Static Fourier transform spectrometers and laser wavelength meters

Darren Steers

Thesis for the degree of

Doctor of Philosophy

School of Physics and Astronomy,

University of St Andrews,

St Andrews, Fife,

Scotland

May 20, 1999

Abstract

This thesis reports the development of novel static Fourier transform spectrometers based upon Wollaston prisms and compact detector arrays. The path difference between orthogonal polarisation states of the input light varies smoothly across the aperture of the prism forming an interferogram in the spatial, rather than temporal, domain that is recorded with the detector array. A Fourier transform of this interferogram gives the spectral distribution of the incident light. The elimination of moving parts from the design makes the recorded interferogram inherently stable.

I have developed an improved spectrometer utilising Wollaston prisms fabricated from materials of opposite sign birefringence. This new instrument has a significantly increased field of view compared with previous Wollaston prism based spectrometers.

Additionally my work has involved the construction and evaluation of spectrometers for operation in the ultraviolet, visible and infrared regions of the spectrum. These spectrometers have been applied to the analysis of gasolines. Important properties such as gasoline brand and octane number have been identified from ultraviolet and near-infrared spectra using principal component analysis.

Finally, I adapted the spectrometer design to make a fibre-coupled laser wavelength meter based on a modified Wollaston prism. For a narrow linewidth source a fringe period measurement technique, used as an alternative to the Fourier transform algorithm, obtains precision measurement (1 part in 10^6) of the centre wavelength. The wavelength meter and spectrometers have numerous applications in the field of general laboratory instrumentation and their robust, compact nature makes them particularly suitable where field based operation is required.

Certification

(i) I, Darren Steers, hereby certify that this thesis, which is approximately 31000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

Date... May 20, 1999 Signature of candidat

(ii) I was admitted as a research student in 1995 and as a candidate for the degree of PhD in physics; the higher study for which this is a record was carried out in the University of St Andrews betv [illegible]

Date... May 20, 1999 Signature of candidate.

(iii) I hereby certify that the candidate has fulfilled the conditions of the resolution and regulations appropriate for the degree of PhD in the university of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date... May 20, 1999 Signature of supervisor

Copyright

Acknowledgements

I would like to thank those who have helped and supported me throughout this project. First and foremost, my wife Fiona, whose understanding and patience know no bounds, and my son Kieran who shows no patience and understanding. Wilson Sibbett for his advice throughout, and my research colleagues who ably assisted with the design, construction, and programming of the spectrometers.

Bill Hirst for his help, support and patience, and of course, the generous financial support of Shell without which this project would not have happened.

However, before I thank the person who has done most to aid the success of this PhD first let me tell you a short story.

SCENE	It's a fine sunny day in the forest, and a rabbit is sitting outside his burrow, tippy-tapping on his typewriter. Along comes a fox, out for a walk.
FOX	"What are you working on?"
RABBIT	"My thesis."
FOX	"Hmm. What's it about?"
RABBIT	"Oh, I'm writing about how rabbits eat foxes."
	(incredulous pause)
FOX	"That's ridiculous! Any fool knows that rabbits don't eat foxes."
RABBIT	"Sure they do, and I can prove it. Come with me."
	They both disappear into the rabbit's burrow. After a few minutes, the rabbit returns, alone, to his typewriter and resumes typing.
	Soon, a wolf comes along and stops to watch the hardworking rabbit.
WOLF	"What's that you're writing?"
RABBIT	"I'm doing a thesis on how rabbits eat wolves."
	(loud guffaws)
WOLF	"You don't expect to get such rubbish published, do you?"
RABBIT	"No problem. Do you want to see why?"
	The rabbit and the wolf go into the burrow, and after a few minutes the rabbit again returns by himself, and goes back to typing.
SCENE	Inside the rabbit's burrow. In one corner, there is a pile of fox bones. In another corner, a pile of wolf bones. On the other side of the room a huge lion is belching and picking his teeth. (The End)
MORAL	It doesn't matter what you choose for a thesis subject.
	It doesn't matter what you use for data.
	What does matter is who you have for a PhD supervisor.

Thank you Miles Padgett, the best PhD supervisor anyone could have.

Contents

Chapter 1

Introduction

1.1 Spectrometry

Lord Rutherford is said to have made a dictum, that in experimental science one should never attempt anything difficult: the modern research student might reply, with just a touch of bitterness, that all the easy things have been done and that the difficult things have been saved up for him. What is implied really is that one should not lightly attempt research in a field where the experimental technique is poorly developed so that all one's time is spent in persuading awkward apparatus to perform. Fortunately there are some people who take a deliberate interest in doing the difficult things: who make it their business to persuade and coax inanimate matter and to advance the experimental art. They too have made their contribution to the advancement of knowledge, despite a certain amount of condescension by the purists. Quite an array of respectable precedents can be quoted for making this a method of research: to choose a technique, to advance it as much as possible and only then to search for a physical problem that it can be applied to.

Spectrometry belongs to this branch of physics and is basically an experimental subject which is concerned with the absorption, emission, or scattering of electromagnetic radiation by atoms or molecules. Electromagnetic radiation covers a wide wavelength range from radio waves to γ-rays and the atoms or molecules may be in the gas, liquid, or solid phase.

Figure 1.1: *Sir Isaac Newton resolving white light into its component colours using a prism in the* 17^{th} *century.*

Perhaps the oldest form of spectrometery is represented by the ability of the human eye to distinguish different colours. The human eye is sensitive to a comparatively narrow range of the electromagnetic spectrum defined as the visible region. This is roughly the wavelength range 400–700 nm. Radiation at the high-wavelength (low-energy) extreme of the visible region appears red whilst that at the low-wavelength (high-energy) extreme appears blue. It is not surprising that the experimental methods of spectrometry began in the more accessible visible region of the electromagnetic spectrum where the eye could be used as the detector. In 1665 Newton had started his famous experiments on the dispersion

of white light into a range of colours using a triangular glass prism as shown in figure 1.1.

However, it was not until about 1859 that Bunsen and Kirchoff began to develop the prism spectroscope as an integrated unit for use as an analytical instrument. The dark lines in the solar spectrum observed by Wollaston in 1802 and Fraunhofer in 1814 were identified by Bunsen and Kirchhoff. Using their instrument they showed that the dark D line corresponded exactly with the yellow line emitted by a flame into which sodium salts had been introduced. The conclusion being the sun emitted light of practically all wavelengths but that sodium vapour present in the atmosphere surrounding the sun absorbed light at certain wavelengths.

In addition to sodium, the visible spectrum of atomic hydrogen had been observed both in the solar spectrum and in an electrical discharge in molecular hydrogen many years earlier, but it was not until 1885 that Balmer fitted the resulting series of lines to a mathematical formula. In this way began the close relationship between experiment and theory in spectroscopy, the experiments providing the results and the relevant theory attempting to explain them and predict results in related experiments.

The quantum theory of the atom and molecule followed the discovery of line and band spectra, and followed at some distance for most of the time. Band spectra were only discovered because someone invented a bigger and better spectrometer so that the 'fluted' spectra, as they had been known, resolved themselves into lines [1].

Spectrometers can be categorised into three simple classes depending upon their principle of operation, which can be dispersion, diffraction or interference. The most powerful of these operate using interference and form the basis of Fourier

transform spectrometers which are widely used throughout science and industry. They are very powerful tools for making spectroscopic measurements of very weak radiation with high resolution. However, the interferometric nature of Fourier transform instruments usually requires precision scanning mechanisms and highly stable designs, which implies high cost and substantial bulk.

At the University of St Andrews we have therefore designed and constructed compact, static Fourier transform spectrometers. These instruments maintain the optical throughput advantange of interference based spectrometers, yet do not have the problems associated with moving parts.

1.2 Static Fourier transform spectrometers

The need for moving parts in a FT spectrometer can be eliminated with an interferometer that forms an interferogram in the spatial rather than the temporal domain. The resulting interferogram can be recorded with a multielement detector connected to a computer for data processing. A number of groups have designed static FT spectrometers utilising birefringent optical components. The operating principles of such instruments will be discussed in detail later. In essence, a Wollaston prism introduces a path difference between orthogonally polarised components of the incident light that varies linearly with lateral position. Consequently, when illuminated between crossed polarisers, interference fringes are produced parallel to the sides of the Wollaston prism. The operating principle results in a robust and compact Fourier transform spectrometer having no moving parts.

This thesis reports the design and construction of several static Fourier transform spectrometers operating in the ultraviolet, visible and infrared regions of the spectrum. These spectrometers have uses not only in general laboratory in-

strumentation but are suitable for incorporation into more complex devices. We have evaluated their use for the analysis of gasolines, pollutant gas and vapour detection. Where the spectrometer's compact size, robust construction, high optical throughput and ease of use make it ideal for the industrial environment.

1.3 Static laser wavelength meters

The measurement of laser wavelength has been a growing area of optical instrumentation in recent years. The development and widespread utilisation of tunable lasers in many scientific and industrial areas has brought the need for instruments that can conveniently and accurately measure their wavelengths. The power and potential of tunable laser sources cannot be fully exploited unless the laser wavelength is accurately known. Unfortunately, conventional methods for measuring optical wavelengths tend to be either inadequate in their resolution and accuracy or else cumbersome and slow.

Using the same optical assembly as the static spectrometer I have constructed a laser wavelength meter. When used as a conventional spectrometer a Fourier transform of the recorded interferogram gives the spectral distribution of the incident light. Unfortunately, the static nature of the instrument limits the number of fringes that can be recorded in the interferogram. The Fourier transform method is problematic with few data points as it yields information about the period of a sinusoid only indirectly and is known to suffer from systematic errors, or bias, which interfere with the real spectrum[2]. Alternatively, for a narrow linewidth source, such as a laser, I have used a fringe period measurement technique to give precision measurement of the centre wavelength.

Chapter 2

Conventional Spectrometers

2.1 Fundamental concepts

There are many ways of classifying spectroscopic instruments. They can be considered according to the physical principles that they embody, whether they depend on differential refraction, diffraction or interference; or alternatively according to the function that they are to carry out. That is, whether they are intended for the accurate measurement of wavelengths in an unknown source, or for searching for the radiation from a feeble astronomical object.

From the point of view of the experimentalist the latter classification is the more useful, since it is according to function that the instrument will be selected. Thus there will be a need to compare the merits only of instruments that are suited to a particular purpose. The instrument designer, on the other hand, may well prefer the former type of classification, since each of his types follow a certain set of laws, and many of the design problems are common to the whole class.

The type of spectrometer that is chosen for a task depends on several factors,

some of which are uncontrollable, such as availability or cost. Technically the instrument should be adequate for its purpose, whether it is to measure a bright source at high resolution, or a very feeble source at a lower, but still fixed resolution.

This chapter contains a short description of the different types of spectrometers that are in common use. It is not intended as a thorough description of each spectrometer type but to put the novel spectrometers designed at St Andrews University in context. More complete and in-depth descriptions can be found in many physics textbooks[1, 3, 4].

Before discussing different types of spectrometers I shall mention some important technical factors which govern users' choices. Other factors to be borne in mind while making a choice, in addition to the technical issues, are simplicity and reliability.

2.2 Conventional spectrometer components and properties

Despite the apparent complexity of many spectrometers, much of the optical system is for auxiliary purposes. The part of the optical system that is concerned with the dispersion of the light or for forming an interference pattern is quite simple.

2.2.1 Components

An idealised spectrometer could consist of some or all of the following components:

An entrance aperture. In a conventional dispersion spectrometer this consists of a long narrow slit. It may be of any shape in fact, and the shape and size depend on the particular means that are used to manipulate the light in the spectrometer.

A collimating element. This is to make parallel all the rays passing through one point of the entry slit or aperture. This element may be a lens or a mirror, or it may be an integral part of the main optics used to manipulate the light.

A dispersing/interference element. This is any device which alters the intensity that passes through the system, in a way that depends on the wavelength.

A focusing element. This is to form an image of the dispersion/interference pattern at some convenient focal plane.

An exit aperture. This is a stop at the focal plane which transmits the light from the image that the focusing system has formed. There need not be a real aperture there, many systems are designed so the sensitive area of the detector forms the exit field-stop.

2.2.2 Resolving power

The resolving power of a spectrometer is given by the relation,

$$R = (\lambda/\Delta\lambda)$$

where $\Delta\lambda$ is the minimum separation of two perfectly monochromatic spectral lines that are just distinguishable.

2.2.3 Resolution

The resolution of a spectrometer can be understood with reference to the instrument's spectral response function (SRF). If a spectrum consists of a single emission line at a wavelength λ_1 with unit intensity, then, when it is examined with a spectrometer, the output from a perfect instrument would be $S(\lambda) = \delta(\lambda - \lambda_1)$. No real instrument is perfect and the spectrum that is actually measured will be of the form,

$$S(\lambda) = K.F(\lambda - \lambda_1),$$

where K is a constant which depends on the transmission (see section 2.2.4) of the spectrometer, and the function $F(\lambda)$ is called the spectral response function (SRF).

It is normal to regard the SRF as having the same form at all wavelengths. This is not generally true, but the change of shape of a spectrum line due to instrumental defects is usually small over any practical range of wavelengths.

If an ideal grating spectrometer has no optical defects, then a finite slit width will give a triangular SRF. Beyond the feet of the triangle the intensity should be zero everywhere. That is to say no stray light passes through the instrument. Unfortunately, in practice this is not the case, and the ratio of the areas of the spectral profile that lie inside and outside the feet of the ideal triangle provide a measure of the quality of the design.

Stray light may be unimportant when emission spectra are to be studied, but in absorption spectroscopy it is of great importance and should be kept to a minimum when high-quality absorption spectra are needed.

The triangular shape of $F(\lambda)$ will be obtained usually at comparatively modest resolutions. When approaching the theoretical resolving power of the system,

diffraction effects will begin to modify the line shape. To take a specific example, suppose that the SRF were of the sinc^2 form, the Fourier transform of this is a triangle. From convolution theory applied to laser systems the Fourier transform of the spectrum is multiplied by the FT of the SRF, leading to a reduction in the higher spatial frequency components. These high spatial frequency components correspond to the fine structure in the spectrum, and are irretrievably lost.

On the other hand the SRF may be of the Gaussian form. When calculating the spectrum it is now necessary, due to the Gaussian nature of the SRF, to divide the experimental data points by very small quantities, and the process can be seriously degraded by any noise present. Even so deconvolution is still possible in principle. The information is still there, but it is necessary to have a most precise knowledge of the SRF in order to be able to decode the signal and obtain higher resolution.

Having discussed the SRF $F(\lambda)$ it is possible to give meaning to the term 'resolution'. There is no unique criterion for resolution that holds for every type of spectrometer, but generally one may say that two spectrum lines are resolved if they can be detected as two lines without the aid of deconvolutions or a knowledge of the SRF.

This implies that if the two curves $F(\lambda - \lambda_0)$ and $F(\lambda - \lambda_1)$ cross over just below the 50% level then the two wavelengths λ_0 and λ_1 are just resolved. An alternative definition, very similar to this, is that the curve which represents the combined intensities of the two lines,

$$\Phi(\lambda) = F(\lambda - \lambda_0) + F(\lambda - \lambda_1),$$

should have a dip in its peak which is 20% of the height of the curve. This 20% is an arbitrary figure, based on visual spectroscopy.

The Rayleigh criterion, which gives results very close to the criterion above, is that

two diffraction-limited profiles, which have zeros on either side of their principal maximum of the other, are just resolved if the zero of one is superimposed on the principal maximum of the other.

Thus the 'resolution' of a spectrometer is defined as,

$$\Delta\lambda = \lambda_0 - \lambda_1.$$

and the resolving power as mentioned in section 2.2.2 $R = \lambda/\Delta\lambda$. All the definitions mentioned above are effectively defined on an arbitrary basis and have no firm mathematical basis.

2.2.4 Ètendue of a spectrometer

Ètendue has no connection with the brightness of a source, but is used instead to describe the '*light-gathering-power*' of an instrument. Other terms, synonymous in this sense, are '*luminosity*' and '*throughput*'. The term denotes the amount of light that passes from the source, through the instrument and on to the detector. The amount will depend on

1. The area of the entrance aperture.

2. The solid angle subtended at the entrance aperture by the source.

The ètendue, E and resolving power, R of a spectrometer can be used to define the 'efficiency' of a spectrometer. The efficiency is given by $E_{ff} = R.E$. For many types of spectrometer the efficiency is a constant or a function of the wavelength that is being measured.

This concept can be used as a means of comparing the relative merits of competing systems. For a given instrument at a given wavelength the efficiency is fixed, but

it is possible to trade resolution for luminosity or vice versa. The limit set to this trading is the theoretical diffraction limit to the resolving power of the system. In a simple grating spectrometer the light throughput can be increased by widening the slit, or the resolution can be increased by closing it. There comes a point, though, when closing the slit further does not increase the resolution, though the luminosity continues to fall.

2.3 Prism spectrometer

As already mentioned, the earliest scientific record of the use of a prism as a dispersing element appears to be that of Newton. The earliest record of the use of a slit and collimating lens to achieve a respectable resolving power is that of Wollaston, who discovered the 'Fraunhofer lines' in the spectrum of the sun. It was Fraunhofer who first studied these lines and gave a correct explanation of their origin. It is also to Fraunhofer that we owe the design of the classical prism spectrometer which is illustrated in figure 2.1.

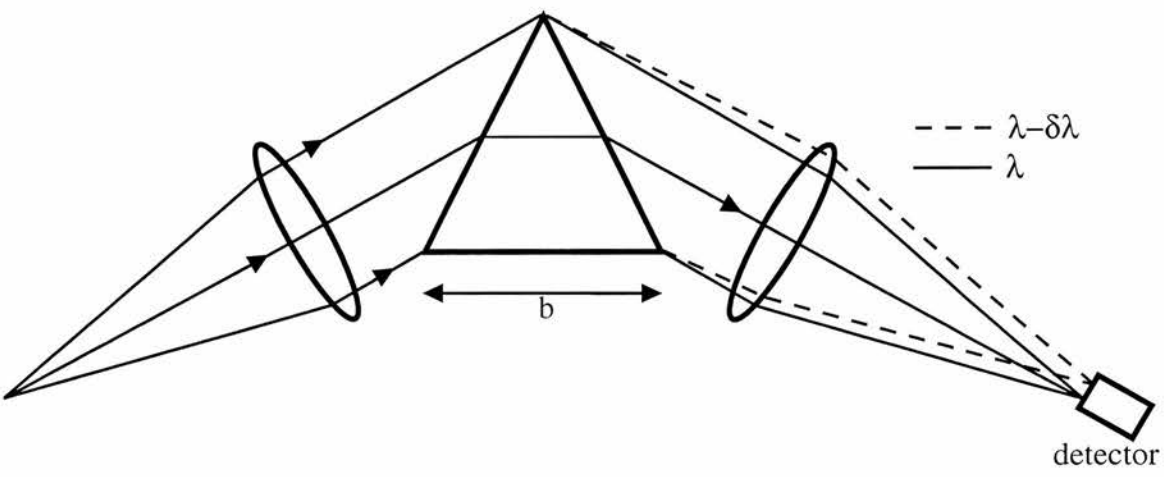

Figure 2.1: *The classical prism spectrometer arrangement.*

For serious work at high resolutions this design is largely superseded by the plane-

grating spectrometer, but for certain applications it possesses some advantages over grating instruments. The absence of overlapping spectral orders and the freedom from 'ghosts' make it suitable for Raman spectroscopy, where a search must be made for faint emission lines very close to a strong parent line. With careful selection of the prism material (to be free from defects) and careful polishing of the refracting surfaces the stray light level can be brought down to a lower level than is possible with the best grating spectrometer, and hence there is still a use for the prism spectrometer in absorption spectrometery. However, the highest resolution that is possible is an order of magnitude less than that which is possible from a grating spectrometer.

The dispersing element usually consists of one or more prisms each with a 60° apex angle. There is no particular magic about the choice of 60°, but it is not possible to exceed this angle greatly since the shortest wavelength that can be refracted corresponds to the condition,

$$\sin\left(\frac{\theta}{2}\right) = \frac{1}{n},$$

where θ is the apex angle of the prism and n is the refractive index associated with this shortest wavelength. In the case of the 60° prism this refractive index is 2, which includes the region of most available prism materials. Ignoring any losses from reflection the resolving power of a prism spectrometer is obtained from the equation[1],

$$R = -b.(\delta n/\delta\lambda), \tag{2.1}$$

where b is the length of the base of the prism, and $\delta n/\delta\lambda$ is the dispersion. The equation is remarkable for the quantities that it does *not* contain. R does not depend on the prism apex angle, and so not on the length of the refracting side

either. Neither does the refractive index appear but only its rate of change with wavelength.

The étendue of a prism spectrometer is obtained by multiplying the exit slit area by the solid angle subtended by the refracting face of the prism at the exit slit. If the length of the exit slit is taken as $1/50th$ of the focal length (the usual criterion) of the focusing lens, the étendue comes to[1],

$$E = -\frac{b.T}{50}\frac{\delta n}{\delta \lambda}.\Delta\lambda, \tag{2.2}$$

where T is the thickness of the prism parallel to the refracting edge. Both the resolving power and the luminosity quoted here are at the maximum possible resolving power that can be achieved.

The efficiency, or LR product is given by [1],

$$E_{ff} = \frac{T.b.b'}{25.\lambda^2} \tag{2.3}$$

2.4 Grating spectrometer

A logical extension of Young's famous double-slit interference experiment is to increase the number of slits from two to a larger number N, i.e. a diffraction grating.

As for a double slit, the intensity pattern that results when monochromatic light of wavelength λ falls on a grating consists of a series of interference fringes. The *angular separations* of these fringes are determined by ratio λ/d, where d is the spacing between the centres of adjacent slits. The relative *intensities* of these fringes are determined by the diffraction pattern of a single grating slit, which

depends on the ratio λ/a, where a is the slit width. It can be shown that increasing N, (i) does not change the spacing between the (principal) interference fringe maxima, (ii) sharpens the principal maxima, (iii) introduces small secondary maxima between the principal maxima.

A principal maximum will occur when the path difference between rays from adjacent slits is given by,

$$d \sin \theta = m\lambda$$

where $m = 0, 1, 2, \ldots$ and is called the order number. The locations of the principal maxima are thus determined only by the ratio λ/d and are independent of N.

Figure 2.2 shows a simple grating spectrometer. The light from source S is focussed by lens L_1 on a slit placed in the focal plane of collimating lens L_2. The parallel light emerging from this lens falls on grating G. Parallel rays associated with a particular interference maximum occurring at angle θ fall on lens L_3, being brought to a focus on the detector D. A symmetrical interference pattern is formed on the other side of the central position, as shown by the dotted lines. The entire spectrum can be viewed by rotating the detector through various angles.

Grating instruments can be used to make absolute measurements of wavelength, because the grating spacing d is known accurately. Several spectra are normally produced in such instruments, corresponding to $m = \pm 1, \pm 2, \ldots$ This may cause some confusion if the spectra overlap. Furthermore this multiplicity of spectra reduces the recorded intensity of any given spectral line because the available energy is divided amoung a number of spectra.

The disadvantage of the grating instrument can be offset by shaping the profile of the grating grooves so that a large fraction of the light is directed into a particular order on a particular side (for a given wavelength). This technique, called *blazing*,

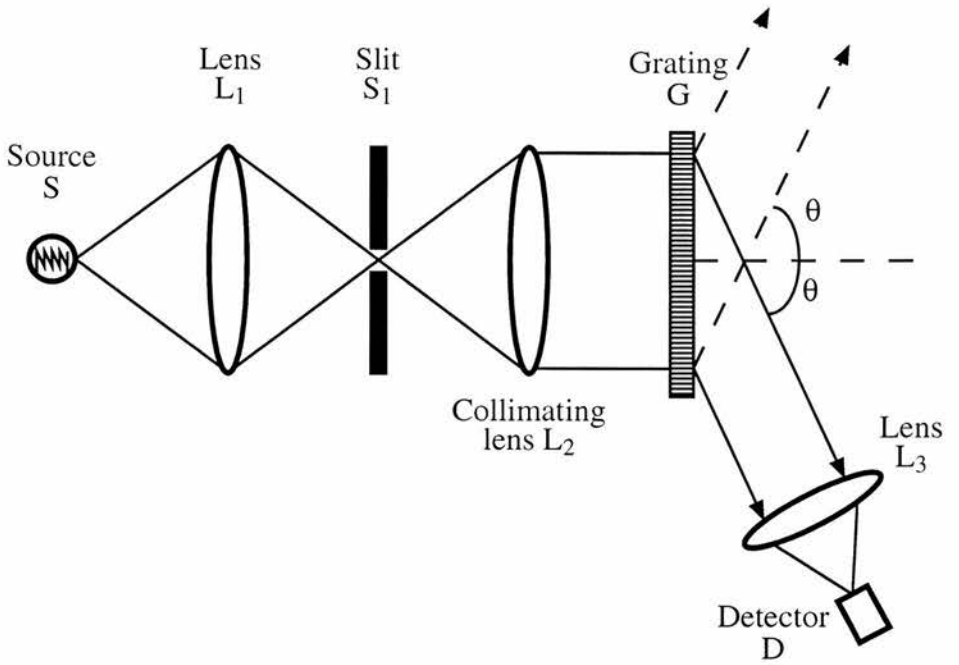

Figure 2.2: *A simple type of grating spectrometer.*

alters the diffracting properties of the individual grooves (by controlling their profiles) such that the light of wavelength λ diffracted by a single groove has a sharp peak of maximum intensity at a selected angle $\theta(\neq 0)$.

To distinguish light waves whose wavelengths are close together, the maxima of these wavelength peaks formed by the grating should be as narrow as possible. To achieve a high resolving power gratings with many rulings are needed. If we use the Rayleigh criterion to determine the resolving power of the grating spectrometer it can be shown that [5],

$$R = Nm, \tag{2.4}$$

where N is the total number of rulings in the grating and m is the order. As expected, the resolving power is zero for the central principal maximum ($m = 0$), all wavelengths being undeflected in this order.

2.5 Fabry-Pérot spectrometer

The Fabry-Pérot étalon is a common device which is explained in nearly every textbook on physical optics. It usually consists of two parallel discs of glass or quartz, with the inner faces worked optically so that the gap between them is constant over the whole area to better than $\frac{1}{100}$th of a wavelength of green light. These flat faces are coated with some material so that most of the light incident upon them is reflected and only a small fraction is transmitted.

If a beam of light is incident upon the étalon from an extended source, the intensity transmitted through it in a direction θ with the normal to the gap is given by the 'Airy formula',

$$I(\delta) = \frac{I(0)}{1 + \frac{4r^2}{(1-r)^2} \cdot \sin^2\left(\frac{\delta}{2}\right)}. \tag{2.5}$$

The independent variable in this equation is $\delta = (2\pi/\lambda)2d.\cos\theta$, where λ is the incident wavelength, d is the optical thickness of the gap, θ is the angle of incidence of the light beam in the gap, and r is the reflection coefficient for intensity of the coatings on the flat surfaces.

By allowing λ, d, or θ to vary while the other quantities are kept fixed, we can use this formula to describe: (i) How the transmitted intensity varies with wavelength when light is incident from a fixed direction. (ii) How the transmitted intensity of monochromatic light varies as the separation of the plates is changed. (iii) How the transmitted intensity of monochromatic light varies with angle of incidence.

The airy function plotted with δ as the variable is shown in Figure 2.3. The half-intensity width of each peak on this graph depends on the value of r, the reflection coefficient, and the peaks become narrower as r increases.

The familiar treatment of the theory of the Fabry-Pérot étalon leads to the de-

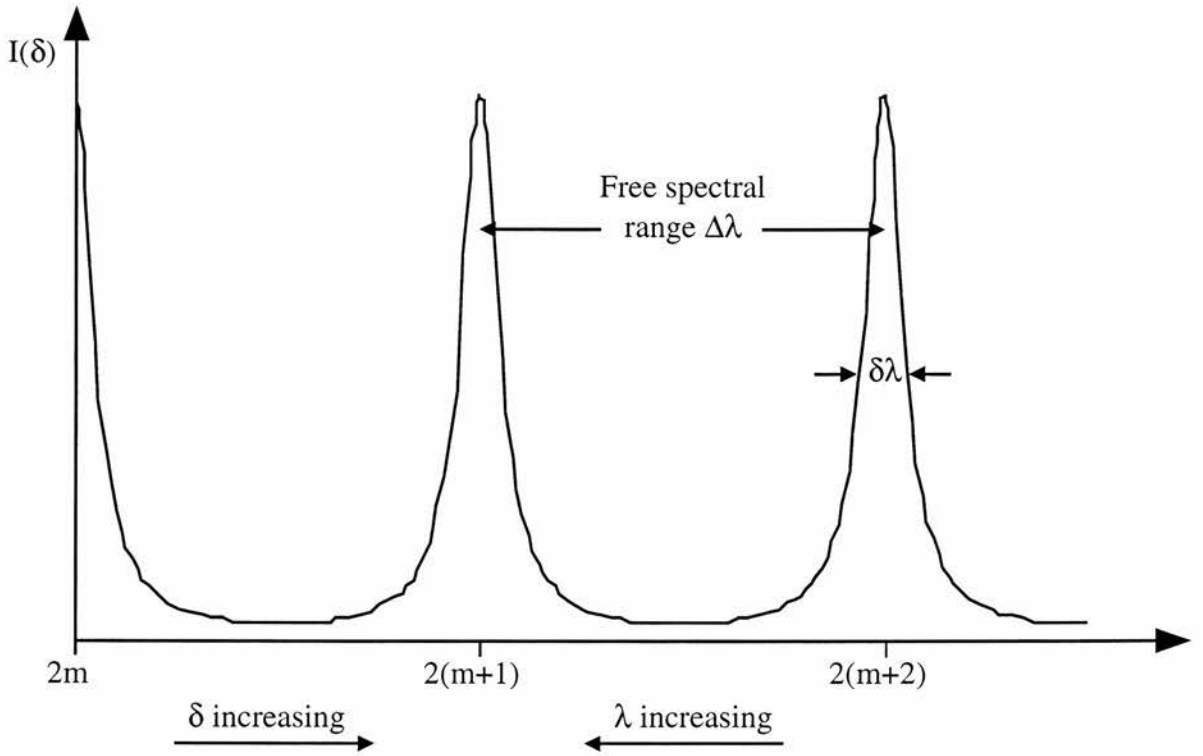

Figure 2.3: *Plot of the Airy function.*

scription of the ring system formed when light is transmitted after coming from an extended monochromatic source. The rings can be seen at infinity by looking through the étalon or they can be projected onto a screen by a lens.

Figure 2.3 shows the instrument profile of the instrument when it is used as a spectrometer. It suffers, like the grating spectrometer, from overlapping orders. We can define,

Resolving power. $R = \lambda/\delta\lambda$ where $\delta\lambda$ is the half-intensity width of one of the peaks.

$$R = \frac{\lambda}{\delta\lambda} = \frac{m\pi\sqrt{r}}{(1-r)}$$

Free spectral range. $\Delta\lambda$ being the distance between neighbouring peaks on the wavelength scale.

$$\Delta\lambda = \frac{\lambda}{m} = \frac{\lambda^2}{2d} = \frac{2d}{m^2}$$

Finesse. F is the ratio $\Delta\lambda/\delta\lambda$, being the number of resolvable spectral elements in one spectral range. It is a function of the reflection coefficient r and the surface flatness of the plates.

$$F = \frac{\Delta\lambda}{\delta\lambda} = \frac{\pi\sqrt{r}}{(1-r)} = \frac{R}{m} = \frac{R\lambda}{2d}$$

Unlike the grating and the prism, the Fabry-Pérot étalon disperses the light passing through it in two directions, both perpendicular to the optic axis. If after transmission the rings are brought to a focus on a screen, different wavelengths produce rings of different radius. Thus the dispersion is radial, and by cutting a circular hole in the screen some wavelengths can be selected for transmission while all the others are rejected. This hole plays the same part as the exit slit in a grating spectrometer.

As a spectrometer, the étalon suffers from a small unambiguous spectral range and the consequent overlapping of orders. Some kind of order-sorting device is needed just as is the case with the grating spectrometer. In this instance a grating spectrometer itself is often used. The effective resolution of the grating spectrometer can be as low as $\Delta\lambda$, the spectral range of the étalon, and so there is a gain in luminosity. The effective resolution of the combined spectrometer and étalon is $\delta\lambda$, the resolution of the étalon. Thus if the two devices are optically matched the RL product is increased by a factor F, the finesse of the étalon plates. The gain in luminosity or resolving power is thus quite considerable, although not as great as the RL product of the étalon itself.

An alternative method of order-sorting is to use interference filters. Filters can be selected for a particular problem with the centre of the pass band sited at a particular wavelength for a particular problem. The basic function of the filter is to convert the spectrum under examination to a line spectrum which consists of a single line about 1 nm wide. The fine structure of this line is what is observed with the étalon. The shape of this line, which is the transmission curve of the filter can be determined using the complete spectrometer when white light is incident upon it. Modifications to this profile are then produced by the unknown spectrum, which can be recovered by simple proportionality.

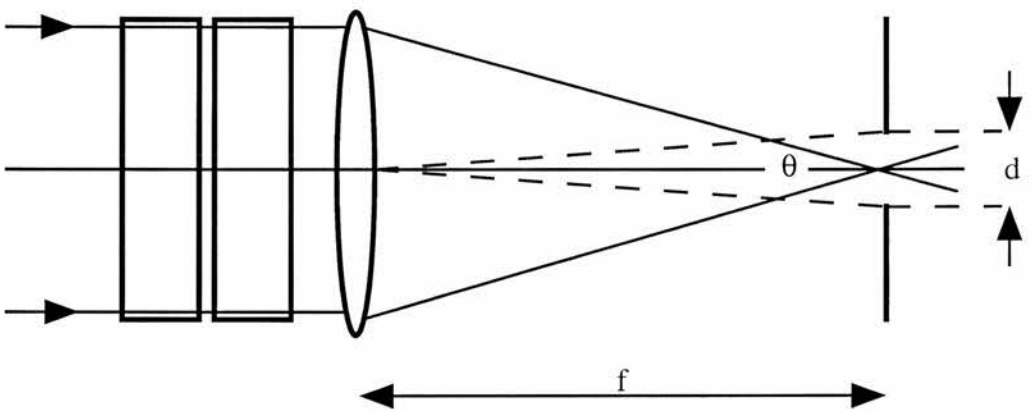

Figure 2.4: *The Fabry-Pérot étalon as a spectrometer.*

The major shortcoming of the Fabry-Pérot spectrometer is that unlike the grating spectrometer, its instrument profile does not fall to zero outside the pass band. Consequently, when white light is incident, not only is the desired wavelength passed in full, but all other wavelengths are present to a finite degree. The integrated effect of all these can be large, even larger than the signal from the desired wavelength [1].

Figure 2.4 shows a basic Fabry-Pérot spectrometer. The actual stop diameter d is given by the equation: $d = 2f\sqrt{2/R}$

For the apparatus to function as a spectrometer it is necessary to be able to change the wavelength at will and in a known way. The three variables in the Airy formula are available, but since λ is the dependent variable, either d or θ must be altered to produce the necessary change. Changing θ implies altering the shape of the field-stop or tilting the étalon bodily. The latter method can be used but results in a drastic loss of luminosity and ultimately resolution too.

Both resolution and luminosity are fully maintained if d, the optical thickness of the gap between the plates, is taken as the variable. Either the physical gap can be changed mechanically or the refractive index of the material in the gap can be changed. The medium between the plates will normally consist of a gas of some sort, and a convenient method of scanning is to alter the pressure of this gas and hence the refractive index. The more common method of changing the gap width is mechanically using piezoelectric scanning.

2.6 Fourier transform spectrometer

Fourier transform spectrometers (FTS) are commonly utilised in many industrial, scientific, and even remote space-based situations [6]. A conventional FTS is equivalent to an interference spectrometer combined with a two-beam interferometer, such as a Michelson interferometer[7] .

The science of Fourier transform spectroscopy was initiated in 1880 when Michelson invented the interferometer[8, 9]. Michelson used the interferometer to measure, with greater precision than anyone before him, the velocity of light in a vacuum. He proved, in the famous Michelson-Morley experiment, that the earth does not move through an ether.

Shortly after the conception of the Michelson interferometer, both Michelson

and Lord Rayleigh recognised that it was theoretically possible to obtain spectra from the interference pattern (*interferogram*) generated by the interferometer through the computation of its Fourier transform. However, performing a Fourier transform on data from polychromatic spectra is an extremely complex and time consuming affair when performed by hand. Since Michelson lacked computers and electronic detectors, he could not exploit the field of Fourier transform spectroscopy.

It was not until about the middle of the 20^{th} century that the Michelson interferometer gained renewed interest with the realisations of Fellgett[10] and Jacquinot[11]. The Fellgett and Jacquinot advantages combine to form the fundamental basis for improved performance of FTS over dispersive spectrometers.

The next step in the acceptance of Fourier-transform spectroscopy was the discovery (or more accurately rediscovery) of the fast Fourier-transform (FFT) algorithm by Cooley and Tukey[12] in 1964. Instead of requiring minutes or even hours to compute a spectrum from an interferogram using the conventional FT, the FFT permitted low-resolution spectra to be computed in a matter of seconds. By the late 1960s, however, two technological developments were made that, when applied to spectrometry, finally brought the FTS into mainstream commercial use. These breakthroughs were the fabrication of microcomputers[13] and small gas lasers. The microcomputer allowed spectra to be computed in the laboratory directly after measurement of the interferogram. He-Ne lasers were used to monitor the travel of the moving mirror of the interferometer, permitting interferograms to be digitised at precisely equal intervals and giving, in effect, an internal wavelength standard for all measurements.

The optical layout of a basic Michelson interferometer is shown in figure 2.5. There are a multitude of variations, but all the different instruments are basically the same[14].

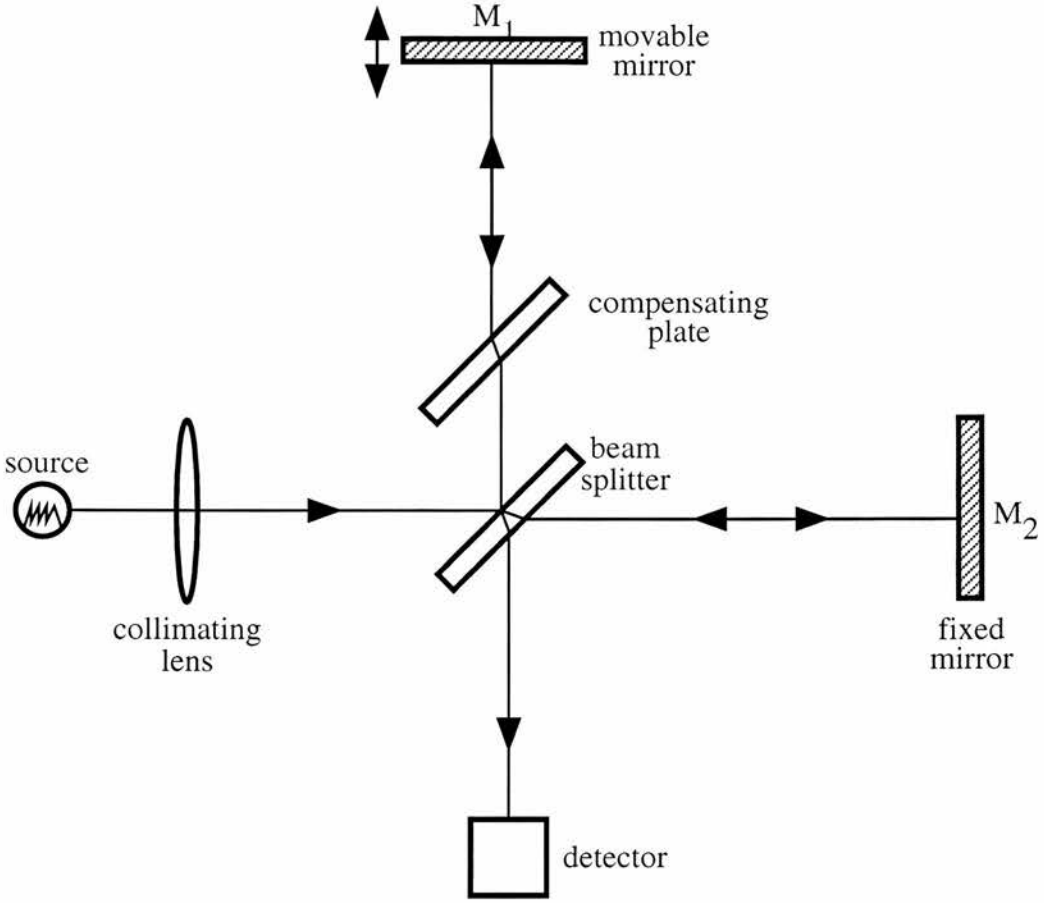

Figure 2.5: *The classical Michelson interferometer arrangement.*

The operating principles of a Michelson interferometer are straight forward. First a source is chosen, which is usually mated with a desired spectral region — mercury lamps for the far-infrared, glowers for the near-infrared, and a variety of lamps for the visible. After the radiation beam is collimated by a mirror or lens, it is amplitude divided at the beam splitter. The beam splitter comprises a semireflecting coating on the surface of a plane-parallel glass plate. The same beam splitter is used to recombine the beams reflected back from the two mirrors. One part of the radiation beam is reflected from the beam splitter to a movable mirror, which reflects the beam back through the beam splitter to a detector. The other part of the radiation beam is transmitted through the beam splitter

to a fixed mirror. This is reflected back to the beam splitter which directs the beam to the detector.

The interference effects are produced by an optical path difference in the two arms. The path difference is twice the arm displacement away from the balanced position. To obtain fringes with a white-light source the optical paths of the two beams must be equal for all wavelengths. Both arms must therefore contain the same thickness of glass having the same dispersion. However, one beam traverses the beam splitter three times while the other traverses it only once. Accordingly, a compensating plate (identical with the beam splitter, but without the semireflecting coating) is introduced into the first beam.

Michelson realized that each wavelength in the interferogram produces its own characteristic interference pattern as the movable mirror is displaced. A monochromatic source yields a cosine variation in the flux of the combined beams at the detector. The period of the cosine function uniquely determines the wavelength and optical path difference for radiation beams in the two arms. Each wavelength has its own characteristic cosine flux pattern with a particular magnitude. The recording of the detected signal versus optical path difference is the interferogram. For a source of many frequencies, the interferogram is the sum of the fluxes of each wavelength pattern. Fourier analysis enables one to convert the interferogram into a spectrum. That is, Fourier analysis of the interferogram picks out the pattern for each frequency and determines the magnitude of the flux at that frequency.

2.6.1 Resolution of a Michelson interferometer

It is useful and fairly simple to illustrate conceptually how the resolution of a spectrum measured interferometrically depends on the maximum retardation or

path length difference of the scan[3]. As an example, consider the case of a spectrum consisting of a doublet, both components of which have equal intensity. Figure 2.6*a* shows the spectrum, and figure 2.6*b* shows the interferogram from each line. Figure 2.6*c* shows the resultant of these curves.

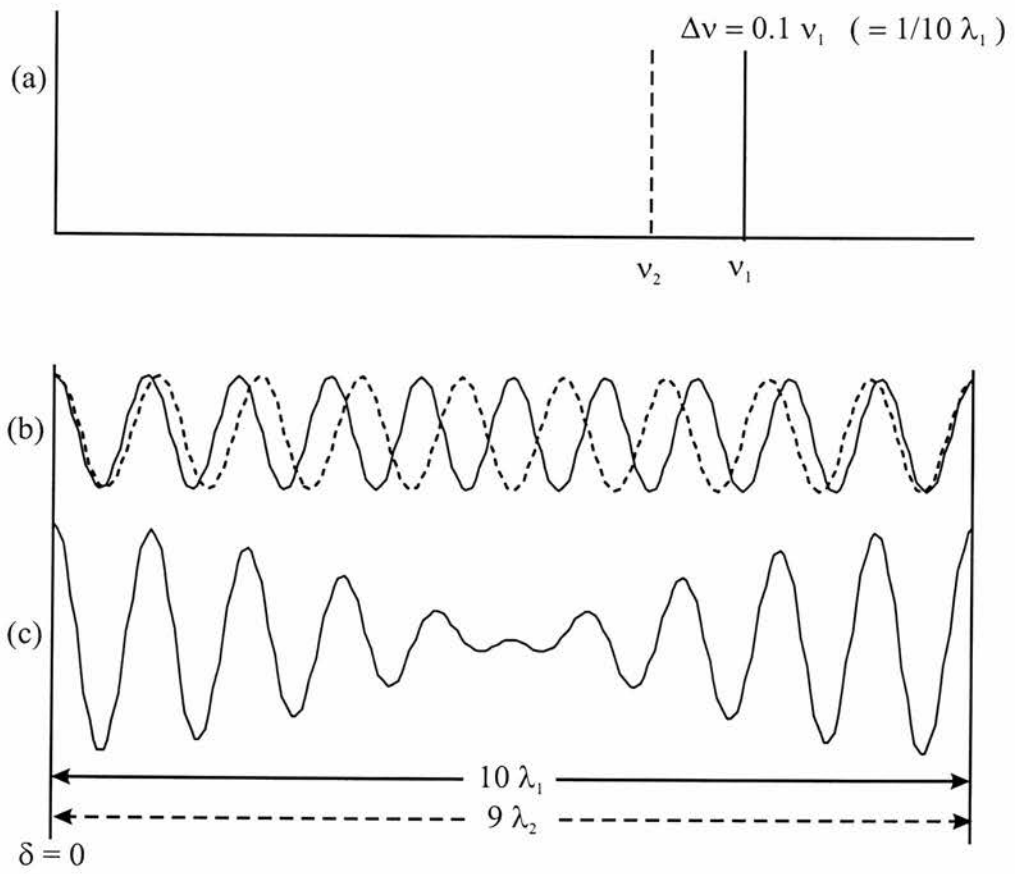

Figure 2.6: *(a) Spectrum of two lines of equal intensity at wavenumbers ν_1 (solid line) and ν_2 (broken line) separated by $0.1\nu_1$. (b) Interferogram for each spectral line shown individually as solid and broken lines, respectively. (c) Resultant interferogram with the first maximum of the beat signal at $10\nu_1$; to resolve these two spectral lines, it is necessary to generate an optical path difference (retardation) of at least this value.*

If the doublet has a separation of $\Delta\nu$ $(= \nu_1 - \nu_2)$, the two cosine waves in figure 2.6*b* become out of phase after an optical path difference of $0.5\,(\Delta\nu)^{-1}$, and

are once more back in phase after a retardation of $(\Delta\nu)^{-1}$. To go through one complete period of the beat frequency, a retardation of $(\Delta\nu)^{-1}$ is therefore required. An interferogram measured to only half this retardation could not readily be distinguished from a source yielding a sinusoidal interferogram within an exponential Lorentzion envelope, for example. The narrower the separation of the doublet, the greater is the optical path difference before the cosine waves become in phase. It is therefore apparent that the spectral resolution depends on the maximum optical path difference of the interferometer. Intuitively it might be concluded that the two lines could just be resolved if the retardation were increased to the point when the two waves became in phase for the first time after zero optical path difference. In the above discussion it was shown that the two waves become in phase for the first time after the zero retardation point when

$$\delta = (\Delta\nu)^{-1} \tag{2.6}$$

Thus, if the maximum optical path difference of an interferometer is Δ_{max}, the best resolution that could be obtained using this interferometer is given by

$$\Delta\nu = (\Delta_{max})^{-1} \tag{2.7}$$

Although this conclusion was arrived at intuitively, the answer proves to be approximately correct. A more rigorous mathematical verification of this conclusion can be found in many textbooks on spectroscopy[7, 3].

The prism, grating, Fabry-Perot and two-beam interferometer spectrometers can be compared in terms of their ultimate resolving power.

The prism has the lowest resolution because its resolving power is limited by the size of the prism. The resolving power is the product of the length of the prism base b and the dispersion of the prism material $dn/d\lambda$. The best resolving powers that can be obtained with a prism-type instrument are of the order of 10^4.

The resolving power of the grating instrument depends on its slit width s, the

focal length F of the collimating mirror in the monochromator section, and the angle of rotation Θ of the grating away from zero order. For good gratings, in the near-IR, resolving powers of the order of 10^5 can be attained. This resolving power has a diffraction limited maximum of Nm, which is the product of the total number of grooves in the grating N and the grating order number m.

Practical Fabry-Pérot instruments achieve resolving powers of approximately 10^6, which, until Fourier transform spectroscopy became widespread was the highest resolution attainable from common spectrometers. However, large arm displacements used with Michelson interferometers allow them to surpass the Fabry-Pérot instruments.

The resolving power depends linearly on the relative arm displacement of the interferometers movable mirror. For example, if a measurement is made at a frequency of $10^5 cm^{-1}$, a displacement of $10cm$ will give a resolving power of 10^6. A $2m$ mirror displacement will give resolving powers of the order of $10^7 - 10^8$. The Michelson interferometer, in its spectroscopic mode of operation, surpasses all other instruments in its ultimate resolving power and in its unhampered wavenumber sorting ability over large frequency ranges.

The principal advantages of Fourier-transform spectrometers over dispersive instruments arise from two major concepts, the Fellgett and Jacquinot advantages.

2.6.2 Fellgett advantage

The Fellgett[10] or multiplex advantage is used to describe the rate at which the spectrometer collects information about the spectrum. If each resolved element of the spectrum is regarded as one information channel then, clearly, a monochromator has a small information capacity compared with a polychromator, since

only one channel is open at one time. By contrast, an interferometer receives information from the entire range of a given spectrum during each time element of a scan. Thus, the interferometer receives information about the entire spectral range during the entire scan, while the grating instrument receives information only in a narrow band at a given time. Although an important principle in maintaining the importance of FT spectroscopy, the widespread use of detector arrays in dispersive spectroscopy has made the Multiplex advantage less relevant now.

2.6.3 Jacquinot advantage

The "Jacquinot advantage"[11] relates to the larger optical throughput of a FTS, which allows the recording of spectra with high resolution at low light levels. The interferometer can be operated with large solid angles at the source and detector of the instrument with no strong limitation on the resolution. However, the resolution of a conventional grating-type spectrometer depends linearly on the instruments slit width, requiring narrow slits to achieve high resolution. Increasing the resolution of a dispersive instrument, to compete with the resolution of an interferometer, by narrowing the slit width reduces the light throughput.

The maximum allowed throughput of a FTS is[3]

$$E_I = 2\pi A_I \frac{\Delta\nu}{\nu_{max}} \, cm^2 sr \tag{2.8}$$

where A_I is the area of the interferometer mirrors being illuminated. The throughput of a grating spectrometer is[3]

$$E_G = \frac{hA_G\Delta\nu}{fa\nu^2} \, cm^2 sr \tag{2.9}$$

where A_G is the area of the grating being illuminated, f is the focal length of the collimating mirror, h is the slit height, and a is the grating constant. Jacquinot's

advantage is therefore given by

$$\frac{E_I}{E_G} = \frac{2\pi A_I f a \nu^2}{A_G h \nu_{max}} \tag{2.10}$$

Quantitatively, the ability of interferometers to collect large amounts of energy at high resolution was expressed by Jacquinot as a throughput or *étendue* advantage of interferometers. For the same resolution, the optical throughput of a FTS based on a Michelson interferometer is at least 190 times greater than that of a dispersive spectrometer [11, 15].

The disadvantages of interferometers are few. For instance, when transmittance or reflectance measurements are made, the results can be in error up to 5% of the absolute value. Fluctuations in the interferogram can also cause this amount of error. However, the interferometric nature of Fourier-transform instruments usually requires precision scanning mechanisms and highly stable designs, which implies high cost and substantial bulk. In addition the temporal resolution is limited by the maximum mechanical scanning rate. This hinders the measurement of transient phenomena, such as short laser pulses and explosion spectra, where integration of many events is required to obtain meaningful results.

2.7 Survey of spectral regions

This section is a brief summary of the experimental techniques that can be used in the different spectral regions, starting with the visible region and proceeding outwards. If we take 'visible' to reach a little way into both the infra-red and the ultra-violet, we have a well defined region from 1000 nm (1 μm) to 200 nm, which is by far the easiest to work. Cheap, easy to use silicon based detectors are sensitive over this region. Quartz is transparent throughout this region and

glass is through most of it. The choice is open between prisms, gratings and interferometers, and there is no difficulty about windows or lenses. A wide range of sources are available for both emission and absorption spectroscopy, and among these one can now include tunable lasers.

Going below 200 nm in wavelength, first oxygen (in air) and then quartz start to absorb. To overcome the first the light path has to be evacuated – hence the name vacuum ultra-violet for this region. Quartz can be replaced by other crystals to extend the transmission range a few tens of nm, as far as 104 nm, which is the lithium fluoride limit, but prism spectrometers are used only for a few rather special applications in this region, and interferometers run into additional problems with surface tolerances and low reflectivity. Grating spectrometers are the instruments of real importance for wavelengths below 180 nm, with interferometers playing a marginal role down to about 120 nm. Lenses have to be replaced by mirrors or eliminated altogether by using concave gratings. Reflectivities decrease with wavelength, and below 35 nm gratings have to be used at grazing incidence to overcome this problem. However, the high energy of the photons makes detection relatively easy and background problems are small compared with those of the infra-red. The crystal transmission limit of 104 nm often defines a jump in experimental complexity because of the additional difficulties of running light sources and absorption cells without windows.

Going the other way, into the infra-red, the effective limit of crystal transmission is about 40 μm, but prisms are little used above a few μm, and mirrors are so efficient that they are nearly always preferred to lenses. Fourier transform spectroscopy is the important technique and the interferometers can be made with thin films or grids to overcome the constraints of crystal beamsplitters. The principal problem of the infra-red is the low intensity of infra-red sources, together with background and noise problems that become increasingly severe as the photon energy approaches kT. It is usually necessary to trade spectral

resolution off against signal-to-noise ratio. Finally, the vacuum requirements in the infra-red are not severe because dry oxygen and nitrogen do not absorb and it is only necessary to eliminate water vapour and carbon dioxide.

2.8 Static Fourier transform spectrometers

The conventional FTS based upon Michelson interferometers, although very powerful in offering both high resolution and signal-to-noise ratio, has two principal drawbacks. The interferometers require high quality mirror scanning mechanisms, implying high cost and substantial bulk, and the temporal resolution is limited by the maximum scanning rate.

Much research has gone into FTSs in order to overcome the poor temporal resolution and reduce the instrument bulk. These problems can be overcome by using an interferometer in which the entire interferogram is formed simultaneously in the spatial domain rather than temporal domain. This eliminates the need for moving parts in a FTS.

2.8.1 Modified Michelson interferometers

There are a variety of interferometers able to produce fringes in the spatial domain. Stroke and Funkhouser[16] have configured a Michelson interferometer with tilted mirrors to give a path difference which varies with lateral position across the exit aperture of the instrument. The wavefronts interfere at small angles and the relative phase difference between phase fronts is then a function of the transverse displacement, so the interferogram appears as a spatial intensity distribution. In this interferometer, since the two beams producing the interference fringes make round trips in different arms, they are independently influenced

by external turbulence and vibration. Therefore, the interferogram produced by the beams is sensitively affected by the turbulence and vibration, especially in the UV and visible region. Even the zero-order fringe is very difficult to obtain.

To overcome this problem, Yoshihara and Kitade[17] have used a triangle common-path interferometer, an example of which is shown in figure 2.7. In this interferometer, since the two beams travel on the same path but towards opposite directions, external disturbances are almost compensated for.

These spectrometers were originally called holographic spectrometers as they recorded the interferogram photographically which was then Fourier transformed optically with coherent light to obtain the wavelength spectrum. Current technology has enabled the photographic film to be replaced by a photodiode array connected to a computer which will perform a Fourier transform[18, 19, 20].

The spectrometer configuration by Okamoto[18] is shown in figure 2.7. At the heart of this spectrometer is a triangle common-path interferometer. In the interferometer the beam from the source is divided into two by the beam splitter. Each of these two beams travels the same path in the opposite direction and is recombined by the beam splitter after reflections at the plane mirrors 1 and 2. These two diverging beams are collimated by the lens, reflected by mirror 3 and form an interference pattern on the photodiode array, which is located at the back focal plane of the lens. The dashed line in figure 2.7 indicates the position for mirror 1 which produces the zero split of beams due to the symmetrical arrangement of two mirrors. The distance d between the dashed line and mirror 1 along the optical axis makes two collimated beams intersect on the detector plane and produces the interference fringe pattern on the photodiode array. The data from the diode array are then read by a computer before performing a Fourier transform to obtain the spectrum of the source.

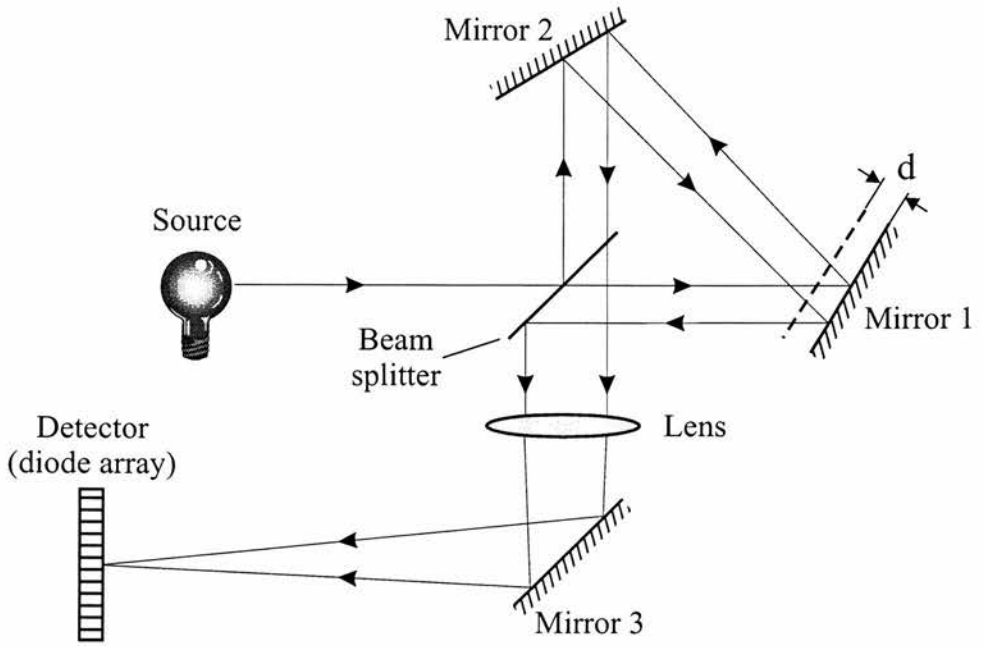

Figure 2.7: *Optical layout of the Fourier transform spectrometer based on a triangle common-path interferometer. It generates an interferogram without moving parts which is then imaged onto a photodiode array before data processing by a computer.*

Since this interferometer needs neither an aperture nor a slit, the dimension of the radiation source may be large. This provides the throughput advantage for widely extended sources. However, a lens of relatively large F number is necessary to form a triangle path which limits the solid angle. In addition, the necessity for a distance as long as the focal length of the Fourier transform lens between the lens and the detector could also become a problem if one wanted to reduce drastically the dimensions of the instrument.

2.8.2 Birefringent interferometer

Recently, novel static Fourier-transform spectrometers(SFTS) based upon birefringent interferometers[21] have been developed which have no moving parts and are extremely robust and compact[22]. An example of Okamoto's use of a birefringent interferometer in a Fourier transform spectrometer is shown in figure 2.8.

These birefringent interferometers have the same advantages of throughput and ruggedness against vibration as the common-path interferometer. In addition, this interferometer can be built on a much smaller scale owing to its in-line formation. The solid angle subtended by the lens aperture can be much larger than the triangle path interferometer.

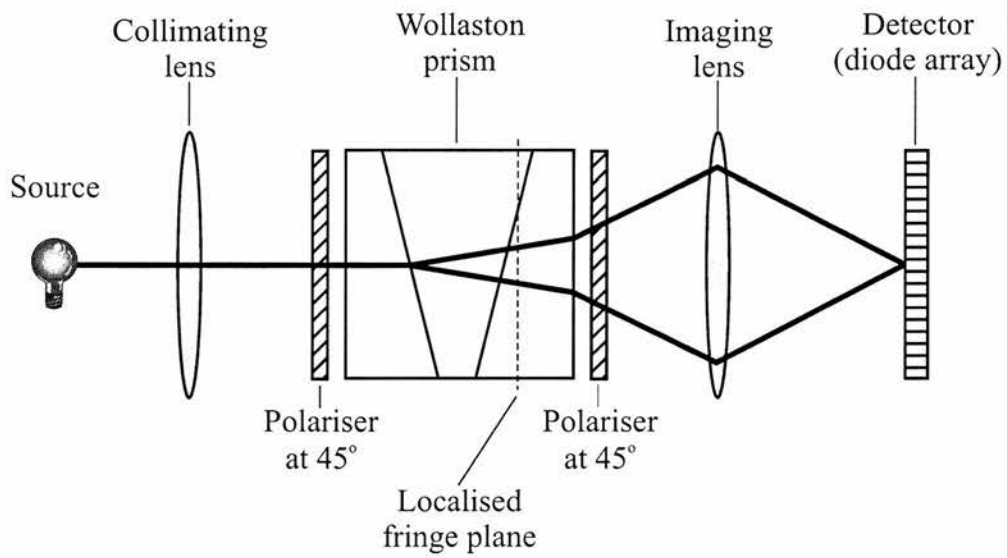

Figure 2.8: *Optical layout of the birefringent interferometer, at its heart is a Wollaston prism. It generates an interferogram without moving parts which is then imaged onto a multi-element photodiode before data processing by a computer.*

At the heart of the birefringent interferometer in figure 2.8 lies a Wollaston prism. The operating principles of a Wollaston prism based spectrometer will be ex-

plained in detail in chapter 3. Suffice it to state here that a Wollaston prism is composed of two wedges of birefringent crystal joined along their hypotenuse and with their optic axes parallel to the front/back faces but mutually perpendicular. The Wollaston prism introduces a path difference between the two orthogonal polarizations which varies linearly with the lateral position of the input ray. When the output rays are analyzed through a polarizer at 45^o then interference occurs. If the aperture of the Wollaston is illuminated with an extended beam, then the whole interferogram is formed simultaneously in the spatial domain. After recording with a detector array, a computer then performs a Fourier transform on the data to give the spectrum of the light.

2.9 Conclusions

The remainder of this thesis concentrates on the development of static Fourier transform spectrometers based on birefringent materials. My project work at the University of St Andrews has been primarily involved in improving the spectrometer by miniaturising the instrument and improving its operating performance. Specifically I have addressed issues such as field of view and operating wavelength. The spectrometer has also been developed for applications such as gas detection, liquid fuel monitoring, and laser wavelength measurement.

Chapter 3

Fourier transform Spectrometers based upon Wollaston prisms

3.1 Introduction

As previously discussed in chapter 2(section 2.6), conventional Fourier transform spectrometers are based on Michelson interferometers. Although these interferometers offer high spectral resolution and an excellent signal-to-noise ratio they have two drawbacks. The necessary high precision scanning mechanism implies substantial bulk and results in an expensive instrument while the temporal resolution is limited by the maximum scanning rate.

Recently, at the University of St Andrews, we have developed static Fourier transform spectrometers based upon Wollaston prisms, enhancing both the design and capabilities of the instruments[23, 24, 25, 26]. This work (awarded the NPL metrology prize in 1994) has resulted in the construction and commercialisation of an ultra-violet instrument used for pollutant gas detection. The work has been further recognised for its commercial potential when I was awarded an Enterprise

Fellowship from the Royal Society of Edinburgh and Scottish Enterprise in 1998.

At the heart of these novel static FTSs lies a Wollaston prism, these prisms were invented by William Hyde Wollaston (1766 – 1828). Wollaston was educated at Charterhouse and Cambridge where he graduated in medicine which he practiced for 11 years in London. While working in London his interest and devotion to various branches of natural science increased and in 1800 he retired from medical practice to concentrate on his scientific pursuits. The chief events in Wollaston's life are his discoveries, which flowed in uninterrupted succession from 1800 through to the time of his death. He published fifty-six papers on pathology, physiology, chemistry, optics, mineralogy, crystallography, astronomy, electricity, mechanics, and botany, and almost every paper marks a distinct advance in the particular science concerned.

As an inventor of optical apparatus Wollaston is highly regarded. In 1802 he described the total-reflection method for the measurement of refractivity. In 1807 he patented the camera lucida, leading indirectly to many discoveries in photography as others tried to improve the image from Wollaston's camera. The reflecting goniometer followed in 1809, which rendered the exact measurement of crystals and determination of minerals.

It was in an 1820 paper entitled 'On the Method of cutting Rock Crystals for Micrometers' that Wollaston described the double-image prism that bears his name. The double image effect can be seen in figure 3.1 where a photograph was taken through a Wollaston prism.

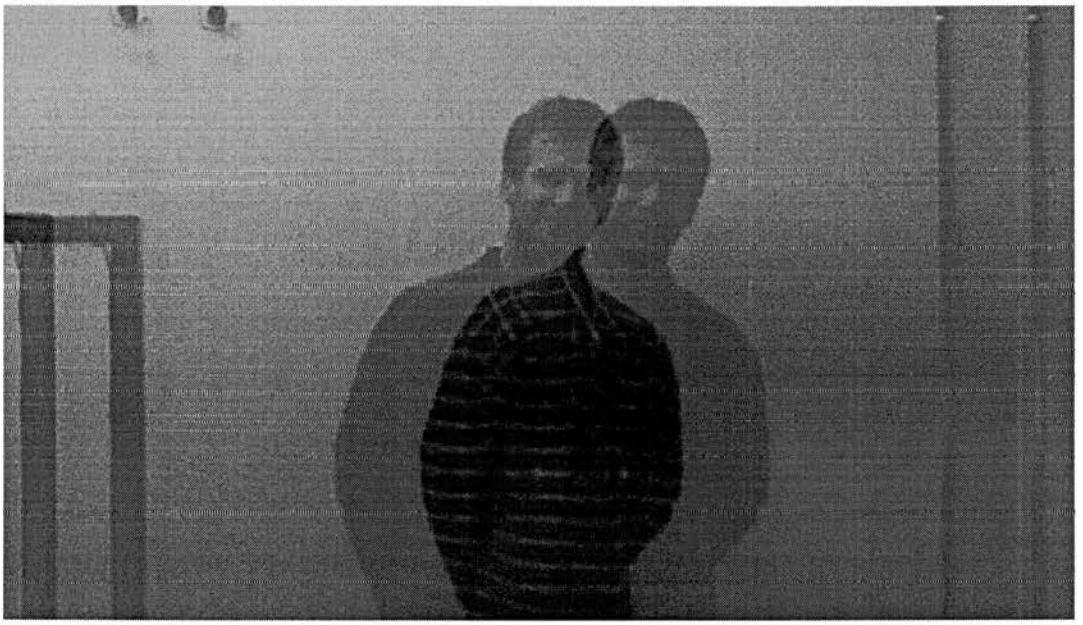

Figure 3.1: *A photograph of Dr Miles Padgett through a Wollaston prism demonstrating the "double image" effect. The double image seen is the result of light rays splitting into two at the wedge interface of the prism. Further investigation is inconclusive as to whether Miles is demonstrating his intelligence by showing that two brains are better than one, or that I had a two headed monster for a supervisor.*

3.2 The Wollaston prism

A Wollaston prism is a common optical component comprising of two geometrically similar wedges of birefringent material joined by their hypotenuse to form a rectangular block, as shown in figure 3.2.

The optic axes within the two wedges are aligned perpendicular to each other and parallel to the entrance–exit faces of the composite block. Input light consisting of two polarisation components undergoes a symmetrical splitting at the wedge interface. This is due to the two polarisation components of the light experiencing a different change of refractive index as they pass the interface between the wedges

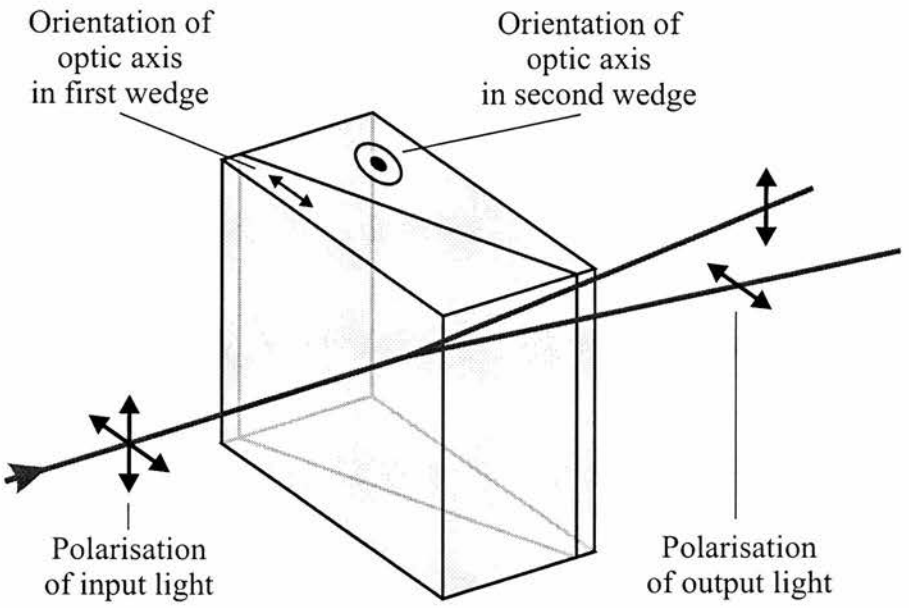

Figure 3.2: *A Wollaston prism.*

of the prism, one traveling through the first wedge as an extraordinary ray, then entering the second wedge as a ordinary ray, and the other alternating from an ordinary ray to an extraordinary ray. Because the angle of refraction at the internal interface of the prism depends upon the polarisation state of the input light the customary use of a Wollaston prism is as a polarising beam splitter.

The deviation angle, α, (see figure 3.3) between the two beams depends on ϑ, the wedge angle, and the birefringence of the prism material, $n_e - n_o$ (see figure 3.3), and is given by equation 3.1[21],

$$\alpha \cong 2(n_e - n_o) \tan \vartheta. \tag{3.1}$$

3.2.1 Principle of operation

The use of Wollaston prisms as polarising beam splitters has been known since their introduction in 1820. However, the reason the Wollaston prism is of interest to the spectroscopist, and the key principle of operation, is the varying wedge thickness across the aperture which introduces a *path difference* between the two orthogonally polarised components of the beam. This is directly analogous to a Michelson interferometer where a scanning mirror introduces a path difference between two beams. Figure 3.3 shows normally incident light resolved in the first wedge of the prism into ordinary and extraordinary components that propagate collinearly with different phase velocities. At the interface between the wedges, the components are interchanged so that the ordinary component propagates as an extraordinary and vice versa. In the centre of the prism the effective path difference introduced between the orthogonal polarisations in the first wedge is cancelled out by an equal and opposite effective path difference in the second wedge. Moving laterally away from the central position the effective path difference, Δ, is proportional to the displacement from the centre of the prism, d, and is given by[21],

$$\Delta \cong 2d\,(n_e - n_o)\tan\vartheta, \tag{3.2}$$

where n_o and n_e are the ordinary and extraordinary refractive indices respectively of the Wollaston material and ϑ is the wedge angle of the prism. It is this path difference introduced by a Wollaston prism between orthogonal polarisations that forms the basis of the FTS design.

Figure 3.4 shows that by placing the Wollaston prism between two polarisers aligned at $\pm 45^o$ to the optic axes of the prism material and illuminating with a light source, a set of straight-line interference fringes will be produced localised to

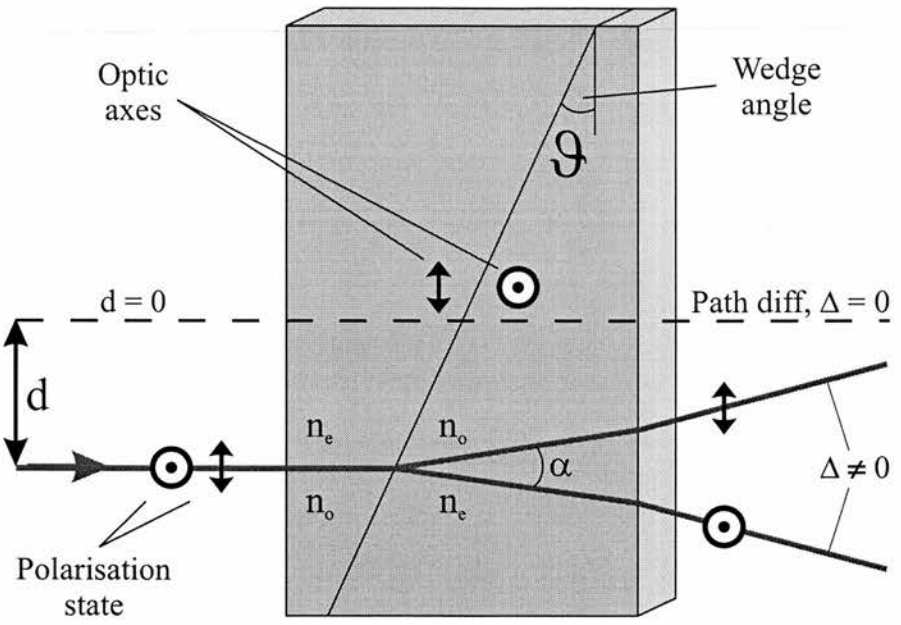

Figure 3.3: *Side view of a Wollaston prism showing an off-centre ($d > 0$) beam traversing the prism. The beam is split into its two orthogonal polarisation components. The vertical polarisation component experiences a small amount of extraordinary and larger amount of ordinary refractive indices. Whereas, the horizontal component undergoes a change from small ordinary to large extraordinary refractive indices. The rays therefore exit the prism at different times, appearing as if they had experienced different path lengths.*

a plane within the prism[21]. The first polariser ensures light entering the Wollaston prism is always at 45^o, which is resolved by the prism into horizontally and vertically polarised components to form two independent beams. The second is an analysing polariser enabling the common components of the two orthogonally polarised beams to interfere. As we move laterally across the face of the prism the path difference between the two beams varies. Assuming a monochromatic light source the interference fringes will vary in a sinusoidal pattern and the period, x_o, of the fringes produced can be determined from equation 3.3

$$x_o = \frac{\lambda}{2\,(n_e - n_o)\tan\vartheta} \tag{3.3}$$

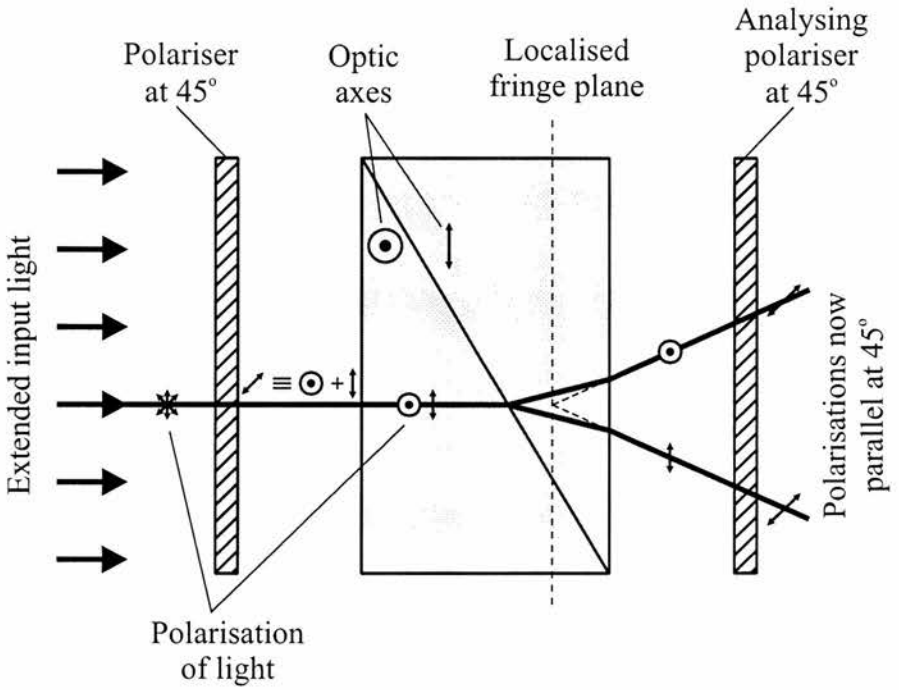

Figure 3.4: *Placing the Wollaston prism between two polarisers aligned at ±45° to the optic axes of the prism material and illuminating with a light source, a set of straight-line interference fringes will be produced localised to a plane within the prism.*

It is now a simple matter to image these interference fringes onto a detector array, the data from which can be recorded by a computer for processing. This can be done simultaneously for all path differences by illuminating the entire aperture of the Wollaston prism and using a detector array positioned in the imaged plane of the ray intersections to record the interferogram. The output of each photodiode is proportional to the integral of the light intensity of a small element of the interferogram during the exposure time of the photodiode array.

3.2.2 Derivation of interferogram from fringes

The fringes generated by the Wollaston prism can be considered equivalent to the fringes generated from a Michelson interferometer. It can be shown that recording the output of a Wollaston prism interferometer as a function of the displacement of the input ray generates fringes that are the Fourier transform of the spectral power distribution[23]. Hence a Wollaston prism can be used as the basis of a Fourier transform spectrometer.

We can consider light propagating in the z direction incident normally upon a Wollaston prism. The light is linearly polarised with its plane of polarisation at an angle of 45^o to the axis of the prism, which is taken to lie in the x direction. The time variation of the incident optical field can be expressed in terms of its x and y components as

$$E(t) = \hat{i} \int_{-\infty}^{\infty} A(\omega) \exp(i\omega t) d\omega + \hat{j} \int_{-\infty}^{\infty} A(\omega) \exp(i\omega t) d\omega, \tag{3.4}$$

where $\hat{i}$ and $\hat{j}$ are unit vectors in the x and y directions, respectively, and $A(\omega)$ is the Fourier amplitude corresponding to angular frequency ω.

At distance d from the central position of the Wollaston prism, the two halves of the Wollaston prism differ in thickness. The difference between the optical path lengths Δ traversed by the x- and y-polarised components is simply written as

$$\Delta \cong 2d(n_e - n_o) \tan \vartheta. \tag{3.5}$$

After passing through the Wollaston prism, the x and y components follow slightly different ray paths before recombining at the detector array. The linear polariser placed after the prism has its axis at an angle of 45^o to the x axis. It transmits the

component of the x and y polarisations at 45^o to the x axis. Taking into account the effect of this linear polariser and the phase delay introduced by the optical path length between the prism and the detector, we find that the amplitude of the optical field at the detector array is[7]

$$E(t) = \frac{1}{\sqrt{2}} \int_{-\infty}^{\infty} A(\omega) \exp\left[i\omega\left(\frac{t - p_x}{c}\right)\right] d\omega + \frac{1}{\sqrt{2}} \int_{-\infty}^{\infty} A(\omega) \exp\left[i\omega\left(\frac{t - p_y}{c}\right)\right] d\omega. \tag{3.6}$$

The signal S recorded by the detector array is proportional to the intensity integrated over the time of observation and it can be shown that[27]

$$\frac{S - S_0}{S_0} = \frac{\int_0^{\infty} \mid A(\omega) \mid^2 \cos\left[\omega \frac{2d(n_e - n_o)\tan\vartheta}{c}\right] d\omega}{\int_0^{\infty} \mid A(\omega) \mid^2 d\omega}. \tag{3.7}$$

Where $S_0 = k_0 \int_0^{\infty} \mid A(\omega) \mid^2 d\omega$, representing the signal for large optical path difference, and k_o is a constant.

In the ideal case, when the birefringence $(n_e - n_o)$ does not depend on frequency ω, the right hand side of equation (3.7) is exactly the normalised Fourier cosine transform of the power spectrum $\mid A(\omega) \mid^2$ of the incident optical field, with the quantity $2d(n_e - n_o)\tan\vartheta/c$ acting as the time variable. Assuming equally spaced detector elements in the array, we can therefore use the fast Fourier transform algorithm to generate an estimate of the optical power spectrum of the incident light.

An example interferogram, recorded using an extended metal halide lamp source is shown in figure 3.5. The prominent fringes near zero path difference result from the broadband nature of the light source, and the fringes at large path differences

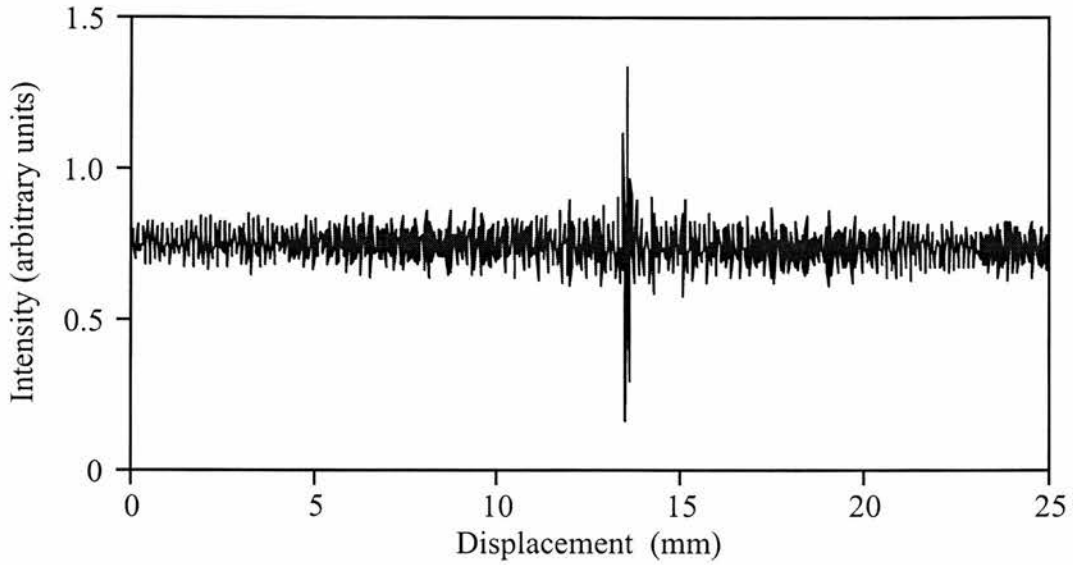

Figure 3.5: *Interferogram obtained from a metal halide white light source. The interferogram is recorded from a Wollaston prism with a 25 mm aperture, imaged onto a 25 mm detector array.*

are due to the discrete emission lines within the spectrum.

The spectral distribution of the input light can be obtained by recording the output of a Wollaston prism interferometer as a function of the displacement of the input ray and performing a Fourier transform on the resulting fringes. A computer reads the interferogram data from the detector array prior to data processing with a fast Fourier transform (FFT)[28]. The Fourier transform of the interferogram from figure 3.5 results in the spectrum shown in figure 3.6 and demonstrates that fringes are formed simultaneously for wavelengths between 400 and 1000 nm. The presence of emission lines illustrates the spectroscopic applications of the instrument. In particular, the xenon line at 823 nm and the unresolved sodium-D lines at 590 nm can be identified clearly.

However, the birefringence $(n_e - n_o)$ in equation 3.7 does depend on the frequency of the light. Therefore, the spectrometer is calibrated at a known wavelength

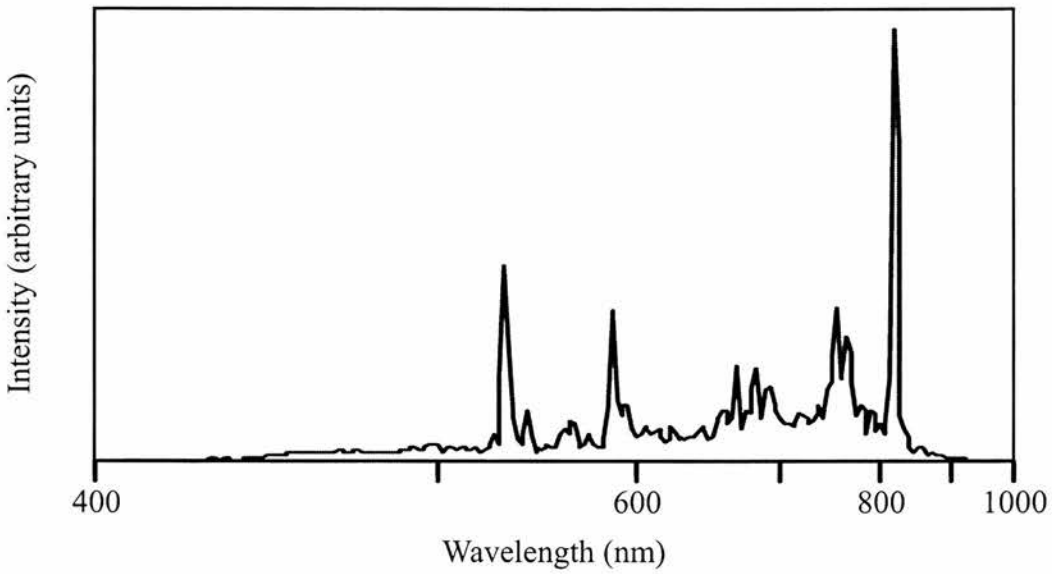

Figure 3.6: *Fourier transform of the interferogram in figure 3.5 showing the spectral properties of the light source.*

within the spectral region of interest, using a knowledge of the variation in the birefringence with frequency the correction factors for neighbouring wavelengths are determined.

3.2.3 Spectrometer calibration of a FTS based upon a Wollaston prism

For light of a given wavelength, the period of the fringes in the localised fringe plane is proportional to wavelength and inversely proportional to the birefringence[23]. The variable in the resulting Fourier spectrum is linear in $k\Delta n$, where k is the wavenumber and $\Delta n = n_e - n_o$ is the birefringence of the Wollaston prism material. To first order, Δn is constant, and therefore the Fourier transform variable is linear with wavenumber, and has an upper wavenumber limit set by the Nyquist criterion. However, to obtain an accurate estimate of the spectrum, both in terms

of frequency and intensity scales, the change in birefringence with wavelength has to be considered.

We calibrate the wavenumber scale of the spectrum with respect to a known wavenumber, k_{cal} (e.g. $k_{cal} = 2\pi/633nm$). Assuming the spectral line associated with k_{cal} occurs at array element i_{cal} in the Fourier-transformed data, we obtain the following implicit expression for $k(i)$, the wavenumber of the ith array element[25]

$$k(i) = \frac{i}{i_{cal}} k_{cal} \frac{\Delta n(k_{cal})}{\Delta n(k(i))}. \tag{3.8}$$

Providing the birefringence is known as a function of wavelength, this relationship can be applied recursively to determine the wavenumber associated with each array element in the Fourier-transformed data. Once calculated, these wavenumber values can be held in a calibration array.

As stated above, the variable in the Fourier spectrum is proportional to $k\Delta n$. In order to plot the spectral intensity against wavenumber, we must make the following change of variable[25]:

$$I_k(i) = I_{k\Delta n}(i) \left| \frac{d(k\Delta n)}{dk} \right|_{k(i)}, \tag{3.9}$$

where I_k and $I_{k\Delta n}$ are the spectral intensity functions in terms of k and $k\Delta n$ respectively. The derivative, which is evaluated at $k(i)$, is given by[25]

$$\frac{d(k\Delta n)}{dk} = \Delta n + k \frac{d\Delta n}{dk}. \tag{3.10}$$

For example, in the case of a calcite Wollaston prism, the second term in the

above expression is negligible over the wavelength range of interest. Therefore, the spectral intensity per unit wavenumber is[25]

$$I_k(i) \approx I_{k\Delta n}(i)\Delta n(k(i)). \tag{3.11}$$

The source code to perform the calibration of the spectrometer can be seen in appendix A. It is written in the C programming language. Additional corrections can be applied to the intensity axis to allow for the spectral response of the camera detector.

The following sections contain several designs of static Fourier transform spectrometers which we have employed at the University of St Andrews. Each of these designs combines the two rays onto suitable detectors to form an interferogram in the spatial domain. As the path difference is introduced along a common path within the Wollaston prism, the design is robust and insensitive to mechanical alignment. Unlike previous static FTS, there are no beamsplitters or mirrors and with a suitable choice of components no power is returned to the source. All optical components are held on a single optical axis, which leads to an extremely compact instrument.

3.3 Single Wollaston prism with imaging lens

This is the first design proposed for a static spectrometer at the University of St Andrews[23]. It is a modification of the original design of Okamoto[22] (shown in section 2.8.2) where the Wollaston prism has been replaced by one of simpler construction. The instrument uses a lens to image the interferogram which is localised within the prism itself to an external fringe plane where a detector array could be placed to record the image. The optical path of this simple and

intuitive method of imaging the interferogram is shown in figure 3.7.

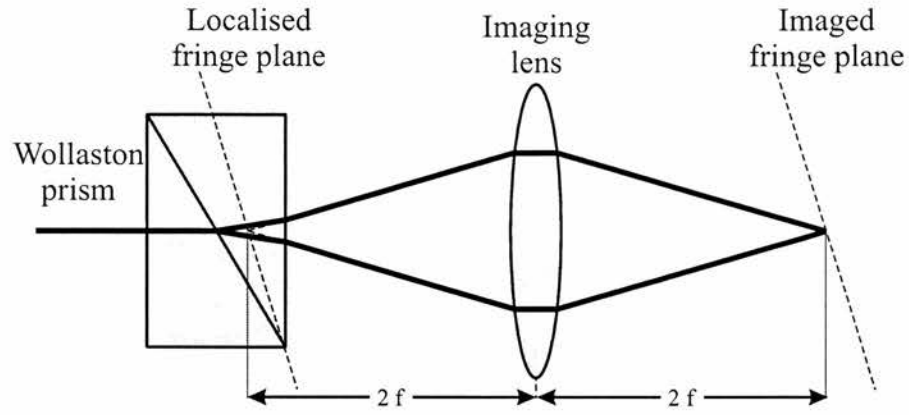

Figure 3.7: *The simple optical path of the single Wollaston prism spectrometer which uses a lens to image the interference pattern onto the detector.*

The design and optical layout of the spectrometer is outlined in figure 3.8. The Wollaston prism was fabricated from calcite with a 20 mm aperture and an internal angle, ϑ, of 3^o. The virtual fringe plane within the prism was imaged onto a 1024 element detector array by using an achromatic lens. The polarisers are of the simple Polaroid sheet type and are placed either side of the Wollaston prism.

The spatial extent and localisation of the fringes means that aberrations in the imaging lens can result in a loss of fringe contrast and a distortion in apparent fringe period. Careful optical design and the use of high quality components are required to limit these effects and the need for low aberration imaging sets a minimum size on the optical system. An alternative system was investigated which would remove the need for a lens, thereby reducing the size of the instrument and the problems of imaging aberration. This new design was based upon two Wollaston prisms where the recombination of the two polarisations was performed by the second prism.

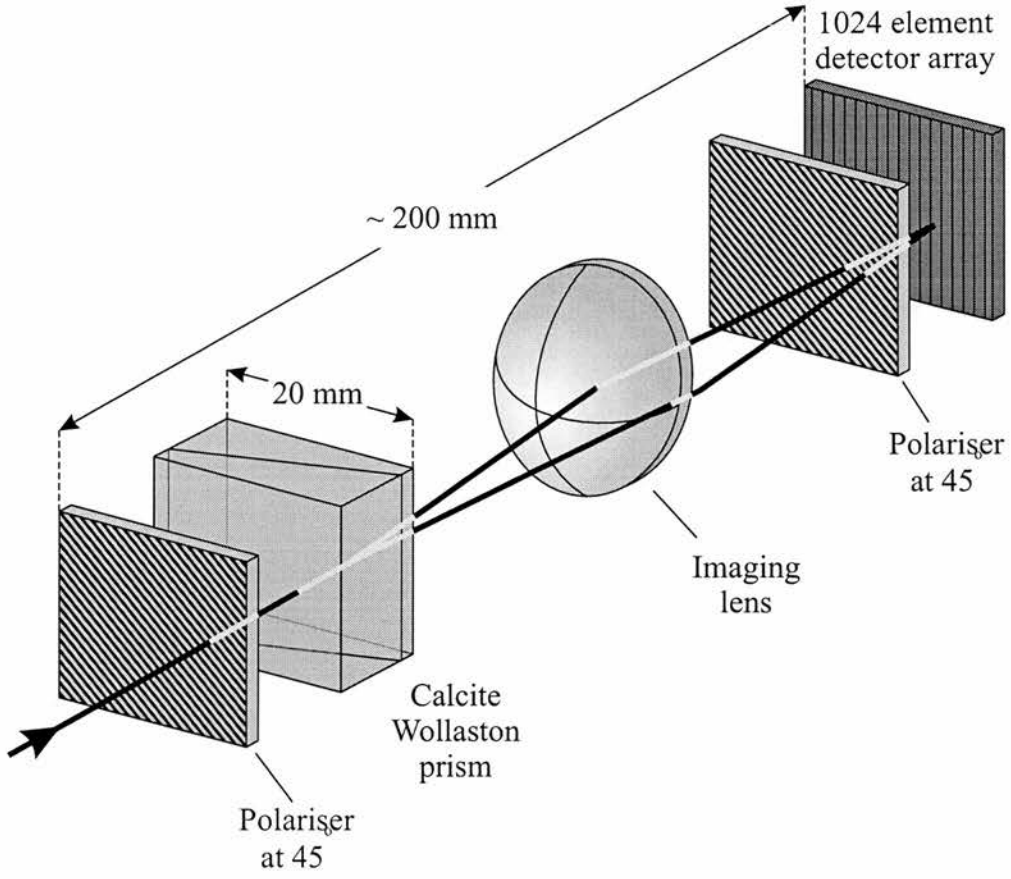

Figure 3.8: *The optical layout of the single Wollaston prism spectrometer which uses a lens to image the interference pattern onto the detector array.*

3.4 Double Wollaston prism spectrometer

A combination of two Wollaston prisms was used to eliminate the imaging problems with the lens based system[24]. The second Wollaston replaced the lens and localised the fringes to the plane of the detector without introducing any of the possible chromatic aberrations present in the previous design[23], giving undistorted fringes. The use of the second prism also resulted in a five fold reduction in the length of the instrument.

The projection of the localised fringe plane to a position outside of the prism pair can be understood with reference to figure 3.9 (i). The first prism splits the light into its two orthogonally polarised components with an angular deviation of α_1, given by equation 3.1. The second prism is oriented to reverse the splitting introduced by the first prism. Providing that $\alpha_2 > \alpha_1$, after transmission through the second prism, the two component rays will converge, the crossing point being the position to which the fringes are localised. Figure 3.9 (ii) shows an alternative configuration for the prisms.

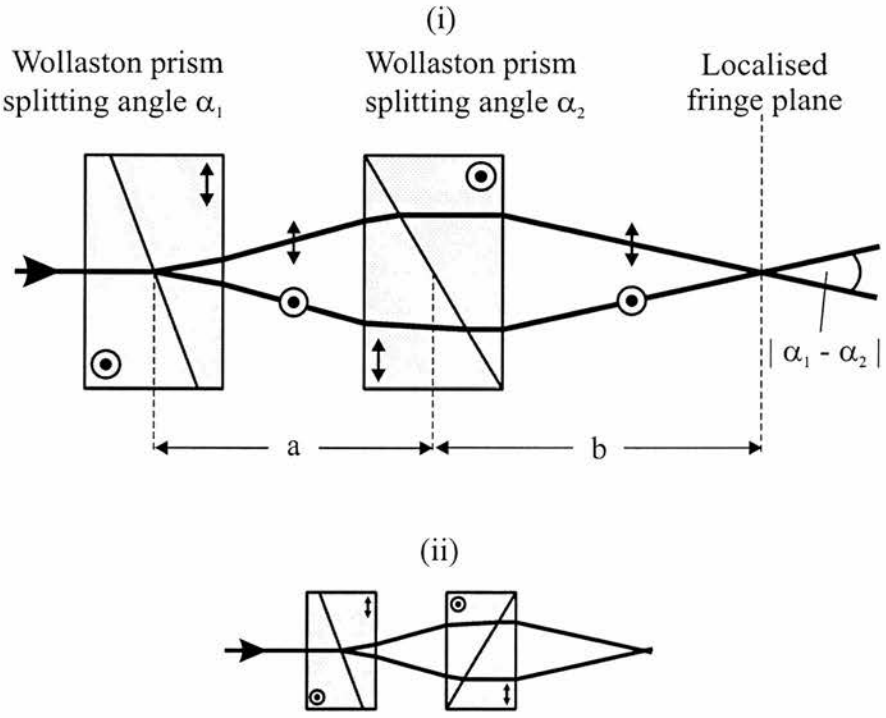

Figure 3.9: *(i) The optical path of the double Wollaston prism spectrometer which uses a second prism to localise the fringe plane to a position outside the Wollaston prisms. (ii) Alternative orientation of the prisms.*

The approximate location of the fringe plane can be calculated from the deviation angles of the Wollaston prisms and their separation, and is given by[24]

$$(\alpha_2 - \alpha_1)\left(b + \frac{t}{2\bar{n}} - \frac{t}{2}\right) \simeq \alpha_1\left(a + \frac{t}{\bar{n}} - t\right), \tag{3.12}$$

where a is the centre-to-centre separation of the prisms, b the distance between the second prism and the fringe plane, t the thickness of the prisms, $\bar{n}$ the mean refractive indices of the prism material. For a spectrometer constructed from two 4 mm thick calcite Wollaston prisms with splitting angles of 2.1^o and 1.3^o equation 3.12 gives the approximate distance of the fringe plane as 3 mm behind the exit face of the second prism. The two polarisers are a simple dichroic sheet located on either side of the prism pair which results in an extremely compact and robust spectrometer measuring $40 \times 40 \times 100$ mm, see figure 3.10. A photograph of the complete instrument is shown in figure 3.11

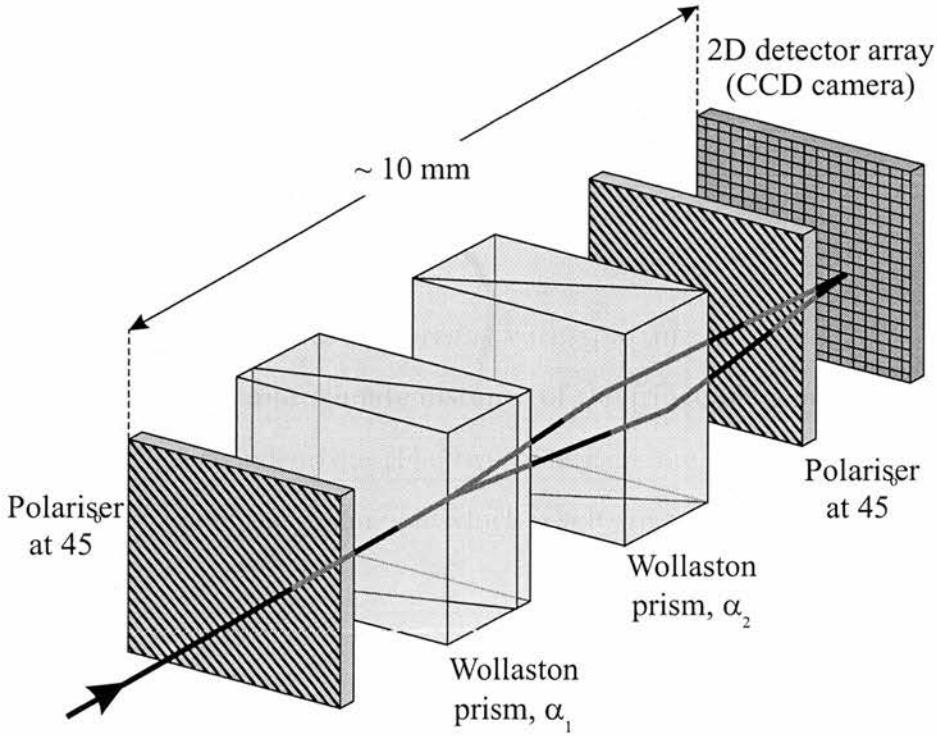

Figure 3.10: *The optical layout of the double Wollaston prism Fourier transform spectrometer.*

The path difference, Δ, introduced between the two orthogonally polarised beams

by the prism pair is given by the sum of the path difference introduced by each individual prism. For a small angle, α, this is given approximately by[24]

$$\Delta \approx d(\alpha_2 - \alpha_1). \tag{3.13}$$

The instrument configuration above gives a nominal splitting angle of 0.8^0, with a prism width of 10 mm, the maximum path difference across the detector array equation 3.13 is approximately ± 50 μm.

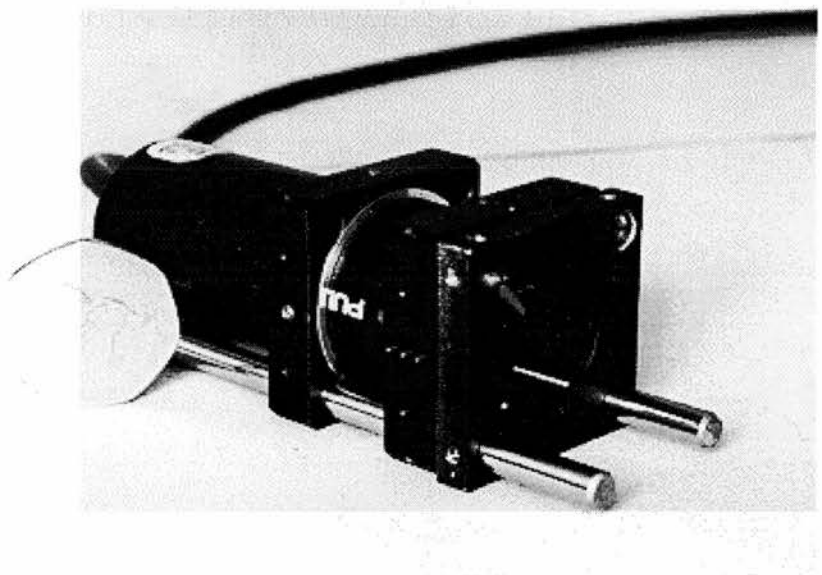

Figure 3.11: *Photograph of the compact static Fourier transform spectrometer based upon two Wollaston prisms.*

A further modification from the original design in section 3.3 is the change from a one dimensional detector array to the use of a two dimensional detector array. The detector array is a 512×767 element standard CCD video camera with an active area of 5×7 mm. Its output is read via a standard frame grabber into a personal computer. The use of readily available CCD cameras and frame grabbers

has enabled the cost of the instrument to be reduced.

The data processing is eased by rotating the CCD array so that the fringes lie perpendicular to the lines of the camera image. Averaging the pixel columns of the image generates a single 767 element array containing the fringe information. The data is zero padded to 1024 elements to enable the use of a fast Fourier-transform algorithm from which the spectral power distribution of the input light is calculated.

3.5 Modified single Wollaston prism (inclined optic axis)

As discussed in section 3.2.1, for a standard Wollaston prism the fringes produced by an extended source are localised within the prism itself. However, it was previously noted[21] that realigning the optic axis of the birefringent material with respect to the prism faces alters the plane of fringe localisation to where it is imaged outside the prism, as shown in figure 3.12. There is no requirement for an imaging lens or a second prism because the fringe pattern is localised behind the exit face of the modified prism and can be recorded directly by a multi-element detector array[25].

The Wollaston prism shown in figure 3.13 shows a modified design such that fringes produced by an extended source are localised behind the exit face of the prism. The two orthogonal polarisations of the beam split at the entrance face of the prism and further refraction at the wedge interface reverses the splitting and the rays interfere at the fringe plane.

This design has several notable advantages. First the component count is re-

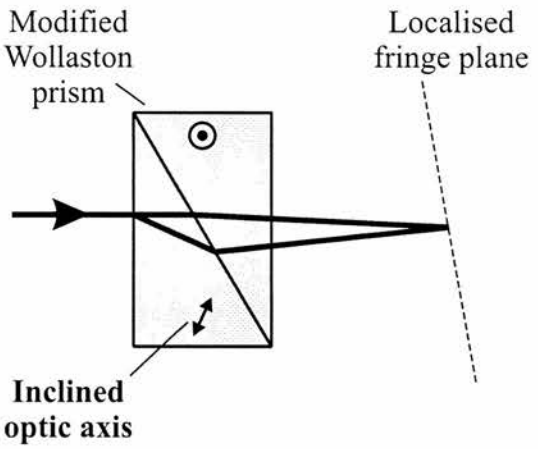

Figure 3.12: *The optical path of the modified Wollaston prism with the inclined optic axis. The beam is split at the entrance face and the wedge interface focuses the fringe plane to a point outside the exit face of the prism.*

duced, thereby reducing the complexity of assembly and cost of components. Second, since the field of view of a Wollaston prism based spectrometer is inversely proportional to the total thickness of the birefringent optical components, the elimination of one prism increases significantly the field of view of the spectrometer. This field of view increase will be discussed in detail in chapter 4.

The spectrometer was implemented with a 10×10 mm aperture, 1.5 mm thick, calcite Wollaston prism. The prism has an internal wedge angle of 1.6^o and the optic axis in the first wedge is inclined at an angle of 12^o to the entrance face (13.8^o to the interface). This prism configuration generates the fringe plane approximately 2.7 mm behind the exit face. From equation 3.2 the path difference as a function of lateral position in the localised fringe plane is 9.3 $\mu m/mm$, or ± 46 μm across the whole of the detector array. The complete optical layout, including polarisers and detector array, of the spectrometer is shown in figure 3.13.

To realise the true potential of the compact design it is obviously important to

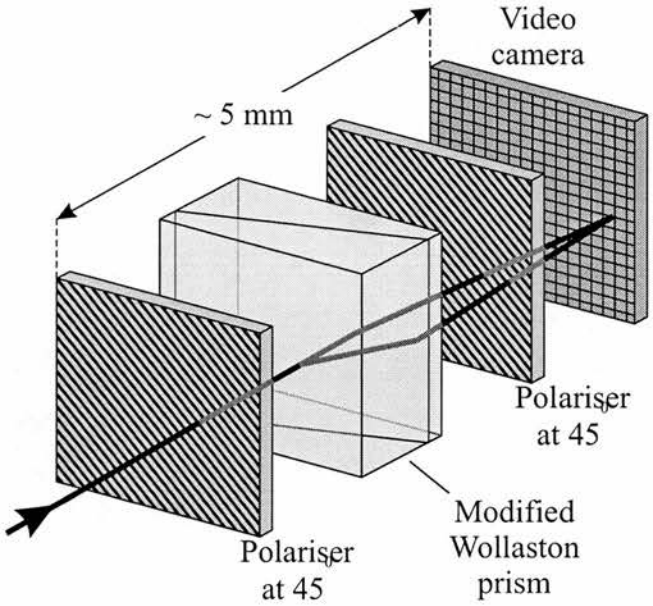

Figure 3.13: *The optical layout of the modified Wollaston prism spectrometer which uses the inclined optic axis to image the interference pattern onto the detector.*

combine the Wollaston prism and polarizer assembly with an appropriately sized detector. The evaluation instrument described here used a highly integrated CMOS VLSI camera, which fits onto a 30 mm diameter PCB and has a total power consumption of less than 150 mW. In addition to a 312×287 array sensor with a total active area of 6.14×4.59 mm, the camera includes the necessary drive circuits to deliver a formatted CCIR composite-video signal. The camera is mounted behind the exit face of the Wollaston prism to coincide with the localised fringe plane. This produces an ultra-compact design and a photograph of the complete spectrometer can be seen in figure 3.14.

A dedicated frame-grabber interface board reads the video output from the camera into a computer. The 2-D image of the interference fringes is averaged across the frame to give a single 287 element data array. The array is zero-padded to 512 elements and the fast Fourier transform of the interferogram data produces

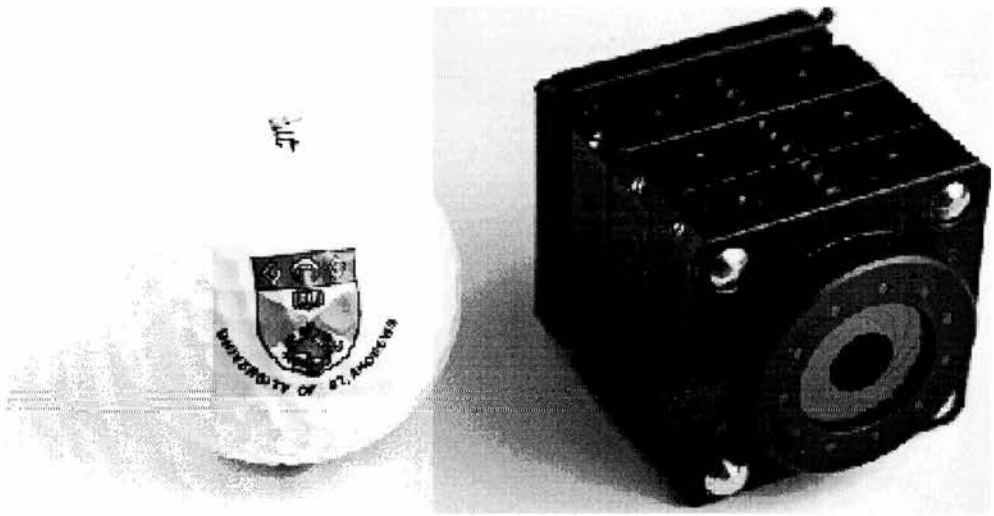

Figure 3.14: *A photograph of the ultra-compact, static, Fourier transform spectrometer based on a single modified Wollaston prism.*

an optical power spectrum, which to first order is linear in wavenumber.

3.6 Instrument properties of Wollaston prism based Fourier transform spectrometers

The key considerations when assessing the performance of the spectrometers relates to the operating wavelength range, resolution, wavelength accuracy, acquisition rate, and field of view.

3.6.1 Wavelength range

The basic property of a spectrometer is its operating wavelength range. If it does not take measurements in the spectral region required by the user, it is useless.

Fundamentally, the most important component is the Wollaston prism. Table 3.1 lists the properties of several birefringent materials which may be used to fabricate the prisms.

Material	Transmission range (μm)	Refractive index at 633 nm $(n_e + n_o)/2$	Birefringence at 633 nm $(n_e + n_o)$
Quartz	0.18 – 3.5	1.55	0.009
Calcite	?	1.57	0.17
Rutile	0.5 – 5	2.59	0.25
Magnesium Flouride (MgF_2)	0.11 – 8	1.38	0.01
Ammonium Dihydrogen Phosphate (ADP)	0.18 – 1.7	1.47	-0.05

Table 3.1: *Wollaston prisms may be manufactured from these common birefringent materials. This table lists the important properties such as useful operating wavelength range and birefringence[29].*

Initial instruments were developed using affordable CCD detector arrays manufactured from silicon which has a wavelength sensitivity in the visible and near infrared of approximately 350–1050 nm. It can be seen from table 3.1 that Wollaston prisms can be fabricated from materials which will operate in the ultraviolet or infrared regions. The development of spectrometers using detector arrays which operate in the ultraviolet or mid-infrared will be discussed in chapter 5.

The final optical components are the polarisers which in the initial spectrome-

ters were of the simple polaroid sheet variety designed to operate in the visible region. These varied in operating wavelength within the range of 400 – 1000 nm. The polarisers were often found to restrict the operating spectral range of the spectrometers as the Wollaston prisms and CCD detector arrays have a greater spectral range. Alternative polarisers would allow access to the full useful range of the detector. Ultimately, the spectral bandwidth of the instrument is limited only by the range over which multielement detector arrays can be obtained.

3.6.2 Resolution

Irrespective of their detailed design, the resolution of all FTSs depends on the maximum path difference, $\Delta_{\max}$, introduced between the interfering rays. The exact relationship depends on the form of apodisation used on the interferogram prior to the Fourier transform. In general, the resolution [full-width half-maximum (FWHM)] in wavenumbers, $\Delta\sigma_{FWHM}$, is approximated by [7],

$$\Delta\sigma_{FWHM} \approx \frac{1}{\Delta_{\max}}. \tag{3.14}$$

The Nyquist criterion requires at least two data points per fringe period for the unambiguous measurement of the period of oscillation. This, coupled with the number of elements on the detector effectively sets an upper limit to the maximum path difference. If the interferogram is symmetrically recorded about zero path difference, the highest resolution attainable depends on the number of elements across the detector, N, and the shortest wavelength of interest, $\lambda_{\min}$, and is given by,

$$\Delta\sigma_{FWHM} = \frac{4}{\lambda_{min} N}. \tag{3.15}$$

3.6.3 Wavelength accuracy

The static nature of the spectrometers necessitates the use of a detector array to record the interferogram. The fixed number of detector elements limits the number of data points recorded in the interferogram, and the fast Fourier transform algorithm yields information only indirectly, suffering from systematic errors or bias, which is accentuated by the finite length of the input data [2].

However, careful calibration and the stability of the interferogram have enabled the production of spectrometers with moderate accuracy suitable for measuring broad-band spectral sources. Using a 1024 element detector array [23] a wavelength accuracy of approximately 0.2 nm at 633 nm has been demonstrated. A later design utilising a 287 element detector array, after confirmation with a selection of narrow-bandwidth, interference filters, demonstrated an accuracy of better than 1 nm over the operating wavelength range of 400–900 nm.

Higher accuracies up to 0.001 nm can be achieved when measurement is upon quasi-monochromatic light, such as a laser source. Operation of the instrument as a high accuracy, laser wavemeter is discussed in detail in Chapter 6.

3.6.4 Data acquisition rate

The speed of camera-computer interface essentially sets the data acquisition rate. Each version of the spectrometer has used a different camera and frame-grabber. Spectrum update rates therefore vary between 50 ms[23] and 0.25 s[25].

The static nature and spatially recorded interferogram of the Wollaston prism spectrometers enables the recording of transient phenomena such as short laser pulses and explosion spectra. A flash lamp synchronised with the spectrometer so that a complete interferogram is recorded for a single flash of the lamp has

been successfully demostrated[23].

Several detectors also enable extremely low light level operation where the camera integration time can be reduced to several seconds.

3.6.5 Field of view

The field of view(FOV) is an instrument property which effects the étendue of spectrometers and is instrumental in the 'Jacquinot advantage' for FTSs. The FOV also modifies the resolution of Michelson interferometer based FTS. Earlier designs of Wollaston prism based spectrometers had significantly smaller FOVs than that of a standard Michelson-based spectrometers with similar resolutions.

3.7 Conclusion

This chapter has reviewed previous designs of static Fourier transform spectrometers based upon Wollaston prisms. Specifically it has explained how many important properties such as resolution and wavelength accuracy can be determined. It has also briefly touched on an important aspect of spectrometer performance, namely the field of view.

A more detailed discussion of the field of view of Wollaston prism based spectrometers can be found in chapter 4, including designs and results for spectrometers with increased FOVs.

Chapter 4

Increasing the field of view of static Fourier transform spectrometers

4.1 Introduction

Many scientific and industrial applications require compact spectrometers with high étendue or optical throughput. The étendue E, as described in section 2.6.3, is defined as

$$E = A\Omega,$$

where A is the instrument aperture, and Ω, the solid angle field of view (FOV). Fourier transform spectrometers utilise this relationship to their benefit; they have a significantly higher étendue than dispersive instruments with similar resolutions[11].

The solid-angle FOV of early Wollaston prism spectrometer designs was poor compared to Michelson interferometer based spectrometers of a similar resolution. The resulting lower étendue lessened the Wollaston prism instrument's claim to

the 'Jacquinot advantage'. At the University of St Andrews we have therefore designed spectrometers with increased fields of view[25, 30].

As part of this process a program was written to model the optical properties of the spectrometers[27]. The FOV, Nyquist wavelength, and position of the fringe plane can be investigated for various Wollaston prism and detector array configurations enabling selection of an optimised spectrometer solution. I utilised this program to design and construct a spectrometer comprising two Wollaston prisms made from different materials with opposite sign of birefringence. The second prism compensates for the angular dependent path difference introduced by the first prism giving a very large FOV. The resulting instrument has an étendue comparable to that of a Michelson interferometer based spectrometer with a similar resolution.

4.1.1 Field of view definition

Ideally, the path difference in a spectrometer is independent of the incident angle of the light, depending solely on the mirror position in the case of a Michelson interferometer, or lateral displacement across the Wollaston prism in the case of a static FTS. In reality, however, the path difference does vary with angle of incidence. As the solid angle Ω_{source}, subtended by a source is increased, the variation in the path difference as a function of angle will decrease, ultimately to zero, the fringe contrast in the interferogram, resulting in a loss of spectral data.

The field of view Ω_{spec}, as shown in figure 4.1, determines the shape and size of the largest light source that does not significantly compromise the visibility of fringes.

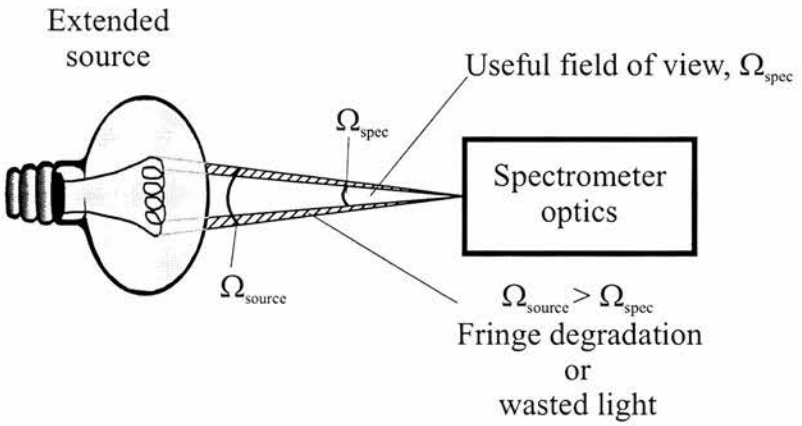

Figure 4.1: *The extended source presents a solid angle of Ω_{source} to the spectrometer. However, the spectrometer only operates with a field of view of Ω_{spec}, which is smaller than Ω_{source}. Therefore, there is interferogram degradation where $\Omega_{source} > \Omega_{spec}$. Alternatively, the spectrometer may operate with an aperture stop, blocking light entering the instrument at an angle greater than Ω_{spec}, the incident illumination from $\Omega_{source} > \Omega_{spec}$ is thus lost light.*

4.2 Field of view of a Michelson interferometer

When using an extended source with a Michelson interferometer based spectrometer, one should think both about rays parallel to and oblique to the optical axis. Consider the effect of a noncollimated beam of monochromatic light passing through the interferometer with a divergence half-angle of α. At zero retardation, the path difference between the central ray passing to the fixed and movable mirrors is zero, and there is also no path difference for the extreme rays. If the movable mirror is moved a distance d as shown in figure 4.2, there is an additional path difference introduced between the central and the extreme ray.

For a given relative mirror displacement, this yields an interferogram which is expanded because of the longer path difference of the oblique rays. This expansion of the interferogram produces a measured spectrum which differs from the true spectrum in a manner which depends on both wavelength and the solid angle

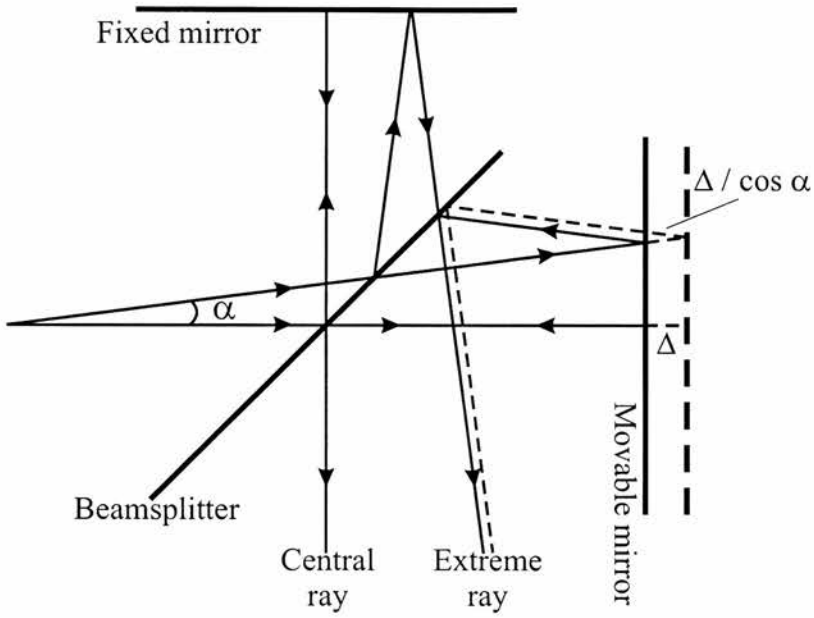

Figure 4.2: *Schematic representation of a diverging beam passing through a Michelson interferometer. The angle between the central ray and the extreme ray is α, and the physical distance moved by the mirror is Δ, corresponding to a nominal retardation of 2Δ for the central ray. The extreme ray experiences a retardation of $\Delta/\cos\alpha$.*

subtended by the source.

4.2.1 Effect of field of view upon the spectrum

When a source is extended in size this has two effects: (i) the computed spectrum is shifted to longer wavelengths by a small amount, introducing a wavelength spread, and (ii) there is a loss of measured power in the spectrum.

Consider the interferograms due to the central and extreme rays from a monochromatic source with reference to figure 4.2. For a particular half-angle α, the path difference between the central and extreme rays x, at a retardation Δ is given

by[3]

$$x = \frac{1}{2}\alpha^2\Delta \tag{4.1}$$

For the extreme ray, there is an increased path difference ($\Delta+x$), and the effective wavelength of this ray is therefore changed to a longer wavelength than that of the central ray.

If α is small, then to a first approximation, the wavelength of the line computed from this interferogram is the mean of the wavelengths of the central and extreme rays. However, a complete derivation leads to a calculated shift of twice the magnitude approximated, because the extreme rays have a larger contribution to the total signal than the central rays.[31, 32]

Assuming an on-axis point source, the interferogram is given by[7]

$$F(\delta) \equiv \int_0^\infty B(\sigma)s(\sigma,\delta)\cos[2\pi\sigma\delta + \varphi(\sigma,\delta)]d\sigma \tag{4.2}$$

When using the interferometer with an extended source the resulting oblique-to-axis optical rays lead to a modification of equation 4.2. It can be shown that the interferogram for some maximum solid angle Ω is given by[7]

$$F(\delta,\Omega) = \int_0^\infty \left\{ B(\sigma)\text{sinc}\left(\frac{\Omega\sigma\delta}{2}\right)\cos\left[2\pi\sigma\delta - \left(\frac{\Omega\sigma\delta}{2}\right)\right]\right\} d\sigma \tag{4.3}$$

The modulation of the interferogram in equation 4.3 by $\text{sinc}(\Omega\sigma\delta/2)$ has interesting results. The sinc function decreases the energy in the spectrum with increasing solid angle Ω. For $\Omega = 2\pi/\sigma\delta$, the sinc function reaches its first zero, and for $2\pi/\sigma\delta < \Omega < 4\pi/\sigma\delta$, the sinc function is negative. The negative values of the sinc function express the fact that the interference fringes have reversed phase. Thus, if $\Omega > 2\pi/\sigma\delta$, energy is effectively lost from the spectrum because of de-

structive interference. Therefore, for a FTS the field of view of the Michelson interferometer limits the solid angle over which light contributes usefully to the power in the interferogram.

4.2.2 Effect of field of view upon the resolution

Increasing the solid angle of the source has a far more fundamental effect on a FTS based upon a Michelson interferometer. In addition to reducing fringe contrast and introducing wavelength spread, the field of view is intimately linked with the resolution of the instrument.

The solid-angle FOV, Ω, for light of wavelength λ depends on the maximum path difference, $\Delta_{\max}$, and is approximated by [7],

$$\Omega \approx \frac{2\pi\lambda}{\Delta_{\max}}. \tag{4.4}$$

It can be seen from equation 4.4 that for a given wavelength, a larger field of view can only be attained when the maximum path difference is small. As discussed in chapter 2 the resolving power of a Fourier transform spectrometer also depends on the maximum path difference introduced between the interfering rays. Large $\Delta_{\max}$ giving high resolution. However, for a Michelson interferometer with a large FOV the maximum path difference is small, from equation 4.4, which leads to a reduction in resolution.

The resolution reduction can be explained intuitively by reference to figure 4.3, as the solid angle of the source is increased the resultant Fourier transformed spectrum suffers two effects. There is a wavelength spread introduced, and a reduction in the perceived power of the spectrum. This is shown in figure 4.3 where the spectrum of two monochromatic sources whose wavelengths are close

can be seen. The two distinct peaks that are generated when the interferometer was used with a small solid angle can be easily resolved from each other. However, when the interferometer was used with an large extended source each of the two peaks not only suffered wavelength spread but a loss in recorded power. This has the effect of making the two separate wavelengths increasingly difficult to resolve apart.

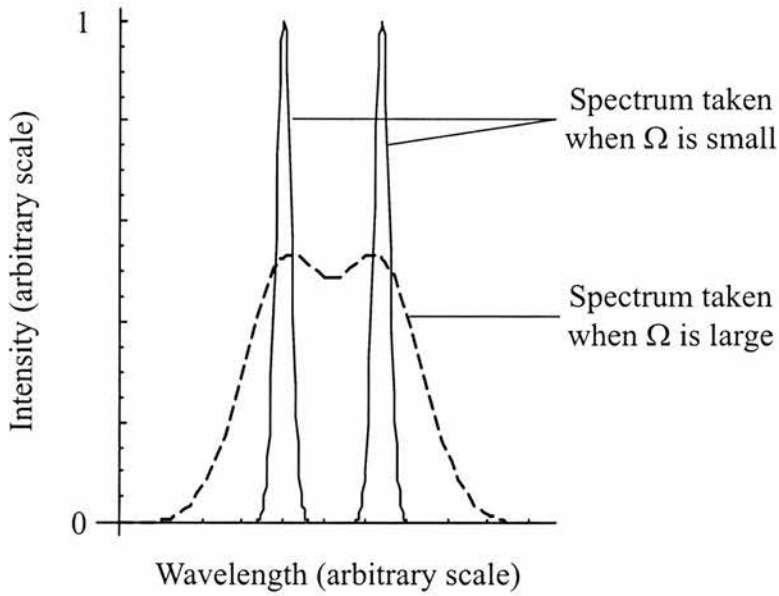

Figure 4.3: *Spectrum of two monochromatic light sources whose wavelengths are close together.*

4.3 Field of view of Wollaston prism spectrometers

For a static Fourier transform spectrometer based upon Wollaston prisms, the field of view of the prism may limit the solid-angle over which light contributes usefully to the power in the interferogram. When using a light source of extended angular size, the case of a Wollaston prism based spectrometer is no different from

the Michelson interferometer discussed earlier. There is an angular dependent path difference which limits the FOV, and if this is exceeded then the fringe contrast is reduced.

The standard derivation of the étendue for a Michelson interferometer based FTS considers the change in the path difference between the two arms of the interferometer as the angle of incidence is increased. Once the additional path difference between the on-axis and oblique rays exceeds $\lambda/2$, the oblique rays are subtracting power from the interferogram. Using a similar FOV criterion with a Wollaston prism based spectrometer the field of view, at a wavelength λ, can be usefully defined as the range of input angles over which the path difference does not vary by more that $\lambda/2$. The solid-angle FOV, Ω, of a Wollaston prism is given by [21],

$$\Omega \approx \frac{2n_o^2 n_e}{n_o^2 - n_e^2} \times \frac{\pi\lambda}{t}, \tag{4.5}$$

where n_e and n_o are the extraordinary and ordinary refractive indices and t is the thickness of the birefringent material.

The field of view of Wollaston prism based spectrometers can be understood visually with reference to figure 4.4. The fringe pattern on the left demonstrates the effect of different incident angles, α and β, on the FOV. The recorded FOV of a double Wollaston prism configuration can clearly be seen as a hyperbolic fringe pattern, where the dark areas correspond to $\lambda/2$ path difference between the rays, and hence a degradation in fringe contrast. The useful FOV is the central region of the plot, which in this example gives acceptance angles of approximately $\pm 2^o$. The hyperbolic pattern can be compared directly to the circular FOV pattern of a Michelson interferometer.

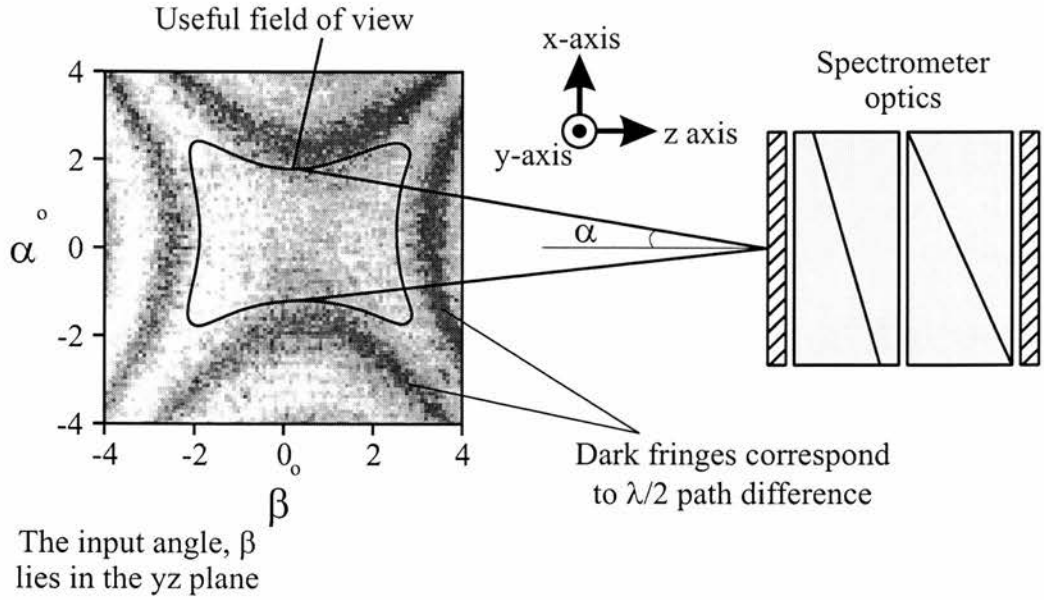

Figure 4.4: *Observed field of view for the double Wollaston prism spectrometer. The useful FOV corresponds with the marked central area. As the angle of incidence (α or β) increases the path difference between the two rays passing through the Wollaston prisms nears $\lambda/2$, the point where destructive interference reduces fringe contrast.*

4.3.1 Measuring the field of view

To determine the angular acceptance of the spectrometer, a light source can be projected, with a variable magnification, onto a ground-glass screen. This image on the glass screen can be used as the input source to the spectrometer. Changing the magnification of the source onto the screen allows the angular extent of the source to be varied without changing the light intensity significantly. The point at which the strength of a spectral line, as measured by the spectrometer, falls by 30% can be taken to be the limit of the useful field of view of the spectrometer.

Measuring the spectral peak strength or fringe contrast as a function of the angular size of the source, although giving qualitative agreement, cannot confirm the detailed structure of the field of view predicted by the model. By suitable

experiment the principle of reversibility[33] can be used to measure the field of view of the Wollaston prisms. Light emanating from a source and interfering at a detector has the same path difference as light emanating from the detector and interfering at the source.

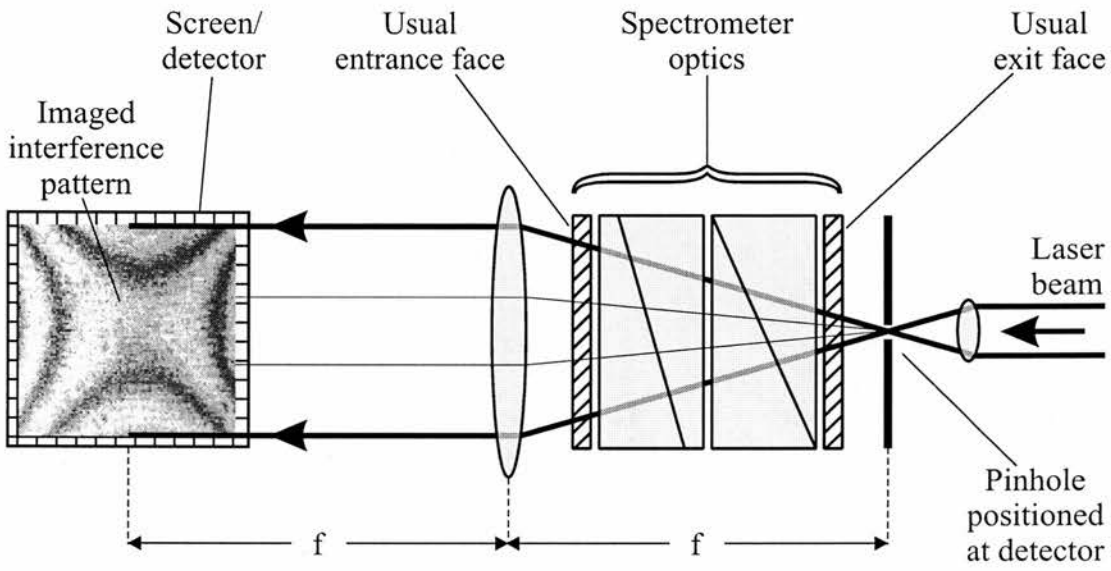

Figure 4.5: *Experimental arrangement for the observation of the field of view of a Wollaston prism spectrometer. The spectrometer optics are illuminated by a pinhole from behind. The resulting hyperbolic interference pattern is then focused onto a screen or detector for observation and measurement.*

The angular dependence of α and β, can be studied by positioning a highly divergent point light source, e.g. a pinhole illuminated by a laser, at a point D and observing the far field interference pattern on the other side of the Wollaston prisms. The experimental arrangement is shown in figure 4.5. In essence, the spectrometer optics are used in reverse, the pinhole is placed at the position of the detector, corresponding to the localised fringe plane, and illuminates the usual exit face of the optics. The interference pattern is the result of a ray traversing back through the optics from the detector.

The observed hyperbolic fringe pattern shows the field of view of the instrument

and confirms the numerical modeling. The separation between adjacent fringes corresponds to a change in the path difference of one wavelength. The dark fringes show the incident angles which cause $\lambda/2$ path difference and loss of fringe contrast. The useful FOV is therefore contained within the dark fringes and is the region we wish to maximise to enable maximum light throughput for the spectrometer.

4.3.2 Modeling the field of view

The operating principles of the Wollaston prism based spectrometer were outlined in chapter 3. However, the optical properties were only studied approximately, always assuming that the extraordinary ray encounters a refractive index $n \approx n_e$. This approximation is valid only as long as the light travels near-normally to the optic axes, which is the case of a Wollaston prism used at near-normal incidence. Consequently, if the effect of an extended light source is to be considered, a more accurate model for ray-tracing in birefringent media needed to be developed. One of the most important requirements is modeling the non-Snell's law refraction of the extraordinary-ray. However, most commercially available ray tracing packages do not deal with refraction of the e-ray in birefringent material.

Within this work at the University of St Andrews we have developed a ray-tracing program in Mathematica[34] that is applicable to birefringent materials[27, 30]. The program incorporates equations derived by Simon and Echarri[35] which describe ray propagation in birefringent media. It enables the calculation of path differences as functions of the angles of incidence, α and β, allowing the investigation of various prism configurations.

The design of the spectrometer is completely specified by the geometry, material and position of the Wollaston prisms and the relative position of the detector

array. Given these parameters, for any position on the detector array and any pair of angles of incidence, α and β we need to model the path difference between two light rays with mutually orthogonal polarisations that originate from the same point source. i.e. calculate the additional path difference as a function of angle for any point on the detector array. For a spectrometer with a large field of view, it is necessary to make the angular dependent path difference as small as possible.

When modeling and optimising the FOV of the Wollaston prism additional complexity arises from the necessity to ensure the geometrical relationship between the prism and the detector array is maintained. Manipulation of the Wollaston prism configuration to obtain a large field of view additionally affects other important parameters such as localisation of the fringe plane and hence positioning of the detector array behind the exit face of the prism. The Nyquist wavelength is also modified by altering the prism geometry and care needs to be taken to ensure the Nyquist limit is below the detector sensitivity or the minimum wavelength of interest, and can be derived from equation 3.3.

The difficulty when modeling the FOV of the spectrometer arises in knowing what is the optimum optical configuration which will give a large field of view, i.e. maintaining a small angular dependent path difference. The optimum Wollaston prism designs are arrived at using an iterative process, the program user inputs various configurations and the program calculates the resultant FOV. When the user feels they have found the design which meets the problem specification the process is stopped. Example FOV output from the program can be seen in figure 4.6.

The program also calculates and displays the position of the detector array relative to the exit face of the Wollaston prism. As an example, figure 4.7 shows the program output for the early double Wollaston prism design [24] comprising two

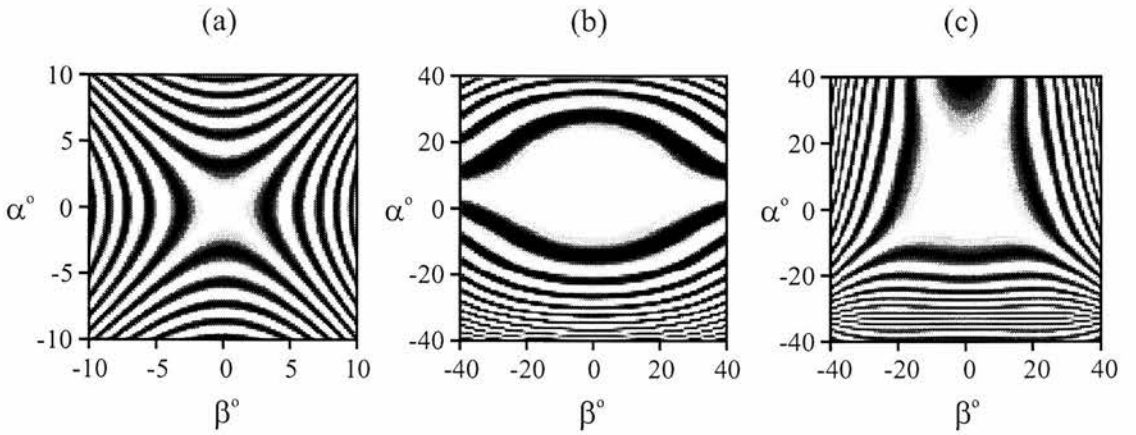

Figure 4.6: *Sample output from the birefringent ray tracing program. Showing the predicted fields of view for 3 different Wollaston prism configurations.*

4 *mm* thick, 10 *mm* aperture calcite prisms with internal wedge angles of 3.6^o and 5.9^o, matched to a 6 *mm* wide detector array. The two prisms are on the left of the fringe plane diagram, the vertical planes are the entrance/exit faces, and the inclined planes are the wedge angles. The right most plane where the fringes are localised, i.e. the required position of the detector array, and in this case is approximately 2 *mm* behind the exit face of the second prism. The FOV for this configuration has been calculated to be approximately $\pm 2^o$ and the hyperbolic fringe pattern is displayed graphically on the right hand side of the figure.

Further detail on the Mathematica model, including a program listing can be found in the thesis by Courtial [27].

Confirmation of the field of view

The accuracy of the numerical model can be confirmed by measuring the field of view of the spectrometer using the method discussed in section 4.3.1.

Confirmation of the model can be seen in figure 4.8. The interference pattern

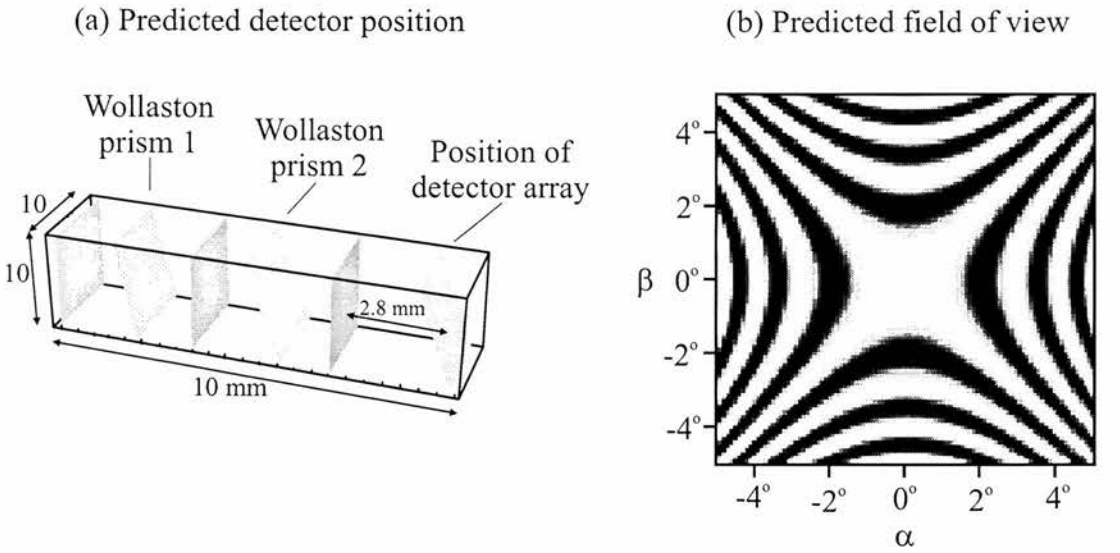

Figure 4.7: *Sample output from the Mathematica modeling program [27]. The double Wollaston prism design [24] has been modeled, calculating the position required for the detector array and the FOV of the Wollaston prisms.*

in 4.8 (a) is the predicted FOV for the two calcite prism spectrometer[24], and 4.8 (b) is the recorded and measured FOV respectively.

4.4 Increasing the field of view of a Wollaston prism based Fourier transform spectrometer

The resolution of a SFTS, although primarily determined by the maximum path difference is closely related to the number of pixels on the detector array and the Nyquist wavelength, as seen in equation 3.15. The path difference from the Wollaston prisms is configured to match the capabilities of the chosen detector, ensuring that the fringe plane which is imaged onto the detector array does not

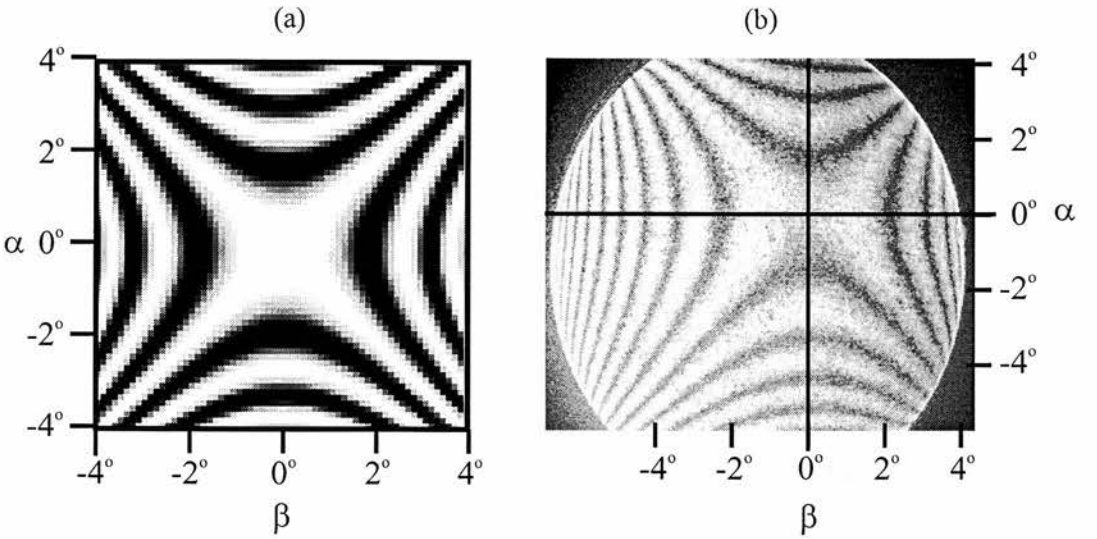

Figure 4.8: *Calculated and measured field of view for the original double Wollaston prism spectrometer.*

violate the Nyquist criterion for the lowest wavelength of interest. It is, therefore, useful to consider how the field of view of the Wollaston prisms could be increased while not compromising the maximum path difference afforded by the detector array and the Nyquist wavelength.

The solid-angle FOV of a Michelson interferometer can be increased by reducing the maximum path difference achievable between the two arms of the interferometer, as shown in equation 4.4. In contrast, as we have already seen in equation 4.5, the FOV of a Wollaston prism is given by

$$\Omega \approx \frac{2n_o^2 n_e}{n_o^2 - n_e^2} \times \frac{\pi\lambda}{t},$$

where n_e and n_o are the extraordinary and ordinary refractive indices and t the thickness of the birefringent material.

Importantly, it can be seen from equation 4.5 that the field of view of a Wollaston

prism based spectrometer is inversely proportional to the total thickness of the birefringent optical material.

Again, using the example of the twin Wollaston prism design [24] discussed in section 4.3.2. The spectrometer comprises two calcite prisms, each 4 mm thick, putting the numbers into equation 4.5 gives an approximate FOV of $\pm 2^o$ at a wavelength of 633 nm.

This prism configuration, combined with a 767 element, 6 mm wide detector array gives a maximum path difference of ± 45 μm. From equation 4.4 a standard Michelson interferometer operating with a maximum path difference of ± 45 μm has a FOV of $\pm 9.6^o$. However, unlike a Michelson interferometer, it is possible for the field of view of Wollaston prism based spectrometers to be increased without significantly reducing the resolution.

4.4.1 Inclined optic axis Wollaston prism

Studying equation 4.5 leads to the obvious and intuitive method for increasing the FOV, simply decrease the thickness of the prisms. The field of view of a Wollaston prism spectrometer is inversely proportional to the total thickness of the birefringent optical components.

A spectrometer design already discussed in detail in section 3.5 improves on this simple idea further. The inclined optic axis Wollaston prism reduces the thickness of the birefringent material by eliminating the need for the second prism. Using a 1.5 mm thick Wollaston prism with a 2.4^o wedge angle and an optical axis inclined at 12.4^o with respect to the wedge interface again gives a path difference of ± 45 μm across a 6 mm wide detector array. The field of view for this spectrometer configuration has increased to approximately $\pm 5^o$, as seen

in figure 4.9.

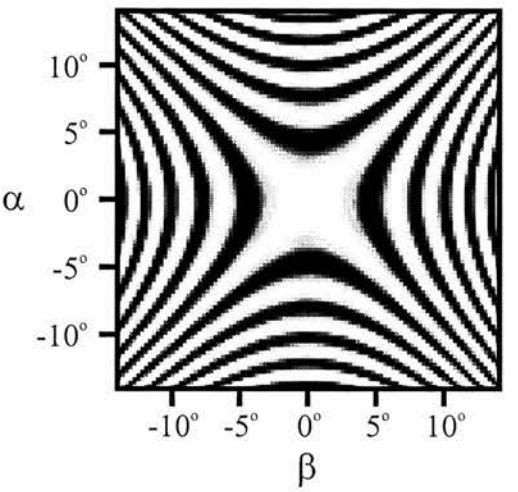

Figure 4.9: *Calculated field of view for the inclined optic axis Wollaston prism spectrometer.*

However, the FOV is not as high as a comparable Michelson interferometer based spectrometer with the same resolution, or the following Wollaston prism configurations. This design of spectrometer does have the additional advantage of reducing the component count and therefore the cost of any instrument.

4.4.2 Two Wollaston prisms and a half-wave plate

The FOV can be increased further by the addition of a compensation plate[21]. In essence, compensation of the angular dependence of the path difference arises if one ensures that the angular path difference that is introduced by the compensation plate is equal in magnitude but opposite in sign to that of the Wollaston prism. This principle will be discussed later in this chapter, suffice it to say here, that using the double calcite Wollaston prism design it is possible to achieve this compensation by inserting a half-wave plate between two Wollaston prisms of the same sign of birefringence. The optical layout of such a system can be seen in figure 4.10.

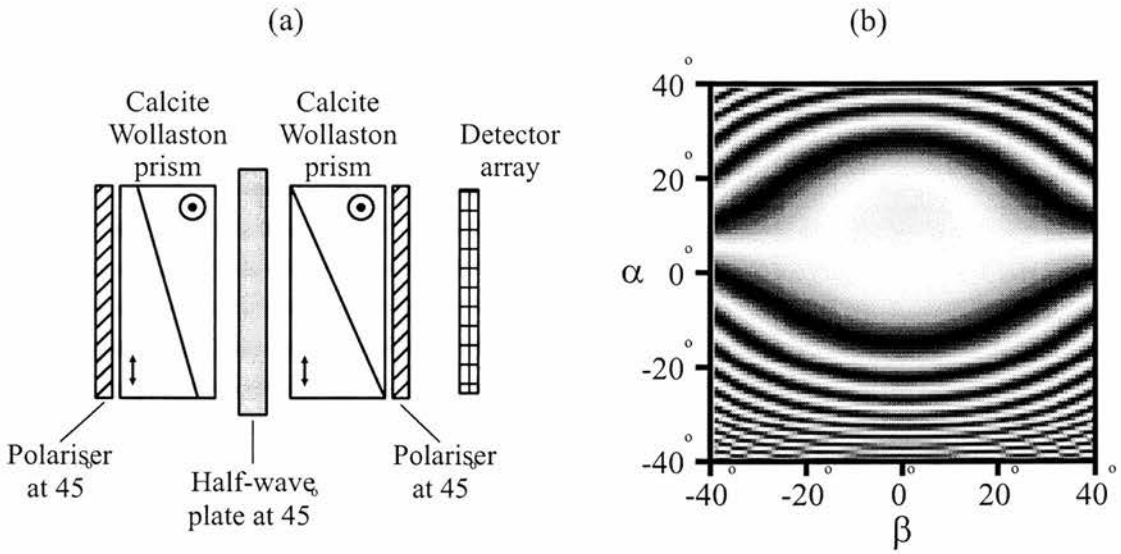

Figure 4.10: *(a) Optical layout of the wide-angle spectrometer based upon two Wollaston prisms and an achromatic half-wave plate. (b) Calculated field of view for the spectrometer.*

The predicted FOV for the two Wollaston prisms and the half-wave plate shows an approximate ten fold increase over the same prisms without the addition of the compensation plate. This can be clearly seen in the modeled FOV in figure 4.10 (b).

4.4.3 Two Wollaston prisms of opposite sign birefringence

The angular dependent compensation can also be achieved by using two Wollaston prisms of opposite sign birefringence. My work on increased FOV Wollaston prism based spectrometers has focused on the design and construction of an instrument using materials with opposite signs of birefringence for the prisms.

4.5 Construction of a wide field of view spectrometer using Wollaston prisms of opposite sign birefringence

Previous designs of Wollaston prism based FTSs did not exhibit the full extent of the "Jacquinot advantage" because of the limited field of view of the Wollaston prism itself. In the remainder of this chapter I describe how two Wollaston prisms fabricated from materials with opposite sign of birefringence can increase significantly the field of view. This in turn increases the optical throughput to a level where it is comparable to a Michelson based FTS of the same resolution. The portability of the instrument was also an important design requirement and is ensured by using a PCMCIA card to interface the Wollaston-based FTS to a laptop computer.

4.5.1 Increasing the field of view using a compensation plate

The field of view of a Wollaston prism can be increased with the addition of a compensation plate. A standard compensation plate resembles a Wollaston prism with no wedge angle and is fabricated from material with opposite sign of birefringence to that of the Wollaston prism itself. Compensation of the angular dependence of the path difference arises by ensuring the path difference introduced by the compensation plate is equal in magnitude but opposite in sign to that of the Wollaston prism. For optimum cancellation, the relative thickness of the Wollaston prism, t_W, and of the compensation plate, t_c is given by[21],

$$\frac{t_W}{t_c} = \left| \frac{\left(\frac{n_e^2 - n_o^2}{n_o^2 n_e}\right)_c}{\left(\frac{n_e^2 - n_o^2}{n_o^2 n_e}\right)_W} \right|. \tag{4.6}$$

In the FTS the second Wollaston prism can be used, not only to localise the fringe plane behind the prism pair, but also to act as a compensation plate for the first prism. For example, two suitable materials are rutile, which has a positive birefringence ($n_e > n_o$) and calcite, which has a negative birefringence ($n_e < n_o$). For operation in the visible region of the spectrum equation 4.6 requires the rutile prism to be 1.6 times that of the calcite prism. Combining these prisms with the 767 element, 6 mm wide detector array used previously gives a predicted FOV of approximately $\pm 20^o$ with the same maximum path difference of $\pm 45 \mu m$, an improvement over a Michelson interferometer based spectrometer. A plot of the calculated FOV can be seen in figure 4.11.

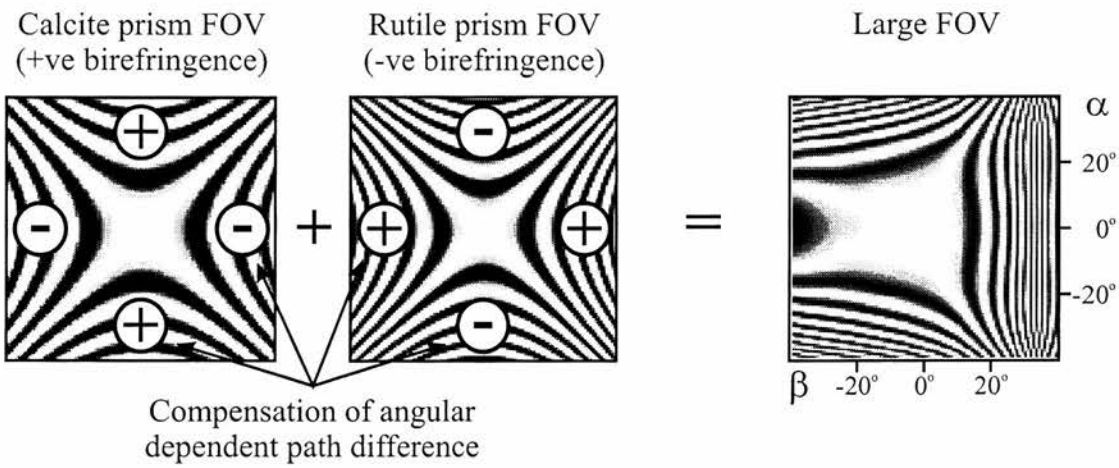

Figure 4.11: *Simple graphical explanation of the compensation of the angular dependent path difference using two Wollaston prisms of opposite sign birefringence. The path difference introduced by the second prism, the compensation plate, is equal in magnitude but opposite in sign to that of the first prism.*

The inclusion of rutile into an optical system has two disadvantages: cost and a short wavelength cutoff at 420 nm. Alternative materials allowing operation

from 180 nm to 1.7 μm are quartz and ammonium dihydrogen phosphate (ADP). Quartz[29] has a positive birefringence of +0.009, which is compensated by the negative birefringence of −0.05 in the ADP[29]. To maximise the field of view in the visible region of the spectrum, equation 4.6 implies that the thickness of the quartz Wollaston prism should be approximately five times that of the ADP prism.

4.5.2 Design of a Fourier transform spectrometer incorporating Wollaston prisms of opposite sign of birefringence

To realise the true potential of the compact design of static FTSs it was important to combine the Wollaston prisms and polariser assembly with an appropriately sized detector. We used a highly-integrated CMOS camera which fits onto a $30mm$ diameter PCB and has a total power consumption of less than $150mW$[36]. In addition to a 312×287 array sensor with a total active area of $6.12 \times 4.59mm$, the camera includes the necessary drive circuits to deliver a formatted CCIR composite-video signal. The camera is mounted behind the exit face of the second Wollaston prism to coincide with the fringe plane.

The computer model[30] as described in section 4.3.2 was used to determine the optimum configuration of the spectrometer. Various designs were investigated based upon the detector specification above and two Wollaston prisms of opposite sign of birefringence. In the optimised design both Wollaston prisms and the polarisers have an aperture of $10mm$[29]. The first prism – fabricated from synthetic quartz – is $10mm$ thick and has an internal wedge angle of 23°. The $2mm$ thickness of the second prism – fabricated from ADP – results in a large FOV of the prism pair. An internal wedge angle of 10.3° gives a path difference of

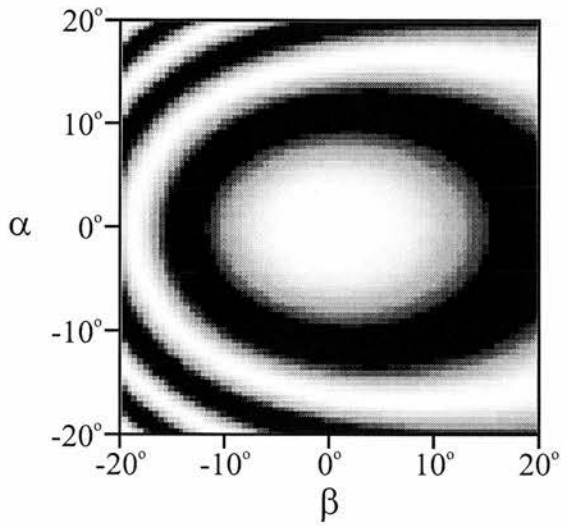

Figure 4.12: *Calculated field of view for the spectrometer based on two Wollaston prisms of opposite sign birefringence.*

$\pm 28\mu m$ across the aperture of the detector. In conjunction with the 312 element detector array the Nyquist wavelength is set to $355nm$. The separation between the two prisms is $0.5mm$, resulting in a localised fringe plane $3.5mm$ behind the exit face of the second prism. The calculated FOV for this optical layout can be seen in figure 4.12 which shows an expected FOV of approximately $\pm 10^o$. For operation in the visible region of the spectrum simple photographic-quality polarisers are adequate. These are aligned with their polarisation axes at 45° to the edges of the Wollaston prism. The optical layout of the spectrometer is illustrated in Figure 4.13.

The camera outputs data via a PCMCIA card to a desktop or laptop computer with a PCMCIA socket and appropriate software. The raw data from the camera is a 2-D image of the interference fringes which is averaged across the frame to give a single 312 element data array. The array is zero-padded to 512 elements and a fast-Fourier-transform of the interferogram data produces an estimate of the optical power spectrum. For ease of use the control and data-processing software has been written for a windows-based[37] environment. Through simple mouse

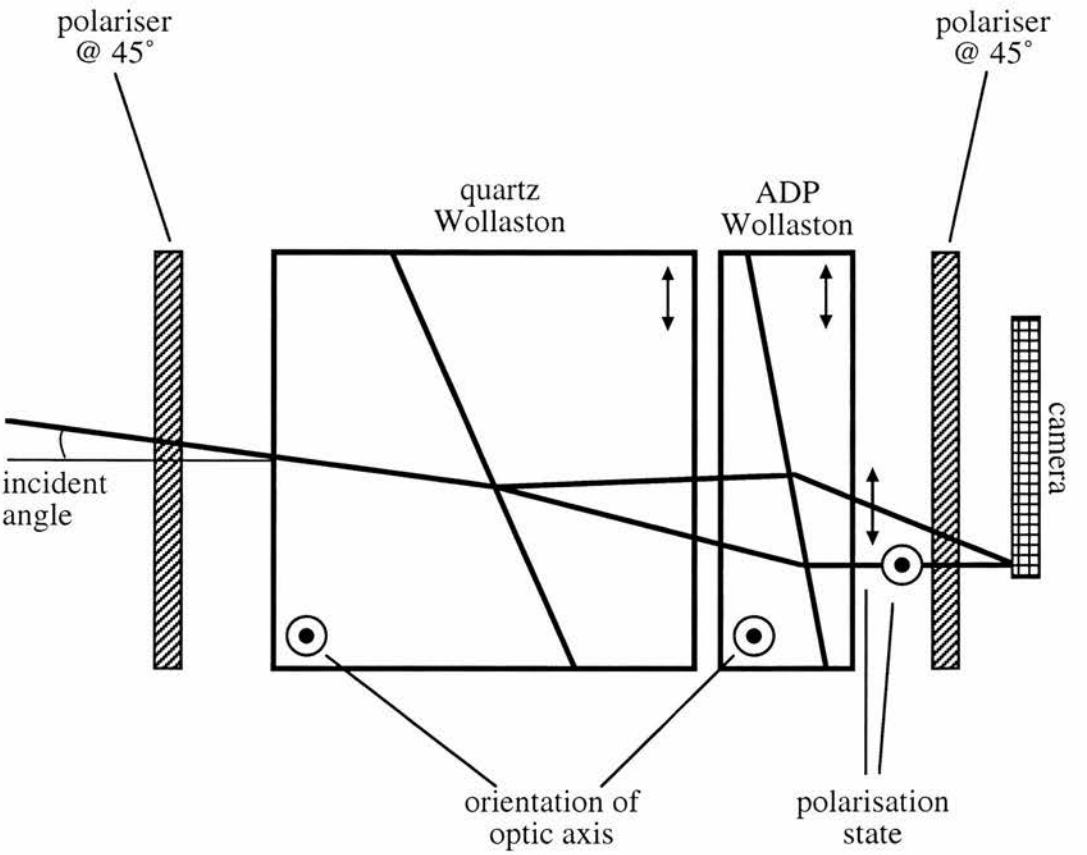

Figure 4.13: *The optical layout of the wide field of view spectrometer based on Wollaston prisms of opposite sign of birefringence.*

button clicks the user can view: the 2-D image of the interference fringes; the raw interferogram data; the apodised interferogram or the optical power spectrum. A photograph of the complete spectrometer and controlling laptop computer is shown in Figure 4.14.

Figure 4.14: *The wide field of view, ultra-compact, static Fourier-transform spectrometer interfaced to a laptop computer.*

4.5.3 Results of a Fourier transform spectrometer incorporating Wollaston prisms of opposite sign of birefringence

The key considerations when assessing the performance of the spectrometer relate to the resolving power, wavelength range, acquisition rate, and field of view.

Using a selection of narrow-band, interference-filters the resolving power of the spectrometer was demonstrated to be approximately $350cm^{-1}$, as dictated by the maximum path difference within the interferogram. The operating spectral range of $400 - 900\ nm$ for the instrument is limited by the specification of the

polarisers. Alternative polarisers would allow access to the full $355 - 1050\ nm$ sensitivity range of the camera. The speed of the camera-computer interface sets the data acquisition rate to approximately four frames per second.

The field of view characteristics can be assessed by measuring the fringe contrast as a function of source size. A laser beam expanded and transmitted through a rotating ground glass screen gives a convenient method for obtaining a monochromatic source of variable angular extent. Using such a system we were able to confirm that the useful field of view of the FTS spectrometer was $\pm 10°$, close to the theoretical maximum for a Michelson based FTS of the same resolution. A FTS based on a Michelson interferometer with a maximum path difference of $28\mu m$ has a FOV at $633\ nm$ of $\pm 12°$. This feature combined with the collection aperture of $6 \times 4.6\ mm$ means that this design leads to a compact, robust spectrometer having a high optical throughput.

Representative performance of the instrument is given in Figure 4.15 which shows the interferogram and corresponding spectrum of a He-Ne laser and a diode laser recorded using the spectrometer.

The modest resolution is the main design limitation. Although the resolution can be improved by increasing the number of elements on the detector array, and appropriate redesign of the Wollaston prisms, this can only be achieved in larger overall package size.

4.6 Conclusions

I have used two Wollaston prisms of opposite sign birefringence, where the second prism acts as the compensator, to produce a compact spectrometer with a large field of view. This spectrometer is not only extremely robust and compact

but has the largest field of view, and hence highest optical throughput, of any Wollaston prism spectrometer so far constructed. Enhancing both the low light level operation and ease of instrument alignment.

The current performance specifications are such that I believe the spectrometer will have numerous applications in the field of general laboratory instrumentation, and particularly where field based operation is required.

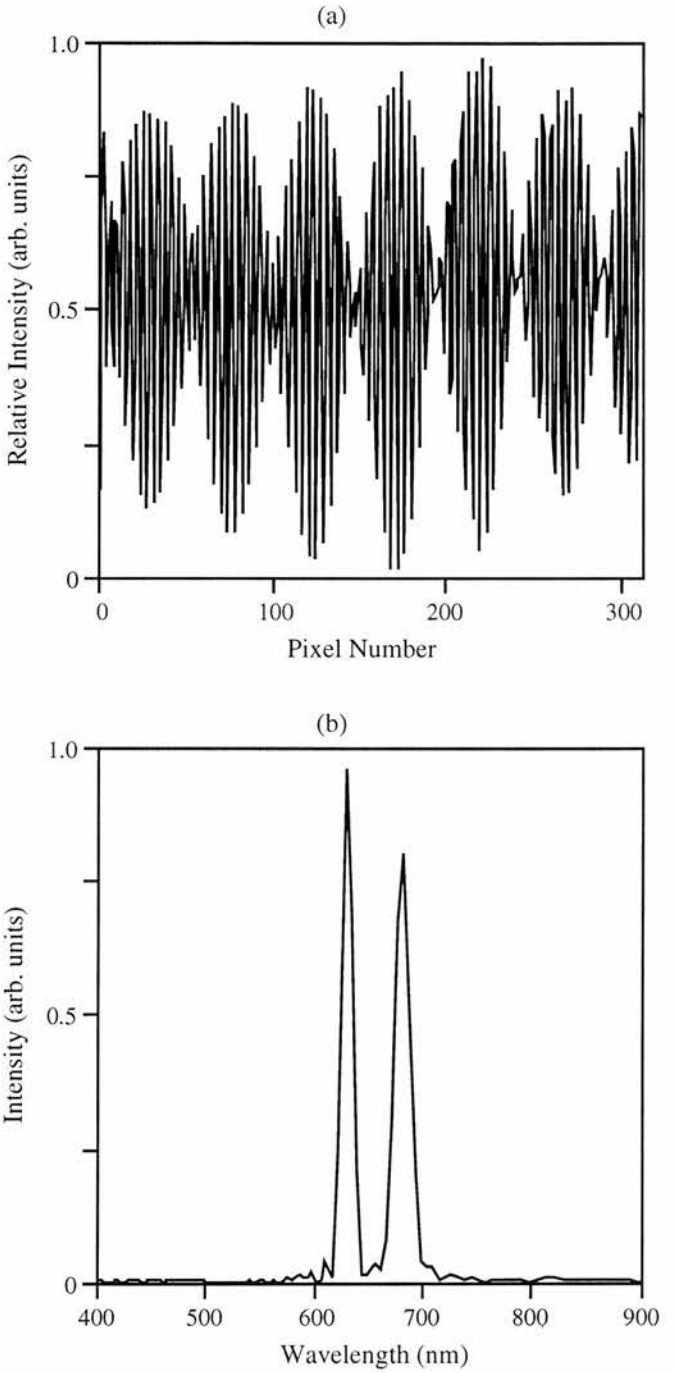

Figure 4.15: *The interferogram (a) and corresponding spectrum (b) of a He-Ne laser(633nm) and a diode laser(670nm).*

Chapter 5

Petroleum measurement utilising Wollaston prism based spectrometers in the near infrared and ultraviolet

5.1 Introduction

Gasoline is a complex mixture of a range of refinery products, blended under strict regulatory controls to produce fuel conforming to exacting supply specifications. In the Western world sophisticated distribution systems ensure product quality is maintained from the refinery to the customers and quality assurance (QA) on the forecourt is rarely required. However, since the majority of the fuel retail price is tax and duty, which varies significantly by country and fuel type, some distributors may be tempted to adulterate gasoline of one type with another or to illicitly obtain fuels when near national borders. This can lead to a lower

quality product for the customer, reputation damage and loss of revenue for the producer and duty authorities. In some circumstances it can therefore be necessary to check the quality of the fuel on the forecourt, which often involves taking a sample and returning it to a central laboratory for analysis. This slow process limits the number of samples that can be tested. The ability to screen fuels on the forecourt would not only speed up this operation but enable the retesting of suspect fuels and further samples to be removed for more detailed analysis at a well equipped laboratory.

Gasoline quality is defined by national and international standards which specify the measurement method to be used. The key property of gasoline is octane number, and the standard for reporting this measurement is an internal combustion engine in which octane is measured by interpolating between the nearest standards above and below the unknown sample. Additional techniques for gasoline monitoring include chromatography[38] and vapour samplers comprising a selection of chemical sensors[39, 40]. All these methods are time consuming, requiring several hours for a measurement, involve expensive and maintenance-intensive equipment, skilled labour, and are usually lab based, making real-time process control and forecourt measurement impractical[41]. In recent years there has been increasing interest in the use of optical measurement systems to predict fuel properties for quality control during manufacture and the identification of adulteration or counterfeiting on the forecourt. To date, such investigations have concentrated on the rotational-vibrational transitions in the mid- and near-infrared regions of the spectrum[42, 43, 44, 45].

5.1.1 Infrared spectra and molecular vibrations

To understand the interpretation of IR spectra we need first to examine the nature of molecular vibrations, for it is these that give rise to the spectra. Spectral

features may be due to an isolated structure in the molecule such as a C=O (carbon-oxygen double bond) or C=C group, a specific type of molecule such as a steroid or carbohydrate, or to a characteristic of a macromolecule, such as protein chain conformation [46]. Within this thesis there is not a lengthy or detailed discussion on molecular vibrations and their relation to infrared spectra, but it should be clear that many of the chemicals in the oil and petroleum industry are well suited to measurement using infrared spectroscopy. A more informative and indepth presentation may be found in the many textbooks on the field e.g.,[47, 3].

Most of the basic aspects of molecular vibrations are displayed by the simplest molecule, a diatomic with formula AB (e.g. CO, N_2). To begin with, let us assume the molecule is distorted from its equilibrium state by an increase in the length of the bond, δr, and that the force tending to restore the molecule to its equilibrium state is proportional to δr. This is a statement of Hooke's law, which applies to any spring if the displacement δr is sufficiently small. It follows from quantum mechanical considerations, that the energy of the molecule resulting from this distortion, the vibrational energy E_{vib}, is given by equation 5.1,

$$E_{vib} = \left(\upsilon + \frac{1}{2}\right) h\nu \tag{5.1}$$

where h is Planck's constant, υ is a quantum number which can take any integral value including zero, and ν is the frequency of vibration (s^{-1}).

If the molecule is made to gain or lose a quantum of energy ΔE equal to $h\nu$, the vibrational quantum number υ will change by one unit. The energy may be directly absorbed (or emitted) as IR radiation of frequency ν giving rise to an IR spectrum. Fundamental vibrations have frequencies in the approximate range 1.0×10^{12} to 1.2×10^{14} s^{-1}, corresponding to wavelengths of $300 - 2.5\mu m$. A non-linear (polyatomic) molecule with N atoms has a total of $3N - 6$ vibrations.

Not all of these vibrations will have different frequencies, but nevertheless a large molecule will have a number of distinct vibration frequencies [46].

5.1.2 Portable mid-infrared gasoline measurement

Optical analysis offers the potential for simple, rapid, non-contact measurement of fuels. With a global industry worth billions of pounds and potential savings of hundreds of millions of pounds for gasoline producers if accurate, convenient measurements could be made, many spectrometer manufacturers have developed mid-infrared instruments dedicated to the analysis of fuels. These 'turn key' multi analysis systems cost from £150,000 upwards. Even single-stream analysers, such as octane number systems cost from £100,000 upwards. The high quality of spectra that can be recorded using these commercial instruments can be seen in figure 5.1.

Unfortunately, despite their ability to accurately determine fuel properties, these instruments are based upon Michelson interferometers making them bulky and lab based. The need for a field portable device that could be used to monitor fuels motivated Shell to develop a simple compact instrument that operated in the MIR. The instrument utilises a rotating filter wheel to record absorbance at 10 carefully selected wavelengths, the wavelengths of the filters can be seen in figure 5.1. A photograph and drawing of the instrument layout can be seen in figure 5.2. This simple instrument measured octane and cetane numbers and is commercially available at a cost effective price of £ 20,000.

However, the analysis of complex mixtures such as gasoline will invariably give overlapping spectral features, which means that absorption measurement at only a few wavelengths may be insufficient to distinguish between different fuels. Selectivity and the ability to determine multiple properties requires a necessary

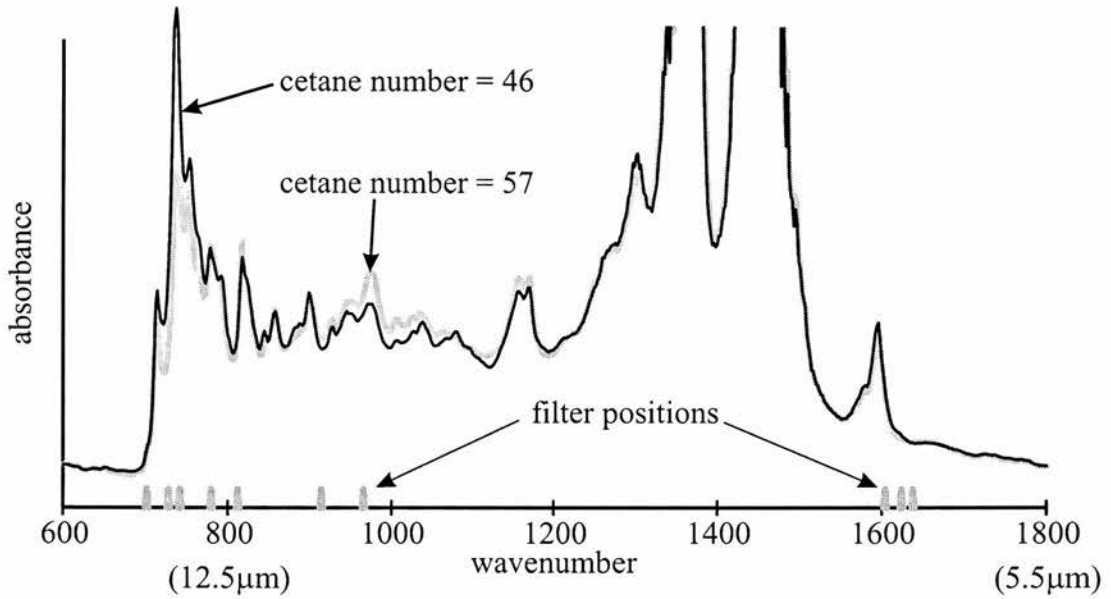

Figure 5.1: *Mid-infrared spectra of two diesel fuels. The fuels differ by their cetane numbers, which are 46 and 57 in this example. The filter positions are those used in the octane/cetane meter in figure 5.2.*

measurement at many wavelengths so multichannel spectrometers provide a better solution.

In this chapter, I report the use of static Fourier transform spectrometers for measuring the near-infrared and ultraviolet absorption spectra of liquid phase gasoline. Subsequent analysis of the spectrum using multivariate statistical methods allows us to determine octane number or identify gasoline brand.

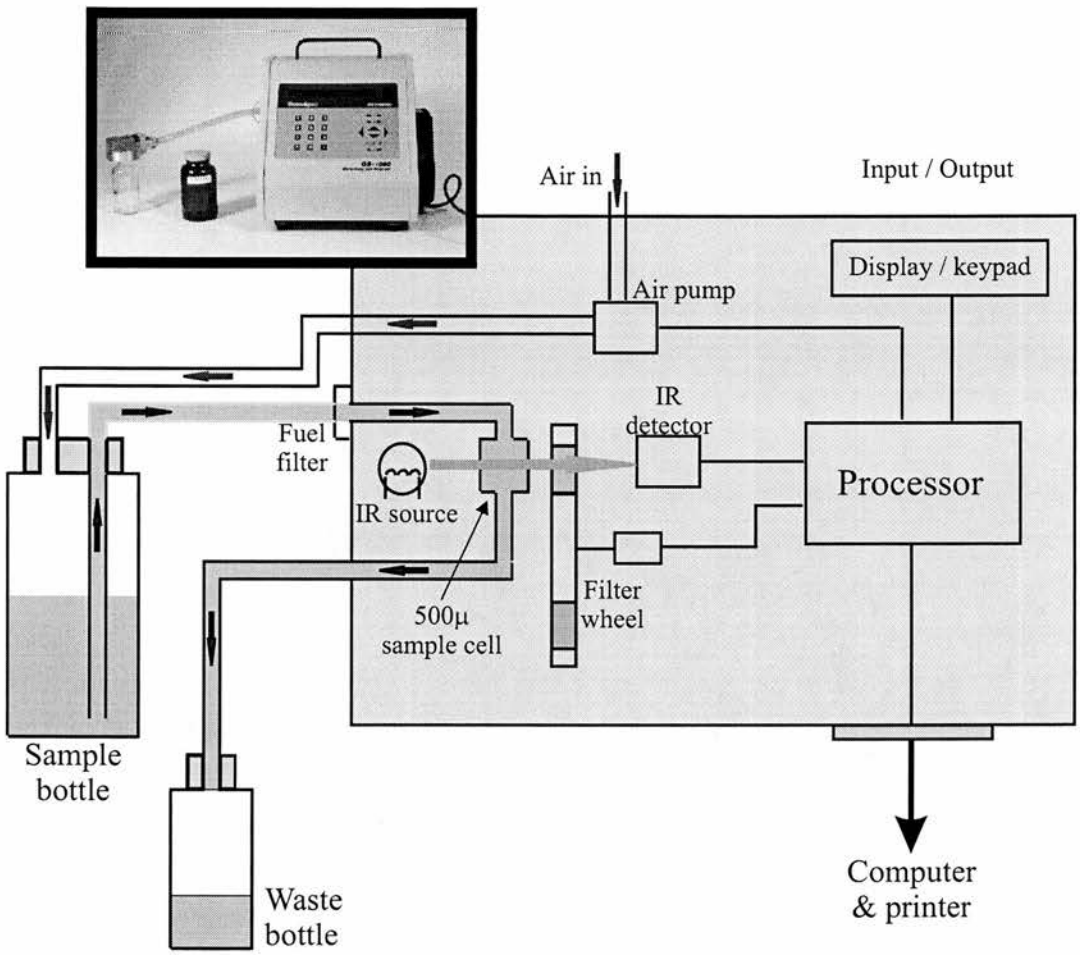

Figure 5.2: *Photograph (inset) and layout of the Shell octane and cetane measurement instrument.*

5.2 Infrared Fourier transform spectrometers based upon Wollaston prisms

As discussed in chapter 2, Fourier transform spectrometers (FTSs) are measurement tools that operate in the ultraviolet, visible, and infrared (IR) regions of the spectrum. However, they are particularly powerful in the IR because of the "Jacquinot advantage", which states that FTSs have a 190 times higher optical throughput compared with dispersive instruments of equivalent resolution[11].

This enables high resolution spectra to be recorded despite the weak intensity of infrared sources and low energy of IR radiation. The ability to work with low intensity radiation and maintain a high signal-to-noise ratio means FTSs are preferred for many spectroscopic applications in the mid-IR region of the spectrum. Beyond a few μm, moving further into the IR, background and noise problems become increasingly severe until FTSs are in fact the only instruments of choice. The domination of instruments in the infrared has lead many scientists and engineers to associate IR spectroscopy with FTSs. Presentations on the UV and visible versions of Wollaston prism based FTS have invariably drawn questions on the feasibility of FT-IR spectrometers based upon the same technology.

5.2.1 Mid-infrared Wollaston prism based spectrometer

The ultra-compact, Wollaston prism based spectrometer discussed in chapter 4, operating in the visible region, was constructed as precursor for a potential IR spectrometer. The opposite birefringence materials increased significantly the acceptance angle, and therefore the light gathering power of the instrument. This is particularly important if the sample and/or light source is close to the entrance aperture of the spectrometer. Consider the case of an in-line monitor with the light source and spectrometer head on opposite sides of a flow pipe. The spectrometer has the large entrance aperture and large field of view required to use all the light from the extended source, giving higher light throughput than grating or fibre based systems. Its robust design is also suited for an industrial environment.

The design of the static FTSs often begins with selection of an appropriate detector array. The most important property is wavelength sensitivity, but others such as size, cost, data acquisition rate, and number of pixels can be equally critical depending upon the application.

The static spectrometers discussed in previous chapters have all operated in the visible region of the spectrum. They have used 1-D or 2-D detector arrays with up to 1024 elements across the aperture. The cost of these arrays ranged from approximately $£100 - £1000$, giving very affordable instrument construction.

Unfortunately, infrared sensitive detectors are manufactured from relatively exotic materials which are more expensive and difficult to engineer. The linear arrays for operation in the mid-infrared were available with a maximum of only 256 elements. As we have discussed earlier in section 4.4 the resolution of the spectrometer, although primarily limited by the maximum path difference, is closely related to the number of pixels on the detector array. The limited resolution achievable with a 256 element detector makes them unsuitable for the work the spectrometers were to be designed for. In addition, at a cost of approximately $£9000 - £35000$ they were also considered too expensive.

Despite the unsuitability of IR detector arrays due to performance or cost constraints it was possible to design a low cost mid-infrared spectrometer based upon Wollaston prisms. Instead of an array, a single element detector could be used in conjunction with a scanning mechanism. We chose to construct a spectrometer with a fixed detector and a movable Wollaston prism, the basic design of which can be seen in figure 5.3.

5.2.2 Design of the scanning Wollaston prism spectrometer for operation in the mid-infrared

A single element detector manufactured from lead selenide (PbSe) with a spectral sensitivity of $1\mu m - 5\mu m$ allows operation in the mid-infrared region of the spectrum. The detector has an active area of $2mm \times 2mm$ and operates at room temperature avoiding the need for complex and bulky cooling mechanisms.

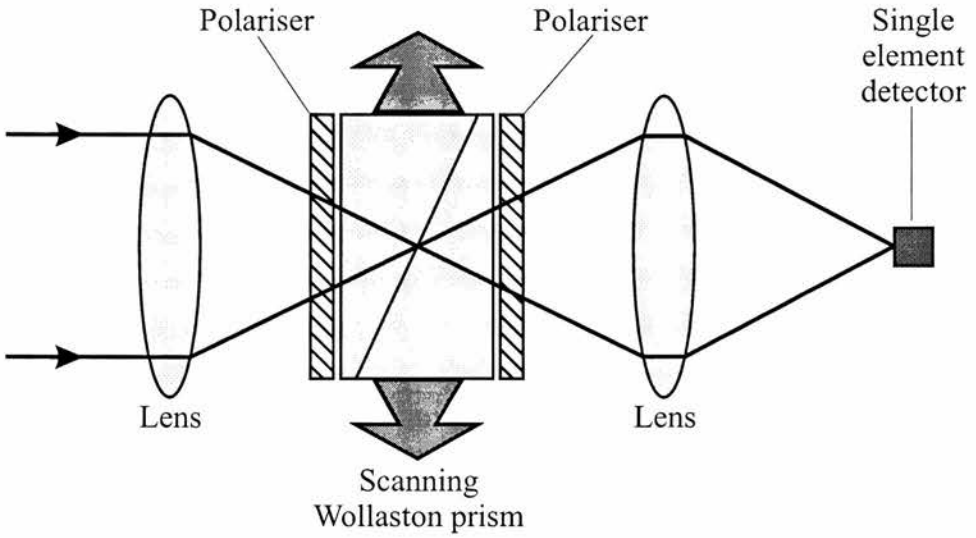

Figure 5.3: *Optical layout of the spectrometer when used with a single element detector. In this design the Wollaston prism is scanned from side to side enabling the light recorded at the single element detector to traverse the prism at various points across the wedge interface.*

The Wollaston prisms are manufactured from magnesium fluoride (MgF_2), which has a birefringence of 0.01 and a $0.11\mu m - 8\mu m$ spectral transmission. In this design we have used two prisms configured to maximise the path difference across the aperture. Each prism is 30 mm long, 15 mm high, and 15 mm thick, with a 26.56^o wedge angle. From equation 3.2 the maximum path difference, Δ_{max}, for one prism is $150\mu m$, used in the two prism setup this doubles to a Δ_{max} of $300\mu m$.

Positioned either side of the Wollaston prisms were IR sheet polarisers with a spectral range of 0.7–2.2 μm. The prism and polariser assembly was attached to a sliding mount and connected to a linear stepper motor. Thus enabling the prisms to be scanned from side to side. The layout of the FT-MIR scanning Wollaston prism spectrometer can be seen in figure 5.4.

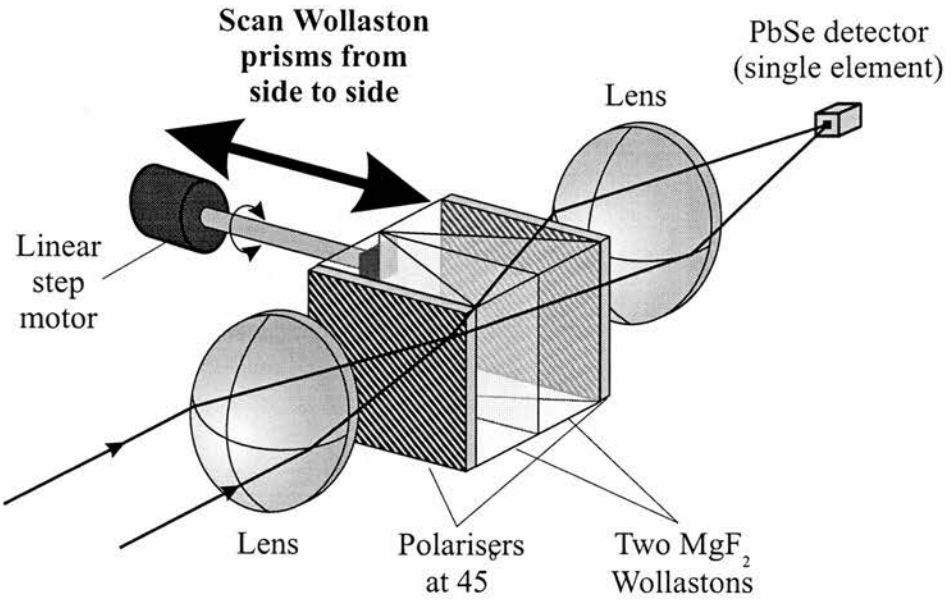

Figure 5.4: *The MgF_2 Wollaston prism scanning FT-MIR spectrometer.*

5.2.3 Mid-infrared results

Initial test spectra were recorded using white light from the visible region of the spectrum. This enabled easy alignment of the instrument by simply replacing the IR detector with a silicon based visible detector and running the tungsten lamp at its rated current.

Although some early results were obtained when operating in the visible region, results with the MIR source and detector were less encouraging. The low brightness of the source within the spectral transmission range of the polarisers resulted in weak fringes which were barely above the noise floor of the detector. This made it almost impossible to obtain useful spectra from the poor interferograms.

These results, combined with the fact that ultimately a single detector in conjunction with a scanning mechanism has little advantage compared to a dispersive instrument, led us to consider the viability of this solution. Consequently further development of the MIR scanning Wollaston prism instrument was abandoned in

favour of investigating alternatives, such as the near-infrared spectral region.

5.2.4 Near-infrared Wollaston prism based spectrometer

The poor performance of the scanning MIR spectrometer led us to consider other approaches. The availability of economical silicon based detector arrays enabled exploration of the near infrared spectral region.

Absorptions in the NIR result from overtones and combination bands which although considerably weaker and broader than the fundamentals, are well defined and potentially useful for quantifying petroleum. The availability of economic NIR detector arrays has brought an increase in the number of compact, grating based NIR spectrometers. Even the poor performance of these instruments relative to FT spectrometers provides useful results. Recording spectra using instruments with moderate resolutions of 4 nm has demonstrated accurate measurement of gasoline octane number.

5.2.5 Construction of a FT-NIR spectrometer based on Wollaston prisms

Component selection for a IR version of the instrument is determined by the required operating wavelength and resolution as discussed in section 5.2.1. Conventional CCD detectors typically have a long wavelength cut-off at approximately 1100 nm. For NIR operation we have therefore chosen to use a simple, 1024 element, 25 μm pitch photodiode detector array.

The obvious requirement of IR transparency for the Wollaston prisms gives a wide choice of birefringent materials including magnesium fluoride and synthetic

quartz. The longer wavelength compared with the visible region require Wollaston prisms with larger wedge angles to ensure that the maximum number of interference fringes are formed across the detector array. We have selected prisms manufactured from MgF_2 and quartz with respective wedge angles of 26^o and 12^o to give a maximum path difference across the array of $\pm 75\mu m$ and a corresponding resolution of 135 cm^{-1}. With a 1024 element array this gives a Nyquist wavelength of 290 nm.

More difficult than the choice of Wollaston prism is the selection of suitable polarisers. Above approximately 800 nm polaroid sheet is more difficult to obtain and performs significantly worse than in the visible. Glan-Taylor prism polarisers manufactured from birefringent material offer extinction coefficients in excess of 10^5, although resulting in a larger overall instrument. We have selected Glan-Taylor polarisers made from BBO that have a field of view of 6^o which is in excess of the Wollaston prism configuration and therefore not a limitation on the étendue of the spectrometer. Figure 5.5 shows the optical layout of the spectrometer.

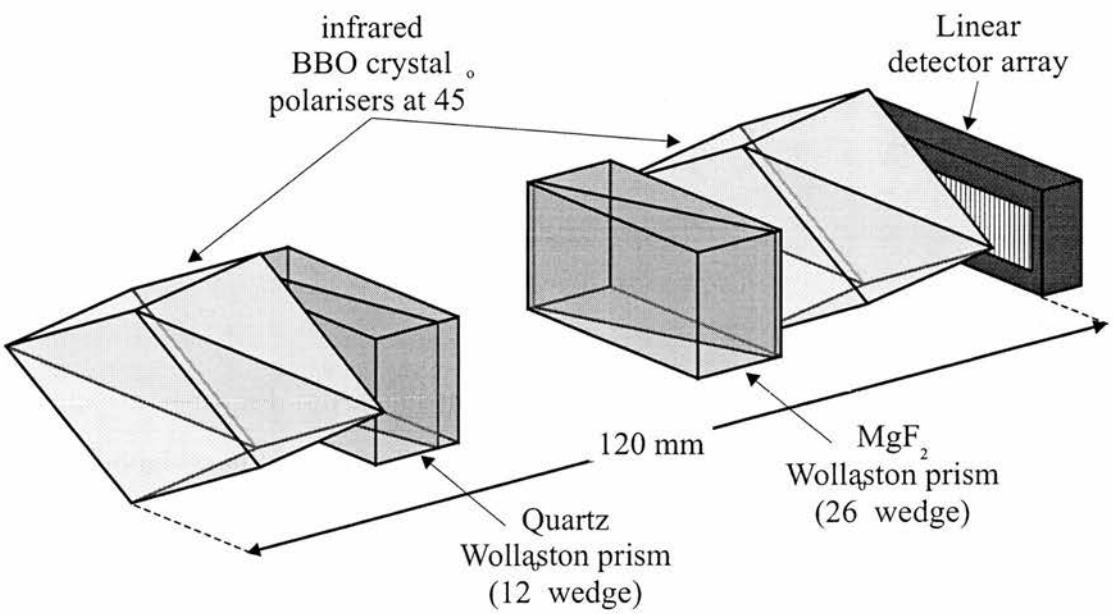

Figure 5.5: *NIR static Wollaston prism based FT spectrometer.*

In the near IR spectral region, an efficient source of broad-band light is a tungsten bulb run below its rated operating current. The liquid fuel is analysed in a 10 mm thick sample cell with fused silica windows. The complete NIR gasoline measuring system can be seen in figure 5.6

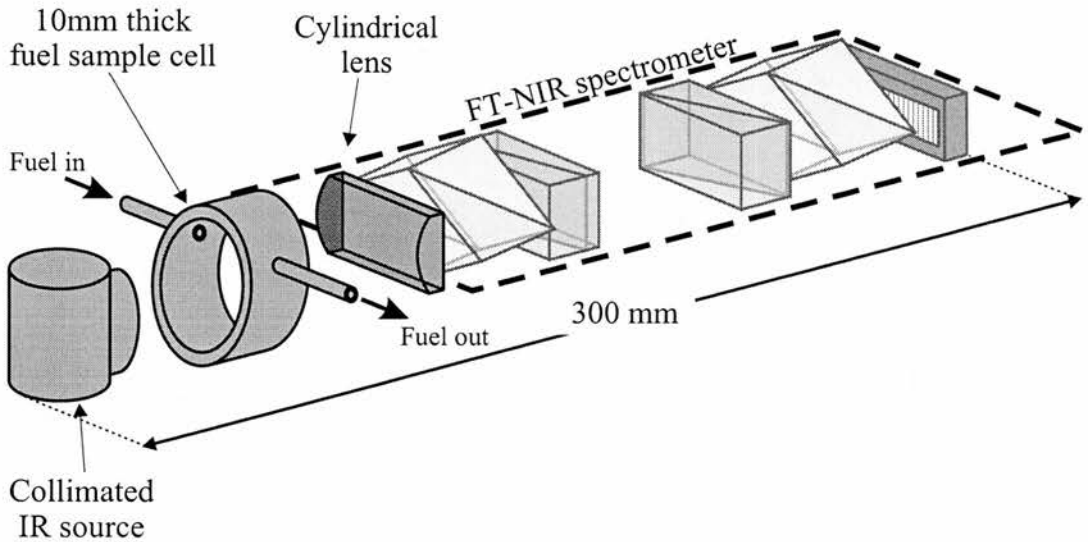

Figure 5.6: *System layout of the FT-NIR gasoline monitor.*

5.2.6 Near-infrared result for fuel octane number measurement

The fuel spectra of nine fuels corresponding to three different octane numbers, recorded using the FT-NIR spectrometer can be seen in figure 5.7. Even in its raw form these data show a clear distinction between the fuel octane rating. The difference can be further emphasised by multivariate statistical analysis of the data.

Principal component analysis is a multivariate statistical technique useful for identifying groupings within multidimensional data sets. In our case, each dimension of the data set corresponds to the absorption at a particular wavelength.

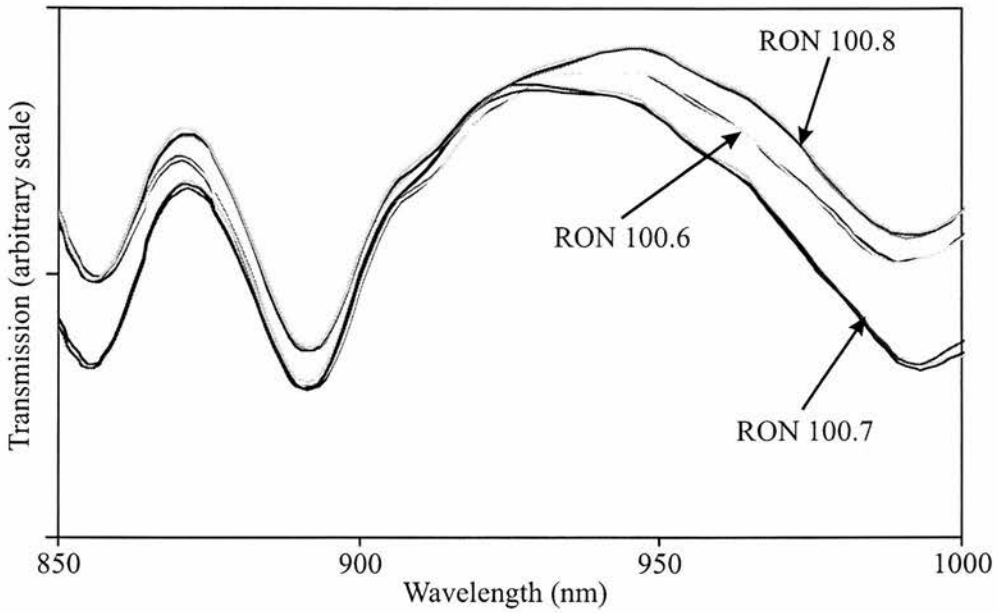

Figure 5.7: *NIR transmission spectra of 9 platformate fuel samples recorded using the Wollaston prism based FT-NIR spectrometer. The 9 spectra consist of 3 samples of* 100.6 *RON, 3 of* 100.7 *RON, and 3 of* 100.8 *RON.*

The principal components are the orthogonal directions in this multidimensional space which show the greatest sample to sample variation. The principal components correspond to the eigenvalues of the co-variance matrix of the data. Often it is found that the first few principal components account for most of the variation between the various samples and that one or two of the components can be used to place the sample in different grouping which have physical significance. A brief description of principal component analysis can be found in appendix B.

Principal component analysis is well suited to a computer based system where the spectral differences can be represented in the form of a few values rather than the complete spectral data. Such representation allows easy grouping of the various samples by respective properties of the fuels. A scatter plot in figure 5.8, using the first two principal components clearly reveals grouping corresponding to the octane number of the nine gasolines.

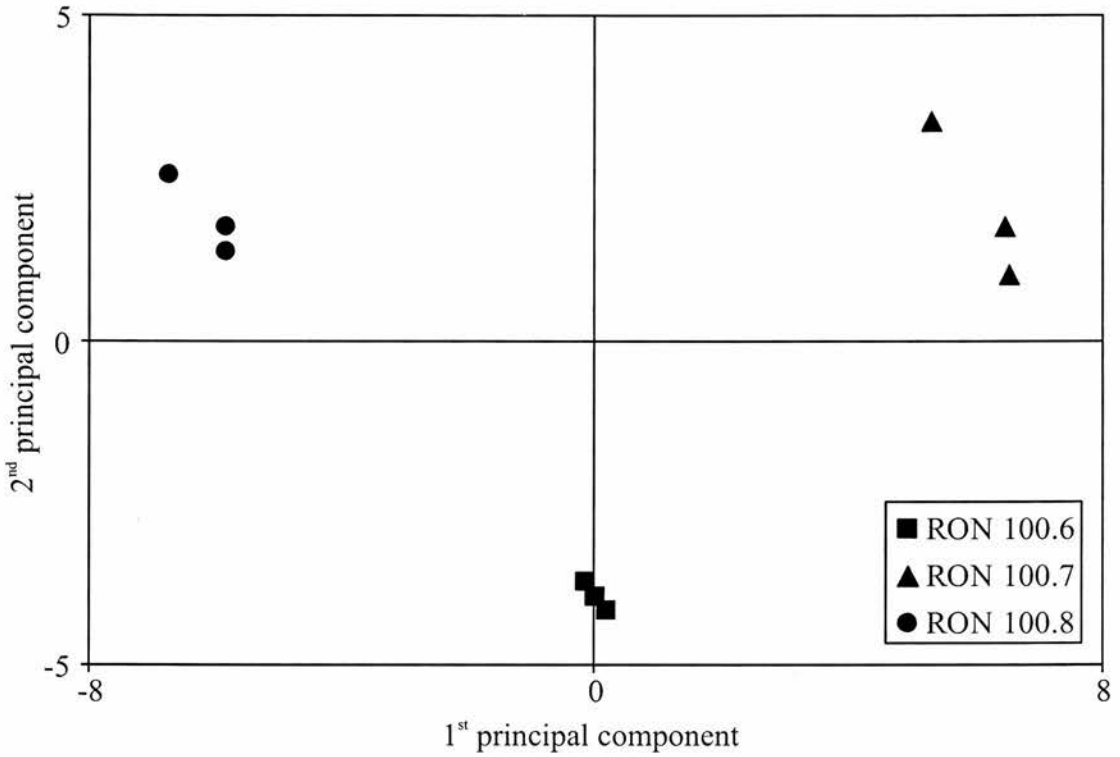

Figure 5.8: *Results of the principal component analysis on the NIR gasoline spectra in figure 5.7. The clusters correspond to different octane numbers.*

These results show clearly that octane number can be distinguished using a near infrared static FTS based on Wollaston prisms. This is not unexpected as NIR instruments based upon dispersive instruments or filters have been shown capable of similar results.

Possible advantages of our approach may be greater wavelength flexibility to cope with new fuel types introduced at a later date. The larger optical throughput over dispersive instruments may result in an instrument less sensitive to alignment, leading to greater reproducibility in spectral measurement.

5.3 Ultraviolet Fourier transform spectrometers based upon Wollaston prisms

Unlike conventional FT spectrometers, which because of mechanical vibration sensitivity are often restricted to use in the infrared, static Wollaston prism based spectrometers show no such limitation. Consequently, at St Andrews we have pioneered the development of a Wollaston prism based Ultraviolet Fourier Transform (FT-UV) spectrometer with spectral coverage extending down to 190 nm[48]. Initial work in the UV focused on the detection of pollutant gases.

In recent years the detection of toxic or flammable gases has been performed in the ultraviolet, predominately using optical open-path techniques[49]. Other spectral regions such as the visible and infrared have a limited suitability. The rotation-vibration bands of the ever present carbon dioxide and water vapour in open-path instruments gives very complex spectral features centered around the IR wavelengths of 2.6 μm, 4.2 μm, 6 μm and 15 μm. The presence of this "interference" in the spectrum from CO_2 and H_2O often makes it very difficult if not impossible to extract spectral information regarding other molecules.

In the visible region most gases and vapours tend to be transparent and the molecules tend not to conform to the common concept of a spectroscopic transition arising from the acquisition of energy by an electron in moving from one atomic orbital to another of higher energy. For molecules, the basic concept is identical, but instead of dealing with atomic orbitals it is often useful to think in terms of molecular orbitals, i.e. energy levels intrinsically associated with the molecule rather than with its component atoms. A detailed treatment of molecular orbital theory is beyond the scope of this thesis, but the following qualitative treatment may help in the understanding of the origin of molecular spectra in the UV region.

5.3.1 Ultraviolet transitions from molecules

For organic molecules, several types of molecular orbital may be distinguished [46]. In saturated compounds, such as methane, ethane, etc., carbon atoms are linked by "single bonds" or σ bonds. The only types of molecular orbitals here are σ orbitals at lowest energy levels and σ^* orbitals at higher energy levels. In the ground state, each σ orbital is usually occupied by a pair of electrons and the σ^* are not occupied. The occupied σ are known as *bonding orbitals* whilst the unoccupied σ^* orbitals are known as *antibonding orbitals.*

In compounds which contain double bonds, such as ethylene, in addition to σ orbitals, other orbitals known as π orbitals (bonding) are occupied by a pair of electrons whilst the higher energy π^* orbitals (antibonding) are not occupied in the ground state. A double bond may therefore be considered in terms of a pair of electrons in a σ orbital and another pair in a π orbital. A further type of orbital, known as an n orbital occurs in molecules containing an atom such as O or N which possesses a lone pair of electrons. Again, the n orbitals contain a pair of electrons, but in this case these electrons may not be involved in chemical bonding and n orbitals are therefore known as *non-bonding orbitals.* In general the energies of these types of orbital increase in the sequence $\sigma < \pi < n < \sigma^* < \pi^*$.

In systems containing *conjugated* or *delocalised* double bonds, the energy difference between the π and π^* orbitals can become quite small. From this simple sequence, therefore, it is possible to deduce that in saturated molecular groupings (e.g. the hydrocarbon chains of fatty acids) the only possible transition is $\sigma \rightarrow \sigma^*$ and this would be at high energy or low wavelength. This explains why such compounds are colourless and do not show any significant absorption above 200 nm. In unsaturated systems, however, and particularly in delocalised systems such as benzene or porphyrins, the $\pi \rightarrow \pi^*$ transition is of sufficiently small energy (long wavelength) to produce an absorption band in the near-UV region (benzene) or

even in the visible region (porphyrins). For experimental reasons it is very difficult to measure spectra below 190 nm and, in practice, this means that, for organic compounds, observable spectra arise only from unsaturated molecules.

5.3.2 FT-UV instrument construction

Ultraviolet component selection, like the infrared, is restricted by the difficulty in manufacturing UV optics and their resulting high cost.

The short wavelength cut-off of conventional CCD detectors is typically 350–400 nm. Silicon detectors have an inherent sensitivity down to 190 nm requiring only careful selection of the material for the protective 'glass' covering the sensitive detector elements. We have therefore used the same 1024 element silicon based detector array that forms the basis of the NIR spectrometer. Using a quartz window on the array enables access to the detectors full sensitivity range of 190–1100 nm.

The choice of UV transparent birefringent material for the Wollaston prisms is limited to magnesium fluoride, ADP, and synthetic quartz. The most robust and economical of these is quartz which covers the spectral range 180–3500 nm.

One of the major difficulties in UV component selection is the polarisers. Below approximately 250 nm it is not possible to obtain polarisers that operate effectively, polaroid sheet type is no longer available. Stacked Brewster plates offer a wide bandwidth but are expensive. An alternative is a single inclined plate manufactured from fused silica which is coated to give an extinction ratio of approximately 4:1 over the spectral range of interest. Unfortunately, the use of inclined plate polarisers limits how close the detector array can be located to the exit face of the second Wollaston prism, which extends the length of the

spectrometer.

The wedge angles for the first and second Wollaston prisms are 6° and 11.8° respectively and the prisms are separated by 60 mm. The resulting maximum path difference of ± 24 μm gives an effective resolution for the spectrometer of approximately 200 cm^{-1}. With a 1024 element array this gives a Nyquist wavelength of 190 nm.

In the UV spectral region, the most efficient source of light is a deuterium arc lamp. Although possessing a high ultraviolet output, the light source also emits strongly in the visible region of the spectrum. Given the higher sensitivity of the camera to light in the visible region, an ultraviolet band pass filter was positioned at the entrance aperture of the spectrometer to avoid saturation of the detector array.

Initially the FT-UV spectrometer was used for open-path pollutant gas detection. The compact and rugged nature making it ideal for inclusion and use in an industrial environment.

5.3.3 Open-path pollutant gas detection

The pollutant gases of interest were hydrogen sulphide, sulphur dioxide, nirogen oxide, ammonia and benzene[48, 50]. Benzene is a listed carcinogen and there is increasing concern over the levels of exposure in everyday life as well as in the work place. Currently, UK legislation and European directives permit concentrations of up to 5% of benzene in unleaded gasolines. There is a strong desire to reduce these concentrations and improve standards of exposure monitoring. Many of these gases absorb over the range of 160 nm–270 nm, however, the short wavelength cut-off for useful open-path atmospheric measurements is effectively

imposed by the oxygen absorptions at 190 nm. Therefore, the operating range of the spectrometer was ideally 190–270 nm.

The optical layout of the open-path gas detection system can be seen in figure 5.9. The UV spectrometer as described above can be identified in the bottom right corner of the diagram. The deuterium lamp source is collimated by a 90^o off-axis parabolic source mirror with a focal length of 60 mm. The collimated beam has a total power of 1 mW over the spectral range of 200-270 nm and is expanded to a diameter of 150 mm using a two-mirror afocal beam-expanding telescope. This gives a beam intensity that is both eye-safe and skin-safe.

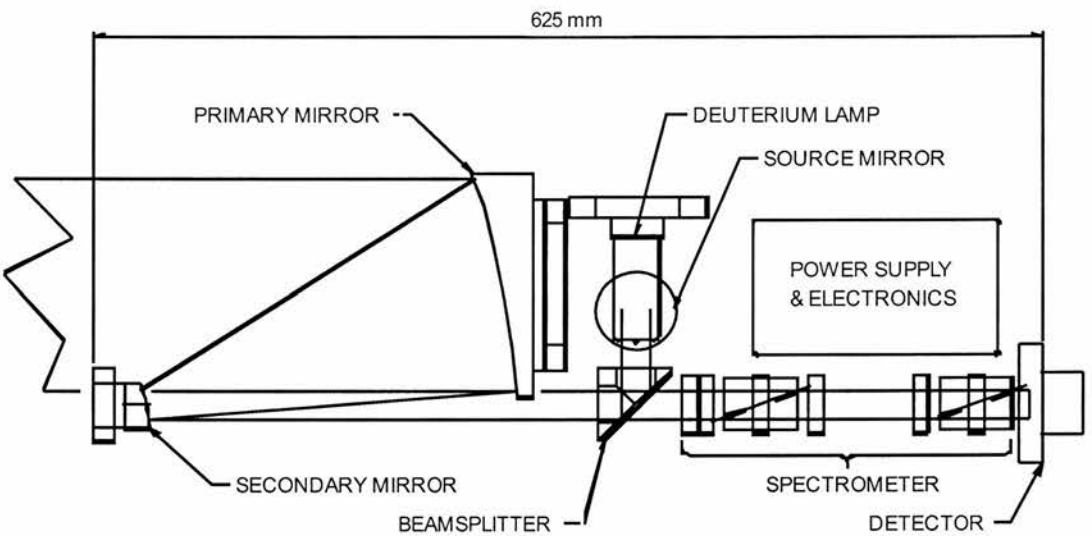

Figure 5.9: *A schematic diagram of the telescope optics for the static FTUV spectrometer.*

The instrument can be configured as a single ended or double ended system. In the single-ended system, the light source and spectrometer are mounted in the same unit and a 50/50 beam splitter allows both to be aligned coaxially with the transceiving telescope, as seen in figures 5.9 and 5.11. The expanded beam is transmitted across the region to be monitored and reflected by a 150 mm diameter array of retro-reflecting corner cubes. After reflection by the corner cubes, the beam returns along the same path and is collected by the transceiving

telescope which reduces the beam to a diameter of 20 mm for coupling into the spectrometer. To eliminate sunlight and other unwanted light a band-pass filter is located between the telescope and the spectrometer.

The double-ended system is based on the same telescope and housing but has the spectrometer and light source mounted in separate units. Removing the beam splitter and eliminating the losses associated with the retro-reflectors increases the received light level, allowing reliable operation at longer ranges or in adverse weather conditions. Both versions of the instrument have no moving parts, such that they are mechanically robust and have low maintenance requirements.

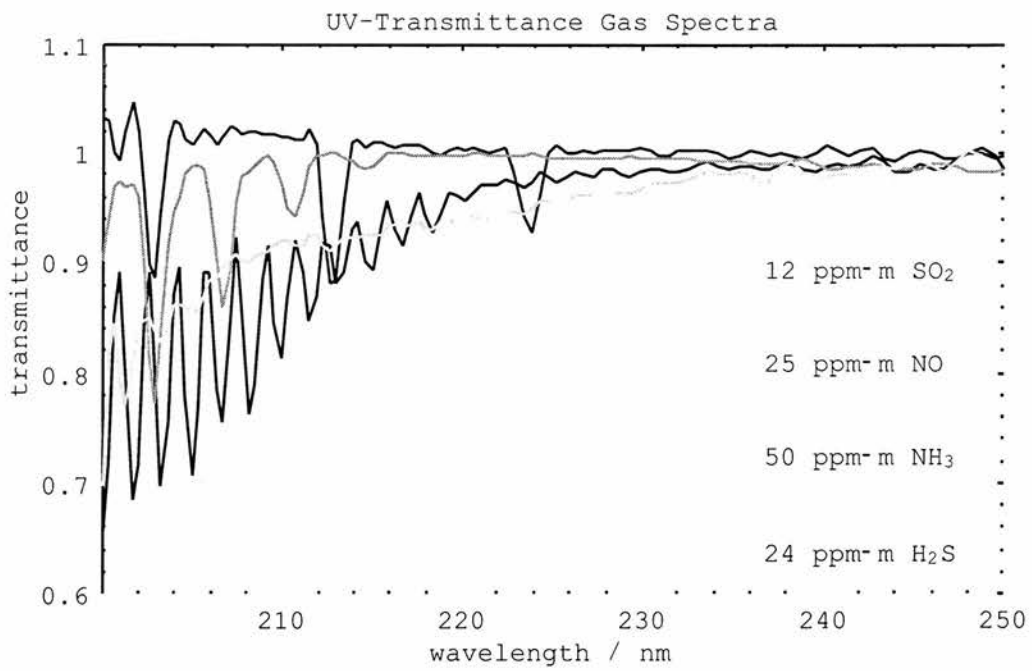

Figure 5.10: *Absorption spectra recorded using the open-path gas-detection system based on a static FT-UV spectrometer.*

5.3.4 Open-path gas detection results

The spectroscopic nature of the instrument makes possible the detection of a wide variety of gases. The gases detected and quantified so far by the open-path FTUV instrument include hydrogen sulphide, sulphur dioxide, ammonia and nitrogen oxide. The recorded spectra of several of the different gases are shown in figure 5.10. When used within a designated test area, the instrument easily detected a liquid benzene spillage of a few mls 10 metres up wind of the monitoring path. Table 5.1 details the minimum detectable concentration for each of the gases investigated to date.

Gas Released	Detection Sensitivity at a range of 100 m (ppm-m)
Hydrogen sulphide (H_2S)	4
Sulphur dioxide (SO_2)	2
Nitrogen oxide (NO)	4
Ammonia (NH_3)	6
Benzene (C_6H_6)	2

Table 5.1: *Minimum gas concentration detectable using the prototype instrument operating at a range of 100 m.*

As part of an optical, open-path, gas detection system it was shown that this type of instrument can satisfy accepted industry safety requirements for the detection of hazardous gases at eye-safe illumination levels. The high instrument étendue can be especially advantageous for the detection of low light levels and/or the realisation of rapid instrument response times. To the best of our knowledge no alternative optical system has been demonstrated that is capable of meeting these requirements. These results have encouraged commercialisation of this system, a photograph of the prototype can be seen in figure 5.11.

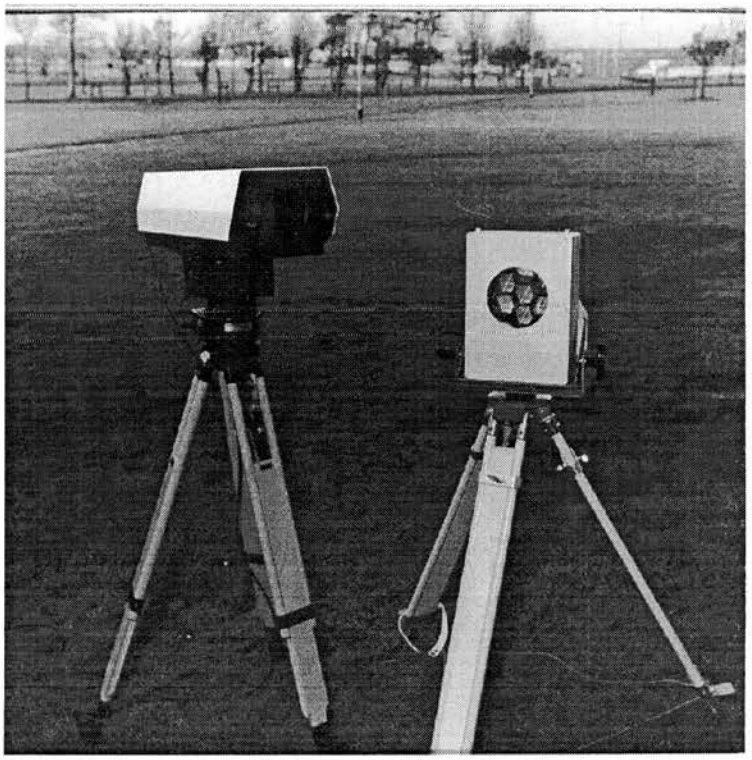

Figure 5.11: *Photograph of the prototype commercial static FTUV spectrometer and telescope (left) and the retroreflector (right).*

5.3.5 Gasoline measurements in the ultraviolet

During the gas detection project, concerns were encountered relating to the viability of optical detection in the UV spectral region. The chief concern is one of absorption of the light by oil mist either in the atmosphere itself or due to the build up of an oil film on any window or mirrors exposed to the environment. Oil film on any of the optical components in the beam path would significantly block much of the ultraviolet light[27], thereby preventing detection of certain gases. Although such a loss of light is a detectable failure mode in a safety monitoring system, it is believed that in oil production environments the requirement for regular cleaning would be excessive.

Despite the problem with the extremely high absorption coefficient of oil based products in the ultraviolet it is still possible to obtain useful UV spectra. Gasoline spectra can be recorded by increasing the power of the light source or using a very thin sample of fuel.

The use of a high optical efficiency (Jacquinot advantage) FT spectrometer also assists the collection of high quality spectra. In this application, the use of an FT spectrometer, removes the need for collimating and focusing optics which would be required to couple the light into the input slit of a dispersive spectrometer. Additionally any small-angle scattering of the light by the sample does not influence either the collection efficiency of the FT spectrometer or the recorded spectrum itself. Although the spectrometer used for gas detection had only a moderate optical efficiency it was possible to increase this further by specifying alternative components.

5.3.6 Ultraviolet spectrometer construction for gasoline measurement

Improving the optical efficiency of the FT-UV spectrometer was a relatively simple task. The chief modification compared with the earlier design is the replacement of the Brewster plate polarisers with Glan-Taylor polarisers fabricated from the synthetic crystal beta barium borate (BBO). The new polarisers give extinction ratios in excess of 10000:1 over a spectral range of 210–400 nm, producing an interferogram with higher fringe contrast than obtained previously, hence improving the signal-to-noise ratio within the processed spectrum.

The remaining components in the spectrometer were identical to those used in the pollutant gas detection system. The Wollaston prisms are fabricated from quartz with wedge angles of 6° and 11.8° respectively and a 60 mm prism separation.

The detector comprises a 1024 element array with a wavelength sensitivity from 190–1100 nm. The optical layout can be seen in figure 5.12.

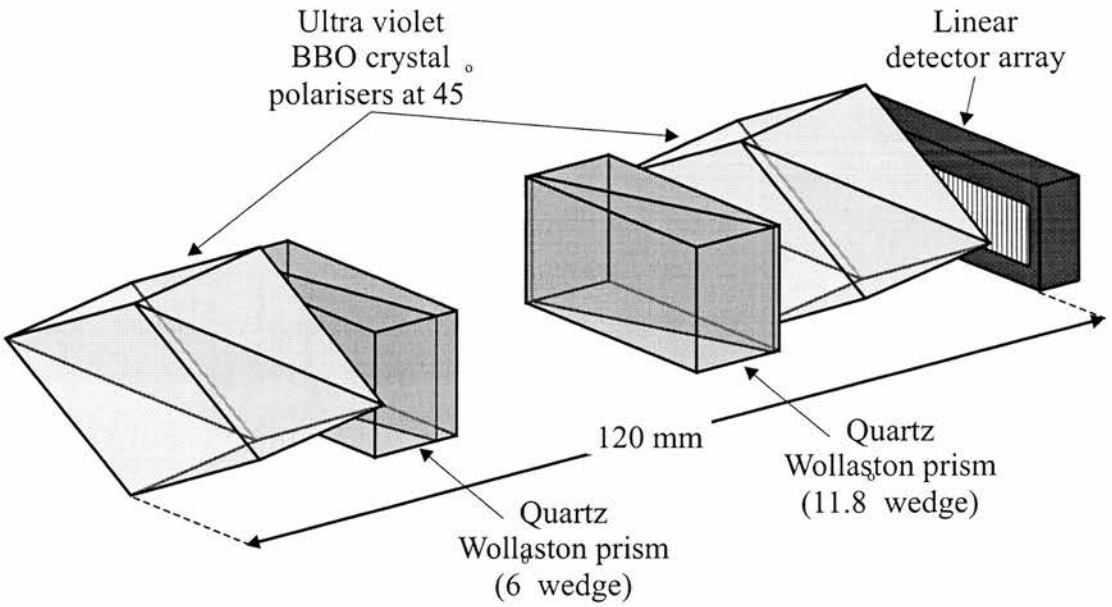

Figure 5.12: *Optical layout of the FTUV spectrometer.*

5.3.7 Gasoline measurement system in the ultraviolet

Optical path length in the gasoline sample was found to be critical to the determination of fuel properties because of the extremely high UV absorption coefficient of oil based liquids. Too extended a sample would effectively block the transmission of the ultraviolet light, very thin samples were difficult to achieve and resulted in large variations in the recorded spectra. After experimenting with several fuel sample cells we determined the optimum thickness of cell that would enable enough transmitted light to record a reproducible spectrum with sufficient detail. In this study we used two silica optical flats with a 5 μm thick aluminium coated Polyethylene terephthalate film for a spacer as the sample cell. Filling the cell simply required a single drop of fuel onto one of the flats which was subsequently sandwiched against the other. We find that this gave extremely

reproducible spectra and was not critical to the identification of the gasoline brand. Figure 5.13 shows a schematic of the deuterium light source, sample cell and spectrometer configuration.

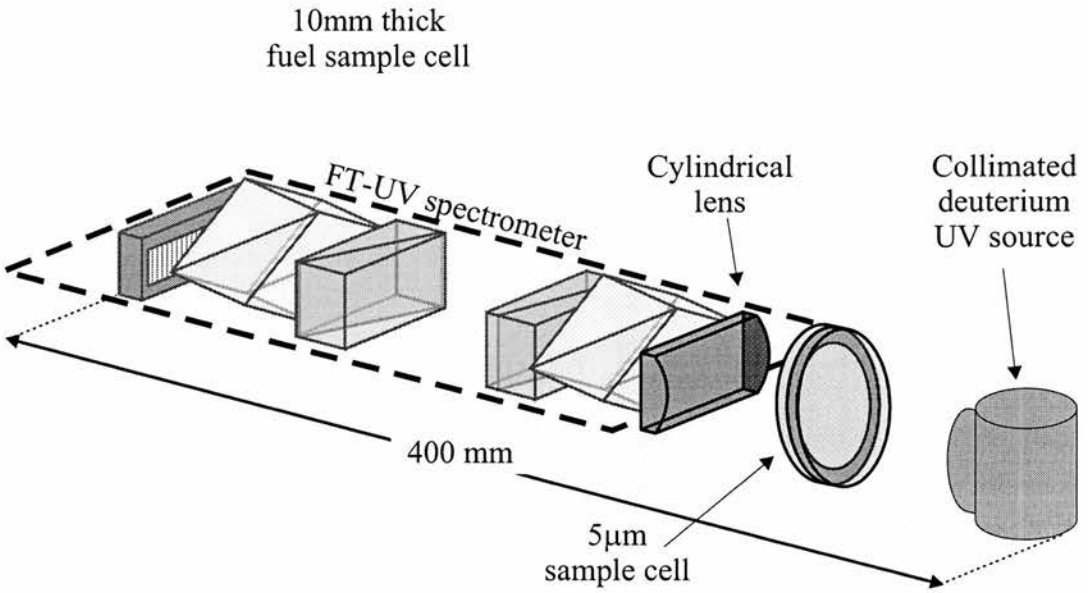

Figure 5.13: *System layout of the FTUV gasoline monitor.*

5.3.8 Results of gasoline measurement in the ultraviolet

A number of super unleaded gasoline spectra, obtained from Fourier-transforms of the corresponding interferograms, are shown in figure 5.14. Although some features distinguish the various fuel brands it is difficult to identify a characterisation criteria from the spectra alone.

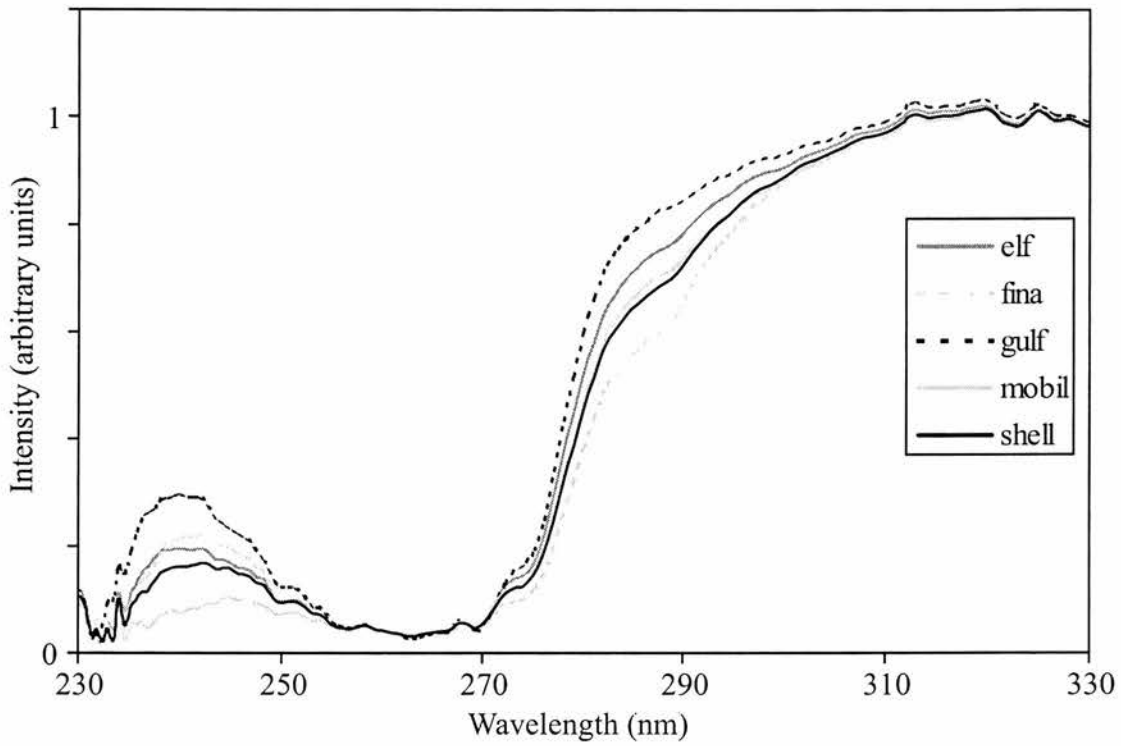

Figure 5.14: *Ultraviolet spectra of 5 super-unleaded gasolines recorded using the FTUV gasoline monitor.*

We have used principal component analysis (PCA) to identifying groupings within the multidimensional data sets. In our case, each dimension of the data set corresponds to the absorption at a particular wavelength in the gasoline spectra.

Often it is found that the first few principal components account for most of the variation between the various samples and that one or two of the components can be used to place the sample in different grouping which have physical significance.

This is the case with the gasoline spectra recorded in the ultra violet. Principal component analysis was performed on fifteen super unleaded spectra and a graph plotted using the first two principal components. The resulting graph can be seen in figure 5.15 and clearly demonstrates the different super unleaded gasoline brands have clustered together, enabling simple brand identification.

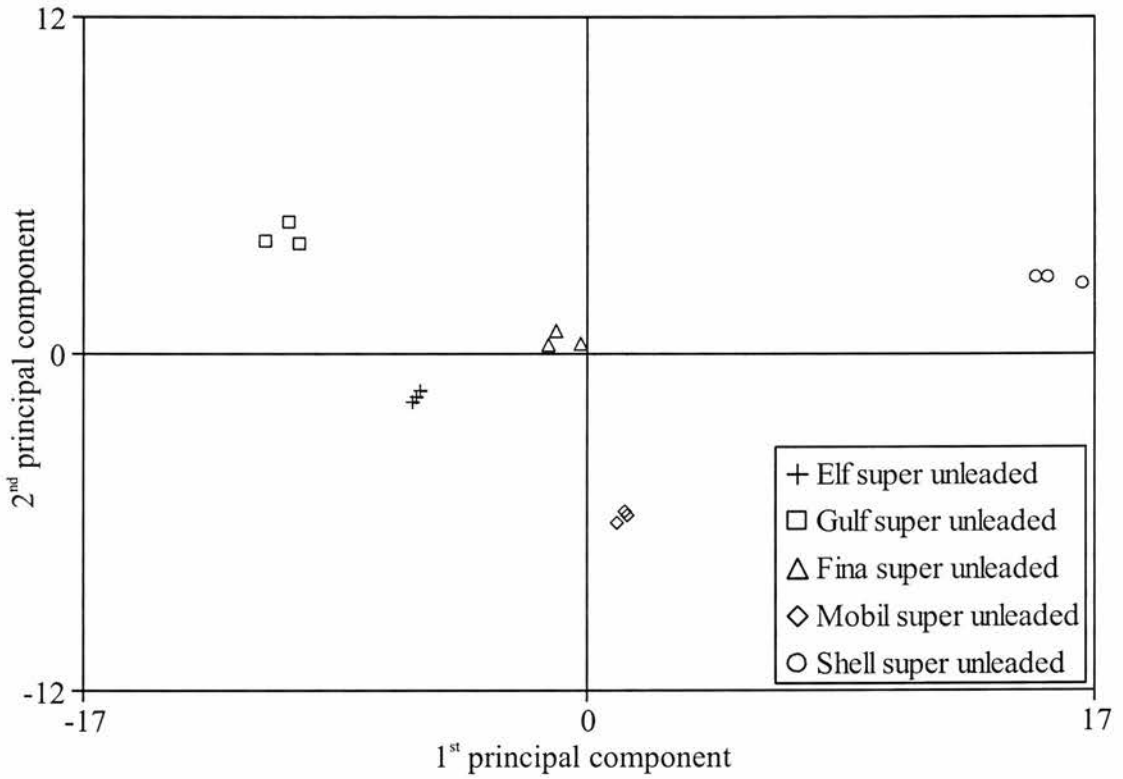

Figure 5.15: *Scatter plot using the first two principal components of five Super unleaded gasolines.*

Performing PCA on fifteen unleaded gasolines reveals a similar result. Figure 5.16 shows a scatter plot of the first two principal components, the five different fuel brands again cluster into easily identifiable groups.

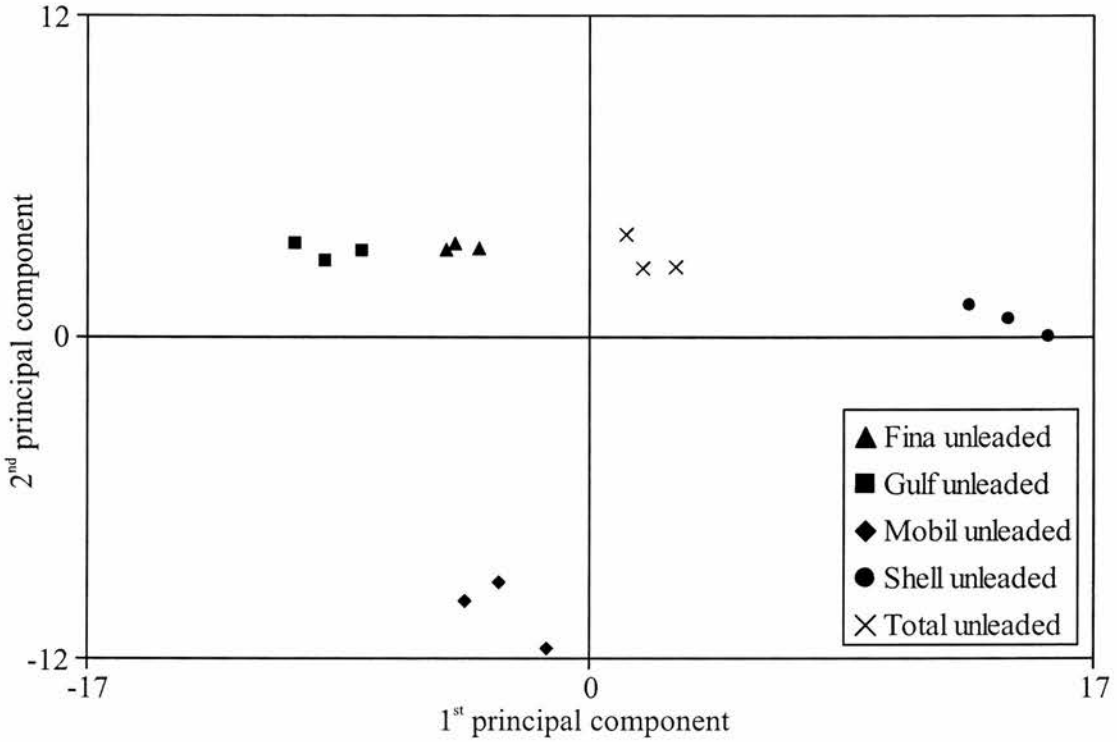

Figure 5.16: *Scatter plot using the first two principal components of five unleaded gasolines.*

Performing a principal component analysis on the combined super unleaded and regular unleaded gasoline spectra enables not only brand identification but fuel type classification. Figure 5.17 is a plot of the first two principal components, it can be seen that not only have the fuels clustered into their different brands but there is also grouping that enables the classification of the fuel type within each brand.

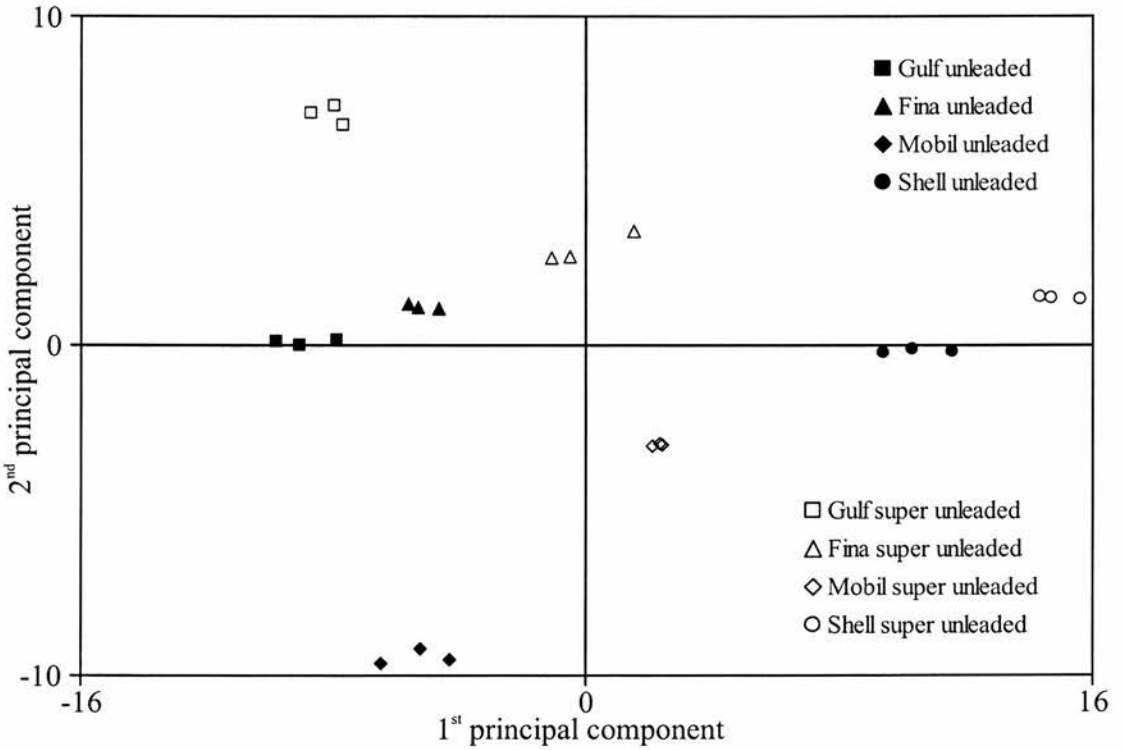

Figure 5.17: *Results of principal component analysis on both unleaded and super unleaded gasolines. The grouping within the Scatter plot reveals not only brand identification but the ability to distinguish between fuel types.*

In addition to brand identification, the octane number of gasolines could also be measured using the static FT-UV spectrometer. The UV spectra of several platformate samples of differing octane number can be seen in figure 5.18.

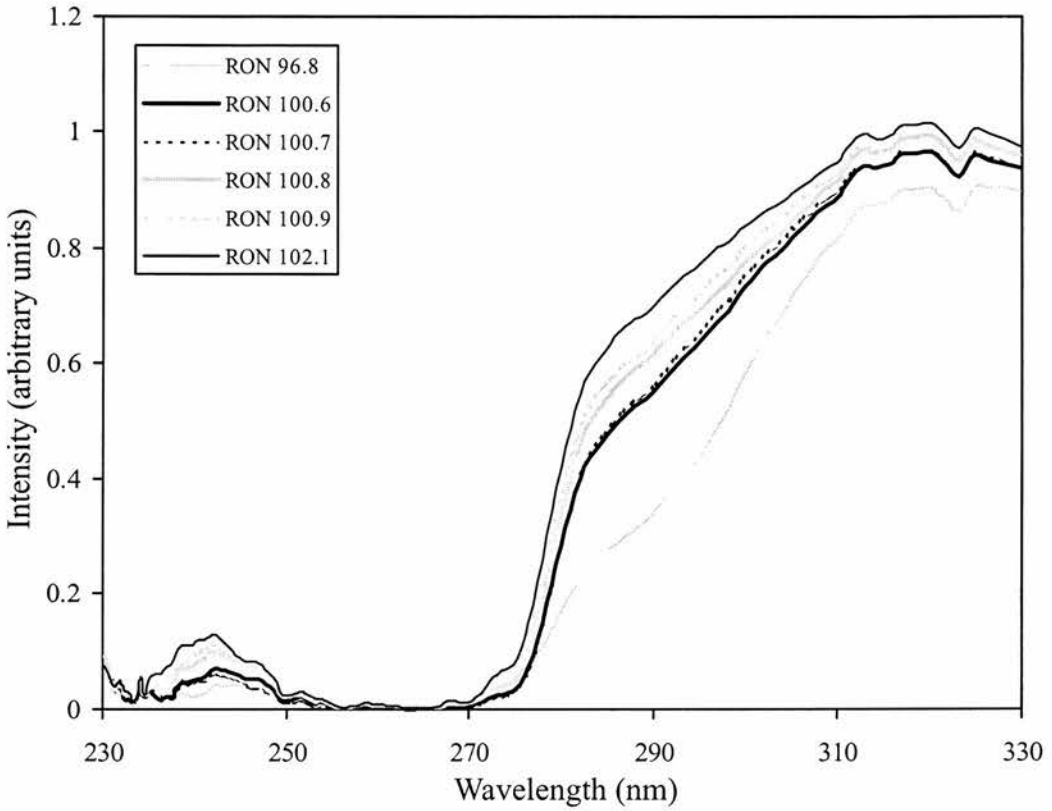

Figure 5.18: *Ultra violet spectra of Platformate fuel samples.*

The results of the principal component analysis on the platformate spectra can be seen in figure 5.19. The clearly visible grouping corresponds to the different octane numbers of the fuel samples. The UV method can distinguish fuels with an octane number difference of 0.1 giving a similar performance to the NIR system.

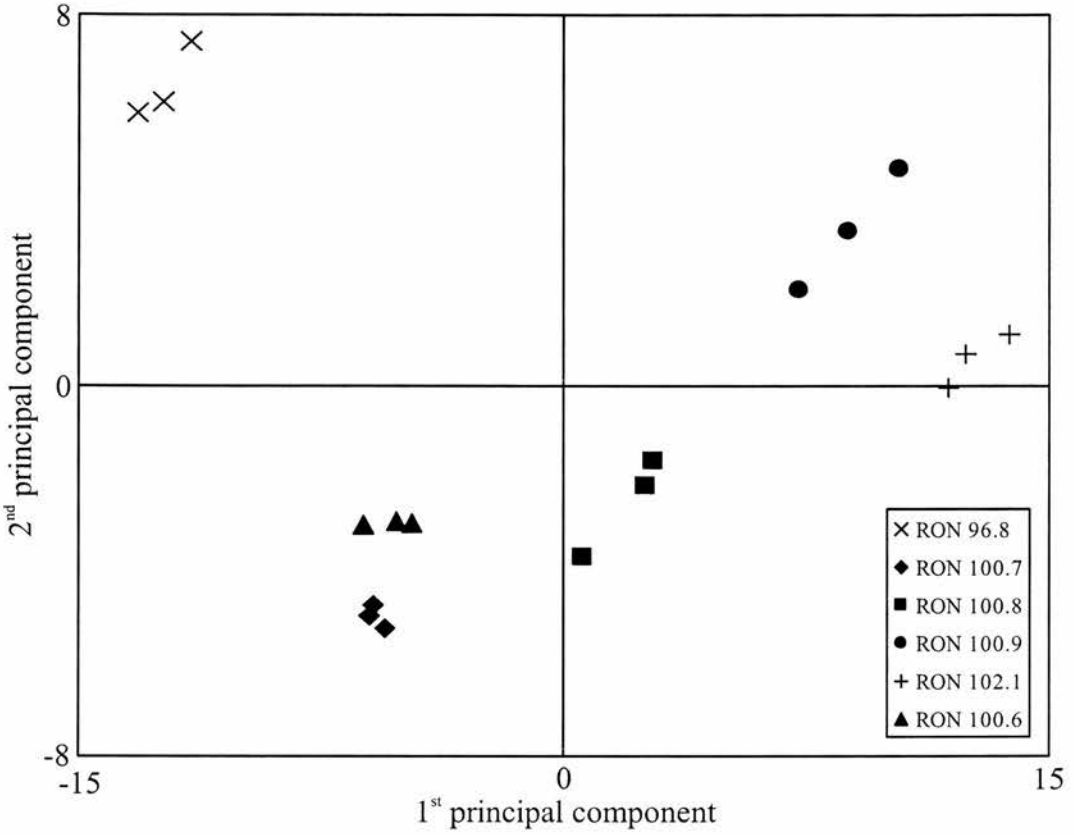

Figure 5.19: *Scatter plot using the first two principal components of six platformate fuels. The groupings enable identification of fuels with an octane difference of* 0.1

5.4 Conclusions

Although spectrometers operating in the mid-infrared spectral region are powerful instruments and have many applications the construction of a static FT spectrometer based upon Wollaston prisms has not been successful. The cost of MIR detector arrays is prohibitive with current technology, however, as materials science improves we may yet see a Wollaston prism based spectrometer operating in the MIR.

The success of the near-infrared instrument in measuring octane number, although encouraging for spectrometers based upon Wollaston prisms in general,

is not entirely unexpected as similar results have been achieved with dispersive instruments. However, our approach may offer advantages such as greater wavelength flexibility, larger optical throughput, and less sensitivity to alignment problems giving more reliable spectra recording.

Unfortunately, attempts at recording diesel spectra was unsuccessful because of the higher density of the fuel which absorbs more strongly in the ultra violet. Even with a 3 μm sample cell there was little or no light transmission to the detector to record meaningful spectra.

The measurement of gasoline properties in the ultraviolet spectral region is promising but it must be stressed that these are preliminary results because of the limited data set of fuels. A more exhaustive data set must be used before firm conclusions can be drawn. However, comprehensive fuel sets are expensive and further work would require significant support from Shell.

Chapter 6

Laser wavelength meter based upon Wollaston prisms

6.1 Introduction

Measurement of the absolute wavelengths or frequencies of light has long been an important area for many scientists. However, with the development and widespread utilisation of tunable lasers in areas such as spectroscopy has come the need for instruments that can conveniently and accurately determine their wavelengths. The power and potential of tunable laser sources cannot be fully exploited unless the laser wavelength is accurately known. Unfortunately, conventional methods for measuring optical wavelengths tend to be either inadequate in their resolution and accuracy or else cumbersome and slow.

The highest accuracy measurements of laser frequency have been achieved using direct frequency measurements. Elaborate frequency synthesis chains have been designed to compare selected laser frequencies to an atomic clock which produce results with uncertainty of only a few parts in 10^{10}[51]. However, the complex-

ity of such frequency chains means that they do not lend themselves to routine laboratory use. Rather than measure the laser frequency directly it is possible to determine frequency from a measurement of the laser wavelength. These instruments, termed laser wavelength meters, or often called wavemeters, are required for reasons of practicality to have accuracy, speed, portability and ease of use for non-technical users.

Despite lasers enabling very accurate direct frequency measurements with complex multistage instrumentation, nearly all actual measurements are still derived from wavelength standards. For measurement of wavelength a primary standard of wavelength is needed, together with a number of secondary standards distributed through the spectrum. A set of tertiary standards, readily reproducible in any laboratory, can then be generated from the secondary standards, and for virtually all purposes these are adequate for the wavelength calibration required by most instruments. Some knowledge of these standards is useful, even if one is primarily concerned with measuring line shapes and widths rather than absolute wavelengths. The saga of the primary wavelength standard is therefore of sufficient interest to warrant a brief account[4].

Until 1983 the primary standard of wavelength was an orange line of krypton emitted under certain specified conditions, and most of the existing secondary standards have been measured interferometrically against this standard in different laboratories with accuracies of order 1 part in 10^7. For the most part they are lines of iron, thorium and the inert gases, which are updated and added to every few years by a commission of the International Astronomical Union. A much larger number of tertiary standards has been derived from these, often with similar interferometric accuracy[52].

The history of the primary wavelength standard is inseparably tied up with the history of the meter. Until 1960 the ultimate standard of length was the standard

meter, the distance between two scratches on a platinum-iridium bar kept in Paris. There were obvious drawbacks to such a standard. The first was geographical and became evident as early as 1884, when it was shown that the great chart of absolute wavelengths compiled by Ångström some 16 years earlier, based on the Uppsala meter, needed correcting by 13 parts in 10^5, the discrepancy between the two meter bars. The second disadvantage became apparent with the development of high resolution spectroscopy at the end of the 19^{th} century: the accuracy of wavelength measurement was actually limited by the finite width of the scratches on the bar. To remedy this it was decided to adopt as an international standard of *wavelength* the red line of cadmium, emitted from a source in certain specified conditions. Once this had been determined as accurately as possible in terms of the Paris meter, all other wavelengths could be measured relative to it. However, the standardisation of the red cadmium line revealed the third defect of the Paris meter: measurements over a period of years showed the bar to be shrinking slowly as a result of metal fatigue. Similar measurements on the Imperial Yard in London revealed an even worse shrinkage. All this led to the decision to adopt a wavelength as the ultimate standard of *length.* The red cadmium line was no longer the best choice because of line broadening arising from unresolved isotope shift. After some years of controversy over the relative merits of the green line of mercury and the orange line of krypton, the latter was chosen in 1960, and the international standard of length was defined to 4 parts in 10^9 by taking 1 m to be 1650763.73 vacuum wavelengths of this line ($\lambda_{vac} = 605.780210\ nm$) emitted by the isotope ^{86}Kr under specified conditions.

Incidentally, the mercury green line which was the runner up in the contest was one of the lines investigated by Michelson as a possible standard in the 1890s, using visibility measurements with his interferometer. He rejected it in favour of the red cadmium line because of what we now recognise as isotope structure.

The 1960 decision is not the end of the saga. Since 1967 the fundamental standard

of time, or frequency, has also been an atomic standard, defined by a transition between the hyperfine structure levels of the ground state of ^{133}Cs. The second is defined as 9192631770 periods of this transition frequency, and the measurement can be made to 1 part in 10^{13}, which corresponds to 1 second in 300000 years. The measurement and comparison of frequencies is the most accurate of all types of measurements, and with the development of lasers it has become possible to transfer this frequency standard down a multistage chain of oscillators and lasers into the near infra-red[51]. With frequencies in the visible known to about 1 part in 10^{10} and the ability to compare laser wavelengths interferometrically to a few parts in 10^{10}, the velocity of light can be measured to an accuracy limited only by the krypton generated meter. Since only two out of the three quantities λ, ν and c can be separately defined, the limitation on wavelength measurements imposed by the krypton standard can be removed by abolishing the meter altogether and fixing the value of c. This happened in 1983 when c was defined as 299792458 ms^{-1}, and the meter has become the distance traveled by light in a fraction, $1/c$, of an 'atomic second'. The wavelength of any line is the reciprocal of its frequency, multiplied by the constant c. Any line that is a frequency standard automatically becomes also a wavelength standard, capable of comparison with another wavelength to a few parts in 10^{10}, an improvement of an order of magnitude over the krypton standard.

6.2 Common interferometric wavelength meters

By far the greatest and most diverse research efforts in wavemeters have been directed at developing instruments employing interferometry. These include wavemeters based on Michelson, Fabry-Perot and Fizeau wedge interferometers, either as dynamic or static instruments.

One of the difficulties encountered with interferometric wavemeters is sensitivity to optical alignment and to wavefront quality. Most wavemeters utilise resolution-enhancing techniques in order to achieve the maximum measurement precision with a limited number of fringes. In effect, these methods measure distances to small fractions of an optical fringe. However, for the measurements to be accurate, the optical alignment and the fringe quality must be very good. In particular this usually implies that both the alignment and the wavefront flatness must be very reproducible under all conditions of use and all wavelengths of the unknown laser.

6.2.1 Michelson wavemeters

Basically, a Michelson wavemeter is a Michelson interferometer (or similar) in which the path difference of the two arms can be changed. The interferometer is illuminated simultaneously by the laser whose wavelength is to be measured, and by a separate reference laser of known wavelength, as shown in figure 6.1. The changing path difference causes the fringe intensity for each laser beam at the detector to oscillate. The ratio of the fringe frequencies of the two lasers is approximately equal to the inverse ratio of their wavelengths. The unknown wavelength can therefore be found by multiplying the known wavelength by the easily measured ratio of the fringe frequencies.

The precision of Michelson interferometers when used as laser wavemeters is set by the number of fringes within the interferogram. Counting 10^6 fringes will give approximately 1 part in 10^6 precision. However, many of these instruments utilise a variety of algorithms and designs to increase the precision up to approximately 10^8 while still maintaining realistic scan lengths and times.

In view of the obvious similarities of a dynamic Michelson wavemeter and a

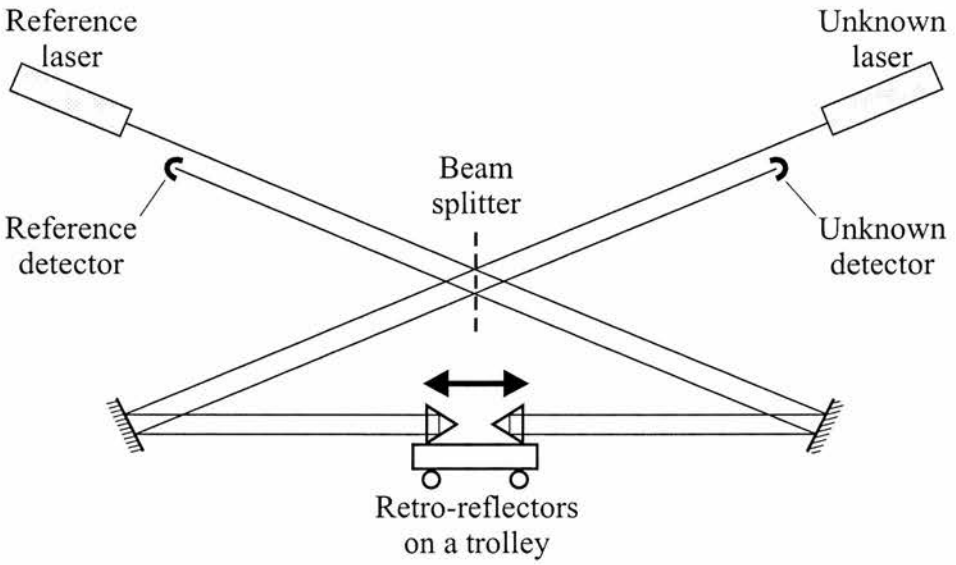

Figure 6.1: *Layout of a dynamic Michelson wavemeter, changing the path length in both arms simultaneously with back to back corner cubes.*

Fourier transform spectrometer it is not surprising that a Fourier transform wavemeter has been reported[53]. The approach offers the advantage of operation with multimode lasers, which are often problematic for many wavemeters. The assumption that a multimode spectrum consists of a few isolated laser lines enables a reduction in the large number of data points required for precision measurements, i.e. the recorded data is under sampled. The reduction of data adds some additional software complexity to alleviate the problems of aliasing and degrades the signal-to-noise ratio of the spectrum, but the signal-to-noise ratio can be tolerated because of the relatively high intensities of laser sources.

Much like their FTS counterparts these Michelson interferometer based instruments usually require high precision scanning mechanisms and very stable designs, which inevitably implies high cost and substantial bulk. This has led to considerable research into designs of instruments having no moving parts, where an interferogram formed in the spatial domain is recorded with a detector array.

These static instruments have the advantage of use with both cw and pulsed lasers.

6.2.2 Fabry-Perot wavemeters

The principal difference between Michelson and Fabry-Perot interferometers is that the former is a two-beam interferometer whereas the latter is a multibeam interferometer. By exploiting the interference of many waves, which have made different numbers of round trips, a Fabry-Perot interferometer can provide much sharper and narrower interference fringes relative to the fringe spacing.

Wavemeters can be constructed from scanning Fabry-Perot interferometers, where the time interval between fringes from a reference laser is compared with the time interval between fringes from the unknown laser when both illuminate a linearly scanning spherical Fabry-Perot interferometer.

Static Fabry-Perot wavemeters that operate with cw or pulsed laser sources have also been constructed. Often they consist of multiple F-P interferometers with free spectral ranges varying by decades from 10 cm^{-1} to 0.01 cm^{-1}. The fringe pattern from each interferometer is recorded by a photodiode array before analysis by computer. A simpler device based upon only a single set of solid fused silica etalon and a 1024-channel diode array, as shown in figure 6.2, was constructed by Koo [54].

The etalon generates 6 fringes (3 fringe pairs) which are recorded by the detector array. The wavelength of the laser light is calculated from equation 6.1 using the 3 diameters of fringe pairs[55].

$$D_p^2 = \frac{4nf^2\lambda}{{n_0}^2 d}(p+\varepsilon) \qquad (p = 0, 1, 2, \ldots) \tag{6.1}$$

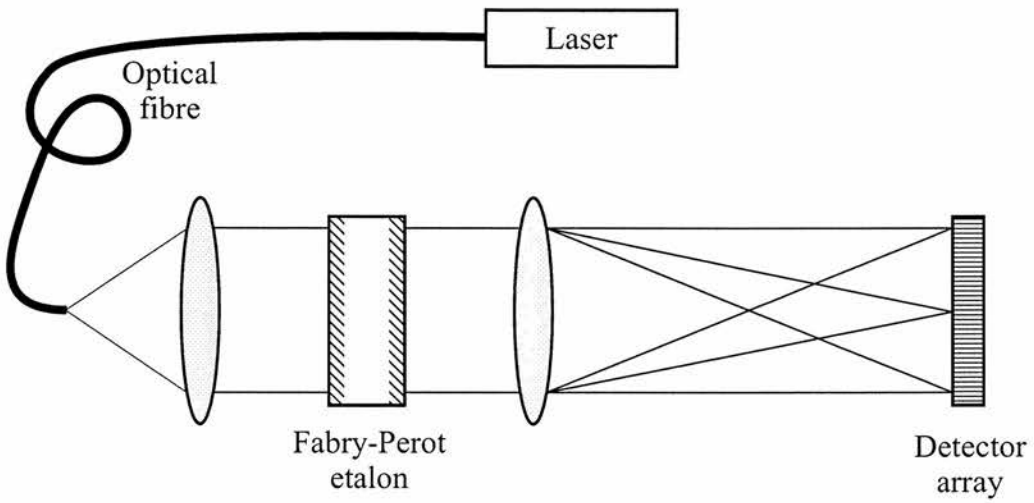

Figure 6.2: *The optical layout of the simple Fabry-Perot etalon based wavemeter.*

where d and n are the thickness and refractive index of the etalon, and f is the focal length of the lens used to image the fringes onto the detector array. The diameter of the pth ring (fringe pair) is given by D_p.

6.2.3 Fizeau wedge wavemeters

Static wavemeters based on the Fizeau interferometer have been developed [56, 57, 58] for use by either cw or pulsed laser sources. The typical components required for a Fizeau wedge interferometer are shown in figure 6.3. The wavelength is determined from a measurement of the CCD recorded fringe pattern formed by reflection of an expanded laser beam from a wedge consisting of two separated uncoated glass plates with a small degree of tilt between them.

A Fizeau interferometer is formed from a pair of uncoated optical flats slightly inclined with respect to one another. Incoming light is focused through a filtering pinhole, and is collimated by either a lens or a parabolic mirror. A portion of the

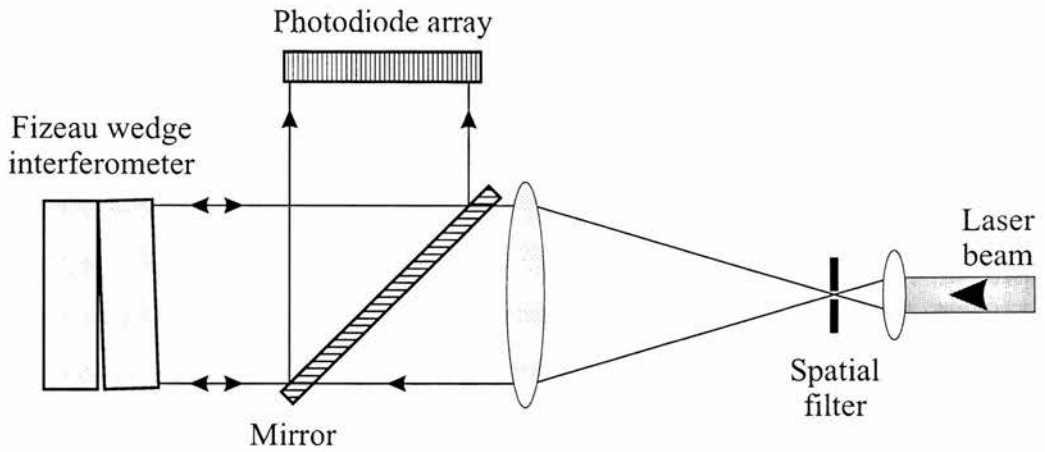

Figure 6.3: *Optical layout of a Fizeau wedge wavemeter. A linear photodiode array monitors the fringe pattern obtained in reflection from an uncoated wedge interferometer.*

light reflects from the front interferometer surface, while a nearly equal portion reflects from the rear interferometer surface which imposes a path difference between the two reflected rays. The path difference varies linearly across the Fizeau wedge and the reflection of plane monochromatic light from these two surfaces produces a series of uniformly spaced parallel fringes whose intensity along an axis perpendicular to the fringes is given by equation 6.2

$$I(x) = \frac{1}{2}\left[1 + \cos\left(\frac{2\pi x}{\Lambda}\right)\right] \tag{6.2}$$

with $\Lambda = \lambda/(2\alpha)$, where x is the distance along the array, Λ is the fringe spacing, λ is the wavelength of the light, and α is the wedge angle between the two surfaces.

These fringes are recorded by a photodiode array connected to a computer which calculates the fringe period and hence the wavelength of the incident light. Measurement accuracy requires initial calibration and depends on the stability of the interferometer, since a reference laser is not used.

6.3 Wollaston prism based laser wavelength meters

Static Fourier-transform spectrometers (SFTS) based upon Wollaston prisms have been developed that have no moving parts and are consequently extremely robust and compact instruments[22, 23, 24, 25, 26]. The operating principles and applications of spectrometers based on Wollaston prisms have been discussed in detail in chapters 3 – 5. However, it is useful to remind the reader that when a Wollaston prism is illuminated between polarisers, interference fringes are produced parallel to the sides of the prism. Static instruments record these fringes using a detector array which sets a limit on the number of fringes within the interferogram. This is due to the Nyquist criterion which requires at least 2 pixels per fringe, combined with the fixed number of pixels on a detector array. This limits the resolution and wavelength precision of the spectrometer. In addition, the fast-Fourier-transform (FFT) method is problematic with so few interferometric data points as it yields information about the period of a sinusoid only indirectly and is known to suffer from systematic errors, or bias, which interfere with the real spectrum[2].

Despite this, large scanning instruments are capable of measuring laser wavelengths with an accuracy of ≈ 1 part in 10^8 using the FFT algorithm[59]. The principle advantage of such systems is their ability to measure the wavelength of several single-mode and multimode lasers simultaneously. Static instruments, although limited when using the FFT algorithm, can increase the precision beyond that dictated by the number of recorded fringes through the implementation of alternative methods.

The key principle that can be exploited is interference fringes from a single monochromatic laser source will form a sinusoidal pattern. Making direct mea-

surements upon these sinusoidal fringes enables higher precision laser wavelength measurement.

A laser wavemeter, proposed by Padgett[24] and based upon a Wollaston prism interferometer, utilised the sinusoidal fringe pattern to increase the accuracy of wavelength determination beyond that of a Fourier transform limited instrument. The wavelength was calculated by finding the maximum value in the power correlation function between the interference pattern and a perfect sinusoid. This correlation function is equivalent to the Fourier transform but can be calculated for sine waves with an arbitrarily small increment in period. Preliminary measurements demonstrated a short term reproducibility of approximately 2 parts in 10^5. The reproducibility is somewhat limited by the temperature dependence of the birefringence. However, it was noted that in conjunction with temperature stabilised Wollaston prisms or a reference wavelength, it would be possible to increase the long term reproducibility.

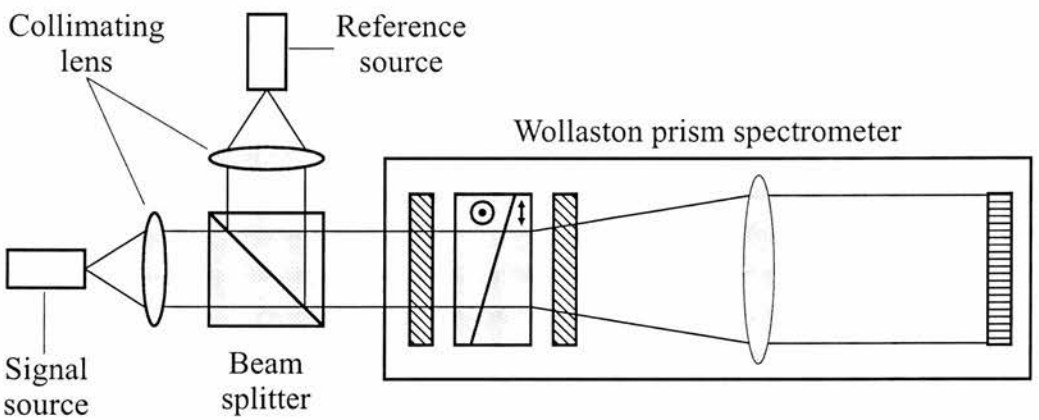

Figure 6.4: *Optical layout of the wavelength meter developed by Jiang[60]. The system incorporates a self-referencing scheme with a standard Wollaston prism interferometer.*

6.3.1 Wollaston prism wavelength meter utilising a reference laser

A method which utilises the SFTS based upon a single Wollaston prism and an imaging lens has recently been proposed by Jiang[60]. The essence of the system involves putting a known reference laser and unknown signal into a standard Wollaston prism spectrometer[23], as shown in figure 6.4. The spatial intensity distribution is a superposition of the two sets of interference fringes from the two wavelengths and thus fringe beating is observed. Using autocorrelation and Gaussian filtering techniques, a variation between the two wavelengths can be measured with a precision up to 0.01 nm. For this system to work the wavelengths of the two light sources needs to be within 15 nm of each other. If a larger measurement range of 110 nm is required, the accuracy drops to approximately 1 nm. Additionally, the interference pattern produced by two light sources can be distorted due to intensity differences and spectral characteristics of the optics.

6.4 Design of a dual purpose wavelength meter and spectrometer

For construction of an ultra-compact wavemeter[61] we based the design upon the birefringent interferometer utilising a single modified Wollaston prism where the optic axis was inclined with respect to the entrance aperture. The operating principles of the interferometer have been discussed previously in section 3.5.

The first point of design for many Wollaston prism based instruments is the detector array. The detector utilised in the proof of principle, modified Wollaston prism spectrometer enabled the instrument to function at moderate resolutions,

having only 312 × 287 pixels. Operation as a high performance laser wavelength meter would require a superior detector array. Increasing the number of pixels across the detector would extend the number of fringes that could be recorded, thus leading to enhanced measurement accuracy. We use a highly-integrated CCD monochrome video camera[62], which has a 752 × 582 array sensor with a total active area of 7.95 × 6.45 *mm*. It also includes the necessary drive circuits to deliver a formatted video signal.

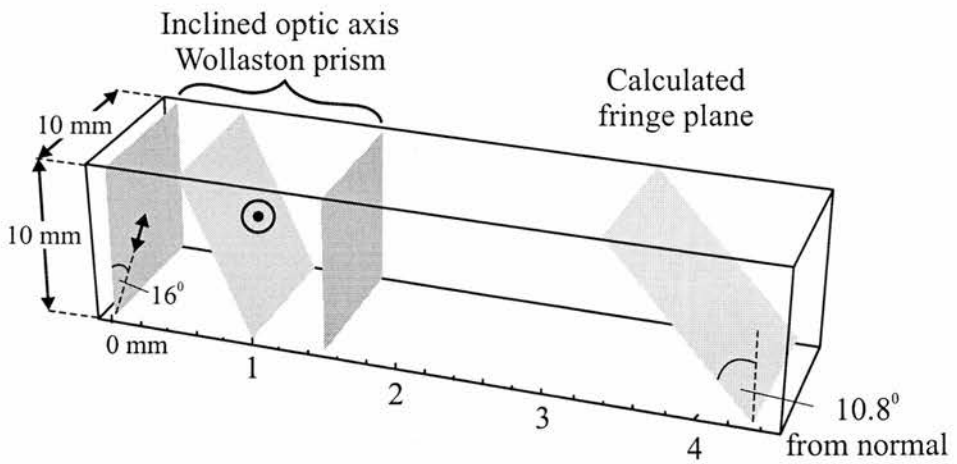

Figure 6.5: *The modified Wollaston prism showing the plane to which the fringes from an extended source are localised.*

The compact nature of the static instrument was maintained by matching an appropriately designed Wollaston prism and polarisers to the detector. In the optimised, computer-modeled[30] design both Wollaston prism and the polarisers are manufactured with an aperture of 10 *mm*. The Wollaston prism, manufactured from calcite[29], is 1.5 *mm* thick and has an internal wedge angle of 2.8°. The optic axis in the first wedge is inclined at an angle of 16.0° to the prism entrance face (18.8° to the wedge interface). By adhering the polarisers to the faces of the prism with their polarisation axes aligned at 45° to the edges gives a single optical component instrument approximately 3 *mm* thick. The fringe

pattern, seen in figure 6.5, forms about 2.0 mm behind the exit face of the prism, coincident with the position of the detector array. This optical arrangement gives a maximum path difference of approximately 50 μm across the aperture of the detector. The Nyquist criterion means that for a 752 element array the shortest measurable wavelength is 320 nm.

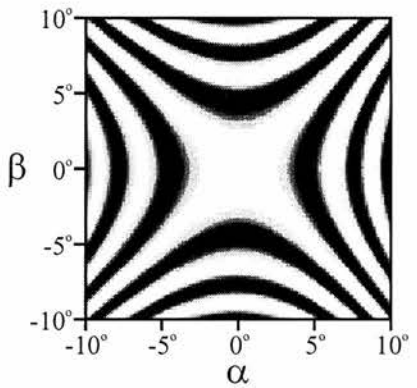

Figure 6.6: *The predicted field of view for the ultra-compact spectrometer and laser wavelength meter.*

One of the main advantages of a Fourier-transform spectrometer over a dispersive instrument is its higher optical throughput or étendue. This can be as much as 190× greater for equivalent resolutions[7]. As discussed in section 2.6.3, étendue is proportional to the field of view of the spectrometer. The field of view for this optical configuration was predicted to be approximately $\pm 5^o$ as shown in figure 6.6.

For laser wavemeter applications optical throughput is usually not an issue. When configured as a wavemeter we use a large 1 mm diameter core fibre to couple light into the instrument. This not only gives a convenient, easily aligned instrument but the fibre coupling ensures uniform intensity across the interferogram, which in turn gives greater measurement stability. The optical layout of the wavemeter / spectrometer is illustrated in figure 6.7.

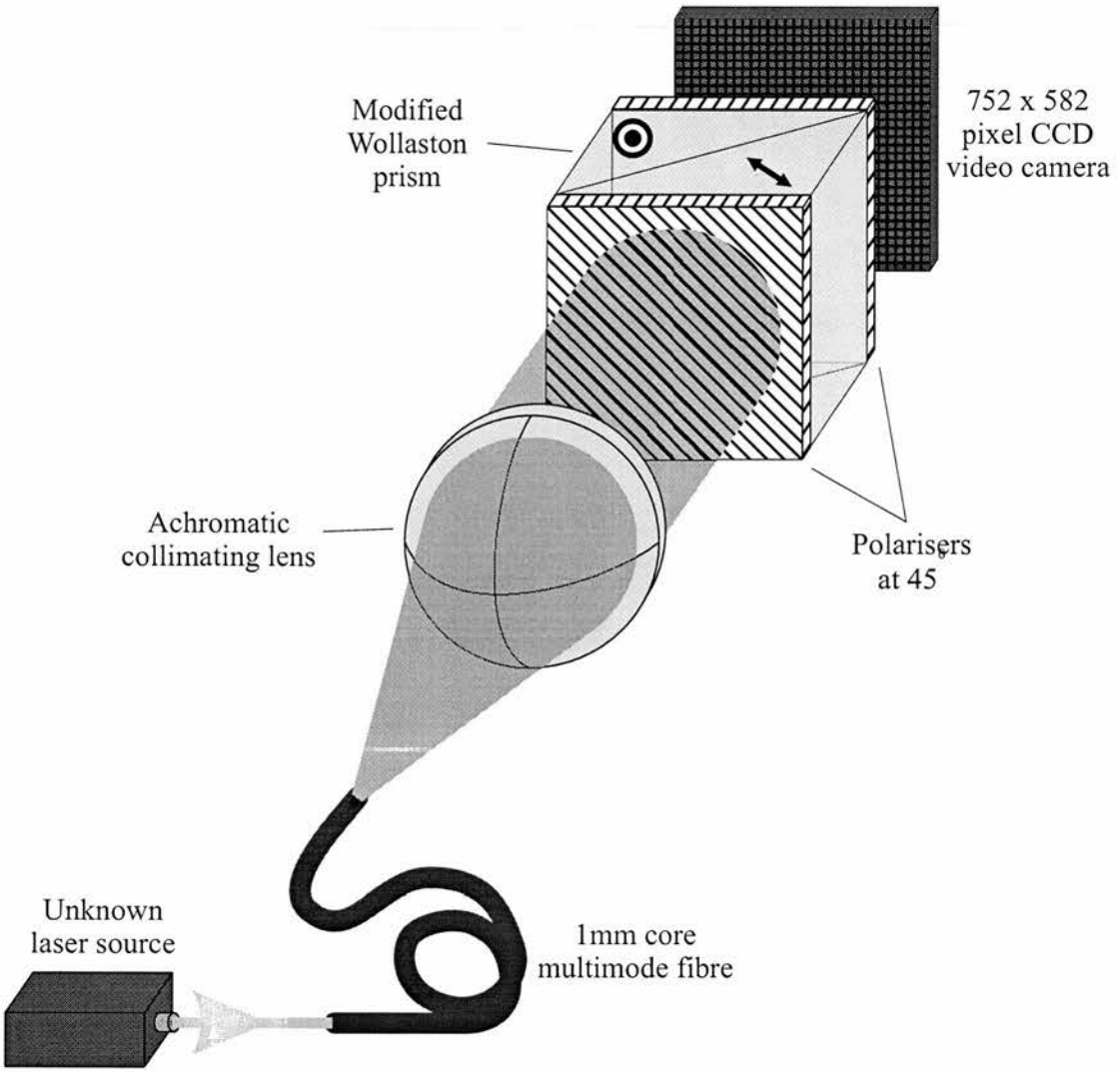

Figure 6.7: *The optical layout of the wavemeter / spectrometer.*

The camera outputs data to a desktop or laptop computer via a frame grabber card and appropriate software. The raw data from the camera is a two-dimensional image of the interference fringes which is averaged down the frame to give a single 752 element data array. Averaging the data over the 582 lines of the array reduces the effect of fixed pattern noise on the resulting interferogram. For ease of use the control and data-processing software was written for a Microsoft Windows environment. Through simple mouse button clicks the user can

control the wavemeter or spectrometer. A photograph of the complete instrument interfaced to a laptop computer is shown in figure 6.8.

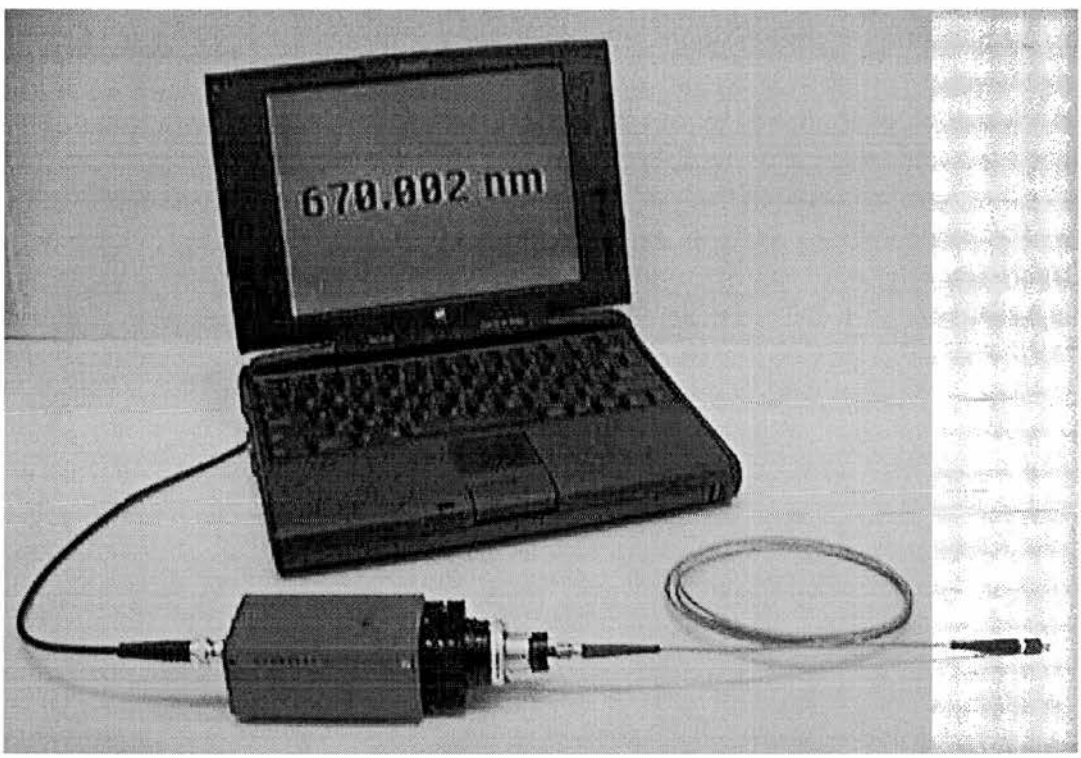

Figure 6.8: *The fibre-coupled, ultra-compact static Fourier-transform spectrometer and wavemeter interfaced to a laptop computer.*

6.5 Wavelength calculation

Fourier transform spectrometers have a restricted wavelength accuracy when few data points are supplied to the FFT algorithm. Limited interferogram data is particularly relevant to static instruments where the fringes are recorded by a detector array. The Nyquist criterion requires two data points per fringe to avoid under sampling problems, restricting the interferogram to a few hundred fringes in most cases.

6.5.1 Fourier transform spectrometer

An example of a He-Ne laser spectrum from the static FTS can be seen in figure 6.9 (i) with a resolution of approximately 5 nm FWHM at 633 nm. Zooming into the He-Ne spectral peak in figure 6.9 (ii) it can be seen that low resolution spectra often produce broad uneven spectral peaks with few data points. The uneven peak coupled with the simple algorithm often employed to determine the peak wavelength can produce results with limited accuracy.

```
WeightedIndex = 0;
Weight = 0;
for (i = (MaxIndex - 5); i <= (MaxIndex + 5); i++)
{
      WeightedIndex = (Wavelength[i] * Intensity[i])
                    + WeightedIndex;
      Weight        = Intensity[i] + Weight;
}
WavelengthValue = WeightedIndex / Weight;
```

Figure 6.9: *Wavelength calculation routine in the C programming language which is used when running the instrument as a Fourier transform spectrometer. The peak intensity index value is determined in the spectrum which is averaged with 5 wavelengths either side which are used proportionately with their intensity.*

The simple method utilised to calculate the peak wavelength when running the instrument as a Fourier transform spectrometer first determines the wavelength value with the highest intensity. Due to the uneven nature of the peak this is often not the actual wavelength of the laser, even though it may appear to be at the peak centre. We therefore use five wavelength data points either side of the maximum intensity wavelength to calculate the average wavelength value, the simple algorithm is shown in figure 6.10. To ensure uneven peaks are accounted for each wavelength value is only used in the averaging calculation in proportion to its intensity. The result of this can be seen in figure 6.10 (ii) where the calculated wavelength (dashed line) is shifted slightly to the right of the spectral data point

with the highest intensity. Intuitively this is correct due to the intensity values on the right side being slightly higher than the intensities on the left side of the spectral peak. Using the instrument as a Fourier transform spectrometer we have demonstrated a wavelength accuracy of 0.2 *nm*.

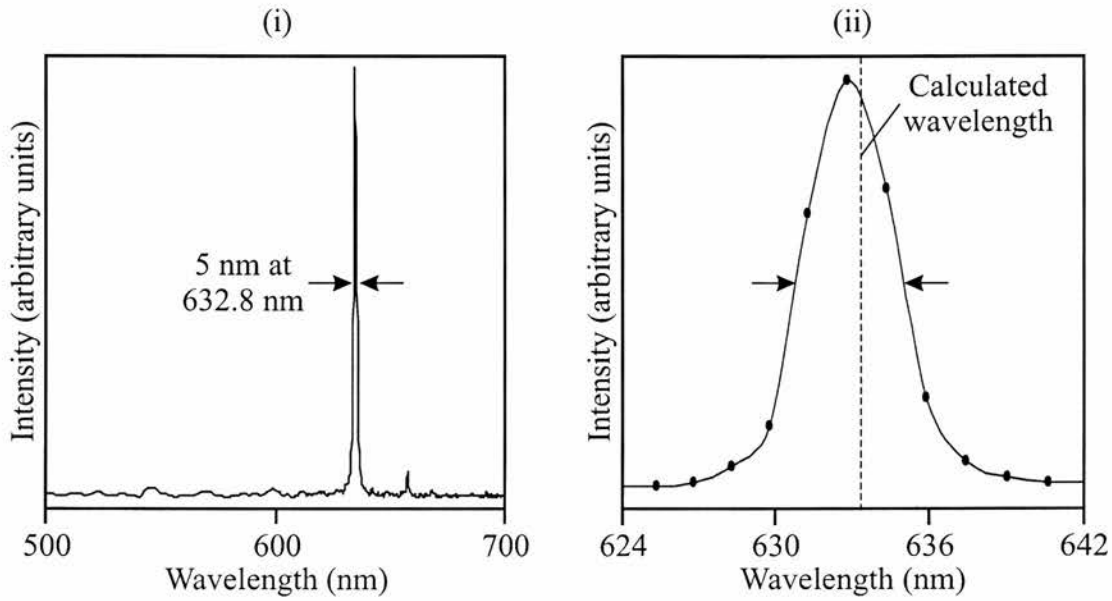

Figure 6.10: *(i) Typical spectrum of a He-Ne laser recorded while operating the instrument as an Fourier transform spectrometer. The peak is approximately 5 nm wide at FWHM. (ii) The spectral peak at 632.8 nm magnified.*

However, the lack of moving parts within the spectrometer itself results in an extremely stable interferogram. This inherent stability can be used to increase the precision beyond that dictated by the number of fringes within the interferogram. If it is assumed that the input light is quasi-monochromatic then a direct measurement of the sinusoidal fringe period gives the wavelength of the source to an arbitrary high accuracy.

6.5.2 Laser wavemeter

Our approach measures the fringe period of the interferogram to determine the wavelength of laser light directly. This method does not require a reference source, which not only simplifies the instrument but also gives a wavelength measurement range of several 100 nm with no loss of accuracy.

We use an algorithm for measuring the fringe period that was proposed by Snyder[2] which calculates the period of a truncated sinusoidal fringe pattern in the presence of noise. The method suffers no apparent systematic errors induced by a nonuniform fringe amplitude or by changes in the initial and final phases of the fringe pattern. The basis of the algorithm is to first smooth the data by processing with a simple adaptive filter that locates the symmetry points of the fringe pattern (i.e., the zero axis crossing points of the sinusoid). Using a least squares method a straight line is then fit to the set of symmetry point positions. The period (or frequency) of the fringe pattern is related directly to the slope of the straight line fit to the data.

The simple method for calculating the fringe period can be understood with reference to figure 6.11. Part (i) shows the raw interferogram data recorded by the instrument. The first operation of the algorithm is to filter the raw data to smooth and centre it about a zero axis. The symmetry filter chosen is a single cycle of a unit amplitude square wave centred at the origin. The period of the filter requires careful selection to avoid asymmetric phase errors. Analysis shows that, although a filter width cannot be chosen for which the phase error is zero, a carefully chosen filter will make the phase error at the maxima of the sinusoid equal and opposite the phase error at the minima, giving an unbiased calculation of the fringe period[2]. Ideally we would like the filter to be as large as possible, to suppress high frequency fluctuations (noise), but small compared with the size of the data set. As a first attempt the filter period is calculated from an estimate of

the fringe period. The symmetry filter is then convolved with the raw data, and the (interpolated) zero crossings of the filter output are calculated. The result of the second phase of the calculation can be seen in part (ii) of figure 6.11. Finally, in part (iii), the symmetry points are plotted with the set of crossing indices as the independent variables. This is processed by a linear least-squares routine to obtain the slope of the fitted line. The fringe period is twice the gradient of the fitted line, from which the wavelength of the incident light can be determined. Appendix C contains the source code for the fringe period calculation in the C programming language.

In general it will be necessary to iterate the algorithm using the new estimate of the fringe period to calculate the period of the filter. In most situations this procedure converges very rapidly, often after one iteration, to the correct value[2].

The scaling between the fringe period and wavelength is set by the calibration of the instrument. Like the Fourier transform spectrometer, the variation in birefringence with wavelength means the relationship is not linear. Therefore, the calibration takes account of this.

The path difference, Δ, introduced by a Wollaston prism between orthogonally polarised components of the incident light was given previously by equation 3.2[21]

$$\Delta = 2d\left[\Delta n\left(\lambda\right)\right]\tan\vartheta,$$

where d is the displacement from the centre of the prism, $\Delta n\left(\lambda\right)$ is the wavelength dependent birefringence of the prism material, and ϑ is the wedge angle. For a monochromatic source, the fringe spacing, S, is therefore given by:

$$S = \frac{\lambda}{2\left[\Delta n\left(\lambda\right)\right]\tan\vartheta} = C\left(\lambda\right).\lambda, \tag{6.3}$$

The denominator in this equation relates to the calibration of the instrument and can be measured explicitly. Rearranging this equation to give the wavelength in terms of a measured fringe spacing gives:

$$\lambda = C'\left(S\right).S \tag{6.4}$$

The calibration of the wavemeter is carried out using several sources of known wavelength and corresponding fringe spacing. A functional form of $C'\left(S\right)$ is obtained. This can then be used to calculate C' for any observed fringes spacing enabling the corresponding wavelength to be determined.

The relationship between fringe period and wavelength is set by the pitch of the CCD array and the wedge angle of the Wollaston prism. Therefore, the calibration of the instrument does not change significantly with time and good performance can be obtained without a reference laser.

6.6 Results

As with earlier designs of the instrument it can be used as a general purpose laboratory spectrometer where the Fourier-transform of the interferogram gives an estimate of the power spectrum of the incident light. The use of a single birefringent element only 1.5 mm thick results in a large angular acceptance of $\pm 5°$. Representative performance of the spectrometer is given in figure 6.12 which shows the recorded interferogram and Fourier-transformed spectrum of a He-Ne laser and 670 nm diode laser simultaneously incident onto the input

aperture. The resolution of the spectrometer is set by the number of pixels and corresponding maximum path difference[24]. Using a selection of narrow-band, interference filters the resolution of the spectrometer was measured to be 5 nm at 633 nm. The operating spectral range of the instrument of 400–900 nm is limited by the specification of the polarisers. The data acquisition rate of approximately 25 frames a second was dictated by the speed of the camera-computer interface.

As discussed earlier, when used as a laser wavemeter the incident light is fibre-coupled into the instrument. This not only gives a convenient instrument but also ensures a uniform and consistent illumination of the Wollaston prism. When operating as a laser wavemeter the key performance criteria are short time scale precision and long time scale stability. Figure 6.13 shows the wavelength of a He-Ne laser measured using a one-second averaging time. This shows the short term stability of the wavemeter to be approximately 1 part in 10^6 which should be compared with the inherent stability of a standard He-Ne laser of approximately 1 part in 10^6. Over a 12 hour period, figure 6.13 shows the observed stability is 2 parts in 10^5.

Figure 6.14 shows the wavelength of a diode laser with a nominal emission wavelength of 670nm recorded as a function of time. Figure 6.14 (i) shows wavelength measurement over a 12-hour period, the change in wavelength arises from the change in ambient temperature. Figure 6.14 (ii) shows the wavelength change arising from deliberate heating and cooling of the laser diode.

6.7 Conclusions

I have described the design, construction and initial evaluation of a static Fourier transform spectrometer and laser wavemeter based on a modified Wollaston prism configuration. The operating spectral range of 400–900 nm is set by the choice

of polarisers but alternative polarisers and detector array selection would allow operation anywhere within the transparency window (220–2200 nm) of the calcite Wollaston prism.

When used as a laboratory spectrometer, the large field of view and large aperture give an instrument with a large optical throughput. Thus the "Jacquinot advantage" of Fourier-transform spectroscopy can be realised in this novel design.

When used as a laser wavemeter, fibre coupling gives convenient use and ensures uniform intensity across the interferogram. Rather than use a Fourier-transform of the interferogram data we use an algorithm to measure the fringe period and hence the wavelength to an arbitrary high precision. By recording the wavelength of a He-Ne laser which is inherently stable to better than 1 part in 10^6 we show that the one-second measurement time stability of our instrument is approximately 1 part in 10^6. The long time-scale stability is 2 parts in 10^5, which we believe to be limited by the temperature coefficient of the birefringence of the calcite. Windows environment software gives a user friendly instrument when used in either spectrometer or laser wavemeter configuration.

Of course it should be noted that a full test of the true accuracy requires a large number of calibration measurements at wavelengths throughout the relevant spectral region in order to sufficiently sample the effects of arbitrary periodic and dispersion-type systematic errors. Unfortunately, we could only find a few sources with sufficient accuracy to calibrate the wavemeter at or below the 10^{-6} level so the instrument is supported by calibration for only a few points.

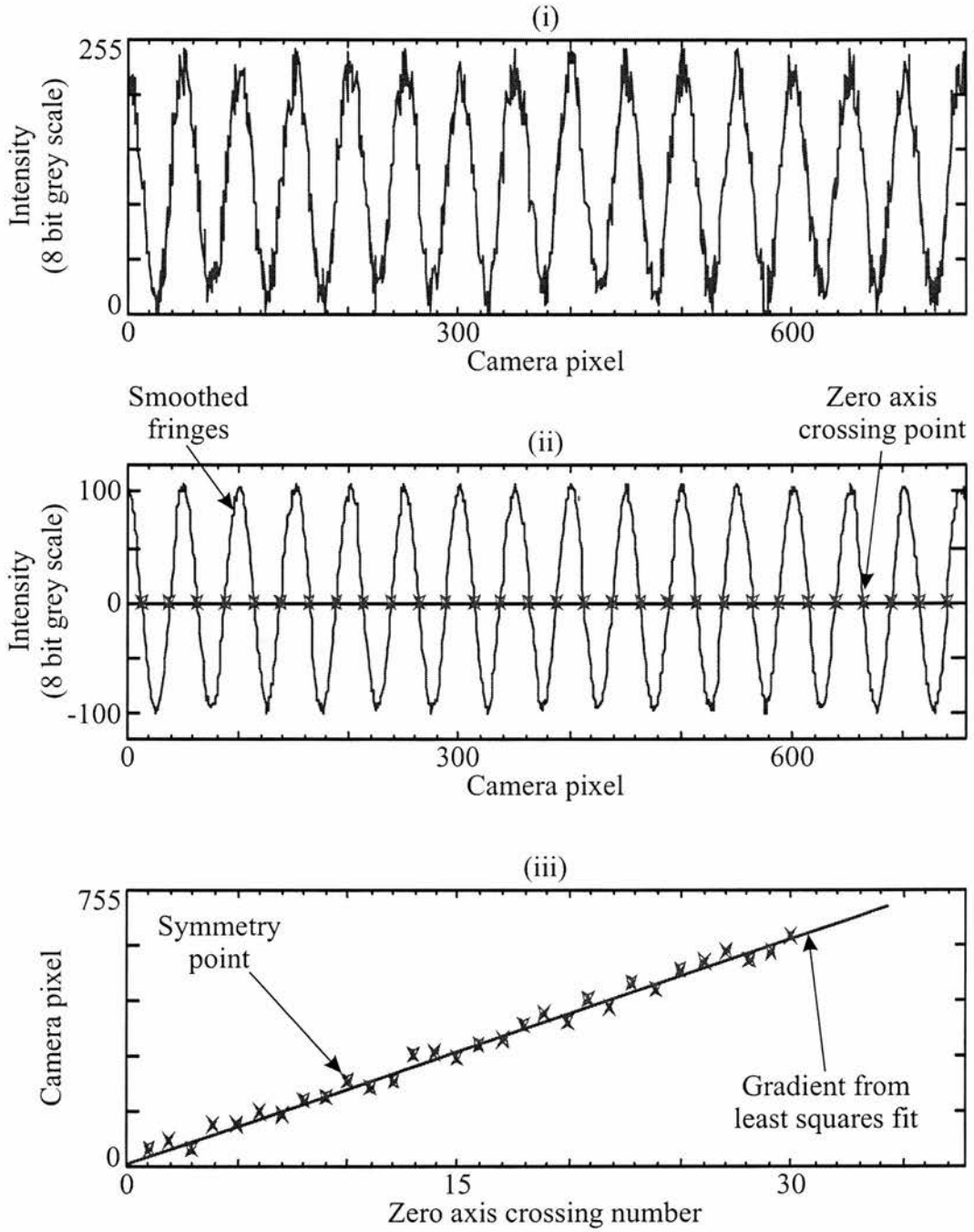

Figure 6.11: *The simple algorithm to determine the fringe period and hence the wavelength of the incident light. (i) The raw fringe data from the instrument showing high frequency noise. (ii) The smoothed fringe data centered about the zero axis. The smoothing filter removes the high frequency noise components from the interferogram, allowing accurate, interpolated determination of the zero axis crossing points of the sinusoid. (iii) Plot of symmetry points showing the best fit straight line. The gradient of the line is proportional to the wavelength.*

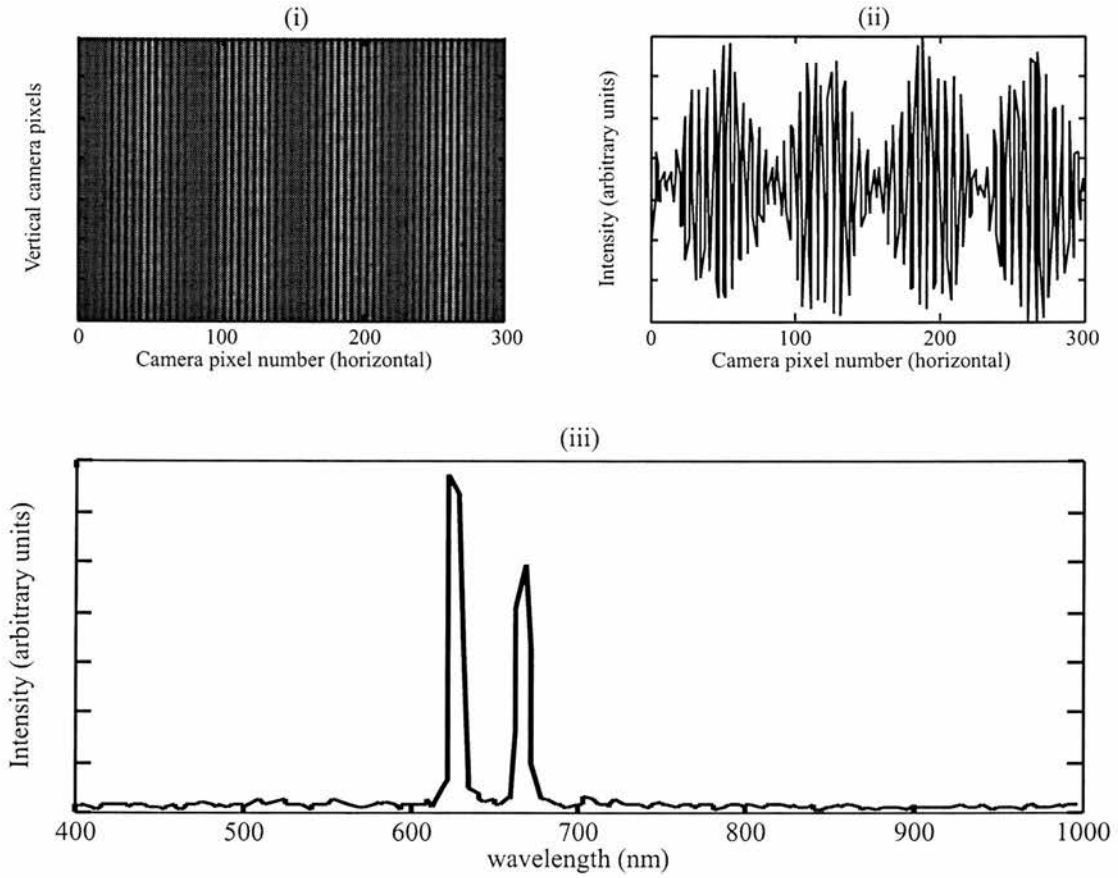

Figure 6.12: *Sample results when using the instrument as a spectrometer while illuminating with He-Ne and diode lasers. Part (i) is a section of the 2-D fringes as recorded by the detector array, part (ii) is the corresponding 1-D interferogram and part (iii) is the Fourier-transformed result.*

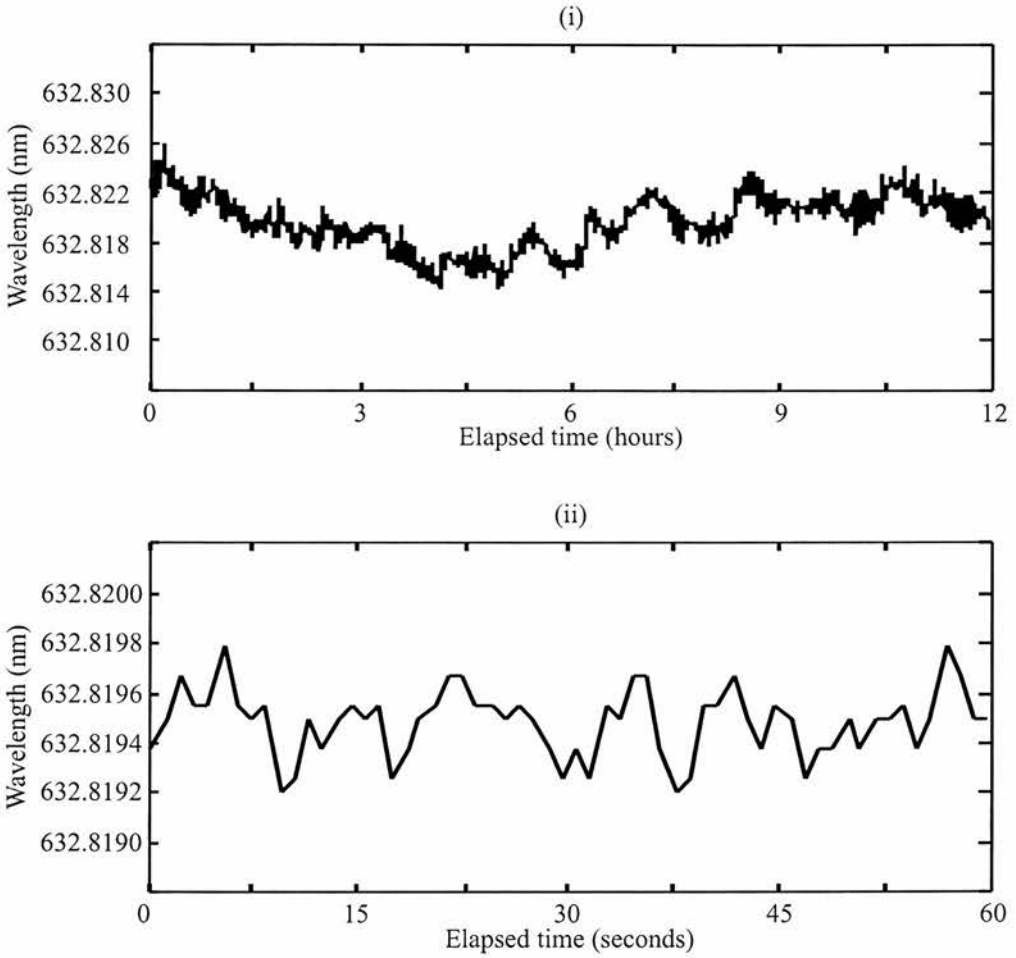

Figure 6.13: *Sample results from the wavemeter. Wavelength measurement was of a commercial He-Ne laser over a (i)* 12 *hour period which demonstrated a stability of the instrument of approximately* 2 *parts in* 10^5. *Part (ii) is a short term recording over* 60 *seconds, demonstrating an accuracy and stability of approximately* 1 *part in* 10^6.

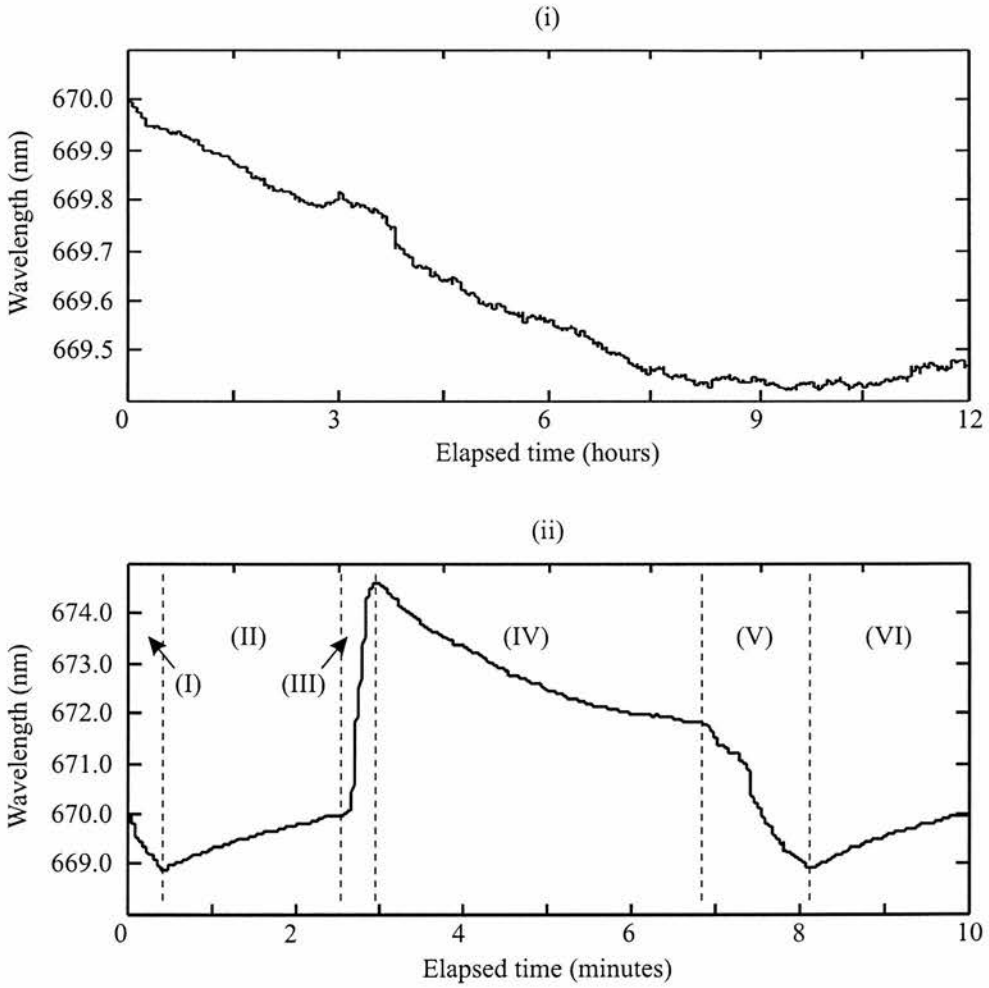

Figure 6.14: *Sample results from the wavemeter for a commercial laser diode operating at a nominal wavelength of 670nm. Part (i) was recorded over a 12 hour period and shows the wavelength drift of the diode as the ambient room temperature changes. Part (ii) was recorded over a period of 10 minutes while the diode was subjected to forced heating and cooling. Section (I) was a period of forced rapid cooling, in section (II) the diode heated up naturally followed in section (III) by a period of forced rapid heating. The diode was allowed to cool unaided in section (IV) until it was rapidly cooled again in section (V), and finally allowed to heat up again to its preferred operating temperature in section (VI).*

Chapter 7

Conclusions

This thesis has reported the development of novel static Fourier transform spectrometers and laser wavelength meters based upon Wollaston prisms and compact detector arrays. Key achievements of this work have been the development of a wide field of view spectrometer, an accurate laser wavelength meter and systems for the analysis of gasoline.

7.1 Wide field of view spectrometer

My work on the spectrometer has increased significantly the acceptance angle, and therefore the light gathering power of the instrument. Using two Wollaston prisms fabricated from materials with opposite sign of birefringence has resulted in a spectrometer with a field of view of $\pm 10^{o}$, an increase in étendue of approximately 4 times previous designs. Thus the "Jacquinot advantage" of Fourier transform spectroscopy can be realised in this novel design.

This is particularly important if the application requires the sample and/or light

source to be close to the entrance aperture of the spectrometer. Consider the case of an in-line monitor with the light source and spectrometer head on opposite sides of a flow pipe. The spectrometer has the large entrance aperture and large field of view required to use all the light from the extended source.

7.2 Laser wavelength meter

Utilising a modified Wollaston prism configuration I have assembled a dual purpose compact Laser wavelength meter and Fourier transform spectrometer. Fibre coupling gives convenient use and ensures uniform intensity across the interferogram. Rather than use a Fourier-transform of the interferogram data I use an algorithm to measure the fringe period and hence the wavelength to a precision of 1 part in 10^6. An increase of 1000 times over the precision of the Fourier transform method.

7.3 Gasoline measurement

The success of the near infrared instrument in measuring octane number, although encouraging for spectrometers based upon Wollaston prisms in general, is not entirely unexpected as similar results have been achieved with dispersive instruments. However, our approach may offer advantages such as greater wavelength flexibility, larger optical throughput, and less sensitivity to alignment problems giving more reliable spectra recording.

The measurement of gasoline properties in the ultraviolet spectral region is promising. However, it must be stressed that these are preliminary results because of the limited data set of fuels. A more exhaustive data set must be used before

firm conclusions can be drawn. However, comprehensive fuel sets are expensive and further work would require significant support from Shell.

7.4 General Instrumentation

In addition to higher étendue, a further advantage of FTSs compared with grating instruments is that FTS do not suffer from overlapping spectral orders. Hence, it is difficult for grating spectrometers to operate over more than one spectral order, i.e. 200–400 *nm* or 400–800 *nm*. By contrast a FTS suffers no technical limitation. Rather it is the availability of polarisers which sets the operating range.

Taking a SFTS design based on an economical silicon detector array, the practical wavelength range is 200–1100 *nm*. Unfortunately current polariser technology can not match this wide range and as in the case of the UV and NIR instruments a choice has to be made. New application areas may be accessible should new polarisers become available. Even extending the work on fuel monitoring may produce more accurate results by enabling simultaneous measurement and analysis in the ultraviolet and near-infrared spectral regions

Each of the spectrometers are controlled with windows based software, giving simple to control, user friendly instruments. It has also been possible to interface the spectrometers to lap-top computers enabling fully portable measurement systems.

The wavelength meter and spectrometers have numerous applications in the field of general laboratory instrumentation and their robust, compact nature makes them particularly suitable where field based operation is required.

The main design limitation is the modest resolution of the spectrometer. Although the resolution can be improved by increasing the number of elements on the detector array, this can only be done with increased cost.

7.5 The future

The spectrometers and wavelength meters have reached the point where commercialisation is now a feasible goal. Following exhibition at several international trade fairs and optics conferences, I have received many enquiries related to marketing or buying the prototype designs, demonstrating the commercial potential of these instruments.

In recognition of this potential I have been awarded an Enterprise Fellowship from Scottish Enterprise and the Royal Society of Edinburgh. The fellowship will provide the ideal platform to transform the prototype instruments into competitive marketable products. The real test of this project work is yet to come..... Can it survive in the commercial environment?

Appendix A

Spectrometer calibration source code

The following source code is used for calibrating the Wollaston prism spectrometers. It requires a knowledge of the wavelength birefringence dependence of the prism material.

```
/*************************************************************************
 Procedure   : CalibrateSpectrometer
 Description : Calibrates the spectrometer by reading in a fresh data set,
               performing an FFT to get the spectrum and finding the peak in
               spectrum.  This peak is the assigned the wavelength that the
               user has supplied and the X axis scale is calibrated with this
               information and a knowledge of the wavelength birefringence
               dependence of the Wollaston prism.
*************************************************************************/
void CalibrateSpectrometer(double CalibLamda)
{
        double MaxPeak;

        // Read in a data set to calibrate the spectrometer
        GetCameraData (RawFringes, XPixelElements, YPixelElements, 100);
        // perform the FFT to produce a spectrum from the raw data.
        PerformFFT (RawFringes, Spectrum, XPixelElements);

        // find the peak in the spectrum
        GetMaxPeak(&MaxPeak, Spectrum, XPixelElements);

        // Use the calibration data to calculate the calibrated
```

```
        // wavenumber/length arrays
        CalcCalibratedWavenumbersArray (CalibWaveNoArray, CalibLamda,
                                        MaxPeak, XPixelElements);
        CalcCalibratedWavelengthsArray (CalibLamdaArray, CalibWaveNoArray,
                                        XPixelElements);
}

/*************************************************************************
 function    : DeltaN
 Description : Evaluates polynomial representation of the birefringence of
               calcite which varies with wavelength.
*************************************************************************/
double DeltaN (double wavenumber)
{
     double calc_value;

     // make sure wavenumber is in a sensible range (1000-320nm)
     if (wavenumber < 9000)
           wavenumber = 9000;

     if (wavenumber > 31000)
           wavenumber = 31000;

     calc_value = -0.380192 - ((19201.3 * pow(wavenumber, 2))/ 1E+14)
               + ((99.5269 * wavenumber) /  1E+7)
               + (1937.01 / wavenumber)
               - (5.69789E+6 / pow(wavenumber, 2));

     return(calc_value);
}


/*************************************************************************
 Procedure   : CalcCalibratedWavenumbersArray
 Description : Calculates the calibrated wavenumbers array.
               Algorithm as described in Patterson et al (1997).
*************************************************************************/
void CalcCalibratedWavenumbersArray (   double kOut[FFT_POINTS],
                                        double CalWaveLength,
                                        double calIndex,  int NoPoints)
{
    int i;
    double CalWaveNo;

    CalWaveNo = (1 / CalWaveLength)* 10000000;

    // This is an iterative process.  After each iteration the
    // the correction required for the birefringence variation is
    // more accurate.  Two or three passes are all that is required
    // but it is performed 7 times to ensure accuracy.
    for (i = 1; i < NoPoints; i++)
    {
          kOut[i] = i * CalWaveNo / calIndex;
          kOut[i] = (i * CalWaveNo / calIndex) * (DeltaN (CalWaveNo)
                                                / DeltaN (kOut[i]));
```

```
            kOut[i] = (i * CalWaveNo / calIndex) * (DeltaN (CalWaveNo)
                                                 / DeltaN (kOut[i]));
            kOut[i] = (i * CalWaveNo / calIndex) * (DeltaN (CalWaveNo)
                                                 / DeltaN (kOut[i]));
            kOut[i] = (i * CalWaveNo / calIndex) * (DeltaN (CalWaveNo)
                                                 / DeltaN (kOut[i]));
            kOut[i] = (i * CalWaveNo / calIndex) * (DeltaN (CalWaveNo)
                                                 / DeltaN (kOut[i]));
            kOut[i] = (i * CalWaveNo / calIndex) * (DeltaN (CalWaveNo)
                                                 / DeltaN (kOut[i]));
            kOut[i] = (i * CalWaveNo / calIndex) * (DeltaN (CalWaveNo)
                                                 / DeltaN (kOut[i]));
    }
}


/************************************************************************
 Procedure   : CalcCalibratedWavelengthsArray
 Description : Uses the Calibrated wavenumbers array to calculate the
               calibrated wavelength X axis array.
************************************************************************/
void CalcCalibratedWavelengthsArray ( double LambdaOut[FFT_POINTS],
                                      double WaveNoIn[FFT_POINTS],
                                           int NoPoints)
{
        int i;

     for (i = 1; i < NoPoints; i++)
     {
         LambdaOut[i] = 10000000/WaveNoIn[i];
     }
}



/************************************************************************
Procedure   : GetMaxPeak
Description : Returns the position in the spectrum array that corresponds to
              to the peak.  This is a real number as a weighted system is used.
              The previous 5 and following 5 points are used to produce a more
              accurate positioning of the centre of the peak.
************************************************************************/
void GetMaxPeak(double *MaxPeak, double Spectrum[FFT_POINTS], int ArraySize)
{
        int i, MaxIndex;
        double tempMax, WeightedIndex, Weight;

        tempMax = Spectrum[1];
        MaxIndex = 0;

        // find the index value with the highest intensity.
        for (i = 1; i < ArraySize; i++)
        {
                if (Spectrum[i] > tempMax)
                {
```

```
                    tempMax = Spectrum[i];
                    MaxIndex = i;
            }
        }
        WeightedIndex = 0;
        Weight = 0;

        // calculate the wavelength taking into account 5
        // wavelengths and their intensities either side of the
        // maximum value already found.
        for (i = (MaxIndex - 5); i <= (MaxIndex + 5); i++)
        {
            WeightedIndex = (i * Spectrum[i]) + WeightedIndex;
            Weight =  Spectrum[i] + Weight;
        }
        *MaxPeak = WeightedIndex / Weight;
}
```

Appendix B

Principal Component Analysis

B.1 Introduction

Principal component analysis (PCA) is a multivariate statistical technique that can be used to reduce a large number of interrelated variables to a smaller set of variables. For example, a survey questionnaire with 100 items may be completed by several thousand respondents. Such an extensive survey will usually contain a lot of information that cannot easily be understood without some sort of summarisation. Although simple descriptive statistics can help one to examine the frequency distribution of responses or correlations on particular pairs of items visualising relationships among items is generally limited to two or three dimensions. As such, it seems almost natural to try and reduce the dimensionality of the data for descriptive purposes but still preserve as much of the underlying structure of the relationships as possible.

The basic approach to this reduction first involves gathering observations from a sample of n people on some set of p variables of interest. This provides a simple data matrix containing observations on p variables from the n sample. In matrix

form, this collection of information can be represented by a matrix X as follows:

$$X = \begin{bmatrix} x_{11} & x_{12} & . & . & . & x_{1p} \\ x_{21} & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ x_{n1} & . & . & . & . & . \end{bmatrix}$$

Where the rows represent the persons questioned and the columns the variables or questions asked. The concern, therefore, is whether the matrix X can be replaced by another matrix with far less columns than p, and yet represent the information in X as closely as possible. A powerful method for achieving this data reduction is by PCA.

Principal component analysis is a method used to reduce a set of observed variables into a relatively small number of components that account for most of the observed variance. This is accomplished by mathematical linear transformations of the observed variables under two conditions. The first condition is that the first component (the principal component) accounts for the maximum amount of variance possible, the second the next, and so on and so forth. The second condition is that all components are uncorrelated with each other.

B.2 A simple example of principal component analysis

The easiest way to illustrate PCA is by way of a simple bivariate case. Assume that two variables X_1 and X_2 are observed from a large sample, and that, for

ease of presentation, their distribution is bivariate normal. Figure B.1 represents a bivariate normal relationship with a positive correlation. Note that it is customary to represent a bivariate normal distribution by drawing cross sections of the surface of a bell-shaped mound, called isodensity contours.

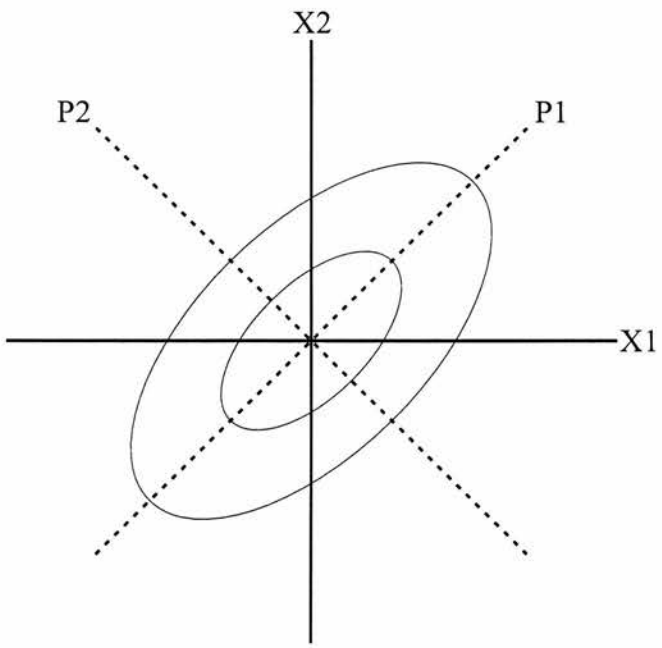

Figure B.1: *Bivariate normal distribution with a positive correlation.*

The contours presented in figure B.1 show that higher values of X_1 tend to be associated with higher values of X_2 (and vice versa). This is the reason most observed cases are piled up along the first and third quadrants. As it turns out, the axes of the observed contours can be graphed as a principal axis (P1) representing the line on which most of the data points are located. Of course, there is a second axis (P2), which can be used to represent the data with the fewest observed cases.

In general terms, PCA can be conceptualised as a technique that attempts to represent the relative position of each observed case on the best dimension or

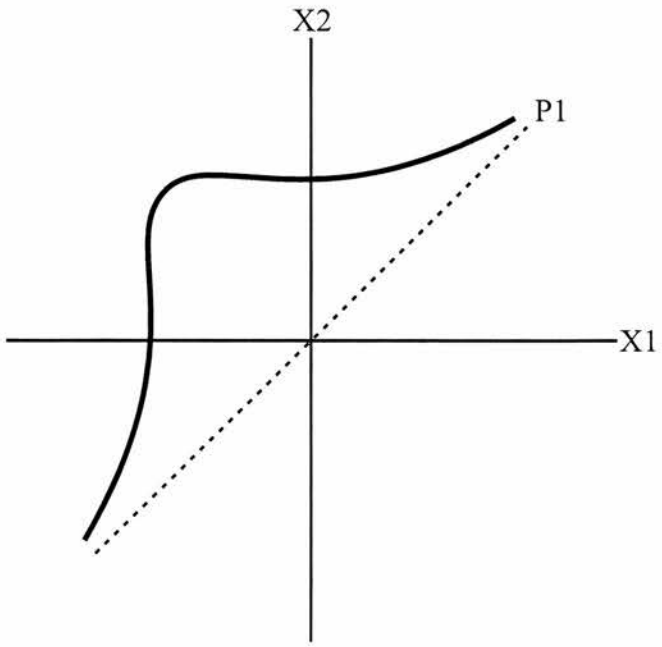

Figure B.2: *Bivariate normal distribution with a perfect correlation between X_1 and X_2.*

axis. Obviously, for this example case, the best choice is the first axis (P1). In the case where the two variables X_1 and X_2 are perfectly correlated (see figure B.2, the first principal axis contains all the information necessary to describe each case. Similarly, if the two variables are uncorrelated (see figure B.3), there is no principal axis that can be used to describe the observed cases. Thus, the principal component analysis is no more than a presentation of cases along some principal axes.

B.3 The principal components

Principal components are mathematical functions of observed variables. The PCA functions do not attempt to explain the intercorrelations among the variables, but to account for the variability observed in the data. Consider an example

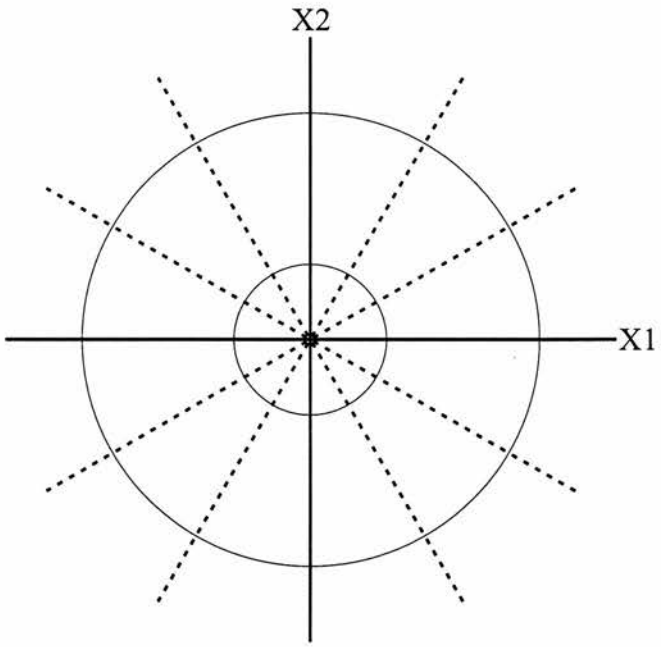

Figure B.3: *Bivariate normal distribution with no correlation between X_1 and X_2.*

where eight variables of interest are observed from a sample (i.e., $x_1, x_2, \ldots, x_8$), and found to have a correlation matrix R. Using these eight observed variables, a PCA will attempt to construct linear combinations as follows:

$$\begin{aligned}
Y_1 &= \Gamma_1' X = \gamma_{11}x_1 + \gamma_{21}x_2 + \ldots + \gamma_{81}x_8 \\
Y_2 &= \Gamma_2' X = \gamma_{12}x_1 + \gamma_{22}x_2 + \ldots + \gamma_{82}x_8 \\
Y_3 &= \Gamma_3' X = \gamma_{13}x_1 + \gamma_{23}x_2 + \ldots + \gamma_{83}x_8 \\
&\vdots \\
Y_8 &= \Gamma_8' X = \gamma_{18}x_1 + \gamma_{28}x_2 + \ldots + \gamma_{88}x_8
\end{aligned}$$

or in matrix notation

$$Y = \Gamma' X$$

The Ys in these equations are the computed principal components, and the γs are the coefficients of the observed variables. thus, the principal component are the uncorrelated linear combinations of Ys whose variances are as large as possible. The first principal component is the linear combination with maximum variance. This basically is the component that maximises the variance of Y_1. However, because the variance of Y_1 can be increased by choosing any Γ' (i.e., any set of variables for $\gamma_{11}, \gamma_{21}, \gamma_{31}, \gamma_{41}, \gamma_{51}, \gamma_{61}, \gamma_{71}$, and γ_{81}), a restriction that the product be equal to unity (i.e., $\Gamma'\Gamma = 1$) is also imposed. Thus, the first principal component is that selection of γs that maximises the variance of Y_1 subject to $\Gamma_1'\Gamma_1 = 1$. The second principal component is the linear combination that maximises the variance of Y_2 subject to $\Gamma_2'\Gamma_2 = 1$, and assures that the two combinations are uncorrelated. The analysis continues until all principal components are generated (e.g., in this case there will be eight components).

This thesis does not attempt a thorough description of PCA, a more rigorous explanation including the methodology of calculation can be found in the many textbooks on multivariate statistics [63, 64, 65, 66].

Unfortunately, the question of how many axes or principal components are important or needed to account for the variance in the data is not straight forward. In other words, given the eigenvalues of a correlation or covariance matrix, how does one decide how many to use in order to account for the observed variance? There really is no correct answer as it depends upon the data set and the application. We have found for the analysis of gasolines the first two principal components have shown sufficient variance to enable identification of key fuel properties.

Appendix C

Wavemeter source code

The following is the source code used by the laser wavemeter to calculate the fringe period. The algorithm assumes the light source is monochromatic and in essence first smooths the data by processing with a simple adaptive filter that locates the symmetry points of the fringe pattern (i.e., the zero axis crossing points of the sinusoid). Using a least squares method a straight line is then fit to the set of symmetry point positions. The period (or frequency) of the fringe pattern is related directly to the slope of the straight line fit to the data. The program extract is written in the C programming language

```
/*****************************************************************************
 Function    : GetWavemeterLambda
 Description: Returns the wavelength (lambda) which is calculated from the
              RawFringeData.
*****************************************************************************/
double GetWavemeterLambda(double RawFringeData[], double RawDataModified[],
                          int NoPoints)
{
    #define MaxNumIterations 10
    #define eps 0.0000001

    int    b, i;
    double lambda, lambdaOld;

    i = 0;
    lambda = 5;      // this is a rough estimation for lambda as starting
```

```
                    // value for the iteration

    // do-while loop performs the iteration until the last calculated
    // wavelength is near the current calculated wavelength
    do {
        i = i + 1;
        b = floor((0.371 * lambda) + 0.5);
        if (b == 0)
        {
                b = 1;
        };
        sym(b, RawFringeData, RawDataModified, NoPoints);
        lambdaOld = l;
        lambda = fabs(lambda (b, RawDataModified, NoPoints));
    }while((fabs(lambda - lambdaOld) / lambda) > eps || i < MaxNumIterations);

    return(lambda);
}


/*************************************************************************
 Function   : sym
 Description: Performs the convolution to smooth the raw data.
*************************************************************************/
void    sym (int b,  double RawFringeData[],  double SmoothedFringes[],
             int NoPoints)
{
    int         x, y;
    double      sum;

    for (x = b; x < (NoPoints + 1 - b); x++)
    {
        sum = 0;
        for (y = (x - b); y <= (x - 1); y++)
        {
                sum = sum - RawFringeData[y];
        }
        for (y = x; y <= (x + b - 1); y++)
        {
                sum = sum + RawFringeData[y];
        }
        SmoothedFringes[x] = sum;
    }
}


/*************************************************************************
Function   : lambda
Description: lambda(b) yields a better (sometimes) value for the
             wavelength lambda of the original input data
             (as opposed to the array modified by procedure sym).
*************************************************************************/
double lambda (int b, double DataIn[], int NoPoints)
{
    int         j, k, maxy, nu, nd;
```

```
bool        plus;
double      xconst, x, xlast, sfd, sfu, sfjd, sfju, au, ad, sf, sfj, a;

// helps avoid catastrophes if not enough fringes
if (b > 500) b = 500;

// true indicates that the last value in the data array was positive
// this is for detecting the zero crossings
plus = (DataIn[b] >= 0);

/*
initialise the variables needed  for the calculation of the linear
least squares fit;
u means up, i.e. the slope of the convolution is positive(minima of the
original data)
d means down (maxima).
fitting the two sets of zero crossings separately and then taking the mean
value seems to be johannes idea of intelligent.
*/

// nu contains the number of zero crossings of the convolution with
// a positive slope;
nu = 0;             // u = up
sfu = 0;            // will become sum ( i^[j] )
sfju = 0;           // will become sum( i^[j] * j )
nd = 0;             // d = down
sfd = 0;
sfjd = 0;

for ( j = b; j < NoPoints + 1 - b; j++)
{
    if (plus)
    {
        if (DataIn[j] < 0)
        {
            // linear interpolation to determine the zero crossing
            x = j - DataIn[j] / (DataIn[j] - DataIn[j-1]);
            nd = nd + 1;
            sfd = sfd + x;
            sfjd = sfjd + x * nd;
            plus = false;
        }
    }
    else
    {
        if (DataIn[j] >= 0)
        {
            // linear interpolation to determine the zero crossing
            x = j - DataIn[j] / (DataIn[j] - DataIn[j-1]);
            nu = nu + 1;
            sfu = sfu + x;
            sfju = sfju + x * nu;
            plus = true;
        }
    }
```

```
    }

    if (nd > 1)
    {
        ad = (sfjd - s1(nd) * sfd / nd ) / (s2(nd) - pow(s1(nd), 2) / nd);
    }
    if (nu > 1)
    {
        au = (sfju - s1(nu) * sfu / nu ) / (s2(nu) - pow(s1(nu), 2) / nu);
    }
    if (nd > 1 && nu > 1)
    {
        return( (au + ad) / 2 );
    }
    else
    {
        if (nd > 1)
        {
                return(ad);
        }
        else
        {
                if (nu > 1)
            {
                return(au);
            }
            else
            {
                //try a smaller value of b to get some zero crossings
                return( b / 0.371 / 0.95 );
            }
        }
    }
}



\begin{verbatim}
/*************************************************************************
 Function   : s1
 Description: Used for the linear least squares fit.
              s1(n) = sum(i, i = 1..n)  (Maple notation)
*************************************************************************/
double s1 ( long int n)
{
    return( (n * (n + 1)) / 2);
}


/*************************************************************************
 Function   : s2
 Description: Used for the linear least squares fit
              s2(n) = sum(i^2, i = 1..n)  (Maple notation)
*************************************************************************/
double s2 ( long int n)
```

```
{
    return( (n * ( n * (2 * n + 3) + 1)) / 6 );
}
```

Bibliography

[1] J. F. James and R. S. Sternberg. *The Design of Optical Spectrometers.* Chapman and Hall Ltd, London, 1969.

[2] J. J. Snyder. Algorithm for fast digital analysis of interference fringes. *Appl. Opt.*, 19:1223–1225, 1980.

[3] P. R. Griffiths and J. A. de Haseth. *Fourier transform infrared spectrometry*, volume 83 of *Chemical analysis: A series of monographs on analytical chemistry and its applications.* Wiley-Interscience, Chichester, 1986.

[4] A. P. Thorne. *Spectrophysics.* Chapman and Hall, London, second edition, 1988.

[5] D. Halliday and R. Resnick. *Physics.* John Wiley, Chichester, third edition edition, 1978.

[6] M. J. Persky. A review of spaceborne infrared Fourier transform spectrometers for remote sensing. *Rev. Sci. Instrum.*, 66(10):4763–4797, 1995.

[7] R. J. Bell. *Introductory Fourier Transform Spectroscopy.* Academic, New York, 1972.

[8] A. A. Michelson. Title. *Phil. Mag. Ser. 5*, 31:256, 1891.

[9] A. A. Michelson. On the application of interference methods to spectroscopic measurements. *Phil. Mag.*, 34:280, 1892.

[10] P. Fellgett. Title. *J. Phys. Radium*, 19(187), 1958.

[11] P. Jacquinot. The luminosity of spectrometers with prisms, gratings, or Fabry-Perot etalons. *J. Opt. Soc. Am.*, 44:761–765, 1954.

[12] J. W. Cooley and J. W. Tukey. An algorithm for the machine calculation of complex Fourier series. *Math Comput*, 19(297), 1965.

[13] H. Yoshinaga, S. Fujita, S. Minami, Y. Suemoto, M. Inoue, K. Chiba, K. Nakano, S. Yoshida, and H. Sugimori. A far infrared interferometric spectrometer with a special electronic computer. *Appl. Opt.*, 5:1374, 1965.

[14] W. H. Steel. *Interferometry.* Cambridge University Press, London, England, 1967.

[15] P. Jacquinot. New developments in interference spectroscopy. *Rep. Progr. Phys.*, 23:267–312, 1960.

[16] G. W. Stroke and A. T. Funkhouser. Fourier-transform spectroscopy using holographic imaging without computing and with stationary interferometers. *Phys. Lett.*, 16(3):272–274, 1965.

[17] K. Yoshihara and A. Kitade. Title. *Japan. J. Appl. Phys*, 6:116, 1967.

[18] T. Okamoto, S. Kawata, and S. Minami. Fourier transform spectrometer with a self-scanning photodiode array. *Applied Optics*, 23(2):269–273, 1984.

[19] T. H. Barnes. Photodiode array Fourier transform spectrometer with improved dynamic range. *Appl. Opt.*, 24:3702 – 3706, 1985.

[20] L. V. Egorova, I. E. Leshcheva, B. N. Popov, and A. Yu Stroganova. Static fast-response Fourier spectrometer having a linear CCD image forming system. *Sov. J. Opt. Technol*, 56(4):220–221, 1989.

[21] M. Françon and S. Mallick. *Polarization Interferometers.* Wiley-Interscience, London, 1971.

[22] T. Okamoto, S. Kawata, and S. Minami. A photodiode array Fourier transform spectrometer based on a birefringent interferometer. *Appl. Spectroscopy*, 40(5):691–695, 1986.

[23] M. J. Padgett, A. R. Harvey, A. J. Duncan, and W. Sibbett. Single-pulse Fourier transform spectrometer having no moving parts. *App. Opt.*, 33(25):6035–6040, 1994.

[24] M. J. Padgett and A. R. Harvey. A static Fourier transform spectrometer based on Wollaston prisms. *Rev. Sci. Instrum.*, 66(4):2807–2811, 1995.

[25] B. A. Patterson, M. Antoni, J. Courtial, A. J. Duncan, W. Sibbett, and M. J. Padgett. An ultra-compact static Fourier-transform spectrometer based on a single birefringent component. *Optics Communications*, 130(9):1–6, 1996.

[26] D. Steers, B. A. Patterson, W. Sibbett, and M. J. Padgett. Wide field of view, ultracompact static fourier-transform spectrometer. *Rev. Sci. Instrum.*, 68(1):30–33, 1997.

[27] J. Courtial. A Fourier transform spectrometer based on a Wollaston prism: theoretical analysis and application to the detection of hydrogen sulphide. Master's thesis, Physics and Astronomy, University of St Andrews, Scotland, 1995.

[28] W. H. Press, S.A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical recipies in C.* Cambridge University Press, Cambridge, England, 2nd edition edition, 1992.

[29] Halbo Optics, 83 Haltwhistle Road, Chelmsford, England, Product Guide 3.

[30] J. Courtial, B. A. Patterson, A.R. Harvey, W. Sibbett, and M.J. Padgett. Design of a static Fourier transform spectrometer with increased field-of-view. *Appl. Optics*, 35(34):6698–6702, 1996.

[31] J. Chamberlain. *Principles of Interferometric Spectroscopy.* Wiley-Interscience, New York, 1979.

[32] A. E. Martin. Infrared interferometric spectrometers. In J. R. Durig, editor, *Vibrational Spectra and Structure*, volume 8. Elsevier, Amsterdam, 1980.

[33] E. Hecht and A. Zajac. *Optics.* Addison-Wesley, Reading, 1974.

[34] MATHEMATICA, Wolfram Research, Inc., Champaign, Ill.

[35] M. C. Simon and R. M. Echarri. Ray tracing formulas for monoaxial optical components: vectorial formulation. *Appl. Optics*, 25:1935–1939, 1986.

[36] VVL1011 PCB, VLSI Vision Limited, Edinburgh, Scotland.

[37] Microsoft Corporation, One Microsoft Way, Redmond, WA 98052.

[38] G. Proticlovasic, N. Jambrec, D. Deursiftar, and M. V. Prostenik. Determination of catalytic reformed gasoline octane number by high resolution gas-chromatography. *Fuel*, 69(4):525–528, 1990.

[39] B. S. Hoffheins and R. J. Lauf. Performance of simplified chemical sensor arrays in a neural network-based analytical instrument. *Analusis*, 20(4):201–207, 1992.

[40] C. W. McCarrick, D. T. Ohmer, L. A. Gilliland, P. A. Edwards, and H. T. Mayfield. Fuel identification by neural-network analysis of the response of vapour-sensitive sensor arrays. *Analytical Chemistry*, 68(23):4264–4269, 1996.

[41] J. M. Andrade, S. Muniategui, P. Lopez, and D. Prada. Costs, laboratory safety, productivity and faster research octane number and motor octane number determinations in industrial chemistry laboratories. *Analyst*, 120(2):249–253, 1995.

[42] M. Ichikawa, N. Nonaka, I. Takada, and S. Ishimori. Estimation of the octane number of automobile gasoline by Fourier-transform infrared absorption spectrometry. *Applied spectroscopy*, 46(6):996–971, 1992.

[43] G. E. Fodor, K. B. Kohl, and R. L. Mason. Analysis of gasolines by FT-IR spectroscopy. *Analytical Chemistry*, 68(1):23–30, 1996.

[44] G. Buttner. The use of NIR analysis for refineries. *Process control and quality*, 9(4):197–203, 1997.

[45] I. LitaniBarzilai, I. Sela, V. Butalov, I. Zilberman, and I. Schechter. On-line remote prediction of gasoline properties by combined optical methods. *Analytica Chimica Acta*, 339(1–2):193–199, 1997.

[46] S. B. Brown. *An introduction to spectroscopy for biochemists.* Academic Press, London, 1980.

[47] F. S. Parker. *Application of infrared spectroscopy in biochemistry, biology, and medicine.* Plenum Publishing, New York, second edition, 1984.

[48] J. Courtial, B. A. Patterson, , W. Hirst, A. R. Harvey, A. J. Duncan, W. Sibbett, and M. J. Padgett. Static Fourier transform ultraviolet spectrometer for gas detection. *Appl. Optics*, 36(13):2813–2817, 1997.

[49] S. A. Reid and S. Gillespie. *Open path detection of hydrogen sulphide.* International Symposium on optical remote sensing for environmental monitoring. Atlanta, Georgia, 1993.

[50] B. A. Patterson, J. P. Lenney, W. Sibbett, B. Hirst, N. K. Hedges, and M. J. Padgett. Detection of benzene and other gases with an open path, static Fourier-transform UV spectrometer. *Applied Optics*, 37(15):3172–3175, 1998.

[51] B. W. Petley. *The fundamental physical constants and the frontiers of measurement.* Hilger, London, 1985.

[52] *Handbook of chemistry and physics.* CRC Press, Cleveland, Ohio, Published annually.

[53] M-L. Junttila, B. Ståhlberg, E. Kyrö, T. Veijola, and J. Kauppinen. Fourier transform wavemeter. *Rev. Sci. Instrum.*, 58(7):1180 – 1184, 1987.

[54] J. Y. Koo and I. Akamatsu. A simple real-time wavemeter for pulsed lasers. *Meas. Sci. Technol.*, 2:54–58, 1991.

[55] M. Born and E. Wolf. *Principles of optics.* Pergamon, Oxford, 1975.

[56] M. B. Morris, T. J. McIlrath, and J. J. Snyder. Fizeau wavemeter for pulsed laser wavelength measurement. *Applied Optics*, 23(21):3862–3868, 1984.

[57] C. Reiser and R. B. Lopert. Laser wavemeter with solid fizeau wedge interferometer. *Applied Optics*, 27(17):3656–3660, 1988.

[58] B. Faust and L. Klynning. Low-cost wavemeter with a solid fizeau interferometer and fibre-optic input. *Applied Optics*, 30(36):5254–5259, 1991.

[59] M-L. Junttila and B. Ståhlberg. Laser wavelength measurement with a fourier transform wavemeter. *Appl. Opt.*, 29:3510–3516, 1990.

[60] X. Q. Jiang, J. Kemp, Y. N. Ning, A. W. Palmer, and K. T. V. Grattan. High-accuracy wavelength-change measurement system based on a wollaston interferometer, incorporating a self-referencing scheme. *Appl. Opt.*, 36:4907–4912, 1997.

[61] D. Steers, W. Sibbett, and M. J. Padgett. Dual-purpose, compact spectrometer and fibre-coupled laser wavemeter based on a wollaston prism. *Appl. Opt.*, 37(4):5777–5781, 1998.

[62] 1100 series rs-170 monochrome camera. Cohu, Inc. Electronics Division, 5755 Kearny Villa Road, San Diego, CA.

[63] J. E. Jackson. *A user's guide to Principal components.* Wiley-Interscience, Chichester, 1991.

[64] I. T. Jolliffe. *Principal component analysis.* Springer-Verlag, Berlin, 1986.

[65] B. Flury and H. Riedwyl. *Multivariate statistics: a practical approach.* Chapman and Hall, London, 1988.

[66] G. A. Marcoulides and S. L. Hershberger. *Multivariate statistical methods: A first course.* Lawrence Erlbaum Associates, Mahwah, New Jersey, 1997.