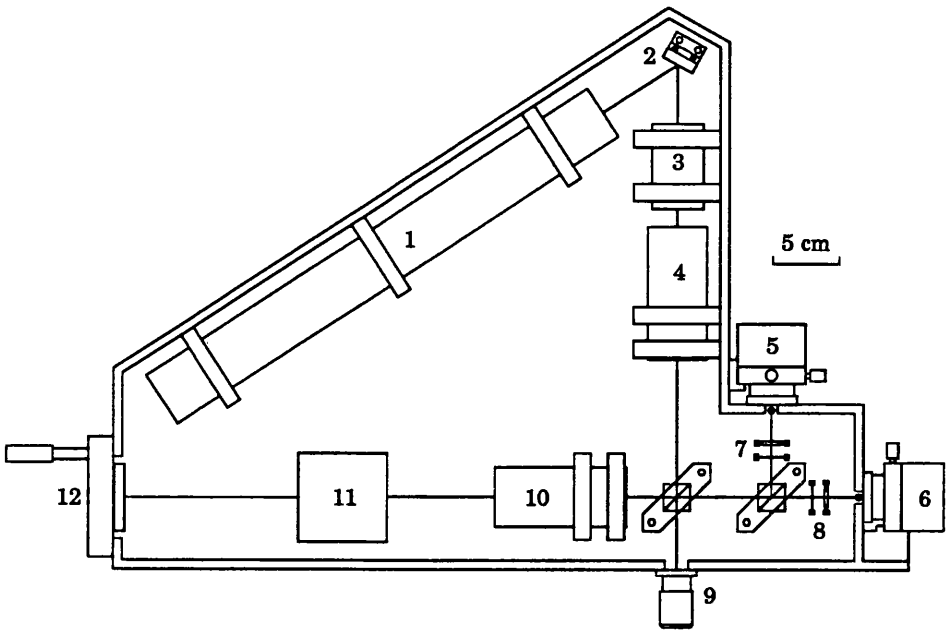
5. The prototype version of the microscope

After demonstrating the principle of the system as a quantitative profilometer, using the bench top set up described in Chapter 3, the next step taken was to design and build a stand alone instrument. The main reasons for this were twofold. Firstly, by placing the optics in a compact, enclosed unit, the effects of microphonics and turbulence are reduced. Furthermore, two dimensional object scanning has been incorporated, in order to produce a true microscope rather than a profilometer.

5.1 The design of the microscope

5.1.1 The optics

The optical components for the microscope were the same as those used in the final version of the bench top implementation and these have already been described in Chapter 3. Simple gimbal mounts were used to hold the majority of these in place, allowing a limited degree of adjustment, but still sufficient to enable the system to be aligned. The layout of the unit containing the optics is shown in Fig. 5.1. The interferometers were separated spatially rather than by polarisation, justified by the results of Chapter 3, and the crosstalk measurements presented in Section 4.3. The optical head was designed to be assembled, aligned and tested horizontally, and then mounted vertically, with the objective over the horizontal scan stage. The scan stage is housed in a sturdy case, which forms the base of the instrument and provides space for the laser power supply, Bragg cell driver, RF amplifiers and the 80 MHz bandpass filter. Photographs of the complete system, with the cover removed are shown in Fig. 5.2. Provision has been made for certain key adjustments to be made to the optics from outside the case. These include tilt of the reference mirror, rotation of the Bragg cell and x,y,z positioning of the photodiodes in each detector head.



1. Laser head.
2. Plane mirror.
3. Faraday isolator.
4. Beam expander (x10).
5. Reference detector.
6. Probe detector.
7. Spatial filter for reference interferometer (aperture matched to objective hole), followed by detector lens.
8. Spatial filter for probe interferometer (opaque spot on glass plate), followed by detector lens.
9. Objective (drilled singlet lens, reflecting objective or projecting objective).
10. Beam reducer.
11. Bragg cell mounted on a rotation stage.
12. Reference mirror mounted on a micrometer tilt stage.

Figure 5.1. Layout of the optical head for the prototype microscope.

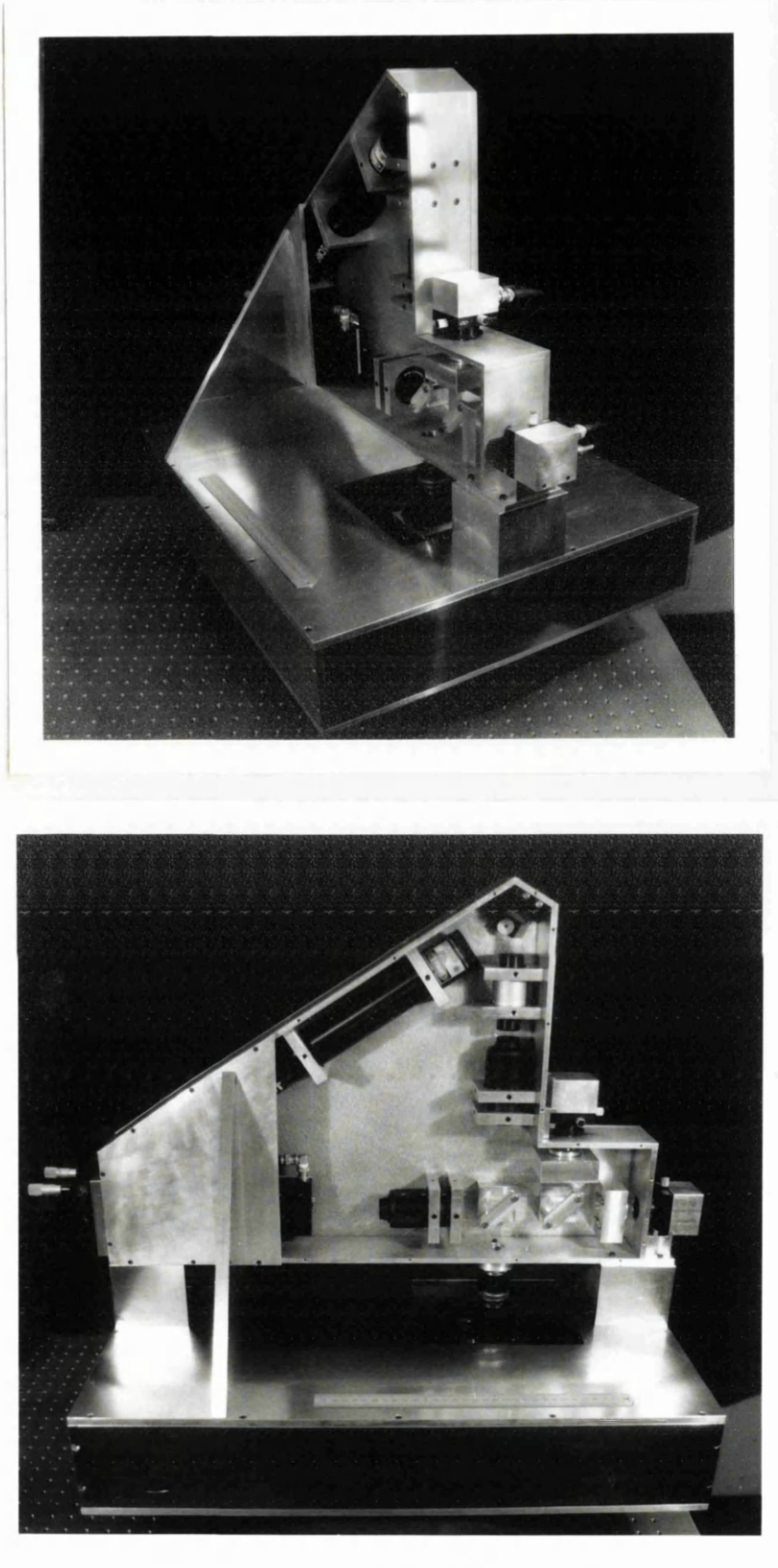


Figure 5.2. Two views of the prototype microscope, with a 30 cm steel rule to give scale.

5.1.2 The scan stage

The scan system chosen for the instrument was a Burleigh Instruments system 7000. The x and y translation stages have crossed roller bearings and are driven by "Inchworm" motors. These motors are composed of two piezoelectric clamps separated by a piezoelectric spacer. Each one drives a shaft which actuates the stage, by following a clamp-unclamp procedure analogous to the motion of a caterpillar. In the initial stage of the cycle, the clamps are both activated, and the spacer is unextended. Motion is effected by unclamping the first clamp with a suitable voltage and extending the piezoelectric spacer. The first clamp is then reactivated and then the second is unclamped, so that the spacer can be made to return to its unextended state. The complete cycle therefore results in a relative motion between the motor and the shaft.

The advantage of the system is that it retains the fine motion control provided by piezoelectrics, whilst also achieving long travel. The main drawback of the system is the low scan speed, which is typically 2 mm/sec. The image acquisition time is therefore proportional to the total distance scanned. For example, a 480x480 pixel bidirectional scan of an area of 1 mm^2 takes about 4 minutes. Position sensing is achieved using Heidenhain optical encoders, each with a base resolution of 0.1 μm. The complete system has its own controller, which generates the voltages to drive the piezoelectric clamps and spacers. Finally, an x,y tilt stage was mounted on the scanner to allow adjustment for sample tilt.

5.1.3 The software

Although the scan system is primarily designed to be operated from an IBM personal computer in terms of the hardware support, there is provision to control the scanner from other machines via the controller, using the RS232 interface. This option was utilised to control the system from an Apple Macintosh IIx computer, since this has excellent

graphics when used with the high-resolution RGB colour monitor. This decision was also governed by the move within the laboratory generally towards Apple computers due to the added advantages of mouse driven software, reliability and their superior desk top publishing capability. Problems were encountered however when programming for this application due to a lack of suitable documentation. As a result, relatively routine operations on the IBM, such as image display have required significant developmental work. The entire software package for controlling the scanning, data acquisition and image enhancement and display was written in the 'C' programming language.[65] The analogue to digital converter (A/D) consisted of a "MacADIOS II" 12 bit A/D and D/A converter from Amplicon Liveline Ltd. This has a maximum data rate of 142 kHz, multiplexed for more than one input and is the fastest currently available for the Macintosh II. Since the speed of scanning is relatively slow, the data rate for two inputs (amplitude and phase) is well within the capabilities of this board.

The algorithm for controlling the scanning and data acquisition is summarised briefly as follows. Firstly, communication between the computer and the controller through the RS232 interface is established and then the required scan area is entered. The entire code to generate this scan in controller language is then downloaded to the controller as an ASCII string and is executed. Each time the controller instructs the scan stage to perform a line scan, it flags the computer, and waits for a return flag from the computer instructing it to continue. When the main 'C' program running on the computer receives the flag from the controller, it sends the continue flag and reads the amplitude and phase data from the A/D. In this way, data acquisition is synchronised with the start of each line scan. The acquisition of each data point within a line scan is not synchronised to the scan position, and therefore relies on the scan velocity being uniform. The reason for this was that the communication from the controller to the scan stage was too slow to

[65] I am indebted to Eric Ngau, John Lyle and Freddie Chin for their help with this work.

permit point by point movements. In practice, the above proved to be acceptable solution given the limited time available to complete the work. A more rigorous approach would be to access the pulses from the optical encoders and feed them directly to the digital inputs of the computer to control the data acquisition point by point.

5.2 Results from the microscope

In this section, images obtained with the microscope for a number of representative samples are presented. For each sample, three separate images have been recorded. The first is the basic scanning optical microscope response i.e. it relies on the intensity of the light reflected from each measurement point on the sample. This will be referred to as the Type I image, in the terminology of Wilson and Sheppard.[66] The second image is derived from the amplitude of the heterodyne interference signal from the probe interferometer. This is the Type II image in Wilson and Sheppard's terminology, and they have shown mathematically that it is equivalent to the image obtained with a confocal microscope. The final image is the absolute phase response. This is the most important image, because it provides a sensitive as well as quantitative contrast mechanism.

The following images were obtained using the 0.65 NA projecting objective, since this gave noticeably better lateral resolution than that obtained using the 0.65 NA drilled reflecting objective. Although the reflecting objective has provision to adjust for a collimated input beam, by adjusting the separation of the mirrors, its vernier scale was not calibrated correctly. In fact, the resolution improved right up to the extreme of this adjustment, indicating that optimum resolution could be obtained if further travel was available.

---

[66] T.Wilson and C.J.R.Sheppard, *Theory and Practice of Scanning Optical Microscopy*, Academic Press, London (1984).

5.2.1 800 Å track sample

This sample is the one profiled in Chapter 3, exhibiting a series of tracks, 32 μm wide with heights of about 800 Å. The sample was fabricated in the manner described in Section 3.5, using a mask originally used to make the electrodes of a surface acoustic wave (SAW) device. Thus the the tracks are in fact connected at the ends as illustrated by the micrographs of the sample shown in Fig. 5.3.

Fig. 5.3(a) is the Type I response and has the lowest contrast of the three micrographs. Most of the detail is the result of scattering from the edges of the tracks, dust particles and scratches. The Type II response of Fig. 5.3(b), derived from the amplitude of the interference signal, exhibits a dramatic improvement in contrast over the previous case. The final micrograph of Fig. 5.3(c) is the phase response, which is the most quantitative of the three, with lighter areas representing the highest points on the sample. Visually however, it is not as spectacular as the Type II response. The reason is that rejection of scattered light effectively gives the Type II response a degree of edge enhancement. Using image processing techniques, the phase picture could easily be edge enhanced if so desired, by differentiation for example. Another point worth noting is the defect in the structure near the top right corner of each micrograph. Although obvious from the two interference images, it would have been difficult to detect it if presented solely with the basic scanning optical microscope response of Fig 5.3(a).

5.2.2 200 Å track sample

This sample was fabricated with the same mask as for the previous case, but the track heights were made only 200 Å high in this instance. This has resulted in severe damage to the structure when the mask was lifted away, illustrated by the micrographs of Fig. 5.4. The two interference images clearly show how the thin deposited chrome layer

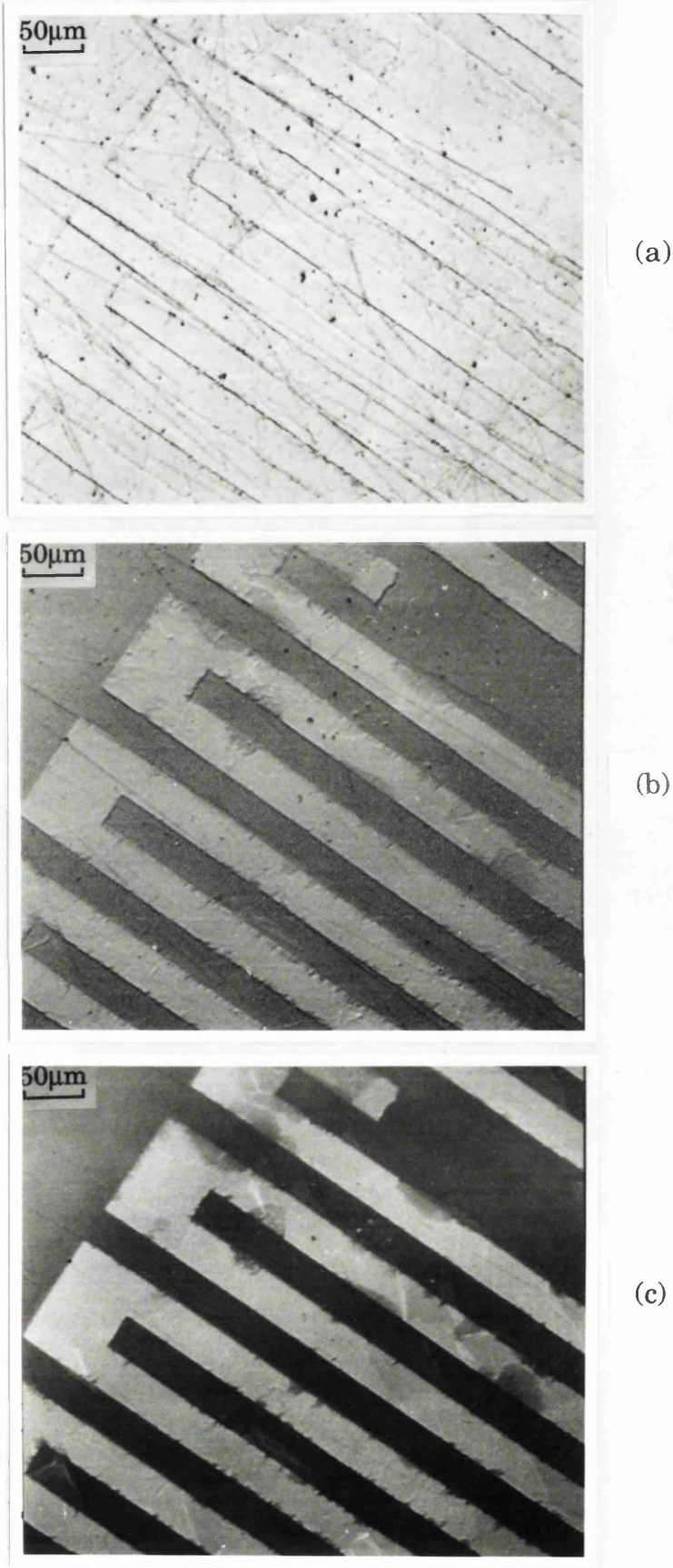


Figure 5.3. Micrographs of the 800 Å track sample. (a) Intensity (Type I), (b) Amplitude of interference signal (Type II) and (c) Phase of interference signal.

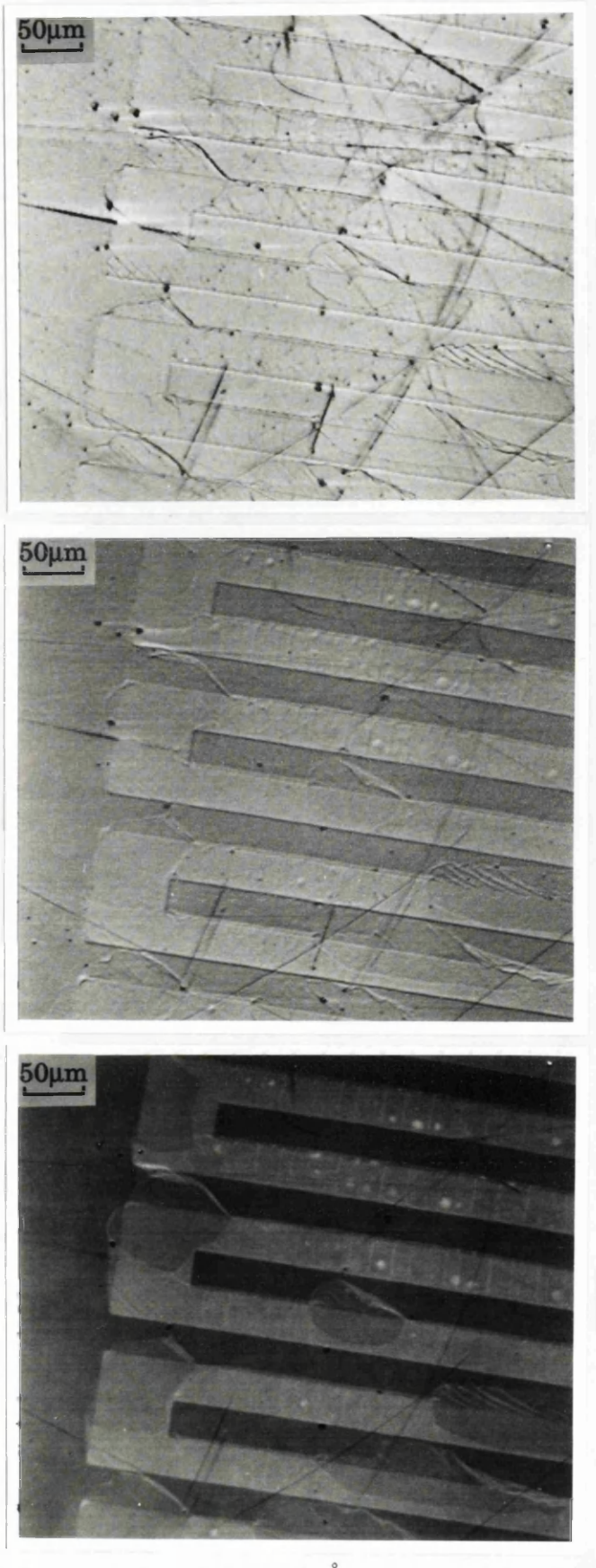


Figure 5.4. Micrographs of the 200 Å track sample. (a) Intensity (Type I), (b) Amplitude of interference signal (Type II) and (c) Phase of interference signal.

has torn in places as the mask was removed, forming flaps which have then folded or wrinkled. In addition, the phase response of Fig 5.4(c) indicates the presence of thin ridges, running across the raised tracks near the top of the picture. These are only just resolved, with a fairly regular spacing of about 20 μm. A possible explanation of their origin is that the chrome layer was inelastically stretched as the mask was lifted, leading to the formation of wrinkles. They are also apparent on the Type I and Type II images, though there would be an ambiguity in the interpretation with these pictures alone, since they could be ridges or troughs.

A number of raised circular features are also present on this part of the sample, with diameters of about 10 μm. One possible explanation is that they result from dust particles on the original glass surface. However, they are restricted to the raised track regions, and thus it is more likely that a blistering effect has occurred, again as the mask was lifted away.

5.2.3 Discrete edge sample

The series of micrographs in Fig. 5.5 were recorded from a plasma etched sample exhibiting a number of steps each about 5 mm apart. This is the same sample as that characterised with the first implementation, the results of which are presented in Section 3.5. These images were recorded in the region of edge H, using the previous notation, for which the step height measured with the bench top implementation was 661 Å (Table 3.1).

The Type I image of Fig. 5.5(a) is rather poor quality, stemming from the lack of contrast with this response. The contrast has been enhanced to the bring out the detail and this has also emphasised the high degree of noise present, since the signal is extracted at DC. Furthermore, a preamplifier with variable gain and offset has been used to optimise the extent of the signal variation relative to the range of the analogue to

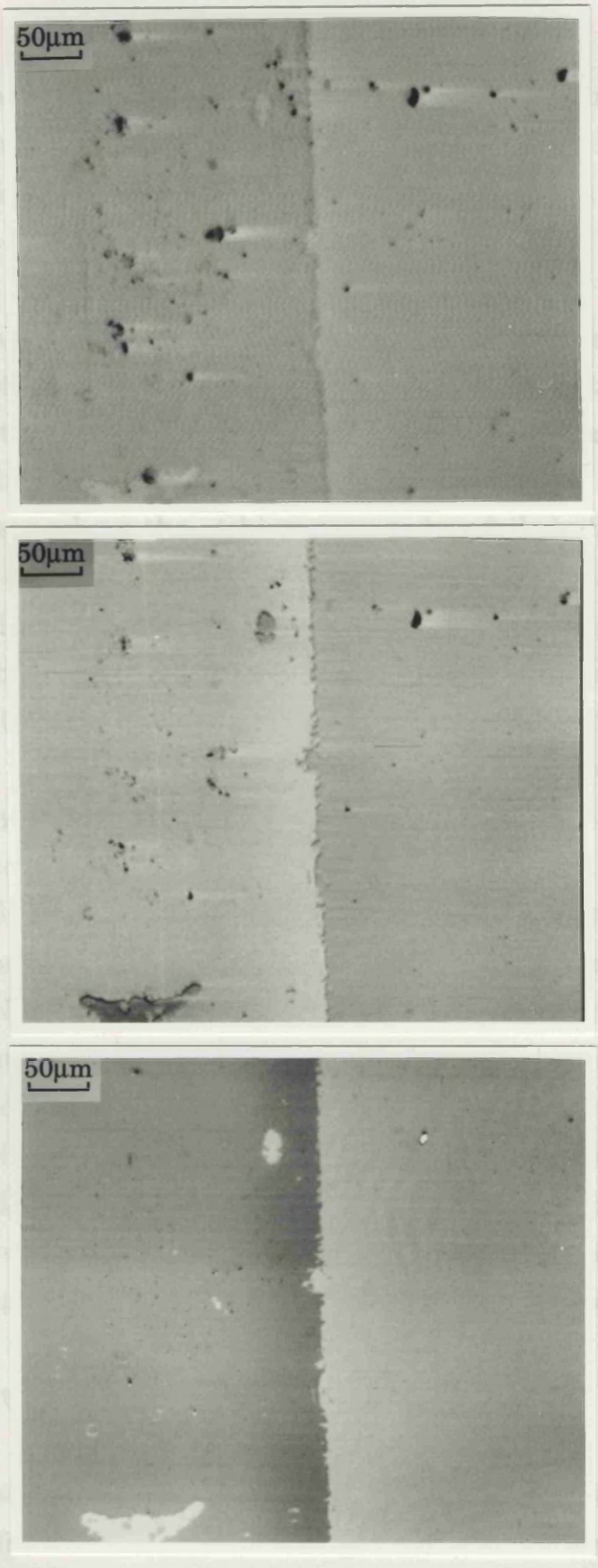


Figure 5.5. Micrographs of the discrete edge sample. (a) Intensity (Type I), (b) Amplitude of interference signal (Type II) and (c) Phase of interference signal.

digital converter. While this is essential to obtain a meaningful picture, the preamplifier has saturated where the system has encountered a dust particle, causing the light streaking effect on the "downstream" side of the particle as it is scanned.

The micrograph using the amplitude of the interference signal shown in Fig. 5.5(b) is far less noisy, but still exhibits the streaking effect after the largest pieces of debris on the surface. The plasma etched region is to the left of the edge and has a number irregular islands, which are clearly areas where the etching process has failed. The phase image of Fig. 5.5(c) demonstrates the limitation of the technique when applied to a discrete edge object. At the edge, the true phase change is preserved, but there is distortion away from it due to the finite area of the sample reference beam. If the region of the edge is blocked (with a finger for example), the two remaining sides are have the same grey level, apparently indicating that they have the same height. The phase reaches the same value on each side of the step after a distance from the edge equal to the radius of the sample reference beam. This radius was 100 μm for this measurement, using the 0.65 NA projecting objective. The 0.65 NA drilled reflecting objective gives a beam radius of 1 mm and consequently, the distortion produced would be much less apparent on the scale of the present micrographs. The effect has been considered theoretically in Section 6.3.2 and the results are compared with a line trace of edge H, also recorded using the projecting objective. This line trace, shown in Fig 6.9, can therefore be directly compared with the phase image, and illustrates the distortion more clearly than the image.

5.2.4 Waveguide sample

This sample consists of a number of waveguides on a silicon wafer, designed for radiation of wavelength 1.3 μm. A cross section through one of these waveguides is shown schematically in Fig. 5.6.

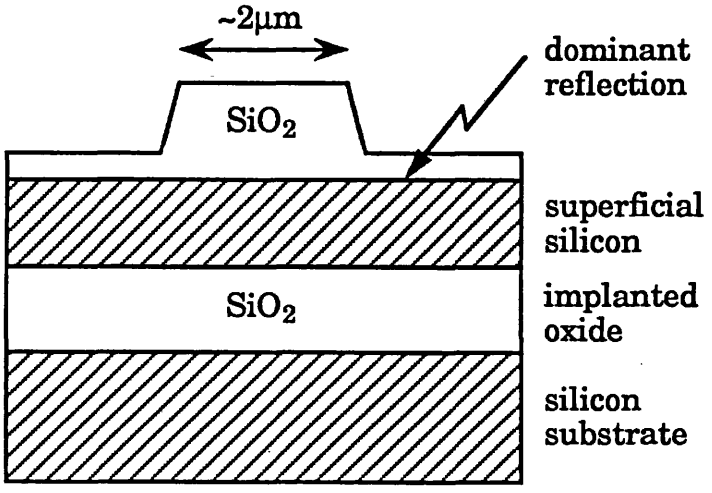


Figure 5.6. Schematic cross section through one of the waveguides on a silicon substrate.

The micrographs of the sample are shown in Fig. 5.7. The most interesting feature of these results is that the phase image seems to indicate that the waveguide is a trough about 650 Å deep,rather a ridge on the surface (light areas indicate high regions and dark areas indicate lower regions). The nature of the structure is well known from the fabrication procedure and from electron microscopy. The observed phase response is a consequence of the fact that the dominant reflection occurs at the uppermost silicon and silicon dioxide (SiO_2) interface, rather than the top SiO_2 surface.

The light which reaches the silicon under the ridge of SiO_2 traverses a greater optical path length than the rest (since SiO_2 has a larger refractive index than air), and this makes it appear as a trough. Using a value of 1.54 for the refractive index of SiO_2, we may now interpret the phase change in terms a ridge 1204 Å high, by dividing the original figure by the difference in the refractive index of SiO_2 and air. The contribution from the top surface SiO_2 has been neglected in this calculation and this can be justified for two reasons. Firstly, the reflection coefficient for SiO_2 is only about 0.04 and also, the depth sectioning property of the interferometer will tend to result in a reduction of interference efficiency for the rays from this surface. In summary then, this sample demonstrates the ambiguity which may arise when interpreting the phase from an interferometer directly in terms of topography.

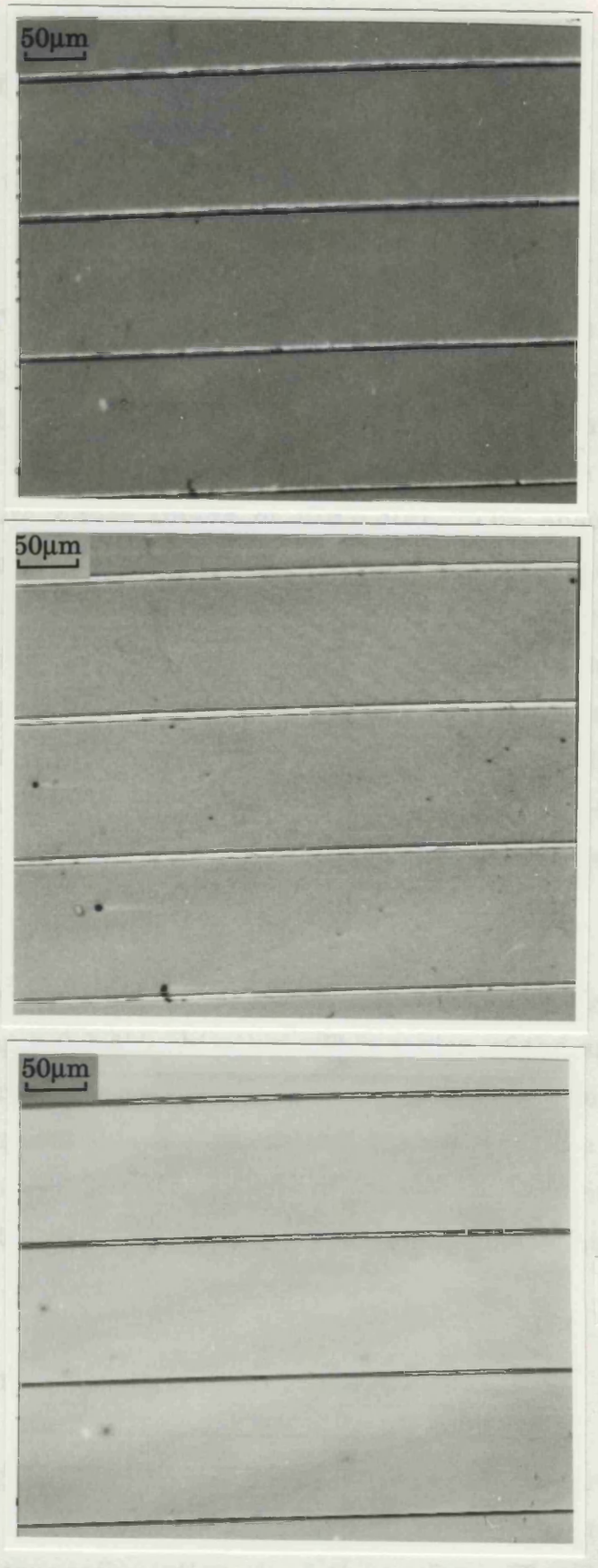


Figure 5.7. Micrographs of the waveguide sample. (a) Intensity (Type I), (b) Amplitude of interference signal (Type II) and (c) Phase of interference signal.

5.2.5 Comparison with other techniques

The track samples used to obtain the results of Sections 5.2.1 and 5.2.2 have also been imaged using a "Leitz" confocal microscope and a Nomarski differential interference contrast (DIC) system (Olympus AH2).[67] Typical results obtained from these systems are shown in Figs. 5.8 and 5.9. Fig. 5.8 is a DIC image of the 200 Å track sample, recorded with a 0.55 NA objective. The edge enhancement provided by the technique shows the tearing and wrinkling damage more clearly than the absolute phase image of Fig.5.4(c). The absolute phase image contains far more quantitative information about the surface however, and if so desired could be differentiated to produce a more visually attractive image. There is insufficient information in the DIC image to perform the inverse operation, i.e. to generate a quantitative absolute image by integration. Another advantage of the absolute phase image is that it can be differentiated in both the x and y directions using image processing methods, whereas the DIC image differentiates along one axis only.

Fig. 5.9 shows a reflection confocal image of the 800 Å track sample, taken using a 0.7 NA objective. This region of the sample exhibits the ridges described in Section 5.2.2 for the 200 Å sample. These ridges are imaged surprisingly well, considering that even the track heights are well within the depth of focus of the objective. As with the DIC system however, it is difficult to extract quantitative height information from this image.

5.3 Summary

The micrographs presented in this chapter demonstrate that the three different responses available from the instrument provide complementary information about the surface. The Type I response,

[67] I would like to acknowledge Unilever Research for permitting the use of these instruments at Port Sunlight Laboratory, under the terms of a CASE award.

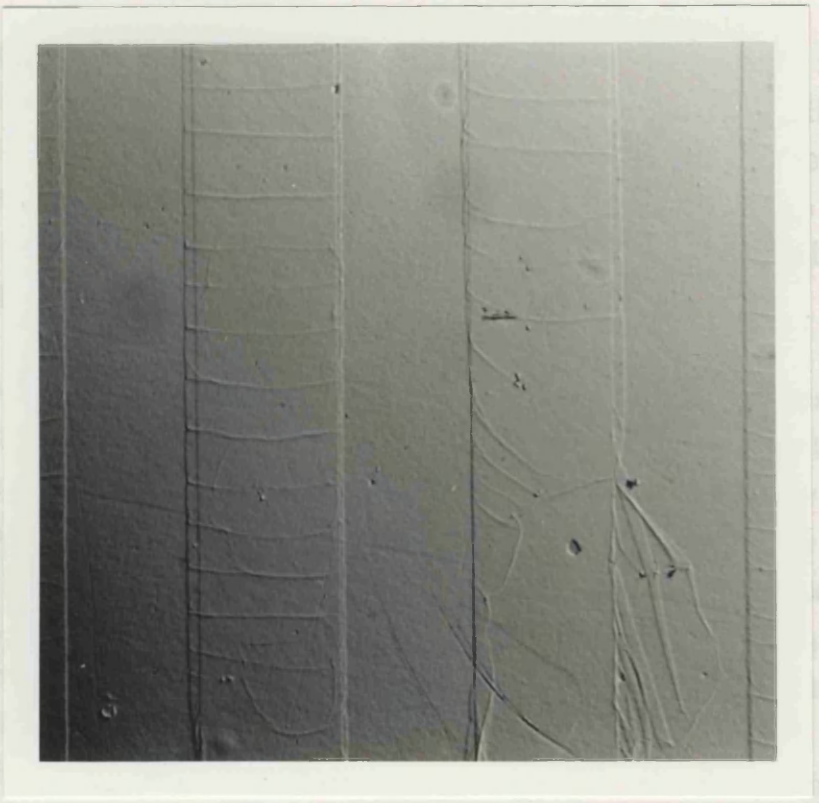


Figure 5.8. Micrograph of the 200 Å track sample taken with a Nomarski DIC microscope.

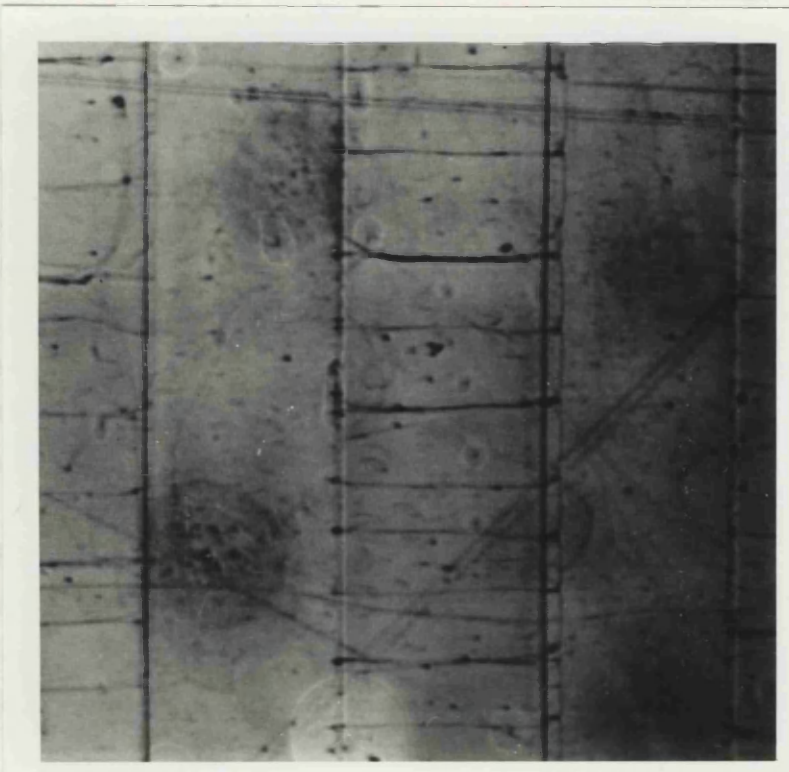


Figure 5.9. Micrograph of the 800 Å track sample taken with a reflection confocal microscope.

although noisy since it is extracted at DC, picks out detail resulting from the highly scattering features, such as scratches, digs and dust particles. The Type II response has a much lower noise level, since it is derived from the heterodyne signal, for which $1/f$ noise is negligible. Good contrast is therefore obtained even for features which are not such strong scatterers.

The phase response represents the most sensitive contrast mechanism of the three available images and also has the advantage that the information is quantitative. Thus numerical analysis may be applied to the raw data to extract values such as rms roughness, or to obtain the object spatial frequency distribution. In addition, as stated earlier, image processing techniques such as edge enhancement may be used to improve the visual impact of the phase pictures. The results from the discrete edge and waveguide samples have highlighted the fact that care must be taken when converting the measured phase change to a surface height variation.

Allowing for the slight differences in the numerical apertures, the lateral resolution of the system is not quite as good as that obtained with the confocal and Nomarski instruments. The most likely explanation for this is that the 0.65 NA microscope objective used to build the projecting objective is designed primarily for a conventional full field microscope with a 160 mm tube length. Ideally, an objective designed for a collimated input beam (known as infinity corrected) should be used. Furthermore, conventional microscope objectives tend to have residual aberrations, which are corrected for by the eyepiece. Such aberrations would therefore limit the resolution if the objective is taken from such a system and used for the present application.

Typical phase stability measurements for the prototype system have been presented in Section 4.2.2. The best stability observed in experiments was about 1 Å for the 0.27 NA lens singlet. This was with the scanner disabled however, and in practice the microphonics caused by the piezo clamps of the scanner proved to be the most severe limitation to the phase images.

CHAPTER 6

IMAGE FORMATION THEORY FOR THE MICROSCOPE

6. Image formation theory for the microscope

The method of describing an optical system as an object spatial frequency filter is an established technique in optical microscopy. The main advantage of this approach lies in the representation of the optical system by an imaging function which is common to all objects. The development of the image formation theory of the conventional optical microscope is mainly attributed to Hopkins.[68] This work was later extended[69,70] to encompass the imaging in scanning homodyne microscopes. Most recently, the technique has been applied to a variety of scanning heterodyne techniques.[71,72] In this chapter, the technique will be applied to the present microscope, in order to gain a better understanding of its imaging characteristics and to determine the types of surface structure that it best suited to, in terms of providing a quantitative absolute phase response. The main adaptations made to the above theory has been to allow for the use of indirect interference and the inclusion of an annular objective in the probe interferometer.

The microscope incorporates two heterodyne interferometers, and for the purpose of the calculation these may be treated separately. The transfer function approach has been applied to the probe interferometer, taking into account the effect of the annular objective lens. In contrast, the response of the reference interferometer has been determined using geometrical optics, since it offers simplicity of formulation and faster

---

[68] H.H.Hopkins, "On the diffraction theory of optical images", Proc. Roy. Soc., **A 217**, pp. 408-432 (1953).

[69] C.J.R.Sheppard and A.Choudhury, "Image formation in the scanning microscope", Optica Acta, **24** (10), pp 1051-1073 (1977).

[70] C.J.R.Sheppard and T.Wilson, "Fourier imaging of phase information in scanning and conventional optical microscopes", Phil. Trans. Roy., **295**, pp. 513-536 (1980).

[71] M.G.Somekh and R.K.Appel,"Image formation in common path differential profilometers", *Proc. SPIE Surface Characterisation and Testing II*, Editors J. E. Greivenkamp and M.Young,**1164**, pp. 99-109 (1989).

[72] M.G.Somekh, "Image formation in scanned heterodyne microscope systems", J. Microscopy, to be published (1990).

computation. This approach is valid since the illuminating beam is collimated, with a diameter of several thousand wavelengths and so diffraction effects are negligible. In any case, it has already been demonstrated in Section 4.3.2 that the heterodyning process tends to reject diffracted light. Furthermore, Section 4.4 demonstrates that the reference interferometer is insensitive to light propagating at an angle to the common reference beam. These effects are important since only the specular reflection contributes to the signal, ensuring the validity of the approach. This is not the case for the probe interferometer because the lens collects the higher diffracted orders from the sample (which carry the high object spatial frequency information) and recollimates them.

The overall phase output is determined by subtracting the phases of the two interferometer signals. Ideally, the phase of the reference interferometer should be constant as the sample is scanned. Its purpose is to carry information concerning the phase error introduced by microphonics. This phase error term is also present in the probe interferometer signal, and is therefore removed by subtraction as described previously. In practice however, there may be a slow variation in the phase of the reference with scan position, which would serve to distort the absolute phase measurement. It is therefore important to quantify this effect for the various object structures likely to be encountered.

In Section 6.1 the principles behind the transfer function approach are introduced. This is applied to the probe interferometer in Section 6.2 to obtain a general expression for its imaging response. This expression is then evaluated in Section 6.3 for the sinusoidal grating and the discrete step. In each case the response from the reference interferometer is determined using geometrical optics and the overall phase output is then calculated.

6.1 Fundamentals of image formation theory

Assuming that scalar diffraction theory is applicable, the intensity variation in the microscope image due to a line structure (i.e. an object with variation along the x direction only) may be expressed as

$$I(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} C(m;p)T(m)T^{*}(p)\exp 2\pi j(m-p)x \, dm\,dp \qquad (6.1)$$

where m and p are spatial frequencies in the object, $C(m;p)$ is the partially coherent transfer function,[73] and $T(m)$ is the Fourier transform of the object transmittance or reflectance $t(x)$ given by

$$T(m) = \int_{-\infty}^{\infty} t(x)\exp -2\pi jmx \, dx \qquad (6.2)$$

$T(m) = \int_{-\infty}^{\infty} t(x)\exp -2\pi jmx \, dx$

The function $C(m;p)$ gives the magnitude of the spatial frequency component $(m\text{-}p)$ in the intensity image. It is a property depending solely on the optical system, irrespective of the object. This is an important advantage of the technique, since Eqn. (6.1) may be evaluated for any object $t(x)$, providing its transform $T(m)$ is known. The need to formulate the microscope response in terms of two object spatial frequencies m and p is a result of the square law nature of the detector. Since the detector only responds to intensity, all possible cross terms between arbitrary spatial frequencies which photomix at the detector surface, must be accounted for.

The above expression may be simplified somewhat if the discussion is restricted to weak objects. A weak object is one which has small variations in amplitude or phase, dominated by a strong background. It may be represented by

$$t(x) = 1 + t'(x) \qquad (6.3)$$

---

[73] T.Wilson and C.J.R.Sheppard, *Theory and Practice of Scanning Optical Microscopy*, Chpt. 3, Academic Press, London (1984).

where $t'(x)$ is the term due to light diffracted by the object with $|t'(x)| \ll 1$. The expression has been normalised such that the term due to the undiffracted light is unity. In this case Eqn. (6.1) reduces to[74]

$$I(x) = C(0;0) + 2Re\left\{\int_{-\infty}^{\infty} C(m;0)T'(m)\exp 2\pi jmx\, dm\right\} \tag{6.4}$$

where $C(m;0)$ is termed the weak object transfer function. The reduction of $C(m;p)$ to $C(m;0)$ is a consequence of the negligible contribution from cross terms between two arbitrary spatial frequencies. The only significant contributions arise from cross terms between each spatial frequency and the large DC term due to specular reflection. The weak object transfer function may be evaluated by considering the effect of a spatial frequency in the object on the light which is focused on to it. Each spatial frequency m will diffract the illuminating beam away from the normal by an angle θ given by

$$\theta = m\lambda \tag{6.5}$$

where λ is the optical frequency. Clearly, there exists a cut off spatial frequency m_{max} for which the beam will be diffracted outside the pupil of the detector lens. This occurs when

$$\theta = \frac{2a}{f} \tag{6.6}$$

which gives

$$m_{max} = \frac{2a}{\lambda f} \tag{6.7}$$

where f is the focal length of the lens and a is the radius of the lens aperture. These ideas form the basis of the Abbe theory of microscope

[74] D.K.Hamilton and C.J.R.Sheppard, "Differential phase contrast in scanning optical microscopy", J. Microscopy, **133** (1), pp. 27-39 (1984).

imaging. The function $C(m;0)$ may now be calculated graphically as the area of overlap of two circles illustrated in Fig. 6.1. One circle represents the lens aperture and the other the beam diffracted by a spatial frequency m. From Eqn. (6.5) we see that $C(m;0)$ is a maximum at $m=0$ and decays monotonically to zero at the cut-off frequency. This is illustrated in Fig. 6.2 where $C(m;0)$ is plotted as a function of normalised spatial frequency.

6.2 Image formation in the probe interferometer

The probe interferometer is a heterodyne Michelson interferometer, incorporating an annular objective lens. The imaging theory for this interferometer has been determined by adapting the work of Sheppard and Wilson on the homodyne scanning interference microscope. In Sections 6.2.1 and 6.2.2 modifications are made to this theory to account for the annular objective lens and for heterodyning at twice the Bragg frequency.

6.2.1 Imaging in the scanning interference microscope

We will take as our starting point the homodyne scanning interference microscope discussed by Sheppard and Wilson.[75] The system is arranged in a similar manner to the Mach Zehnder interferometer to give signals corresponding to sum and difference of the two beams, denoted I_+ and I_- respectively. The expression they derive for the detected intensity is given by

$$I_{\pm}(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} C_{11}(m;p)T_1(m)T_1^*(p)\exp 2\pi j(m-p)x\, dm\, dp$$
$$+ |W|^2 C_{22}(0;0) \pm 2\mathrm{Re}\left\{W^* \int_{-\infty}^{\infty} C_{12}(m;0)T_1(m)\exp 2\pi jmx\, dm\right\} \quad (6.8)$$

The first term in the above expression is the response from a

[75] C.J.R.Sheppard and T.Wilson, "Fourier imaging of phase information in scanning and conventional optical microscopes", Phil. Trans. Roy., **295**, pp. 513-536 (1980).

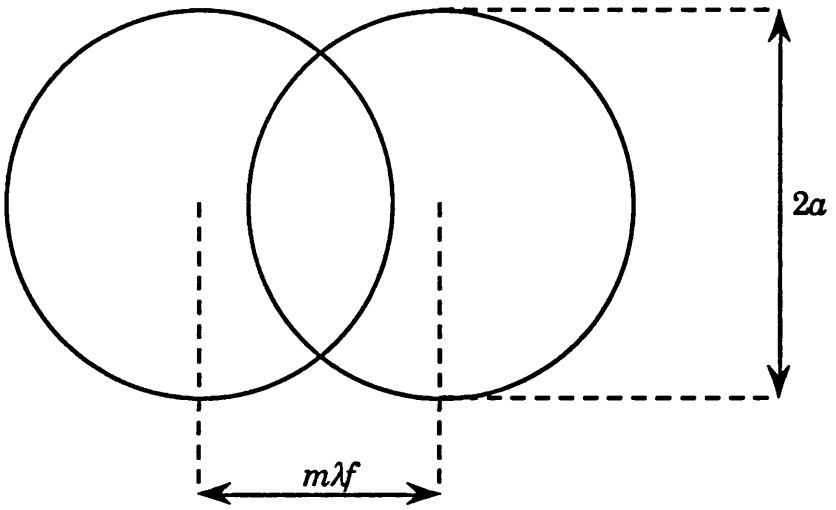


Figure 6.1. Calculation of the weak object transfer function $C(m;0)$, as the area of overlap of two displaced circles.

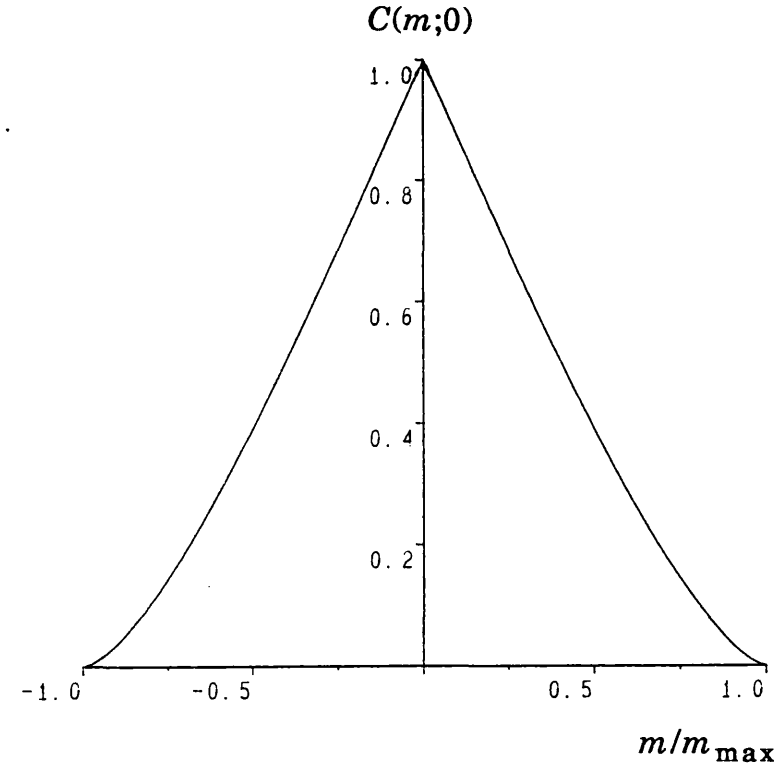


Figure 6.2. The weak object transfer function $C(m;0)$ versus object spatial frequency m normalised to the cut off value.

conventional scanning optical microscope as introduced in Section 6.1, and similarly the second is due to the reference beam which has constant amplitude and phase denoted by W. The final term is the interference term which may be isolated by subtracting the sum and difference signals, I_+ and I_-. $C_{12}(m;0)$ is the partially coherent transfer function for the interference term, given by

$$C_{12}(m;0) = \int_{-\infty}^{\infty} P_3(\lambda fm - \xi_2)P_4(\lambda fm - \xi_2)P_3^*(-\xi_2)P_4'(\xi_2)\,|P_3'(\xi_2)|^2\,d\xi_2 \tag{6.9}$$

Where P_3 is the entrance pupil function, P_3' is the detector lens pupil function, and P_4 and P_4' are the pupil functions for the objective lens for the incident and returning beams respectively. It is important to note that the transfer function for the interference term is of the form $C(m;0)$ rather than $C(m;p)$. This is an approximately true for weak objects with the non-interference microscope, but in this case it is exact, even for strong objects. The reason is simply that only cross terms between a particular diffracted spatial frequency and the reference beam are responsible for the interference signal. A significant feature of Eqn. (6.9) is that the detector lens pupil function P_3' is only present as its modulus squared. This means that the transfer function is independent of aberrations in this lens, since their effect is to introduce a phase factor into the pupil function.

We are now in a position to discuss the image formation in the probe interferometer. Eqns. (6.8) and (6.9) are applicable, provided some minor modifications are made. Firstly, the constant reference W must include an exp $2j\omega_B t$ dependence to allow for heterodyning at twice the Bragg frequency ω_B. In addition, P_4 and P_4' in Eqn. (6.9) are replaced by annular pupil functions to account for the annular objective lens. There is no longer any need to subtract sum and difference signals to extract the interference term, since it is now an AC signal. The interference term for the probe interferometer is therefore given by

$$I_p(x) = 2\mathrm{Re}\left\{ W^* \exp 2j\omega_B t \int_{-\infty}^{\infty} C_{12}(m;0) T_1(m) \exp 2\pi j m x \, dm \right\} \qquad (6.10)$$

In order to evaluate this expression for particular objects, the transfer function $C_{12}(m;0)$ is required. This is considered in the following section.

6.2.2 Transfer function for an annular aperture

Since the pupil functions are everywhere either unity or zero, the transfer function $C_{12}(m;0)$ given by Eqn. (6.9) may be reduced to a convolution of two equal annular pupils (i.e. we must calculate the area of overlap as one annulus is displaced relative to the other). The geometry of the calculation is shown in Fig. 6.3. This is analogous to the calculation of the weak object transfer function in Section 6.1, where the area of overlap of two displaced circles was determined. An analytical expression for the convolution of two annular pupils was derived by O'Neill.[76] Thus the normalised transfer function may be written as

$$C_{12}(m;0) = (1-\varepsilon^2)^{-1}[A + B + C] \qquad (6.11)$$

where

$$A = \frac{2}{\pi}\left\{ \cos^{-1}\Omega - \Omega\left[1-\Omega^2\right]^{\frac{1}{2}} \right\} \qquad 0 \le \Omega \le 1$$

$$= 0 \qquad \Omega > 1 \qquad (6.11a)$$

$$B = \frac{2\varepsilon^2}{\pi}\left\{ \cos^{-1}\frac{\Omega}{\varepsilon} - \frac{\Omega}{\varepsilon}\left[1-\left(\frac{\Omega}{\varepsilon}\right)^2\right]^{\frac{1}{2}} \right\} \qquad 0 \le \frac{\Omega}{\varepsilon} \le 1$$

$$= 0 \qquad \frac{\Omega}{\varepsilon} > 1 \qquad (6.11b)$$

---

[76] E.L.O'Neill, "Transfer function for an annular aperture", J. Opt. Soc. Am., **46** (4), pp. 285-288 (1956).

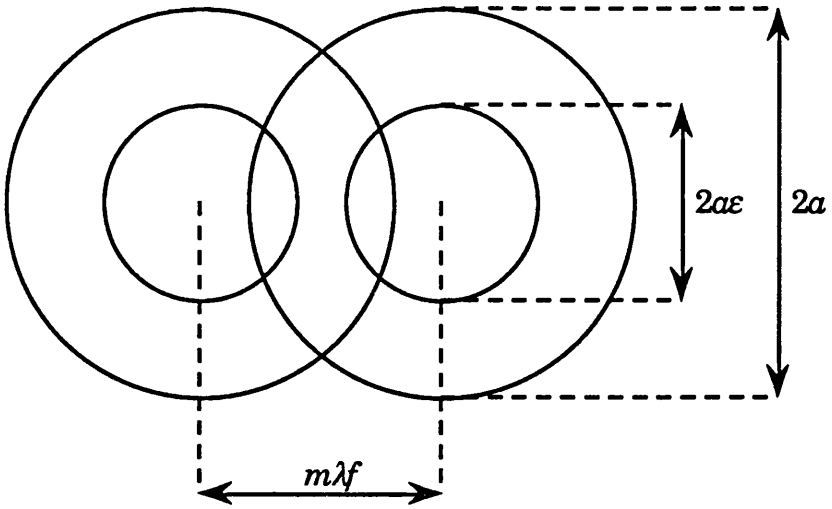


Figure 6.3. Calculation of the transfer function $C_{12}(m;0)$ for an annular aperture, as the area of overlap of two displaced annuli.

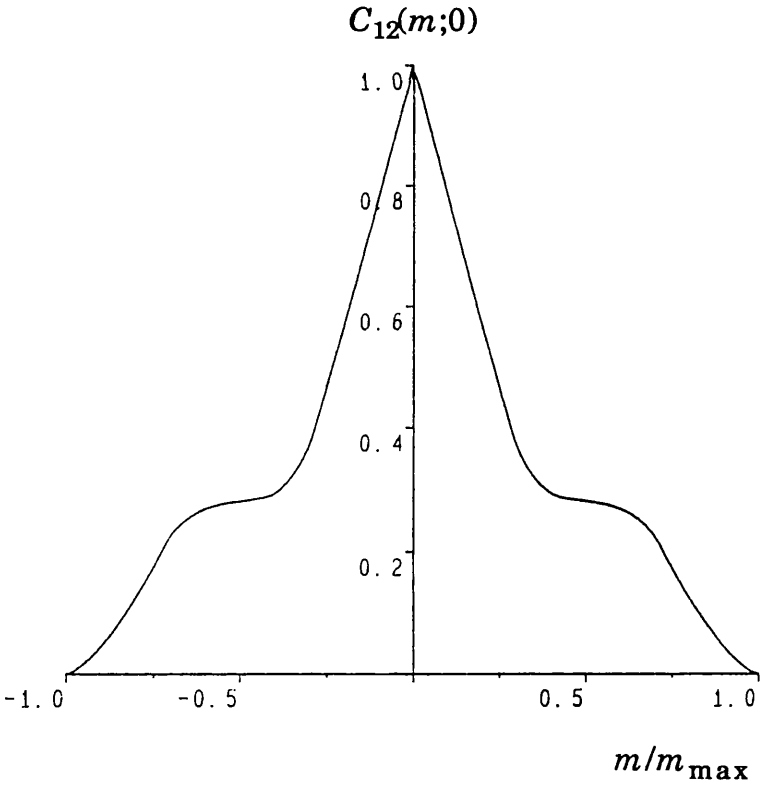


Figure 6.4. The transfer function $C_{12}(m;0)$ for an annular aperture with ε=0.44 versus object spatial frequency m normalised to the cut off value.

$$C = -2\varepsilon^2 \qquad 0 < \Omega < \frac{1-\varepsilon}{2}$$

$$= -2\varepsilon^2 + \frac{2\varepsilon}{\pi}\sin\phi + \left(\frac{1+\varepsilon^2}{\pi}\right)\phi$$

$$- \frac{2(1-\varepsilon^2)}{\pi}\tan^{-1}\left[\left(\frac{1+\varepsilon}{1-\varepsilon}\right)\tan\frac{\phi}{2}\right] \qquad \frac{1-\varepsilon}{2} \le \Omega \le \frac{1+\varepsilon}{2}$$

$$= 0 \qquad \Omega > \frac{1+\varepsilon}{2} \qquad (6.11c)$$

and where

$$\phi = \cos^{-1}\frac{(1-\varepsilon^2-4\Omega^2)}{2\varepsilon}; \qquad \Omega = \frac{m}{m_{max}} \qquad (6.11d)$$

The above expression differs slightly from that quoted in the paper as it contains a typographical error. It has been evaluated for the 0.65 NA reflecting objective used in the experiments. This has a value of ε of 0.44, where ε is the annulus parameter defined previously. The results are shown graphically in Fig. 6.4.

6.3 Response to typical features

In this section the response of the common path interferometer to various objects will be considered. This is achieved by evaluating Eqn. (6.10) for the probe interferometer and by using geometrical optics for the reference interferometer. The overall phase response is then calculated by subtracting the phases of these two signals.

6.3.1 The sinusoidal grating

A weak grating like object may be represented as

$$t(x) = 1 + a\cos 2\pi m x \qquad (6.12)$$

where in general a is complex. The Fourier transform of this object is

given by

$$T_1(m) = \delta(0) + \frac{a}{2}\delta(m) + \frac{a}{2}\delta(-m) \tag{6.13}$$

Substituting for $T_1(m)$ in Eqn. (6.10) we obtain and performing the integration we obtain

$$I_p(x) = 2\mathrm{Re}\{C_{12}(0;0)W^* \exp 2j\omega_B t$$

$$+W^*\frac{a}{2}C_{12}(m;0)\exp 2\pi jmx \ \exp 2j\omega_B t$$

$$+ W^*\frac{a}{2}C_{12}(0;m)\exp -2\pi jmx \ \exp 2j\omega_B t\} \tag{6.14}$$

Using the relations $W = |W| \exp -j\theta$ and $C_{12}(m;0) = C_{12}(0;m)$ this reduces to

$$I_p(x) = 2|W|\,\mathrm{Re}\left\{\left[C_{12}(0;0)+aC_{12}(m;0)\cos 2\pi mx\right]\exp(2j\omega_B t + \theta)\right\} \tag{6.15}$$

The phase of the above signal is extracted by comparing it with the signal from the reference interferometer I_r. The reference interferometer probes the sample with a broad collimated beam, to provide an average reference phase. The spatial wavelength of the sinusoidal object is much smaller than the diameter of the collimated beam, and we may therefore take the phase of I_r to be constant as this sample is scanned (i.e. for periodic objects the reference interferometer provides a truly constant reference phase). In this instance then, the overall output from the system is simply the phase of $I_p(x)$. This is not always the case; for certain objects (eg. a discrete step) there will always be some variation in the phase of the reference signal, as long as the field of view is comparable or greater than the diameter of the collimated beam.

If a is imaginary (i.e. a pure phase grating) the measured phase ϕ_p is

given by

$$\phi_p = \tan^{-1}\left[\frac{C_{12}(m;0)}{C_{12}(0;0)} \mid a \mid \cos 2\pi m x\right] \tag{6.16}$$

and thus

$$\phi_p \approx C_N(m;0) \mid a \mid \cos 2\pi m x \tag{6.17}$$

where $C_N(m;0)$ is the is the normalised transfer function for the interference term. This coefficient serves to reduce the measured value of the peak to peak phase change of the grating. The system will therefore only give quantitative phase information when the spatial frequency of the grating is well below the cut off m_{max}. The degree of degradation of the response for a particular grating is given directly by the transfer function of Eqn. (6.11), plotted for the 0.65 drilled reflecting objective in Fig. 6.4.

6.3.2 The discrete step

A step object is defined simply as a sharp discontinuity in the sample surface properties. This may take the form of a topographic step (i.e. a pure phase object), or a boundary between different materials. The latter example will in general give a discontinuity in both amplitude and phase. A pure phase object may be represented as

$$t(x) = \exp jH(x)\theta \tag{6.18}$$

Where $H(\theta)$ is the Heaviside unit step function and θ is the height of the step. The response of the probe interferometer to this object has been evaluated numerically using Eqn. (6.10) The response for the 0.65 NA reflecting objective is shown in Fig. 6.5(a) and exhibits a slight ringing effect due to the transfer function for the annular aperture. For objectives which are not centrally obstructed such ringing is less

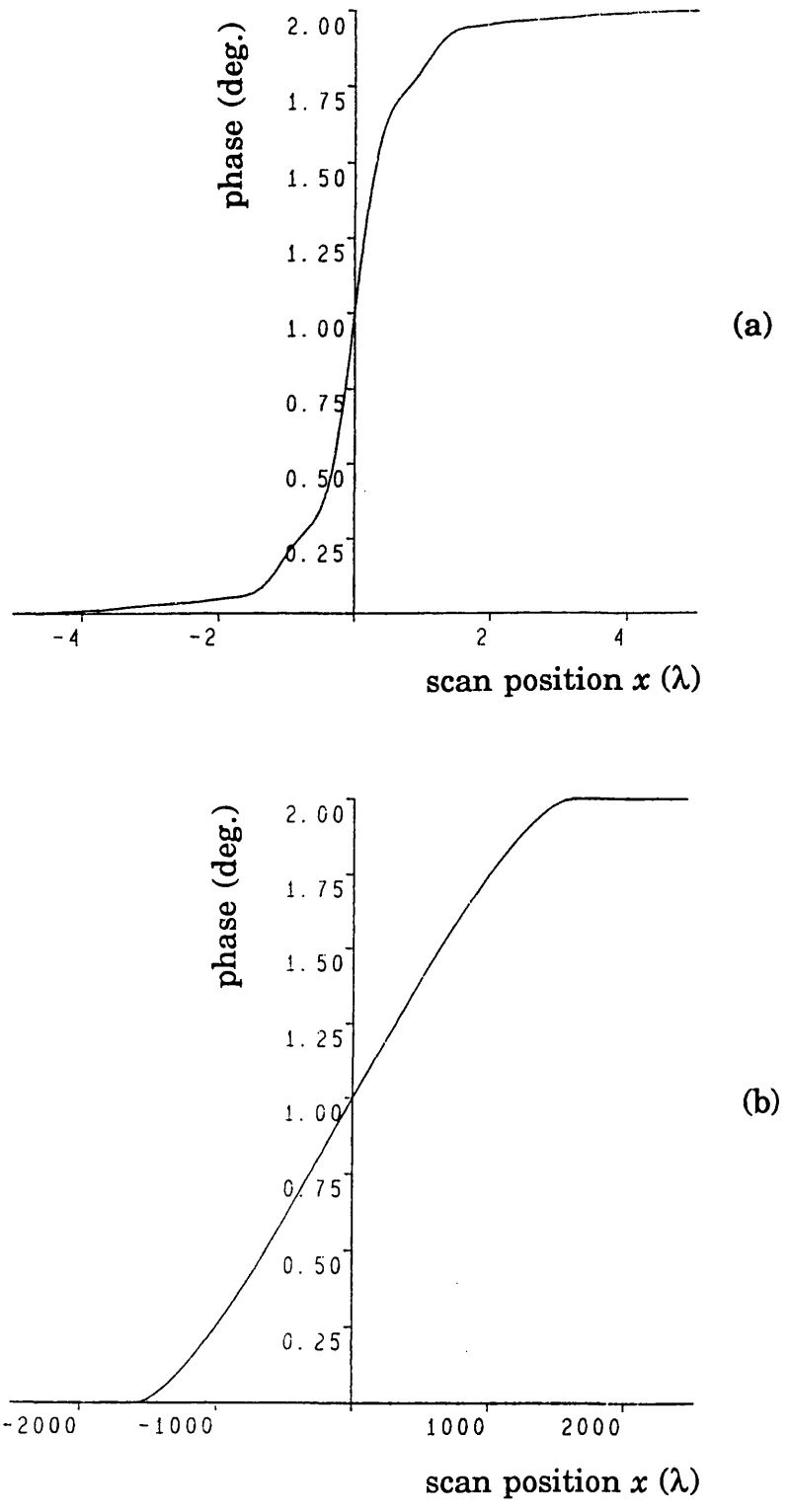


Figure 6.5. Response to a 2 degree phase step for (a) the probe interferometer and (b) the reference interferometer.

apparent, since the smooth decay of the transfer function (shown in Fig. 6.1) has an apodising effect. The ringing need not be as severe for the drilled refracting objective, since the annulus parameter can be typically made as low as 0.25. The annulus parameter for the reflecting objective has a minimum value of 0.44 set by its mirror arrangement and not the drilled hole.

In addition, in order to calculate the overall response, we require the variation in the signal from the reference interferometer as the step is scanned. Fig. 6.6 shows the geometry of the calculation for an arbitrary scan position x. The beam illuminates an area $A(x)$ on the side of the step of phase θ_1. The remaining portion of the beam $\{\pi r^2\text{-}A(x)\}$, where r is the radius of the reference beam, illuminates an area of phase θ_2.

The phase of the returning beam is given by the vector sum of two phasors representing the light from either side of the step. The magnitude of each phasor is weighted by the fraction of the total beam cross section illuminating the appropriate side of the step as shown in Fig. 6.7.

The resultant phase is therefore

$$\theta_R(x) = \tan^{-1}\left\{\frac{A(x)\sin\theta_1 + \left[\pi r^2 - A(x)\sin\theta_2\right]}{A(x)\cos\theta_1 + \left[\pi r^2 - A(x)\cos\theta_2\right]}\right\} \tag{6.19}$$

where

$$A(x) = r^2\cos^{-1}\left(\frac{-x}{r}\right) + x\sqrt{r^2 - x^2} \tag{6.20}$$

The response given by Eqn. (6.19) is plotted in Fig. 6.5(b), as the sample reference beam is scanned completely over the step. It shows a broad, almost linear variation as the step is scanned. The overall system response is given by the difference in the above probe and reference

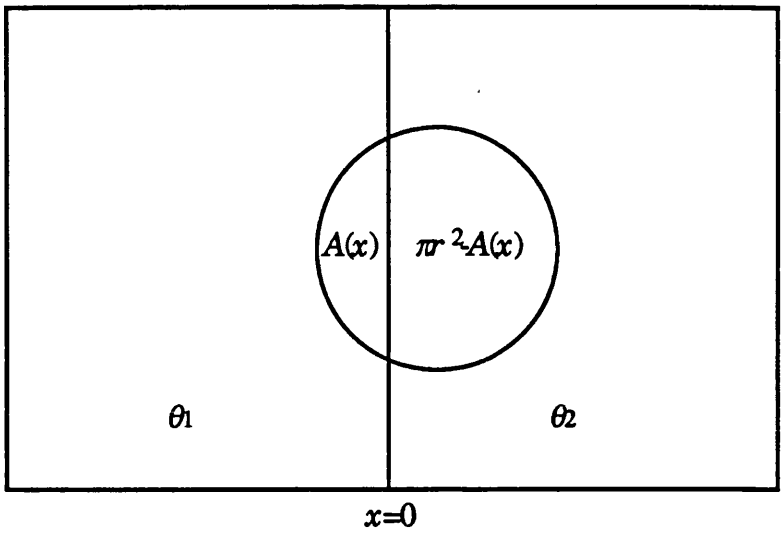


Figure 6.6. Calculation of the phase of the reference interferometer as it scans over a phase step centred on x=0.

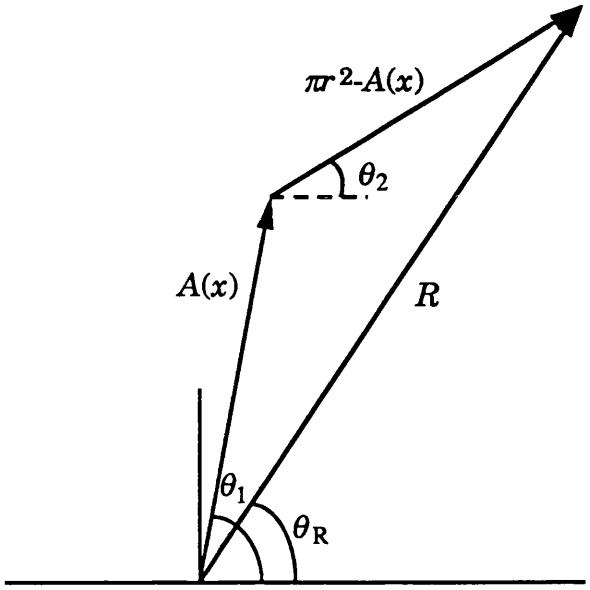


Figure 6.7. Phasor diagram to calculate the resultant phase of the reference interferometer at a scan position x over a step.

signals. This is plotted on two scales in order to fully describe the response. Fig. 6.8(a) shows the response on the scale of the reference interferometer variation and illustrates the limitation of the system in terms of providing the true profile at scan positions far from the step. However in the local region of the step, shown in Fig. 6.8(b) the variation of the reference interferometer is negligible and the system gives a good reproduction of the step. Allowing for the difference in scale, the response shown in Fig. 6.8(a) shows very good agreement with the experimental step response shown in Fig. 6.9. This trace was recorded for edge H on the plasma etched track sample, which was also used for step height measurements presented in Section 3.5 and the images of Section 5.2.3. The 0.65 NA projecting objective was used for the measurement and this gave a collimated sample reference beam diameter of about 0.2 mm, defining the extent of the broad variation of the overall phase response in the region of the edge. This result clearly demonstrates the applicability of the ray theory used to evaluate the response of the reference interferometer.

The true phase step is not quite centred within variation from the reference interferometer and this is result of the hole in the lens being drilled slightly off centre. The value of the phase change indicates a step height of 535 Å, and this compares with a value of 661 Å recorded with the lens singlet (from Table 3.1). Part of this discrepancy is undoubtedly genuine, since variations of over 40 Å have been observed between measurements at different positions along the step. However, there is also a clear systematic difference between the measurements recorded using the low power singlet and the higher power reflecting or projecting objectives. This is a consequence of the obliquity effect discussed previously in Section 4.2.1.

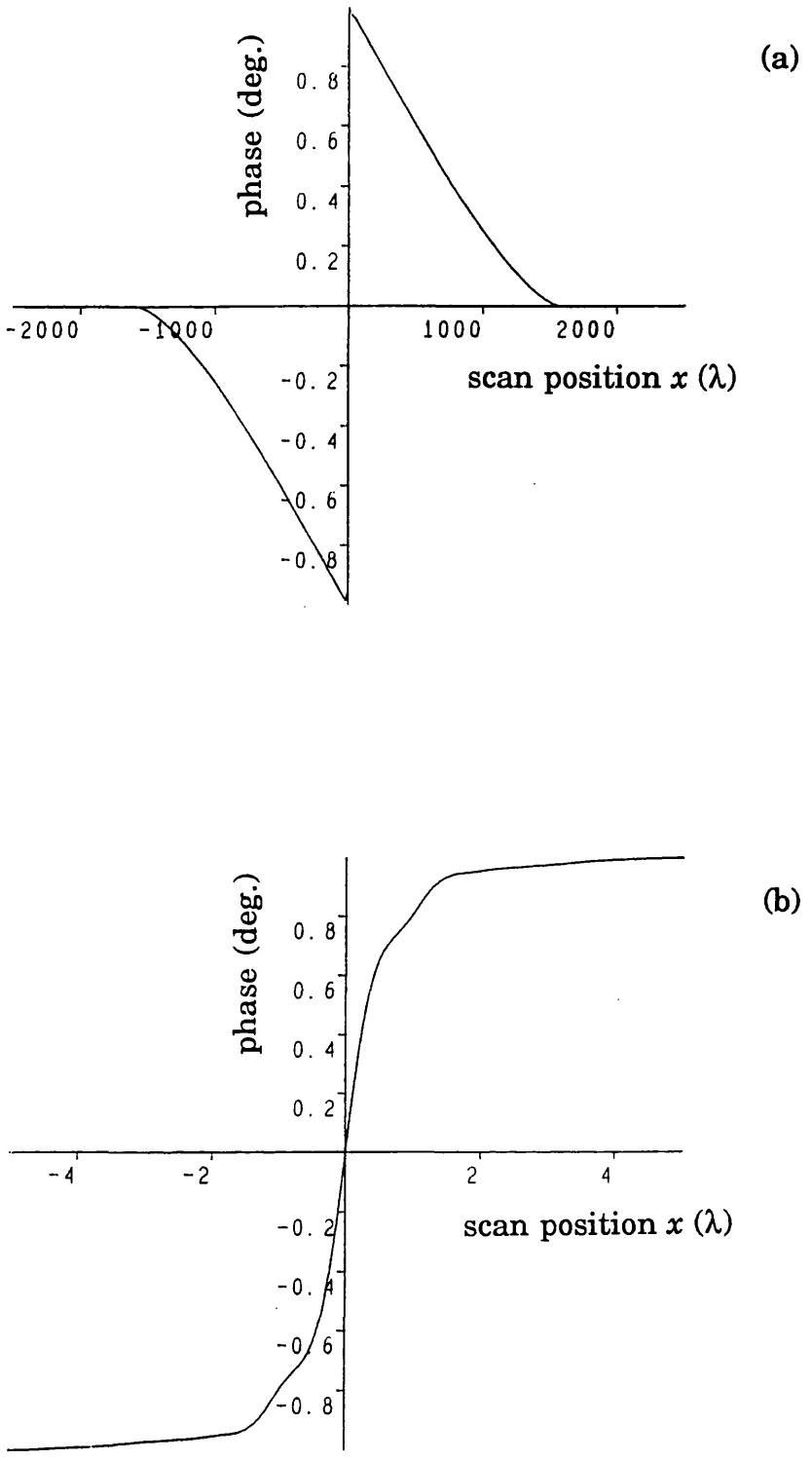


Figure 6.8. Overall system response to 2 degree phase step, (a) over a scan distance comparable with the diameter of the sample reference beam and (b) in the local region of the step.

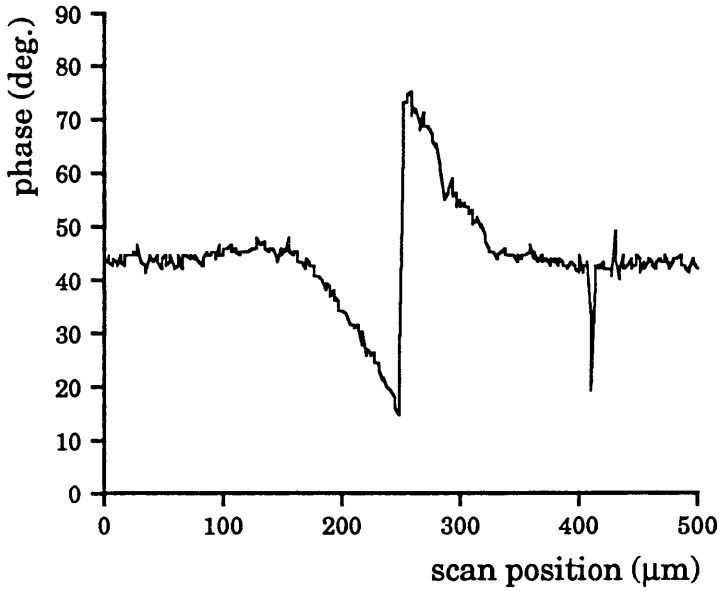


Figure 6.9. Experimental phase response to a topographic step on a plasma etched silicon wafer.

6.4 Summary

The responses for the grating and the step illustrate the extremes of the performance of the reference interferometer, in terms of providing a constant reference phase. For surfaces with fine periodic or random structure the reference phase does remain constant and so the phase measurement is truly absolute. A discrete step does produce significant distortion the reference phase when the system scans far from the edge. This is not a severe limitation, since many applications only require accuracy in the local region of the step. The system is therefore ideally suited to the measurement of step heights, as demonstrated by the experimental results of Chapter 3. However, care must be taken in the interpretation of the phase response when using high power objectives due to the obliquity effect, described in Section 4.2.1. Each objective should be calibrated against standard step samples in order to obtain the most reliable quantitative measurements. The calibration coefficient, known as the NA factor may determined using Eqn. (4.22).

A structure of small lateral dimensions (compared to the diameter of the sample reference beam) on a relatively flat background is imaged well by the system for two reasons. Firstly, the amount of light reflected from the structure is swamped by the rest of the sample reference beam and therefore, there is very little change in its phase as the sample reference beam encounters the structure. Secondly, once the structure is within the field of view of the sample reference beam its position is immaterial and the overall reference phase is virtually constant. Thus there is a near perfect absolute phase response for scan positions within a circular region, centred on the small structure of radius equal to that of the sample reference beam. At the perimeter of this region there is a slight change in phase as the structure leaves the field of view of the sample reference beam. This argument may be extended to an array of such structures. If their spacing is less than the radius of the sample reference beam, as many of these structures will enter its field of view as those that leave. This results in a true absolute response for scan

distances even beyond the diameter of the sample reference beam. This is a significant result, since it means that the system will cope well with regular structures such as a series of tracks or ridges. These are typical features fabricated for optoelectronic and semiconductor devices.

The effect of the annular objective is to enhance the higher spatial frequencies at the expense of those in the mid-band. This apparent improvement in response over a full lens is offset by the reduction in total optical power, due to the absence of the central region. In the practical system the annulus parameter ε is typically equal to about 0.25 and in this case the transfer function illustrated in Fig. 6.4 is quite similar to that for a full aperture plotted in Fig. 6.2.

CHAPTER 7

CONCLUSIONS AND FURTHER WORK

7. Conclusions and further work

7.1 An overall assessment of the instrument

The primary objective at the start of this work was to develop a sensitive absolute phase profilometer/microscope for the examination of fine surfaces. Originally, an idea was pursued which was a scanning heterodyne system based on the Mirau interferometer. This was eventually abandoned, since the Mirau interferometer is not common path, and consequently the system was sensitive to sample microphonics. The experience gained from this system led to the present design, which gives much more immunity to this limitation, by the use of an on sample reference beam.

The design presented in this thesis is comprised of two scanning heterodyne Michelson interferometers in parallel, facilitated by a specially constructed objective lens. The first, referred to as the probe interferometer illuminates the sample with a tightly focussed beam and the local phase structure is encoded as the phase of an AC signal. The second, designated the reference interferometer is designed to be insensitive to the sample properties. Its purpose is to register the phase noise generated in the optics, which is common to the probe interferometer signal. This noise is then subtracted out in the phase sensitive detection process.

Since the two sample beams are not interfered directly, the constraint that their wavefronts are well matched is eliminated. This leads to the single most important advantage of the technique over existing systems. Since the sample areas illuminated by the probe and reference interferometers are virtually independent (within the constraints of the objective lens) the reference area can be made arbitrarily large. This is required for the reference interferometer to be insensitive to sample phase structure. Thus the microscope provides a faithful absolute phase response with a sensitivity approaching that of a common path

interferometer.

The system is not without its limitations however and these are summarised as follows:

(i) Heterodyne interferometry is one of the most direct methods of extracting sample phase information, but the present lack of a suitable full field phase sensitive detector dictates that some form of scanning is required. In this system, mechanical object scanning is employed, but this is relatively slow compared to full field techniques. Alternatively, optical scanning using galvanometer mirrors may be used, but care must then be taken to reduce off axis lens aberrations.

(ii) In the present configuration, the most important factor limiting the sensitivity of the phase measurement is sample microphonics. This is still a problem , despite the use of an on sample reference beam, because it does not illuminate the sample at the same angle as the probe beam. The effect is more serious for higher power objectives, as the average probe beam angle increases. This limitation is not fundamental however and an objective design which overcomes it is proposed in Section 7.2.1.

(iii) The use of a collimated sample reference beam gives the advantage of a reliable absolute phase measurement, but this has also led to a significant drawback. It makes the system sensitive to sample slope and this can result in a significant phase error as well as an attenuation of the interference signal. A method of overcoming the problem of broad changes in slope (the cause of the phase error) is proposed in Section 7.2.2. Surface roughness with spatial wavelengths below the diameter of the sample reference beam are not a problem in terms introducing a phase error, but do result in an attenuation of the reference signal. This is because the reference rays which are deviated by the sample do not interfere efficiently, i.e. the heterodyning process mainly picks out the undeviated rays. Unfortunately, this has been the most important factor

limiting the type of sample which is suitable for the examination with the technique. The situation could be improved somewhat by separating the two interferometers according to polarisation, rather than spatially, in the manner described in Section 3.1. This overcomes the 3 dB loss in optical power at the second beamsplitter, since virtually all of the light in the two sample beams is directed to the correct detector, instead of half being wasted at each spatial filter. More importantly however, is the control it permits over the amount of common reference beam power that is directed to each detector. This is achieved by adjusting the polarisation direction of the common reference beam using a rotatable quarter wave plate. In this way, extra power may be diverted to the reference detector in order to boost the reference signal when examining rough surfaces. With birefringent surfaces, the spatial filters would also have to be used, in order to eliminate crosstalk. This is advisable anyway, as they also intercept light from the part of the common reference beam that does not contribute to the interference signal at each detector, thus minimising the level of shot noise.

In summary then, the system in its present form is best suited to highly reflective surfaces whose relief is such that a reasonable proportion of the incoming collimated sample reference beam is reflected back along the same path. Examples include mirrors, polished metal surfaces and measurements of structures on silicon wafers such as tracks, step heights and waveguides for optoelectronics. The technique does not rely on polarisation and is therefore suited to the characterisation of birefringent surfaces. The present initiatives in the field of nanotechnology should also result in important applications of this instrument in the future.

7.2 General conclusions

From the foregoing chapters, the following conclusions may be made as to the various methods of achieving common path performance:

(i) In order to eliminate the effect of piston type microphonics, it is not essential to have a true common path interferometer with the associated constraints on how absolute the phase measurement is. The rule to follow when designing such an instrument is that there should be at least two beams propagating through all components. As far as possible, the configuration should be arranged so that both beams are affected equally by the vibration. Methods of distinguishing the two beams may include polarisation, optical frequency and spatial separation. If the beams are from the same interferometer, the phase noise is removed optically by the common path effect. If the beams are from different interferometers, the phase noise is detected and subsequently removed electronically by subtraction. This rule covers the three possible scenarios of (a) pure optical subtraction by a single common path interferometer such as the Downs technique discussed in Chapter 2, (b) the parallel interferometer technique described in this thesis which is virtually pure electronic subtraction, (c) the combined use of optical and electronic subtraction to give overall immunity to microphonics, such as the Huang system, also described in Chapter 2. A system should, as far as possible be designed to subtract out the phase noise optically, since the electronic method requires that the detection electronics of each interferometer have exactly the same transfer functions. However as demonstrated in this thesis, electronic subtraction does allow more freedom to design a system which provides a more absolute phase measurement.

(ii) There are two ways of providing a sample reference beam whose phase does not vary as the probe beam scans the surface. The first way is to keep it in a fixed position on the sample. This has the advantage that it provides a true constant reference phase, independent of the sample structure. Also, if the reference beam is focused by the same objective as the probe beam,as in the systems of Sommargren and Jungerman from Chapter 2, there is no obliquity effect to give sensitivity to sample microphonics. However, both of these systems have their drawbacks: the circular scanning in the Sommargren system and

problems with the signal processing and alignment with the Jungerman technique. The second approach is to use an on sample reference beam which averages the surface structure over a large area and thus its phase remains constant as it scans the surface. Previously, the methods of achieving this were (a) to have both the probe and reference beams focused on the sample but with different numerical apertures i.e. the reference beam has a larger Airy disk, as in the Huang system, and (b) to have the reference beam out of focus at the sample surface so that it illuminates a larger area, as in the Downs system. Both these techniques only provide a sufficiently large reference area, to allow accurate examination of very smooth surfaces. We may now add a third category whereby this limitation is overcome, i.e. the use of a collimated on sample reference beam.

We have still not achieved the ultimate goal of a general purpose absolute phase interferometer with complete immunity to phase noise generated in the optics. However, the technique has highlighted the concept of indirect interference, which was used in present design to overcome the fact that the two sample beams do not overlap spatially. This principle has also been applied recently to a simultaneous differential intensity and phase system[77] which is time division multiplexed. Without the use of indirect interference, the two beams would not interfere because they do not overlap temporally. Furthermore, indirect interference has been employed in the design proposed in Section 7.4, as a means of detecting the interference signal at a fixed heterodyne frequency, when using an acousto-optic device to scan the beam. In previous systems which have used this technique, the performance has been limited by the problem of extracting the phase of a heterodyne signal whose frequency is constantly changing. Thus the idea of indirect interference has emerged as a general principle, which relieves many of the traditional limitations on the design of high

[77] M.S.Valera and M.G.Somekh, "A common path differential intensity and phase profilometer using time division multiplexing", Elect. Lett., submitted for publication (1990).

performance interferometers for microscopy.

7.3 Improvements to the instrument

7.3.1 An objective with immunity to sample vibration

In this section, details of an objective are proposed which should overcome the present limitation of sample microphonics caused by the obliquity effect as discussed in Section 4.2.1. The idea is simply to arrange for the collimated sample reference beam to illuminate the sample at the average illumination angle of the sample probe beam. The suggested lens design to achieve this is shown cross section in Fig. 7.1(a). Basically, an annular region of the curved surface of a planoconvex is straightened off, so that this region acts as a prism which refracts light at the average angle to a large area spot on the sample. The position and angle of the straight surface is determined by a tangent to the curved surface at a radial distance defined by the incident ray which is focused by the lens at the required average angle.

There are two factors which must be considered in order to put the idea into practice. These are firstly. the fabrication of such a lens and secondly the determination of the average angle. The fabrication of the lens is certainly possible as a Fresnel lens or by moulding, the only question then being the quality of the final product. It may however be possible to make such a lens in glass. If not, an alternative would be to take a conventional planoconvex lens and polish an annular region flat as shown in Fig. 7.1(b). Although this would work, it is an approximate solution, since some light entering the lens near the edges of the flat region would encounter the curved surface of the lens on the return journey and would not be recollimated. This effect is not shown in Fig. 7.1(b) for the sake of clarity. The stray light would not be a problem in terms of crosstalk, since it would not interfere efficiently with the common reference beam. However, there will be an slight loss of optical power in the sample reference beam. For a low NA lens, it should not be

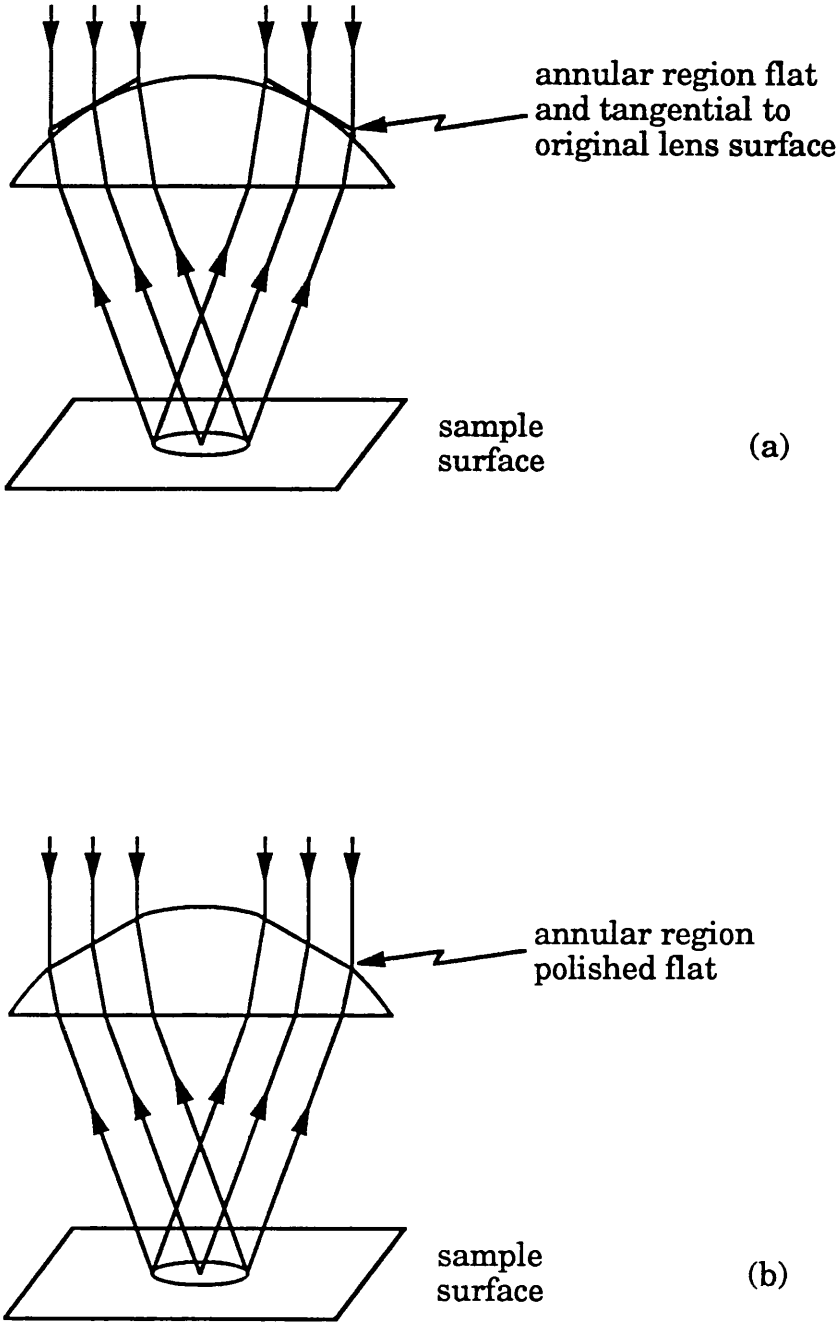


Figure 7.1. An objective with complete immunity to sample microphonics (a) the ideal lens design (b) an approximate solution which may be easier to fabricate.

a problem and this is desirable anyway to reduce spherical aberration. Higher powers may be achieved by projection, in the manner already demonstrated for the drilled lens in Section 3.6.2.

The determination of the position of the average illumination angle requires careful calculation. It is dependent on the width of the straightened region and since the average angle is determined by the remaining curved region of the lens the calculation would need to iterate to the best position. Therefore, the problem would have to be solved numerically, using theory similar to that formulated for the NA factor of an annular lens presented in Appendix A4.1. In its present form, this theory does not account for a Gaussian input beam and this should be included to achieve complete immunity to sample vibration. Experiments could also be performed to help in the determination of the average angle and to gauge the effectiveness of a completed lens by comparing step heights measured with the instrument to the true values. This allows the determination of the NA factor as described in Section 4.2.1.

The diameter of the area illuminated by the sample reference beam is dependent on the radial width of the annular prism region. It is therefore a factor of two less than the diameter provided by a lens with a central hole of the same radius. Since the beam profile is Gaussian however, the width of the annular region should be increased significantly to maintain equal power division between the two sample beams, so this should compensate. In addition, the reference area will be increased slightly due the oblique illumination angle introducing a $1/\cos\theta$ factor. In summary then, an objective has been proposed which is capable of achieving complete immunity to sample microphonics in the axial direction. This is highly significant, since this is the factor currently limiting the sensitivity of the phase measurement.

7.3.2 A mechanism to correct for sample slope

This method involves the use of servo mechanism to control a piezo driven tilt stage, which would correct for slope as the sample is scanned. A third beamsplitter is required to tap off a portion of the returning sample reference beam to a split detector. The idea is that any change in slope as the sample is scanned will cause beam walk off at the split detector, resulting in an error signal, which may be used to control the tilt stage. The optimum configuration to correct for tilt in both the x and y directions would be to use a detector divided into three segments, each at 120^{o}. This provides three error signals from each pair of adjacent segments, and these can be used to control the three piezo legs of a tilt table which are also 120^{o} apart. If a three segment detector is not available, a similar arrangement could be devised using a quadrant detector.

It might be argued that this mechanism cannot be justified in view of the fact that others systems are available which are not as sensitive to tilt. The advantage this system offers however, is the quality of the absolute phase measurement and this, combined with a tolerance to slope would be useful for a number of applications. Examples include ball bearings and parabolic mirrors, where there is a fine surface finish superimposed on a dramatic curvature. Using the above technique, it should also be possible to extract the broad height variation by processing the servo signals to the tilt stage. This could then be combined with the surface roughness measurement from the interferometer. Essentially, this combines the large dynamic range offered by geometrical techniques (as discussed in Chapter 2) with the sensitivity of common path interferometry.

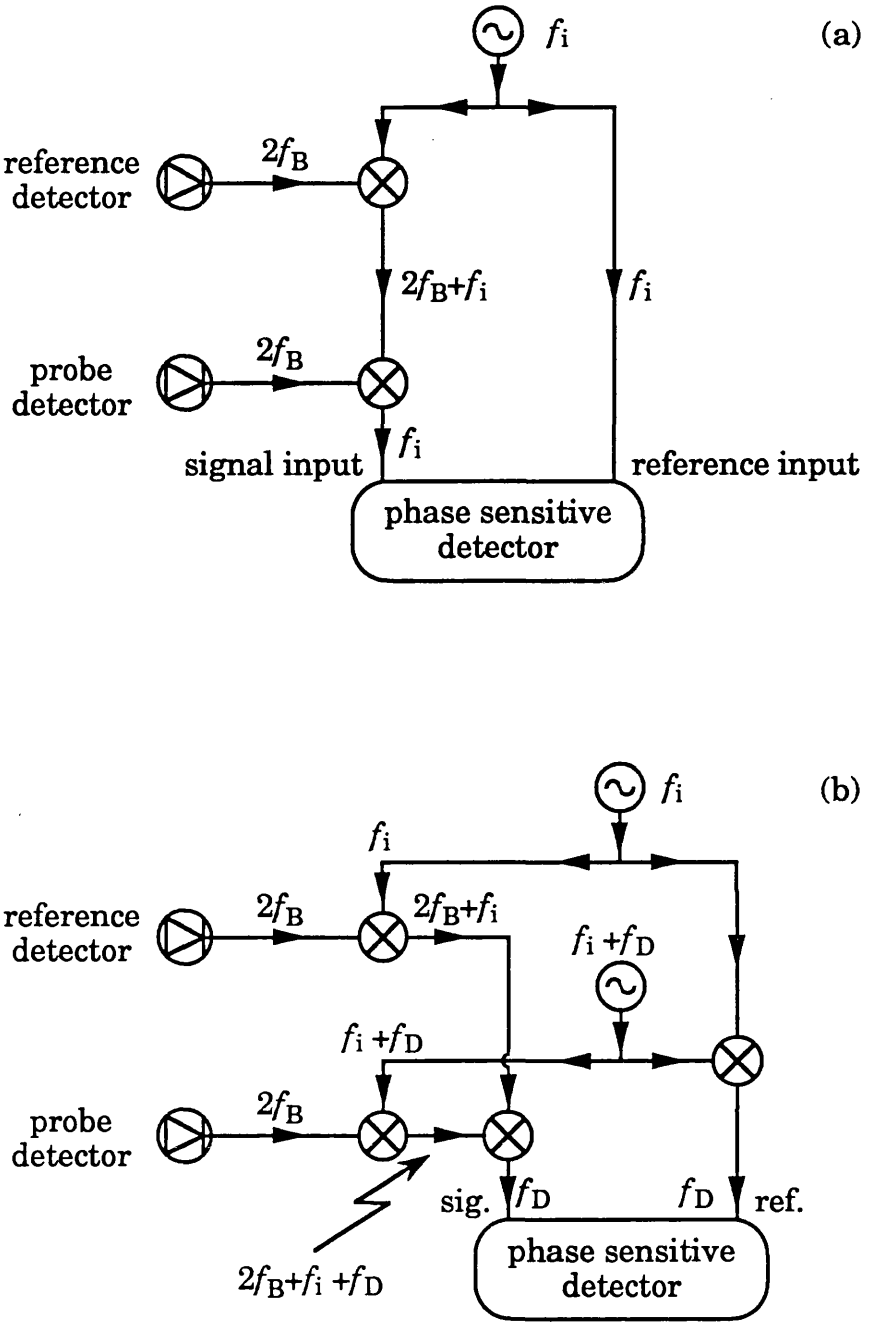
7.3.3 An improved signal processing scheme

All of the experimental results from the interferometer presented in this

thesis have been recorded using a vector voltmeter and a lock-in amplifier to extract the phase. The vector voltmeter is used to beat the two detected heterodyne signals from 80 MHz down to 20 kHz, which can then be applied to the lock-in amplifier. Although this is a convenient solution, it is not ideal, since the optically generated phase noise is not subtracted out until a late stage of the processing scheme. This increases the likelihood that the two noise terms will not be entirely correlated and consequently, the cancellation is not as efficient as it could be.

The rule to follow therefore, must be to perform the subtraction as early as possible, to achieve maximum common mode rejection. Two similar schemes which do this are outlined in Fig. 7.2. The simplest method, shown in Fig 7.2(a) uses a single oscillator to provide an intermediate frequency f_i. This is multiplied with one of the heterodyne interferometer signals using an upper sideband mixer, to produce a resultant at frequency $(2f_B+f_i)$, where f_B is the Bragg cell drive frequency. This is then multiplied with the remaining heterodyne signal using a lower sideband mixer. The subtraction brings the frequency of the resultant down to the intermediate frequency, f_i and also removes the optically generated phase noise. This noise is replaced with the electronic noise from the oscillator however, though it is much less severe. By performing the phase sensitive detection (in the usual manner described above) using the oscillator as the reference, this noise is also removed.

The second method allows the signals to mixed down to a frequency low enough to enable them to be applied directly to the lock-in amplifier (i.e. ~100 kHz), rather than use the vector voltmeter. This is difficult with the previous method, since the sidebands produced by the mixing would be only 100 kHz each side of the 80 MHz heterodyne frequency. The arrangement shown in Fig. 7.2(b) does not require such fine filtering. Each heterodyne signal is first mixed with separate oscillator frequencies f_i and (f_i+f_d) where f_d is the required final frequency.



Note: the filters used to isolate the labelled frequency components have not been shown, for the sake of clarity.

Figure 7.2. Improved signal processing schemes (a) Simple approach that removes optically generated phase noise (b) A method which also beats the signals down to a frequency compatible with a lock-in amplifier.

The resultant upper sidebands are then multiplied together, using a lower sideband mixer, to perform the subtraction down to frequency f_d. As before this cancels the optically generated phase noise. The two oscillator signals are also mixed together to provide the reference signal to the lock-in amplifier, at their difference frequency. In this way f_i may be chosen to optimise the separation of the sidebands in order to improve the quality of the filtering.

7.4 Further work: A new configuration

This section is concerned with a suggestion for a novel heterodyne interferometer which has a number of important attributes. The optical configuration is shown in Fig. 7.3. The basic structure consists of a Michelson interferometer with identical Bragg cells in each arm. Each Bragg cell is driven with two signals, one at a fixed frequency and one whose frequency follows a staircase profile (for reasons discussed later). This produces two first order diffracted beams from each Bragg cell, one at a fixed angle and one whose angle varies with the staircase signal. A lens in each arm focuses each pair of beams to adjacent points on the sample and reference mirror. Thus, both the sample and reference mirror are illuminated by two focused beams one fixed and one which scans the surface. All four beams pass back through their respective Bragg cells to a single photodetector.

The fixed sample beam (ψ_{fs}) and the fixed reference beam (ψ_{fr}) interfere to form what will be termed the reference interferometer. Similarly, the scanned sample beam (ψ_{ss}) interferes with the scanned reference beam (ψ_{sr}) to form the probe interferometer. Thus the system is comprised of two heterodyne Michelson interferometers. Since these interferometers use the same detector, they must be separated by arranging for them to have different heterodyne frequencies. The way this achieved is best explained by considering the theory for the signal extraction.

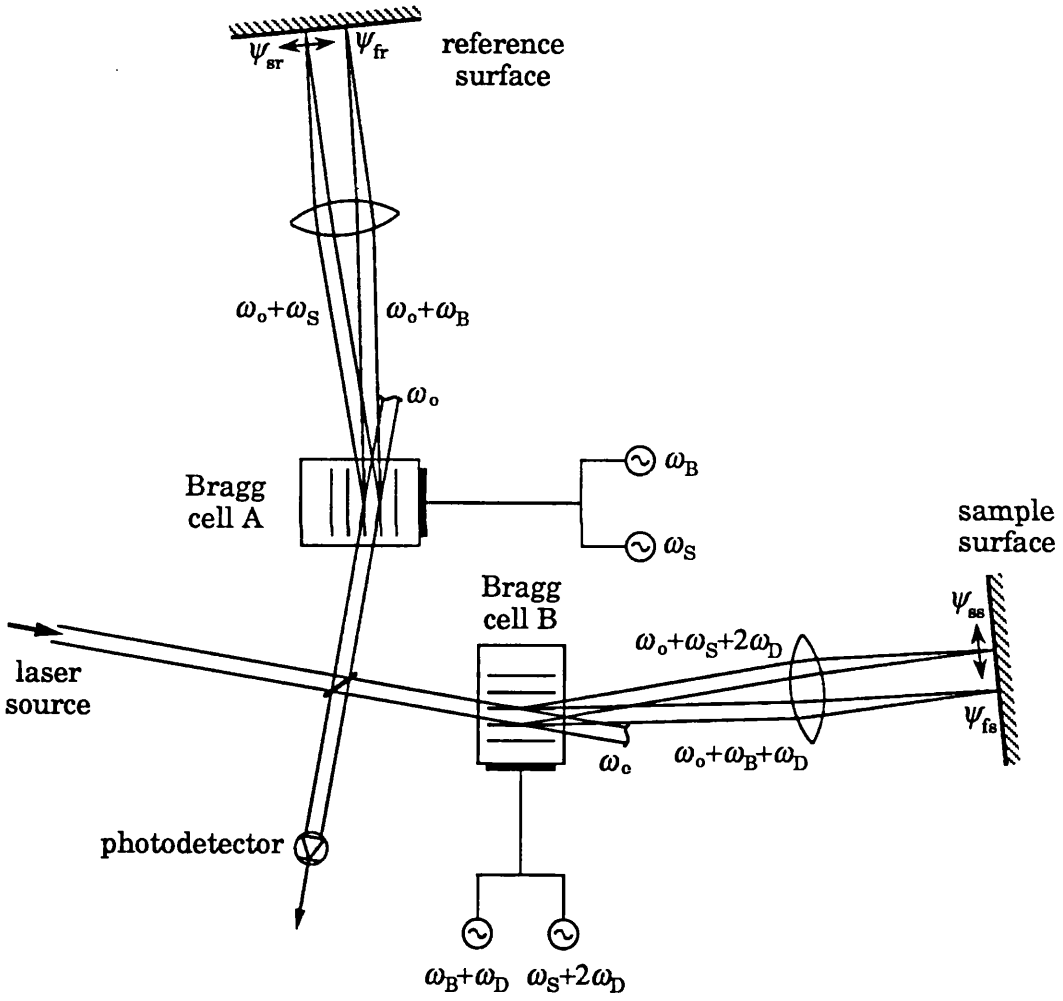


Figure 7.3. Optical configuration of the proposed heterodyne interferometer.

Three signal sources are used to derive the four Bragg cell drive signals, in the manner depicted in Fig.7.4. Source S1 has a fixed frequency ω_B and drives the Bragg cell A in the reference arm to produce the fixed reference beam ψ_{fr}. This signal is also multiplied with the second source S2 at frequency ω_D, using a single sideband mixer. The resulting signal at frequency at $(\omega_B+\omega_D)$ is applied to Bragg cell B in the sample arm to generate the fixed sample beam ψ_{fs}. In addition, source S3 which provides the frequency staircase ω_S, represents the second signal applied to the Bragg cell A and produces the scanned reference beam ψ_{sr}. Finally, the signal from S2 is frequency doubled and then multiplied with that from S3, resulting in a signal at $(\omega_S+2\omega_D)$. This signal, applied to Bragg cell B, generates the scanned sample beam ψ_{ss}. Noting the fact that each of the four sample beams pass through their respective Bragg cells twice, thereby doubling the optical frequency shifts, the four beams may written as

$$\psi_{fr} = A_{fr}\exp j\left[(\omega_o + 2\omega_B)t + \theta_{fr} + m_{fr} + 2n_B\right] \tag{7.1}$$

$$\psi_{fs} = A_{fs}\exp j\left[(\omega_o + 2\omega_B + 2\omega_D)t + \theta_{fs} + m_{fs} + 2n_B + 2n_D\right] \tag{7.2}$$

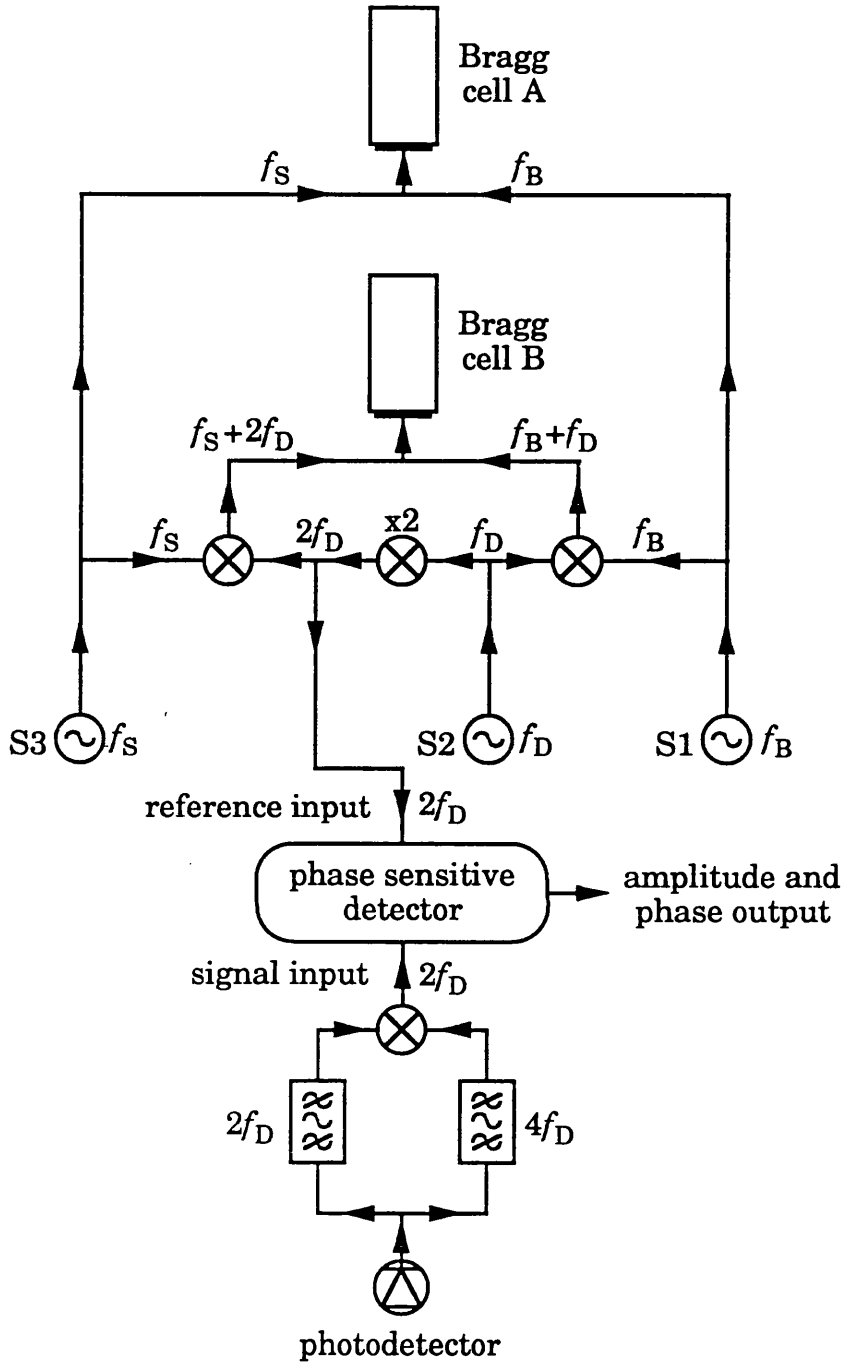
$$\psi_{sr} = A_{sr}\exp j\left[(\omega_o + 2\omega_S)t + \theta_{sr} + m_{sr} + 2n_S\right] \tag{7.3}$$

$$\psi_{ss} = A_{ss}\exp j\left[(\omega_o + 2\omega_S + 4\omega_D)t + \theta_{ss} + m_{ss} + 2n_S + 4n_D\right] \tag{7.4}$$

Where ω_o is the original optical frequency, each θ represents the phase imposed by the reference mirror or sample, each m the optically generated phase noise and each n the electronic phase noise arising from the signal sources. Thus the detected heterodyne reference signal I_r arises from interference of the beams given by Eqns. 7.1 and 7.2

$$I_r \propto A_{fs}A_{fr}\cos\left[2\omega_D t + (\theta_{fs}-\theta_{fr}) + (m_{fs} - m_{fr}) + 2n_D\right] \tag{7.5}$$

Similarly, the detected heterodyne probe signal I_p arises from



Note: the filters used to isolate the labelled frequency components have not all been shown, for the sake of clarity.

Figure 7.4. Signal processing arrangement for the proposed heterodyne interferometer.

interference of the beams given by Eqns. 7.3 and 7.4

$$I_p \propto A_{ss}A_{sr}\cos\left[4\omega_D t + (\theta_{ss}-\theta_{sr}) + (m_{ss} - m_{sr}) + 4n_D\right] \tag{7.6}$$

It should be noted that these signals are both at fixed frequency, despite the use of the frequency ramp to produce a scanning probe beam. The two signals are separated with appropriate filtering and the phase is extracted in the manner detailed in Fig 7.4. Firstly, the two signals are multiplied together to produce a signal I_{sig} given by

$$I_{sig} \propto A_{ss}A_{sr}A_{fs}A_{fr}\cos\left[2\omega_D t + (\theta_{ss}-\theta_{fs}) + (\theta_{fr}-\theta_{sr}) + (m_{ss} - m_{fs}) + (m_{fr} - m_{sr}) + 2n_D\right] \tag{7.7}$$

This multiplication stage removes the optically generated noise, since $m_{ss}\approx m_{fs}$ and $m_{sr}\approx m_{fr}$. The resultant phase may be extracted by phase sensitive detection using the frequency doubled source S2 as the reference frequency, I_{ref}. This is given by

$$I_{ref} \propto \cos(2\omega_D t + 2n_D) \tag{7.8}$$

Thus the remaining electronic noise is removed at this stage and the final phase output θ_{out} may be written

$$\theta_{out} = (\theta_{ss}-\theta_{fs}) + (\theta_{fr}-\theta_{sr}) \tag{7.9}$$

The phases of the two fixed beams θ_{fs} and θ_{fr} remain constant and by using a flat reference mirror together with a low power lens in the reference arm, θ_{sr} is also kept virtually constant. Thus the phase output is dependent on θ_{ss}, the phase of the scanned sample beam and nearly all the phase noise, both optical and electronic, is removed. The phase output from the system should therefore be truly absolute, even for structures such as a discrete edge, since the on sample reference beam remains at a fixed position. It is clearly far easier to maintain a

constant phase for the beam which scans the reference mirror, where the optics and surface quality may be chosen arbitrarily, than to rely on the phase of a scanning on sample reference remaining fixed.

The idea of using a Bragg cell to perform the scanning in a microscope has been used previously by Jungerman et al., in the system already discussed in Section 2.1.2. This system has a number of limitations however, and these are overcome with the proposed configuration. Firstly, the detected heterodyne signal is not at a fixed frequency, caused by the frequency ramp used to generate a scanning beam. This makes the signal processing difficult, and probably accounts for the relatively low sensitivity of the phase response. In addition, a flat area of the sample is required, which can be illuminated with a second set of beams to provide an optical reference. By the use of indirect interference in the proposed design, the signals are detected at a fixed frequency, and the system only requires a flat reference mirror, rather than a flat region on the sample. Furthermore, there is more spatial overlap of the beams, which gives greater immunity to phase noise due to thermal effects.

One problem highlighted by the above workers is that the acoustic delay in the Bragg cell, causes the optical frequency shift of the diffracted beam to lag behind the electrical drive frequency as it is ramped. For this reason, they could not use the electrical signal as a reference. In the suggested system, the limitation is not as severe since only the difference in the delays of the two Bragg cells is important. Thus identical Bragg cells may be used, and the distance from the transducer that each beam enters the respective Bragg cell may be adjusted to minimise the difference. The problem of any residual delay is solved by stepping the frequency in a staircase fashion, rather than using a continuous ramp.[78] As long as the delay is less than the time over which the frequency is kept constant, there will be a time over which the two scanned beams will be separated by the intended optical frequency difference of $4\omega_D$. As an example, two lead molybdate Bragg cells with

[78] I would like to acknowledge Dr. S.Valera at the Department of Electronic and Electrical Engineering,University of Nottingham, for making this suggestion.

acoustic velocity 3630 m/s and centre frequency 80 MHz might be used. A difference in the propagation distance of the acoustic wave in each cell of 10 μm would produce a time delay of 2.8 ns, between the staircases in optical frequency of the two scanning beams. It is reasonable to assume that the distance from the transducer to the traversing beam in each Bragg cell can be made equal to within 10 μm. The resulting difference in propagation distance for the acoustic wave in each cell would produce a time delay of 2.8 ns between the staircases in optical frequency of the two scanning beams. The acoustic delay should not therefore, be a severe limitation on the rate at which the frequency staircase is stepped.