Demonstration and Investigation of the Spatial Heterodyne Spectroscopy Technique

Sarah M. Eddes

University of New South Wales at the Australian Defence Force Academy

Spatial Heterodyne Spectroscopy (SHS) is a developing interference spectroscopy technique which provides the Fourier transform of a selected spectral interval centred on a specified wavelength, without the need for scanning. The technique delivers high spectral resolution results with large etendue in a compact and robust instrument, thereby providing a wider range of practical applications than conventional interferometers. This thesis describes the construction and implementation of a spatial heterodyne spectrometer, and culminates in the analysis of successful outputs. The process of determining the parameters of the system from the interferograms obtained is demonstrated. The accuracy of outputs is examined with respect to potential errors emanating from processing techniques and noise within the system. The effectiveness of the spectrometer which was constructed is compared with the theoretical range of a true SHS system with the same controls, and found to be inferior due to the conditions under which it is trialled. Recommendations are made to eliminate or reduce these differences. In addition, proposals are made in respect of the successful use of a SHS spectrometer in the molecular fingerprint region. This could be achieved by the use of advanced apparatus and incorporating the all-reflection SHS configuration.

Contents

I. Introduction ........................................................................................................................................... 3
   A. General Terminology .......................................................................................................................... 3
   B. Motivation .......................................................................................................................................... 3
   C. Aim ..................................................................................................................................................... 3

II. Background .......................................................................................................................................... 4
   A. Theory of Spatial Heterodyne Spectroscopy .................................................................................... 4
   B. Evaluation of Spatial Heterodyne Spectroscopy Technique ............................................................. 5
   C. Practical Uses of Spatial Heterodyne Spectroscopy ......................................................................... 5

III. Construction ....................................................................................................................................... 6
   A. Experimental Preconditions ................................................................................................................ 7
   B. Apparatus .......................................................................................................................................... 7
   C. Construction Considerations ............................................................................................................. 7
   D. Outputs ........................................................................................................................................... 8

IV. Post Processing Analysis ................................................................................................................... 8
   A. Central Process ................................................................................................................................ 8
   B. One-Dimensional vs. Two-Dimensional Analysis .......................................................................... 9
   C. Precision of Peak Positions .............................................................................................................. 11
   D. Noise Analysis ................................................................................................................................. 13

V. Spectrometer Parameters .................................................................................................................. 16
   A. Theoretical Spectrometer Parameters .............................................................................................. 16

1 SBLT, School of Engineering & Information Technology. ZEIT4500

Final Project Report 2011, UNSW@ADFA
B. Restrictions on Theoretical Parameters of Spectrometer .................................................. 17
C. Practical Parameter Results .................................................................................................. 18
D. Advancement to Fingerprint Region .................................................................................... 20
VI. Conclusions .......................................................................................................................... 21
VII. Recommendations ............................................................................................................... 21
Acknowledgements ...................................................................................................................... 22
References ........................................................................................................................................ 22
APPENDICES
Appendix A. List of Apparatus Used .......................................................................................... A1
Appendix B. Interferogram Outputs ........................................................................................... B1
Appendix C. Fringe Orientation Relationship Between Spatial and Frequency Domain .......... C1
Appendix D. Dark Noise Image and Corresponding FFT ........................................................... D1
Appendix E. Images Taken with Decreasing Exposure Time ....................................................... E1
Appendix F. Method Used to Determine Range of Orders Detectable .......................................... F1

Nomenclature

\[\begin{align*}
SHS & = \text{Spatial Heterodyne Spectroscopy} \\
\lambda & = \text{wavelength (m)} \\
\sigma & = \text{wavenumber (m}^{-1}\text{)} \\
h & = \text{Plank’s constant; } 6.626 \times 10^{-34} (\text{m}^2\text{kg/s}) \\
c & = \text{speed of light; } 299.8 \times 10^6 (\text{m/s}) \\
E & = \text{energy; (J)} \\
L_{i} & = \text{lens} \\
A_{i} & = \text{aperture} \\
G_{i} & = \text{grating} \\
\sigma_{L} & = \text{Littrow wavenumber (m}^{-1}\text{)} \\
\theta_{L} & = \text{Littrow angle (deg)} \\
\gamma & = \text{angle of the output wavefront normal to the optical axis (deg)} \\
m & = \text{order of diffraction} \\
l/d & = \text{groove density (g/mm)} \\
\delta\sigma & = \text{wavenumber difference from Littrow (m}^{-1}\text{)} \\
W & = \text{width of detector (pixels)} \\
f_{x} & = \text{spatial frequency of fringe pattern in x-plane (cycles/m)} \\
B(\sigma) & = \text{spectral density /brightness distribution} \\
I(\sigma) & = \text{intensity distribution} \\
x & = \text{location on detector array in x-plane in reference to centre} \\
\sigma_{i} & = \text{first distinguishable wavenumber from Littrow} \\
R & = \text{resolving power} \\
\Delta\sigma & = \text{minimum wavenumber variance which can be separated from a wavenumber } \sigma \\
W_{G} & = \text{width of gratings illuminated (m)} \\
n & = \text{number of distinguishable output fringes} \\
N & = \text{number of detector elements} \\
\alpha & = \text{angle of grating rotation (deg)} \\
y & = \text{location on detector array in y-plane in reference to centre} \\
f_{y} & = \text{spatial frequency of fringe pattern in y-plane} \\
HeNe & = \text{Helium-Neon} \\
BS & = \text{beam-splitter} \\
TGI & = \text{Twyman-Green Interferometer} \\
CCD & = \text{charge coupled device} \\
MCT & = \text{Mercury Cadmium Telluride}
\end{align*}\]
I. Introduction

A. General Terminology
Interferometry is the broad technique of evaluating the properties of electromagnetic waves by means of their resultant integration and interference. Interference spectroscopy uses optical interferometry to allow a spectral analysis of an incoming light source from an interference fringe pattern. Fourier-transform spectroscopy is particularly relevant to this thesis. This refers to the group of spectrometry techniques which requires the resultant direct data to undergo a Fourier transform in order for its spectral properties to be analysed. (Smith 1996).

In the fields of spectroscopy it is more common to refer to wavelength as wavenumber, where wavenumber is the reciprocal of wavelength as per Eq. (1) below. Here $\lambda$ represents wavelength and $\sigma$ represents wavenumber.

$$\sigma = \frac{1}{\lambda} \quad (1)$$

The SI units for wavenumber are reciprocal meters ($m^{-1}$) which are used in calculations, however, wavenumber in text is usually referenced as reciprocal centimetres ($cm^{-1}$). Both wavelength and wavenumber will be used throughout this document in consideration of this direct relationship. Care must be taken in calculations, however, that the substitution is made early, such that the full relationship exchanges can be accounted for. (Griffiths and de Haseth 2007).

B. Motivation
Recent developments in the availability of highly accurate detector arrays and advanced computer processors have allowed for more advanced spectroscopy techniques to be established, such as Spatial Heterodyne Spectroscopy (SHS) (Englert, Harlander, Cardon, Roesler 2004). This evolving technique introduces several advantages over other mainstream interference spectrometers, such as the Michelson and Fabry-Perot interferometers, owing to its potentially high-spectral-resolution diffuse-source results with simplified and robust equipment (Englert, Harlander, Cardon, Roesler 2004). One such advantage is that the spatial heterodyne spectrometer has no moving parts, allowing the equipment to be more compact, robust and easily cooled. This provides means for the apparatus to be used in more practical situations such as applications in satellites to measure atmospheric winds. Further advantages include the ability to be field widened, a larger etendue, hence larger interferometric throughput, and a multiline capability, whereby multiple emission lines can be measured simultaneously. (Englert, Babcock and Harlander 2007).

Despite obvious advantages, SHS has lacked the same depth of thorough research and growth over the past two decades, in comparison to the fields of spectroscopy and heterodyne spectroscopy, with very few authors publishing all SHS papers released to date. In spite of the limited research, there is much scope for improvement and further development of the SHS technique. These include field hardening and miniaturisation, improving sensitivity, field calibration techniques and the range of wavelengths which the spectrometer can detect. (Englert, Babcock and Harlander 2009). Of particular interest is the technique’s sensitivity within the molecular fingerprint region to accurately detect the presence of molecules.

The molecular fingerprint region is the range of wavenumbers at which infrared light will impart vibrational energy upon most molecules and is considered to be between 1500 $cm^{-1}$ and 500$cm^{-1}$. (Chimento 2008). According to quantum mechanics, wavenumber ($\sigma$) is related to energy ($E$) through the relationship with Planck’s constant ($h$) and the speed of light ($c$) by the equation $E = h\sigma$. When infrared energy interacts with certain molecules, it can be absorbed, causing the chemical bonds in the material to vibrate. The exact amount of energy absorbed is a characteristic of each molecule depending on its structure. Given the relationship between energy absorbed and wavenumber, unknown molecules can therefore be identified if the wavenumber at which they absorb infrared radiation can be established. (George and McIntyre 1987).

The ability of SHS to accurately identify molecules, combined with the aforementioned advantages of the technique, provides potential for the development of robust, accurate molecule detection equipment in a practical environment which is superior to conventional interferometry techniques.

While SHS development has yielded significant results, there is little literature available which outlines the exact experimental data and processes used to obtain these outcomes. Thus, a more thorough understanding of the apparatus, theory, processes and relationships within the system would be of benefit to future development within a localised environment.

C. Aim
The aim of this thesis is to model the SHS technique with the apparatus available in order to demonstrate the associated procedure, and characterise outputs with corresponding wavelength. This includes considerations of accurately performing the technique and evaluating the consequent limitations of the spectrometer.

The research conducted in this project will allow for further UNSW@ADFA project development, in particular, expansion of the model for use with a mid-infrared laser source to measure spectra in the molecular fingerprint region.

II. Background

A. Theory of Spatial Heterodyne Spectroscopy

Similar to the Michelson Interferometer, the SHS technique employs a two beam dispersive interferometer. A beam splitter is used to divide a collimated incoming wavefront into two interferometric arms. Figure 1 illustrates the basic assembly of a SHS spectrometer. Lens L1 is used to collimate the wavefront from aperture A1. The input wavefront is shown by the vertical straight line after L1. SHS differs from the Michelson Interferometer in that each interferometer arm is fixed at a symmetric length and the mirror component at the end of the interferometer arm is replaced by a fixed and angled reflective grating, seen as G1 and G2. These gratings have static angles which are selected to fulfil the Littrow condition for a specific incoming wavenumber \( \sigma_L \) (Harris, et al. 2005).

The Littrow condition refers to the configuration at which specific incoming light of \( \sigma_L \) is diffracted back from the grating in the direction from which it came. The angle of the grating which allows this to occur is referred to as the Littrow angle and follows from the general grating equation:

\[
m \lambda = d(\sin \theta_i + \sin \theta_r)
\]

(Newport Co. 2005) which is modified in Littrow when the angle of incidence is equal to the angle of diffraction to become:

\[
m \lambda = 2d \sin \theta_r
\]

(Newport Co. 2005) where \( m \) is the order of diffraction; \( \lambda \) is the wavelength of the incoming light; \( d \) is the groove spacing of the grating (inverse of the groove density \( 1/d \)); \( \theta_i \) is the angle of incidence and \( \theta_r \) is the angle of diffraction.

Recalling from Eq. (1) that wavenumber is the inverse of wavelength and that in SHS the grating angle remains constant, it follows that for any SHS system there is a specific fundamental wavenumber. This is referred to as the Littrow wavenumber, \( \sigma_L \).

For any wavelength entering the system, \( \sigma \), after having diffracted from the gratings in each interferomic arm, both wavefronts recombine at the beam splitter. These two coherent plane wavefronts are labelled as 1 and 2 in Fig. 1. The angle at which the wavefronts cross after combining at the beam splitter is given by:

\[
2\gamma = 4 \left(1 - \frac{\sigma_L}{\sigma}\right) \tan \theta_L
\]

(Harlander, Reynolds and Roesler 1992) where \( \gamma \) is the angle at which the wavefront is dispersed from the grating normal to the optical axis; \( \sigma \) is the wavenumber of the light entering the system; \( \sigma_L \) is the Littrow wavenumber and \( \theta_L \) is the (Littrow) angle of the gratings corresponding to the wavenumber \( \sigma_L \).

In respect of Fig. 1, \( 2\gamma \) is the angle between wavefronts 1 and 2, whereby wavefront 1 is considered at an angle \( +\gamma \) to the horizontal and wavefront 2 at \( -\gamma \). For the Littrow wavenumber, given the angle of incidence is equal to the angle of diffraction there is no tilt in either wavefront, and thus \( \gamma = 0 \).

The wavefronts consequently interfere with each other and are focussed onto a position sensitive detector array to produce a wavenumber-dependent spatial frequency Fizeau fringe pattern. The generation of this fringe pattern corresponds to the adjusted grating equation whereby:

\[
\sigma \left[ \sin \theta_L + \sin (\theta_L - \gamma) \right] = \frac{m}{d}
\]

(Roesler and Harlander 1993) where \( \sigma \) is the wavenumber of the light entering the system; \( \gamma \) is the angle at which the wavefront is dispersed from the grating normal to the optical axis; \( m \) is the order of diffraction and \( 1/d \) is the groove density of the grating.
These recombining wavefronts are parallel and coincident for only the Littrow wavenumber ($\sigma_L$) and under these circumstances, assuming the system is aligned such that the wavefronts are parallel to the detector, the wavefronts interfere constructively and the Fizeau fringes produce a constant intensity fringe pattern across the detector. (Englert, Harlander, Cardon, Roesler 2004). All wavenumbers ($\sigma$) other than $\sigma_L$, however, will have their wavefronts disperse at a different angle from the gratings in accordance with Eq.(5), rotate in opposite directions and cross after the beam-splitter to alter the resulting fringe pattern. If the difference between a wavenumber and the Littrow wavenumber is close, the spatial frequency of the fringes is directly proportional to the wavenumber difference $\delta \sigma$, where $\delta \sigma = \sigma_L - \sigma$. (Englert, Harlander, Cardon, Roesler 2004). This is illustrated in Fig. 2, where the relationship between wavenumber, wavefront and intensity of the fringe pattern is demonstrated.

Presuming magnification from the collimating lenses is at unity for the following theory, the spatial frequency of the resultant Fizeau fringe pattern at the detector is given by:

$$f_x = 2\sigma \sin \gamma \approx 4(\sigma - \sigma_L)\tan \theta_L$$  (6)

(Harlander, Reynolds and Roesler 1992)

The spectral density for the incident radiation, $B(\sigma)$, can be informally considered as the brightness distribution of the fringe pattern. The intensity distribution $I(\sigma)$ of the fringe pattern as a function of the position, $x$, on the detector is given by the equation below:

$$I(x) = \int B(\sigma) \cdot (1 + \cos(2\pi(\sigma - \sigma_L)x \tan \theta_L)) d\sigma$$  (7)

(Harris, et al. 2005), where $x$ is the location on the detector array referenced to the centre, as per Fig.1.

The input spectrum is derived by taking the Fourier transform of $I(x)$. This may be physically achieved from analysis of the Fourier domain image.

B. Evaluation of Spatial Heterodyne Spectroscopy Technique

An advantage of the SHS technique is that there is no requirement for any element to be mechanically scanned and no dynamic components in its construction, thereby allowing for more robust and compact equipment. A successful result is still made possible given that, unlike the established Fourier Transform Spectroscopy method, zero spatial frequency does not correspond to zero wavenumber, but to the Littrow wavenumber, which may be selected. (Roesler and Harlander 1993). This allows for the SHS technique to record the ‘Fourier transform of the spectrum on a position-sensitive detector (simultaneously) without scanning and (to) heterodyne the interferogram with a frequency corresponding to the Littrow wavenumber of the gratings.’ (Harlander, Reynolds and Roesler 1992) It is effectively the gratings and the wavefronts’ interaction at the output of the beamsplitter which heterodyne the spatial frequencies in relation to the Littrow wavenumber, thereby reducing the detected fringe frequency but maintaining the same level of high resolution and throughput as other techniques such as the Michelson and Fabry-Perot Interferometers. (Harlander, Reynolds and Roesler 1992).

The passband and resolution potential of SHS is also an important element to consider given the restriction on the resolution of the imaging detector. This must be fully explained in conjunction with the resolving power, $R$, which is usually a dimensionless quantity and gives an indication as to the ability of the technique to illuminate and differentiate adjacent spectral lines. By definition:

$$R \equiv \frac{\lambda}{\Delta \lambda} \equiv \frac{\Delta \sigma}{\sigma}$$  (8)

(Harris, et al. 2005) where $\Delta \sigma$ is the minimum variance in wavenumbers, which can be separated from a wavenumber $\sigma$. For the SHS technique, the resolving power may be expressed as:

$$R = 4 W_G \sigma \sin \theta_L$$  (9)

(Roesler and Harlander 1993) where $W_G$ is the width of the illuminated gratings.

Using the definition of $R$ in Eq. (8) and the derived resolving power of the system in Eq. (9), a single resolved spectral element may be expressed as:

$$\Delta \sigma = 4 W_G \sigma^2 \sin \theta_L$$  (10)
The number of resolvable fringes is related to the size of the imaging detector. The highest spatial frequency achievable is restricted by the resolution of the imaging detector elements, given the constraint of the Nyquist frequency to avoid aliasing. ‘For N detector elements...N/2 spectral elements may be recovered.’ (Harlander, Reynolds and Roesler 1992). Given the SHS technique allows for the central (Littrow) wavenumber to be selected and produces a heterodyned interferogram with respect to the Littrow condition, high resolution is possible with comparatively few samples. This high resolution, however, restricts the results to a narrow spectral range. (Harlander, Reynolds and Roesler 1992 & Harris, et al. 2005). Care must thus be taken that the resolution is kept low enough so that a usable passband is achieved. (Dawson and Harris 2008).

The passband may be improved by a factor or two without aliasing, by introducing the two-dimensional SHS technique. (Harris, et al. 2005). This is achieved by rotating one of the original gratings around the plane of interference by a small angle, α. This discontinues the symmetry of the grating such that it is no longer symmetrical around τi where (τ-τi) and -(τ-τi) are identical. Rather, the fringes are orientated in the detector plane, which is considered x-y plane, rather than being perpendicular to the x-axis as in the one-dimensional format. (Harlander, Reynolds and Roesler 1992). This allows a ‘nominally wavelength independent spatial modulation in the y direction as well as the strongly wavelength-dependent modulation in the x direction’. (Roesler and Harlander 1993). Wavenumbers greater than τi are rotated counter-clockwise. (Roesler and Harlander 1993).

This result in the y-dimension is constrained to low resolution because the spatial frequencies are not heterodyned in the y-dimension using this technique. High resolution in the x-dimension, however, is still maintained. (Harlander, Reynolds and Roesler 1992). It follows that a two-dimensional Fourier transform of the result is required to produce the input spectrum displaying the separated spectra.

This two-dimensional format can also be used in conjunction with the gratings being used in multiple orders, to further increase the spectral range of the spectrometer, as per the echelle spectrogram technique. In this SHS configuration, it can be considered that there are multiple Littrow wavenumbers for an individual Littrow grating angle, given the possible range of orders. Equation (11) below is derived from Eq. (3) to produce:

$$\sigma_{\text{Lm}} = \frac{m}{2d \sin \theta_L}$$  \hspace{1cm} (11)

(Harlander, Reynolds and Roesler 1992), whereby $\sigma_{\text{Lm}}$ is the Littrow wavenumber for a particular order; m is the order of diffraction; d is the groove spacing of the gratings and $\theta_L$ is the Littrow angle of the gratings. To avoid aliasing and adequate resolution of fringes in both dimensions, the number of spectral elements within a particular order is restricted to the number of pixels in the x-dimension of the detector and the number of orders is restricted to the number of pixels in the y-dimension divided by four. (Roesler and Harlander 1993).

The Fourier transform output of the two-dimensional SHS technique is shown in Fig. 3, where it can be seen that frequency is read from the x-dimension and order is read from the y-dimension. This two-dimensional technique also allows for uncertainty in apparently present wavelengths (referred to as ghost spectra) within a particular order to be resolved, given the dispersion in the y-dimension, and is therefore also useful in single-order analysis. (Roesler and Harlander 1993).

**Figure 3. Two-Dimensional Fourier Transform of an Arbitrary Fringe Pattern.** (Roesler and Harlander 1993).

**C. Practical Uses of Spatial Heterodyne Spectroscopy**

Unlike other instruments of similar resolving power, construction of a SHS system is highly compact and robust, given it has no moving parts and has no requirement for mechanical scanning elements. This, combined with its exceptional throughput capability, ability to be field-widened and ‘tolerance of optical imperfection and misalignment’ (Lawler, et al. 2009), allows SHS instruments to be used in a wide range of practical applications. To date, research of applications include a high resolution version for hydrogen Lyman-alpha line profile measurements in planetary atmospheres, UV field-widened version for measurement of the diffuse galactic emission line background, visible field-widened version for ground-based aeronomy (Roesler and Harlander 1993), high spectral resolution space-based remote sensing (J. Harlander 2003), Doppler-shift measurements of atmospheric emission lines (Englert, Babcock and Harlander 2007), remote sensing of diffuse UV-visible emission line sources in the solar system (Harris, et al. 2005), examination of diffuse interstellar emission lines at far-ultraviolet wavelengths (Harlander, Reynolds and Roesler 1992) and long wave IR sensing (Englert, Babcock and Harlander 2009).
Limitations on the worth of the SHS system, however, are based on several characteristics, which include spectral resolution, resolving power, throughput and sensitivity. The constraints of these sources have been the basis of much research on the technique over the past decade, and continue to be a field of interest for improvement.

III. Construction

A. Experimental Preconditions

One type of laser was available for the duration of this project. To adjust for this whilst still allowing an analysis of the technique, the traditional experimental configuration was altered, such that a range of wavelengths entering the system was simulated. This was achieved by considering the relationship between the angle of the gratings and wavenumber entering the system as per Eqn’s 3-6. Explicitly, in traditional SHS the angle of the gratings remains constant whilst the input wavelength changes. Given the relationship between grating angles and input wavelength, each wavefront will diffract off the gratings at a different angle, depending on its wavelength, and hence produce a change in the Fizeau fringe pattern. This can be simulated if the wavelength is constant by slightly changing the angle of the gratings.

B. Apparatus, Configuration and Procedure

The apparatus used is listed in Appendix A. Whilst the fundamental equipment was able to be sourced, some of the minor supplementary equipment was found to be less than ideal for accurate construction of the spectrometer. One of the more prominent issues was the lack of variety in stands and poles to adjust the height of the apparatus. Where the correct stands were unavailable, such as for the collimator, other objects such as textbooks and magazines were used to compensate for the height adjustment. This was undesirable in respect of possible vibration affects albeit, given the considerable weight of the collimator and the solidness of the makeshift stands, this was considered acceptable. Where available, magnetized stands were able to be obtained and hence positioning of the apparatus was made relatively easy for system alignment.

The detector obtained was a PULNiX TM-1020 series, which was a high resolution, high speed 1 inch charge couple device (CCD) camera. This permitted full vertical and horizontal resolution with exceptionally high shutter speed between 1/60th second 1/16000th of a second. It had a resolution of 1008 by 1018 pixels, which were 9μm square, allowing full dimensional analysis to be obtained. The associated software facilitated manual adjustment of the camera settings such as shutter rate and gain between exposures.

The apparatus was set up as per Fig. 4 in accordance with theory. Initially, the experiment was conducted with tilted mirrors placed at the ends of the interferometric arms rather than the gratings, to allow for ease of alignment. Once acceptable results were obtained, these mirrors were replaced with the gratings to allow full analysis of the pure SHS technique.

C. Construction Considerations

In selecting and positioning equipment, consideration was given to the effect of vibrations which propagated from the base of the apparatus stands and maximised any free movement within the system. To minimise these effects, the equipment was set as low as possible and ensured to be tightly screwed to the laser table. The greatest concern with vibrations occurred when images were attempted to be taken, where any movement in the room caused visual disruption in the image. Care was taken to capture images when all movement had ceased.

Limiting light entering the system other than from the source was an important consideration. The theory outlined in Chapter II only covers the situation for an ideal monochromatic source. If multiple wavenumbers enter the system they too are dispersed from the gratings at their relative angles and contribute to the fringe pattern. By itself, a small change between input wavenumbers would ordinarily cause a slight shift in the pattern, rather than a large change in the wavenumber which would alter the actual number of fringes which appear. Thus, several wavenumbers entering the system simultaneously within a small range of the true input wavenumber would likely blur or widen the fringes slightly. This is given the resolution and resolving power of the spectrometer is sufficient to differentiate between wavenumbers of this close proximity. If many

Figure 4. Apparatus Configuration for SHS Helium Neon laser at left, enters spatial filter and collimator. Output of collimator enters beamsplitter (BS) which splits light to gratings G1 and G2. Diffracted light from the gratings returns to BS where directed towards detector at bottom right of image.

Final Project Report 2011, UNSW@ADFA
wavenumbers enter the system, such as with white light, many overlapping and interfering wavefronts are produced across the detector. Although less intense than the main input source, these wavefronts effectively fade the overall pattern. This essentially becomes noise across the detector and relatively lowers the signal strength. This was avoided by ensuring that the room was made as dark as possible while the exposures were being taken and that the lights in the room were extinguished. Altering the shutter speed and gain of the detector was additionally useful in revealing better quality fringe patterns. By variously adjusting these two parameters, over-saturation could be avoided in high light situations and low intensity fringes could be distinguished in low light situations.

Integrity of the apparatus was also considered a prominent concept. As in any optical system, high accuracy between components was necessary, given the system’s reliance on very small angle changes and the relative interference between wavefronts. Any small misalignments or unwanted reflections such as dust, fingerprints and reflections between components can cause unmeasurable effects throughout the system. Whilst the former are relatively easy control, reflections within the system were much harder to regulate.

There were observed reflections between the components of the collimator, between the laser and the collimator and between the beam-splitter and the collimator. These were reduced as much as possible through adjustment and alignment of the components, but not all could be avoided. To reduce these reflections further, additional lens components could be added; however, as the risk of causing more interference and attenuation of the signal was high, this course was not pursued. Additionally, reflection error was introduced to the system through the camera itself. Unfortunately, the nature of the camera’s construction causes reflections between its internal components, which also interfere with and distort the incoming fringe intensity pattern. Other than ensuring the detector was aligned as close as possible with the correct axis, these reflections cannot be avoided.

D. Outputs
Upon completion of the construction phase, multiple fringe pattern exposures were successfully taken for the SHS technique. In all instances, as the angle of the gratings was altered, the number and rotation of the fringes in the interferogram altered accordingly. The gratings allowed two planes of movement, a combination of which were trialled to obtain a wide range of results. To fully analyse the system, a series of several experiments were conducted. These included obtaining the largest and smallest fringe frequency outputs, conducting noise analysis, testing the resolving power and capturing several orders diffracted from the gratings. Several outputs, including the specific experimental results mentioned above, are shown in Appendix B.

IV. Post Processing Analysis

A. Central Process
The most fundamental post-processing procedure required for SHS is to find the fringe frequency of the output interferogram. The fringe frequency may then be entered into Eq. (6) to solve for the input wavenumber of the system if the Littrow angle and wavenumber are known. Similarly, if the input wavenumber and the fringe frequency are known, Eqs. (3-6) may be manipulated to find the relative Littrow wavenumber and Littrow grating angle for the calculated fringe frequency of a given SHS system. This is possible on the assumption that the spectrometer was constructed using the traditional technique whereby the gratings have common angles.

**Figure 5.(a) Example Detected Output Image**

**Figure 5.(b) Scaled Log Magnitude of Detected Output Image**
An example of the output of the interferometer is shown in Fig. 5(a). As observed, there are relatively clear fringes visible with only minor discontinuities affecting their intensity in remote areas of the image. This is common amongst the interferograms taken, as can be observed in Appendix B. These small discrepancies do not affect the ability to determine the fringe frequency although they do contribute some degree of noise. All post-processing of the image was conducted using the computer program Matlab.

Initially, the fast Fourier transform of the image was evaluated and its scaled log-magnitude shifted version displayed as per Fig. 5(b). Through careful analysis of this image alone, many of the system’s parameters could be evaluated. The most important features of the image are the two bright peaks located diagonally left above and diagonally right below the centre point of the image. These are the greatest indicators of the fringe orientation and frequency in the spatial domain. Their prominence and relative magnitude to the brightest central pixel is clearer when viewed in 3D format, as is shown in Fig. 6. The axis displayed provides an indication of the distance of the peaks in the x and y dimensions from the centre. The relationship between the angle and number of fringes in the spatial domain, and the angle and distance of the bright peaks from the centre in the frequency domain is highlighted in Appendix C. This contains several interferograms captured in the spatial domain with their associated frequency domain images. As the number of fringes becomes larger, the distance of the bright spots from the centre also becomes larger. Likewise, the angle of the bright peaks relative to the centre indicates the orientation of the fringes. Specifically, the angle of the bright peaks from the centre may be rotated 90 degrees clockwise to determine the angle of the fringes in the spatial domain.

The distance of the bright peaks from the centre may be used to determine the spatial frequency of the fringes. Determination of this distance is found (in pixels) using the position of the peaks in x- and y-coordinates. The number of pixels along a line which is mapped across the image joining the centre and two brightest peaks is then found. This line is illustrated in Fig. 7 where the centre and two brightest peaks are marked with an ‘X’ and the adjoining line highlighted in blue. The ratio of the peak distance from the centre and the length of this line produces a value for cycles per pixel. This pixel value may then be converted to distance in meters with the given pixel dimensions of the detector. In this instance, each pixel is 9μm by 9μm. This allows for a value of cycles per meter to be established. This is the spatial frequency of the Fizeau fringes of the interferogram, the value of which may then be entered into Eq. (6). This is the vital information required to solve for the wavenumber entering the system or the Littrow angle and wavenumber variables. This process is confirmed with the one-dimensional analysis outlined in the next section.

In terms of post-processing, it is therefore essential to find the accurate coordinates of these two bright peaks. Simple trigonometry can then be used to find the distance of the bright spots from the centre pixel and the number of pixels along their adjoining line. The process to find the fringe frequency can thus be applied.

B. One-Dimensional vs. Two-Dimensional Analysis

Section A outlines the general process of finding the fringe frequency purely from information gained in the Fourier domain. This process is accurate for both one and two-dimensional systems, since the bright peak pattern is constant for any correct linear fringe pattern observed. Despite this method being sufficient, there is a simplified process for finding the fringe pattern in a one-dimensional system. In the spatial domain image, the pixel values along a detector line perpendicular to the fringe’s main orientation must be taken. When these values are plotted the result is sinusoidal. The Fourier transform of this sinusoidal wave effectively determines the number of fringes across the image. This value may then be converted to fringe frequency by considering the distance over which that number
of fringes was spread. In this instance, the distance is the length of the line taken to determine the sinusoidal values, once converted into meters.

Since the fringes are not exactly horizontal or vertical the line which is used to take the values must be rotated such that it is perpendicular to the fringes. To achieve this, the angle of the fringes must first be found, which is best achieved through the Fourier domain. The angle of the bright spots relative to the centre vertical axis is 90 degrees anti-clockwise from the true angle of the fringes. Finding the angle between the bright spots and the centre can therefore be used to determine the angle of the fringes. More conveniently, the angle between the bright spots is the angle at which the line needs to be designated in order to be perpendicular to the fringes. This is the line shown in Fig 7.

A line may therefore be drawn in the spatial domain image and the values of pixels along it used to plot the one-dimensional sinusoidal pattern of the fringes. This line is shown in Fig. 8 with respect to the fringes. Trials in conducting this method confirmed that intervals of less than one pixel should be used to increase accuracy for the technique. The line values are then plotted in the time domain and their Fourier transform derived. This is shown in Fig. 9 and Fig 10 respectively. The number of fringes in the spatial domain may then be read off the frequency axis of Fourier domain image in Fig. 10.

The small error related to this method is the conversion between the frequency and spatial domains in ensuring that the angle obtained for the line is correct, and the discrete samples taken will not always give accurate values which truly depict the values along that path. This is why small sub-pixel intervals should be taken. For fringes which are very close to vertical or horizontal alignment, it could be considered that taking the perpendicular row or column would be of equivalent error. For example, since the fringes in Fig. 5 are close to horizontal, one detector column could be taken rather than creating the line. In this instance, the output only produced a difference of less than a tenth of a fringe. If the angle was substantial, however, this abbreviated method has potential to exclude entire or multiple fringes from the count, which is less than ideal in an accurate system. The general method should therefore be based on an angled line.

The fundamental problem with this method is that it cannot be used reliably when there is more than one set of fringes present on the interferogram, particularly when they are of different frequency and orientation. Multiple fringes would occur if several similar wavelengths entered the system from the same source, or multiple orders were overlapping or close enough to be captured on the same image. The latter is of greater concern in regards to this technique. An example of close proximity orders produced in the SHS system can be seen in Fig. 11. The fringes in the top half of the image are in a different orientation than fringes in the bottom half of the image. When converted to the Fourier domain, as in Fig 12, multiple bright peaks occur according to the orientation of each of the fringe patterns present. These are represented by the groups of blue crosses plotted in the Fourier domain image. Ambiguity then comes in defining the orientation of the values-line required, which is based on the dominant fringe. Whilst the partial values of both fringes will be detected and displayed on the one-dimensional Fourier output, there is potential for significant error in the output for the less dominant fringe, given alignment of the line is based around the more dominant fringe.
Two-dimensional analysis is therefore preferred because it accommodates this type of variation in results. Bright peaks can easily be recognised, and their distance from the centre, and their angle, can be accounted for on an individual basis. Throughout this project, however, both methods were used and the results found to match consistently. Both techniques rely on the Fourier domain analysis and, therefore, by further investigating the accuracy of the two-dimensional method, the accuracy of the one-dimensional method was also be informally incorporated. In particular, both methods rely on the accuracy of the bright peak locations.

C. Precision of Peak Positions

It is important that the bright peak locations obtained are accurate so that precise fringe frequency may be achieved. Given the discrete nature of the image detector, it cannot be assumed that the centre of the brightest pixel is the location of the true brightest intensity point of the incoming light. Consider if a light beam was to enter the detector at the exact intersection of four pixels. Given that the detector’s measurement capacity is not a continuous function and cannot measure light between pixels, the intensity of the entering light would be spread to all surrounding pixels. This means that the magnitude of any given pixel is actually an accumulation of the light which directly hits the pixel and the intensity of light spread to it from adjoining edges. This also means that the centre of the brightest pixel is not necessarily the brightest point of input intensity or indeed within the brightest pixel at all. This logic may be adapted to the Fourier domain situation of SHS, whereby the centre of the brightest pixel found may not be an accurate representation of the true point of brightest intensity. Any error in this measurement would therefore introduce error into the calculations for input wavenumber, due to the relationship of the pixel’s distance from the centre of the image with fringe frequency and hence calculation of the input wavenumber.

Similarly, if the location of the brightest point of intensity could be found within a fraction of a pixel, this could improve the accuracy and range of the overall system. Furthermore, if the error in measurement can be estimated in terms of the standard deviation or other such means, then this can give an indication of the error in determining the fringe frequency. This then transfers through the relevant equations to give an error estimate of the input wavenumber. This of course would have to be considered in conjunction with the resolving power of the system, which also incorporates a wavelength range measurement, given other physical system parameters.

The method of finding the intensity of the brightest pixel from previously must therefore be adjusted accordingly. At present, it has been described as the centre of the brightest pixel. This can now be considered an estimate of the region of brightest intensity rather than a set peak value. Careful observation of the brightest peaks in the Fourier domain image (in Fig. 5(b)) and its 3D version (in Fig. 6) clearly shows that the surrounding pixels are of similar height to the brightest peaks. If the right hand area of the brightest peak were to be brighter than the left hand area, for instance, the true brightest point of intensity may be further to the right from the centre of the brightest pixel.

The method used to analyse this was to define an area around the brightest peak and to calculate its intensity-weighted centroid, thereby incorporating the magnitude of each pixel within the defined area. The function of an intensity-weighted centroid may be informally described as a means of finding the ‘centre of mass’ of the defined area in terms of its intensity. The output of the intensity-weighted centroid was the x and y location of this point within the specified area. This coordinates could then be found in terms of the image to give an overall location of the brightest peak to an accuracy of one tenth of a pixel.

A trial was originally conducted, such that a single area was defined with the brightest peak in the centre. This was a 9 by 9 pixel box and slightly displaced the apparent brightest peak by a fraction of a pixel. It was found, however, that if a slightly different box was defined in terms of position or size then the centroid would move quite significantly given the range of pixel magnitudes which surrounded it. It was considered that in an ideal situation the brightest peak would stand alone and would have very low magnitude pixels surrounding it. In this situation, any position of the box used to find the centroid would still produce the same overall result and locate the same bright spot unambiguously within the total image.

It was consequently decided that a single centroid output could not be used in itself. A more accurate method was devised which allowed multiple centroids to be found from constantly moving box locations. The size of the box was held constant and the location moved randomly, always ensuring, however, that the brightest pixel was located inside the area of

Figure 13. The centroid of each box is found and plotted with a green cross. Average centroid position is marked with a purple circle.
the box. Initially, the perimeter and consequent centroid of each box was plotted as per Fig. 13 to display the effectiveness of the technique and the range of intensity peaks found. The black boxes in Fig. 13 indicate the area in which the centroid was found and the green crosses indicate the location of the centroid with respect to the overall image for each of the boxes. The blue cross indicates the centre of the brightest pixel. The location of the average centroid was consequently calculated, which provided a better indication of the brightest point. This is indicated by the purple circle in Fig. 13. In this instance, a box with dimensions 15 pixels by 15 pixels was used and its position reassigned 20 times.

The entire box iteration process described above was subsequently run a further 25 times to gain an indication of the range of average bright peaks found. This is shown in Fig. 14. The box perimeters are not displayed and the green crosses again indicate the position of the centroid calculated from each individual box. As can be seen, they are in a reasonably well-defined square surrounding the middle pixel. The purple circles are the average points of each iteration of 20 boxes, mainly within a pixel of the brightest pixel. The black square shown is the average of these 20 averages and is within the brightest pixel as expected. This output would be more accurate still if the purple outlier on the far left was removed from the calculation. This black square is the most accurate position of the brightest peak which can be reasonably found. Of course, this process may be repeated several times, however, at the inconvenient cost of processing time and power.

The ideal procedure for any given Fourier domain image would be the initial process of moving the box the initial 20 iterations and averaging the centroid positions within the image to give one point. The error of this final location in relation to repeating the procedure a further 25 times may be avoided if the final output is given with a potential error margin. To accomplish this, the standard deviation of the points was found with respect to the black square location calculated previously, as per Fig. 14. The standard deviation in terms of distance between points was thus found to be 0.29 for the top peak and 0.23 for the bottom peak for a given trial. Analysis of the method and the variation in results between peaks suggests that this method cannot be defined as strictly accurate, as the standard deviation may only be calculated based on an averaged average peak location. Another indication of error in the technique was that the bright peaks found, using this centroid method, were not exactly the same distance from the centre as one another. According to theory, both peaks should be equal in magnitude, and equidistant and symmetrical around the central pixel. This was confirmed manually between the brightest pixels, but not true for the final precise bright peaks found.

This method was also conducted with respect to the central pixel and led to a more accurate estimate of standard deviation. It was determined that the magnitude of all of the pixels in the vicinity of the central pixel were symmetrical about the central pixel to a range of at least 20 pixels. Given this symmetry, the location of the true brightest peak in this region would in fact be the centre of the central pixel. When the centroid method was conducted for the centre pixel, however, there was a slight offset in even the averaged average value as seen in Fig. 15 where the black square again indicates the final value for the bright peak using the method discussed, and the blue cross is the centre of the centre pixel. Given that the true value is known in this instance, a more accurate standard deviation value was thus calculated and found to be 0.25. The fact that the bright peak was not directly on the centre point indicated that there was some slight error present in the methodology, which was expected, given that the standard deviation was based on an averaged value and the boxes selected did not follow a symmetrical pattern around the centre point.

A final concept to consider from this analysis is the potential to extend the measurable fringe frequency of the detector. The standard deviation calculated using this method indicates that there is potential to define the accuracy of the brightest peak to less than a pixel inclusive of error, which is the necessary condition for expansion of the brightness peak range, and therefore by relation, expansion of the fringe frequency and
detectable wavenumber range or resolution. The method of pinpointing the brightest peak and its standard deviation must be improved or consolidated before this can accurately be assessed with confidence, given the sources of error exposed through the analysis. This factor must also be considered in terms of the resolving power of the system since both provide an indication of error from a given measurement. The combination of these would therefore determine if there was potential in this extension aspect of the analysis.

D. Noise Analysis
The signal to noise ratio within the system was considered to be relatively high in that all fringes could be well defined and there were no significant limitations on the post analysis of the interferograms analysis due overarching noise. Use of the spectrometer after nightfall substantially reduced the amount of background noise entering the system, such that when images were taken with the laser switched off there was no apparent signal captured. Conversion of these dark noise images into the Fourier domain, however, indicated that despite the background signal being low, noise contributed to the low frequency region of the overall signal output. This is seen in Appendix D where the Fourier domain of the dark noise signal is displayed and a low magnitude ‘cloud’ is present around the centre of the image.

Despite this, fringes were able to be well defined in most cases with manipulation of the detector’s shutter and gain features. As can be observed from the variety of fringe patterns in Appendix B, there was some constant detector distortion in the form of transparent circular patterns superimposed on the top of the fringes. This was likely due to dust on one or more of the elements. Likewise, when the components were misaligned or the gratings were angled too far, mismatching of the two returning wavefronts occurred, which caused slight disruption in the output fringe pattern. An example of this is seen on the right hand side of Fig. 5(a). Whilst these physical ailments and misalignments within the system caused disturbance in the spatial domain, their effect was relatively minor in the frequency domain.

The quality of the output fringes was initially addressed in the noise analysis of the system. Several attempts were made to reduce the noise within the system due to undefined fringes, despite obtaining sufficient results for an accurate analysis. Many of these attempts, however, were later proven to be ineffectual, unnecessarily
intrusive or indeed damaging to results. The intention of these noise-reducing measures was to effectively lower the overall noise floor and thus produce more prominent bright peaks in the Fourier domain, thereby displaying sharper fringes in the spatial domain. This could provide more accuracy within the post-processing analysis process. A band-pass angular filter and a Gaussian point filter were separately implemented on the interferogram in Fig 5(a). The Fourier and spatial results of these are shown in Fig’s 16 and 17 respectively. In both instances, it can be seen that there is a significant improvement in the clarity of the fringes which depicts the success of the designed filters. Both of these methods of filtering, however, are quite obtrusive and have potential to destroy relevant data. For instance, it was considered that the haziness in these fringes in the spatial domain could in fact suggest the presence of secondary fringes overlaid and slightly offset from the first. This would cause secondary (or numerous) sets of bright peaks in the Fourier domain, of varying magnitude which could be destroyed through filtering. This cannot be avoided in the angular filter method but may be compensated for in the Gaussian filter if multiple points were selected and the width of the Gaussian function was increased. Thus, despite this second filter being more aggressive than the first, passband regions may be chosen more intelligently and thus potentially less important information is lost using this method.

The advantage of both techniques is that the DC component of the signal has effectively been removed, in addition to any noise entering the system from outside the allocated region. The observed fringes are clean and clear along the entire image and the effects of the circular pattern in the top left corner and the darkening on the left side are all but gone. The one-dimensional Fourier transform of the first filtered fringe pattern is seen below in Fig. 18. In comparison with its unfiltered version in Fig.’s 9 and 10 there is a significantly reduced DC input to the system, which also better defines the spectral peaks in the Fourier domain. There was a difference of less than 0.1 fringes between the filtered and unfiltered results, however, suggesting that this method did not overly improve or reduce accuracy.

This point led to a further investigation into the sources and types of noise inherently captured by the detector. The main sources of noise in a system such as this are thermal noise, shot noise and speckle noise. Thermal noise is introduced by random thermal disturbance of electrons in the detector. This is particularly relevant for long exposure times where the elements are susceptible to heating. Alternatively speckle noise is introduced by into the system by random phase variations in the wavefront causing an irregular signal. To investigate the sources and types of noise present in the system the signal to noise ratio was observed under two different conditions. The first of these was with varying exposure time and the second with varying signal power. It was thought from this analysis that an indication of the relationship between noise and signal could be established to determine whether noise was dependant or independent of the signal.

A method of measuring the signal to noise ratio was thus devised. The magnitude of the signal component was taken from the point established using the Centroid method discussed in Chapter IV Section C. The noise component, however, was deliberated. It was originally thought that the magnitude of the brightest peak could be compared to those of the surrounding pixels, given that any lessening in its magnitude would cause the surrounding pixels to become more dominant. It was found, however, that as the magnitude of the intensity of the input signal was lowered, the surrounding high magnitude pixels were lowered in proportion. This suggested that they were in fact artefacts of other close frequency fringes present in the time domain and not noise contributors. The wavenumbers associated with these close frequency fringes were likely to have been from the laser source. Instead, the average noise was taken from four areas within the Fourier domain at the same radius from the centre as the bright peaks, but rotated 60 degrees. This was an area made up of generally speckled signal and was not considered a contributor to the fringe pattern. The constant radius also avoided the effects of roll-off, whereby is was previously observed that noise lessens at the outer perimeters of the Fourier domain image. These four areas are shown in Fig. 19, in the areas surrounded by the black boxes. The signal to noise ratio was then found in terms of amplitude and decibels, where the decibels measurement was used consistently throughout consequent analysis. It was found that when using this method, however, the signal to noise ratio varied widely, depending on the area of the noise.

Figure 18. One-dimensional Fourier analysis conducted on filtered result.

Figure 19. Boxed areas are where noise component was taken.

14
Final Project Report 2011, UNSW@ADFA
measurement selected. This suggested that perhaps this method was not entirely accurate. For the purposes of the following experiments, the area and position of the noise area selected remained constant. This allowed the signal to noise ratio relative to one another between trials to remain accurate, despite the numerical output being questionable.

The first signal to noise experiment was conducted with varying exposure times, where exposure 0 represented the longest shutter opening time and exposure 8 represented the shortest shutter time where signal was still observable in the Fourier domain. The nine spatial domain images produced from this testing are found in Appendix E, and display the range of visualisation between exposure limits.

The signal to noise ratios were then found for each of these exposure instances and plotted as per Fig. 20. It was initially noted that there is an exponential trend between exposure time and signal to noise ratio. The fact that this is not a linear plot indicates in the first instance that both signal and noise vary between measurements. Reading from the right hand side of the graph, as exposure time increases, signal to noise ratio initially increases and then levels out after a certain shutter time. The initial increase in signal to noise ratio shows that the rate of signal increase is faster than the rate of increase in noise as the shutter time increases. This makes sense if the situation is considered practically. If the signal remains constant as shutter time increases, the detector would 'accumulate' the signal in those areas of the detector, thus making the signal stronger. Alternatively, thermal noise could be building in the detector with extended shutter times. If the noise was constant in magnitude and position, it too would become stronger at the same rate as the signal, meaning that the signal to noise ratio would vary very little. This trend then suggests that the noise is not constant and varies in position and magnitude independent of signal. In addition, the levelling of the signal to noise ratio at low shutter speeds, also suggests that noise is not dependent on signal. The source of this noise could then be considered as perhaps random background noise, thermal or detector noise.

The second section of this signal to noise investigation required the power of the signal to be varied. A polariser was obtained and used at varying angles in the beam path to reduce signal power. Several images were taken with the polariser at varying angles and thus variable signal power. The results of these trials are shown in Fig. 21. In general, this shows a generally constant trend between signal power and signal to noise ratio. This would suggest a dependence of noise levels on the power of the signal. There were several errors within this experiment, however, which deemed it untrustworthy, including the inability to obtain appropriate apparatus. Further testing and results will be needed to confirm results. It was considered that the only true comparison which could be made was with the polariser included, and then excluded from the beam path. The results between these two situations showed a greater signal to noise ratio when the glass was removed than when it was inserted in the beam path. This suggested that either the signal dropped while the noise remained the same, or that the signal dropped proportionally more than noise. This may again suggest that noise is independent of signal. More thorough testing, however, would be required to confirm this.

Although the accuracy of both signal to noise tests above was questionable, both suggested that noise was independent of input signal. Methods to reduce this noise were thus considered, given that the filtering methods previously outlined were not able to be justified. In a situation where the noise is random it is generally independent of the signal, is spread across the whole space, and varies between images taken. It can thus be reduced by averaging several images of the same arrangement where the signal remains constant. This means that if random signal, such as noise, is not constantly present between the images it will be reduced in magnitude. A trial was conducted, whereby six images were taken within close proximity to one another and averaged to produce a final composite image. Its Fourier transform and signal to noise ratio was calculated in comparison to one of the original images and found to have only improved by less than 1%. This could suggest several notions. Firstly, if all was ideal and the signal to noise calculation methods had been proven accurate, this result would suggest that there is very little random dynamic noise in the image. The noise present could therefore be constant from the detector, reflections, an outside source or, in fact, be vital artefacts of other
signals present. This could also mean that there was some experimental error present, such as slight vibrations causing the fringes to be misaligned between exposures, thereby increasing noise in the averaged signal section of the composite image which was not present in the first. To reduce the chance of this happening, a small film could have been taken with the detector and the subsequent frames averaged rather than taking individual shots at a greater time difference.

V. Spectrometer Parameters

A. Theoretical Spectrometer Parameters

The apparatus selected for any given SHS spectrometer instils inherent restrictions in its use. The range and accuracy of the wavenumbers able to be defined by the system are described by the characteristics of its parts and their orientation within the system. Thus, if a known range of wavenumbers is required to be detected by the device at an accepted accuracy, appropriate parts must be chosen to fulfil these requirements. Similarly, for a given system with pre-determined equipment, the parameters of the system must be understood in order to allow appropriate selection of the orientation of apparatus for the desired Littrow condition. This project presented the former of these options, whereby the apparatus was chosen prior to analysis and adjusted accordingly. Therefore, parameters such as the potential range of wavelengths able to be detected and the resolution of the system were able to be sought.

The first information sought was the potential wavenumber range of the system. This was found through analysis of the equations from Chapter II with respect to the known variables of the equipment being used. The wavenumber range is primarily based on the number of fringes which the detector can physically resolve and on the Littrow condition selected. The most extreme cases of both of these conditions provide an indication of the theoretical detectable input wavenumber range.

The highest and lowest fringe frequency can be calculated by determining the theoretical maximum and minimum number of fringes possible across the detector. This, in turn, is based on the number of pixels across the detector. Theoretically, the maximum number of distinguishable fringes possible would occur if every second row of pixels was white and every other row of pixels was black. This would mean a spatial period of two pixels from which the fringe frequency can be obtained. The smallest number of cycles that could be reliably detected would be one period across the entire length of the detector’s largest dimension. Combining this information, the range of fringe frequencies theoretically possible is therefore between 109.15 and 55.55×10^7 cycles/meter.

The next set of parameters to be considered is the range of Littrow conditions which can exist with the chosen elements. Any combination of angles and wavenumbers which satisfy the Littrow grating equation for order m and grating of groove spacing d may therefore qualify as a possible Littrow configuration. Since the equipment is constant in the given situation, the variable d is constant and is set at 150000 m^-1. With respect to the grating equation, the next consideration is the number of possible orders which the system can detect. Specifically, “the number of orders which can be recovered unambiguously (is restricted) to N_4/4” (Harlander, Reynolds and Roesler 1992), where N_4 is the number of pixels in the y-dimension of the detector. With the detector being used for this experiment, this value is 252. The final consideration in determining the range of Littrow configurations is the physical restriction of θ_L. For the purposes of this theoretical analysis, the Littrow angle can be set between 0.1 degrees and 89.1 degrees. In reality, however, at extreme angles there would not be sufficient path difference between the grooves of the grating to have a practical effect on the outcome. A simple Matlab program was used to find the range of corresponding Littrow wavenumbers (σ_L) for the theoretical system, the results of which are stored in Table 1.

<table>
<thead>
<tr>
<th>Table 1. Theoretical Parameters for the Given Spectrometer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Maximum</td>
</tr>
<tr>
<td>Minimum</td>
</tr>
</tbody>
</table>

The information accumulated thus far was sufficient to calculate the range of wavenumbers which the spectrometer could theoretically detect. By considering the range of fringe frequencies, orders of diffraction, and Littrow configuration parameters with the relevant equations the most extreme maximum and minimum parameters allowed for the range of possible input wavenumbers to be determined. The information was plotted in a four dimensional plot in terms of both possible input wavenumbers and possible input wavelengths by incorporating Eq. (1) into the relevant calculations. The latter displayed the more visually comprehensive result and is thus shown in Fig. 22.
Each input wavelength range is calculated for each order of diffraction across the range of possibilities, and are displayed in apparent layers in the graph. It can be observed that as the orders increase, the range of possible input wavelengths for each order becomes narrower as each layer decreases in height in comparison to the one before it. It can also be seen that there is a non-linear trend between increasing Littrow angle input wavelength, given that the layers flatten out after a certain point within each layer. This may suggest that gratings placed at too great an angle are ineffective. Solving for the highest and lowest input wavelength and wavenumber produced results tabulated in Table 1.

B. Restrictions on Theoretical Parameters of Spectrometer

The above analysis in Section A outlines the most ideal theoretical range of the spectrometer based only on the gratings used and the size of the detector. There are several restrictions on the potential of the spectrometer, however, which must be considered. These include transmission and detection spectral ranges of the apparatus and the resolution of the system. Furthermore, the defining parameters are further restricted in a true SHS system, given that once equipment and a Littrow configuration have been selected and the device has been built, there is no adjustment allowed in lieu of destroying the ‘no moving parts’ advantage which the technique boasts. This latter point, however, is an accepted restriction and therefore has no requirement to be analysed in these sections.

In terms of the spectrometer overall, the transmission and detection ranges of the equipment must be considered. In this configuration, it was found that the beam-splitter and detector had the most restrictive spectral ranges within the system. The combination of these limited the total range of detectable input wavenumbers of the system to between 700nm and 1100nm. When these limitations were incorporated into the model for input wavenumbers there were obviously significant restrictions in the results. Figure 23 illustrates the total range of scenarios in which input wavelengths can be detected. Orders beyond 20 do not produce results as their wavelength range does not fall within the limits specified by the components of the system. This is a more reasonable expectation from the range of the spectrometer.

Another important factor to consider is the resolution of the system. This may be expressed through the resolving power, which is essentially the ability and accuracy of the system in differentiating between spectral lines. It provides the smallest variance in wavenumbers which can be differentiated from the input wavenumber as described through Eq.’s (8) and (9). In this respect, this variance can also be considered as the potential error in determining the input wavenumber. From a wider perspective, if there is variance in the input wavenumber, there would also be variance in fringe frequency and Littrow angle, given all are related. This can be more easily appreciated with respect to Fig. 22(a) where all variables are integrated on the one plot.

The resolving power again relies on the components of the system. Specifically, the width of the illuminated grating contributes to the resolving power of the system in conjunction with the variables of input wavenumber and Littrow grating angle. With the defined components for
this system, the size of the illuminated grating surface is restricted to a maximum of 2.5cm. The resolving power of the system and the variance in wavenumber can thus be investigated with respect to change in Littrow configuration, input wavelength and varying grating illumination width. These relationships are plotted in Fig.’s 24 and 25. Figure 24 illustrates that the resolving power of the system is sinusoidal with Littrow angle and increases linearly with increasing input wavenumber. This was expected given the structure of the equation from which it is based. The sinusoidal shape can be interpreted with respect the relationship between Littrow number and Littrow wavelength. As Littrow angle changes so too does the Littrow wavenumber. The relative resolving power pattern suggests better resolution with respect to the proximity to Littrow wavelength. The input wavenumber variation plotted in Fig. 25 also illustrates that with increasing angle and increasing input wavenumber, the variation lessens. This stands to reason, given that under these same circumstances, the resolving power is increasing.

As can be observed, the overall shape of the graphs remained the same with varying grating illumination, which was expected, given its variable features as a scaling factor within both relevant equations. The variation in respect of resolving power in Fig. 24 shows that the magnitude of the sinusoid shaped function lowers with lowering illumination width. This is observed through the layers which can be seen within the sinusoidal peaks and troughs. Likewise, the variation in input wavenumber is decreased as illumination width is decreased.

This analysis allows a balance between resolution and bandwidth to be recognised. The relevant equations for resolving power and wavenumber variation are based on variables which allow scaling of the outputs. Thus, if a high resolution is required, the parameters of the system may be adjusted accordingly. A high resolution however, comes at the sacrifice of a lower achievable bandwidth for the given configuration. A balance must then be decided upon with respect to the intended purpose of the instrument.

C. Practical Parameter Results

The analysis undertaken thus far on the parameters of the system has been based on theoretical and equipment based restrictions. Whilst these outcomes are reasonable for an otherwise ideal system, it is of benefit to compare these outcomes with the practical results obtained from the physical system itself. The difference between practical and theoretical results may then give an indication of the environmental, conditional and unmeasurable restrictions on the system such as noise, vibrations and misalignments. Similarly, they do not consider possible advantageous extensions such as the expansion of the detectable fringe frequency range due to sub-pixel accuracy.

The first practical element considered was the maximum and minimum number of fringes which could be produced. By conducting the appropriate computations on each of their relevant images, it was determined that the minimum number of fringes achievable from the practical system was 2.5 and the maximum number of fringes was 288.8. These equated to fringe frequencies of 307.47 and 26.9×10^3 cycles per meter respectively. These parameters were significantly reduced from the theoretical fringe frequency range of 109.15 to 55.55×10^3 cycles per meter.
The next parameter tested was the number of detectable orders. The procedure for determining this range is outlined in Appendix F and successfully ascertained that only the first four orders could be detected. This is a reasonable difference from the theoretical 20 orders expected of a system with the given parameters. Numerous factors are considered to contribute to this depletion, however, including the wavenumber of light entering the system, the imprecise alignment of components, and the less than ideal lighting in the room.

The wavenumber range simulation from the first part of Section B above was thus re-simulated to reflect these further restrictions of order and fringe frequency. For this simulation, input wavenumber and Littrow angles were kept at the ranges stipulated previously. The result is plotted in Fig. 26. As can be seen, a huge restriction exists in the results compared to pure and restricted theoretical simulations, the main restrictors of which are the number of orders and range of fringe frequencies detectable and the spectral ranges of the detector and beam-splitter. The overall result was expected to be quite different to theoretical section in any case, given the less than ideal and inaccurate conditions under which the experiments were held.

The passband from the practical system could also be determined from the results obtained, by assuming the He-Ne laser input was the Littrow wavenumber. This means that the theoretical Littrow grating angle could be derived. This information could then be substituted into the fringe frequency equation with the maximum fringe frequency obtained to derive the range of wavenumbers detectable either side of the Littrow wavenumber. The results of this calculation suggested a bandpass for wavenumbers of $1.58 \times 10^6 \pm 1.43 \times 10^5$ m$^{-1}$. Again, these results must be considered very carefully for their likely accuracy in terms of experimental error. For the theoretical range of fringe frequencies, the bandpass of the system should have been $1.58 \times 10^6 \pm 2.92 \times 10^5$ m$^{-1}$. There is therefore a clear reduction in bandpass, which was due to the conditions of the experiment, as previously discussed. This was considered a reasonable reduction in the prevailing circumstances.

The resolving power of key wavenumbers was also briefly considered. Using Littrow wavenumber and assuming 2.5cm of gratings were being illuminated, the resolving power was calculated to be 7499. This is low for a SHS system and is based on the Littrow configuration parameters used, since the width of the illuminated gratings remained constant throughout the experiments. The resolving power was then further investigated by observing the effects of reducing the width of the illuminated gratings. This was conducted by reducing the area of the light entering the beam-splitter and reaching the gratings. The reduced grating width illuminated was measured to be 7mm from the 25mm which was previously achieved. As a result, the incoming wavefronts did not fill the entire space of the detector, as can be seen in the comparison between Fig. 27 (a) and (b).

The Fourier transform of the two images was then obtained and their principal central regions examined as shown in Fig. 28(a) and (b). The blue crosses indicate pixels with high magnitudes and the red crosses show the three pixels with the highest magnitudes within the image. In Fig. 28(a) there are three distinct high magnitude groups represented by the blue crosses and there is a red cross in each of the high magnitude groups. The groups of high magnitude are also relatively small in area. In Fig. 28(b) there are again three high magnitude groups, however, they are more spread out in size, in comparison with the previous image. Also, the three
pixels of highest magnitude are grouped within the central high magnitude group, rather than being spread out over each of the groups. These results correspond well with the theory associated with resolving power. A reduction in the width of the illuminated gratings causes a reduction in the resolving power of the system. In turn, a reduction in the resolving power leads to more ambiguity in determining a specific input wavenumber. This translates to an uncertainty in the fringe frequency which can be observed in the Fourier domain by a magnitude spreading around the bright peak, rather than there being a distinct point. With uncertainty in the fringe frequency, it follows that uncertainty exists in the input wavelength. Given all variables within the system remained the same, except for the grating illumination width, this spreading in the Fourier domain is shown to be a cause of reduced resolving power. In particular, it was noted that there was more spreading around the central pixel compared with the other two high magnitude groups in the second image. This was caused by the dark area surrounding the interference pattern, which would have caused a larger DC component to be present in the overall signal, and emphasised in the Fourier domain.

D. Advancement to Fingerprint Region

A brief exploration of the technique’s potential in the fingerprint region was conducted. It was originally intended that the current spectrometer be characterised such that the relevant adjustments in components and their orientation allowed an input wavelength range of 6μm. In lieu of investigating the spectrometer’s parameters, however, it was found that serious restrictions of such potential exist. The predominant issue is not the technique itself but the restriction of the components spectral transmission and detection range. The beam-splitter being used has a spectral range of 700nm and 1100 nm. Therefore, it does not accommodate the intended 6000nm input wavenumber. One option is to use a cesium iodide beam-splitter which has a published spectral range of 1600nm to 50μm. (Griffiths and de Haseeth 2007). The other transmission components, such as the collimator, must then also have appropriate anti-reflection coatings which will accommodate this range of wavelengths.

To avoid the use of specialist optical elements with appropriate transmission spectrums, another configuration of SHS has been developed. This is referred to as the all-reflection SHS technique, as is illustrated in Fig. 29. It makes use of a symmetrically ruled diffraction grating which acts as a beam-splitter, labelled as G₀. Light enters the system through S which is an elliptical aperture. The mirror M₄ acts to collimate the light before being reflected onto G₀ by mirror M₃. The grating G₀ then splits the light into the two interferometric arms, where it reflects off mirrors M₁ and M₂ at a wavelength dependent angle. The light is then diffracted back to G₀ where it is recombined and directed towards M₅, which in turn directs it to the exit aperture of S and onto the detector I. The mirrors in this instance can be replaced by gratings to allow a wider range of angles whilst still keeping the instrument relatively small. Resolving power may also be enhanced with the input of gratings. The grating configuration

Figure 28(a) Fourier transform of interferogram without reduced aperture
(b) Fourier transform of interferogram with reduced aperture
Red crosses indicate three brightest pixels. Blue crosses indicate groups of high magnitude pixels

Figure 29. All-reflection SHS configuration
(Harlander, Reynolds and Roesler 1992)
is seen in the bottom half of Fig 29. (Harlander, Reynolds and Roesler 1992).

The trials for this technique have been limited, however, appear to be successful thus far for one-dimensional analysis. Given the restrictions on this project, this technique was not explored further, but its usefulness noted in solving the dilemma of spectral range limitations.

Whether the traditional or all-reflection SHS configuration is used for the 6μm range, the reflection coatings of the gratings and the type of detector must also be altered to suit this region. The ideal reflective coating for the gratings would protective gold with a 97% accuracy range between 0.8μm and 10μm. (Edmund Optics n.d.). In terms of the detector, a mercury cadmium telluride (MCT) detector could be used, which can detect a spectrum range between 1μ and 10μm, depending on the type used. (Single Channel Detectors 2011). The cost of these machines, however, is very high and must be considered carefully in respect of budget restrictions associated with the spectrometer.

VI. Conclusions

This report has outlined the concept, advantages, impact and motivation in investigating the newly developed SHS technique. A SHS spectrometer was built in accordance with theory, and successful Fizeau fringe interferograms were obtained. The central post-processing technique of determining the fringe frequency was investigated and successfully demonstrated for one and two-dimensional analysis. Fringe frequency outputs were applied to the appropriate equations to find the theoretical input wavelength and Littrow conditions of the system.

The accuracy of finding the true fringe frequency by means of the two-dimensional method was investigated and an error margin established. An investigation of the sources and types of noise within the system was also undertaken, and elementary methods were designed to reduce noise in the process.

The effectiveness of the spectrometer built was compared with the theoretical range of a true SHS spectrometer with the same parameters. In particular, the range of wavelengths detectable, passband and the resolution of both ideal and practical systems were compared. The spectrometer which was built fell well short of an ideal system due to inaccuracies in construction and the less than ideal conditions under which the experiments were undertaken.

Expansion of the model for use with a mid-infrared laser source to measure spectra in the molecular fingerprint region was additionally investigated. Owing to the spectral restrictions and quality of apparatus used in the model, it was found to be unsuitable for this use. Application of the all-reflection SHS configuration in conjunction with alternative apparatus with more suitable spectral characteristics was recommended.

Additional recommendations include further analysis of noise and error, for extended accuracy of the technique. Construction of an all-configuration SHS spectrometer would also make an interesting extension project.

VII. Recommendations

Although the project was considered successful in demonstrating and characterising SHS techniques and processes, certain recommendations are made to enhance the outcomes obtained, if the project was to be extended. Based on the spectrometer analysis conducted thus far, it is recommended that methods to enhance accuracy of true bright peak detection and reduce noise effects within the system be further investigated.

The centroid-based method used to find the true bright peak location allowed for a standard deviation margin to be established which indicated the range of possible error in the derived fringe frequency value. It is recommended that this error be investigated further and that the method be rectified so that the standard deviation value obtained relies primarily on the uncertainty in fringe frequency, based on the surrounding pixel values and not on the technique itself.

A more exhaustive noise analysis should also be performed so that noise may be reduced in the system itself, rather than in post-processing. This would involve a more thorough and successful extension of the signal to noise trials conducted in this project. In particular, the trial to reduce the signal power should be repeated but with more advanced apparatus, which would allow more confidence in the results obtained.

If this project were to be extended, it is recommended that the new spectrometer be constructed in a more controlled environment. Ideally, this would include a room in which light could be fully controlled and vibrations reduced. This would provide results with increased accuracy and a better comparison with theory. It was found that there is limited information available on the SHS technique, and that several of the publications contradict one another. Only wider research will enable more accurate information.

A broader recommendation is to extend this project to explore the all-reflection SHS configuration to provide an analysis of its advantages and disadvantages. An outcome of this project would be characterisation of the model for use in the molecular fingerprint region.
Acknowledgements

The author wishes to thank the following people who provided support and guidance throughout the course of this project. A/Prof Charles Harb, for his direction, patience and encouragement throughout this year, which led to the success of the project. Thank you for your support, especially under the difficulties associated with long distance communication. Dr Andrew Lambert, for his technical advice and steadfast patience in discussing the principles of this new technique. The outcomes achieved would not have been possible without your generous assistance and valuable time. A/Prof Don Fraser, for sharing his wealth of knowledge in image processing and clever coding techniques. Your advice was invaluable in conducting the analysis of this project’s results. Thank you for your kind assistance even after your course had been completed.

References

Chimento, Philip F. Narrowband CARS Spectroscopy in the Molecular Fingerprint Region. Enschede: Optical Sciences Group, 2008.
Fraser, D. “Programs: fringetes, gausscircle, logshow.” not published, 2011.