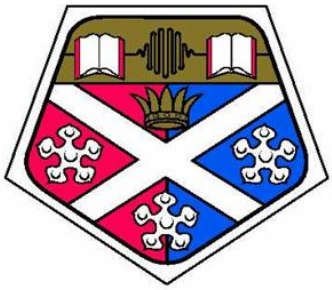
Aspects of Laser Absorption Spectroscopy in the Mid-Infrared and Visible

Robert Neil Joseph Campbell



Department of Physics
Department of Electronic and Electrical Engineering

A thesis presented in the fulfillment of the
requirements for the degree of
Doctor of Engineering
2016

University of Strathclyde Regulation 20.6.6 Declaration:

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Signed:

Date:

Strathclyde Regulation 3.49:

Where the author has not lodged with the Librarian a written objection to the copying of their thesis, the Librarian may give permission for single copies of that thesis, in whole or in part, to be made for the purposes of research or private study, or for deposit in the British Library.

Dedication

For mum and dad, for their constant love and support.

Acknowledgements

This project was carried out after preparatory Masters level study at the

School of Physics and Astronomy, St Andrews University

And funded by

EPSRC

Facilities for the programme were provided by the:

Department of Physics, Strathclyde University

Cascade Technologies Ltd, Stirling

Edinburgh Business School, Heriot-Watt University

Department of Electronic and Electrical Engineering, Strathclyde University

The programme was coordinated by the

EPSRC Centre for Doctoral Training in Applied Photonics, Heriot-Watt University

And completed under the auspices of the

Faculty of Science, Strathclyde University

Thesis proofreading and editorial comments were provided by Mrs Maureen Lee.

Abstract

Laser absorption spectroscopy can be used to identify and quantify gas analytes. The sponsor company's present systems operate in the mid-infrared using room temperature pulsed quantum cascade lasers (pulsed-QCL's). These systems use the noise reduction / sensitivity enhancing technique of sweep integration (SI). In this work, an extension of measurement capabilities is sought in two ways. Firstly, sensitivity enhancement is pursued. The noise reduction technique of wavelength modulation spectroscopy (WMS) is applied using a room temperature continuous wave (cw) QCL spectrometer. Secondly, molecular oxygen is added to the list of measurable analytes. This molecule's near-infrared and visible transitions are addressed with a wavenumber prototype semiconductor diode laser.

The sensitivities of the SI and WMS techniques are compared for the cw-QCL spectrometer, and compared to the SI sensitivity of a typical company pulsed-QCL system. New analysis and modeling software was written to facilitate the thesis work and to carry it forward. A thorough analysis of a pulsed-QCL CT3000 analyzer is undertaken to minimize a reduction in capability - should an oxygen measuring laser replace one of its pulsed-QCL's.

The experimental work was constrained by time and budget - particularly with regard to the cw-QCL spectrometer's AC-coupled detection. Using AC-coupled detection had cost and integration advantages, but posed a number of problems - including electronic incompatibility issues. Nevertheless, the outlook is positive, and a modest sensitivity improvement was found for WMS over sweep integration (0.017 absorbance units (a.u.) in 102.4s compared to 0.080 a.u. in 51ms). Both sensitivities are some way behind the present sweep integration performance of the company's pulsed spectrometers (0.004 a.u. in 10ms). However, the sensitivities are comparable to earlier stages of development. In the case of oxygen spectroscopy, the prototype diode laser's thermal stability was an issue, but several spectral regions were found to be suitable for single or multimode spectroscopy.

Contribution to Knowledge and Innovation

The sponsor company's gas analyzing technology is based upon using the intra-pulse technique (a form of sweep integration (SI)) with pulsed-quantum cascade lasers (pulsed-QCL's). A sensitivity enhancement was sought by combining a (continuous wave) cw-QCL with the noise reduction technique of wavelength modulation spectroscopy (WMS).

It was not known how the cw-QCL laser, designed for stability, would respond to current sweeping and modulation. The characterization and tuning work was therefore of value as a precursor to WMS, and of value to the sponsor company.

Prior to this work, the limitations and potential problems of using AC-coupled detection had not been appreciated. This detection had posed no problems for the sponsor company with pulsed-QCL's. The work on elucidating these limitations was therefore of value since there are cost and integration advantages in retaining these detectors. Establishing practical parameters for Sweep Integration (SI), and WMS, was an aid to the electronic team's planning, since cw-QCL's have different timing and electrical requirements to the pulsed-QCL's.

Establishing the absorbance Limit of Detections (LoD's) (noise floors) of SI and WMS allowed Cascade Technologies to estimate the future potential for improving concentration sensitivities. The LoD's were similar to the pulsed-QCL systems at an earlier stage in their development.

An asymmetry in the feet of the 2f-WMS lineshape was found. By modeling, this can be attributed to the simultaneous 1f intensity modulation of the laser output. In general, this asymmetry is much larger than that seen in the literature for diode lasers. However, it was established that this did not affect the experimentally found linearity of the calibration relationship between sweep integration and 2f-WMS peak heights (at least at low pressures).

The performance of a semiconductor diode laser, and its potential for (visible) oxygen spectroscopy was also investigated. This laser was a (wavenumber) prototype for the manufacturer / supplier, and it was the first time that the sponsor company had sought to move their measurement capabilities beyond the mid-infrared – to measure molecular oxygen. The strengths and weaknesses of the laser design / growth was therefore of interest to both parties.

A Cascade Technologies CT3000 gas analyzer was constructed and thoroughly characterized. The end use for this sensor was uncertain, so the primary aim was to compensate, as much as possible, for a future swapping out of one or more of the four component lasers. Detailed information was taken on LoD's, cross-interferences, and the relative noise of measuring the same gas on different lasers. The standard wavenumber fitting windows in some cases were expanded.

During this work a catastrophic loss of signal strength occurred and was traced to a contaminating residue in the Herriott gas cell. This appears to have come in from the manufacturer. As well as being useful knowledge for the future, a general diagnostic test sheet was developed to eliminate other possibilities.

A significant contribution to process efficiency was made by deciphering, documenting the workings of, and modifying a linearity testing spreadsheet. This spreadsheet was constantly in use, but the workings of the Microsoft VBA code on which it ran had passed out of company memory.

A number of software programs were developed - mainly coded in LabVIEW - adding capability, functionality, and process efficiency which was not otherwise available. For instance, there was a software gap in processing spectra and estimating concentration values from pre-recorded data or data straight from an oscilloscope. An automated and interactive data analysis program was written. A Gaussian and Lorentzian modeler was integrated with LabVIEW's Levenberg-Marquardt fitter to extract number densities and concentrations. A Voigt profile and pseudo-Voigt

modeler was also developed, and some work done on the integration of this with LabVIEW's data fitter. (The sponsor company had previously restricted themselves to the mid-infrared where Gaussian or Lorentzian profiles are used). An Allan variance program was made for the optimization of the number of averages, in case a means had become available for taking continuous concentration measurements. A full 2f-WMS lineshape model was coded and compared to a similar small absorbance model from literature, and compared to experimental data.

Contents

Detailed Contents

Appendices

Chapter 1

Introduction

Tier 1	Tier 2	Tier 3
Solid electrolyte	Paramagnetism	UV fluorescence
Wet electrolyte	Photoionisation	Chemiluminescence
Catalytic metal oxide	Flame ionisation	Photoacoustic
Infrared absorption	UV absorption	Ion mobility
	Thermal conductivity	TDLAS (tunable diode laser absorption spectroscopy)
		CRDS (cavity ring down spectroscopy)
		LIDAR (light detection and ranging)
		Semiconducting polymers

Table 1-1: Some major gas sensing technologies. The tiers refer to groupings of technologies according to the decreasing number of sensors based upon those technologies in the market. The list was compiled from two articles by Bogue (2011a, 2015a).

1 Introduction

1.1 Gas sensing applications and technologies

There are many application areas for gas sensing, served by many different detection technologies. Table 1-1 lists some of these technologies, as given by Bogue (2011a, 2015a). Bogue notes that the choice of technology depends upon 'factors such as the gas species involved, the concentration range, required resolution, sensitivity, specificity and response time, the operating environment (temperature, humidity, interfering species, etc.), size, cost and other considerations such as whether the requirement is for a single point measurement or over an open path'. Tiers 1, 2, and 3 in Table 1-1 form three broad categories, ordered according to the decreasing number of sensors based upon these technologies in the total market.

Three of the four detection technologies in Tier 1 are non-optical. Probably the most ubiquitous gas sensors are solid electrolyte 'lambda sensors'. These are used in the exhausts of petrol engines to measure molecular oxygen. Lambda sensors control the air / fuel ratio, so that combustion is complete and harmful exhaust gases can be minimized. The most common design consists of a zirconium oxide thimble, coated inside and out with micro-porous platinum electrodes. The zirconium oxide acts as the heated electrolyte as the O^{2-} ions can diffuse through the structure. The inside of the thimble is filled with a concentration of oxygen at higher pressure than that of the exhaust. Diffusion then occurs according to the pressure gradient (hence exhaust concentration), creating a voltage across a resistance between the electrodes (Jahnke, 2000). Since the device reacts to a pressure difference, it is non-selective to oxygen molecules, and so significant amounts of other combustion products can lead to erroneous readings.

Gas sensors are based on a plethora of different physical and chemical mechanisms. For instance: a conductive material may change its resistance when exposed to a gas; a piezoelectric oscillator may change its mass, and hence oscillation frequency, due

Optical gas sensing techniques

Principle	Specific techniques
Absorption	Non-dispersive infra-red (NDIR)
	UV absorption
	Fourier transform infra-red (FTIR)
	Differential optical absorption spectroscopy (DOAS)
	Photoacoustic spectroscopy (PAS)
	Cavity ring-down spectroscopy (CRDS)
	Tuneable diode laser absorption spectroscopy (TDLAS)
	Differential absorption LIDAR* (DIAL)
	Gas cloud imaging
Luminescence/fluorescence	Chemiluminescence (CL)
	UV fluorescence (UVF)
Ionisation	Photoionisation detection/detector (PID)

Note: *LIDAR – light detection and ranging

Table 1-2: Some currently important optical gas sensing techniques from an industrial perspective. (Taken from Bogue, 2015a).

to gas adsorption; chemiluminescence (in Tier 3) occurs if there is a chemical reaction between the target gas and some compound which emits light, which can then be detected. In general, the relatively inexpensive solutions - while of robust design because of their simplicity - tend to be non-selective or only partially selective to the target gas, and may be prone to contamination. They therefore tend to be application specific, and used in situations where the potential gas mixtures are well known.

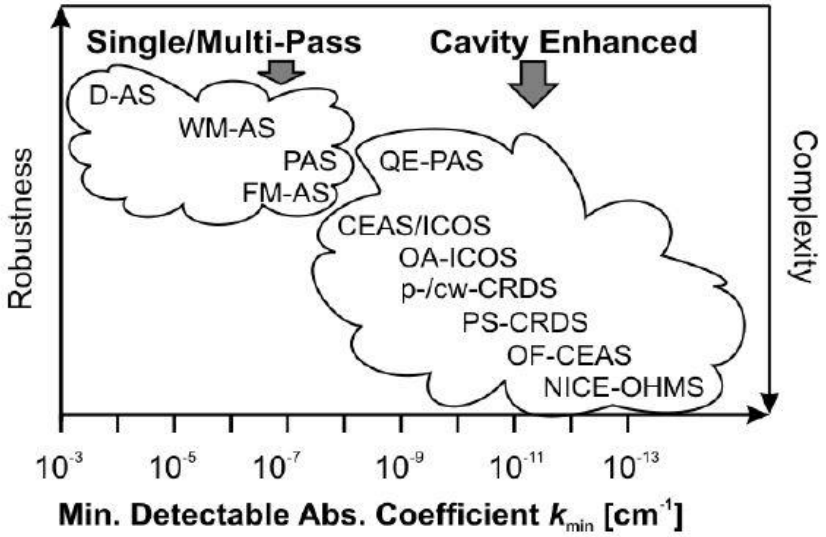
Of the optical techniques, infrared absorption is widespread, since many molecules of industrial and environmental interest undergo transitions in this spectral region. In the mid-infrared, such transitions are due to rotation-vibration changes from the ground vibration-rotation state. Many molecules show narrow and uncongested absorption lines in this region giving a 'molecular fingerprint'. In the near-infrared, weaker overtone and combination transitions are accessible.

There are a large variety of optical techniques (see for example Table 1-2). In terms of the number of sensors produced, most do not use lasers. For instance, non-dispersive infrared (NDIR) sensors use a filament lamp as a broadband source. A major use of these is to measure CO_2 – to which the non-optical Tier 1 techniques are not well suited. The simplest NDIR sensor collects the broadband radiation through two filters – one letting through wavelengths in a non-absorbing spectral region, and the other letting through radiation attenuated by absorption. The resulting differential signal is then related to the concentration of the analyte (Jahnke, 2000, Bogue, 2015a). Fourier transform infrared (FTIR) spectroscopy (Section 4.5.3.1) uses a broadband thermal source, such as a Nernst glowbar. High resolution, broad bandwidth spectra can be recorded, but on significant time scales, and with bulky equipment.

Turning to tunable diode laser absorption spectroscopy (TDLAS), until the advent of the (pulsed-) Quantum Cascade Laser (Faist et al., 1994) and its subsequent commercialization, the choice for TDLAS was generally two-fold. In the

near-infrared, thermoelectrically cooled III-V semiconductor diode lasers, such as GaAs/AlGaAs systems, could be used. In the mid-infrared, cryogenically cooled II-IV lead-salt semiconductor diode lasers were used. As mentioned, near-infrared transitions tend to be weaker and more congested, but the III-V lasers had undergone extensive development due to their use in the telecommunications industry. Mid-infrared II-IV lasers accessed stronger transitions, but tended to have relatively broad linewidths, were multimode (Fabry-Perot resonators), and subject to mode hopping. Quantum cascade lasers (QCL's), although having a fundamentally different operating principle to diode lasers (Section 2.9.3), nevertheless benefitted from III-V semiconductor growth technology. Compared to lead-salt lasers, QCL's have been found to have narrower linewidths, higher output powers, increased measurement precision, and better spectral stability (Webster et al., 2001). Since QCL's can be operated at room temperature, by thermoelectric cooling, unattended gas sensor operation is possible. Single mode operation is usually achieved by a distributed feedback (DFB) grating etched above the device's active layer. This does, however, make the overall wavenumber coverage much less than that of a lead-salt diode laser (several wavenumbers for the QCL to hundreds of wavenumbers (albeit with incomplete coverage) for the lead-salt laser) (Welzel et al., 2010). QCL's must therefore be matched more carefully to the intended absorptions. An alternative approach uses a non-DFB anti-reflection coated QCL structure as the gain medium in an external cavity arrangement (Section 5.11). This maintains a broad tuning range.

In a typical TDLAS direct absorption experiment, the starting wavenumber of a scan is set by the laser substrate temperature, and the injection current is ramped to tune over an absorption feature or features, measuring one or more gases in a scan. Gas identification can be made from the spectral position of the transitions, and the concentration can be obtained from first principle calculations (Section 2.6). The signal to noise ratio is usually improved by 'sweep integration' – averaging sequential scans. Sensitivities of 10^{-2} to 10^{-4} absorbance units can be achieved in this way (Welzel, 2009).



Acronym	
D-AS	Direct absorption spectroscopy
WM-AS	Wavelength modulation absorption spectroscopy
PAS	Photo-acoustic absorption spectroscopy
FM-AS	Frequency modulation absorption spectroscopy
QE-PAS	Quartz enhanced photo-acoustic absorption spectroscopy
CEAS / ICOS	Cavity enhanced absorption spectroscopy / Integrated cavity output spectroscopy
OA-ICOS	Off-axis integrated cavity output spectroscopy
p-/cw-CRDS	Pulsed- / continuous wave- cavity ring down spectroscopy
PS-CRDS	Phase shift cavity ring down spectroscopy
OF-CEAS	Optical feedback cavity enhanced spectroscopy
NICE-OHMS	Noise immune cavity enhanced optical heterodyne molecular spectroscopy

Figure 1-1: Some laser based optical gas sensing techniques, from a research / laboratory perspective. The figure (from Welzel et al., 2010) indicates broadly where the techniques fall in terms of relative robustness and simplicity for industrial use. Also indicated are the sensitivities which have been obtained. The 'minimum detectable absorption coefficient' is equivalent to the absorbance limit of detection per unit length.

Other noise reduction techniques can be applied to enhance the measurement sensitivity for trace gas analysis. Figure 1-1 places some of these techniques, in broad terms, according to the technique's complexity and the best (per unit path length) sensitivities achieved, generally with diode lasers. The choice of technique will depend upon factors such as the associated sensor's commercial development and its cost. The figure indicates a general trade-off between the achievable detection sensitivity and the robustness and simplicity of a given technique. Of the noise reduction techniques, after sweep integration, wavelength modulation spectroscopy (WMS) is probably the easiest to implement. In general, the laser tuning is oscillated about the absorbance lineshape as overall the tuning scans across the line. This moves the signal to a higher (lower noise) frequency. Derivative-like signals result which can be related back to the direct absorption, and hence to the gas concentration (Sections 1.2 and 5.3). In WMS, the amplitude of the tuning modulation is larger than the absorbance half-width, and the modulation frequency is much smaller than the half-width. Sensitivities in WMS are usually limited by laser excess noise and etalon effects in the beam path. Typical sensitivities of about 10^{-5} absorbance units are obtainable with diode lasers (Kluczynski and Axner, 1999, Welzel, 2009).

WMS is usually coached in terms of the radiation intensity (proportional to the square of the electric field). Alternatively, in terms of the electric field, a great many sidebands are created when the laser's centre (carrier) frequency is subject to the modulation frequency. This comb of sidebands sweeps across the absorption lineshape as the laser centre frequency tunes. Frequency Modulation Spectroscopy (FMS) is essentially the same technique, but occurs when the modulation is small compared to the absorbance-half width, and the modulation frequency is much larger. Under these circumstances, the spacing between sidebands is large, and generally only one pair of sidebands will sweep across the absorbance profile in addition to the carrier, as the carrier frequency tunes. The resulting signal is the sum of the two beat signals of the carrier with the sidebands. In the absence of absorbers these beats occur in anti-phase and a DC intensity signal results. In the presence of absorbers, as a sideband sweeps across the absorbance profile, its beat signal

becomes unbalanced with its counterpart - by absorption or dispersion - and an intensity change results. FMS's higher frequency, compared to WMS, moves the signal into a lower noise region of the spectrum. This is particularly useful if laser excess noise is a limiting factor. Sensitivities of 10^{-6}-10^{-7} absorbance units are typical for diode lasers, limited by etalon fringing and a background 'residual amplitude modulation' signal arising from simultaneous modulation of the laser amplitude. FMS needs more complicated equipment than WMS, and the higher frequencies place much higher bandwidth requirements on the detection system and processing electronics (Kluczynski and Axner, 1999). 'Two-tone' FMS (Janik et al., 1986) lowers the bandwidth requirement by modulating the carrier at two closely spaced frequencies. The signal is then processed at the resulting beat frequency.

Another way of improving the signal to noise ratio is by increasing the absorption path length, via the use of multi-pass absorption gas cells. The astigmatic Herriott cell (Section 2.8.2) can produce path lengths up to about 200m. Alternatively, cavity enhanced techniques can achieve path lengths of a few kilometers (Welzel et al., 2008, Berden and Engeln, 2009, Gagliardi and Loock, 2014). Thus, these techniques have the potential for sensitive measurements. Figure 1-1 shows their sensitivity *per unit path length*. In these terms, cavity enhanced methods seem to have higher sensitivities than those obtained in multi-pass cells. However, literature values of their absolute sensitivities (in absorbance units) can be one or two orders of magnitude less favourable (Welzel, 2009).

In cavity enhanced methods, radiation is coupled into the resonant modes of a cavity formed by two high finesse mirrors. In cavity enhanced ring down spectroscopy (CRDS), the radiation in the cavity is built up using a pulsed laser or an interrupted continuous wave laser. The subsequent exponential temporal decay of the radiation leaking from the cavity is measured, preferably from a single cavity mode. The radiation decay occurs more quickly according to the concentration of absorbing gas molecules present. In cavity enhanced absorption spectroscopy (CEAS) or integrated cavity output spectroscopy (ICOS), many cavity modes are excited by a continuous

Some applications of gas sensing
Process optimization and control: industrial plant, engines
Pollution monitoring: industrial plant, shipping, atmosphere, air quality control
Environmental and atmospheric chemistry: greenhouse gases, agricultural processes, water processing
Homeland security: explosives, narcotics
Leak detection: safety (fire, explosion, and asphyxiation prevention), aerosol leaks, pipeline leaks
Breath analysis: medical diagnostics, breathalysers
Plasma diagnostics: monitoring, testing models of reactions and cleaning processes

Table 1-3: Some application areas of gas sensing technologies.

wave laser (CEAS) or a continuously pulsed laser (ICOS). The laser frequency is then repeatedly tuned over an absorption profile, resulting in an averaged congested spectrum, related to the absorption encountered. Noise immune cavity enhanced optical heterodyne molecular spectroscopy (NICE-OHMS) is a hybrid technique, combining optical cavities with mode locking and FMS (Ye et al., 1998, Foltynowicz et al., 2008). Sensitivities as low as 10^{-9} absorbance units have been obtained (Welzel, 2009).

Applications of gas sensing technologies are diverse, and Table 1-3 summarizes some important categories. Pollution control and leak detection are, perhaps, obvious examples. Monitoring gaseous products from chemical reactions can also optimize those reactions – for instance in maintaining or achieving a stoichiometric ratio. Monitoring of the atmosphere (McManus et al., 2010) helps in applications such as controlling urban pollution, measuring greenhouse gases, and in indoor air quality management. Breath analysis can offer a non-invasive form of diagnosis – since the gases in breath can be indicative of disease (Risby and Tittel, 2010, Shorter et al., 2010). Other applications include homeland security (such as explosive detection) (Normand et al., 2010, Bogue, 2011b, 2015b), and plasma diagnosis (Ma et al., 2009, Welzel, 2009, Welzel et al., 2010). (See also Section 1.3).

1.2 Wavelength Modulation Spectroscopy

The two noise reduction methods used in this thesis are Sweep Integration (Section 4.3) and Wavelength Modulation Spectroscopy (WMS) (Section 5.3). Sweep integration involves taking sequential scans of an absorption feature, and averaging to improve the signal to noise ratio. Although a technique in itself, it is also used to supplement other methods. It is perhaps the most obvious technique, but has relied on the development of fast digitizers and computer processors for its success - since being pioneered by Jennings (Jennings, 1980, Cassidy and Reid, 1982).

Wavelength Modulation Spectroscopy modulates a parameter of an experiment, moving the frequency of the detected signal to a lower noise spectral region. Phase sensitive detection is then used to isolate the signal. In this thesis, the injection current to the laser is sinusoidally modulated. As the laser frequency slowly tunes over an absorption profile, the faster modulation oscillates the tuning about the absorption profile. The detector then measures a signal whose amplitude oscillates at the modulation frequency and its harmonics. After processing the signal through a lock-in amplifier, derivative-like signals result. For oscillations that are very small in comparison to the absorption linewidth, the signal is the n^{th} derivative of the line profile when detected at the n^{th} harmonic of the modulation frequency. The ratio of modulation amplitude to the absorption halfwidth is known as the 'modulation depth'. In practice, to improve the signal to noise ratio, the oscillation amplitude is generally made larger than the absorption linewidth, and the signals are only derivative-like. Therefore, straight forward integration can not be used to recover the direct absorption signal – and hence the gas' concentration.

The theoretical modeling of WMS signals began in the context of nuclear magnetic resonance (nmr) and electron spin resonance (esr) experiments, and has been extended in the context of spectroscopy with injection current modulated semiconductor diode lasers. In this latter case, a complicating factor upon the signal shape is the simultaneous modulation of the laser intensity with frequency tuning. Intensity variations at the detector are therefore caused by both the laser intensity modulation and from oscillations of the laser frequency about the absorption profile.

In early nmr / esr modeling, the signals were represented by Taylor series expansions – strictly valid only for small modulation depths (see for example: Myers and Putzer, 1959, and Russell and Torchia, 1962). Wilson (1963) applied Fourier analysis to make his approach valid for all modulation depths. He used numerical integration to calculate the first three harmonic derivative-like signals arising from Lorentzian and Gaussian lineshapes. Wahlquist (1961) and Arndt (1965) derived equivalent

analytical expressions for the harmonics of a Lorentzian lineshape, and gave explicit expressions for the first two harmonics.

Many WMS experiments use the 2nd harmonic signal (2f-WMS), which produces 2nd-derivative-like signals (Section 5.3.3). Reid and Labrie (1981) showed that the nmr models available were useful in the context of experimental 2f-WMS semiconductor diode laser spectroscopy. However, these models did not take into account the mixing of harmonics due to simultaneous laser intensity modulation. (In terms of Section 5.9, they were essentially calculating only the H_2 term). Reid and Labrie compared theory with experiment for Arndt / Wahlquist's analytical 2f-WMS model derived from a Lorentzian lineshape. Similarly, Wilson's 2f-WMS numerical method was used for the 2f-WMS H_2 signals resulting from Gaussian and Voigt lineshapes. In particular, they showed that the maximum 2f-WMS was obtained theoretically and experimentally at a similar modulation depth for each lineshape (modulation depth ≈2.2).

The effect of simultaneous laser intensity modulation was considered in the Taylor approximation by Olson et al. (1980), and the 2f-WMS signal was modeled by Philippe and Hanson (1993) via Fourier analysis. Unlike previous approaches, the intensity multiplier in the Beer-Lambert Law was treated as a sinusoidally oscillating function, leading to a mixing of intensity modulations from the laser and the absorption line – giving the detected signal. An expression for the lock-in output can then be derived by multiplying the detected signal by the lock-in's reference sinusoid and averaging. The WMS signal is, in general, still dominated by the nth derivative-like signal - for detecting at the nth harmonic - but the character of other derivatives can also influence the signal shape.

This approach was continued by Kluczynski and Axner (1999) and expanded upon by Kluczynski et al. (2001). They presented a general theoretical treatment of WMS signals based on Fourier analysis, for any order of harmonic detection. They also added a nonlinear (quadratic) term in the intensity term of the Beer-Lambert Law. In

turn, this analysis was built on by Li et al. (2006) to form explicit practical expressions for the magnitude of a lock-in output for the 1f-WMS and 2f-WMS cases - the magnitudes being formed from the relevant X and Y lock-in outputs. (The X output is derived and used in Section 5.9 of this thesis). Li et al. paid particular attention to the phase differences between the linear and nonlinear laser intensity oscillations and the frequency tuning.

All of the above models are formulated in terms of the radiation intensity at the detector (proportional to the square of the electric field). Some treatments had cast WMS in the same theoretical structure as that used for Frequency Modulation Spectroscopy (FMS) - that is, in terms of the amplitude and the phase of the radiation's electric field (Section 1.1). The physical principles of both techniques are the same, and switching to the FMS theory can also handle dispersion signals through the retention of phase information. Bomse et al. (1992) and Silver (1992) considered a noise analysis of WMS and FMS techniques in the FMS formalism. Supplee et al. (1994) wrote a paper to guide experimentalists in mapping general 1f-WMS signal trends in a two dimensional parameter space according to the chosen modulation depth and the modulation frequency. They also provided an expression for the 1f-WMS signal at low modulation depth, derived from the FMS formalism.

However, unlike in the FMS case, the low modulation frequency of WMS implies that a very large number of frequency side-bands interact with the absorption feature, and therefore need to be summed over. Kluczynski and Axner (1999) concluded that it was 'neither necessary nor convenient' to use the FMS approach. In addition, the conventional FM approach assumes that the radiation's electric field is directly modulated, that is, there is a linear relationship between the modulating signal and the modulation of the field. This is the case during external modulation of an otherwise fixed frequency laser beam, via, say, an acousto-optic or electro-optic modulator. However, Zhu and Cassidy (1997) pointed out that injection current modulation of a diode laser directly modulated its intensity (proportional to the square of the electric field) – since the physical mechanism is the conversion of

electron-hole pairs into photons. For this case, then, the square root of the field becomes appropriate in the electric field based FM formalism. The result was validated by the predicted absence of a 2f residual amplitude modulation, when the laser intensity modulation contained no quadratic or higher order components (or the transmission contained no non-analyte wavelength dependent effects, such as from etalons (Kluczynski and Axner, 1999)).

Returning to the intensity approach of WMS, several 'calibration free' methods have appeared in the literature. In principle, a direct absorbance calculated from first principles (via the Beer-Lambert Law) is calibration free. However, a 'factory calibration' against known gas concentrations is still advisable, if possible, to account for instrumentation effects. In the methods described below, Henningson and Simonsen (2000) used interpolation formulae to link a 2f-WMS model to the direct absorbance, while Stewart et al. (2011) extracted the direct absorbance from the 1f-WMS signal. Reiker et al. (Rieker et al., 2009) do not concern themselves with the direct signal, and seek to remove calibration factors be ratioing their 2f-WMS signal with a simultaneously recorded 1f-WMS signal, and comparing to the pressure (concentration) dependent theoretical relationship.

From first principles, Henningsen and Simonsen (2000) gave a formula relating the peak of the direct absorbance to the peak to trough height of their expected 2f-WMS signal (based on calculating H_2). Their formula was valid in the small absorbance approximation where the two peak heights are directly proportional to each other. To make their formulae apply to arbitrary modulation depth, they characterized corresponding widths and heights of the calculated direct and 2f-WMS signals as functions of modulation depth – using interpolation formulae to fit the trend of their calculations. The accuracy of the approach was better than 10% for their external cavity semiconductor diode laser set-up. As a general method, the Henningsen and Simonsen approach is limited to isolated absorbance lines so that the 2f-WMS line is undistorted. The laser's wavenumber scan also needs to capture enough baseline for normalizing out the effect of the changing laser intensity. This may be difficult to

achieve with pressure broadened lines, and in spectral regions subject to interference from other lines. The environmental conditions of temperature, pressure, and path length also need to be known. For cases where large instrumental broadening occurs, the instrumental factors would need to be calibrated beforehand. If these factors changed with time in-situ they would negate this 'calibration free' approach.

In a series of papers, the Optical Gas Sensing group, in the Electronic and Electrical Engineering Department at Strathclyde University, developed two calibration free approaches (Duffin et al., 2007, McGettrick et al., 2008, McGettrick et al., 2009, Bain et al., 2011, Stewart et al., 2011). These were developed for process control in harsh environments utilizing semiconductor diode lasers. The idea is to extract a direct absorption signal from first harmonic detection. This allows calibration issues to be avoided, particularly where the pressure is variable and unknown, while retaining the noise reduction advantages of phase sensitive detection. The method is not well suited to trace gas detection, however. The signal can be strong – but rides on a high background signal – reducing the sensitivity. The technique also requires the laser tuning to be wide enough to encompass a background. This is needed to normalize out the effect of the slowly changing laser intensity as the laser tunes across the absorption feature.

In the first calibration free technique, the 'RAM approach' exploits the phase difference, due mainly to thermal lag, between the 1st derivative-like component of the 1f-WMS signal and the 'residual amplitude modulation' ((1f-)RAM) component. The 1st-derivative-like signal arises from wavenumber oscillation about the absorbance profile as the laser slowly tunes across the profile. This becomes an intensity amplitude oscillation at the detector. The RAM signal arises from simultaneous laser intensity oscillations caused by injection current modulation of the laser's wavenumber. In general the 1f-WMS signal is composed of an infinite number of terms caused by the mixing of these effects. The RAM signal has the phase of the intensity modulation, which also provides the lock-in reference frequency. The 1st derivative-like component has the same frequency but is delayed

in phase. Other terms, proportional to higher derivatives of the absorbance lineshape, appear at the same frequency but with integer multiples of this phase delay. The signal components can conveniently be considered in a phasor diagram approach.

In a conventional 1f-WMS experiment, the lock-in amplifier's measurement axis is along the RAM signal. That is, at zero phase – the phase of the reference signal. This tends to maximize the signal to noise ratio. The 1f-WMS signal then consists of the RAM signal, the projection onto this axis of the 1st derivative-like signal, and the smaller projections of the higher order terms.

The RAM signal itself encounters the absorbance lineshape and is reduced proportionately. The direct absorbance is therefore contained within this signal. The RAM approach extracts this signal by first choosing the lock-in measurement axis 90° to the 1st derivative-like signal. This can be done, for instance, by minimizing the asymmetrical influence on the symmetrical direct absorption. This nulls the 1st derivative-like signal and allows a projection of the RAM signal, and hence the absorbance profile, to be measured. However, the projections of the remaining terms cause the direct lineshape to be underestimated. But a 'normalized shape distortion function', calculable in advance, can be applied to reconstruct the full direct profile.

If the bandwidth of the system is sufficient, a modulation frequency can be found which causes the phase delay to become 90°. The measurement axis therefore aligns the RAM signal and so the full signal is measured. The correction factor still needs to be applied to account for the other projections.

The second technique, the 'phasor decomposition method' (PDM), is an extension of the RAM approach allowing the full RAM signal to be measured at any modulation frequency - reducing the system bandwidth needed for a maximized RAM signal. This method requires two available lock-in amplifier channels. The lock-in's x-axis is again aligned along the RAM component, and the y-axis is simultaneously recorded. The dominant 1st derivative-like signal can then be removed by

trigonometrical manipulation, regardless of its phase angle, provided that this angle is known or measured. The correction function is then applied as before.

The High Temperature Gasdynamics Laboratory at Stanford University circumvented the need for the laser scan to capture the whole absorption line (Li et al., 2006, Rieker et al., 2009). By normalizing their 2f-WMS signal with a simultaneously measured 1f-WMS signal, there was no need for the 2f-WMS signal to also include a baseline for extrapolation. For small absorbances (<0.05) the 1f-WMS absorbance is a small fraction of the background 1f-WMS signal, and so has little effect during normalization. The 2f line centre is also at the zero of the 1f signal. For the analysis to be successful, the line centre needs to be captured and the gas pressure needs to be reasonably well known. Measuring and modeling the 2f/1f ratio removes the calibration factors between measured and simulated spectra, such as those due to intensity fluctuations and lock-in amplifier gain. If the temperature is not known, it can be found from the ratio of the measured 2f/1f signal. In the small absorbance approximation, the ratio of the 2f/1f measured signal to the 2f/1f simulated signal, for a single absorbance line, is in the same ratio as the actual partial pressure to the assumed simulated partial pressure. The mixing ratio can then be obtained from the actual partial pressure.

In a conventional 2f-WMS experiment, the lock-in reference phase needs to be adjusted to zero - to maximize the signal - and at large modulation depths inaccuracy is introduced because the output is a sum of terms with different phases. Rieker et al.'s (2009) preferred approach is to measure and simulate the lock-in's X and Y outputs to calculate the signal magnitude and remove the dependence on reference phase. The magnitude of the 1f signal is used to normalize the raw X and Y 2f outputs. The normalized measured (or simulated) X and Y backgrounds are then subtracted from the respective normalized raw X and Y components. The resulting X and Y components can then be added in quadrature and square rooted to get the final 2f/1f ratio. The measured signals thus come from four lock-in amplifier channels – X and Y outputs for both the 2f and 1f signals.

1.3 Contextual History: The Strathclyde Molecular Spectroscopy Group, Cascade Technologies Ltd., and selected other developments

The first cryogenically cooled pulsed-QCL was realized in 1994 (Faist et al.). With improved thermal management in the designs, Peltier cooled room temperature versions appeared shortly afterwards (Faist et al., 1996b, 1996a). In general, the pulsed-QCL's could also be operated in continuous wave mode at cryogenic temperatures, and the cw mode became the specific focus of some designs (Sirtori et al., 1996). In general, QCL's need higher operating voltages than their diode laser counterparts, to align the energy levels of the device (Section 2.9.3). This slowed the development of Peltier cooled cw-QCL's, but a laboratory version of a device was presented in 2001 (Hofstetter et al.). Room temperature operation was demonstrated shortly afterwards (Beck et al., 2002).

A sub-group of the 'Experimental Quantum Optics and Photonics Group', at Strathclyde University, began using a commercially available room temperature pulsed-QCL in 2001. For convenience, this sub-group will be referred to hereafter as the 'Molecular Spectroscopy (MS) Group'. These studies led to the patenting of intra-pulse spectroscopy and the formation of Cascade Technologies in 2003.

The first QCL publication from the Strathclyde MS Group (Normand et al., 2001) noted the potential to measure a whole high resolution spectrum across a single chirped pulse, using a DFB pulsed-QCL. This potential is a consequence of the pulsed-QCL maintaining a narrow emission spectrum at any particular time during the wavenumber chirp. The MS Group observed this using the pulsed laser as the source in a high resolution Fourier Transform Spectrometer (FTS). As the injection current heats the QCL's active region, the narrow linewidth laser emission downchirps across an absorption line. The temporal profile of the overall radiation pulse therefore has a simple mapping onto the frequency domain. Indeed, Normand et al found the time to frequency relationship to be linear (constant tuning rate) for $0.6cm^{-1}$ of a $1cm^{-1}$ (106ns) pulse. In addition, there was also a linear mapping of the

amplitude of the radiation intensity to the amplitude of the applied current pulse. The detailed temporal structure of the current pulse reaching the laser was measured with a Rogowski coil around the cabling between pulse generator and laser.

Faist et al (1997) had earlier found this narrow instantaneous linewidth as a pulsed-QCL chirped. Using a boxcar integrator, they followed the temporal evolution of a 100ns pulse in frequency space. They found their measurement of the radiation linewidth during each 3ns gate, spaced 10ns apart, to have a maximum limit bounded by their Nicolet FTS resolution of $0.125cm^{-1}$ (3.7GHz). Normand et al were able to use a higher resolution Bomem DA-3 FTS with maximum resolution of $0.005cm^{-1}$ (150MHz) - mapping more detail of the temporal evolution of the pulse in the frequency domain. Rather than using Faist et al's time resolved method, they continuously scanned their FTS (with the pulsed-QCL as source) with a pulse repetition frequency greater than the FTS sampling frequency. This allows the laser to be 'seen' as a quasi-continuous narrowband source, forming, in effect, a simulated broadband FTS frequency source when considering the sweep as a whole, as the laser downchirps across absorption features of fixed frequency.

For a well behaved top hat current pulse, the intra-pulse laser frequency emission spectrum is approximately trapezoidal in profile (Figure 4-45). In general, a reduction in intensity occurs across the pulse as the laser becomes less efficient with rising active layer temperature, and non-radiative mechanisms such as phonon scattering compete for energy. In Normand et al's study the impedance matching of the current to the laser was imperfect - producing oscillations in the injection current reaching the laser. These oscillations, however, benefitted the study by allowing the time evolution of the frequency structure within the overall radiation envelope to be recorded – because of narrow instantaneous emission and the near linear relationship between time and laser centre frequency during the downchirp. They were also able to confirm that the fast tuning was due to the injection current in the laser's active region, due to a lack of chirp rate dependence upon substrate temperature. The temperature induced wavenumber downshifts were characterized as a function of

increasing duty cycle and pulse lengths. In a following paper in collaboration with Glasgow University (Garcia et al., 2003), the Strathclyde MS Group used their continuous scanning method to record the frequency emission envelopes of a multimode Fabry-Perot pulsed-QCL. They noted that the multimode behaviour could be advantageous in finding appropriate tuning start points for interrogating the widely spaced lines of NH_3.

From here, the work of the MS Group concentrated on building intra-pulse mid-infrared spectrometers with DFB pulsed-QCL's, and using them for studies and applications in the linear and non-linear regime. The linear (Beer-Lambert Law) regime was also the province of Cascade Technologies.

In the linear regime, the MS Group used their first QCL spectrometer to study the $\nu(9)$ transitions of the relatively heavy gas 1,1-difluoroethylene (CF_2CH_2) (Duxbury et al., 2002, Normand et al., 2003). This gas has many closely spaced absorption lines, due to its high moment of inertia, and could therefore test the new spectrometer's sensitivity and resolution for sufficient identification and measurement capabilities. (1,1-difluoroethylene also acts as a proxy for similar hydroflourocarbons emitted into the atmosphere (McCulloch et al., 2003)). The QCL spectrometer's performance was evaluated against a high resolution FTS spectrum ($0.0015cm^{-1}$ from a Bruker IFS 120HR at the Rutherford Appleton Laboratory), and a strong correlation was found between the recorded spectra. An estimate of the instantaneous linewidth of the laser was made by comparing FTS frequency domain spectra, with, respectively, the chirped QCL acting as the effective broadband FTS source, and a traditional broadband globar source. No significant broadening was found beyond that due to the FTS instrument function, indicating that the chirping QCL had had an instantaneous linewidth no more than that of the FTS resolution of $0.0015cm^{-1}$ (45MHz). Outside of the FTS context, a QCL intra-pulsed spectrometer uses the laser downchirp to record a spectrum in time. For this method, the effective resolution of the system is larger than the instantaneous linewidth and set by the time resolution of the acquisition system (Section 2.10.3). Normand et al. (2003)

estimated the resolution of their spectrometer's chirped spectrum by multiplying the time resolution – limited by the oscilloscope bandwidth - with the chirp rate, giving 0.013cm^{-1} (390MHz). The spectrometer's sensitivity was estimated by diluting the gas towards the baseline noise level. A concentration limit of 500ppb was estimated for CF_2CH_2, for the accessible linestrengths, and 30ppb for COF_2.

This first QCL based spectrometer used a Peltier cooled pulsed-QCL at a nominal wavenumber of 10.26μm. (This was the same laser used in Normand et al's original paper (2001) but with the proper impedance matching). An in-house pulse generator could supply top-hat current pulse durations of 3ns to 200ns. Lack of significant broadening due to averaging indicated that the pulse chain was highly reproducible. Typical operating parameters for this spectrometer were 4.6A 200ns pulses (~1.4cm^{-1} coverage). The 0.75Torr gas sample was sealed in a White cell (White, 1942). This is an arrangement of a larger concave spherical mirror facing two smaller ones on pivots – in this case configured to give 16 beam passes through the gas, giving an absorption length of 9.6m. The detector was a high speed cryogenically-cooled photovoltaic MCT detector connected to a 500MHz oscilloscope.

This work formed the basis of a patent application (Langford et al., 2003) and the founding of Cascade Technologies Ltd in 2003 by Normand, Howieson (two alumni of the MS Group), and Cooper (see Cascade Technologies Ltd, 2014a).

The MS Group continued improving and testing the capabilities of their intra-pulse spectrometer (McCulloch et al., 2003). The sensitivity was increased by replacing the White cell with an astigmatic Herriott cell (Herriott et al., 1964, Herriott and Schulte, 1965, McManus et al., 1995) (Section 2.8). This arrangement of two identical concave spherical mirrors allows a large number of passes, with each mirror having two perpendicular radii of curvature - allowing complicated Lissajou patterns of non-overlapping beam spots. The two patterns used here produced path lengths of 18m and 101m, with mirrors designed earlier by Howieson (1997). Finding suitable spot patterns was facilitated by an in-house ray tracing program based upon McManus'

method (McManus et al., 1995). Electrical shielding was introduced to minimize distortion in the recorded signal caused by electrical pick-up between the pulse generator and detector. This was not a problem with the longer pathlength due to the time difference between current pulse and radiation detection. The cryogenically cooled fast MCT photovoltaic detector used was from Kolmar Technologies (KV104-0.25A).

The time to wavenumber relationship was characterized via the Fabry-Perot fringes of a germanium etalon (equally spaced in frequency). The linear downchirp for 100ns pulses was found to need a small quadratic correction for 200ns pulses and a small cubic correction for 300ns pulses. The latter pulses could cover a spectral range of ~2.5cm^{-1}. Typical operating parameters were 4-5A 200ns pulses (~1.4cm^{-1}) with 5kHz repetition rate, and 4096 averages. Most experiments were conducted at less than 0.1Torr. The specificity of the spectrometer was considered with atmospheric trace gas sensing in mind. It was found that the spectrometer could adequately distinguish between the similar dense molecular fingerprints of 1,1-difuoroethylene (CF_2CH_2) and carbonyl fluoride (COF_2). Spectrometer sensitivity was tested qualitatively by recording ambient laboratory air for weak linestrength water vapour (H_2O) and carbon dioxide gas (CO_2). The gas pressures in the cell were conditioned to be between 50 and 200Torr. The reduced pressure broadening increases selectivity while maintaining the mixing ratios. The change of CO_2 signal with breath blown into the cell could be clearly measured.

This paper (McCulloch et al., 2003) also introduced a formula (Equation 2-70) showing that the best resolution attainable was proportional to the square root of the chirp rate, and that the intra-pulse method, as well as the inter-pulse method, could get close to this limit. The effect of a reduced chirp rate on resolution was demonstrated by lowering the amplitude of the top-hat current pulse, which reduces the heating of the active layer and slows the chirp. Longer pulses with a non-linear time-wavenumber relation will also have better resolution at the end of a pulse, as the heating effect slows towards equilibrium.

The attainable resolution leading to Equation 2-70 is limited by a balance between the chirp rate and the bandwidth theorem (Section 2.10). These considerations came out of Kosterev's inter-pulse work (Kosterev et al., 2000). However, the formula is applicable to any frequency scanned system (although the chirp rate will not always be the limiting factor on resolution). A statement of this formula is found much earlier in the literature in the context of Pb-salt diode laser spectroscopy (Gorshunov et al., 1979). Because the time resolution obtainable is limited by the bandwidth of the detection system, McCulloch et al used a higher specification oscilloscope of 1.1GHz bandwidth. This oscilloscope also allowed a maximum of 4096 averages for noise reduction.

Meanwhile, Cascade Technologies had released their first QCL system in 2004. In 2005, trials for marine emission monitoring began, demonstrating on-stack continuous emission monitoring (CEM) capability in harsh environments. Demonstrations also revealed the QCL sensor's ability to detect explosives (Cascade Technologies Ltd, 2014c).

By 2005 (McCulloch et al.), the gas cell had been replaced with one which allowed gas flow - allowing real-time measurements to be taken as gas concentrations changed. The gas outlet and inlet were located away from the end mirrors. This design gives the cell mirrors some protection from harsh gases but produces a non-uniform gas distribution in the cell – the concentration reading is therefore only a spatial average of the concentration in the cell. The dwell time of gases in the cell was estimated to be 2s judged by the mass flow rate and cell volume. The Lissajou patterns of the beam spots gave absorption path lengths of 66m and 110m. The acquisition system had become computer controlled, using a Femto HS-Y-1-40 1.1GHz amplifier and a 500MHz 2Gs Acqiris AP200 digitizer capable of 64,000 averages. As well as the usual pulsed-QCL at 10.26μm, another at 7.84μm was also used. This latter laser had a slower chirp rate and hence the potential for higher resolution.

This spectrometer was first used with the 10.26μm laser and a high chirp rate – hence low resolution ($0.12cm^{-1}$) – to measure ethylene (ethene, C_2H_4). Cell pressures were 2.4mTorr and at 35Torr. At the lower pressure, non-linear effects were reported (see below). The sensitivity for this gas was found to be 1.5ppm (66m path length). This was estimated both by dilution with nitrogen and in noting observable weak absorption lines – those with small dipole transition moments or those due to hot bands (transitions other than from the ground state). The ethylene sensitivity was projected to improve to 250ppb by increasing the path length and averaging. The 7.84μm laser was used to investigate nitrous oxide lines. A near $\sqrt{N}$ improvement was shown by taking averages up to 64,000, and a sensitivity of 0.03ppm estimated (110m path length). Both ethylene and nitrous oxide sensitivities had the potential to be improved by using lasers with tuning ranges covering larger linestrength transitions. The paper went on to study car exhaust emissions with the 10.26μm laser - measurements with applicability to trace gas atmospheric pollution and engine control and optimization. The evolution of ethylene and carbon dioxide, in different age and capacity engines, could be followed in real-time. Typical operating parameters for these latter experiments were 5A 140ns pulses ($1.4cm^{-1}$ coverage) at a repetition frequency of 5kHz and effective resolution of $0.014cm^{-1}$. 5,000 averages were used, and the cell pressure was 35Torr to quench lower pressure nonlinear effects. 2005 also saw the MS Group publish a tutorial review paper (Duxbury et al.) This summarized inter-pulse and intra-pulse QCL spectroscopy and covered some spectroscopic applications and nonlinear effects.

The system described above was used to measure trace gases in the ground level atmosphere, exhaled breath components, petrol vehicle emissions, and products in tobacco vapour (McCulloch, 2006). Atmospheric results were in broad agreement with accepted ambient levels. The 10.26μm laser measured water and carbon dioxide (115m path length, 50Torr; 232ns 5kHz pulses, 64,000 averages), and the 7.8μm laser measured methane, water, and nitrous oxide (101m path length, 1270ns and 1500ns 20kHz pulses, 64,000 averages). The 7.8μm laser was used for the non-invasive analysis of four patients for methane, carbon dioxide, water, and nitrous

oxide (115m path length; 50Torr; 1270ns 20kHz pulses, 20,000 averages). Exhaled breath was transferred to the spectrometer by balloon, but if a spectrometer attachment was made, these experiments demonstrate the potential of producing a real-time non-invasive breath analysis. For tobacco vapour, identification was carried out by comparing the pulsed-QCL measured vapour to pulsed-QCL spectra or FTS spectra of the pure component gases. The 10.26μm laser measured ethylene, ammonia, carbon dioxide, and water (66m path length, ~20Torr to 2Torr during an evacuated 'puff' (begun by briefly opening a valve to the measurement cell), 140ns 10-20kHz pulses, 5000-10,000 averages). The 7.8μm laser measured formaldehyde, methane, acetylene, water, and carbon dioxide (115m path length, ~20Torr to 2Torr during an evacuated 'puff'; 1400ns 20kHz pulses, 20,000 averages). The intensity between the background and spectrum baselines also gives information about the particulate concentration – that is, the scattering of radiation out of the beam path.

The spectrometer was used to measure gases from the exhaust of various petrol and diesel vehicles. The 10.26μm laser covered ethylene, water, and carbon dioxide lines (66m path length, 35-50Torr; 140ns 5kHz pulses, 5,000 averages). The 7.8μm laser covered methane, nitrous oxide, carbon dioxide, water and acetylene lines (115m path length, 35-50Torr; 1270ns 20kHz pulses, 20,000 averages). Experiments effectively tested the efficiency of the catalytic converters in real-time as they heated up to operating temperature.

This work was followed by the building of a spectrometer mounted onto an aluminium plate (optical breadboard) (Wright et al., 2006). The plate was designed for attachment to the seating rails of a Dornier 228 aircraft flown out of Oxford by NERC's Airborne Research and Survey Facility (NERC ARSF, 2012). The 7.84μm laser was utilized to measure methane, water vapour, and nitrous oxide from this plane - all within the same wavenumber downchirp. Ground based measurements were also undertaken near the city centre of Glasgow. The detector used was a cryogenically cooled reverse biased photovoltaic MCT of bandwidth 500MHz. A new astigmatic Herriott Cell was used, set to a path length of 28m, with gas inlet and

output ports at either end faces of the cell. (The optical part of this arrangement was available to be brought back into service and added to during Chapters 4 and 5 of the present thesis). Gas input was controlled via a mass flow constant-pressure controller and solenoid valve. This valve also had an input for a 'zero air' supply (dry air filtered to minimize hydrocarbon pollutants). The zero air was used to purge the cell periodically for background and etalon spectra to be taken. The gas cell pressure was 100Torr to balance the effects of pressure broadening on the absorption lines, and allowing them to be fitted with a relatively simple Lorentzian lineshape, while quenching low pressure nonlinear effects. It was evident that methane and nitrous oxide could be satisfactorily measured within the same dynamic range – as the smaller methane transition moment was compensated for by its larger concentration in the atmosphere. Pulses were 1,300ns (~3.4cm^{-1} coverage) at a repetition rate of 10kHz. 20,000 raw spectra were averaged at a time and the averaged spectra stored for post-flight processing. The signal was amplified by an AC-coupled 2GHz amplifier and fed to the Acqiris digitizer used previously. The aircraft spectra suffered from radio frequency feedback, thought to be from an electrical ground loop, and this needed to be digitally filtered out in post-processing.

Three more similar flight tests took place in 2006 (Hay et al., 2008), consisting of the legs of a return flight and a circuit. The spectrometer was substantially the same but with better electrical isolation, a 100m path length, and a cryogenically cooled photovoltaic MCT detector - which didn't require reverse biasing to obtain a high bandwidth (a higher bandwidth Kolmar Technologies KV104). The repetition frequency was increased to 20kHz, with, again, 20,000 averages per stored spectrum. Repetitive noise was found in the spectra, attributed to digitization noise from the Acqiris card, and this was removed in post-processing. The measured water concentrations correlated well with the weather conditions along the flight path, and the methane concentrations spiked around pollution hotspots such as airports and major roads. It was also noted that the speed of spectra acquisition led to near vibration immunity in the signal to noise ratio, comparing well with laboratory spectra. In addition, this system was used for measuring nitrous oxide in the

atmosphere at a few hundred ppb levels, before and after the spreading of nitrate fertilizers on nearby fields.

Cascade Technologies presented their work on 10ms detection of trace amounts of explosive vapours in ambient air (Normand et al., 2006). These explosives were nitroglycerine (NG, $C_3H_5N_3O_9$) and its breakdown products, and the detection taggant ethylene glycol dinitrate (EGDN, $C_2H_4N_2O_6$) – a compound added to many explosive chemicals to make them easier to detect by slow evaporation. This technology was placed into a people screening portal (similar to an airport metal detecting portal, but blowing ambient air across the target into the spectrometer), and tested at Glasgow airport (Lindley et al., 2007). A review article was written highlighting the state of QCL growth technology (Normand and Howieson, 2007), as companies began commercializing the lasers for standardized volume production. The company's collaboration with manufacturers to meet standard and quality control targets was highlighted as important in building customer confidence in the new QCL sensors against incumbent technologies.

In 2007 Cascade Technologies released the CT2000 4 laser pulsed-QCL system for industrial gas analysis (Cascade Technologies Ltd, 2014c). In these systems, the four laser chips were soldered directly onto the main control board. The MCT detectors operated at room temperature by thermoelectric (Peltier) cooling.

Early 2008 also saw the MS Group's cryogenically cooled detectors replaced by thermoelectrically cooled MCT photovoltaic detectors from VIGO Systems (Hay, 2010). These higher bandwidth detectors had an immersion lens and built-in (AC-coupled) transimpedance amplifier - giving an improvement in signal to noise ratio. The thermoelectric (Peltier) cooling reduces instrument size - since no liquid nitrogen dewar is needed. It also opens up the possibility of long term unattended operation.

A further flight out of Gloucester airport measured methane, nitrous oxide, carbon dioxide, carbon monoxide and water vapour (Hay, 2010). This was done with a Cascade Technologies CT2400 4 laser pulsed-QCL system, each laser pulsing at 5kHz, with a 76m path length astigmatic Herriott cell. The cell was operated between 100 and 150Torr, controlled by a pressure regulator and built-in pump. The detector was a thermoelectrically cooled VIGO with built in transimpedance amplifier. Digitization used a Cascade AV1500 digitizer. The inbuilt software used 500 averages per concentration value, displaying these concentration values in real-time.

This system had not been designed with aircraft measurements in mind, and problems with the 'on-the-fly' readings were found due to laser power fluctuations at the detector. These were mainly attributed to the laser being influenced by large temperature changes within the casing - causing mismatches between the intensity levels of background and absorption measurements - from which the gas mixing ratios are derived. Vibrations and thermal creep were also thought to lead to mirror misalignment, hence intensity fluctuations at the detector and randomized mixing ratios. The relative weakness of the internal pump also smoothed out fast concentration changes due to longer gas dwell times in the cell. However, post-processing of data improved the results. This CT2400 was also used to test the exhausts of various diesel bus engines. Gases measured were nitrogen dioxide, carbon monoxide, carbon dioxide, nitrous oxide, methane, formaldehyde (H_2CO), nitric oxide, and water vapour (Hay, 2010).

By 2009 Cascade Technologies were supplying large companies with multiple sensors: for instance in industrial plant, marine emissions, the automotive industry, and in the production of domestic boilers. They also received the ISO 9001 Quality Management Accreditation System (QMS) certification.

The MS Group also collaborated with the 'Laser Chemistry, Spectroscopy and Dynamics Group' at Bristol University (Cheesman et al., 2006, Ma et al., 2009). A 7.8μm DFB pulsed-QCL was used to make line-of-site measurements across a

microwave reactor - used for the plasma enhanced chemical vapour deposition of diamond. The laser was mounted in a development head and driver developed by Cascade Technologies. The radiation was collected, in the first paper, by a cryogenically cooled photovoltaic detector (Kolmar Technologies KV104-0.1-1A-3). The signal was processed by a fast amplifier (Femto Messtechnik HAS-Y-1-40) and a fast digitizer (Acqiris DP210). In the second paper, a thermoelectrically cooled photovoltaic detector was used, with a built-in transimpedance amplifier (VIGO). The integrated absorbances of methane (CH_4) and acetylene (C_2H_2) lines were measured. These line-of-site measurements do not spatially resolve the gas distribution along the line-of-sight. However, these stable molecules are expected to mainly reside in the cooler outer regions of the plasma reactor. When combined with models of the whole reactor gas distribution and its thermo-diffusion, the measurements can provide insight into the chemistry occurring throughout the reactor, and test the validity of the modeling. The real-time nature of intra-pulse spectroscopy also allows the rise to, and fall from, the steady state gas distribution to be monitored.

In 2010, Cascade Technologies signed an exclusive worldwide supply agreement with a major company in the automotive test market. The feedback from this company had helped to develop the CT3000 system (Chapter 7) (which was marketed generally in 2011). This is another 4 laser pulsed-QCL system with thermoelectrically cooled detectors. However, each laser is housed in its own module for ease of on-site laser replacement - for repair or swapping out with a different laser to detect different gases. By this time, Cascade Technologies was also supplying an aerosol container leak detection system to processing plants, for conveyer belt use. Cascade Technologies also gave a presentation highlighting their systems' capabilities for explosive and narcotic detection (Normand et al., 2010). A review article was written (Howieson, 2011) highlighting efforts to increase the reliability of QCL lasers. The lasers and cooling elements were now being housed in robust hermetically sealed packages. A brief outline of the CT3000 gas analyzer, its method of line fitting, and its applications was given. Cascade Technologies

highlighted their ability to make their own spectral databases via a high resolution FTS ($0.004cm^{-1}$) - to make up for deficiencies in publicly available databases – particularly at high temperatures.

The MS Group conducted open path measurements for carbon monoxide, carbon dioxide, and water vapour, through the exhaust plume of a turbojet engine (Duxbury et al., 2011). For these, a 4.86μm DFB pulsed-QCL was used at 20kHz repetition frequency with 12V 1300ns pulses (~$3.8cm^{-1}$ coverage). Chirp rate dependent resolutions varied from $0.013cm^{-1}$ to $0.005cm^{-1}$. Instead of a gas cell, the beam was sent through the plume and back via a corner cube retroreflector giving a physical path length of 1m. Detection was via the Peltier cooled VIGO and its transimpedance amplifier to the Acqiris AP200 digitizer taking 2600 points per spectra. Variously, 8000 averages, 0.4s acquisition time, or 4000 averages, 0.2s acquisition time, were used to produce each final spectrum. The composition of the exhaust plume and the time evolution of the three combustion gases were followed as a function of engine speed rpm, engine temperature, and horizontal height through a cross section of the plume.

The second avenue of research, pursued by the Strathclyde MS Group, concerned the non-linear regime where the Beer-Lambert Law does not hold (Equation 2-30). In this regime, the lower state population is significantly reduced by the radiation so that the absorbance is no longer linearly proportional to the intensity of further radiation. Non-linear effects can include: rapid passage, (macroscopic) molecular alignment, power (or saturation) broadening, and self focusing.

Rapid passage can occur with the high chirp rates of pulsed-QCL's, in low pressure gases. In this case, the rate of population transfer to the excited state can be much faster than the collisional relaxation time. The population of excited molecules is then significant and those molecules are then not available to absorb further incoming radiation. As an absorption line is swept over, the resulting asymmetrical signal consists of an absorption dip followed by an emission spike - as the excited

population relaxes and adds photons to the incoming radiation. This is followed by optical nutation – a dampened oscillatory structure of absorption and emission (Allen and Eberly, 2003). Relatively slow chirp rates, which are still faster than the relaxation time, can, however, enhance the population transfer through a longer interaction time between the field and the transition. This also promotes a macroscopic polarization as the molecule's electric dipole moments have time to align with the (linearly polarized) field. Alignment may also occur during the slow tuning ramp. The consequent de-phasing of the macroscopic polarization can result in free induction decay (FID) oscillations in the signal.

The above type of non-linear signals were previously seen in the context of nuclear magnetic resonance (nmr) spectroscopy (Ernst, 1966). They could be analysed in that context by coupled differential equations relating the components of a rotating 'Bloch vector'. The three Bloch vector components correspond to the three components of the system's magnetization (the sum of the individual nuclei's magnetic dipole moments per unit volume) (Siegmann, 1986). A molecular gas can be analysed in a similar fashion with the 'Optical Bloch equations'. These relate to a vector constructed of components proportional to the population difference, the component of the system's polarization in-phase with the radiation field (the dispersive component), and the in-quadrature component of the polarization (the absorptive component), respectively (Allen and Eberly, 2003). The 'wiggles' in the optical case are more suppressed than in the nmr case, and this has been modeled by treating the gas' molecular oscillators as a phased array (Duxbury et al., 2005, McCulloch et al., 2006). The Doppler inhomogeneous broadening implies that different velocity groups of molecules are excited sequentially as the laser frequency chirps.

Another non-linear phenomenon is power broadening, favoured by strong radiation fields and large molecular dipole moments. The field strength dependent Rabi flopping frequency (Equation 2-16, Section 2.4) can be faster than the collisional relaxation rate. A fast flopping frequency can enhance the upper state population,

particularly when combined with a relatively slow chirp rate, but also implies that there is a shortened time before a downward shift in population. This reduces the average lifetime of the upper state, resulting in an apparent frequency broadening of the transition (Bernath, 2005, Demtröder, 2008a).

The MS Group first noted nonlinear absorption in their 2005 paper (McCulloch et al.). They observed the characteristic rapid passage (emission) overshoot in a spectrum of strong ethylene transitions at 2.1mTorr with a chirp rate from their 10.26μm pulsed-QCL of ~200MHz ns^{-1} (6,671cm^{-1} ms^{-1}). This behaviour was explored further in the Group's tutorial review article (Duxbury et al., 2005) and laid out in more detail later (McCulloch et al., 2006), building on the work of Ernst and others. 5A pulses were used to increase the laser chirp rate to ~300MHz ns^{-1} (10,007cm^{-1} ms^{-1}) covering 1.4cm^{-1} with a 140ns electrical pulse. They switched to nitrous oxide for these later studies since its spectrum is less congested than those of ethylene. The lineshape was found to be a function of chirp rate, transition dipole moment, and gas pressure (collision rate). They postulated self focusing of the laser beam through the gas, to account for the large field strength, and hence Rabi flopping frequency, needed to reproduce the lineshapes seen.

Self-focusing of a laser beam can occur when the field strength is sufficiently strong to make the molecular oscillations large enough to be significantly anharmonic. For the majority of materials, the production of secondary radiation wavelets leads to a reduced phase velocity of the resultant wave (an increase in refractive index). For a Gaussian beam, the intensity profile is more intense at the line centre. This slows the beam more than at the edges – resulting in a curved wavefront, akin to passing through a focusing lens. An electromagnetic wave is still subject to diffraction and consequent expansion as it travels, even in vacuum, due to the production of secondary wavelets, so the wave will not totally collapse unless the field strength is above a critical parameter.

A further study on acetylene was presented in 2008 (Duxbury et al.), a review article written (Duxbury et al., 2009), and presentations given (Duxbury et al., 2007, 2010). Further detailed studies of non-linear effects using pulsed-QCL's were undertaken in collaboration with the chemistry departments of the Università Ca' Foscari Venezia (Tasinato et al., 2010a, 2010b), Oxford University (Northern et al., 2010, 2011), and with the Pacific Northwest National Laboratory (Duxbury et al., 2012a, 2012b, 2012c).

1.4 Thesis motivation and structure

Overall, this thesis contains a representative sample of work from the research side of the EngD programme. Much unfunded work was also undertaken on the analysis and programming side, to form a coherent body of work, and place the experiments in a form in which they could be taken further. The primary focus of the work was, firstly, Cascade Technologies' desire to seek a limit of detection (noise floor) improvement - in order to measure lower gas concentrations. This was sought by using a continuous wave quantum cascade laser (cw-QCL), with the noise reduction techniques of sweep integration and wavelength modulation spectroscopy. Secondly, the company wished to add oxygen to its range of measurable molecules. This was pursued with a wavenumber prototype semiconductor diode laser using sweep integration.

Chapter 2 sets the scene by introducing basic spectroscopic modeling equations and touches on some spectroscopic theory. Also covered is an introduction to the two types of lasers used in this thesis – quantum cascade lasers and semiconductor diode lasers, along with their detectors – a Mercury Cadmium Telluride photovoltaic semiconductor detector and an avalanche photodiode detector, respectively. A consideration is given of the astigmatic Herriott cell – a way of increasing absorption - and hence concentration sensitivity - by making the laser beam traverse multiple path lengths through a given gas sample.

There was a software gap for the automation of processing absorption spectra and extracting concentration values from pre-recorded data or data straight from an oscilloscope. The group at Strathclyde University was not orientated to this task, and therefore produced spectra on an ad hoc basis. The software of that group's associated company – Cascade Technologies – was tied to its real time pulsed-QCL systems.

Chapter 3 therefore lays out a data analysis program which was developed gradually in LabVIEW. It started as an exercise in working out the method of calculating concentration values, and became invaluable in extracting the information contained in Chapters 4, 5, and 6. The program therefore added greatly to the capability, functionality and process efficiency of what was otherwise available. As it stands, the program is also not far from being converted to a basic spectroscopic simulator. The fitting of Voigt profiles to experimental data is also not far off. Other useful software was developed and is introduced as appropriate.

Chapters 4 and 5 concern the noise reduction techniques of Sweep Integration and Wavelength Modulation Spectroscopy (WMS), respectively, as conducted with a continuous wave quantum cascade laser (cw-QCL). The experimental part of this work was done at Strathclyde University.

Cascade Technologies' gas analytic systems are based upon the technology of intra-pulse spectroscopy (a form of Sweep Integration) with pulsed Quantum Cascade Lasers (pulsed-QCL's) (Section 2.9.5.2). One continuous wave QCL was available as being suitable for use with the WMS technique - the continuous beam facilitates the relatively slow modulation frequencies used in WMS. The WMS technique is regularly used with semiconductor diode lasers in the near infrared – giving a sensitivity enhancement over the Sweep Integration technique (e.g. Bain et al., 2011). The ability to measure smaller concentrations confers competitive advantage by allowing customers enhanced process control and monitoring, and some future-proofing against legislation.

It would have been academically natural to use DC-coupled detection with a continuous wave laser. However, none suitable for integration in Cascade Technologies' systems were available. From a commercial point of view, there is advantage in using the simplest techniques and changes of equipment necessary. This keeps costs down and aids integration with existing systems. There is therefore some logic in using the detector which was available – an AC-coupled VIGO detector of the same type as used in the pulsed-QCL systems. Nevertheless, integration into existing pulsed systems would still be difficult. The cw laser has different electrical requirements to the pulsed lasers and the scanning and modulation frequencies are much reduced for the cw version.

The interest of Cascade Technologies was in the noise floor of the WMS system – that is, the absorbance Limit of Detection (LoD). However, it soon became clear that calculating absorbances would be affected by AC-coupled detection (Section 4.6.1). Moreover, the detector response to low WMS frequencies and electronic incompatibility issues also likely affected the noise floor. The approach taken in Chapters 4 and 5 was therefore to delineate the issues arising from the use of AC-coupled detection and the other available equipment.

It was also unknown how the cw-QCL structure - designed for stability - would respond to modulation. QCL's lack the fast tuning mechanism of current density tuning available to semiconductor diode lasers operating in the near-infrared or visible (Sections 2.9.2 and 2.9.4) – although diode laser tuning is itself dominantly thermal at low modulation frequencies. Chapter 4 therefore characterized the cw-QCL response via etalon fringes using the data analysis program of Chapter 3. Chapter 4 also established the noise floor of the Sweep Integration technique as a basis of comparison with the WMS of Chapter 5. Three data sets at decreasing number densities were taken to provide calibration of WMS voltages to absorbance units.

A model for the 2f-WMS lineshape from literature was coded into LabVIEW and is presented in Chapter 5. This could be used to compare how applicable the semiconductor diode laser model is to the greater asymmetry found in the cw-QCL WMS lineshapes. An Allan variance program was also constructed and coded and presented in Section 4.8. Because of a software gap, however, there was no way of taking continuous concentration readings, so the more rigorous technique of Allan variance could not be used to optimize the number of averages for minimum noise. (This was also limited by oscilloscope recording). Instead, Sweep Integration and WMS (with averaging) were compared for practical settings.

WMS needed supplemented with averaging - partly due to electronic compatibility issues with the AC-coupled transimpedance amplifier (Section 4.6.1). Discussions with several electronic engineers and technicians found that the resulting anomalous oscilloscope scaling (Sections 4.6.1.4 and 5.7) was outwith their experience. There was no opportunity to pursue this incompatibility experimentally, due to an abrupt loss of bench space. Unfortunately, the less noisy and more organized later data sets were contaminated by repeated and undeclared use of the equipment. Much time was lost into looking for non-existent leaks, and identifying what was spuriously thought to be sporadic outgassing from the components of the Herriott gas cell.

The experimental part of Chapters 6 and 7 were done at Cascade Technologies' premises. A semiconductor diode laser had been loaned which was a prototype for that laser manufacturer. Experimental time with the laser was short, but the laser's performance characteristics were explored, along with its potential for oxygen detection in the visible via sweep integration. Until then, Cascade Technologies had been exclusively concerned with molecules having their 'fingerprint' transitions in the mid-infrared – for which Lorentzian lineshape modeling was appropriate. In the visible, better accuracy is given by a Voigt profile fit. To this end a pseudo-Voigt model from literature was coded and compared to the measured spectra.

Chapter 7 concerns the setting up, construction of system specifications for, and the testing of a Cascade Technologies CT3000 CEM's sensor. There was no clear end use for the system given. However, it could have conceivably been later used with the fitting of an oxygen detecting laser, or using it for atmospheric measurements aboard a meteorology test plane. The approach taken was therefore to catalogue more fully what the detector could measure with the four pulsed-QCL component lasers available. In this way, if one or more lasers were replaced later, as much compensation as possible could be made with the remaining lasers. During this work, an apparently unique problem of Herriott cell contamination led to the design of a diagnostic check list. This could be used in more general circumstances (Section 7.5.2 and Appendix A). A contribution to process efficiency was also made by deciphering and re-writing a much used linearity testing spreadsheet (Section 7.6.1 and Appendix B).

Conclusions and suggestions for further work are made by chapter. Chapter 8 rounds off by summarizing the thesis' principle findings and the work's future outlook.

Chapter 2

Background Material

2 Background Material

2.1 Introduction

This thesis concerns laser absorption spectroscopy in the linear regime. This regime is characterized by radiation weak enough, and whose wavenumber tuning rate is slow enough, to not significantly alter the population of the gas molecules' lower energy level. Further incoming radiation therefore essentially finds the gas in its initial absorption state. It follows that the absorption is linearly proportional to the intensity of radiation – and is described by the Beer-Lambert Law (Section 2.6). Choosing a radiation source is often a choice between using semiconductor diode lasers in the visible / near-infrared, or quantum cascade lasers in the mid-infrared. In general the mid-infrared contains well delineated strong absorption lines – giving a 'molecular fingerprint' of a molecule. In the near-infrared, the lines tend to be weaker and less distinct – but the lasers tend to be well developed due to their use in the telecommunications industry. Chapters 4, 5 and 7 concern the mid-infrared, while Chapter 6 concerns the visible region neighbouring the near-infrared boundary.

This Chapter lays out several background topics relevant to the rest of the thesis. It begins by giving a definition of the spectral regions of interest. The succeeding section covers the P and R, and sometimes Q, branches seen in an absorption band - of which the laser's operating range only covers a small portion. This section also takes a foray into the subtle - and often hard to decipher - world of spectroscopic transitions and notation – those which are most relevant to the thesis.

This Chapter then progresses to discuss spectral lineshapes and basic modeling equations - in a form convenient to the computer modeling and fitting demonstrated in the rest of the thesis. The Chapter concludes with general notes on some of the equipment used in the thesis – lasers, detectors, and the astigmatic Herriott cell.

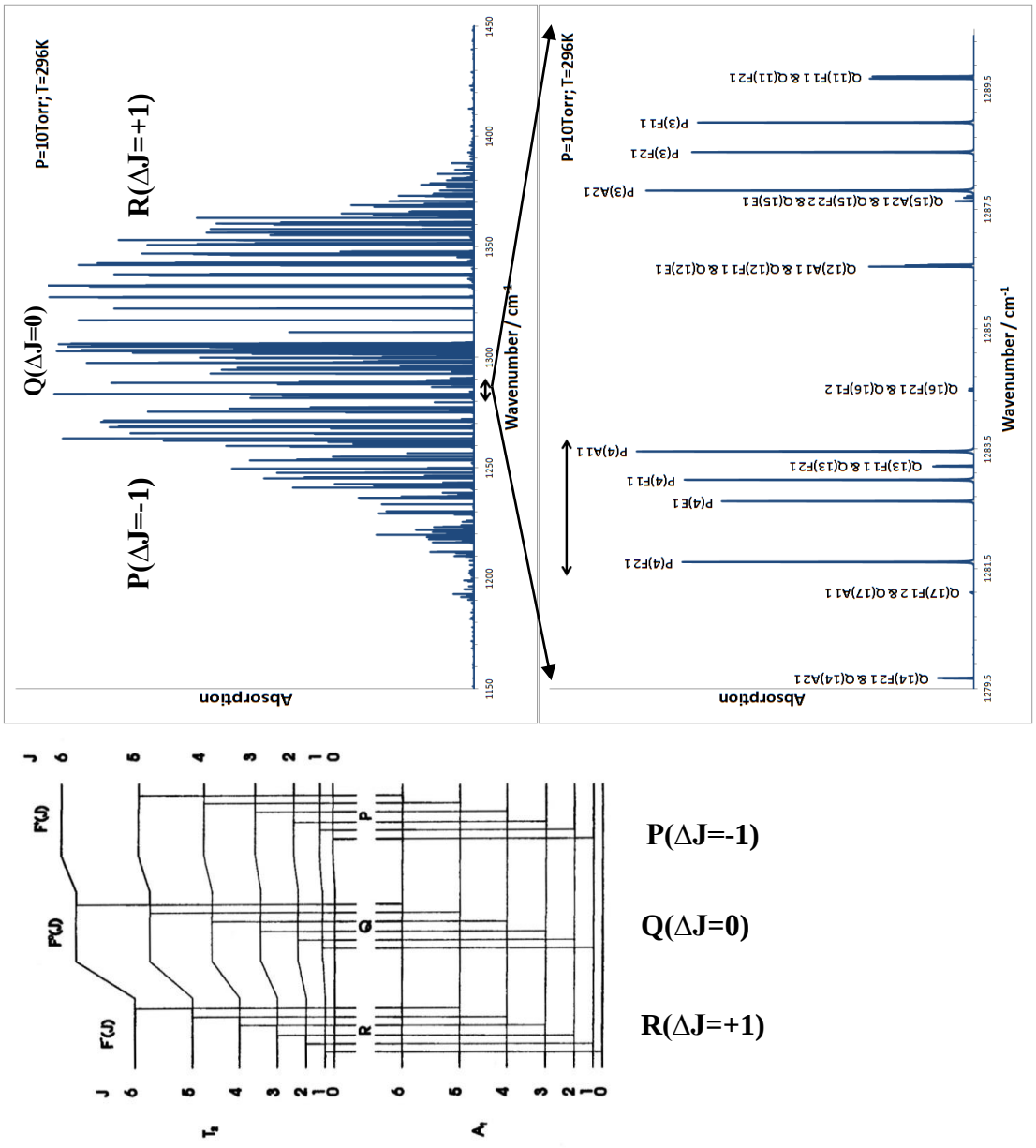


Figure 2-1: **Diagram illustrating a vibration-rotation band structure. At the bottom of the figure is a schematic diagram of the energy levels for a $T_2 \leftarrow A_1$ (or $F_2 \leftarrow A_1$) transition in a spherical top molecule such as $^{12}CH_4$. (Diagram from Bernath, 2005). At the top left of the figure are the resulting absorption transitions (at 10Torr and 296K) produced from HITRAN on the Web (V E Zuev IAO). The top right plot shows the small portion of this band which falls within the full operating range of the cw-QCL. The range used in Chapter 4 is shown by the horizontal arrow. The range covered in a single current sweep is significantly less.**

2.2 The molecular transitions

Since the electromagnetic spectrum is continuous, there is a degree of arbitrariness in its division into wavenumber regions. These divisions tend to relate to how the radiation can be detected and produced. Following the wavenumber choices of Hollas (2009) gives:

Mid-infrared:	400cm^{-1} to 4,000cm^{-1} (12THz to 120THz; 25μm to 2.5μm)
Near-infrared:	4,000cm^{-1} to 13,000cm^{-1} (120THz to 390THz; 2.5μm to 769nm)
Visible:	13,000cm^{-1} to 26,000cm^{-1} (390THz to 779THz; 769nm to 385nm)

The molecular transitions occurring in Chapters 4, 5, and 7 fit comfortably into the mid-infrared region. In this region absorption is due to transitions between quantized vibration-rotation energy levels, where each vibrational level has a stack of rotational energy levels associated with it.

The vibration-rotation band 0001 $^1F_2 \leftarrow$ 0000 $^1A_1$ of methane gas in Chapter 4 extends from 926cm^{-1} to 1,641cm^{-1}, with the main grouping of strong lines between 1,200cm^{-1} and 1,400cm^{-1}. Of these, the full operating range of the laser, set by substrate temperature and injection current, only samples a small proportion (Figure 2-1). The operating range actually used in Chapter 4 covered fewer lines, and the coverage in a single sweep is even less – dependent on the scan rate, and hence the laser's ability to tune with increasing injection current. The υ_4 $P(4)_{A2\ 1 \leftarrow A1\ 1}$ transition - used to compare noise reduction techniques - falls at 1283.458791cm^{-1} (Rothman et al., 2009).

In this mid-infrared spectral region, a quantum cascade laser (Section 2.9.3) was used as radiation source, with a HgCdTe photovoltaic detector (Section 2.11).

The Oxygen A-band in Chapter 7 falls between 12,849cm^{-1} and 13,340cm^{-1}, and thus straddles Hollas' boundary between the near-infrared and visible, with the main

grouping of lines in the visible between 13,000cm^{-1} and 13,340cm^{-1} (Figure 6-15). This places the R21Q22 transition, studied in Chapter 6, in the visible at 13162.50095cm^{-1} (Rothman et al., 2009). These higher transition frequencies are due to transitions between the electronic levels of the molecule, with associated vibrations and rotations.

2.2.1 Recognising band structure and notes on notation

The notation used for spectroscopic bands and transitions relates to the characteristics of the transition, and to the symmetry of the wavefunctions which describe the energy levels bounding the transition. Accounts of spectroscopy and the role of symmetry can be found in many sources – though there are many subtleties to be appreciated. The following account has been mainly constructed by reference to Banwell, Walton, Bernath, Hollas, Willock, and Atkins and Friedman (1966, 2004, 2005, 2009, 2009, 2011).

The molecular energy levels can be calculated from the Schrödinger equation, $\hat{H}\psi = E\psi$. A symmetry operation $\hat{R}$ belonging to the point group of the molecule brings the molecule into a physically equivalent state, and therefore one with the same energy. So, if ψ represents an energy level of the molecule with energy E, $\hat{R}\psi$ represents an equivalent energy state of energy E. Therefore $\hat{H}\left(\hat{R}\psi\right) = E\left(\hat{R}\psi\right)$. So, $\hat{R}^{-1}\hat{H}\hat{R}\psi = E\psi$ implying that $\hat{R}^{-1}\hat{H}\hat{R}\psi = \hat{H}\psi$ and $\hat{H}\hat{R}\psi = \hat{R}\hat{H}\psi$, which implies $\left[\hat{H},\hat{R}\right] = 0$. The Hamiltonian and symmetry operators therefore commute and a common set of energy and symmetry eigenfunctions can be found. That is, for a non-degenerate level, $\hat{H}$ and $\hat{R}$ have simultaneously the same eigenfunction ψ. For a degenerate level, a particular linear combination of the constituent levels' ψs is an eigenfunction of both $\hat{H}$ and $\hat{R}$.

This has two immediate consequences. Firstly, only wavefunctions which transform, under the symmetry operations of the point group, in a way which mimics the combinations of symmetry operations themselves are solutions to Schrödinger's equation. This mimicking can be achieved in more than one way – albeit a small number of ways. These ways appear as distinct rows in the point group's 'character table'. Secondly, since the eigenfunctions are common to the symmetry operations and the Hamiltonian operator, the energy levels can be labeled by the row in the character table that captures their transformation properties under the point group operations.

For instance, the $^{12}CH_4$ methane molecule has a high degree of symmetry and belongs to the tetrahedral point group T_d (or F_d). In the overall vibrational transition, $0001\ ^1F_2 \leftarrow 0000\ ^1A_1$, of the vibration-rotation band of Chapter 4, the transition occurs from the non-degenerate ground state of transformation symmetry A_1 to the triply degenerate vibration energy level of transformation symmetry F_2. (The '1' pre-superscripts distinguish between states which might otherwise have the same symmetry).

Another piece of information needed to understand the spectroscopic notation used in this thesis, concerns the molecule's ground state vibration. In general, a given molecule's ground state vibration will comprise a complicated Lissajou pattern. However, this can be decomposed into a finite number of 'normal modes' of vibration. Within a particular normal mode, the molecule's nuclei (to a first approximation) undergo harmonic motion with the same frequency, in phase or anti-phase, but with generally different amplitudes. Methane has 9 normal modes of vibration, but when degeneracy is taken into account, 4 distinct sets of vibrations remain.

The notation of the vibration-rotation band $0001\ ^1F_2 \leftarrow 0000\ ^1A_1$ of methane gas, in Chapter 4, indicates that a photon has been absorbed into the fourth vibrational mode of the molecule ($0001 \leftarrow 0000$). This mode keeps its frequency, and nuclei phase

relationships, but the amplitude of vibration increases, thus increasing the total energy of the vibration.

The whole of the band in the top right plot of Figure 2-1 is due to this same vibrational transition. The detail of which transition in the band is excited is due to which rotational levels are simultaneously involved. As mentioned, there are stacks of rotational levels associated with each vibrational level (top left plot of Figure 2-1). The structure of the band is largely determined by the change in total angular momentum quantum number J during the transition.

The changes in quantum number which can occur are determined by 'selection rules'. These come from a consideration of the matrix element of the perturbing Hamiltonian $\hat{H}'$ between the energy states involved. For interaction with a plane electromagnetic wave (Schiff, 1968, Bohm, 1989):

$$\left\langle \psi_n \left| \hat{H}' \right| \psi_m \right\rangle = \int \psi_n^* e^{i\mathbf{kr}} \nabla \psi_m \, d\tau \tag{2-1}$$

Here, $d\tau$ represents an infinitesimal volume element, and the $e^{i\mathbf{kr}} = 1 + i\mathbf{k}.\mathbf{r} - \frac{1}{2!}(\mathbf{k}.\mathbf{r})^2 + \ldots$ term comes from the perturbing radiation field. Symmetry arguments involving ψ_n and ψ_m can also help here, to determine when this integral is non-zero.

Since the wavelength (λ) of the electromagnetic field is usually much larger than the spatial extent of the wavefunctions describing the molecule, the **k.r** terms are usually small compared to unity ($k = 2\pi/\lambda$) where ψ_n and ψ_m are appreciable in the integral. Retaining only the unity term, using $\mathbf{p} = -i\hbar\nabla = m\left(\frac{d\mathbf{r}}{dt}\right)$ (where $\mathbf{r}$ is now a position operator for the molecule), and $\frac{d}{dt} < \psi_n \,|\, \mathbf{r} \,|\, \psi_m > = i\omega_{nm} < \psi_n \,|\, \mathbf{r} \,|\, \psi_m >$

derived from the Schrödinger equation; multiplying and summing by charges and dropping multiplicative constants, finally gives the electric dipole 'transition moment':

$$\mathbf{R}^{nm} = \int \psi_n^* \boldsymbol{\mu} \psi_m d\tau \qquad \boldsymbol{\mu} = \sum_i q_i \mathbf{r}_i \tag{2-2}$$

$\boldsymbol{\mu}$ is the electric dipole moment operator, where q_i and $\mathbf{r}_i$ are the charge and position of the i th electron or nucleus.

'Allowed' transitions refer to a non-zero value of the 'transition moment' in this electric dipole approximation - coupling to the radiation field's electric vector takes place via the molecule's dipole moment. 'Forbidden' transitions refer to zero values of the transition moment. However, allowed transitions may also have 'accidentally' zero or very low intensities. Transitions 'forbidden' by electric dipole selection rules may also occur by re-adding more terms in the $e^{i\mathbf{k}..\mathbf{r}}$ expansion - but with much reduced line strengths. The electronic transitions of the O_2 molecule in Chapter 6 are due to the magnetic dipole transitions. Both magnetic dipole and electric quadrapole transitions arise from including the $i\mathbf{k}.\mathbf{r}$ term in the plane wave expansion.

Selection rules for the ('spherical top') methane molecule are:

$$\begin{aligned} \Delta n &= \pm 1, \pm 2, \pm 3, \ldots \\ \Delta J &= 0, \pm 1 \end{aligned} \tag{2-3}$$

Δn refers to the transitions between vibrational energy states, with the transition between neighbouring states $\Delta n = \pm 1$ overwhelmingly the most probable for most molecules. ΔJ refers to which associated rotational levels can be involved.

In spectroscopy, changes in angular momentum quantum numbers are often denoted by letters. Those of relevance to this thesis are: $P \Rightarrow -1$, $Q \Rightarrow 0$, and $R \Rightarrow +1$. The $\Delta J = 0, \pm 1, \ldots$ spherical top selection rule gives rise to three sets of transitions within the vibration-rotation band: the P branch ($\Delta J = -1$), a relatively narrow Q branch ($\Delta J = 0$), and an R branch ($\Delta J = +1$). Across the P and R branches, the spacing between transitions declines as the wavenumber increases. This is due to anharmonicity. The vibrations themselves are anharmonic, and the rotational energy stacks associated with each are not equivalently spaced. The rotational spacings associated with a particular vibrational level are also influenced by centrifugal distortion. As the angular velocity increases, the nuclei separate. For a real molecule the restoring force is anharmonic, so the restoring potential is shallower than the harmonic case – closing up the energy levels as J increases.

The intensity of each transition is determined by the square of the transition moment, the degeneracy of the lower state, and the Boltzmann factor giving the probability of a molecule being in the lower state. The degeneracy and Boltzmann factor have strong opposing tendencies with increasing J. It is the balance between these two factors which determines the relative intensity distribution seen in the branches.

The specific vibration-rotation transition υ_4 $P(4)_{A2\ 1 \leftarrow A1\ 1}$ of Chapter 4 can be interpreted in the following way. The vibrational part of the transition has already been characterized as 0001 $^1F_2 \leftarrow$ 0000 $^1A_1$. The υ_4 similarly refers to this transition involving the fourth normal mode of vibration. The 4 in parentheses indicates that the lower rotational state has $J = 4$, and since the P indicates a $\Delta J = -1$ transition, the upper rotational energy level has $J = 3$. A_1 refers to the wavefunction of the lower state transforming according to the A_1 symmetry of the tetrahedral point group character table. Similarly, A_2 refers to the higher rotational state having A_2 symmetry in the character table. The additional '1's are 'multiplicity indices' needed to further distinguish between transitions which otherwise have the same labels.

A doublet in the 1110←0110 vibration-rotation band of nitrous oxide, $^{14}N_2{}^{16}O$, is explicitly identified in Chapter 4. This molecule has $C_{\infty\upsilon}$ point group symmetry, and is a heteronuclear linear polyatomic molecule with no inversion symmetry. The rotational selection rule for the transition considered here is $\Delta J = \pm 1$, so this band only shows P and R branches. N_2O has 4 normal modes of vibration, two of which form a doubly degenerate bending mode, leaving 3 distinct modes. The second vibrational mode is the degenerate bending mode, and the third number in the four number HITRAN notation refers to the modulus of its vibrational quantum number $|l|$ - which implies both $+l$ (clockwise) and $-l$ (anti-clockwise) This particular transition is therefore often seen in written form as $11^10 \leftarrow 01^10$. Here, a photon has been absorbed into the first vibrational mode. This notation allows the states for the transitions in this point group to be distinguished without recourse to symmetry labels.

The rotational part of the two doublet lines are denoted P9e and P9f respectively. The P puts the transition in the P branch ($\Delta J = -1$) and the number implies a transition from the total angular momentum quantum number $J = 9$ to $J = 8$. The 'e' and 'f' refer to the parity of the states' wavefunctions - that is, how the wavefunctions change sign, or not, under inversion.

The oxygen transitions of Chapter 6 are electronic transitions in the visible. Each electronic energy level has a stack of vibrational levels associated with it. The electromagnetic forces associated with a change in a molecule's electronic configuration will, in general, change the molecule's vibrational state as the nuclei adjust. Each vibrational level has a stack of rotational levels associated with it. If the vibration causes a change in the molecule's moment of inertia, through the average nuclear spacing, the rotational angular velocity may also change. The oxygen A-band concerns a vibrational transition $0 \leftarrow 0$, implying that the vibration part of the wavefunction changes from the ground state vibration of the lower electronic level, to the ground state vibration of the higher. The particular electronic transition itself is

denoted $b^1\Sigma_g^+ \leftarrow X^3\Sigma_g^-$. The $\Sigma^\pm$ symbols refer to the relevant symmetry transformation properties of the states - captured in the relevant row of the molecule's point group character table. The $^{16}O_2$ molecule is in point group $D_{\infty h}$, and is a homonuclear linear diatomic molecule with inversion symmetry. The rotational angular momentum, $\mathbf{N}$, selection rule for this transition is $\Delta N = \pm 1$, so that only P and R branches are seen. (See below for the meaning of the notation change). In addition, only transitions with odd values of N are seen, which relates to the symmetry of bosonic nuclei exchange. The g symbol (as opposed to u) implies symmetry to inversion through the molecule's centre (as opposed to anti-symmetry). (Note that g and u refer to the electronic state label as a whole, while the e and f labels above referred to rotational states).

The labeling of the electronic states themselves refers to a convention for diatomic molecules. The ground electronic state is labeled X, here X^3 denoting a triplet spin state ($S = 1$, $M_S = 0, \pm 1$). Higher energy states are labeled $A, B, C, \ldots$ in order of increasing energy - if they have the ground state's multiplicity (S value), otherwise $a, b, c, \ldots$. Here the higher electronic level is a singlet spin state, b^1 ($S = 0$, $M_S = 0$). HITRAN simply denotes this transition as $b0 \leftarrow X0$ where the zeroes denote that there are no quanta in the normal mode of vibration of this molecule. That is, the two vibrational ground states involved in the transition vibrate with their zero point energies only.

The interaction of the fields produced by the molecule's various angular momenta and charges help to determine the molecular energy levels which are possible in the solution of the Schrödinger equation. In quantum mechanics, the angular momenta add vectorially such that the total angular momentum, $\mathbf{J}$, has magnitude $|\mathbf{J}| = \sqrt{J(J+1)}\hbar$ and its projections on a measurement axis are restricted to integer or half-integer values in units of $\hbar$:

$$J_z = \pm J, \pm (J-1), \pm (J-2), \ldots, 1/2 \text{ or } 0 \tag{2-4}$$

From a reference frame moving with an orbiting electron, the charged nuclei orbit the electron. The nuclei's coulombic field appears as a magnetic field at the electron's position. The electron's magnetic moment, associated with its spin, interacts with this field - an interaction known as spin-orbit coupling. Since a nucleus' coulombic field is proportional to the nuclear charge, Ze, spin-orbit coupling is often weak and neglected for small and medium sized molecules. This is the case for molecular oxygen, whose coupling of angular momenta is modeled by 'Hund's case b'. The electrons are kept in orbit by the charges on the nuclei. In Hund's case b, the total electronic orbital angular momentum, $\mathbf{L}$, is strongly coupled to the coulombic field of the internuclear axis - about which it precesses rapidly. The component along the internuclear axis $\Lambda\hbar\mathbf{k}$ is well defined, but rapid precession means that the perpendicular component, and hence the magnitude of $\mathbf{L}$, is not well defined. The electrons follow the rotation of the nuclear framework, whose angular momentum is $\mathbf{R}$. $\mathbf{R}$ and $\Lambda\hbar\mathbf{k}$ are therefore coupled to give the angular momentum $\mathbf{N} = \Lambda\hbar\mathbf{k} + \mathbf{R}$. The total electron spin angular momentum, $\mathbf{S}$, then couples to $\mathbf{N}$ to give the total angular momentum $\mathbf{J}$.

The oxygen A-band has, by definition, a particular electronic transition and a particular vibrational transition. Since all the lines in the A-band involve a transition from triplet to singlet state, $\Delta S = -1$ for all lines. Since spin-orbit coupling is weak, the P ($\Delta N = -1$) and R ($\Delta N = +1$) rotational branches have a similar overall appearance to the vibration-rotation bands in the mid-infrared already discussed. However, the vibrational change associated with the electronic transition is sufficiently large that the reduction in transition spacing across the R branch goes to near zero at a 'band head' transition. After this, the energy gaps between higher N rotational transitions, associated with the different vibrational levels, then correspond to lower wavenumbers.

The individual lines are labeled by the relevant changes in the quantum numbers N and J and their lower state values N'' and J''. The format is: $\Delta N\, N''\, \Delta J\, J''$, where $\Delta N = \pm 1$ (only odd N allowed) and $\Delta J = 0, \pm 1$.

One of the oxygen lines explicitly identified in Chapter 6 is denoted R21Q22d. According to the convention above, this implies $\Delta N = +1, N'' = 21, \Delta J = 0, J'' = 22$. (Note that there is no Q-branch, which would occur for zero change in *rotational* angular momentum, $\Delta N = 0$). The 'd' suffix implies that this transition is due to a magnetic dipole interaction with the radiation (as opposed to the much stronger electric dipole interactions, or the even weaker electric quadruple interactions). The weakness of the lines is compensated for in Chapter 6 by the high abundance of oxygen in air at ≈20%.

2.3 Lorentzian, Gaussian, Voigt and pseudo-Voigt lineshapes

Accounts of the lineshapes relevant to this thesis - Lorentzian, Gaussian, and Voigt - can be found in spectroscopy textbooks (Thorne et al., 1999, Bernath, 2005, Demtröder, 2008a, Hollas, 2009). In addition, a popular approximation to the Voigt profile is covered below. Each of these lineshapes is defined by their centre wavenumber and full width at half maximum (FWHM). The centre wavenumber as a function of pressure is discussed within the section on the Lorentzian lineshape, and the FWHM's are discussed in Section 2.5 below.

2.3.1 The Lorentzian lineshape: natural linewidth and collisional broadening

All spectral lines have a finite width, including those due to a single molecule. At a fundamental level, 'natural broadening' is due to the wave nature of matter and radiation. The interaction of radiation with the energy levels of the molecule - giving rise to the absorption - can be described as a forced and damped harmonic oscillation of the interaction's dipole moment. The frequency dependence of the absorbance is

given by the imaginary part of the refractive index – and is Lorentzian. A Fourier transform to the time domain shows an exponentially decaying sinusoidal time signal as a photon is absorbed. Only an infinitely long sinusoid oscillates at a single frequency. Finite time signals can be considered to be synthesized from a spread of infinite sinusoids – this is Fourier's theorem (Hecht, 1987).

The classical bandwidth theorem links the characteristic frequency spread to the time which characterizes the duration of the signal. In terms of wavenumber:

$$\Delta\tilde{\nu}\Delta\tau \geq \frac{1}{200\pi c_0} \tag{2-5}$$

(Where c_0 is the numerical value of the speed of light in vacuum, if measured in ms^{-1}).

$\Delta\tilde{\nu}$ is often taken as the FWHM of the frequency lineshape, and $\Delta\tau$ is the time constant of the exponentially decaying signal. Since this interaction is identical for the same molecular transition, under the same conditions, it is termed homogenous broadening. The overall lineshape will therefore be a sum of the identical average contribution of each molecule.

The Heisenberg Uncertainty Principle gives the same result. The frequency spread is attributed to the finite width of the energy levels themselves, since the states they describe are not infinitely long lived. In both classical and quantum cases, the natural broadening is usually dominated by other broadening mechanisms.

The average lifetime of the molecule in the upper state of the υ_4 $P(4)_{A2\ 1\leftarrow A1\ 1}$ transition of methane, used in Chapter 4, is the reciprocal of the Einstein A_{lu} coefficient available from the HITRAN database (V E Zuev IAO), $\Delta\tau = 1/A_{lu} = 0.452s$. From Equation 2-5, $\Delta\tilde{\nu} = 1.17\times10^{-11}cm^{-1}$.

'Collisional broadening' also leads to a homogenous Lorentzian frequency spread. A full treatment would be very involved, needing to take account of the intermolecular potentials as a function of distance and time during a collision. However, the main effect can be considered via Equation 2-5. Since the collisions cause interaction between the molecules, they lessen the lifetime of a molecular state and thus also lead to an increased FWHM (frequency spread) of the transition's Lorentzian lineshape.

The form of a unit area normalized Lorentzian lineshape, $g_l(x-x_0)$, is given in terms of the centre wavenumber, x_0, and the FWHM, ω_l:

$$g_l(x-x_0)=\frac{2}{\pi\omega_l}\frac{1}{1+4\left(\frac{x-x_0}{\omega_l}\right)^2} \tag{2-6}$$

The values of x_0 and ω_l for a particular transition, for a particular gas, under particular conditions, is considered in Section 2.5.

How to calculate the FWHM of the line is given in Equation 2-27 below. ω_l is a function of pressure - as the higher the pressure the more collisions occur, and the shorter the lifetime of an energy state. Collisional broadening is therefore also known as pressure broadening – even though Equation 2-27 also has an explicit temperature dependence.

In the mid-infrared, at the pressures of 200Torr and above used in Cascade Technologies' systems (e.g. Chapter 7) the Lorentzian lineshape dominates over the Gaussian lineshape (Section 2.3.2). However, due to a Gaussian FWHM dependence on centre frequency (Equation 2-20), in the visible region (Chapter 6) the Gaussian FWHM is comparable to the Lorentzian FWHM at 200Torr, and a Voigt profile becomes appropriate (Section 2.3.3).

As well as broadening due to pressure, there is also a pressure shift of the centre frequency. For many transitions this shift is towards a smaller frequency and is linearly proportional to the pressure (Equation 2-19). This can be accounted for by considering two molecular energy levels involved in a transition. When the molecule is perturbed during a collision, both energy levels may have their potential energies increased or decreased in the same direction, depending on whether the interaction is attractive or repulsive. However, the higher energy level, in general associated with a more outlying state, may be more greatly affected – narrowing the gap between levels, and hence lowering the frequency of any absorbing transition.

2.3.2 The Gaussian lineshape: Doppler broadening

The Gaussian lineshape arises from inhomogeneous broadening – where groups of molecules in the same gas are affected differently. A common cause, in an otherwise common environment, is the thermal velocity distribution of the molecules. Since the molecules have a spread of velocities, even at thermal equilibrium, the Doppler effect ensures that the effective frequency of the radiation seen by a molecule will reflect the relative velocity distribution between the molecules and the radiation. Thus, in general, transitions occur at a wider range of frequencies than if the molecules were stationary with respect to the radiation. This is known as 'Doppler broadening' of the absorbance lineshape.

The appropriate form of the unit area normalized Gaussian distribution, $g_g(x-x_0)$, can be derived from the one dimensional Boltzmann distribution. This distribution is peaked at zero velocity, which coincides with the case of a molecule stationary with respect to the radiation – reflecting the fact that, in an isotropic gas, a molecule is as likely to be moving towards the laser beam as away from it.

$$g_g(x-x_0) = \frac{2}{\omega_g}\sqrt{\frac{\ln 2}{\pi}}e^{-4\ln 2\left(\frac{x-x_0}{\omega_g}\right)^2} \tag{2-7}$$

The FWHM, ω_g, (Equation 2-20) depends upon the centre frequency, x_0, and temperature. As the temperature increases there is more total energy available for distribution through molecular collisions, and the spread of speeds broadens. The Doppler shift itself, $\Delta\tilde{\nu}$, and therefore the spread of frequencies which can be absorbed, is proportional to the laboratory centre frequency of the radiation, $\tilde{\nu}_0$:

$$\Delta\tilde{\nu} = \frac{v_z}{c_0}\tilde{\nu}_0 \tag{2-8}$$

At low pressures, for instance the <2Torr used in Chapters 4 and 5, the Gaussian lineshape dominates the Lorentzian lineshape.

2.3.3 The Voigt lineshape

In the visible region, for the oxygen spectroscopy of Chapter 6, the centre frequency dependence of the Gaussian FWHM makes it comparable to the Lorentzian FWHM at 200Torr, so that neither lineshape dominates the other. In this regime it is usual to assume the Doppler and collisional broadening mechanisms are independent of each other. This makes a convolution of the two lineshapes an appropriate model. The resulting lineshape is known as the Voigt profile.

Convolution can be viewed as treating one of the lineshapes as being made up of delta functions at each wavenumber. The other lineshape's centre is then placed at each delta function and weighted. Since the lineshape is not limited to its centre frequency, this leads to a great many overlapping lineshapes at any wavenumber. The overall Voigt lineshape at any wavenumber is then the weighted sum of component lineshape contributions at that wavenumber.

Another, mathematically equivalent, way of viewing convolution seems physically less satisfying. This is to start by reflecting one of the lineshapes according to wavenumber (a trivial operation for symmetrical lineshapes). The reflected lineshape

is then given a wavenumber offset which is used to slide it across the other lineshape. The contribution to the Voigt profile at each wavenumber of the sliding offset is then the area of overlap of the two component lineshapes occurring for that offset. A detailed account of both points of view is given in Hecht (1987).

The unit area normalized Voigt profile, $g_v(x-x_0)$, is:

$$g_v(x-x_0) = (g_g * g_l)(x-x_0) = \int_{-\infty}^{\infty} g_l(x'-x_0)\,g_g(x-x')\,dx' = \int_{-\infty}^{\infty} g_g(x'-x_0)\,g_l(x-x')\,dx'$$

$$\Rightarrow g_v(x-x_0) = \frac{4}{\pi w_g w_l}\sqrt{\frac{\ln 2}{\pi}}\int_{-\infty}^{\infty} \frac{\exp\left[-4\ln 2\left(\dfrac{x'-x_0}{\omega_g}\right)^2\right]}{1+4\left(\dfrac{x-x'}{w_l}\right)^2}\,dx' \qquad (2\text{-}9)$$

In Equation 2-9 the Gaussian lineshape has been kept fixed about x_0, and the Lorentzian origin x' made variable so that the Lorentzian lineshape can be distributed along the wavenumber axis, while being weighted by the Gaussian function at x'. Equation 2-9 gives one value of the Voigt function for the single wavenumber x. That is, the integral needs to be repeatedly performed over all x' for each x value in the integrand - to build up the Voigt profile over all values of x.

Equation 2-9 has no closed form solution, but it is a relatively simple programming task to perform the integration, say by the trapezoidal rule. Denoting the integrand as $f(x'-x_0)$, and digitizing the integrand into N points of interval $\Delta x'$, implies the Voigt profile value for wavenumber x can be calculated as:

$$g_v(x-x_0) \approx \frac{4}{\pi w_g w_l}\sqrt{\frac{\ln 2}{\pi}} \times$$

$$\times \frac{\Delta x'}{2}\left[f(x_1-x_0)+ f(x_N-x_0)+ 2\{f(x_2-x_0)+...+ f(x_{N-1}-x_0)\}\right] \qquad (2\text{-}10)$$

Here $x' = x_1,\ldots,x_N$ represent the wavenumbers to be integrated over, and x_0 represents the centre frequency of the Gaussian distribution.

Although Equation 2-9 can thus be computed numerically, this gets more and more costly in time as the number of data points used for integration increases. A nested For Loop is needed – the inner to perform the integral for each Voigt point x, and the outer to run over the N wavenumber x values - so that the processing time scales as N^2.

2.3.4 A pseudo-Voigt lineshape

Since the numerical evaluation of the Voigt profile is computationally expensive, approximations to this lineshape are often used. A popular pseudo-Voigt model is given by Whiting (1968) for each wavenumber of the lineshape as:

$$g_v(x-x_0) = g_{v\,MAX}\left\{\left[1-\frac{w_l}{w_v}\right]\exp\left[-4\ln 2\left(\frac{x-x_0}{w_v}\right)^2\right]+\left[\frac{w_l}{w_v}\right]\frac{1}{1+4\left(\frac{x-x_0}{w_v}\right)^2}+\right.$$

$$\left.+\,0.016\left[1-\frac{w_l}{w_v}\right]\left[\frac{w_l}{w_v}\right]\left(\exp\left[-0.4\left|\left(\frac{x-x_0}{w_v}\right)\right|^{2.25}\right]-\frac{10}{10+\left|\left(\frac{x-x_0}{w_v}\right)\right|^{2.25}}\right)\right\} \qquad (2\text{-}11)$$

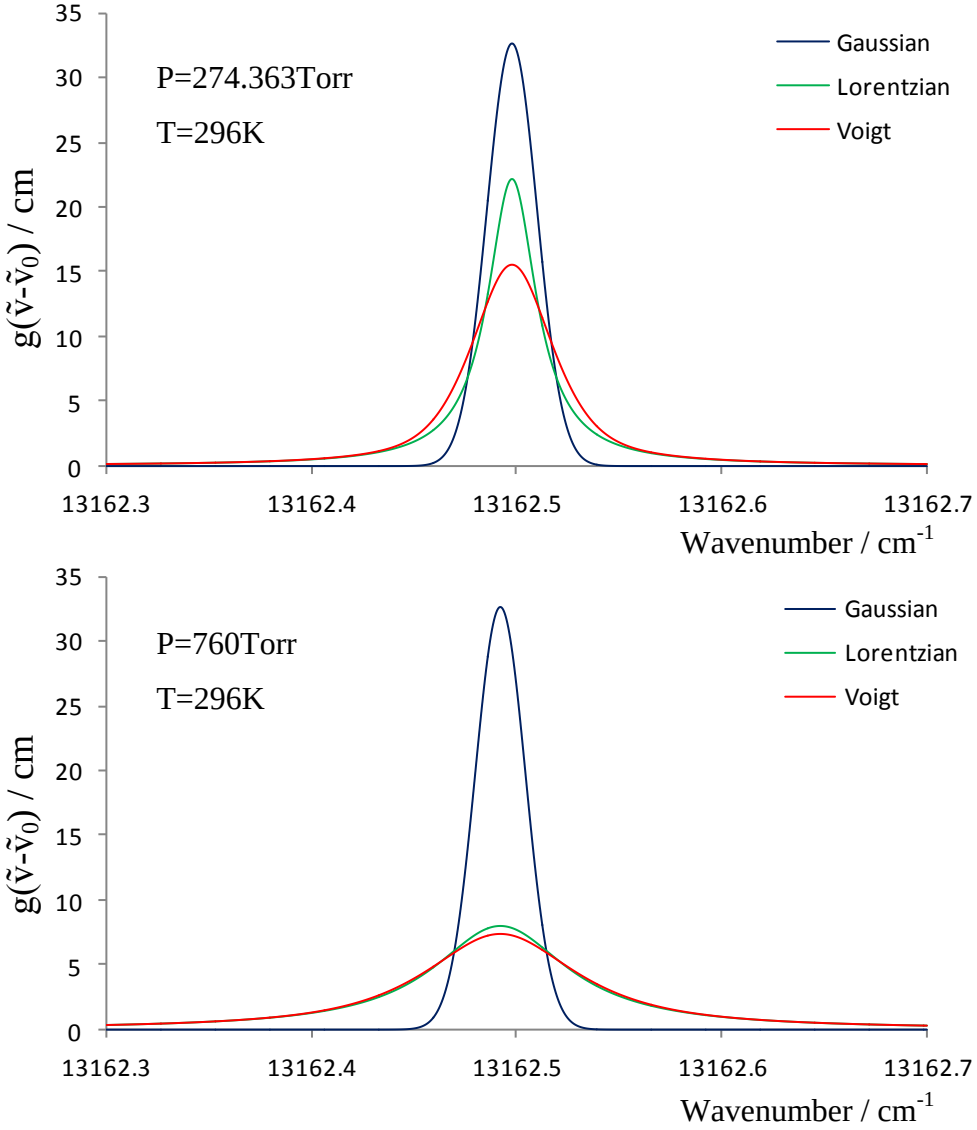


Figure 2-2: **Calculated unit area normalized Gaussian, Lorentzian, and Voigt lineshapes (for the A-band R21Q22 line of O_2 gas in laboratory air) at 296K and ≈274Torr and 760Torr, respectively.**

In the top plot at ≈274Torr, the Gaussian and Lorentzian lineshapes have the same FWHM. The Voigt lineshape has a FWHM less than the sum, due to the use of Equation 2-13 for FWHM. The Voigt lineshape has an intermediate shape, Lorentzian in the wings, but rising earlier than the Lorentzian near line centre. (Figure 2-6 models a similar result at the experimentally used 200Torr).

In the bottom plot, at atmospheric pressure, the pressure broadened FWHM of the Lorentzian lineshape is ≈2.77 times the Gaussian FWHM (which has only a temperature dependence). This gives a broadened Voigt lineshape to which the Lorentzian lineshape approximates well at this pressure.

Where:
$$g_{vMAX} = \frac{1}{w_v\left[0.058\left(\frac{w_l}{w_v}\right)^2 + 0.447\left(\frac{w_l}{w_v}\right) + 1.065\right]} \quad (2\text{-}12)$$

It also seems popular to combine Whiting's model with the empirical formula for the Voigt FWHM provided by Olivero and Longbothum (1977) (Section 2.5):

$$\omega_v = 0.5346\omega_l + \sqrt{0.2166\omega_l^2 + \omega_g^2} \quad (2\text{-}13)$$

Going back to consider Equation 2-11, the first two terms replace the Voigt convolution with a linear sum of Gaussian and Lorentzian lineshapes weighted according to their FWHM's. The last two terms are corrections for the lineshape wings.

2.3.5 A comparison of lineshapes

The top plot of Figure 2-2 shows unit area normalized Gaussian, Lorentzian, and Voigt profiles near 274Torr (T=296K). At this pressure the Lorentzian and Gaussian FWHM's are equal. The Voigt FWHM is less than the sum of these FWHM's due to the use of Equation 2-13 for ω_v. It can be seen that the Gaussian lineshape is more narrowly peaked around the centre wavenumber. The Lorentzian lineshape allows considerable absorption in the line wings. The convolution of these two lineshapes, the Voigt profile, takes an intermediate shape at 274Torr – rising earlier than the Lorentzian near line centre, but also broad like the Lorentzian in the wings. By 760Torr, the bottom plot shows that the Lorentzian lineshape is dominant. Since the Gaussian FWHM has not broadened with pressure, its relative influence on the Voigt lineshape has diminished. On the scale of Figure 2-2 the Voigt and pseudo-Voigt lineshapes are essentially overlapping.

2.4 Two other broadening mechanisms: transit time broadening and power broadening

Another broadening mechanism is transit time broadening. The effect is small but is included here to rule it out as a significant broadening mechanism in Section 4.6.4. The interaction time of a molecule with the laser beam is truncated if it travels at an angle to the laser beam. Maximum truncation occurs for molecules travelling perpendicular to the beam. The effect can be approximated from the bandwidth theorem, which implies a shortened interaction time is equivalent to, at least, a minimum increase in the apparent frequencies of the beam. For a monochromatic beam, assuming a boxcar intensity profile, gives an apparent sinc function broadening of the monochromatic frequency. An assumed Gaussian intensity profile gives an apparent Gaussian broadening of the monochromatic frequency.

The velocity, v, of an average methane molecule can be estimated from the one dimensional Boltzmann distribution (Waldram, 1989):

$$\mathrm{v} = \sqrt{\frac{k_B T}{m}} \tag{2-14}$$

k_B is Boltzmann's constant, $T = 296K$, and m is the mass of a methane molecule, giving $\mathrm{v} = 419.3ms^{-1}$. The minimum interaction time occurs at the TEM_{00} cavity mode beam waist. The beam waist, w_0, can be calculated as (Quimby, 2006):

$$w_0^2\left(\frac{L}{2}\right) = \frac{\lambda L}{2\pi}\sqrt{\frac{2r}{L} - 1} \tag{2-15}$$

$\lambda = 7.8\mu m$ is the nominal laser wavelength, $L \approx 50cm$ is the mirror to mirror spacing, and r is a radius of curvature of the mirrors. Each identical mirror had two perpendicular radii of curvature: $r_1 = 77.274cm$ and $r_2 = 67.425cm$ (Hay, 2010). r_2

gives the smaller beam diameter, $2w_0 = 1.80mm$. The transit time, $\Delta\tau = 2w_0/\mathrm{v}$, then allows the FWHM to be calculated, $\Delta\tilde{\nu} = 0.89/100c_0\Delta\tau = 6.99\times10^{-6}cm^{-1}$, where the bandwidth theorem applicable to a 'boxcar' beam function is $\Delta f\,\Delta\tau \approx 0.89$ (Bernath, 2005). For a Gaussian beam, $\Delta f\,\Delta\tau = 0.441$ (Demtröder, 2008b), giving $\Delta\tilde{\nu} = 1.71\times10^{-6}cm^{-1}$.

Power or saturation broadening is a nonlinear effect. When an electromagnetic wave, of field strength E, interacts with the dipole moment, μ_{ul}, of a (two level) molecule, the oscillating field cycles the probability of finding the molecule in the excited state (Bernath, 2005). The frequency and amplitude of this probability oscillation depends on how far the electromagnetic field frequency is detuned from resonance. On resonance, the probability of finding the molecule in the excited state cycles between 0 and 1 – the latter probability implying complete population inversion. And on resonance, the probability cycles at the Rabi frequency, $\nu_R = \mu_{ul}E/h$, where h is Planck's constant (in SI units). In the linear regime, molecular collisions dampen out these oscillations, giving a constant excited state probability of $1/2$ on resonance. However, if a large electric field strength causes the probability oscillations to be fast enough to beat the collision time, a molecule will on average only spend a short time in the excited state before a field induced transition. This shortening of the average upper state lifetime is associated with an increased frequency spread for the transition. The magnitude of the broadening can be approximated using the bandwidth theorem, $\Delta\tilde{\nu} \approx 1/(200\pi c_0\Delta\tau)$ (Equation 2-5). Following Bernath (2005), the average upper state lifetime, $\Delta\tau$, can be approximated by the reciprocal of the Rabi frequency:

$$\Delta\tau = \frac{1}{\nu_R} = \frac{h}{\mu_{ul}E} \tag{2-16}$$

In general, the Rabi frequency is not equal to the frequency of the electromagnetic radiation, nor the transition's centre frequency. For the υ_4 $P(4)_{A2\ 1\leftarrow A1\ 1}$ methane

transition in Section 4.6.4, the laser frequency and transition frequency are ≈38.5THz, and the calculated Rabi frequency is 1.95MHz.

The dipole moment (in Debye, 1D=3.33564×10^{-30}Cm) can be calculated using parameters available from the HITRAN database (V E Zuev IAO) and Equation 2-17 below (Rothman et al., 1998):

$$\left|\mu_{ul}\right|^2 = \frac{3hg_u A_{lu}}{64\pi^4 \tilde{\nu}_{ul}^3} \times 10^{36} \tag{2-17}$$

In this formula, h is measured in $erg\,s$ ($1erg\,s = 1\times10^{-7}\,J\,s$). g_u is the statistical weight of the upper level, A_{lu}, is the Einstein A coefficient, and $\tilde{\nu}_{ul}$ is the transition's centre wavenumber. The radiation's electric field strength can be calculated from the laser power, P, the beam waist, w_0,.and the standard equation relating intensity, I, to field strength (Quimby, 2006):

$$I = \frac{P}{\pi w_0^2} = \frac{1}{2} n c_0 \varepsilon_0 E^2 \tag{2-18}$$

n is the refractive index, which is ≈1 for a dilute gas. Power broadening calculations are laid out in Table 4-2 of Section 4.6.4, finding that $\Delta\tilde{\nu} \approx 10^{-5} - 10^{-6} cm^{-1}$ for the transitions and parameters considered there.

2.5 Basic pressure shift and FWHM modeling

To model the Gaussian, Lorentzian, and Voigt spectra in this thesis, line list parameters were taken from HITRAN on the Web (V E Zuev IAO). The basic spectroscopic modeling equations can be found in a number of sources (Spectralcalc.com, Rothman et al., 1998, Killinger and Plutov, 2009).

Firstly, the pressure shifted transition wavenumber (see Section 2.3.1) is calculated from its listed (zero pressure) transition wavenumber, $\tilde{\nu}_{0i}$:

$$\tilde{\nu}_0\left(cm^{-1}\right)=\tilde{\nu}_{0i}\left(cm^{-1}\right)+\delta\left(cm^{-1}\right)\frac{P}{P_0} \tag{2-19}$$

Here δ corresponds to HITRAN's 'Ashift' – 'the air broadened pressure shift' (Rothman et al., 2009). This single parameter is used whether the gas is pure or diluted in nitrogen. P_0 is a reference pressure with the same units as P. In HITRAN P_0 is always taken as $1atm$ and not shown explicitly. Accordingly the units for δ in HITRAN are cm^{-1}/atm, rather than cm^{-1} as here. The magnitude of δ there is the same as here.

The other parameter needed to reproduce the lineshape is the full width at half maximum (FWHM) - ω_g, ω_l, or ω_v. The Gaussian FWHM is a function of the centre wavenumber and temperature (discussed in Section 2.3.2):

$$\omega_g\left(cm^{-1}\right)=2\tilde{\nu}_0\sqrt{\frac{2(\ln 2)k_BT}{mc^2}}\approx 7.2\times10^{-7}\left(K^{-1/2}\right)\tilde{\nu}_0\left(cm^{-1}\right)\sqrt{\frac{T(K)}{M}} \tag{2-20}$$

m is the mass of a single molecule and M is the molecular mass number.

The collisional (pressure) broadened Lorentzian FWHM is modeled according to the proportionality:

$$\omega_l \quad \alpha \quad \left(\frac{P}{P_0}\right)\left(\frac{T_0}{T}\right)^{\gamma} \tag{2-21}$$

Here P_0 and T_0 are the reference pressure and temperature (T_0=296K in HITRAN). γ corresponds to 'Texp' in HITRAN – 'the coefficient of temperature dependence of the air-broadened halfwidth'.

The form of this proportionality can be found from a re-focusing of the usual classical mean free path argument of kinetic theory (Waldram, 1989, Kassander, 2009, Nave, 2014). Considering the gas molecules as hard spheres of diameter d, a given molecule will collide with molecules whose centres are d away from its own. The molecule therefore has an effective cross-sectional area of collision of πd^2. If the average relative velocity of the molecule is $\bar{\upsilon}_{rel}$ and the average time between collisions is τ_c, the effective volume swept out by the molecule until a collision is on average $\pi d^2 \bar{\upsilon}_{rel} \tau_c$. This will be equal to the average volume per molecule (the reciprocal of the number density):

$$\pi d^2 \bar{\upsilon}_{rel} \tau_c = \frac{V}{N} \tag{2-22}$$

So, from the bandwidth theorem and the ideal gas law:

$$\omega_l = \frac{2\pi}{\tau_c} = \frac{2\pi^2 d^2 \bar{\upsilon}_{rel} N}{V} = \frac{2\pi^2 d^2 \bar{\upsilon}_{rel} P}{k_B T} \tag{2-23}$$

The average velocity component of a gas molecule, of mass m, approaching a stationary molecule is given by Maxwell's one-component velocity distribution as:

$$\bar{\upsilon} = \sqrt{k_B T / m} \tag{2-24}$$

A correction factor is needed to give the relative average velocity since both molecules are in motion:

$$\bar{\upsilon}_{corrected} = \sqrt{\overline{\upsilon_2^2} + \overline{\upsilon_1^2}} = \sqrt{2}\bar{\upsilon} \tag{2-25}$$

$\bar{\upsilon}$ is the same for both molecules, and $\overline{\boldsymbol{\upsilon}_2.\boldsymbol{\upsilon}_1} = 0$ for uncorrelated motions. Substituting Equation 2-24 and Equation 2-25 into Equation 2-23 gives:

$$\omega_l = \frac{\sqrt{2}\left(2\pi^2 d^2\right)}{\sqrt{mk_B}} \frac{P}{\sqrt{T}} \tag{2-26}$$

It is worth noting that although collisional broadening is also known as pressure broadening, Equation 2-26 also includes an explicit temperature dependence. From this simple model the temperature dependence is $T^{-\gamma}$ where $\gamma = 0.5$. In practice, γ is found to vary from this fraction, and is line specific.

HITRAN splits the proportionality constant needed in Equation 2-21 into a sum of two constants. The first contribution to a finite pressure broadened linewidth comes from the air-broadened halfwidth, α_{air}. The methane gas of Chapters 4 and 5 was essentially pure and did not involve this factor (though this gas was sometimes contaminated, see Section 4.6.3)). The oxygen of Chapter 6 was in air, and the gases of Chapter 7 were diluted in Nitrogen (the main component of air). The second broadening coefficient is the self-broadened halfwidth, α_{self}, taking account of the collisions of the gas molecules responsible for the transition with each other. The relative influence of these factors is given by $q = N_i / N_T$ - the fraction of the component gas molecules to that of the total.

So, the Lorentzian FWHM is:

$$\omega_l\left(cm^{-1}\right) = 2\left[(1-q)\alpha_{air}\left(cm^{-1}\right) + q\alpha_{self}\left(cm^{-1}\right)\right]\left(\frac{P}{P_0}\right)\left(\frac{T_0}{T}\right)^{\gamma} \tag{2-27}$$

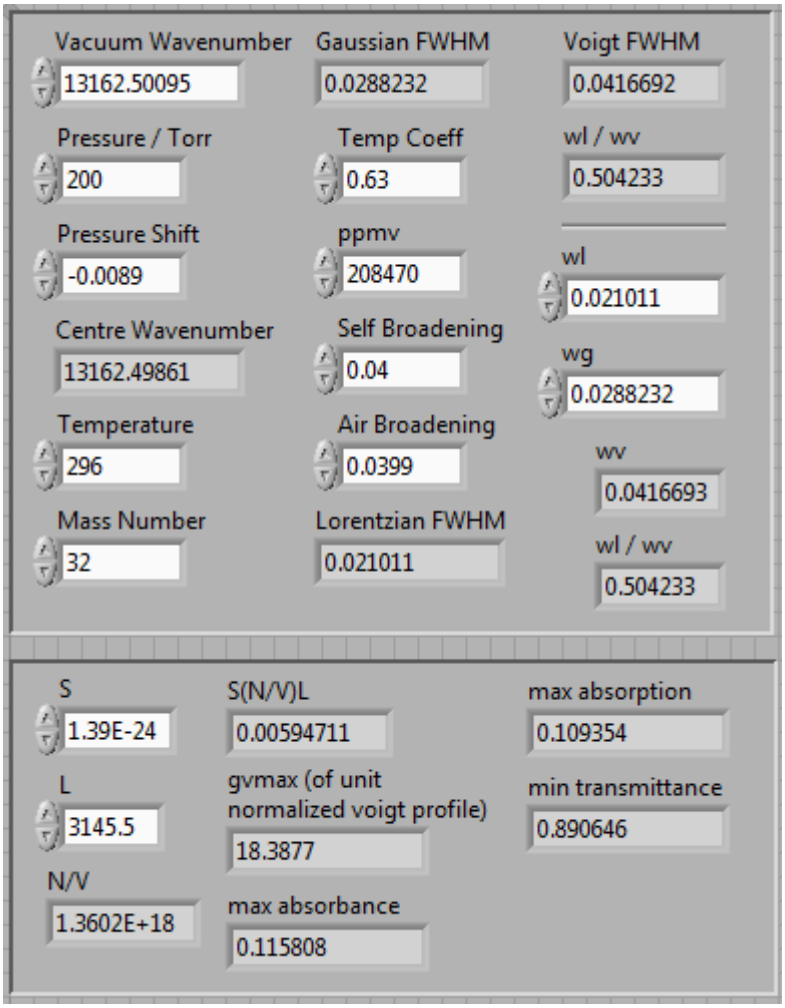


Figure 2-3: **The front panel of a custom made calculator for performing the lineshape calculations of Section 2.5 – producing the centre wavenumber and FWHM information to model, Gaussian, Lorentzian and Voigt profiles from HITRAN line list data. The number density is calculated via Equation 2-51 and the user entered concentration value in ppmv. The bottom panel displays the absorbance multiplier of the unit area normalized lineshapes, and other parameters for the unit area normalized Voigt lineshape. The data shown is for the R21Q22 A-band O_2 line in air at 296K and 200Torr.**

The convolution of two Gaussian lineshapes is another Gaussian lineshape of FWHM $\omega_g = \sqrt{\omega_{g1}^2 + \omega_{g2}^2}$. And the convolution of two Lorentzian lineshapes is another Lorentzian lineshape of FWHM $\omega_l = \omega_{l1} + \omega_{l2}$ (Vaughan, 1989). However, the Voigt lineshape – the convolution of a Gaussian and Lorentzian lineshape has no closed form solution, and its FWHM must be estimated from numerical modelling. A popular choice is due to Olivero and Longbothum (1977):

$$\omega_v = 0.5346\omega_l + \sqrt{0.2166\omega_l^2 + \omega_g^2} \tag{2-28}$$

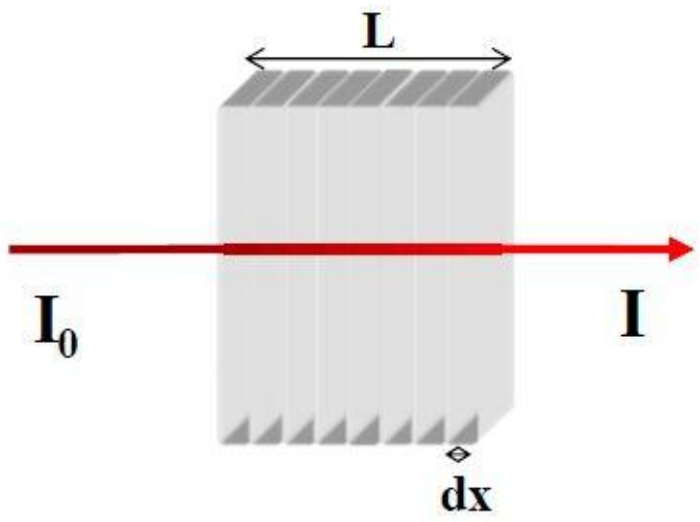
Equation 2-28 gives an exact result at one extreme of a pure Gaussian profile, but a small error for a pure Lorentzian profile. At a given temperature, Equation 2-20 implied that the Gaussian FWHM will stay constant with increasing pressure (apart from the line shift contribution through $\bar{\nu}_0$). The FWHM of the Voigt lineshape will therefore grow with the Lorentzian contribution of Equation 2-27. Since the FWHM also determines the lineshape height (Equation 2-9), the Voigt profile takes on an increasingly Lorentzian aspect as the pressure grows.

2.5.1 Lineshape Parameter Calculator

The formulae for FWHM's were made into a lineshape parameter calculator in LabVIEW (Figure 2-3). Users can enter HITRAN parameters, and pressure, temperature, and concentration parameters. The FWHM's are then returned. The number density and integrated absorbance (Section 2.6.5) are also returned.

2.6 The Beer-Lambert Law

The work horse of linear absorption spectroscopy is the Beer-Lambert Law. This relates the intensity of the radiation entering a gas to that leaving it. From this gas number density and concentration can be derived. The Beer-Lambert Law rests on a



$$dI_x(\tilde{\nu}) = -\sigma(\tilde{\nu}) I_x(\tilde{\nu}) \frac{N}{V} dx$$

$$I(\tilde{\nu}) = I_0(\tilde{\nu}) \exp[-\sigma(\tilde{\nu}) \frac{N}{V} L]$$

Figure 2-4: **The Beer-Lambert Law leads to an exponential decay of incident radiation, as the radiation encounters successive layers of gas. The gas contains a number density N/V of absorbing molecules, each having an absorption cross section of $\sigma(\tilde{\nu})$.**

number of assumptions well realized in an ideal gas, subject to reasonably weak intensity radiation.

In the linear regime, where the Beer-Lambert Law applies, it is assumed that the radiation does not significantly change the energy level population of the sample as a whole. If it does, raising a significant number of molecules to the excited state, the non-linear regime is entered, where further radiation is not linearly absorbed due to a reduced lower state population. The absorbers are also assumed to act independently of each other, so that the absorption of radiation by one molecule does not affect the likelihood of absorption by any other. Scattering of radiation is assumed to be negligible, so that all radiation received, after correcting for noise and laser power fluctuations, is due to absorption. It is further assumed that the gas homogenously fills the sample volume so that the concentration of target molecules in any thin slice perpendicular to the propagation direction of the radiation is constant. For the same reason each ray of radiation is assumed to have the same path length through the gas.

The Beer-Lambert Law is derived by considering the absorption of radiation within infinitesimal slabs of unit area and thickness dx (Figure 2-4):

$$dI_x(\tilde{\nu}) = -\sigma(\tilde{\nu}) I_x(\tilde{\nu}) \frac{N}{V} dx \tag{2-29}$$

$\sigma(\tilde{\nu})$ is known as the absorption cross section, $I_x(\tilde{\nu})$ is the initial radiant intensity upon the slab of frequency $\tilde{\nu}$ and N/V is the total number density of absorbers (number of absorbing gas molecules per unit volume). This form of the Beer-Lambert Law includes the effect of spontaneous and stimulated emission within the definition of $\sigma(\tilde{\nu})$ - so that $\sigma(\tilde{\nu})$ is combined with the total number density regardless of whether the molecules are in the lower or upper level of the transition. (This can be contrasted with the Bernath definition discussed below). Integrating this expression yields the Beer-Lambert Law:

$$I(\tilde{\nu}) = I_0(\tilde{\nu})\exp[-\sigma(\tilde{\nu})\frac{N}{V}L] \tag{2-30}$$

L is the radiation's path length through the gas.

2.6.1 Line Strength and lineshapes

The absorption cross section, $\sigma(\tilde{\nu})$, accounts for the average absorbance due to a single molecule for a specific transition. It can be split into two factors – firstly, a unit area normalized lineshape, $g(\tilde{\nu}-\tilde{\nu}_0)$, and a scaling factor – known as the line strength, S:

$$\sigma(\tilde{\nu}-\tilde{\nu}_0) = Sg(\tilde{\nu}-\tilde{\nu}_0) \tag{2-31}$$

Examples of the lineshapes, $g(\tilde{\nu}-\tilde{\nu}_0)$, appropriate to this thesis were covered in Section 2.3 – namely, Gaussian, Lorentzian, Voigt and pseudo-Voigt. The line strength S can be found by calculation or experiment, and for the molecules and transitions in this thesis the line strengths are listed in the HITRAN database (V E Zuev IAO, Rothman et al., 2009). (It should be noted that the definitions of σ and S, and hence their interrelationship, can vary markedly in different texts. For instance, Bernath (2005) deals with degeneracy, and spontaneous and stimulated emission from the upper level, outside of his definition of σ. His definition of σ is therefore combined with the number density of the molecules separated into those in the lower transition state from those in the upper state. He retains the HITRAN definition of S).

The line strength relates to the Einstein coefficients of absorption and stimulated emission between levels. It is proportional to the expectation value of the dipole moment operator between those levels. The line strength also takes account of the degeneracies and number densities of the levels, and has a temperature dependence (Bernath, 2005). The formula for relating the line strength S_0 given at 296K in the

HITRAN database to other temperatures can be found in the same references for the FWHM modeling of Section 2.5. (Spectralcalc.com, Rothman et al., 1998, Killinger and Plutov, 2009):

$$\frac{S(T)}{S(T_0)} = \frac{\exp\left(-100hc_0\tilde{E}_L/k_BT\right)}{Q(T)} \frac{1}{\frac{\exp\left(-100hc_0\tilde{E}_L/k_BT_0\right)}{Q(T_0)}} \left(\frac{1-\exp\left(-100hc_0\tilde{\nu}_0/k_BT\right)}{1-\exp\left(-100hc_0\tilde{\nu}_0/k_BT_0\right)}\right) \quad (2\text{-}32)$$

The first two factors account for the changing (Boltzmann) probability of absorption with temperature. The lower state energy, $\tilde{E}_L$, is another parameter included in HITRAN, and given in cm^{-1}. Here $Q(T)$ is the total partition function – the probability normalizing factor equal to the sum of Boltzmann factors over all of the energy levels of the system, at that temperature. The experiments of Chapters 4, 5, and 6 were conducted at room temperature, and the HITRAN line intensity at 296K was adequate. However, HITRAN does provide a tabulation of 'total internal partition sums', as a function of temperature, in the file *parsum.dat*. For the experiments of Chapter 7, the line strength adjustment was handled within the pulsed-QCL sensor by embedded Cascade Technologies software.

The final ratio in Equation 2-32 arises from the different reductions of the absorption linestrengths at the two temperatures - due to spontaneous emission into the lower state. $\tilde{\nu}_0$ is the centre wavenumber of the transition.

Theoretically, a general expression for the linestrength, in the electric dipole approximation, for an (open) two level transition, can be written in Gaussian-cgs units as:

$$S_{line} = \frac{8\pi^3}{3hc}\tilde{\nu}_{ul}\left[\frac{(N_l/V)}{g_l} - \frac{(N_u/V)}{g_u}\right]\left|\mu_{ul}\right|^2 \quad (2\text{-}33)$$

where $$\left|\mu_{ul}\right|^2 = \sum_u \sum_l \left|\int \psi_u^* \mu \psi_l d\tau\right|^2 . \quad (2\text{-}34)$$

Here, $\left|\mu_{ul}\right|^2$ is the electric dipole transition moment squared and represents the strength of interaction of a gas molecule with the radiation field for this transition. Equation 2-33 shows that the linestrength due to the collection of gas molecules is proportional to this quantity. The linestrength is also proportional to the wavenumber of the transition ν_{ul}. A given incident radiation intensity, I_0, will be made up of less photons when the wavenumber is higher, $n = I_0 / h\nu_{ul}$. A given molecule's absorptive transition will therefore cause a greater fractional change in intensity, I/I_0, and hence absorption, $(1 - I/I_0)$, the higher the frequency of the transition. Equation 2-33 also shows the gas' linestrength is proportional to the difference between the number density of the gas' molecules in the lower state, which can hence absorb, and the number density of those gas molecules in the higher state – which can reduce net absorption due to stimulated and spontaneous emission. The double sum in Equation 2-34 takes account of any degeneracy in the upper and lower levels and includes all of the transitions between them. The $(N_l/V)/g_l$ term in Equation 2-34 gives an equal population weighting to each lower degenerate level - similarly for the $(N_u/V)/g_u$ term.

Bernath (2005) provides some detailed sections on calculating linestrengths, albeit with different notation. Information on notations and equations applicable to the HITRAN database can be found in Appendix A of Rothman et al (1998) and 'HITRANonline' (Rothman).

To get to the 'per molecule' linestrength, S, given in the HITRAN database, Equation 2-33 is divided by the gas' total number density N/V. This also allows the per molecule linestrength to be given independent of a specific total number density since the resulting ratios, N_l/N and N_u/N, can be replaced with the Boltzmann distribution under conditions of local thermodynamic equilibrium:

$$\frac{N_i}{N} = \frac{g_i}{Q(T)} \exp\left(-\frac{hc\tilde{E}_i}{k_B T} \right) \tag{2-35}$$

Substituting Equation 2-35 into Equation 2-33, gives the expression used in the HITRAN database (for electric dipole transitions):

$$S = \frac{8\pi^3}{3hc} \frac{I_a g_l}{Q(T)} \tilde{\nu}_{ul} \exp\left(-\frac{hc\tilde{E}_l}{k_B T} \right) \left[1 - \exp\left(-\frac{hc\tilde{\nu}_{ul}}{k_B T} \right) \right] \left\{ \frac{1}{g_l} |\mu_{ul}|^2 \right\} \times 10^{-36} \tag{2-36}$$

Here, $\nu_{ul} = E_u - E_l$ and E_l are in cm^{-1}. I_a is a factor HITRAN includes to adjust the gas' total concentration to that due to the constituent isotopologue responsible for the transition - according to the isotopologue's abundance in the natural terrestrial environment. For other abundances, this factor can be modified. The factor of 10^{-36} reduces the HITRAN quoted value in braces in (Debye)^2 units - the weighted dipole transition moment squared – to its value in units of (ergs cm^3); 1Debye = 10^{-18}ergs cm^3. The term in braces is $\Re_{ul}$ in HITRAN notation.

In the databases after 2000, the Einstein coefficient for spontaneous emission, A_{lu}, is used instead of $|\mu_{ul}|^2$.

For electric dipole transitions (of Chapters 4, 5 and 7) (Rothman, Simeckova et al., 2006):

$$A_{lu} = \frac{64\pi^4}{3h} \tilde{\nu}_{ul}^3 \frac{g_l}{g_u} \Re_{ul} \times 10^{-36} \tag{2-37}$$

For magnetic dipole transitions (for the O_2 transitions of chapter 6):

$$A_{lu} = \frac{64\pi^5}{3hc} \tilde{\nu}_{ul}^3 \frac{g_l}{g_u} \Re_{ul} \times 10^{-36} \tag{2-38}$$

Giving an expression applicable to both cases with the relevant A_{lu}:

$$S = I_a \frac{A_{lu}}{8\pi c \tilde{\nu}_{ul}^2} \frac{g_u}{Q(T)} \exp\left(-\frac{hc\tilde{E}_l}{k_B T}\right)\left(1-\exp\left(-\frac{hc\tilde{\nu}_{ul}}{k_B T}\right)\right). \quad (2\text{-}39)$$

2.6.2 Experimental Signals, Transmittance, Absorption, and Absorbance

An important point is that, in general, the experimental signal recorded, $I(\tilde{\nu}-\tilde{\nu}_0)$, does not have the same shape as the absorbance lineshape. Nor, in general, is it linearly proportional to the concentration of the gas. The reason for this is that two things are going on at once. Not only are the gas molecules absorbing radiation with a frequency spread proportional to their lineshapes, but the radiation is being reduced at each layer of gas (Figure 2-4) - proportionately to the intensity making it through from the proceeding layer. This gives the exponential form of the measured signal of the Beer-Lambert Law, reproduced here:

$$I(\tilde{\nu}-\tilde{\nu}_0) = I_0(\tilde{\nu})\exp[-Sg(\tilde{\nu}-\tilde{\nu}_0)\frac{N}{V}L] \quad (2\text{-}40)$$

It can be seen that the quantity which is directly proportional to number density (and concentration) is the value of the exponent – and this value is termed the 'absorbance':

$$\text{Absorbance} = \text{Abs} = Sg(\tilde{\nu}-\tilde{\nu}_0)\frac{N}{V}L \quad (2\text{-}41)$$

The experimental signal is generally recorded as intensity (or power, or voltage) against time. The time axis is then converted to a frequency axis, and the intensity normalized via its baseline (Chapter 3 and Figure 4-1) - the 'transmittance' then results:

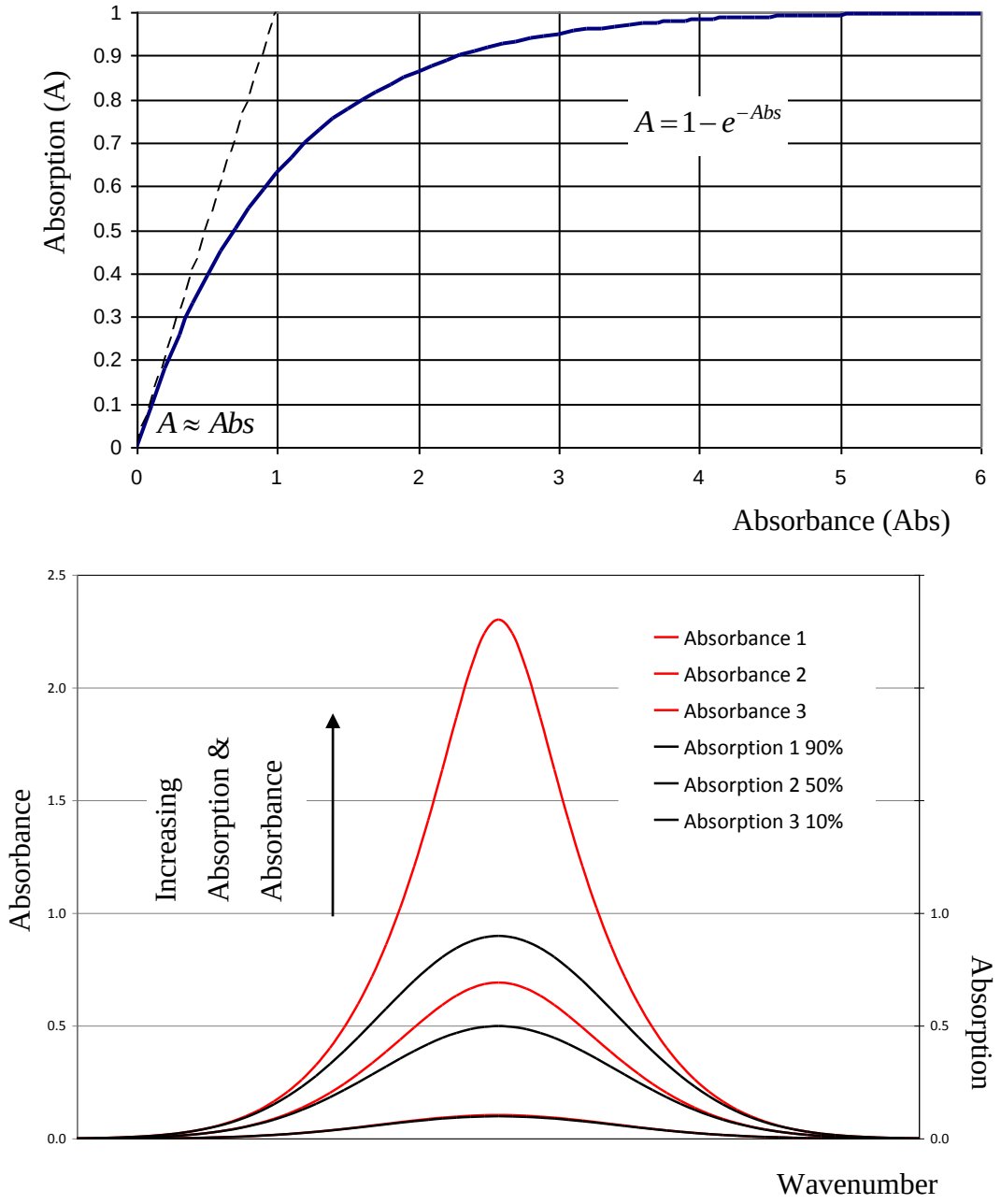


Figure 2-5: **The top plot shows the functional form of absorption versus absorbance. Absorption essentially comes from the experimentally measured signal, while the absorbance calculated from it is proportional to the gas' number density and concentration. As the absorption increases it falls behind the absorbance and hence the concentration of the gas. This is due to the faster exponential decay of the detected radiation through the higher concentration gas.**
The bottom plot shows increasing absorption plots (in black) from Equation 2-43. The underlying absorbances are of Gaussian line shape (in red). Again, as the absorption increases it falls behind the absorbance - and hence underestimates the concentration of the gas causing the absorption.

$$\text{Transmittance} = T = \frac{I(\tilde{\nu} - \tilde{\nu}_0)}{I_0(\tilde{\nu})} = \exp\left[-Sg(\tilde{\nu} - \tilde{\nu}_0)\frac{N}{V}L \right] = e^{-\text{Abs}} \qquad (2\text{-}42)$$

It can be noted that 'absorption' is used for one minus the transmittance, and so is not in general equal to the absorbance:

$$\text{Absorption} = A = 1 - T = 1 - e^{-\text{Abs}} \qquad (2\text{-}43)$$

In general the absorbance, and hence number density, is found by calculation from the transmittance:

$$\text{Abs} = -\ln(T) = -\ln(1 - A) \qquad (2\text{-}44)$$

However, if the absorbance is small (<10%), because the gas is 'optically thin', the absorption becomes equal to the absorbance. The experimental signal then becomes proportional to the number density, and the experimental profile reflects the underlying lineshape (when the abscissa is in wavenumbers rather than time). From Equation 2-41, the gas does not absorb as much radiation when any of the linestrength, number density, or path length is sufficiently small, making the gas optically thin. (This is equivalent to using Equation 2-29 rather than Equation 2-30 for the Beer-Lambert Law).

So, when Abs <10%:

$$\text{Absorption} = A = 1 - e^{-\text{Abs}} \approx 1 - (1 - \text{Abs}) = \text{Abs} = \text{Absorbance} \qquad (2\text{-}45)$$

And $$\text{Absorbance} = -\ln(1 - A) \approx -((1 - A) - 1) = A = \text{Absorption} \qquad (2\text{-}46)$$

The top plot of Figure 2-5 shows the functional relationship between absorption and absorbance. (Figure 7-6 shows experimental data illustrating these relationships). The bottom plot of Figure 2-5 shows some Gaussian absorption profiles against their

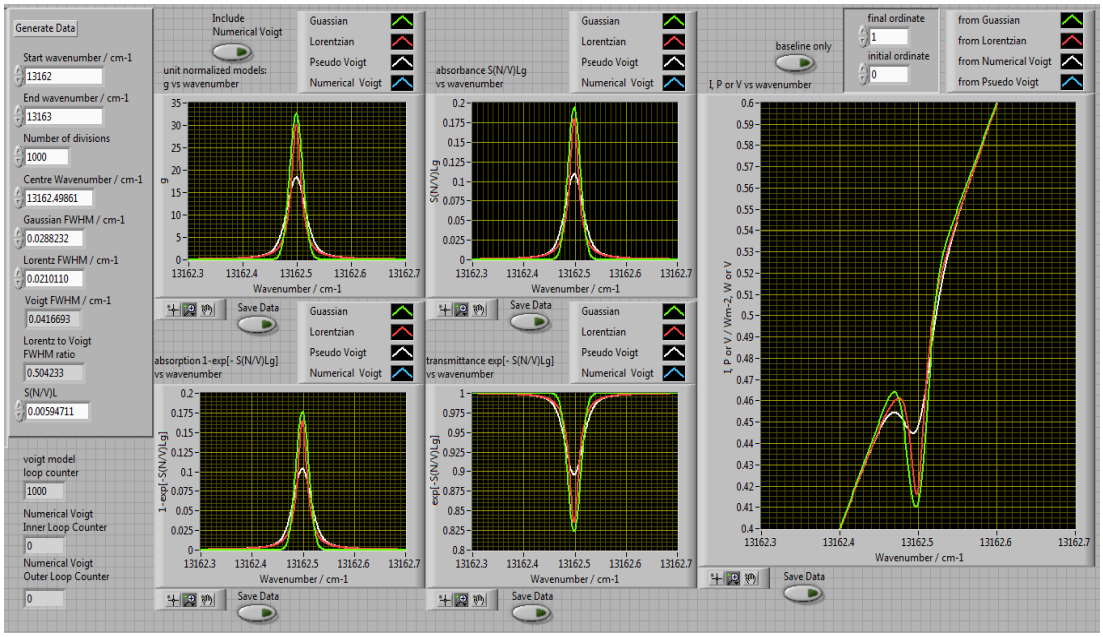


Figure 2-6: **The front panel of a custom made program for calculating Gaussian, Lorentzian, and Voigt lineshapes. The lineshape parameters used here come from those in the parameter calculator of Figure 2-3. That is, for the R21Q22 A-band O_2 line in air at 296K and 200Torr. The unit area normalized lineshapes are first calculated and displayed. They are then scaled to absorbance by a user entered multiplier (via the parameter calculator). Then the absorption and transmittance are calculated and displayed. The large panel allows a linear baseline to be added to the transmittance spectrum. Note that this makes the perfectly symmetrical line profiles appear asymmetric. Also, on the diagram's scale, the Voigt and pseudo-Voigt profiles are essentially overlapping.**

underlying absorbance lineshapes. In both plots the absorption and the absorbance diverge as the concentration increases.

2.6.3 Modeling Program for Lineshape, Absorbance, Transmittance, and Absorption

Figure 2-6 shows the front panel of a modeler constructed in LabVIEW. This displays the noiseless signals expected from entered FWHM values – often from those calculated via the parameter calculator of Section 2.5.1 and Figure 2-3. The program served as a useful starting point for fitting models to experimental data (Chapter 3).

Figure 2-6 shows Gaussian, Lorentzian, and overlapping Voigt and pseudo-Voigt profiles for the Chapter 6 R21Q22 A-band O_2 line in air at 296K and 200Torr. The program starts by displaying the unit area normalized lineshapes. It then weights these absorbances by the integrated absorbance (Equation 2-49) for a particular molecular transition and number density. Then the absorbance and transmittance are displayed. Finally, the user may enter a linearly sloping baseline to get an idea of how the experimental signal may look. It might be noted that the slope in the large panel causes the perfectly symmetrical line profiles to appear asymmetrical.

2.6.4 Intensity, Power, and Voltage

The Beer-Lambert Law (Equation 2-40) is usually couched in terms of intensity. Since $I = P/\text{Area}$, if all the power P falls on the detector, intensity can be replaced with power. In this thesis the detector signals are measured in terms of voltage.

The DC-coupled Avalanche Photodiode Detector (APD), used for oxygen detection in Chapter 6, gave an excellent linear response of Voltage produced to Power received. The AC-coupled VIGO photovoltaic detector used for methane sensing in

Chapters 4 and 5 also provides a linear response. (This contrasts with the standard text book logarithmic response of photovoltaics).

The AC-coupled nature of the VIGO detector made it difficult to characterize its response statically (unswept) – for instance, chopping the beam was too slow to avoid detector time constant effects (Section 4.6.1.2). However, the static Power vs Current measured on a power meter, and Wavenumber vs Current results from Fourier Transform Spectroscopy, showed a near linear behaviour over a typical current sweep range. This near linear response is observed in the voltage output of the VIGO detector – in both the baseline of the spectral scans at 10kHz and the amplitudes of etalon fringes (e.g. Figure 4-21). The reasonableness of the extracted number densities also lends weight to the assumption used throughout this work, that:

$$\frac{I(t)}{I_0(t)} = \frac{P(t)}{P_0(t)} = \frac{V(t)}{V_0(t)} = \exp\left[- Sg(\tilde{\nu}(t))\frac{N}{V} L \right] \tag{2-47}$$

2.6.5 Line strength units, integrated absorbance and absorbance peak height

Rearranging Equation 2-40 gives:

$$\text{Absorbance} = Sg(\tilde{\nu} - \tilde{\nu}_0)\frac{N}{V} L = -\ln\left(\frac{I(\tilde{\nu} - \tilde{\nu}_o)}{I_0(\tilde{\nu})} \right) \tag{2-48}$$

$\Rightarrow$ Integrated Absorbance

$$= S\frac{N}{V} \mathrm{L} \int \mathrm{g}(\tilde{\nu} - \tilde{\nu}_0) d\tilde{\nu} = S\frac{N}{V} L = -\int \ln\left(\frac{I(\tilde{\nu} - \tilde{\nu}_0)}{I_0(\tilde{\nu})} \right) d\tilde{\nu} \tag{2-49}$$

$$\Rightarrow S = -\frac{1}{\frac{N}{V} L} \int \ln\left(\frac{I(\tilde{\nu} - \tilde{\nu}_0)}{I_0(\tilde{\nu})} \right) d\tilde{\nu} \tag{2-50}$$

since $\int g(\tilde{\nu}-\tilde{\nu}_0)d\tilde{\nu}$ =1 by definition of a unit area normalized function.

It might be noted that the area under the lineshape depends both on the number density (or concentration) and the line strength. Thus two very similar looking lines, with very different line strengths, represent very different concentrations.

The line strength S has units derived from the units chosen in Equation 2-50. In HITRAN N/V has units of molecules cm^{-3}, L in cm, and $\tilde{\nu}$ in cm^{-1}. HITRAN units are therefore (cm^{-1}/molecules cm^{-2}).

Equation 2-50 also shows how the number density can be found from the area of the integrated absorbance:

$$\Rightarrow \frac{N}{V} = -\frac{1}{SL}\int \ln\left(\frac{I(\tilde{\nu}-\tilde{\nu}_0)}{I_0(\tilde{\nu})}\right)d\tilde{\nu} \tag{2-51}$$

However, Equation 2-48 allows the possibility of calculating number density from any point on the curve – say at the maximum value where SNR is expected to be highest. (In Section 4.6.5, a useful definition of signal to noise ratio is the signal peak height divided by twice the standard deviation of the baseline noise). From Equation 2-48:

$$\frac{N}{V} = -\frac{1}{SL}\frac{\ln\left(\frac{I(\tilde{\nu}-\tilde{\nu}_o)}{I_0(\tilde{\nu})}\right)\Big|_{\max}}{g(\tilde{\nu}-\tilde{\nu}_0)|_{\max}} \tag{2-52}$$

In many experiments, the gas is able to be conditioned so that the signal strength is ideally only affected by the number density (or mixing ratio). For instance, the gas cell will have fixed temperature and pressure, and the gas will have the volume of the cell, or be drawn through the cell at a constant volume flow rate.

However, the signal may be noisy, and fitting a lineshape may well be necessary to locate the value of the peak. (Figure 4-32 shows an experimental example of this). In any case, taking the area may be more robust. Small fluctuations between cell temperature and pressure, for instance, will lead to a change of FWHM and height where the Lorentzian lineshape is a significant contribution. However, the area will still reflect a constant concentration. Gas collisions in the linear regime are area preserving. Although the excited state lifetime is reduced - sending molecules into their lower state more quickly with increasing pressure - the lower state population remains essentially unchanged and ready to absorb radiation – the defining characteristic of the linear regime.

A possible downside of taking the area is deciding on what range of wavenumbers should be integrated over in the case of the broad wings of a Lorentzian lineshape. This is particularly so if the baseline is noisy or subject to interference from other lines. In harsh environments, where many gas parameters can't be controlled, the lineshape area is usually needed.

In Chapter 5 absorbance peak height is used for a calibration - since at low pressure the Gaussian lineshape is appropriate and is less susceptible to pressure fluctuations than the Lorentzian and Voigt lineshapes. Nevertheless, a calibration using area produced similar results. In the data analysis program of Chapter 3, the area method can be used.

2.6.6 Measures of concentration: number density and mixing ratio

An important distinction can be made between the two measures of concentration used in this thesis: number density and mixing ratio. The number density of a component gas is an absolute value. For instance, the number density of component i is N_i/V, measured in, for instance, molecules per cm^3. Mixing ratio is a pure number - implying the amount of one component in relation to the whole. A mixing ratio is the number of molecules of the gas species i, to the total number of

molecules in the gas, N_i/N_T. This is also equal to the ratio of species number density to the total number density.

In Chapters 4 and 5 pure methane gas is used. The mixing ratio is therefore always 100%, or equivalently 1,000,000 parts per million. Different strength signals are produced by pumping on the gas cell, which changes the number density (and hence pressure) – changing the number of absorbers, and hence the absorption.

Similarly, in Chapter 6 oxygen spectroscopy is conducted on laboratory air flowing through a gas cell at 760Torr and 200Torr. The oxygen mixing ratio in air is the same in both cases, at approximately 20%, but the number density is different since the volume flow rate is adjusted to change the total pressure.

In Chapter 7, different concentrations of gas are prepared by diluting the target gas with differing amounts of nitrogen gas. So this time the mixing ratio is variable, but the overall number density remains the same as the temperature, pressure and volume flow rate of the gas through the cell are well controlled.

The following formula allows conversion between number density in molecules cm^{-3} to concentration (mixing ratio) in parts per million by volume (ppmv) - if the total cell pressure in Torr and the gas temperature in Kelvin are known. This formula is used in the data analysis program of Chapter 3.

$$x(ppmv) = \left(\frac{N_i}{V_{T_{cm^3}}}\right)\left(\frac{molecules}{cm^3}\right)\frac{760\times10^{12}}{101325}\frac{k_B T_K}{P_{T_{Torr}}}\left(\frac{cm^3}{molecule}\right)(ppmv) \qquad (2\text{-}53)$$

2.7 Basic Signal to Noise Ratio Improvement

To improve the signal to noise ratio (SNR), the lowest electronic noise combination of the equipment available can be sought. Optically, avoiding unwanted parallel surfaces in the beam path mitigates against the unwanted forming of etalon fringes.

The temperature of the laboratory can be controlled to prevent expansion of components and therefore change in beam path. The gas suspension of optical benches can reduce vibration, and thus alignment changes causing changes in power on the detector. Ambient light can be minimized. The temperature, pressure, and volume of the gas can be controlled to prevent fluctuations of the lineshape not due to concentration changes. The power of radiation falling on the detector after absorption can also be maximized, short of saturating the detector, to improve the SNR.

After this, specific experimental noise reduction techniques can be used, and the techniques of Sweep Integration and Wavelength Modulation Spectroscopy are discussed in Chapters 4 and 5 respectively.

The equation for absorbance (Equation 2-48) - the quantity proportional to concentration - also indicates ways of enlarging the signal. Using a transition with relatively large linestrength S will increase the amount of absorption, and hence enlarge the absorbance signal. However, the linestrength must be such that the absorbance stays small enough over the range of concentration values expected to be measured. This avoids lineshape distortion and consequent deviation from the Beer-Lambert Law. Another means of increasing the absorption signal is by increasing the path length through the gas, so that more absorbers encounter the radiation, as below.

2.8 The Astigmatic Herriott-cell

To achieve an acceptable sensitivity, the path length of the laser beam must be sufficiently long in order to achieve sufficient absorption through the gas. This is achieved in a compact space by repeated reflection off opposing mirrors. Compactness is both convenient and necessary in order to allow efficient gas exchange and consequently fast update times. Path length is limited by Fresnel losses at the mirrors, scattering and absorption losses, and the achievable density of spot pattern to beam size on the mirrors. To avoid Fabry-Perot interference, which here

constitutes noise, the beam enters the cavity at an angle to the optic axis and follows a pattern designed so that beam overlap does not occur. If the beam's radius and spot size are similar to the cavity's resonant modes, continual refocusing also occurs – minimizing diffraction losses and distortion of the beam profile.

2.8.1 The Herriott cell

The Herriott cell (Herriott et al., 1964, Herriott and Schulte, 1965) is used for collimated sources such as lasers. It consists of two – usually matching - concave spherical mirrors facing each other. Cells are close to confocal as this results in minimal cavity mode sizes and permits the pattern to be more compact and the cell smaller.

Often the beam enters and exits in the centre of the same mirror but at a different angle. For this to happen, solutions are found for a 're-entrant' condition. That is, the entrance hole coincides with where the beam pattern would 're-enter' its own pattern within the cavity and start to repeat itself.

The beam enters at a skew angle to an optical plane, in general forming an elliptical pattern of beam spots on both mirrors. The size and shape of the pattern is determined by the beam's initial entry position and angle. Whether the pattern is re-entrant depends on the ratio of mirror focal length (f) to mirror spacing (d). These parameters are adjusted in the Herriott cell to produce re-entrant paths of various lengths. This is why the Herriott cell is *close to* being confocal rather than depending on *being* confocal.

If the optic axis is denoted z in a Cartesian coordinate system, the ray paths in the cell can be analyzed by considering consecutive ray positions (x_n, y_n) and gradients ($x_n^{'}, y_n^{'}$) at the x-y plane immediately following the n^{th} mirror. The cell can be modeled by ray matrix methods as equivalent to a periodic system of equally spaced thin lenses. The results are (Herriott et al., 1964, Yariv and Yeh, 2007):

$$x_n = x_0 \cos n\theta + \sqrt{\frac{d}{4f-d}}(x_0 + 2fx_0^{'})\sin n\theta = A\sin(n\theta + \alpha) \tag{2-54}$$

$$\text{with} \qquad 0 < \frac{d}{f} < 4 \qquad \text{for a stable system,} \tag{2-55}$$

with the 'advance angle' θ given by $$\cos\theta = 1 - \frac{d}{2f} \tag{2-56}$$

$$\tan\alpha = \frac{\sqrt{\frac{4f}{d} - 1}}{1 + 2f\frac{x_0^{'}}{x_0}} \tag{2-57}$$

and $$A^2 = \frac{4f}{4f-d}\left(x_0^2 + dx_0x_0^{'} + dfx_0^{'2}\right) \tag{2-58}$$

where A is the maximum excursion of the ray in the x-direction.

Equivalent results hold for $y_n = A\sin(n\theta + \beta)$. (2-59)

Although the parameter θ is included for analytical convenience, it has geometrical significance. If d and f are such that $\nu 2\theta = \mu 2\pi$, where ν and μ are integers, the elliptical pattern is re-entrant in position and gradient after ν round trip traversals of the cell. That is, two beam spots coincide in position and direction after the elliptical spot pattern is 'circled' μ times.

2.8.2 The astigmatic Herriott cell

By introducing a small astigmatism in one or both mirrors, by design or by stressing in a mirror mount, the mirror takes on two unequal radii of curvature in its x and y

directions (R_x and R_y) (Herriott et al., 1964). Since $\theta = \theta(d, f) = \theta(d, R)$, θ becomes θ_x and θ_y. The beam spots again trace out sinusoidal paths on the mirrors, but now the x and y frequencies of complete 2π trips around the pattern are different. This results in complex and dense Lissajou patterns of beam spots, allowing more passes and a longer path length to be achieved for the same mirror areas. Two independent re-entrant conditions must now be satisfied in the x and y directions: $\nu 2\theta_x = \mu 2\pi$ and $\nu 2\theta_y = \gamma 2\pi$. That is, the x-component of the beam spot's oscillation pattern and its y-component, both bring the beam spot an integer number of complete round trips round the beam spot pattern after ν round trip passes between the mirrors. Each pattern can be described by the set of integers $\{\nu,\mu,\gamma\}$.

McManus has analyzed this system in terms of 4x4 matrices for determining $x_n, y_n, x_n^{'}$ and $y_n^{'}$ (McManus et al., 1995, McManus, 2007). He has shown that a combination of mirror separation and mirror twist can compensate for errors in manufactured mirror radii by effectively adjusting R_x and R_y. These same adjustments allow different re-entrant paths and hence path lengths to be achieved with the same mirrors.

Without rotation, and identical mirrors, the 'coordinates' **Z** are given by the ray matrix:

$$\mathbf{Z} = \mathbf{C}^{i}\mathbf{Z_0} \tag{2-60}$$

where $$\mathbf{C(R,D) = D(d)R(R_x,R_y)D(d)R(R_x,R_y),} \tag{2-61}$$

here i is the number of round trip traversals across the cell and back,

$$\mathbf{Z} = \begin{pmatrix} x_n \\ x_n^{'} \\ y_n \\ y_n^{'} \end{pmatrix} \qquad \mathbf{Z_0} = \begin{pmatrix} x_0 \\ x_0^{'} \\ y_0 \\ y_0^{'} \end{pmatrix} \qquad \mathbf{D} = \begin{pmatrix} 1 & d & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & d \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2-62}$$

$$\mathbf{R}(R_x, R_y) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -\frac{2}{R_x} & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -\frac{2}{R_y} & 1 \end{pmatrix} \tag{2-63}$$

With rotation of one of the mirrors by θ_t, the **R** matrix is modified by coordinate transformation to:

$$\mathbf{R}'(R_x,R_y) = \mathbf{T}(-\theta_t)\mathbf{R}(R_x,R_y)\mathbf{T}(\theta_t) \tag{2-64}$$

where

$$\mathbf{T}(\theta_t) = \begin{pmatrix} \cos\theta_t & 0 & \sin\theta_t & 0 \\ 0 & \cos\theta_t & 0 & \sin\theta_t \\ -\sin\theta_t & 0 & \cos\theta_t & 0 \\ 0 & -\sin\theta_t & 0 & \cos\theta_t \end{pmatrix} \tag{2-65}$$

and the advance angles θ_x and θ_y are modified.

Useful graphs can be produced from this analysis and used to predict the behaviour of the cell and the parameter values needed for re-entrant paths of varying lengths. x_n and y_n pairs give the expected pattern seen on a mirror for a required path length. By plotting θ_x against θ_y zones of nearly re-entrant paths can be mapped. If the number of traversals are calculated for a range of mirror separations d and mirror rotation angles θ_t, 'stability' graphs can be produced to show where solutions for a required number of traversals can be found.

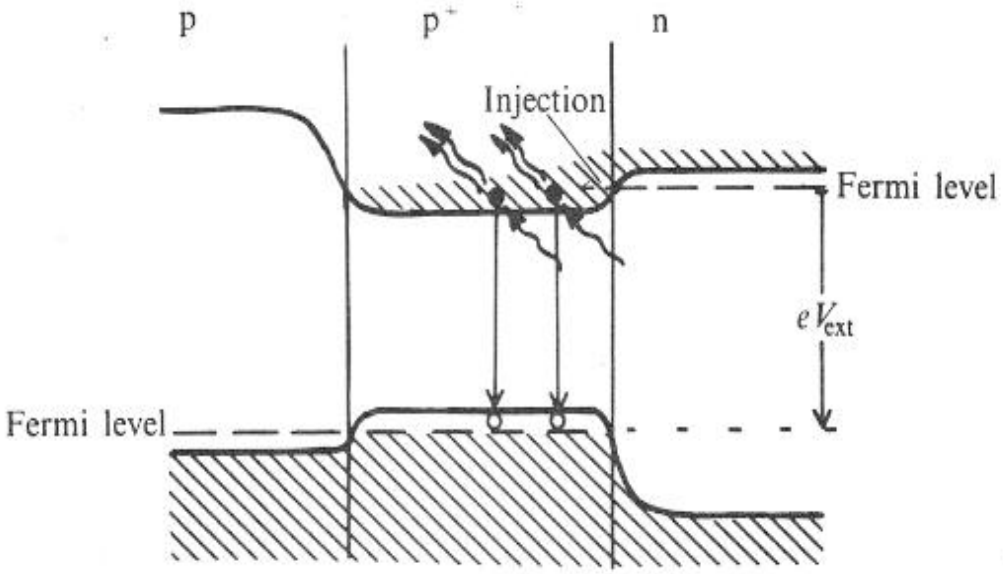


Figure 2-7: Semiconductor heterojunction diode laser band diagram, under bias V_{ext}. (Diagram from Fraser, 1990).

2.9 The Laser Types

2.9.1 Semiconductor Diode Lasers

Semiconductor diode lasers have become ubiquitous in modern life since their realization in the early 1960's. The essence of a semiconductor diode laser is a p-n junction with a band gap determined by the materials of the device – for instance, an AlGaAs / AlGaInP double-heterojunction for wavelengths near 760nm (suitable for O_2 sensing). Figure 2-7 illustrates the band structure of a biased double-heterojunction design. (Illustration from Fraser, 1990). Radiation is produced by recombination of electron-hole pairs. The device is heavily doped, so that when biased, electrons stack up due to the Pauli Exclusion Principle in the conduction band of the middle p-AlGaAs layer. This creates the necessary population inversion. Also the refractive index difference of the two materials helps to confine photons to the active region - improving amplification.

Single mode operation is often achieved via Distributed Feedback (DFB). That is, a diffraction grating is etched into the device's guiding layer. The laser's resonant cavity is formed by cleaving and polishing the device's end faces. The length of the device is kept small to decrease the number of cavity modes supported and increase their Free Spectral Range (FSR).

2.9.2 Semiconductor diode laser tuning

The frequency emerging from a semiconductor diode laser is a function of the active layer's band gap energy, cavity length, and refractive index. It has two tuning mechanisms – relatively slow temperature tuning and relatively fast refractive index tuning due to injection current density. (In contrast Quantum Cascade Lasers (QCL's), below, only have temperature tuning).

In the case of temperature, relatively slow control is provided by Peltier control of the laser's substrate temperature, while faster temperature tuning comes from the

Joule heating of the active layer via current. Temperature affects the active layer's band gap, affects its refractive index through expansion, and cavity length through expansion. In the case of a DFB narrow stripe design, the period of the grating also changes – although band gap and refractive index change are the dominant mechanisms.

In addition, increasing the injection current changes the refractive index due to carrier density build up at the p-n junction. This provides the faster tuning response to modulation - which finds use in the telecommunications industry. Temperature tuning causes more change in wavenumber, but takes place on a μs to ms timescale. Injection current tuning takes place on a ns time scale and becomes the dominant mechanism above a few hundred MHz modulation frequency.

The gain profile of the laser is much wider than the cavity modes supported. (A few THz to ~100GHz). These devices (in common with QCL's and other semiconductor lasers) therefore show a 'staircase effect' in tuning. That is, they tune continuously over a relatively small wavenumber range before suffering discontinuous 'mode hops' in wavenumber – before resuming continuous tuning again. This is caused by the relative movement of the gain profile (with bandgap energy) and the cavity modes. Since both the gain profile and the cavity modes tune - not in general by the same wavenumber - mode hops will not necessarily be between adjacent cavity modes.

2.9.3 Quantum Cascade Lasers

The idea for the quantum cascade laser (QCL) was proposed by Kazarinov and Suris (1971) following the work of Esaki and Tsu on superlattices (1970). An actual device was first realised by Faist *et al* (1994).

By making the width of a heterojunction smaller than an electron's de Broglie wavelength, a quantum well is formed in the conduction band. Here, quantum

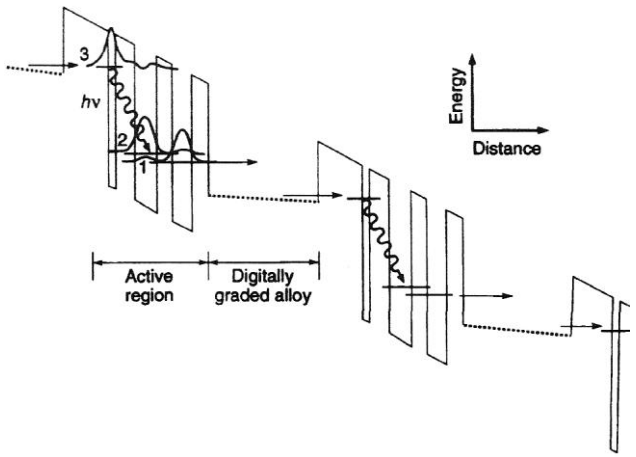


Figure 2-8: **The band structure of two active regions in a quantum cascade laser (QCL) under voltage bias. ((Diagram from Faist et al., 1994)). An electron from an injector region tunnels into the first quantum well, where it makes a radiative transition. The wavefunction of the electron in the lower energy level is strongly coupled to that in an adjoining well, allowing rapid non-radiative depopulation of the lower level, preventing photon re-absorption.**

A 'digitally graded alloy' contains a constant period made up of the two heterojunction materials composing the device. However, the width assigned to either material within a period changes linearly across the injector region. This is found to give a flat band under the appropriate bias – preventing carrier accumulation which would prevent the electrons cascading.

confinement produces discrete energy levels analogous to those of an atom. These levels, within a quantum well, are known as *subbands* - discrete in the direction of confinement and quasi-continuous in the directions parallel to the confining semiconductor layers.

A quantum cascade laser is a heterostructure device consisting of many nm scale semiconductor layers making up sequences of paired injector and active regions (Figure 2-8). The active regions are made up of multiple quantum wells or superlattices while the injector regions are usually superlattices. (A superlattice is a sequence of quantum wells with ultrathin barrier layers (<1nm). This allows corresponding subbands to couple to each other via tunnelling and broaden into *minibands*, analogous to the broadening of corresponding atomic levels upon formation of a crystal). QCL's are usually grown by MBE (Molecular Beam Epitaxy) or MOCVD (Metal-Organic Chemical Vapour Deposition) (Saleh and Teich, 2007). Modern designs often use AlInAs/InGaAs quantum wells on InP substrates or AlGaAs/GaAs technology, and can have as many as 20 to 100 'injector – active region' stages.

Kazarinov and Suris proposed and calculated it possible for an electron to make a radiative transition between the levels in a quantum well – an *intersubband* transition – and then to be 'recycled' by tunnelling through to an adjoining well, and so on. By applying a suitable design and voltage, it is possible to line up the bottom level of one stage with the top of the next stage. In this way a single electron can 'cascade' through a number of stages, producing many photons as it goes and making possible a quantum efficiency greater than one and a high output power (mW to hundreds of mW).

The design is fundamentally different to that of the semiconductor diode laser. There, only one photon is produced per recombination event. The recombination of an electron in the conduction band and a hole in the valence band removes these carriers so that they can not be recycled. The diode laser also relies on a build-up of carrier

density for its fast refractive index tuning. In contrast, the quantum cascade laser relies on recyclable electrons making transitions within the conduction band. (QCL's are known as *unipolar* – single charge carrier - devices). The accumulation of carriers (as at a p-n junction in a diode laser) would be detrimental to the cascading process, and is avoided. In the bipolar diode laser, the conduction and hole bands have opposite curvature. In contrast, the subbands of the QCL have the same curvature and this allows a reduced bandwidth of emitted radiation.

The bandgap of a QCL, and thus the centre frequency of the light it produces, is also independent of the material properties of the active semiconductor. The QCL bandgap is engineered. This is not the case with diode lasers whose bandgap is set by the difference in energy of its valence and conduction band. In contrast, a QCL's bandgap is a function of the size quantization set by the width of the active material, rather than the material's own bandgap. This allows longer wavelengths to be accessed than would be possible with a small bandgap material - whose population inversion would be subject to thermal wash-out and non-radiative phonon scattering.

Many variations of QCL design now exist (Gmachl et al., 2001), but generally the system is four level. The superlattice injector lines up with the first level of the active region, which does *not* line up with the injector region of the succeeding stage. The lower level does line up with the succeeding injector region, and in the original device its wavefunction was strongly spatially coupled to a lower level, spaced in energy to facilitate depopulation by optical phonon scattering. Thus a population inversion can be achieved with lifetimes determined by barrier thicknesses and wavefunction overlaps.

The injector region is normally 'digitally graded', meaning that the period of the two repeating layers of the superlattice is constant, but the widths taken up by the two layers within a period vary linearly. This design has been shown to produce a linear ramp of potential within an injector region and, hence, overall, one period of a sawtooth potential for the device (Faist et al., 1994). When linearly biased the

injector region becomes nearly flat – quasi-neutral – reducing the accumulation of carriers and device failure. The injection layer is also strongly n-doped to combat space-charge effects caused by carrier injection.

The QCL avoids many of the problems inherent with the lead salt diode lasers which preceded them in the mid-infrared. The power output can be up to a hundred times greater (up to hundreds of mW), meaning longer absorption paths can be used allowing greater sensitivity, the linewidth is narrower leading to higher resolution and thus selectivity, and QCL's have a larger wavelength tuning range. QCL's also do not need to be cryogenically cooled, being available in both room temperature pulsed and continuous wave (cw) versions. This opens the possibility of long unattended continuous operation for applications like pollution monitoring or process control. Temperature tuning is achieved around room temperature via Peltier elements with the help of forced air or liquid coolers.

2.9.4 Quantum Cascade Laser tuning

Because a QCL design avoids carrier build up to keep its electrons cascading, it lacks the fast refractive index tuning due to carrier density, available to semiconductor diode lasers. In the QCL, the tuning is by temperature - either by relatively slow temperature control of the laser's substrate, or relatively fast Joule heating of the active layer with injection current. The tuning is mainly by temperature induced change of the device's refractive index. To a lesser extent, tuning also occurs by expansion of the device's Bragg grating - in order to oscillate in a single mode, the QCL laser often comes with a Bragg grating etched into the guiding layer above the active QC layer. The evanescent part of the radiation field couples to this, and this is sufficient to act as a distributed mirror to reflect the whole wave field by Bragg reflection, for wavelengths in a narrow bandwidth of the Bragg wavelength. This is known as 'distributed feedback' (DFB-QCL). Whether by refractive index or grating expansion, an increase in temperature implies that a longer centre wavelength is

reflected and hence the radiation frequency is red shifted (shifted to lower frequency).

2.9.5 Use of pulsed-Quantum Cascade Lasers in Absorption Spectroscopy

The Cascade Technologies' CT3000 gas analyzer constructed and tested in Chapter 7 runs with up to four pulsed quantum cascade lasers (pulsed-QCL's). These operate in the intra-pulse mode – the second mode discussed below. Applying an injection current pulse causes heating, and, as explained above, leads to a frequency downchirp.

2.9.5.1 Inter-Pulse (Short Pulse) Technique

Two main techniques have emerged in applying pulsed-QCL's to absorption spectroscopy. In the inter-pulse (short pulse) method the frequency chirp is minimized by using very short current pulses (~3-10ns), along with either a fast scan, by using a subthreshold current ramp (achieving a spectral microwindow of $0.22cm^{-1}$ in Namjou *et al*'s paper (1998)), or a slow scan (achieving $2.5cm^{-1}$), by temperature rise of the laser substrate. A frequency range is therefore interrogated by sequentially incrementing the laser's centre frequency. This approach was demonstrated by Namjou (Namjou et al., 1998) and developed by Rice University (Kosterev et al., 2000) and Aerodyne (Nelson et al., 2002). These techniques use relatively complex control and processing setups - phase sensitive detection was used by Namjou and sweep integration and pulse normalization was used in the two latter systems. Pulse normalization uses a reference pulse split from the main beam and bypassing the gas. It arrives at the detector very close in time to the signal pulse, and hence is a suitable divisor to negate amplitude variations in signal pulse heights. (Typical tuning ranges are $1\text{-}2cm^{-1}$, at repetition frequencies of tens of Hz to kHz (Howieson et al., 2005)).

2.9.5.2 Intra-Pulse (Long Pulse) Technique

In contrast, the intra-pulse method utilizes the almost linear frequency downchirp by using longer current pulses (50-1500ns) to fast scan a range of frequencies (hence absorption features) essentially simultaneously over a window of typically 2.5cm^{-1} but up to 4-6cm^{-1} (Howieson et al., 2005). Up to 100kHz repetition rates are typical. The method requires a fast broadband detector and amplifier, and fast digitiser, but does not require the relatively complex control and processing setups of the inter-pulse technique.

In both methods it is necessary to repeat a large number of scans for averaging (typically $\geq$ 500, and often as many as 10,000) in order to achieve an acceptable signal to noise ratio. The inter-pulse (short pulse) method does suffer some limitations, however, not found in the long pulse system. To minimize the chirp in the short pulse case, threshold current is used to minimize the heating of the device. This limits the intensity of the laser and hence the absorption path length through the gas. The aversion to heating also limits the duty cycle of the device. In contrast, the amplitude of the current pulse applied in the long pulse method, is deliberately several amps above threshold to cause heating and to obtain a rectangular pulse. Pulse-to-pulse shot noise is also a limitation of the short pulse technique.

2.10 The limit of analogue resolution in frequency swept systems

2.10.1 Introduction

The intrinsic FWHM linewidth of an unswept QCL has been found to be approximately 4×10^{-7}cm^{-1} (12kHz) - using a frequency stabilized cryogenically cooled cw-QCL (Williams et al., 1999). However, a picture of a swept intrinsic linewidth is not a good model for the resolutions achieved. This is due to a given technique's methodology and the influence of the instruments. The time to wavenumber relationship found from etalon fringes during a sweep (Section 3.3.3

and Figure 4-1), is essentially a relationship between time and centre wavenumber, giving no information about the frequency spread of the measurement.

Nelson et al. (2006) used the apparent broadening of nitric oxide lines during a 5kHz current sweep to calculate an effective laser FWHM of $0.0008cm^{-1}$ - for a room temperature cw-QCL - swept with standard current and temperature controllers (ILX Lightwave DC power supply and Alpes Lasers TC51 cooler). This effective linewidth is below the Doppler widths of the CH_4 and N_2O transitions studied at low pressure in this thesis (approximately $0.0040cm^{-1}$ and $0.0024cm^{-1}$ respectively). The resolution of Nelson's system would therefore be Doppler limited (at low pressure) for these transitions. In contrast, the resolution obtainable from pulsed-QCL systems tend to be limited by their comparatively large effective laser linewidths – related to their high chirp rates.

In general, a best resolution estimate of any frequency swept system can be made, and is proportional to the square root of the laser's chirp rate (Equation 2-70). This relation appears in a paper by Gorshunov et al. (1979) in the context of semiconductor diode laser spectroscopy. It also appears in McCulloch et al.'s (2003) intra-pulse work based on Kosterev et al.'s (2000) considerations on optimum pulse length in inter-pulse spectroscopy. The best resolution formula can be used for interpulse, intrapulse and cw-QCL swept spectroscopy. However, the interpretation of the bandwidth and time duration in the formula's derivation differs between the first and the latter cases.

In the following discussion it is helpful to clearly separate the meaning of two bandwidths associated with photovoltaic detectors. The first bandwidth is the range of radiation frequencies which can activate the detector – that is, those which create electron-hole pairs by promoting electrons across the semiconductor bandgap. These optical oscillation frequencies are too fast for the detector to resolve (~7.8μm. = 38.4THz for the cw-QCL used in Chapters 4 and 5). Instead, the detector 'sees' an average intensity of the oscillation amplitude. This 'activation bandwidth' is

particularly relevant to the inter-pulse method (Section 2.9.5.1) when considering the concepts of a laser's instantaneous linewidth and the chirp rate at which the average optical intensity's frequency content changes during a sweep.

The second bandwidth relates to the ability of the detector to follow temporal changes in the average optical intensity as its frequency content sweeps – for instance, intensity changes due to molecular absorptions. For the VIGO detector and transimpedance amplifier used in Chapters 4 and 5, this time resolution determining bandwidth is ~1kHz to many MHz. This bandwidth is particularly relevant to the time resolution of the intra-pulse and continuously swept methods.

2.10.2 The effect of chirp rate on the resolution in inter-pulse spectroscopy

As discussed (Section 2.9.5.1), inter-pulse spectroscopy uses a single laser pulse for each frequency datum. Resolution will be degraded by the activation bandwidth of the pulse which activates the detector. This spread of frequencies can be estimated by the magnitude of the change in laser centre frequency during the chirp. If the resulting radiation pulses each have duration τ the frequency width due to the laser chirp rate $d\nu/dt$ will be:

$$\Delta\nu_{chirp} = \frac{d\nu}{dt}\Delta\tau \tag{2-66}$$

Frequencies within this single datum pulse will therefore not be resolved. The driving current pulses are therefore kept short, typically 3-10ns. Successive radiation pulses are then tuned across an absorption profile by means of an added current ramp to record the whole spectrum. When using QCL's, this added current ramp is kept at sub-threshold values to further limit heating and thus limit the laser chirp rate.

The minimum spread of frequencies in a laser pulse can also be estimated from the bandwidth theorem. The product of a pulse's frequency bandwidth, $\Delta\nu_{bandwidth}$, and duration, $\Delta\tau$, can always be formed. In general, this product will not be minimal. According to Fourier's theorem (Hecht, 1987), any time signal can be considered to be composed from a sum of infinite duration (monochromatic) sinusoids - for a periodic signal, or an integral of infinite duration sinusoids - in the general case. The components which have significant amplitude depend upon the pulse shape and duration, τ . There are an infinite number of ways in which the various components can be assembled according to the number of significant components and their frequencies, amplitudes and phases, in order to form the resultant time signal. However, there is a minimum frequency spread, $\Delta\nu_{min\text{-}bandwidth}$, which can make up a time signal of duration τ , forming a product C - this is the bandwidth theorem. The bandwidth theorem implies:

$$\Delta\nu_{min-bandwidth} = \frac{C}{\Delta\tau} \tag{2-67}$$

C is a constant which depends upon the signal shape and the particular definitions used for $\Delta\nu$ and $\Delta\tau$. For a rectangular pulse in the time domain, which has a sinc function in the frequency domain, and using their FWHM's implies that C =0.886 (Demtröder, 2008b). For a Gaussian pulse in the time domain, which has a corresponding Gaussian function in the frequency domain, defined by their FWHM's, C =0.441. A rectangular time pulse implies an instantaneous turn on and turn off of the pulse. Fourier analysis implies that even a monochromatic wave when truncated is composed of a spread of infinite sinusoids. A Gaussian time window implies a particular build up and reduction in the pulse's signal amplitude with time. C =1 can be taken as a limiting approximation in Equation 2-67.

Equation 2-67 implies that reducing the pulse duration, to minimize chirping of the centre frequency, $\Delta\nu_{chirp}$ (Equation 2-66), causes the minimum significant frequency components which make up the pulse, and which activate the detector, to spread,

$\Delta\nu_{min-bandwidth}$. Likewise, increasing the pulse duration to minimize the Fourier spread increases the chirp of the centre frequency. The best that can be done is therefore to balance these competing effects by making $\Delta\nu_{chirp} = \Delta\nu_{min-bandwidth} = \Delta\nu_{optimum}$. The optimum resolution obtainable is therefore calculable from Equation 2-66 and Equation 2-67:

$$\Delta\tau = \frac{\Delta\nu_{optimum}}{\left|\frac{d\nu}{dt}\right|} = \frac{C}{\Delta\nu_{optimum}} \tag{2-68}$$

So that

$$\Delta\nu_{optimum} = \sqrt{C\left|\frac{d\nu}{dt}\right|} \tag{2-69}$$

and

$$\Delta\nu_{optimum}(Hz) = \sqrt{C\left|\frac{d\nu}{dt}\left(\frac{Hz}{s}\right)\right|} \tag{2-70}$$

and

$$\Delta\tilde{\nu}_{optimum}\left(cm^{-1}\right) = \sqrt{\frac{10C}{c_0}\left(cm^{-1}ms\right)\left|\frac{d\tilde{\nu}}{dt}\left(\frac{cm^{-1}}{ms}\right)\right|} \tag{2-71}$$

Therefore the optimum pulse resolution is a function of the chirp rate. For instance, in Kosterev et al.'s paper (2000), they found an optimum resolution of 0.0097cm^{-1} (290MHz). This is larger than typical Doppler widths. In general, inter-pulse systems tend to be chirp rate limited when measuring low pressure gases.

2.10.3 The effect of chirp rate on the resolution of intra-pulse spectroscopy

In intra-pulse spectroscopy, no attempt is made to minimize the pulse duration and a whole spectrum is recorded within the pulse. The time to frequency relationship, determined from etalon fringes, maps the time abscissa to a laser centre frequency abscissa. This gives a similar looking spectrum in the time and frequency domains.

The resolution is set by the time resolution, $\Delta\tau$, of the detector / preamplifier / digitizer system. If the time signal can not be resolved, or is distorted, within a duration $\Delta\tau$ on the time axis, the portion unresolved on the frequency axis, $\Delta\nu_{chirp}$, is:

$$\Delta\nu_{chirp} = \left|\frac{d\nu}{dt}\right|\Delta\tau \tag{2-72}$$

The chirp rate, $d\nu/dt$, depends on the particular laser (primarily heat dissipation from the active layer) and the operating parameters. The chirp rate also affects the spread of Fourier frequency components in the time signal. For a given steepness of slope of an absorption line in the frequency domain, the steepness of slope in time depends on the chirp rate. From Fourier's theorem, more rapidly changing signals are synthesized from significant higher frequency components. This affects how small $\Delta\tau$ can be made. The detection system needs a fast enough rise time to capture the full signal. In the frequency domain, its bandwidth must be wide enough to respond to the high frequency content of fast time changes. If the detection bandwidth is too small, components making up the steep edges will be lost, and the measured signal flattens and broadens.

The minimum spread in Fourier frequency components, $\Delta\nu_{line}$, of a given spectral feature measured in time can be found from the bandwidth theorem:

$$\Delta\nu_{line} = \frac{C}{\Delta\tau} \tag{2-73}$$

$\Delta\nu_{line}$ indicates the minimum bandwidth the detection system needs for a time resolution of $\Delta\tau$. To lower the demands on a detection system it is advantageous to sacrifice time resolution by lengthening $\Delta\tau$, lowering the bandwidth needed, $\Delta\nu_{line}$. Lengthening $\Delta\tau$, however, reduces the resolution on the frequency axis by

lengthening $\Delta\nu_{chirp}$. Therefore, a compromise which optimizes the resolution (limits the increase in $\Delta\nu_{chirp}$), while mutually optimizing the detection bandwidth requirement (decreasing $\Delta\nu_{line}$), occurs when $\Delta\nu_{optimum} = \Delta\nu_{line} = \Delta\nu_{chirp}$ (and $\Delta\tau$ becomes $\Delta\tau_{optimum}$). This argument then gives Equation 2-70 for $\Delta\nu_{optimum}$.

Equation 2-70 is useful as a limiting case in the intra-pulse case because the high chirp rates of cw-QCL's generally imply that the detection bandwidth is not large enough to reduce $\Delta\tau$ to $\Delta\tau_{optimum}$. (Equation 2-68 implies $\Delta\tau_{optimum} = \sqrt{C\left(\left|dv/dt\right|\right)^{-1}}$). In principle, with infinite detection bandwidth, $\Delta\tau$ could be made arbitrarily small until the resolution becomes limited by the instantaneous frequency of the laser. However, detection bandwidths, $\Delta\nu_{sys-band}$, are finite. In general the time resolution is inversely proportional to the detection bandwidth:

$$\Delta\nu_{sys-band}\,\Delta\tau = D \tag{2-74}$$

For a linear first order system D would be ≈0.35 (Buzuayene, 2008). $D = 1$ can be taken as a limiting value.

As a numerical example, a detector rise time of 8ns implies a detector bandwidth of ~125MHz (Equation 2-74 with $D = 1$). A (median) pulsed-QCL chirp rate from Figure 4-46 is 81MHz ns^{-1}. The 125MHz detection bandwidth is therefore less than the optimum 285MHz required by Equation 2-70 for this chirp rate. Taking 8ns as an approximation to the smallest resolvable time period, this time resolution corresponds to a 648MHz (0.0216cm^{-1}) centre frequency resolution (Equation 2-72). The optimum resolution of 285MHz (0.0095cm^{-1}) implies an optimum time resolution of 3.5ns. (Both resolutions are larger than, say, the methane Doppler width of 0.004cm^{-1}).

The chirp rate varies substantially during the long pulses of the intra-pulse technique. The optimum resolution can be taken as a guide to the required detection bandwidth at a particular time during the pulse. Alternatively, for a given detection bandwidth, the optimum resolution indicates where in the pulse the spectrum is likely to be adequately resolved.

2.10.4 The effect of chirp rate on the resolution of cw swept spectroscopy

The arguments of the last section can be applied to the case of a general swept system – in particular the spectra resulting from current sweeping a cw-QCL. In principle, in this case, the chirp rates can be made arbitrarily small. This implies that the time resolution, $\Delta\tau$, can be reduced to $\Delta\tau_{optimum}$ with modest detection bandwidths. The sweep used in Chapter 4 takes place at 10kHz and a median chirp rate is 0.14MHz ns^{-1} (Figure 4-24). This implies an optimum detection bandwidth of 0.38MHz ($1.3\times10^{-5}cm^{-1}$) (Equation 2-70 with $C=1$). This is smaller than a typical Doppler width and smaller than typical effective QCL linewidths when subject to fluctuations induced by standard current and temperature controllers (Section 4.6.4). The implication is that the chirp rate is unlikely to be the limiting factor on resolution for most moderately swept systems.

2.11 The Detector Types

For the mid-infrared spectroscopy in Chapters 4, 5, and 7, Peltier cooled VIGO HgCdTe photovoltaic detectors were used. Absorbed photons create electron-hole pairs in the depletion region of the detector's p-n junction. These are swept through the junction, creating a current in an external circuit. The bandgap of the alloy determines which wavenumbers can be detected. In photovoltaic mode, the diode is connected across a high-impedance load and the changing voltage measured. A photovoltaic set-up provides a fast response, as does a high electron mobility in the alloy. Using the detector with AC-coupling avoids much low frequency noise. Unlike

the textbook logarithmic response (Quimby, 2006), the detectors from VIGO have a linear response over most of their range (Section 2.6.4).

In Chapter 6, in the visible, an Avalanche Photodiode Detector (APD) was used. Photons create electron-hole pairs in the depletion region of an n^+-p-i-p^+ semiconductor structure. Under a sufficiently high reverse bias, the accelerated carriers create more and more electron-hole pairs in an accumulative process. Thus, signal amplification results.

In mid-infrared Fourier Transform spectroscopy (Chapter 4), the spectrometers are customarily equipped with liquid nitrogen cooled bolometric detectors. Incident photons heat a fine wire or metal strip, causing a measurable change in resistance.

Chapter 3

Data Analysis Program

3 Data Analysis Program

3.1 Abstract

A software gap existed in the automation of producing absorbance spectra from pre-recorded data or data straight from an oscilloscope. To plug this gap, a suite of interactive LabVIEW programs was developed - bringing raw spectra in Volts vs Time format through to Absorbance vs Wavenumber format for Sweep Integration (SI) or Volts vs Wavenumber for Wavelength Modulation Spectroscopy (WMS). The program has been written to incorporate in-built LabVIEW algorithms where possible. The processed SI spectra are then fitted with a Gaussian or Lorentzian line shape in order to extract a gas concentration. The 2f-WMS spectra are peak fitted to obtain signal amplitudes. A measurement of the noise levels of the respective techniques facilitates a sensitivity comparison to be made.

3.2 Introduction

No automated software was available for producing absorbance spectra and extracting concentration values from pre-recorded voltage vs time data, or data taken straight from an oscilloscope. Cascade Technologies' software was tied to their 'on-the-fly' pulsed-QCL systems.

What started as an exercise in obtaining a gas concentration, developed into an interactive data analysis program written in LabVIEW. As well as general data processing, the program's main application, in this thesis, is to allow data from the Wavelength Modulation Spectroscopy (WMS) technique (Section 5.3) to be processed and calibrated against data processed from the Sweep Integration (SI) technique (Section 4.3). A comparison of the sensitivities of the techniques can then be made from their respective absorbance noise floors (Section 5.8).

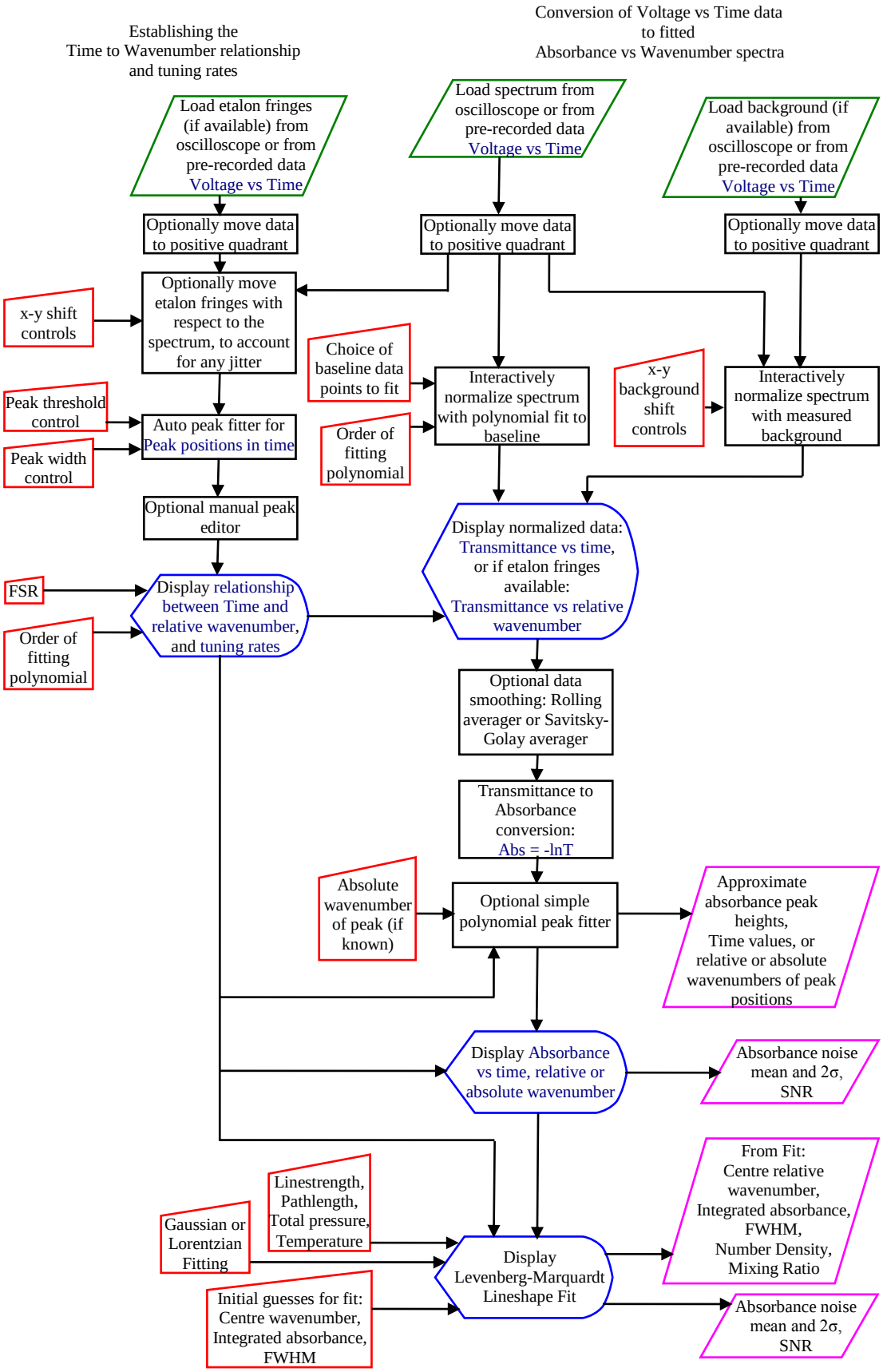


Figure 3-1: Data flow diagram for the Sweep Integration data processing program.

The program's aim in the Sweep Integration (SI) case is to take the experimental voltage versus time data and transform it to an absorbance spectrum – that is, absorbance versus wavenumber. In the linear regime, where an unmodified Beer-Lambert Law holds (Equation 2-51), and at sufficiently low or high pressures respectively, a Gaussian or Lorentzian fit to this spectrum allows the integrated area of the lineshape to be found. The number density is then known if the line strength (a function of temperature) and the radiation path length through the absorbing gas are known. In situations where the gas pressure is also stable and known, the number density can also be extracted from the fit's peak height (Equation 2-52) and, more generally, number densities can be converted to mixing ratios (Equation 2-53). The program also allows the absorbance noise mean, 2σ, and signal to noise ratio to be read off.

Figure 3-1 shows the data flow for the SI analysis. In order to exclude intensity changes due to the laser from those due to molecular absorption, the voltage versus time data is normalized. This can be done interactively by division with a movable measured background or via a polynomial fit to the data's baseline. After normalization the data's ordinate measures transmittance. This can then be converted to absorbance, proportional to the gas's number density, via Equation 2-44. Data smoothing is offered as an option.

If no data is available on the laser's tuning, the program continues with the data's abscissa as a time axis. This is equivalent to direct proportionality between time and relative wavenumber with gradient k=1. This still allows the absorbance peak height and noise data to be collected, and a first approximation to be made to the lineshape's integrated area. If etalon fringe data is available, the actual relationship between time and relative wavenumber can be determined via automatic and manual peak fitting, and tuning rates displayed. If the wavenumber of a peak is known, the relative wavenumber axis can be shifted to an absolute wavenumber axis. Figure 4-1 (Chapter 4) provides a graphical representation of the data processing.

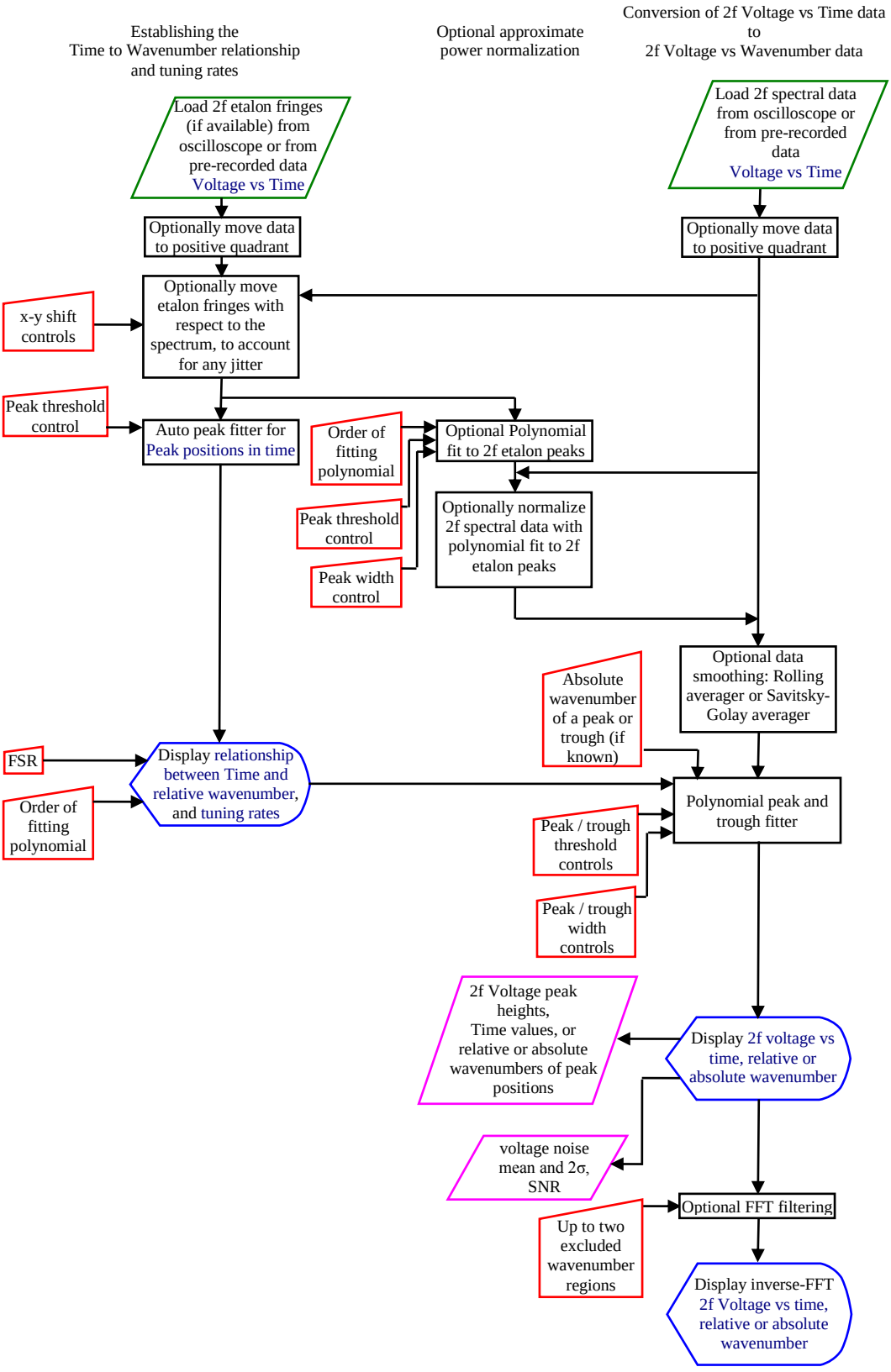


Figure 3-2: Data flow diagram for the Wavelength Modulation Spectroscopy data processing program.

The data flow for the 2nd harmonic Wavelength Modulation Spectroscopy (2f-WMS) processing is shown in Figure 3-2. In this technique, the baseline is relatively flat and so intensity normalization is not required when studying a single absorption line.

Voltage peak heights (for calibration with sweep integration absorbances) and noise data are of most interest here. The program allows these to be quickly obtained and read off. Data smoothing is an option. Periodic noise, say from unwanted etalon fringing in the experimental beam path, can also be removed by post-acquisition Fourier transform filtering of the relevant frequencies. Intentional etalon fringe data, if available, can be used to display laser tuning rates and convert the time abscissa to relative wavenumber. This is achieved, as in the sweep integration case, by peak fitting the (abscissa) time positions of the etalon peaks and mapping them to a constant free spectral range. Fitting a polynomial to the trend of the (ordinate) etalon peak voltages can be used to intensity normalize the heights of two or more peaks in the same WMS data set. This is useful when different sweep parameters have been used for the sweep integration, so that the SI baseline cannot be used.

The data analysis program was central to processing and analyzing the data in Chapters 4, 5 and 6. The program therefore added greatly to the efficiency and functionality of what was available, and was particularly helpful in processing large numbers of data sets - particularly in weeding out sets affected by electronic compatibility problems seen in Chapters 4 and 5. The task of programming also lead to a deeper understanding of the equations, models and processes involved.

In the remainder of this chapter, a description of the program's functionality and workings is laid out by tab, with screenshots shown. Table 3-1, at the chapter's end, provides a look up table of the program's functionality by tab.

Some help on the basics of LabVIEW programming was provided by Travis and Kring's book (2009) and LabVIEW help files. However, most of the program was

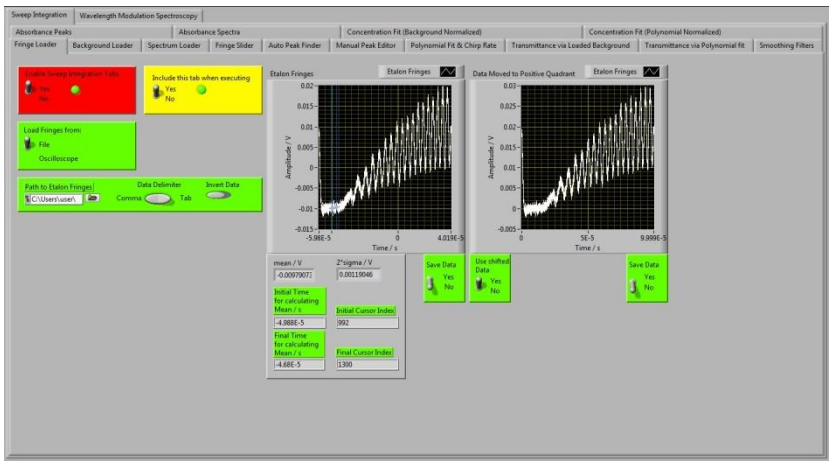


Figure 3-3: Etalon Fringe Loader Tab. The ac data can be optionally shifted to the positive quadrant via the mean value between a pair of slidable cursors.

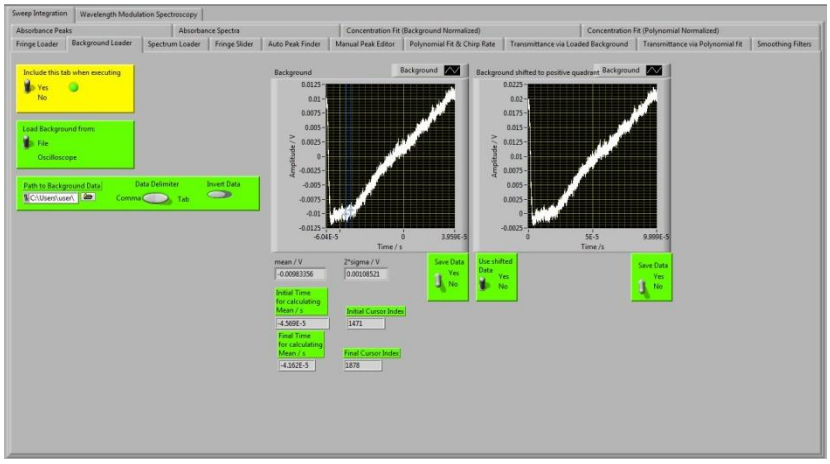


Figure 3-4: Background Loader Tab. The ac data can be optionally shifted to the positive quadrant via the mean value between a pair of slidable cursors.

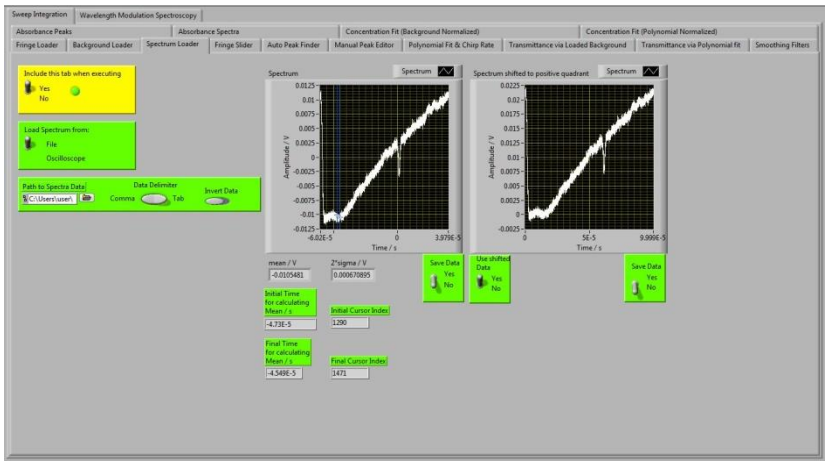


Figure 3-5: Spectrum Loader Tab. The ac data can be optionally shifted to the positive quadrant via the mean value between a pair of slidable cursors.

constructed by trial and error on LabVIEW's functions, structures and data array manipulations.

3.3 Sweep Integration by Tab

3.3.1 Fringe, Background and Spectrum Loader Tabs

See Figure 3-3, Figure 3-4, and Figure 3-5. The program is set up to take pre-recorded or oscilloscope voltage vs time data. If only the spectrum is available, it can be worked through to absorbance vs time - via a polynomial fit to the baseline. If a background is available, this can be used to normalize the spectrum. To convert the time axis to a wavenumber axis - and hence calculate concentration values - etalon fringes are needed.

The first three tabs of the sweep integration analysis are used to load the fringes, background and spectra - as available. Pre-recorded data is accepted in either tab or comma delimited form, and can be inverted to account for the way it was recorded on the oscilloscope. Data could also be loaded straight from the Tektronix oscilloscope. This bit of programming was stripped down to just what was needed from a publicly available LabVIEW program for running the whole oscilloscope (National Instruments).

In Section 4.6.1.1 the below threshold output of a cw-QCL is used to approximate a 0V level - to circumvent the loss of this information due to ac-coupling. The loading tabs facilitate using this below threshold level to shift the data to the positive quadrant if desired - and thus avoid divide by zero errors. A 'property node' on the program's block diagram stores the data row numbers of two movable cursors on the data plot. The average of the data between the cursors is then calculated and used to shift the average to zero, and the initial time value is used to start the time axis from zero.

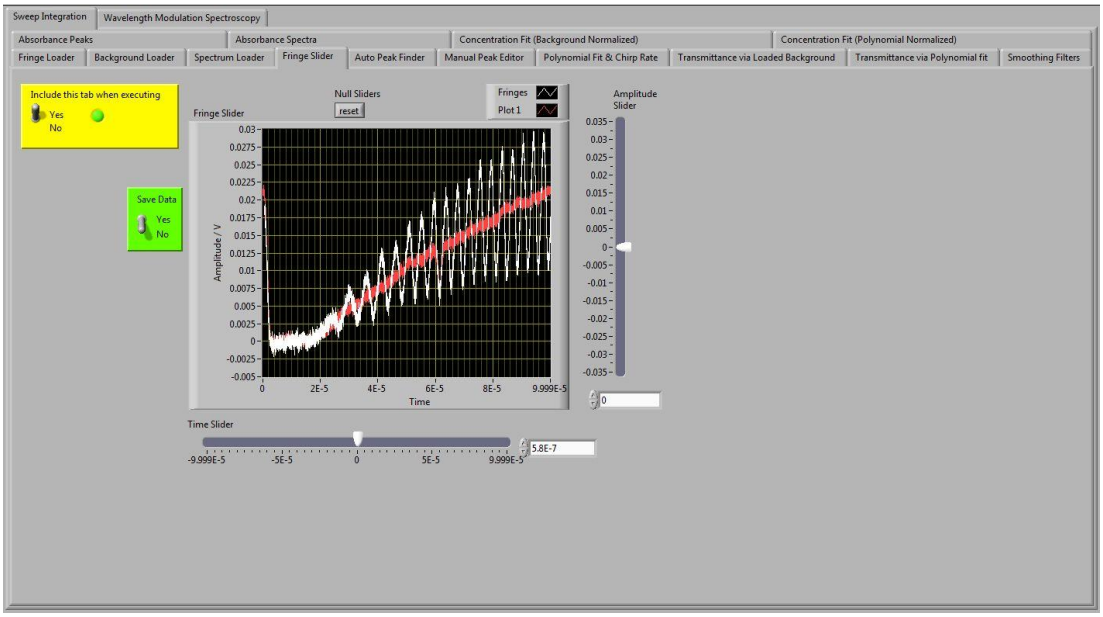


Figure 3-6: Fringe Slider Tab, via two sliders.

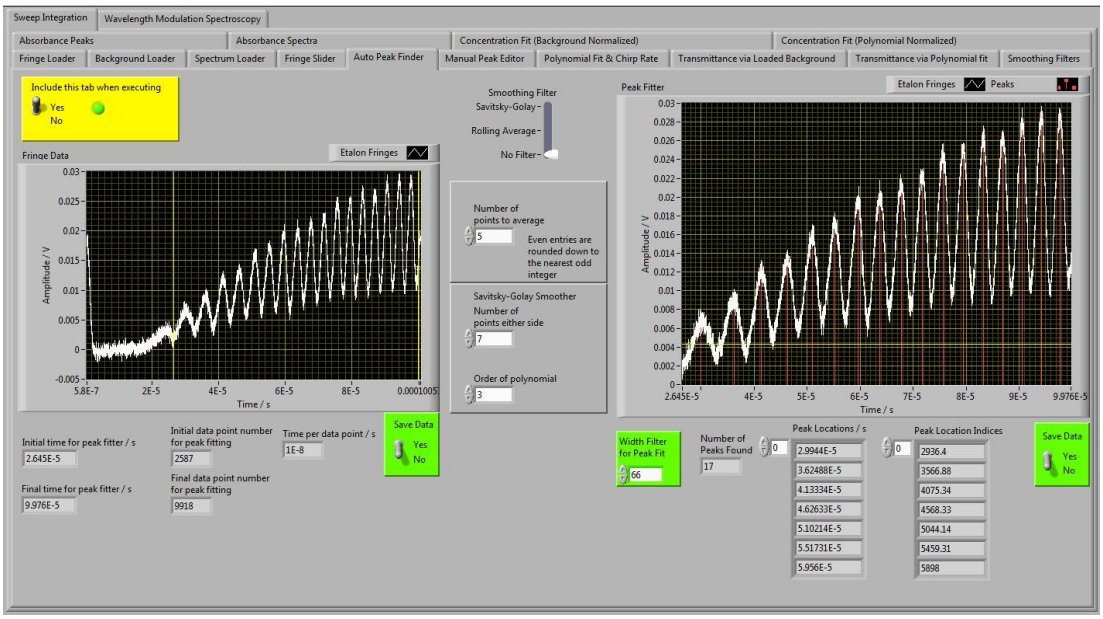


Figure 3-7: Automatic Peak Finder Tab. The peak recognition is via LabVIEW's quadratic fit peak finder and is controlled by a threshold slidable cursor and a peak width control.

3.3.2 Fringe Slider Tab

See Figure 3-6. Occasionally, jitter needed to be taken account of, or it might be desirable to use etalon fringes taken some time apart from spectra, with different oscilloscope settings. A plot was therefore produced of slidable fringes superimposed on a spectrum. The plot axes are fixed to fit both data sets in the same window to begin with. X and y sliders add or delete constants from the fringe plot to move it in relation to the spectrum. The range of the sliders is set at plus and minus the length of the respective axis, and the increment is $1/512^{th}$ of this – approximately two y-axis data points. (512 is the y-axis resolution on the oscilloscope (9-bit = 2^9)). A reset button takes the sliders back to zero.

3.3.3 Auto Peak Finder Tab

See Figure 3-7. As a prelude to converting the time axis to a relative wavenumber axis, the etalon fringes - equally spaced in wavenumber but not in time - are peak fitted. (This constancy in the wavenumber spacing is estimated in Appendix C to be as good as 3 parts per million). The etalon fringes are displayed and the row numbers of two slidable vertical cursors are used to select a data range for peak fitting. The selected data and the peak fits are presented in a second plot. The peak fitting is done by an inbuilt LabVIEW routine which fits a quadratic polynomial to a number of consecutive data points. The number of points is specified by a user controlled 'width' function. The value of the peak must also be above a 'threshold' and the value for this is taken from a slidable horizontal cursor on the plot. Usually, these two controls are sufficient to correctly pick out the wanted peaks. However, smoothing filters (discussed in Section 3.3.8) are also made available here, for processing the data prior to peak fitting.

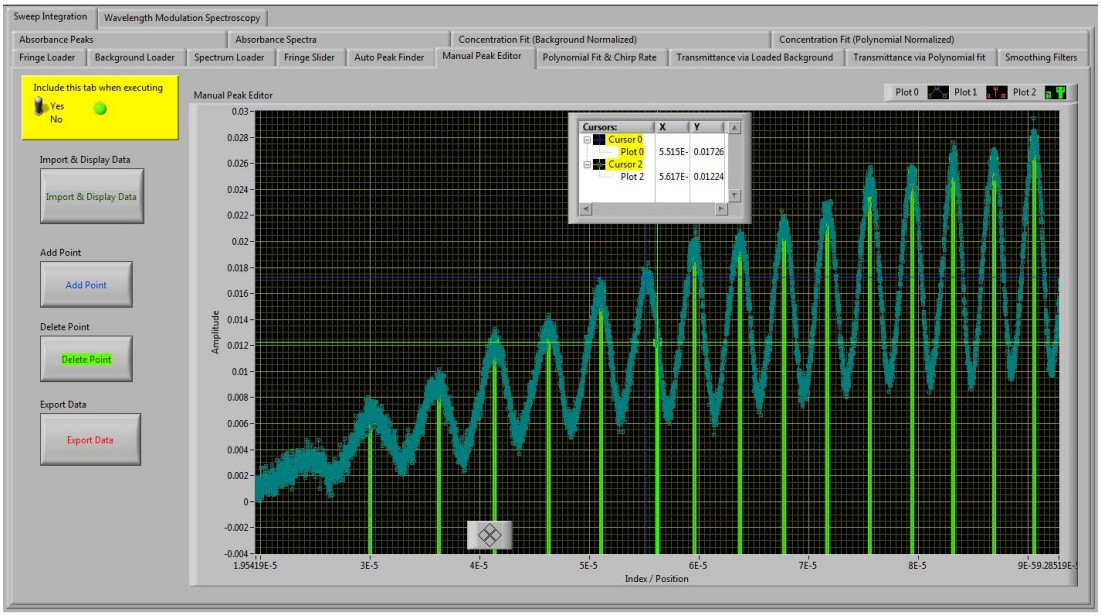


Figure 3-8: **Manual Peak Editor Tab. This part of the program is controlled by an 'event structure' triggered by the four buttons on the left. The first button inputs the results from the automatic peak fitter (Figure 3-7). Peaks can be added or deleted by the positions of respective cursors when pressing the relevant buttons. When finished the last button sends the resulting peaks to the Polynomial Fit & Chirp Rate tab (Figure 3-9).**

3.3.4 Manual Peak Editor Tab

See Figure 3-8. In the case of the low quality etalon fringes found during the oxygen spectroscopy of Chapter 6, it was convenient to have a way of manually adding and deleting peaks. This was achieved by using a LabVIEW 'event structure' - in one version of the program. Pressing the 'Import & Display Data' button passes on the fringe data and peak fitted points from the previous tab.

Peaks can be deleted by placing the green cursor over a peak and pressing the 'Delete Point' button to activate the event. The 'single plot' cursor jumps from peak to peak so that no other points interfere with the deletion process. The cursor row index is used to delete the point from the peak position data array.

Adding peaks occurs through activating the event by pressing the 'Add Point' button. This programming is more involved than that for deletion. The slidable blue cursor moves from point to point in the full fringe data set. The 'added' point is first checked to see if it is already in the peak array, if so nothing changes. If the point is new its amplitude and time are concatenated onto the peak position array. Some array manipulation is then needed to gain access to LabVIEW's inbuilt functions - to sort the peaks into the correct order. This was done by turning the peak array into a 'cluster' - where each element of the cluster contains both the amplitude and time location elements for a single peak. The cluster could then be sorted by LabVIEW's 1D array sorter by time location. The cluster was then returned to an array using an element by element converter inside a loop.

A shift register held the updated peak array for the start of the next event structure iteration. When all the required peaks were added or deleted, the 'Export Data' button closed the event structure and set the remainder of the program looping for interactivity.

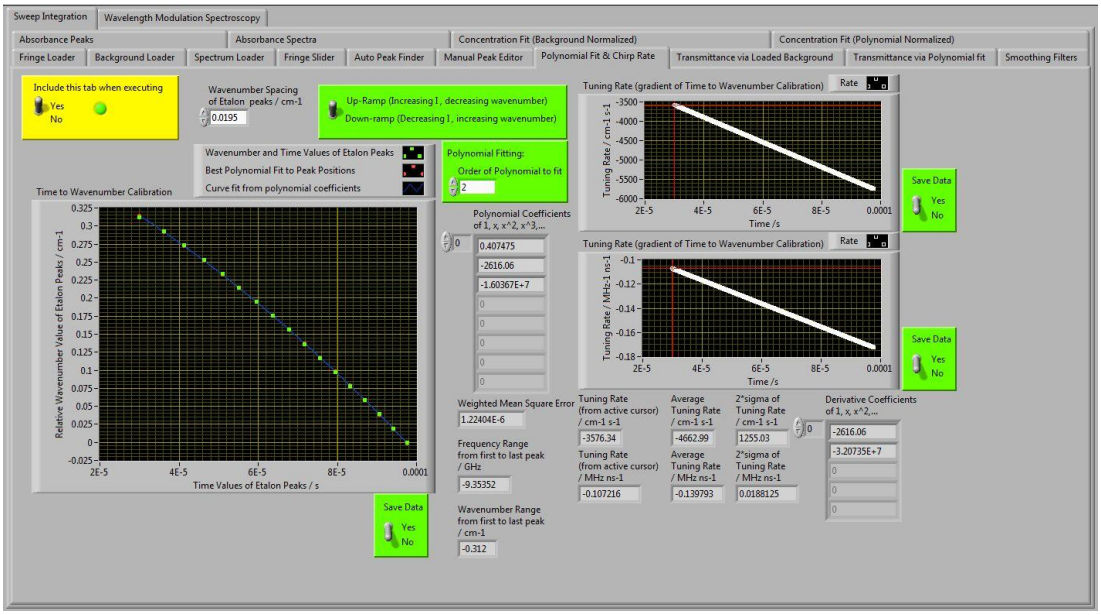


Figure 3-9: Polynomial Fit & Chirp Rate Tab. The user enters the etalon FSR here. The user chooses the order of the polynomial fit for the time to wavenumber conversion of the data. The derivative of this polynomial is displayed as the laser chirp rate in $cm^{-1}s^{-1}$ and $MHz^{-1}ns^{-1}$.

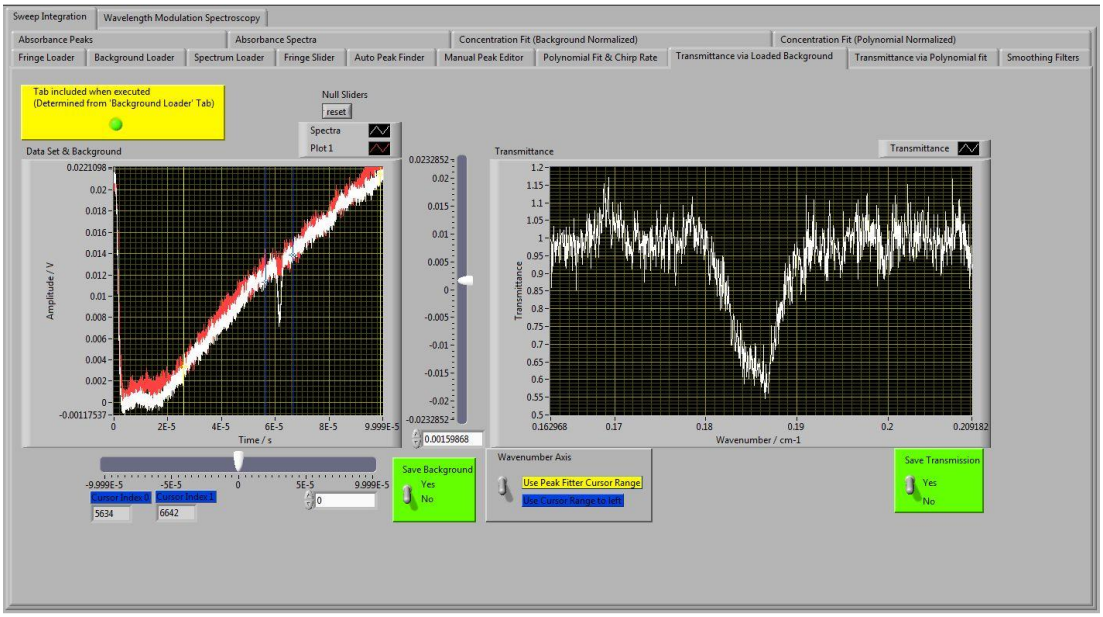


Figure 3-10: Transmittance via Loaded Background Tab. Normalization of the baseline is interactive using the x-y sliders.

3.3.5 Polynomial Fit & Chirp Rate Tab

See Figure 3-9. The peak positions are plotted. Their time values are on the x-axis and the evenly spaced relative wavenumbers are on the y-axis. The value of the wavenumber spacing is the user entered Free Spectral Range (FSR) of the etalon used. The user then chooses the order of polynomial to best fit the wavenumber – time relationship. The resulting fit line is displayed along with where the peak position time values fall on this line. This polynomial will then be used to convert all the time values of the spectrum data points to relative wavenumber.

The derivative of the fit line is displayed as the tuning rate in $cm^{-1}s^{-1}$ and $MHz^{-1}ns^{-1}$. A movable cursor displays the tuning rate for any particular time value. The polynomial fitting is done by LabVIEW's inbuilt routine. This is set to the least squares method of residue minimization on the assumption of Gaussian distributed noise.

3.3.6 Transmittance via Loaded Background Tab

See Figure 3-10. Two sliders and a reset button are used here to move a background, if available, against a spectrum – working in the same way as described in Section 3.3.2. The fringe peak fitter data range, or a range set by the user via two cursors, then picks out the data to be normalized into a transmittance plot. If no fringe data is available, transmittance is shown against time - otherwise, transmittance is shown against relative wavenumber. The interactivity provided by the looping program is useful in adjusting the background so the transmittance baseline falls at an average value of one.

3.3.7 Transmittance via Polynomial Fit Tab

See Figure 3-11. If desired, or if a background is not available, normalization of the spectrum to a transmittance vs time or relative wavenumber plot can be achieved via

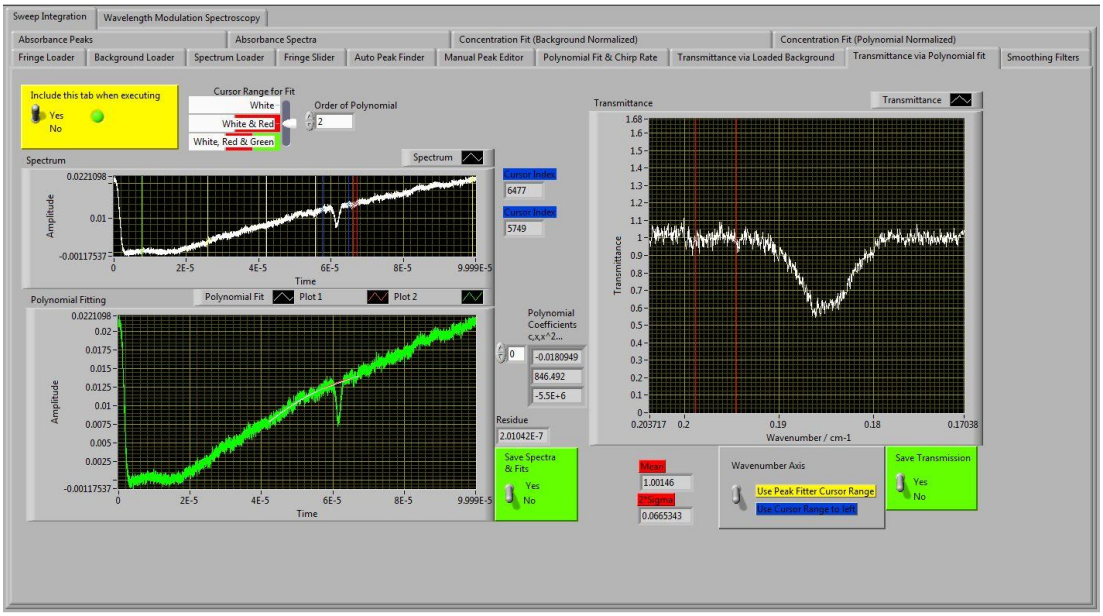


Figure 3-11: Transmittance via Polynomial Fit Tab. Up to three sections of the baseline can be used for fitting a baseline polynomial. Sliding the cursors in the first plot adjusts the polynomial, of user defined order, allowing interactivity of the normalization seen in the large plot. The quality of the polynomial fit itself can be seen against the green spectrum.

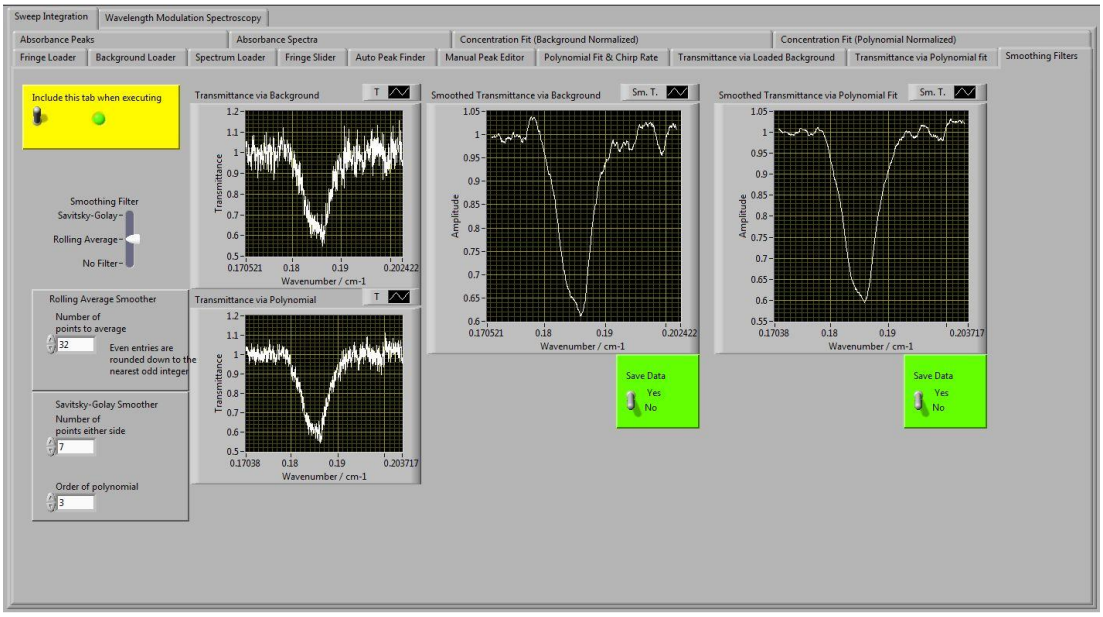


Figure 3-12: (Optional) Smoothing Filters Tab, with custom made rolling averager and a LabVIEW Savitsky-Golay filter.

LabVIEW's polynomial fitter - with the user's choice of polynomial order. For entering baseline data points into the fitter, three pairs of cursors are provided - allowing one, two or three baseline sections to contribute to the overall baseline polynomial fit. The returned polynomial coefficients are used to plot the baseline fit. This fit and the resulting transmittance spectrum are displayed to allow interactive adjustment.

3.3.8 Smoothing Filters Tab

See Figure 3-12. Although the data presented in this thesis is unsmoothed, it can be useful to observe or process a line profile without high frequency noise. Smoothing achieves this at the expense of some broadening and flattening of the feature.

A choice of two smoothers is given – a rolling averager which was custom written, and one based around the popular Savitsky-Golay routine provided by LabVIEW.

In the custom written rolling averager, most data points in the original set are replaced by the simple numerical average of a user defined window - centered on that data point. The routine first checks for an odd number of points N for averaging, or rounds down to the nearest odd integer. This facilitates a symmetrical window about the data point to be replaced. The first point to be replaced by the rolling average is the $(N+1)/2^{\text{th}}$ point - that is, the first point the averaging window can be placed around symmetrically. (This point has an array index of $[(N+1)/2]-1$ since LabVIEW counts from 0 not 1). Likewise if X is the number of points in the original data set, the last point to be fitted is the $X-[(N+1)/2]^{\text{th}}$ point. So, in the original data set the first and last $(N-1)/2$ points are unaveraged and therefore not displayed in the averaged plot.

A Savitsky-Golay filter is a more complex form of sliding filter which can preserve more of the underlying line profile when the profile is not too much narrower than the filter width (Press et al., 2007). The window need not be symmetrical about the

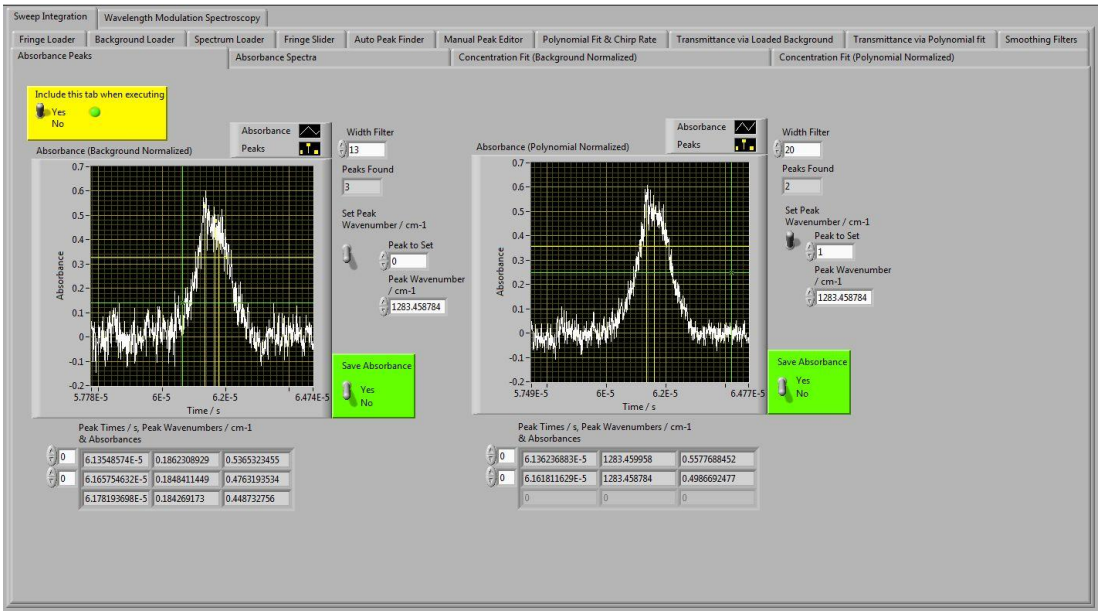


Figure 3-13: Absorbance Peak Finder Tab, via LabVIEW's quadratic polynomial peak fitter routine. The relative wavenumber axis can be optionally shifted to an absolute wavenumber axis here.

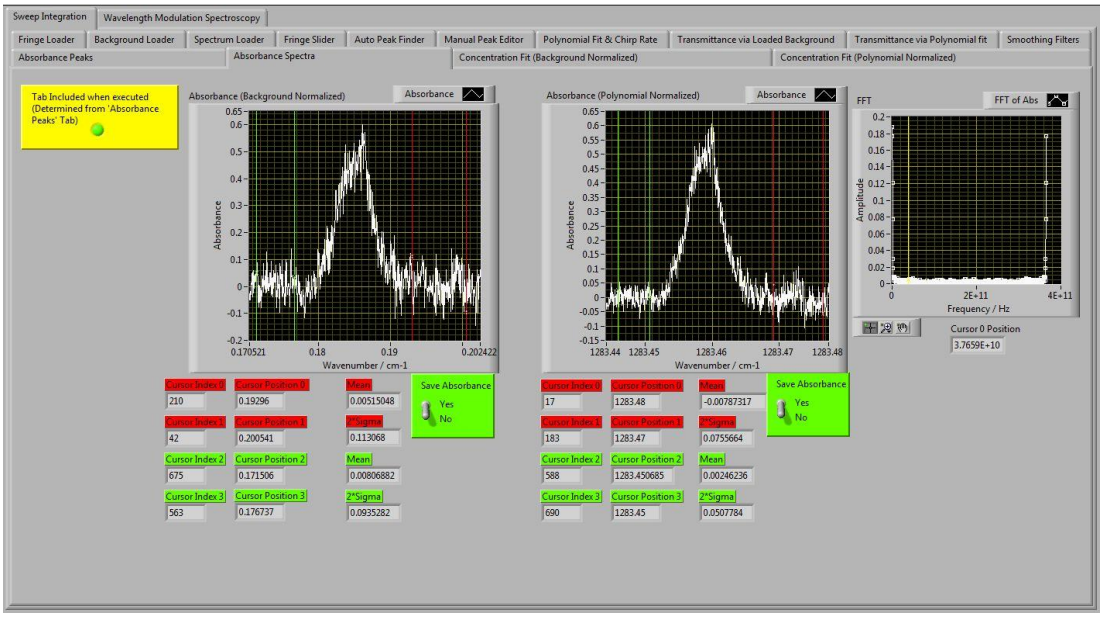


Figure 3-14: Absorbance Spectra Tab, for background and / or polynomial normalized spectra, as available. The mean and 2σ of the absorbance noise can be measured here. (Also shown is the two-sided complex FFT output of the polynomial fitted absorbance. The wavenumber axis approximates an axis of equally spaced time points for this purpose).

point of interest. In essence, the filter fits a polynomial of user defined order about each data point, and replaces the point of interest with its value on the polynomial.

3.3.9 Absorbance Peaks Tab

See Figure 3-13. This tab uses LabVIEW's peak fitter to approximate the absorbance peak line centres. The data analysis program allows a particular peak to be chosen and set to a user defined wavenumber. In the programming, the difference between this wavenumber and the peak's relative wavenumber is used to shift the whole absorbance spectrum to absolute wavenumber. This can also be done from the concentration fit tab using the centre position of a fitted line profile (Section 3.3.11).

3.3.10 Absorbance Spectra Tab

See Figure 3-14. This displays the finished absorbance spectra from background and polynomial normalization, as available. Two pairs of slidable cursors allow the baseline either side of a feature to be selected for mean and 2σ baseline noise measurements. An overall mean and 2σ value is also calculated from the combined raw data.

3.3.11 Concentration Fit Tabs

See Figure 3-15 and Figure 3-16. These tabs display the background or polynomial normalized absorbance spectra for fitting with a line profile. As in the previous tab, baseline mean and 2σ values can be displayed. The line fitting is done via LabVIEW's Levenberg-Marquardt fitter – which is a nonlinear least squares method (Press et al., 2007). The fitter requires an input of the fitting model via a 'static reference'. This reference refers to a sub-program in which the model is made, taking each wavenumber value and outputting the corresponding model value. The adjustment in the fitting is done by inputting a cluster of parameters which can be varied in the model. In the case of Gaussian and Lorentzian fits these are chosen as

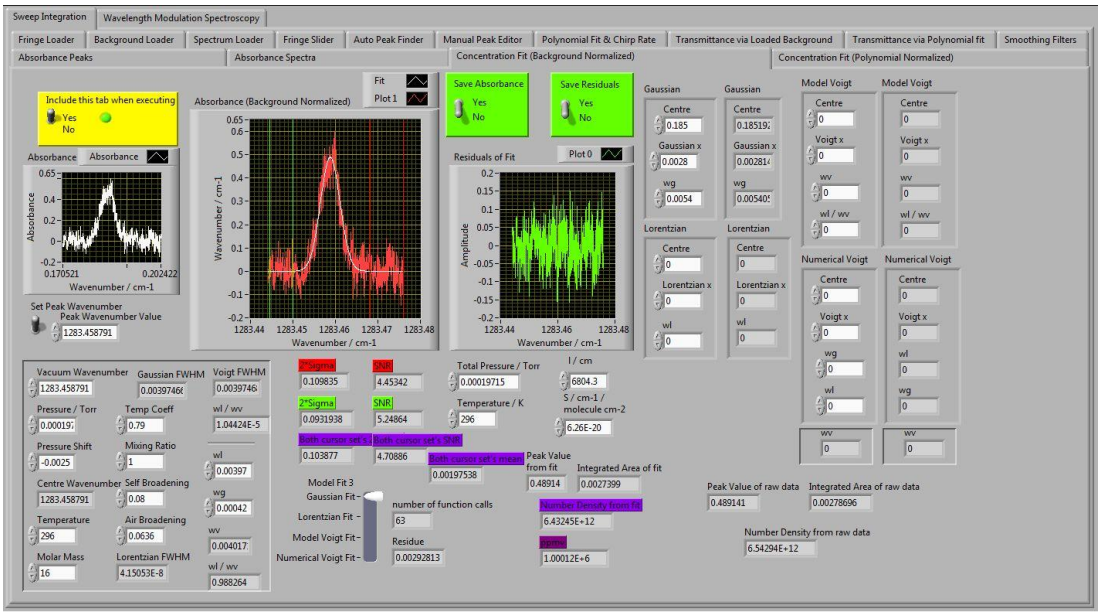


Figure 3-15: Concentration Fit Tab, for background normalized data. Gaussian and Lorentzian fits are currently fully functional. The mean and 2σ of the absorbance noise can be measured here.

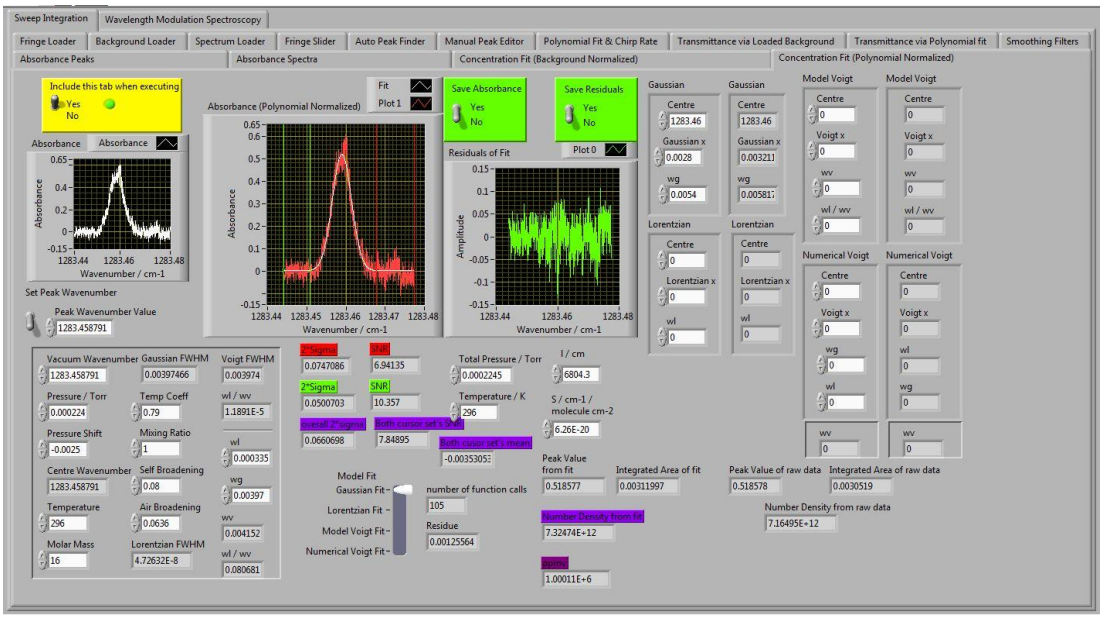


Figure 3-16: Concentration Fit Tab, for polynomial normalized data. Gaussian and Lorentzian fits are currently fully functional. The mean and 2σ of the absorbance noise can be measured here.

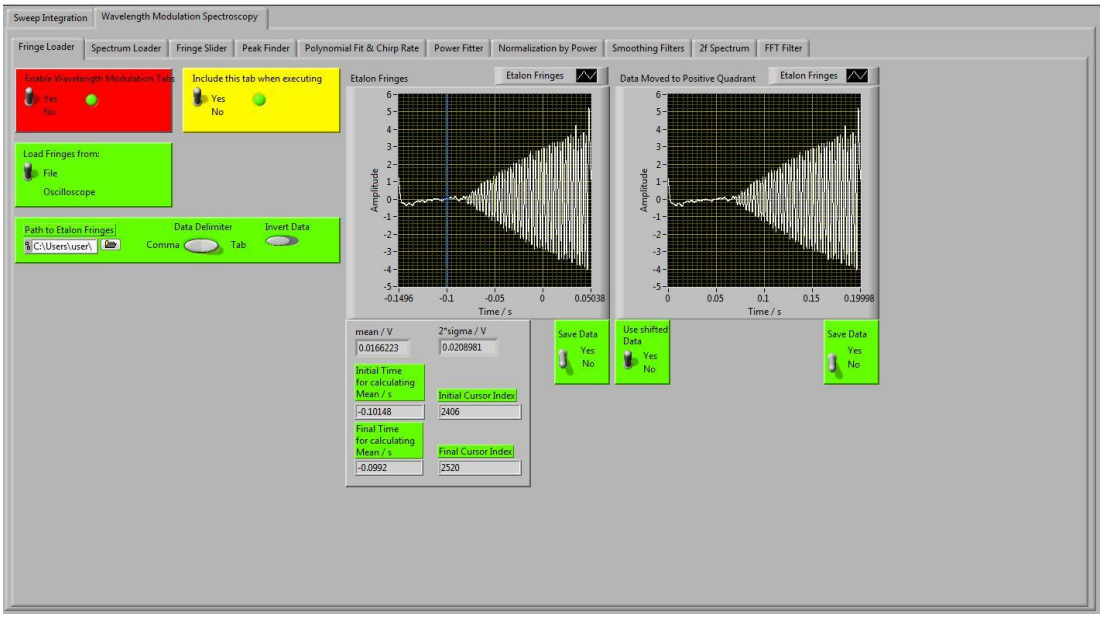


Figure 3-17: Etalon Fringe Loader Tab. The fringes are from the 2f-WMS signal of the etalon fringes. The time axis can be shifted to positive values, and the mean value between two slidable cursors can be used to refine the baseline about 0V.

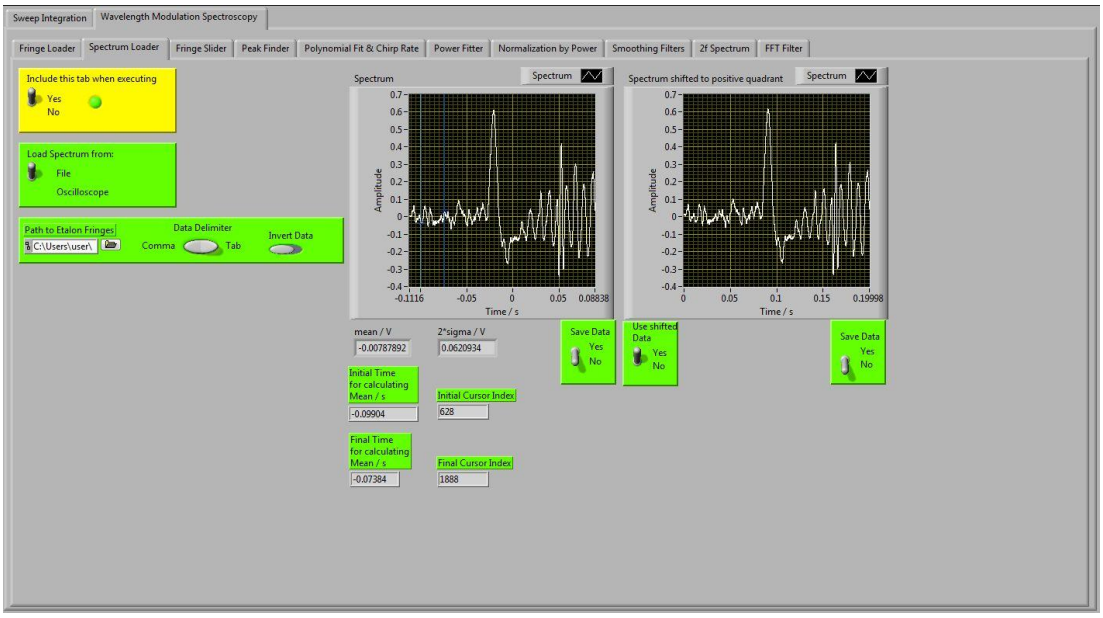


Figure 3-18: 2f-WMS Spectrum Loader Tab. The large peak in the data corresponds to the onset of the current ramp. This particular data set shows a slightly misaligned optical system with unwanted etalon fringing of the same order of magnitude as the 2f signal. (This data set will illustrate filtering in the frequency domain in Figure 3-26).

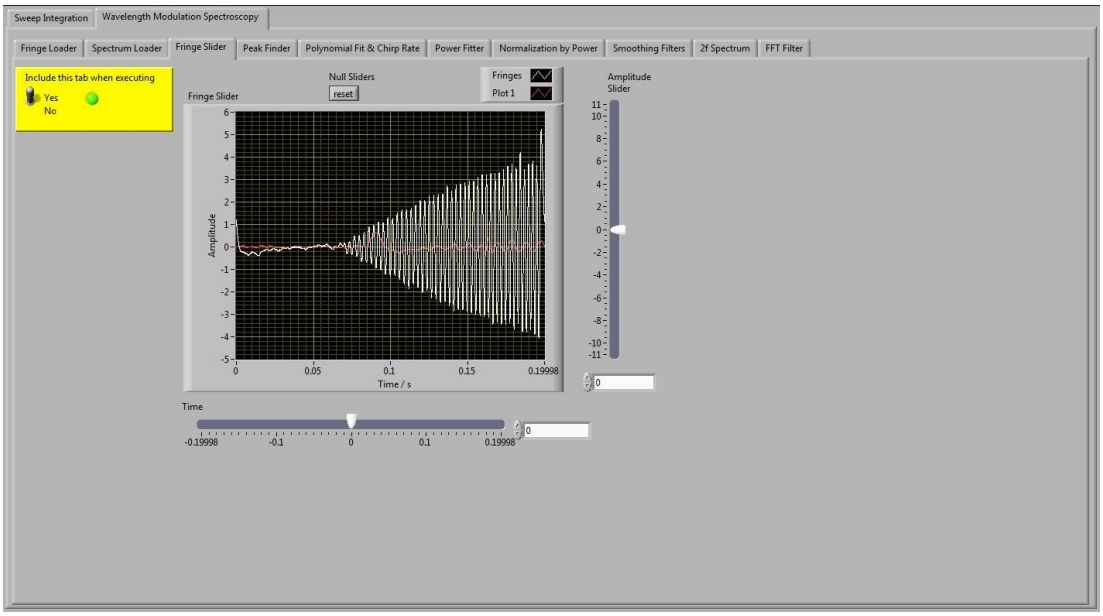


Figure 3-19: Etalon fringe Slider Tab, via two sliders.

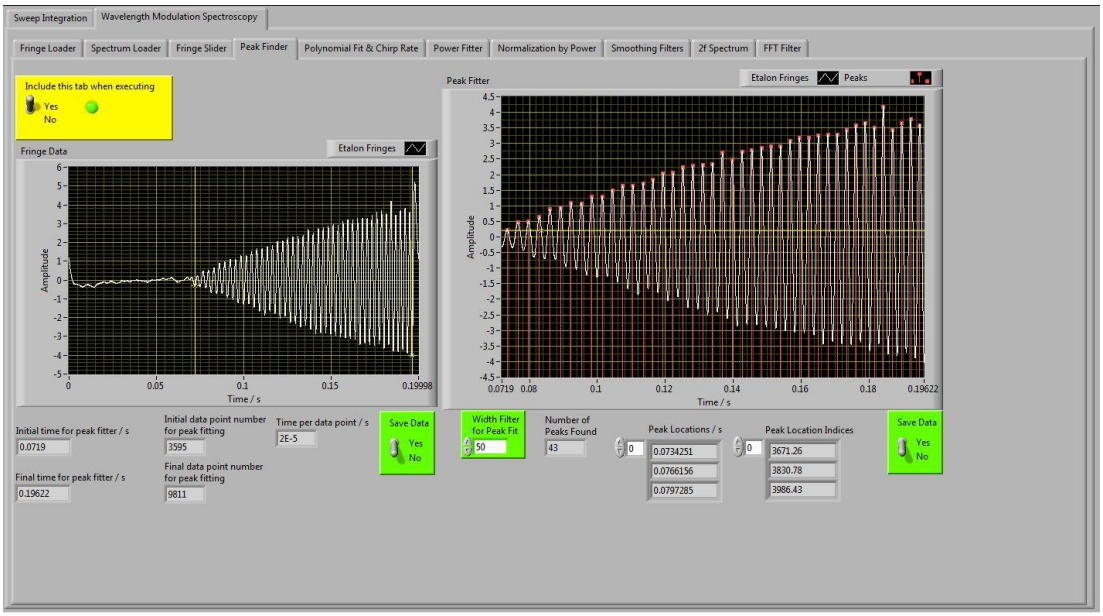


Figure 3-20: Peak Finder Tab, applied to the 2f-signal of the etalon fringes. The peak recognition is via LabVIEW's quadratic fit peak finder and is controlled by a threshold slidable cursor and a peak width control.

the centre position, FWHM, and the amplitude of the lineshape function. The resulting best fit parameters are displayed and the lineshape they define plotted against the measured data. At the time of writing, a pseudo-Voigt fit (Section 2.3.4) was being developed. This fit is more involved due to the greater choice of possible variable parameters and the need to stop some parameters going negative.

Once the best lineshape has been defined, its area can be found via, for instance, the trapezoidal rule. Knowing the path length and linestrength gives the number density (Equation 2-51). If the pressure and temperature are known the ppmv concentration can be obtained (Equation 2-53). The fit's centre wavenumber, if known, can be used for shifting the relative wavenumber scale to an absolute one. A calculator is also provided to show the centre wavenumber and FWHM expected from HITRAN parameters (based on the equations of Section 2.5).

3.4 Wavelength Modulation Spectroscopy by Tab

3.4.1 Fringe and Spectrum Loader Tabs

See Figure 3-17 and Figure 3-18. The etalon fringe and spectrum loading tabs work in the same way as described in Section 3.3.1. However, no background or baseline polynomial fit is needed for baseline normalization. The second harmonic (2f) signals are proportional to the second derivative of the signal (Section 5.3.3). Hence, since the baselines here are of gentle curvature the 2f baseline flattens out without normalization. The averaging done by the data shifter on this tab can be used to refine the baseline about zero. The time axis can be shifted to positive values.

3.4.2 Fringe Slider, Peak Finder, and Polynomial Fit & Chirp Rate Tabs

See Figure 3-19, Figure 3-20 and Figure 3-21.The functionality of these tabs is the same as in Sections 3.3.2, 3.3.3 and 3.3.5 - providing wavenumber coverage and

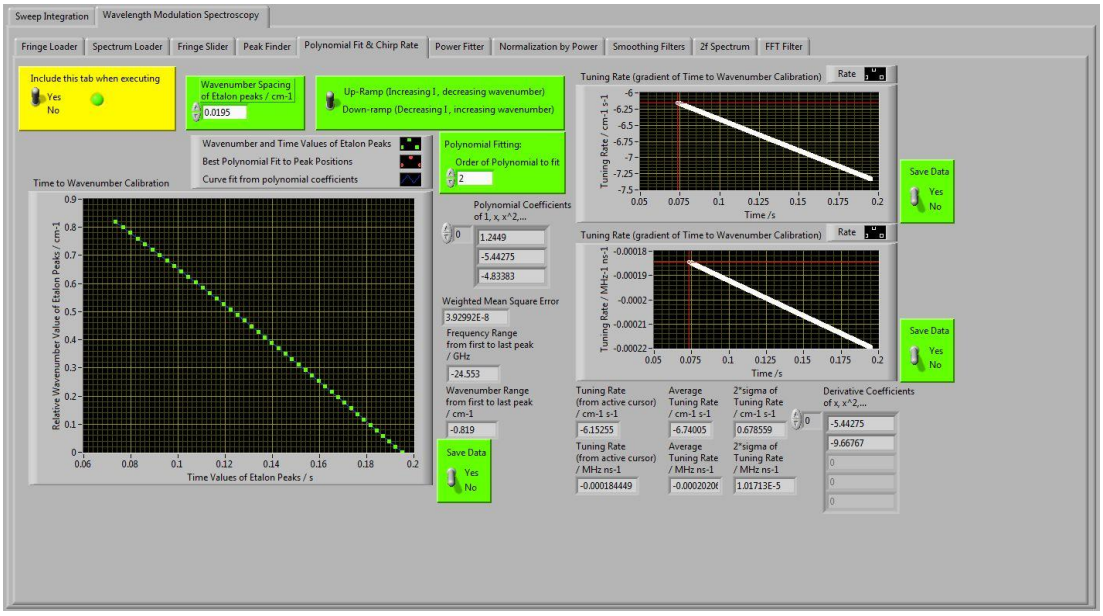


Figure 3-21: Peak Finder & Chirp Rate Tab. The user enters the etalon FSR here. The user chooses the order of the polynomial fit for the time to wavenumber conversion of the data. The derivative of this polynomial is displayed as the laser chirp rate in $cm^{-1}s^{-1}$ and $MHz^{-1}ns^{-1}$.

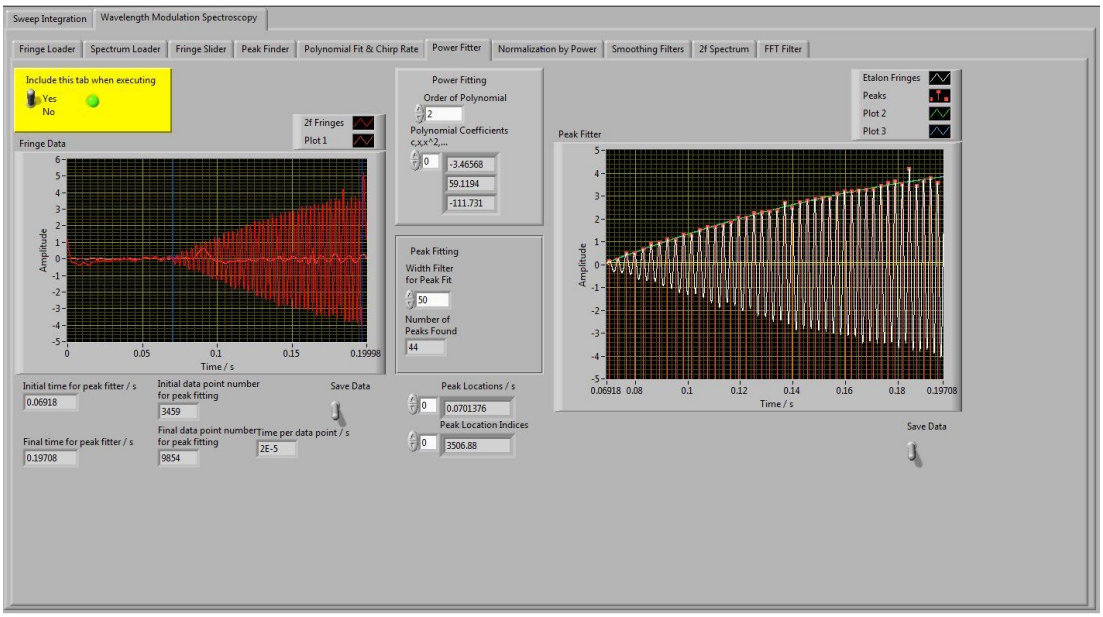


Figure 3-22: Power Fitter Tab. This tab aims to generate a polynomial through the peaks of the etalon fringes as an approximation to the changing laser power. This polynomial can then be used in the following tab (Figure 3-23) to 'power normalize' the 2f-spectrum – particularly for use if more than one spectral feature is present.

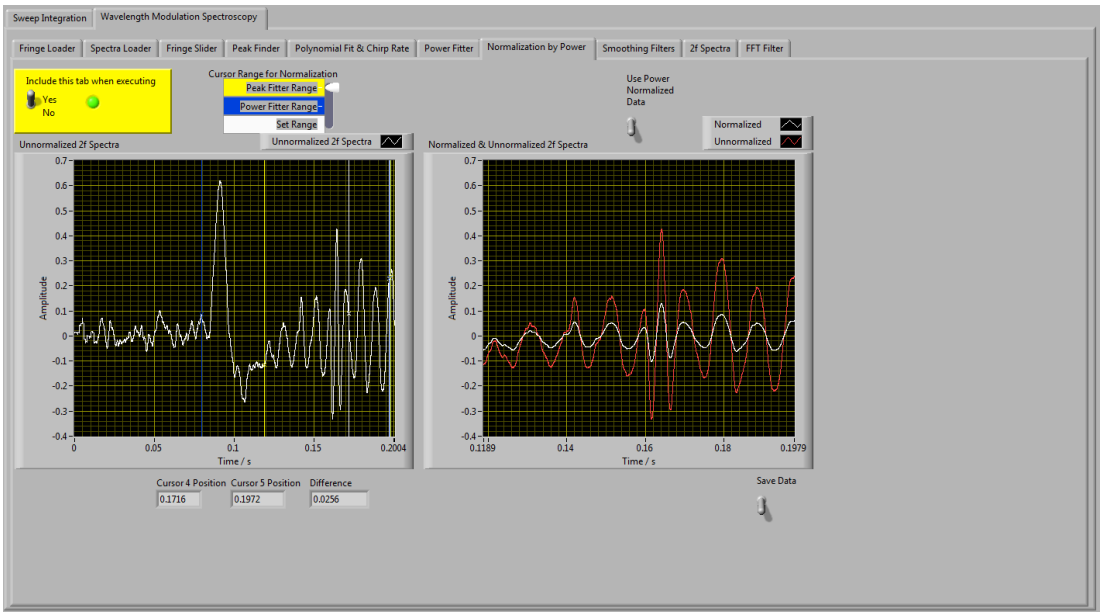


Figure 3-23: Normalization by Power Tab. This tab gives the option of using a spectrum which has been 'power normalized' by the polynomial found in the previous tab (Figure 3-22).

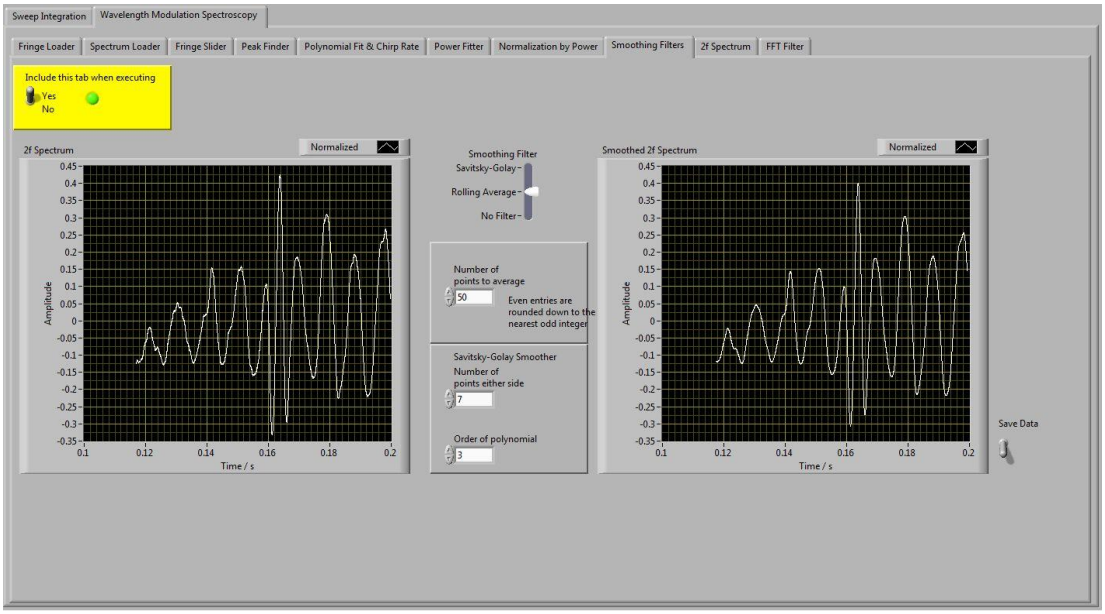


Figure 3-24: (Optional) Smoothing Filters Tab, with custom made rolling averager and a LabVIEW Savitsky-Golay filter.

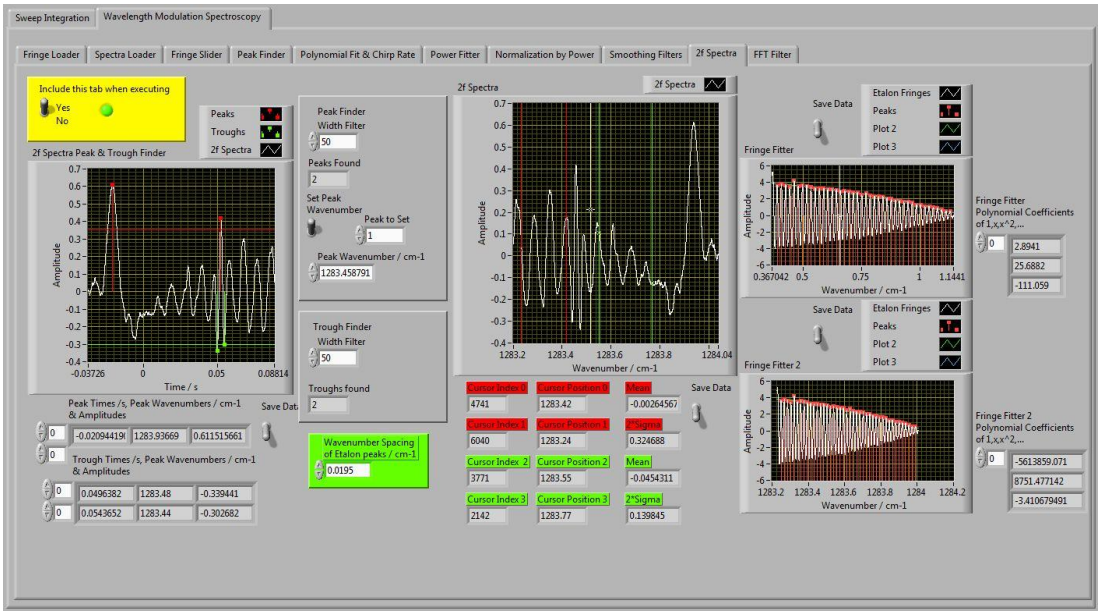


Figure 3-25: 2f-WMS Spectrum Tab. Here the peak and trough positions and amplitudes can be found via LabVIEW's peak fitter, the relative wavenumber scale can be converted to an absolute wavenumber scale, and the mean and 2σ of the noise can be taken.

The tab also allows the polynomial coefficients of the power fit to the etalon fringes to be found. The abscissa wavenumber values are slidable, allowing different parts of the power curve to be used as the baseline about line centre. This adds flexibility if the spectrum and fringes are taken from different experimental runs.

tuning rate data on the laser - and a means to convert the spectrum time axis to relative wavenumber.

3.4.3 Power Fitter and Normalization by Power Tabs

See Figure 3-22 and Figure 3-23. Had the cw-QCL laser used in Chapters 4 and 5 been available for longer, it would have been useful to process WMS spectra containing multiple absorption lines. For this, normalization to take account of changing laser power would be desirable. In Section 4.6.1.5 it will be seen that, since only an AC-coupled detector was available, different operating parameters were needed (in recording a given spectral feature) for sweep integration and WMS. This meant that the changing sweep integration baseline could not be used to normalize WMS features accurately.

To get around this, an approximation to a power normalization baseline was made available from a polynomial fit to the peaks of the 2f etalon fringe signal. (The normalization is approximate since the peak heights depend on the beam's angle of incidence on the etalon, and there is usually a (constant) phase difference between laser output power and wavenumber tuning).

3.4.4 Smoothing Filters Tab

See Figure 3-24. As in the sweep integration part of the program (Section 3.3.8), the rolling averager and a Savitsky-Golay filter are provided as an option.

3.4.5 2f Spectrum Tab

See Figure 3-25. Here, the processed WMS spectrum is displayed. The data is fed to a peak and trough finder provided by LabVIEW - each with its own threshold horizontal cursor and width control. In this way the amplitude and position of the 2f lineshape peak can be found and displayed, along with the two 2f troughs in the lineshape. The relative wavenumber axis can be shifted to absolute wavenumber by

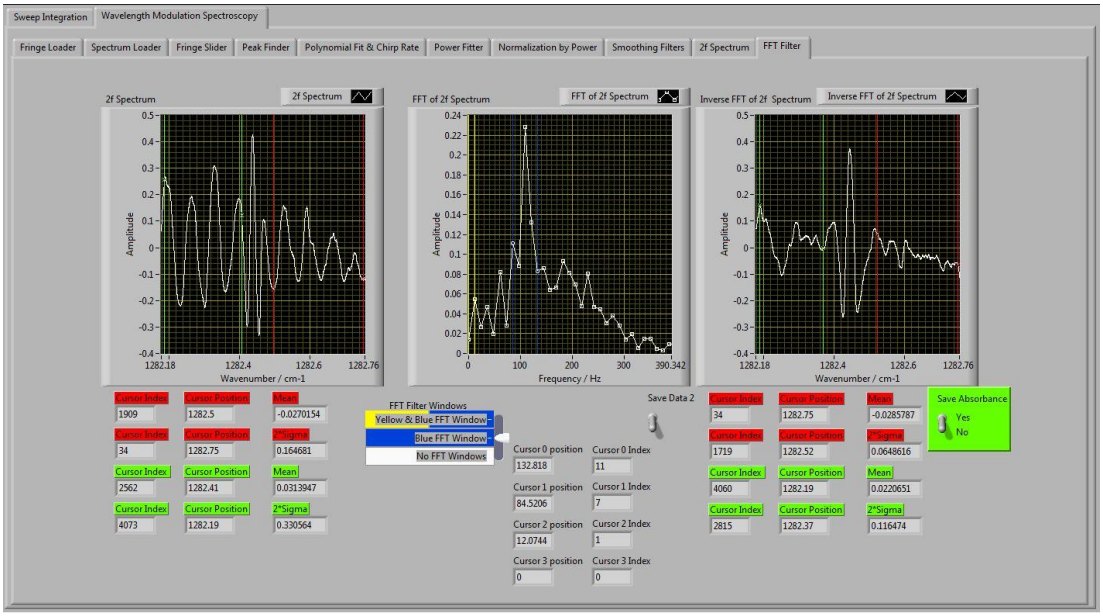


Figure 3-26: FFT Filter Tab. Here up to two frequency regions can be removed from the frequency content of the 2f-WMS spectrum. The figure illustrates a data processing suppression of unwanted etalon fringing.

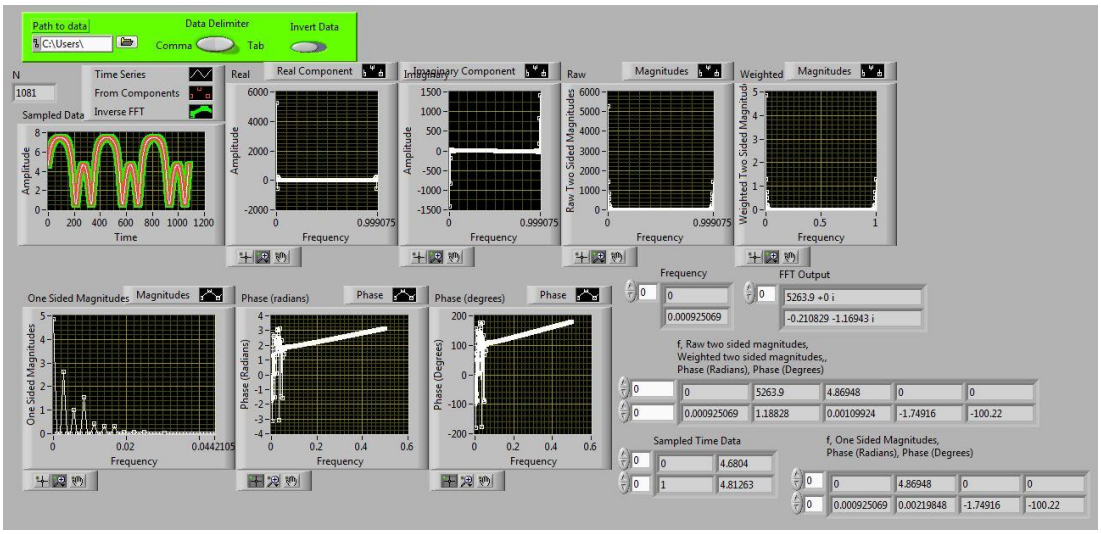


Figure 3-27: A stand alone Fourier analysis program based on LabVIEW's FFT routine – illustrated here with the 2nd harmonic WMS detector signal about an absorbance line centre. The two-sided complex output arising from the Fourier transformed input is converted to the more intuitive one-sided amplitudes of Fourier frequency components within the signal and their associated phases. Two checks are provided, by synthesizing and plotting the original signal from the one-sided components, and also using LabVIEW's inverse-FFT routine on the two-sided components.

entering the peak's wavenumber if known. Two sets of slidable cursors allow the baseline mean and 2σ to be read off.

In Section 5.9.1.4 the effect of a slowly changing direct signal baseline upon the 2f-WMS signal is modeled. For a quadratic baseline, the polynomial coefficients of the etalon fringe power fit can be used, and these can be obtained from this 2f Spectrum tab. Flexibility was provided to use different parts of the power curve as the baseline about the line centre, as it was sometimes necessary to use fringes and spectral data taken from different experimental runs. (The fit coefficients depend on the wavenumber values of the abscissa). A slidable cursor on the spectrum plot could be moved to the position of the line centre. This cursor's relative wavenumber would then shift the relative wavenumber scale on the fringe plot, by giving its wavenumber to another slidable cursor on the fringe plot. The fringe plot could then be shifted to absolute wavenumber to read off the correct best fit coefficients.

3.4.6 FFT Filter Tab

See Figure 3-26. It is often useful to look at the frequency content of a time signal, and this can be achieved by using LabVIEW's Fast Fourier Transform (FFT) – an efficient implementation of the Discrete Fourier Transform (DFT). Figure 3-26 shows a situation where unwanted Fabry-Perot etaloning in the optical path is of the same order of magnitude as the 2f-WMS signal. The long term fix for this is to tweak the optical alignment in an attempt to minimize the fringing. However, in the short term, the frequency of the etaloning can be isolated and removed in data processing - if the etalon frequency is sufficiently different from that of the main frequency content of the signal. Here, an FFT is applied to the spectrum, up to two frequency regions are removed between two pairs of slidable cursors, and the remaining frequencies are subject to an inverse-FFT to produce the original spectrum with less etaloning. The DFT normally accepts equally spaced time points, though in the cases of Chapter 5, the FFT → inverse-FFT process on the weakly quadratically spaced wavenumber points produces similar results with little distortion.

LabVIEW's help notes were not particularly helpful on using the two-sided complex output of its FFT routine. In exploring this, another program was written to convert the FFT output to the more intuitive one-sided real amplitude versus frequency format (Figure 3-27). Two useful sources of information on the DFT, of which the FFT is an efficient implementation, are Smith (1997) and Dorran (2012).

3.5 Conclusions

A data analysis program was successfully written in LabVIEW. This can take pre-recorded voltage vs time data, or data straight from a Tektronix oscilloscope, and process the data through to absorbance versus absolute wavenumber with a high level of efficient interactive automation. A quick look up table of the program's functionality is given in Table 3-1 at the end of this chapter.

3.6 Suggestions for further work

The automated Lorentzian and Gaussian fits could be further refined, calibrated and tested. This would need a pressure gauge of greater sensitivity than that available in Chapter 4 for an independent pressure measurement. Test gas concentrations would need to be controlled, perhaps via the Environics dilution system of Chapter 7.

The pseudo-Voigt fitter could be completed. The model itself has been constructed in the program presented in Section 2.6.3, and could be used for trial and error fitting. However, automation requires inputting this model into a LabVIEW 'static reference' - which in turn is fed into LabVIEW's Levenberg-Marquardt fitter. This is more complex than the Gaussian and Lorentzian cases – a greater choice of variable parameters are possible, and some of these need to be kept positive during fitting. This would add significant functionality - Cascade Technologies has hitherto been concerned only with fitting Lorentzians - appropriate to relatively high pressure (≥200Torr) mid-infrared spectra. The visible oxygen spectroscopy of Chapter 6

benefits from a Voigt profile since the component Gaussian and Lorentzian FWHM's are of the same order.

The gas temperature range of the models and fits could be extended. The work in this thesis was conducted at room temperature, so the default HITRAN linestrengths at 296K were adequate. To move beyond this, linestrengths at other temperatures would be calculated from Equation 2-32. The relevant data would be located and input from the relevant HITRAN molecular line list. Likewise, the needed partition functions are provided in the HITRAN file *parsum.data*.

At this stage the data analysis program would be close to a functioning spectral modeling and fitting program. The production of simulated spectra, and the fitting of multiple measured lines, could be accomplished by data extraction from the above databases and looping the program until a user defined wavenumber region had been covered.

Ultimately, it would be desirable to rewrite the parts of the program that use established LabVIEW routines – such as the peak finder, polynomial fitter, and Levenberg-Marquardt fitter. This would free the program from being platform specific.

Table 3-1: Functionality of the data analysis program.

Tab	**Description**
Sweep Integration	
Fringe Loader	Loads etalon fringe data from oscilloscope or text file, two interactive cursors take a mean of the data between them to shift ac data to positive values.
Background Loader	Loads background data from oscilloscope or text file, two interactive cursors take a mean of the data between them to shift ac data to positive values.
Spectrum Loader	Loads spectral data from oscilloscope or text file, two interactive cursors take a mean of the data between them to shift ac data to positive values.
Fringe Slider	Allows x and y adjustment of etalon fringe data to account for any jitter between the fringe and spectrum data sets.

Auto Peak Finder	Built around LabVIEW's in-built peak finder to associate a time value with each etalon fringe peak (equally spaced in wavenumber). Two interactive cursors set the time range to use, a horizontal active cursor and a width control filters out unwanted peaks.
Manual Peak Editor	Allows the peak points from the auto fitter to be deleted manually, and data points added manually as peaks.
Polynomial Fit & Chirp Rate	Built around LabVIEW's in-built polynomial fit to build a conversion relationship between time and relative wavenumber based on the wavenumber free spectral range of the etalon. The order of the time to wavenumber relationship's polynomial fit can be interactively changed and tested. The tuning rate in cm^{-1}/s and MHz/ns is automatically calculated and plotted.
Transmittance via Loaded Background	Provides an x-y shifter to line up a background data set with the spectra to account for any jitter between data sets. The resulting transmittance is automatically calculated and plotted. If fringe data is available the x-axis is converted from time to relative wavenumber in the range used for peak fitting or another range is set with active cursors.
Transmittance via Polynomial Fit	Instead of a background, a polynomial fit can be used as a background. Up to 3 sets of interactive cursors specify ranges for LabVIEW's in-built polynomial fit to extrapolate from. The order of the polynomial used is interactively set and the resulting transmittance is displayed interactively. If fringe data is available, the x-axis is converted from time to relative wavenumber in the range used for peak fitting or another range. Two more interactive cursors allow calculation of the mean and standard deviation of the data between them.
Smoothing Filters	Allows smoothing of the data (reduction of high frequency noise) via a custom written rolling average filter of interactive size, or via LabVIEW's in–built Savitsky-Golay filter (a widely used type of weighted rolling average).
Absorbance Peaks	The transmittance data is converted into Absorbance vs Relative Wavenumber (or time, if fringe information is not available). Then, LabVIEW's inbuilt peak fitter (based on parabolas) is used to provisionally find the centre of any absorbance lines. If the line centre is known (say, from the HITRAN database) the relative wavenumber axis can be shifted to an absolute wavenumber axis.
Absorbance Spectra	Displays the final Absorbance spectra (background and / or polynomial normalized). Two sets of interactive cursors allow the mean and standard deviation to be calculated for the data between them.
Concentration Fit	Fits a Gaussian or Lorentzian lineshape to the (room temperature) data built around LabVIEW's inbuilt Levenberg-Marquardt non-linear regression algorithm. This allows extraction of a concentration value. Data such as line centre and FWHM is calculated and displayed.

Wavelength Modulation Spectroscopy	
Fringe Loader	Loads etalon fringe data from oscilloscope or text file.
Spectrum Loader	Loads WMS data from oscilloscope or text file.
Fringe Slider	Allows x and y adjustment to etalon fringe data to account for any jitter between the data sets.
Peak Finder	Built around LabVIEW's in-built peak finder to associate a time value with each peak (equally spaced in wavenumber), two interactive cursors set the time range to use, a horizontal active cursor and a width control filters out unwanted peaks.
Polynomial Fit & Chirp Rate	Built around LabVIEW's in-built polynomial fit to build a conversion relationship between time and relative wavenumber based on the wavenumber free spectral range of the etalon. The order of polynomial to use can be simply entered and tested. The tuning rate in cm^{-1}/s and MHz/ns is automatically calculated and plotted.
Power Fitter	The amplitudes of the etalon fringes are used to interactively fit a polynomial via LabVIEW's in-built peak fitter. This approximates to a background to capture the change in laser power during a scan. The range to use is interactively set with active cursors.
Normalization by Power	Allows WMS spectra to be normalized by the power background from the previous tab. This allows two features in the same scan to be compared, for instance during temperature measurements.
Smoothing Filters	Allows smoothing of the data (reduction of high frequency noise) via a custom written rolling average filter of interactive size, or via LabVIEW's in–built Savitsky-Golay filter (a widely used type of weighted rolling average).
2f Spectrum	Displays the processed WMS spectrum. Two sets of interactive cursors allow the calculation of mean and standard deviation for the data between them. LabVIEW's in-built peak finder (based on parabolas) is used to return the relative wavenumber and / or time values of the peaks. If the line centre is known (say, from the HITRAN database) the relative wavenumber axis can be shifted to an absolute wavenumber axis. Polynomial coefficients of the etalon power fit are also returned, These are used in separate modeling of the affect of a slowly changing direct signal baseline on the 2f-WMS signal (Section 5.9.1.4).
FFT Filter	Built around LabVIEW's in-built FFT algorithm, allowing the display of the spectrum's power spectrum. Up to two sets of interactive cursors are used to remove ranges of unwanted frequencies. The inverse transform is then performed to re-display the noise reduced spectrum.

Table 3-1: Functionality of the data analysis program.

Chapter 4

Sweep Integration with a continuous wave quantum cascade laser

4 Sweep Integration with a continuous wave quantum cascade laser

4.1 Abstract

Sweep Integration was performed on $^{12}CH_4$ gas with a cw-QCL, of nominal mid-infrared wavelength 7.84μm, and the signal detected with an AC-coupled HgCdTe photovoltaic detector. It was necessary to sweep at a rate of, at least, 10kHz to avoid time constant distortions from the built-in AC-coupled transimpedance amplifier. The limit of detection was 0.080 absorbance units when limited to the oscilloscope's averaging of 512 averages, corresponding to ≈51ms integration time. Although comparable to Cascade Technologies' pulsed-QCL spectrometers at an earlier stage in their development, this result is someway behind the current routine performance of the CT3000 system of Chapter 7 - of 0.004 absorbance units in ≈10ms. Increasing the effective number of averages to 7680, ≈0.77s, brings the cw-QCL spectrometer down to 0.020 absorbance units. Further improvement seems possible by increasing the number of averages still further, and increasing the power falling upon the detector. The slow chirp rates attainable with a cw-QCL mean that the spectrometer's resolution is not limited by chirp rate, but by the Doppler width of the transitions themselves - further broadened by a symmetric Gaussian broadening. Through the elimination of other possibilities, this further broadening was attributed to the limitations of the system's electronics - particularly wavenumber fluctuations and wavenumber jitter.

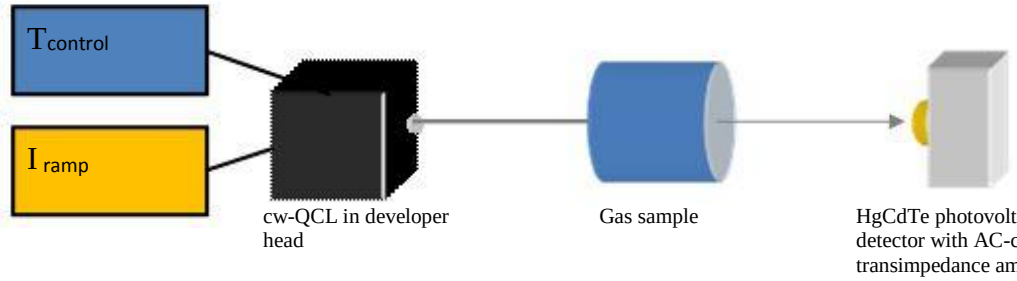
4.2 Introduction

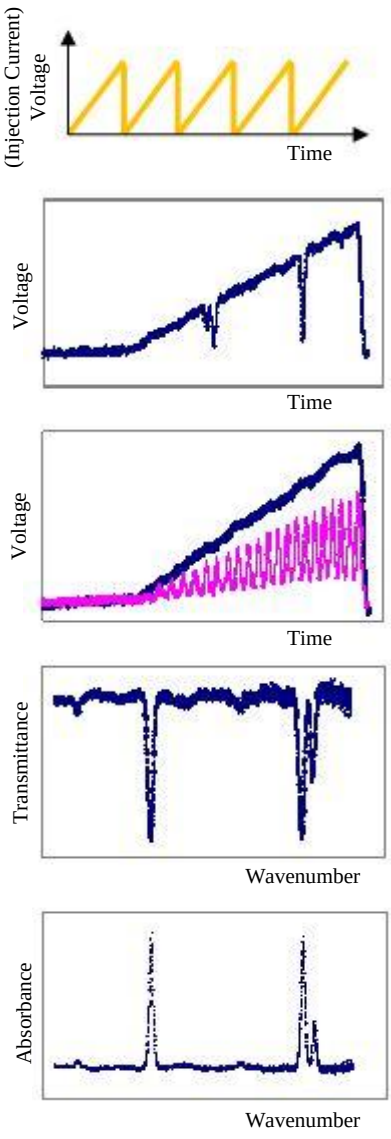
One of Cascade Technologies' directors had had prior experience applying the noise reduction technique of Wavelength Modulation Spectroscopy (WMS) (Chapter 5), using a near-infrared semiconductor diode laser (Howieson, 1997). The company had one mid-infrared cw-QCL available for the purpose of ascertaining its potential for increasing the sensitivity of gas measurement by WMS.

The company's current technology is based around mid-infrared pulsed-QCL's, and so the company's mid-infrared detectors are AC-coupled – limiting 1/f noise. There is a commercial pressure to use the simplest techniques possible for a given task, with the minimum change of equipment. This keeps costs down and aids integration with existing systems. So, although it would be natural to use a DC-coupled detector with a cw laser, none was available. Using an AC-coupled detector posed some problems relating to the distortion of low frequency signals, the need to include the below threshold noise level in a scan – as an approximation for V_0 in Beer-Lambert's Law, and the transimpedance amplifier's compatibility with other electrical equipment. These issues are discussed further in Section 4.6.1.

The cw-QCL was operated from a development head designed for pulsed-QCL's, so the laser's power and tuning response was first characterised, to enable specific absorption lines to be identified and swept over. Despite literature showing cw-QCL's being modulated well into the MHz range, resistance to the line of research was encountered. Accordingly, the cw-QCL's response to applied voltages of various forms and frequencies was investigated, within the limitations of the available electronics.

Prior to performing Wavelength Modulation Spectroscopy (WMS) (Chapter 5), Sweep Integration was performed. This allowed the voltages associated with the WMS technique to be calibrated to absorbance units (and hence, if required, to number densities and concentrations). The noise reduction offered by Sweep Integration is also of interest in itself. Sweep Integration (Jennings, 1980) has improved in its noise reducing capability, relative to WMS, as computer processing and fast digitizers and detectors have evolved.





Tuning

The starting wavenumber is set by temperature control of the laser's substrate. A current ramp repetitively tunes over an absorption feature. The wavelength change is induced by Joule heating of the laser's active layer, causing an optical path length change, mainly through a change of refractive index.

Averaging raw spectra

The average of the N spectra is taken to improve the SNR proportional to $\sqrt{N}$, producing a single averaged spectrum. This noise reduction is possible up to a 'maximum integration time', after which longer term drifts begin to dominate the random noise.

Preparation for normalization

A background is taken, or a polynomial fit to the averaged spectrum's baseline is prepared. This is used as a divisor to eliminate distortions caused by the laser's power changing with tuning. The equal frequency spacing of etalon fringes is used to convert the time axis to relative wavenumber. This eliminates distortions of spectral shape caused by the changing wavenumber tuning rate. (This constancy in the wavenumber spacing is estimated in Appendix C to be as good as 3 parts per million).

Transmittance Spectrum

After normalization, a transmittance spectrum results with a baseline of 1. The transmittance profiles represent the fractional change in intensity caused by absorbers, as required for use in the Beer-Lambert law.

Absorbance Spectrum

The absorbance spectrum is obtained from the transmittance spectrum via the Beer-Lambert law. Since the number density and concentration of the gas is proportional to absorbance, these parameters can be recovered – in the general case from the areas under the absorbance features. Considering the wavenumber coverage of the scan, gas species can be identified, and the relative wavenumber axis converted to an absolute wavenumber scale.

Figure 4-1: Schematic diagram of the Sweep Integration technique as used in this chapter. The SNR is proportional to the square root of the number of scans - until longer term drifts begin to dominate over random noise.

4.3 The Sweep Integration Technique

Sweep Integration is a noise reduction technique, performed by repeatedly tuning the laser over an absorption line. The resulting spectra are added and averaged. Since the signal of interest is expected to add each time, the signal grows in proportion to the number of scans N. The absorption feature is then recovered after averaging by the number of scans N. Due to the fluctuating nature of random noise, when the spectra are averaged the random noise is expected to only grow in proportion to the square root of the number of scans $\sqrt{N}$. Thus, when averaged by N, the noise is reduced in proportion to $1/\sqrt{N}$, and the SNR increases by $\sqrt{N}$. This $\sqrt{N}$ proportionality can also be considered in terms of the square root of the 'integration time' τ, since the time to perform the technique increases in proportion to the number of averages.

Mathematically, the $\sqrt{N}$ result arises from considering the true signal's N corresponding noise readings for any particular time, *t*, (and wavenumber, $\bar{\nu}$) across N sweeps. The random noise values are assumed to be normally distributed and 'the standard deviation of the mean' of the readings is calculated at the time *t*. (See for example Taylor, 1997).

The sweep integration technique does not deal with systematic noise, such as etalon fringing from un-intentionally parallel surfaces in the optical path – since systematic noise grows on addition just like the signal. Differences between scans, such as un-corrected timing jitter - affecting the position of the signal - will also degrade the averaged spectra. Also, after a 'maximum integration time', longer term drifts in the system, say from laser power fluctuations or movement of the etalon fringes, will set a limit on the maximum number of averages which are noise reducing. Figure 4-1 shows a schematic diagram of the essentials of the sweep integration technique, with some outline notes on the data processing.

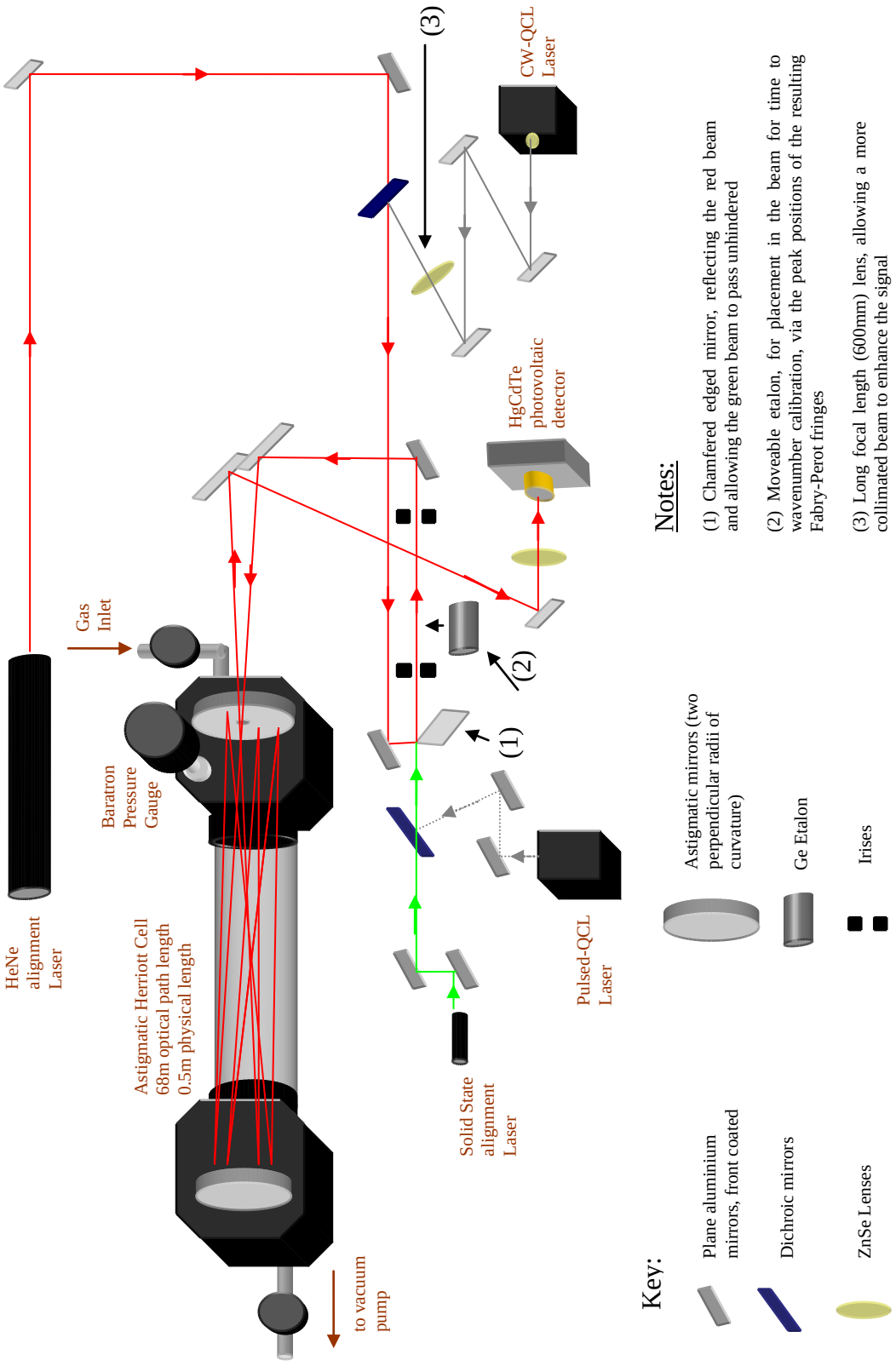


Figure 4-2: Optical Set-up for Direct (Sweep Integration) and Wavelength Modulation Spectroscopies with the cw-QCL.

4.4 Experimental Setup

4.4.1 Optical Setup

The optical part of a previous pulsed-QCL laser spectrometer was brought back to working order (Wright et al., 2006, Hay et al., 2008, Hay, 2010). The mirror alignment of its astigmatic Herriott cell was tweaked to reproduce the Lissajou spot patterns which had given a path length of 68m. This pathlength was confirmed via the delay time of a pulse from the pulsed-QCL (part of, and driven by, a Cascade Technologies LS-03-D quantum cascade laser driver system). The cw-QCL's radiation was coupled into the system as shown in Figure 4-2. Optical mounts along the right page margin of the diagram prevented a simpler optical path. The steering mirror losses were heavy, so that about 3mW of power entered the cell, and the signal on the detector was in the μW region. The VIGO detector–built-in transimpedance amplifier-combination was easily saturated (when producing a voltage above 1V). This low power was therefore not regarded as a particular problem – though the signal could be further optimized in future by using a different optical arrangement.

4.4.2 Electrical Setup

The lowest noise combination out of the equipment and cabling available was looked for by trial and error. The final electrical setup is summarized by Figure 4-3. The RT-DFB-CW-QCL, of nominal wavelength 7.8μm, came from one of Cascade Technologies' suppliers and was mounted in a Cascade Technologies' LM-03-D laser development head. The laser's substrate temperature was operated about room temperature by means of a thermoelectric (2-stage Peltier) cooler, controlled by an ILX Lightwave LDT-5980 temperature controller, and assisted by water cooling. Injection current control was via an ILX Lightwave LDX-3232 current controller.

This current controller set the 'base current' onto which the sawtooth (or 'serrodyne') ramp signal was added through the controller's analogue BNC input, for

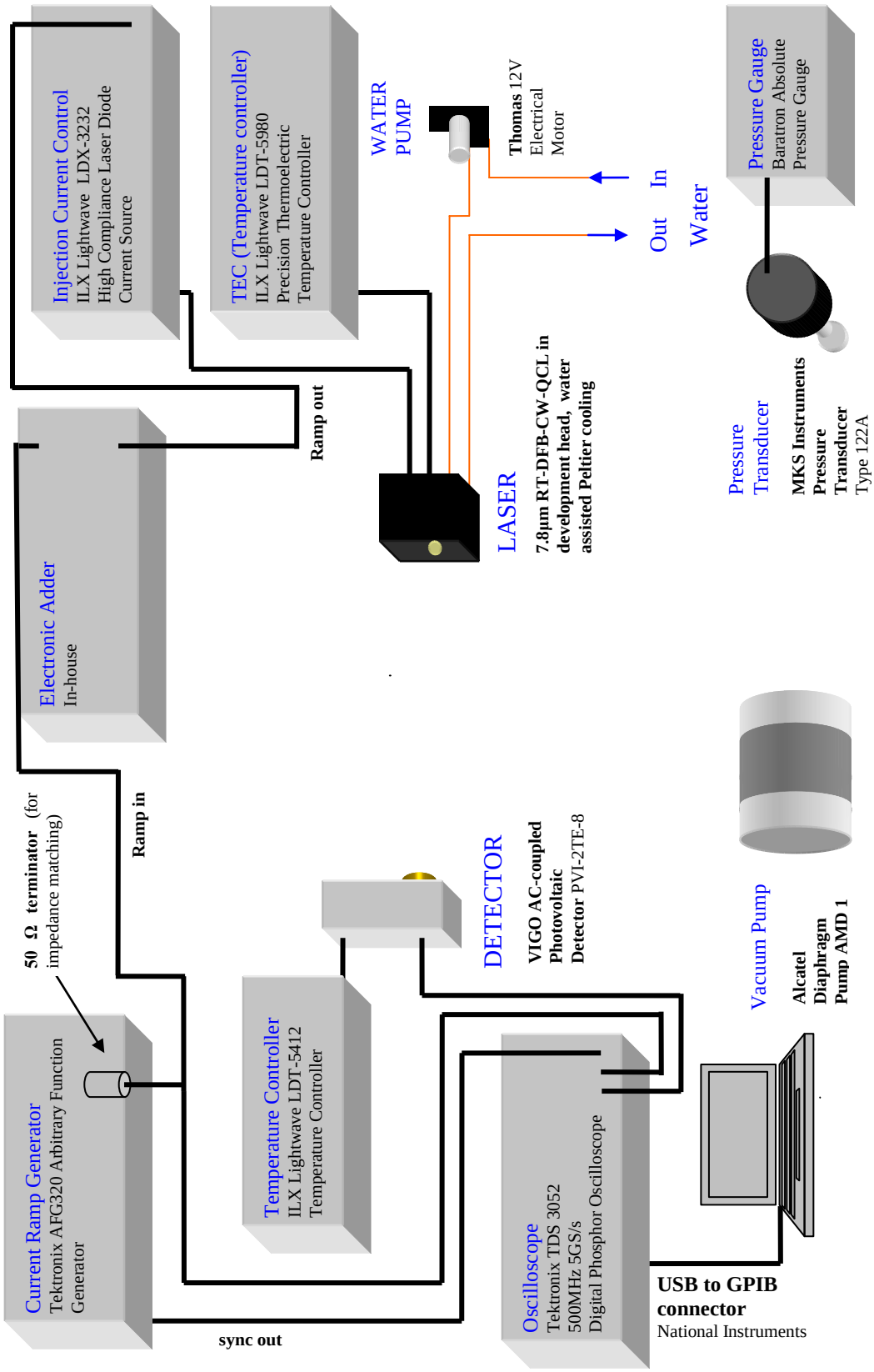


Figure 4-3: Electrical Set-up for Direct (Sweep Integration) Spectroscopy with the cw-QCL.

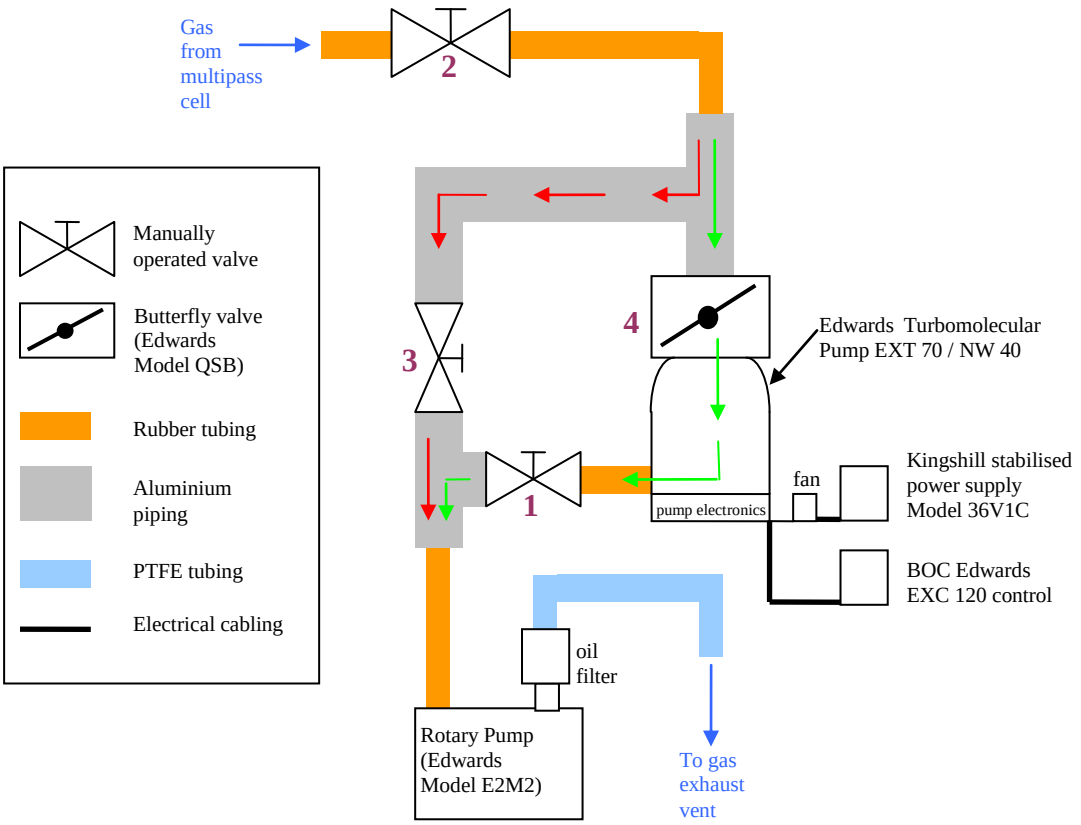
sweep integration. The ramp signal was generated by a Tektronix AFG320 arbitrary function generator. On route to the current controller, the ramp signal was fed as the only input through an in-house electronic adder. This made the electrical setup as close as possible to that used for a comparison with the Wavelength Modulation Spectroscopy technique (Chapter 5). As a precaution, a LabVIEW program was written to slowly sweep the LDX-3232 current controller digitally. However, digital ramps were hardware limited to $\Delta t = 20ms$ and $\Delta I = 0.1mA$, which was not practical for the experiments. The digital and analogue ramps were tested on an LED prior to connecting up the laser.

The optically immersed photovoltaic detector was a VIGO System's PVI-2TE-8 HgCdTe (MCT) photodiode. These types of detectors are supplied to Cascade Technologies in TO packages with BaF_2 windows, and with impedance matched AC-coupled transimpedance amplifier already attached. Unlike the 'text book' logarithmic response, these detectors produce a linear response of output voltage verses received laser power over much of their range.

The sweep integration signals were analyzed on a Tektronix TDS 3052 500MHz 5GS/s digital oscilloscope. This set a digital resolution of 10,000 data points on the time axis and a maximum averaging of 512 sweeps.

4.4.3 Pumping System

Through much of the experimentation, coarse control over the Herriott cell pressure, and hence gas number density, was provided by a diaphragm pump (Alcatel AMD 1). Later, a turbomolecular pumping system (Edwards EXT 70 / NW 40) backed by a rotary pump (Edwards Model E2M2) became available. This allowed quicker and more sensitive control of gas number density. Unfortunately, the Pressure Transducer available was not sensitive enough for the vacuum reached. This prevented an independent measurement, which would have been useful for checking



Start Up Procedure (from pumps off and valves 1,2,3 & 4 all closed)

1) Turn rotary pump on
2) Power on turbo's fan and electronics
3) Open valve 1 to rough out turbo
4) Start turbo and wait for it to reach operational rpm
5) Close valve 1
6) Open valves 2 & 3 to rough out the gas cell and pipe work (down to about 1 Torr) [Red arrow route]
7) Close valve 3
8) Open valve 1
9) Open Butterfly valve 4 to turbo [Green arrow route]

Shut Down Procedure

1) Close butterfly valve 4 to turbo
2) Turn turbo's electronic control box off
3) Wait for turbo to spin down before turning turbo itself off at its back switch
4) Turn off fan and its own power supply
5) Close valve 1
6) Close cell valve 2
7) Turn rotary pump off

Figure 4-4: The turbomolecular pumping system which replaced simple diaphragm pumping.

the fitting of the data analysis program (Chapter 3). The pumping system setup and some operating notes are given in Figure 4-4.

4.5 Preliminaries: Static Tuning

Before applying current sweeps to the laser, the laser's threshold and power response were characterised statically (point by point) as a function of injection current and laser substrate temperature. Then, similarly, the laser's wavenumber region and coverage was characterised by Fourier Transform Spectroscopy (FTS). This aided the later adjustment of the sweep integration parameters to cover specific absorption lines.

4.5.1 Operating Parameters

The laser's specification sheet indicated that this cw-QCL had been operated from -30º → +50ºC, and from approximately 260mA → 750mA. The QCL's nominal wavelength was 7.8μm (≈1282cm^{-1}), and coverage 1279.2cm^{-1} (7.817μm) → 1290.4cm^{-1} (7.750μm).

The development head was designed for pulsed-QCL use. So, to facilitate the different electrical requirements of the cw-QCL, wiring had been connected by Cascade Technologies through the housing. Accordingly, the head could not be evacuated, and the lowest temperature used was +8ºC – comfortably above the dew point of water to prevent condensation damage. Initially, during the FTS characterization, the Peltier cooler was backed by forced air cooling - but this proved inadequate. As a precaution against overheating, the highest temperature used was +32ºC. And this range proved adequate after a water cooler had been obtained.

The threshold current at +8ºC was ≈400mA, and although the current went to 750mA for FTS and power characterization, it was limited to 700mA during wavenumber

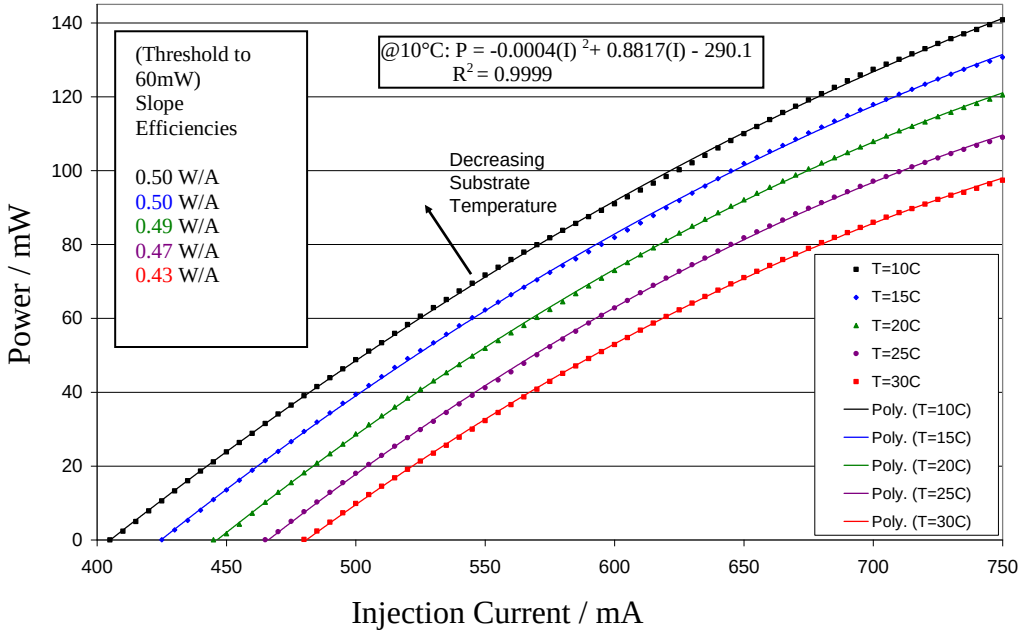


Figure 4-5: **Measured output power of cw-QCL / power meter combination as a function of injection current for fixed laser substrate temperatures. Quadratic fits are shown, with the fit equation at 10°C given explicitly, and linear slope efficiencies up to 60mW.**

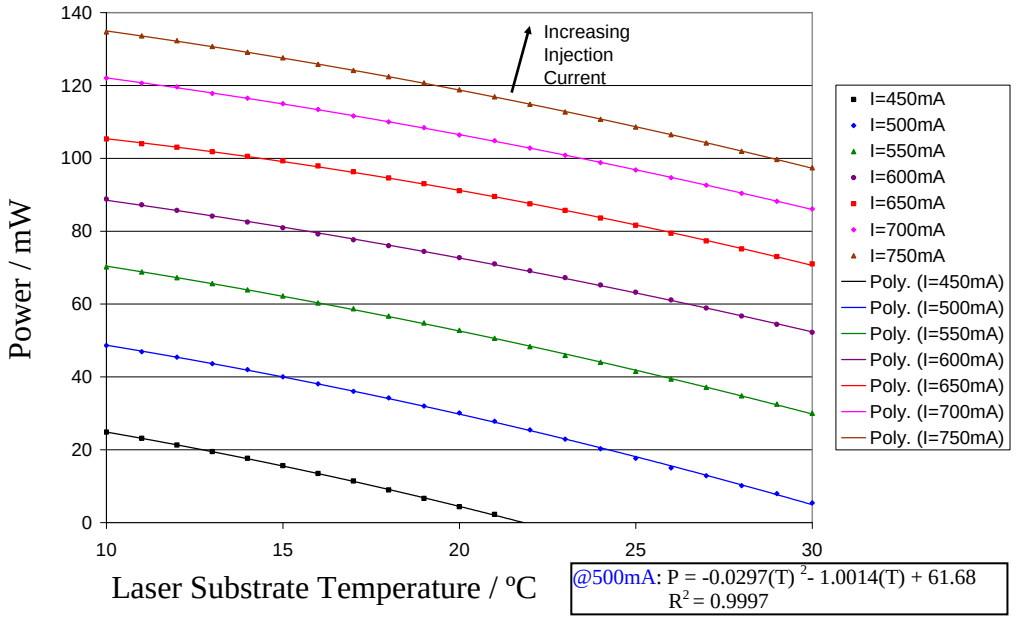


Figure 4-6: **Measured output power of cw-QCL / power meter combination as a function of laser substrate temperature for fixed injection currents. Quadratic fits are shown, with the fit equation at 500mA explicitly given.**

sweeping - as a precaution against overheating and of exceeding the laser's current limit.

4.5.2 Output Power Characterisation

The cw-QCL's threshold and power output were investigated with a power meter (Ophir ORION-TH-RoHS) and thermopile sensor (OPHIR-3A-SH-RoHS). The results for this cw-QCL / power sensor combination are displayed in Figure 4-5 and Figure 4-6.

Figure 4-5 shows the measured power against injection current for fixed substrate temperatures. The lasing threshold is seen to increase with increasing substrate temperature, from about 400mA (10ºC) to about 480mA (30ºC), as non-radiative relaxation processes compete more efficiently with lasing. Between laser threshold and about 60mW the power output is fairly linear, with slope efficiencies from 0.50W/A (10ºC) to 0.43W/A (30ºC).

Beyond 60mW the output power is a quadratic function of the injection current. In the later sweep integration experiments (Section 4.6.2), the swept current response showed a linearly increasing baseline (when sweeping fast enough to avoid AC-coupling time constant effects). However, due to the lag in QCL response to injection current when swept, it is not certain that the output power was reaching the above 60mW level (at the laser end of the experiment) – even though the current was swept to 700mA.

There is a slight power dip in the tuning characteristic at 10ºC between 600mA and 700mA. This was a reproducible feature of this laser / power sensor combination's behaviour.

Figure 4-6 offers more temperature data and confirmed that hysteresis was not a problem with this laser.

4.5.3 Fourier Transform Spectroscopy

4.5.3.1 Technique

Fourier Transform Spectroscopy (FTS) generally uses a broadband thermal infrared source, such as a Nernst Glowbar to illuminate a LN_2 cooled bolometric detector. All source frequencies impact the detector simultaneously. Pre-detector, the source radiation is collimated, split and recombined in an interferometric arrangement – often a Michelson arrangement. One interferometer mirror remains fixed, while the other is movable. At zero optical path difference there is a large centre burst of radiation corresponding to zero phase difference. As the mirror moves, an optical path difference opens up, and an interferogram is recorded by the detector as a function of time – its pattern depending on the phase differences between the component source frequencies. These source frequencies can then be obtained by a Fast Fourier Transform (FFT) of the interferogram into the frequency domain. When an absorber is present, the relevant frequencies are reduced in amplitude and are marked by absorption dips in the frequency spectrum.

The resolution of the system is a function of the distance that the movable mirror scans. Any truncation loses information. In the time domain, the closer two frequencies are together the slower is their beat frequency. It therefore takes more time, that is scan distance, for their interference pattern to be readily distinguishable. In the frequency domain, truncation effectively broadens each frequency component by convolving it with the truncation function, the broadening being inversely proportional to the scan length, making each component less distinguishable after the FFT.

Resolution also needs maintaining by decreasing an iris in the beampath as the scan length is increased. The detector pattern is extended – a concentric set of rings in the Michelson case – and there would be no overall change in intensity if the pattern were not restricted, usually to the central fringe. The diameter of the whole pattern decreases with the optical path difference of the interferometer's mirrors. The

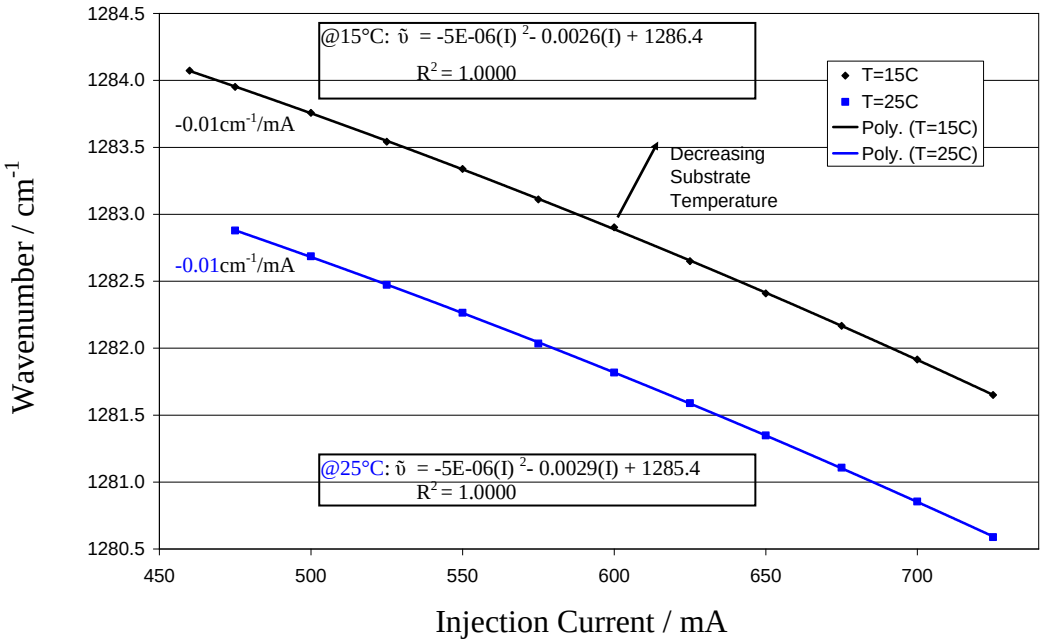


Figure 4-7: **FTS Temperature Tuning Data (via injection current): cw-QCL wavenumber versus injector current for fixed substrate temperatures. Quadratic fits and equations are shown, as are tuning rates.**

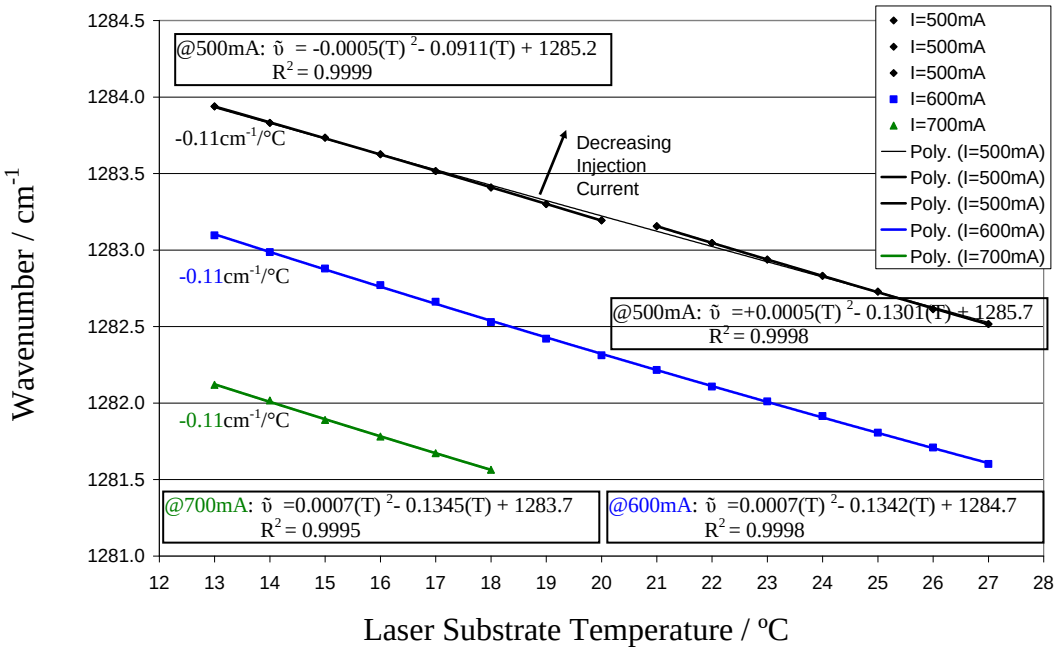


Figure 4-8: **FTS Temperature Tuning Data (via laser substrate temperature): cw-QCL wavenumber coverage versus laser substrate temperature for fixed injection currents. Quadratic fits and equations are shown, as are tuning rates.**

restriction is necessary to limit the solid angle to approximately where the divergence between the central and extreme rays become π out of phase, marking the edge of the central fringe. A detailed account of FTS is given by Griffiths and de Haseth (1986).

The need to scan the mirror can make Fourier Transform spectrometers bulky and make scan times long – sometimes hours as opposed to seconds for laser spectroscopy. However, the wavenumber coverage can be hundreds of times broader. The need for LN_2 cooling of the detectors interferes with long term unattended operation.

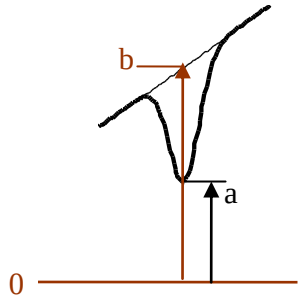
4.5.3.2 Experimental

The cw-QCL radiation was coupled into an existing setup based around an ABB-Bomem DA3.002 FTS, with resolution $0.025cm^{-1}$. Figure 4-7 shows the point by point tuning of wavenumber against injection current for two laser substrate temperatures. Figure 4-8 shows wavenumber versus laser substrate temperature for three injection currents.

The data is relatively sparse for two reasons. Firstly, the FTS could be temperamental in producing useful frequency spikes from the interferogram, presumably due to sensitivity to alignment. Secondly, before obtaining a water cooler for the laser, forced air cooling was inadequate and led to long settling times even with optimized PID settings on the controllers.

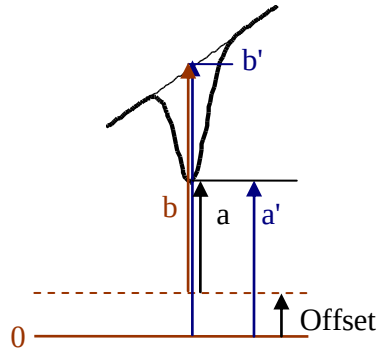
Nevertheless, enough useful data was obtained to allow tuning over specific $^{12}CH_4$ absorption lines. This retrospectively confirmed that the FTS was sufficiently well calibrated.

Figure 4-7 and Figure 4-8 show that the wavenumber tuning is weakly (small curvature) quadratic with injection current and laser substrate temperature, and show the associated wavenumber downchirp with increased heating. The average tuning



Usual point by point calculation of DC transmission spectrum, 0V and oscilloscope's zero coincide

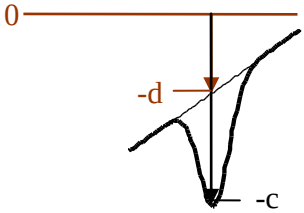
$$\frac{a}{b}$$

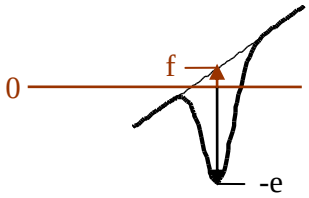


DC spectrum with offset of the oscilloscope's zero from true 0V

$$\frac{a'}{b'} = \frac{a + offset}{b + offset} \neq \frac{a}{b}$$

$$\frac{a'}{b'} \to 1 \text{ as } offset \to \infty$$





AC signal: height of the *same* feature in the calculated transmission spectrum becomes an artificial function of where it's scanned in relation to oscilloscope's zero

e.g. $\frac{-c}{-d} \neq \frac{-e}{f}$

Figure 4-9: Illustration that the transmittance spectrum becomes an arbitrary function of its oscilloscope position, without knowledge of 0V and hence I_0. (Also illustrates the apparent asymmetry of a perfectly symmetric Gaussian shape due to a linear power ramp).

rates (unswept) of –(0.0100 ± 0.0001)cm^{-1}/mA and –(0.1100 ± 0.0003)cm^{-1}/°C were determined from linear fits (gradient from best fit ± propagation of manufacturer specified errors in wavenumber, current or temperature). The results show that the QCL can potentially tune over $\Delta\tilde{\nu} \approx$-2.4cm^{-1} for $\Delta I \approx$+260mA and $\Delta\tilde{\nu} \approx$-1.5cm^{-1} for $\Delta T \approx$+14ºC. Overall coverage of the data collected here is ≈1280.5cm^{-1} to ≈1284.0cm^{-1}.

In the case of tuning via injection current, Section 4.6.2 will show that the heating (tuning) lags the injection current when swept. The potential tuning range becomes very much less and decreases as the current sweep frequency increases. For example, at a 10kHz sweep rate, wavenumber coverage was typically $\Delta\tilde{\nu} \approx 0.35 - 0.40$cm^{-1}. Figure 4-8 shows a probable mode hop at 500mA between 20°C and 21°C – the two sets of 500mA data breaking from an overall quadratic fit.

4.6 Sweep Integration

4.6.1 Issues associated with AC-coupled detection

4.6.1.1 Need for I_0 calibration

In order to accurately calculate a transmittance spectrum, the baseline of the signal needs to be known. The magnitude of the baseline signal above 0V, gives $V_0=V_0(t)$ which is proportional to $I_0=I_0(t)$ in the Beer-Lambert Law. I_0 is the intensity of the radiation in the absence of absorbers, from which the fraction of radiation absorbed can be calculated.

When used with Cascade Technologies' usual pulsed-QCL's, the AC-coupled VIGO detectors had posed no problems in calculating transmittance spectra. The trapezoidal signals of duration 500-1500ns and ≈12ns rise time, are fast enough to pass the capacitive part of the transimpedance amplifier with negligible distortion. The noise level at ≈0V, and the signal level are both clearly seen.

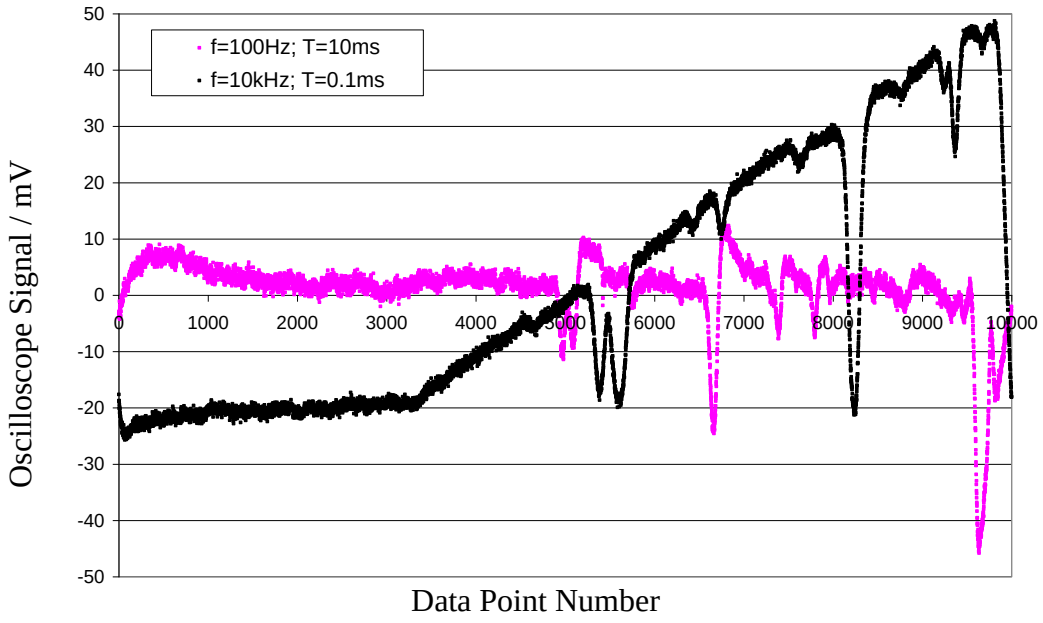


Figure 4-10: Illustration of the transimpedance amplifier's 180µs time constant effect at 100Hz and 10kHz. The difference in time scales represented is ×100. I_{base}=420mA, $I_{base\text{-}pk}$=280mA, $T_{substrate}$=10°C, 512 averages. (See Section 4.6.3 for notes on attempted line identifications).

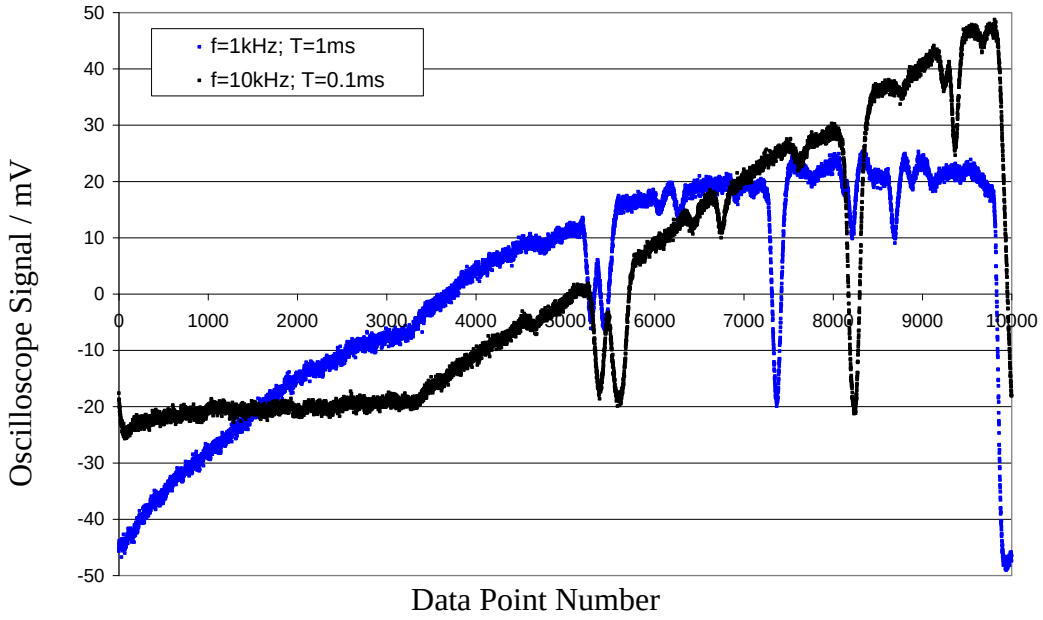


Figure 4-11: Illustration of the transimpedance amplifier's 180µs time constant effect at 1kHz and 10kHz. The difference in time scales represented is ×10. I_{base}=420mA, $I_{base\text{-}pk}$=280mA $T_{substrate}$=10°C, 512 averages.

However, for the continuous wave QCL and the relatively slow changing sawtooth ramps used here, the DC component is removed by the transimpedance amplifier's capacitance. This leaves the remaining signal averaging about the oscilloscope's zero level. To avoid divide by zero errors in generating a transmittance spectrum, the signal must be offset. But without knowing the 0V level, the calculated transmittance becomes an arbitrary function of the offset chosen – tending to zero transmittance as the offset tends to infinity. Figure 4-9 illustrates these points.

To circumvent this problem, the sawtooth scans were started below laser threshold. The laser's sub-threshold noise level, including spontaneous emission, was then used as an approximation to 0V. That this was a reasonable approximation is shown in Section 4.6.1.3, from the sub-threshold level of a rectified signal effectively delineating 0V.

The downsides of including the sub-threshold level in a sweep, is a restricted scan range, more contrivance of sweep parameters to arrange lasing close to the spectral feature of interest, and a degradation of SNR with less laser power over the spectral feature.

4.6.1.2 Transimpedance amplifier time constant effects

Early sweeps at low frequency showed a curving of the baseline, reminiscent of detector saturation. However, further investigation showed this to be a consequence of the high pass response of the AC-coupled transimpedance amplifier.

The capacitive part of the transimpedance amplifier will affect the signal whenever the signal changes slowly enough for an opposing capacitive voltage to build up. Different sweep rates were tried, and 10kHz was the slowest rate to show a linear, as opposed to curved, response. Figure 4-10 and Figure 4-11 show a 10kHz comparison against sweep rates of 100Hz and 1kHz respectively. An exponential fit to the below threshold curve at 1kHz (and 500Hz) revealed a transimpedance amplifier time constant of 180µs. (The detector element's rise time was specified as 8ns). Using the

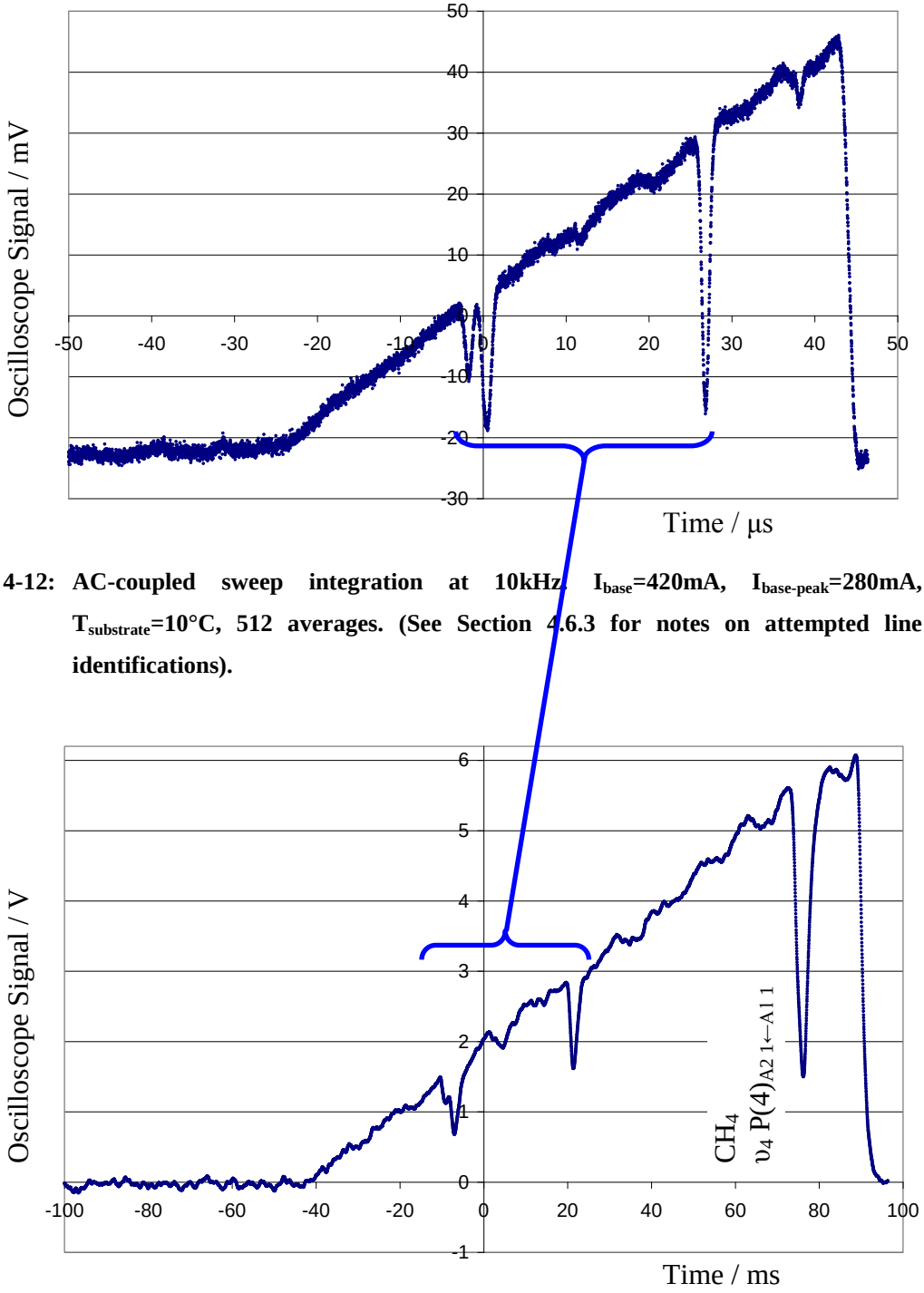


Figure 4-12: **AC-coupled sweep integration at 10kHz. I_{base}=420mA, $I_{base\text{-}peak}$=280mA, $T_{substrate}$=10°C, 512 averages. (See Section 4.6.3 for notes on attempted line identifications).**

Figure 4-13: **Chopped and rectified sweep integration: mechanical chopper at 2kHz, ramp at 5Hz, lock-in time constant =1ms. I_{base}=420mA, $I_{base\text{-}peak}$=280mA, $T_{substrate}$=10°C, 512 averages.**

first order low pass filter approximation, $f_c = 1/2\pi RC$, 180μs gives f_c=0.88kHz. This is not too dissimilar from the nominal 1kHz advertised cut off.

4.6.1.3 Exploring an alternative to simple AC-coupled detection: chopping and rectifying

To avoid the need for a below laser threshold portion in the scan, the possibility of converting the signal to DC was tried. Some loss of simplicity was entailed by adding a mechanical chopper and lock-in amplifier to the setup. (Bentham Instruments 218 variable frequency chopper with Type 218H mechanical chopping head, and a Princeton Model 5209 lock-in amplifier).

Figure 4-12 and Figure 4-13 show a comparison between the AC-coupled sweep integration at 10kHz and the chopped and rectified signal. This latter was performed with a chopper frequency of 2kHz and a 5Hz ramp (near the limit of a stable oscilloscope trace). The chopped signal is seen to have much poorer resolution – both in distinguishing the doublet and in recording the heights of the features. At first the reason for this was not clear. Although the slower ramp covers a greater spectral range, there are still enough data points to have distinguished the features.

The difference in time scales was negated by converting to absorbance versus relative wavenumber, and lining up the beginning of each main feature (Figure 4-14 and Figure 4-15). This made clear that the distortion was again due to the transimpedance amplifier time constant effect. The detector sees a changing amplitude of a square wave (after chopping) and at 2kHz this was too slow to avoid capacitive opposition, consistent with the earlier results of Section 4.6.1.2.

The respective areas under the doublet happen to be nearly equal. However, this is a coincidence since those under the single line are not. The distortion at 2kHz is therefore too great to allow extraction of useful concentration values. Had a faster chopper system been available, it may have been more worthwhile pursuing this

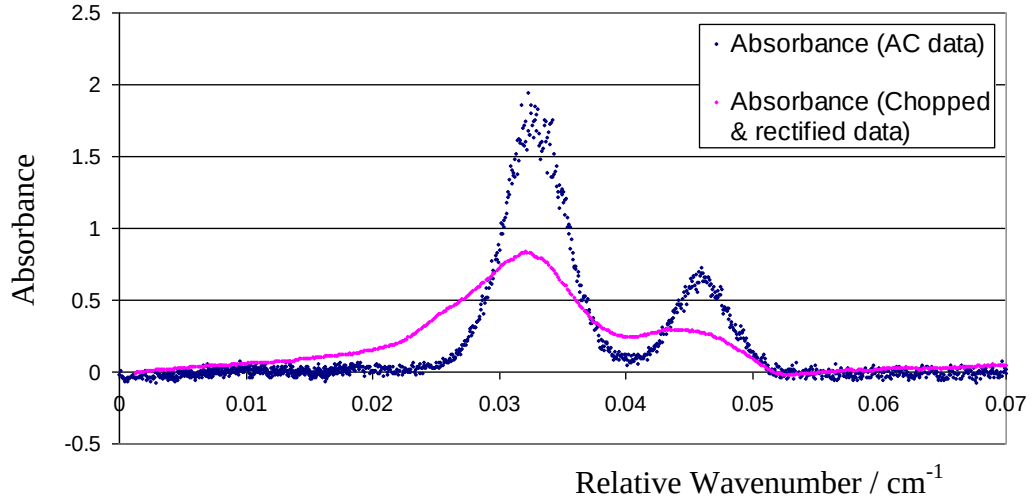


Figure 4-14: Illustration of the transimpedance amplifier's 180µs time constant effect on resolution of doublet. The data was processed from that in Figure 4-12 and Figure 4-13. (See Section 4.6.3 for notes on attempted line identifications).

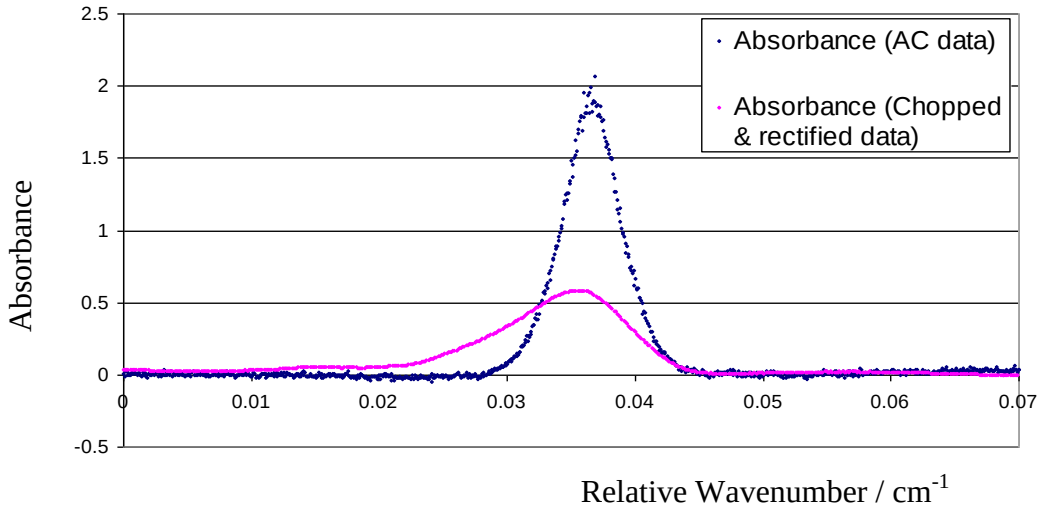


Figure 4-15: Illustration of the transimpedance amplifier's 180µs time constant effect on single line. The data was processed from that in Figure 4-12 and Figure 4-13.

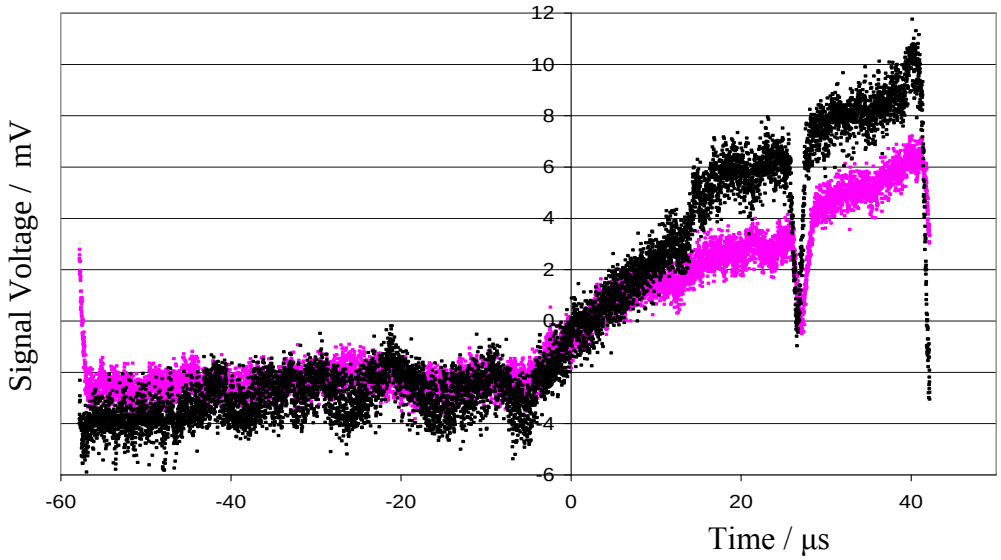


Figure 4-16: Anomalous scaling between transimpedance amplifier and oscilloscope of what should be overlapping signals for the $^{12}CH_4$ υ_4 P(4)F_1 3←F_2 1 line at a wavenumber of 1281.610588cm^{-1}. Here, the recorded scan at 100mV per division (black) is compared to that at 5mV per division (pink). Sweep integration parameters: 10kHz, I_{base}=500mA, $I_{base-peak}$=180mA, $T_{substrate}$=31°C, 128 averages.

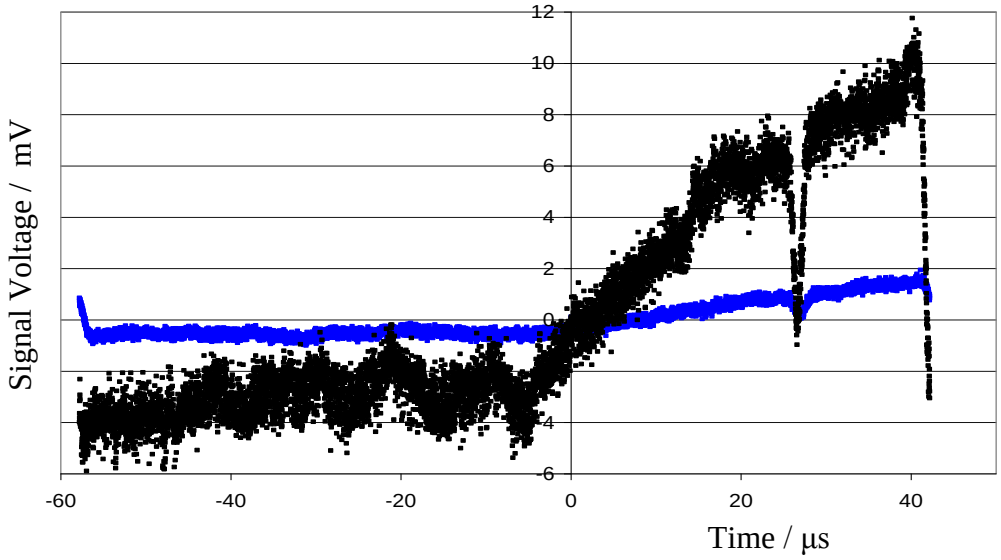


Figure 4-17: Anomalous scaling between transimpedance amplifier and oscilloscope of what should be overlapping signals for the $^{12}CH_4$ υ_4 P(4)F_1 3←F_2 1 line at a wavenumber of 1281.610588cm^{-1}. Here, the recorded scan at 100mV per division (black) is compared to that at 1mV per division (blue). Sweep integration parameters: 10kHz, I_{base}=500mA, $I_{base-peak}$=180mA, $T_{substrate}$=31°C, 128 averages.

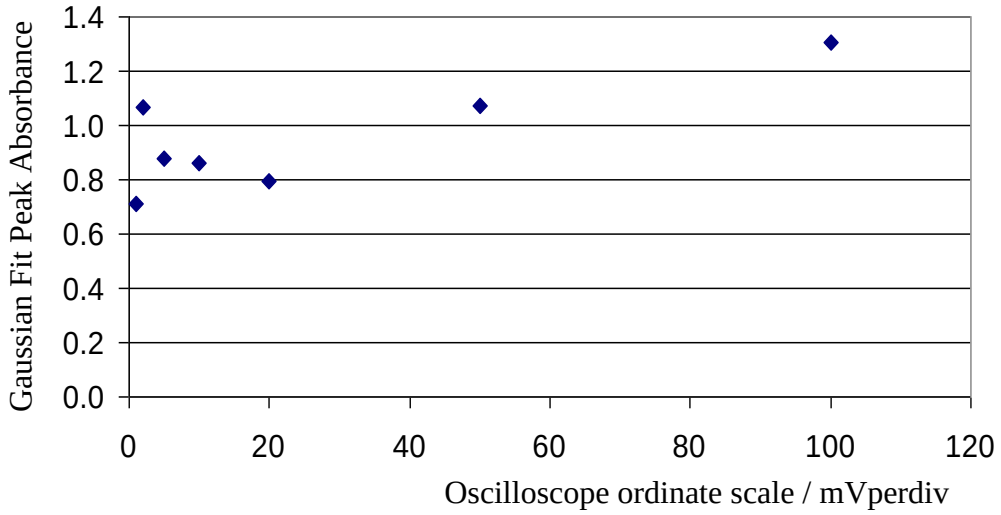


Figure 4-18: Anomalous scaling of the $^{12}CH_4$ υ_4 P(4)F_1 3←F_2 1 line's absorbance height under the same conditions but with different oscilloscope scale settings.

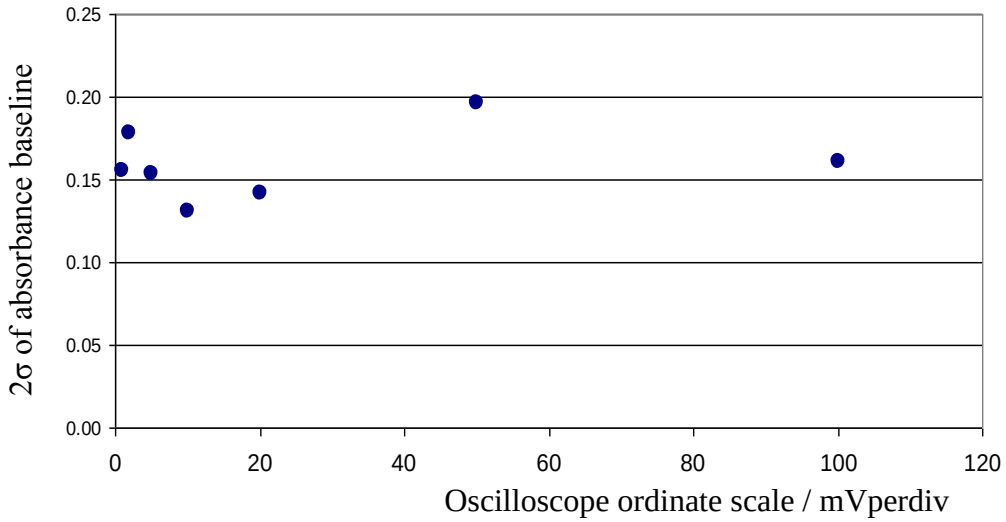


Figure 4-19: Anomalous scaling of the $^{12}CH_4$ υ_4 P(4)F_1 3←F_2 1 line's 2σ absorbance noise floor under the same conditions but with different oscilloscope scale settings.

approach. Nevertheless, the previous graphs of Figure 4-12 and Figure 4-13 do show that the below laser threshold noise level is close enough to a noise level at 0V to be a useful approximation.

4.6.1.4 Scaling problem between transimpedance amplifier and oscilloscope

Near the end of the tenure over the laser, it became apparent that many of the calculated ratios (Chapter 5) were not as expected. This was traced to an incorrect scaling of both the direct and 2f (wavelength modulated) signals on the oscilloscope.

Although the ramp signal coming straight from the function generator scaled correctly, the signal via the detector-preamp or from the detector-transimpedance amplifier-lock-in signal chain did not. When the mV per division was changed on the oscilloscope the size of the signal on screen remained approximately the same (rather than, say, doubling when the mV per division was halved). This meant that a recording of the spectra gave very different sized signals for what should have been the same overlapping signal (for example, Figure 4-16 and Figure 4-17). This anomalous scaling was also non-linear, as shown by the calculated absorbance peak height and the 2σ of the baseline around the absorbance feature (Figure 4-18 and Figure 4-19). Therefore, the ordinate values for a spectral feature from scan to scan could not be simply compared if the scales had been changed – as, in general, there was both an unknown number density and scale difference. This meant that much data had to be abandoned.

However, the data used in calculating the sweep integration and WMS Limits of Detection (LoD) was taken without changing the respective scales of the direct and 2f spectra.

A sister Tektronix oscilloscope also showed the same anomalous scaling. Given that the ramp signal, when fed to the oscilloscope directly, scaled properly, it is likely that this anomalous scaling is an electronic load effect between the AC-coupled

transimpedance amplifier and the oscilloscope. Unfortunately, losing tenure of the laser gave no time to investigate this effect thoroughly.

4.6.1.5 Different parameters for sweep integration and wavelength modulation spectroscopy

The sweep integration had to be run at, at least, 10kHz to avoid the transimpedance amplifier time constant effects. The modulation frequency used for wavelength modulation spectroscopy was of the same order of magnitude, ≈10kHz (Chapter 5). Accordingly, the ramp had to be significantly slower to allow the modulation to oscillate sufficiently about a reasonably small set of wavelengths before moving on.

So, in general, different operating parameters were needed for sweep integration and wavelength modulation spectroscopy. The implication in this work was that care was needed in mapping and matching each wavelength modulated signal to its corresponding sweep integration signal - for the purpose of calibration and comparison.

Different parameters also meant that the sweep integration background or baseline could not be used to accurately normalize the wavelength modulated signal. (The background ramp is suppressed in the (2nd harmonic) modulated spectrum). Had it been desired to compare two modulated signals in the same scan this would have posed a problem. An approximate solution to such a problem is to use the power variation of modulated etalon fringes – although the height of these is influenced by any non-normal incidence of the laser beam onto the etalon. A LabVIEW routine was drawn to perform such a normalization (Section 3.4.3).

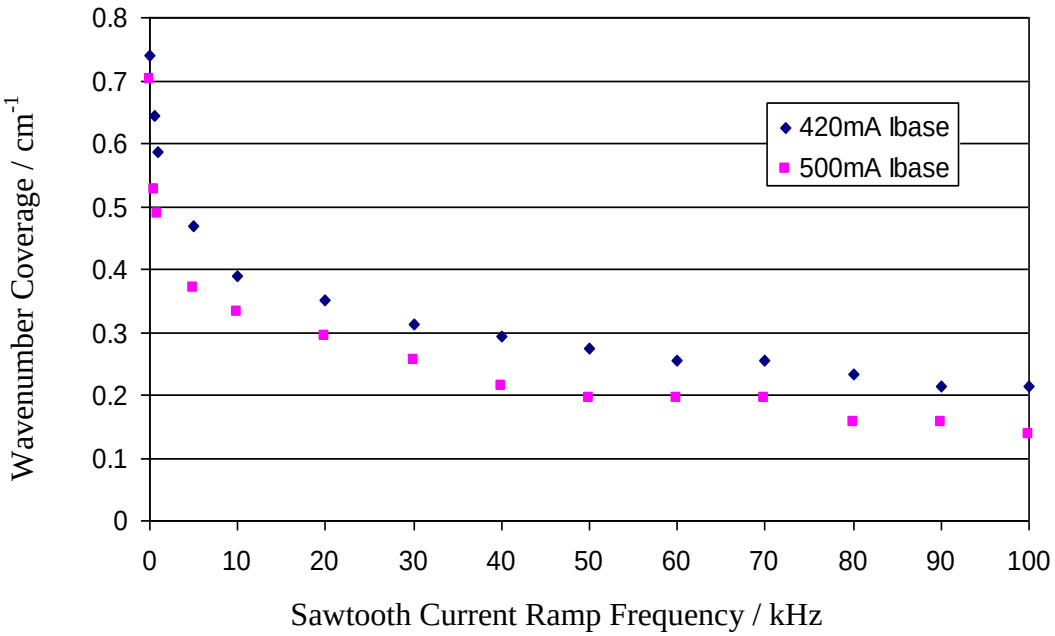


Figure 4-20: **cw-QCL wavenumber coverage for various sweep rates at two different sawtooth ramp base currents. Sweep integration parameters: (blue) I_{base}=420mA, $I_{base-peak}$=280mA, $T_{substrate}$=10°C; (pink) I_{base}=500mA, $I_{base-peak}$=200mA, $T_{substrate}$=10°C.**

4.6.2 Dynamic Tuning

4.6.2.1 Current sweep tuning characterisation: wavenumber coverage

The relative wavenumber coverage while applying a current ramp was determined by passing the cw-QCL radiation through a Germanium etalon of FSR=0.0195cm^{-1}. Figure 4-20 shows the wavenumber coverage for different sweep rates. The higher base current gives less coverage, being nearer to the maximum current, even though this base current has the advantage of starting the sweep nearer to laser threshold.

Although the spectral range is higher below a 10kHz sweep rate, Section 4.6.1.2 showed that such spectra are distorted by the transimpedance amplifier's high pass time constant. Given fast enough electronics, a high sweep rate can be desirable to achieve as much $\sqrt{N}$ averaging before drifts become significant. However, consideration of averaging in section 4.6.6 will show this not to be the limiting factor here, with the oscilloscope's limit of 512 averages being more of a limitation than the 10kHz sweep rate. Accordingly, 10kHz was used as the standard sweep rate to maximize spectral coverage, making line identification easier.

Figure 4-21 shows the data sets at 10kHz, 70kHz, and 100kHz. There is a 0.24µs delay between laser and detector, including the optical pathlength delay. (The pathlength within the Herriott cell was measured with a pulsed-QCL laser. This gave the transit time of a light beam through the cell , section 4.4.1, which includes signal delay in the cabling and averaging time, plus the out of cell continuous wave beam path). The current ramp voltage (in green) used to trigger the oscilloscope is therefore displaced earlier from the spectral data, since this signal is cabled straight from the function generator. This 0.24µs displacement (minus a small time due to cabling length differences) is fairly negligible on the time scales shown.

The wavenumber coverage is seen to decrease from both ends of a scan as the repetition frequency increases. A relatively small absorption dip remains centrally

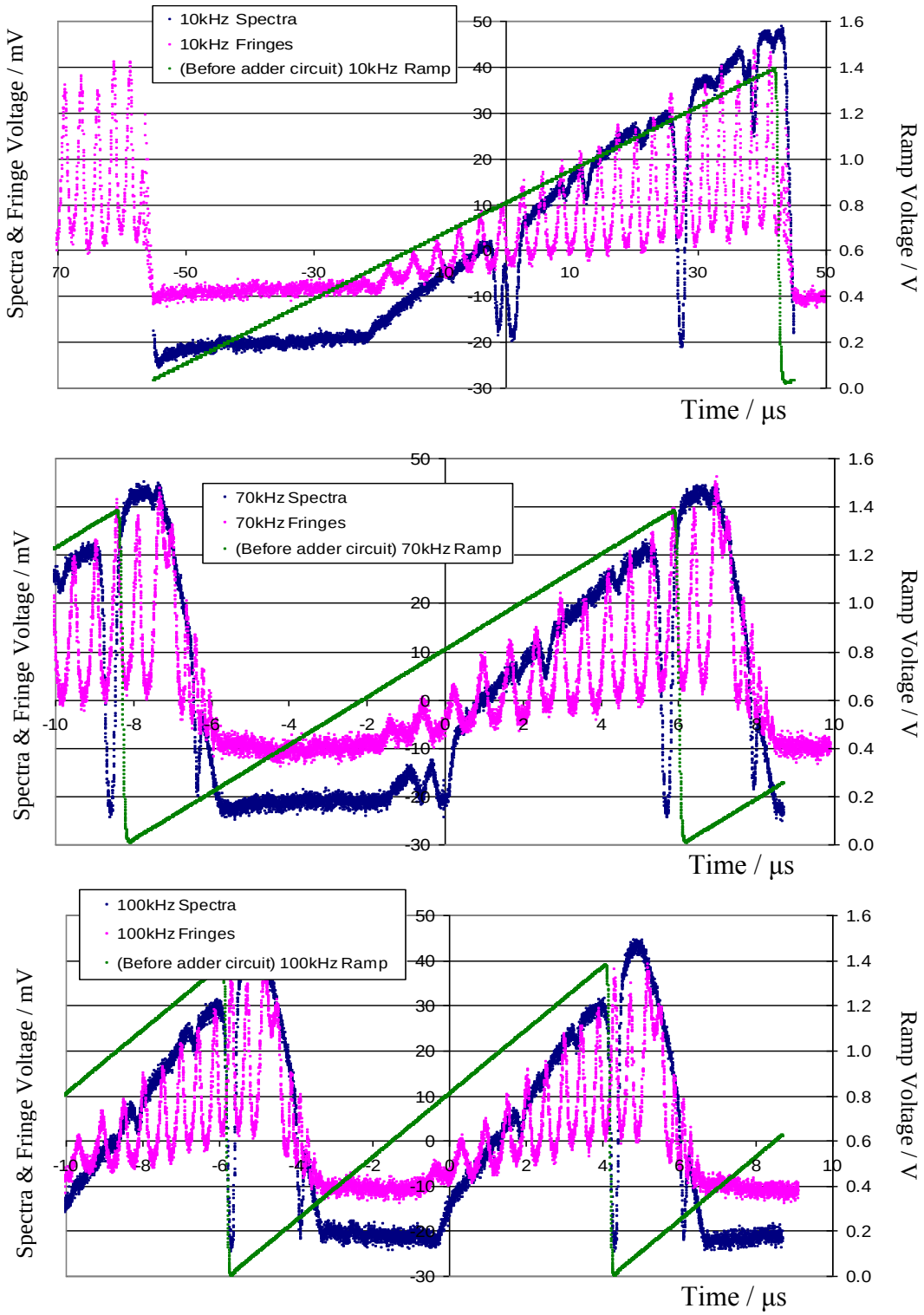


Figure 4-21: Sweep integration spectra (after passing through the 68m path length astigmatic Herriott cell) and Ge etalon fringes of FSR=0.0195cm^{-1}, with pre-adder circuit ramp voltages for repetition frequencies of 10kHz, 70kHz and 100kHz. I_{base}=420mA, $I_{base-peak}$=280mA, $T_{substrate}$=10°C, 512 averages.

placed. This behaviour is expected. As the ramp frequency increases, the Peltier cooler is no longer able to keep the laser substrate at the steady state temperature of its set-point. Heat extraction therefore becomes less efficient and the cw-QCL reaches and then oscillates about a steady state temperature that is primarily caused by the current ramp rather than the Peltier cooler. The thermal delay in wavenumber tuning implies that the amplitude of the wavenumber oscillation falls further behind the amplitude of the applied current ramp as the ramp frequency increases. This restricts the tuning range from both ends, as seen in Figure 4-21.

The long maximum to minimum fall times in Figure 4-21, of about 3μs to 1.5μs, were initially puzzling on the basis of the trigger ramp signal (in green) – which was recorded before passing to the laser through the in-house adder circuitry. An overestimate of ring down times can be made by treating the laser cavity and etalon as empty $(\alpha = 0)$ Fabry-Perot cavities, of length L and refractive index n_1, with Fresnel reflection losses, R, from both ends. After one round trip the intensity exiting a face is $I = I_0\left(Re^{-\alpha}\right)^2 = I_0\left(e^{\ln R}e^{-\alpha}\right)^2$. After m round trips the output intensity is $I_0\left(e^{\ln R}e^{-\alpha}\right)^{2m}$. m can be substituted for by using the radiation travel time, $t = m2Ln_1/c_0$. Going over to the continuous limit, the $1/e$ cavity ring down time from one facet is, after setting $\alpha = 0$ (Yariv and Yeh, 2007):

$$\tau = \frac{n_1 L}{c_0\left(-\ln R\right)} \tag{4-1}$$

where

$$R = \left(\frac{n_2 - n_1}{n_2 + n_1}\right)^2 \tag{4-2}$$

Taking L = 1mm and n_1 =3.8 (gallium arsenide) to approximate the quantum cascade laser, with n_2 = 1 for air, gives a cavity ring down time of 12ps. Taking L = 6.3cm and n_1 = 4.11 for the germanium etalon (Appendix C), with n_2 = 1 for air, gives a

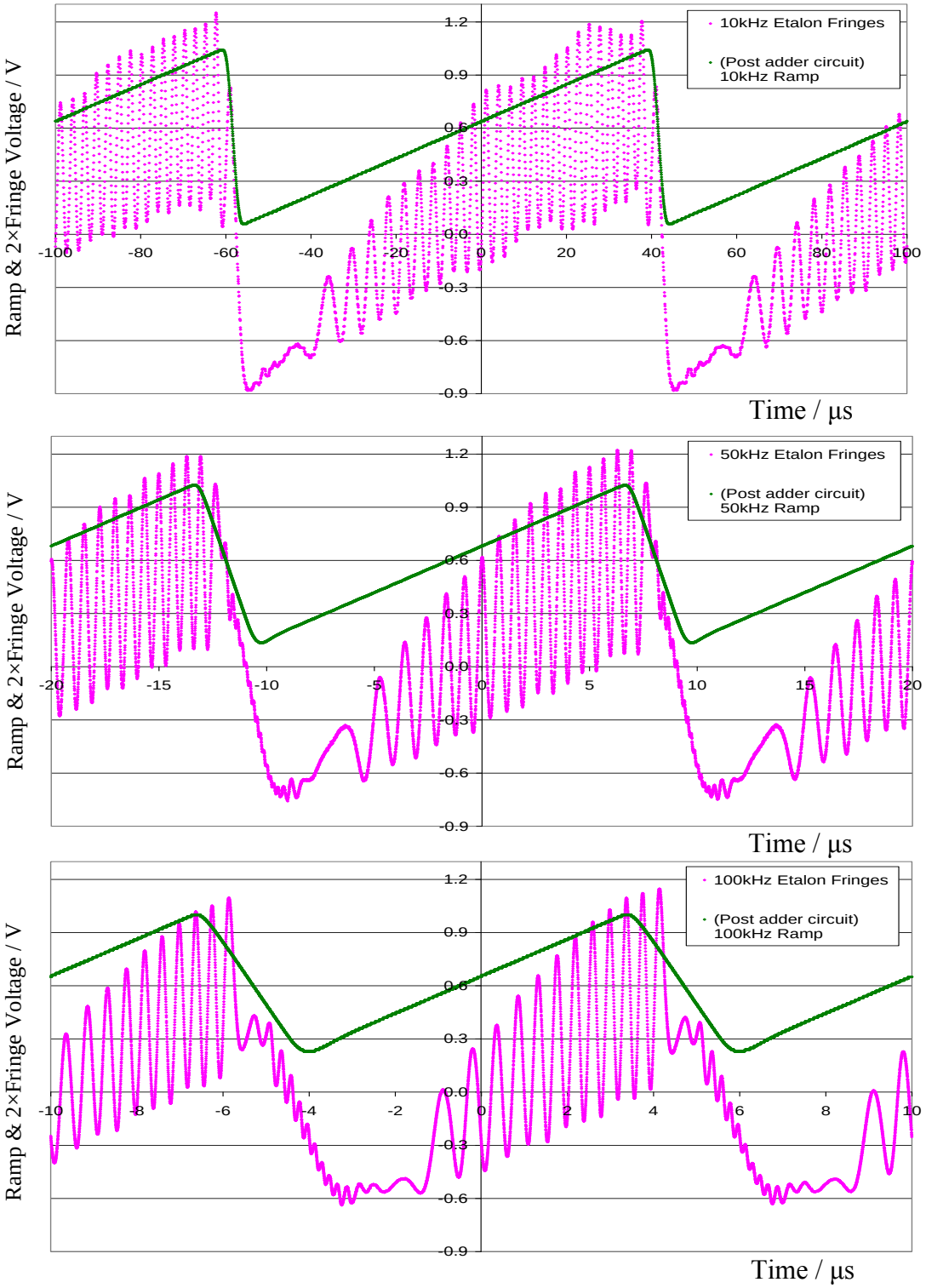


Figure 4-22: Signals without the Herriott cell in place. Ge etalon fringes of FSR=0.0195cm^{-1}, and post-adder circuit ramp voltages, for repetition frequencies of 10kHz, 50kHz and 100kHz. I_{base}=500mA, $I_{base\text{-}peak}$=200mA, $T_{substrate}$=20°C, 512 averages. (The 10kHz data was taken at a lower digital resolution per ramp period).

cavity ring down time of 0.9ns. Both cavity ring down times therefore have a negligible effect on Figure 4-21.

A look back at earlier characterization results, Figure 4-22, shows that ramp distortion is added by the in-house adder circuitry. Unlike the rapid sawtooth fall in injection current from the function generator (Figure 4-21), the fall after passing through the adder takes up a larger proportion of the total ramp period, as the ramp frequency increases – 4% at 10kHz, 15% at 50kHz, and 24% at 100kHz. The effect of this distortion is to prevent the injection current falling completely to its base current before rising again. This is gentler on the laser than the abrupt sawtooth fall, and since the laser still falls below threshold (Figure 4-22), this ramp distortion does not restrict the tuning range. This restriction is still due to the steady-state temperature reached, as explained above.

With the benefit of hindsight, Figure 4-22 was an early indication of the oscilloscope scaling problem in the detector-transimpedance amplifier-oscilloscope chain (Section 4.6.1.4). The etalon fringes were recorded as seen in the figure relative to the ramp signal – and with a common ordinate scale. Greater experience with the VIGO detectors, with different associated equipment, such as in Chapter 7, indicated that these detectors saturate when producing a voltage above 1V. The fringe voltage is therefore labeled as '2×' on the ordinate axis.

4.6.2.2 Current sweep tuning characterisation: wavenumber tuning rate

The tuning rates implied by the spacing of the etalon fringe peaks were mapped for sweep frequencies from 500Hz to 100kHz. The polynomial fit for 10kHz is given in Figure 4-23, relating the equal frequency spacing of the peaks to their occurrence in time. For all the sweep frequencies, a quadratic polynomial with a small degree of curvature was an excellent fit to this relationship. Accordingly, the tuning rates (the 1st derivatives of these relationships) were close to linear – implying a constant rate of change of the tuning rates with time (2nd derivative of the calibration curve), for a

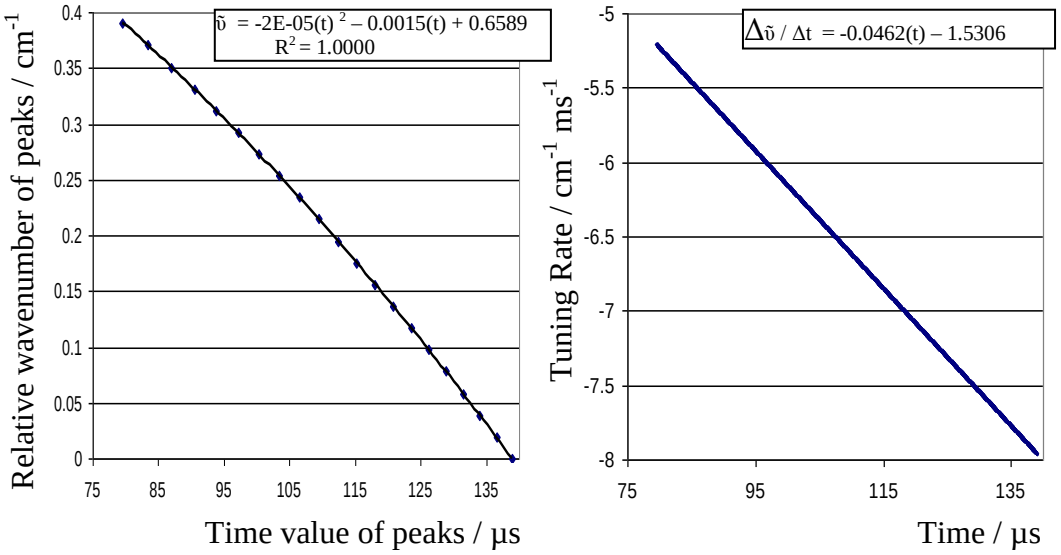


Figure 4-23: 10kHz sawtooth ramp, I_{base}=420mA, $I_{base-peak}$=280mA, $T_{substrate}$=10°C: relative wavenumber to time spacing relationship of Ge etalon fringe peaks with FSR=0.0195cm^{-1}, and this relationship's derivative (the tuning rate). The fit equations are shown.

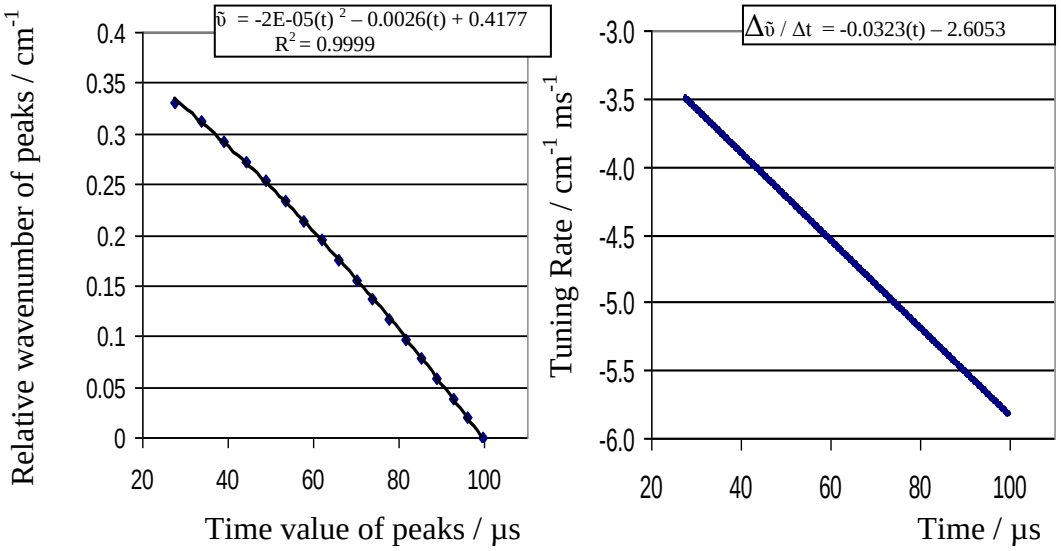


Figure 4-24: 10kHz sawtooth ramp, I_{base}=500mA, $I_{base-peak}$=180mA, $T_{substrate}$=13°C: relative wavenumber to time spacing relationship of Ge etalon fringe peaks with FSR=0.0195cm^{-1}, and this relationship's derivative (the tuning rate). The fit equations are shown.

given sweep frequency. The average tuning rate in Figure 4-23 is -6.58$cm^{-1}ms^{-1}$ with wavenumber coverage of 0.39cm^{-1}.

Figure 4-24 shows the similar 10kHz tuning data for the laser parameters used in section 4.6.5, of average tuning rate -4.65$cm^{-1}ms^{-1}$ and 0.33cm^{-1} wavenumber coverage. These are applicable to the data set used to quantify the limit of detection and to compare with WMS in Section 5.8.

Figure 4-25 shows the linear tuning rates and their envelopes for different sweep rates. Overall, the tuning rates as a function of sweep frequency show a roughly quadratic trend. This indicates that as the sweep rate increases the temperature tuning is less and less able to keep up with the applied voltage – consistent with the lower wavenumber coverage in the previous section.

4.6.3 Absorption Line Identification

With the guidance of the FTS tuning data, section 4.5.3, $^{12}CH_4$ absorption lines were tuned over and identified by comparing to 'Hitran on the Web' data (V E Zuev IAO). This data itself comes from the HITRAN molecular spectroscopic database (Rothman et al., 2009).

Due to the AC-coupled detection, in order to calculate the transmittance spectrum, the noise level – approximating for the 0V level – needs to be included in the same sweep as the spectral line (Section 4.6.1.1). Having found a line for one set of sweep integration / laser parameters, the threshold noise level can be brought on screen by increasing the substrate temperature (moving along the temperature axis in Figure 4-5). This increases the threshold current, and so can increasingly place more of the current ramp, of fixed base current, below threshold. With higher temperature, the sweep starts at a lower wavenumber (downchirp) for a given current, and so a spectral feature of fixed wavenumber (lower than the starting wavenumber) is encountered sooner. The spectral feature therefore tends to move off to the left of the

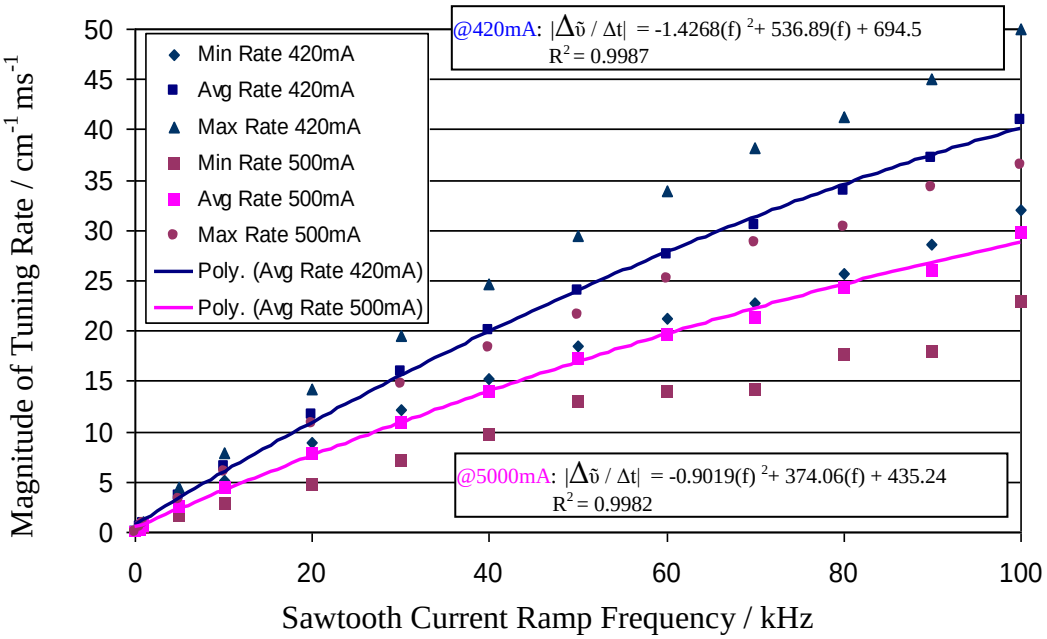


Figure 4-25: Tuning Rate Magnitudes from the linear variation of tuning rates with time, as illustrated in Figure 4-23 and Figure 4-24. The mid tuning rate and the envelopes are given for various sawtooth ramp frequencies, at two different base currents. The variation of the tuning rates with ramp frequency is roughly quadratic, and such a fit is given for the mid tuning rates. Sweep integration parameters: (blue) I_{base}=420mA, $I_{base\text{-}peak}$=280mA, $T_{substrate}$=10°C; (pink) I_{base}=500mA, $I_{base\text{-}peak}$=200mA, $T_{substrate}$=10°C.

$^{12}CH_4$ Identification	Wavenumber / cm^{-1}	Line Strength / cm^{-1}/molecule cm^{-2}	Sweep Integration: Laser Substrate Temperature / °C
υ_4 P(4)A_2 1←A_1 1	1283.458791	6.26×10^{-20}	13
υ_4 Q(13) F_1 4←F_2 1	1283.217895	2.13×10^{-21}	15
υ_4 Q(13) F_2 4←F_1 1	1283.207805	2.13×10^{-21}	15
υ_4 P(4) F_2 2←F_1 1	1282.984175	3.74×10^{-20}	17
υ_4 P(4) E 2←E 1	1282.624469	2.49×10^{-20}	21
υ_4 P(4) F_1 3←F_2 1	1281.610588	3.71×10^{-20}	31
Sweep Integration Parameters	Sawtooth Current Ramp: 10kHz, 500mA to 680mA		

Table 4-1: Laser substrate temperatures for finding the six strong $^{12}CH_4$ lines covered with sweep integration parameters: 10kHz sawtooth ramp, I_{base}=500mA, $I_{base\text{-}peak}$= 180mA.

oscilloscope screen. This movement can be counteracted by lowering the sweep base current (moving down an isotherm (quadratic fit) on Figure 4-5) which upchirps the sweep's starting wavenumber.

It was found possible to keep the sweep integration parameters common for all of the six strong methane lines found, with just a change of substrate temperature. The lines are identified in Table 4-1.

Some absorption lines were harder to identify. Repeated and undeclared use of the equipment complicated or negated many spectra. Much time was lost trying to identify what was spuriously thought to be intermittent outgassing from cell components or inward leaks. Many of the other absorption lines identified were water vapour or the common laboratory test gas nitrous oxide (for example see Figure 4-26 and Figure 4-27). Attempts were made to identify the lines initially appearing in Figure 4-10 and thereafter. Although they are in a similar wavenumber region to those in Figure 4-26, to the high wavenumber side of the CH_4 line at 1283.458791cm^{-1}, their identities have remained resistant to modeling in Spectralcalc and HITRAN on the Web (Spectralcalc.com, V E Zuev IAO). Complicating the analysis is the sensitivity of the line spacing to the etalon FSR chosen. The FSR is a function of etalon length and thus of ambient temperature. To be sure of the FSR at least two identifiable lines need to be in a spectrum – this was not the case in Figure 4-10 and Figure 4-13.

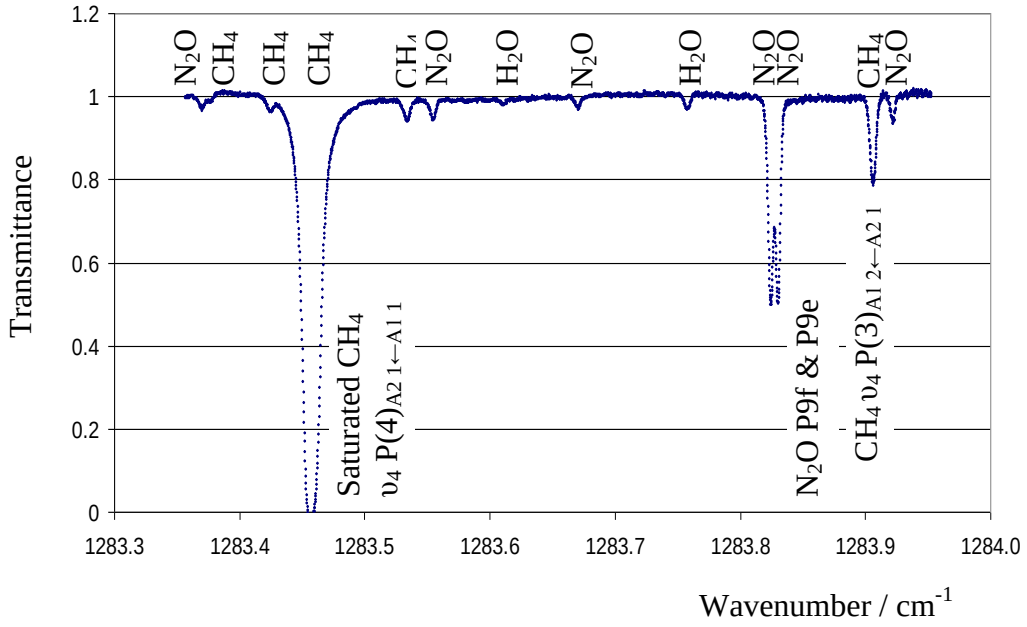


Figure 4-26: Transmittance spectrum of the N_2O P9f P9e doublet of spacing $\Delta\tilde{\nu} \approx 0.00585 cm^{-1}$

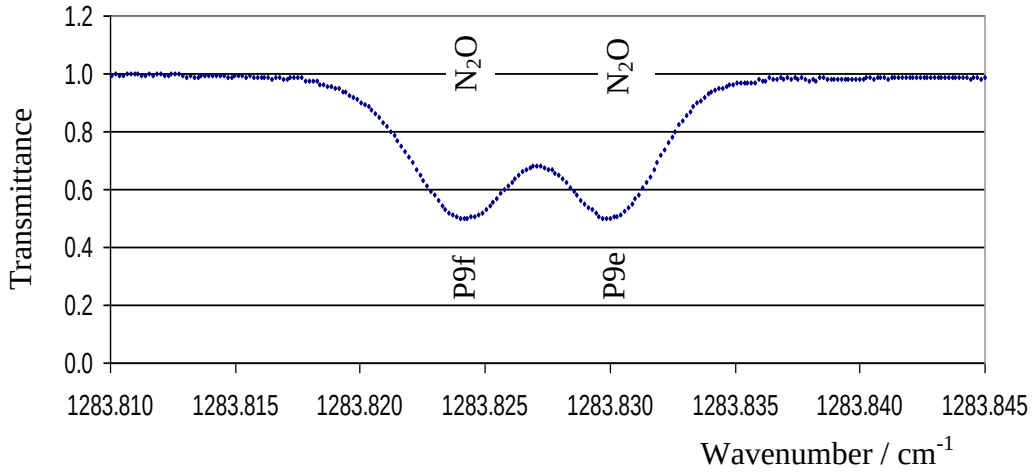


Figure 4-27: Close-up of a section of Figure 4-26.

4.6.4 System Resolution, Linewidths and Instrumental Broadening

The doublet featured in Figure 4-26 and Figure 4-27 was identified as the P9f and P9e lines of the $11^10 \leftarrow 01^10$ vibration-rotation band of N_2O at 1283.824000cm^{-1} and 1283.829850cm^{-1} (Rothman et al., 2009). This doublet, spaced by 0.00585cm^{-1}, provided a first measure of the system's resolution. To refine the measurement, a program was written to fit the sum of two Gaussians to the data via LabVIEW's Levenberg-Marquardt fitter (top panel of Figure 4-28). To obtain an acceptable fit it was necessary to fix the (known) centre spacing. This was done by tweaking the value for the etalon FSR in the time to wavenumber conversion. The peak heights and linewidths were then allowed to vary. Once fitted, the P9e line's centre frequency was translated in wavenumber so that the saddle point in the P9f and P9e sum was only just visible (bottom panel of Figure 4-28). This provided a better estimate of system resolution, when measuring N_2O at low pressure, of 0.0041cm^{-1}.

All the measured lines in this chapter exhibited symmetric broadening - significantly greater than their calculated Doppler widths. In each case, the measured lines were better fitted by Gaussian profiles - rather than Voigt profiles. Although Gaussian and Voigt profiles are not too dissimilar for the measured broadenings (for example Figure 4-29), the sources of Lorentzian broadening are too small to account for the measured FWHM's. Estimates of these are laid out below and summarized in Table 4-3. The symmetric non-Doppler Gaussian broadening is therefore mainly attributed to the system's electronics. The underestimate of the line wings in Figure 4-28 is attributed to the limitations of the LabVIEW fitting of two Gaussians - the fitting of the N_2O doublet was very inaccurate if the line separation was not fixed - rather than from any significant Lorentzian broadening – particularly as the fitting of isolated CH_4 lines (as in Figure 4-31) was Gaussian.

The Doppler FWHM's for P9f and P9e of N_2O both round to 0.00240cm^{-1} (296K) (Equation 2-20). The fitted FWHM's are much broader at 0.00411cm^{-1} for P9f and

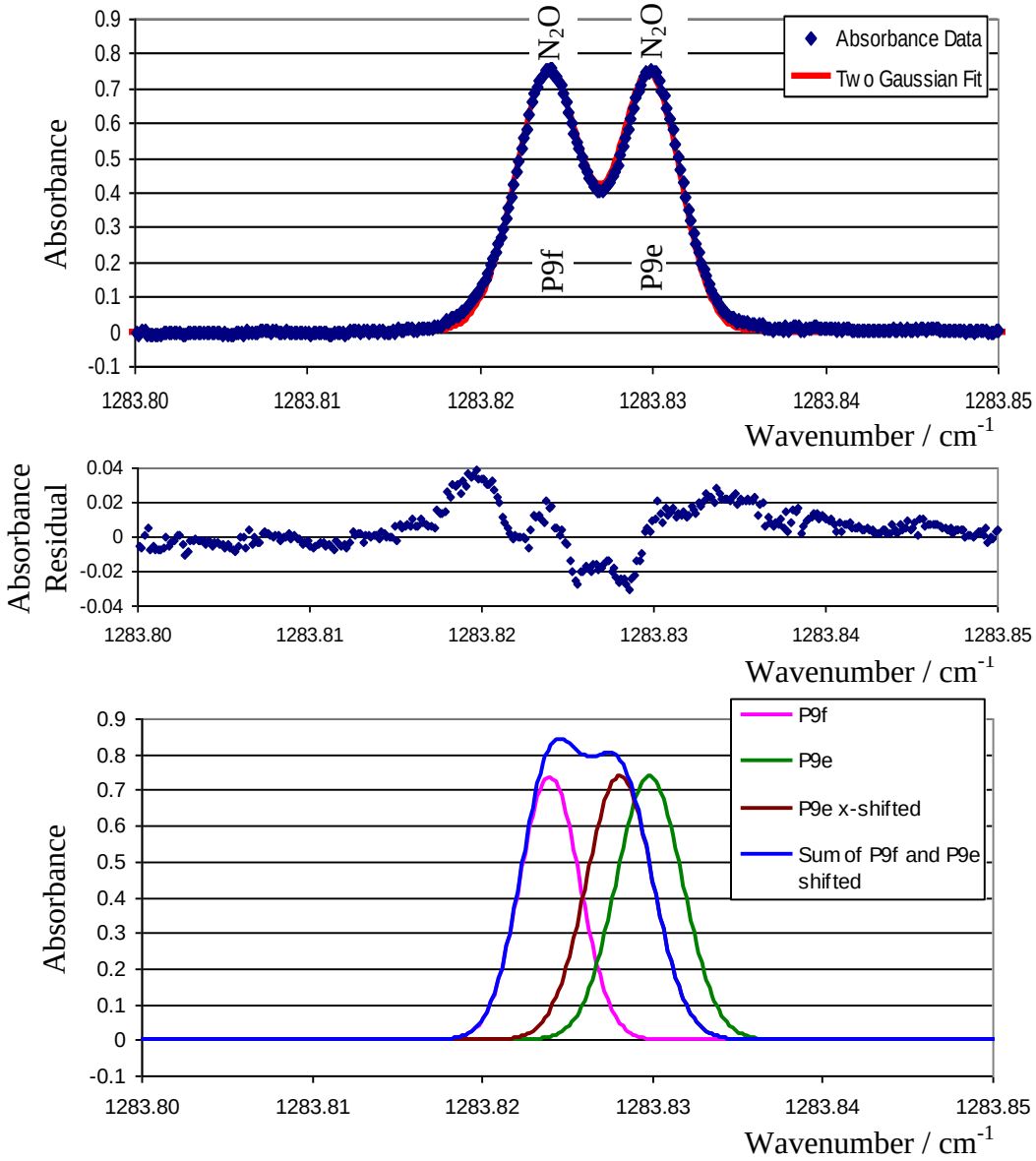


Figure 4-28: **Top panel: A two Gaussian Levenberg-Marquardt fit to a spectrum of the P9f and P9e lines of nitrous oxide, separated by 0.00585cm^{-1}.**

Middle panel: Residuals of the top panel fitted data (observed data minus the fit data).

Bottom panel: The sum of P9f and a wavenumber shifted P9e, so that the saddle point in the sum is just visible, giving a resolution of 0.0041cm^{-1}. This is comparable to the best resolution of Cascade Technologies' pulsed-QCL laser systems of 0.004cm^{-1}.

The Doppler FWHM of both lines is ≈0.00240cm^{-1} (296K), whereas the fitted FWHM's are 0.00411cm^{-1} for P9f and 0.00460cm^{-1} for P9e. This implies broadening (in quadrature) of 0.0033cm^{-1} and 0.0039cm^{-1} respectively.

0.00460cm^{-1} for P9e. An estimate of the non-Doppler contribution to broadening can be estimated since Gaussian FWHM's convolve 'in quadrature'. Pythagoras' theorem can therefore be used to deconvolve the contributions, giving non-Doppler broadening FWHM's of 0.00334cm^{-1} for P9f and 0.00392cm^{-1} for P9e - 39% and 63% larger than the Doppler widths respectively. The ratios of the non-Doppler line width to the measured line width are 81% and 85% respectively. If a Voigt profile had been appropriate, Lorentzian FWHM's can be calculated by rearranging Equation 2-28 and using the quadratic root formula for ω_l. This gives 0.00265cm^{-1} for P9f and 0.00329cm^{-1} for P9e. These FWHM's are also too large to be accounted for by the broadening mechanisms discussed below.

As an example of an isolated line, the Gaussian fit to be shown in Figure 4-32 has a measured FWHM of 0.00663cm^{-1} for $^{12}CH_4$ υ4 P(4)A_2 1←A_1 1 at 1283.458791cm^{-1}. The calculated Doppler FWHM is 0.00397cm^{-1} (296K). This implies a non-Doppler contribution of 0.00531cm^{-1}, 34% larger than the Doppler width. The ratio of the non-Doppler line width to the measured line width is 80%. (A Voigt profile giving the same measured FWHM would have ω_l = 0.00414cm^{-1}). Sliding two Gaussians together in the manner of the bottom plot of Figure 4-28, of the measured FWHM of 0.00663cm^{-1}, implies a resolution of 0.0058cm^{-1}.

Since the resolution is limited by the measured FWHM, the exact resolution of the system will depend on the lines, the gas being measured, and the operating conditions. The Doppler width is a function of the molecular mass, gas temperature, and line centre wavelength (Equation 2-20).

One possible source of broadening from the electronics is timing jitter - meaning that the starting times of equivalent scans do not coincide - from the triggering of voltages applied to the laser or in the measuring of the detected signal. Another source of broadening is wavenumber fluctuations and jitter in the laser's output - caused by instability from the laser's current and temperature controllers. Much analysis was necessarily carried out after the experimental period, so investigation of

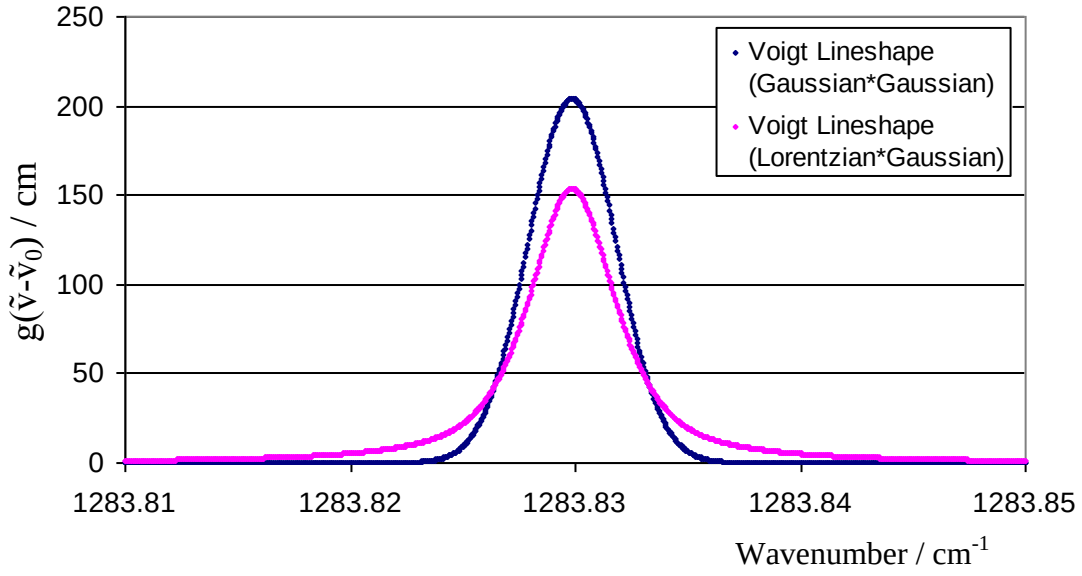


Figure 4-29: **Unit area normalized absorbance profiles for N2O P9e at 296K. The total FWHM of both lineshapes is the measured 0.00460cm^{-1}. The Doppler FWHM of both profiles is 0.00240cm^{-1}. For the overall Gaussian profile, the difference in widths is made up by convolving the Doppler Gaussian with a Gaussian of FWHM 0.00392cm^{-1}. For the Voigt profile, the Doppler Gaussian is convolved with a Lorentzian broadening of FWHM 0.00329cm^{-1}.**

this effect was limited. If a suitable laser is available, these effects could be investigated from spectral and etalon fringe data - for the different number of averages the oscilloscope can record (2^n, where n is an integer from 0 to 9 inclusive). The variation in line centre position could therefore be determined from single shot data, and the changing effective broadening measured as a function of the number of averages.

However, some indirect data was available from earlier characterization work. For Section 4.6.6, 15 sets of 512 averages on the $^{12}CH_4$ υ4 P(4)A_2 1←A_1 1 transition at 1283.458791cm^{-1} were taken for further averaging. Judging by Figure 4-33, and the fuller data set this figure was produced from, timing jitter does not appear to be an issue. However, the mean and standard deviation of the line centre position, pertaining to any particular 512 average set, is (1283.4588±0.0004)cm^{-1}. Also, the FWHM calculated mean and standard deviation is (0.0071±0.0008)cm^{-1}. (Errors in the data processing to fit the lineshapes are an order of magnitude smaller). These variations may imply that the underlying single shot Doppler profile is fluctuating in position while being averaged. This is plausible. As a rough approximation, a $^{12}CH_4$ Doppler broadened line of FWHM 0.004cm^{-1}, spaced in two positions 2×0.0004cm^{-1} apart (that is, plus and minus the standard deviation), has a superposition of FWHM ≈0.011cm^{-1}. (It can be noted that the pressure broadened FWHM's of Chapter 7 are an order of magnitude larger than these Doppler widths. Therefore, electronic and laser instability will have a larger proportional effect on the Doppler widths).

The assumed broadening from the system's electronics is in addition to the asymmetric AC-filter broadening from the transimpedance amplifier discussed in Section 4.6.1.2. The latter was not significant at a 10kHz sweep rate. At the low pressures used, pressure broadening is also negligible. Unfortunately, the available pressure gauge was not sensitive enough to measure the vacuum attainable. However, all pressures were significantly below 2Torr. This pressure was therefore used to place an upper estimate on the pressure broadening, to show that it would not

Line	$\tilde{\nu}$ / cm^{-1}	A_{lu} / s^{-1}	g_u	$\lvert\mu_{ul}\rvert$ / D	ν_R / MHz	$1/\nu_R$ / ns	$\Delta\nu$ / kHz	$\Delta\tilde{\nu}$ / cm^{-1}
N_2O P9f	1283.824000	6.55	153	1.23	5.84	171	929	3.10×10^{-5}
N_2O P9e	1283.829850	6.51	153	1.22	5.79	172	921	3.07×10^{-5}
CH_4 ν_4 P(4) A2 1 ← A1 1	1283.458791	2.21	35	0.41	1.95	514	314	1.05×10^{-5}

Table 4-2: Calculated power broadening for a 3mW laser beam with a cavity beam diameter of 1.80mm, giving an intensity of 1,179Wm^{-2} and an electric field of 942.5Vm^{-1} at the beam waist.

Broadening Mechanism	for N_2O P9f / cm^{-1}	for N_2O P9e / cm^{-1}	for CH_4 ν_4 P(4)$_{A2\ 1\leftarrow A1\ 1}$ / cm^{-1}
Natural	10^{-11}	10^{-11}	10^{-11}
Transit time	10^{-5}	10^{-5}	10^{-6}
Power broadening	10^{-5}	10^{-5}	10^{-5}
Chirp resolution	0.0005	0.0005	0.0004
Instantaneous linewidth, broadened by laser controllers (from literature)	0.0008	0.0008	0.0008
*Pressure FWHM	0.0006	0.0006	0.0004
Sqrt(Quadrature Sum)	0.0011	0.0011	0.0010
Addition Sum	0.0019	0.0019	0.0016
Doppler FWHM	0.0024	0.0024	0.0040
Measured FWHM	0.0041	0.0046	0.0066
Implied non-Doppler Broadening	0.0033	0.0039	0.0053

Table 4-3: Estimated contributions from various line broadening mechanisms for the lines shown in Figure 4-28 and Figure 4-31, showing a shortfall with the measured non-Doppler broadening. The majority contribution to the additional broadening is therefore attributed to the electronics.
***The pressure FWHM, calculated at 2Torr, is a particular overestimate.**

account for the measured broadenings. The tuning rate at P9f's centre position was -7.50cm^{-1} ms^{-1} and at P9e's centre position was -7.55cm^{-1} ms^{-1}. The chirp rate contribution to broadening at these centre positions is therefore ~0.0005cm^{-1} (Equation 2-70 with C=1). From literature values, the instantaneous linewidth of the laser is likely to be more than this estimate - broadened by fluctuations from standard current and temperature controllers (0.0008cm^{-1}, Section 2.10.1). The chirp rate at the CH_4 line centre was -5.06cm^{-1} ms^{-1}, giving a chirp rate contribution to broadening of 0.0004cm^{-1}.

The natural broadenings, estimated from the reciprocal of the Einstein A_{21} coefficients and the bandwidth theorem (Equation 2-5), are on the order of 1×10^{-11}cm^{-1}. Transit time broadenings from geometrical considerations are of the order of 1×10^{-5}-10^{-6}cm^{-1} (Section 2.4), and also not significant. The laser power entering the Herriott cell was about 3mW. Power broadening, described in Section 2.4, at 3mW with a cavity beam diameter of 1.80mm, gives an electric field of 942.5Vm^{-1} at the beam waist. These give FWHM's on the order of 1×10^{-5}cm^{-1}. (At 100mW the broadenings would be 1×10^{-4}-10^{-5}cm^{-1}). The power broadening calculations are laid out in Table 4-2, making use of the formulae in Section 2.4. $\tilde{\nu}_{ul}$, A_{lu}, and g_u are taken from HITRAN data. Both the square root of the quadrature sum (convolution of Gaussian FWHM's) and the larger addition sum (convolution of Lorentzian FWHM's) are given.

Leaving the pressure broadening overestimate and the power broadening aside, gives an unaccounted for line broadening of 0.0032cm^{-1} for N_2O P9f, 0.0038cm^{-1} for N_2O P9e, and 0.0063cm^{-1} for the CH_4 line (calculated in quadrature).

In summary, the chirp rate of the cw-QCL swept at 10kHz is 10^3 slower than the typical pulsed-QCL chirp rates. The chirp rate therefore is not the limiting factor on the cw-QCL spectrometers' resolution. Instead, the system's resolution is limited by the Doppler width of the absorption lines themselves, further broadened by (a majority contribution from) a symmetrical Gaussian broadening. Through the

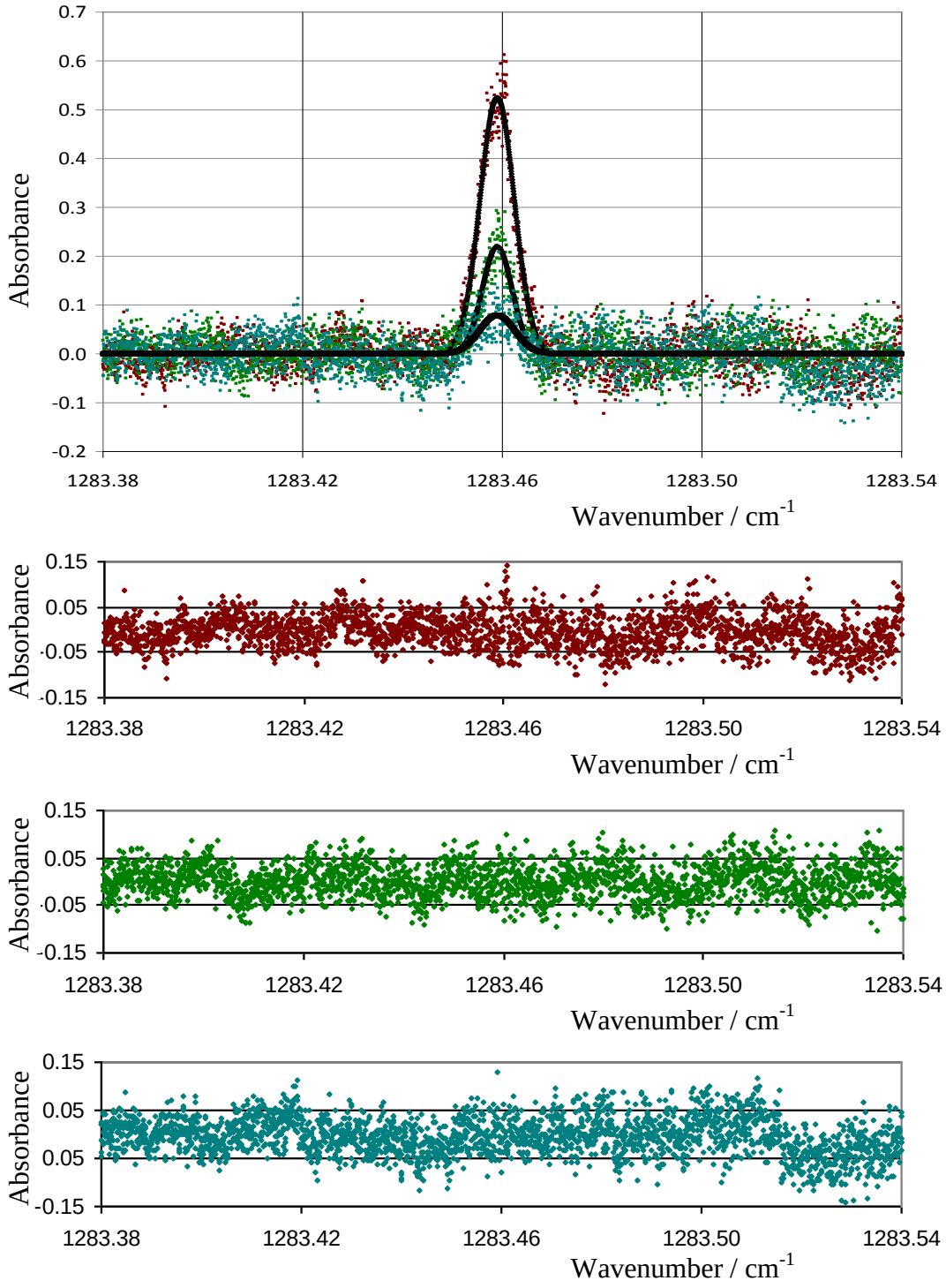


Figure 4-30: Absorbance spectra taken on the υ4 P(4)A_2 1←A_1 1 line of pure $^{12}CH_4$ at 1283.458791cm^{-1}, for three different number densities. Gaussian fits via LabVIEW's Levenberg-Marquardt fitter are also shown. Sweep integration parameters: 10kHz, I_{base}=500mA, $I_{base-peak}$=180mA, $T_{substrate}$=13°C. The fit residuals are also shown.

elimination of other likely possibilities, this is attributed to the system's electronics.

4.6.5 Sweep Integration data set for limit of detection and comparison with Wavelength Modulation Spectroscopy

The Herriott cell was sequentially pumped down to produce absorbance spectra of the $^{12}CH_4$ υ_4 P(4)A_2 1←A_1 1 transition at 1283.458791cm^{-1} (Rothman et al., 2009) for three different number densities (of pure CH_4). These spectra were used for calibration and comparison with the wavelength modulation technique (Chapter 5). Their baseline also indicates the limit of detection (or 'noise floor') for the sweep integration technique with this cw-QCL / spectrometer combination and integration time.

An absorbance spectrum can be calculated from a sweep integration by division of background or by division by a polynomial fit to the baseline of the sweep. In some circumstances, division of background could be a noise reduction technique, benefitting from the same partial cancellation of random noise as leads to the $\sqrt{N}$ reduction of averaging by sweep integration. However, in practice, due to instability of the baseline (section 4.6.6), background division only approached the noise level obtained by polynomial division. Background division, in this case, usually increased the noise level, and baseline instability between background and sweep integration could 'wash out' small features. Accordingly, polynomial division was used in this section.

Figure 4-30 shows the three absorbance spectra on a wavenumber scale convenient for comparison with the wavelength modulation technique. Gaussian fits were made to the data sets via LabVIEW's Levenberg-Marquardt fitter. Figure 4-31 shows the data and fits in close up, with the data points joined to indicate the noise envelopes.

It is common to see the limit of detection defined as one, two or three standard deviations of the baseline noise. Here, twice the standard deviation was chosen as

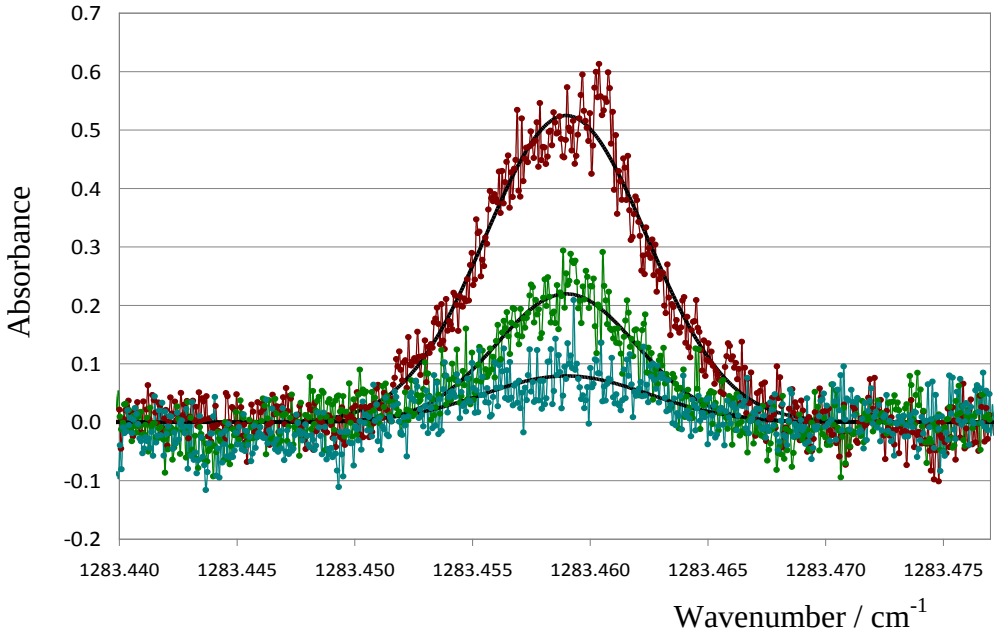


Figure 4-31: Close up of Figure 4-30 data showing the Gaussian fits and with data points joined to indicate noise envelopes. The noise deviation near the largest peak is likely due to cross-talk with the oscilloscope's trigger channel.

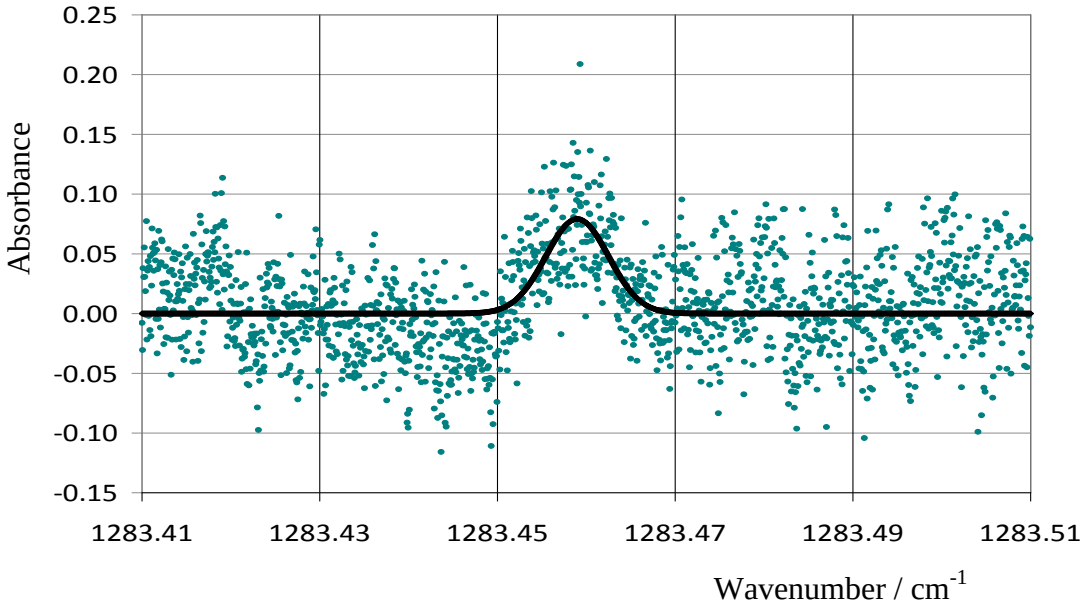


Figure 4-32: Data points from Figure 4-30 for the smallest pure $^{12}CH_4$ number density taken, and their Gaussian fit. The SNR is ~1. The baseline region for calculating the sweep integration technique's 2σ limit of detection (or noise floor) is shown.

this gave a SNR ≈1 for the lowest number density in Figure 4-30. That is, a signal which has a maximum absorbance equal to that of the noise floor, is considered to be the minimum absorbance and associated number density that the spectrometer can measure.

Figure 4-32 shows the smallest absorbance spectra recorded and the baseline points used to calculate the 2σ limit of detection. The data points to the left of the absorbance average around a slope. This was caused by baseline instability, and could have been straightened out in processing, but with minimal effect on the numbers extracted here. Both the peak height and the limit of detection are 0.080 absorbance units. The integration time for a 10kHz sweep integration is (512 averages × 0.1ms period =) 0.0512s, plus a small overhead for the oscilloscope averaging and display.

4.6.6 Potential for improvement by more averaging and increasing power to the detector

There is expected to be a $\sqrt{N}$ noise reduction up to a maximum integration time, at which point longer term drifts prevent further improvement. To test if the oscilloscope's maximum averaging of 512 sweeps was a limitation, 15 sweep integrations of 512 averages were recorded back to back (with 10,000 time points per set). A cumulative average signal (in multiples of 512 averages) was calculated. (The average of averages gives the overall average if the number of data points in each set is the same).

Figure 4-33 shows the average of all 15 sequential sweep integrations, consisting of 7680 averaged points per time value. This average was centred on most of the sweep integrations, but two wayward sets of 512 averages are displayed. This illustrates baseline instability from one sweep integration to the next.

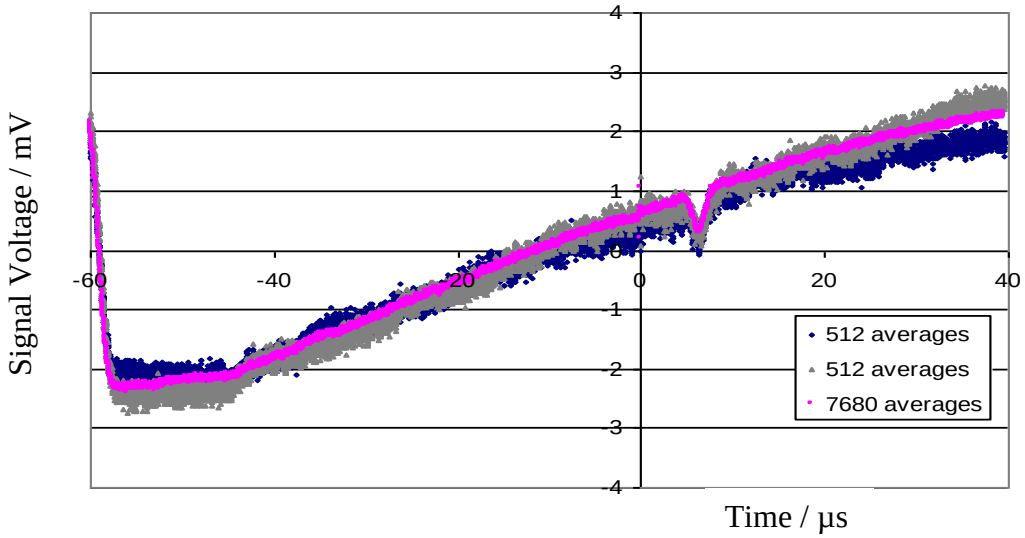


Figure 4-33: An average of 15 sweep integrations across the υ_4 P(4)A_2 1←A_1 1 line of $^{12}CH_4$ at wavenumber 1283.458791cm^{-1}. Each sweep integration is itself averaged 512 times. The average (pink) is centred on most of the sweep integrations, but two wayward runs are shown to illustrate baseline instability. Sweep integration parameters: 10kHz, I_{base}=420mA, $I_{base-peak}$=280mA, $T_{substrate}$=10°C.

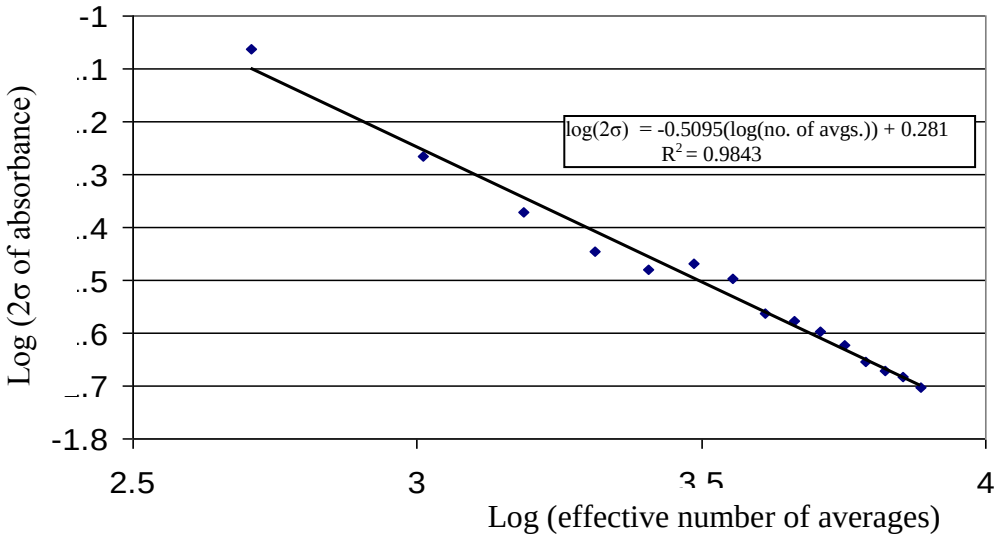


Figure 4-34: Log-log plot of the absorbance limit of detection against the effective number of averages (in multiples of 512), derived by cumulatively adding and averaging 15 sequential sweep integrations (of 512 averages each). The linear trend line shows the $\sqrt{N}$ reduction in baseline noise.

The 2σ noise floor was calculated about the absorbance line for the cumulative average data set, from 512 averages to 15×512=7680 averages.

Figure 4-34 displays a log-log plot of the 2σ noise floor against the cumulative number of averages. It is the voltage level of the transmittance signal which is averaged by the oscilloscope to produce the $\sqrt{N}$ noise reduction. However, the voltage baseline is equivalent to a weak absorption level, and the absorption to absorbance approximation of Equation 2-45 applies. Thus, the voltage averaging of Figure 4-34 also shows the expected 0.5 gradient of the averaged absorbance relationship.

Figure 4-34 indicates that the noise is still reducing at 7680 averages, indicating that more averaging to reduce noise is possible. At 7680 averages the absorbance limit of detection has improved to 0.02 absorbance units in an integration time of ≈0.77s. (Compared to 0.08 absorbance units after ≈51ms).

It also seems likely that the noise floor could be improved by optimizing the power falling on the detector – thus benefitting from a decrease in optical noise with increasing power. It was prudent in the time available to make use the existing (pulsed) spectrometer optics (section 4.4.1), but this could be dismantled or rearranged to decrease the number of optical components and hence power losses.

4.7 Other Ramp Shapes

To gain insight into the behaviour of the cw-QCL to injection current, early experiments were undertaken with different injection current waveforms. Figure 4-35 illustrates the cw-QCL response to a triangular injection current waveform. The repetition frequency is 10kHz and the base current for the ramp begins above laser threshold at 520mA. The two highest power fringes in Figure 4-35 represent the transition of the tuning downchirp (current up-ramp) to that of the upchirp (current down-ramp). Even though the injection current is falling by the time of the maximum

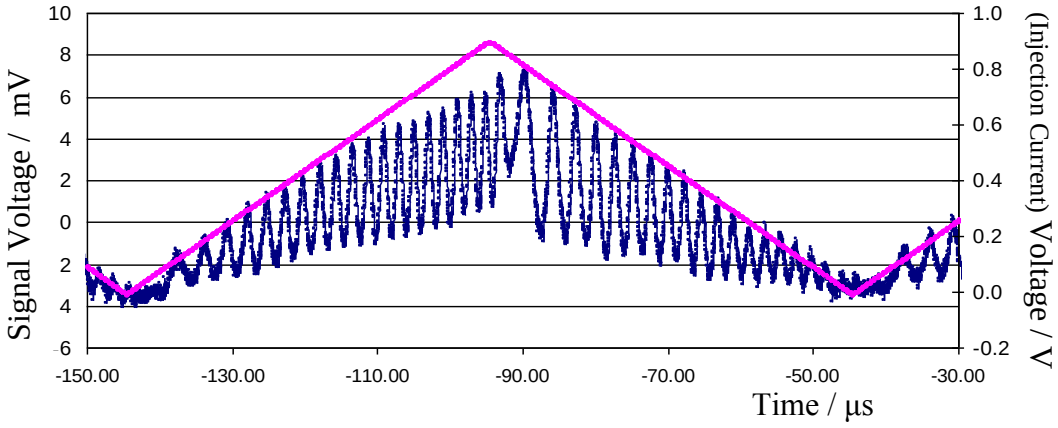


Figure 4-35: cw-QCL response to a 10kHz triangular injection current ramp, as illustrated by passing the laser's radiation through a Ge etalon of FSR=0.0195cm^{-1}. I_{base}=520mA, $I_{base\text{-}peak}$=180mA, $T_{substrate}$=10°C.

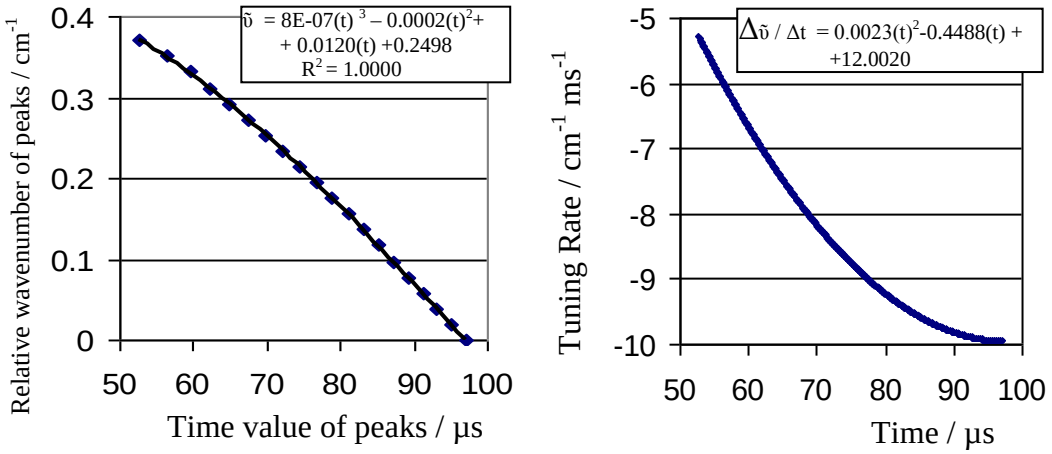


Figure 4-36: Up-Ramp, 10kHz triangular, I_{base}=520mA, $I_{base\text{-}peak}$=180mA, $T_{substrate}$=10°C: relative wavenumber to time spacing relationship of Ge etalon fringe peaks with FSR=0.0195cm^{-1}, and this relationship's derivative (the tuning rate). The fit equations are shown.

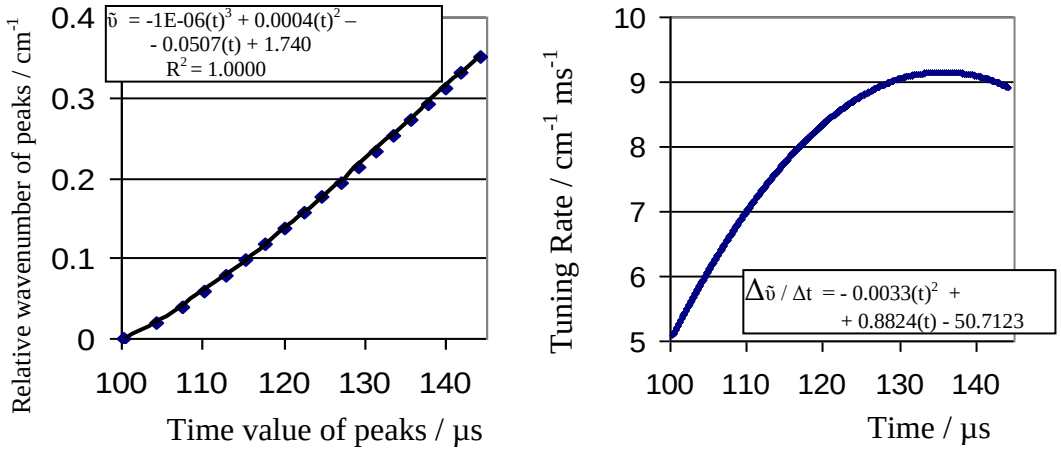


Figure 4-37: Down-Ramp, 10kHz triangular, I_{base}=520mA, $I_{base\text{-}peak}$=180mA, $T_{substrate}$=10°C: relative wavenumber to time spacing relationship of Ge etalon fringe peaks with FSR=0.0195cm^{-1}, and this relationship's derivative (the tuning rate). The fit equations are shown.

power fringe, this down-ramp initial power increase is consistent with cooling suppressing non-radiative relaxation mechanisms.

The response to the up-ramp and down-ramp is a little asymmetrical. This fits with the injection current temperature tuning being more efficient, as it heats the active layer directly. Upon decreasing this current, heat extraction relies on the Peltier cooler trying to maintain the laser's substrate at the temperature set-point.

Both ramps are well fitted by cubics. The up-ramp (Figure 4-36) downchirps the laser's tuning over 20 fringes, ≈0.39cm^{-1}, from a rate of -5.4cm^{-1}ms^{-1} to -10.0cm^{-1}ms^{-1}. The current down-ramp (wavenumber upchirp) covers about 18 fringes, ≈0.35cm^{-1} before the power flattens out between up- and down-ramps. The down-ramp's tuning rate is from +5.2cm^{-1}ms^{-1} to +9.2cm^{-1}ms^{-1} (Figure 4-37). The asymmetry can be mainly attributed to the time for the Peltier cooler to start bringing the temperature down, shown by the increased etalon spacing just beyond the top of the ramp.

Figure 4-38 shows the cw-QCL tuning response to a sinusoidal injection current wave at 10kHz and 420mA base current. The heating – cooling asymmetry is more pronounced here due to the larger base to peak current. Spectral coverage during the up-ramp, induced by the direct injection current heating of the active layer, is about 27 fringes, ≈0.53cm^{-1}, and during the down-ramp, by Peltier extraction of heat via the substrate, 17 fringes, ≈0.33cm^{-1}. For the sinusoidal injection current, the rate of change of current goes through a maximum between peaks. This is reflected in the up-ramp's quadratic tuning rate – starting at -17.6cm^{-1}ms^{-1}, reaching a maximum at -20.4cm^{-1}ms^{-1}, then slowing to -4.5cm^{-1}ms^{-1}. The time for the Peltier to start bringing the laser temperature down, is evident in the spacing of the fringes near the maximum current. The downward tuning rate is more straight forward, beginning at +4.8cm^{-1}ms^{-1} and flattening out at +20.1cm^{-1}ms^{-1}.

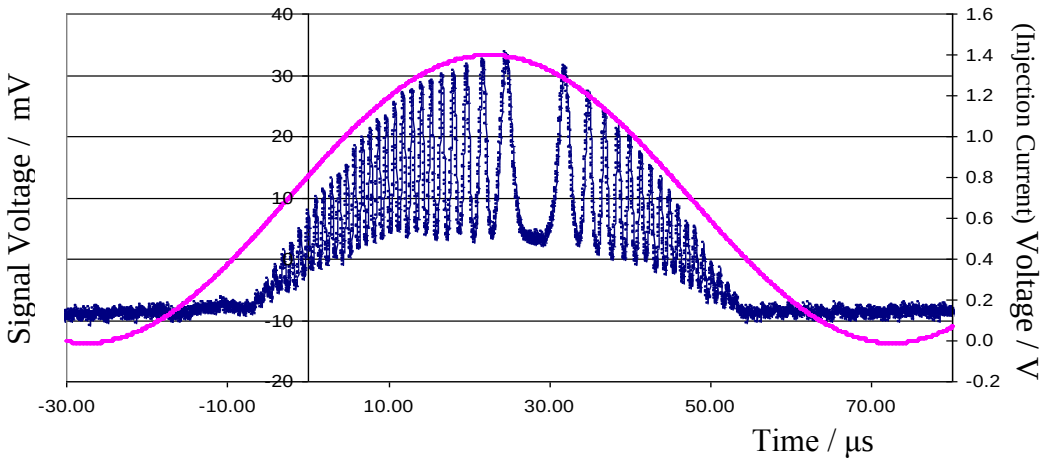


Figure 4-38: cw-QCL response to a 10kHz sinusoidal injection current wave, as illustrated by passing the laser's radiation through a Ge etalon of FSR=0.0195cm^{-1}. I_{base}=420mA, $I_{base\text{-}peak}$=280mA, $T_{substrate}$=10°C.

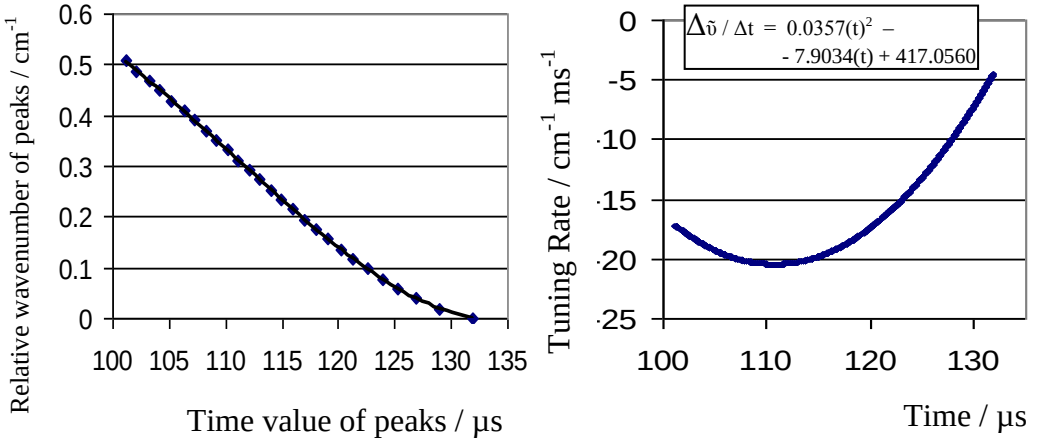


Figure 4-39: Up-Ramp, 10kHz sine, I_{base}=420mA, $I_{base\text{-}peak}$=280mA, $T_{substrate}$=10°C: relative wavenumber to time spacing relationship of Ge etalon fringe peaks with FSR=0.0195cm^{-1}, and derivative (tuning rate). The fit equations are shown.

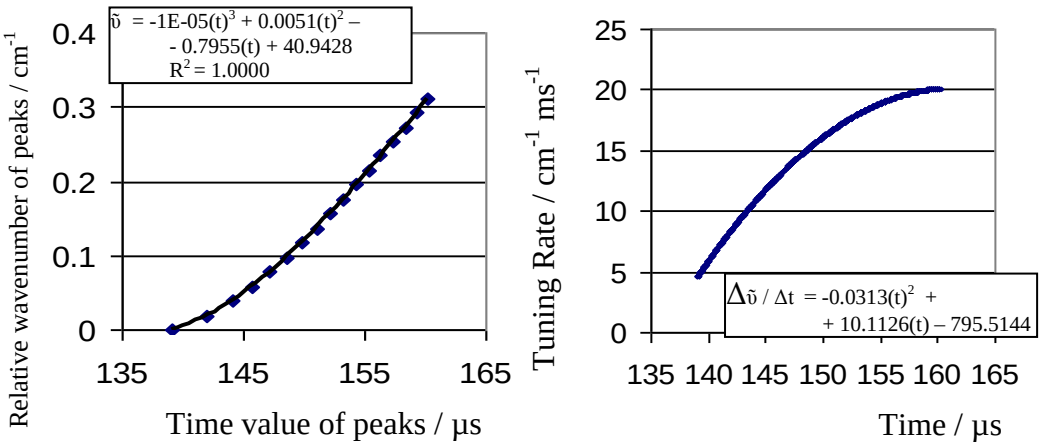


Figure 4-40: Down-Ramp, 10kHz sine, I_{base}=420mA, $I_{base\text{-}peak}$=280mA, $T_{substrate}$=10°C: relative wavenumber to time spacing relationship of Ge etalon fringe peaks with FSR=0.0195cm^{-1}, and derivative (tuning rate). The fit equations are shown.

Of some interest was the cw-QCL's response to a square wave of laser input voltage. Figure 4-41 shows the downchirp response to a 10kHz square wave. The applied square wave had an amplitude of 140mA straddling the base current of 420mA. The downchirp is reminiscent of the behaviour of pulsed-QCL's. The time scale though is very different, with the pulsed version typically being run with a pulse width of 500-1,500ns and a duty cycle of 2.5% to 7.5%. Here, the cw-QCL has a pulse width of 50μs and duty cycle of 50%. Nevertheless, the potential to take spectra using the same method as the pulsed systems may present integration synergies - particularly in adapting the software and fitting routines of current systems.

There is an optical path length delay of 0.24μs between laser output and detection in Figure 4-41. This is fairly negligible on the timescale of the figure. Beyond this, a phase delay is seen to exist between voltage application and laser output.

Sections of the cw-QCL etalon fringe pattern are amenable to fitting with low order polynomials. The initial rise has a very linear wavenumber – time relationship, which implies a constant tuning rate of -200$cm^{-1}ms^{-1}$ over the first 14 fringes (≈0.25cm^{-1}) (Figure 4-42). The 10% to 90% intensity rise time is ≈1.2μs. The minimum to maximum rise and fall times are about 2μs each. These are well accounted for by the adder circuitry distortion discussed in Section 4.6.2.1 while taking the laser power from below threshold to its maximum. (At 10kHz the sharp edge of a 1V sawtooth waveform was distorted to 4.2 μs). The tuning rate corresponding to the power 'roll over' of the square waveform (fringes 10 to 22, ≈0.21cm^{-1}) is well described by a quadratic polynomial (Figure 4-43). By fringe 22, the tuning rate has slowed to -62$cm^{-1}ms^{-1}$, as the time the Peltier has to react to the thermal load increases. The nearly 14 fringes (≈0.27cm^{-1}) making up the flat power section could be reasonably fitted with a 6th order polynomial, but with a large fit error between the last two fringes. The last 7 fringes (≈0.13cm^{-1}) fit a cubic polynomial well, with the tuning rate starting at -8.2$cm^{-1}ms^{-1}$, before flattening out at -2.5$cm^{-1}ms^{-1}$ (Figure 4-44).

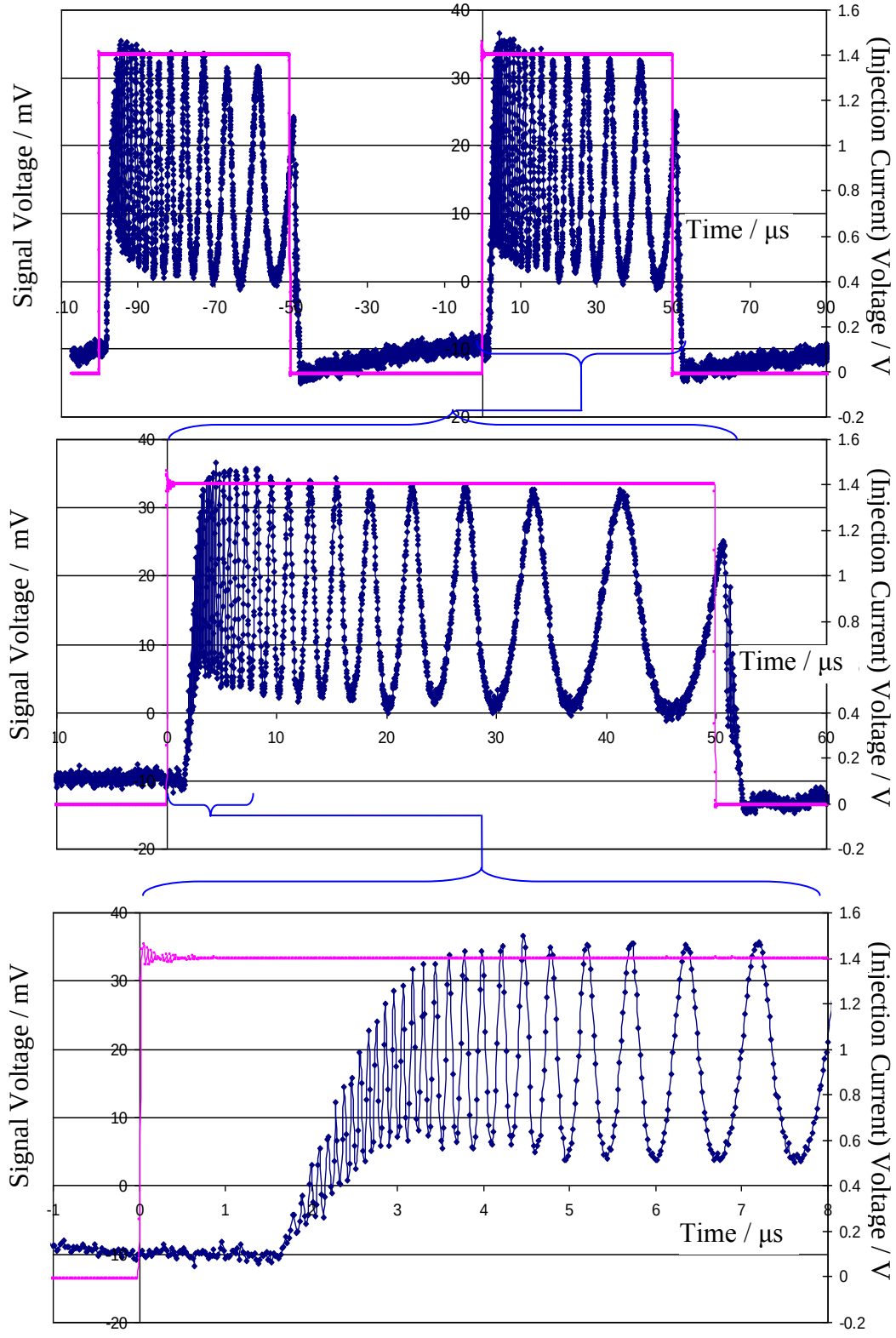


Figure 4-41: cw-QCL response to a 10kHz injection current square wave, as illustrated by passing the laser's radiation through a Ge etalon of FSR=0.0195cm^{-1}. I_{base}=420mA, $I_{base\text{-}peak}$=±140mA, $T_{substrate}$=10°C.

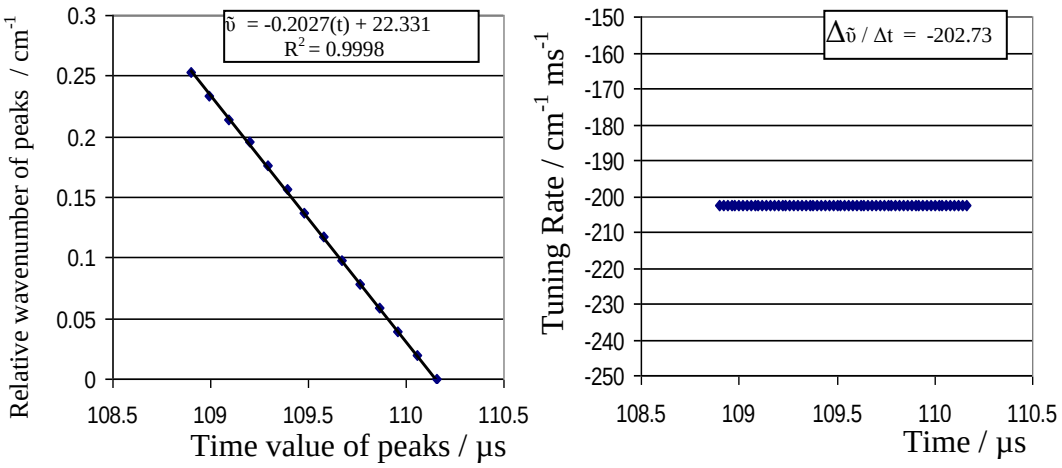


Figure 4-42: Linear fit of the relative wavenumber to time value of the etalon peaks of Figure 4-41, for the first 14 of 38 peaks. And this relationship's derivative (the tuning rate).

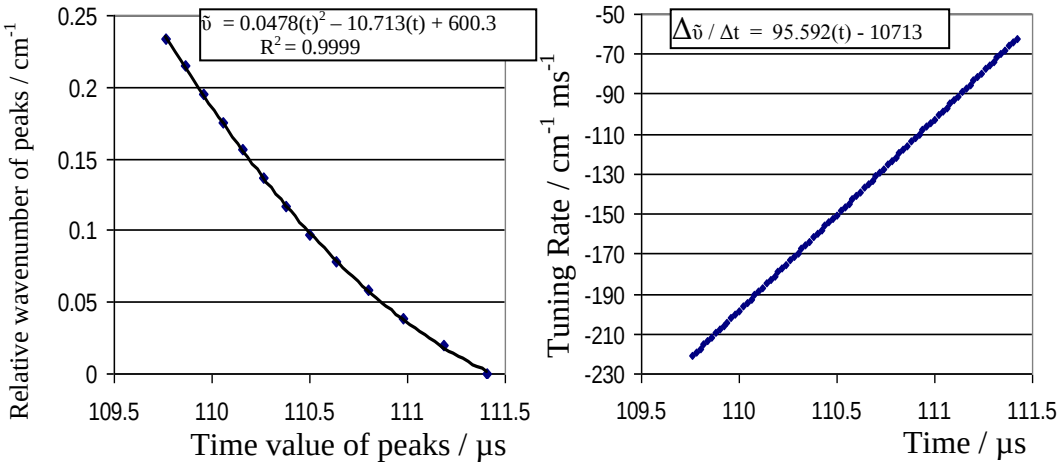


Figure 4-43: Quadratic fit of the relative wavenumber to time value of the etalon peaks of Figure 4-41, for the 'roll over' of power from peak 10 to 22 of 38 peaks. And this relationship's derivative (the tuning rate).

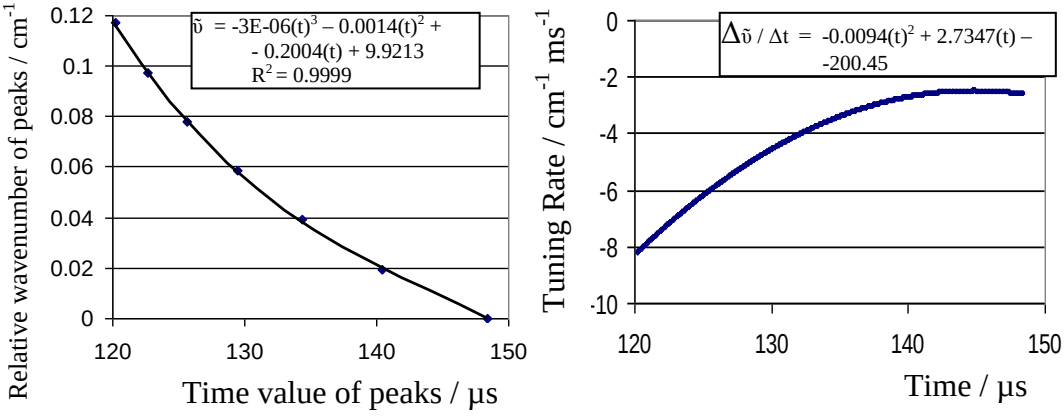


Figure 4-44: Cubic fit of the relative wavenumber to time value of the etalon peaks of Figure 4-41, from peak 32 to 38 of 38 peaks. And this relationship's derivative (the tuning rate).

Overall wavenumber coverage is nearly 39 fringes, about (39-1)×0.0195cm^{-1} = 0.74cm^{-1}, with the flat part of the power coverage being somewhat over a third at 0.27cm^{-1}. In comparison the Gaussian FWHM at room temperature for methane is 0.004cm^{-1}, using Equation 2-20.

Unlike for sawtooth, triangular and sinusoidal waveforms, slowing to 5kHz makes little difference to the number of fringes (41 fringes ≡ 0.78cm^{-1}). The laser still sees the same abrupt current step as at 10kHz. And, although the flat top of the pulse goes on for twice as long, the Peltier cooler's opposition to heating significantly reduces the chirp rate by the 'second half' of the pulse – covering only 3 more fringes. The minimum chirp rate at 5kHz repetition frequency is about -1.3cm^{-1}ms^{-1} – about half the minimum tuning rate at 10kHz.

At 5kHz or 10kHz, reducing the square wave amplitude reduces the respective chirp rates, and hence reduces the number of fringes. The heating effect is less pronounced with a smaller current step, and the Peltier is more able to oppose it.

By way of comparison, the etalon fringes from a pulsed laser, as used in the Cascade Technologies' spectrometer of Chapter 7 are shown in Figure 4-45 and analysed in Figure 4-46. The FSR of the Ge etalon here is some ×5 greater at 0.096cm^{-1}. The rise time is ≈12ns, spectral coverage 1.63cm^{-1}, and tuning rates about 10^3 faster compared to the slow tuning end of the cw-QCL pulse. Nearly all the top section of the pulsed-QCL pulse is amenable to a cubic fit of relative wavenumber to time. The tuning rate varies between -3.9cm^{-1}μs^{-1} and -1.6cm^{-1}μs^{-1} with an average of -2.5cm^{-1}μs^{-1}.

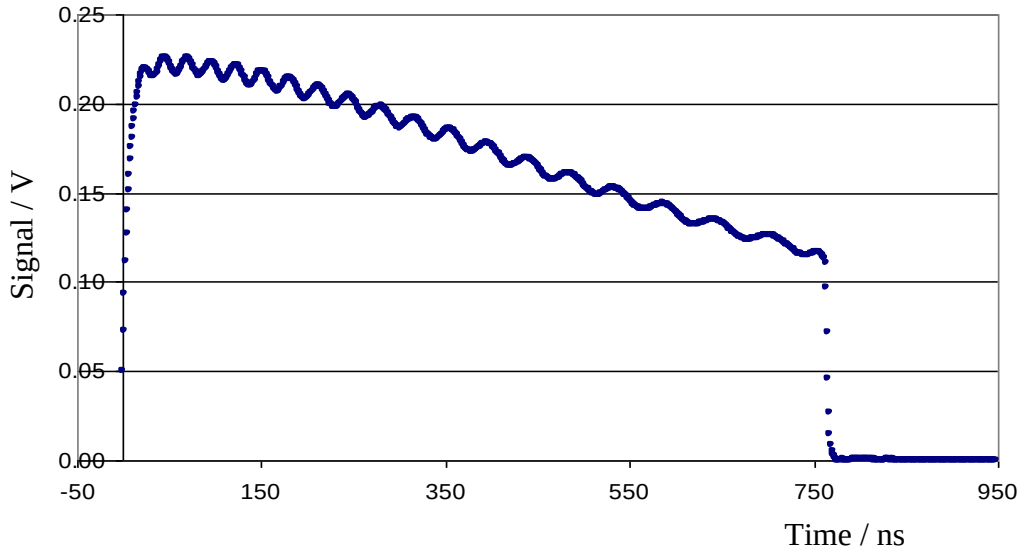


Figure 4-45: Pulsed-QCL response to a square wave voltage of 750ns and 3.75% duty cycle. Ge etalon FSR=0.096cm^{-1}, 1000 averages.

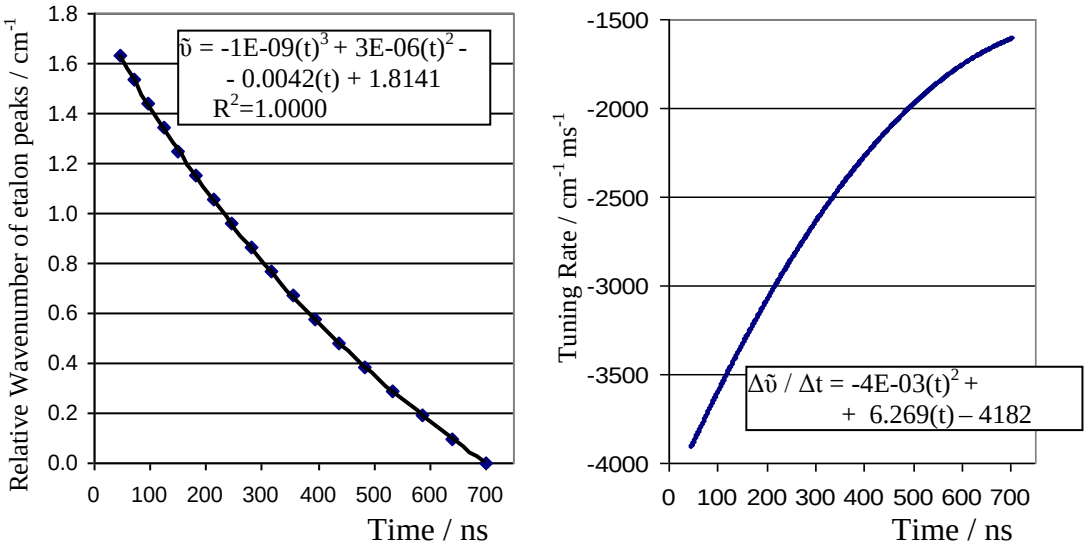


Figure 4-46: Cubic fit of the relative wavenumber to time value relationship of the pulsed-QCL's etalon peaks, and this relationship's derivative (the tuning rate).

4.8 Allan variance program

No software was available for taking continuous parameter values - such as concentration readings. Cascade Technologies' software was tied to their pulsed-QCL systems. In preparation for this software gap being filled, and in the event of a resumption of this line of research, a LabVIEW program was written to calculate the Allan variance from a continuous set of data.

Allan variance is a statistical technique for optimizing the number of averages for a given system. The written program is based on the equation found in Werle et al's paper (1993), reproduced as Equation 4-3 below. An interpretation of this equation is given below and in the diagrams.

As used here, Allan variance considers a single continuous time series of N unaveraged data points measuring some parameter characteristic of the system. For a spectrometer, this is usually the concentration readout. The data set of N points is sub-divided into 'bins' of k data points. Each bin is averaged to give a concentration value for that bin.

The concentration readout usually comes via calculating the area under an absorbance curve. Consider that if this underlying signal is constant, the concentration readout is also constant. However, consider that both the signal, and therefore the concentration readout, are subject to random experimental noise (no drift). The concentration in this case fluctuates about some average, and this average A_s becomes a more accurate approximation to the underlying concentration as more concentration values are averaged. That is, by taking a larger bin size of k data points, more data points are averaged, and the SNR is improved by the well known $\sqrt{k}$ factor (Taylor, 1997). This implies that, as the bin size is increased (equivalent to increasing the averaging ('integration') time τ), the average in each bin will be converging towards the 'true' underlying concentration value. The difference between these bin averages will therefore decrease with increasing bin size. The

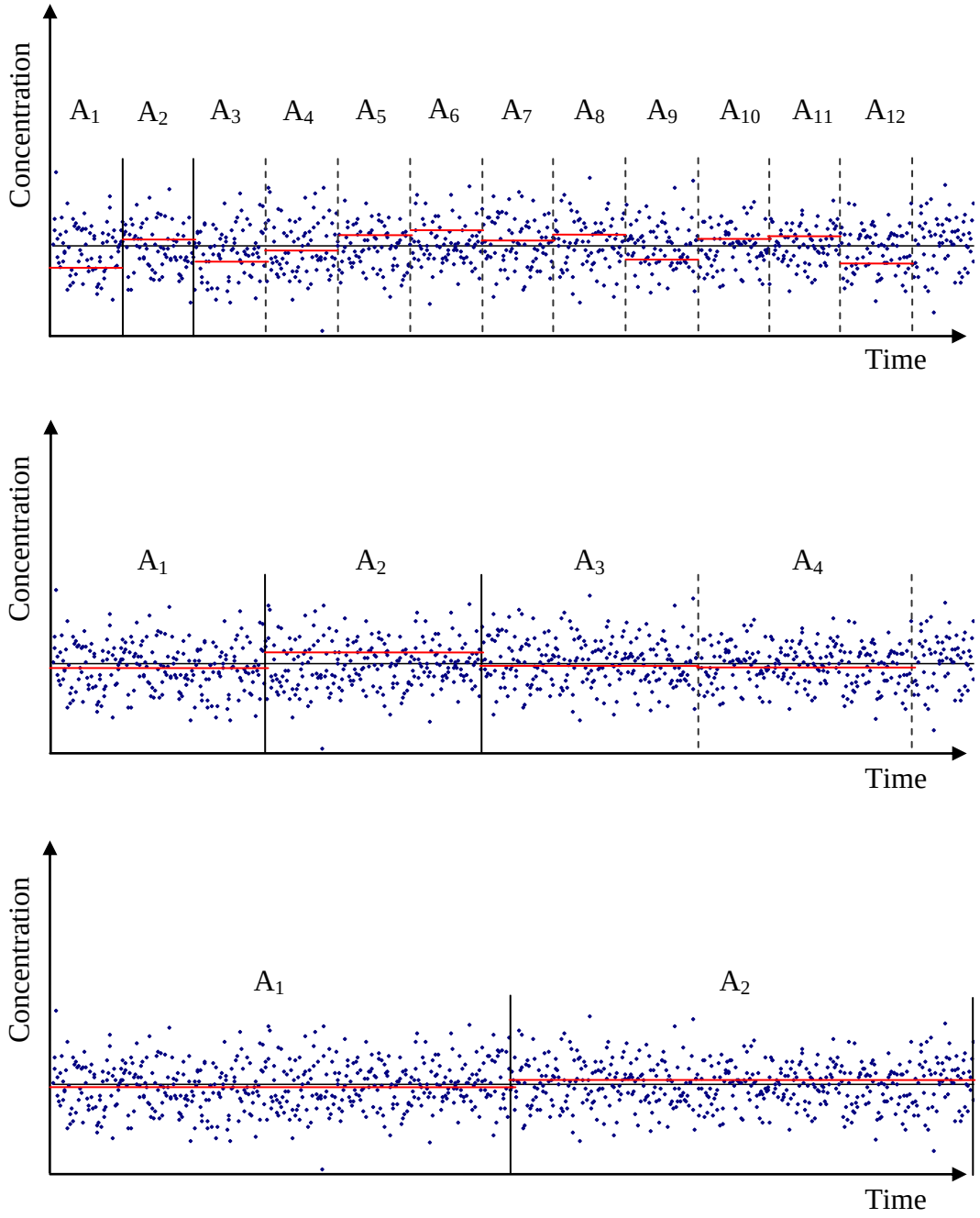


Figure 4-47: Schematic diagram illustrating the application of Allan variance to the same concentration versus time data - using different data bin sizes for averaging. The noise follows a Gaussian distribution about a constant concentration value. The calculated bin means (in red) are also Gaussian distributed about the constant concentration. As the number of data points k in a bin grows the standard deviation of the mean grows smaller in proportion as sqrt(k). Therefore, for random noise, the Allan variance – based on the difference between the averages for a pair of adjacent bins – follows a reducing trend with increasing bin size, as the averages converge to the true concentration.

Allan variance, the difference of adjacent averages, therefore decreases as the bin size increases. (The suffix s on the average A_S keeps track of which bin is being averaged).

Figure 4-47 shows a schematic diagram illustrating this idea. Each plot contains the same series of concentration readings. The dark vertical lines delineate the first pair of adjacent bins, from which an Allan variance can be calculated. However, since calculating the Allan variance, for this bin size, from an average of all the adjacent pairs of bins, gives a more accurate estimate, this average is made. (This does mean, however, that as the bin size of k points grows, for a given sized data set of N points, there are less total pairs of adjacent bins to average).

The time averaged Allan variance for a bin size k is given by the formula:

$$\left\langle \sigma_A^2(k) \right\rangle_t = \frac{1}{2m} \sum_{s=1}^{m} \left[A_{s+1}(k) - A_s(k) \right]^2 \tag{4-3}$$

where

$$A_s(k) = \frac{1}{k} \sum_{l=1}^{k} x_{[(s-1)k]+l} \tag{4-4}$$

Here, Equation 4-4 is the average of the k concentration values, x, in the bin denoted by s. The indices on the x term can be explained as follows. There are N values of concentration in the whole continuous time series of data, x_1 to x_N. $[(s-1)k]$ gives the first concentration index in a given bin, s, and the index l locates the concentration value within that bin. For instance, if the average is over the first bin of, say, $k = 10$ points, $s = 1$ and the first concentration in the bin is at $l = 1$, $x_{[(1-1)\times 10]+1} = x_1$. Similarly, the 5th concentration value in the 3rd bin is $x_{[(3-1)\times 10]+5} = x_{25}$, the 25th concentration data point in the time series.

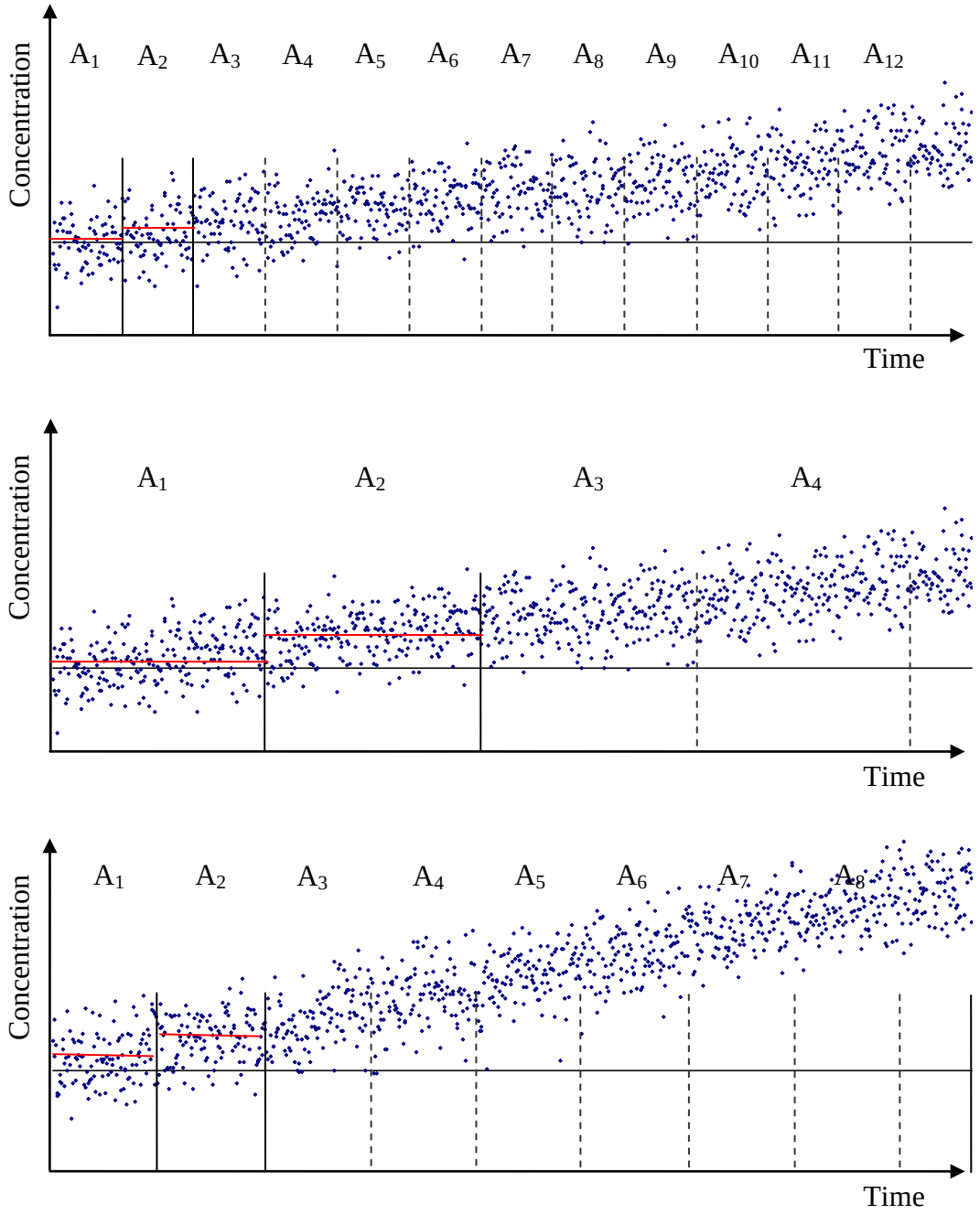


Figure 4-48: **Schematic diagram illustrating the application of Allan variance to Gaussian noise and exaggerated linear drift. The top two diagrams show the same data set. In the top plot the bin size is relatively small compared to the drift. Whether the averages of adjacent bins (first two in red) converge or diverge with increasing bin size (near this bin size), depends on the relative trend of Gaussian noise reduction with averaging to the trend of divergence of the underlying concentration value itself with drift. In the second plot, the bin size is large enough to be dominated by drift, so the concentration averages in adjacent bins are divergent, so that the Allan variance based on the difference between them will increase. The bottom plot shows a data set with twice the rate of linear drift – implying the averages, and hence Allan variance, will be divergent at a smaller bin size.**

The summation term in Equation 4-3 calculates the difference in averages in an adjacent pair of bins, the value of which is expected to grow smaller with more averaging – that is, a greater bin size. The summation averages over pairs of adjacent bins to get a more accurate estimate. Plotting the time averaged Allan variance $\left\langle \sigma_A^2(k) \right\rangle_t$ versus bin size k, should therefore reveal a decreasing trend, when the signal is subject only to random noise.

However, the signal, and the resulting concentration value, will also usually be subject to longer term drifts. For instance, the signal may be distorted by unwanted etalon fringes moving through the spectrum if the spacing of optical components changes – due, for instance, to vibrations or changes in temperature or pressure. The laser's wavenumber tuning may also drift in the longer term, placing the absorption in varying positions during a given number of spectral sweeps. The composition or environment of the gas itself may change over time, due to changing environmental conditions or leaks.

So, when the signal is subjected to drift, the concentration values would not be expected to converge to a steady underlying value in the longer term. Figure 4-48 illustrates this idea for a data set subject to an exaggerated linear drift. The average of adjacent bins do not converge, and therefore their difference – the Allan variance – does not converge. The quicker (steeper) the drift, and the larger the bin size, the more divergent the bin averages are.

The technique of Allan variance seeks to find where these competing trends are equal. That is, where the convergence of concentration values due to averaging out random noise is about to be overcome by the divergent effect of drift. This marks the optimum number of averages and integration time for a particular system. A plot of $\log\left(\left\langle \sigma_A^2(k) \right\rangle_t\right)$ versus $\log(k)$ makes the change from one trend to the other stand out sharply and produces the 'Allan-plot'.

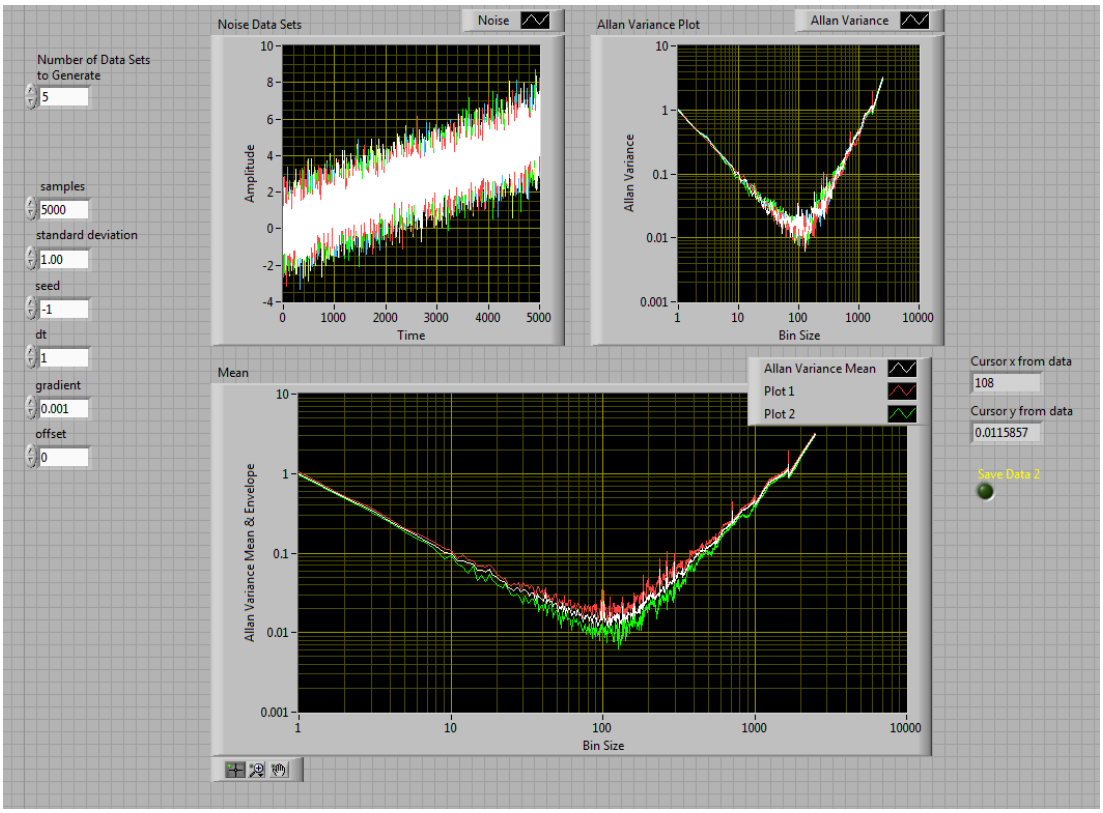


Figure 4-49: **Screenshot of the Allan variance program. In this version, Gaussian noise and linear drift is simulated for multiple data sets. The top two panels show the sets and their Allan variances. The bottom plot shows the average of the data sets' Allan variances plotted against the bin size. The envelope of the data sets' Allan variances is also shown by the green and red traces.**

In a second, similar version, multiple experimental data sets can be loaded automatically from a given folder, and their Allan variances calculated and plotted in the same way as above.

A screenshot of the LabVIEW program for automating the lengthy procedure of the Allan variance calculation is shown in Figure 4-49. In the version shown, a number of continuous data sets can be generated consisting of user defined levels of LabVIEW generated random (Gaussian white) noise riding on a linear drift. If only one time series is used, a single Allan variance plot is produced. If more than one time series is generated, the time series and Allan variances are displayed in the top two panels. The point by point bin size average of the Allan variances is displayed in the lower panel along with their envelope.

A second version of this program allows single or multiple sets of experimental continuous data to be input. It does this by using LabVIEW's 'Folder Path', 'List Folder', and 'Concatenate Strings' functions to load each data set, within a given folder, sequentially into a custom Allan variance routine. The Allan variance for each bin size, for each data set, exits the loop when all the sets have been processed. From there, the inner loop of a nested For Loop averages all the Allan variances, $\left\langle \sigma_A^2(k) \right\rangle_t$, collected from the different data sets, corresponding to the same bin size k. The maximum and minimum Allan variances for this k are also recorded. The outer loop steps through all of the bin sizes. The resulting data is output to a graph of the Allan variance mean and envelope as a function of k.

The custom written Allan variance routine itself, calculates the Allan variance for all bin sizes for $k=1$ to $k=\lfloor N/2 \rfloor$ - for a particular set of concentration versus time data from the folder. The last k value implies that the largest bin size for an even data set is found by dividing the number of data points in two. If the set contains an odd number of points, the routine removes the last data point from the array - making an even array. In terms of the original N points, the maximum bin size is found by 'flooring' $N/2$ to the highest preceding integer. A nested For Loop then calculates the Allan variance. The outer loop iterates from $k=1$ to $k=\lfloor N/2 \rfloor$. For each particular k value, the inner loop calculates the average Allan variance implied by Equation 4-3. The number of inner loop iterations is found by dividing the floor of

N, that is $\lfloor N \rfloor$, by twice the bin size, $2k$. This is because two adjacent bins are needed to form each summand value. If this quotient is not an integer, then $2k$ times the fractional remainder of concentration data points go unused, for that set of bin size inner loop iterations.

So by the end of the loop the whole summation and average is completed for that bin size, giving a single $\left\langle \sigma_A^2(k) \right\rangle_t$ value for that particular k (Equation 4-3). The outer loop repeats the calculation for all the possible bin sizes, and forms the array $\left\langle \sigma_A^2(k) \right\rangle_t$ against k. The routine then takes in the next series of concentration versus time data from the folder.

4.9 Conclusions

The cw-QCL spectrometer had a limit of detection of 0.080 absorbance units in an integration time of 51ms. This is comparable to the performance of Cascade Technologies' pulsed-QCL spectrometers at an earlier stage of their development. However, this performance is some way behind their current specifications. The CT3000 Cascade Technologies spectrometer used in Chapter 7 has a limit of detection of ≈0.004 absorbance units in 10ms.

By increasing the effective number of averages to 7680, the cw-QCL spectrometer's limit of detection became 0.02 absorbance units in 0.77s. Further improvement seems possible by increasing the number of averages further, and by optimizing the power falling on the detector.

A number of difficulties were found with AC-coupled detection. To record V_0 for use in the Beer-Lambert law, a sub-threshold noise level was included in the sweep. This reduced the laser power over an absorption feature, degrading SNR. The spectrometer also had to be operated at, at least, 10kHz to avoid transimpedance amplifier time constant effects. This reduced the spectral wavenumber coverage per

scan. Since the chirp rate is one contribution to the wavenumber resolution, the associated faster tuning rates have the potential to degrade resolution, although they allow more noise reduction by averaging within a given maximum integration time.

However, in the cw-QCL spectrometer of this chapter the chirp rate is typically 10^3 times slower than the pulsed-QCL spectrometers. The system's resolution is therefore limited by the Doppler widths of the transitions themselves - further broadened by a symmetrical Gaussian broadening. The practical resolution is therefore line, gas, temperature and pressure specific. By eliminating other possibilities, the majority contribution to the additional broadening is attributed to the system's electronics. The resolution is estimated to be $0.0058cm^{-1}$ for the $^{12}CH_4$ transition $\upsilon 4$ $P(4)A_2$ $1 \leftarrow A_1$ 1 transition at $1283.458791cm^{-1}$ at low pressure, and $0.0041cm^{-1}$ for the $^{14}N_2{}^{16}O$ P9f and P9e transitions at $1283.824000cm^{-1}$ and $1283.829850cm^{-1}$ at low pressure.

The greatest AC-coupling difficulty, in terms of noise, may be the seeming electronic incompatibility of the detector-transimpedance amplifier with the oscilloscope used. Chapter 5, on wavelength modulation spectroscopy, will show that this incompatibility was adding random noise. It is therefore likely that the sweep integration noise levels are also worse than need be.

The cw-QCL's response to sawtooth voltage ramps of various frequencies was investigated. It was found that the temperature tuning of the laser (by injection current or substrate temperature) was less and less able to follow the voltage waveform as this waveform's frequency increased. Spectral coverage was reduced from both wavenumber ends as the Peltier cooler oscillated the substrate temperature round a steady state temperature set more by the injection current, rather than its own set-point temperature.

The cw-QCL's tuning response to symmetric triangular and sinusoidal voltage waveforms was found to be asymmetrical – with spectral coverage and tuning rates

smaller on the down ramp. This is attributed to the asymmetry of the heating – cooling cycle, with the injection current heating the laser's active layer directly and the Peltier cooler operating by means of the temperature of the laser's substrate.

4.10 Suggestions for further work

A faster beam chopping system would allow the rectification approach to be pursued. This would negate the need for an approximately 0V sub-threshold in a sweep. This would allow a higher laser power and enhanced SNR. Likewise, rearranging the optical set-up would allow optimization of the laser power falling on the detector. If a DC-coupled detecting system became available, this would be better suited to a continuous wave laser – allowing greater spectral coverage and obviating the need to include the threshold baseline in a scan. Excluding 1/f noise might, however, become an issue.

An independent digitizer would allow more averages to be taken for enhanced SNR, at the expense of increased integration time. The LabVIEW data analysis programme of Chapter 3 was designed for pre-recorded data, and Cascade Technologies' real time systems were tied to their own pulsed sensors. If software became available, however, a continuous time series for a constant parameter – such as concentration – could be recorded for Allan Variance analysis with the software prepared in Section 4.8. This would give an indication of the maximum integration time for random noise reduction, after which long term drifts degrade the SNR.

Resolving the issue of incompatibility between the detector-transimpedance amplifier electronics and the oscilloscope might improve the SNR. Chapter 5 will show that this added much random noise to wavelength modulation spectroscopy measurements.

Chapter 5

Wavelength Modulation Spectroscopy with a continuous wave quantum cascade laser

5 Wavelength Modulation Spectroscopy with a continuous wave quantum cascade laser

5.1 Abstract

An electrical arrangement for Wavelength Modulation Spectroscopy (WMS) was set up. This was used to compare the absorbance sensitivity of this technique to that of Sweep Integration (SI), using the cw-QCL of Chapter 4. The sweep integrations used 512 averages. The WMS signal was also averaged 512 times – due to high levels of noise in the signal chain. A linear relationship between sweep integration and WMS peak heights was established, for $^{12}CH_4$ at low pressure. This relation was used to calibrate the WMS noise floor from voltage to absorbance units and hence establish the absorbance Limit of Detection for the system. The WMS technique combined with 512 averages had a modestly improved sensitivity over that of sweep integration with 512 scans - of 4.7 times at 0.017 absorbance units. This came at the expense of the need for calibration, a more complex electronic setup, and a 2000 fold increase in integration time to 102.4s. Had the WMS averaging not been necessary, the increase in integration time would have been a modest 3.9 times. In addition, 2f-WMS lineshape models were coded into LabVIEW, and these models were shown to be useful in the context of WMS using cw-QCL's.

5.2 Introduction

Lower noise in a measurement technique allows lower gas concentrations to be discerned – hence, providing process optimization on a finer scale, competitive advantage, and some future proofing of a system against the tightening of legislation.

Wavelength Modulation Spectroscopy (WMS) predates Sweep Integration as a practical noise reduction technique. WMS modulates an experimental signal – here using a function generator to add a sinusoidal current (hence wavenumber oscillation) to the cw-QCL laser. This allows frequency selective filtering of the signal from surrounding noise by phase sensitive detection (Section 5.3). However,

the derivative-like signals which result need to be related back to their direct absorption equivalents in order to extract concentration values.

The repeated averaging of the Sweep Integration technique (Chapter 4) has gained ground on the WMS technique with ongoing improvements in fast processing and digitization, and fast, high bandwidth detectors – allowing a $\sqrt{N}$ SNR improvement against random noise before longer term drifts negate the noise reduction. Nevertheless, for diode lasers, WMS is still a widely used technique (e.g. Bain et al., 2011). The purpose of this chapter is to establish and quantify if this WMS advantage also pertains to the cw-QCL spectrometer built here.

A qualitative description of the WMS technique is given. Four freely available online resources contributed to this understanding. Dyroff (2008) provides a useful account and includes a diagram similar to the middle plot of Figure 5-6. This is a key diagram for visualizing the technique. A prerequisite for a good understanding of WMS is an understanding of the lock-in amplifier. Armen (2008) sketches a diagram showing how the lock-in's DC output voltage can be positive or negative - due to the phase difference between signal sinusoid and reference sinusoid. A fuller diagram is given in this work as Figure 5-4. The top plot of Figure 5-1 was inspired by a similar noise diagram found in technical notes (Bentham Instruments Ltd). The fourth online resource gives useful lock-in amplifier notes, and notes on general noise sources in opto-electronic experiments (Stanford Research Systems).

A 2f-WMS (2^{nd} derivative-like) lineshape model from literature (Li et al., 2006) was coded into LabVIEW and is presented in Section 5.9. This model is based on harmonic analysis beyond the derivative limit.

The experimental setup is described (Section 5.4), as is the electronic noise in the signal chain (Section 5.7). The sensitivity comparison of the sweep integration technique to the WMS technique is made in Section 5.8. As was discussed in Section 4.6.1.4, there was an oscilloscope scaling problem – most likely from some

electronic incompatibility between the AC-coupled transimpedance amplifier and the other equipment. This might also contribute to the high amplitude low frequency noise seen in the lock-in signal (Section 5.7). Unfortunately, tenure over bench space and the cw-QCL laser and its drive electronics ended before this could be investigated properly. The result here was that the WMS technique needed to be combined with averaging to produce a usable signal.

Nevertheless, the combined WMS and averaging technique was found to have a modest sensitivity advantage on the Sweep Integration technique as used in Chapter 4. However, this comes at the expense of the need for phase sensitive detection and calibration. The integration time is also much larger – mainly due to the averaging of the slow tuning ramp used for WMS.

Had there been a way of taking continuous concentration values, a more thorough Allan variance analysis could have been undertaken (using the program of Section 4.8) with each technique evaluated at its optimum performance for the system. Nevertheless, the approach taken was to compare the techniques using convenient, usable parameters on the equipment to hand – such as the maximum oscilloscope averaging of 512 averages.

The AC-coupling may not be a problem in itself for WMS, because the laser radiation is modulated at a relatively high frequency – here at 10kHz, which in Chapter 4 was relatively free from transimpedance amplifier distortion. However, this is a relatively low frequency for the VIGO detector to handle, and this may therefore create additional noise (Section 5.7).

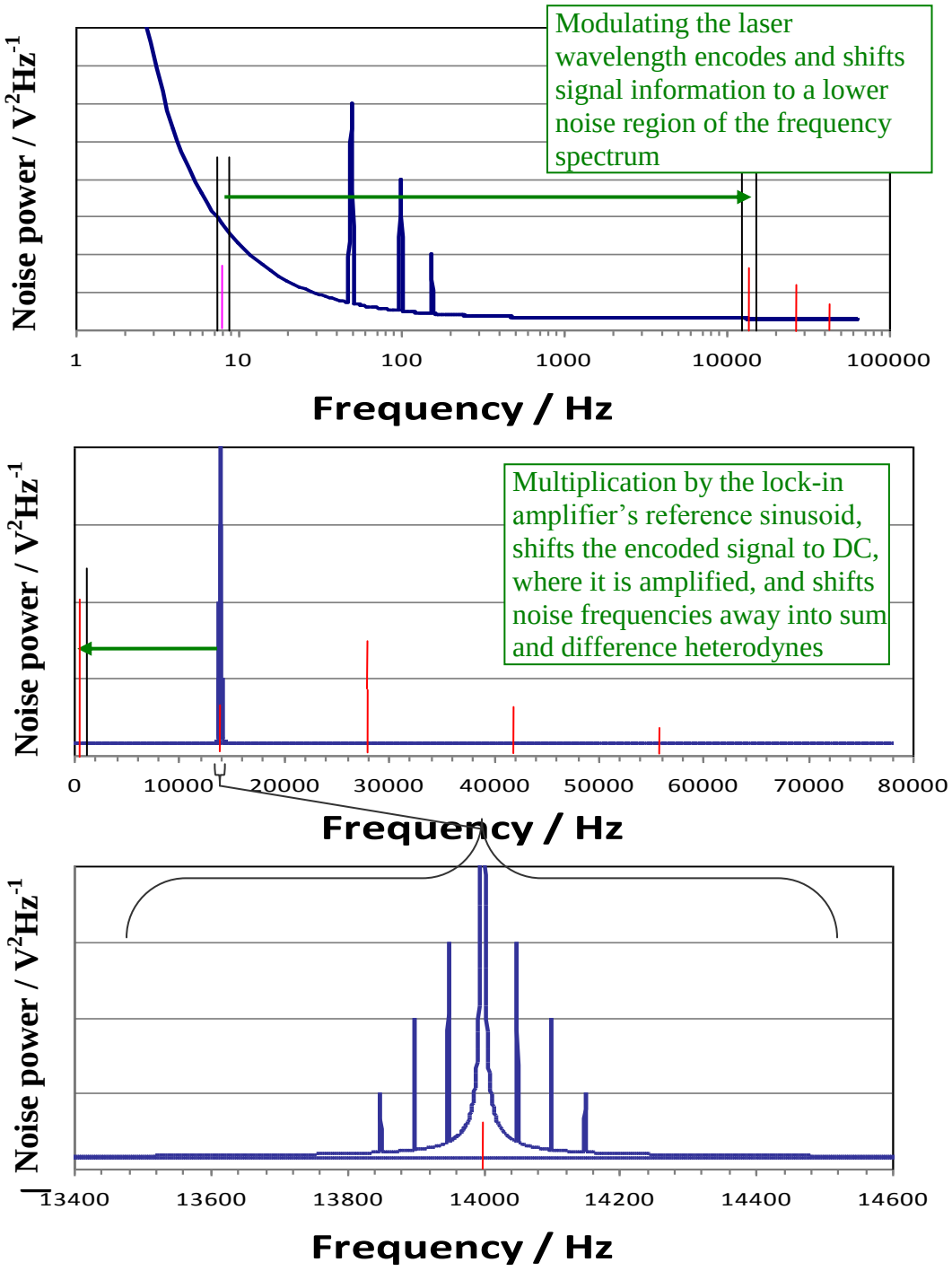


Figure 5-1: **Schematic diagram of frequency shifting with the lock-in amplifier - for noise reduction. Top plot: Modulating the experiment moves signal information, so that the information occurs at the modulation frequency and its harmonics - in a lower noise part of the spectrum. Middle plot: Multiplying with a reference sinusoid shifts the selected information to DC for low pass filtering, and noise frequencies are shifted into pairs of sum and difference AC heterodynes - centered on the greater of the original noise frequency or the reference frequency.**

5.3 The Wavelength Modulation Spectroscopy Technique

5.3.1 Encoding and shifting signal information to lower noise frequencies

The top plot of Figure 5-1 shows a schematic of a frequency noise spectrum on a logarithmic abscissa. For example, 1/f noise is illustrated along with noise spikes for transformer and mains noise. The plot indicates the difficulty of making a measurement of a low frequency signal – here indicated by one of its significant Fourier components, nominally at 8Hz. Such a low frequency signal might arise from slow tuning over an absorption feature – as in the sweep integration technique for an 8Hz ramp frequency. Periodically scanning over the absorption signal will produce a detected time signal with a fundamental Fourier component at the scanning frequency - and its harmonics.

Without any filtering, all the noise frequencies within the bandwidth of the detection system will be present at any given measurement time - swamping the signal. In Figure 5-1, electronic filtering about the low frequency signal (represented by the black vertical lines) also does not reveal the signal – since the noise within the bandpass filter has greater amplitude. (Although averaging through the sweep integration technique of Chapter 4 might help).

What is needed is to move the signal information to a lower noise region of the spectrum. In absorption spectroscopy, if resolution can be sacrificed, this shift can be achieved by using a faster scanning frequency – increasing the Fourier components of the detected direct absorption signal. However, this requires fast digitizers, fast processors and fast detectors of high bandwidth (to capture the higher frequency components). These are conditions met in the sweep integration averaging method of Chapter 4. This method has been gaining ground as a noise reduction technique as the technology of electronics and computing has advanced. However, the older method of WMS is still often found to have a noise reduction advantage – though the

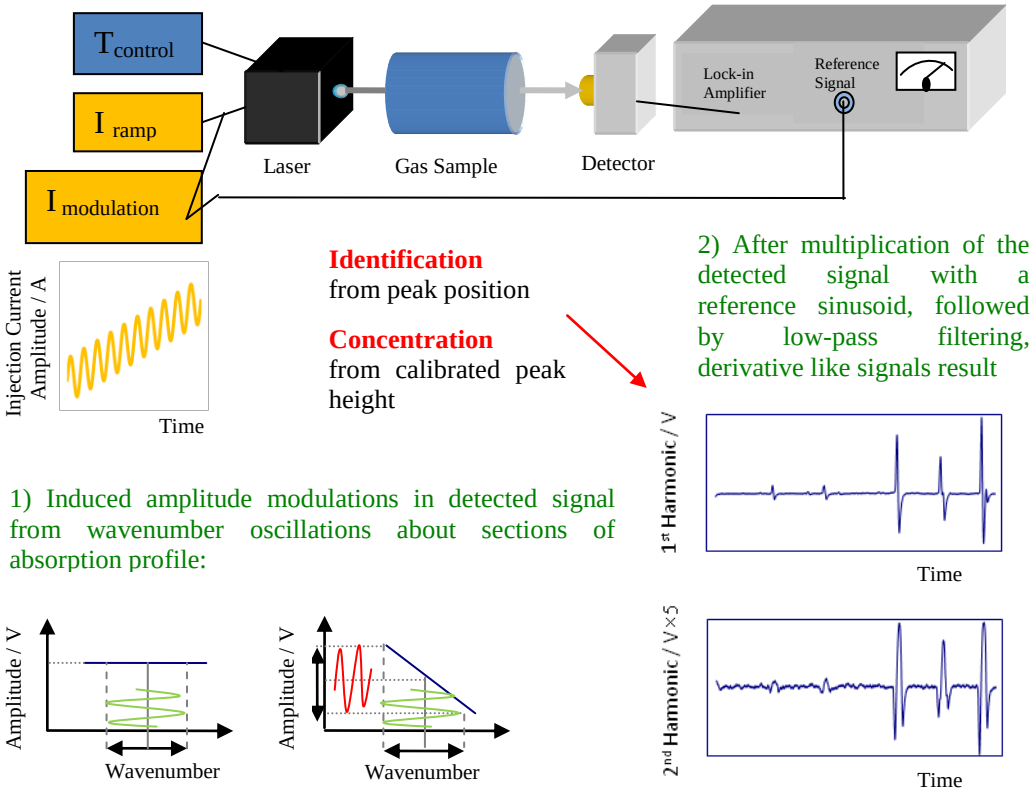


Figure 5-2: Schematic diagram of the wavelength modulation spectroscopy (WMS) technique.

information it reveals, in its typical and simplest implementation, is indirect and needs calibration.

In the WMS technique, as used in this work, the laser's radiation is sinusoidally modulated and signal information is encoded and shifted to the modulation frequency and its harmonics. This occurs through a conversion of the laser's wavenumber modulation to an amplitude modulation in the detected signal - as an absorption profile is tuned over. How absorption information is encoded into the amplitude modulation at the detector is taken up in Section 5.3.3 below.

5.3.2 Extracting the signal information via multiplication and low pass filtering

Modulating the laser's current while tuning produces signal information at the modulation frequency and its harmonics in the detected time signal (Section 5.3.3). This was shown schematically in the top plot of Figure 5-1 by means of an original signal frequency component at a nominal current ramp scanning frequency of 8Hz, and the amplitude modulated components at a nominal modulation frequency of 14kHz, plus harmonics at 28kHz and 42kHz (etc.). Filtering at these higher frequencies would now reveal signal information, although a typical electronic filter could not compensate satisfactorily for drifts in the modulation frequency.

A more satisfactory method of extracting the signal information is to use a lock-in amplifier. The lock-in amplifier is essentially a very low noise, frequency selective, AC-voltmeter - returning a DC signal proportional to the amplitude of a selected frequency component within the received signal. Figure 5-2 shows a schematic of the WMS technique. As in the Sweep Integration technique of Chapter 4, a sawtooth current ramp is applied to the cw-QCL to tune across an absorption feature. However, this tuning ramp is relatively slow and a relatively fast sinusoidal current modulation is added - its modulation frequency, ω_{mod}, is also fed to the lock-in

amplifier as the lock-in's external reference frequency. From this, the lock-in produces an internal reference sinusoid at the angular frequency $\omega_{ref} = n\omega_{\mathrm{mod}}$, $n = 1,2,3,\ldots$. The lock-in therefore receives the modulation frequency directly and also the detected signal containing this frequency and its harmonics. In this way the lock-in can stay locked to the signal's frequencies even if the modulation frequency drifts.

The operation of the lock-in amplifier can be thought of as being a *multiplier* followed by a very frequency selective *low pass filter*. And this is how the signal information is extracted. The lock-in produces a reference sine wave of amplitude V_{ref} and adjustable phase θ_{ref}, at the modulation frequency or one of its harmonics $\omega_{ref} = n\omega_{\mathrm{mod}}$, $n = 1,2,3,\ldots$. By multiplying the detected signal by the reference sinusoid, every Fourier component of the WMS signal, at amplitude V_{comp}, angular frequency ω_{comp}, and phase θ_{comp}, becomes a sum and difference heterodyne:

$$\begin{aligned} V_{multiplier} &= \frac{GV_{comp}V_{ref}}{2}\cos\left(\omega_{ref}t + \theta_{ref}\right)\cos\left(\omega_{comp}t + \theta_{comp}\right) \\ &= \frac{GV_{comp}V_{ref}}{2}\cos\left(\left[\omega_{ref} - \omega_{comp}\right]\mathrm{t} + \theta_{comp} - \theta_{ref}\right) + \\ &+ \frac{GV_{comp}V_{ref}}{2}\cos\left(\left[\omega_{ref} + \omega_{comp}\right]\mathrm{t} + \theta_{comp} + \theta_{ref}\right) \end{aligned} \tag{5-1}$$

If the lock-in returns the component's root mean squared value, as many do, then the substitution $2 \to \sqrt{2}$ can be made. Alternatively, this reduction can be accounted for within the factor G which represents any gain added by the lock-in amplifier.

It can be noted that any component that was originally DC will be shifted by this multiplication to the reference frequency and filtered out. Further, for each originally AC component, multiplication splits the amplitude at a given frequency ω_{comp} into two half amplitudes oscillating at $\left(\left|\omega_{comp} - \omega_{ref}\right|\right)$ and $\left(\omega_{comp} + \omega_{ref}\right)$. These can be

considered to be equally spaced about the larger of ω_{comp} or ω_{ref} (bottom two plots of Figure 5-1). The mid-point angular frequency is given by:

$$\begin{aligned}
&\left(\left|\omega_{comp}-\omega_{ref}\right|\right)+\frac{\left(\omega_{comp}+\omega_{ref}\right)-\left(\left|\omega_{comp}-\omega_{ref}\right|\right)}{2} \\
&\quad=\omega_{comp} && \text{if} \quad \omega_{comp}>\omega_{ref} \\
&\quad=\omega_{ref} && \text{if} \quad \omega_{comp}<\omega_{ref} \\
&\quad=\omega_{comp}=\omega_{ref} && \text{if} \quad \omega_{comp}=\omega_{ref}
\end{aligned} \tag{5-2}$$

The middle and bottom plots of Figure 5-1 illustrated the bunching of low frequency noise about the modulation frequency - after multiplication by a reference sinusoid at this frequency (here 14kHz).

In general, each signal component's heterodynes are AC and can be filtered out by the low pass filter – giving no lock-in output. However, the difference term of the component which matches the reference sinusoid frequency is DC, and will pass through the low pass filter – along with some low frequency noise within the filter's bandwidth set by its time constant. Ideally, the filter's time constant is set just short enough to allow the capacitive part of the circuit to respond to changes in the wanted DC signal as the centre wavenumber of the modulation is scanned over the absorption profile. Higher frequencies reverse in sign too quickly for a significant capacitive voltage to build-up, and hence the filter circuit does not pass these signals. From Equation 5-1 with $\omega_{ref}=\omega_{comp}$ and after filtering:

$$V_{lock-in}=\frac{GV_{comp}V_{ref}}{2}\cos\left(\theta_{comp}-\theta_{ref}\right) \tag{5-3}$$

Since the reference phase is adjustable, it can be made equal to the component sinusoid's phase in order to maximize the signal.

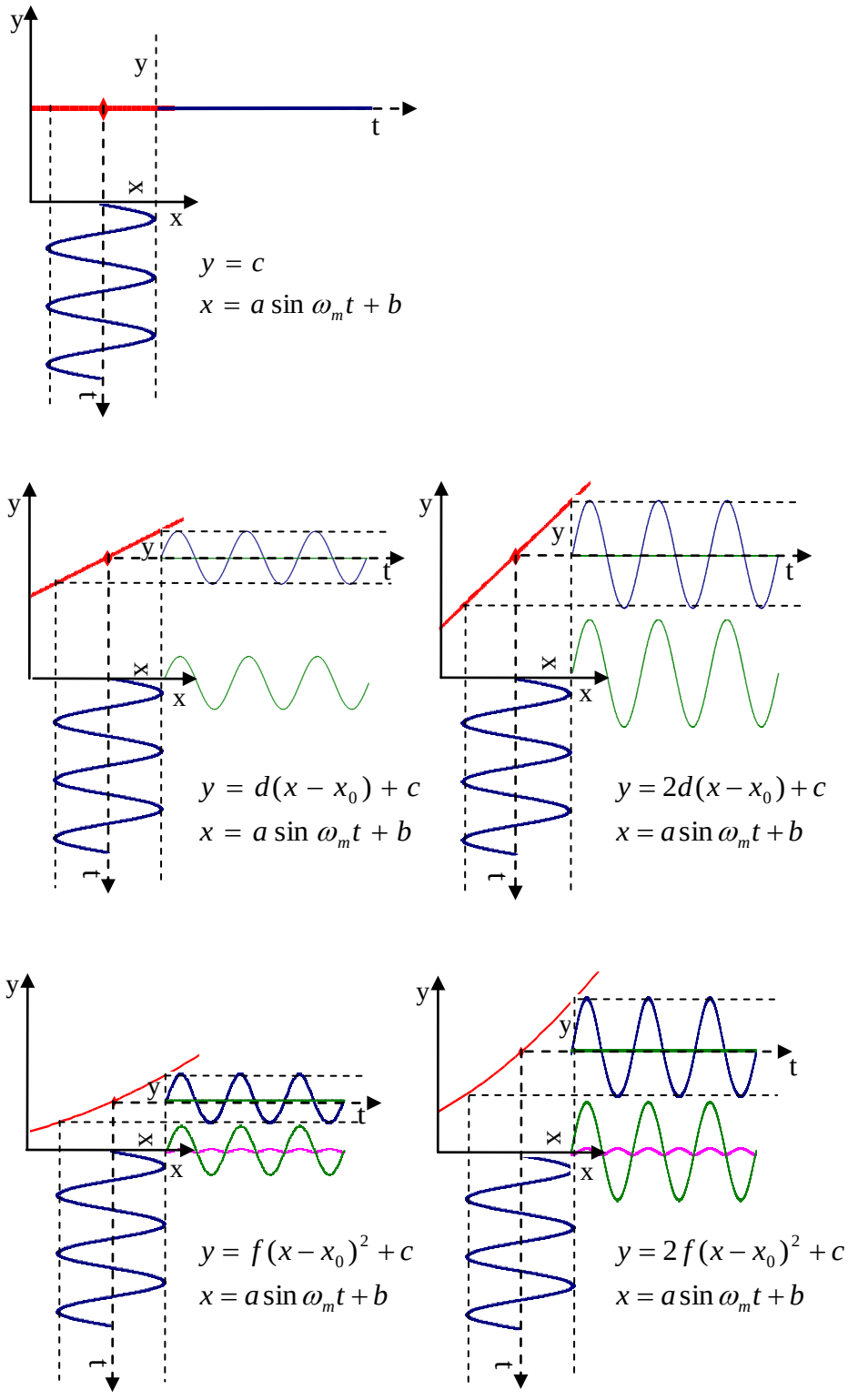


Figure 5-3: **A schematic diagram illustrating wavenumber oscillations (at bottom against x-axis) about a point on constant, linear and quadratic sections of an absorption profile (red traces). The detected amplitude oscillations are shown on the y-axis in blue. The constant and 1f Fourier components of the detected signal are in green, and the 2f component is in pink. (Simultaneous laser intensity oscillations are neglected).**

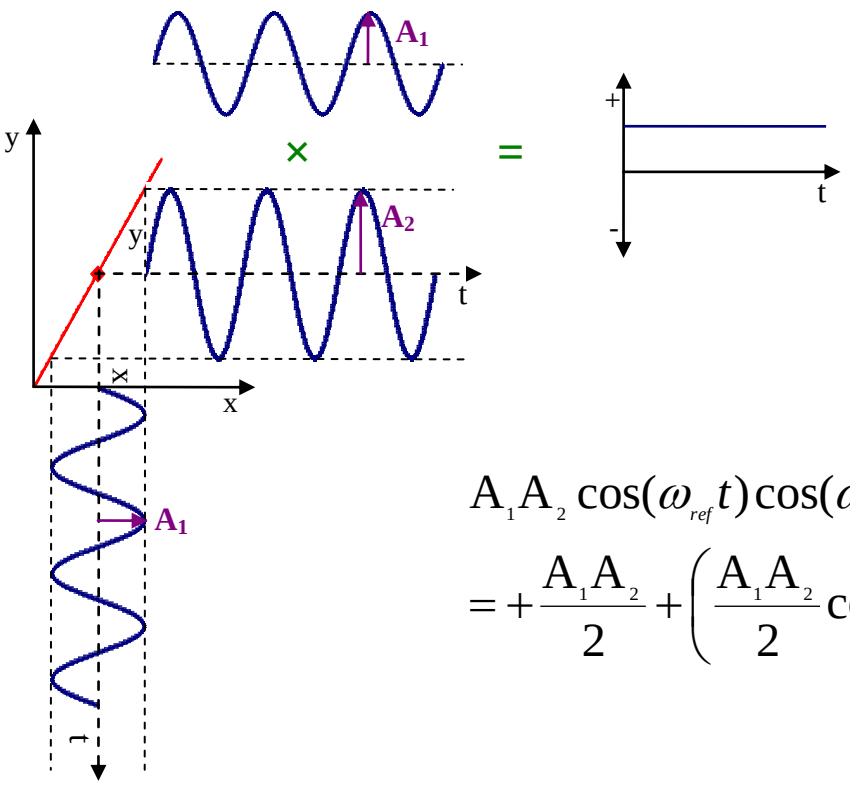
5.3.3 Encoding: How wavelength modulation spectroscopy creates detected signals containing absorption line profile information

The keys to understanding the WMS technique are Figure 5-3 and Figure 5-6. The technique is easier to visualize starting from the assumption of a significantly higher frequency sinusoidal current modulation than the current ramp tuning frequency. With this assumption, as each wavenumber point on an absorption profile is tuned over, it becomes associated with several cycles of the wavenumber modulation.

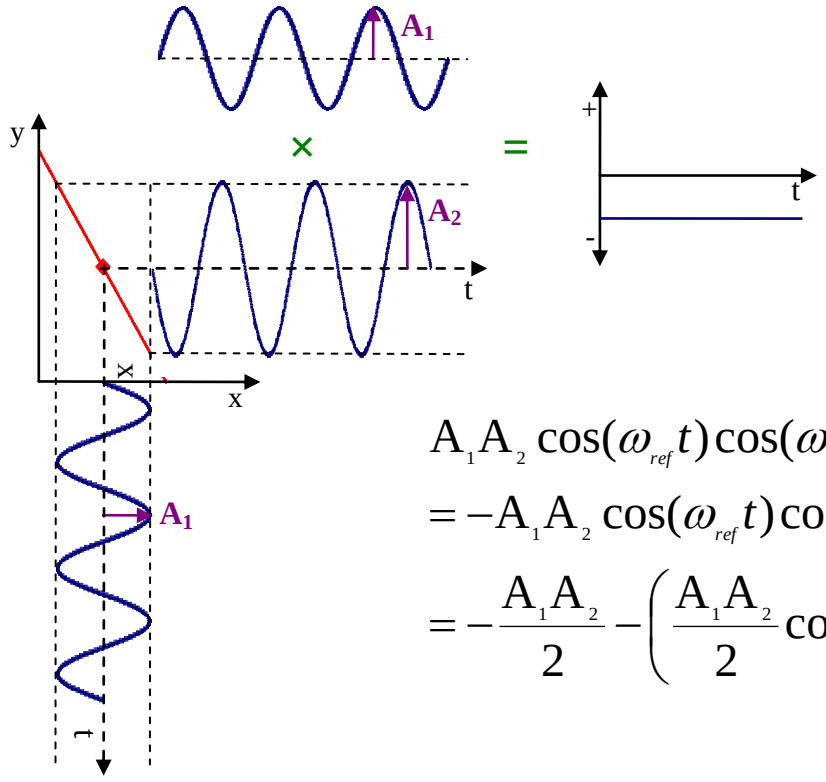
Figure 5-3 (middle plot) shows this oscillating wavenumber tuning on the x-axis when a hypothetical straight line section of absorption profile is encountered. (Simultaneous laser intensity oscillations are neglected). For sections like these, the wavenumber oscillation becomes an amplitude oscillation in the detected signal at the modulation frequency (shown on the y-axis). It can be seen that the amplitude of these detected oscillations is proportional to the gradient of the absorption profile about that point. In this way, information about the direct absorption profile's shape – in this case its gradient about a point - is being encoded into the amplitude of the detected oscillations.

In the derivative limit of very small wavenumber oscillations, $\Delta\tilde{\nu} = a\cos\omega_m t$ with $a << 1$, a Taylor expansion about a point, $\tilde{\nu}$, on the absorption profile confirms that the detected signal has a component at the modulation frequency - 1f – the component's amplitude being proportional to the gradient (1st derivative) at that particular point (and to any higher derivatives of odd order):

$$\begin{aligned} A(\tilde{\nu}+\Delta\tilde{\nu}) &= A(\tilde{\nu}) + \left.\frac{dA(\tilde{\nu})}{d\tilde{\nu}}\right|_{\tilde{\nu}} (\Delta\tilde{\nu}) + \frac{1}{2!}\left.\frac{d^2A(\tilde{\nu})}{d\tilde{\nu}^2}\right|_{\tilde{\nu}} (\Delta\tilde{\nu})^2 + \ldots \\ &= A(\tilde{\nu}) + \left.\frac{dA(\tilde{\nu})}{d\tilde{\nu}}\right|_{\tilde{\nu}} a\cos\omega_m t + \frac{1}{2!}\left.\frac{d^2A(\tilde{\nu})}{d\tilde{\nu}^2}\right|_{\tilde{\nu}} \frac{1}{2}a^2(1+\cos 2\omega_m t) + \ldots \end{aligned} \tag{5-4}$$



$$A_1 A_2 \cos(\omega_{ref} t)\cos(\omega_{ref} t)$$
$$= + \frac{A_1 A_2}{2} + \left(\frac{A_1 A_2}{2} \cos(2\omega_{ref} t) \right)$$



$$A_1 A_2 \cos(\omega_{ref} t)\cos(\omega_{ref} t + \pi)$$
$$= -A_1 A_2 \cos(\omega_{ref} t)\cos(\omega_{ref} t)$$
$$= -\frac{A_1 A_2}{2} - \left(\frac{A_1 A_2}{2} \cos(2\omega_{ref} t) \right)$$

Figure 5-4: **Schematic indicating that the sign associated with the derivatives of the absorption profile affects the phase relationship between the detected signal and the lock-in's reference signal. And whether multiplication of these signals will lead to a positive or negative DC voltage signal after low pass filtering.**

Multiplication of the detected signal by the lock-in's reference sinusoid at this frequency produces a DC signal for that absorption profile point – proportional to the amplitude of the 1f component in the detected signal – and therefore also proportional to the direct absorption profile's gradient at that point. Therefore, as the current ramp tunes across the whole absorption profile, the DC amplitude from the lock-in traces out a signal proportional to the first derivative of the direct line shape.

Similarly, any curvature at an absorption point (2^{nd} derivative) introduces a component at twice the modulation frequency, 2f, into the Taylor expansion and detected signal. By multiplying by a reference sinusoid at twice the modulation frequency this component's amplitude is moved to DC. As the whole profile is scanned over, the lock-in traces out the 2^{nd} derivative of the direct absorption profile, and so on.

So, in the derivative limit, each point on the absorption profile creates a DC signal proportional to the order of the derivative at that point, corresponding to the order of harmonic chosen – 1^{st} derivative if using as reference the modulation frequency (the fundamental frequency or 1^{st} harmonic), 2^{nd} derivative if using the 2^{nd} harmonic, and so on. So overall, as the current ramp tunes over the whole absorption profile, the derivatives of this profile become available as lock-in outputs.

The lock-in amplifier outputs a rectified and therefore (varying) DC signal. Nevertheless, the lock-in's output voltage can be positive or negative as the line profile is scanned. This occurs because the sign of the derivatives of the lineshape leads to a phase shift between the detected sinusoidal Fourier component and the reference sinusoid. Figure 5-4 illustrates this and is shown for the case of the lock-in generated reference being in phase with the laser output intensity. In practice, there is usually an intrinsic phase shift between the laser's current modulation and intensity output. However, this is easily compensated for within the lock-in by adjusting the phase of the generated reference signal, bringing it in phase with the laser intensity output, and hence maximizing the signal (Equation 5-3).

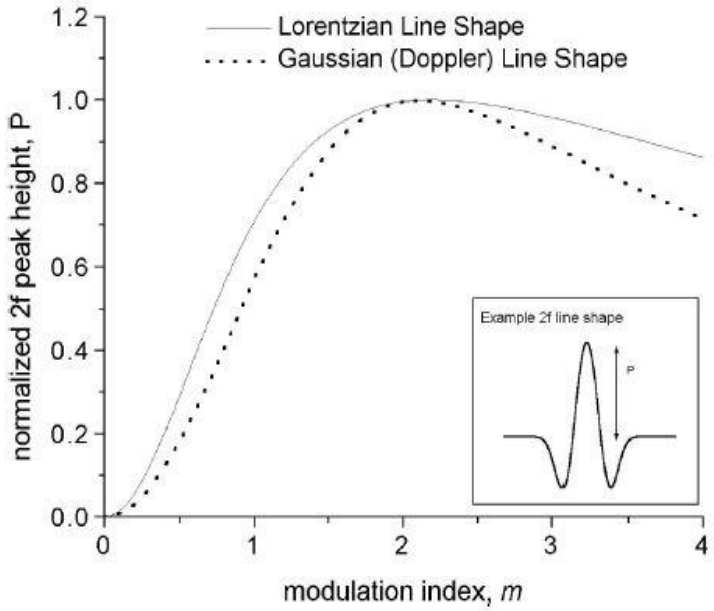


Figure 5-5: **2f-WMS H_2 peak heights versus modulation depth derived from Gaussian and Lorentzian lineshapes via Equation 5-20 for k=2 and a single line. The figure is normalized so that the maximum peak height equals 1 for each lineshape. (Figure from Liu et al., 2004a).**

In practice, it is usually the '2nd derivative' – 2f signal - which is chosen for WMS studies. Modulating a QCL or diode laser via injection current simultaneously modulates both its wavenumber tuning and its output intensity. The lock-in amplifier produces a varying DC signal proportional to the sum of detected intensity oscillations at a given harmonic - due to both the absorption profile and changes in the laser intensity. Since the laser intensity will in general have its largest intensity oscillation at the modulation frequency ($1\omega_m$), 1f-WMS will in general have a large DC background level in addition to that due to the absorption. This background will also be sloped as the laser power increases across the current ramp, and will need to be normalized out in processing. Noise on this large background signal will interfere with the absorption 1f-WMS signal, and may generally require system electronics with a wide dynamic range.

2f-detection may avoid or reduce these problems. If there is a nonlinear component to the laser intensity modulation (oscillating at 2f), this will in general create a background DC level – known as the residual amplitude modulation (RAM) (Equation 5-24) - though this is generally much smaller than the equivalent background signal at 1f. 2f-WMS therefore in general produces much flatter baselines without the need for normalization. Similarly, broad spectral features with small curvature (small 2nd derivatives) are suppressed in 2f-WMS, and 2f detection can provide increased discrimination between closely spaced sharp features. The 2nd derivative is still a relatively strong signal, and it has a maximum amplitude at the same wavenumber as the absorption profile. (In the derivative limit there is also a simple relation between the signal's zero crossings and the direct profile's FWHM, $\omega_g / \sqrt{2\ln 2}$ for a Gaussian, and $\omega_l / \sqrt{3}$ for a Lorentzian).

It might be noted that the sawtooth intensity ramp itself has Fourier components at the ramp repetition frequency and its harmonics. The ramp or modulation frequency

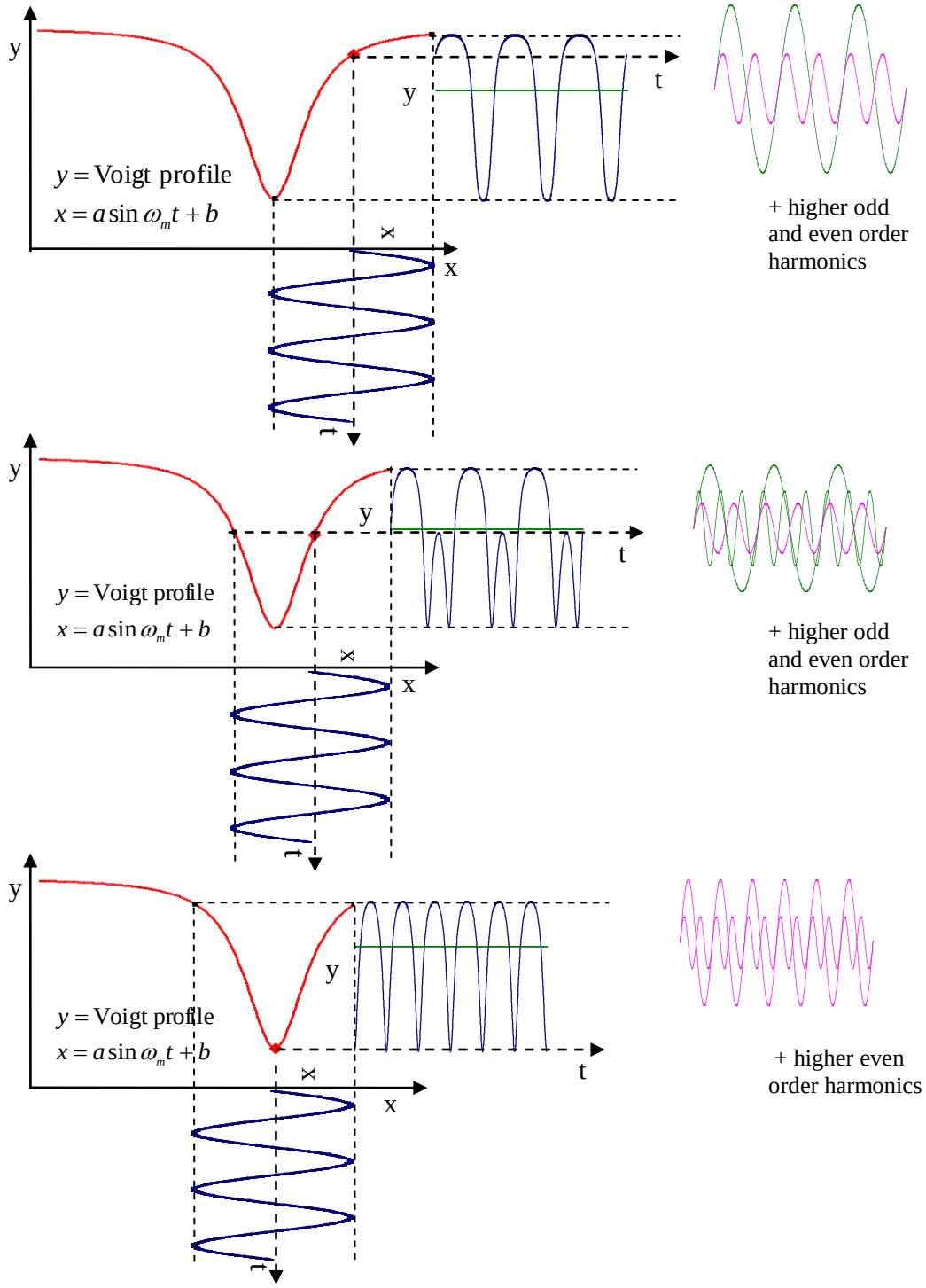


Figure 5-6: **x-axis wavenumber oscillations about different points on a Voigt absorption profile (red traces), leading to detected amplitude oscillations (y-axis). The modulation depth is 2.2. The constant, and some high order Fourier components of the detected signal are shown (DC, and AC odd order components in green, AC even order components in pink. AC components displaced for clarity). The 1f, 2f, etc. signals can be brought to DC and amplified by the lock-in amplifier. As the current ramp tunes across each point in the absorption profile, derivative-like lock-in signals result. (Simultaneous laser intensity oscillations are neglected).**

should therefore be chosen so that a ramp component does not coincide or nearly coincide with the chosen harmonic of the reference frequency. If such coincidence does take place, a further DC level or a beat oscillation will pass through the low pass filter after multiplication.

In practice, the derivative limit must be greatly exceeded in order to maximize the signal. A characteristic dimensionless parameter is the 'modulation depth'. This is the ratio of the modulation amplitude, a, (the maximum wavenumber excursion introduced by the modulation) to the halfwidth, $\omega_{g,l}$, of the particular absorbance lineshape: $m = a/(\omega_{g,l}/2)$. The modulation depth which maximizes the 2f-WMS signal for a Lorentzian lineshape is 2.20 (Reid and Labrie, 1981). Section 5.9 calculates the maximizing modulation depth for a Gaussian lineshape as 2.11. Although far beyond the derivative limit of $m \approx 0$, the lock-in signals retain their derivative like forms qualitatively. Beyond the 2f-WMS signal maximum 'over modulation' occurs and the signal reduces in peak height and its width further broadens as the modulation depth increases. Figure 5-5 shows the 2f-WMS peak height as a function of modulation depth derived for Gaussian and Lorentzian lineshapes (from Liu et al., 2004b). The figure is normalized so that the maximum peak value is 1 for each lineshape. (These trends, and equations which can model them, will be taken up in Section 5.9).

Figure 5-6 shows detected signals arising from a modulation index of ~2.2 for a selection of points on a Voigt profile. (Simultaneous laser intensity oscillations are neglected). The first plot's detected signal has a DC, 1f, 2f, etc. component. The 1f, 2f, and higher harmonic components can be extracted by a lock-in amplifier, even though the overall signal at the detector oscillates at the modulation frequency. It can be seen graphically - as in the middle plot of Figure 5-6 - that the 2f component becomes explicit in the detected signal when the minimum of the absorption profile falls within the modulation envelope. In the lowest plot, centred on the minimum, the 2f signal has a maximum, and there is no 1f component. These observations remain valid for Gaussian and Lorentzian profiles, since they are similarly shaped.

As well as being far beyond the derivative limit, in practice, some asymmetry in the tuning modulation is expected from heating and cooling hysteresis in the cw-QCL. Models also normally assume that the wavenumber, $\tilde{\nu}$, on the absorption profile, about which the modulation, $\Delta\tilde{\nu} = a\cos\omega_m t$, takes place is fixed for at least one modulation cycle. However, if the modulation frequency is not very much greater than the ramp frequency, the ramp will move the centre of oscillation along the absorption profile before one modulation cycle is completed. This is considered further in Section 5.9.2. Nevertheless, again, the lock-in signals retain their derivative-like forms qualitatively – broadening as the modulation depth increases. The peak height of the 2nd derivative (in volts) can still be calibrated with direct signals (in absorbance) of known concentration, in order to extrapolate to concentrations smaller than those where the direct absorption profiles have diminished into noise. This is the methodology followed through experimentally in this chapter.

5.3.4 A note on the difference between using the lock-in amplifier for rectification of a chopped signal and for wavelength modulation spectroscopy

In Section 4.6.1.3 the continuous laser beam was mechanically chopped to form an AC square wave signal which the AC-coupled detector could respond to. The lock-in amplifier was then used to rectify this signal back to DC to take advantage of the lock-in's noise discrimination and for ease of display. The difference between this technique and that of WMS is that encoding the beam's intensity with frequency information, with the chopper, is separate from the tuning of the laser. Since the oscillations of the optical field are too fast for the detector to follow, it sees an average intensity chopped into a square wave. A square wave is equivalent to the sum of a sinusoid at the square wave's frequency and smaller amplitude sinusoids at higher odd order harmonics. Therefore multiplication within the lock-in, with a 1f reference sinusoid for example, will pick out a sinusoidal harmonic in the signal which is proportional to the intensity of the square wave. As the laser tunes across an

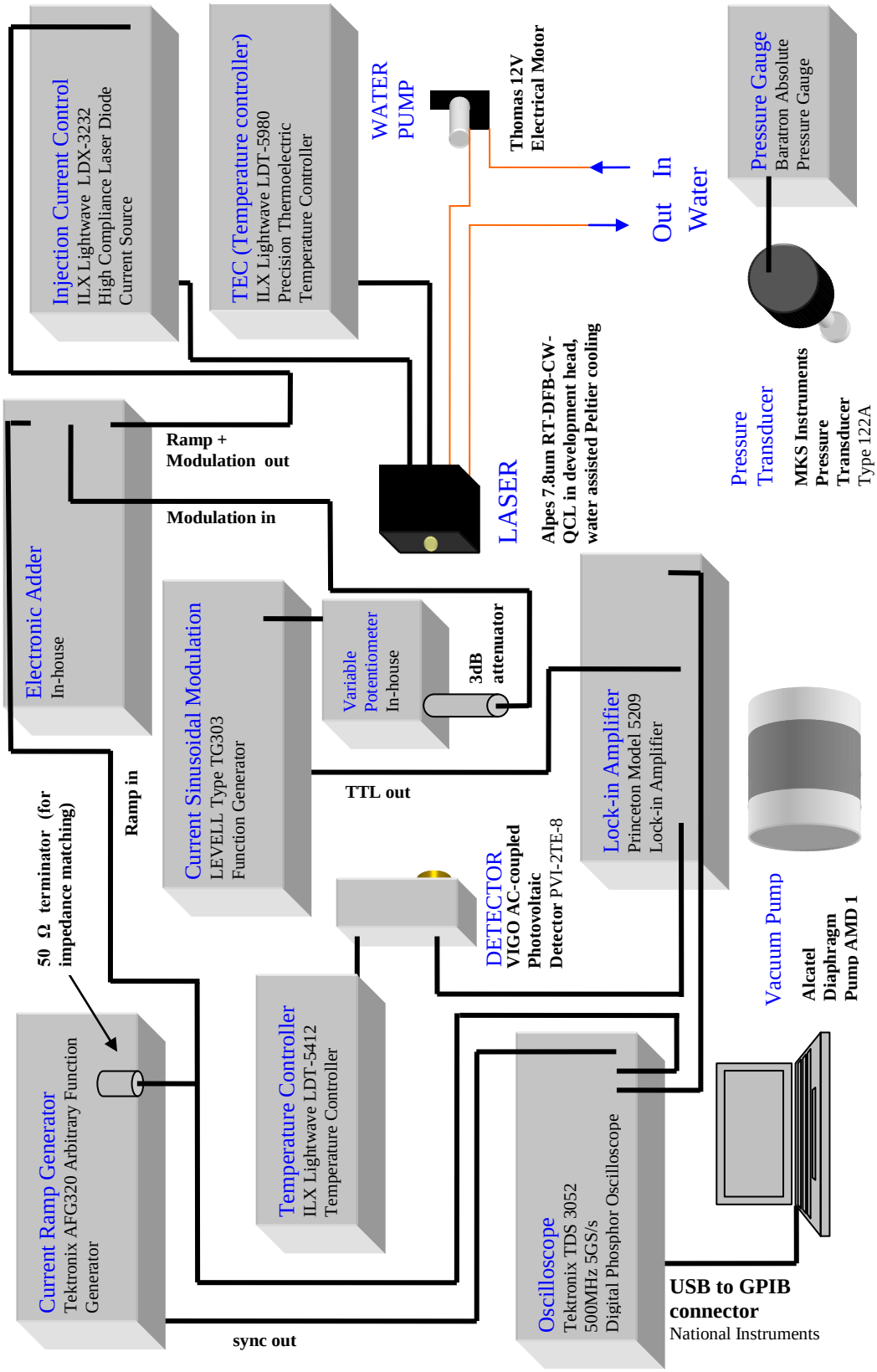


Figure 5-7: Electrical Setup for Wavelength Modulation Spectroscopy with the continuous wave quantum cascade laser.

absorption feature, the amplitude of the square wave changes in proportion to the absorption. Hence, the harmonic components also change in proportion.

In Wavelength Modulation Spectroscopy the frequency encoding simultaneously tunes the laser. As was shown in the previous section (Section 5.3.3), this tuning results in wavenumber oscillations about each point on the absorption profile. This in turn results in a periodically changing intensity signal for each point. In general this consists of a fundamental at the modulation frequency and its even and odd harmonics. Multiplication within the lock-in then rectifies these oscillations for the chosen harmonic. Considering the resulting signal overall for all the absorption profile points, none of the Fourier components is directly proportional to the absorption profile amplitude at that point. Instead, the Fourier components are qualitatively similar to the corresponding derivatives of the absorption profile (1f similar to 1^{st} derivative, 2f similar to 2^{nd} derivative etc). As Section 5.3.3 showed, the rectified voltage signal from WMS can be positive or negative, whereas chopping and rectifying produces a signal potentially between a positive maximum (no absorption) and zero (complete absorption).

5.4 Experimental Setup

The optical setup remained the same as that presented in Section 4.4.1 and Figure 4-2. The electrical setup of Section 4.4.2 and Figure 4-3 was retained and added to. The resulting full electrical setup for Wavelength Modulation Spectroscopy is shown in Figure 5-7.

The current ramp continued to be provided by a Tektronix AFG320 Arbitrary Function Generator. This had dual outputs, but to prevent cross-talk and the possibility of supplying the wrong settings to the laser, the sinusoidal current modulation was added from a separate LEVELL Type TG303 Function Generator. The output of this generator provided both the sinusoidal waveform to the cw-QCL laser at the modulation frequency, and a TTL output at the same frequency to the

$^{12}CH_4$ **Identification**	**Wavenumber / cm^{-1}**	**Line Strength / cm^{-1}/molecule cm^{-2}**	**Sweep Integration: Laser Substrate Temperature / °C**	**WMS: Laser Substrate Temperature / °C**
υ_4 P(4)$_{A2\ 1\leftarrow A1\ 1}$	1283.458791	6.26×10^{-20}	13	11
υ_4 Q(13)$_{F1\ 4\leftarrow F2\ 1}$	1283.217895	2.13×10^{-21}	15	13
υ_4 Q(13)$_{F2\ 4\leftarrow F1\ 1}$	1283.207805	2.13×10^{-21}	15	13
υ_4 P(4)$_{F2\ 2\leftarrow F1\ 1}$	1282.984175	3.74×10^{-20}	17	15
υ_4 P(4)$_{E\ 2\leftarrow E\ 1}$	1282.624469	2.49×10^{-20}	21	19
υ_4 P(4)$_{F1\ 3\leftarrow F2\ 1}$	1281.610588	3.71×10^{-20}	31	29
Sweep Integration Parameters	Sawtooth Current Ramp: 10kHz, 500mA – 680mA			
WMS Parameters	Sawtooth Current Ramp: 5Hz, 450mA – 690mA Sinusoidal Modulation: 10kHz, 4.4mA amplitude Lock-in amplifier's RC time constant: 1ms			

Table 5-1: Operating parameters and $^{12}CH_4$ line positions for sweep integration and WMS.

lock-in amplifier's reference channel. The lock-in was a Princeton Model 5209. The ramp and modulation signals were combined via an in-house electronic adder, and subsequently passed through a variable potentiometer and 3dB attenuator, prior to the laser, as an extra current limit precaution. (The whole setup was also tested on an LED, prior to involving the laser). Setting the modulation index to ~2.2 was a simple matter in practice, by adjusting the sinusoidal amplitude and observing the 2f peak height on an oscilloscope. The lock-in output is also dependent on the phase difference between the detected signal component and the lock-in's reference wave form (derived from the TTL input). If the lock-in's output baseline is not too noisy, this is best maximized by nulling the signal and adding 90° of phase.

5.5 Operating Parameters

Due to AC-coupled detection, different parameters were needed for WMS than sweep integration (Section 4.6.1.5). The current ramp for overall wavenumber tuning was run at 5Hz, which was near the lower limit of a stable oscilloscope trace - while fitting a whole ramp on screen. This allowed the modulation frequency to be relatively fast, while allowing high resolution and high wavenumber coverage for line identification.

Modulation frequencies between 10kHz and 50kHz were tried, with similar results. In this chapter, the particular 2f spectra - used for comparison with the sweep integration results of Section 4.6.5 - was taken at 10kHz.

The sawtooth current ramp ran from 450mA to 690mA. The added sinusoidal oscillation at 10kHz was found to give a maximum 2f signal (modulation index ~2.2) at an amplitude of 4.4mA.

The lock-in amplifier's RC time constant was set at 1ms. This controls the bandwidth of the low pass filter, with a long time constant letting through less low frequency noise, although the time constant needs to be short enough to allow changes in the

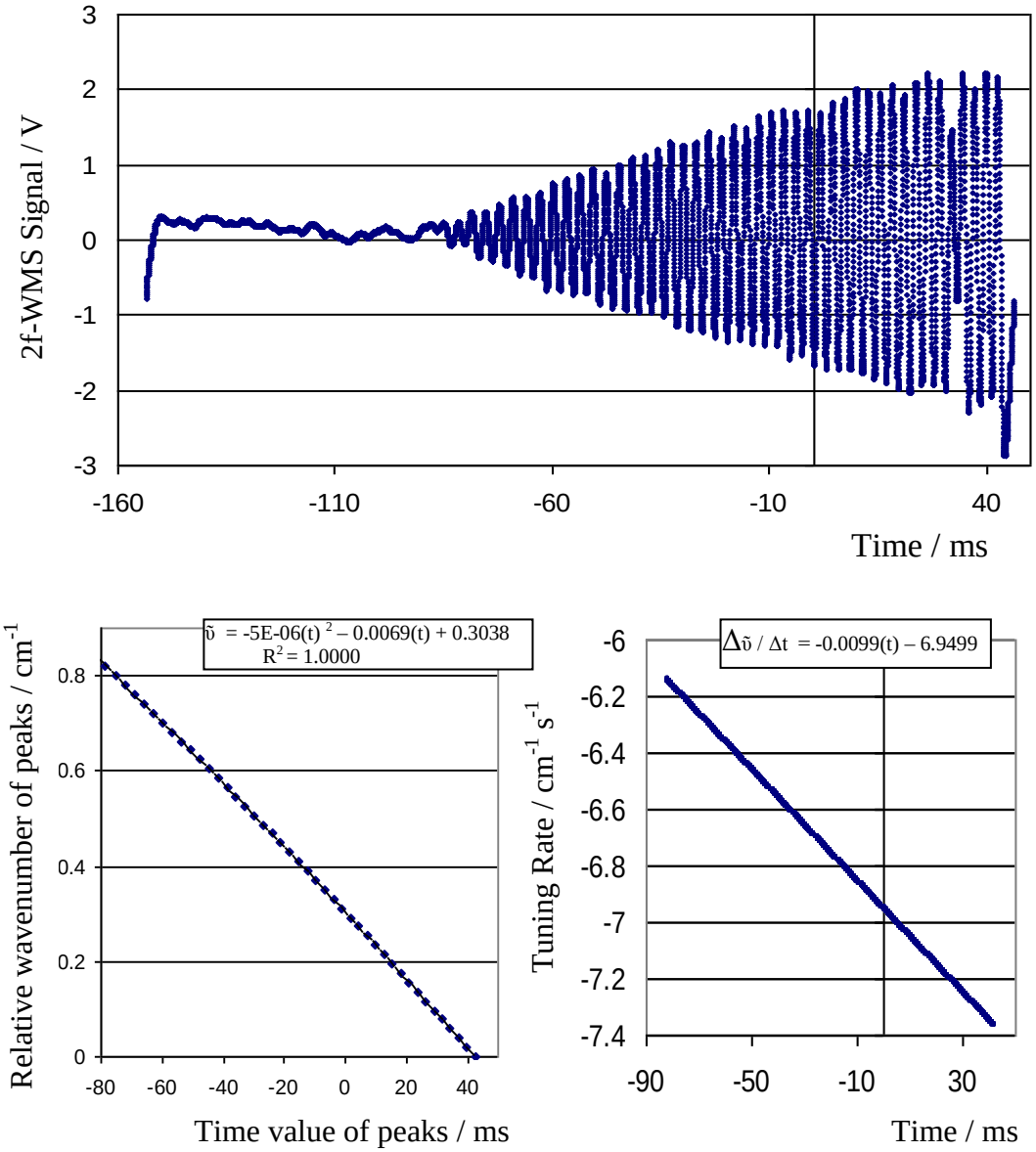


Figure 5-8: **2f-WMS signal of Ge etalon fringe pattern, FSR=0.0195cm^{-1} (top plot). Also shown are the implied time to relative wavenumber calibration plot and the associated tuning rate. 5Hz sawtooth current ramp with 10kHz sinusoidal modulation, modulation amplitude =4.4mA, I_{base}=450mA, $I_{base-peak}$=240mA, $T_{substrate}$=11°C, RC=1ms.**

DC magnitude to be properly recorded. With these WMS parameters, the 6 strong CH_4 lines from Chapter 4 could be found by reducing the laser substrate temperature by 2°C from its sweep integration equivalent. Table 4-1 is updated as Table 5-1 to reflect this.

5.6 Tuning Rate and Wavenumber Coverage

Since the 2nd derivative peaks of the Fabry-Perot fringes occur at the same frequency spacing as the fringes themselves, the 2f-WMS signal can be used for wavenumber calibration of the 5Hz current ramp. The top plot of Figure 5-8 shows such peaks with the WMS settings used in this chapter (Section 5.5). Although the ramp's upper limit was fixed, it would have been possible to increase the current ramp's coverage some more by bringing the base current up closer to threshold. However, including a section below threshold in the scan was useful to be sure of line identification, and these WMS settings facilitated the -2°C relation to the sweep integration laser substrate temperature seen in Table 5-1, while placing the absorption feature at the high power / lower noise end of the scan.

The wavenumber coverage in Figure 5-8 is $0.84cm^{-1}$, with a current range of 240mA. This is consistent with the slower tuning ramp allowing the cw-QCL to follow the current induced temperature tuning more closely. Figure 4-23 showed a coverage of $0.39cm^{-1}$ at 10kHz for a similar current range of 280mA, and Figure 4-24 showed a coverage of $0.33cm^{-1}$ at 10kHz for the 180mA range used for the sweep integration comparison. The quadratic fit of Figure 5-8's 5Hz time to wavenumber calibration is a little more linear than that of the 10kHz calibrations.

By definition, the 5Hz scan is 2000 times slower than the 10kHz scans, and shows an average tuning rate of $-6.67cm^{-1}s^{-1}$. This is about 1000 times slower than Figure 4-23's $-6.58cm^{-1}ms^{-1}$ and about 700 times slower than Figure 4-24's $-4.65cm^{-1}ms^{-1}$.

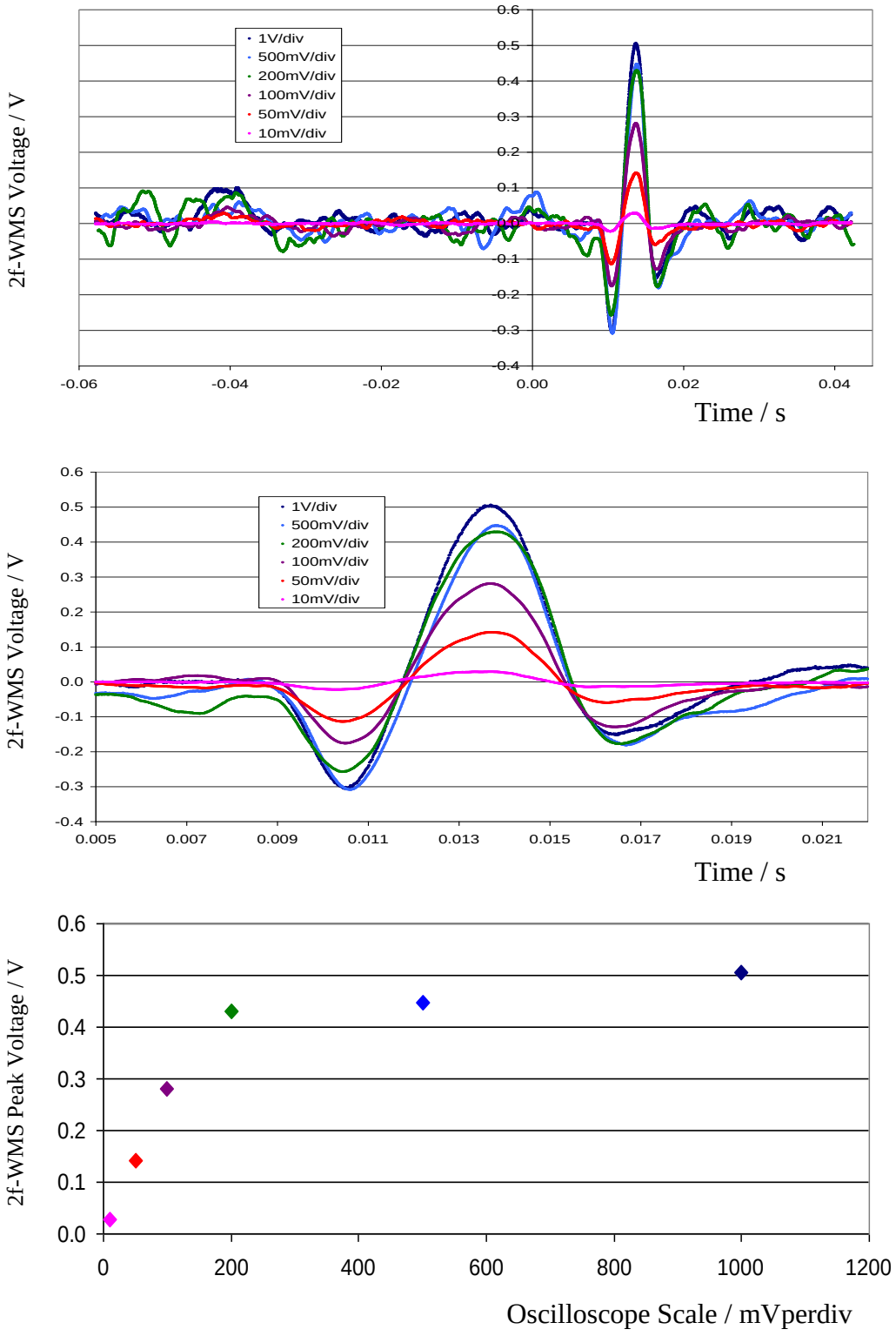


Figure 5-9: **Electronic incompatibility in the detector-transimpedance amplifier-lock-in-oscilloscope signal chain. What should be overlapping signals are artificially affected by the chosen oscilloscope scale. The bottom plot shows the 'saturation' of peak height with oscilloscope scale, and this greater consistency at larger oscilloscope scales is taken as giving a more reliable reading.**

5.7 Electronic incompatibility issues

As discussed in Section 4.6.1.4, there was an issue of incorrect and non-linear oscilloscope scaling through the detector-transimpedance amplifier-oscilloscope and detector-transimpedance amplifier-lock-in-oscilloscope signal chains. This meant that, what should be overlapping signals, artificially changed magnitude when recorded with different oscilloscope scales. (Top and middle plots of Figure 5-9). This discovery occurred too close to loss of tenure of the laser to explore experimentally. So the approach taken was the same as previously. Since the required information – here the 2f-WMS peak height – showed a growing consistency of reading with larger oscilloscope scales (bottom plot of Figure 5-9), these readings were taken as being the more reliable. They also produced the more conservative estimates of the Limit of Detection (LoD).

The presence of electronic problems in the signal chain probably contributes to the large noise in the lock-in amplifier's output. Although averaging with sweep integration supplements many techniques, it was particularly necessary here. High noise in the signal chain made the use of the full 512 oscilloscope averages necessary to obtain a usable signal. The high noise also meant that the Princeton 5209 lock-in could not autophase, and the signal had to be maximized by adjusting the phase manually. This remained true for the two other lock-in amplifiers available – a Stanford SR510 and a Stanford SR850.

Figure 5-10 shows the 2f-WMS signal obtained using 2, 8 and the maximum 512 oscilloscope averages. This averaging significantly increases the integration time – 0.2s for a single 5Hz scan, and 102.4s (plus a small averaging and display overhead) for 512 averages. This compares to 0.0512s for 512 10kHz sweep integration scans of Chapter 4.

Although the noise signals coming from the lock-in have large amplitudes, the Fourier components of the 2 average scan shows that it is primarily low frequency noise – consistent with an RC time constant of 1ms (Figure 5-11). There does seem

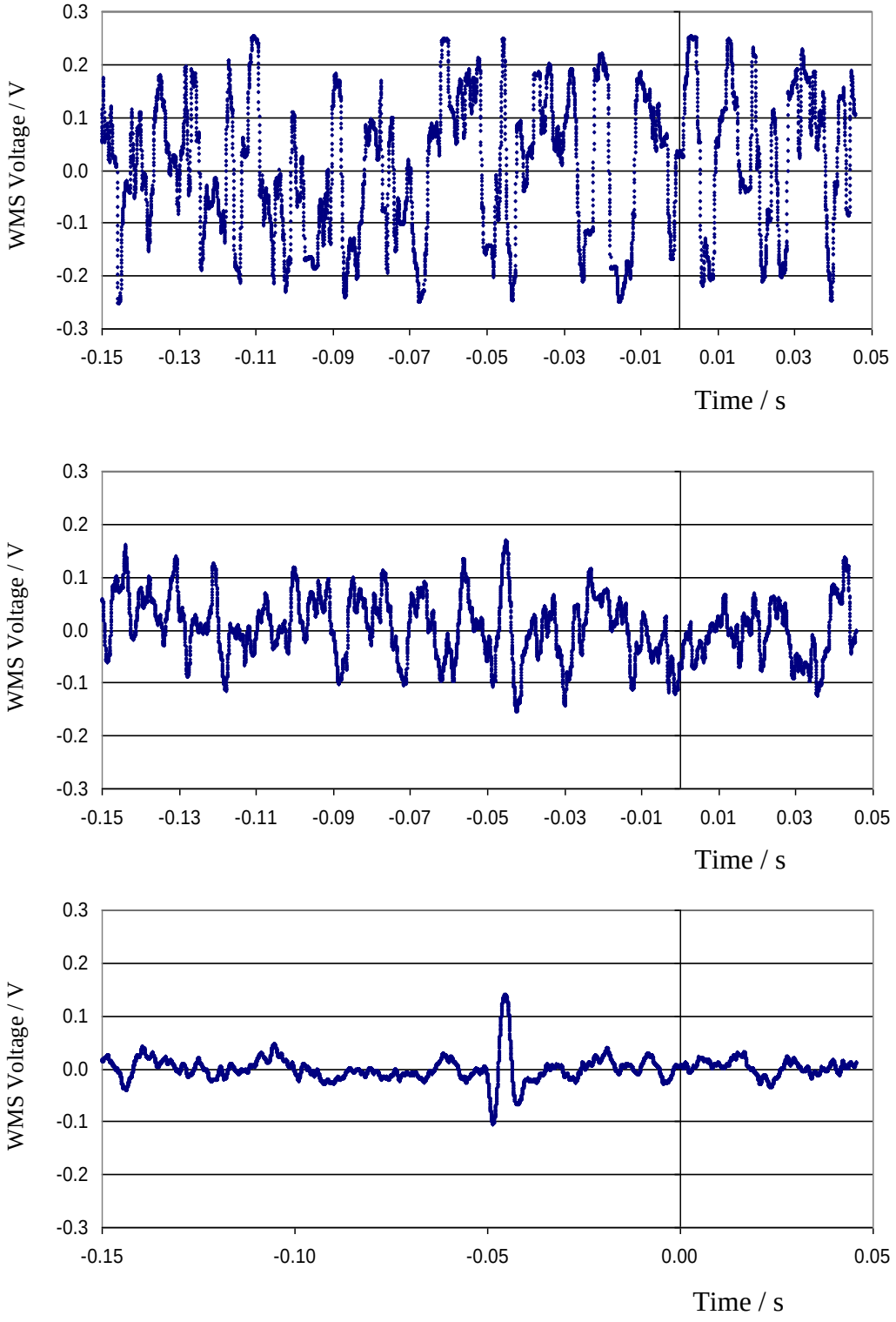


Figure 5-10: 2f Wavelength Modulation Spectroscopy spectrum of pure $^{12}CH_4$ υ_4 P(4) $_{F1\ 3\leftarrow F2\ 1}$ at 1281.610588cm^{-1} for 2, 8 and 512 oscilloscope averages, respectively. Ramp frequency =5Hz, modulation frequency =10kHz, modulation amplitude =4.4mA, lock-in time constant =1ms, I_{base}=500mA, $I_{base\text{-}peak}$=200mA, $T_{substrate}$=29°C.

to be structure in the top two plots of Figure 5-10, having FWHM's comparable to the 2f-WMS lineshape. (The time to wavenumber calibration is approximately linear). However, if this were due to etalon fringing in the system, it would be systematic and not average out so readily. It might be noted that a 10kHz modulation is at the lowest end of a linear response for this detector (Section 4.6.1.2). The VIGO detector is easily capable of measuring MHz frequencies, so the recordings here may be at the noisier end of its response curve.

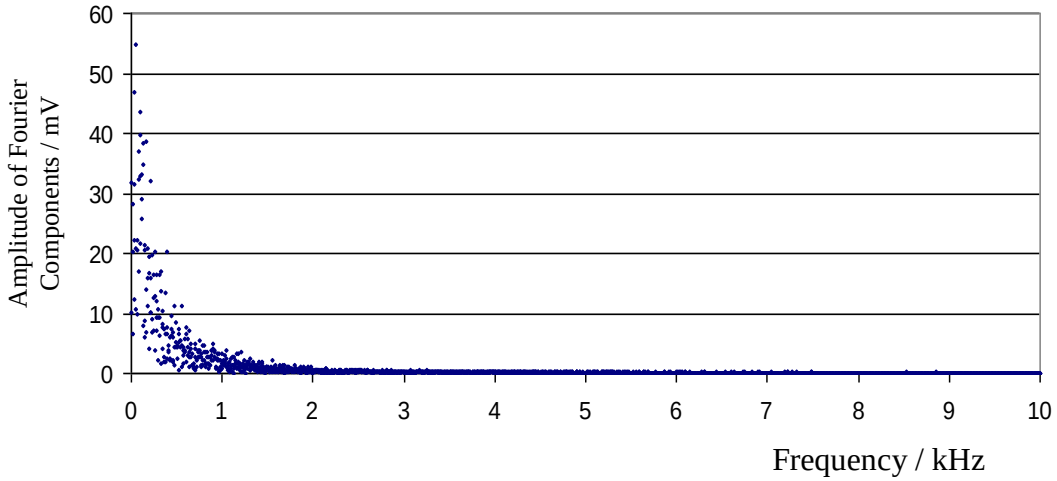
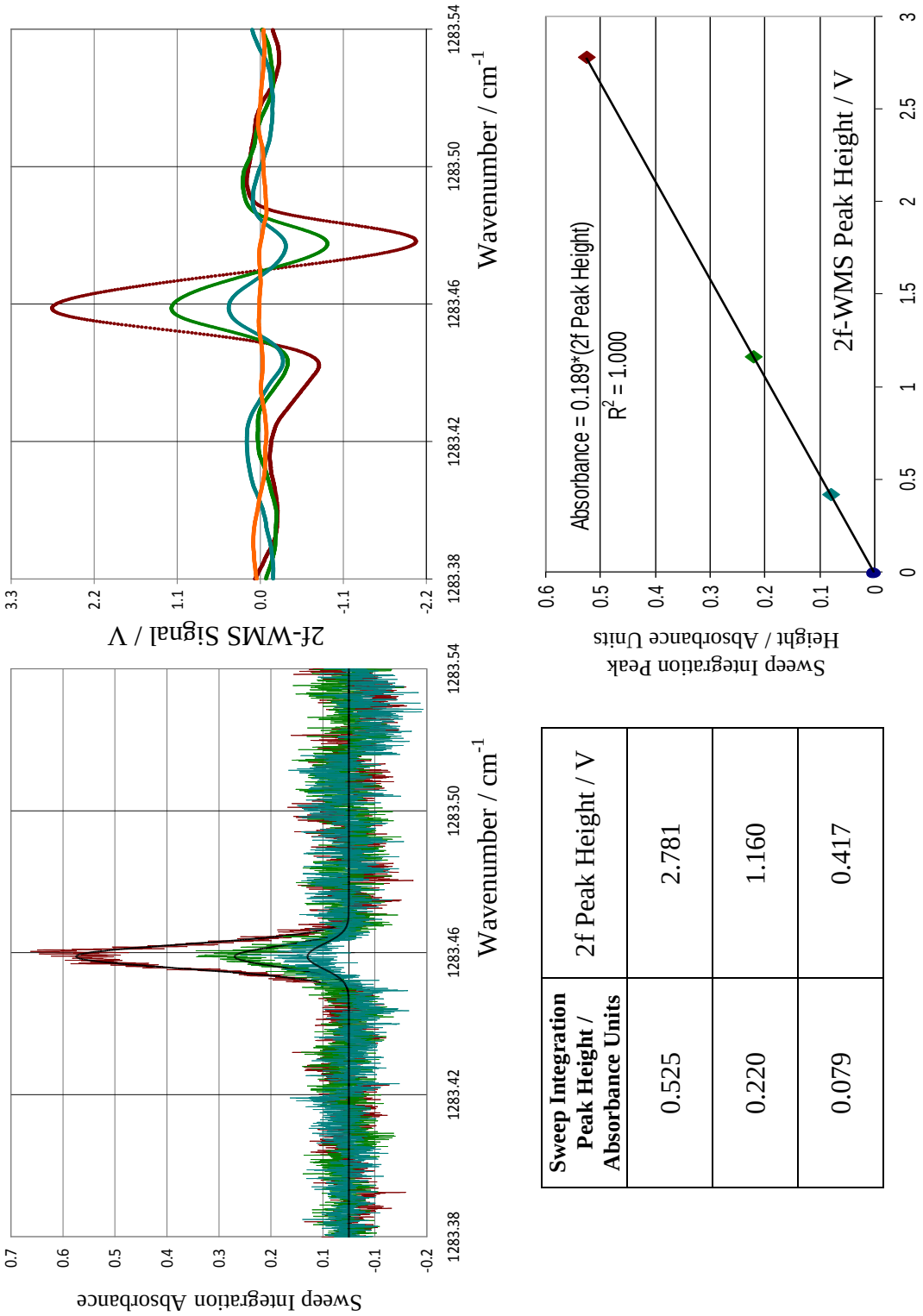


Figure 5-11: Fourier amplitudes of the frequency components of the top plot of Figure 5-10. The noise amplitudes are consistent with a low-pass RC time constant of 1ms.



Sweep Integration Peak Height / Absorbance Units	2f Peak Height / V
0.525	2.781
0.220	1.160
0.079	0.417

Figure 5-12: Sweep Integration and WMS (combined with 512 averages) calibration and noise floor determination, on a sample of pure CH_4, using the $^{12}CH_4$ υ_4 P(4)A_2 1←A_1 1 transition at 1283.458791cm^{-1}. The 2f-WMS peak height voltages are calibrated to absorbance values using their equivalent sweep integration peak absorbances. The 2σ of the WMS noise floor (orange trace) is then entered into this relationship to find the WMS noise floor in absorbance units.

5.8 Limit of Detection

The Wavelength Modulation Spectroscopy signals are produced in terms of voltage. They therefore need calibration to absorbance values. Absorbance values are proportional to number density and / or mixing ratio. They also allow the Limit of Detection (LoD) to be specified at the fixed value of absorbance noise for the system, with the operating parameters used. This fixed absorbance noise level can then be used to specify the concentration LoD, which depends on the particular gas and absorption line chosen.

In Chapter 4, Section 4.6.5, the Herriott cell was sequentially pumped down to produce absorbance spectra at three different number densities, using a fixed mixing ratio (100%) of $^{12}CH_4$ for the υ_4 $P(4)_{A2\,1\leftarrow A1\,1}$ transition at vacuum wavenumber 1283.458791cm^{-1} (Rothman et al., 2009). The noise floor of these measurements – taken as twice the standard deviation of the absorbance noise surrounding the absorbance signal - allowed the Sweep Integration limit of detection to be found in absorbance units.

The trend of Sweep Integration peak heights, in absorbance units, allows a calibration to their equivalent 2f-WMS peak heights in volts. Therefore an arbitrary 2f-WMS voltage level can then be entered into this relationship to find its equivalent absorbance. The calibration trend is shown in Figure 5-12. It is made up from the three pairs of corresponding peak heights and forcing the fit through zero - for zero concentration. The fit is linear, with an R^2 of essentially 1. The linear calibration equation is therefore:

$$Absorbance = 0.189\left(2f\ WMS\ peak\ height\right) \tag{5-5}$$

Taking 2σ of the surrounding 2f-WMS noise floor after removal of systematic noise (see below), taking 2σ from a scan less affected by systematic noise, or adding 90° of phase to null the signal to measure 2σ of the remaining signal (the orange trace in Figure 5-12) at the signal position, gives a voltage level of 0.090V. From the

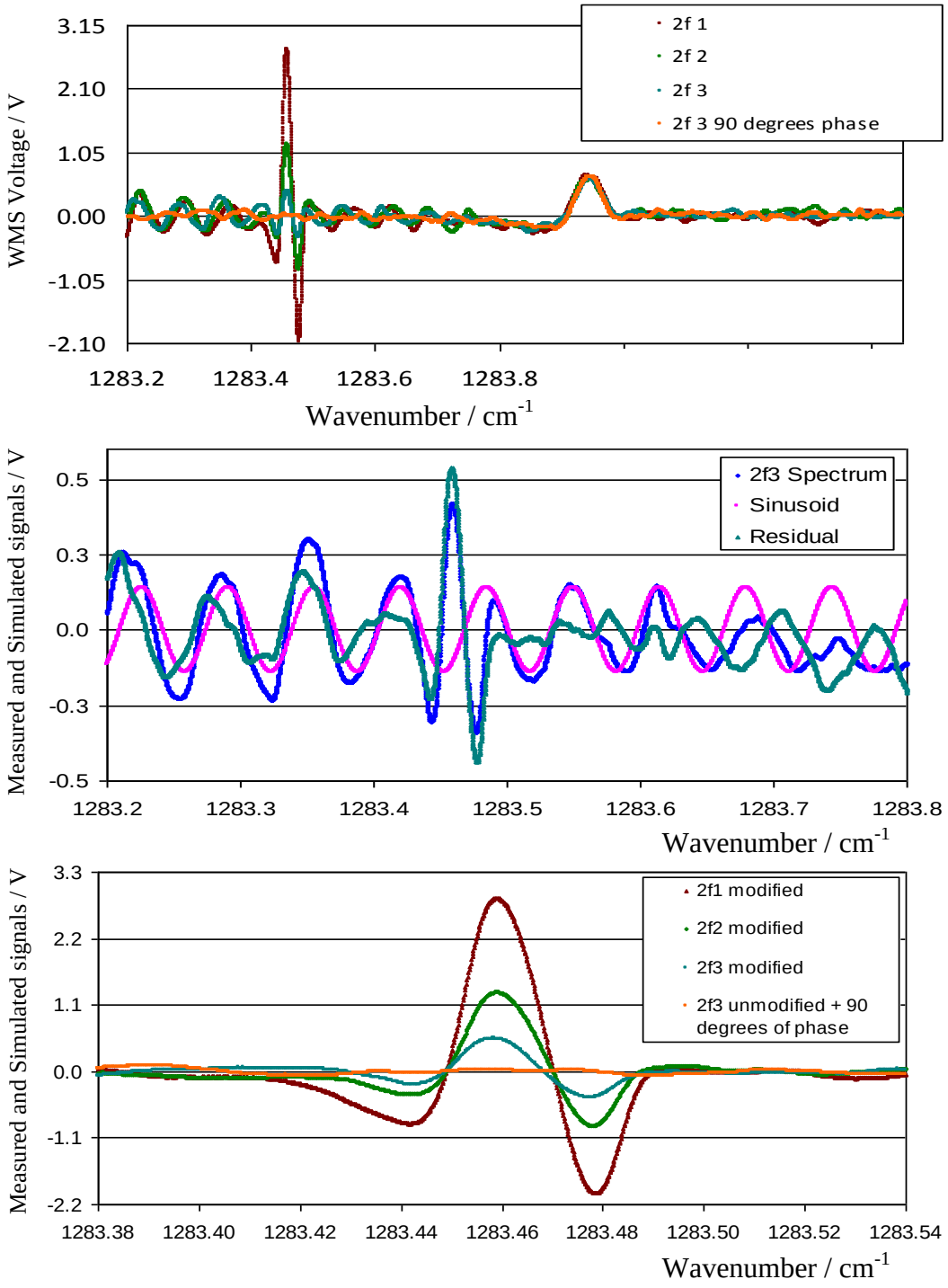


Figure 5-13: **Top plot: The full data sets from which the calibration data was taken for Figure 5-12. Systematic noise is evident.**
Middle Plot: The 2f3 data showing systematic noise. The noise is modeled in the region of the spectral feature as a constant amplitude sine wave. The residual of these data sets gives a modified 2f3 data set - with increased peak height and feet asymmetry.
Bottom plot: The modified 2f1, 2f2, and 2f3 calibration data sets, plotted against the noise signal after nulling the unmodified 2f3 signal.

calibration equation (Equation 5-5) this voltage corresponds to an absorbance noise floor for the 2f-WMS technique (combined with 512 averages) of 0.017 absorbance units in an integration time of 102.4s. This compares with 512 sweep integration averages giving 0.080 absorbance units in 51ms.

The 2f-WMS technique, combined with 512 averages, therefore improves the absorbance sensitivity by a factor of 4.7, but at the expense of needing calibration, phase sensitive detection, and an increase of 2000 in the integration time. (Had averaging not been necessary, the WMS integration time would have been 3.9 times larger than 512 sweep integrations).

Calibrating against the areas of the direct absorbance fits gave a very similar linear best fit line, but with a slightly degraded spread (R^2=0.994).

5.8.1 Discrepancy with the predicted trend

The linear calibration trend of direct and 2f-WMS peak heights is convenient, but is not the trend predicted. For the absorbance range of 0 to 0.525 the predicted trend is quadratic, with the rise of the 2f-WMS signal weakening with increasing direct absorbance (Section 5.9.1.3). This is not a problem, however, as the trends agree as the absorbance decreases (see for example Figure 5-20) towards the detection limit, and, in any case, with better pressure (number density) control, calibration can be done nearer to the small absorbance limit (α<0.1) where both predicted trends are linear.

The deleterious aspects of the experiment meant that only the early calibration set used in the above analysis was suitable. This data set was subject to periodic noise (top plot of Figure 5-13), most likely from unwanted etalon fringes. This was eliminated as a cause of the linear trend by approximating the noise with a sine wave over the 2f-WMS lineshape. The residual then gave a largely undistorted version of the 2f-WMS signal (middle plot of Figure 5-13). The procedure had little effect on

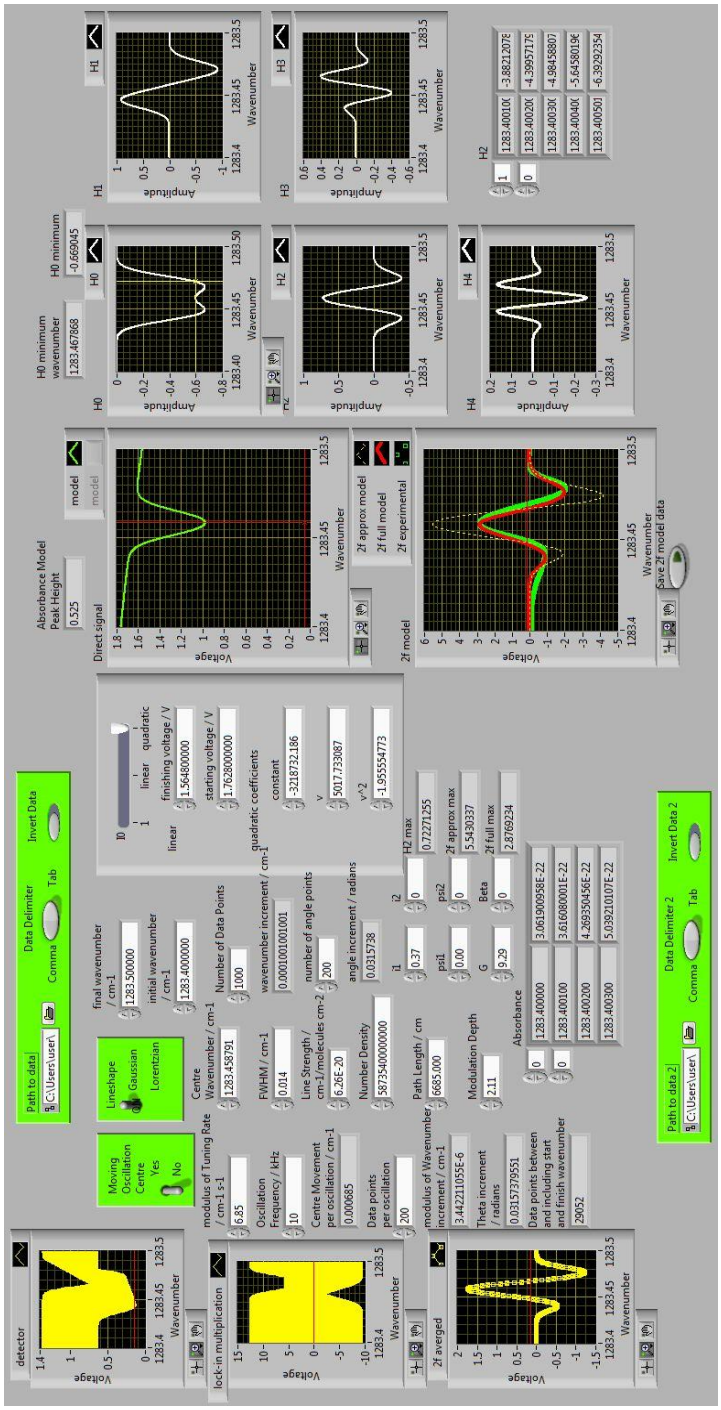


Figure 5-14: Screenshot of the 2f-WMS lineshape modeler for the full and approximate models. The Fourier amplitudes which contribute to the overall signal in the model of Li et al. (2006) are derived from the direct absorbance and plotted as H_0 to H_4. The mix of their contributions is determined by the adjustable parameters i_1, i_2, ψ_1, and ψ_2.

the two larger signals, but increased the smallest signal's peak height, and restored a feet asymmetry towards the ratio seen in the larger two sets. The adjusted set is shown as the bottom plot of Figure 5-13. The trend of the new plot was still linear, however, with a slightly degraded spread of R^2= 0.9883 (R^2= 0.9779 against direct fit area). It would seem that the electronic broadening (Section 5.7) is somehow compensating for the changes in peak height with number density.

A further check on the noise floor was provided by the FFT tab of the data analysis program (Section 3.4.6), by nulling the systematic noise. However, this technique could not be used for general noise reduction for the purpose of calibration, since the noise period is very similar to the width of the 2f WMS lineshape – and hence to the period of its main Fourier component. Therefore, in this case, reducing noise by this method distorts the signal itself.

5.9 2f-WMS Modeling

Two 2f-WMS lineshape models based on Fourier series were encoded into LabVIEW. A screenshot is shown in Figure 5-14. The first model was a custom written numerical simulation valid for arbitrarily large signals. This allows all of the lines in a calibration (Section 5.8) to be modeled. The second model was coded from equations 1 to 8 of Li's paper (2006), itself based on Kluczynski's extensive account (2001). This second model gives the 2f-WMS lineshape in terms of five Fourier amplitudes, and is useful for absorbances ≤10%. The model is therefore appropriate for low concentration measurements, since this regime is where the noise reduction advantages of WMS are most needed. The model also provides insight into the mixing of terms describing the detected intensity signal. The detected intensity changes are due to sweeping over, and oscillating about, wavenumbers on the absorbance profile, and the simultaneous changes in the laser output intensity.

Both models begin by assuming that the laser's wavenumber modulation is perfectly sinusoidal in time about a point $\tilde{\nu}$ on the absorption profile.

$$\tilde{\nu}_t(t,\tilde{\nu},a) = \tilde{\nu} + a\cos(\omega t) \tag{5-6}$$

It is also assumed, in the first instance, that the modulation is fast enough that the centre of oscillation, $\tilde{\nu}$, remains fixed for at least one cycle. (The effect of relaxing this assumption is discussed later (Section 5.9.2)). The injection current also causes a modulation in laser intensity, of amplitude $i_1\bar{I}_0$. The intensity modulation is (nearly) in phase with the injection current, but advanced relative to the wavenumber modulation by ψ_1. (That is, the QCL's temperature (wavenumber) tuning lags the injection current). On the fast time scale of the modulation, the mean intensity $\bar{I}_0$ output of the laser is taken as fixed, though it will vary as the centre of oscillation is slowly tuned. The laser intensity is therefore modeled by Equation 5-7:

$$I_0(t) = \bar{I}_0\left[1 + i_1\cos(\omega t + \psi_1) + i_2\cos(2\omega t + \psi_2)\right] \tag{5-7}$$

The final term in Equation 5-7 represents a quadratic nonlinearity in the laser output intensity – oscillating at the second harmonic frequency, 2ω (since $\cos^2(\omega t + \psi_2') = (1/2)(1+\cos(2\omega t + 2\psi_2'))$, and the DC factor of $1/2$ becomes part of $\bar{I}_0$. This implies that, if $\cos\omega t$, representing the wavenumber modulation, peaks at $\omega t = 2\pi n$ where $n = 0, \pm 1, \pm 2, \ldots$, then $\cos(2\omega t + 2\psi_2')$, representing the quadratic intensity modulation, peaks when $2\omega t + 2\psi_2' = 2\pi n$, so that $\omega t = \pi n - \psi_2'$ $= \pi n - \psi_2/2$). Higher harmonics can be added if necessary. The subscript of I_0 implies that this is the intensity coming from the laser before encountering the absorption line profile. i_1, ψ_1, i_2, and ψ_2 are properties of the particular laser used.

In terms of Figure 5-6, the x-axis wavenumber oscillations about the absorption profile would still be represented by the oscillation at the fundamental frequency as in Equation 5-6 and shown in Figure 5-6. The y-axis oscillations of Figure 5-6 are drawn for no laser intensity modulation, so that all the intensity modulation is due to the radiation encountering the absorption line shape, and is due to the $\alpha(\tilde{\nu}_t)$ term (in Equation 5-11 below). That is, if a horizontal section of line profile was encountered a constant intensity would be detected (top plot of Figure 5-3). The last two terms of Equation 5-7 mix with the intensity oscillations due to the absorbance. These additional terms have phases ψ_1 and ψ_2 respectively to the 1f and 2f wavenumber oscillations, and, for instance, the 2f frequency might therefore become explicit in the overall signal - even when the absorbance maximum is not within the modulation amplitude. (This contrast with the absorbance only intensity oscillations about points on a Voigt profile: Figure 5-6 in Section 5.3.3).

The custom written numerical model is based upon multiplying the signal at the sequential stages of: detection; multiplication by the lock-in amplifier's reference sinusoid; and low pass filtering. At the detection stage the signal is given by the Beer-Lambert Law, $I_\tau(\tilde{\nu},\theta,a) = I_0(\theta)\exp\left[-\alpha(2\tilde{\nu}_0 - \tilde{\nu},\theta,a)\right]$. The intensity is a function of time, $\theta = \omega t$, as the laser tunes – due to the laser output intensity, $I_0(\theta)$, and the absorbance profile, $\alpha(2\tilde{\nu}_0 - \tilde{\nu},\theta,a)$.

The argument of the absorbance (lineshape) has been written as $2\tilde{\nu}_0 - \tilde{\nu}$, rather than $\tilde{\nu}$, to indicate a subtlety of calculations involving the direct lineshape in the wavenumber domain – when addressing wavenumber downchirps. The three plots on the left hand side of Figure 5-15 illustrate the experimental situation. The direct absorbance line is swept over in the time domain. The negative gradient of this lineshape is encountered first as the transmittance begins to dip. For a wavenumber downchirp this negative gradient occurs at the high wavenumber part of the sweep. The generally largest asymmetrical component of the resulting 2f-WMS signal is proportional to this gradient - captured by the term known as H_1 (see later). A time

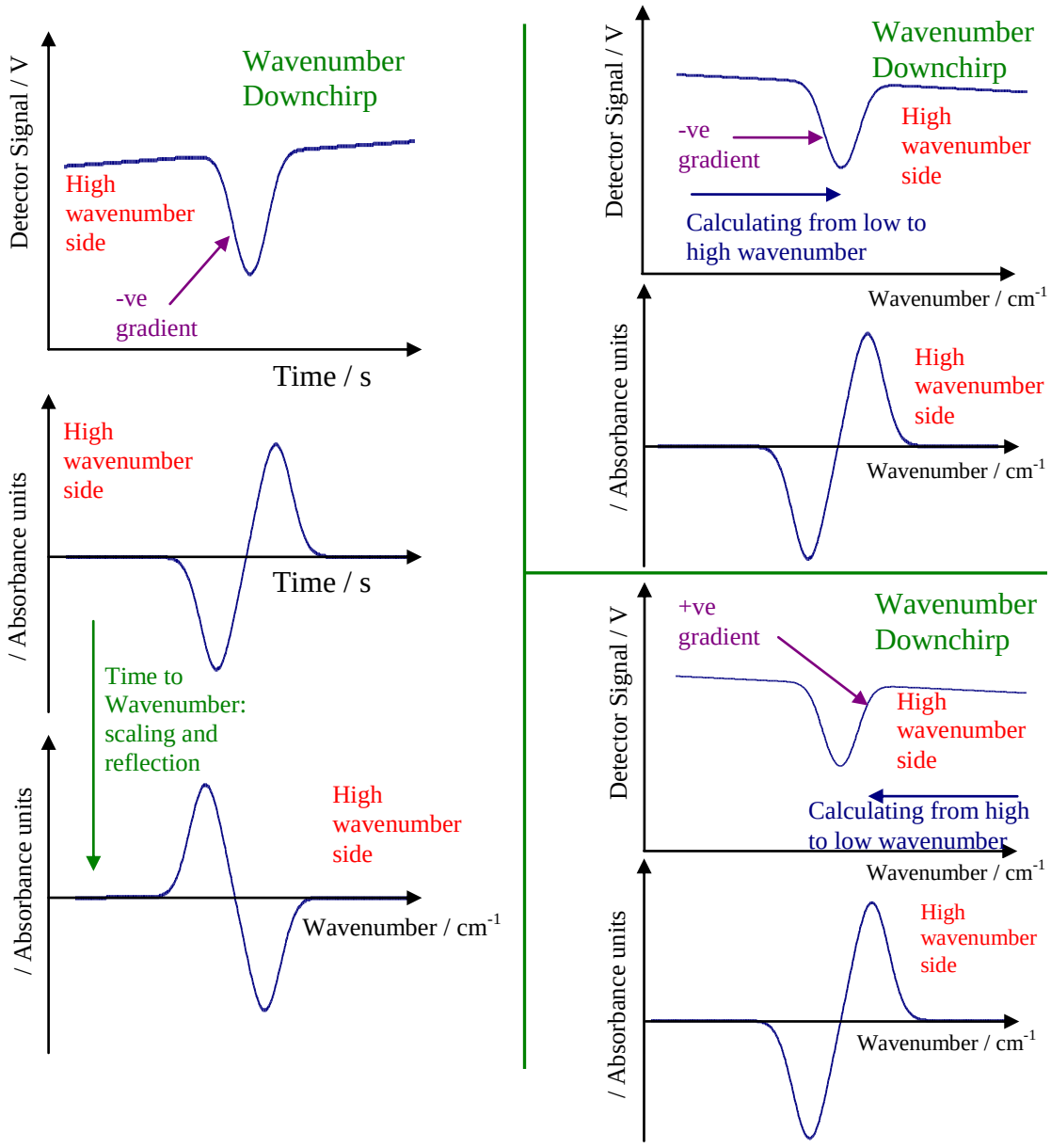


Figure 5-15: **Schematic diagram illustrating a subtlety in modeling the asymmetrical terms in a 2f-WMS signal, in the wavenumber domain. The three left hand side plots show the experimental situation where the signal is measured in the time domain. The axis is then reflected and scaled to convert from time to wavenumber. Since the negative gradient of the direct signal is encountered at the start of the absorption dip, this conversion correctly places the negative part of the asymmetrical H_1 component at the high wavenumber end.**

The two plots in the upper left panel illustrate the spurious placement of the negative part of H_1 when calculating from low to high wavenumber in the wavenumber domain.

The two plots in the lower left panel illustrate the spurious asymmetry of H_1 when working from high to low wavenumber in the wavenumber domain, since a positive gradient is first encountered on the high wavenumber side.

Making the substitution $\tilde{\nu} \to 2\tilde{\nu}_0 - \tilde{\nu}$ in the direct lineshape function, when calculating in either direction, fixes this problem. (See also Figure 5-16).

to wavenumber conversion can then be used to scale and reverse the abscissa, correctly placing the negative part of the H_1 term at high wavenumber.

However, when working with the direct lineshape in the wavenumber domain, a negative gradient will be encountered on different sides of the profile depending on whether the calculation is performed from low to high, or from high to low wavenumber. The two plots in the upper right panel of Figure 5-15 illustrate the low to high wavenumber calculation. Similarly the lower right panel shows the high to low case. The low to high order of calculation is usual, but it can be seen that the encountering of the negative gradient incorrectly places the negative part of H_1 at low wavenumber. Programming the order of calculation from high to low is no solution, however. In this case, a positive direct lineshape gradient is encountered on the high wavenumber side. It was found that the transformation of $\tilde{\nu} \to 2\tilde{\nu}_0 - \tilde{\nu}$ in the direct lineshape allowed the modeling equations, laid out below, to be used otherwise unaltered – while calculating in either wavenumber direction. Overall in the lineshape function, this reflects the wavenumber centre of oscillation about the line centre, while keeping the phase of oscillation the same: $\tilde{\nu} - \tilde{\nu}_0 + a\cos\theta \to -(\nu - \nu_0) + a\cos\theta$. The even ordered symmetrical parts of the 2f-WMS signal are unaffected. Wavenumber axis considerations which enable this transformation to automatically reflect the asymmetrical components about the line centre position are illustrated in Figure 5-16.

Returning to the models, the lock-in amplifier's reference sinusoid can be written as $G\cos(2\theta + \beta)$ where G accounts for any gain and β is an adjustable phase constant provided by the lock-in amplifier. Low pass filtering is usually modeled by averaging over an oscillation cycle. (In practice, however, the integration time is a property of the overall electronic system and will often be longer – giving further smoothing and broadening of the output signal). The 2f-WMS signal from the lock-in amplifier's output can therefore be modeled by:

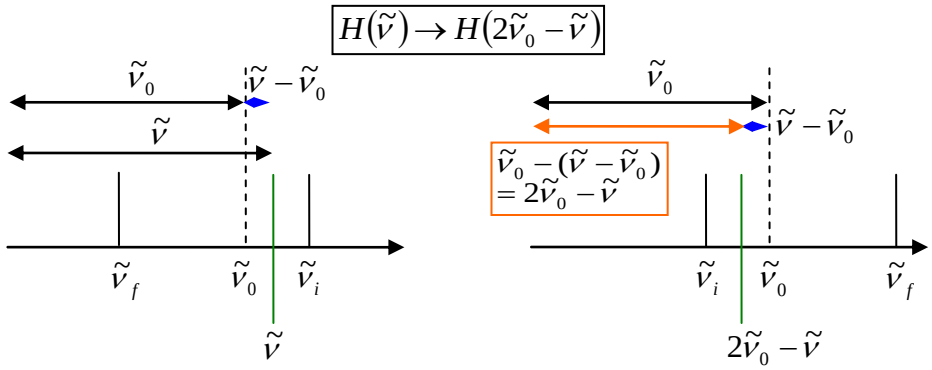


Figure 5-16: The wavenumber transformation needed to reflect the odd WMS components, H_n with n odd, about the lineshape centre frequency $\tilde{\nu}_0$. This is needed to correctly capture the asymmetry of the 2f-WMS downchirp signals in the modeling equations. The higher initial wavenumber is $\tilde{\nu}_i$, and the lower final wavenumber is $\tilde{\nu}_f$. In the left hand plot, the wavenumber difference $\tilde{\nu} - \tilde{\nu}_0$ (blue arrow) of a general wavenumber $\tilde{\nu}$ from the line centre $\tilde{\nu}_0$, is used to reflect the corresponding function value about $\tilde{\nu}_0$, as indicated in the right hand plot. The needed transformation formula follows from this.

$$X_{2f}(\tilde{\nu},a)=\frac{G}{2\pi}\int_{-\pi}^{\pi}I_{\tau}(2\tilde{\nu}_0-\tilde{\nu},\theta,a)\cos(2\theta+\beta)\,d\theta \tag{5-8}$$

The explicit expression for a 2f-WMS signal resulting from a Gaussian absorbance is:

$$\begin{aligned}X_{2f}(\tilde{\nu},a)=&\frac{G\bar{I}_0(\tilde{\nu})}{2\pi}\int_{-\pi}^{\pi}\left[1+i_1\cos(\theta+\psi_1)+i_2\cos(2\theta+\psi_2)\right]\times\\&\times\exp\left[-\frac{N}{V}LS\frac{2}{\omega_g}\sqrt{\frac{\ln 2}{\pi}}\left(\exp\left(-\frac{4\ln 2}{\omega_g^2}(2\tilde{\nu}_0-\tilde{\nu}+a\cos\theta-\tilde{\nu}_0)^2\right)\right)\right]\cos(2\theta+\beta)\,d\theta\end{aligned} \tag{5-9}$$

This expression includes all the frequency components in the detected signal, and picks out the 2f component by multiplication with the $G\cos(2\theta+\beta)$ term, and subsequent averaging. This form is particularly useful when programming the effect of allowing the centre of oscillation, $\tilde{\nu}$, to move during an oscillation. Equation 5-9 can, however, be usefully written in an equivalent form (Equation 5-16). This latter form deals with the 2f component in the detected signal explicitly. The derivation below also gives more insight into the intermediate stages of the signal processing.

The detected 2f-component can be written:

$$X_{2\theta\mathrm{det}}(\tilde{\nu},\theta,a)=A_{2\theta}(2\tilde{\nu}_0-\tilde{\nu},a)\cos 2\theta \tag{5-10}$$

where $\theta=\omega t$. The amplitude, $A_{2\theta}(2\tilde{\nu}_0-\tilde{\nu},a)$, can be obtained from the full detected signal using the orthogonality of sinusoids, as is usual for Fourier series. As above the full detected signal is given by the Beer-Lambert Law:

$$I_{\tau}(\tilde{\nu},\theta,a)=I_0(\theta)\exp\left[-\alpha(2\tilde{\nu}_0-\tilde{\nu},\theta,a)\right] \tag{5-11}$$

Here again, the absorbance, $\alpha(2\tilde{\nu}_0 - \tilde{\nu}, \theta, a) = (N/V)LSg(2\tilde{\nu}_0 - \tilde{\nu}, \theta, a)$, is a function of time, $\theta = \omega t$, as the modulation oscillates the laser tuning about a wavenumber on the absorbance profile.

It follows from Fourier analysis that:

$$\begin{aligned} A_{2\theta}(2\tilde{\nu}_0 - \tilde{\nu}, a) &= \frac{1}{\pi}\int_{-\pi}^{\pi} I_\tau(2\tilde{\nu}_0 - \tilde{\nu}, \theta, a)\cos 2\theta d\theta \\ &= \frac{1}{\pi}\int_{-\pi}^{\pi} I_0(\theta)\exp\left[-\alpha(2\tilde{\nu}_0 - \tilde{\nu}, \theta, a)\right]\cos 2\theta d\theta \end{aligned} \tag{5-12}$$

Here, $\cos 2\theta$ extracts the amplitude of the 2f-component, $A_{2\theta}(2\tilde{\nu}_0 - \tilde{\nu}, a)$, and is distinct from the $\cos 2\theta$ in Equation 5-10. Using Equation 5-7, Equation 5-12 can be written explicitly for Gaussian and Lorentzian absorbance profiles, respectively, as:

$$\begin{aligned} A_{2\theta}(2\tilde{\nu}_0 - \tilde{\nu}, a) = \frac{\bar{I}_0(\tilde{\nu})}{\pi}\int_{-\pi}^{\pi} &\left[1 + i_1\cos(\theta + \psi_1) + i_2\cos(2\theta + \psi_2)\right]\times \\ \times \exp\left[-\frac{N}{V}LS\frac{2}{\omega_g}\sqrt{\frac{\ln 2}{\pi}}\left(\exp\left(-\frac{4\ln 2}{\omega_g^2}(2\tilde{\nu}_0 - \tilde{\nu} + a\cos\theta - \tilde{\nu}_0)^2\right)\right)\right]&\cos 2\theta d\theta \end{aligned} \tag{5-13}$$

$$\begin{aligned} A_{2\theta}(2\tilde{\nu}_0 - \tilde{\nu}, a) = \frac{\bar{I}_0(\tilde{\nu})}{\pi}\int_{-\pi}^{\pi} &\left[1 + i_1\cos(\theta + \psi_1) + i_2\cos(2\theta + \psi_2)\right]\times \\ \times \exp\left[-\frac{N}{V}LS\frac{2}{\pi\omega_l}\frac{1}{1 + \frac{4}{\omega_l^2}(2\tilde{\nu}_0 - \tilde{\nu} + a\cos\theta - \tilde{\nu}_0)^2}\right]&\cos 2\theta d\theta \end{aligned} \tag{5-14}$$

The multiplication of the detected signal, $X_{2\theta\mathrm{det}}(\tilde{\nu}, \theta, a) = A_{2\theta}(2\tilde{\nu}_0 - \tilde{\nu}, a)\cos 2\theta$, by the lock-in amplifier's reference sinusoid with phase difference β, $\mathrm{X}_{ref}(\theta) = G\cos(2\theta + \beta)$, gives:

$$X_{mult}(\tilde{\nu},\theta) = \mathrm{X}_{2\theta\,\mathrm{det}}(\tilde{\nu},\theta,a)\mathrm{X}_{ref}(\theta) = GX_{2\theta\,\mathrm{det}}(\tilde{\nu},\theta,a)\cos(2\theta+\beta)$$
$$= GA_{2\theta}(2\tilde{\nu}_0 - \tilde{\nu},a)\cos 2\theta\cos(2\theta+\beta) = \frac{GA_{2\theta}(2\tilde{\nu}_0-\tilde{\nu},a)}{2}(\cos\beta+\cos(4\theta+\beta)) \quad (5\text{-}15)$$

Low pass filtering removes the AC term (at 4θ). This can be represented mathematically by averaging over a cycle, giving the lock-in amplifier's 2f output, $X_{2f}(\tilde{\nu})$:

$$X_{2f}(\tilde{\nu},a) = \frac{1}{2\pi}\int_{-\pi}^{\pi} X_{mult}(\tilde{\nu},\theta,a)\,d\theta$$
$$= \frac{GA_{2\theta}(2\tilde{\nu}_0-\tilde{\nu},a)}{4\pi}\int_{-\pi}^{\pi}(\cos\beta+\cos(4\theta+\beta))d\theta = \frac{GA_{2\theta}(2\tilde{\nu}_0-\tilde{\nu},a)}{2}\cos\beta \quad (5\text{-}16)$$

where the trigonometric identity $\sin(u\pm v) = \sin u\cos v \pm \cos u\sin v$ has been used in the evaluation of the last term of the integral.

Li et al.'s approximate model (2006) provides insight into how the intensity terms in the 2f-WMS signal mix – those due to wavenumber oscillation about points on the absorption profile and those due to the simultaneous changes in laser power. The model represents the exponential function in the Beer-Lambert law by its infinite power series expansion. However, only the absorbance, $\alpha(\tilde{\nu},t,a)$, to the first power is retained - meaning that terms oscillating at 2f arising from higher powers of the absorbance function are neglected and don't contribute to the final rectified 2f-WMS signal. The approximation is valid for small absorbances (≤10%) and hence also to the Limit of Detection of the WMS technique.

The detected (transmittance) signal is given by the Beer-Lambert law (Section 2.6):

$$I_\tau(\tilde{\nu},t,a) = I_0(t)\exp\left[-\alpha(\tilde{\nu},t,a)\right]$$
$$\approx I_0(t)\left[1-\alpha(\tilde{\nu},t,a)\right] = I_0(t) - I_0(t)\frac{N}{V}L\sum_j S_j g(\tilde{\nu},t,a)_j \quad (5\text{-}17)$$

Equation 5-17 makes the small absorbance approximation and the sum over j allows for the possibility of more than one absorbance line.

The detected intensity changes about $\bar{I}_0$ due to changes in the laser output are contained in the $I_0(t)$ term, while changes in the detected intensity due to wavenumber oscillation about the absorbance profile at $\tilde{\nu}$ are due to the $\alpha(\tilde{\nu}(t))$ term. In order to see how these simultaneous changes in intensity interact, it is useful to expand the absorbance $\alpha(\tilde{\nu}(t))$ into its Fourier components. This will also allow the overall terms at 2f to be calculated.

Choosing to describe the wavenumber oscillation in terms of cosine, the Fourier expansion of the detected intensity modulation due to $\alpha(\tilde{\nu}(t))$ simplifies to an even function in time, and its Fourier expansion is a cosine series (Equation 5-18). That is: $\alpha[\tilde{\nu} + a\cos(\omega t)] = \alpha[\tilde{\nu} + a\cos(-\omega t)]$, and the detected signal looks the same going backwards or forwards in time from the time origin. Now:

$$-\alpha[\tilde{\nu} + a\cos(\omega t)] = -\frac{N}{V} L \sum_j S_j g[\tilde{\nu} + a\cos(\omega t)]_j = \sum_{k=0}^{\infty} H_k(\tilde{\nu}, a)\cos(k\omega t) \quad (5\text{-}18)$$

where the negative sign preceding α is thus absorbed into the definitions of H_k for convenience:

$$H_0(\tilde{\nu}, a) = -\frac{\frac{N}{V}L}{2\pi} \int_{-\pi}^{\pi} \sum_j S_j g_j(2\tilde{\nu}_0 - \tilde{\nu} + a\cos\theta) d\theta \quad (5\text{-}19)$$

for $k = 0$; and for $k \geq 1$:

$$H_k(\tilde{\nu}, a) = -\frac{\frac{N}{V}L}{\pi} \int_{-\pi}^{\pi} \sum_j S_j g_j(2\tilde{\nu}_0 - \tilde{\nu} + a\cos\theta)\cos k\theta d\theta \quad (5\text{-}20)$$

H_0 is the DC component in the signal reaching the detector at point $\tilde{\nu}$, after absorbance (and neglecting the effect of simultaneous laser intensity modulation). The H_k terms, $k \geq 1$, pick out the amplitude of the fundamental and harmonic components of the detected signal due to the absorbance about the point $\tilde{\nu}$. It will be shown by Equation 5-23 that the second two terms of the intensity modulation of Equation 5-7 mix these derivative-like signals in the overall lock-in output. Without laser intensity modulation, during wavenumber modulation, the only term contributing to the 2f-WMS signal in the small absorbance approximation would be H_2, leading to a symmetrical second derivative-like lock-in signal. However, mixing causes H_0, H_1, H_2, H_3 and H_4, and thus their associated derivative-like forms, to contribute to the overall 2f shape. Asymmetry is introduced through the H_1 and H_3 terms leading to asymmetrical 1st-derivative and 3rd-derivative-like content. However, since H_1 and H_3 are zero at line centre, they do not affect the 2f-WMS peak height.

Substituting Equation 5-18 into Equation 5-17 and expanding the sum, and substituting for $I_0(t)$ with Equation 5-7, implies:

$$\begin{aligned} I_\tau(\tilde{\nu},t,a) = {} & \bar{I}_0\left[1+i_1\cos(\omega t+\psi_1)+i_2\cos(2\omega t+\psi_2)\right]+ \\ & +\bar{I}_0\left[1+i_1\cos(\omega t+\psi_1)+i_2\cos(2\omega t+\psi_2)\right]\times \\ & \times\{H_0+H_1\cos\omega t+H_2\cos 2\omega t+H_3\cos 3\omega t+H_4\cos\omega t+\ldots\} \end{aligned} \quad (5\text{-}21)$$

The multiplication in the last term of Equation 5-21 represents the simultaneous intensity modulation of the laser and its interaction with the intensity modulation caused by wavenumber oscillation about the absorption line profile. This causes the mixing of terms, so that, in particular, several occur at 2f:

$$\begin{aligned} I_\tau(\tilde{\nu},t,a) = {} & \bar{I}_0\left[1+i_1\cos(\omega t+\psi_1)+i_2\cos(2\omega t+\psi_2)\right]+ \\ & +\bar{I}_0\left[i_2H_0\cos(2\omega t+\psi_2)+\frac{i_1H_1}{2}\cos(2\omega t+\psi_1)+H_2\cos(2\omega t)+\frac{i_1H_3}{2}\cos(2\omega t+\psi_1)+\frac{i_2H_4}{2}\cos(2\omega t+\psi_2)+\ldots\right] \end{aligned} \quad (5\text{-}22)$$

The multiplication stage of the lock-in amplifier can be represented, for extracting the 2f signal, by multiplying Equation 5-22 by the reference sinusoid $G\cos(2\omega t+\beta)$, where G is the signal gain and β is an adjustable phase angle provided by the lock-in. This multiplication creates many sum and difference heterodynes. The low-pass filtering stage of the lock-in is represented by retaining only the DC terms resulting from this multiplication. This gives for the final 2f-WMS signal:

$$X_{2f}(\tilde{\nu},a)=\frac{G\bar{I}_0}{2}\left[H_2\cos\beta+\frac{i_1}{2}(H_1+H_3)\cos(\psi_1+\beta)+i_2\left(1+H_0+\frac{H_4}{2}\right)\cos(\psi_2+\beta)\right] \quad (5\text{-}23)$$

Equation 5-23 applies at a given oscillation centre wavenumber in the scan. The overall 2f-WMS signal results from the oscillation centre moving to the slower time scale of the current tuning ramp. On this scale, $\bar{I}_0$, and the H's are all functions of the ramp time, so that the overall signal comes from calculating X_{2f} at increasing ramp times, as the whole absorption profile is scanned over. Once the signal has been found in terms of time, etalon fringes can be used to calibrate to wavenumber.

Equation 5-19 and Equation 5-20 show that the H terms are fixed for a given absorbance $\alpha(\tilde{\nu})$. So, it is the 'mixing' parameters i_1, i_2, ψ_1, and ψ_2, all properties of the particular laser used, which determine different 2f-WMS lineshapes. In particular, with the amplitude of the linear intensity modulation, $i_1\bar{I}_0$, fixed by a wavenumber modulation index of 2.11, ψ_1 - the phase angle between the wavenumber tuning and the laser intensity modulation at the same frequency – can have a pronounced effect on the asymmetry of the line profile (even though the small absorbance approximation of Equation 5-17 is being used in the model). $\bar{I}_0$ gives an asymmetric intensity skew as the laser power increases across a current ramp.

It can be noted from Equation 5-18 that zero absorbance implies that all the H terms are zero. However, since the nonlinear intensity term from Equation 5-7 is at the 2f

frequency, this modulation will be detected at 2f, whether or not there is absorption - creating its own 2f-WMS signal via the lock-in's multiplication. This is known as Residual Amplitude Modulation (RAM) noise, and is calculated from Equation 5-23 for no absorbance (H terms zero), giving:

$$X_{2f}^{0} = \frac{G\bar{I}_0 i_2}{2}\cos\psi_2 \tag{5-24}$$

The implementation of the 2f-WMS model in LabVIEW, starts with entering the parameters required to specify a Gaussian or Lorentzian absorbance profile. These are the centre wavenumber, FWHM, line strength, number density and path length. The initial and final wavenumbers are entered for the 2f-WMS lineshape, along with the number of wavenumber points to consider between and inclusive of these points. The program subsequently calculates the wavenumber increment in order to calculate the wavenumber values to be used.

Once the absorbance is specified, the modulation depth of the wavenumber oscillations to be modeled is chosen, along with the number of angle points to be used in calculating and plotting the Fourier amplitude H terms. For each wavenumber value on the absorbance lineshape, a point on each H curve is calculated and displayed. The value of the integrands in Equations 5-19 and 5-20 are calculated on a point by point basis, and these values summed by the trapezoidal rule to get the integral. The formula for the absorbance lineshape is used in the integrands so that oscillation beyond the defined wavenumber limits near the end points can be accommodated.

The next wavenumber from the absorbance profile is then fed in, and another point on the H curve calculated and so on. Once the H curves are found, Equation 5-23 gives the 2f-WMS lineshape interactively with the choice of values for the mixing parameters: i_1, i_2, ψ_1, and ψ_2.

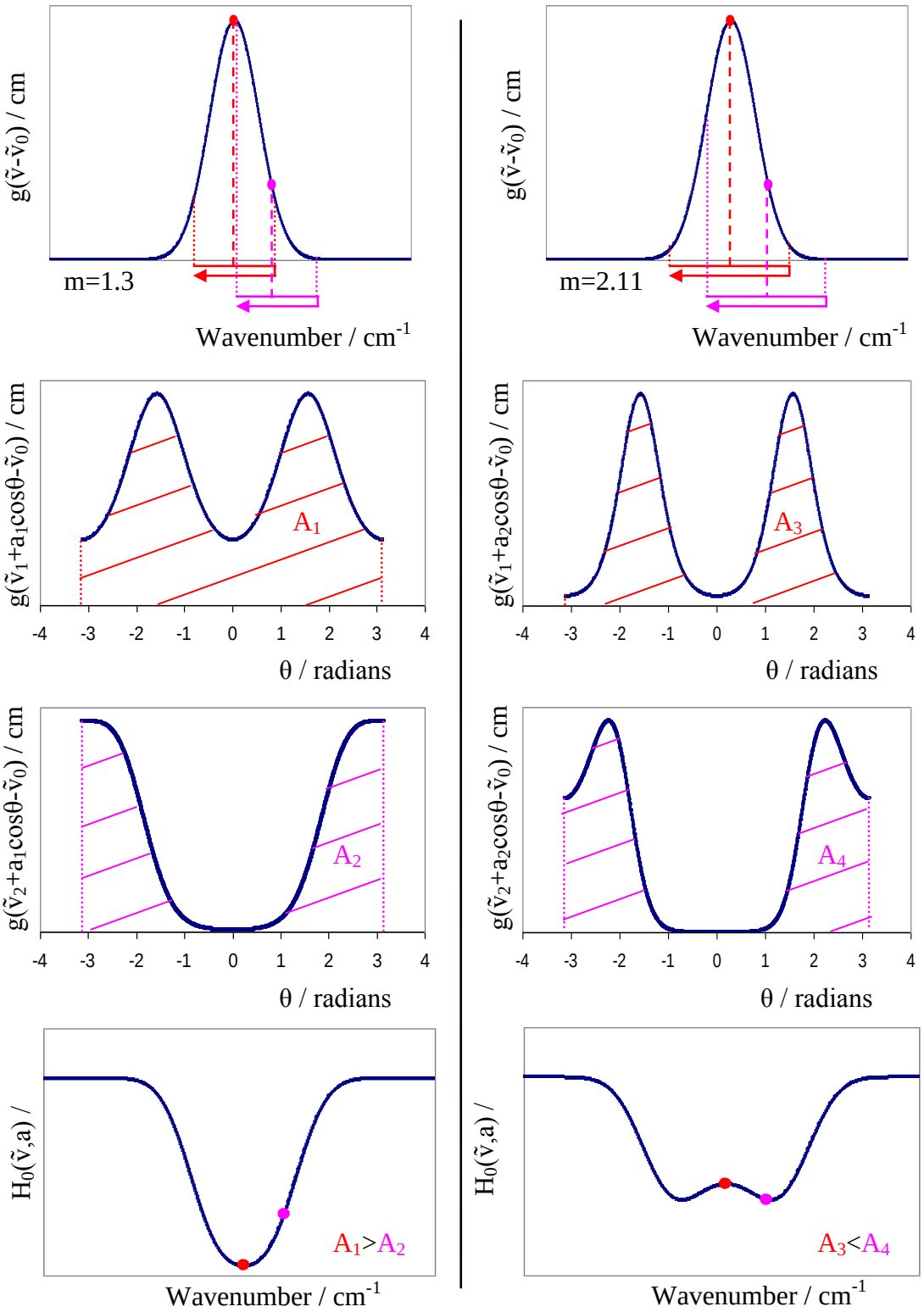


Figure 5-17: Calculating each point of the $H_0(\tilde{\nu},a)$ term involves integrating the curve swept out with respect to θ during one oscillation about a wavenumber on the absorbance lineshape. For Gaussian and Lorentzian lineshapes a saddle point forms in the $H_0(\tilde{\nu},a)$ curve for a modulation index approximately >1.55. This implies that oscillation about the absorbance line centre produces less area than about neighbouring points. See text for detail.

5.9.1 Modeling results: Some general trends and observations

5.9.1.1 The H_0 Term

The H_0 term in the approximate model is worthy of a comment. This is the amplitude of the DC part of the signal arriving at the detector due to the absorbance (Equations 5-17 and 5-18). Without mixing, the lock-in amplifier would not pick out this term. However, with mixing, this part of the signal contributes to the overall 2f signal through the nonlinear (quadratic $i_2 \neq 0$) intensity modulation. The modeling shows that this term develops a saddle point, for both Gaussian and Lorentzian absorbances, above a modulation index of approximately 1.54, which deepens as the modulation index increases. This property of the curves was not *a priori* obvious.

Equation 5-19 picks out the DC part of the detected signal by finding the average of the detected absorbance over a modulation cycle. The integration limits of $-\pi$ to π takes $\tilde{\nu}$ in Equation 5-6 from the lowest wavenumber of the modulation, $\tilde{\nu} - a$, to the highest, $\tilde{\nu} + a$, and back again. The integration therefore gives the area under a curve traced out from the absorbance profile as the wavenumber sweeps between its limits. The saddle point implies that the area formed from oscillating about the absorbance line centre is less than the area formed from around neighbouring points - for modulation indices, m, above approximately 1.54. Explicit calculation confirms this conclusion. Figure 5-17 shows simulations for two points on a Gaussian absorbance profile for modulation indices of 1.3 and 2.11. The equivalent points on the H_0 curves are proportional to the swept out areas through Equation 5-19.

5.9.1.2 The H_2 Term

Turning to the H_2 term, this term is normally dominant in the 2f-WMS signal and so gets particular attention in the literature. H_2 is the amplitude of the 2f component of the signal at the detector. If there were no mixing from the laser's intensity modulation with tuning, H_2 would be the only Fourier component output from the

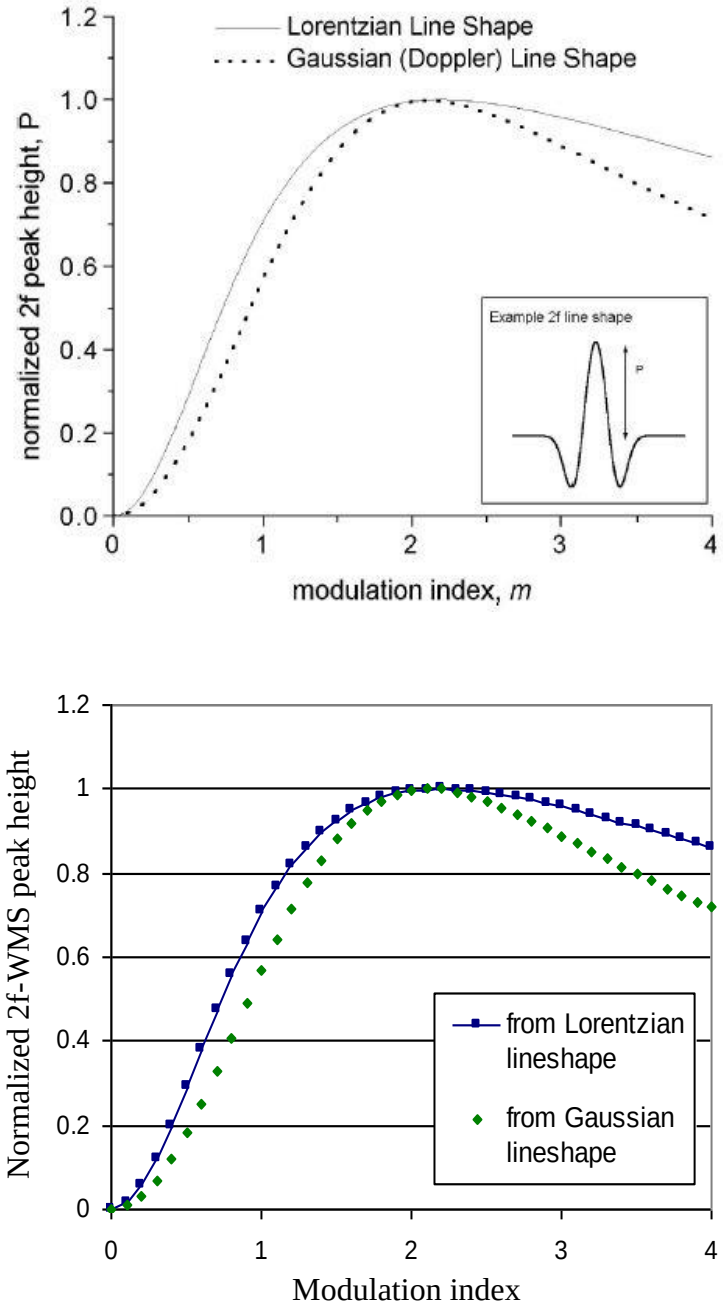


Figure 5-18: Top Plot: (Reproduction of Figure 5-5). 2f-WMS H_2 peak heights versus modulation depth derived from Gaussian and Lorentzian lineshapes via Equation 5-20 for k=2 and a single line. The figure is normalized so that the maximum peak height equals 1 for each lineshape. (Figure from Liu et al., 2004a).

Bottom Plot: Normalized 2f-WMS H_2 peak heights versus modulation depth, produced from the custom written modeling program. The general trends of Liu's plot are reproduced. The maximum 2f-WMS signal occurs in the Lorentzian case for a modulation depth of 2.20, and in the Gaussian case for a modulation depth of 2.11.

lock-in amplifier in the small absorbance approximation. The term is calculated from Equation 5-20 with $k = 2$. Wilson (1963) used (trapezoidal) numerical integration to calculate H_2 for Gaussian and Lorentzian absorbance lineshapes. Reid and Labrie (1981) followed this approach for Gaussian lineshapes, and used an analytical expression for the Lorentzian lineshapes due to Arndt (1965). Reid and Labrie stated that the modulation depth that maximized the H_2 peak height for both lineshapes was $m = 2.2$. Liu et al's paper (2004a), from which the top plot of Figure 5-18 was taken, is a little less definite, writing that the 2f-WMS peak height was maximized at $m \sim 2.2$. In this thesis, trapezoidal numerical integration was used to evaluate H_2 for both lineshapes. In the case of the Lorentzian absorbance, the maximizing m was found to be 2.20. However, for a Gaussian absorbance, the H_2 maximizing modulation index was found to be 2.11. Nevertheless, the bottom plot of Figure 5-18 was generated from the 2f-WMS modeling program and reproduces the general trends seen in Liu et al's diagram (top plot). Also, the model's $m = 2.11$ unnormalized (Gaussian) H_2 peak height, when prefactors are taken out to match Reid and Labrie's definition of H_2 , occurs at their value of 0.438.

5.9.1.3 Comparing the approximate and full 2f-WMS models

Figure 5-19 shows a comparison of curves generated by the programmed approximate and full 2f-WMS models, when there is no mixing of the absorbance lineshape with the laser's intensity modulation, for a Gaussian absorbance of FWHM =0.004cm^{-1}. Residuals are also shown. The approximate model shown is Equation 5-23 with $G = \bar{I}_0 = 1;\ i_1, i_2, \beta = 0$, and is hence equivalent to $H_2/2$. The full model shown is Equation 5-16 with $G = \bar{I}_0 = 1;\ i_1, i_2, \beta = 0$. Agreement between the approximate and full models can be seen to be very good for a peak Gaussian absorbance of 0.01 (top plot), and a reasonable agreement is obtained at 0.1 absorbance units (middle plot). As the absorbance grows to 1 (bottom plot), and beyond, the approximate model increasingly over-estimates the 2f-WMS peak height (and hence the concentration).

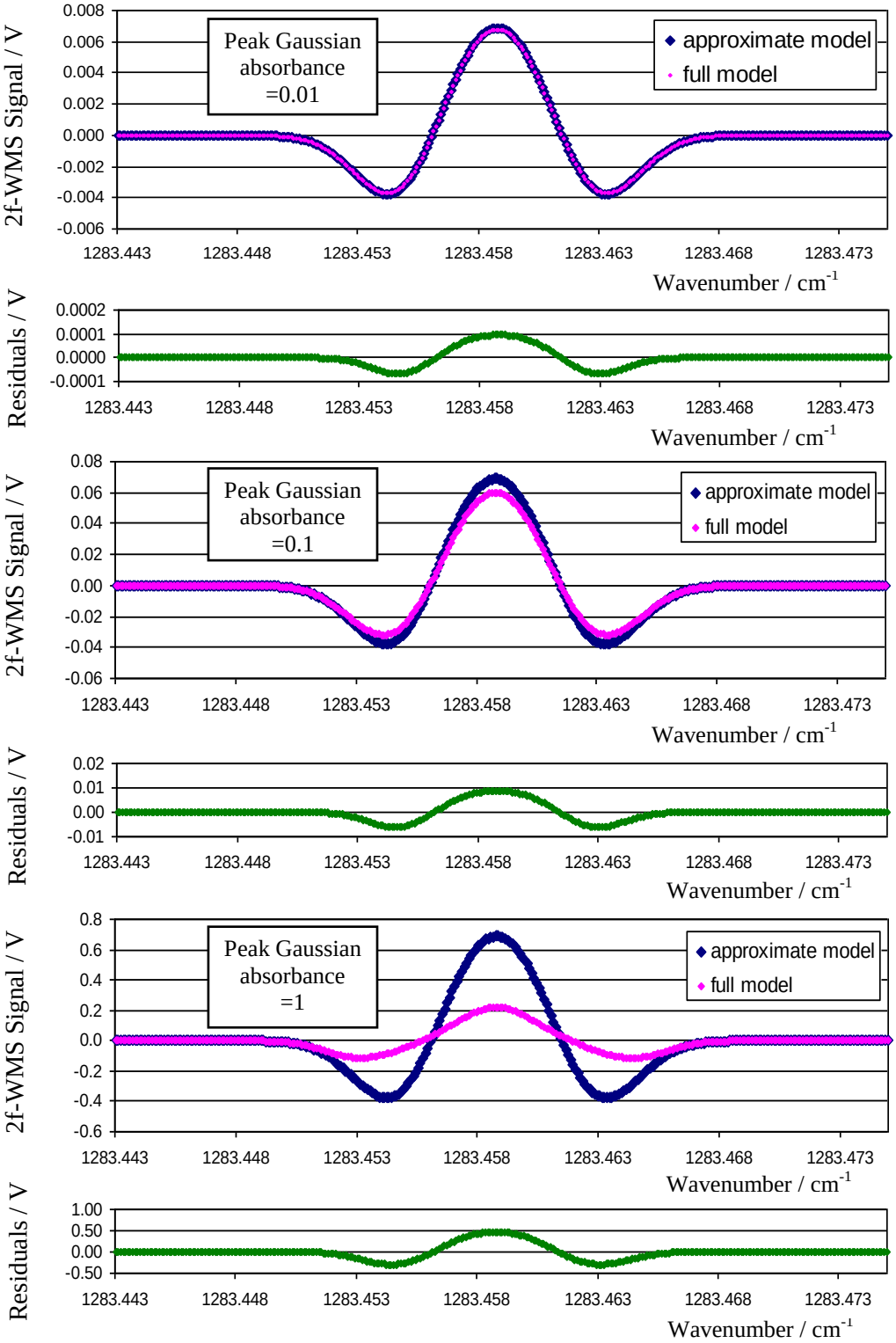


Figure 5-19: A comparison of the approximate and full 2f-WMS models for a Gaussian absorbance FWHM of 0.004cm^{-1} and peak absorbances of 0.01, 0.1, and 1. The residuals are also shown: approximate minus full model voltages. The approximate model shown is Equation 5-23 (with $G = \bar{I}_0 = 1,\ i_1 = i_2 = \beta = 0$) and the full model shown is Equation 5-16 (with $G = \bar{I}_0 = 1,\ i_1 = i_2 = \beta = 0$).

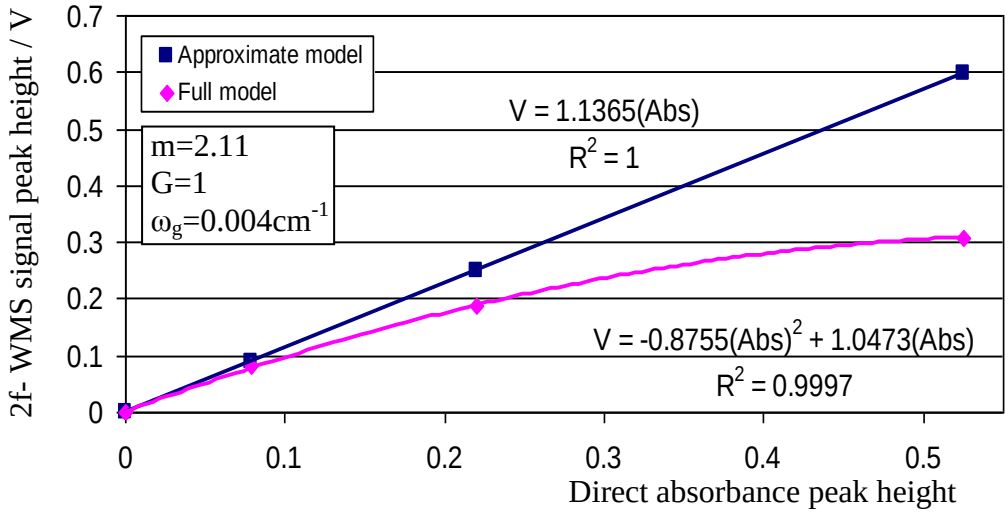


Figure 5-20: **The predicted approximate and full model calibration trends – linear and quadratic, respectively – taken over the range seen in Figure 5-12. The direct signal from which the 2f-WMS model lineshapes were calculated included a realistic quadratic direct baseline:**

$$I_0(\tilde{\nu}) = -1.955554773(\tilde{\nu})^2 + 5017.733087(\tilde{\nu}) - 3218732.186$$

Other model parameters: $\omega_g = 0.004 cm^{-1}$, $m = 2.11, G = 1$; $i_1, i_2, \psi_1, \psi_2, \beta = 0$

Note that the direct absorbance peak height is the dependent variable, and is thus displayed on the abscissa. Figure 5-12's calibration axes are displayed the other way around – since that seems a natural way to present a look-up calibration chart, for converting measured 2f-WMS values to absorbances.

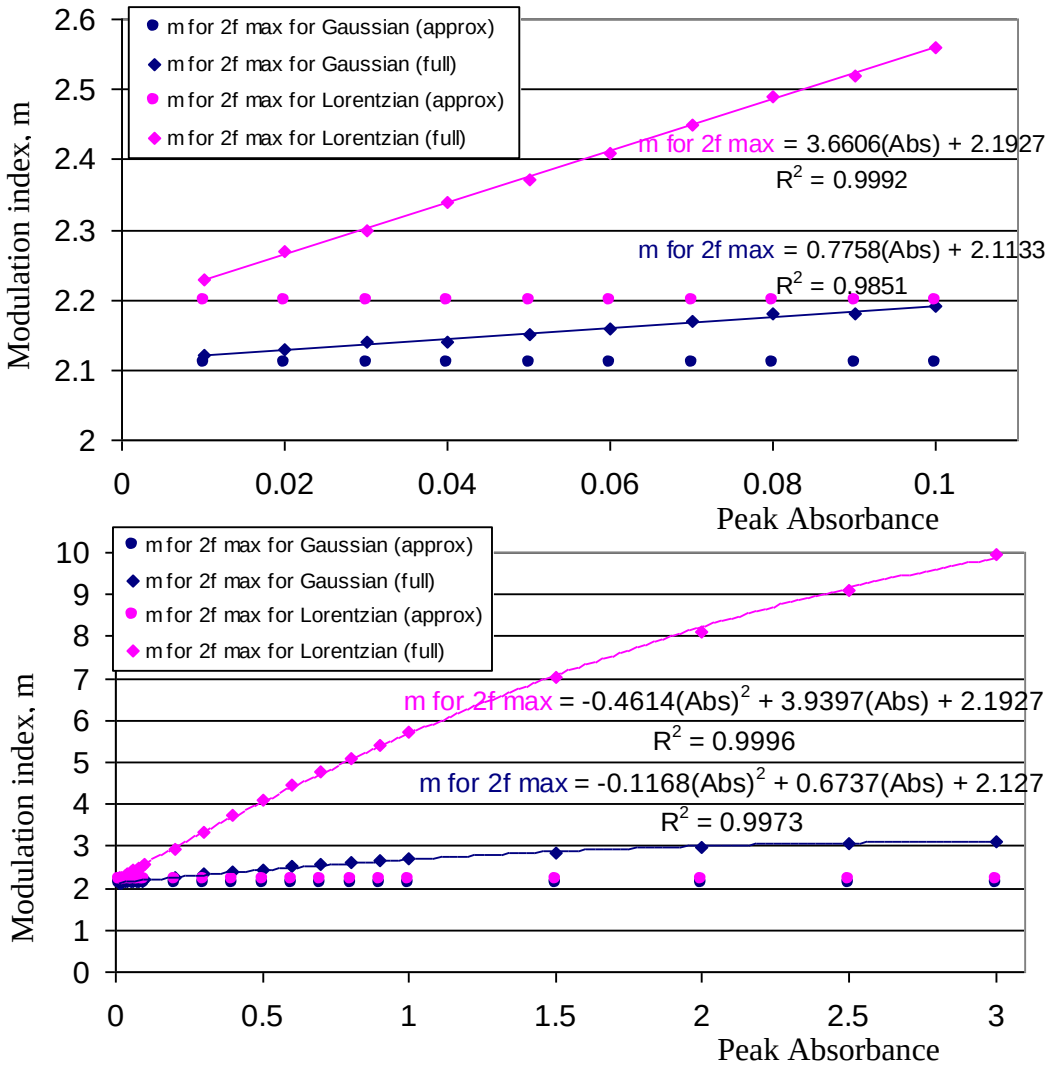


Figure 5-21: Plots showing the trend of the modulation index, m, needed to maximize the 2f-WMS signal, against the peak absorbances of the Gaussian or Lorentzian lineshapes that the WMS signal is derived from.

Top plot: Peak absorbances from 0.01 to 0.1 absorbance units. In the approximate model of Equation 5-23 (with $G = \bar{I}_0 = 1$, $i_1 = i_2 = \beta = 0$), the modulation index needed for maximizing the 2f signal is 2.11 for the signal derived from a Gaussian, and 2.20 for a signal derived from a Lorentzian - for any peak absorbance in the range. In the full model of Equation 5-16 (with $G = \bar{I}_0 = 1$, $i_1 = i_2 = \beta = 0$), the maximizing modulation index increases linearly with the Gaussian or Lorentzian peak absorbance.

Bottom plot: An extended view of the top plot showing a peak absorbance range from 0.01 to 3 absorbance units. The approximate model continues the constancy of the 2f signal maximizing modulation indices. The full model shows quadratic trends in the 2f signal maximizing modulation index against peak absorbance.

The predicted calibration trends are shown in Figure 5-20 (for $m = 2.11$, $\omega_g = 0.004 cm^{-1}$), over the direct absorbance range used experimentally in Section 5.8. The approximate (small absorbance) model predicts a linear relationship overall all ranges. The full model predicts a quadratic trend, although this is approximately linear for small absorbances.

Figure 5-21 shows the modulation index, m, needed to maximize the 2f-WMS signal resulting from a peak direct absorbance between 0.01 to 3 absorbance units. The cases shown are for changes in intensity due to the absorbance profile – that is, for no mixing with the changing output laser intensity. The approximate model of Equation 5-23; predicts that m remains constant, independent of peak absorbance. The full model of Equation 5-16 exhibits a more involved behaviour. Between 0.01 and 0.1 absorbance units (top plot) the 2f-WMS maximizing m value increases linearly. The divergence of the full model from the approximate model is larger for the Lorentzian than the Gaussian profile. Between 0.01 and 3 absorbance units (bottom plot) the trend in increasing 2f-WMS maximizing m values, for the full model, is well described by quadratic polynomials.

5.9.1.4 The effect of a slowly changing baseline

The program allows the effect of a slowly changing baseline to be modeled through the $\bar{I}_0(\tilde{\nu})$ term in Equations 5-9, 5-13, 5-14, and 5-23. The baseline can be set to a constant value. Alternatively, a linearly sloping baseline between the required wavenumber range can be chosen by entering corresponding start and finishing voltages. Finally, a quadratic baseline can be entered by choosing the three constants specifying a quadratic curve. Practical values for these constants can be found from the power fitting of etalon fringes in the data analysis program of Chapter 3 (Section 3.4.5).

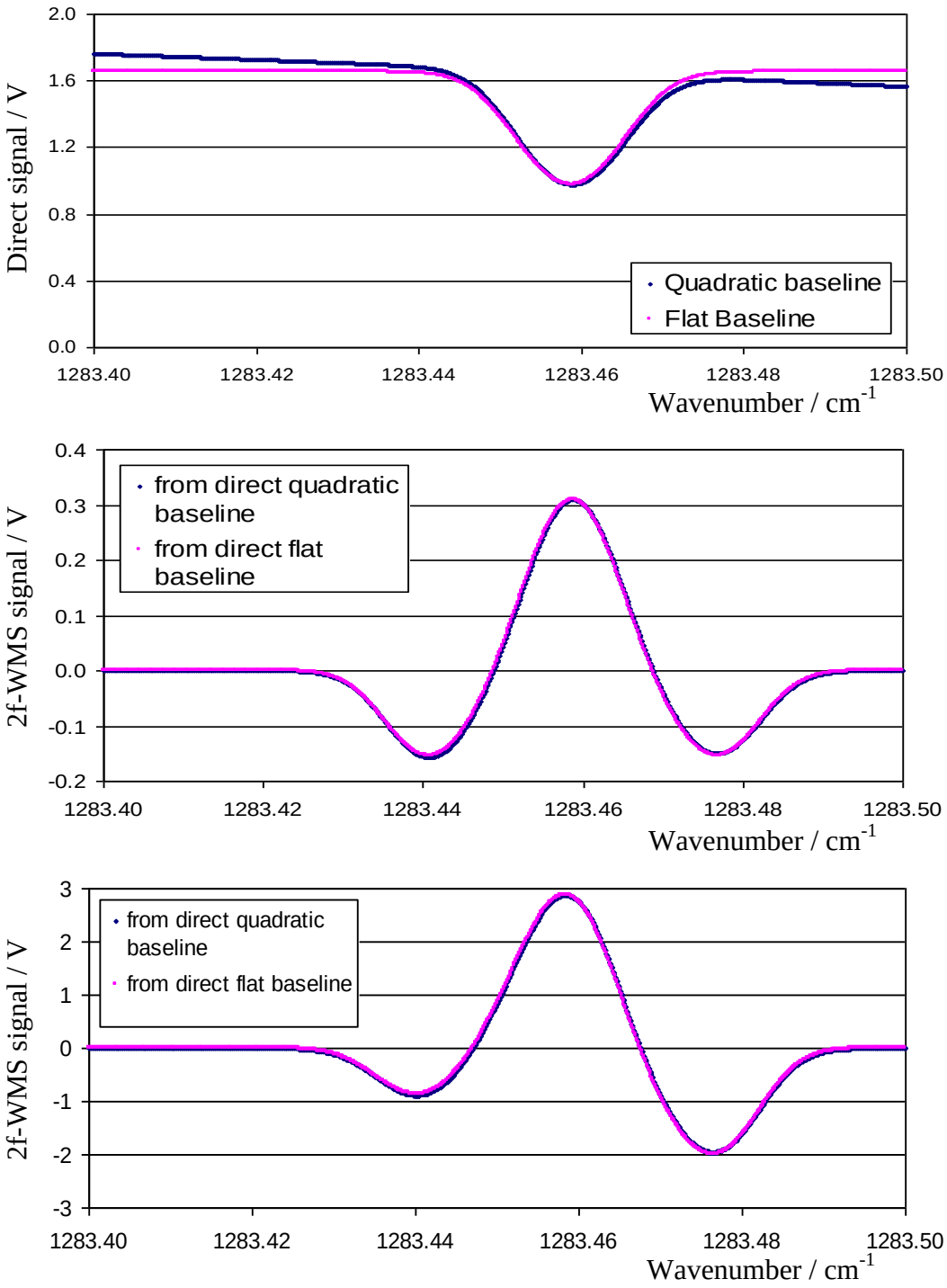


Figure 5-22: Plots illustrating the negligible effect of direct baseline curvature. Each plot shows two spectra for the same absorbance profile (top plot): for a quadratic direct baseline (blue), and a flat direct baseline (pink). The values of the baselines are the same at line centre. All plots: direct absorbance peak height $=0.525$**,** $m=2.11$, $i_2,\psi_1,\psi_2,\beta=0$**. Middle plot:** $i_1=0$, $G=1$**. Bottom plot:** $i_1=0.37$, $G=9.29$**.**

Quadratic baseline: $I_0(\tilde{\nu})=-1.955554773(\tilde{\nu})^2+5017.733087(\tilde{\nu})-3218732.186$**.**

For a wavenumber downchirp with the cw-QCL, the laser intensity increases as the injection current increases to increase temperature and down tune the laser. The baseline therefore enhances the low wavenumber side of the 2f-WMS signal relative to the high wavenumber side. This trend works against the asymmetry introduced by the H_1 term (Section 5.9.3). However, for this laser, the baseline effect is negligible. It can be discerned as a very small change in feet asymmetry (blue trace) in the middle and bottom plots of Figure 5-22. The middle plot concerns an otherwise symmetrical lineshape (pink trace), and the bottom plot concerns a 2f-WMS lineshape influenced by the H_1 term (pink trace).

5.9.2 Modeling the time signal at the detector

Throughout this chapter, the modeling assumption is made that the oscillation centre does not move significantly during the WMS oscillation period. This was modeled with a nested For loop. The inner loop effectively calculates the detector signal over an oscillation, for a fixed wavenumber, multiplied by the reference sinusoid, and averages over an oscillation. The outer loop steps the wavenumber to the next value. This is Equation 5-9. The number of angle points and the number of wavenumber points can be set independently.

In practice, however, the wavenumber of the oscillation centre does move during the oscillation. This section presents modeling of the detector signal under these conditions. The data can then be averaged over an oscillation, and used to estimate the changes to the 2f-WMS signal which might result.

The detector signal, when the oscillation centre is moving, was simulated by rewriting the nested For loop. The reasonable assumption is made that the tuning rate of the oscillation centre is constant near the absorption line. This rate, at the ramp position of the absorbance used for the cw-QCL during the calibration experiment, can be found from the bottom right plot of Figure 5-8, and is -6.85cm^{-1} s^{-1}. A key parameter is the wavenumber range covered per oscillation. This is the product of the

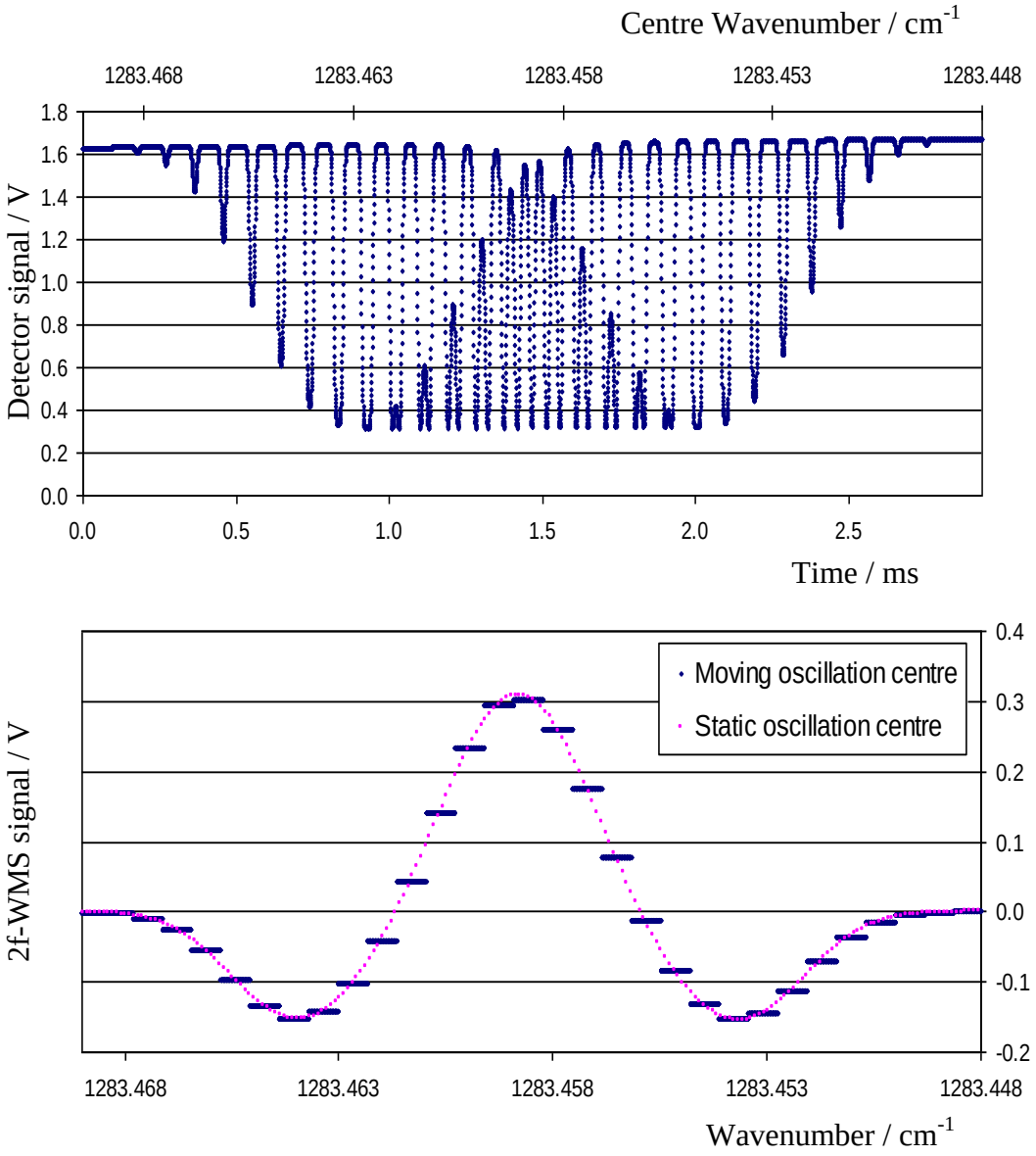


Figure 5-23: **Top plot: The simulated detector signal for a 10kHz, m=2.11, modulation with a centre of oscillation wavenumber moving at -6.85cm^{-1}s^{-1} (Direct signal FWHM=0.004cm^{-1}).**

Bottom plot: The detector signal's average per oscillation, associated with all the centre wavenumber values in that oscillation. This creates a step function 2f-WMS signal which is compared to the signal for (nearly) fixed oscillation centre.

Modeling Parameters: Wavenumber tuning rate = -6.85cm^{-1} s^{-1}, ramp frequency 5Hz (0mA to 420mA above threshold), modulation frequency 10kHz, direct absorbance peak height=0.525, $\omega_g = 0.004 cm^{-1}$, $m = 2.11$, $G = 1$, $i_1, i_2, \psi_1, \psi_2, \beta = 0$,

$I_0(\tilde{\nu}) = -1.955554773(\tilde{\nu})^2 + 5017.733087(\tilde{\nu}) - 3218732.186$.

tuning rate and the oscillation period, -0.000685cm^{-1} for the 5Hz ramp of Figure 5-8. The inner loop still steps through an oscillation cycle of angle values, from $-\pi$ to π, but the number of angle (time) points used, N_θ, is made equal to the number of wavenumber points used, $N_{\tilde{\nu}}$, over the wavenumber coverage per oscillation. In this way, the wavenumber advances the correct amount per oscillation. The outer loop steps the inner loop's starting wavenumber by the wavenumber coverage per oscillation. This calculates the detector signal, which is the integrand in Equation 5-9 (without the $G\cos(2\theta+\beta)$ multiplication or the averaging over a cycle). The angle increment, $\Delta\theta$, and wavenumber increment, $\Delta\tilde{\nu}$, are fixed by the number of data points set per oscillation, $N_{\tilde{\nu}} = N_\theta$:

$$\Delta\theta = \frac{2\pi}{N_\theta - 1} \tag{5-25}$$

$$\Delta\tilde{\nu} = \frac{(\text{wavenumber tuning rate}) \times (\text{oscillation period})}{N_{\tilde{\nu}} - 1} \tag{5-26}$$

The top plot of Figure 5-23 shows the simulated detector signal. A similar signal would result for a (nearly) fixed wavenumber centre of oscillation, but would be much more densely packed with oscillations. To get an idea of the change in 2f-WMS signal which might result, the simulated detector signal (for a moving oscillation centre) was averaged over each oscillation, and the appropriate average associated with each centre of oscillation wavenumber in the oscillation. This creates a step function 2f-WMS signal which is compared to the fixed oscillation centre signal in the bottom plot of Figure 5-23 (from a direct absorbance FWHM $\omega_g = 0.004 cm^{-1}$). A more sophisticated averaging would be needed to smooth the data. However, a visual inspection of the likely envelope of distortion, will indicate that the influence of the moving oscillation centre is modest compared to the very large instrumental broadenings which will be seen in Section 5.9.3.

5.9.3 Modeling results: Comparison with experimental data

The primary purpose of the 2f-WMS modeling was to enable the body of work to be carried forward. The modeling was necessarily completed after the experimental proportion of the work. It will be seen below that the (semiconductor diode laser based) models are useful in that they can be made to provide reasonable fits to the cw-QCL experimental data. The results are illustrative, rather than definitive, since there are many unknown and mutually interacting free parameters. In particular, to properly test the models the laser parameters - i_1, ψ_1, i_2, ψ_2 - would need to be known. Li et al. (2006) demonstrates a way in which these parameters could be measured. They split their laser beam onto two detectors, one path going through an etalon. The path without the etalon records the laser intensity (in volts), which is a sum of the ramp intensity plus the modulated sine intensity, as a function of time. The second beam path records the etalon fringes. Plotting the relative spacing of these fringes as a function of time reveals how fast the wavenumber is tuning. The intensity – wavenumber phase difference, for a particular modulation depth and modulation frequency, can be taken from equivalent points on theses plots (say, their maxima). The intensity plot can be fitted with a 1f sinusoid. The i_1 term can be found from this as the ratio of the sinusoidal amplitude to the ramp intensity average - near the wavenumber where an absorption feature will be introduced. The residual of the intensity variation and fit gives the intensity modulation from nonlinear contributions. The largest of these is usually the 2f modulation, so a 2f sinusoidal fit to the residual allows i_2 and ψ_2 to be found.

The largest measured 2f-WMS signal from the calibration results was chosen to demonstrate the models. This signal is denoted '2f1 modified' in Figure 5-13, Section 5.8. Similar conclusions apply to the smaller signals since they show similar shapes. The full model is used, but the approximate model gives the same results below approximately 0.1 absorbance units. The '2f1' transition is $^{12}CH_4$ υ_4 $P(4)A_2$ $1 \leftarrow A_1$ 1 at 1283.458791cm^{-1} (100% mixing ratio). The measured low pressure Gaussian FWHM for the direct signal corresponding to '2f1' was 0.00663cm^{-1}. Since

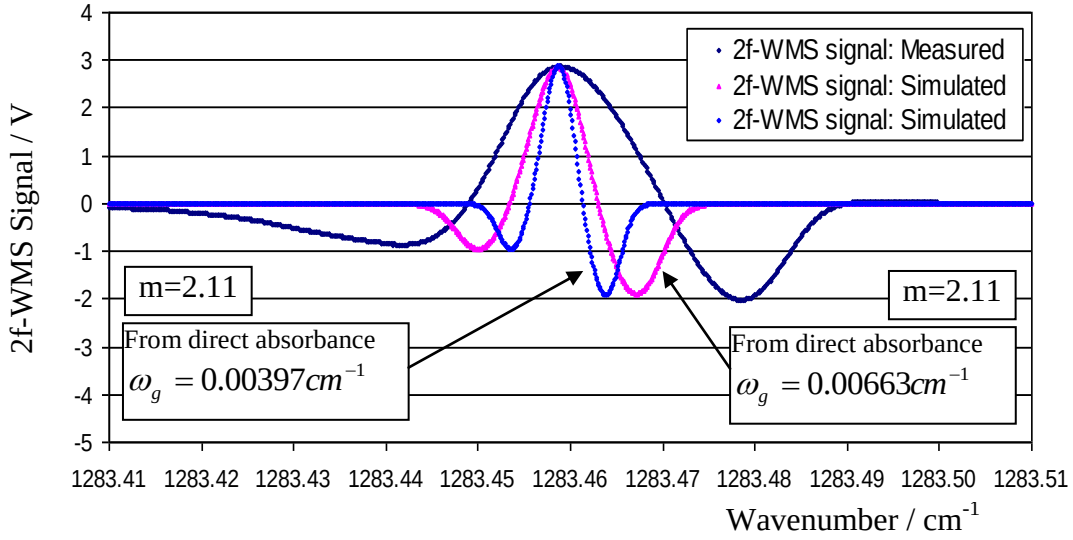


Figure 5-24: Measured versus simulated 2f-WMS spectra. Although the feet asymmetry is well captured, the 2f-WMS signal broadening is not. The simulated spectra were calculated for an assumed direct absorbance peak height of 0.525 absorbance units. The assumed direct FWHM's were those calculated and measured for the $^{12}CH_4$ υ_4 P(4)A_2 1←A_1 1 transition at 1283.458791cm^{-1}.
Model parameters:
$\omega_g = 0.00397cm^{-1}$, $i_1 = 0.32$, $G = 9.30$ **and** $\omega_g = 0.00663cm^{-1}$, $i_1 = 0.32$, $G = 9.22$.
Other model parameters: $m = 2.11$; $i_2, \psi_1, \psi_2, \beta = 0$.

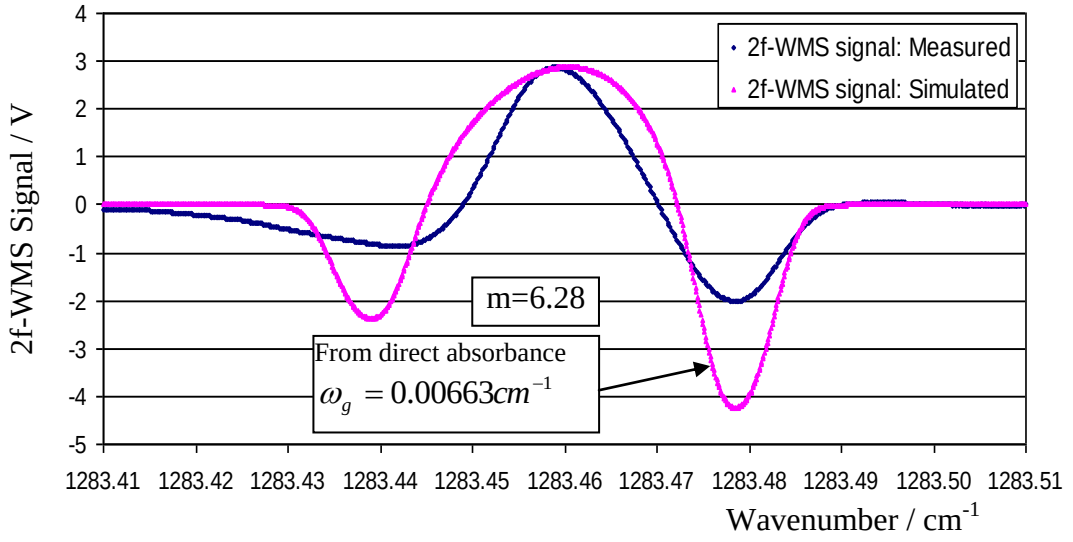


Figure 5-25: Measured versus simulated 2f-WMS spectra. The spectrum was calculated for a large over-modulation. The spectrum shape and the asymmetry of the feet are not good fits. The simulated spectrum was calculated for an assumed direct absorbance peak height of 0.525 absorbance units. The assumed direct FWHM was that measured for the $^{12}CH_4$ υ_4 P(4)A_2 1←A_1 1 transition at 1283.458791cm^{-1}.
Model parameters:
$\omega_g = 0.00366cm^{-1}$, $i_1 = 0.32$, $G = 15.58$, $m = 6.28$; $i_1, \psi_1, \psi_2, \beta = 0$.

this is broadened from the theoretical value of 0.00397cm^{-1} (Equation 2-20), the true direct peak height will be larger than the measured 0.525 absorbance units. The true peak height could be calculated from the true FWHM using Equation 2-41, with Equation 2-7 and $\tilde{\nu} = \tilde{\nu}_0$. However, the pressure gauge was not sensitive enough to measure the needed number density. For the purpose of modeling, the direct absorbance peak height was therefore taken as 0.525 absorbance units. The simulated 2f-WMS peak height was then made to agree with the measured signal by adjusting the lock-in amplifier gain term, G. (The G value needed doesn't reveal the peak height, since the direct and 2f-WMS signals are both considerably broadened). This lack of knowledge only affects the ordinate scaling, however, and the signals' line widths are of more interest in the analysis below.

Figure 5-24 shows the measured 2f-WMS spectrum against the simulated spectra for direct FWHM from theory (0.00397cm^{-1}) and measurement (0.00663cm^{-1}). It is immediately obvious that the 2f-WMS signal is significantly broader than either simulation. The electronic issues of noise in the signal chain and broadening (Section 5.7 and Section 4.6.4) therefore seem at play here. It may be that these instabilities are worse at lower sweep rates, as it takes the system longer to interrogate the lineshape. Another disadvantage of AC-coupled detection is that the direct signal can not be observed with a ramp of 5Hz, as used here.

However, the asymmetry in the feet can be captured by the modeling. The i_1 and ψ_1 terms, in Equations 5-9 and 5-23, work in tandem to produce the bulk of the asymmetry. i_1 sets the maximum asymmetry possible for a given ψ_1, and this phase difference between the intensity and wavenumber modulation determines how much of the maximum asymmetry is realized. For the purpose of Figure 5-24, ψ_1 was set at zero to give the maximum asymmetry possible through adjusting the i_1 term. The i_1 term was then adjusted to try to match the feet depths.

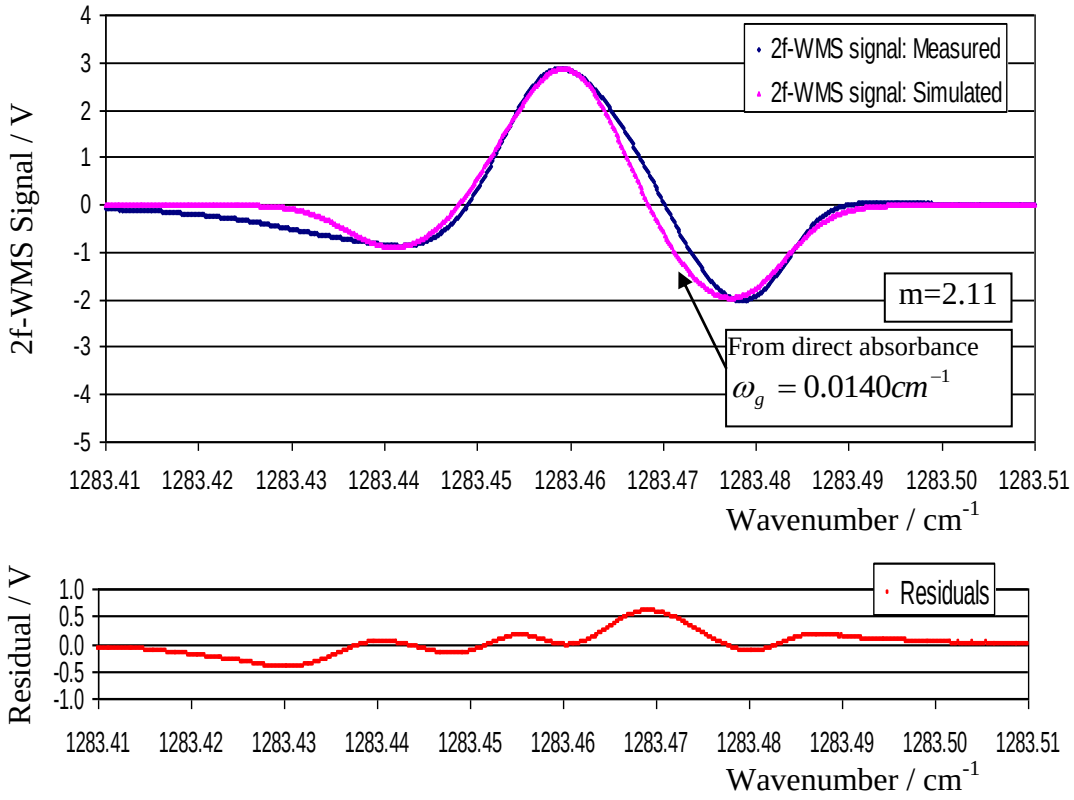


Figure 5-26: Measured versus simulated 2f-WMS spectra, and residuals. An assumed direct FWHM broadened to ω_g=0.0140cm^{-1} provides a reasonable fit to the measured 2f-WMS lineshape, and the asymmetry in the feet is captured. The simulated spectrum was calculated for an assumed direct absorbance peak height of 0.525 absorbance units. The transition is $^{12}CH_4$ υ_4 P(4)A_2 1←A_1 1 at 1283.458791cm^{-1}.

Model parameters:

$\omega_g = 0.0140cm^{-1}$, $i_1 = 0.37$, $G = 9.29, m = 2.11$; $i_1, \psi_1, \psi_2, \beta = 0$.

Another general observation is that the measured mean 2f1 baseline is close to zero (-0.028V). This remains so after subtracting a simulated sine wave to reduce the systematic noise (Section 5.8.1 and Figure 5-24). One effect of the i_2 term is to add a DC offset (RAM signal) (Equation 5-24). Therefore, the non-linearity in this laser's intensity output over an oscillation is therefore small, and negligible compared to the linear intensity modulation from i_1.

Returning to the broadness of the 2f-WMS measured signal, overmodulation (modulation depth >2.11) was ruled out as a likely source. It was always unlikely as the roll over of the signal strength at line centre (Figure 5-18) could be seen experimentally, by adjusting the dial controlling the modulation current. Figure 5-25 shows the very large overmodulation (m=6.28) needed to get close to the feet spacing for a direct signal FWHM of 0.00663cm^{-1}. No choice of m value correctly matched the feet spacing, and no choice of i_1 could capture the feet depth asymmetry. The shape of the simulated signal is also wrong. Similar considerations hold for considering a direct FWHM of 0.00397cm^{-1}.

A far more satisfactory fit was obtained by assuming an effective, broadened, direct FWHM of 0.014cm^{-1} at the optimum modulation depth of m=2.11. Figure 5-26 shows that the feet spacing and depth asymmetry is much more satisfactory, and the lineshape is much more closely matched. The modeling is therefore shown to be useful in the context of WMS with cw-QCL's.

5.10 Conclusions

A comparison was made between the Limit of Detection (noise floor) of the Sweep Integration and WMS techniques using the cw-QCL laser of Chapter 4. Because the technique of Allan variance was not available, the comparison was made using convenient parameters for the techniques with the equipment available (Table 5-1). Because of electronic incompatibility in the signal chain, and possibly due to using low kHz frequencies compared to the detector bandwidth, high noise dictated that WMS be combined with averaging to produce a useable signal.

Sweep Integration using 512 averages had a Limit of Detection of 0.080 absorbance units with an integration time of 51ms (Chapter 4). WMS combined with 512 averages had a Limit of Detection which was a modest 4.7 times better at 0.017 absorbance units. This comes at the expense of needing calibration and phase sensitive detection, and a 2000 fold increase in integration time to 102.4s – caused by the averaging of a slow tuning current ramp.

These Limits of Detections are somewhat behind the present pulsed-QCL systems of Cascade Technologies, but nevertheless comparable to those systems at an earlier stage of their development. The noise floor of the CT3000 pulsed-QCL sensor of Chapter 7 is 0.004 absorbance units in 10ms.

The calibration of WMS voltage to absorbance units, established by comparing Sweep Integration peak heights to WMS peak height's was linear (R^2~1). This was established using the $^{12}CH_4$ υ_4 $P(4)_{A2\ 1\leftarrow A1\ 1}$ transition at the low pressure of ~2Torr.

Additional programs were written in LabVIEW for modeling 2f-WMS lineshapes and for calculating Allan variances (Section 5.9 and Section 4.8). The lineshape modeling was illustrative, rather than definitive, due to the number of free parameters in this particular experiment. Nevertheless, the models were shown to be useful in the context of WMS with a cw-QCL laser. The 2f-WMS signal maximizing

modulation index, m, was found to be 2.11 in the Gaussian case, contrary to the 2.2 literature value.

5.11 Suggestions for further work

It would be useful to further investigate the electronic problems in the signal chain. This might lead to a lower noise floor and enable the WMS technique to be used without as much averaging – dramatically cutting the integration time.

If software became available for continuous concentration measurements with the cw-QCL, an Allan variance analysis could be carried out for optimizing the integration time for both techniques (using the program of Section 4.8). A more thorough comparison of the two techniques' Limits of Detections (noise floors) could then be made.

The analysis could also be carried out at higher pressures closer to the 200Torr to 760Torr used in many of Cascade Technologies' systems. Line broadening would imply lower peak heights, and hence the noise floor would be reached at higher concentrations for a given absorption line.

The 2f-WMS model for diode lasers could be used in future work - as the starting point for describing the 2f-WMS lineshape for both diode lasers and QCL's.

Of perhaps more academic interest than immediate commercial interest, would be avoiding the complication of simultaneous intensity modulation – at the expense of more complex equipment and alignment sensitivity.

For instance, a QCL can be used as the gain medium in an external cavity arrangement, EC-QCL, by antireflection counting the laser chip. Reflection of a constant intensity beam then takes place between two external mirrors. Tuning of a

cavity mode under the QCL's gain profile is achieved via the rotation of a diffraction grating behind the non-output mirror, feeding a particular wavenumber back into the cavity. Very large continuous tuning ranges can be achieved by this method. (>120cm^{-1} mode hop free and >500cm^{-1} overall (Daylight Solutions, 2015)). Wavenumber modulation can be achieved by modulating the current to the QCL chip (Hancock et al., 2009), which modulates the intensity of the beam. However, modulation can also be achieved by using piezoelectric transducers to dither the cavity length via the mirror separation or dithering the grating angle, or a combination of methods (Wysocki et al., 2005, Karpf and Rao, 2009, Chao et al., 2012).

Alternatively, a Stark cell uses a modulated field between two parallel electrical plates to modulate the energy spacings, and hence frequency, of the gas' molecular transitions. The frequency of the laser beam remains fixed.

Acousto-optic modulation provides another possible approach, using a radio frequency wave of angular frequency Ω in a piezoelectric crystal, such as lithium niobate. The crystal diffracts a fixed frequency laser beam, of angular frequency ω, into a number of spatially separated beams of different frequency. The -1st order beam has an angular frequency of $\omega - \Omega$. The beam can therefore be modulated by controlling Ω.

A relatively new technique making use of acousto-optic modulation and a cw-QCL is 'Chirped Laser Dispersion Spectroscopy' (CLaDS) (Wysocki and Weidmann, 2010). The frequency of the laser beam, ω, is chirped over an absorption feature by fast changes of injection current, while Ω remains fixed. The zeroth and minus first order diffracted beams, of frequencies ω and $\omega - \Omega$, are recombined onto a detector. The beat note, $f(\omega)$, of the resulting heterodyne signal contains dispersion information due to the molecular absorption. In the general case, both beams pass through the gas cell. The beat note is demodulated into a voltage amplitude

proportional to the beat notes' instantaneous frequency (the rate of change of phase divided by 2π), and recorded as a function of changing laser angular frequency, ω. For a linear frequency sweep (implying a constant chirp rate), the instantaneous beat frequency is:

$$f(\omega)=\frac{1}{2\pi}\left[\Omega+\frac{\Delta L}{c}\frac{d\omega}{dt}-\frac{L_C}{c}\omega\frac{d\omega}{dt}\left(\left.\frac{dn}{d\omega}\right|_{\omega-\Omega}-\left.\frac{dn}{d\omega}\right|_{\omega}\right)\right] \tag{5-27}$$

L_c is the geometrical path length of the beam through the cell, and L is the out of cell path length L. Any difference in out of cell path length between the two laser beams is accounted for by ΔL, which can be eliminated by balancing the arms of the spectrometer.

Equation 5-27 against ω is essentially a difference of dispersion signals, $dn/d\omega$. The technique's sensitivity enhancement comes from the dependence of the dispersion signal on the laser chirp rate, $d\omega/dt$, which can be high with QCL's. The absorbance, and hence gas concentration, responsible for the signals can be found by fitting a dispersion model.

For a dilute medium the Kramers-Kronig relations reduce to a single expression which depends on the absorbance:

$$n(\omega)=1+\frac{c}{\pi}\int_0^{+\infty}\frac{\alpha(\omega')}{\omega'^2-\omega^2}d\omega' \tag{5-28}$$

where

$$\alpha(\omega')=\frac{N}{V}LSg(\omega'-\omega_0) \tag{5-29}$$

α is adjusted to find n and $dn/d\omega$ until an acceptable fit is obtained. The method shows a high immunity to changes in optical power, and like 2f-WMS, reduces the

background the signal rides on – increasing sensitivity and reducing dynamic range issues.

A version of the technique, in some ways analogous to WMS, has also been implemented, by sinusoidally modulating the chirp rate of the laser by means of the injection current. This is known as ‘Chirp Modulated – ClaDS’ (CM-CLaDS) (Nikodem et al., 2012, Nikodem and Wysocki, 2012).

Chapter 6

Proof of principle oxygen spectroscopy with a wavenumber prototype semiconductor diode laser

6 Proof of principle oxygen spectroscopy with a wavenumber prototype semiconductor diode laser

6.1 Abstract

The suitability of a continuous wave DFB semiconductor diode laser for oxygen spectroscopy was evaluated. This laser's wavenumber range was new to the manufacturer / supplier, and exhibited instability – particularly to a change of substrate temperature. The laser showed sidemode / multimode competition over much of its spectral coverage. However, the absorption due to the R21Q22 line at 13162.50095cm^{-1} was single mode and gave the most stable signal. Two other large absorbance spectral regions contained three absorption lines – which could be modeled through multi-mode superposition. Many desirable characteristics of performance have been captured in the present design. These include suitable wavenumber coverage, linear power and wavenumber tuning with injection current (at fixed substrate temperatures), and constant wavenumber tuning rates using triangular injection current ramps.

6.2 Introduction

Cascade Technologies wished to expand their capabilities beyond the mid-infrared, primarily to measure oxygen gas – a combustion product and an indicator of asphyxiants through mixing ratio. To this end, a continuous wave DFB semiconductor diode laser of nominal wavelength 759.7nm (13,163.09cm^{-1}) was provided by an existing supplier-manufacturer. However, the wavelength range was new to this manufacturer, so the laser's evaluation served both to guide Cascade Technologies' requirements and to provide feedback to the manufacturer on operational strengths and weaknesses to modify future designs.

The specification sheet was tentative, so FTS data was first collected to map the spectral coverage of the laser and to aid in identifying oxygen lines for spectroscopy.

FTS and opto-electrical data was also taken on the device's stability and operating characteristics, particularly after dual mode lasing became apparent.

The laser was set up for oxygen spectroscopy through a 30m optical path length Herriott cell. Three regions of large absorption in laboratory air were interrogated and compared to Voigt simulations at 760Torr and 200Torr. These regions consisted of a single mode spectrum of R21Q22 at approximately 1283.5cm^{-1}, and two multi-mode regions containing R17R17 (13156.623203cm^{-1}) and R19Q20 (13160.818147cm^{-1}), respectively. Each also contained a doublet formed from R19R19 and R17Q18 centered at about 13158.7cm^{-1}.

6.3 Fourier Transform Spectroscopy and opto-electrical tuning and stability data

The Fourier Transform Spectroscopy (FTS) technique was outlined in Section 4.5.3.1. Data was taken with a Mirmat 8025 FTS operated by Cascade Technologies, with the results processed and analyzed contemporaneously, so that the tuning data taken could be influenced and extended. In particular, since this diode laser was a prototype, the measurements were extended to include current, temperature, and wavenumber stability data.

The FTS data were taken at the instrument's maximum resolution of 0.5cm^{-1}, with boxcar apodization, 40 averages per interferogram – each consisting of 842 points, and an 'oversampling' of ×4 beyond the Nyquist frequency to improve resolution.

6.3.1 FTS Tuning Data

The supplier's tentative specifications set a substrate temperature range of +10°C to +20°C. The lowest integer temperature which remained stable on application of injection current was, however, +13°C. This foreshadowed stability issues, but was

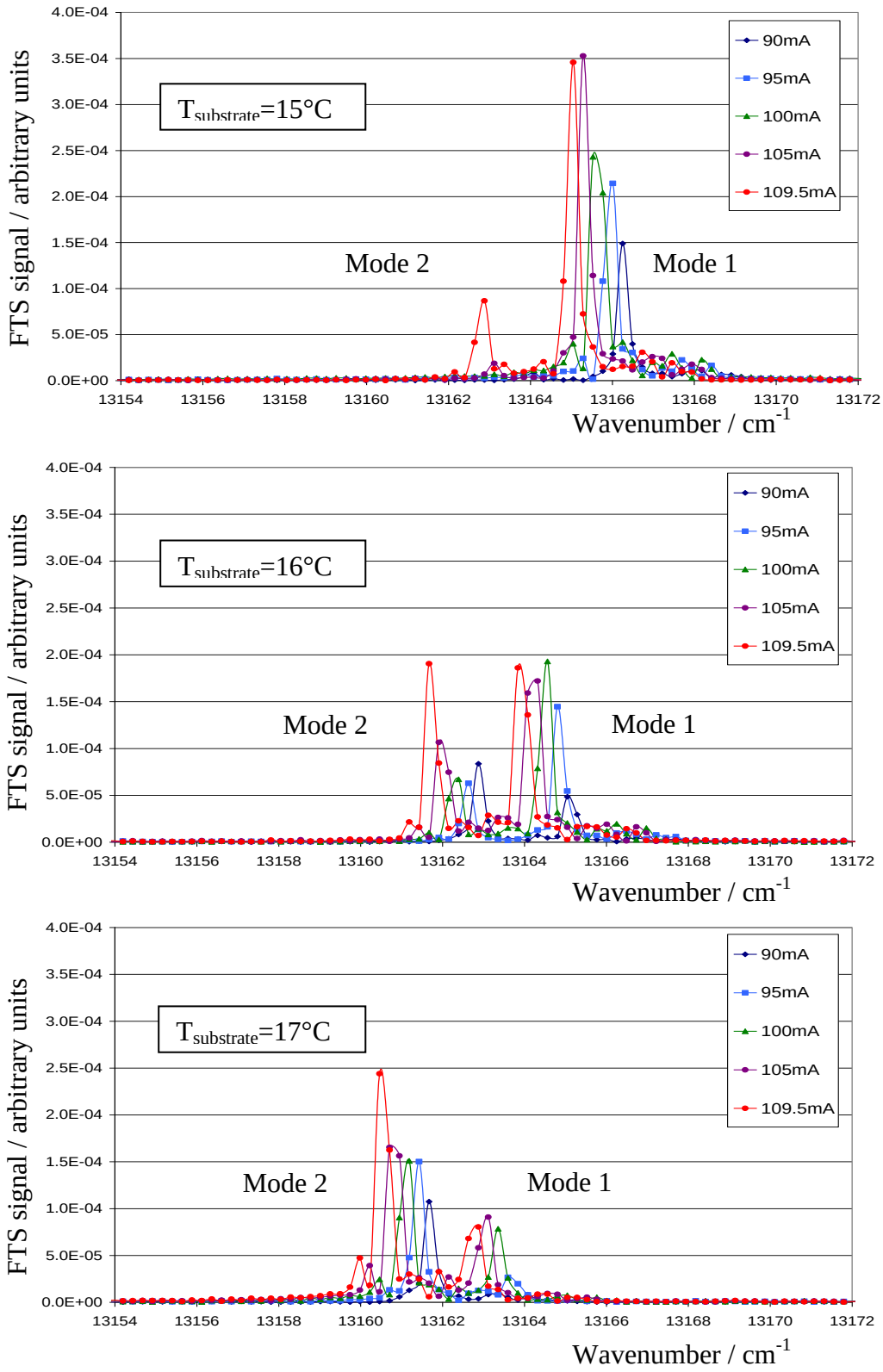


Figure 6-1: Dual mode behaviour at a substrate temperature of 16°C and the switching of the dominant mode between 15°C and 17°C, from FTS data. The data points are joined to aid visual interpretation.

no real restriction in terms of wavenumber as +10°C to +14°C turned out to be beyond the band head of the Oxygen A-band for the operating currents used (Section 6.7, Figure 6-15). Hence, this tuning range covered spectral regions containing no absorbances. The FTS data were taken from +13°C to +18°C (though for later spectroscopy this was extended to +19.5°C). The maximum specified injection current was 120mA. Threshold was found to be approximately 85mA, depending on substrate temperature. FTS data was taken at fixed integer temperatures with injection current in 5mA steps between 80mA and 110mA.

The data sets showed two modes in competition around +16°C, as the cavity modes tune under the laser's gain profile, with each mode exhibiting the linear tuning typical of a diode laser.

Figure 6-1 shows this behaviour. 'Mode 1' at higher wavenumbers dominates below +16°C, and 'Mode 2' at lower wavenumbers dominates from +17°C and above. At +16°C dual mode behaviour is apparent. The specification sheet's recommended continuous wave operation of +16°C and 106mA was therefore the least advantageous for basing a tuning sweep upon (Figure 6-7). Spectroscopy is still possible in multimode. If only one mode encounters an absorption feature during tuning, the time signal will be essentially a single mode one. If multiple modes encounter spectral features during tuning, the resulting time signal will be a superposition from different wavenumber regions (Section 6.8.3).

Figure 6-2 shows the FTS peak positions for each mode as a function of injection current for fixed substrate temperatures. Both modes' tuning rates drop in the mode change over region – bounded by and including the green trend lines at +16°C – to an average chirp rate of (-0.0492±0.0003)cm^{-1}/mA (mean ± standard deviation). Outside of this region Mode 1's chirp rate averages (-0.062±0.001)cm^{-1}/mA, and Mode 2 averages (-0.0638±0.0004)cm^{-1}/mA.

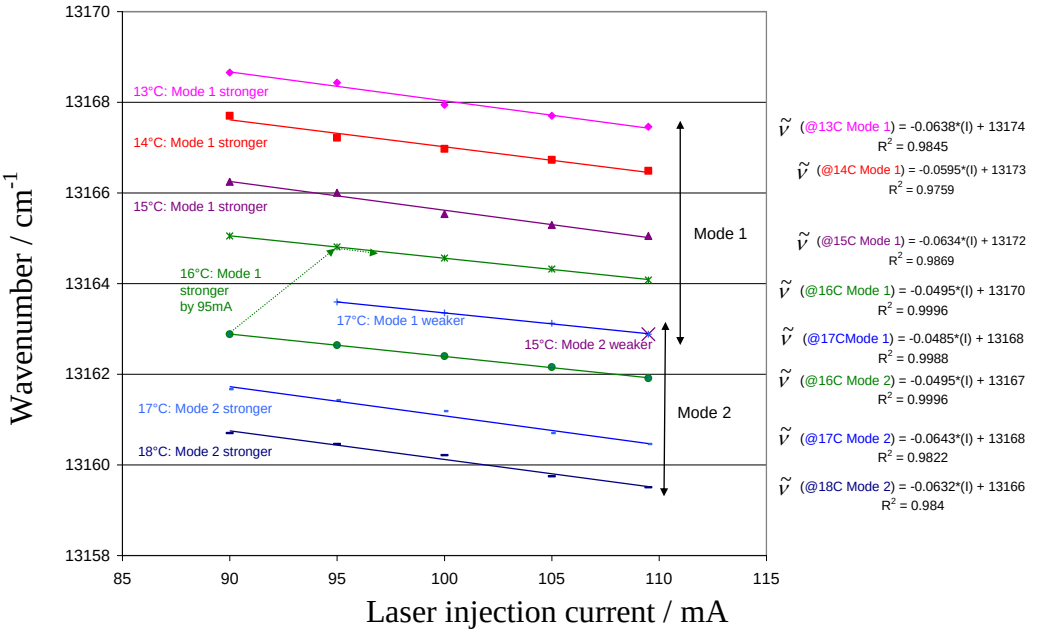


Figure 6-2: Static tuning data of wavenumber vs injection current at fixed substrate temperatures, with linear fits - summarized from FTS plots similar to those of Figure 6-1.

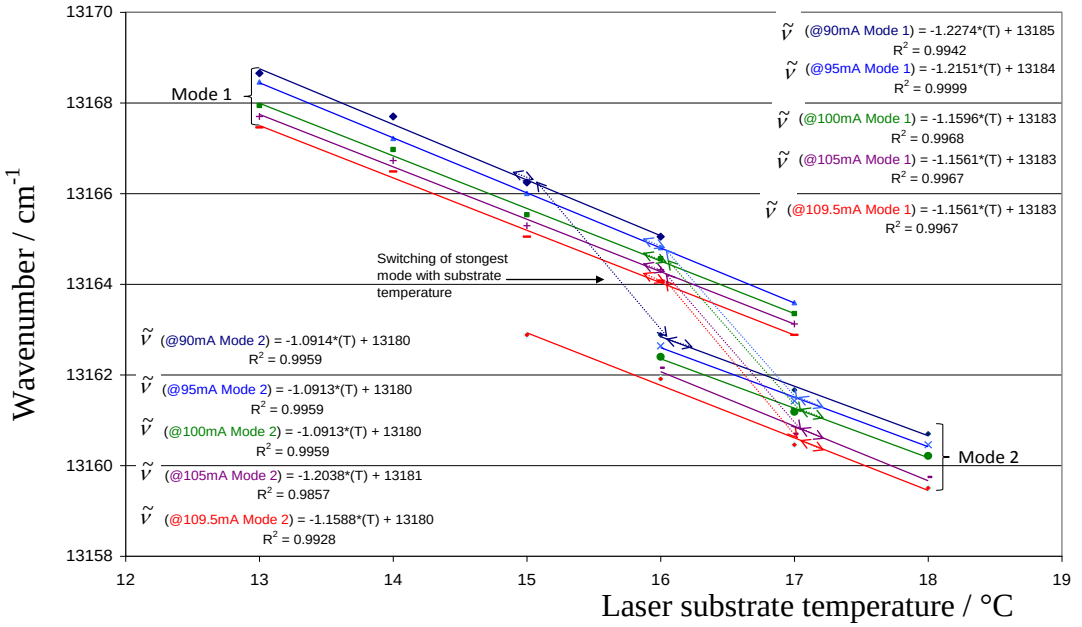


Figure 6-3: Static tuning data of wavenumber vs laser substrate temperature at fixed injection currents, with linear fits. Note this is data extrapolated from Figure 6-2. It makes the presence of two lasing modes clearer but fails to capture the instability, nonlinearity and hysteresis with substrate temperature present in Figure 6-9 and Figure 6-11.

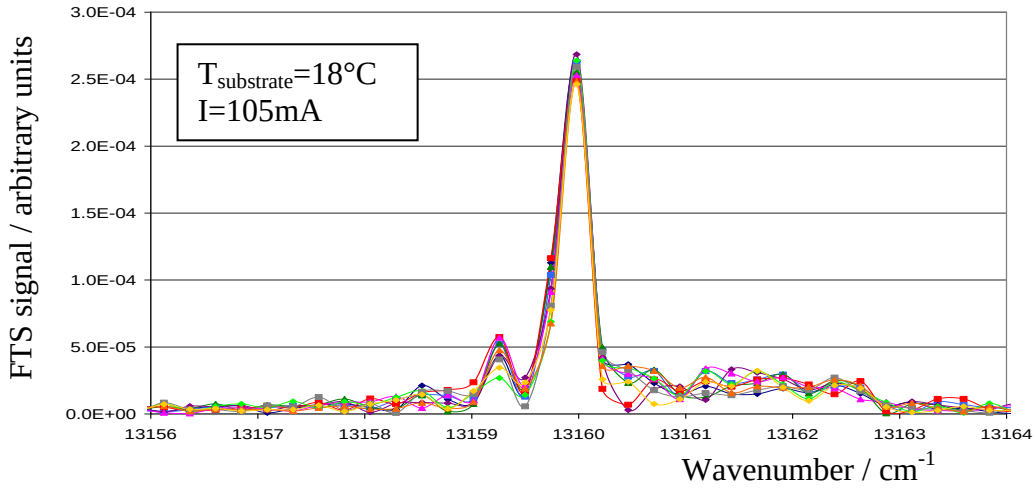


Figure 6-4: **Ten consecutive FTS spectra from single shot interferograms, showing a strong side mode and broadband output over a range of wavenumbers. Data points are joined for ease of visual interpretation.**

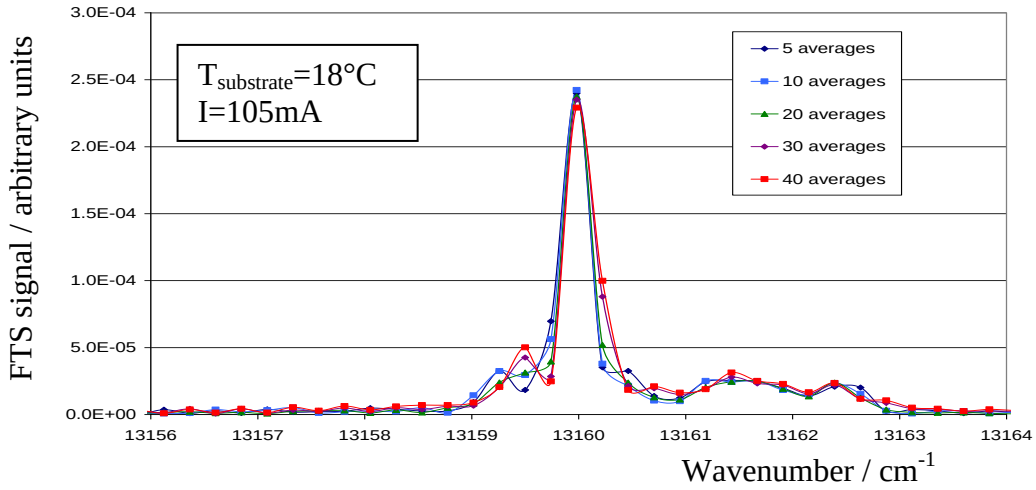


Figure 6-5: **Five FTS spectra from increasing averages of interferograms, showing a strong but variable side mode and broadband output over a range of wavenumbers. Data points are joined for ease of visual interpretation.**

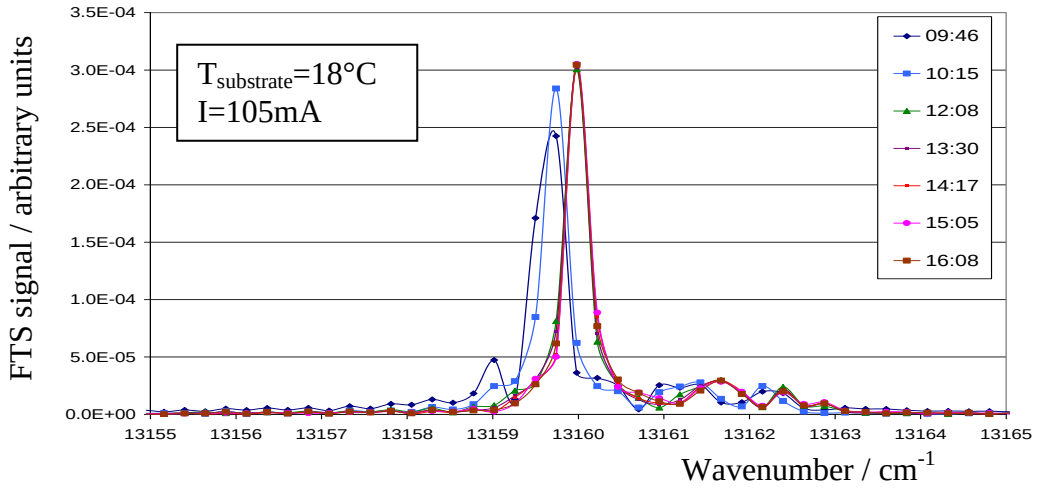


Figure 6-6: **FTS stability data at fixed substrate temperature and injection current where one main laser mode is dominant, taken at various times throughout one day. A side mode and broadband output can also be seen. 40 interferogram averages per spectra. Data points are joined for ease of visual interpretation.**

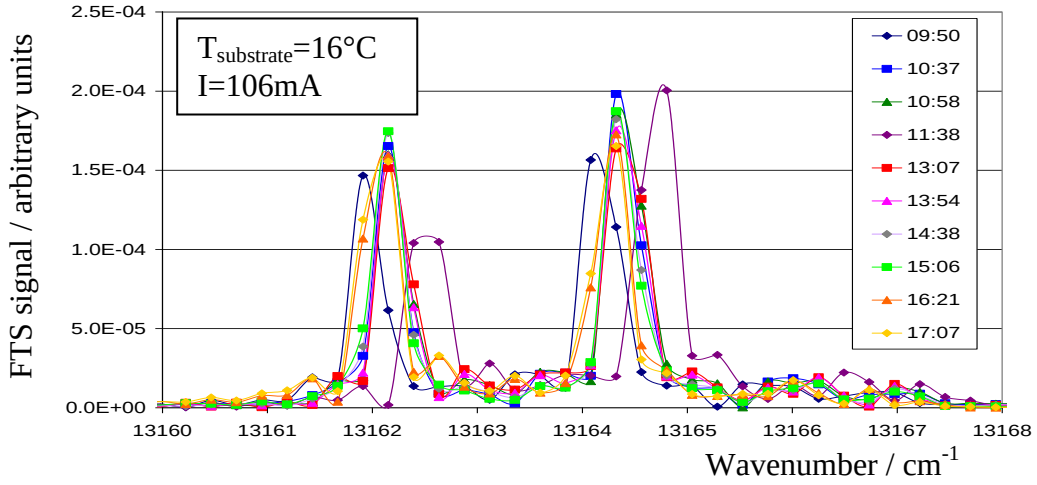


Figure 6-7: **FTS stability data at fixed substrate temperature and injection current where dual mode competition is most intense, taken at various times throughout one day. Side mode and broadband output can also be seen. 40 interferogram averages per spectra. Data points are joined for ease of visual interpretation.**

It would have been preferable to take separate FTS data on wavenumber versus substrate temperature at fixed currents. The foreshadowing of instability issues made hysteresis in the tuning likely, and this indeed proved to be the case when opto-electronic measurements were taken (Section 6.3.2). However, separate FTS data was not taken. Nevertheless, extrapolating the FTS data of Figure 6-2 gives a useful plot (Figure 6-3) which makes the presence of the two lasing modes more apparent.

Figure 6-4 shows ten FTS spectra taken back to back from single shot interferograms. Figure 6-5 shows spectra produced from averaging different number of interferograms. Both Figures show that the competition from the main side mode is variable. No information was available on the linearity, or otherwise, of the Mirmat 8025 FTS's amplitude scale. If linearity is assumed, the red trace of Figure 6-5 would suggest a minimum side mode suppression ratio (SMSR) of:

$$SMSR = 10\log_{10}(\text{amplitude of main peak}/\text{amplitude of side peak}) = 6.6dB \quad (6\text{-}1)$$

This would compare unfavourably with the 30dB typically seen in manufacturer's specification sheets for semiconductor diode lasers.

6.3.2 FTS stability data and opto-electrical stability and hysteresis data

Figure 6-6 through Figure 6-9 show laser output stability data. Spectra were taken at intervals during one day (40 interferogram averages per spectra). Figure 6-6 shows FTS spectra where one mode dominates at a substrate temperature of 18°C and injection current of 105mA, and Figure 6-7 shows where dual mode behaviour is strongest at 16°C and 106mA. Evidently, there is a problem with the device's thermal management, as both Figures are consistent with a settling time of up to 2 hours, after which the power and wavenumber output become relatively stable. Since the temperature controller indicated a stable substrate temperature at each setting,

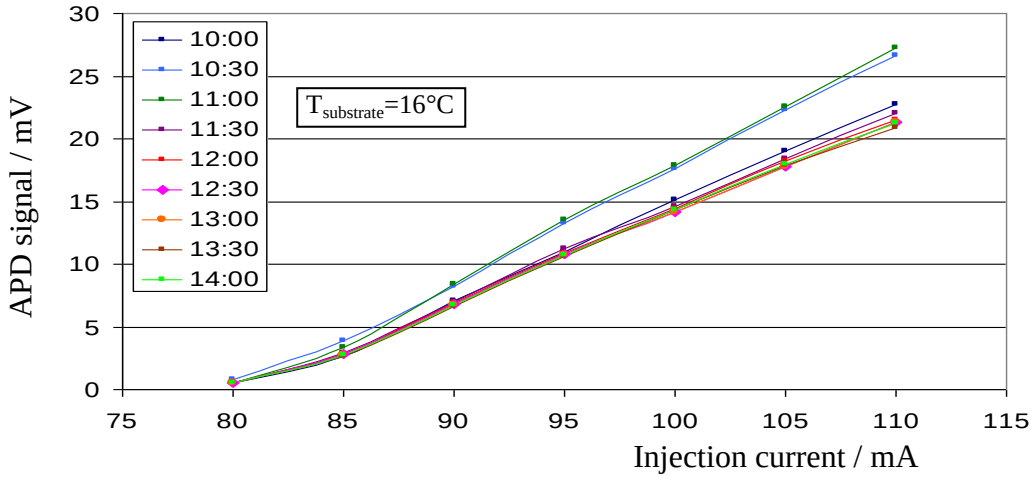


Figure 6-8: **Opto-electric stability data at a fixed substrate temperature, taken at various times throughout one day. Instability is shown particularly over the first 1½ hours. Data points are joined for ease of visual interpretation.**

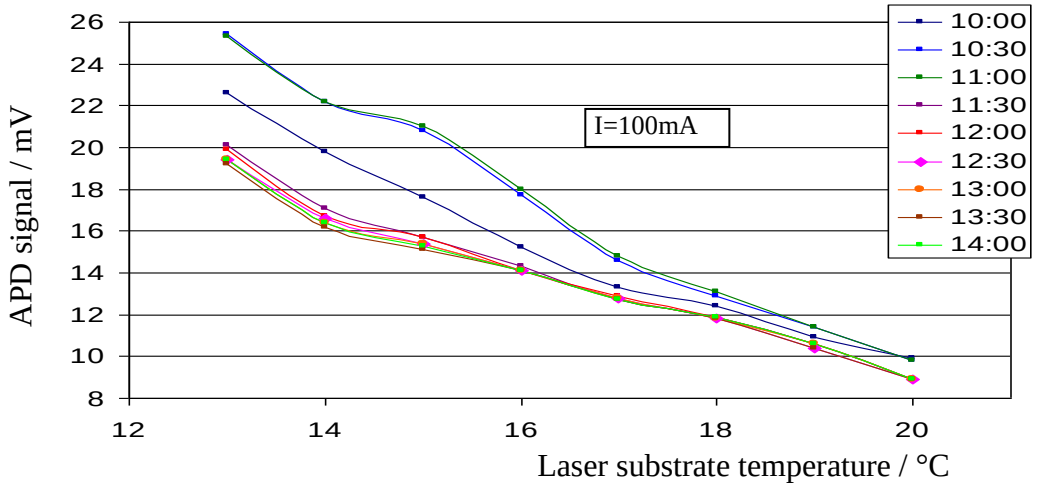


Figure 6-9: **Opto-electric stability data at a fixed injection current, taken at various times throughout one day. Instability and nonlinearity is shown particularly over the first 1½ hours. Data points are joined for ease of visual interpretation.**

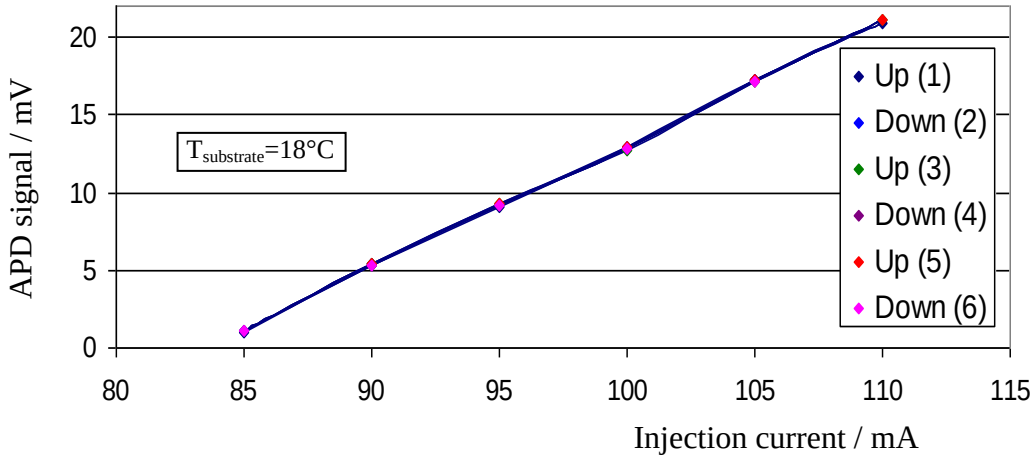


Figure 6-10: **Opto-electrical hysteresis data shown via consecutive readings from repeatedly increasing and decreasing the laser injection current at fixed substrate temperature. The data points overlap on this scale showing minimal hysteresis. Data points are joined to aid visual interpretation.**

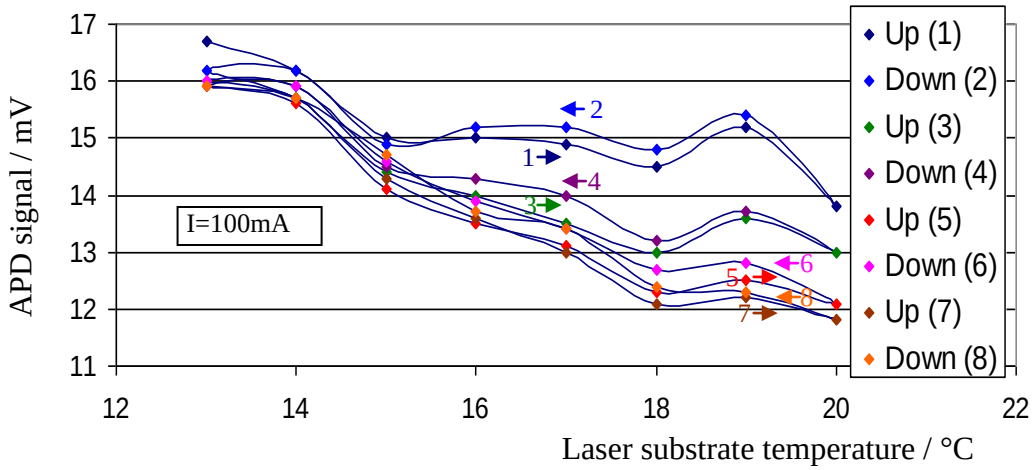


Figure 6-11: **Opto-electrical hysteresis data shown via consecutive readings from repeatedly increasing and decreasing the laser substrate temperature at fixed injection current. Hysteresis is present. Data points are joined to aid visual interpretation, and the numbers and arrows give a sense of the order in which data points were taken.**

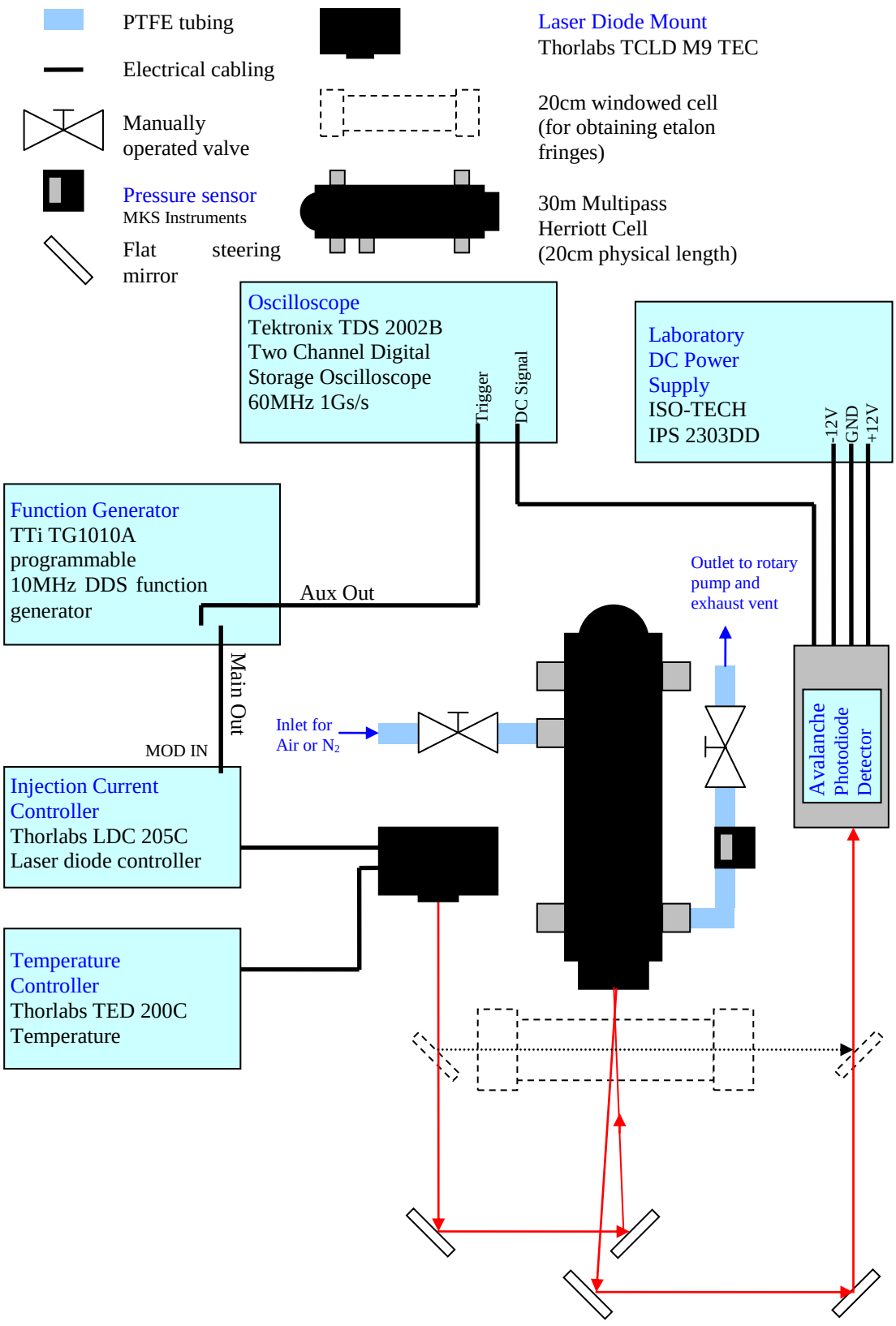


Figure 6-12: Optical, electrical and plumbing setup for oxygen spectroscopy.

Figure 6-9 suggests inefficiency in heat transport between the laser's substrate and active layer.

Figure 6-8 and Figure 6-9 show opto-electronic data taken after the laser had been set up for spectroscopy (Section 6.4). Both figures indicate a settling time of 1½ hours before the output characteristics, while not stable, are relatively closely bunched. It is noteworthy that at a fixed substrate temperature and changing injection current, the laser output behaves linearly (Figure 6-8), but that a fixed injection current and changing substrate temperature gives a less well behaved characteristic (Figure 6-9).

Inefficiency in transport *between* substrate and active layer is also supported by the hysteresis data of Figure 6-10 and Figure 6-11. In both figures readings were taken continuously, starting at the low abscissa value to the high, and sequentially back down again before repeating, and so on. At fixed substrate temperature and changing injection current (in the active layer) (Figure 6-10), the data points overlap on this scale - showing characteristically linear tuning. However, for fixed injection current and changing substrate temperatures, Figure 6-11 clearly shows hysteresis and non-linear behaviour.

6.4 Experimental Setup for diode laser spectroscopy

The optical, electrical and plumbing setup for oxygen spectroscopy is shown in Figure 6-12. The continuous wave semiconductor diode laser was mounted in a Thorlabs TCLD M9 TEC laser diode mount. The injection current was controlled with a Thorlabs LDC 205C laser diode controller, and the substrate temperature controlled via a Thorlabs TED 200C temperature controller. The tuning current ramp for the sweep integration technique came from a TTi TG1010A function generator. The laser's manufacturer also supplied an avalanche photodiode detector (APD) whose signal was recorded on a Tektronix TDS 2002B digital oscilloscope with a maximum of 128 averages.

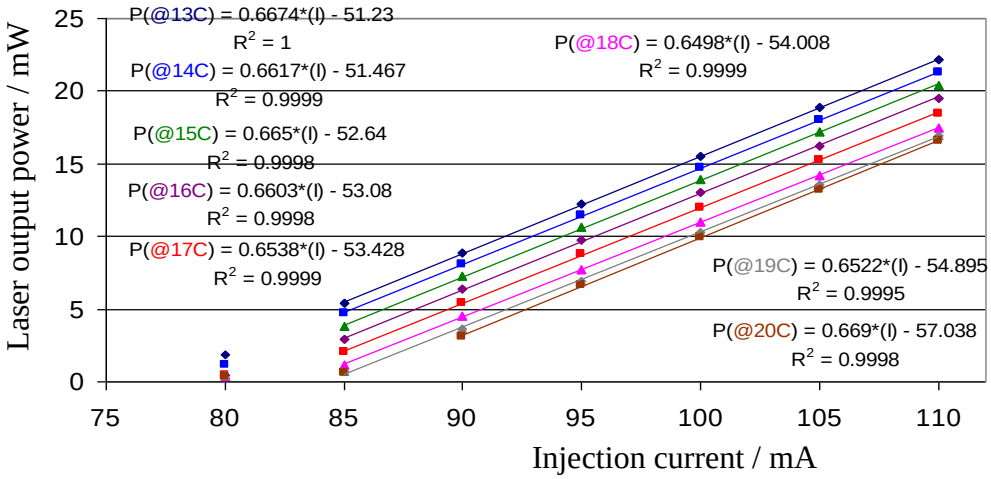


Figure 6-13: Laser power output vs injection current at fixed substrate temperatures, with linear fits. Over time the gradients would change due to output instability, but the linearity remained.

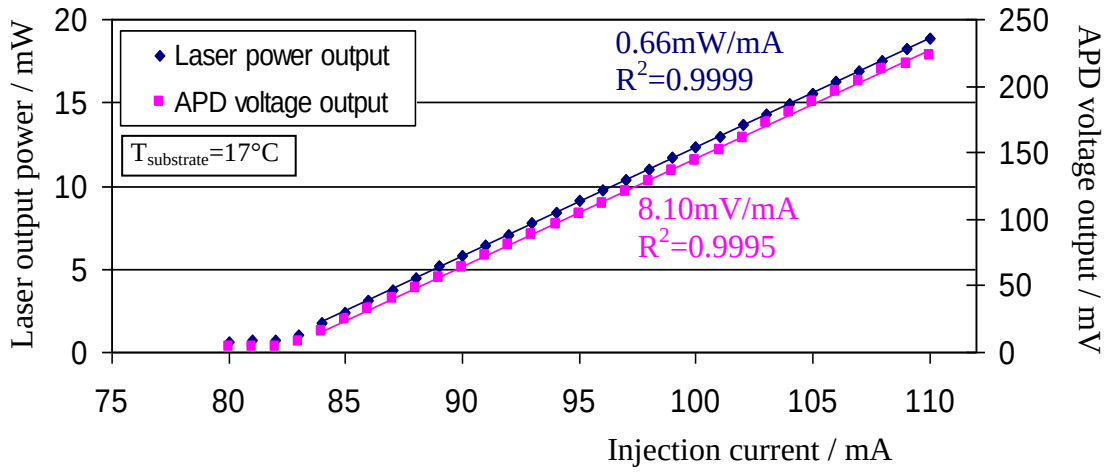


Figure 6-14: Example laser power output and APD voltage output vs laser injection current at a fixed laser substrate temperature. Linear fits are shown and imply a linear relation between laser output power and recorded APD voltage output of 12.3mV/mW. Over time the gradients would change due to laser output instability but the linearity remained.

An astigmatic Herriott cell was used to increase the oxygen signal obtained from laboratory air. This cell was of the same type as used in Cascade Technologies' CT3000 series gas analysers (Chapter 7), with a physical spacing between the mirrors of 20cm and an optical path length of 30m. The out of cell path length was 1.455m. The pressure in the cell was controlled by restricting the flow through the cell to a rotary pump and exhaust vent.

6.5 Laser and Avalanche Photodiode Detector characterisation

The diode laser's output power versus injection current was measured for integer substrate temperatures and was linear (Figure 6-13) (measured with a Thorlabs PM100D power meter with S302C thermal power head). The voltage from the avalanche photodiode detector (APD) was also recorded as a function of injection current for integer temperatures. A typical example is shown in Figure 6-14, for a substrate temperature of 17°C, showing a power gradient of 0.66mW/mA and a voltage gradient of 8.10mV/mA. These trends imply a conveniently linear relationship between laser power and APD output, here at 17°C of 12.3mV/mW.

The instability of the laser meant that the gradient values would change over time, but the linear relationships still remained. The oxygen spectra shown later have linear baselines, showing laser stability during the averaging of 128 50Hz triangular sweeps (≈2.6s), but slightly different gradients between different spectral datasets.

6.6 Operating Parameters

Sweep Integration oxygen spectra were recorded from the air in the laboratory at room temperature and at two pressures: atmospheric (760Torr) and 200Torr. The latter pressure is a typical operating pressure of a Cascade Technologies sensor. 200Torr allows reasonably straightforward interfacing from the pressures

encountered in measurement environments, whilst delineating the absorption features through their narrowing with reduced pressure (while maintaining mixing ratios).

Given the instability upon change of substrate temperature in Figure 6-9 and Figure 6-11, it was sensible to fix the laser substrate temperature to give emission near a spectral feature of interest, then tune the laser over the feature by sweeping the injection current. By scanning sequentially, three regions containing the strongest absorption features were identified for detailed analysis, with a base injection current of 100mA, these were found when setting the substrate temperature to approximately 15.8°C, 17.3°C and 19.2°C.

A (triangular) waveform was chosen, centred at 100mA injection current so that the sweep range could be expanded about a point almost mid-way between threshold and the current limit. A 5mA amplitude (10mA peak to peak) provided sufficient wavenumber coverage to record the feature found when setting the substrate temperature to 15.8°C. The amplitude needed extending to 10mA to include sufficient baseline for the other two strong absorption regions appearing at substrate temperatures of 17.3ºC and 19.2ºC. The triangular waveform allowed the asymmetries of the wavenumber downchirp (while heating) and upchirp (while cooling) spectra to be observed.

The sweep could be performed at a slow frequency since the laser was continuous wave and the detector was DC. Resolution enhancement of slow scanning (Section 2.10) has a trade off with the time needed for acquisition and averaging. The spectra here were taken at 50Hz (specification sheet maximum was 1kHz).

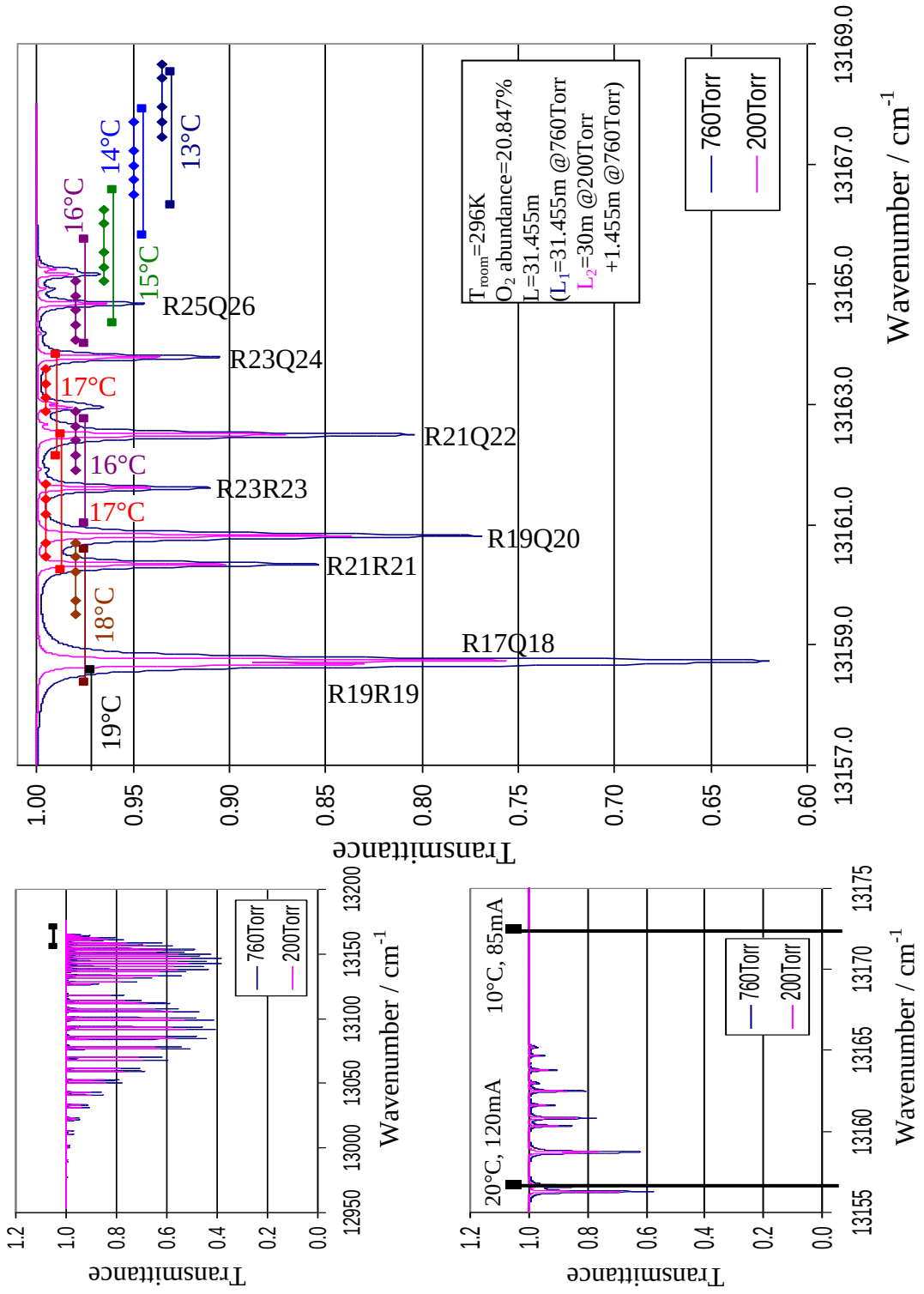


Figure 6-15: Laser wavenumber coverage against the oxygen A-band at 760Torr and 200Torr for 296K (V E Zuev IAO). The diamond shaped points plot the wavenumber tuning data from Figure 6-2 (arbitrary amplitudes). The square data points plot the extrapolated tuning between 90mA and 110mA at the relevant substrate temperatures. The average static tuning of 1.3cm^{-1} is twice the tuning during sweeping of 0.65cm^{-1}. Laser instability makes the ranges approximate. The insets show the laser's full extrapolated wavenumber coverage, 13156.72cm^{-1} to 13172.51cm^{-1}, determined by threshold and the specification sheet limits.

6.7 Spectral Coverage

The large display in Figure 6-15 displays as diamond shaped data points the FTS tuning data of Figure 6-2 and Figure 6-3, set against transmission spectra from the HITRAN database (V E Zuev IAO) at atmospheric pressure and at 200Torr. In each case the mixing ratio is the same, but the number density in the gas cell reduces in the 200Torr case. The laser wavenumber coverage is seen to overlap the band head of the oxygen A-band – and so some spectral coverage is beyond the absorbance lines and thus redundant. Also shown between square data points are the approximate sweep integration coverages for the 20mA sweeps centred on 100mA (90mA to 110mA). These ranges were extrapolated from the trends in Figure 6-2 and Figure 6-3. The instability of the laser makes these trends only rough guides to actual tuning, but gave a first expectation of the likely multi-mode spectra. As is usual, the average static (implied) tuning range of 1.3cm^{-1} is larger than the dynamic tuning range of 0.65cm^{-1} (Section 6.8.3.3). The two insets show the laser's full extrapolated wavenumber coverage, 13156.72cm^{-1} to 13172.51cm^{-1}. This range is determined by the static operating limits imposed by laser threshold (85mA at 10ºC) and the specification sheet's maximum current limit (120mA), and the substrate's allowed temperature range (10º to 20ºC).

Line	Line Centre Wavenumber $\tilde{\nu}_{0\ vac\ /\ cm^{-1}}$	Centre Wavenumber $\tilde{\nu}_{0\ 200\ /\ cm^{-1}}$ | $\tilde{\nu}_{0\ 760\ /\ cm^{-1}}$		FWHM $\tilde{\omega}_{v\ /\ cm^{-1}}$		Absorbance Peak $S\frac{N}{V}Lg_{max}$	
	0Torr	200Torr*	760Torr	200Torr*	760Torr	200Torr*	760Torr
R21Q22	13162.500954	13162.498499	13162.492054	0.043	0.090	0.110	0.179
R19Q20	13160.818147	13160.815807	13160.809647	0.044	0.093	0.173	0.294
R21R21	13160.343498	13160.341213	13160.335198	0.043	0.090	0.102	0.165
R17Q18	13158.743620	13158.741480	13158.735820	0.045	0.096	0.257	0.465
R19R19	13158.679074	13158.677060	13158.671774	0.044	0.093	0.160	0.268
R17Q18 +R19R19				0.100	0.142	0.267	0.457
R17R17	13156.623203	13158.621421	13156.615703	0.044	0.096	0.240	0.348

Table 6-1: **Voigt profile parameters for the likely strong lines in the three regions of the largest absorptions. At 760Torr (and 0Torr) the parameters come from the Parameter Calculator (Section 2.5.1) based on HITRAN line data (Rothman et al., 2005) and the equations of Chapter 2, for a gas temperature of 296K, optical path length of 3145.5cm and oxygen abundance of 20.847%.**

***At 200Torr the Voigt profile for 3000cm path length had a 760Torr Voigt profile added to it for an out of cell path length of 145.5cm. The parameters were then found graphically from the overall profile.**

6.8 Proof of Principle Oxygen Spectroscopy

6.8.1 Preliminaries

Initial experiments showed that absorbances were present within the wavenumber ranges expected. This was enough to show that the laser design was on the right track. To investigate the strengths and weaknesses of the present design more thoroughly, measured absorbance spectra were compared to expected Voigt profiles generated in the program of Section 2.6.3.

The absorbance spectra were calculated from the measured detector signal in the data analysis program of Chapter 3. As in previous chapters, this was done via a polynomial (linear) fit to the baseline. Division by an actual measured background would have increased noise due to two reasons: firstly, from non-random noise on the baseline from unintentional etalon fringing in the optical path; secondly, from laser instability causing a small change in gradient of baseline and spectra so that they did not match over a full scan.

The theoretical parameters for the absorbance lines were calculated from the calculator based on the equations of Chapter 2, using HITRAN parameters for a standard temperature of 296K. At 760Torr, the outside cell pathlength of 1.455m was added to the 30m inner cell path length. For the Voigt profile at 200Torr and 30m, a Voigt profile of 1.455m at 760Torr was added, and the parameters found graphically from the resulting profile. The expected concentration was taken as 20.946% (Encyclopædia Britannica, 2015), weighted by the $^{16}O^{16}O$ isotopic abundance of 0.995262 (V E Zuev IAO), i.e. 20.847%. The resulting Voigt parameters are recorded in Table 6-1.

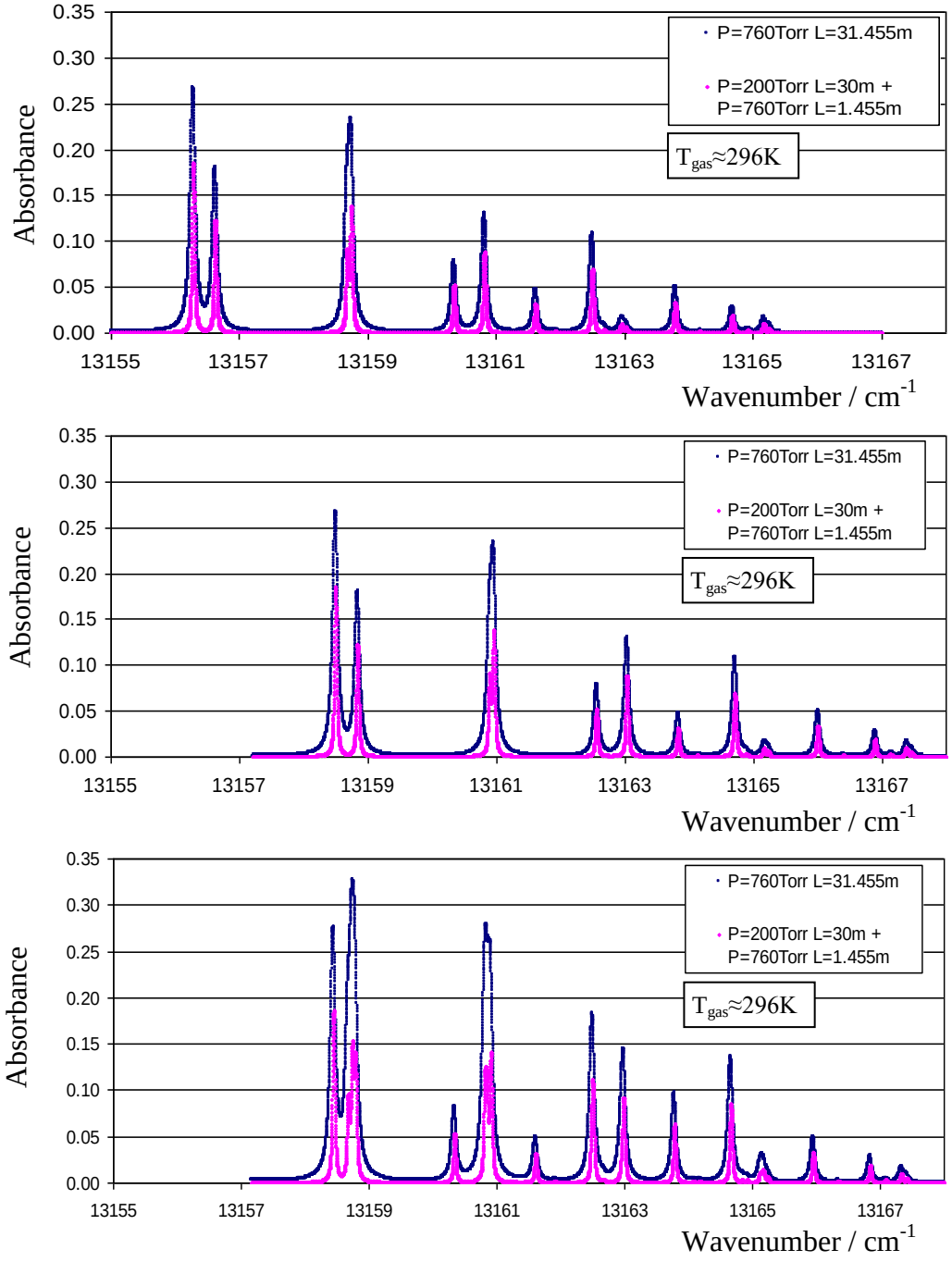


Figure 6-16: **Top plot: Absorbance spectra near the band head of the oxygen A-band (weighted by 0.5), for Herriott cell pressures of 760Torr and 200Torr. Data for these plots was generated from HITRAN on the Web (V E Zuev IAO). The mixing ratio is 20.847%.**

Middle plot: The same spectra as the top plot but displaced by a laser cavity mode spacing of 2.22cm^{-1}.

Bottom plot: The sum of the spectra in the top and middle plots, simulating the multimode spectra of two equal intensity delta function cavity modes spaced by 2.22cm^{-1}.

6.8.2 Modeling

At least two of the three large absorbance regions display multi-mode behaviour. One way of analyzing such spectra, with a single multi-mode laser and single detector, was developed by Ewart's group at Oxford, and known as Multi-Mode Absorption Spectroscopy (MUMAS) (Arita and Ewart, 2008, Arita et al., 2008, Hamilton et al., 2010). In their method, the laser cavity mode output is modeled by a comb of Airy functions. These are then convolved with a larger function – representing the change in intensity of the cavity modes as they tune under the laser gain profile, or pass through an external filter. This overall laser output is then convolved with a simulated gas transmission spectrum to simulate the normalized detected signal. The observed and simulated spectra can then be matched by adjusting the gas and laser parameters. The final gas parameters then indicate the gas concentration. The examples given for the use of this technique used well behaved and characterized multi-mode lasers.

The mode structure of the laser studied here has been shown to be subject to drift and broad side mode behaviour, even under conditions of constant substrate temperature and injection current (Section 6.3). Another complicating factor in a detailed analysis is the uncertain FSR of the etalon fringing. In two of the large absorption regions, the known spacing between the peaks of a doublet formed from R19R19 and R17Q18 can be used to set the wavenumber scale. However, in the absence of a Voigt profile fitter, a visual fit introduces error. The exact concentration of O_2 in the laboratory air is also not known.

Nevertheless, a simplified model is very useful in line identification and in drawing general conclusions about the laser mode structure and the reasonableness of the fits to an assumed O_2 concentration of 20.847%. The laser output is therefore modeled as two delta function modes. This is equivalent to one absorbance spectrum being superimposed upon a version of itself transposed along the wavenumber axis by the laser cavity mode spacing. The bottom plot of Figure 6-16 illustrates such a superposition spectrum near the band head of the oxygen A-band for a cavity mode

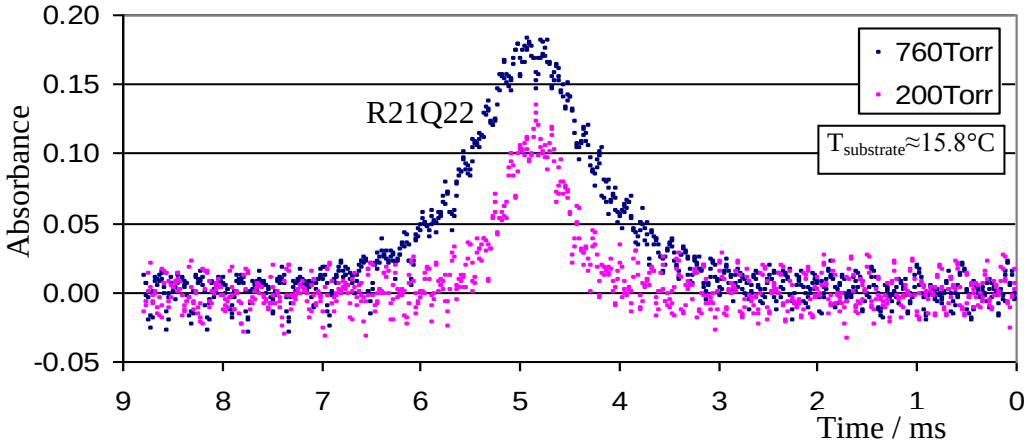


Figure 6-17: Current upsweep (wavenumber downchirp) room temperature oxygen absorbance spectra at 760Torr and 200Torr. The time axis has been reversed to reflect lower wavenumber values occurring on the left of the altered abscissa. Operating parameters: 50Hz triangular current waveform of 10mA amplitude centred on 100mA at a substrate temperature of ≈15.8°C. 128 oscilloscope averages.

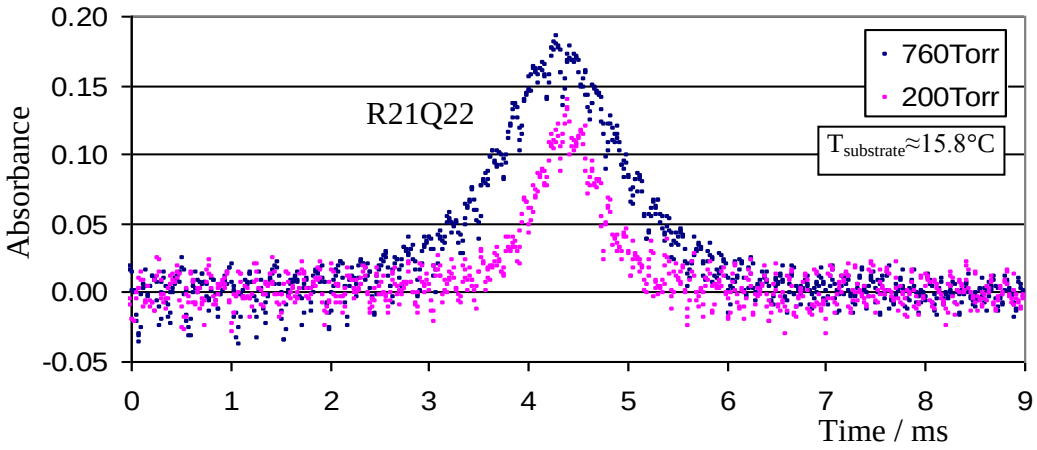


Figure 6-18: Current downsweep (wavenumber upchirp) room temperature oxygen absorbance spectra at 760Torr and 200Torr. Operating parameters are as in Figure 6-17.

spacing of 2.22cm^{-1} - with each of the modes assumed to carry half the laser intensity. Using the R19R19 / R17Q18 doublet at 200Torr to set the wavenumber axis in two of the wavenumber regions, the wavenumber of the remaining line from the other laser mode, minus its known wavenumber, gives the cavity mode spacing. The upchirp and downchirp spectra in these regions at 200Torr give four spectra which all agree on a cavity mode spacing of 2.22cm^{-1} (Figure 6-23 and Figure 6-24, and Figure 6-26 and Figure 6-27). This is similar to the spacing of the absorption lines near the O_2 A-band band head (Figure 6-16). Each spectral region is therefore potentially multi-mode – but whether this is realized in practice depends on the strength of the two laser modes.

In the modeling below, the individual Voigt profiles were prepared in the Voigt modeler of Section 2.6.3, then summed and adjusted in Microsoft Excel.

6.8.3 Results

6.8.3.1 Laser substrate temperature ≈ 15.8°C

The first of the three large spectral features appeared for a substrate temperature of ≈15.8°C. The absorbance spectra for the current upsweep (wavenumber downchirp) are shown in Figure 6-17. The time axis is reversed to place lower wavenumbers on the left of the altered abscissa. This figure can therefore be compared to the current downsweep (wavenumber upchirp) in Figure 6-18, which can be seen to be very similar.

Figure 6-19 shows that the peak heights of the measured spectra are a reasonable fit to a Voigt model of R21Q22. (No Voigt fitter was available, so the experimental and calculated models are simply superimposed in the figure). No etalon fringe data of definite FSR could be obtained (discussed below). The time to wavenumber relationship therefore had to be found approximately. The FSR of the etalon could be approximated by choosing its value so that the experimental FWHM's of, say, the 760Torr R21Q22 absorbance matched the modeled value. This is approximate since

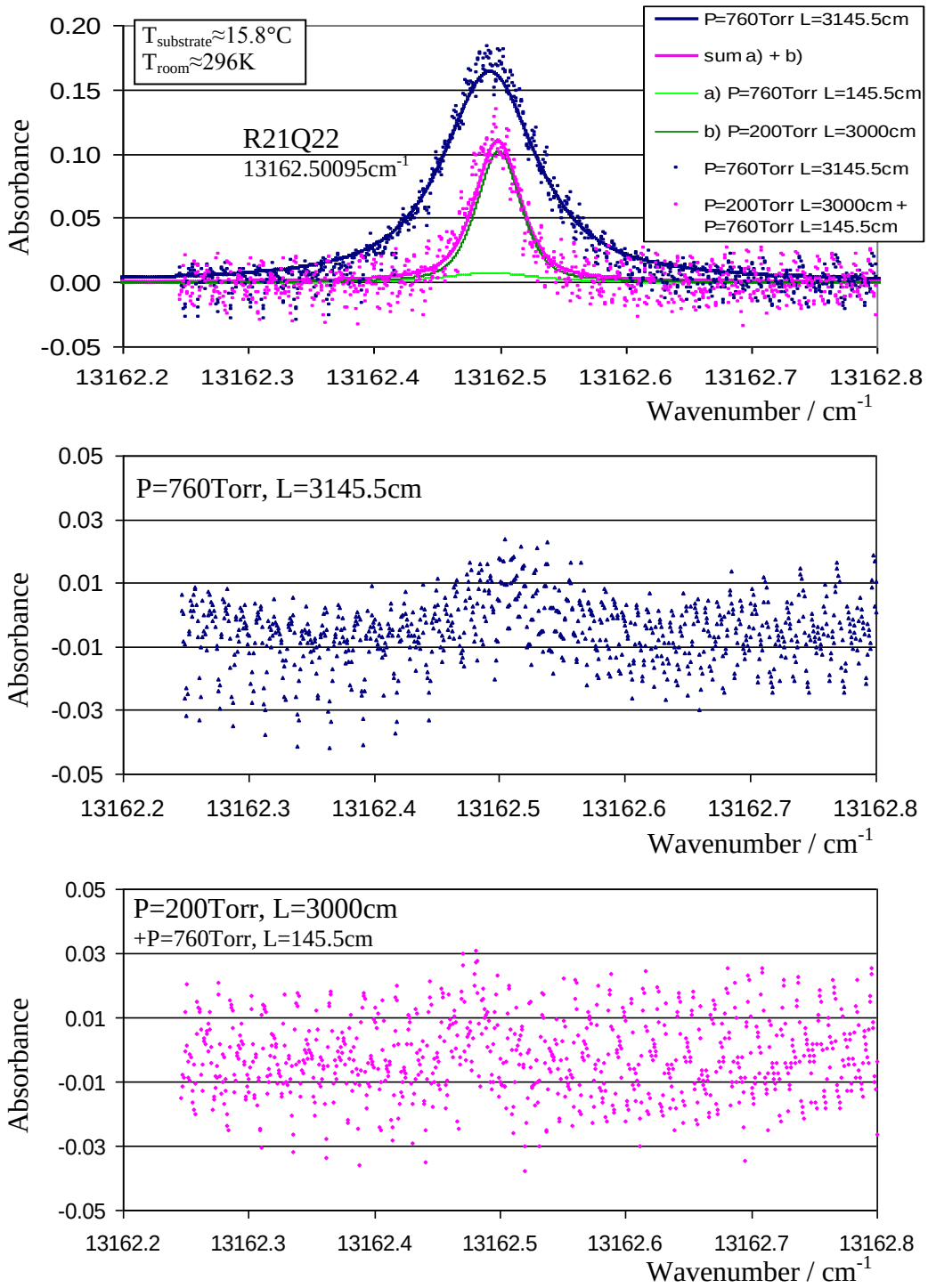


Figure 6-19: **Top plot: Absorbance spectra from Figure 6-17 plotted against single mode Voigt models for R21Q22 (Mode 1) for a gas temperature of 296K and an oxygen abundance of 20.847%. The out of cell path length of 145.5cm has been accounted for in both Voigt models. The abscissa of the measured data has been linearly scaled and shifted to match the FWHM of the Voigt model at 760Torr.**
Middle Plot: Residuals of the 760Torr, 3145.5cm visual line fit.
Bottom Plot: Residuals of the 200Torr, 3000cm (plus 760Torr, 145.5cm) visual line fit.

the model assumes a mixing ratio of 20.847%, which may not be true in the particular sample of laboratory air. Since the etalon fringes, equally spaced in frequency, were also nearly equally spaced in time (see below), the time to wavenumber relationship was linear. This linearity allows a simplification of the FSR procedure, which becomes equivalent to a linear scaling of the time axis – without the need to explicitly work with the etalon data. Such a linear scaling was applied to the 760Torr experimental data of Figure 6-17 to produce the 760Torr data in the top plot of Figure 6-19. A linear shift added to the linear scaling converts the relative wavenumber axis to an absolute one, by matching a wavenumber to a known absorbance peak position.

The linear shift and scaling appropriate to the 760Torr data was also applied to the 200Torr spectrum. The ability of this scaling to reproduce the calculated (and different) FWHM of $0.043cm^{-1}$ of the 200Torr R21Q22 experimental absorbance profile, implies that the laboratory mixing ratio was indeed close to the standard 20.847%. While not a perfect match, the scaled 200Torr experimental FWHM is not too dissimilar to the calculated $0.043cm^{-1}$. A possible source of error would be a change in the gradient of the linear time to wavenumber gradient, in the time between scans, due to laser fluctuations and drift. However, both measured traces are seen to be reasonably close to those expected without too much further tweaking of concentration, pressure and temperature values. R21Q22 is therefore a good candidate for more general concentration fitting and calibration - if spectroscopy is to be pursued with this laser.

The pressure induced downshift in peak position is also discernible between the 200Torr and 760Torr data (Figure 6-17, Figure 6-18, and the top plot of Figure 6-19). The peak position wavenumber difference between the Voigt models at 200Torr and 760Torr is $-0.00645cm^{-1}$.

Another possibility for the measured spectra of Figure 6-17 and Figure 6-18 was a dual-mode superposition of R21Q22 from Mode 1 and R21R21 from Mode 2. For

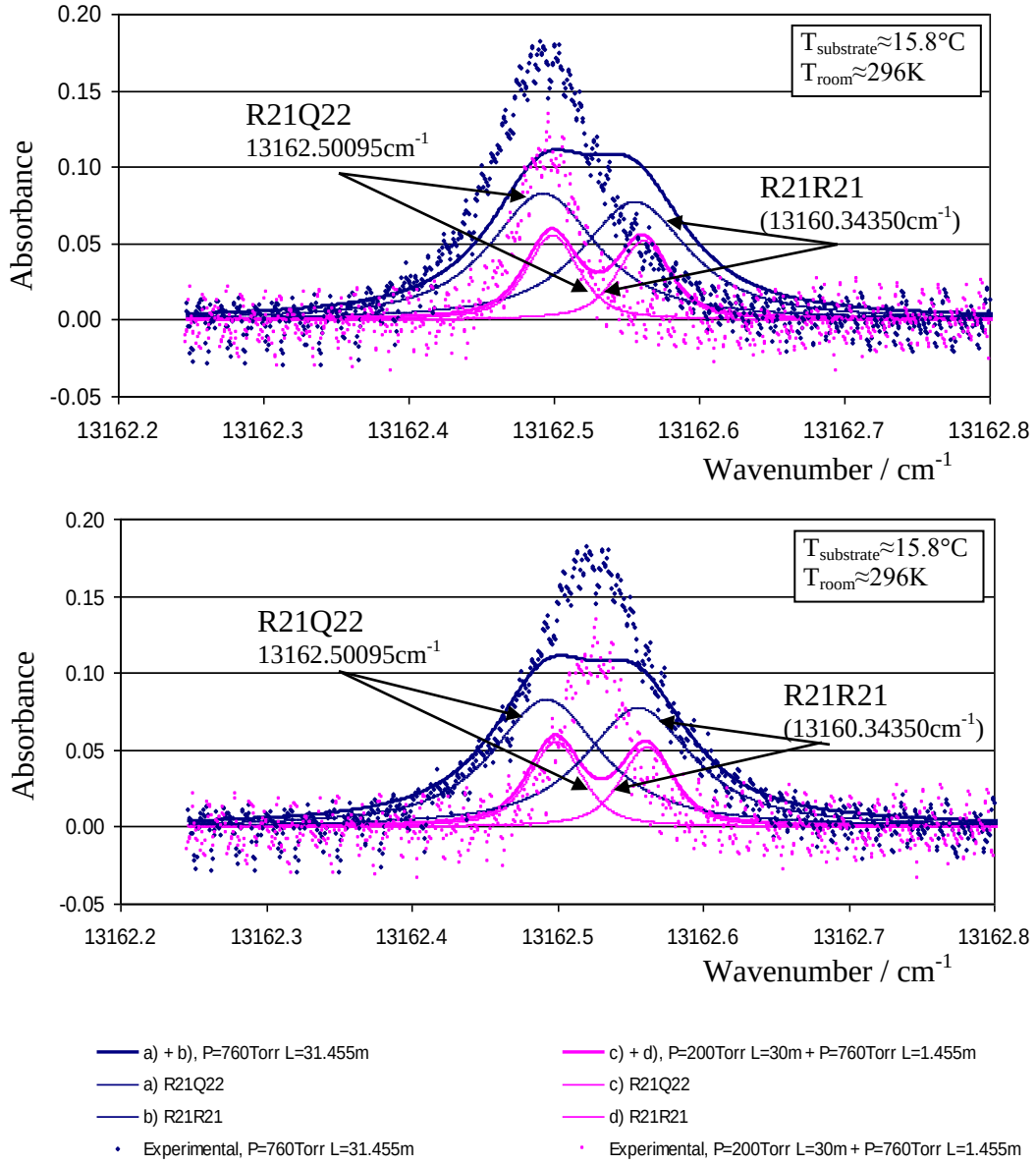


Figure 6-20: Absorbance spectra from Figure 6-17 plotted against a multi-mode sum of Voigt profiles for R21Q22 (Mode 1) and R21R21 (Mode 2), with a cavity mode spacing of 2.22cm^{-1}. The simulations assume that each mode carries half the full laser intensity. A lack of fit is demonstrated – indicating that the experimental absorbances are caused by a single strong laser mode (Mode 1), as in Figure 6-17. The simulations use a gas temperature of 296K and an oxygen abundance of 20.847%. The out of cell path length of 145.5cm has been accounted for.

The models are the same in both plots.

Top plot: The abscissa of the measured data has been linearly scaled and shifted as in Figure 6-17 for the single mode fitting.

Bottom Plot: The wavenumber axis of the experimental data is linearly scaled so that the 760Torr data fits the wings of the Voigt doublet.

this to be plausible the lines would need to be close together. However, the cavity mode spacing was fixed by the other two spectral regions (below) at 2.22cm^{-1}. The top and bottom plots of Figure 6-20, with this mode spacing, rule out the dual mode possibility. The top plot uses the wavenumber abscissa for the measured data that gave the single mode fit in Figure 6-19. The bottom plot stretches the abscissa so that the wings of the superposition match the measured data at 760Torr.

Both R21Q22 and R21R21 have similar lineshapes. However, the absorbance is attributed to Mode 1 – tuning over R21Q22. Mode 1 (higher wavenumbers) was dominant over Mode 2 at substrate temperatures less than 16°C in the FTS tuning data (Figure 6-1). Mode competition occurred at higher substrate temperatures. The spectra at substrate temperatures of 17.3°C and 19.2°C (below) are definitely dual mode. They also show upchirp and downchirp asymmetry in the cavity mode strengths. Pronounced asymmetry is not present in the 15.8°C case of Figure 6-17 and Figure 6-18.

6.8.3.2 System Limit of Detection

The Limit of Detection (LoD) of the spectrometer can be estimated from the baseline about the absorbance of the 200Torr Voigt profile (Figure 6-17). (The 760Torr profile still has significant wings over the wavenumber region shown, due to pressure broadening). 2σ of the absorbance noise floor gives 0.023 absorbance units - when averaging 128 50Hz triangular sweeps (≈2.6s). As a particular example, adjusting the mixing ratio in the parameter calculator, of Section 2.5.1 and Figure 2-3, to an R21Q22 absorbance peak height of 0.023, indicates that this spectrometer configuration could measure 2.9% oxygen in a general gas mixture at 760Torr, using this line.

The LoD for this (visible region) diode laser / APD system is comparable to the (mid-infrared) cw-QCL / photovoltaic detector system of 0.020 absorbance units for 7,680 averages (0.77s). The diode / APD system produced its spectra within the

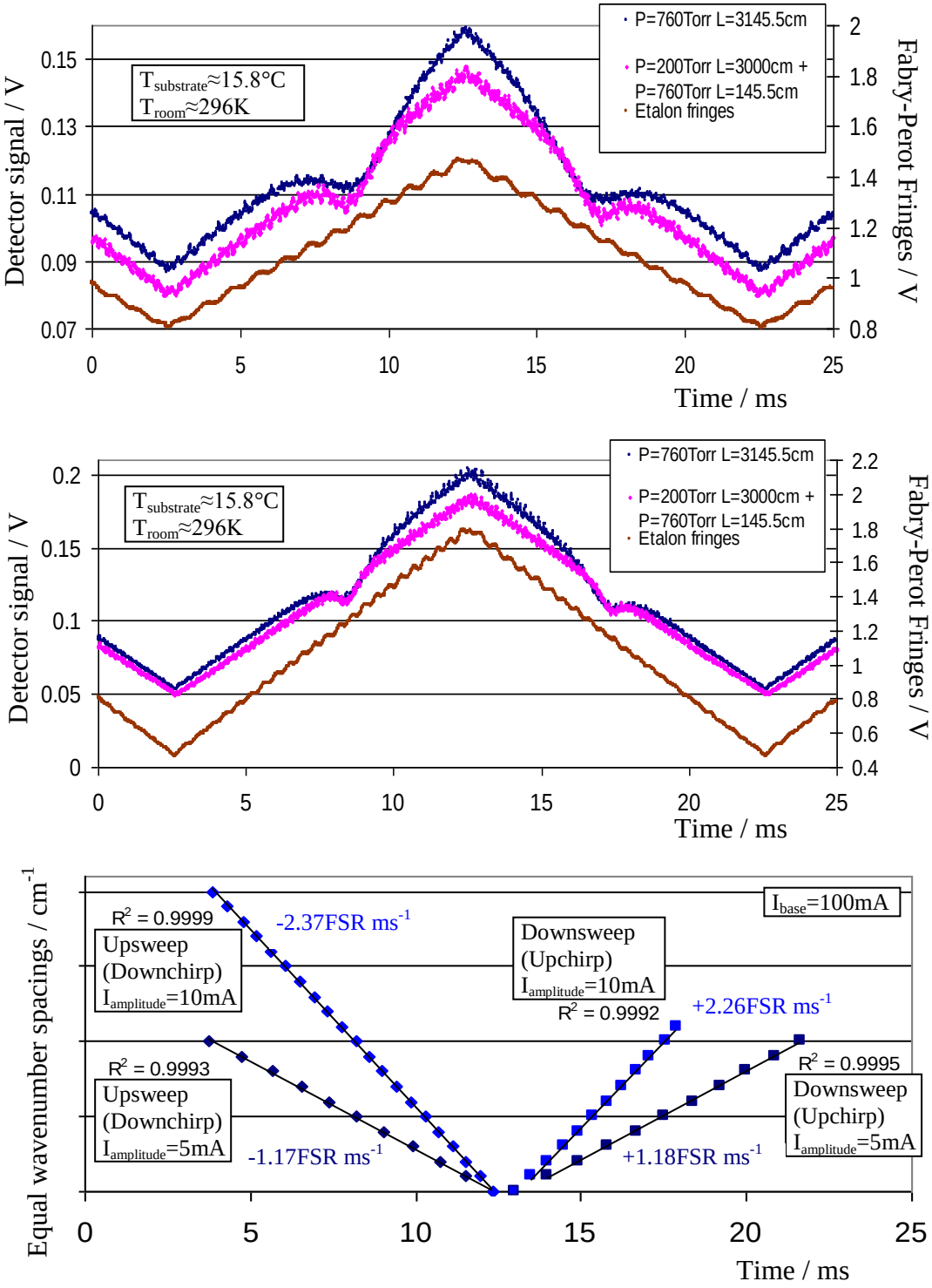


Figure 6-21: Three plots illustrating the linear time to wavenumber relationship via etalon fringing and the resulting constant tuning rates. The top and middle plots show R21Q22 (Mode 1) being swept over with a 50Hz triangular current ramp centred on 100mA, of 5mA and 10mA respectively. The laser substrate temperature was ≈15.8°C and 128 oscilloscope averages were taken. The exact etalon spacing is unknown, but is likely 1.7mm by using the fringes to reproduce the FWHM via time to wavenumber conversion.

averaging limitations set by the oscilloscope, with 60 times less averaging than the spectrometer of Chapter 4. The longer integration time of the present system (3.4 times) may be mainly a function of the slow 50Hz scan rate. An experiment testing noise against frequency sweep rate would be needed to confirm this. Also, the scan rate could easily be doubled by changing to a sawtooth waveform - removing the gradual downsweep. The two spectrometers do, of course, measure different spectral regions. This is an example of the enhanced performance of semiconductor diode lasers and their detectors, developed for the telecommunications industry – however, they do not cover the fundamental vibration-rotation transitions in the mid-infrared, where most molecules (but not oxygen) have their strongest transitions.

6.8.3.3 Etalon Fringes and possible tuning rates

As mentioned above, obtaining etalon fringes for an independent time to wavenumber calibration was not possible with the equipment available. This was due to uncertainty in the etalon length of the fringes obtained, and sometimes the quality of the fringe pattern.

Nevertheless, a sealed demonstration cell did produce observable fringes by tweaking its alignment around near normal incidence (Figure 6-21). These fringes allowed the linear tuning to be observed. The top and middle plots of Figure 6-21 show the fringes for the 5mA and 10mA amplitude spectra about an injection current of 100mA. During the first downsweep fringe, the tuning slows slightly as the current reverses, but then linear tuning resumes. This temporary slowdown is reflected in the downsweep absorption line appearing slightly later after the start of the downsweep than the upsweep.

The bottom of Figure 6-21 shows the tuning rates in terms of free spectral range per millisecond for those fringe peaks which were easily discernible. The tuning rates (apart from the slight slowdown in switching from current upsweep to downsweep

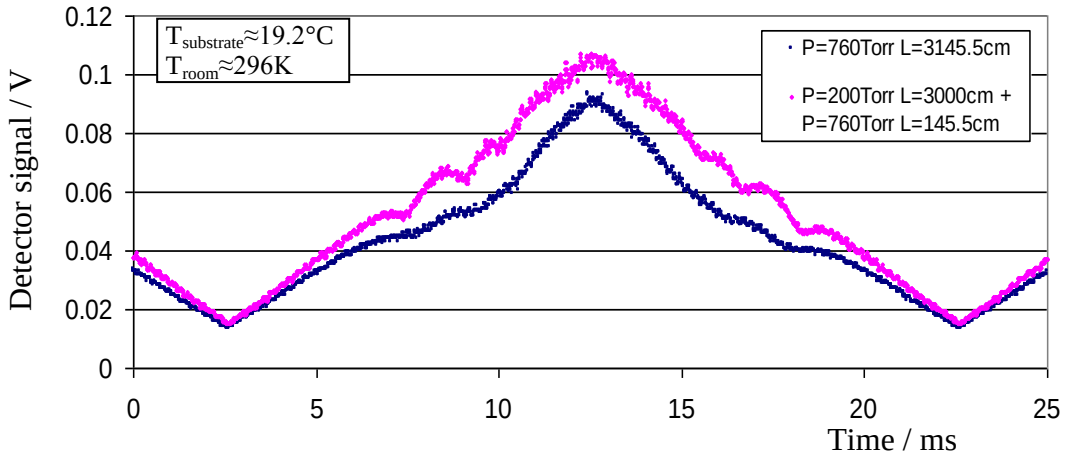


Figure 6-22: The detector signal showing oxygen absorptions for a laser substrate temperature of ≈19.2°C. Mode 1 is responsible for absorption dips due to R19R19 and R17Q18, and Mode 2 brings in R17R17 (see Figure 6-23 and Figure 6-24). Operating parameters: 50Hz triangular current waveform of 10mA amplitude centred on 100mA. 128 oscilloscope averages.

are constant. The double amplitude sweeps (10mA) occur at about twice the constant tuning rates of the 5mA current sweeps.

To produce the FWHM of 0.090cm^{-1} for R21Q22 at 760Torr, the fringe spacing would have to be ≈0.0295cm^{-1} (varying slightly each time the demonstration cell was moved in and out and aligned, and over time with laser fluctuations and drift). This corresponds to an etalon optical path length of 1.7mm (Equation C-4). This may be compatible with the impression of the time, that fringes were being formed by the partial overlap of the beam between the KBr window and the back of an internal aluminium entrance plate, before entering the demonstration cell fully. This FSR implies tuning rates of ≈±0.07cm^{-1}ms^{-1} (10mA amplitude) and ≈±0.035cm^{-1}ms^{-1} (5mA amplitude) at 50Hz. The scan ranges for these parameters are ≈0.32cm^{-1} (5mA amplitude) and 0.65cm^{-1} (10mA amplitude).

6.8.3.4 Laser substrate temperature ≈ 19.2°C

At 19.2°C, the spectra capture three features during both the current upsweep and downsweep (Figure 6-22). These are a doublet formed by R19R19 and R17Q18 from Mode 1 and R17R17 from Mode 2. Figure 6-23 and Figure 6-24 show the wavenumber downchirp (current upsweep) and wavenumber downchirp, respectively. The measured data was scaled by visually fitting the measured wavenumber difference between the doublet peaks to the simulated difference at 200Torr. The wavenumber of the remaining R17R17 peak, minus its known wavenumber, then gave the cavity mode spacing of 2.22cm^{-1}.

The 760Torr spectrum is further stretched to provide a better fit to the simulation. The need for this extra scaling can be accounted for by a changing gradient of baseline between 200Torr and 760Torr scans. Although a linear wavenumber to time relationship holds, the scaling factor has changed between scans – so that the 760Torr data needs a different scaling factor to that of the 200Torr data.

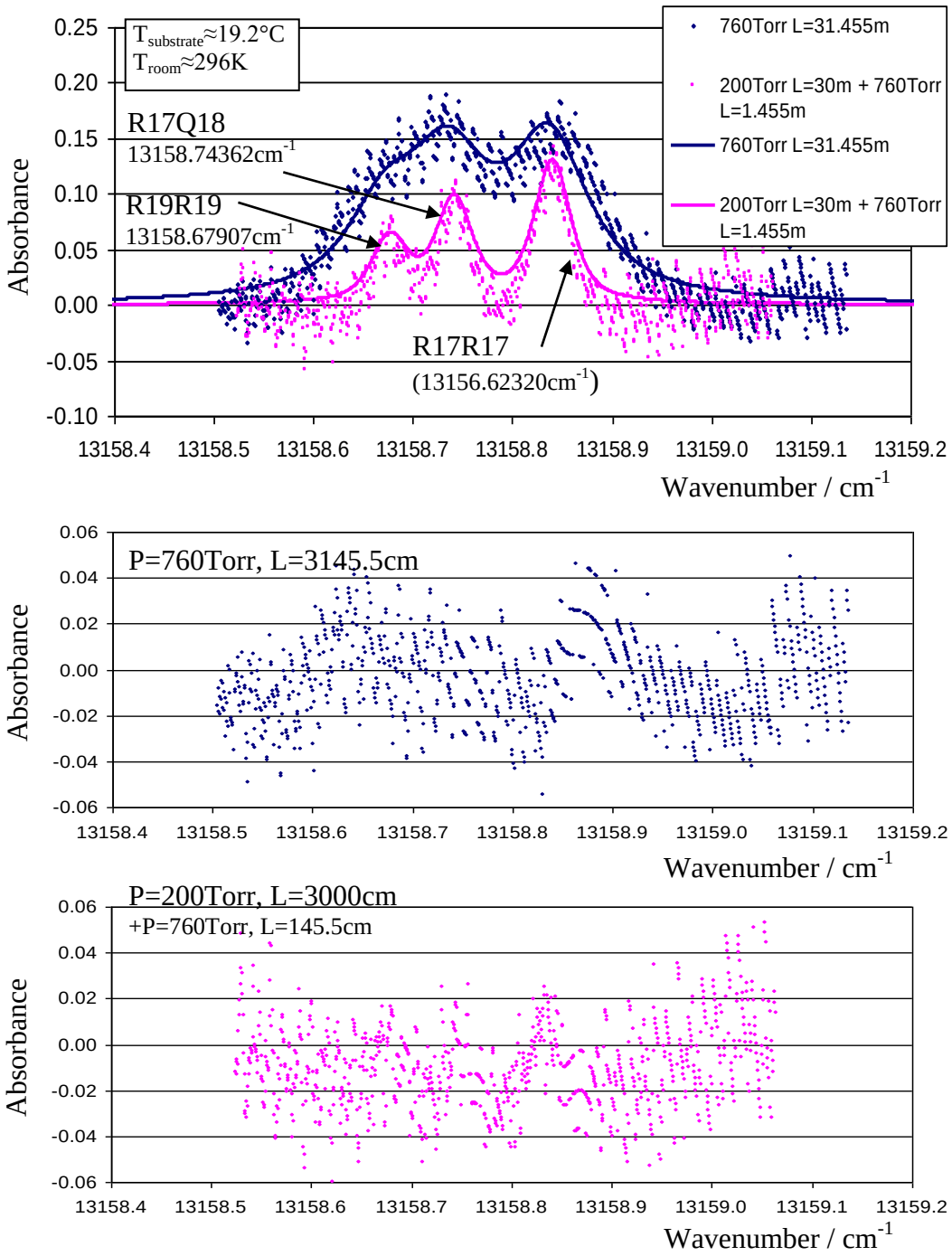


Figure 6-23: **Top plot: Absorbance spectra from the wavenumber downchirp part of Figure 6-22 plotted against a multi-mode sum of Voigt models for R19R19 and R17Q18 (Mode 1) and R17R17 (Mode 2). The simulations use a gas temperature of 296K and an oxygen abundance of 20.847%. The out of cell path length of 145.5cm has been accounted for. The absorbance weightings are given in Table 6-2. These weightings account for the different intensities of the laser cavity modes. The abscissa of the 200Torr measured data has been linearly scaled to match the known peak spacing between R19R19 and R17Q18. The 760Torr data has been stretched further to obtain a better fit.**
Middle Plot: Residuals of the 760Torr, 3145.5cm visual line fit.
Bottom Plot: Residuals of the 200Torr, 3000cm (plus 760Torr, 145.5cm) visual line fit.

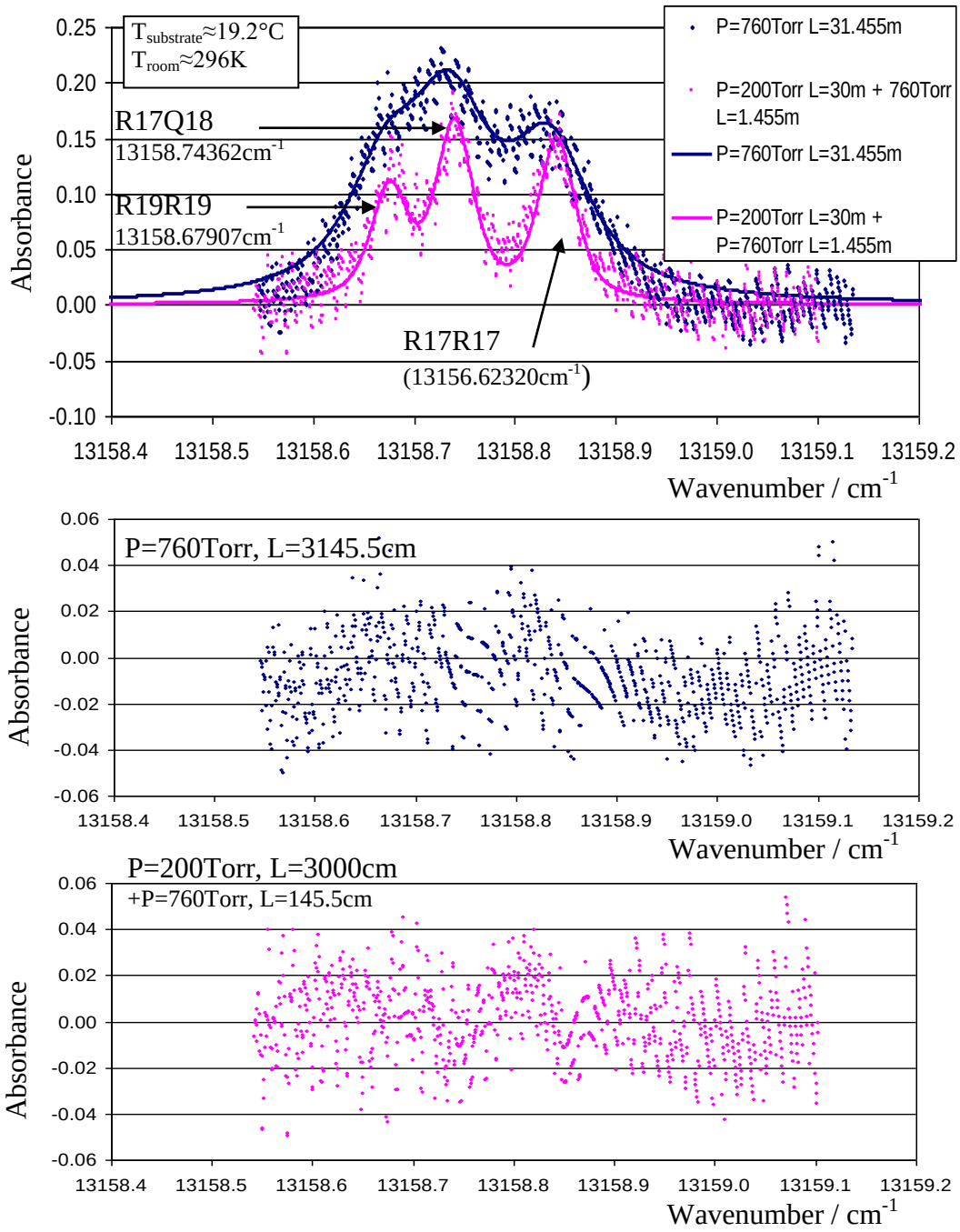


Figure 6-24: **Top plot: Absorbance spectra from the wavenumber upchirp part of Figure 6-22 plotted against a multi-mode sum of Voigt models for R19R19 and R17Q18 (Mode 1) and R17R17 (Mode 2). The simulations use a gas temperature of 296K and an oxygen abundance of 20.847%. The out of cell path length of 145.5cm has been accounted for. The absorbance weightings are given in Table 6-2. These account for the different intensities of the laser cavity modes. The abscissa of the measured 200Torr data has been linearly scaled to match the known peak spacing between R19R19 and R17Q18. The 760Torr data has been stretched further to obtain a better fit.**
Middle Plot: Residuals of the 760Torr, 3145.5cm visual line fit.
Bottom Plot: Residuals of the 200Torr, 3000cm (plus 760Torr, 145.5cm) visual line fit.

Substrate Temperature / ºC	Wavenumber downchirp or upchirp	Line	Absorbance weighting (P=200Torr)	Absorbance weighting (P=760Torr)
19.2	downchirp	R19R19	0.33	0.30
		R17Q18	0.33	0.30
		R17R17	0.49	0.40
	upchirp	R19R19	0.60	0.41
		R17Q18	0.60	0.41
		R17R17	0.60	0.37
17.3	downchirp	R19Q20	0.67	0.60
		R19R19	0.10	0.05
		R17Q18	0.10	0.05
	upchirp	R19Q20	0.32	0.25
		R19R19	047	0.35
		R17Q18	0.47	0.35

Table 6-2: Absorbance weightings used to simulate the laser cavity mode strengths resulting in Figure 6-23 and Figure 6-24, and Figure 6-26 and Figure 6-27.

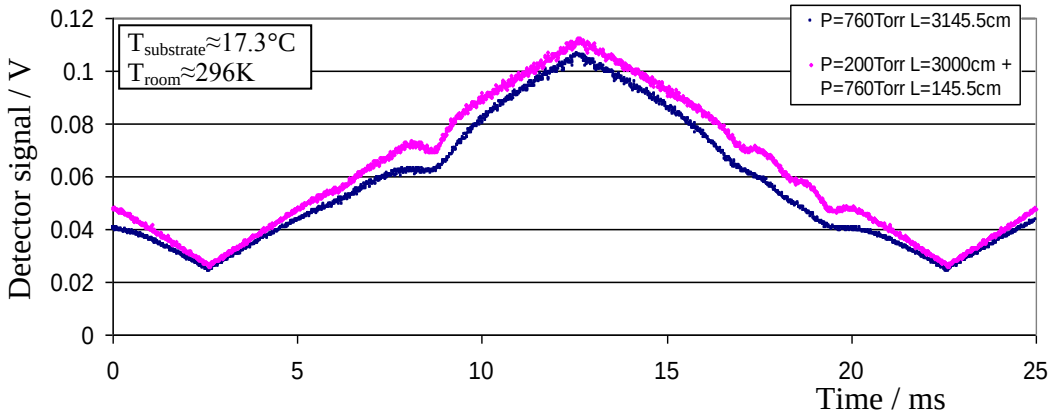


Figure 6-25: **The detector signal showing oxygen absorptions for a laser substrate temperature of ≈17.3°C. Mode 1 is responsible for an absorption dip due to R19Q20, and Mode 2 brings in R19R19 and R17Q18 (see Figure 6-26 and Figure 6-27). A large asymmetry in current upsweep and downsweep is present. Operating parameters: 50Hz triangular current waveform of 10mA amplitude centred on 100mA. 128 oscilloscope averages.**

The absorbance lines in the simulations, due to Mode 1 and Mode 2, were then weighted with scaling factors to produce a good visual fit to the measured data. Thus, this scaling takes into account the relative strengths of Mode 1 and Mode 2. The absorbance scaling factors are given in Table 6-2. This table and Figure 6-23 and Figure 6-24 demonstrate asymmetry in the relative mode strengths during heating (wavenumber downchirp) and cooling (wavenumber upchirp). These factors also take into account that a divisor from the detected baseline intensity would in general be too large. $I(\tilde{\nu})/I_0(\tilde{\nu})$ is required in the Beer-Lambert Law at a particular frequency. However, when more than one mode is lasing, part of the detected intensity comes from another frequency region.

The fits show that the spectra obtained, at this substrate temperature, have a simple structure and are thus amenable to multi-mode spectroscopy and, thus, oxygen measurement. However, a practical sensor built around this particular laser would be limited by the frequent need for recalibration due to instability and drift.

6.8.3.5 Laser substrate temperature ≈ 17.3°C

Figure 6-25 shows the detected signal for a current sweep with a laser substrate temperature of ≈17.3°C. Figure 6-26 and Figure 6-27 show the resulting wavenumber downchirp and upchirp absorbance spectra against multi-mode Voigt modeling, using the absorbance weightings in Table 6-2. The asymmetry in downchirp and upchirp is dramatic between these spectra. The R19R19 / R17Q18 doublet is again present, this time brought in by Mode 2. Mode 1 picks out the R19Q20 absorbance line (and the wavenumber axis is denoted by the Mode 1 values). During the downchirp the doublet is largely suppressed, while during the upchirp Mode 2 competes much more effectively for laser gain.

Again, the spectra obtained at this substrate temperature are amenable to oxygen spectroscopy. However, drift and instability are an issue with this particular laser.

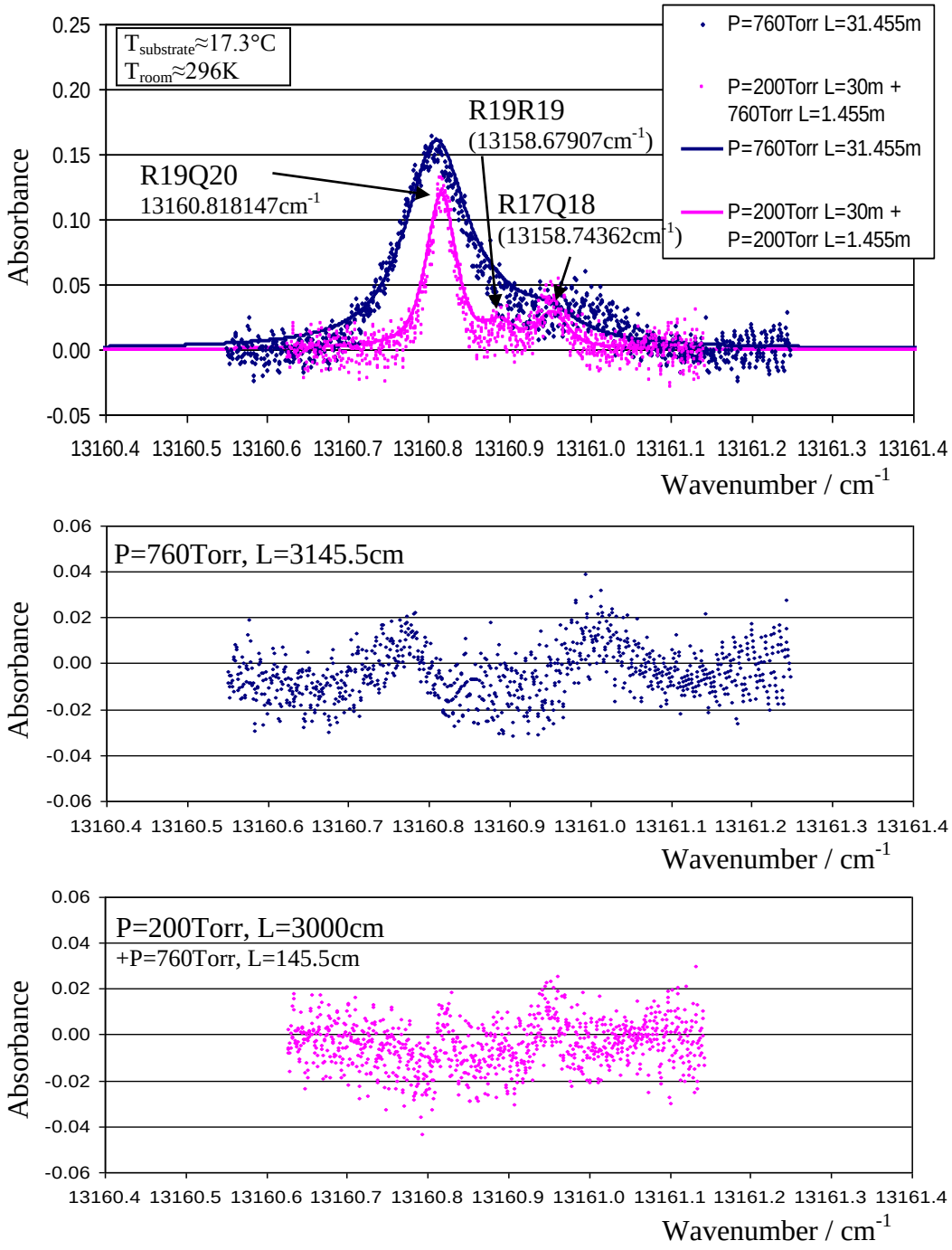


Figure 6-26: **Top plot: Absorbance spectra from the wavenumber downchirp part of Figure 6-25 plotted against a multi-mode sum of Voigt models for R19Q20 (Mode 1) and R19R19 and R17Q18 (Mode 2). The simulations use a gas temperature of 296K and an oxygen abundance of 20.847%. The out of cell path length of 145.5cm has been accounted for. The absorbance weightings are given in Table 6-2. These weightings account for the different intensities of the laser cavity modes. The abscissa of the 200Torr measured data has been linearly scaled to match the known peak spacing between R19R19 and R17Q18. The 760Torr data has been stretched further to obtain a better fit.**
Middle Plot: Residuals of the 760Torr, 3145.5cm visual line fit.
Bottom Plot: Residuals of the 200Torr, 3000cm (plus 760Torr, 145.5cm) visual line fit.

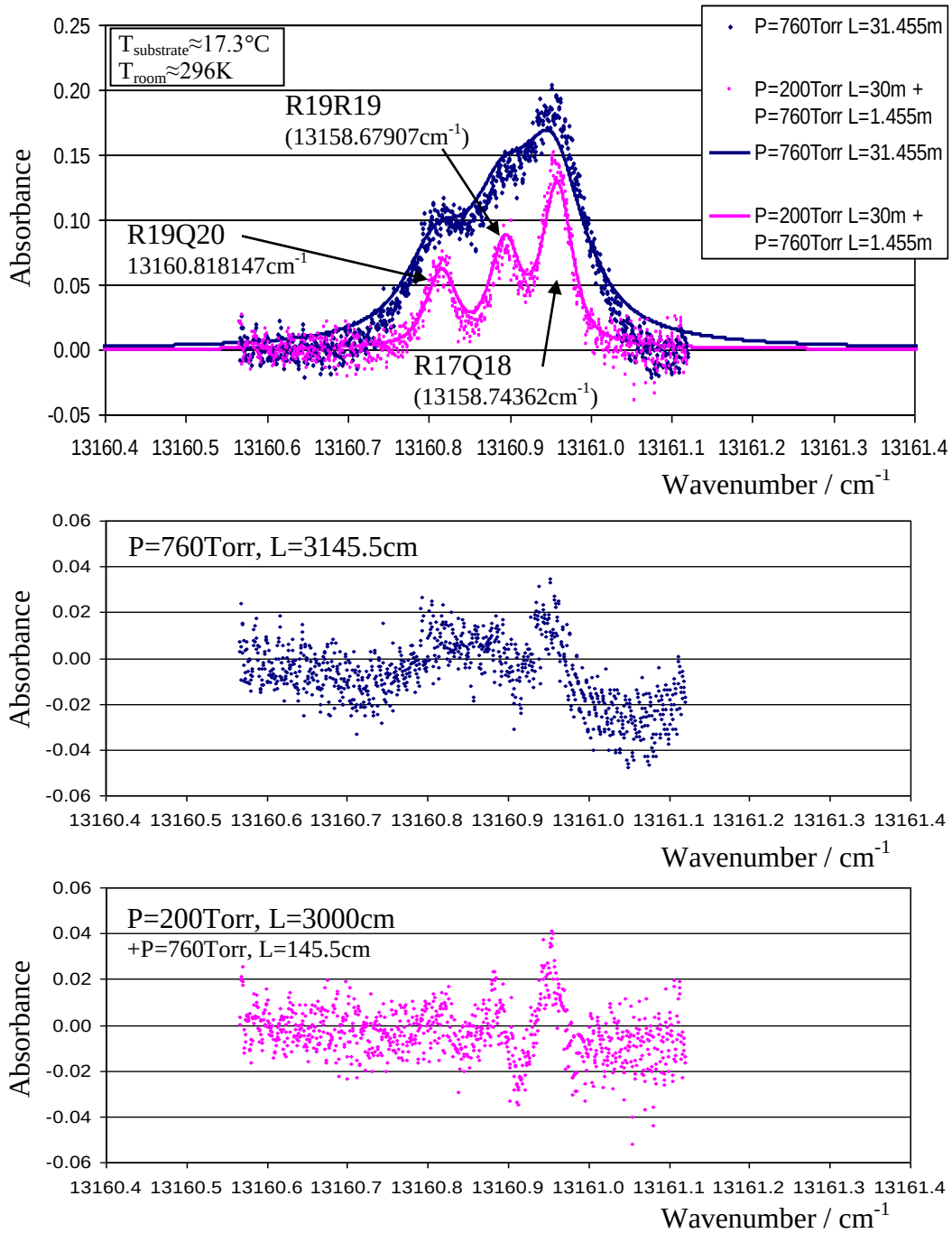


Figure 6-27: Top plot: Absorbance spectra from the wavenumber upchirp part of Figure 6-25 plotted against a multi-mode sum of Voigt models for R19Q20 (Mode 1) and R19R19 and R17Q18 (Mode 2). The simulations use a gas temperature of 296K and an oxygen abundance of 20.847%. The out of cell path length of 145.5cm has been accounted for. The absorbance weightings are given in Table 6-2. These weightings account for the different intensities of the laser cavity modes. The abscissa of the 200Torr measured data has been linearly scaled to match the known peak spacing between R19R19 and R17Q18.
Middle Plot: Residuals of the 760Torr, 3145.5cm visual line fit.
Bottom Plot: Residuals of the 200Torr, 3000cm (plus 760Torr, 145.5cm) visual line fit.

6.9 Conclusions

Although imperfect, the semiconductor diode laser's design is on the right track for visible oxygen spectroscopy.

A design strength is that the wavenumber coverage includes some fairly strong oxygen transitions, with R21Q22 at 13162.50095cm^{-1} a candidate for single mode spectroscopy with this laser. Two other regions of large absorbance have been shown to be amenable to multi-mode spectroscopy. The laser's output power is linear with injection current and stable over the time needed to produce the spectra displayed in this chapter (≈2.6s). The time to wavenumber tuning is also conveniently linear (giving a constant tuning rate). The scan range is also sufficient to capture broadened lines at 760Torr. The spectrometer's LoD was found to be 0.023 absorbance units, from averaging 128 50Hz scans.

The laser output in power and wavenumber, however, is unstable over time, particularly within the first two hours of operation. There is also hysteresis and nonlinearity upon change of substrate temperature. In terms of wavenumber coverage, much of the coverage is wasted beyond the band head of the oxygen A-band where there are no absorbances.

6.10 Suggestions for further work

A second batch laser could be analysed as in this chapter to identify improvements, particularly in regard to stability and single mode behaviour. An absorption line could then be chosen for ease of concentration measurements between a maximum and minimum set by a given application environment. The tests to be outlined in Chapter 7, for example for linearity and cross interference, could then be conducted on fixed oxygen concentrations, as a prelude to commercial applications. The continuous nature of the beam and DC detection also makes wavelength modulation spectroscopy a suitable candidate for sensitivity enhancement.

Integration of visible oxygen detection into existing mid-infrared systems would be a challenge. The semiconductor diode needs its own detector and has different electrical requirements. Timing is also an issue with the maximum repetition rate being 1kHz for the present laser. The pulsed-QCL's are generally operated with a chain of four 1000ns pulses being repeated at 50kHz.

Chapter 7

CT3000 series gas analyser construction, MCERTS testing and constructing specifications

7 CT3000 series gas analyser construction, MCERTS testing and constructing specifications

7.1 Abstract

A Cascade Technologies' CT3000 gas analyser was constructed and tested based on company procedures. The end use of this sensor was ill-defined, so a much fuller record of its capabilities than normal was established and recorded. In this way, maximum compensation could be made if particular lasers were swapped out in future for different end applications. For two of the four component lasers, the typical wavenumber windows used for gas fitting were extended by 25% and 50% respectively. This allowed a fuller idea of the compromises necessary to measure more gases with those lasers. In particular, more detailed information was collected on cross interferences and limits of detection. Along the way, a diagnostic test sheet was developed, and a much used company macro for running the linearity test spreadsheet was reverse- and re-engineered for greater process efficiency.

7.2 Introduction

The CT3000 series gas analysers have two optical paths between different mirrors in the same astigmatic Herriott gas cell. These are of 30m for low concentration measurements and 84cm for higher concentration measurements. The shorter path length reduces the absorption and allows higher concentrations to be measured without 'concentration saturation' - deviation from the Beer-Lambert Law due to line distortion (Section 7.5.3). Ideally, the concentrations to be measured for a given gas will overlap the coverage of the two paths – giving a large and continuous range of measurement.

The CT3000 analyser accommodates up to 4 mid-infrared pulsed-QCL lasers, each measuring one or more gas concentrations. The intra-pulse technique is used with 'sweep integration' - repeatedly tuning over absorption features of interest (Sections 2.9.5.2 and 4.3). Typically, N =500 pulses are averaged for a $\sqrt{N}$ enhancement of

SNR, and an absorbance spectrum is calculated from the resulting transmittance spectrum.

Fitting Lorentzian profiles to the absorbance spectrum allows the integrated absorbance to be calculated - and hence the concentration to be obtained (Equations 2-51 and 2-53). Each laser operates at 50kHz with a pulse length typically between 500 and 1000ns. The signals are digitized at 750MHz.

Lasers '2', '3' and '4' were supplied to be included in constructing and testing the analyser. The main gases probed with these lasers are CO_2 (laser 2), N_2O and CH_4 (laser 3), and NH_3 (laser 4). An SO_2 measuring laser was added as 'laser 1'.

The intended application was not given - although it could be used for atmospheric measurements, perhaps onboard a plane, or it might be used along with a visible oxygen measuring diode laser for combustion measurements. (The latter application could follow on from the work of Chapter 6). Due to this uncertainty, the approach taken was to test and document the analyser's full measuring capability with the four component lasers available. As well as documenting useful information, this data would allow maximum possible measuring capability to be retained if one or more of the lasers were later swapped out. Therefore, the maximum number of useful concentration readouts from different gases was sought per laser wavenumber fit window – sometimes by increasing the size of the windows. Concentration read outs for a typical configuration and for this analyzer can be compared by comparing Table 7-1, pA228 and Table 7-2, pA246ff.

The operating gas pressure in the cell was chosen as 200Torr, which provides a reasonable, and typical, compromise between narrow distinguishable absorption lines for many small gas molecules, and a customer's conditioning of the gas between *in situ* environment and Herriott cell. The gas temperature within the cell can be up to 190°C without causing damage, but here 50°C was chosen. This would make combustion less likely if the analyser was later used for oxygen measuring. It is also

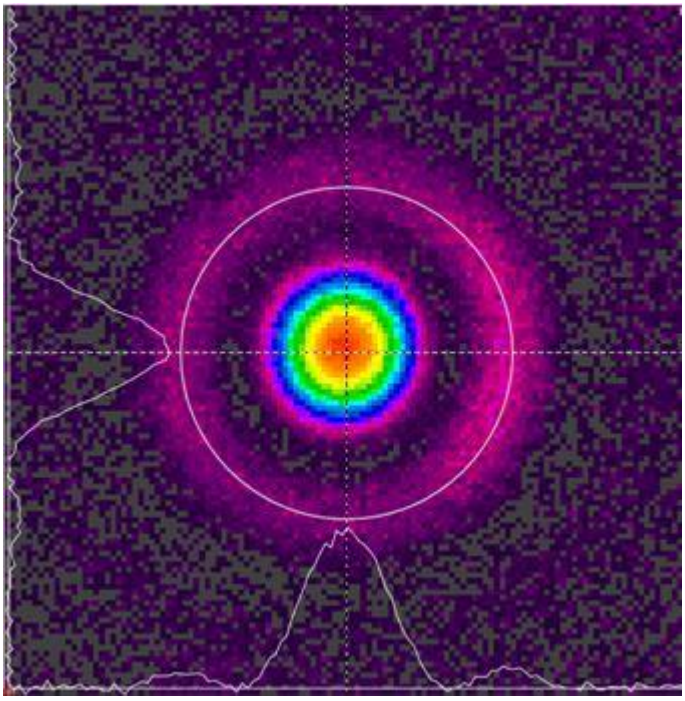


Figure 7-1: **Typical intensity Airy diffraction pattern as seen on the Pyrocam III camera, arising from the near Gaussian pulsed-QCL beam profile and diffraction of the beam through the laser's focusing optics.**

sufficiently above ambient room temperature to allow a stable operating temperature to be maintained and to prevent water condensation.

In this chapter, the construction and alignment of the analyser is first described. Next described is constructing the analyser's concentration specifications and test parameters, and testing of the analyser itself. These parameters were influenced from the pass / fail criteria embedded in the company's test spreadsheets. These criteria were deciphered and adapted to the greater number of gases tested here.

The day after the thermal test, the pulse strengths had dropped severely. A diagnostic test sheet was therefore created to eliminate various parts of the analyser as the cause. Also, the workings of the Visual Basic for Applications (VBA) macro which ran the linearity test spreadsheet had passed out of company memory, so it was deciphered and re-engineered to improve data processing efficiency. The 'concentration saturation' behaviour of the analyzer was also recorded – recording deviations from the Beer-Lambert Law due to lineshape distortions (Section 7.5.3). This allowed a rationale to be recorded for taking the maximum used absorbance of a line as ≈3 absorbance units. A fuller than usual gas measuring capability results sheet was prepared. This is summarized in Table 7-2, pA246ff.

7.3 Gas Analyser Construction

7.3.1 Laser alignment

Each pulsed-QCL laser was operated at 50kHz, duty cycle ≈3.5%, and, initially, with generic temperature, voltage and pulse length settings. These settings approximate the wavenumber windows required for concentration fitting – the settings to be adjusted during the calibration stage (Section 7.3.2 below).

Each laser beam was collimated along the length of a Spiricon PyroCam III camera rail, via the lateral and longitudinal adjustment of a ZnSe lens mounted inside the

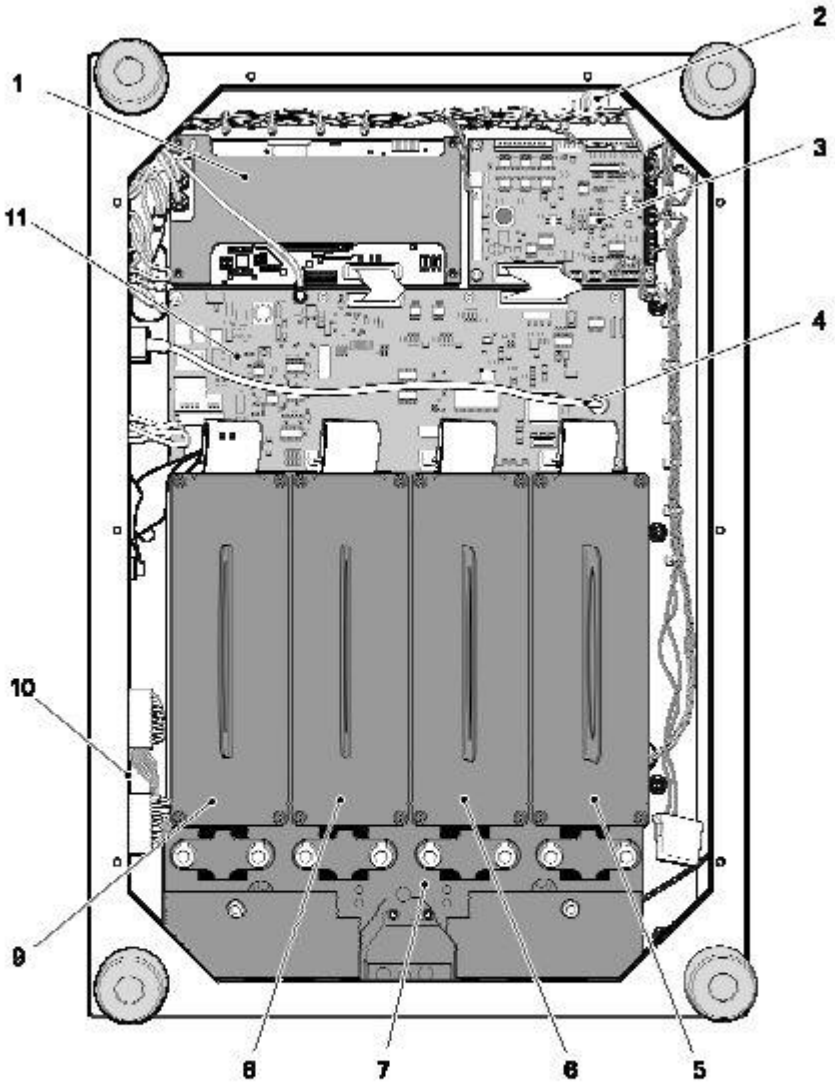
rectangular laser housings. The observed Airy intensity patterns were matched to charts displaying patterns known to give good laser collimation through the CT3000 CEM analyser. A typical pattern is shown in Figure 7-1.

7.3.2 Laser calibration

In order to facilitate the measurement of a gas' concentration, the generic operating parameters for the particular lasers were adjusted. The laser must tune over a particular wavenumber window to capture a specific absorption line (or group of lines), with a line strength which will enable the full range of concentrations to be measured when in the intended application environment. Also, for these concentrations to be measured, the particular polynomial converting time to wavenumber for each particular laser must also be determined and recorded from etalon fringes.

Therefore, each laser was calibrated on a company test setup, with each laser and a dedicated VIGO detector run from electronics equivalent to those in the CT3000 system. The beam goes through a 20cm gas cell containing a relatively high known concentration of the relevant gas at 200Torr - the typical operating pressure of the CT3000. A retractable Ge etalon in the beam path allows etalon fringes to be recorded with nitrogen gas flowing in the cell.

The laser's pulse duration and voltage were tweaked so that the pulse length and wavenumber chirp rate produced the correct number of etalon fringes for the desired window. The laser's substrate temperature was adjusted to place the absorption feature over the correct fringe within this window. Cascade Technologies' software was then set to adjust the (typically quadratic) time to wavenumber polynomial via adjusting the etalon fringe FSR and the plot's x-shift - so that the Lorentzian fit to the line shape produces the known gas concentration. The fit was further improved by manually adjusting the FSR and x-shift (wavenumber shift) and by adjusting the polynomial coefficients themselves. The laser operating parameters were juggled if



1	Digitiser PCA	7	Laser Manifold
2	LED PCA	8	Laser Module
3	Peripheral PCA	9	Laser Module
4	Ethernet cable	10	Detector TEC PCA
5	Laser Module	11	Motherboard
6	Laser Module		

Figure 7-2: Bottom of CT3000 series analyser with cover removed, showing the 4 laser modules bolted into a combiner block. The block directs radiation from the 4 pulsed-QCL lasers (and optionally a visible solid state alignment laser) up to a beam splitter in the top of the analyser. (PCA = Printed Circuit Assembly). (Diagram from the CT3000 operation manual (Cascade Technologies Ltd., 2012)).

necessary to keep the trapezoidal shaped laser pulses relatively square (i.e. relatively flat topped). This ensures that the laser is not over stressed.

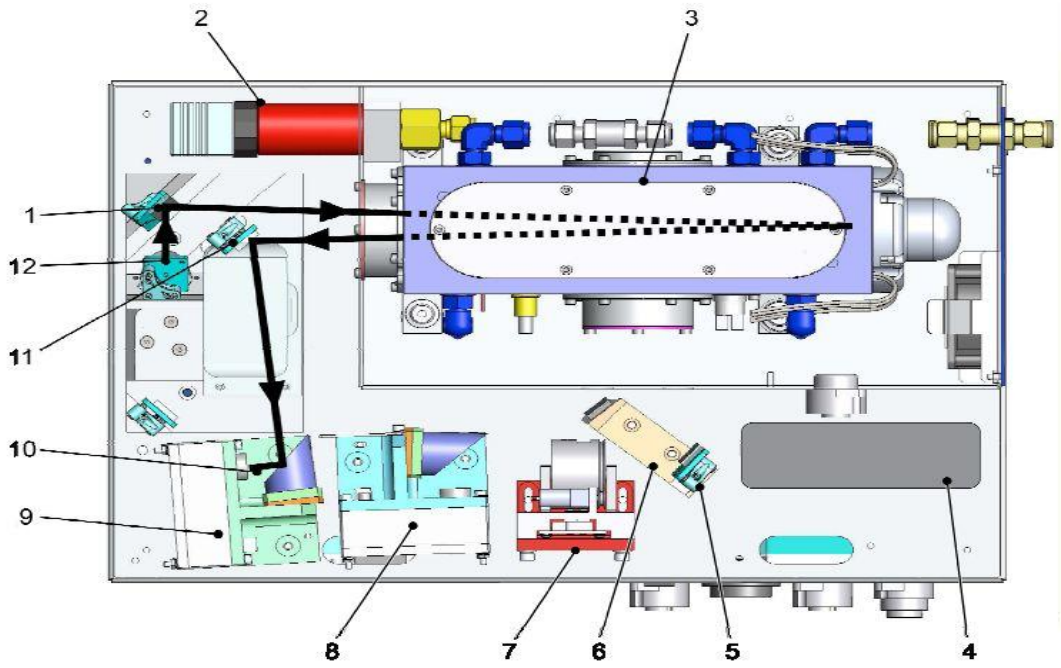
7.3.3 Combiner block alignment

Once calibrated, the lasers' rectangular housings were bolted into a single combiner block. The block provides a steering mirror for each laser, guiding the beams at near right angles onto a central prism. A reflective coating on the outside of the prism allows the beams to be sent up through the CT3000 once the block is aligned. A hole in the middle of the prism allows the QCL beams to be bunched around the beam from a red solid state alignment laser.

The whole block was then mounted onto the Spiricon alignment system and the steering mirrors adjusted, so that the QCL beams became arranged in a square, with the alignment laser at centre. The whole pattern was offset slightly from normal incidence, to compensate for the known deviation it will encounter through the ZnSe beamsplitter of the CT3000 system.

7.3.4 CT3000 series analyser alignment

Figure 7-2 shows the underside of the CT3000 with the bottom cover removed. The four lasers' rectangular housings were bolted to the combiner block and their beams sent upwards by the block's prism into the upside of the analyser. Figure 7-3 shows the analyser's long path, after being collimated by two OAP's before reaching a beamsplitter. Portions of the beams which come up through the ZnSe beamsplitter, are angled down by a mirror at (12) and sent into the 30m path length Herriott cell by the input mirror at (1). The astigmatic mirror at the back of the cell was translated and rotated until the known Lissajou pattern for 30m was achieved with the red alignment beam. The exiting beam was then focused by an OAP mirror onto a VIGO detector at (9).



1	Mirror	7	Etalon Assembly
2	Pressure Sensor	8	Short Path Detector Assembly
3	Analysis Cell	9	Detector
4	Relay box	10	Long Path Detector Assembly
5	Mirror	11	Mirror
6	Short Path Steering Assembly	12	Beam Steering Block Assembly

Figure 7-3: Top of CT3000 series analyser with the cover removed, showing the optical long path from the bottom of the analyser to the VIGO detector. The path length within the 20cm Herriott cell is 30m, implying 75 radiation round trips between entering and leaving the cell. The beam splitter sits under the mirror marked at (12). (Diagram from the CT3000 operation manual (Cascade Technologies Ltd., 2012)).

The Herriott cell was aligned at room temperature for convenience, and the position of the exit beam on the exit mirror at (11) noted. The cell was then brought up to operating temperature (50°C in this case) and the displaced beam brought back into alignment.

Figure 7-4 shows a portion of the beams being reflected by the ZnSe beamsplitter horizontally to the analyser base at (1), onto a steering mirror at (9). From here, the beams were steered onto an 84cm path length Lissajou pattern, with the exit beam being focused by an OAP mirror onto a VIGO detector at (8). The short path also contains a retractable Ge etalon at (7) for *in situ* calibration of the lasers and gas fits.

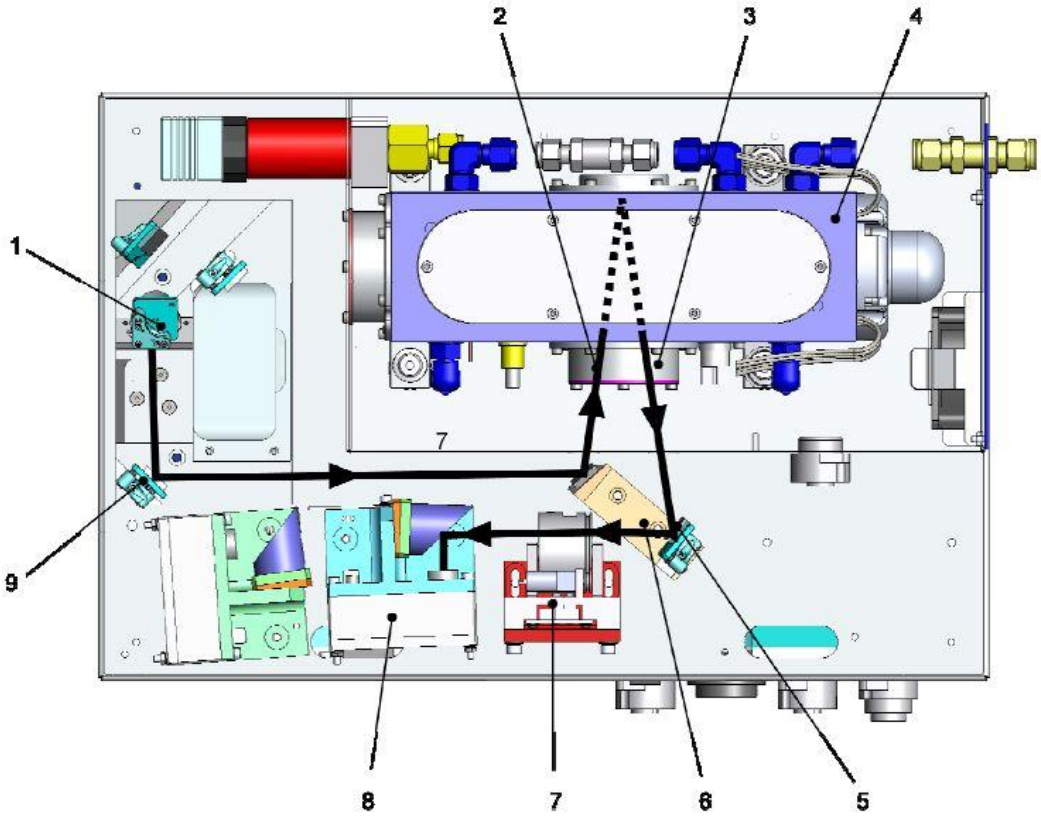
7.4 Drop test

After the alignment was satisfactory, the steering mirror mounts were locked down with araldite and the mirror alignment screws were secured with Loctite. The analyser was then dropped from a few centimetres to ensure that the alignment, and the rest of the system, was sufficiently robust.

7.5 MCERTS testing

Having been constructed, the analyser needed setup for final use, and tested to prove the reliability, accuracy, and repeatability of its concentration measurements. The tests that Cascade Technologies use are based on standards such as the UK Environment Agency's Measurement and Certification Scheme (MCERTS).

The tests consist of: thermal drift; zero noise and maximum (span) concentration noise - on both the long and short paths, 24 hour drift of the zero and span concentration readings; pressure drift; linearity; limit of detection (LoD); and cross interference testing.



1	Beam Splitter Block Assembly	6	Short Path Steering Assembly
2	Optical window (IN)	7	Etalon Assembly
3	Optical window (OUT)	8	Short Path Detector Assembly
4	Analysis Cell	9	Mirror
5	Mirror		

Figure 7-4: Top of CT3000 series analyser with the cover removed, showing the optical short path from the bottom of the analyser to the VIGO detector. The physical path length across the Herriott cell is 7cm with a Lissajou pattern for 84cm, implying 6 radiation round trips between entering and leaving the cell. The beam splitter sits under the mirror marked at (1). (Diagram from the CT3000 operation manual (Cascade Technologies Ltd., 2012)).

7.5.1 Thermal testing

The analyser's surrounding temperature has some affect on alignment and laser tuning, so it is important that the machine should still function correctly within its advertised range. For this reason, the analyser was placed in a thermal cabinet, and its performance over a period of hours assessed at each of the surrounding temperatures of 25°C, 45°C, and back down to 25°C. The change in laser pulse strengths should remain within acceptable margins - indicating that changes of alignment are within acceptable bounds. A bad failure can sometimes indicate that there had been unrelieved mechanical strains in the system.

The Lorentzian fitting software expects to find a particular spectral window. Since the surrounding temperature affects the laser's tuning, the Lorentzian fitting routine can become disturbed. This leads to noisy and inaccurate concentration readings. Two parameters per laser were therefore recorded into the analyser's configuration file. These parameters compensate for the x-shift and stretch of the simulated fitting spectra at 45°C, and indicate the linear interpolation needed at other surrounding temperatures. Since there are only two parameters per laser, they were taken on the main gas in the fit window for that laser. So that the linear correction is adequate, the concentration readings at 25°C and 45°C should be within ±4% of each other prior to correction.

7.5.2 Troubleshooting pulse intensity failure

Although the analyser passed the thermal test, the next day the intensity of one laser pulse was found to be much reduced, and the other three were almost non-existent. A visual inspection of the Herriott cell by removing its lid revealed nothing obviously wrong. There was also much company resistance to issuing another cell, or taking the mirrors out of the existing cell for cleaning – due to expense and the possibility of damage.

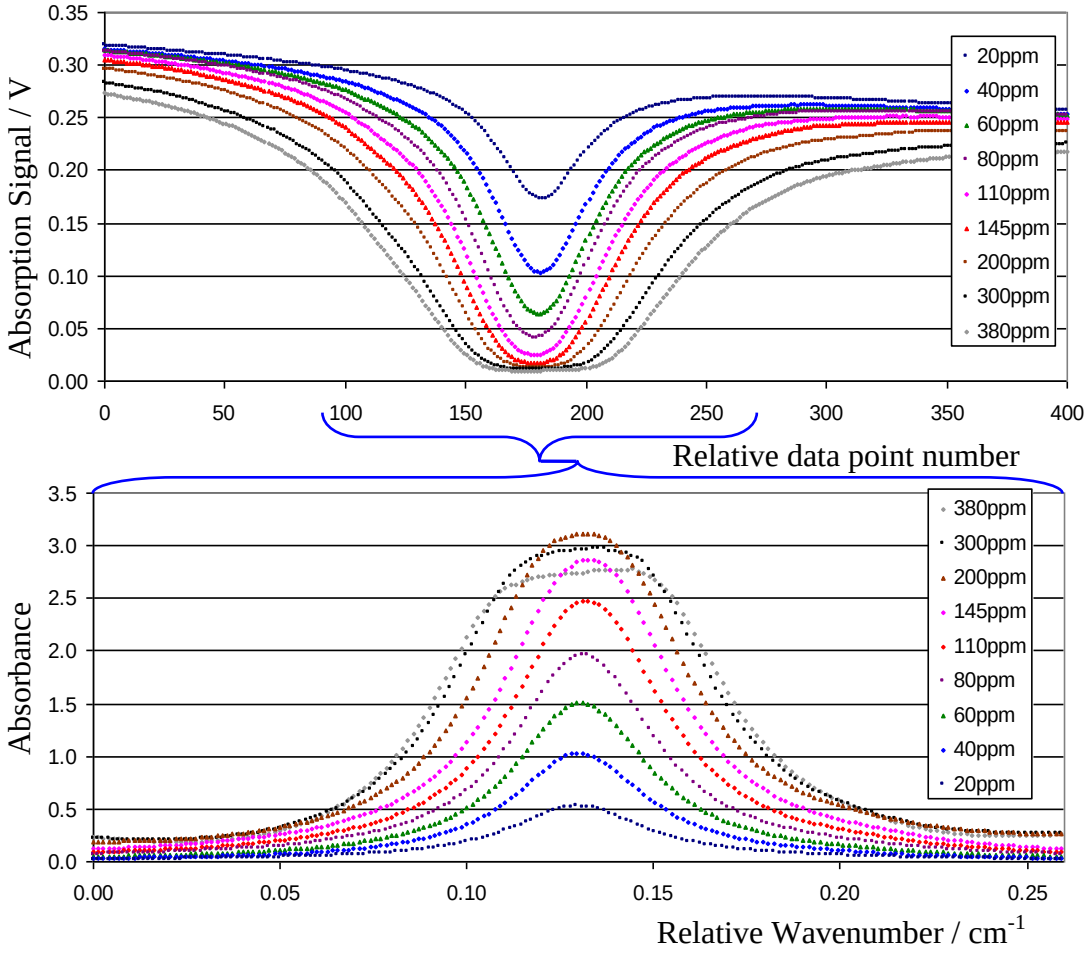


Figure 7-5: **Top plot: Absorption spectrum vs time taken from the detected radiation pulse. As the concentration in the gas cell increases, the absorption increases. Eventually, all the radiation available at the centre frequency is absorbed or nearly absorbed. However, radiation at frequencies in the absorption wings can still readily absorb – giving line shape distortion.**

Bottom plot: Absorbance plot from the absorption data. At first there is a linear relation between absorption peak height and integrated absorbance (bottom plot of Figure 7-6), the latter being proportional to the measured concentration in the CT3000 software. After an absorbance ≈2 - here at a supplied concentration of 80ppm - the peak absorbance lags the integrated absorbance. The integrated absorbance (and measured concentration) shows a linear relationship to the supplied concentration until a supplied concentration of 145ppm is reached (top plot of Figure 7-6). This supplied concentration corresponds to a peak absorbance of ≈3 (middle plot of Figure 7-6).

Due to these circumstances, a diagnostic test sheet was developed to eliminate the possibility of other faults - by showing that the rest of the analyser was working correctly. The test sheet can therefore be used in more general circumstances and is given in Appendix A. It partly consists of re-doing the alignments and replacing optical components, prior to the thermal testing stage, and bypassing various electronic boards by using those on the calibration rigs, and using software known to be working.

Realignment made no substantial difference to the signals, nor did operating the analyser from other computers. Eventually, permission was granted to swap out the analyser's Herriott cell with one known to be working from another analyser. This new cell functioned correctly within the analyser.

The mirrors in the faulty cell were then cleaned, which partially, but not satisfactorily, restored some of the signal strength. A swab of the inside of the cell was taken which found the cell to be coated with a thick black residue – which it was communicated was not normal. The metalwork had been baked for 24 hours upon arrival from the manufacturers - for the purposes of cleaning and out-gassing. This was redone, with the residue reducing but not clearing – indicating that the dirt had come in from the manufacturer, rather than the cell having missed its initial baking.

A new Herriott cell was issued and the analyser alignment of Section 7.3.4 redone, this time the analyser having no problems after the thermal test. The previous failure was likely due to the release of the residue when thermally tested - into the cell and onto the mirrors. It would be highly scattering / absorbing in the mid-infrared. The fact that one pulse was less affected than the others indicates a wavenumber dependence to this scattering. A sample of the residue was kept in case Cascade Technologies had been interested in its analysis.

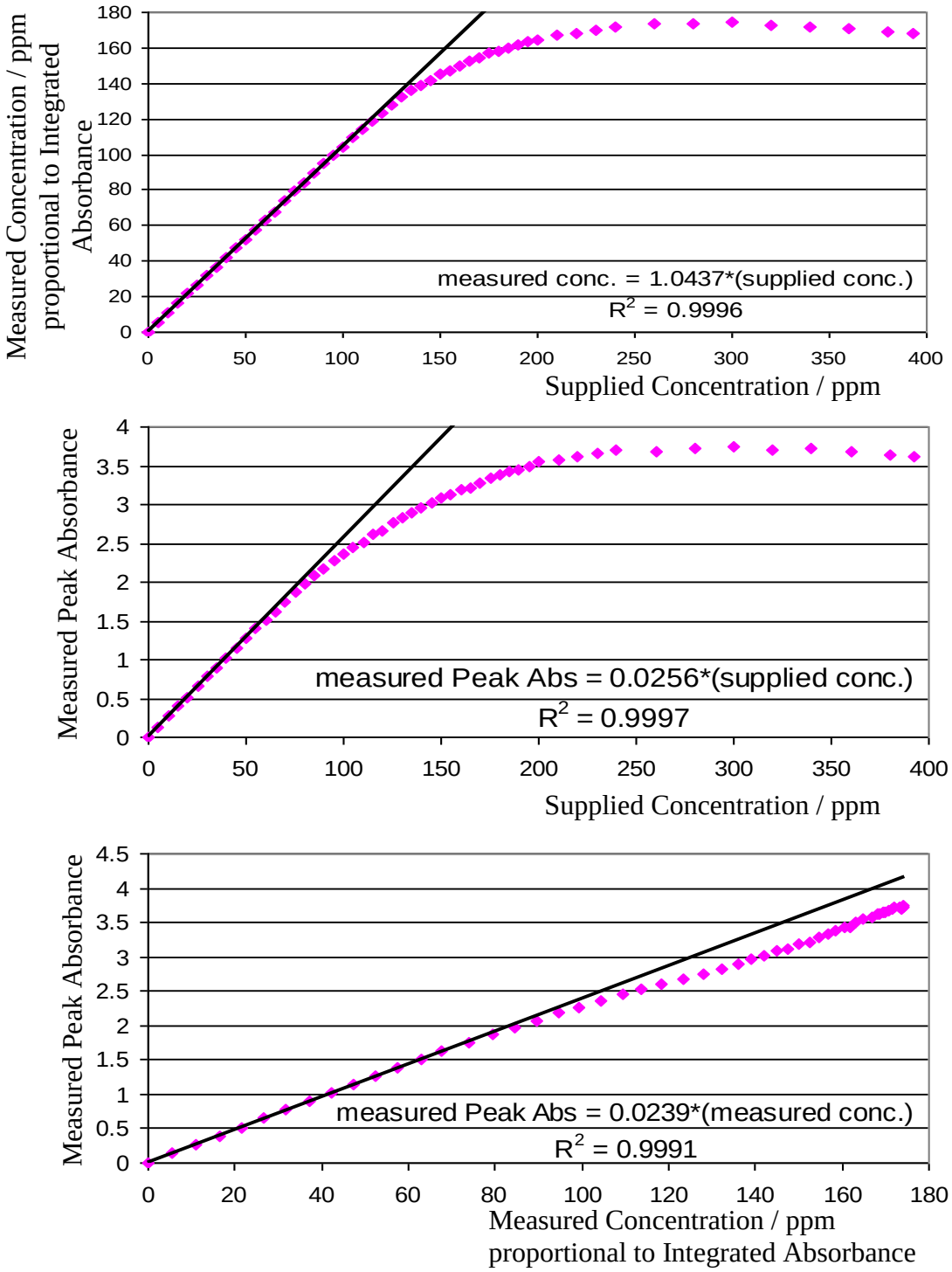


Figure 7-6: **The top plot shows the measured concentration (and hence the integrated absorbance which is calculated from it) having a linear relation to the supplied concentration up to ≈145ppm. The middle plot shows this concentration occurring at a peak absorbance of ≈3. The bottom and middle plots show that the peak absorbance begins to lag the integrated absorbance (and measured concentration) after an absorbance of ≈2.**

7.5.3 CT3000 analyser concentration saturation behaviour

The known maximum absorbance noise for the CT3000 analyser under normal operation is 0.004 absorbance units. The maximum peak absorbance used for extracting a concentration value (from the integrated absorbance) is around 3 absorbance units. No records had been kept on the rationale for this upper limit, so it was worth digressing to explore the analyser's non-linear behaviour with increasing concentration.

The top plot of Figure 7-5 shows a main absorption line captured within a laser pulse. As the laser tunes, each time point corresponds to a frequency point (or more correctly a small group of frequencies which make up the effective laser linewidth). As the gas concentration increases, more and more of the laser's radiation is absorbed. Eventually, most of the radiation available at line centre has been absorbed. However, the molecules can still absorb radiation with frequencies corresponding to those in the lineshape wings. This flattens out the absorption peak and raises the wings – giving line distortion from the simple (Lorentzian) profile.

The bottom plot of Figure 7-5 shows the absorbance plot calculated from the absorption spectra. Since the measured concentration in the CT3000 software comes from the integrated absorbance, the distortion in line shape with increased concentration breaks the linear relationship with supplied concentration.

The top plot of Figure 7-6 shows that the linear relation between integrated absorbance (and hence measured concentration) and supplied concentration holds up to a concentration of ≈145ppm. The middle plot shows this supplied concentration to correspond to a peak absorbance of ≈3 (which corresponds to a transmittance of ≈5%).

Taken together, the top and middle plots therefore indicate that the integrated absorbance is a more linear measure of supplied concentration – compared to the peak absorbance - over a larger concentration range. The peak absorbance is linear

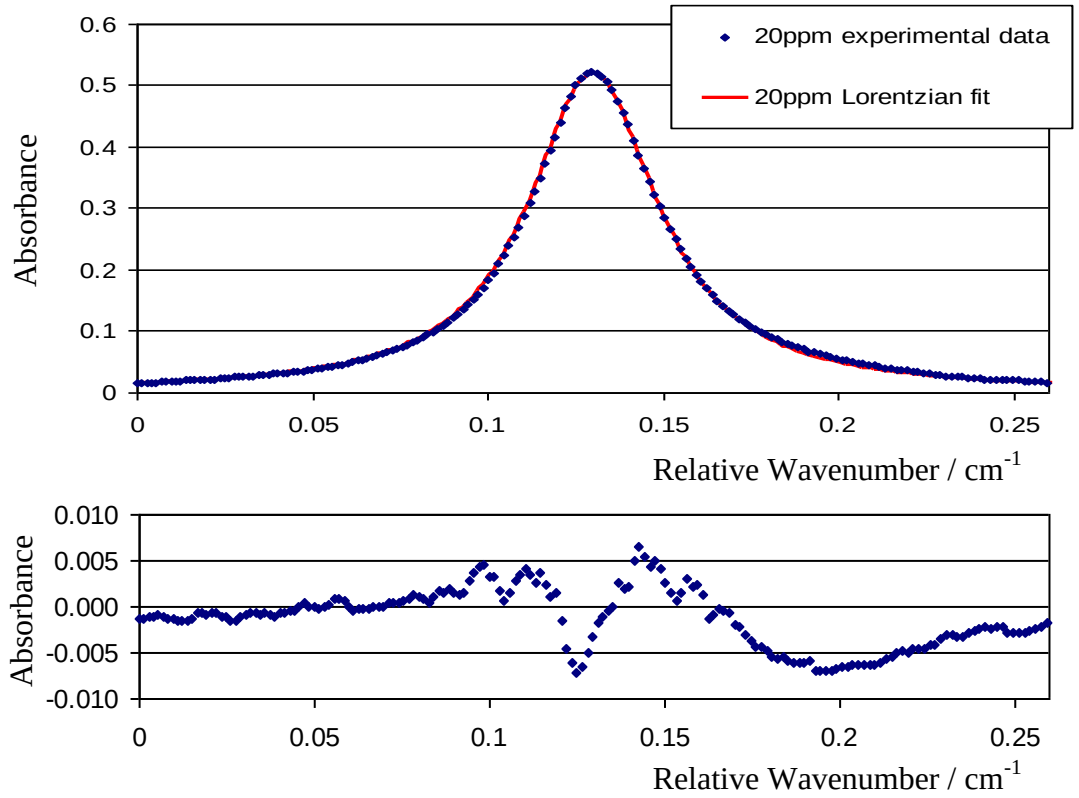


Figure 7-7: **Top plot: Measured absorbance for a gas absorption line at 20ppm, diluted in nitrogen gas. The total gas pressure is 200Torr and its temperature is 50°C. Also shown is a Lorentzian fit to the data.**
Bottom plot: Residuals of the fit from the top plot: observed minus calculated data.

	Gas	Max Conc.	LoD
Laser 1	SO_2	2000ppm	0.5ppm
Laser 2	CO_2	15%	0.001%
Laser 3	N_2O CH_4	5000ppm 10,000ppm	0.005ppm 0.5ppm
Laser 4	NH_3	2000ppm	0.1ppm
Repeatability: ±2%, Accuracy: ±2%, Linearity: R^2 >0.999			

Table 7-1: **One common advertised specification for measuring gases on the four lasers which were installed in this particular CT3000 analyser. The concentrations measured vary depending on the particular fitting windows and absorption lines chosen to suit a particular application or gas environment. (Extracted from Cascade Technologies Ltd, 2014b)**

with supplied concentration to an absorbance of ≈2, after which line distortion has a greater affect on it than on integrated absorbance. The bottom plot shows the peak absorbance lagging the integrated absorbance after a peak absorbance of ≈2 (transmittance ≈13.5%).

7.5.4 Comments on system resolution

Figure 7-7 shows a Lorentzian fit to the 20ppm pressure broadened line of Figure 7-5 (20ppm N_2O diluted in N_2), at 200Torr and 50ºC. The fit was performed in a cut down version of Chapter 3's analysis program, allowing data already in absorbance vs wavenumber format to be fitted. A FWHM value ~0.045cm^{-1} was obtained. The Lorentzian FWHM shown in Figure 7-7 is 2% larger than the calculated value of ~0.044cm^{-1} (from the parameter calculator of Section 2.5.1). This compares particularly favourably to the circa 67% to 92% broadenings seen with the spectrometer of Chapter 4. The improved performance of the CT3000 is likely due to optimized electronics, but also due to the pressure broadened (Lorentzian) lines at 200Torr having FWHM's an order of magnitude larger than the Doppler widths seen in Chapter 4. Instability in the line centre position due to electronic and laser fluctuations will therefore have proportionately less affect on the measured line width.

Sliding two Lorentzian lines together, of the FWHM seen in Figure 7-7, allows the disappearance of the saddle point in the superposition to be evaluated (as was illustrated in Section 4.6.4 for two similar Gaussian line profiles). This gives a practical measure of resolution of 0.027cm^{-1} for this line, at 200Torr and 50ºC.

In Chapter 4, lines at much lower pressure (<2Torr) were analysed, and their smaller FWHM's, even with the large percentage broadenings, gave an effective resolution of ~0.004cm^{-1}. The contribution to this line width due to cw-QCL chirp rate broadening was 0.0005cm^{-1}, from relatively slow chirp rates of a few cm^{-1}ms^{-1}. The quoted figure for best CT3000 instrument resolution is an order of magnitude larger at 0.004cm^{-1}. This is a value at the end of a long pulse limited by the chirp rate of the

system's pulsed-QCL's, of a few $cm^{-1}\mu s^{-1}$. This resolution value is larger than the Doppler width of the line considered here (~$0.0024cm^{-1}$), so that the CT3000 resolution is chirp rate limited if used to measure low pressure lines for general spectroscopy. The sensors design, however, is for a more specific purpose - to measure the concentrations of particular gases in particular gas mixtures. It does this by targeting known lines or groups of lines known to be isolated, or with manageable levels of cross-interference from other known gases. The standard cell pressure is 200Torr (Section 7.2) giving FWHM's limited by pressure broadening. General resolution is therefore a less important criterion for the CT3000 than sensitivity (0.004 absorbance units).

7.5.5 Constructing the analyser specifications

If working from scratch, details on the gases and range of concentrations in a given customer's environment would be obtained. Suitable gas fit windows would need to be matched to available lasers. Ideally, the absorption lines chosen would have line strengths such that the concentration range to be measured matched the difference between the analyzer's minimum (0.004) and maximum (3) absorbances (Section 7.5.3). If conversely, the absorbance feature changed only a little over the full concentration range, small changes in absorbance, maybe due to noise, would give large errors in concentration reading. Additionally, a line or group of lines would be sought which gave a smooth and overlapping concentration reading at the switch over from the long to the short path. The measured lines should also be as free as possible from cross interference from other gases in the measurement environment.

In the present case, however, several wavenumber fitting windows were well established for each of the main gases on the four lasers available - these windows already suiting the typical needs of customers. One such grouping of windows gives the advertised specifications seen in Table 7-1.

Firstly, from these well established choices, a wavenumber window was chosen for each laser which well matched the lasers' operating parameters. There are two ranges

associated with each wavenumber window. The narrower range homes in on the wanted measured absorption lines within the laser pulse. The second is wider and encompasses the first, and is used solely for extending the simulated fit spectrum. This allows lines outside the measured range to be simulated so that their affect upon the baseline within the measured range can be correctly modeled. This allows the transmittance and absorbance spectra to be correctly calculated. (However, the converse was found to not work at all well. That is, it was not possible to get an accurate concentration of an unmeasured line outside the fitting window, by fitting to the line's tail within the fitting window).

Next, for the gases most usually fitted on these lasers, the likely limit of detections (LoD's) and the maximum (span) concentrations (Section 7.5.8) were estimated. These estimates came from the equivalent concentrations for a line's absorbance peak value - modeled on an absorbance plot in Cascade Technologies' Lorentzian spectral modeler - 'Simspec' – within the absorbance range of ≈0.004 to ≈3 absorbance units. These LoD's and spans were matched as closely as possible to the concentrations of the available test gas cylinders.

Since at least one laser might be swapped out for an oxygen laser at a later date, a fuller set of gases was tested on each of the four lasers. This allows as much of the analyser's capability to be retained as possible. In some cases, reduced performance criteria resulted, and these were noted. Also for the fits on Lasers 1 and 2, the usual fitting windows were expanded by 25% and 50% respectively. The gases tested and notes on performance compromises can be found in Table 7-2, pA246ff.

Plots of multiple gases were used in Simspec to find which gases would likely interfere with each other, and at what concentration level. Once this data had been collected, figures for the pass or fail criteria could be drawn up. The pass and fail criteria were deciphered from the templates of the company's test Excel spreadsheets. The spreadsheets were then expanded to suit these fuller tests. The pass / fail criteria are given in the accounts of the individual tests below.

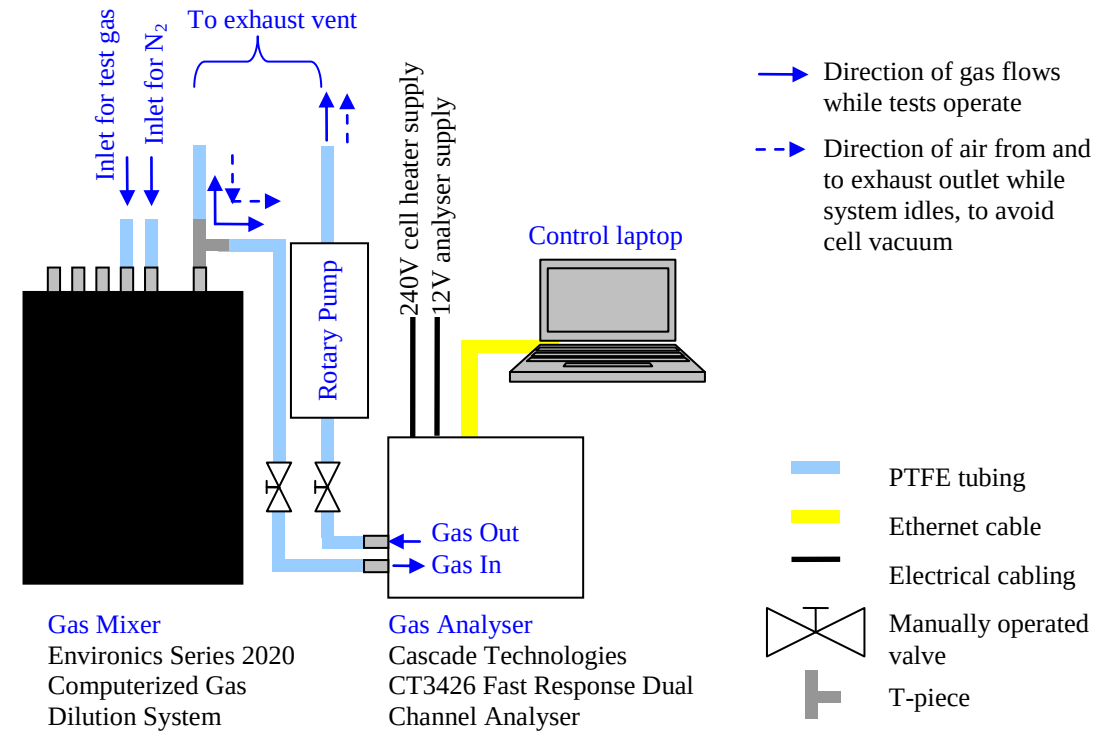


Figure 7-8: Test setup for most MCERTS testing.

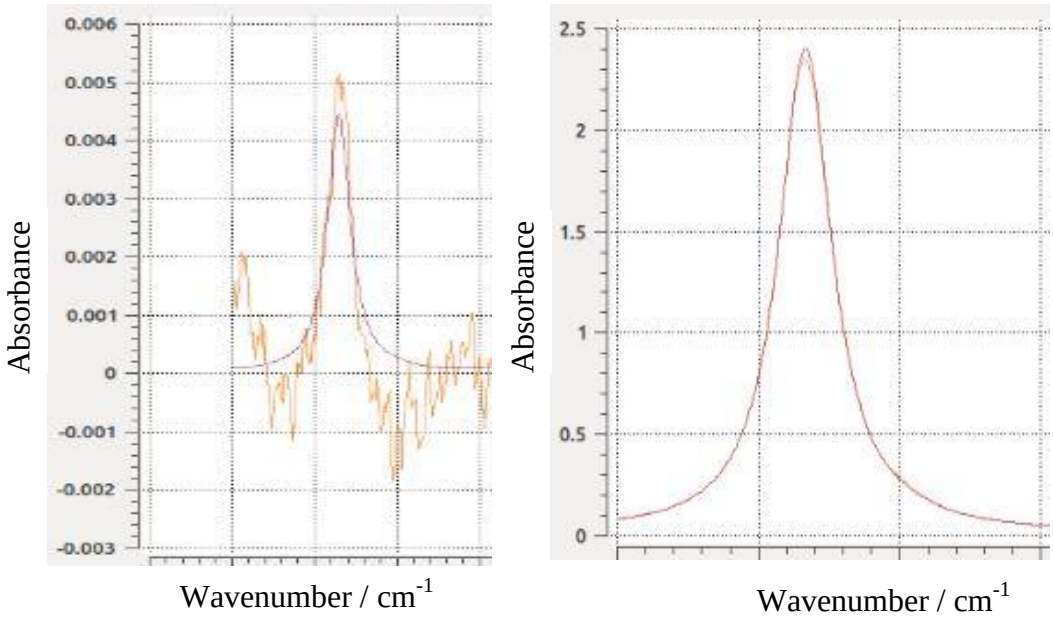


Figure 7-9: Lorentzian fits in purple, at relatively low (left) and high (right) concentrations, with absorbances in orange calculated from measured absorption lines.

7.5.6 Test setup for remaining tests

A test setup was put together according to Cascade Technologies' instructions. A schematic of this is shown in Figure 7-8. The test gases, already diluted with nitrogen, are 'cut' further with more nitrogen to the required concentrations via an Environics Series 2020 gas mixer. It is important that the Herriott cell is not evacuated to avoid the possibility of distortion of its astigmatic mirrors. For this reason, when no test gases are flowing, the rotary pump draws air from the exhaust vent through the cell. When the test gas flows from the mixer, it is set to do so at $6Lmin^{-1}$. The pump draws at most $4Lmin^{-1}$ through the cell, so the remaining $2Lmin^{-1}$ of test gas is T-pieced straight to the exhaust - preventing exhaust gas inflow while testing. The pressure in the cell is maintained at 200Torr by controlling the gas flow with the manually operated valves.

7.5.7 Distinguishing between absorbance and concentration noise

It is helpful to be clear on the distinction between absorbance noise and concentration noise before considering the tests. The absorbance noise can be observed on the absorbance vs wavenumber plot for each fit. In the CT3000 system this consists of random noise riding on larger and more slowly varying unwanted etalon fringing (e.g. Figure 7-13, Section 7.6.2). To get each single concentration vs time data point, 500 sequential absorbance spectra are averaged. A fitting routine then iterates in order to fit a Lorentzian lineshape (or lineshapes) to the average spectra. The integrated absorbance of this simulated spectrum is then used to get a (single) concentration value. Figure 7-9 shows illustrative absorbance fits to a relatively low and high concentration over time. The linearity testing graphs of Figure 7-10 through Figure 7-12 will show concentration vs time plots. As the supplied concentration increases, the peak of, and area under, the Lorentzian absorbance lineshape increases. Each concentration data point is calculated from integrating the area under the fit of the absorbance spectrum, like those shown in Figure 7-9.

Thus, the spread of concentration values - for a given supplied concentration - depends on the stability of the Lorentzian fit to the spectra over time.

Although the stability of the fit is affected by the absorbance noise, it is possible, for instance, to have large amplitude absorbance noise but a smaller spread in concentration values, depending on how the absorbance noise affects the fit stability. In summary:

Absorbance noise: the baseline noise on the absorbance versus wavenumber plot used for fitting a (Lorentzian) absorbance profile. The most useful measure is maximum absorbance noise, due to random noise riding on unwanted etalon fringes.

Concentration noise: the noise on the concentration versus time plot. Each concentration value comes from a Lorentzian fit on the absorbance versus wavenumber plot. This noise is a spread of values caused by the fluctuations of the Lorentzian fit. The most useful measure is the 2σ value of the spread about a mean value.

7.5.8 Zero and span calibration, zero and span noise, and 24 hour drift (repeatability) testing

A well aligned CT3000 analyser has a known peak absorbance noise of ≤ 0.004 absorbance units. The absorbance spectra for each fit window is therefore inspected and recorded to show that the absorbance noise meets this criterion at zero concentration. There are normally 8 such fit windows to check - two per laser, one for the long path and one for the short path. For this particular analyzer, the peak absorbance noise for these fit windows ranged from 0.0025 to 0.004 absorbance units. When using wider windows to fit more gases, the peak absorbance noise could increase to 0.006, and occasionally 0.008 on the edges of the windows.

Calibrating the zero and span concentrations can be done separately or treated as part of the analyser's linearization (Section 7.6 below):

Zero concentration: the concentration reading, for a particular gas on a particular path and fit window, when the gas cell has only nitrogen gas flowing through it.

Span concentration: the maximum concentration that the analyzer is configured to measure, for a particular gas on a particular path and fit window, to have a guaranteed measurement accuracy of ±2% of the supplied concentration.

The end points of the measured concentration range - zero and span concentration – are fixed to agree with the supplied concentrations. Firstly, the gas cell is purged with nitrogen gas, and the nitrogen gas flow maintained. The gas fitting routine will in general output a non-zero value, partly due to noise and partly due to the assumption of a minimum concentration – which prevents the fitting routine 'freezing'. This concentration reading then provides a linear shift to be applied to all measured concentration readings. This means that the measured value at zero target gas concentration does indeed output a mean of zero. Secondly, the maximum (span) concentration of the target gas is applied to the gas cell. To make the measured span reading agree with the supplied concentration, a multiplication factor is applied to all concentration readings. This multiplication leaves the mean of the zero concentration reading unchanged at zero.

At this stage, 2σ of the zero concentration readings should be below the Limit of Detection (LoD) – though this will be revisited during LoD testing (Section 7.6.2). The average of the zero concentration reading is also taken at least 24 hours later. The pass criterion is that the reading will have drifted by less than the LoD, and thus not be giving a false positive reading. Both the long and short paths have their own LoD's.

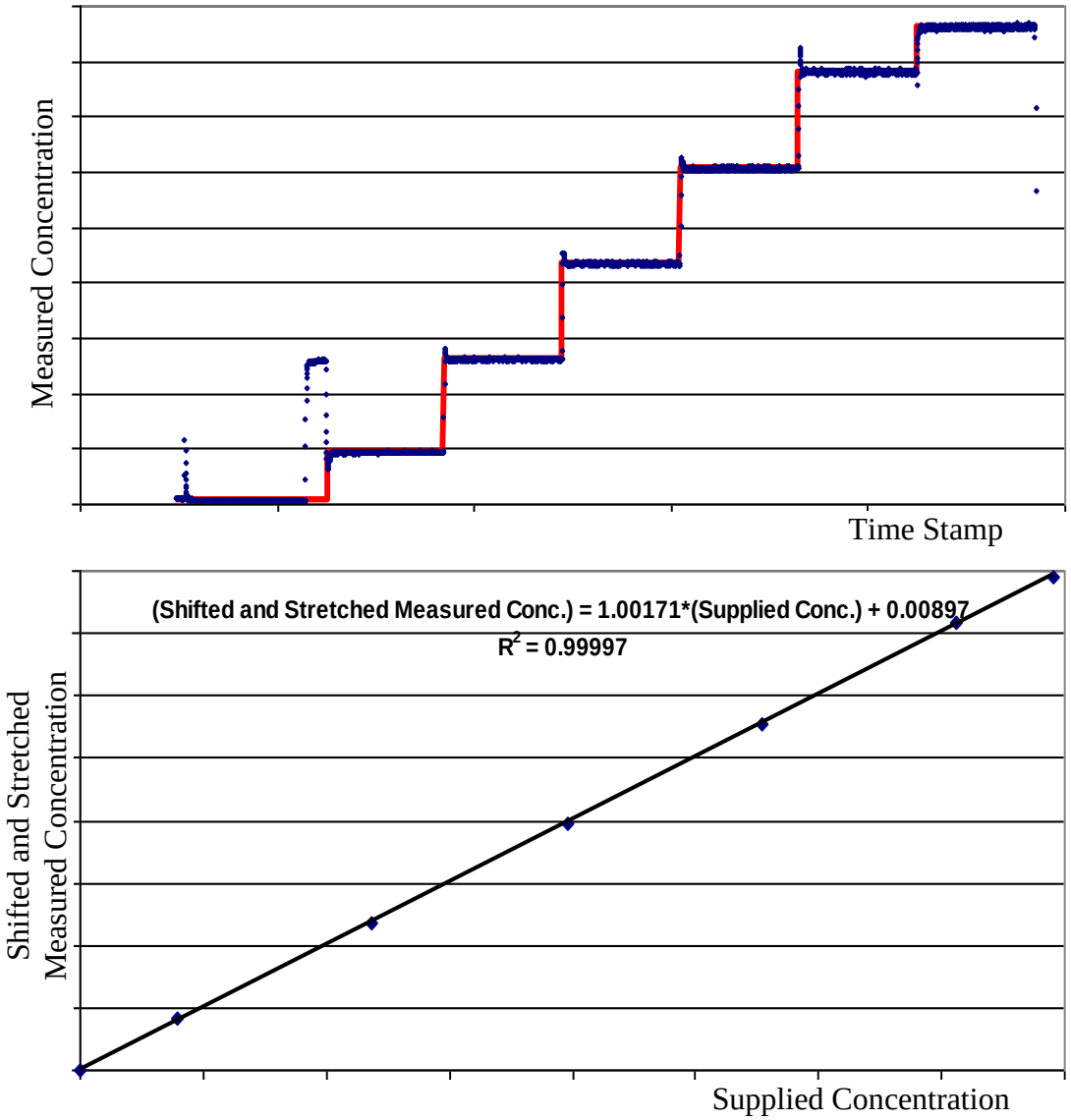


Figure 7-10: **Top graph: The typical form of measured concentration data, on the long or short optical path, during linearity and accuracy testing. The concentration readings come from the fitting of a Lorentzian lineshape to the measured absorbance spectra. Typically, the absorbance line (or strongest lines of a group) within the fit window reaches an absorbance of ≈3 when its concentration reaches span. In this way a change in concentration gives a significant change in integrated absorbance, minimizing the noise in the concentration reading.**
The horizontal sections of the red trace mark the average concentration of the steps.
Bottom graph: The average measured concentrations from the top figure, shifted and stretched, so that the end points match the supplied concentration. A line of best fit is then plotted from the 7 data points.

The 2σ noise and 24 hour drift of the average is similarly taken at the span concentration on both the long and short paths. The pass criteria here are that both drift and noise should be less than 2% of the average span concentration. The drift proves the 'repeatability' claim of ±2%. The drift and span noise, along with linearity, pressure, and corrected thermal testing (x-shift and stretch correction), gives the accuracy claim of ±2%.

7.5.9 Pressure Testing

In this test the cell pressure is varied by ±10%, which for this analyser is at 180Torr and 220Torr. It is normally sufficient to test the variation of just one laser's main gas fit on the long and short paths at span concentration. To pass, the measured concentrations should stay within ±2% of the concentrations at 200Torr.

7.6 Linearization with linearity and accuracy testing

For this test, the gas of interest is supplied in concentration steps between zero and the span. The steps are at 0, 10, 30, 50, 70, 90 and 100% of span, and each step lasts for 5 minutes. An average is taken of each step. All the concentrations are then shifted so that at 0% the analyser does indeed measure zero. All the average concentrations are then stretched (or contracted) by a factor (span concentration / shifted measured concentration reading at 100% of span) so that the concentration at 100% of span does indeed coincide with the span concentration. The end points therefore agree with the supplied concentration. The shift and stretch coefficients are then entered into the analyser's configuration file for use in the field.

To pass the linearity test for normal use, a linear fit through the 7 average concentrations, of the shifted and stretched data, must have an R^2 coefficient ≥ 0.999. In addition, one of the accuracy criteria is taken from the deviation of the shifted and stretched measured concentration from the line of best fit. To pass, the deviation for each point must be within 2% of the supplied concentration.

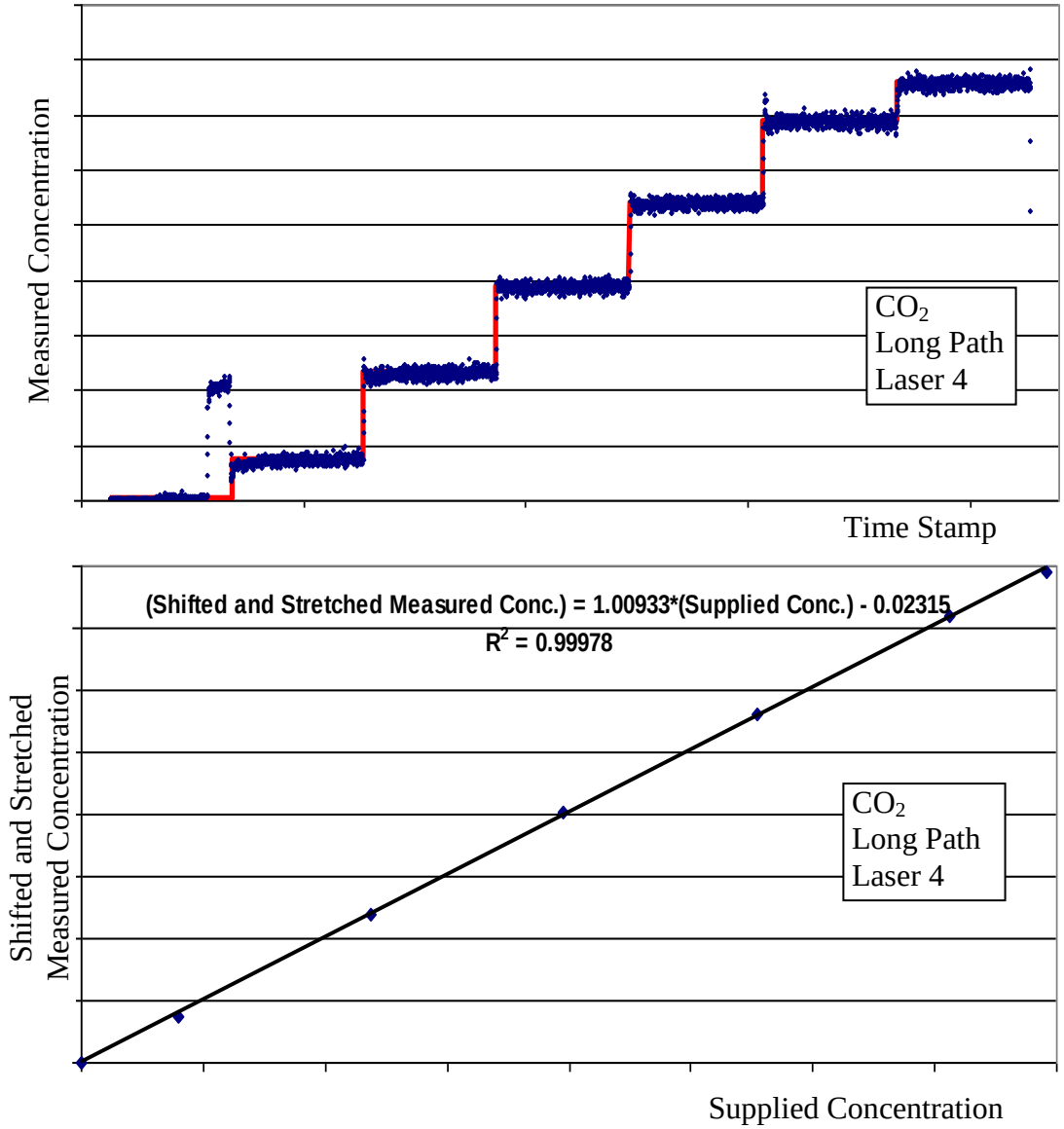


Figure 7-11: Linearity and accuracy data as in Figure 7-10. Here, however, the maximum absorption line in the fit window would reach 100% concentration in about 1/6th of the usual absorbance range. Small differences in absorbance profile are therefore used to calculate relatively large changes in concentration, leading to increased noise on the readings.

Figure 7-10 shows the typical form of the linearity and accuracy testing data. In the top plot, each concentration data point comes via Cascade Technologies' operating software, from Lorentzian fits to the gases' absorbance profile within the fit window (e.g. Figure 7-9, Section 7.5.7). Here, the change of absorbance between the absorbance limits of ≈0.004 and ≈3 is well matched to the concentration range. A change in concentration then leads to a reasonably significant change in absorbance profile, minimizing the noise on the concentration reading calculated from the integrated absorbance.

Each of the analyzer's four lasers is optimized to measure a particular target gas within its fitting window. Figure 7-11 and Figure 7-12 show possible compromises which may be made by measuring a gas in a fitting window optimized for a different gas. These compromises may be required, for instance, to continue the analyzer's mid-infrared measurements on a particular gas, if the primary laser for measuring it is swapped out to enable measurement of oxygen in the visible region.

Figure 7-11 shows concentration readings calculated from a Lorentzian fit to CO_2 absorbance profiles which are not as well matched to the concentration range expected, setup for the primary target gas NH_3. Here, instead of span concentration being at an absorbance of 3 (which it is for the NH_3 line), the maximum possible concentration of CO_2 (100%) does not reach an absorbance of 3 at all. The 100% concentration corresponds to only $1/6^{th}$ of the usual absorbance range. A small change in absorbance profile is therefore mapped to relatively large differences in concentration, leading to increased noise in the concentration readings. Figure 7-12 shows a more extreme version of this effect, where a small N_2O line strength means that 100% concentration is reached at a small fraction of the normal absorbance range, in a fit window optimized for SO_2.

Whether either of the last two gas fits, by Figure 7-11 and Figure 7-12, are useful will depend on the application - in particular, whether an accurate instantaneous concentration reading is needed rather than an average.

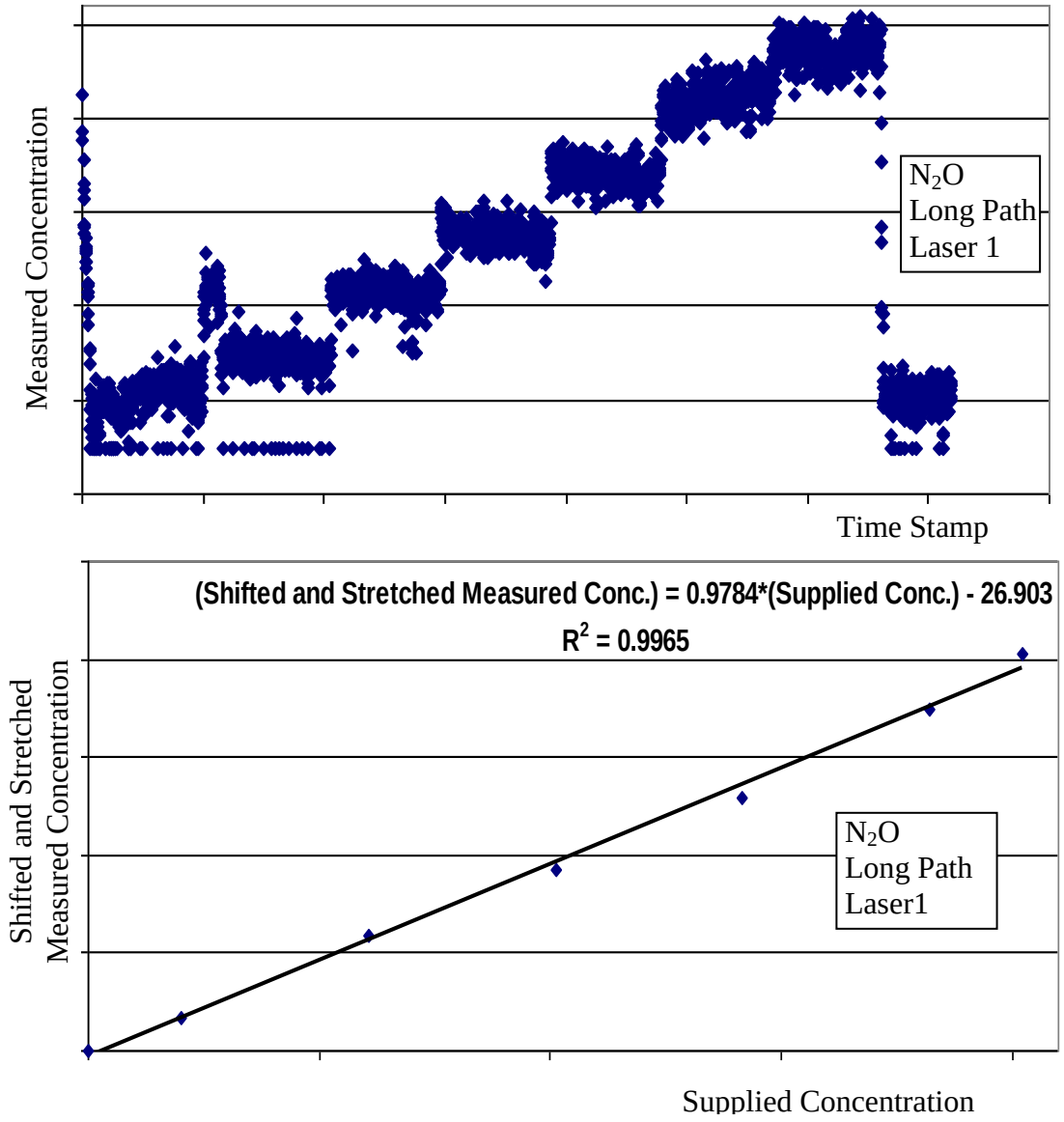


Figure 7-12: Linearity and accuracy data, as in Figure 7-10 and Figure 7-11. Here, however, the concentration data is taken from a Lorentzian absorbance fit which would reach 100% concentration in a small fraction of the usual absorbance range. Very small differences in absorbance profile are therefore used to calculate relatively large changes in concentration, leading to greatly increased noise on the readings.

7.6.1 Reverse- and re-engineering the linearity test spreadsheet

As shown in Figure 7-10, the linearity test is usually performed from low to high concentration. This minimizes the time needed for the concentration on the concentration steps to settle – particularly for sticky gases such as ammonia (NH_3), which would take considerable time to pump down working from high to low concentration.

The company spreadsheet, which automates the test from the raw concentration data, therefore also expects low to high concentration. This presents an inconvenience experimentally. The gas mixer struggles to stabilize the 10% of span mix from below, but not from above. Therefore, after the zero step, 30% of span is applied for one minute before taking the 5 minute 10% readings, as seen in Figure 7-10. This means that this data must be located and deleted from the company spreadsheet to function correctly. When large numbers of tests need to be processed this becomes more and more inefficient.

The detailed working of the VBA (Microsoft 'Visual Basic for Applications') code which ran the linearity spreadsheet had passed out of company memory. There were no comments attached to the program, and the purpose of the program's variables was not immediately obvious. It was therefore worthwhile to decipher the workings of the program and to re-write it for greater efficiency.

The top plots of Figure 7-10 and Figure 7-11 show some of the increased functionality. Where the original macro needed the strict pattern of 7 strictly increasing levels of concentration data, the re-written macro uses the same step detection method to simply ignore the 1 minute at 30% of span. Additionally, the new macro calculates standard deviations for each step. It also updates chart scales,

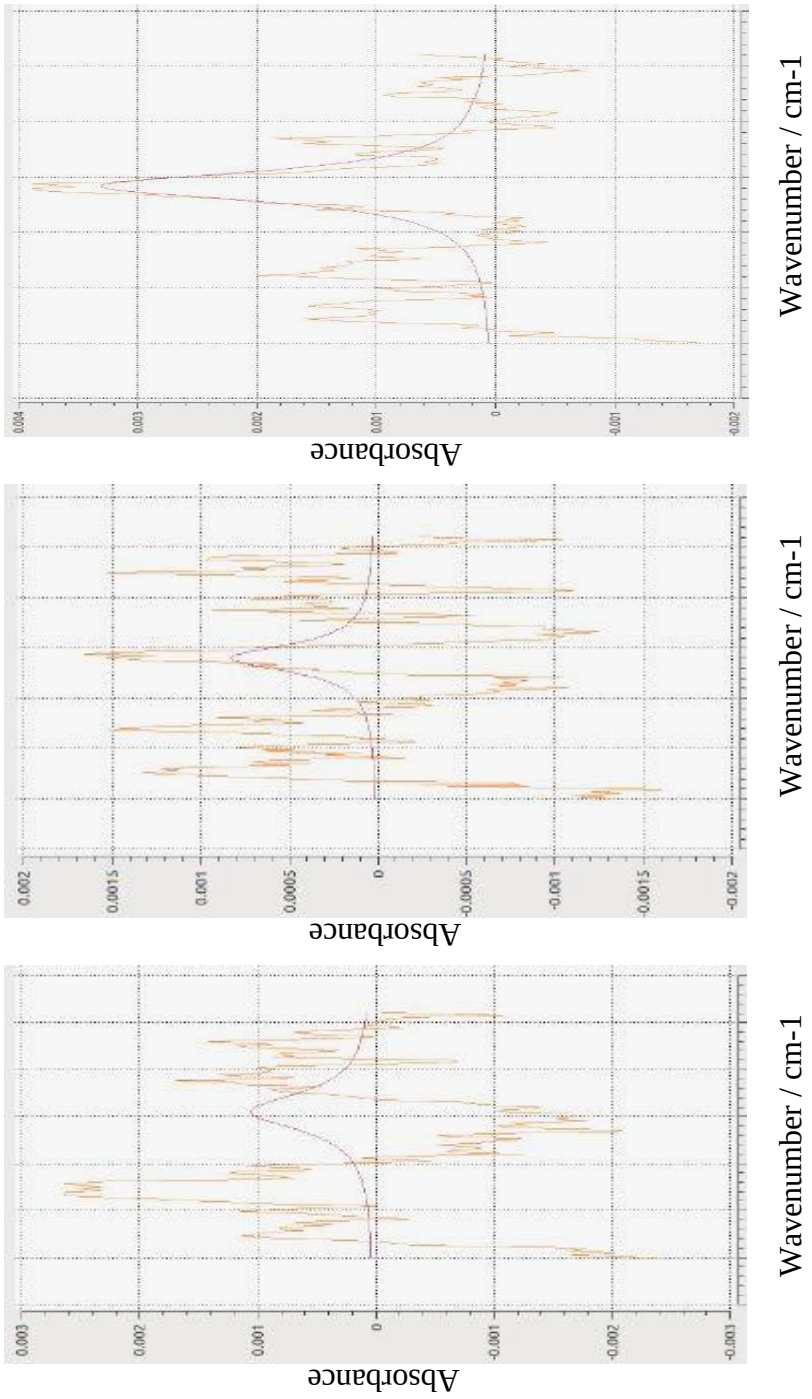


Figure 7-13: Absorbance noise related to LoD considerations. In the bottom plot, random noise rides on etalon fringes. The middle plot shows the beginnings of a definite absorbance line which is relatively stable compared to the noise. Taking the peak of the absorbance line leads to a definite, although noisy, concentration step above zero, and so could be used as a LoD concentration. The top plot at higher N_2O concentration gives a larger and less noisy concentration step above zero, and is still below the generally used absorbance noise of ≈0.004 absorbance units. (Note that the absorbance scales are expanding from plot to plot).

trendlines and units according to the selection of % or ppm. More detail of the workings of the original and modified macros is given in Appendix B.

7.6.2 Meaning of, and Testing for, the Limit of Detection

The purpose of this test is to demonstrate that each gas fit can indeed measure the claimed Limit of Detection (LoD). Here, the long path LoD is significant as it sets the analyser's LoD. With overlap between the short and long path concentrations, the short path LoD is less important.

Even for the long path, in a high concentration environment, it is often the case that the actual LoD is not particularly important. In these cases, claiming a LoD of 10% of span would be adequate and easily proved from the definite step change seen during the linearity test from zero concentration (e.g. Figure 7-10). A more focused consideration takes for the LoD the concentration which makes the peak of the fitted absorption line equal to the known maximum peak absorbance noise of the system of ~0.004 absorbance units. This maximum allows the LoD for any well aligned analyser to be guaranteed. However, the equivalent mixing ratio (for ~0.004 absorbance units) is often significantly increased before testing – making the rounded up LoD easier to demonstrate. In this work, however, the equivalent mixing ratio was tested, without rounding, to provide more detail. The actually achieved minimum mixing ratio for each gas on the long path was also found. This value can be influenced by the particular noise pattern near the fitted absorbance line. Nevertheless, in general, the generic and measured LoD's were found to be very similar. These mixing ratios are given in Table 7-2 on pA246ff.

The general form of noise in the CT3000 system is random noise riding on larger and more slowly varying unwanted etalon fringing. This can be seen in the orange traces of Figure 7-13. The purple trace is the Lorentzian fit to an absorbance line. In the bottom plot the software still tries to line fit, even though there is no line present at zero concentration. A minimum fit concentration needs to be assumed to prevent the iterations of the fitting program getting stuck in a loop near zero concentration.

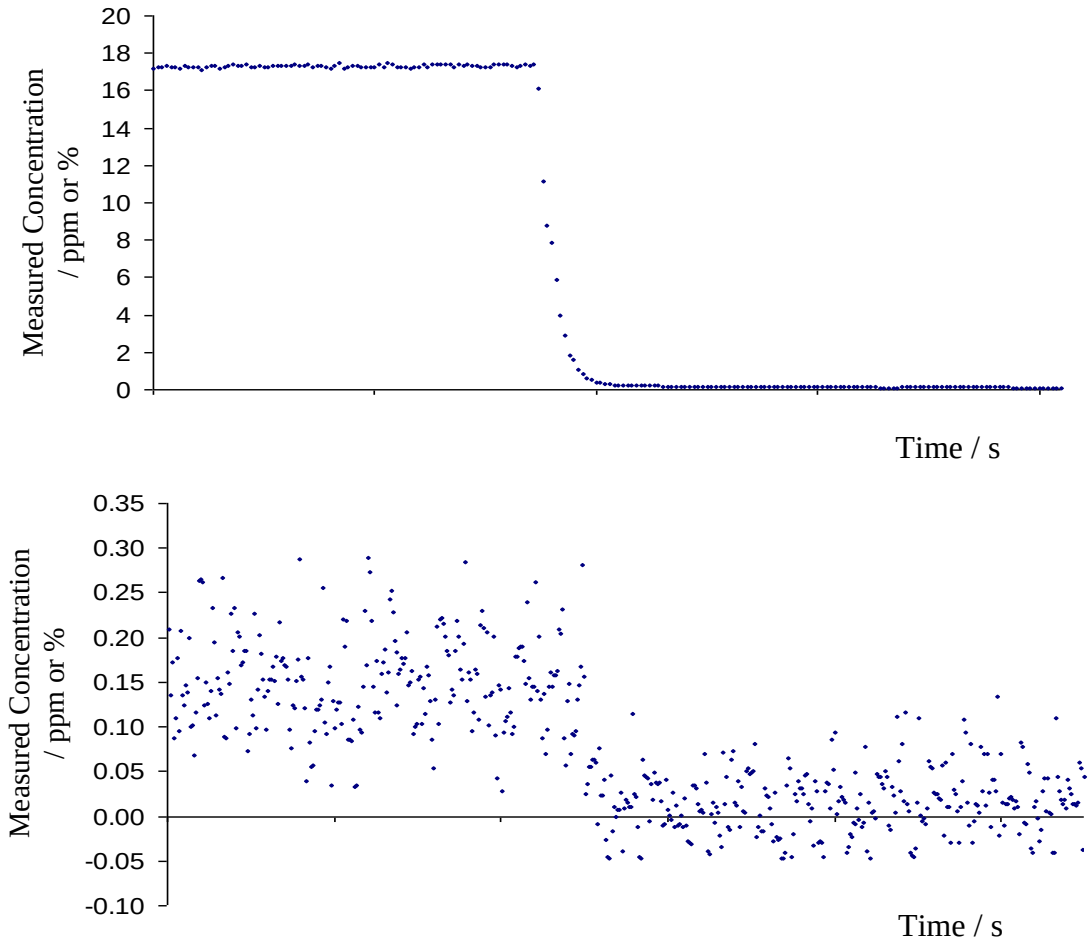


Figure 7-14: **Two general forms of measured concentration data, derived from fits to absorbance spectra. Both show data passing the LoD test. That is, both concentrations are reasonably near the supplied concentration, and both LoD average concentrations exceed 3σ of the zero noise level. The spread of concentration points around a level is determined by how much the absorbance noise causes the Lorentzian absorbance fit to change.**

A pass criterion for the LoD test, whatever definition is chosen, is that the step from zero to the LoD concentration should be greater than 3σ of the concentration noise. (The factor of 3, instead of the more usual 2, accounts for variations in international standards). In addition, the step change should be within a reasonable percentage of the supplied concentration – reasonableness being determined by the intended application and how important the LoD is.

In recording a meaningful LoD for each gas fit - smaller than the maximum absorbance noise of 0.004 - it was necessary to observe the detail of the absorbance spectra. The etalon fringing makes the LoD criterion used in Chapters 4 and 5 less useful here. There, the LoD was taken as the absorbance peak value which equalled 2σ of the (random) absorbance noise and gave a SNR of 1. Here instead, the measured LoD was taken as the concentration value when the fitted line first became unambiguous through the noise and was relatively stable to shifts in the etalon fringes.

The middle plot of Figure 7-13 shows such a case, although the spread of resulting concentration values from fitting fluctuations will be higher than data obtained with moderately greater concentration (as shown in the top plot).

Working at these low LoD's, considering the concentration plots alone would lead to inaccuracy. A concentration step change that meets the pass criteria might be caused by shifting absorbance noise which shifts the fitting routine, or a shift in interferent gas concentration may cause a shift in the absorbance amplitude without being caused by the target gas of interest – if not accounted for in cross interference testing.

Two types of concentration data which were found to pass the current LoD test are shown in Figure 7-14. In the top plot the concentration step change is much larger than 3σ of the concentration noise floor. It was found that there were two cases for this type of data. The data could be due to the LoD being set at an unnecessarily high level. But it could be that the line strength of the fit line is weak enough that it takes

a relatively large concentration for the line to become discernible on the absorbance plot at 0.004 absorbance units (or at whatever LoD absorbance level has been chosen). The bottom plot shows a case where the LoD concentration is much closer to 3σ of the concentration noise floor. Again the reason for this can be two-fold. The data set may be a false positive due to shifts in the absorbance noise. But it may be that the fitted linestrength is strong enough that it takes a relatively low concentration for the line to be discernible above the noise floor of the absorbance plot.

It is also worth noting that the LoD is also used as an upper limit pass / fail criterion in other tests, so that the chosen LoD definition will affect the outcome of these other tests. This is the case for the zero concentration noise test, zero average concentration 24 hour drift test, and cross interference testing. For instance, if an interferent causes a false target gas reading by indicating a concentration above the target gas' LoD, a higher (less sensitive) LoD could be claimed - if the revised LoD concentration level is still useful – and hence not breach the cross interference test.

7.6.3 Cross-Interference Testing

The likely gases which may cause a false concentration reading (≥LoD) in the various target gas readings need to be identified. The gases present in the measurement environment may be specified by a customer, or likely gases in more general environments may be anticipated and avoided with specific windows, or by including the interferent gas in the fitting. However, relatively large concentrations of interferent may still give a false reading on a target gas of interest.

Normally, the cross interference tests are just performed on a pass / fail basis since a particular commercial environment is in mind. The interferent at short path span concentration is entered into the Herriott cell. This is larger than or equal to the highest concentration of interferent that the analyser is expected to encounter. The interferent line shows up in the long and short path absorbance plots for that wavenumber fit window.

A target gas false positive is looked for primarily on the long path. The long path is where the target gases' absorbance spectra will be most affected - since an absorbance line reaching span concentration on the short path will substantially raise the baseline with associated noise riding on it when measured on the long path. In general, the short path absorbance plot is less affected because the shorter path results in less absorbance for the same concentration, and thus the interferent line may have less impact in its fit window, depending on where it falls in relation to the target gas lines. Also, since the short path LoD is a less critical parameter (Section 7.6.2), its LoD can be defined higher to avoid false positives.

Since the ultimate application of this analyser was not known, a fuller examination of the cross interferences of each gas against the other measured gases was undertaken. To more fully test the analyser's capability and compromises, cross-interferent failure was specified by a set of up to eight numbers in the following detail, working from low to high concentration of cross interferent in each of (1) to (4b), and high to low concentration in (5):

1/2aM3(2b)/4a(4b)//5

where:

1) The lowest concentration of cross-interferent which breaches the LoD for the target gas.

2) The highest false concentration of target gas on the long path (2a). This indicates how bad the test failure is in relation to the target gas' long path span concentration, in brackets (2b). If 2a is a small proportion of 2b, it may be satisfactory to simply increase the LoD claimed to negate the test failure.

3) 'M' label if this interferent causes the target gas concentration to reach its long path span concentration - that is, the maximum concentration reading

accurate within the analysers' long path specification. Exceeding this concentration tends to loop the iterations of the fitting program until a significantly lower concentration ensues. Following the 'M' label is the lowest value of cross interferent to 'max out' the analyzer (3a).

4) The false concentration reading of the target gas, on the short path, at cross interferent span concentration (if applicable) (4a). The short path target gas span is given in brackets (4b). The comparison of 4a to 4b indicates how serious the short path failure is.

5) The highest false target gas concentration reading, on the long path - working from high to low concentration of cross interferent. This takes account of asymmetry in fitting behaviour. Sometimes, from high to low concentration, the iterations of the fitting program are more inclined to fit noise as a line if a line (interferent or target gas) has initially been present.

To make it worthwhile going further, a specific gas environment would be needed. The effect of the expected range of cross interferent gases on the expected range of the target gas could then be analysed.

7.6.4 Water Cross Interference Testing

Water vapour is usually measured on a laser not present in the current analyser. Water concentration could be measured on this analyser's lasers 1 and 2, but with increased noise. Modifications to the typical fit windows needed to do this on these lasers were noted. Cascade Technologies does not carry equipment to supply known water concentrations. Instead, a fixed concentration of water vapour is pushed through the cell with nitrogen gas from a bubbler (Grant Sub Aqua 12 Plus). The bubbler's temperature is set to 20°C below the Herriott cell operating temperature to avoid condensation within the cell.

The analyser is checked for reading a constant water concentration when the bubbler has become stable. Any false readings of non-present target gas concentrations are noted.

7.6.5 Speed Optimization of Analyser Processing

If an analyser needs to be intensely optimized for speed, only the gas fits and particular lines absolutely necessary would be loaded into the analyser's configuration file. This minimizes the line fitting that the software needs to do. The settings used to produce a concentration reading can also all be juggled for a balance between accuracy and processing time. These settings include the number of averages of absorbance spectra to take before Lorentzian fitting, the number of fitting iterations, the maximum time of the fit, and the required convergence of concentration values from successive fitting iterations. The time between taking pressure and / or temperature readings can be increased to leave more processing time for updating the gas concentration readings from the fits.

For this analyser, it was sufficient to create separate configuration files for the usual gases measured (Table 7-1, pA228) and the whole expanded set (Table 7-2, pA246ff). Having set up the whole expanded set, it becomes a simple matter to choose the normal or broadened fitting windows, and gases to fit can be removed or added back in quite easily to suit a future application.

7.7 Overall Results

The full quantitative results were recorded in Excel spreadsheets based on Cascade Technologies' templates - but modified to suit the larger number of gas fits and fitting windows. Table 7-2, at the end of this chapter, presents zero and span concentration noise and 24 hour drift test results quantitatively. It can be seen that the noise and drift is often significantly better than the claimed ±2%. The linearity R^2 coefficients are also given.

Cross interference results are given qualitatively (to protect company fitting windows). In many cases a minor breach of the limit of detection on a non-present target gas occurs with a high concentration of interferent. In these cases, a pass can be obtained in the presence of cross interferent by increasing the recorded limit of detection - to what is still a small percentage of the target gas span on that optical path. Cross interference data on CO interference was less detailed than for other gases, because high concentration cylinders also contained other interferents.

None of the spectral fitting windows are ideal for measuring water vapour. The maximum concentrations are reached at low long path absorbances, making for higher concentration noise, particularly on laser 3. Two other gas fits, on lasers 1 and 2, were found to be too erratic and noisy for inclusion - these were N_2O on laser 1's long path in the usual fitting window, and CO_2 on laser 4's short path in the usual fitting window. An attempt to measure an outlying CO concentration from just the baseline inside the experimentally fitted window gave very inaccurate results.

7.8 Conclusions

A Cascade Technologies' CT3000 series gas analyser was successfully constructed and tested. The choices of standard spectral fitting windows - for two of the four incorporated pulsed-QCL lasers - were expanded to allow the possibility of more gas fits. This information might be used to compensate for the swapping out of lasers for other purposes – such as making room for near infrared oxygen measuring diode lasers in the future.

A standard fit window, used with laser 1, was expanded by 25%. This allowed the measurement of N_2O on the long path, with no major loss in the ability to measure SO_2 and CH_4 on the long and short path. The N_2O reading is, however, considerably more noisy than usual (compare Figure 7-10 and Figure 7-12, Section 7.6) – but may be useful in environments where only an average reading is required. The N_2O fit is vulnerable to cross interference from the tail of a NH_3 line at significant NH_3

concentrations. The noise and span for CH_4 on laser 1 is very similar to that on the usual laser for its measurement - Laser 3. There is the possibility of taking a reasonable H_2O measurement in the broadened fit window - but at moderately higher noise than the laser usually used for this task.

Laser 2 is primarily for measuring CO_2. Extending the standard fit window by 50% allows CO to be fitted. Large concentrations of CO will cause a cross interference problem with the broad window CO_2 reading. Small spans of N_2O can also be measured on the long path with standard levels of noise. The N_2O reading is fairly insensitive to the presence of CO_2 and moderate concentrations of CO. However, the presence of N_2O has a major influence on the long path CO_2 and CO readings, while the short path remains insensitive. Expanding the standard window by 83% allows H_2O to be measured at relatively higher noise than using the usual laser for the task.

Laser 3 is used for up to relatively high concentration N_2O readings. It can also be used for up to high concentration CH_4. When the CH_4 reaches particularly high concentrations, it causes cross interference problems with the long path N_2O reading - the short path being insensitive to this cross interference.

Laser 4 is usually dedicated to NH_3, although the long path can also sensibly measure high concentrations of CO_2 (comparable to the short path CO_2 readings on laser 2). The associated noise was found to be moderately higher than normal (Figure 7-11, Section 7.6), due to a relatively small absorbance range covering the whole concentration span.

Two contributions to process efficiency were made during this analyser work. Firstly, a diagnostic test sheet was constructed to isolate a problem analyser component. Secondly, the linearity test sheet was deciphered and re-written to automate processing of the experimental linearity data seen in this chapter.

7.9 Suggestions for further work

By the end of the work in this chapter, the CT3000 series gas analyser was ready for industrial use. It could also be used for the detection of gases at atmospheric abundance. However, atmospheric abundances are only a small fraction of the spans in an industrial environment. Therefore, separate span factors were recorded during linearity testing over a smaller concentration range to provide enhanced accuracy. If the analyser was to be dedicated to atmospheric measurements, it would be worthwhile choosing absorption lines with greater line strength. This would better match a line's absorbance range to the concentration range, giving reduced concentration noise.

For instance, CO_2 is an abundant gas in many industrial environments. Therefore, a line of weak line strength is chosen so that the full span of encountered concentrations does not 'concentration saturate' the analyzer, giving deviations from the Beer-Lambert Law (Section 7.5.3). For atmospheric measurements (≈400ppm CO_2) it would be worthwhile using a higher line strength line, obtainable with the same laser, or with another laser.

The rewritten macro for the linearity testing spreadsheet provides greater efficiency by ignoring the experimentally expedient - but numerically unwanted - first concentration step. It does this by using the method of the original program - to identify the end of the first step from an undershoot of concentration caused by the rotary pump and dilution system when the latter changes concentration. However, there are tests on other systems where such undershoots do not occur, and the number of concentration steps is more variable. To make the macro applicable to these tests it would need a complete revamp. One approach would be to de-emphasize edge behaviour by basing a program on step length and average concentration.

Table 7-2: Summary of the test spreadsheet results for noise, linearity and cross interference on the long and short optical paths of the CT3000 series analyser - for the particular four lasers and spectral fitting windows used.

Upper Repeatability Limit (from span conc. drift): ±2%; Upper Accuracy Limit (from linearity, span conc. noise, pressure testing, and thermal test after correction): ±2%; Linearity: R^2 >0.999 – unless otherwise stated in the notes

Wherever a minor x-int failure is recorded for a target gas, the LoD on that optical path can be increased by a small % of the target gas' span to obtain a pass in the presence of the interferent

b1 → standard fit window for laser 1 broadened by 25%

b2 → standard fit window for laser 2 broadened by 50%

eb2 → standard fit window for laser 2 'extra' broadened by 83%

b3 → the broader of two standard fit windows for laser 3 used, to include CH_4 as well as N_2O (fit window is 154% wider than the narrower version)

conc. => concentration; LP => long path; SP => short path;

LoD => generic concentration limits of detection are given for a generic noise level. Specification sheet values are often rounded up to guarantee the LoD can be met for a given sensor. The concentration long path LoD actually achieved for this sensor (500 averages) is given in brackets.

x-int => cross interference. Cross interference values are given in the format laid out in Section 7.6.3, in ppm unless otherwise stated.

Table 7-2 continued:

	Gas	Notes
Long Path		
Laser 1	SO_2	Span conc. noise 0.61%, span conc. drift 1.03% R^2 =0.99960 for linear fit, R^2 >0.9999 for quadratic fit LoD (ppm): 0.2 (0.2) Minor x-int failure to H_2O Major x-int failure with CH_4: 0/210(150)/180(5,500)
	SO_2 b1	Span conc. noise 0.56%, span conc. drift 0.30% R^2 =0.99879 for linear fit, R^2 >0.9999 for quadratic fit LoD (ppm): 0.2 (0.4) Major x-int failure with CH_4: 0/120(150)/370(5,500)//120
	CH_4	Span conc. noise 0.42%, span conc. drift 0.68% R^2 =0.99977 for linear fit, R^2 >0.9999 for quadratic fit LoD (ppm): 0.3 (0.3) Minor x-int failure from small conc. SO_2: 22/1.2(240)/28(8,600)//1.2
	CH_4 b1	Span conc. noise 0.42%, span conc. drift 0.35% R^2 =0.99970 for linear fit, R^2 >0.9999 for quadratic fit LoD (ppm): 0.3 (0.3) Minor x-int failure from small conc. SO_2 x-int with SO_2: 5/6(240)/95(8,600)//6
	N_2O b1	Span noise 4.65%, span drift 20%, useful for average conc. reading, but high noise on instantaneous reading R^2 =0.9965 for linear fit, R^2 >0.999 for cubic fit LoD (ppm): 1,200 (800) Major x-int failure, but only to large conc. of NH_3: 1,400/1,950(1,230)//1,950* x-int to large conc. CH_4: 5,100/5,460M1,900(2,024)
	H_2O b1	

Table 7-2 continued:

Laser 2	CO_2	Span conc. noise 0.70%, span conc. drift 0.75% R^2 =0.99955 for linear fit, R^2 >0.9999 for quadratic fit LoD (ppm): 21 (15) Major x-int failure from small conc. N_2O: 40/33,100M1,000(17,500)/1,800(650,000)//33,100 x-int pass for small conc. CO; possibility of minor x-int failure from large conc. CO
	CO_2 b2	Span conc. noise 0.70%, span conc. drift 1.15% R^2 =0.99990 for linear fit, R^2 >0.9999 for quadratic fit LoD (ppm): 21 (60) Major x-int failure from small conc. N_2O: 7/32,969M170(17,500)/1,915(650,000)//32,969 Likelihood of major x-int failure to large conc. CO

	CO b2	Span conc. noise 0.67%, span conc. drift 0.16%
		R^2 =0.99911 for linear fit, R^2 >0.9999 for quadratic fit
		LoD (ppm): 0.5 (0.55)
		Minor x-int failure from small conc. CO_2:
		40/6(465)/77(16,500)//20
		Major x-int failure from small conc. N_2O:
		1/725M170(465)/42(16,500)//725
		Minor x-int failure, but only to large conc. of CH_4:
		5,050/2.56(465)/-1.55(16,500)

	N_2O	Span conc. noise 0.39%, span conc. drift 0.18% R^2 =0.99994 for linear fit, R^2 >0.9999 for quadratic fit LoD (ppm): 0.03 (0.01) Minor x-int failure, but only to large conc. of CH_4: 5,050/0.35(21)-0.95(790)//0.35 x-int pass for small conc. CO; minor x-int failure from large conc. CO

Table 7-2 continued:

	N_2O b2	Span conc. noise 0.40%, span conc. drift 0.47%
		R^2 =0.99993 for linear fit, R^2 >0.9999 for quadratic fit
		LoD (ppm): 0.03 (0.01)
		Minor x-int failure from small conc. SO_2:
		0/0.04(22)/0.9(790)//0.06
		Minor x-int failure from moderate conc. CO_2:
		9,900/0.03(22)/-0.3(790)//0.25
		Minor x-int failure from small conc NH_3:
		0/0.03(22)/0.54(790)//0.03
		Minor x-int failure, but only to large conc. of CH_4:
		5,050*/0.2(21)/-0.65(790)//0.2
		Likelihood of moderate x-int failure from large conc. CO

	H_2O eb2	If eb2 window is used, H_2O at bubbler conc. will cause minor x-int failures on the laser 2 readings for CO and CO_2
Laser 3	N_2O b3	Span conc. noise 0.47%, span conc. drift 0.13% R^2 =0.99929 for linear fit, R^2 >0.9999 for quadratic fit LoD (ppm): 0.1 (0.2) Moderate x-int failure, but only to large conc. of CH_4: 5,050/4.85(95)/18(3,400)//4.85
	CH_4 b3	Span conc. noise 0.43%, span conc. drift 0.05% R^2 =0.99916 for linear fit, R^2 >0.9999 for quadratic fit LoD (ppm): 0.3 (0.4)
Laser 4	NH_3	Span conc. noise 0.44%, span conc. drift 0.43% R^2 =0.99898 for linear fit, R^2 =1.0000 for cubic fit LoD (ppm): 0.05 (0.1) Minor x-int failure from small conc. CH_4: 350/0.48(41)/2(1,480)//0.48

Table 7-2 continued:

	CO_2	Span conc. noise 1.98%, span conc. drift 1.31%, moderately higher noise than usual (near limit of pass/fail criteria) R^2 =0.99978 for linear fit, R^2 =1.0000 for cubic fit LoD (%): 0.5 (0.5) Major x-int failure from small conc. NH_3: 4ppm/25%(100%)/-1.64%(100%)//30%

Short Path		
Laser 1	SO_2	Span conc. noise 0.98%, span conc. drift 1.65% R^2 =0.99982 for linear fit, R^2 >0.999 for quadratic fit LoD (ppm): 7
	SO_2 b1	Span conc. noise 0.45%, span conc. drift 0.38% R^2 =0.99948 for linear fit, R^2 =0.9999 for cubic fit LoD (ppm): 7
	CH_4	Span conc. noise 0.62%, span conc. drift 0.10% R^2 =0.99931 for linear fit, R^2 >0.9999 for quadratic fit LoD (ppm): 10 Minor x-int failure from small conc. SO_2
	CH_4 b1	Span conc. noise 0.57%, span conc. drift 0.19% R^2 =0.99967 for linear fit, R^2 >0.9999 for quadratic fit LoD (ppm): 10 Minor x-int failure from small conc. SO_2
Laser 2	CO_2	Span conc. noise 0.58%, span conc. drift 0.40% R^2 =0.99997 for linear fit, R^2 >0.9999 for quadratic fit LoD (ppm): 70,000 Minor x-int failure from span conc. N_2O

Table 7-2 continued:

	CO_2 b2	Span conc. noise 0.80%, span conc. drift 0.32% R^2 =0.99993 for linear fit, R^2 >0.9999 for quadratic fit LoD (ppm): 70,000 Minor x-int failure from span conc. N_2O
	CO b2	Span conc. noise 0.65%, span conc. drift 1.01% R^2 =0.99965 for linear fit, R^2 >0.9999 for quadratic fit LoD (ppm): 20 Minor x-int failure from span conc. CO_2 Minor x-int failure from span conc. N_2O
	N_2O	Span conc. noise 0.38%, span conc. drift 1.79% R^2 =0.99958 for linear fit, R^2 >0.9999 for quadratic fit LoD (ppm): 1
	N_2O b2	Span conc. noise 0.45%, span conc. drift 0.11% R^2 =0.98895 for linear fit, R^2 >0.999 for cubic fit LoD (ppm): 1 Minor x-int failure from span conc. SO_2 or NH_3 Possibility of minor x-int failure to high conc. CO
Laser 3	N_2O b3	Span conc. noise 0.53%, span conc. drift 0.06% R^2 =0.99987 for linear fit, R^2 >0.9999 for quadratic fit LoD (ppm): 4 Minor x-int failure at span conc. of CH_4
	CH_4 b3	Span conc. noise 0.53%, span conc. drift 0.08% R^2 =0.99953 for linear fit, R^2 >0.9999 for quadratic fit LoD (ppm): 10
Laser 4	NH_3	Span conc. noise 1.04%, span conc. drift 0.53% R^2 =0.99984 for linear fit, R^2 >0.9999 for quadratic fit LoD (ppm): 2

Table 7-2: Summary of the test spreadsheet results for noise, linearity and cross interference on the long and short optical paths of the CT3000 series analyser - for the particular four lasers and spectral fitting windows used.

Chapter 8

Summary and outlook

Sweep Integration Parameters (mid-infrared)	Sawtooth Current Ramp: 10kHz, 500mA – 680mA 7.8μm (1282.052cm^{-1}) cw-QCL, HgCdTe photovoltaic detector, 68m astigmatic Herriott cell
Absorbance	0.080 in 51ms (512 averages), possibility of achieving 0.020 in 0.77s (7680 averages)
WMS Parameters (mid-infrared) combined with 512 averages	Sawtooth Current Ramp: 5Hz, 450mA – 690mA Sinusoidal Modulation: 10kHz, 4.4mA amplitude Lock-in amplifier's RC time constant: 1ms 7.8μm (1282.052cm^{-1}) cw-QCL, HgCdTe photovoltaic detector, 68m astigmatic Herriott cell
Absorbance	0.017 in 102.4s (512 averages)
Sweep Integration Parameters (visible)	Triangular Current Ramp: 50Hz, 10mA amplitude centered on 100mA 759.7nm (13163.091cm^{-1}) prototype semiconductor diode laser, APD detector, 30m astigmatic Herriott cell (84cm short path)
Absorbance	0.023 in 2.56s (128 averages)
Sweep Integration (CT3000 intra-pulse gas analyzer)	~750ns Current Pulses, duty cycle ~3.75% (50kHz) various mid-infrared wavenumbers depending on component lasers and intended application up to four mid-infrared pulsed-QCL lasers, two HgCdTe photovoltaic detectors, 30m astigmatic Herriott cell (84cm short path)
Absorbance	0.004 in 10ms (500 averages), limited primarily by unwanted etalon fringing

Table 8-1: Absorbance Limits of Detection (LoD's) (noise floors) for the systems in this thesis.

8 Summary and outlook

Conclusions and suggestions for further work were laid out by chapter. Here, an overview of the thesis' main findings is given:

The sponsor company's current gas analysers use pulsed-QCL's in intra-pulse mode, with sweep integration as the sensitivity enhancing technique. A further sensitivity enhancement was sought by using a (continuous wave) cw-QCL with the wavelength modulation spectroscopy (WMS) technique. Methane was used as the primary test gas. Molecular oxygen was also potentially added to the list of measurable gases by testing a (wavenumber) prototype semiconductor diode laser using sweep integration. New analytic and processing software was written to aid in taking the work forward. This included Gaussian, Lorentzian, and Voigt lineshape modeling, and modeling of the 2f-WMS lineshape.

The work with the cw-QCL was limited by time and budget. A particular limitation appears to be the use of AC-coupling, where electronic noise in the signal chain was large and seemed to be the cause of an anomalous scaling of oscilloscope data. Regarding resolution, the slow sweep of the cw-QCL has the potential to increase resolution compared to the pulsed laser system. However, methane and nitrous oxide Gaussian linewidths were found to be significantly broader than their calculated Doppler widths. By eliminating other possibilities, this broadening is also attributed to noise in the electronic system.

Table 8-1 lists the systems used in this thesis along with their LoD's (limits of detection), known as the noise floors. These are defined here as twice the standard deviation of the baseline noise. The continuous number of averages of the cw-QCL data was limited to 512 by the oscilloscope. A modest sensitivity improvement was found by using WMS over sweep integration, 0.017 absorbance units in 102.4s against 0.080 in 51ms. The inability to take continuous concentration readings meant that the optimum number of averages for each technique could not be found.

However, Allan variance software was written to quickly find this optimum, should continuous recording be added in the future.

For the sponsor company, these LoD's were of an intermediate order of magnitude between their proceed / abandon criteria. The LoD's are comparable to those of the pulsed systems at an earlier stage of development. The typical pulsed LoD is 0.004 absorbance units (limited by unintentional etalon fringing). Improvements were based around the design of specialist electronics, increased temperature and pressure stability of the gas cell, and the optical specifications and design. There seems no reason why further improvements to the cw-QCL approach would not be forthcoming – particularly as most problems seem to stem from the single issue of the AC-coupled signal chain. Future development would likely move from the bread-board set-up, used in this thesis, towards the typical components used by the sponsor company – such as used in their CT3000 system.

However, integrating a cw-QCL into an existing system, whilst also running pulsed-QCL's, would be challenging. The CT3000 system, for instance, contains two AC-coupled detectors, one each for a short and a long optical path. If a DC-detector were used for one path, this would limit the pulsed lasers to the other path. If two AC-coupled detectors were retained, the different electrical and timing requirements of the cw-QCL would need to be incorporated into the control and processing software and electronics.

The sweep integration work with the (wavenumber) prototype semiconductor diode laser showed that single mode and multi-mode oxygen spectroscopy was possible. Since the laser contained a distributed feedback grating, multi-mode behaviour was unintended. This indicates to the manufacturer / supplier that the growth of this layer was imperfect. This particular laser could be used for oxygen spectroscopy but would be unsuitable in a commercial sensor because of its instability, particularly to changes in substrate temperature. The device evidently has difficulty reaching thermal equilibrium, suggesting that phonon transport is hampered within the

structure. Without more detail of the device design, it is not possible to speculate further.

Integrating a diode laser and its APD detector into an existing pulsed system would raise similar issues to incorporating a cw-QCL. In addition, the diode laser wavelength is in the near-infrared and visible, so the existing mid-infrared optics are not optimal. For measuring large concentrations of oxygen, the work in this thesis indicates that the cell optics are not an issue. However, a separate beamsplitter is likely to be needed.

In order to retain as much pulsed-QCL performance as possible - if a pulsed laser was swapped out - a CT3000 dual path gas analyzer was constructed and thoroughly analysed, using a configuration of four typical pulsed-QCL's. The standard fitting windows were expanded in two cases, and the maximum number of gases measured by the four lasers were considered. Detailed information was taken on the minimum mixing ratios obtained and the effects of cross interferences. A diagnostic test sheet was constructed to investigate a unique problem of cell contamination. The linearity testing software was deconstructed and modified for future use.

Appendices

Appendix A : Diagnostic test checklist for the CT3000 series gas analyser

Purpose: To investigate why the CT3000 series laser pulses had dropped overnight by 50% on the short path and by 90% on the long path, subsequent to successful thermal testing (Section 7.5.2). It was necessary to convince Cascade Technologies to allow mirror cleaning and ultimate replacement of the astigmatic Herriott cell. The checklist was orientated to this task, but is easily orientated to more general circumstances.

Procedures:

Unsuccessful Diagnostics tried:

Optical:

1) Prism replaced, all three lasers re-aligned individually and together (2 with new optics), and strengths observed on the Pyrocam III camera.
2) Top optics not unglued from their bases, but all realigned.
3) Attenuation removed from the long path output mirror.
4) One laser temporarily replaced in the sensor to see if pulses were a different magnitude – i.e. to test if something had degraded the original laser strengths over time.
5) Beam splitter rotated.
6) Detector placed in front of the long path window to the cell – signal strength was enough to saturate the detector-digitizer combination for each pulse at 1V.

7) Tried unsuccessfully to maximize another short path beam at the expense of the relatively strong one in case the beams were not ‘tight’ around the visible laser.

Electrical:

1) Both detectors temporarily replaced.
2) All three SMA cables temporarily replaced.
3) All connections and electronic boards bypassed by using the alignment-calibration setup in Product Development.
4) Cell heater run from a 240V supply instead of a 110V supply.

Software:

1) Cell temperature tested with a remote infrared detector to see if it was correct, so as to rule out the software adjusting cell temperature for its fit.
2) Sensor run off 4 different computers, therefore using generic dual path configuration files as well as its own.

Mechanical:

1) One cell base screw was found to be slightly loose. This was tightened.

Optical / Electrical Swaps:

1) Combiner block and its lasers were placed in a sensor known to be working. The signals increased dramatically even though that sensor was attenuated.
2) The faulty sensor’s detectors were tried in the working sensor and they were found to work correctly.

3) Upon replacement of the parts back in the faulty sensor, the VIGO ribbon cables and all 3 SMA cables were temporarily replaced.

Given the above 3 tests and test 3 in 'Electrical' (above) which bypassed all electronics and left the lasers, cell and detectors, it would seem that the cell is at fault.

Successful Diagnostics:

1) Mirrors were cleaned (this is prone to scratch the mirrors so is not usually done). The signals improved markedly.
2) A swab on a cotton bud was taken of the inside of the cell. It was very dirty, being coated with a black sooty type residue – apparently this had not occurred before.
3) Another swab was taken after the cell had been baked for 48hrs. The residue was still there although reduced to about a third. The continued presence of contamination indicated that the cell had probably not missed its initial decontamination baking.

Results:

Cleaning the mirrors increased the detected pulse strengths substantially, although this would have likely been a temporary solution as the inside of the cell was very dirty with a sooty type substance. A sample of the residue was kept in case Cascade Technologies had been interested in its identification.

Outcome:

The most likely explanation for the overnight pulse drops is that thermal testing evaporated some of this black sooty type residue into the body of the cell, and overnight this settled on the mirrors.

Since the mirrors only looked ‘normally’ dusty, a thin layer of this residue is likely to be highly scattering and / or absorbing in the mid-infrared. Interestingly, the short path, highest wavenumber, QCL pulse was much less affected than the others – indicating a wavelength dependency to the scattering and/or absorbance. The cell in the sensor was replaced (from the matt finish type to one of the gloss finish types). Since the problem cell still had significant residue after 48hrs baking, it is likely this residue came in from the metalworkers as general dirt or as incorrectly finished powder coat.

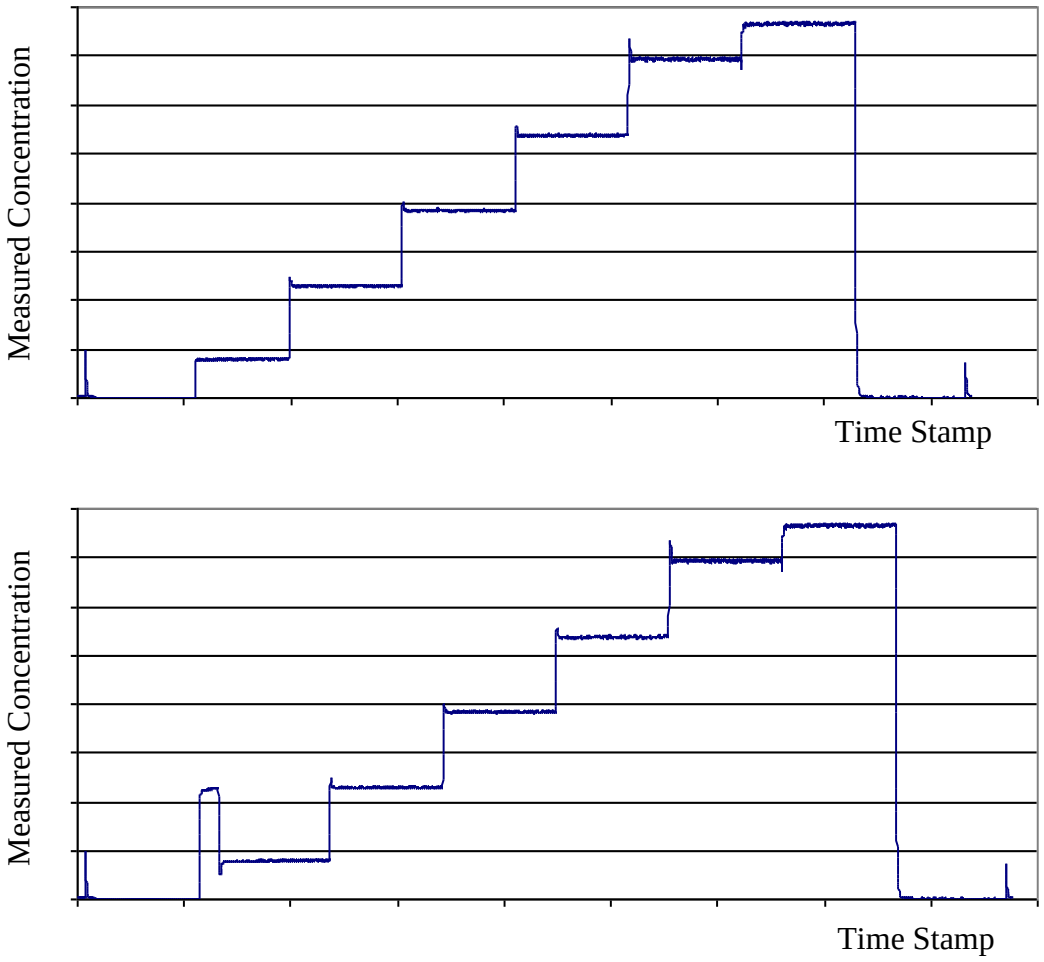


Figure B-1: **Top plot: the seven concentration steps expected by the original VBA linearity testing macro.**
Bottom plot: the form of concentration data taken experimentally.

Appendix B : Reverse- and re-engineering the linearity testing spreadsheet

B.1 Introduction

Linearity testing (Section 7.6) involves taking 5 minutes of data at each of the following percentages of maximum (span) concentration – 0, 10, 30, 50, 70, 90, 100, 0%. However, it is experimentally easier for the Environics gas mixer to approach the low 10% concentration from above. For this reason, an additional 1 minute 30% of span step is included between the 0 and 10% steps. (Conducting the test backwards from 100% concentration downwards is undesirable because of the additional time needed to remove the higher concentrations).

The Cascade Technologies' linearity testing spreadsheet was run from a Microsoft VBA (Visual Basic for Applications) macro. Although the spreadsheet was extensively used, its detailed working had passed out of company memory. The macro contained no comments and the purpose of the program's variables was not immediately obvious. It was therefore worthwhile to decipher the method of the program.

And also, in general, the macro would only function correctly with a specific form of concentration data in seven steps, shown in the top plot of Figure B-1. The bottom plot shows the general form of recorded data. Prior to macro modification, the first 30% step had to be located by hand in the data set and deleted – leading to inefficiency when a large number of sets needed processed.

The top two plots of Figure B-2 show how the macro handles a data set in which the first 1 minute at 30% of span has not been pre-deleted. That step is counted as one of the seven expected, so that the 100% of span step is missed, and the resulting linearity plot based on the step averages is not meaningful.

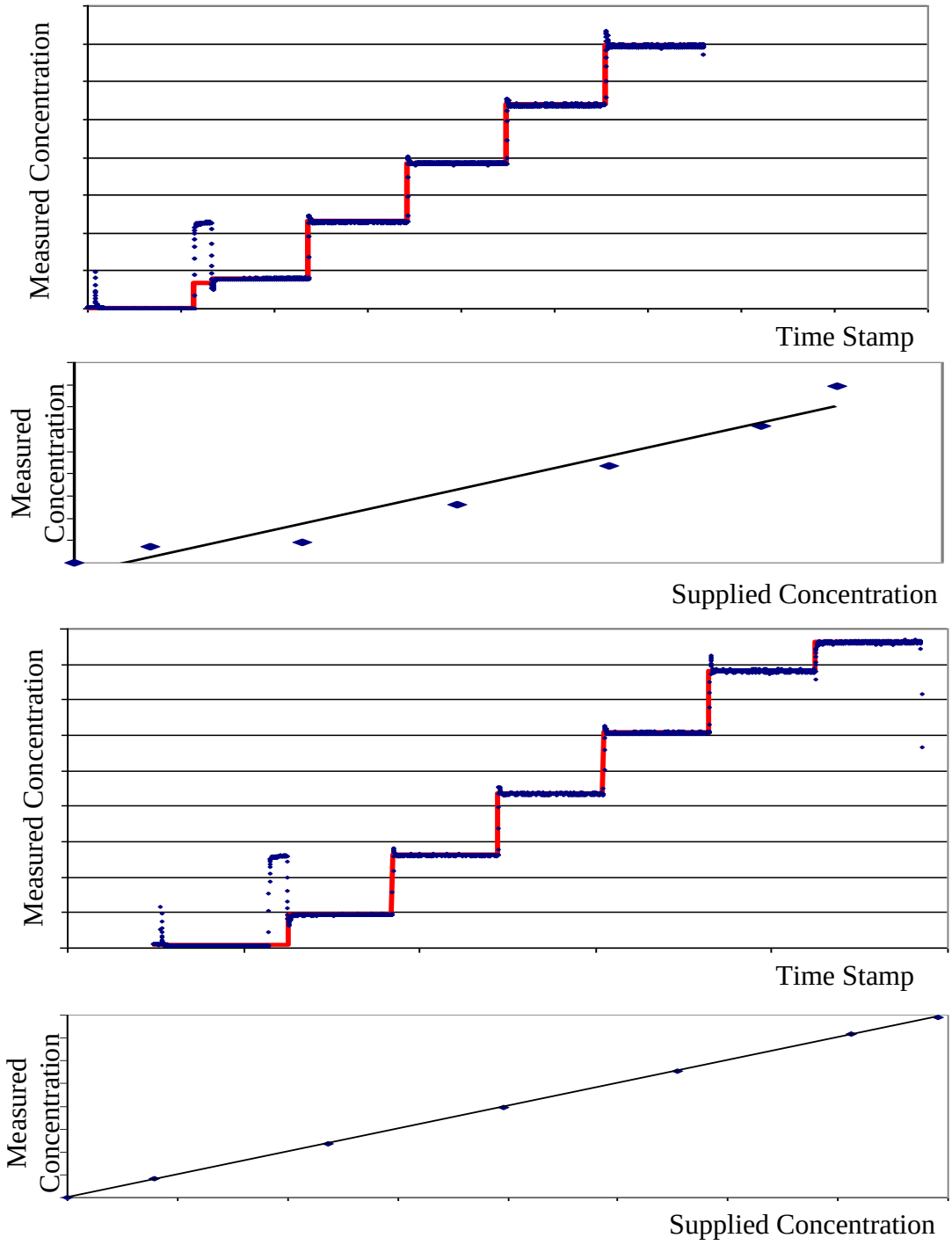


Figure B-2: **Top two plots: the way the original macro handles data sets with the 1 minute at 30% of span not deleted. That step is counted as one of the seven expected, so that the 100% of span step is missed, and the resulting linearity plot based on the step averages is not meaningful.**
Bottom two plots: the macro after modification ignores the 1 minute at 30% of span data as if the data set was the top plot of Figure B-1 resulting in concentration averages being taken from the correct seven steps and producing a correct linearity test plot.

The bottom two plots of Figure B-2 show the modified macro. This macro ignores the 1 minute at 30% of span data as if the data set was the top plot of Figure B-1. This results in concentration averages being taken from the correct seven steps and the production of a correct linearity test plot.

B.2 How the original macro works

The macro notes the time value and, more importantly, the row number (A4) of the first spreadsheet reading for use as the start of the zero 'step'. This starting point will be used towards calculating the average concentration of this zero step and for the start of the red plot seen in Figure B-2. The plateaus of the red plot mark the average concentrations and allow a visual interpretation of how well the macro is working.

The macro essentially compares each concentration value to the one before it, looking for a jump in value which marks the transition between the current concentration step and the next. To eliminate jumps in concentration value due to noise, the current concentration value is compared to the previous concentration value *plus* a user defined 'threshold'. (This threshold is normally entered as 0.006 – that is, 0.6% of the maximum supplied concentration).

The first time a concentration value is greater than its 'previous concentration plus threshold', the program holds the value of this 'previous concentration plus threshold' fixed as a 'test value' and compares subsequent concentrations to this test value. The program also provisionally notes the time value and row number of this test value datum as marking the possible end of the concentration step it was on.

A counter ('flag' in the program) begins and increments in value for each concentration above the fixed test value. The user defines how many consecutive points (generally 50) must be above the test value for the concentration rise to be permanently regarded as a concentration step (this is 'pthres' points in the macro). Should a concentration fall below the test value before this number of points has

been reached, the program resets the counter (flag) to zero, forgets the time value of the test value, and begins comparing the current concentration to the previous concentration (plus threshold) again on a rolling basis. In this way, the short lived spikes in Figure B-1 and Figure B-2 are not recorded as concentration steps.

If the consecutive number of concentration values above the test value reaches the user defined number of points, the concentration step is regarded as genuine. The time value and row number of the test value (now pthres points back) is now permanently recorded as marking the end of a concentration step and the time value and row number of the subsequent concentration is deemed to mark the start of the next concentration step. These time values are entered in a spreadsheet column and used to generate the time axis of the red plot. The row numbers of these time values are used in calculating the average of the steps, as explained below.

The program now continues through the data, repeating its method by comparing each new concentration value to the 'previous concentration plus threshold' - looking for the next step.

This continues until 'pthres' points of step 7 have been reached. The algorithm will then look for a substantial drop to record the row number and time value at the end of step 7. If this drop is not encountered, the macro will stop at the end of the data set. Perhaps unintentionally, a satisfactory stop is also caused by what the macro interprets as an 8^{th} step, as this also causes the end of step 7 to be recorded - its row number being placed in 'pcounter'.

The 'duration' is then calculated as the average number of rows per step. Since each expected step lasts for 5 minutes, the 'duration' should be a good measure of rows per step. This allows streamlining the program instead of using step specific lengths for each iteration. (Note that the average recorded for the included first 30% of span step of the top plot of Figure B-2 is low because the 'duration' for it is too large. Hence, the macro has included a portion of the zero step in its average).

The algorithm now calculates the average concentration of each step by using the row number of the beginning and end of the relevant step, whilst using 'duration' to ignore the first 35% of data points and the last 5% of data points per step – in order to avoid the transitioning concentrations between steps. This also excludes the spikes sometimes seen at the start of a step, as in Figure B-1 and Figure B-2, and prevents them interfering with the step's average concentration value. For the 10% step, the macro also presents 3 times that step's standard deviation as the 'lower detection limit' (LDL) which is used with some systems.

For each step, the average concentration is paired with the times for the beginning and end of the step to generate the red average plot. These average concentration values appear as the 'Measured average before correction' in the 'front of house' spreadsheet. From here the spreadsheet itself carries through the linearity test of Section 7.6. Namely, the concentration averages are shifted and stretched so that the end points agree with the supplied concentrations. Then Excel's SLOPE and INTERCEPT functions generate a line of best fit, and the deviations of the seven average concentrations from this line are calculated. For a pass, each deviation should be <2% of the relevant supplied concentration.

B.3 Purpose of the Macro's Variables

The purposes of the macro's main variables are described below, and a graphical representation of some of them is given in Figure B-3.

complete → a marker/flag to indicate when the program has recorded the 7^{th} concentration step in full and thus when it has finished its task.

prev → starts with the first concentration value, then becomes the previous concentration value to the current value on a rolling basis, unless the program is holding it fixed at a particular previous value while looking for a concentration step.

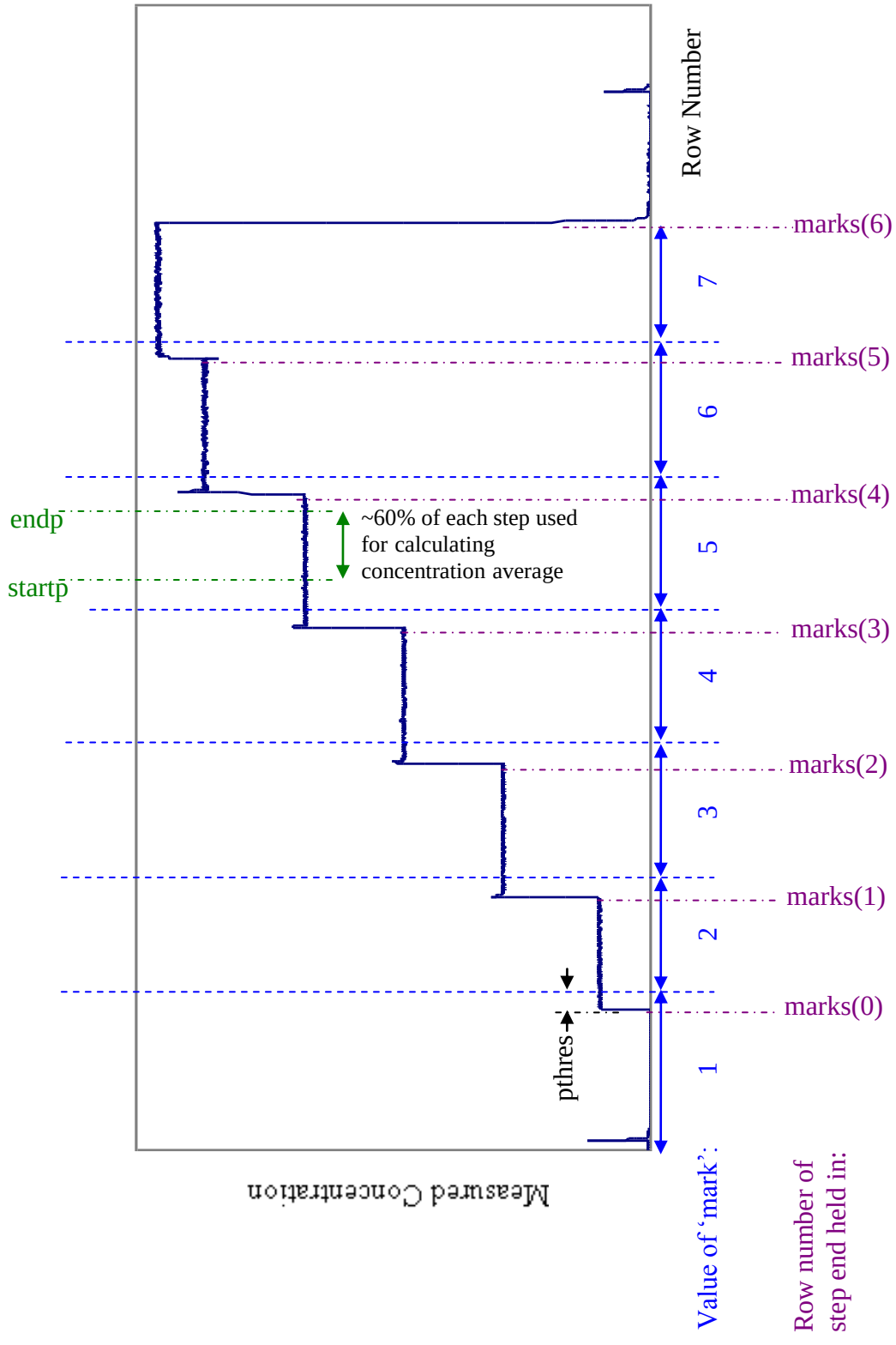


Figure B-3: Some of the linearity macro's variables presented graphically.

counter → counts the number of iterations of the program, and is also the row number of the current concentration value datum in the data set.

pcounter → used in plotting the red trace in the chart, recording the time values of the start and finish of the concentration steps in column O and the corresponding average concentrations in column P.

flag → a counter which starts incrementing when the current value is greater than the previous value plus a user defined threshold. At this point the previous value plus threshold is fixed as a test value with which to compare subsequent concentrations. If a concentration fails to be larger than this test value before flag (essentially the number of consecutive concentrations above the test value) becomes equal to pthres, the flag is reset and no step is permanently recorded. If pthres is reached, the flag is reset and a step is permanently recorded.

pthres → the number of consecutive points above the fixed test value ('previous' plus threshold) that there must be for a concentration step to be permanently recorded.

mark → increments each time flag = pthres. Essentially it records which concentration step the program has reached, and is used in the array 'marks' to record the row number of the end of a step.

marks(mark - 1) → an array marks(0) to marks(6) which records the row number of the end of a step, and is used in 'duration' to calculate the average number of rows per step. This in turn is used to define a portion of the step for the calculation of its concentration average via 'startp' and 'endp'.

duration → = [marks(6) – marks(0)]/6, and is the average number of rows per concentration step.

values(i) → declared but unused array.

B.4 Modifications to the linearity macro and spreadsheet

The major modification to the original macro begins with a shortened step detection loop - controlled by a separate threshold for step detection. This loop only looks for the 1 minute at 30% of span step. It takes advantage of the concentration undershoot between this step and the following 10% of span step - in order to retain the step detection method of the original macro. From this initial loop, the row numbers and time values for the start and end of the first 30% step are found and recorded in the new two element array 'pexclude'. For information, the excluded data points are coloured in the spreadsheet, and the main detection loop starts.

Each iteration, and moving through the datum row numbers, is incremented in the original macro's 'counter' variable. In the new main loop, the start of each iteration begins by checking if the start of the excluded rows has been reached. If it has, the counter variable is jumped over the excluded rows, and the detection loop continues from the next datum. In this way the first 30% of span step is ignored. Later, the excluded points are also kept out of the calculation for the average and standard deviation of the zero step.

In addition, the modified macro displays 2σ values for all of the concentration steps, and also a 3σ value for the zero step. Upon running the modified macro, the value of a '% or ppm' control on the spreadsheet changes the chart scales, trendlines, and titles and axis labels appropriately.

Appendix C : Change in refractive index and Free Spectral Range with wavenumber for a germanium etalon

As part of the data processing in this thesis, the time axis of a spectrum is often converted to relative wavenumber. To do this, the simplifying assumption that the etalon fringes are equally spaced in frequency is made. That is, the Free Spectral Range (FSR) of the etalon is a constant. The program in this appendix was primarily to test the accuracy of this assumption – which turns out to be excellent.

The program uses the Sellmeier equation, with coefficients taken from literature (Frey et al., 2006), to calculate the change of germanium refractive index. The temperature of the etalon was assumed to be at room temperature for simplicity (the cw-QCL beam is ≤ 100mW). The wavenumber of the beam was modeled as varying over $1cm^{-1}$ near $1283.5cm^{-1}$. Once the change in refractive index is calculated from the Sellmeier equation, this change can then be used to calculate the change in FSR. The range of Frey et al's data fitting, from which their coefficients were derived, was $20K \le T \le 300K$ and $1.9\mu m \le \lambda \le 5.5\mu m$. This is short of the $7.8\ \mu m$ of the cw-QCL used in Chapter 4. Nevertheless, these coefficients are assumed to be adequate to enable a ball-park estimate of FSR constancy to be made.

The Sellmeier equation is:

$$n^2(\lambda,T)-1=\sum_{i=1}^{m}\frac{S_i(T)\lambda^2}{\lambda^2-\lambda_i^2(T)} \tag{C-1}$$

where m is an integer. Often m is taken as 3, and the coefficients $S_i(T)$and $\lambda_i(t)$ are taken as cubic functions of temperature, giving:

$$n^2(\lambda,T)-1=\sum_{i=1}^{3}\frac{S_i(T)\lambda^2}{\lambda^2-\lambda_i^2(T)} \tag{C-2}$$

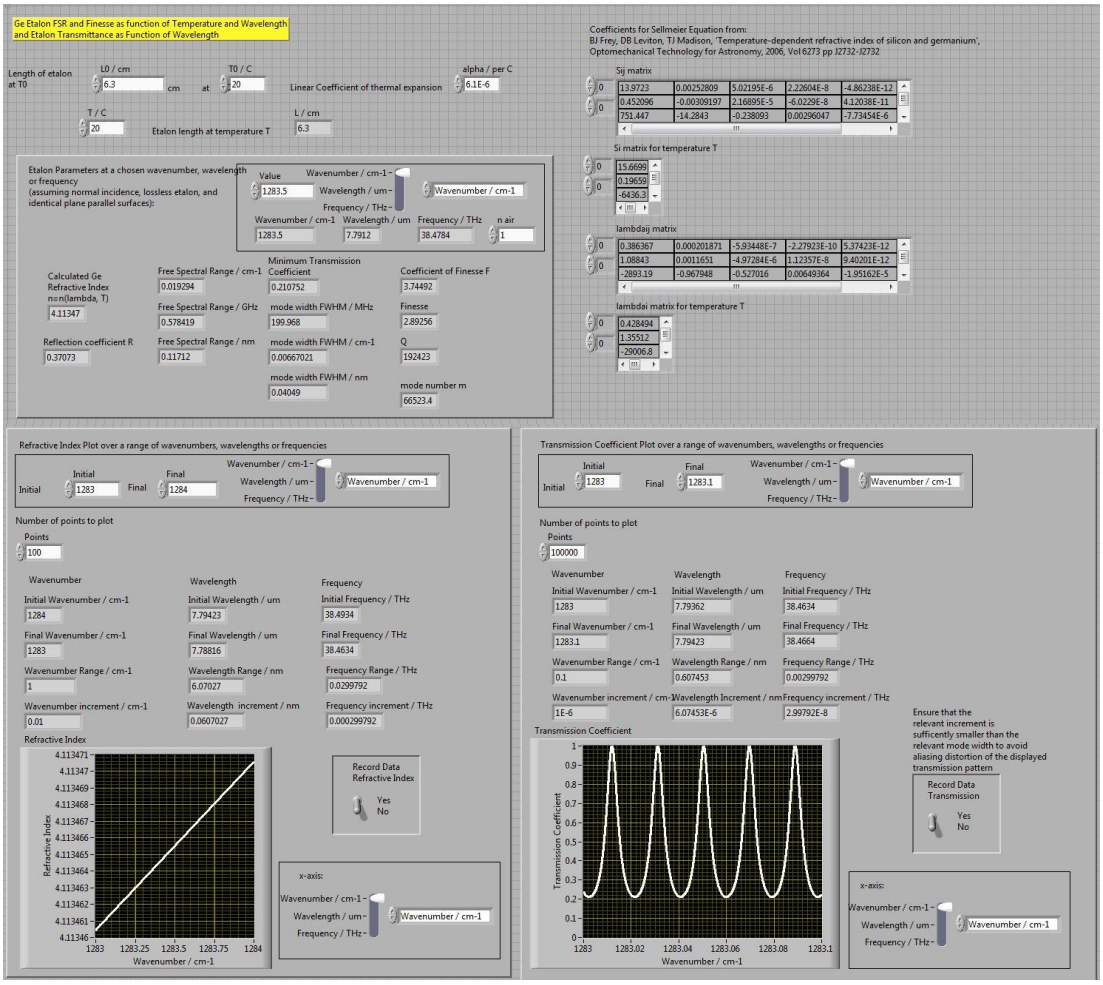


Figure C-1: **In the top panel, the Ge etalon refractive index and FSR is calculated for a given temperature and wavenumber - via use of the Sellmeier equation. Other parameters for characterizing the etalon's performance are also calculated.**

In the bottom left panel, the refractive index is plotted over a user defined range of wavenumbers at a given temperature.

In the bottom right panel, the etalon's expected transmittance is plotted, using the Airy function for Fabry-Perot fringes.

where $S_i(T) = \sum_{j=0}^{4} S_{ij} T^j$ and $\lambda_i(T) = \sum_{j=0}^{4} \lambda_{ij} T^j$ (C-3)

The coefficients from literature, S_{ij} and λ_{ij}, are stored in the program, and they are used with matrix multiplication to calculate the three $S_i(T)$ and three $\lambda_i(T)$ coefficients - for a user defined temperature. The equation is always carried through in wavelength according to Equation C-2. However, in the program, data may be entered and read off, and the abscissa chosen, in any of wavelength (μm), wavenumber (cm^{-1}), or frequency (THz).

Figure C-1 shows a screenshot of the program. The top panel calculates the refractive index from the Sellmeier equation for a fixed, user defined wavelength. Other parameters which characterize the (normal incidence Fabry-Perot) etalon are also calculated and displayed. These parameters can be found in most optics textbooks, such as that by Hecht (1987). Once the refractive index is obtained, and the length of the etalon entered, the FSR follows immediately from:

$$FSR_{cm^{-1}} = \frac{1}{2nL_{cm}} \quad \text{(C-4)}$$

The bottom left panel of the screenshot shows the variation of the refractive index with wavenumber, at a fixed temperature. The variation of the FSR across these wavenumbers is then calculated from Equation C-4 (with L =6.3cm, and n varying near 4.113). The change in refractive index was $\approx 1\times10^{-5}cm^{-1}$ across the $1cm^{-1}$ shown – giving a FSR change of $\approx 5\times10^{-8}cm^{-1}$. So:

$$\frac{\text{Change in FSR}}{FSR} \approx \frac{5\times10^{-8}cm^{-1}}{0.0195cm^{-1}} \approx \frac{3\times10^{-6}cm^{-1}}{1cm^{-1}} = \frac{3cm^{-1}}{10^6 cm^{-1}} \quad \text{(C-5)}$$

Therefore, the wavenumber FSR is constant to 3 parts per million over a $1cm^{-1}$ wavelength range at room temperature near a wavenumber of $1283.5cm^{-1}$. (The $^{12}CH_4$ $\nu(4)$ $_{A2\ 1 \leftarrow A1\ 1}$ line of Chapter 4 is at $1283.458791cm^{-1}$ (Rothman et al., 2009)).

The bottom right panel of the screen shot displays the predicted etalon transmittance (Airy function), neglecting changes in laser power during a scan, as a function of wavenumber over the first $0.1cm^{-1}$ of the $1cm^{-1}$ range above.

Addendum: EngD Masters level coursework transcript

EPSRC
Pioneering research and skills

Centre for Doctoral Training in Applied Photonics

15 May 2015

Robert Campbell

Registration No.: 091512352

The assessment results for the Engineering Doctorate in Optics and Photonics Technologies taught courses are as follows:

Alison Low

Centre Administrator

Module	*Title*	*Grade*	*Credits*
B21SL	Photonics Experimental Laboratories	A	15
B21DN	Displays and Nonlinear Optics	A	10
B21LP	Laser Physics	A	20
B21SP	Photonics Applications	A	15
B21LR	Literature Review	A	15
B21OI	Optical Metrology	A	5
H17EC	Economics	A	20
H17MK	Marketing	B	20
H17PR	Project Management	A	20
H17AC	Accounting	A	20
H17FI	Finance	A	20

Total number of modules: 11
Scottish Credit and Qualifications Framework (SCQF) Credit points: 180

Grade Banding
Grade A = above 70%
Grade B = 60-69%
Grade C = 50-59%
Grade D = 40-49%
Grade E = 30-40%

Centre for Doctoral Training in Applied Photonics
Administration ▸ Heriot-Watt University ▸ School of Engineering and Physical Sciences ▸ Edinburgh Campus ▸ EH14 4AS ▸ United Kingdom
Telephone: +44 (0)131 451 3792 ▸ www.cdtphotonics.hw.ac.uk ▸ Email: cdtphotonics@hw.ac.uk

HERIOT WATT UNIVERSITY · University of St Andrews · University of Strathclyde Glasgow · University of Glasgow · University of Dundee

References

References

ALLEN, L. & EBERLY, J. H. 2003. *Optical resonance and two-level atoms,* New York, Dover Publications Ltd. ISBN: 978-0-486-65533-8.

ARITA, Y., STEVENS, R. & EWART, P. 2008. Multi-mode absorption spectroscopy of oxygen for measurement of concentration, temperature and pressure. *Applied Physics B-Lasers and Optics,* 90(2)**,** 205-211.

ARITA, Y. & EWART, P. 2008. Multi-mode absorption spectroscopy. *Optics Communications,* 281(9)**,** 2561-2566.

ARMEN, B. G. 2008. *Phase sensitive detection: The lock-in amplifier* [Online]. Knoxville: Department of Physics and Astronomy, The University of Tennessee. Available: http://www.phys.utk.edu/labs/modphys/lock-in%20amplifier%20experiment.pdf [Accessed: 08/04/2015].

ARNDT, R. 1965. Analytical line shapes for lorentzian signals broadened by modulation. *Journal of Applied Physics,* 36(8)**,** 2522-2524.

ATKINS, P. & FRIEDMAN, R. 2011. *Molecular quantum mechanics (5th ed.),* Oxford, Oxford University Press. ISBN: 978-0-19-954142-3.

BAIN, J. R. P., JOHNSTONE, W., RUXTON, K., STEWART, G., LENGDEN, M. & DUFFIN, K. 2011. Recovery of absolute gas absorption line shapes using tunable diode laser spectroscopy with wavelength modulation-part 2: Experimental investigation. *Journal of Lightwave Technology,* 29(7)**,** 987-996.

BANWELL, C. N. 1966. Fundamentals of molecular spectroscopy. London: McGraw-Hill Publishing Company Ltd.

BECK, M., HOFSTETTER, D., AELLEN, T., FAIST, J., OESTERLE, U., ILEGEMS, M., GINI, E. & MELCHIOR, H. 2002. Continuous wave operation of a mid-infrared semiconductor laser at room temperature. *Science,* 295(5553)**,** 301-305.

BENTHAM INSTRUMENTS LTD. *225 lock-in amplifier - bentham* [Online]. Bentham Instruments Ltd. Available: http://www.bentham.co.uk/pdf/F225.pdf [Accessed: 08/04/2015].

BERDEN, G. & ENGELN, R. (eds.) 2009. *Cavity ring-down spectroscopy: Techniques and applications,* Chichester: John Wiley & Sons, Ltd. ISBN: 978-1-4051-7688-0.

BERNATH, P. F. 2005. *Spectra of atoms and molecules (2nd ed.),* Oxford, Oxford University Press. ISBN: 0-19-517759-2.

BOGUE, R. 2011a. The gas sensing industry: Look at the 1 percent. *Sensor Review,* 31(1)**,** 5-5.

BOGUE, R. 2011b. Terrorism and military actions pose the ultimate challenge to gas sensing. *Sensor Review,* 31(1)**,** 6-12.

BOGUE, R. 2015a. Detecting gases with light: A review of optical gas sensor technologies. *Sensor Review,* 35(2)**,** 133-140.

BOGUE, R. 2015b. Detecting explosives and chemical weapons: A review of recent developments. *Sensor Review,* 35(3)**,** 237-243.

BOHM, D. 1989. *Quantum Theory,* Mineola, New York, Dover Publications. ISBN: 0-486-65969-0.

BOMSE, D. S., STANTON, A. C. & SILVER, J. A. 1992. Frequency-modulation and wavelength modulation spectroscopies - comparison of experimental methods using a lead-salt diode-laser. *Applied Optics,* 31(6)**,** 718-731.

BUZUAYENE, M. 2008. *Rise time vs. Bandwidth and applications* [Online]. Germantown Pike, PA: Interference Technology. Available: http://www.interferencetechnology.com/rise-time-vs-bandwidth-and-applications/ [Accessed: 15/08/2016].

CASCADE TECHNOLOGIES LTD. 2014a. *Cascade Technologies Ltd website* [Online]. Stirling: Cascade Technologies Ltd. Available: http://www3.emersonprocess.com/Cascade-Technologies/ [Accessed: 17/02/2016].

CASCADE TECHNOLOGIES LTD. 2014b. *CT3000 OEM gas analyser* [Online]. Stirling: Cascade Technologies Ltd. Available: http://www.cascade-technologies.com/media/1026/ct3000-oem-gas-analyser.pdf [Accessed: 06/06/2015].

CASCADE TECHNOLOGIES LTD. 2014c. *History* [Online]. Stirling: Cascade Technologies Ltd. Available: http://www3.emersonprocess.com/Cascade-Technologies/our-company/history/ [Accessed: 14/02/2016].

CASCADE TECHNOLOGIES LTD. 2012. *CT3000 OEM MODULE operation manual,* Stirling, Cascade Technologies Ltd.

CASSIDY, D. T. & REID, J. 1982. High-sensitivity detection of trace gases using sweep integration and tunable diode lasers. *Appl Opt,* 21(14)**,** 2527-30.

CHAO, X., JEFFRIES, J. B. & HANSON, R. K. 2012. Wavelength-modulation-spectroscopy for real-time, in situ no detection in combustion gases with a 5.2 mu m quantum-cascade laser. *Applied Physics B-Lasers and Optics,* 106(4)**,** 987-997.

CHEESMAN, A., SMITH, J. A., ASHFOLD, M. N. R., LANGFORD, N., WRIGHT, S. & DUXBURY, G. 2006. Application of a quantum cascade laser for time-resolved, in situ probing of ch4/h-2 and c2h2/h-2 gas mixtures during microwave plasma enhanced chemical vapor deposition of diamond. *Journal of Physical Chemistry A,* 110(8)**,** 2821-2828.

DAYLIGHT SOLUTIONS. 2015. *Core technology: Quantum cascade lasers: Linewidth & tunability* [Online]. San Diego, CA: Daylight Solutions. Available: http://www.daylightsolutions.com/technology/qcl_technology.htm [Accessed: 13/08/2016].

DEMTRÖDER, W. 2008a. *Laser spectroscopy: Volume 1: Basic principles (4th ed.),* Berlin, Springer. ISBN: 978-3-540-73415-4.

DEMTRÖDER, W. 2008b. *Laser spectroscopy: Volume 2: Technology (4th ed.),* Berlin, Springer. ISBN: 978-3-540-74952-3.

DORRAN, D. 2012. *Discrete fourier transform explained* [Online]. Available: https://www.youtube.com/watch?v=B2iUDBZzBpY&list=FLgGp4PzlC2DpMlu3UwIUy5Q&index=4 [Accessed: 13/04/2015].

DUFFIN, K., MCGETTRICK, A. J., JOHNSTONE, W., STEWART, G. & MOODIE, D. G. 2007. Tunable diode-laser spectroscopy with wavelength modulation: A calibration-free approach to the recovery of absolute gas absorption line shapes. *Journal of Lightwave Technology,* 25(10)**,** 3114-3125.

DUXBURY, G., NORMAND, E. L., LANGFORD, N., MCCULLOCH, M. T. & WALKER, S. 2002. Highly sensitive detection of trace gases using pulsed quantum cascade lasers. *In:* FRIED, A. (ed.) *Diode lasers and applications in atmospheric sensing.* 4817, 14-21, ISBN: 0-8194-4585-1.

DUXBURY, G., LANGFORD, N., MCCULLOCH, M. T. & WRIGHT, S. 2005. Quantum cascade semiconductor infrared and far-infrared lasers: From trace gas sensing to non-linear optics. *Chemical Society Reviews,* 34(11)**,** 921-934.

DUXBURY, G., LANGFORD, N. & IEEE 2007. *Quantum cascade laser spectroscopy, trace gas analysis to non-linear optics*. ISBN: 978-1-4244-1237-2.

DUXBURY, G., LANGFORD, N. & HAY, K. 2008. Delayed rapid passage and transient gain signals generated using a chirped 8 mu m quantum cascade laser. *Journal of Modern Optics,* 55(19-20)**,** 3293-3303.

DUXBURY, G., LANGFORD, N., HAY, K. & TASINATO, N. 2009. Quantum cascade laser spectroscopy: Diagnostics to non-linear optics. *Journal of Modern Optics,* 56(18-19)**,** 2034-2048.

DUXBURY, G., TASINATO, N., HAY, K., WILSON, D. & LANGFORD, N. 2010. Collisional effects on quantum cascade laser induced molecular alignment. *In:* LEWIS, J. K. C. & PREDOICROSS, A. (eds.) *20th international conference on spectral line shapes.* 1290, 194-203, ISBN: 978-0-73540-845-6.

DUXBURY, G., HAY, K. G., LANGFORD, N., JOHNSON, M. P. & BLACK, J. D. 2011. Real-time diagnostics of a jet engine exhaust using an intra-pulse quantum cascade laser spectrometer. *Molecular Physics,* 109(17-18)**,** 2131-2142.

DUXBURY, G., KELLY, J. F., BLAKE, T. A. & LANGFORD, N. 2012a. Observation of infrared free-induction decay and optical nutation signals from nitrous oxide using a current modulated quantum cascade laser. *The Journal of Chemical Physics,* 136(17)**,** 174317.

DUXBURY, G., KELLY, J. F., BLAKE, T. A. & LANGFORD, N. 2012b. The ac stark effect in nitric oxide induced by rapidly swept continuous wave quantum cascade lasers. *Journal of Chemical Physics,* 136(17)**,** 174318.

DUXBURY, G., KELLY, J. F., BLAKE, T. A. & LANGFORD, N. 2012c. Sub-doppler spectra of infrared hyperfine transitions of nitric oxide using a pulse modulated quantum cascade laser: Rapid passage, free induction decay, and the ac stark effect. *Journal of Chemical Physics,* 136(17)**,** 174319.

DYROFF, C. 2008. *Tunable diode-laser absorption spectroscopy for trace-gas measurements with high sensitivity and low drift.* Doktor Ingenieurs, Universität Karlsruhe. Available: http://digbib.ubka.uni-karlsruhe.de/volltexte/documents/703149 [Accessed: 15/04/2015].

ENCYCLOPÆDIA BRITANNICA. 2015. *Air: Atmospheric gas* [Online]. London: Encyclopædia Britannica. [Accessed: 10/06/2015].

ERNST, R. R. 1966. Sensitivity enhancement in magnetic resonance. *In:* WAUGH, J. S. (ed.) *Advances in magnetic resonance.* New York: Academic Press, ISBN: 9781483254401.

ESAKI, L. & TSU, R. 1970. Superlattice and negative differential conductivity in semiconductors. *IBM Journal of Research and Development,* 14(1)**,** 61-65.

FAIST, J., CAPASSO, F., SIVCO, D. L., SIRTORI, C., HUTCHINSON, A. L. & CHO, A. Y. 1994. Quantum cascade laser. *Science,* 264(5158)**,** 553-556. Available: http://www.researchgate.net/profile/Jerome_Faist/publication/235237550_Quantum_Cascade_Laser/links/0912f50f8f3227ff3c000000.pdf [Accessed: 05/06/2015].

FAIST, J., CAPASSO, F., SIRTORI, C., SIVCO, D. L., HUTCHINSON, A. L. & CHO, A. Y. 1996a. Room temperature mid-infrared quantum cascade lasers. *Electronics Letters,* 32(6)**,** 560-561.

FAIST, J., CAPASSO, F., SIRTORI, C., SIVCO, D. L., BAILLARGEON, J. N., HUTCHINSON, A. L., CHU, S. N. G. & CHO, A. Y. 1996b. High power mid-infrared (lambda greater than or similar to-5 mu m) quantum cascade lasers operating above room temperature. *Applied Physics Letters,* 68(26)**,** 3680-3682.

FAIST, J., GMACHL, C., CAPASSO, F., SIRTORI, C., SIVCO, D. L., BAILLARGEON, J. N. & CHO, A. Y. 1997. Distributed feedback quantum cascade lasers. *Applied Physics Letters,* 70(20)**,** 2670-2672.

FOLTYNOWICZ, A., SCHMIDT, F. M., MA, W. & AXNER, O. 2008. Noise-immune cavity-enhanced optical heterodyne molecular spectroscopy: Current status and future potential. *Applied Physics B,* 92(3)**,** 313-326.

FRASER, D. A. 1990. *The physics of semiconductor devices (4th ed.),* Oxford, Oxford University Press. ISBN: 0-19-851866-8.

FREY, B. J., LEVITON, D. B. & MADISON, T. J. 2006. Temperature-dependent refractive index of silicon and germanium - art. No. 62732j. *In:* ATADETTEDGUL, E., ANTEBI, J. & LEMKE, D. (eds.) *Optomechanical technologies for astronomy, pts 1 and 2.* 6273, J2732-J2732, ISBN: 0-8194-6338-8, Available: http://arxiv.org/ftp/physics/papers/0606/0606168.pdf [Accessed: 22/05/2015].

GAGLIARDI, G. & LOOCK, H. (eds.) 2014. *Cavity-enhanced spectroscopy and sensing,* Berlin: Springer. ISBN: 978-3-642-40002-5.

GARCIA, M., NORMAND, E., STANLEY, C. R., IRONSIDE, C. N., FARMER, C. D., DUXBURY, G. & LANGFORD, N. 2003. An AlGaAs-GaAs quantum cascade laser operating with a thermoelectric cooler for spectroscopy of NH3. *Optics Communications,* 226(1-6)**,** 39-43.

GMACHL, C., CAPASSO, F., SIVCO, D. L. & CHO, A. Y. 2001. Recent progress in quantum cascade lasers and applications. *Reports on Progress in Physics,* 64(11)**,** 1533-1601.

GORSHUNOV, B. M., SHOTOV, A. P., ZASAVITSKY, II, KOLOSHNIKOV, V. G., KURITSYN, Y. A. & VEDENEEVA, G. V. 1979. Cw and pulsed pb1-x snx se diode-laser spectroscopy of (NH3)-n-14 and (NH3)-n-15 molecules. *Optics Communications,* 28(1)**,** 64-68.

GRIFFITHS, P. R. & de HASETH, J. A. 1986. *Fourier transform infrared spectrometry,* New York, John Wiley & Sons. ISBN: 0-471-09902-3.

HAMILTON, M. L., RITCHIE, G. A. D., ARITA, Y. & EWART, P. 2010. Multi-mode absorption spectroscopy, mumas, using wavelength modulation and cavity enhancement techniques. *Applied Physics B-Lasers and Optics,* 100(3)**,** 665-673.

HANCOCK, G., VAN HELDEN, J. H., PEVERALL, R., RITCHIE, G. A. D. & WALKER, R. J. 2009. Direct and wavelength modulation spectroscopy using a cw external cavity quantum cascade laser. *Applied Physics Letters,* 94(20)**,** 201110.

HAY, K. G., WRIGHT, S., DUXBURY, G. & LANGFORD, N. 2008. In-flight measurements of ambient methane, nitrous oxide and water using a quantum cascade laser based spectrometer. *Applied Physics B-Lasers and Optics,* 90(2)**,** 329-337.

HAY, K. G. 2010. *Gas sensing using quantum cascade lasers.* PhD, University of Strathclyde.

HECHT, E. 1987. *Optics (2nd ed.),* Reading, Massachusetts, Addison-Wesley Publishing Company. ISBN: 0-201-11611-1.

HENNINGSEN, J. & SIMONSEN, H. 2000. Quantitative wavelength-modulation spectroscopy without certified gas mixtures. *Applied Physics B-Lasers and Optics,* 70(4)**,** 627-633.

HERRIOTT, D., KOMPFNER, R. & KOGELNIK, H. 1964. Off-axis paths in spherical mirror interferometers. *Applied Optics,* 3(4)**,** 523-526.

HERRIOTT, D. R. & SCHULTE, H. J. 1965. Folded optical delay lines. *Applied Optics,* 4(8)**,** 883-889.

HOFSTETTER, D., BECK, M., AELLEN, T., FAIST, J., OESTERLE, U., ILEGEMS, M., GINI, E. & MELCHIOR, H. 2001. Continuous wave operation of a 9.3 mu m quantum cascade laser on a peltier cooler. *Applied Physics Letters,* 78(14)**,** 1964-1966.

HOLLAS, J. M. 2009. *Modern spectroscopy (4th ed.),* Chichester, John Wiley & Sons, Ltd. ISBN: 0-470-84416-7.

HOWIESON, I. F. 1997. *Near-infrared tunable diode laser absorption spectrometer for trace gas detection.* PhD, University of Strathclyde.

HOWIESON, I. F., NORMAND, E. & MCCULLOCH, M. T. 2005. Quantum-cascade lasers smell success. *Laser Focus World,* 41(3)**,** S3-S6. Available: http://www.laserfocusworld.com/articles/print/volume-41/issue-3/optoelectronics-world/quantum-cascade-lasers-smell-success.html [Accessed: 05/06/2015].

HOWIESON, I. F. 2011. Quantum-cascade lasers disrupt gas measurement technologies. *Laser Focus World,* 47(2)**,** 33-38. Available: http://www.laserfocusworld.com/articles/print/volume-47/issue-2/features/quantum-cascade-lasers-quantum-cascade-lasers-disrupt-gas-measurement-technologies.html [Accessed: 06/08/2016].

JAHNKE, J. A. 2000. *Continuous emission monitoring (2nd ed.)*, New York, John Wiley & Sons, Inc. ISBN: 978-0-471-29227-3.

JANIK, G. R., CARLISLE, C. B. & GALLAGHER, T. F. 1986. 2-tone frequency-modulation spectroscopy. *Journal of the Optical Society of America B-Optical Physics,* 3(8)**,** 1070-1074.

JENNINGS, D. E. 1980. Absolute line strengths in nu(4), (12)ch(4): A dual-beam diode laser spectrometer with sweep integration. *Appl Opt,* 19(16)**,** 2695-700.

KARPF, A. & RAO, G. N. 2009. Absorption and wavelength modulation spectroscopy of no2 using a tunable, external cavity continuous wave quantum cascade laser. *Applied Optics,* 48(2)**,** 408-413.

KASSANDER, A. R. 2009. *Spectral lineshapes* [Online]. Tucson: Department of Atmospheric Sciences, The University of Arizona. Available: http://lasersci.blogs.rice.edu/files/2014/05/OE49_Nov2010TR_FKT_Oct29_2010.pdf [Accessed: 05/06/2015].

KAZARINOV, R. F. & SURIS, R. A. 1971. Possibility of the amplification of electromagnetic waves in a semiconductor with a superlattice. *Soviet Physics: Semiconductors,* 5(4)**,** 707-709.

KILLINGER, D. & PLUTOV, D. V. 2009. *HITRAN-pc 4.0 user manual* [Online]. Tampa FL: University of South Florida. Available: http://www.cas.usf.edu/lidarlab/hitranpc/Hitran-PC_4.0_Manual.pdf [Accessed: 11/05/2015].

KLUCZYNSKI, P. & AXNER, O. 1999. Theoretical description based on fourier analysis of wavelength-modulation spectrometry in terms of analytical and background signals. *Applied Optics,* 38(27)**,** 5803-15.

KLUCZYNSKI, P., GUSTAFSSON, J., LINDBERG, A. M. & AXNER, O. 2001. Wavelength modulation absorption spectrometry - an extensive scrutiny of the generation of signals. *Spectrochimica acta. Part B, Atomic spectroscopy,* 56(8)**,** 1277-1354.

KOSTEREV, A. A., TITTEL, F. K., GMACHL, C., CAPASSO, F., SIVCO, D. L., BAILLARGEON, J. N., HUTCHINSON, A. L. & CHO, A. Y. 2000. Trace-gas detection in ambient air with a thermoelectrically cooled, pulsed quantum-cascade distributed feedback laser. *Applied Optics,* 39(36)**,** 6866-6872.

LANGFORD, N., DUXBURY, G. & NORMAND, E. 2003. *Semiconductor diode laser spectrometer arrangement and method*. WO patent application PCT/GB2003/001510. 31/12/2003.

LI, H. J., RIEKER, G. B., LIU, X., JEFFRIES, J. B. & HANSON, R. K. 2006. Extension of wavelength-modulation spectroscopy to large modulation depth for diode laser absorption measurements in high-pressure gases. *Applied Optics,* 45(5)**,** 1052-1061.

LINDLEY, R., NORMAND, E., HOWIESON, L., MCCULLOCH, M., BLACK, P., LEWIS, C. & FOULGER, B. 2007. Quantum cascade laser based screening portal for the detection of explosive precursors - art. No. 67410p. *In:* LEWIS, C. (ed.) *Optics and photonics for counterterrorism and crime fighting iii.* 6741, P7410, ISBN: 978-0-8194-6899-4.

LIU, J. T. C., JEFFRIES, J. B. & HANSON, R. K. 2004a. Wavelength modulation absorption spectroscopy with 2f detection using multiplexed diode lasers for rapid temperature measurements in gaseous flows. *Applied physics. B, Lasers and optics,* 78(3-4)**,** 503-511.

LIU, J. T. C., JEFFRIES, J. B. & HANSON, R. K. 2004b. Large-modulation-depth 2f spectroscopy with diode lasers for rapid temperature and species measurements in gases with blended and broadened spectra. *Applied Optics,* 43(35)**,** 6500-6509.

MA, J., CHEESMAN, A., ASHFOLD, M. N. R., HAY, K. G., WRIGHT, S., LANGFORD, N., DUXBURY, G. & MANKELEVICH, Y. A. 2009. Quantum cascade laser investigations of ch4 and c2h2 interconversion in hydrocarbon/h-2 gas mixtures during microwave plasma enhanced chemical vapor deposition of diamond. *Journal of Applied Physics,* 106(3)**,** 033305.

MCCULLOCH, M. T., NORMAND, E. L., LANGFORD, N., DUXBURY, G. & NEWNHAM, D. A. 2003. Highly sensitive detection of trace gases using the time-resolved frequency downchirp from pulsed quantum-cascade lasers. *Journal of the Optical Society of America B-Optical Physics,* 20(8)**,** 1761-1768.

MCCULLOCH, M. T., LANGFORD, N. & DUXBURY, G. 2005. Real-time trace-level detection of carbon dioxide and ethylene in car exhaust gases. *Applied Optics,* 44(14)**,** 2887-2894.

MCCULLOCH, M. T. 2006. *Quantum cascade laser based absorption for trace gas molecular detection and nonlinear optics.* PhD, University of Strathclyde.

MCCULLOCH, M. T., DUXBURY, G. & LANGFORD, N. 2006. Observation of saturation and rapid passage signals in the 10.25 micron spectrum of ethylene using a frequency chirped quantum cascade laser. *Molecular Physics,* 104(16-17)**,** 2767-2779.

MCGETTRICK, A. J., DUFFIN, K., JOHNSTONE, W., STEWART, G. & MOODIE, D. G. 2008. Tunable diode laser spectroscopy with wavelength modulation: A phasor decomposition method for calibration-free measurements of gas concentration and pressure. *Journal of Lightwave Technology,* 26(1-4)**,** 432-440.

MCGETTRICK, A. J., JOHNSTONE, W., CUNNINGHAM, R. & BLACK, J. D. 2009. Tunable diode laser spectroscopy with wavelength modulation: Calibration-free measurement of gas compositions at elevated temperatures and varying pressure. *Journal of Lightwave Technology,* 27(15)**,** 3150-3161.

MCMANUS, J. B., KEBABIAN, P. L. & ZAHNISER, W. S. 1995. Astigmatic mirror multipass absorption cells for long-path-length spectroscopy. *Applied Optics,* 34(18)**,** 3336-3348.

MCMANUS, J. B. 2007. Paraxial matrix description of astigmatic and cylindrical mirror resonators with twisted axes for laser spectroscopy. *Applied Optics,* 46(4)**,** 472-482.

MCMANUS, J. B., ZAHNISER, M. S., NELSON, D. D., SHORTER, J. H., HERNDON, S., WOOD, E. & WEHR, R. 2010. Application of quantum cascade lasers to high-precision atmospheric trace gas measurements. *Optical Engineering,* 49(11).

MYERS, O. E. & PUTZER, E. J. 1959. Measurement broadening in magnetic resonance. *Journal of Applied Physics,* 30(12)**,** 1987-1991.

NAMJOU, K., CAI, S., WHITTAKER, E. A., FAIST, J., GMACHL, C., CAPASSO, F., SIVCO, D. L. & CHO, A. Y. 1998. Sensitive absorption spectroscopy with a room-temperature distributed-feedback quantum-cascade laser. *Optics Letters,* 23(3)**,** 219-221.

NATIONAL INSTRUMENTS. *Tektronix tds3052 oscilloscope labview drivers* [Online]. Available: http://sine.ni.com/apps/utf8/niid_web_display.model_page?p_model_id=169 [Accessed: 11/05/2015].

NAVE, C. R. 2014. *Mean free path* [Online]. Atlanta, Georgia: HyperPhysics, Department of Physics and Astronomy, Georgia State University. Available: http://hyperphysics.phy-astr.gsu.edu/hbase/kinetic/menfre.html [Accessed: 05/06/2015].

NELSON, D. D., SHORTER, J. H., MCMANUS, J. B. & ZAHNISER, M. S. 2002. Sub-part-per-billion detection of nitric oxide in air using a thermoelectrically cooled mid-infrared quantum cascade laser spectrometer. *Applied Physics B-Lasers and Optics,* 75(2-3)**,** 343-350.

NELSON, D. D., MCMANUS, J. B., HERNDON, S. C., SHORTER, J. H., ZAHNISER, M. S., BLASER, S., HVOZDARA, L., MULLER, A., GIOVANNINI, M. & FAIST, J. 2006. Characterization of a near-room-temperature, continuous-wave quantum cascade laser for long-term, unattended monitoring of nitric oxide in the atmosphere. *Optics Letters,* 31(13)**,** 2012-2014.

NERC ARSF. 2012. *Airborne research and survey facility aircraft* [Online]. Natural Environment Research Council (NERC). Available: http://arsf.nerc.ac.uk/aircraft/ [Accessed: 13/03/2016].

NIKODEM, M., WEIDMANN, D. & WYSOCKI, G. 2012. Chirped laser dispersion spectroscopy with harmonic detection of molecular spectra. *Applied Physics B-Lasers and Optics,* 109(3)**,** 477-483.

NIKODEM, M. & WYSOCKI, G. 2012. Chirped laser dispersion spectroscopy for remote open-path trace-gas sensing. *Sensors,* 12(12)**,** 16466-16481.

NORMAND, E., DUXBURY, G. & LANGFORD, N. 2001. Characterisation of the spectral behaviour of pulsed quantum cascade lasers using a high resolution fourier transform infrared spectrometer. *Optics Communications,* 197(1-3)**,** 115-120.

NORMAND, E., MCCULLOCH, M., DUXBURY, G. & LANGFORD, N. 2003. Fast, real-time spectrometer based on a pulsed quantum-cascade laser. *Optics Letters,* 28(1)**,** 16-18.

NORMAND, E., HOWIESON, I., MCCULLOCH, M. & BLACK, P. 2006. Quantum cascade laser (qcl) based sensor for the detection of explosive compounds - art. No. 64020g. *Optics and Photonics for Counterterrorism and Crime Fighting II,* 6402**,** G4020.

NORMAND, E. & HOWIESON, I. 2007. Quantum-cascade lasers enable gas-sensing technology. *Laser Focus World,* 43(4)**,** 90-92.

NORMAND, E. L., STOKES, R. J., HAY, K., FOULGER, B. & LEWIS, C. 2010. Advances in quantum cascade lasers for security and crime-fighting. *In:* LEWIS, C., BURGESS, D., ZAMBONI, R., KAJZAR, F. & HECKMAN, E. M. (eds.) *Optics and photonics for counterterrorism and crime fighting vi and optical materials in defence systems technology vii.* 7838, 0A, ISBN: 978-0-8194-8356-0.

NORTHERN, J. H., RITCHIE, G. A. D., SMAKMAN, E. P., VAN HELDEN, J. H., COCKBURN, J. & DUXBURY, G. 2010. Rapid passage signals induced by chirped quantum cascade laser radiation: K state dependent-delay effects in the nu(2) band of NH3. *Optics Letters,* 35(16)**,** 2750-2752.

NORTHERN, J. H., RITCHIE, G. A. D., SMAKMAN, E. P., VAN HELDEN, J. H., WALKER, R. J. & DUXBURY, G. 2011. Chirped quantum cascade laser induced rapid passage signatures in an optically thick gas. *Applied Physics B-Lasers and Optics,* 102(1)**,** 37-42.

OLIVERO, J. J. & LONGBOTHUM, R. L. 1977. Empirical fits to voigt line-width - brief review. *Journal of Quantitative Spectroscopy & Radiative Transfer,* 17(2)**,** 233-236.

OLSON, M. L., GRIEBLE, D. L. & GRIFFITHS, P. R. 1980. 2nd derivative tunable diode-laser spectrometry for line-profile determination .1. Theory. *Applied Spectroscopy,* 34(1)**,** 50-56.

PHILIPPE, L. C. & HANSON, R. K. 1993. Laser-diode wavelength-modulation spectroscopy for simultaneous measurement of temperature, pressure, and velocity in shock-heated oxygen flows. *Applied Optics,* 32(30)**,** 6090-6103.

PRESS, W. H., TEUKOLSKY, S. A., VETTERLING, W. T. & FLANNERY, B. P. 2007. *Numerical recipes: The art of scientific computing (3rd ed.),* Cambridge, Cambridge University Press. ISBN: 978-0-521-88068-8.

QUIMBY, R. S. 2006. *Photonics and lasers: An introduction,* Hoboken, New Jersey, John Wiley & Sons. ISBN: 0-471-71974-9.

REID, J. & LABRIE, D. 1981. 2nd-harmonic detection with tunable diode-lasers - comparison of experiment and theory. *Applied Physics B-Photophysics and Laser Chemistry,* 26(3)**,** 203-210.

RIEKER, G. B., JEFFRIES, J. B. & HANSON, R. K. 2009. Calibration-free wavelength-modulation spectroscopy for measurements of gas temperature and concentration in harsh environments. *Applied Optics,* 48(29)**,** 5546-5560.

RISBY, T. H. & TITTEL, F. K. 2010. Current status of midinfrared quantum and interband cascade lasers for clinical breath analysis. *Optical Engineering,* 49(11)**,** 111123. Available: http://lasersci.blogs.rice.edu/files/2014/05/OE49_Nov2010TR_FKT_Oct29_2010.pdf [Accessed: 05/06/2015].

ROTHMAN, L. S. *HITRANonline: Definitions and units* [Online]. Cambridge, Massachusetts: Atomic and Molecular Physics Division, Harvard-Smithsonian Centre for Astrophysics. http://hitran.org/docs/definitions-and-units/ [Accessed: 10/01/2016].

ROTHMAN, L. S., RINSLAND, C. P., GOLDMAN, A., MASSIE, S. T., EDWARDS, D. P., FLAUD, J. M., PERRIN, A., CAMY-PEYRET, C., DANA, V., MANDIN, J. Y., SCHROEDER, J., MCCANN, A., GAMACHE, R. R., WATTSON, R. B., YOSHINO, K., CHANCE, K. V., JUCKS, K. W., BROWN, L. R., NEMTCHINOV, V. & VARANASI, P. 1998. The HITRAN molecular spectroscopic database and HAWKS (HITRAN atmospheric workstation): 1996 edition. *Journal of Quantitative Spectroscopy and Radiative Transfer,* 60(5)**,** 665-710. Available: https://www.cfa.harvard.edu/hitran/ [Accessed: 13/02/2016].

ROTHMAN, L. S., JACQUEMART, D., BARBE, A., BENNER, D. C., BIRK, M., BROWN, L. R., CARLEER, M. R., CHACKERIAN, C., CHANCE, K., COUDERT, L. H., DANA, V., DEVI, V. M., FLAUD, J. M., GAMACHE, R. R., GOLDMAN, A., HARTMANN, J. M., JUCKS, K. W., MAKI, A. G., MANDIN, J. Y., MASSIE, S. T., ORPHAL, J., PERRIN, A., RINSLAND, C. P., SMITH, M. A. H., TENNYSON, J., TOLCHENOV, R. N., TOTH, R. A., VANDER AUWERA, J., VARANASI, P. & WAGNER, G. 2005. The HITRAN 2004 molecular spectroscopic database. *Journal of Quantitative Spectroscopy & Radiative Transfer,* 96(2)**,** 139-204.

ROTHMAN, L. S., GORDON, I. E., BARBE, A., BENNER, D. C., BERNATH, P. F., BIRK, M., BOUDON, V., BROWN, L. R., CAMPARGUE, A., CHAMPION, J. P., CHANCE, K., COUDERT, L. H., DANA, V., DEVI, V. M., FALLY, S., FLAUD, J. M., GAMACHE, R. R., GOLDMAN, A., JACQUEMART, D., KLEINER, I., LACOME, N., LAFFERTY, W. J., MANDIN, J. Y., MASSIE, S. T., MIKHAILENKO, S. N., MILLER, C. E., MOAZZEN-AHMADI, N., NAUMENKO, O. V., NIKITIN, A. V., ORPHAL, J., PEREVALOV, V. I., PERRIN, A., PREDOI-CROSS, A., RINSLAND, C. P., ROTGER, M., ŠIMEČKOVÁ, M., SMITH, M. A. H., SUNG, K., TASHKUN, S. A., TENNYSON, J., TOTH, R. A., VANDAELE, A. C. & VANDER AUWERA, J. 2009. The HITRAN 2008 molecular spectroscopic database. *Journal of Quantitative Spectroscopy and Radiative Transfer,* 110(9–10)**,** 533-572. Available: https://www.cfa.harvard.edu/hitran/ [Accessed: 13/02/2016].

RUSSELL, A. M. & TORCHIA, D. A. 1962. Harmonic analysis in systems using phase sensitive detectors. *Review of Scientific Instruments,* 33(4)**,** 442-444.

SALEH, B. E. A. & TEICH, M. C. 2007. *Fundamentals of photonics (2nd ed.),* Hoboken, New Jersey, Wiley. ISBN: 978-0-471-35832-9.

SCHIFF, L. I. 1968. *Quantum mechanics (3rd ed.),* London, McGraw-Hill. ISBN: 07-Y85643-5.

SHORTER, J. H., NELSON, D. D., MCMANUS, J. B., ZAHNISER, M. S. & MILTON, D. K. 2010. Multicomponent breath analysis with infrared absorption using room-temperature quantum cascade lasers. *Ieee Sensors Journal,* 10(1)**,** 76-84.

SIEGMANN, A. E. 1986. *Lasers,* Sausalito, CA, University Science Books. ISBN: 0-935702-11-3.

SILVER, J. A. 1992. Frequency-modulation spectroscopy for trace species detection - theory and comparison among experimental methods. *Applied Optics,* 31(6)**,** 707-717.

SIMECKOVA, M., JACQUEMART, D., ROTHMAN, L. S., GAMACHE, R. R. & GOLDMAN, A. 2006. Einstein a-coefficients and statistical weights for molecular absorption transitions in the HITRAN database. *Journal of Quantitative Spectroscopy & Radiative Transfer,* 98(1)**,** 130-155.

SIRTORI, C., FAIST, J., CAPASSO, F., SIVCO, D. L., HUTCHINSON, A. L., CHU, S. N. G. & CHO, A. Y. 1996. Continuous wave operation of midinfrared (7.4-8.6 mu m) quantum cascade lasers up to 110K temperature. *Applied Physics Letters,* 68(13)**,** 1745-1747.

SMITH, S. W. 1997. *The scientist & engineer's guide to digital signal processing,* San Diego, California Technical Publishing. ISBN: 978-0966017632. Available: http://www.dspguide.com/ [Accessed: 13/04/2015].

SPECTRALCALC.COM. *Calculation of molecular spectra with the spectral calculator* [Online]. Newport News VA: GATS, Inc. Available: http://www.spectralcalc.com/info/CalculatingSpectra.pdf [Accessed: 11/05/2015].

STANFORD RESEARCH SYSTEMS. *About lock-in amplifiers: Application note #3* [Online]. Standford Research Systems. Available: http://www.thinksrs.com/downloads/PDFs/ApplicationNotes/AboutLIAs.pdf [Accessed: 08/04/2015].

STEWART, G., JOHNSTONE, W., BAIN, J. R. P., RUXTON, K. & DUFFIN, K. 2011. Recovery of absolute gas absorption line shapes using tunable diode laser spectroscopy with wavelength modulation-part i: Theoretical analysis. *Journal of Lightwave Technology,* 29(6)**,** 811-821.

SUPPLEE, J. M., WHITTAKER, E. A. & LENTH, W. 1994. Theoretical description of frequency-modulation and wavelength modulation spectroscopy. *Applied Optics,* 33(27)**,** 6294-6302.

TASINATO, N., DUXBURY, G., LANGFORD, N. & HAY, K. G. 2010a. An investigation of collisional processes in a dicke narrowed transition of water vapor in the 7.8 mu m spectral region by frequency down-chirped quantum cascade laser spectroscopy. *Journal of Chemical Physics,* 132(4)**,** 044316.

TASINATO, N., HAY, K. G., LANGFORD, N., DUXBURY, G. & WILSON, D. 2010b. Time dependent measurements of nitrous oxide and carbon dioxide collisional relaxation processes by a frequency down-chirped quantum cascade laser: Rapid passage signals and the time dependence of collisional processes. *Journal of Chemical Physics,* 132(16)**,** 164301.

TAYLOR, J. R. 1997. *An introduction to error analysis : The study of uncertainties in physical measurements,* Sausalito, Calif., University Science Books. ISBN: 0935702423.

THORNE, A. P., LITZÉN, U. & JOHANSSON, S. 1999. *Spectrophysics: Principles and applications,* Berlin, Springer. ISBN: 3-540-65117-9.

TRAVIS, J. & KRING, J. 2009. *Labview for everyone: Graphical programming made easy and fun (3rd ed.),* Boston MA, Prentice Hall. ISBN: 978-0-13-185672-1.

V E ZUEV IAO. *Hitran on the web* [Online]. Tomsk: V E Zuev Institute of Atmospheric Optics (IAO). Available: http://hitran.iao.ru [Accessed: 16/04/2015].

VAUGHAN, J. M. 1989. *The fabry-perot interferometer: History, theory, practice and applications,* New York, Taylor & Francis Group. ISBN: 0-85274-138-3.

WAHLQUIST, H. 1961. Modulation broadening of unsaturated lorentzian lines. *Journal of Chemical Physics,* 35(5)**,** 1708-1710.

WALDRAM, J. R. 1989. *The theory of thermodynamics,* Cambridge, Cambridge University Press. ISBN: 0-521-28796-0.

WALTON, P. H. 2004. *Beginning group theory for chemistry,* Oxford, Oxford University Press. ISBN: 0-19-855964-X.

WEBSTER, C. R., FLESCH, G. J., SCOTT, D. C., SWANSON, J. E., MAY, R. D., WOODWARD, W. S., GMACHL, C., CAPASSO, F., SIVCO, D. L., BAILLARGEON, J. N., HUTCHINSON, A. L. & CHO, A. Y. 2001. Quantum-cascade laser measurements of stratospheric methane and nitrous oxide. *Appl Opt,* 40(3)**,** 321-6.

WELZEL, S., LOMBARDI, G., DAVIES, P. B., ENGELN, R., SCHRAM, D. C. & ROPCKE, J. 2008. Trace gas measurements using optically resonant cavities and quantum cascade lasers operating at room temperature. *Journal of Applied Physics,* 104(9).

WELZEL, S. 2009. *New enhanced sensitivity infrared laser spectroscopy techniques applied to reactive plasmas and trace gas detection,* Berlin, Logos Verlag Berlin. ISBN: 978-3-8325-2345-9.

WELZEL, S., HEMPEL, F., HUBNER, M., LANG, N., DAVIES, P. B. & ROPCKE, J. 2010. Quantum cascade laser absorption spectroscopy as a plasma diagnostic tool: An overview. *Sensors,* 10(7)**,** 6861-6900.

WERLE, P., MUCKE, R. & SLEMR, F. 1993. The limits of signal averaging in atmospheric trace-gas monitoring by tunable diode-laser absorption-spectroscopy (tdlas). *Applied Physics B-Photophysics and Laser Chemistry,* 57(2)**,** 131-139.

WHITING, E. E. 1968. An empirical approximation to voigt profile. *Journal of Quantitative Spectroscopy & Radiative Transfer,* 8(6)**,** 1379-1384. Available: http://ir.hfcas.ac.cn/bitstream/334002/11949/1/An%20empirical%20approximation%20to%20the%20Voigt%20profile%20.pdf [Accessed: 30/05/2015].

WILLIAMS, R. M., KELLY, J. F., HARTMAN, J. S., SHARPE, S. W., TAUBMAN, M. S., HALL, J. L., CAPASSO, F., GMACHL, C., SIVCO, D. L., BAILLARGEON, J. N. & CHO, A. Y. 1999. Kilohertz linewidth from frequency-stabilized mid-infrared quantum cascade lasers. *Optics Letters,* 24(24)**,** 1844-1846.

WILLOCK, D. J. 2009. *Molecular symmetry,* Chichester, John Wiley & Sons, Ltd. ISBN: 978-0-470-85348-1.

WILSON, G. V. H. 1963. Modulation broadening of nmr and esr line shapes. *Journal of Applied Physics,* 34(11)**,** 3276-3285.

WRIGHT, S., DUXBURY, G. & LANGFORD, N. 2006. A compact quantum-cascade laser based spectrometer for monitoring the concentrations of methane and nitrous oxide in the troposphere. *Applied Physics B-Lasers and Optics,* 85(2-3)**,** 243-249.

WYSOCKI, G., CURL, R. F., TITTEL, F. K., MAULINI, R., BULLIARD, J. M. & FAIST, J. 2005. Widely tunable mode-hop free external cavity quantum cascade laser for high resolution spectroscopic applications. *Applied Physics B,* 81(6)**,** 769-777.

WYSOCKI, G. & WEIDMANN, D. 2010. Molecular dispersion spectroscopy for chemical sensing using chirped mid-infrared quantum cascade laser. *Optics Express,* 18(25)**,** 26123-26140.

YARIV, A. & YEH, P. 2007. *Photonics: Optical electronics in modern communications (6th ed.),* Oxford, Oxford University Press. ISBN: 978-0-19-517946-0.

YE, J., MA, L. S. & HALL, J. L. 1998. Ultrasensitive detections in atomic and molecular physics: Demonstration in molecular overtone spectroscopy. *Journal of the Optical Society of America B-Optical Physics,* 15(1)**,** 6-15.

ZHU, X. & CASSIDY, D. T. 1997. Modulation spectroscopy with a semiconductor diode laser by injection-current modulation. *Journal of the Optical Society of America B-Optical Physics,* 14(8)**,** 1945-1950.