Pacific Northwest
National Laboratory
Operated by Battelle for the
U.S. Department of Energy

Ultra-Trace Chemical Sensing with Long-Wave Infrared Cavity-Enhanced Spectroscopic Sensors

M. S. Taubman   R. M. Williams
T. L. Myers   J. F. Schultz
B. D. Cannon

February 2003

Ultra-Trace Chemical Sensing with Long-Wave Infrared Cavity-Enhanced Spectroscopic Sensors

M. S. Taubman   R. M. Williams
T. L. Myers   J. F. Schultz
B. D. Cannon

February 2003

Summary

The infrared sensors task of Pacific Northwest National Laboratory's (PNNL's) Remote Spectroscopy Project (Task B of Project PL211) is focused on the science and technology of remote and in-situ spectroscopic chemical sensors for detecting proliferation and countering terrorism. Missions to be addressed by remote chemical sensor development include detecting proliferation of nuclear or chemical weapons, and providing warning of terrorist use of chemical weapons. Missions to be addressed by in-situ chemical sensor development include countering terrorism by screening luggage, personnel, and shipping containers for explosives, firearms, narcotics, chemical weapons, or chemical weapons residues, and mapping contaminated areas. The science and technology is also relevant to chemical weapons defense, air operations support, monitoring emissions from chemical weapons destruction or industrial activities, law enforcement, medical diagnostics, and other applications.

Sensors for most of these missions will require extreme chemical sensitivity and selectivity because the signature chemicals of importance are expected to be present in low concentrations or have low vapor pressures, and the ambient air is likely to contain pollutants or other chemicals with interfering spectra. Cavity-enhanced chemical sensors (CES) that draw air samples into optical cavities for laser-based interrogation of their chemical content promise real-time, in-situ chemical detection with extreme sensitivity to specified target molecules and superb immunity to spectral interference and other sources of noise.

PNNL is developing CES based on quantum cascade (QC) lasers that operate in the mid-wave infrared (MWIR - 3 to 5 microns) and long-wave infrared (LWIR - 8 to 14 microns), and CES based on telecommunications lasers operating in the short-wave infrared (SWIR - 1 to 2 microns). All three spectral regions are promising because smaller molecular absorption cross sections in the SWIR are offset by the superior performance, maturity, and robustness of SWIR lasers, detectors, and other components, while the reverse is true for the MWIR and LWIR bands. PNNL's research activities include identification of signature chemicals and quantification of their spectroscopy, exploration of novel sensing techniques, and experimental sensor system construction and testing. In FY02, experimental QC laser systems developed with DARPA funding were used to explore continuous-wave (cw) CES in various forms culminating in the NICE-OHMS technique [1-3] discussed below. In FY02 PNNL also built an SWIR sensor to validate utility of the SWIR spectral region for chemical sensing, and explore the science and engineering of CES in field environments. The remainder of this report is devoted to PNNL's LWIR CES research.

During FY02 PNNL explored the performance and limitations of several detection techniques in the LWIR including direct cavity-enhanced absorption, cavity-dithered phase-sensitive detection and resonant sideband cavity-enhanced detection. This latter technique is also known as NICE-OHMS, which stands for Noise-Immune Cavity-Enhanced Optical Heterodyne Molecular Spectroscopy. This technique, pioneered in the near infrared (NIR) by Dr J. Hall and coworkers at the University of Colorado, is one of the most sensitive spectroscopic techniques currently known. In this report, the first demonstration of this technique in the LWIR is presented. The noise-equivalent absorption sensitivities (NEAS) achieved in these experiments are promising, but must be translated into parts-per-billion volume sensitivities to specific chemicals at the air intake in the presence of interferents to assess practical LWIR CES capabilities and identify the most promising CES techniques, configurations, and modus operandi. While NEAS are relatively easy to measure in the laboratory, the translation to practical sensitivity and selectivity limits involves numerous challenging science and engineering questions. PNNL's goals for FY03 include improving the NEAS performance obtained with LWIR CES in FY02, quantification of chemical sensitivities, identification of the most promising sensing techniques, and formulation of plans for MWIR and LWIR field experiments.

Acknowledgments

Contents

Figures

1.0 Introduction and Technical Background

The most basic absorption-based chemical point sensor consists of a light source, a sample, and a detector observing the light transmitted through the sample. A simple laser realization of a trace-gas absorption sensor is illustrated in Figure 1.1 below. A laser beam (also referred to as a laser field) passes through a gas cell containing an unknown sample at a specific pressure. Let us assume that this sample contains unknown concentrations of chemicals of interest (analytes) that absorb at particular frequencies in the laser tuning range. Light exiting the cell is incident on a transmission detector, the electrical output of which is recorded as the laser frequency is scanned over one or more absorption features. We begin in this introduction using this sensor as an illustration of what a trace-gas chemical sensor may be required to do, and what is important in the design and characterization processes. An outline of the remaining sections of this report is then given.

DIRECT ABSORPTION SENSOR
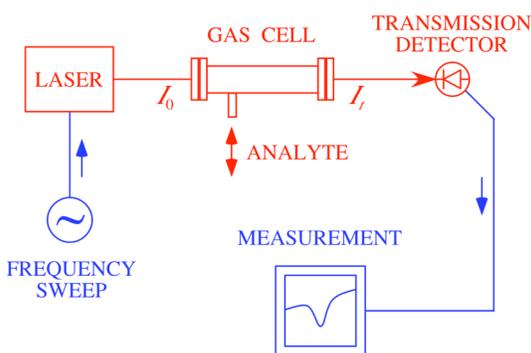
Fig. 1.1. Direct absorption sensor. The output of a laser passes through a gas cell and is then detected. The transmitted intensity is measured as the laser frequency is swept through molecular transitions in the gas sample. Incident and transmitted intensities I_0 and I_t are indicated.

The ultimate purpose of a sensor such as that shown in Figure 1.1 is to indicate the presence of, and to measure the concentrations of particular gaseous species in the unknown sample. This process involves several steps, the most fundamental of which is to measure and record the absorption of the laser field due to the gas sample as a function of optical frequency. The resulting spectrum may consist of features or "lines" from many absorbing chemicals. The second step is to identify these chemicals. Such a sensor ideally records not only the strengths of the absorption features in a gas sample, but also the frequencies at which they occur to within some uncertainty. Since particular chemicals absorb at specific frequencies, a given feature in a measured spectrum could be attributed to certain possible chemicals. A more positive identification can occur when multiple absorption features measured over a specific frequency region can be attributed to the same chemical. However, since these features are measured with a certain spectral width, overlap of features arising from different chemicals can also readily occur. Even so, the known strengths of the absorption lines of the chemicals in question (assuming they have been sufficiently well studied) can be used in a statistical analysis to yield the concentrations of the analytes in the gas sample. This is an example of a chemometric analysis [4] and completes the third step in the chemical sensing process.

In this report, we focus on understanding the performance and limitations of several chemical sensor architectures developed at PNNL for operation in the infrared. There are several elements to the development and characterization of a chemical sensor. Firstly, we must understand what the detected signals in a sensor imply about the absorbing medium. To illustrate, the optical intensities entering and exiting the gas cell in the direct absorption sensor shown in Figure 1.1 are related by the Beer-Lambert law

$$I_t = I_0 \exp(-\alpha L) \qquad (1.1)$$

where α is known as the extinction (or absorption) coefficient of the medium or analyte, and L is the optical path length of the laser field within the analyte. Since L is known for a given gas cell, α can be readily determined by fitting (1.1) to data taken with this sensor. The extinction coefficient α, is the absorption per unit length due to the analyte, and usually has units of cm^{-1}. As we shall see, the way in which α is related to the signals measured in a sensor depends strongly on the sensor architecture, and will be different to (1.1) for more complicated sensors. Understanding this is not only crucial for correctly calibrating the sensor, but also for predicting its performance limitations.

Chemical sensors are usually employed to measure weak absorption features and it is important to know what minimum levels of absorption can be observed with a given sensor. The signal we measure is essentially a change in the detector output (an electrical signal) as the laser frequency is scanned. Fluctuations not originating from scanning over absorption features of an analyte will mask and interfere with measurements of those that are. These spurious signals constitute noise, the major consequence of which is a minimum detectable intensity fluctuation $(\Delta I)_{min}$. For a given sensor, this can be translated (see Section 2.1) into a minimum detectable extinction coefficient α_S – one of several useful metrics for comparing the potential sensitivity of different chemical sensors. For the direct absorption sensor in Figure 1.1, this is:

$$\alpha_S = \frac{(\alpha L)_{min}}{L} = \frac{(\Delta I)_{min}}{I_0 L} \qquad (1.2)$$

It is important to understand that this quantity is a characteristic of the sensor and not of a chemical that we may wish to detect. It is also important to note that the value we obtain for α_S depends on how long we take to measure it. This is because the longer the measurement time, the more we average out any random noise contributions and the smaller values of ΔI we can measure. Consequently, α_S is specified as being measured over a certain averaging time or within a certain bandwidth. In order to compare values measured with different bandwidths, they are usually normalized to one second, or a 1 Hz bandwidth. Consequently, in this report α_S is specified in units of $/\mathrm{cm}/\sqrt{\mathrm{Hz}}$.

1.1 Summary of Work Performed in FY02

During FY02, researchers at PNNL investigated and refined several different sensor architectures, leading up to the first full demonstration of the NICE-OHMS technique in the LWIR, which used quantum cascade lasers (QCLs) operating at 8.5 microns. The

resulting sensor routinely achieves sensitivities α_S down to the mid $10^{-11}/\text{cm}/\sqrt{\text{Hz}}$. The inventors of this technique, John Hall, Jun Ye, and Long-Shen Ma of JILA, University of Colorado, Boulder, CO, demonstrated in 1998 that it gave a factor of 1000 improvement over the best previous cavity enhanced technique, cavity ring-down, and with this technique they still hold the record for the best absorption sensitivity of $10^{-14}/\text{cm}/\sqrt{\text{Hz}}$.

The work conducted at PNNL over FY02 has involved the development of specialized optical cavities and gas cells, and has specifically involved the refinement of the production of high quality mirrors in the LWIR. Traditionally, it has been difficult to find mirrors having high reflectivity, low losses and good surface figures in this wavelength region because of the lack of demand for such optics and the thickness of the coatings. However, through the co-operation of certain coating houses, we have been able to obtain in FY02 mirrors allowing cavities with finesse values of over 2000 (see Section 2.3). This has represented a considerable step forward in LWIR sensor development. In addition to the development and design of optical systems, considerable electronics has been built to facilitate the experimental work performed over FY02. This includes the construction and refinement of low-noise constant-current power supplies for QCLs, servo control systems for locking the QCLs to optical cavities (see Section 3.1) and low-noise transimpedance amplifiers for operation with the mercury-cadmium telluride detectors used in these chemical sensors.

1.2 Layout

The theory of operation of the chemical sensors under development at PNNL will be discussed in detail in Section 2. This will begin with a more in-depth discussion of the direct absorption sensor presented in this introduction, along with its performance limitations due to noise. This forms a basis for the presentation of cavity-enhanced sensors, which are then introduced. The impact of noise in chemical sensors also naturally leads into a discussion of frequency modulation (FM) and the standard signal recovery technique known as phase-sensitive detection. This section finishes with a discussion of the resonant sideband recovery scheme known as NICE-OHMS, or Noise-Immune Cavity-Enhanced Optical Heterodyne Molecular Spectroscopy.

In Section 3, experimental considerations are discussed. The importance of laser stabilization for cw cavity-enhanced chemical sensors is introduced, and a summary of our laser stabilization work for DARPA is given. We discuss in detail the work performed and the results obtained over FY02 at PNNL on cavity-enhanced chemical sensors in the LWIR using quantum cascade lasers. The NICE-OHMS sensor developed at PNNL contains elements from a variety of techniques including cavity-enhancement, frequency modulation and phase-sensitive detection. Consequently, we discuss the operation and advantages of individual techniques and improvements in order of increasing complexity, culminating in a discussion of the NICE-OHMS demonstration.

Section 4 is an outlook of further work to be performed in FY03 and further out. In particular, methods and problems associated with the conversion of absorption information into concentrations are addressed. This is important because firstly, while

absorption sensitivity values are excellent tools for comparing different sensors, they do not give a good idea of the utility of a specific sensor for a particular analyte. Secondly, two different sensors having the same absorption sensitivity may not yield the same possible concentration sensitivities because of logistical differences in sensor operation. An example of this shall be seen in Section 3.3 when discussing the detection of sub-Doppler features because of the limited range of gas pressures required for these features to be seen.

2.0   Theory of Cavity-Enhanced Sensors

In this section, we begin by revisiting the direct absorption sensor, and introducing some of the standard nomenclature used in the field and discussing the topic of noise in some detail. We then extend from this basis as far as possible to the different variants of chemical sensors developed at PNNL over FY02.

2.1 Sensitivity and Figures of Merit

The direct absorption sensor shown in Figure 1.1 is well described by the Beer-Lambert law (1.1), rewritten here with explicit frequency dependence:

$$I_t(v) = I_0(v)\exp(-\alpha(v)L) \qquad (2.1)$$

where $I_0(v)$ and $I_t(v)$ are the optical intensities incident upon and exiting the gas cell, and $\alpha(v)$ is the extinction coefficient in cm^{-1} as introduced in Section 1.0. It may seem strange that I_0 has a frequency dependence, but in fact this is nearly always the case. As with most semiconductor lasers, the frequency of the QCLs we use is altered primarily by varying the drive current. This consequently changes the intensity of the beam incident on the gas cell, causing a sloping background in the resulting absorption traces, a typical example of this being shown in Figure 2.1.
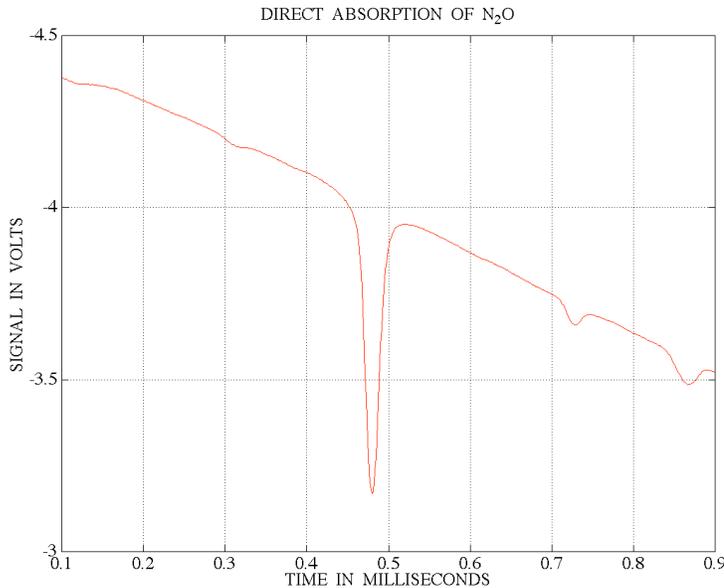
Fig. 2.1. Typical direct absorption scan of nitrous oxide showing several absorption features. A linear current ramp is applied to the QCL, which results in a linear change in laser frequency with time. The sloping background is due to the changing laser power that accompanies the varying laser current that scans the laser frequency.

This sloping background can be removed by normalizing by this varying incident intensity:

$$\frac{I_t(v)}{I_0(v)} = \exp(-\alpha(v)L) \qquad (2.2)$$

For many cases of interest, including gases at low pressure, we can expand the extinction coefficient $\alpha(\nu)$ as

$$\alpha(\nu) = \sum \sigma_i(\nu)N_i \qquad (2.3)$$

where $\sigma_i(\nu)$ is the absorption cross-section in cm^2/molecule of the ith chemical at ν, N_i is its number density in molecules/cm^3, and the sum is over all chemicals present. For gas samples, it is often convenient to use an alternate expression for $\alpha(\nu)$ where partial pressures replace concentrations and extinction coefficients per unit partial pressure, $\alpha_{ip}(\nu)$, replace cross sections. For tunable single-frequency lasers and low-pressure gas samples, laser frequencies can often be found that reduce this sum to a single term, $\sigma(\nu)N$ or $\alpha_p(\nu)P$, involving only one chemical because only its cross-section is significant at the particular frequency ν. We will assume this to be valid throughout the rest of this report and will often use the relation

$$\alpha L = \alpha_p PL = \sigma NL \qquad (2.4)$$

To investigate the performance limitations of the direct absorption sensor in Figure 1.1, we consider the limit of scanning the sensor over small absorption features. The correspondingly small intensity fluctuations at the detector can be written

$$\Delta I = I_0 - I_t \qquad (2.5)$$

Since we are assuming that $\Delta I << I_0$, we can expand and retain only the constant and linear terms of the exponential in (2.2) giving

$$\frac{\Delta I}{I_0} = \alpha L \qquad (2.6)$$

As mentioned earlier, there is a minimum detectable value of ΔI because of the inherent noise sources in the sensor. We now treat this slightly more rigorously than in Section 1.0 by defining a hypothetical optical intensity fluctuation, ∂I_N, equal in size to that which would be produced by all the combined noise sources as if they occurred due to scanning over a weak absorption feature of the analyte (in an otherwise noiseless sensor). Since it is customary to quote the minimum detectable value of ΔI as being that which corresponds to a "signal-to-noise ratio" of unity, we thus write:

$$(\Delta I)_{min} = \partial I_N \qquad (2.7)$$

This yields equation (1.2) for the minimum detectable extinction coefficient, which we extend here to be

$$\alpha_S = \frac{(\alpha L)_{min}}{L} = \frac{(\Delta I)_{min}}{I_0 L} = \frac{\partial I_N}{I_0 L} \qquad (2.8)$$

This quantity is the potential absorption sensitivity of the instrument, and is one figure of merit for judging and comparing sensor performance. In theory, we could use (2.4) to find the minimum detectable number density of a specific well-resolved analyte:

$$N_{min} = \frac{\alpha_S}{\sigma} \qquad (2.9)$$

This would then allow an estimation of the concentration (parts per billion, etc) by dividing this quantity by the total number density of the gas sample. However, this opens the topic of concentration sensitivity in which great care must be taken when quoting the sensor performance in this manner. This discussion is taken up in Section 4.2.

Another common measure of sensor performance similar to α_S but which originates in the remote sensing community where the optical path length may not in fact be known, is the Noise Equivalent Absorbance Sensitivity, or NEAS. The NEAS is based on another often-used quantity, the absorbance:

$$Absorbance \equiv -\ln\left(\frac{I_t}{I_0}\right) = \sigma N L \qquad (2.10)$$

which is linear in the analyte concentration even when very little light is transmitted. In a similar manner to the derivation of α_S above, for small fluctuations in the direct absorption sensor we see that

$$-\ln\left(\frac{I_t}{I_0}\right) = -\ln\left(\frac{I_0 - \Delta I}{I_0}\right) \cong \frac{\Delta I}{I_0} \cong \sigma N L \qquad (2.11)$$

Using (2.8) this gives

$$NEAS = \alpha_S L = \frac{\partial I_N}{I_0} = (\sigma N L)_{min} \qquad (2.12)$$

For a given point sensor of known length, the discussion of which figure of merit may be more useful becomes largely esoteric and opinions vary within the community. Ultimately, α_S, which we refer to as the "absorption sensitivity" or simply the "sensitivity," gives the absolute potential performance of a given sensor, and allows a direct comparison of the sensitivity of two particular sensors. The NEAS on the other hand, is particularly useful when discussing the limitations due to noise on a particular class of sensors or a specific technique of signal recovery, without having to discuss cell length. This will be seen when the shot-noise limited performance of the direct absorption sensor is calculated in Section 2.2. In this report we use both where we can, and stress one over the other where appropriate.

2.2 The Impact of Noise

The sensitivity of any chemical sensor is limited by noise. This is a central issue since the community is continually trying to push the performance limits of sensors for trace-gas detection, i.e., to make them capable of detecting smaller number densities N of a particular analyte. From (2.9) it's clear that for a given absorption sensitivity α_S, choosing absorption lines with the largest available values of absorption cross-section σ makes a lot of sense. While our control over this parameter is limited by what absorption lines are available, we have essentially satisfied this criterion by using QCLs operating in the MWIR and LWIR. These spectral regions contain some of the largest absorption cross-sections for molecular vibrations, which offer both good chemical specificity and sensitivity. Another advantage of QCLs is that they offer much higher tunable powers than previously available in these spectral regions compared to their direct competitors, the lead-salt diode lasers. [5] Provided the molecular transitions being probed don't become saturated, increased optical power reduces the impact of nearly all sources of noise, thus decreasing the value of α_S as can be seen from (2.8), making the sensor more "sensitive."[a] This is not true for noise sources that scale directly with the optical power such as optical fringing (see below).

Having chosen optimum values of σ and optical power, equation (2.8) also shows that α_S can be further reduced for a given sensor by minimizing the total noise (represented by ∂I_N) or maximizing the interaction length L of the analyte with the laser field. In the laboratory, it is far easier to increase the cell length than in the field. Point sensors are often required to be small and portable, limiting possible cell sizes and path lengths. Folded paths such as Herriot cells are one solution to needing a long path length in a small sensor, as are the use of optical cavities, which will be introduced in Section 2.3.

To minimize ∂I_N we need to understand the different sources of noise. We first consider "white" noise such as shot noise in the detector photocurrent and Johnson noise in the detector and pre-amplifier resistors. [6] Shot noise is due to the quantum or particle nature of light, and manifests itself in our sensor upon the conversion of light incident on the transmission detector to electrical current. For a given current from the detector i_{DC}, the shot noise is found to be

$$i_{sh} = \sqrt{2ei_{DC}} \quad \text{Amps}/\sqrt{\text{Hz}} \tag{2.13}$$

where e is the charge of the electron. For a LWIR detector, this current contains contributions from the laser photocurrent, the thermal blackbody photocurrent and the detector leakage current. With QCLs and our sensors, the laser photocurrent dominates and the resulting relative current fluctuations are

---

(a) A quantity that expresses sensitivity often indicates better performance by taking on smaller values. This leads to the seeming contradiction of the term "more sensitive" corresponding to a lower value of "sensitivity". This is usually avoided by stating it in other ways, such as "better sensitivity values" or "better performance", etc.

$$\frac{i_{sh}}{i_{DC}} = \sqrt{\frac{2e}{i_{DC}}} \, /\sqrt{\text{Hz}} \qquad (2.14)$$

For example, if we detect a photocurrent of 65 μA (a reasonable number for a operational chemical sensor in the LWIR), the value of limiting relative current fluctuations due to the shot noise is 7.0×10^{-8} $/\sqrt{\text{Hz}}$. Shot noise is unavoidable,[a]although its impact is eased somewhat by increasing laser power as mentioned earlier, but optical power is often limited by other factors, and the benefit is limited due to the square-root dependence on photocurrent in (2.14). (In contrast, with the low powers available with lead-salt diode lasers, the thermal blackbody photocurrent often dominates and sensitivity improves linearly with increased laser power.)

Johnson noise is due to the thermal agitation of electrons within any resistive element, and resistors are an essential part of the amplifier circuitry used in conjunction with optical detectors. The Johnson current noise is given by

$$i_R = \sqrt{\frac{4kT}{R}} \; \text{Amps}/\sqrt{\text{Hz}} \qquad (2.15)$$

where k is Boltzman's constant, T is the temperature, and R is the resistance. The feedback resistor in a transimpedance amplifier [7] used in conjunction with the detector element is a typical source of limiting Johnson noise. Given the photocurrent of 65 μA used above as an example, 42 kΩ is a reasonable resistance for the preamplifier. The relative Johnson current noise i_R/i_{DC} under these conditions is 9.6×10^{-9} $/\sqrt{\text{Hz}}$. Clearly, the shot noise for this photocurrent, calculated above, dominates the Johnson noise from the feedback resistor, although this need not always be the case. Other sources of electronic noise include the input noise of the operational amplifiers used in the transimpedance amplifier, and the Johnson noise of other surrounding resistors in the circuit. Careful design is required to ensure that the measurement is not dominated by Johnson noise or input amplifier noise.

Another sort of noise is "1/f" noise, which has the property that in a fixed measurement bandwidth centered at a frequency f, the measured noise increases as f decreases. This noise reduces to values below the white noise at higher frequencies. This is the most common kind of noise after white noise discussed above, and is due to a variety of physical processes. Sources of 1/f noise in a chemical sensor include the semiconductor components inside the detector preamplifier, the laser power supply and the laser itself. Methods of circumventing 1/f noise include strategic choices of semiconductors, careful

---

(a) This is not strictly true. While the total quantum noise cannot be reduced, its impact on some variables can be reduced at the expense of increased impact on others. This alluring effect known as "squeezing", has been extremely well covered in the literature, and is still an active area of research. However, it is often difficult and complex. As such, it is not considered here as applicable to mainstream chemical sensing.

electronic design, implementing modulation techniques (Section 2.4) and laser frequency stabilization (Section 3.1).

So far, we have been discussing random or statistical noise sources. There are also systematic noise sources. One such source peculiar to lasers is optical fringing, which is often the largest noise source in a chemical sensor in the LWIR. Optical fringing results from interference between reflected or scattered light from different surfaces, such as cell and detector windows, lenses, and other optics. This interference changes the laser power reaching the detector and changes with laser frequency and the motion or vibration of any optical component that changes the path length difference. We will see direct evidence of this noise in Section 3.2, where we will also discuss its effects at some length. We will also see possible ways around it, including selecting absorption features that have a fundamentally different size to those caused by optical fringing in order to allow them to be easily distinguished.

An effect that can couple both systematic and random noise into the detection signal of a chemical sensor is residual amplitude modulation (RAM). When we modulate the frequency (or phase) of a laser field, there is usually a small component of amplitude modulation (AM) present. How this occurs depends on the modulation method. In our case, we frequency modulate the QCL by applying radio frequency (RF) signals to the drive current. In a similar way to the sweep of the laser current causing an amplitude change in addition to the desired frequency scan, which yields the sloping background shown in Figure 2.1, an RF modulation current produces some AM along with the desired FM. While the FM techniques we employ circumvent the 1/f noise, the RAM imposed via this very same technique can actually re-couple this noise back into the signal. This will become clearer after a discussion of FM modulation in Section 2.4. There are other more subtle effects involving RAM that will be touched on when we discuss locking and stabilization in Section 3.1.

In Section 2.1, we introduced the noise equivalent intensity fluctuation ∂I_N as that hypothetical optical signal in an equally hypothetical noiseless sensor that produces the same size measured electrical signal corresponding to the sum of all the noise sources. We can in fact define a similar noise equivalent intensity fluctuation for each noise source. Provided these quantities are random and independent, they are related to the total equivalent noise intensity fluctuation ∂I_N by the quadrature sum

$$\partial I_N^2 = \sum \partial I_i^2 \tag{2.16}$$

where each of the noise sources is expressed in the same bandwidth. It is also important to note, that since these quantities include the square root of bandwidth in their units, their numerical size becomes smaller for smaller bandwidth or longer averaging times.

2.2.1 The Ultimate Limit

As explained above, shot noise cannot be avoided. Consequently, it is often regarded as the ultimate limit. The effects of all other noise sources can in theory be eliminated or reduced to an arbitrarily small value. This does not mean that it is easy or even practical

to do so. Reducing the effects of noise involves significant engineering and optimization. Many of the complexities and innovations inherent in the more advanced sensors under development at PNNL and elsewhere have been implemented specifically to reduce the effects of noise. Despite this, it is rare that a chemical sensor operates at the shot noise limit. Nevertheless, it is useful to calculate the shot noise limit for such a sensor. Since photocurrent is proportional to optical intensity, we assume shot noise to be the only noise source, and write

$$\frac{\partial I_N}{I_0} = \frac{\partial I_{sh}}{I_0} = \frac{i_{sh}}{i_{DC}} \tag{2.17}$$

Substituting this and equation (2.14) into (2.12) we get the shot noise limited version of the NEAS, or the "SNEAS" for the direct absorption sensor shown in Figure 1.1:

$$SNEAS = \left(\alpha_S L\right)_{sh} = \sqrt{\frac{2e}{i_{DC}}} / \sqrt{\text{Hz}} \tag{2.18}$$

For our example of 65 μA above, this gives $7.0 \times 10^{-8} / \sqrt{\text{Hz}}$. It must be stressed here however, that this noise level will be almost impossible to attain for this sensor due to other noise sources, in particular, 1/f noise due to the laser. As we discuss other sensors in this report, we will review this calculation.

2.3   Enhancement of Absorption Sensitivity Using Optical Cavities

Consider the simple cavity-enhanced absorption sensor depicted in Figure 2.2.

CAVITY-ENHANCED ABSORPTION
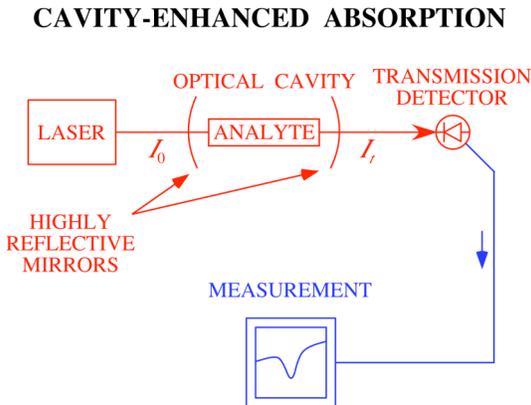
Fig. 2.2. Cavity-Enhanced Absorption. Highly reflective mirrors are placed around the analyte forming an optical cavity. Like any resonator, its sensitivity to loss allows much increased sensitivity to optical absorption.

This is the same direct absorption sensor as depicted in Figure 1.1, but with an optical cavity placed around the gas sample, or analyte. An optical cavity, also known as an optical resonator, consists of two precisely aligned highly reflective curved mirrors where light in the cavity retraces the same path each round-trip between the mirrors. Constructive interference is obtained only in narrow, discrete spectral regions for which the total optical path lengths are approximately integer multiples of the wavelength of the light. These narrow spectral regions are called longitudinal modes, and their separation

in frequency is known as the free spectral range (FSR) of the cavity. When the laser frequency matches a longitudinal mode frequency the constructive interference couples much of the incident laser field into the cavity. The high reflectivity of the mirrors can trap this light for effectively thousands of round trips and results in intensities in the cavity thousands of times higher than that of the incident beam. Placing such a cavity around a gas sample allows the molecular transitions to be probed by the intra-cavity circulating field. Since, as in any resonator, this power-buildup is extremely sensitive to loss, a relatively small increase in absorption due to the analyte causes a dramatic reduction in cavity transmission. For small amounts of absorption, the multiple round trips of light inside an optical cavity present an effective path length much greater than the cavity's physical length. The result is an instrument with potentially much improved absorption sensitivity compared to direct absorption detection. This is illustrated by the following derivation.

The ratio of transmitted to incident power for a resonant optical cavity is [8]

$$\frac{I_t}{I_0} = M \frac{T^2}{\left[1 - R \cdot \exp(-\alpha L)\right]^2} \tag{2.19}$$

where T and R are the intensity transmission and reflection coefficients of both cavity mirrors and L is both the mirror separation and the interaction length in the analyte. The factor M is the fraction of the laser beam that is matched to the spatial profile of the optical cavity mode. Following a treatment similar to Gianfrani et al. [2,9], if αL is much less than one, then

$$\exp(-\alpha L) \approx 1 - \alpha L \tag{2.20}$$

and we can rewrite (2.19) as

$$\begin{aligned} \frac{I_t}{I_0} &\approx M \frac{T^2}{\left[1 - R + R\alpha L\right]^2} \\ &= M \frac{T^2}{\left[1 - R\right]^2} \frac{1}{\left[1 + \alpha L \dfrac{R}{1 - R}\right]^2} \end{aligned} \tag{2.21}$$

If we now use the empty cavity transmission intensity I_{t0} as a reference rather than the incident intensity I_0, we have

$$\frac{I_t}{I_{t0}} = \frac{1}{\left[1 + \alpha L \dfrac{R}{1 - R}\right]^2} \tag{2.22}$$

Since R is only slightly less than one in typical optical cavities for this application, $R/(1-R)$ is large, usually between one hundred and ten thousand, and this large factor multiplies the single pass absorption αL. Hence, a small absorption can have a large effect on the cavity transmission. To illustrate, the cavity transmission falls to one quarter of its maximum value when the single pass absorption αL equals $1-R$. For a numerical example, if the mirror reflectivity R is 99.9% and the cavity length L is 50 cm, this condition occurs for and extinction coefficient α of 2×10^{-5}/cm.

Provided condition (2.20) holds, equations (2.21) and (2.22) are valid for intra-cavity absorption levels sufficient to extinguish the cavity transmission altogether. However, for absorption features that cause relative changes in the cavity transmission that are much less than one, αL must be much less than $1-R$. This allows us to use the binomial expansion to write

$$\frac{I_t}{I_{t0}} \approx \left[1 - \alpha L \frac{2R}{1-R}\right] \qquad (2.23)$$

From this, the relative transmission intensity change caused while scanning over such a weak absorption feature is

$$\frac{\Delta I_t}{I_{t0}} \cong \alpha L \frac{2R}{1-R} \qquad (2.24)$$

Notice this has the same form as for weak direct absorption, equation (2.6) but where the single pass absorption αL has been increased by $2R/(1-R)$.

The common figure of merit for an optical cavity is the finesse, defined as the ratio of the free spectral range to the linewidth of a cavity mode. The mode linewidth, and hence the finesse, are affected by any losses to the cavity mode including mirror transmission and intra-cavity absorption. Since in the final approximation in used to arrive at equation (2.23), i.e., the intra-cavity losses are very small; the cavity finesse is approximately equal to the empty cavity value:

$$F \approx F_{empty} = \frac{\pi \sqrt{R}}{1-R} \qquad (2.25)$$

Since $R \cong \sqrt{R} \cong 1$, by substituting (2.25) into (2.24) we can express the size of small relative fluctuations in cavity transmission as

$$\frac{\Delta I_t}{I_{t0}} \cong \alpha L \times \frac{2F}{\pi} \qquad (2.26)$$

which is the common expression for the enhancement of absorption by a cavity. Similarly to the treatment in Section 2.1, if we now consider the noise equivalent optical

intensity noise fluctuations on the transmitted optical field ∂I_{tN}, then the absorption sensitivity becomes

$$\alpha_S = \frac{\partial I_N}{I_{t0}L} \times \frac{\pi}{2F} \tag{2.27}$$

and similarly, the noise equivalent absorbance sensitivity becomes

$$NEAS = (\sigma NL)_{min} \cong \frac{\partial I_{tN}}{I_{t0}} \times \frac{\pi}{2F} \tag{2.28}$$

Alternatively, the ratio of the noise to the full DC value of the transmission signal now becomes

$$\frac{\partial I_{tN}}{I_{t0}} \cong (\sigma NL)_{min} \times \frac{2F}{\pi}$$
$$= (\sigma NL_{eff})_{min} \tag{2.29}$$

From this we can see that the effect of the resonant cavity is a much-increased effective path length:

$$L_{eff} = L \times \frac{2F}{\pi} \tag{2.30}$$

Note however, that (2.23) through (2.30) are only valid in the limit of small relative changes in cavity transmission, or, for absorption levels much less than the mirror transmission. Calculating the theoretical shot-noise limited performance in the same manner as in Section 2.2, we see that from (2.28) and (2.18) that we have

$$SNEAS = (\alpha_S L)_{sh} = \sqrt{\frac{2e}{i_{DC}}} \times \frac{\pi}{2F} \ /\sqrt{Hz} \tag{2.31}$$

We stress again however, that since we are detecting at low frequencies, this noise level is highly unlikely to be attained due to the presence of 1/f noise, the most common solution for which, is presented in the next section.

2.4 Frequency Modulation and Phase-Sensitive Detection

The sensors we present later in this report harness not only the absorption sensitivity enhancement of optical cavities discussed in Section 2.2, but also frequency modulation (FM) and detection techniques, which will now be discussed in detail. Significant improvements in signal-to-noise are achieved by probing a gas sample with an optical field having an imposed RF frequency modulation. These include immunity to low frequency technical (1/f) noise, and selectivity of features of a particular spectral width.

When the optical field used to interrogate an absorbing medium is frequency or phase modulated, information pertaining to absorption features is encoded on the transmitted field at all harmonics of the modulation frequency to varying degrees. This allows the detection process to be carried out at one of these harmonics rather than at DC. If these RF signals are amplified and then demodulated back to DC (or "base-band") by either using a lock-in amplifier or an RF mixing process, this affords a much-increased signal-to-noise ratio than direct detection. Since the demodulation process is phase-sensitive, access is retained to the phase information in the signal, allowing a choice between absorption and dispersion information from the analyte, the latter cause phase shifts rather than amplitude changes. After some filtering, this yields features (dispersion-shaped or symmetric depending on the technique,) on a relatively flat background, with increased immunity to low frequency noise. This signal recovery technique is known as phase-sensitive or synchronous detection.

There are two principal methods of modeling this modulation and detection process: a time-domain approach, or a frequency-domain approach. In the former, the modulation is modeled by letting the frequency or wavelength of the optical field vary in time, and then calculating the response of the system to these variations. In the latter, the modulated field is broken into its Fourier components, the response of the system applied to each of these components or "sidebands" individually, and the resulting signal calculated. While both these treatments yield the same results, the latter is usually easier computationally, while the former can sometimes provide more insight. We employ the latter technique in this report.

To illustrate the derivation that follows, we refer to Figure 2.3 below, depicting a phase-sensitive FM absorption detector.

FM AND PHASE-SENSITVE DETECTION
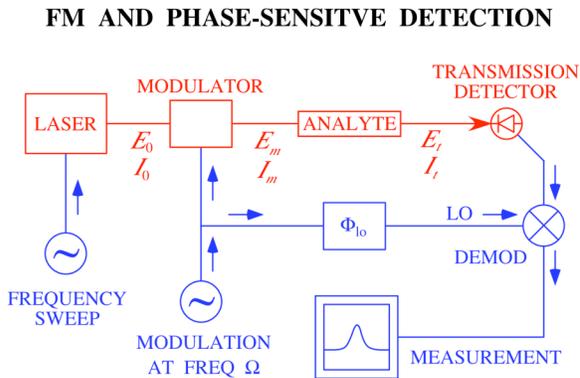
Fig. 2.3. Light from a laser passes through a modulator where it is frequency (or phase) modulated. The light then passes through the analyte and is incident on the transmission detector. The modulation signal is also used to demodulate the detected signal. A phase shifter is used to give the correct local oscillator (LO) phase shift Φ_{lo}. The laser is swept over the absorption features of the analyte and the demodulated signal measured and recorded. Electric field amplitudes and optical intensities are shown for the laser field, modulated field and transmitted fields.

We begin with the unmodulated field exiting the laser, represented by complex optical electric field

$$E_0 = E_{00}\exp(i\omega_0 t) \qquad\qquad (2.32)$$

where E_{00} is the electric field amplitude (taken to be real here) and ω_0 is the angular optical frequency. Only the real part of this quantity is physically meaningful, but retaining the full complex form greatly simplifies algebraic manipulation in the development of the model.

Next, we apply a phase modulation (PM) of index β at a frequency Ω, this frequency being much less than the optical frequency ω_0. This yields the modulated field:

$$E_m = E_{00} \exp\left(i\omega_0 t + i\beta \sin \Omega t\right) \qquad (2.33)$$

(Note that for the purposes of this discussion the presence of residual amplitude modulation (RAM) has been ignored. This has been dealt with in the literature. [10,11]) We can expand this using Bessel functions of the first kind to give

$$E_m = E_{00} \sum_{k=-\infty}^{\infty} J_k(\beta) \exp\left(i(\omega_0 + k\Omega)t\right) \qquad (2.34)$$

This result indicates not just one frequency, but multiple sidebands separated by the modulation frequency Ω. The remaining component at the original unmodulated optical frequency is referred to as the carrier. For applications where the phase modulation index β is around one or less, this infinite series can be truncated to two or three terms including the carrier. However, for the large modulations, and the correspondingly large values of β used in some of our sensors, this is not the case. Generally, approximately β sidebands either side of the carrier must be retained in the expansion of (2.34) to correctly model the modulation. Retaining this many terms is not a problem computationally, although it does not readily produce analytical solutions. We will deal with the case of large modulation index for the rest of this section and the case of small modulation in the next, where we introduce the concepts of resonant sideband detection and NICE-OHMS.

Note that we have begun with a model that modulates the phase (PM) rather than the frequency, even though it is often referred to as FM. This is because FM and PM are two mathematical representations of the same phenomenon. This is easily illustrated by considering a frequency modulation of depth $\partial\omega$ again at a frequency Ω:

$$E_m = E_{00} \exp\left[i\left(\omega_0 + \partial\omega \cos \Omega t\right)t\right] \qquad (2.35)$$

To show the relationship between these two representations, we rewrite the argument of the exponential in equation (2.33) as a cumulative phase function

$$\Phi(t) = \omega_0 t + \beta \sin \Omega t \qquad (2.36)$$

Since frequency is the time-derivative of phase (exactly the angular equivalent of velocity being the derivative of position), the instantaneous frequency $F(t)$ is the derivative of $\Phi(t)$.

$$F(t) = \omega_0 + \beta\Omega\cos\Omega t \qquad\qquad (2.37)$$

Comparing this to equation (2.35), it is immediately obvious that the frequency modulation depth equals the phase modulation index multiplied by the modulation frequency:

$$\partial\omega = \beta\Omega \qquad\qquad (2.38)$$

It is also significant that the modulation function in (2.33) is $\sin\Omega t$ while that in (2.35) is $\cos\Omega t$, meaning that there is a 90-degree phase shift between the two modulation representations. This is to be expected because of the derivative relationship between frequency and phase.

In general, workers in different fields have their own conventions and names for the types of modulations discussed above, and an exhaustive list of these is beyond the scope of this report. To make matters worse, these names may vary depending on the size of the modulation indices relative to the modulation frequencies. This latter distinction is useful however, as it describes different regimes of modulation. To this end, we introduce our own convention below. (For comparison, reference [11] presents a study of various modulation techniques, resulting sensitivities and analysis methods using an alternative nomenclature.)

When the modulation frequency is significantly greater than the linewidths of the features being analyzed, the optimum value of the phase modulation index β approaches a constant value that is of the order of one to two radians depending on the exact demodulation technique being used. The frequency modulation depth $\partial\omega$ however, is not constant, but changes with modulation frequency Ω as seen from the relation (2.38). It also loses any intuitive value since it is the same size or larger than the modulation frequency, which is already significantly greater than the linewidth of the absorption features. In this report, we refer to this as the "PM regime". The nature of detection in this regime is determined by the fact that the sidebands are so far apart that they only interact one at a time with the analyte. The result is that when the local oscillator phase is optimized to obtain the largest signals, we are actually measuring the dispersion of the medium rather than the absorption, the two being related. This is the typical regime for Pound-Drever-Hall (PDH) laser-cavity locking [12-14], and for the resonant sideband detection of NICE-OHMS. [1-3]

On the other hand, when the modulation frequency Ω is less than or of the order of the linewidth of the features being detected, the optimum value of the frequency modulation depth $\partial\omega$ approaches a value that is approximately the same as the linewidth, the exact value depending again on the demodulation technique. The phase modulation index β on the other hand must change according to (2.38) for this to happen, and can consequently take on quite large values (if Ω is small) that have little intuitive meaning, except to fix the number of sidebands in our computational model. Detection in this regime is most sensitive to the change in relative amplitude of the many sidebands as they are scanned through an absorption feature. Alternatively, it is also intuitive to think of the optical

field as scanning the absorption feature at the modulation frequency. We refer to this as the "FM regime". (Note however, that despite the careful definition of these two regimes, unless drawing a specific distinction between them, the terms "frequency modulation" or "FM" will often be used to refer to the general technique.)

These two regimes are illustrated in Figure 2.4, the results of which have been generated by the model we are currently describing. The particular case modeled here is that of an isolated absorption with a Lorentzian lineshape interacting with a modulated optical electric field, the transmission of which through the medium is detected and demodulated at this same frequency. This is known as "1-f" detection. For the curves in this figure, the demodulation phase and the phase modulation index at each frequency value have been optimized for maximum slope of the resulting base-band signal. This calculation was performed by one of us (Taubman) at JILA, the University of Colorado for the purposes of optimizing n-f signals for the purposes of optical clocks [15] for which maximum slope is preferable. Slightly larger signal values are obtained for low modulation frequencies if the modulation index is optimized for maximum signal size rather than maximum slope, but the signal but the resulting curves equivalent to those in Figure 2.4 are discontinuous. This discontinuity is due to significant signal complexity in the transition region between the FM and PM regimes. In any case, the existence and location of these two regimes is not significantly affected by the choice of optimization condition.

The linewidths of the sub-Doppler absorption features we detect (see Section 3.3) are of the order of 700 kHz to 2 MHz. The modulation frequency we use is between 10 and 70 kHz, corresponding to values of Ω from 0.005 to 0.1 FWHM. This places us well in the FM regime to the left of Figure 2.4, and requires values of β between 35 and 200.
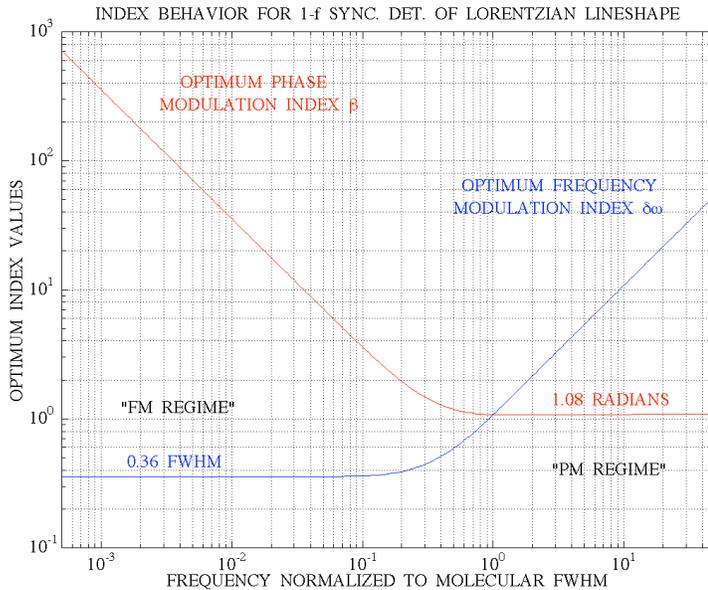
Fig. 2.4. Optimum modulation indexes for phase-sensitive detection of a Lorentzian absorption feature using 1-f detection. At frequencies much below the FWHM linewidth of the absorbing medium, the optimum frequency modulation depth is constant. At frequencies above the linewidth, the optimum phase modulation index is constant. We denote these the PM and FM regimes respectively.

We are now in a position to describe what happens when the modulated optical electric field E_m is passed through the absorbing medium. As mentioned before, we must find the effect of the analyte on each individual sideband of the optical field as it passes through the analyte. To do this, we represent the absorbing medium by a complex transmission function $D(\omega)$, the real part representing the absorption and the imaginary part representing the dispersion of the medium. [16] Because each sideband differs in frequency, the magnitude and phase of this transfer function are different for each sideband. The field transmitted through the medium can thus be written

$$E_t = E_{00} \sum_{k=-\infty}^{\infty} J_k(\beta) D(\omega_0 + k\Omega) \exp(i(\omega_0 + k\Omega)t) \qquad (2.39)$$

In practice, the introduction of $D(\omega)$, be it a complex Lorentzian-, Gaussian- or Voigt-based transfer function usually relegates equation (2.39) to the domain of numerical analysis. To our knowledge, it is only under very limited conditions that models in which the phase modulation index is large have been analytically solved using either Gaussian or Lorentzian lineshapes. [17]

We now turn to the process of detection and demodulation. Our photodetector does not respond to periodic oscillations of the optical electric field, but to average optical intensity, which is given by

$$I_t = \frac{c\varepsilon_0}{2} E_t E_t^* \qquad (2.40)$$

where c is the speed of light, ε_0 is the permittivity of free space and the asterisk represents complex conjugation. At this point, it is customary to convert to photocurrent before continuing the analysis. In fact, to do this analysis properly from an experimental point of view, one should also apply a transimpedance gain to obtain a voltage signal before demodulating. However, since we are interested in signals normalized to the unmodulated field, all these factors will cancel. Consequently, we can normalize the signals while still in terms of optical intensity and work with generic normalized signals that are valid regardless of the physical form. We use (2.32) to find the intensity of the unmodulated field E_0:

$$I_0 = \frac{c\varepsilon_0}{2} E_0 E_0^* = \frac{c\varepsilon_0}{2} E_{00}^2 \qquad (2.41)$$

which is a DC value as expected. For the modulated field E_m the product in (2.40) has the square of the number of terms as the field equation (2.39) because it is the product of two summations. Happily, this is simplified by selecting only those products that oscillate at the frequency at which we wish to demodulate. For example, choosing only those terms that give zero frequency, we get the DC intensity component:

$$I_{t\_DC} = \frac{c\varepsilon_0}{2} E_{00}^2 \sum_{k=-\infty}^{\infty} J_k(\beta)J_k^*(\beta)D(\omega_0 + k\Omega)D^*(\omega_0 + k\Omega) \qquad (2.42)$$

Dividing by equation (2.41) the normalized signal is

$$S_{t\_DC} = \sum_{k=-\infty}^{\infty} J_k(\beta)J_k^*(\beta)D(\omega_0 + k\Omega)D^*(\omega_0 + k\Omega) \qquad (2.43)$$

The quantity S_{t_DC} is actually the normalized sum of the squares of all the field amplitude sidebands transmitted through the medium, and when plotted looks similar to the direct absorption spectrum of the medium as discussed in Section 1.0. It is indeed exactly this when β is set to zero, because

$$J_k(0) = \begin{cases} 1 & \text{for } k = 0 \\ 0 & \text{for } k \neq 0 \end{cases} \qquad (2.44)$$

Notice however, that for non-zero β, the sum is taken over products of the molecular transfer function $D(\omega)$ having different frequency offsets due to the presence of the multiple sidebands of different magnitudes. This effectively convolves the absorption feature with the envelope of the sidebands of the modulated laser field whose width and shape depend on the size of β. The effects of this convolution can range from a slight broadening and flattening, to splitting into multiple features, all of which are readily experimentally observed. In a similar fashion, we can also select only those terms giving signals at Ω, yielding the normalized 1-f component.

$$S_{t\_\Omega} = \sum_{k=-\infty}^{\infty} J_{k+1}(\beta)J_k^*(\beta)D(\omega_0 + (k+1)\Omega)D^*(\omega_0 + k\Omega)\exp(i\Omega t) \qquad (2.45)$$

Notice this is still a complex quantity and is "rotating" at frequency Ω. This is the normalized sum of products (or beats) of all pairs of adjacent sidebands. Similarly, we can extend this to detection at 2Ω, 3Ω or in fact generalize it to $n\Omega$ for any positive integer n.

$$S_{t\_n\Omega} = \sum_{k=-\infty}^{\infty} J_{k+n}(\beta)J_k^*(\beta)D(\omega_0 + (k+n)\Omega)D^*(\omega_0 + k\Omega)\exp(in\Omega t) \qquad (2.46)$$

Note also that if we set $D = 1$ corresponding to zero absorption and dispersion or equivalently the absence of an analyte, this reduces to

$$S_{t\_n\Omega}^{D=1} = \sum_{k=-\infty}^{\infty} J_{k+n}(\beta)J_k^*(\beta)\exp(in\Omega t) = \begin{cases} 1 & \text{for } n = 0 \\ 0 & \text{for } n \neq 0 \end{cases} \qquad (2.47)$$

where we have used the identity [18]

$$\sum_{k=-\infty}^{\infty} J_{k+n}(\beta) J_k^*(\beta) = \begin{cases} 1 & \text{for } n = 0 \\ 0 & \text{for } n \neq 0 \end{cases} \qquad (2.48)$$

This says that if there is no absorption or dispersion feature, the only term that the detector sees is a DC term. Moreover, this DC term is exactly the same as that of the unmodulated optical intensity, since the normalized value is unity. In other words, without an analyte or some other frequency discriminator, FM makes no difference to the detection process. In the absence of anything to change their relative magnitudes or phases, all the beat contributions from the sidebands cancel out on the detector and all the DC components add up to the original unmodulated value. On the other hand, when the individual sidebands are perturbed differently by the action of the absorbing medium either in amplitude or in phase, this gives rise to a non-zero signal at harmonics of the modulation frequency. The overall power in the optical field is conserved by a corresponding reduction in signal at DC. This is in fact a "discrimination" process, converting FM (or PM) to amplitude modulation (AM) at the modulation frequency, which is detectable. This is exactly why RAM (introduced in Section 2.2) on the laser field interferes with the detection of absorption features in the analyte: it is directly detectable without interacting with the analyte and hence produces offsets and false signals. Furthermore, the signals due to RAM are directly proportional to the laser intensity and consequently couple low frequency intensity fluctuations (1/f noise that we've been avoiding by using FM techniques) back into the demodulated signal.

We now demodulate this signal. Since we are interested in real quantities, we need to sum all terms rotating at both $n\Omega$ and at $-n\Omega$. We then multiply this sum by a real local oscillator (LO) function of frequency $n\Omega$ and phase Φ_{lo} relative to the original modulation. We choose this function to be $\sin(n\Omega t + \Phi_{lo})$ for convenience.[a] The resulting normalized base-band signal is

$$
\begin{aligned}
X_{n\Omega} &= \left[S_{t\_n\Omega} + S_{t\_n\Omega}^* \right] \sin(n\Omega t + \Phi_{lo}) \\
&= 2\operatorname{Re}(S_{t\_n\Omega}) \sin(n\Omega t + \Phi_{lo}) \\
&= 2|A_n| \cos(n\Omega t + \arg A_n) \sin(n\Omega t + \Phi_{lo}) \\
&= |A_n| \sin(\Phi_{lo} - \arg A_n)
\end{aligned}
\qquad (2.49)
$$

where $\qquad A_n = \sum_{k=-\infty}^{\infty} J_{k+n}(\beta) J_k^*(\beta) D(\omega_0 + (k+n)\Omega) D^*(\omega_0 + k\Omega) \qquad (2.50)$

Additive terms at twice the optical frequency have been ignored (rotating wave approximation) since they are not detected by the photodetector. Equation (2.49) is

---

easily evaluated for a variety conditions using a computer. Figure 2.5 shows typical traces for 1-f, 2-f and 3-f detection calculated using this model for an absorption feature of Lorentzian line shape. In this case, unlike that of the calculations for Figure 2.4, these features have been optimized for maximum size.[a]

The value of modulation frequency Ω of 0.05 times the FWHM Lorentzian linewidth was chosen, which puts us well into FM regime as discussed previously and as indicated in Figure 2.4. The unmodulated direct absorption feature is also included for comparison, and the relative sizes are written on the plot. These calculated values are important for calculating the impact of the modulation technique on the sensitivity α_S and the NEAS of a chemical sensor using these FM techniques. Clearly the higher the demodulation order, the less signal is recovered. However, this reduction is often greatly outweighed by the advantages such as increasing immunity to low frequency noise, and removal of sloping background effects.
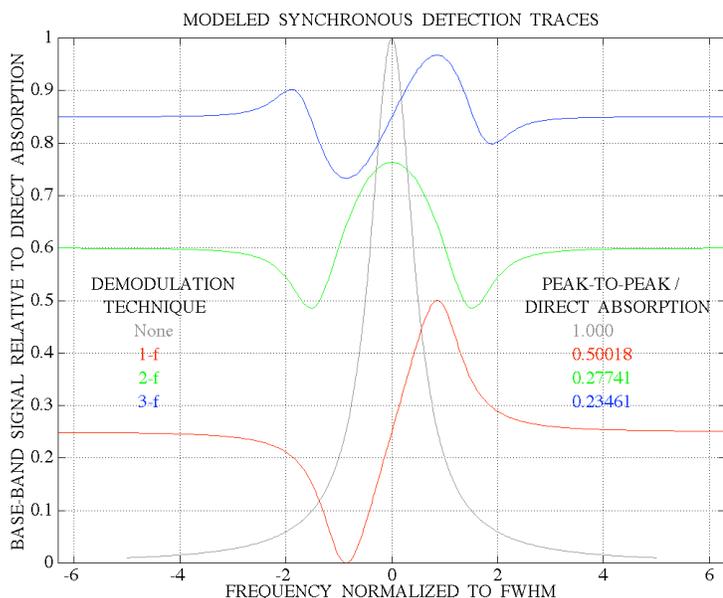
Fig. 2.5. Theoretical absorption signals using 1-f, 2-f and 3-f demodulation, compared to the absorption signal of direct detection. Features are optimized for maximum size. The frequency is scaled in units of molecular FWHM. The vertical axis is normalized to the direct absorption feature. A peak absorption of 1% chosen for this calculation also corresponds to the vertical scale. The line shape used is Lorentzian, and the Modulation frequency is $\Omega = 0.05$ FWHM.

Using this model, we can now evaluate trends in the modulation and recovery performance. Figure 2.6 below shows how the recovered signal size changes as the modulation frequency is varied relative to the linewidth of the absorption feature we are measuring. This can also be viewed as the sensitivity of the technique to features of different spectral widths for a given modulation frequency and frequency modulation depth.

(a) Since the calculation involves only one chosen frequency, the continuity problems discussed regarding optimizing Figure 2.4 for maximum signal are not an issue here.

SIGNAL RECOVERY VS FREQUENCY MOD DEPTH

EXPERIMENTAL CONDITIONS
MOD. FREQ $\Omega = 0.05$ MOL. FWHM
PHASE MOD. DEPTH $\beta = 7.1269$

(y-axis) PEAK-TO-PEAK 1-f SIGNAL RELATIVE TO DIRECT ABSORPTION

(x-axis) FREQUENCY MOD DEPTH NORMALIZED TO MOLECULAR FWHM
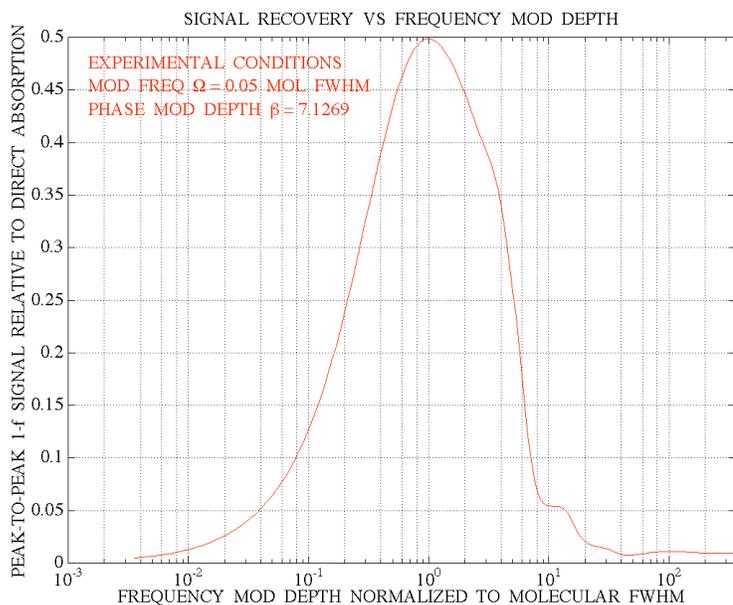
Fig 2.6. Level of signal recovered using 1-f demodulation relative to that of direct absorption, plotted versus modulation frequency. The chosen parameters are Ω = 0.05 FWHM and β = 7.13 radians, giving $\partial\omega = 0.36$ FWHM. Demodulation phase Φ_{lo} is optimized at all frequencies. The result is that features both much narrower and broader than the feature of interest are attenuated.

The response is optimized for the modulation frequency of 0.05 times the FWHM Lorentzian linewidth. As the modulation frequency is varied, the detection phase is optimized to observe the optimum signal, but the phase modulation frequency is kept constant. This clearly demonstrates the ability of phase-sensitive detection to favor absorption features of a specific size, although we do point out that the frequency scale is logarithmic. Nevertheless, we shall see a direct benefit of this effect in our experimental results in Section 3.3 where a larger Doppler profile is suppressed while a much smaller sub-Doppler feature is optimally demodulated. An advantage of higher order detection techniques, such as 2-f and 3-f demodulation, is that they have even greater selectivity. There is an optimum choice however, which is often 1-f detection.

2.4.1 Combining Optical Cavities and Modulation Techniques

In this section, we have modeled FM and phase-sensitive detection as applied to an isolated gas sample or analyte. However, as mentioned before, the chemical sensors discussed in this report combine the absorption sensitivity enhancement of optical cavities with the noise-reduction benefits of the FM techniques described here. The specific applications of these techniques we use are those of cavity-dither modulation and resonant sideband detection. The latter technique is the distinguishing feature of NICE-OHMS and is discussed separately in Section 2.5. The former is also used in the full NICE-OHMS sensor in as well as in the more simple FM cavity-enhanced sensors.

We depict the cavity-dither technique below in Figure 2.7. The frequency of the laser and that of the corresponding longitudinal cavity mode (discussed in Section 2.2) are moved rapidly or "dithered" together in frequency. Ideally, during this process the laser frequency remains coincident with the peak of the cavity mode, and in fact the two are actually locked together to facilitate this as discussed in Section 3.1. This has the effect

CAVITY DITHER MODULATION
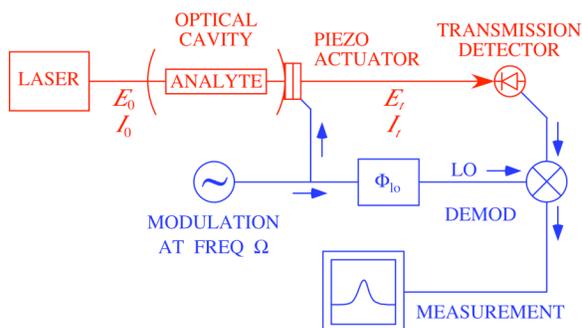
Fig. 2.7. Cavity-dither arrangement. The cavity length is modulated by a small amount using a piezo actuator on one mirror. The laser tracks the resulting frequency dither because of its lock to the optical cavity mode. (Explained in Section 3.1) Demodulation occurs exactly as described for Figure 2.3. The electronics used to scan the cavity and laser across absorption features in the analyte are omitted for simplicity at this stage.

of allowing the intra-cavity analyte to be exposed to an optical field that is more deeply frequency-modulated than the cavity linewidth, while maintaining the sensitivity advantages of the optical cavity.

The combining of an optical cavity with frequency modulation in this way causes several complications. Firstly, as will be seen in Section 3.1, the way the frequency scan is now performed consists of scanning the optical cavity length, and letting the laser (locked to the cavity mode) follow this length scan. In an empty cavity, the frequency of the mode is inversely proportional to its length. However, we are scanning across an absorption feature. As mentioned above when we introduced the transfer function of the molecular absorption feature $D(\omega)$, and as discussed in reference [16], such a feature introduces dispersion as well as absorption. The result is that effective path length changes as the cavity mode is scanned across the feature. Consequently, the cavity length scan no longer corresponds to a linear frequency scan, but is distorted in the region of the molecular absorption. This in turn means that the resulting spectra do not have a linear frequency scale. This can be readily modeled by introducing a correction function in the spectrum after the absorption has been calculated. This is a theory task for FY03.

A more difficult problem arises in such a cavity-dithered sensor. The dynamics and control of an optical cavity that is rapidly changing in length can become very complicated if the mirror being dithered is moving fast enough such as to cause a Doppler shift of the order of the cavity linewidth. [19] The classic results for circulating, reflected and transmitted fields [8] are based on the assumption of a static optical cavity. If the cavity length is modulated this rapidly, the static cavity assumption breaks down. In the worst case, the various cavity signals including those used to lock the laser and cavity together begin to oscillate, making the control of the cavity difficult or impossible. A complete theory for a cavity-dithered absorption sensor has not been developed to the best of our knowledge. An investigation of these complications is also a task for FY03.

However, the model of FM and phase-sensitive detection introduced in this section in the absence of an optical cavity allows a good understanding of our current experimental results, and allows us to approximate the theoretical performance of our FM cavity-enhanced chemical sensors. We can approximate the shot noise-limited performance of large modulation index cavity-dither signal recovery sensors for each of 1-f, 2-f and 3-f

demodulation techniques using our calculations above. There is one other factor entering into the shot noise calculation at this stage. Since this particular noise source appears as photons are converted to photocurrent by the detector (assumed to be noiseless), both the noise and signal are amplified by the transimpedance gain and other gains before or after the demodulation process. Consequently, it is the ratio of the shot noise to the root-mean-square (RMS) signal that matters, multiplicative factors down stream of this make no difference to the shot noise-limited performance. Since, unlike the direct absorption sensor, we are now dealing with a signal oscillating at a certain frequency, we pick up a factor $1/\sqrt{2}$ corresponding to the RMS magnitude of the oscillating function. The resulting shot-noise limited performance is given by

$$\left(SNEAS\right)_{\text{FM}} = \left(\alpha_S L\right)_{sh} = \sqrt{\frac{2e}{i_{DC}}} \times \frac{\pi}{2F} \times \frac{\sqrt{2}}{K} \ /\sqrt{\text{Hz}} \qquad (2.51)$$

where
$$K = \begin{cases} 0.500 & \text{for 1-f} \\ 0.277 & \text{for 2-f} \\ 0.234 & \text{for 3-f} \end{cases} \qquad (2.52)$$

This can also be used for phase-sensitive detection without an optical cavity by removing the factor of $2F/\pi$.

2.5   Resonant Sideband Detection and NICE-OHMS

Resonant sideband detection is the technique on which NICE-OHMS is based and is depicted below in Figure 2.8. This is very similar to Figure 2.3 except that an optical cavity is placed around the analyte, and the modulation frequency is set to be equal to the cavity FSR, setting the interval between adjacent modulation sidebands to equal that between adjacent cavity modes. When locked to the cavity, both the carrier and sidebands of the laser field pass into the cavity, thus allowing phase-sensitive detection to be carried out while benefiting from the sensitivity enhancement of the optical cavity.
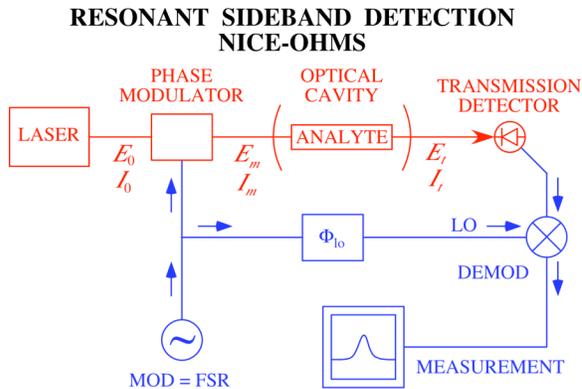
Fig. 2.8. Resonant sideband modulation with phase-sensitive detection. Light from a laser passes through a phase modulator where sidebands are added to the optical field at a frequency equal to the FSR of the cavity that follows. The carrier of the circulating cavity field interacts with the analyte, the resulting field beating with the sidebands to produce AM signals on the transmission detector.

Except for extremely long optical cavities, the FSR is MHz to GHz. For example, a 38.5 cm long cavity has a FSR of 389.3 MHz. In the LWIR, Doppler broadened absorption features of typical analytes at room temperature have widths of around 70 MHz, and the sub-Doppler features of interest at most 2 MHz. Since the modulation frequency is larger than the linewidth, this puts us into the PM regime as defined in Section 2.4, resulting in small optimum phase modulation depth β. Consequently, only one spectral component of the modulated field can interact with a given absorption feature at any one time during a scan, making dispersion and phase shift the key mechanisms for producing the signal. The effect of dispersion is to change the optical path length via a small change in the refractive index. A field traveling a fixed physical length will then see a changing optical path length and hence suffer a phase change as a dispersion feature is encountered. Since this is occurring inside an optical cavity, this optical path length change is multiplied by the number of effective round trips of the cavity, greatly enhancing this effect. Another way to view this is that the presence of the dispersion actually shifts the cavity mode in frequency, the resulting detuning with the optical carrier (or sideband) producing a phase shift and thus an AM signal. This is illustrated below in Figure 2.9.
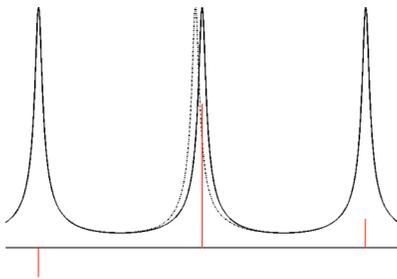
Fig. 2.9. Carrier and Sidebands correspond to adjacent modes of the optical cavity. The presence of a molecular resonance alters optical path length within the cavity, thus shifting or "pulling" the frequency of the cavity mode. The resulting detuning from the corresponding component of the optical field causes a phase change to that component causing a detectable signal. Common mode amplitude and phase changes to all sidebands due to noise between the laser and cavity do not produce detectable signals.

NICE-OHMS [1-3] stands for Noise-Immune Cavity-Enhanced Optical Heterodyne Molecular Spectroscopy. The "noise-immune" part of the name refers to relative frequency noise between the laser and the optical cavity. Such noise can be detrimental in cavity-enhanced sensors, which is why frequency stabilization and cavity locking are so important as discussed in Section 3.1. A change in this relative frequency will change the amplitude and phase of the sidebands both passing through and reflected from the cavity. This is one reason why optical cavities are so useful; they are in essence very sensitive frequency or phase discriminators. However, differential changes to the sidebands of a frequency or phase modulated field will produce amplitude modulation as discussed in section 2.4, thus producing detectable signals. In resonant sideband modulation however, the carrier and sidebands are all passed into the cavity via identical adjacent cavity modes. Any fluctuations due to relative frequency noise between the laser and the cavity will thus be common to all sidebands, and not produce a signal. Changes to the phase or amplitude of one sideband in particular however, such as the dispersion of an absorption feature of the analyte as discussed above, will produce a signal. To finish the discussion of the acronym NICE-OHMS, the terms "cavity-enhanced" and "molecular spectroscopy" are obvious, and the term "optical heterodyne"

refers to the fact that we rely on the beat between the carrier and the high frequency sidebands in the resonant sideband modulation arrangement.

To derive the sensitivity of this technique we consider again FM and phase-sensitive detection without an optical cavity. For this purpose, we refer back to Figure 2.3, which is essentially Figure 2.8 without the cavity. Since we are in the PM regime as discussed above, we consider the limit in which the modulation frequency is large compared to the linewidth of the molecular absorption feature of interest in the analyte. Again, we have the equation for the field, phase modulated at frequency Ω with depth β

$$E_m = E_{00} \exp\left(i\omega_0 t + i\beta \sin\Omega t\right) \qquad (2.53)$$

which we expand using the Bessel functions of the first kind to give

$$E_m = E_{00} \sum_{k=-\infty}^{\infty} J_k(\beta) \exp\left(i(\omega_0 + k\Omega)t\right) \qquad (2.54)$$

The optimum modulation index for this regime is about 1.06 radians. This means we can truncate the series to consider only the carrier and the two first order sidebands either side of it.

$$E_m = E_{00}\left\{J_0 \exp(i\omega_0 t) + J_1\left[\exp\left(i(\omega_0 + \Omega)t\right) - \exp\left(i(\omega_0 - \Omega)t\right)\right]\right\} \qquad (2.55)$$

where we have dropped the explicit appearance of β, and used the fact that $J_{-1}(\beta) = -J_1(\beta)$. As mentioned above, because the modulation frequency is large only one spectral component of the field interacts with the analyte at any one time. We choose that to be the carrier, although similar effects are obtained if either sideband is chosen. We can thus model the effect of the analyte as producing a phase change ϕ to the optical carrier and write the field transmitted through the analyte as

$$E_t = E_{00}\left\{J_0 \exp(i\omega_0 t + i\phi) + J_1\left[\exp\left(i(\omega_0 + \Omega)t\right) - \exp\left(i(\omega_0 - \Omega)t\right)\right]\right\} \quad (2.56)$$

The resulting normalized intensity incident on the transmission detector is

$$\begin{aligned}
S_t &= E_t E_t^* / E_{00}^2 \\
&= J_0^2 + 2J_1 + 2J_0 J_1\left\{\cos(-\Omega t + \phi) - \cos(\Omega t + \phi)\right\} \qquad (2.57) \\
&+ \text{terms of order } 2\omega_0
\end{aligned}$$

The first two terms form the DC contribution, which is seen decrease slightly as β is increased corresponding to power being taken from the carrier and put into the sidebands, and seen to return to unity in the absence of modulation. The terms oscillating at twice the optical frequency are ignored as usual. Similar to the analysis in Section 2.4, we select only those terms oscillating at Ω, and obtain the signal

$$S_{t\_\Omega} = 2J_0J_1\{\cos(\Omega t - \phi) - \cos(\Omega t + \phi)\}$$
$$= 4J_0J_1\sin\Omega t\sin\phi \qquad (2.58)$$
$$\approx 4J_0J_1\sin\Omega t \times \phi$$

where we have assumed that the dispersive phase shift will be small and that $\sin\phi \approx \phi$. Demodulating this with a local oscillator, we obtain

$$X_\Omega = 4J_0J_1\phi\sin(\Omega t)\sin(\Omega t + \Phi_{lo}) \qquad (2.59)$$

We can see by inspection that the optimum detection phase Φ_{lo} is going to be zero, giving

$$X_\Omega = 4J_0J_1\phi\sin^2\Omega t$$
$$= 2J_0J_1\phi(1 - \cos 2\Omega t) \qquad (2.60)$$

Filtering out all but the DC term gives

$$X_\Omega = 2J_0J_1\phi \qquad (2.61)$$

A comparison with the derivation shown in Section 2.4 will show that the above analysis to be identical, except that here it is performed for a specific case. To illustrate, look at equation (2.49) reproduced below for 1-f detection:

$$X_\Omega = |A_1|\sin(\Phi_{lo} - \arg A_1) \qquad (2.62)$$

where,

$$A_1 = \sum_{k=-\infty}^{\infty} J_{k+1}(\beta)J_k^*(\beta)D(\omega_0 + (k+1)\Omega)D^*(\omega_0 + k\Omega)$$
$$= J_0J_{-1}\exp(i\phi) + J_1J_0\exp(-i\phi)$$
$$= J_0J_1(\exp(-i\phi) - \exp(i\phi)) \qquad (2.63)$$
$$= -2iJ_0J_1\sin\phi$$
$$\approx -2iJ_0J_1\phi$$

giving

$$|A_1| = 2J_0J_1\phi$$
$$\arg A_1 = -\frac{\pi}{2} \qquad (2.64)$$
$$\text{and thus} \quad X_\Omega = 2J_0J_1\phi$$

Consequently, the model we developed in Section 2.4 can also be used to plot the expected base-band feature for this regime, shown below in Figure 2.10. This detection process has no selectivity of features according to their spectral width because the modulation frequency is so much larger than the linewidth. Consequently, features such as this will often appear on steep backgrounds due to other effects or broader features in the medium. Consequently, in the full NICE-OHMS experiment, this resonant sideband detection technique is combined with the cavity-dither technique described in section 2.4. Since the signals are already amplified after detection at high RF frequencies, this second modulation/demodulation process affords no more noise immunity, but serves only to increase selectivity. Using this dual modulation technique, the resulting features will resemble the 2-f detection shown in Figure 2.5. A further reduction in signal size of 0.5 will result, but no reduction in signal-to-noise ratio will be suffered.
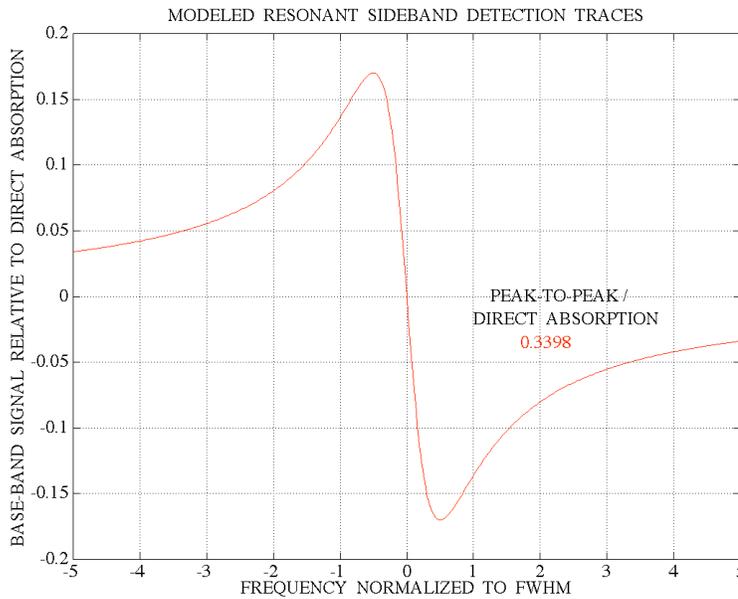
Fig. 2.10. The line shape produced by detection of an absorption feature in the PM regime where the modulation frequency is large compared to the FWHM of the feature.

We next calculate the shot noise-limited performance of the NICE-OHMS system. The phase change caused by the analyte can be written in terms of change in refractive index

$$\phi = \Delta n \frac{\omega_0 L}{c} \tag{2.65}$$

where the peak change in refractive index is related to that of the extinction coefficient via the Kramers-Kronig dispersion relations [16]

$$(\Delta n)_{pk} = \alpha_{pk} \frac{\lambda}{4\pi} \tag{2.66}$$

from which we obtain a relationship between the peak values of ϕ and α:[a]

$$\phi_{pk} = \frac{\alpha_{pk}L}{2} \tag{2.67}$$

Taking the RMS value of (2.58) (giving a division by $\sqrt{2}$ and the removal of the oscillating factor,) at the peak of the feature appearing at the detector, we obtain

$$(S_{t\_\Omega})_{rms} = 2\sqrt{2}J_0 J_1 \phi \tag{2.68}$$

The minimum detectable phase change due to the shot noise limit is then found to be

$$\phi_{sh} = \sqrt{\frac{2e}{i_{DC}}} \times \frac{1}{2\sqrt{2}J_0 J_1} \tag{2.69}$$

Using (2.67) we then find that the shot noise-limited performance of phase sensitive detection of an absorption feature in the PM regime is

$$\left(SNEAS\right)_{PM} = \left(\alpha_S L\right)_{sh} = \sqrt{\frac{2e}{i_{DC}}} \times \frac{1}{\sqrt{2}J_0(\beta)J_1(\beta)} \ /\sqrt{Hz} \tag{2.70}$$

Adding an optical cavity to this, simply increases the sensitivity by the effective path length enhancement factor as derived in Section 2.3 of $2F/\pi$, giving the full NICE-OHMS shot noise performance to be

$$\left(SNEAS\right)_{NO} = \left(\alpha_S L\right)_{sh} = \sqrt{\frac{2e}{i_{DC}}} \times \frac{\pi}{2F} \times \frac{1}{\sqrt{2}J_0(\beta)J_1(\beta)} \ /\sqrt{Hz} \tag{2.71}$$

It should be noted that at present, there is a disagreement of a factor of 2 between our derivation and that of the authors [1], which is currently unresolved.

(a) Strictly speaking this relationship is only as stated here for zero saturation; it varies by a numerical factor as the saturation parameter of the interaction changes. This is ignored for the purposes of this derivation.

3.0   Cavity-Enhanced Sensors in the LWIR Using QC Lasers

The theory has now been covered for direct absorption sensors, unmodulated cavity-enhanced sensors, FM signal recovery sensor with both large and small modulation depths, and finally NICE-OHMS sensors with resonant sideband detection. We now turn to experimental considerations results and details. We begin with laser stabilization and cavity locking, which itself involves modulation and demodulation techniques. We then describe the various sensors investigated over FY02 and describe the results. This section culminates in a discussion of the NICE-OHMS sensor operating at PNNL, and a detailed calculation of the ultimate sensitivity obtained.

3.1   Laser Stabilization and Locking to Optical Cavities

In order to gain many of the benefits of the optical cavity, the laser frequency must be locked or stabilized to one of the cavity modes. This is done using the Pound-Drever-Hall (PDH) technique. [12-14] This is depicted in Figure 3.1, showing a laser, a phase modulator and an optical cavity, detectors monitoring both the transmitted and the reflected field from the cavity, and the associated optics and the necessary electronics.

An RF phase modulation, in our case at a frequency between 10 and 50 MHz, is applied to the optical field before it is incident on the optical cavity. Note that this modulation is not the same as any discussed previously in this report. Because the frequency of modulation is chosen to be considerably larger than the cavity linewidth, but less than half the cavity FSR, these sidebands are reflected from the cavity. Changes in the relative frequency between the laser field and the cavity mode cause changes in the phase and amplitude of the carrier of the optical field. This modified reflected carrier beats with the reflected sidebands on the reflection detector. Since the sidebands are a long way from a cavity resonance, they remain relatively unaffected by changes in the relative frequency of the laser and cavity. Consequently changes in the phase[a] of the reflected carrier cause the conversion of the incident PM to AM at the modulation frequency, which is then demodulated in exactly the same way as other signals described in this report (see Figures 2.3, 2.7 and 2.8), to produce what is known as the PDH error signal shown in Figure 3.2. This error signal is passed to stabilization electronics to adjust the laser drive current to keep its frequency coincident with the cavity mode. As long as it remains locked, the laser now tracks the frequency of that particular cavity mode. The frequency of the sensor can now be swept by changing the cavity length with a piezo-electric element attached to one of the cavity mirrors, illustrated in Figure 3.1. The transmission signal now responds to changing cavity parameters, in particular to intra-cavity absorption, as the frequency is scanned.

(a)   Note again that the modulation frequency is much larger than the linewidth of the cavity, placing this case in the PM regime, making detection sensitive to phase.

LASER / CAVITY LOCKING
PDH TECHNIQUE
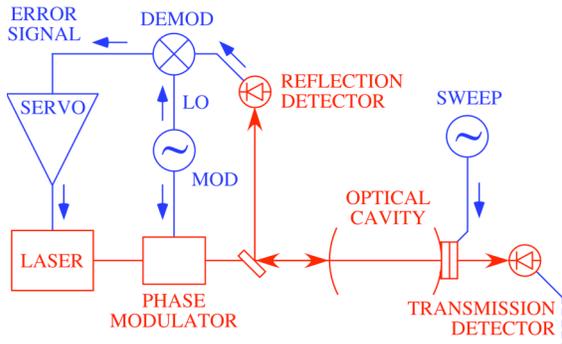
Fig. 3.1. Pound-Drever-Hall (PDH) locking scheme, showing a laser locked to a mode of an optical cavity. Phase modulation sidebands are imposed on the laser field incident on an optical cavity. The reflected field is sampled, detected and then demodulated using an RF local oscillator (LO) at the modulation frequency. The resulting "error" signal is passed to a servo system, which in turn controls the laser frequency, forming a closed loop. Filtering of the signal, phase shifting the LO, and other details are left out for simplicity.
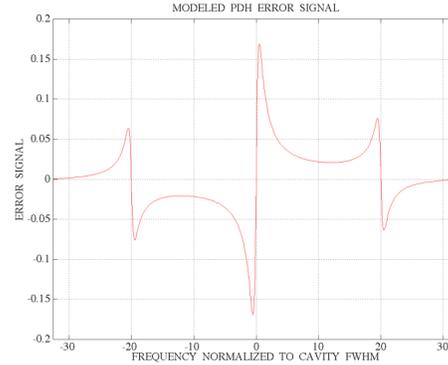
Fig. 3.2. Typical error signal seen after demodulation of the detected reflection signal in a PDH locking scheme depicted in Figure 3.1. The middle (zero-point) of the steep central dispersion-shaped feature is the normal target for the servo system, which maintains the relative frequency of the laser and cavity such that the error signal stays at this point. Servo systems often have gains of 100s of thousands at low frequencies, forcing the laser and cavity to be tightly locked from DC up to a considerable bandwidth, which in our case is around 1.5 MHz.

The process of locking QCLs to optical cavities is a significant component of chemical sensors in the LWIR under development at PNNL. The free-running QCL linewidth is around 160 kHz [20], which is similar to the cavity mode linewidth. The difficulty is that the laser frequency must be tightly constrained within a small region around the peak of the cavity mode - in practice to between 1% and 0.1% of the linewidth. As mentioned previously, this is because an optical cavity behaves like a sharp frequency discriminator. Differential frequency fluctuations between the laser and the optical cavity are converted directly into amplitude fluctuations and can thus be seen on the transmitted signal. This is especially true if the laser frequency is tuned to the side of the cavity resonance as illustrated in Figure 3.3.

To avoid this noise conversion process, the QCL is tightly locked to the peak of the cavity mode. In related experiments, [21] after optimization of the cavity locking the residual fluctuations have been calculated to yield a laser linewidth of less than 1 Hz, and have been measured with heterodyne beat techniques to yield a linewidth of approximately 5.6 Hz. The experimental work performed at PNNL leading to the demonstration of cavity-stabilized QCLs and in particular the above results, was funded by the Defense Advanced Research Projects Agency's Microtechnology Office, DARPA/MTO.
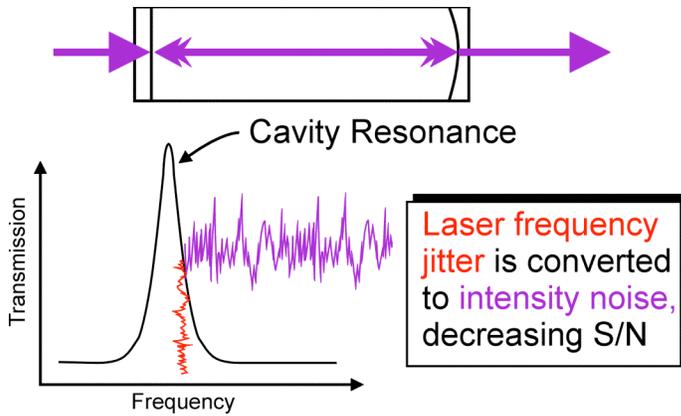
Fig. 3.3. If the laser frequency drifts to a point on the side of a cavity mode rather than the peak, the steep discriminator slope converts relatively small frequency fluctuations into large amplitude fluctuations in the cavity transmission. This interferes directly with cavity absorption measurements.

3.1.1 Additional Effects of RAM

Residual amplitude modulation (RAM) was introduced in Section 2.2 as the small amount of AM present as a result of imposing FM onto an optical field, and that the principal impact of this is to re-couple 1/f noise back into measurements, which would have otherwise been avoided by the FM technique itself. This is because RAM produces signals unrelated to the analyte on the detector that are proportional to the laser intensity, and hence the laser noise. This was further illustrated as the FM theory was developed in Section 2.4. Here, we consider an additional effect specifically relating to locking and stabilization, which in turn aggravates this problem.

RAM often produces an offset in a signal produced by phase-sensitive detection such as the locking error signal above in Figure 3.2. When used in a locking servo system, this offset causes the QCL to be locked to the side of the cavity resonance rather than the peak, resulting in excess transmission noise as illustrated in Figure 3.3. If this effect were constant, it could easily be compensated and removed. In general however, there are many reasons why these offsets vary. In our case, the principal reason is optical fringing also introduced in Section 2.2. Undulations in the signal obtained from the reflection detector in Figure 3.1 cause periodically varying offsets in the error signal, and thus periodic variations in the lock point to the cavity mode as the frequency of the sensor is scanned. This results in excess noise appearing periodically with frequency in sensor measurements.

The effects of RAM described above are dealt with in two ways. The first of these is to attempt to select a modulation frequency for the stabilization, at which the relative phase of the RAM is such that its effect is at a minimum when the local oscillator phase is optimized to that required by the PDH error signal. This is not always possible, and varies laser to laser. The second method is to reduce the optical fringing by carefully aligning the optics to do so. Neither of these techniques however will eliminate this effect altogether.

3.2 Simple Cavity-Enhanced Sensor.

The first experimental chemical sensor we present is a simple cavity-enhanced absorption sensor (Figure 3.4) as discussed theoretically in the previous sections.

SIMPLE CAVITY-ENHANCED SENSOR
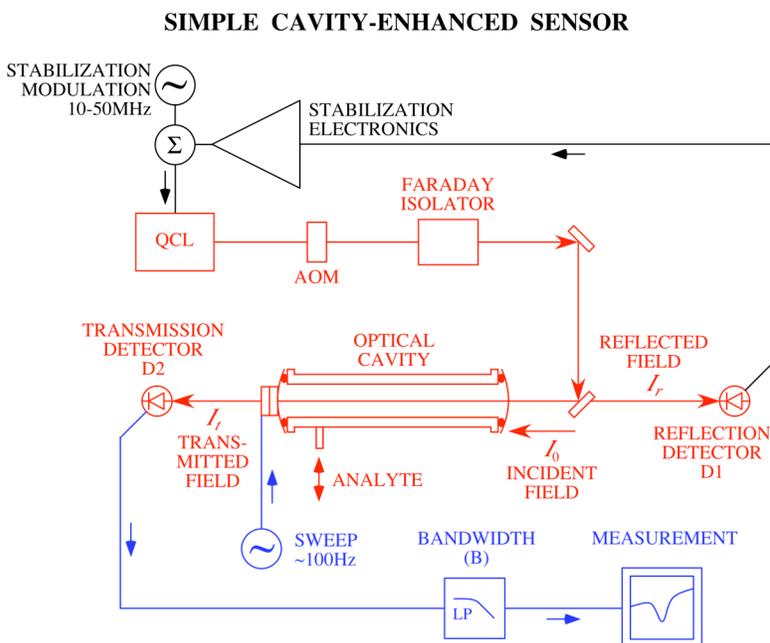
Fig. 3.4. Cavity-Stabilized QCL used as a chemical sensor. Absorption due to the intra-cavity analyte causes dramatic changes in the cavity transmission. Optical elements are in red, stabilization in black, measurement electronics in blue. Incident, reflected and transmitted cavity fields of intensities I_0, I_r and I_t are indicated.

The optical field from a QCL is coupled into an optical cavity via an acousto-optic modulator (AOM) and a Faraday isolator in order to minimize the optical feedback into the QCL from back reflections. Reflected and transmitted light from the cavity are observed using detectors D1 and D2 respectively. The QCL is stabilized to the optical cavity using the PDH technique explained in Section 3.1, the electronics for this being compressed into one block called "stabilization electronics." (For more details, see Figure 3.1.) The optical cavity is made from vacuum fittings forming a chamber allowing an analyte to be introduced at low pressures. A piezo electric element in contact with the transmission mirror allows the optical cavity length to be scanned, thus changing the cavity mode frequencies and that of the QCL since it is locked to one of these modes. As the frequency of the sensor is scanned, the signal from detector D2, is filtered at some bandwidth B, measured on an oscilloscope and then recorded.

An example of the signals we see from this detector is shown in Figure 3.5. The reflectance of the cavity mirrors used for these results was 99.87%, giving an empty cavity finesse of 2415. The resulting effective cavity path length is 592 meters, but is not applicable here, since the absorption value is large compared to the mirror transmission as we shall see below. The traces in this figure show the intensity of the cavity transmission versus time in milliseconds as the piezo element is scanned. The two features in the center of the scans in Figure 3.5 correspond to two nitrous-oxide (N_2O) absorption lines at 1174.8283 cm^{-1} and 1174.8333 cm^{-1}. Traces taken at different pressures are depicted in different colors. The line strengths S (in cm/molecule) of these
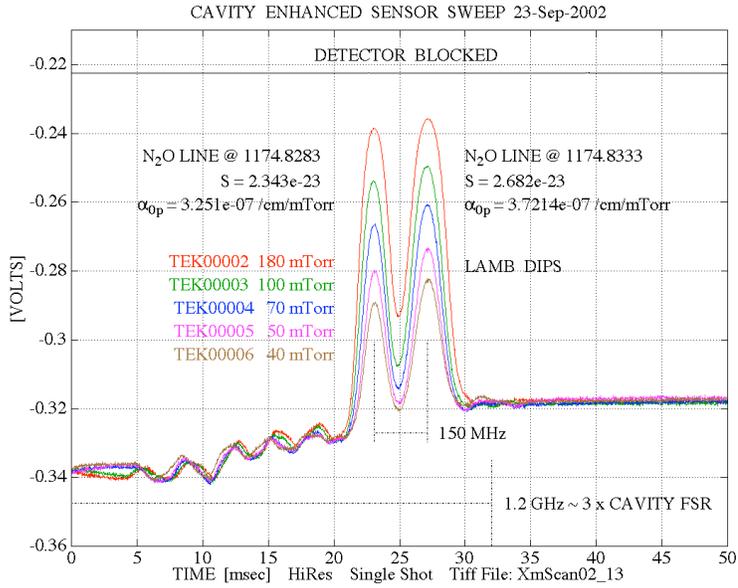
Fig. 3.5. Sample output from the Simple Cavity-Enhanced Sensor shown in Figure 3.4. Two adjacent absorption lines in N_2O are shown. Traces taken at different analyte pressures are shown in different colors.

absorption features as listed in the HITRAN database, and the consequent pressure-dependent absorption coefficients at Doppler line center α_{0p} (in cm^{-1}/mTorr) are shown on the diagram. The relationship between these two quantities as will be derived in Section 3.4, is

$$\alpha_{0p} = \frac{Sc}{\delta v_D} \times \frac{1.2523 \times 10^{-7}}{kT} \qquad (3.1)$$

where c is the speed of light in cm/s, δv_D is the Doppler width of the molecular transition in Hz, T is temperature in Kelvins and k is Boltzmann's constant.

Consider the first absorption feature for the data taken at a pressure of 180 mTorr, measured using a Convectron gauge. Assuming an error in this pressure measurement of 10%,[a] the single pass peak absorption, $\alpha_{0p}PL = \sigma NL$ (see Section 1), is .0022±0.0002 [22] for a 38.5 cm cavity. This is of the same order as our measured value of $T = 1 - R$ of 0.0013. As discussed in section 2, this means we must use equation (2.22) rather than (2.23), giving a reduction in the transmission value to 0.134±0.013 of the empty cavity value. Estimating from the red trace in Figure 3.5 above, the experimental value for this is 0.17±0.02, showing agreement to within two standard errors.

The results in Figure 3.5 show a scan width of up to 1.2 GHz in fully locked mode. This is currently the limit of the length change of the piezo electric actuator used to change the

---

(a) The literature from the manufacturer gives an accuracy of ±5% for air at pressures below 1 Torr, but the response changes with gas composition. The closest gas to N2O for which the correction is given in the manual is CO2 for which the pressure reading is about 10% below the actual pressure.

cavity length. If scans taken by locking the QCL to a sequence of cavity modes, it would be possible to piece these scans together to produce cavity-enhanced absorption spectra across the continuous tuning range of the QCLs, which at present is up to 1.6 cm^{-1}.

As mentioned in Section 2.2, one of the limitations of chemical sensors is fringing. The small ripples to the left of the principal absorption features in Figure 3.5 are an example of this. These ripples are common to all traces, indicating a process that is independent of the analyte pressure. It is clear from these results that the fringing features have about the same horizontal (and hence frequency) scale as the absorption features themselves. Consequently, in a practical situation fringing often constitutes a lower limit to the instrument sensitivity. To illustrate, the relative magnitude of the fringing features to the full transmission signal (measured from the "Detector Blocked" line) in Figure 3.5 is about 0.055. From equation (2.12) this gives

$$
\begin{aligned}
(\alpha L)_{min} &= \frac{(\partial I_t)_{min}}{I_{t0}} \times \frac{\pi}{2F} \\
&= 0.055 \times \frac{\pi}{2 \times 2415} \\
&= 3.6 \times 10^{-5}
\end{aligned}
\tag{3.2}
$$

$$
\text{giving} \quad \alpha_{min} = 9.3 \times 10^{-7}/\text{cm}
$$

for a cavity length of 38.5 cm. Note that we haven't used the term NEAS (or absorption sensitivity α_S) because both of these include the unit $/\sqrt{\text{Hz}}$. We cannot use this unit here because the fringing is not a random process on the time scales of the measurement and thus does not average away. Consequently, this becomes an absolute limit, rather than one that depends on bandwidth. Clearly, this is highly undesirable for a chemical sensor and many approaches to reducing the effects of fringing have been reported.

Assume now that we have arranged for a better alignment or a different setup that has removed the fringing features. The NEAS determined by the relative size of the residual noise on the traces in Figure 3.5 with respect to the full DC transmission level is

$$
\begin{aligned}
NEAS &= 0.007 \times \frac{\pi}{2 \times 2415} \\
&= 4.6 \times 10^{-6}/\text{cm into a 20kHz bandwidth} \\
&= 3.2 \times 10^{-8}/\sqrt{\text{Hz}}
\end{aligned}
\tag{3.3}
$$

$$
\text{giving} \quad \alpha_S = 8.4 \times 10^{-10}/\text{cm}/\sqrt{\text{Hz}}
$$

The shot noise equivalent absorbance sensitivity for this sensor from equation (2.31) in Section 2.3 for a 65 µA detector photocurrent is

$$SNEAS = \sqrt{\frac{2e}{65 \times 10^{-6}}} \times \frac{\pi}{2 \times 2415} \, /\sqrt{Hz}$$

$$= 4.6 \times 10^{-11} \, /\sqrt{Hz} \qquad\qquad (3.4)$$

$$\text{giving} \quad \alpha_{sh} = 1.19 \times 10^{-12} \, /cm/\sqrt{Hz}$$

Despite being nearly three orders of magnitude worse than the shot noise limit, (3.3) actually represents good performance for a sensor operating at base-band frequencies, because as discussed in Section 2.3 we would expect this sensor to be limited by 1/f noise rather than shot noise. However, for all practical purposes, until something is done to reduce the impact of the optical fringing, this sensor is limited to the performance level given by equation (3.2).

There are several ways to reduce interference due to fringing. One way is to amplitude stabilize the optical field entering the cavity. Since the QCL drive current is already used for frequency control, another method must be found. A means available to us is to use the AOM shown in Figure 3.4, the transmission of which can be varied via the video input of the AOM's electronics. A control loop using this video input and an additional detector can achieve this stabilization. This is one of the next stages of development of all the LWIR cavity-enhanced chemical sensors under development at PNNL and is planned for mid FY03.

Amplitude stabilization is expected to reduce the effects of fringing that arise from optical components before the cavity that are independent of the existing frequency control loop. It will not correct for effects caused by periodically changing locking offsets as described in Section 3.1, or from fringing involving the transmission detector. These latter problems are more difficult, and can only be fully dealt with by redesigning the optics in the reflection and transmission pathways to avoid the effects of fringing entering via the cavity-locking loop and the transmission detector respectively. Temporary improvements can always be obtained by studiously re-aligning the reflection optics to optimize for minimum fringing, although any gains here can be easily lost. [23,24]

As pointed out earlier, one reason why optical fringing is so limiting is that as seen in Figure 3.5, the fringes have a similar frequency scale as the Doppler-broadened features of typical analytes. An obvious way to decrease the impact of such an effect is to look for features of a markedly different size so they are easily distinguished. The Lamb dip is one such feature. In a point sensor where the pressure of the sample may be varied, the degree of absorption can be controlled to some extent. This in turn affects another parameter, the saturation [25]. When a molecular transition becomes saturated, this means that a second probing beam would see a reduced level of relative absorption by comparison to the first. However, this does not occur if the two beams interrogate different velocity classes of molecules within the analyte. Since Doppler broadening is an inhomogeneous mechanism, the velocity classes probed by the two beams are different for all detunings from the center of the molecular transition except zero. At an optimum level of saturation, narrow features appear as the frequency scan causes the two

probed velocity classes to coincide. These features are sub-Doppler, and are only broadened by the actual lifetime of the molecular transition and homogeneous effects such as transit-time, pressure and power broadening.

In cavity experiments such as ours where only one beam is used, these features are known as Lamb dips and examples are shown in Figure 3.6 below. (When this is performed with two independent lasers, this phenomenon is known as spectral hole burning. [26])
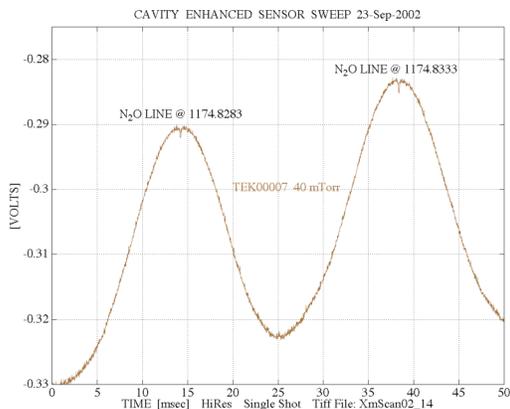
Fig. 3.6. Cavity-Enhanced Absorption features of Figure 3.5 but showing the Lamb dips at the peaks of the Doppler-broadened profiles.

Although these features are smaller in magnitude (see Section 3.3 below) than the Doppler-limited features, they are much sharper. For example, in Figure 3.6, the Doppler-broadened features have a width of 65.7 MHz, while the Lamb dips have a width of 2 MHz or less. They are unmistakably different from the features due to fringing. This means that the presence and strength of absorbing chemicals can be characterized using Lamb dips despite the presence of fringing that would otherwise confuse the detection of Doppler-broadened features. Another advantage of Lamb dips derives directly from their much narrower signatures. If several chemicals were simultaneously present in a sample, the narrower Lamb dips could allow sufficient resolution to distinguish them whereas the Doppler-broadened features may not.

Despite the fact that Lamb dips may allow increased sensitivity and selectivity, a disadvantage is that actually finding them amongst the larger Doppler-broadened features and fringing is difficult with our simple cavity-enhanced chemical sensor shown in Figure 3.4. To facilitate the detection of these narrower features, we now turn to the implementation of the techniques of FM modulation and phase-sensitive detection discussed in Section 2.4.

3.3   FM Cavity-Enhanced Sensor with Phase-Sensitive Detection.

Many of the cavity-enhanced sensor experiments performed at PNNL in the LWIR use phase-sensitive detection. A schematic of such a sensor is shown below in Figure 3.7, which is the experimental realization of the cavity-dither arrangement shown previously in Figure 2.7. Again, the QCL is locked to the peak of the cavity mode and the cavity

FM RECOVERY CAVITY-ENHANCED SENSOR
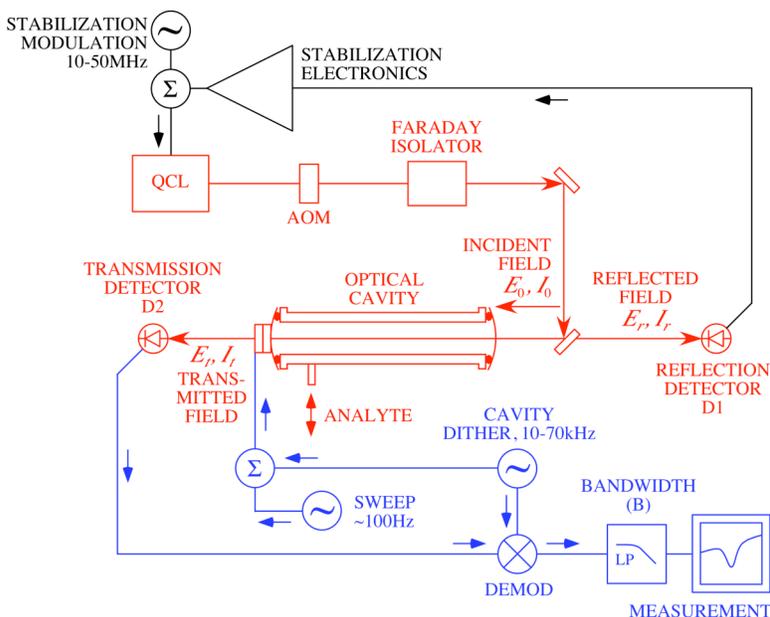
Fig. 3.7. FM Recovery Cavity-Enhanced Sensor. The QCL is locked to the cavity as before. However, in addition to the cavity length being scanned, it is also dithered or modulated. The transmission signal from D2 is demodulated accordingly, then filtered and recorded as before. Intensities and electric field amplitudes are indicated for incident, reflected and transmitted fields.

length is scanned or swept to move the frequency of this cavity mode and that of the laser across molecular absorption features. The difference is that in addition to be being scanned, the cavity length is also rapidly "dithered" using a modulation frequency in the 10s of kHz. This cavity modulation signal (independent of that applied to the laser in order to lock it to the cavity,) is applied directly to the piezo element in addition to the signal used to produce the frequency sweep. Since this modulation frequency is well within the bandwidth of the laser-locking loop, the laser frequency follows this cavity dither in the same way as it follows a cavity sweep. The signal from the transmission detector is then demodulated at this dither frequency, after which, it is filtered and recorded on a digital oscilloscope as before.

Ideally, the QCL lock to the optical cavity would be unaffected by the cavity dither, and the laser and the cavity would move together in frequency causing the modulation to be applied to the analyte. Consequently, regardless of their spectral distribution, all the sidebands would be coupled directly into the cavity. Indeed, we operate under this assumption when applying the theory developed in Section 2.4 for phase-sensitive detection to the cavity-dither technique. However, as also mentioned in that section, this is not exactly true due to disruptions of the cavity mode due to the cavity's length modulation. [19] Nevertheless, we continue under this assumption to make theoretical comparisons, since a more complete theory does not yet exist for this sensor.

Typical traces from the FM cavity-enhanced sensor are shown in Figure 3.8. The upper trace in red corresponds to 1-f detection, or simple FM spectroscopy of the Lamb dip in N_2O at 1174.901 cm^{-1}. These results were taken with an early version of our optical cavity in which the mirror reflectivity was only 99.1%, giving a finesse of 338. Also, our experiment was not fully optimized at this stage. Consequently, the sensitivity of these signals is not better than the results of the simple cavity-enhanced sensor shown in

LAMB DIP FM RECOVERY

LINE AT k = 1174.901 /cm
S = 4.112e-21
QCL CURRENT I = 478 mA
FSR = 560MHz   F = 338

1-f DETECTION
2-f DETECTION
3-f DETECTION

VOLTS

TIME [msec]   Av20   N$_2$O@17mTorr   Dither 19.5kHz   LF Cav
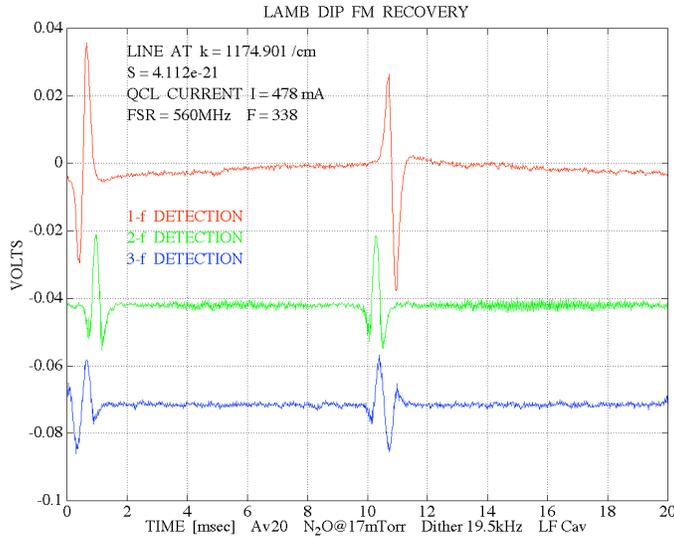
Fig. 3.8. 1-f, 2-f, 3-f detection of Lamb dips in N$_2$O. (These traces were taken with an earlier cavity design having a finesse of 338 rather than 2415.)

Figure 3.5, and are not presented here. The purpose of Figure 3.8 is to illustrate the powerful selectivity of the FM cavity-dither technique. While the 1-f Lamb dip feature is quite sharp, the gentle sloping of the trace to either side is all that remains of the Doppler-broadened line and any fringing effects. (A comparison to the much poorer contrast of the Lamb dip in Figure 3.6 emphasizes the power of this selectivity, although it must be pointed out that the Lamb dips in Figure 3.6 were not optimized.) As we use higher order detection techniques such as 2-f and 3-f detection shown in green and blue respectively, the sloping due to the original Doppler broadened traces is no longer visible on the plot.

3.4   Cavity-Enhanced Sensors Using NICE-OHMS.

The next step we have taken in the development of cavity-enhanced sensors was to use the technique of resonant sideband detection, or NICE-OHMS [1-3], discussed theoretically in Section 2.5. Figure 3.9 shows the NICE-OHMS experimental arrangement.

NICE-OHMS CAVITY-ENHANCED SENSOR
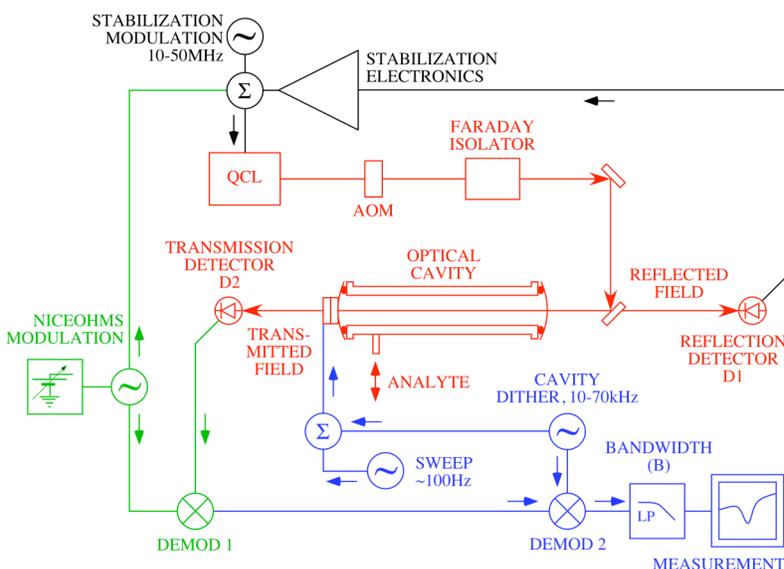
Fig. 3.9. Nice-Ohms recovery sensor. A high frequency modulation equal to the cavity FSR is imposed on the laser in addition to that required for the cavity locking. This high frequency is equal to the spacing between adjacent cavity modes. The resulting transmission signal is first demodulated at this frequency, then secondly at the cavity modulation frequency as explained for the FM recovery cavity-enhanced sensor.

In addition to the modulation required for the locking of the QCL to the optical cavity, a modulation is applied at the cavity FSR, which was about 390 MHz for these experiments. The resulting sidebands present on the incident field coincide exactly with separate optical cavity modes and thus enter the optical cavity. The field transmitted from the cavity is detected and demodulated first at the FSR frequency, and then at the cavity-dither modulation frequency, these processes being represented by demod1 and demod2 respectively. The resulting signal is then filtered at some bandwidth and monitored during cavity scans as before.

The NICE-OHMS sensor configuration shown here potentially combines the best features of all the previous techniques. To begin with, the resonant sideband technique gives immunity to noise in the lock between the QCL and the optical cavity, becoming more important as we use cavities with higher finesse, which are harder to lock. The high frequency modulation required to achieve NICE-OHMS results in the FM detection occurring at a frequency where the technical noise on the laser is virtually non-existent, giving the best possible immunity to 1/f noise. The cavity-dither technique also employed here allows the selectivity of the instrument to be increased as previously discussed. The result is an instrument with highly optimized sensitivity and selectivity.

The relative immunity to noise as seen with our sensor is illustrated by the pair of Figures 3.10 and 3.11. In Figure 3.10, we see a direct absorption trace (as taken using the simple unmodulated sensor configuration of Figure 3.4), the 1-f signal of the FM recovery sensor shown in Figure 3.7, and the full NICE-OHMS signal taken using the configuration of Figure 3.9. In Figure 3.10, these same signals are shown under the best possible circumstances, whereas in Figure 3.11, the lock between the QCL and the optical cavity has been degraded. A visual inspection of these two figures shows that the signal-to-noise ratio of the NICE-OHMS trace is only degraded by a factor of about 2, whereas those of the other techniques by approximately a factor of 10 or more. As higher
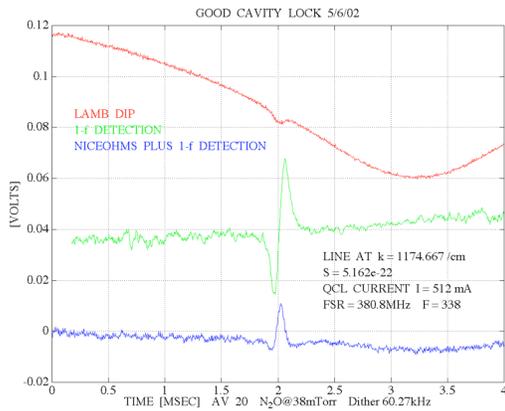
Fig. 3.10. Traces showing direct absorption and the Lamb dip (red), 1-f recovery trace (green) and full Nice-Ohms recovery signals (blue). The Nice-Ohms trace has a symmetric shape rather than the dispersion shape of the 1-f signal due to the extra order of demodulation.
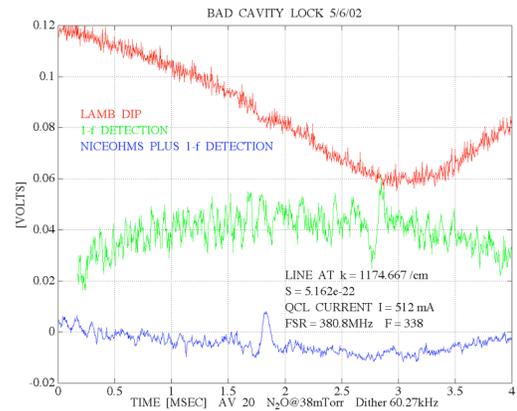
Fig. 3.11. These are the same traces as shown in Figure 3.10, except that the lock between the QCL and the optical cavity was deliberately degraded. The direct absorption (red) and the 1-f recovery (green) traces are degraded more than the Nice-Ohms trace, which is only marginally affected.

cavity finesses are achieved in the LWIR, cavity locking will become proportionately susceptible to noise, and this advantage will become of prime importance, especially for robust deployable sensors.

An example is now given of the calculation of the sensitivity of the NICE-OHMS sensor. A typical piece of data taken with the NICE-OHMS sensor is shown in Figure 3.12. It should be noted that we are using a very weak absorption line for these measurements in order to approach the sensitivity limit of this sensor while still having pressures that are high enough to measure with some accuracy using the Convectron gauge on the optical cavity. This data appears here in exactly the same format as it appears in the log-books of the scientists taking the measurements, hence the extra annotation.

We now proceed to do a detailed calculation the values of the NEAS and sensitivity α_S for the above result as an illustration of how this is done. We do this here rather than in the theory section, because it is not until now that we have covered all the relevant concepts necessary for a full appreciated of this calculation. For FM and NICE-OHMS sensors, this is more complicated than for the direct absorption case dealt with in Section 3.2 because there is no inherent DC transmission level in the results to use as a reference; we must use the strength of the observed line instead. It also must be understood that in doing this, we usually rely on a database for this information, in this case HITRAN.

NICEOHMS PLUS 1-f DETECTION 10-Jul-2002

NICEOHMS PLUS 1-f DITHER
LINE AT k = 1174.9515 /cm
S = 4.029e-23
B = 10kHz,   SR560 G = *20
Femto G = 20dB
FSR = 387.5MHz   F = 2415

TIME [msec]   Av20   Sw95Hz   N$_2$O@16.7mTorr   Dither 57.4kHz
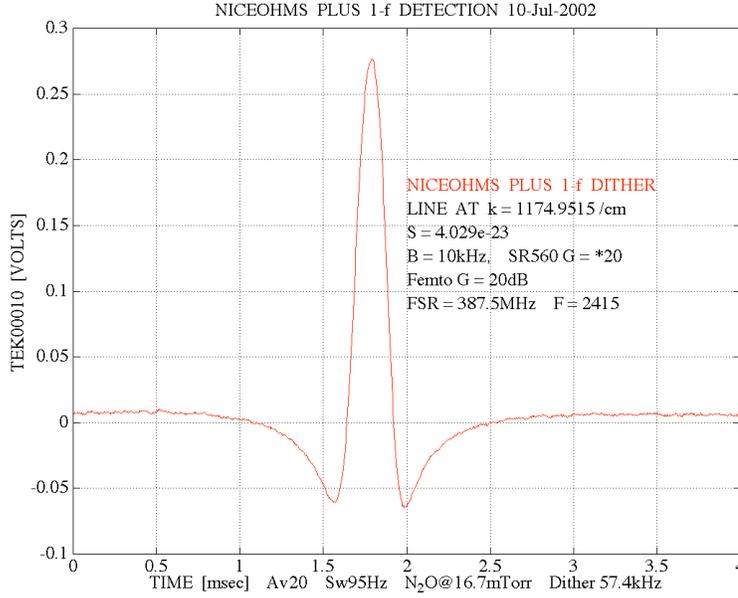
Fig. 3.12. Result taken with optimized NICE-OHMS sensor shown in Figure 3.9. The various annotations on the plot shows the line strength *S*, the bandwidth *B*, the number of averages (20), the pressure (16.7 mTorr), dither frequency and other factors. The vertical scale is in volts because the measurement was made on a digital oscilloscope. The file name is also put on the vertical axis to allow easy identification.

From the given line strengths *S* in cm/molecule, we calculate the absorption cross section

$$\sigma(\nu) = S \cdot c \cdot g(\nu) \qquad (3.5)$$

where *c* is the speed of light in cm/s and $g(\nu)$ is the lineshape function, which is normalized to have a unit area when integrated over frequency. For gases at low pressure, $g(\nu)$ is determined by Doppler broadening. At line center this gives

$$g(\nu_0) = \frac{2\sqrt{\ln 2}}{\sqrt{\pi}\,\delta\nu_D} \qquad (3.6)$$

where $\delta\nu_D$ is the full width at half maximum (FWHM) due to Doppler broadening, which is 65.7 MHz for this transition in nitrous oxide at room temperature, giving $g(\nu_0) = 1.43 \times 10^{-8}\, s$. Substituting (3.5) into (3.6) and including our experimental values, we obtain

$$\sigma(\nu_0) = S \cdot c \cdot g(\nu_0)$$
$$= \frac{2\sqrt{\ln 2}}{\sqrt{\pi}} \times \frac{S \cdot c}{\delta\nu_D} \qquad (3.7)$$
$$= 1.73 \times 10^{-20}\ \text{cm}^2/\text{molecule}$$

The number density in molecules per cubic centimeter is given by

$$N = \frac{1}{1000} \times \frac{P}{kT} \times \frac{101.3}{760 \times 10^3}$$

$$= 9.652 \times 10^{15} \times \frac{P}{T} \qquad (3.8)$$

where P is the pressure in mTorr, T is the temperature in Kelvin and k is Boltzmann's constant. The numerical factors are for the conversion of liters to cubic cm, and kilopascals to mTorr. We can now find the extinction coefficient at line center:

$$\alpha_0 = \sigma(\nu_0)N$$

$$= 1.2523 \times 10^{-7} \times \frac{Sc}{\delta\nu_D} \times \frac{P}{kT} \qquad (3.9)$$

For a pressure of 16.7 mTorr and temperature of 298K, this gives an extinction coefficient of 9.45×10^{-6} /cm. If we divide (3.9) by the pressure P, we obtain the extinction coefficient per mTorr as used in Section 3.2 (see equation (3.1))

$$\alpha_{0p} = \frac{Sc}{\delta\nu_D} \times \frac{1.2523 \times 10^{-7}}{kT} \qquad (3.10)$$

At this point, we recall that the feature in Figure 3.12 is actually a Lamb dip. The maximum possible size of a Lamb dip is 13% of the size of the Doppler broadened feature. [27] (This is because the contrast of a Lamb dip is a function of the saturation parameter, which has an optimum value.) In order to obtain the maximum size Lamb dip signal, the laser power, pressure or some other variable must be optimized. In our case, this optimization is done via the pressure. Consequently, we can write the "maximum saturated absorption contrast" as

$$\Delta\alpha_{sat\_max} = 0.13\alpha_0 \qquad (3.11)$$

which for the above line is 1.2×10^{-6} /cm. This is the absorption level corresponding to the peak-to-peak size of the feature in Figure 3.12.

We now estimate what absorption level corresponds to the noise to either side of the feature. By looking closely, the contrast (signal-to-noise ratio) is seen to be around 75. The corresponding noise-limited absorption judging from this figure is 1.62×10^{-8}/cm. This trace was taken with a filter bandwidth of 10 kHz and 20 averages, giving an effective bandwidth of 500 Hz. Thus the absorption sensitivity corresponding to Figure 3.12 is

$$\alpha_S = \frac{\Delta\alpha_{sat\_max}}{(S/N)} / \sqrt{B/A}$$

$$= 7.2 \times 10^{-10} / cm / \sqrt{Hz} \qquad (3.12)$$

$$\text{giving} \quad NEAS = \alpha_S L = 2.8 \times 10^{-8} / \sqrt{Hz}$$

where S/N is the signal-to-noise ratio, B is the bandwidth and A is the number of averages.

The data shown in Figure 3.12 was chosen because although not the best data taken with this sensor, the noise fluctuations in the to either side of the main feature are clearly visible, facilitating the above illustration. The best performance to date of the NICE-OHMS sensor is

$$\alpha_{S\_best} = 8.5 \times 10^{-11} / cm/\sqrt{Hz}$$

$$\text{giving} \quad (NEAS)_{best} = 3.3 \times 10^{-9} / \sqrt{Hz}$$

(3.13)

For comparison the shot noise equivalent absorbance sensitivity for our NICE-OHMS sensor using the calculation of the inventors [1] for a 65 µA detector photocurrent is

$$\left(SNEAS\right)_{NO} = \sqrt{\frac{2e}{65 \times 10^{-6}}} \times \frac{\pi}{2 \times 2415} \times 4.16 \ /\sqrt{Hz}$$

$$= 1.9 \times 10^{-10} / \sqrt{Hz}$$

$$\alpha_{sh} = 4.9 \times 10^{-12} / cm/\sqrt{Hz}$$

(3.14)

This shows that we are just over an order of magnitude away from shot noise limited performance with this sensor.

4.0 Outlook

Over the past 12 months, we have developed three distinct continuous-wave long-wave infrared cavity-enhanced chemical sensors using quantum cascade lasers, including a simple cavity-locked absorption sensor and an FM recovery cavity-locked sensor. Moreover, we have given the first demonstration of the NICE-OHMS technique using quantum cascade lasers. We have examined in detail the advantages of the additional techniques employed in each stage of the development of these sensors at PNNL, including the use of high finesse optical cavities and the use of frequency modulation and phase-sensitive detection. Combined with the resonant sideband modulation technique of NICE-OHMS to add immunity to any remained frequency jitter between the laser and optical cavity, the resulting combination of techniques gives the final sensor potentially the best sensitivity and selectivity available in the LWIR. In addition to presenting the experimental results obtained over the past 12 months, we have given particular attention to developing the understanding of cavity-enhanced sensor operation. To this end, an in depth discussion was given of the theory of cavity-enhanced chemical sensors, with particular emphasis on the impact of noise on sensor performance. In this concluding section, we now turn to the work intended for FY03, and address some of the outstanding issues faced in taking prototype sensors from the laboratory table out into the field, In particular, issues required to proceed from absorption sensitivities to analyte concentrations. It should be pointed out that more work in this area has been done at PNNL regarding the short-wave infrared (SWIR) sensor development using cavity ringdown, and we refer the reader to the SWIR FY02 report for more information.

4.1 Further Work for FY03

Several improvements to the cavity-enhanced sensors are already underway. FY03 will see the construction of better optical cavities and gas handling systems for use in our chemical sensor experiments. This will involve better mirrors, better acoustic damping, better vacuum pumps, better pressure measurements, and lower contamination from out gassing and leaks for a more complete exploration of the performance of the NICE-OHMS sensor. This is important, because the most convincing demonstration of the sensitivity of a sensor is with absorptions near the noise limit and not by extrapolating by two ordrs of magnitude based on signal-to-noise ratio. Mirrors of higher reflectivity and very low loss are now becoming more available in the LWIR due in part to PNNL's interaction with certain coating houses. The resulting increase in cavity finesse will increase our access to weak absorption lines and enable us to detect lower concentrations of analytes in test samples. New mounting arrangements using better piezo electric actuators will hopefully allow longer scan ranges and increase the utility of the instrument.

Certain fundamental changes in design will also be tested including the implementation of amplitude stabilization as discussed in 3.2 to aid in the reduction of fringing. To completely eliminate the effects of optical fringing, redesign of the optics in the path ways between the cavity and both the reflection and transmission detectors will be necessary in addition to amplitude stabilization. It is unlikely a complete redesign of the

optics will be completed in FY03, but at least work with a ring cavity will be started. A change to higher reflectivity cavity mirrors will likely require addition of an additional servo control loop to maintain the correct NICE-OHMS RF modulation frequency. This has not yet been necessary because the cavity bandwidth has been large enough to allow adjustment of this modulation frequency, but manual adjustment may no longer be adequate with higher finesse values. For example, tuning the cavity over 1.2 GHz or three times the FSR changes the cavity length by 1 part in 30,000. This is small relative to the 1 part in 2415 that the cavity bandwidth is as a fraction of the FSR. However, with increased finesse and/or increased scan lengths, the change in the FSR with scanning will become comparable or larger than the cavity bandwidth.

There are disadvantages to continuously increasing the cavity finesse, and in FY03 we may be at the point where we can investigate these limits. The extra servo loop mentioned above will operate at a low frequency, probably in the 100s of Hz. This will decrease the required measurement bandwidth, or conversely, increase the necessary averaging time to remove the ripple resulting from this extra modulation. The resulting longer measurement times may actually prove disadvantageous. In addition to this, the frequency of the cavity-dither modulation may have to be reduced because of the increased cavity finesse and interference with the cavity mode via the Doppler effect of the moving mirror as discussed in Section 2.4.1, forcing the frequency of the extra modulation described above to be reduced even further. It will be of prime importance to see what sensitivity the instrument has, and what concentrations of which species are available at this sensitivity when this limit is reached. The experimental investigation of the performance of the sensor with higher finesse mirrors will be accompanied by an attempt to gain a better understanding of the theory of the disruption of the cavity mode due to mirror motion, although it is unclear whether a full theory will be developed by the end of FY03.

4.2   The Meaning of Sensitivity

It is often asked when good sensor performance is reported, as to how many "parts per billion" or "parts per trillion" can a sensor detect if it is quoted as having a given sensitivity, as defined in section 1.0. Unfortunately, there is no solid answer to this question, but rather, more questions. In particular: detection of what, in the presence of what, and over how long? This prompts us to ask another question: What does sensitivity mean? Here we explore why these questions are important and how we might find a pathway forward to finding solutions to these issues.

We begin this discussion with a basic observation. The sensitivity (and the NEAS) values calculated in this report are based on what absorption levels correspond to the noise present on the traces of features that are hundreds of times larger than this noise. Consequently, this estimation is an extrapolation at best. If we were required to actually measure a feature corresponding to the resulting noise-limited sensitivity values, this implies a measurement with a signal-to-noise value of unity. Consequently, the feature would appear only slightly larger than the noise fluctuations themselves. If a researcher knows what he or she is looking for and where to look, then this kind of detection is

possible, but it is not generally considered a reliable measurement. This raises the question as to what level of signal-to-noise should be chosen as a reliable lower limit for a chemical sensor in the field. To the best of our knowledge, there is a consensus among the analytical chemistry community, that the smallest detectable quantity is three times the standard deviation of the noise. However, a decision on a value for this quantity in the case of an operating chemical sensor would depend on many factors, and hence may change depending on the circumstances. For example, what the sensor is used for may directly impact the acceptable probability of false positives or negatives, and hence the acceptable level of signal-to-noise ratio. These issues become part of a full chemometric analysis and in many cases, it may not be possible to make sensible judgments about the limitations of sensor performance without extensive field tests of prototype sensors, yielding a large body of test data from which certain conclusions could then be drawn.

Another issue to discuss is the modus operandi of a sensor and how this may affect utilizing the potential sensitivity of the instrument. This is best illustrated by comparing two sensors discussed in this report. The simple cavity-enhanced sensor of Section 3.2 detects the absorption level due to the Doppler-broadened features of an analyte inside the optical cavity. For this sensor, it makes no difference if the analyte is pure and at a very low pressure, or is mixed with a buffer gas at some higher pressure that has no interfering absorptions. If the resulting absorption level is the same, and the pressure is not sufficiently large as to change the absorption profile (which at room temperature for nitrous oxide for example is about 10 Torr), the resulting signal will be the same. Consequently, in theory the absorption from species X can be calculated from the signal, and the number density obtained using equation (2.9) to give

$$N_X = \frac{\alpha_X}{\sigma_X} \qquad (4.1)$$

The concentration of this species C_X could then be found by diving by the total number density of the test sample N_T found from the test sample pressure:

$$C_X = \frac{N_X}{N_T} \qquad (4.2)$$

The fact that this process is relatively straightforward in theory makes this simple sensor very attractive. However as illustrated in Section 3.2, there are technical problems that severely limit this sensor's performance. This is one reason why cavity ringdown techniques are extremely useful, because they present a similar simplicity in theory, but allow some of the technical difficulties to be obviated. They do however present their own challenges, and a more detailed discussion of cavity ringdown techniques can be found in the SWIR report.

Let us now compare the simple cavity-enhanced sensor performance with that of the NICE-OHMS sensor discussed in Section 3.4. Like the FM recovery sensor, the NICE-OHMS sensor uses Lamb dips as the target features for detection. Using the selectivity

of FM and phase-sensitive detection, this allows the effect due to fringing, the principal limitation of the simpler non-modulated sensor, to be all but obviated. However, Lamb dips are saturation features, meaning that they can only be observed below a certain pressure for a given analyte and for a given laser power. In this case, it clearly matters whether the analyte is pure or is in the presence of a buffer gas. This is because even if the absolute absorption level of a given feature remains unchanged, the increased pressure due to a buffer gas may completely wash out the Lamb dips due to pressure broadening. On the other hand, reducing the pressure of an already tenuous analyte into a region acceptable for the observation of Lamb dips may decrease the absorption values below the sensitivity of the instrument. Consequently, it is difficult to convert a given extrapolated sensitivity of this instrument to given concentration levels at pressures other than those observed.

One final point for discussion here, is the rate at which measurements are required to be taken for a practical sensor, and how gas handling systems may affect the speed of measurements. Again this problem has been much more thoroughly dealt with in the SWIR sensor studies carried out at PNNL, but it is so important that it must be mentioned here. Having an absorption sensor with a reputable sensitivity is only one piece of the task of developing a chemical vapor sensor. Gas samples must be moved in and out of the sensor without disrupting its operation or affecting its sensitivity. It must be known that after a given sample is evacuated, that no remnants remain to interfere with the next sample. This is difficult to ensure in any general way, because the nature of different analytes varies enormously. Some gases are known to be "sticky" meaning that they have an affinity for the walls of the chamber if it is not made from certain substances, or some analytes will stick to the mirrors used in the optical cavity, etc. This would necessitate the process of flushing an instrument to remove such stubborn analytes. Flushing such an instrument between successive measurements will take extra time, and must utilize a clean flushing gas such as purified air or dry nitrogen. The continual flow of gas through a cavity-enhanced chemical sensor also presents the possibility of mirror contamination. Some analytes may even be corrosive, and even if harmless to humans at the concentration levels of interest here, could easily sully or even damage the high finesse mirrors used in such a sensor. This is because the mirror coatings are made from extremely high quality dielectric layers, and even a small amount of dust or contamination can multiply the losses in a high finesse optical cavity by tens to hundreds of times, drastically degrading the sensitivity of the instrument.

Over FY03, further study and consideration will be given to these issues in addition to continuing to improve the performance of the cavity-enhanced sensors developed at PNNL.

5.0 References

[1] J. Ye, L.-S. Ma, and J.L. Hall, "Ultrasensitive detections in atomic and molecular physics: demonstrations in molecular overtone spectroscopy," J. Opt. Soc. Am. B, **15**(1), 6(1998).

[2] L. Gianfrani, R.W. Fox, and L. Hollberg, "Cavity-enhanced absorption spectroscopy of molecular oxygen," J. Opt. Soc. Am. B, **16**(12), 2247(1999).

[3] C. Ishibashi and Hiroyuki Sasada, "Highly sensitive cavity-enhanced sub-Doppler spectroscopy of a molecular overtone band with a 1.66 μm tunable diode laser," Jpn. J. Appl. Phys. **38,** 920(1999).

[4] "Chemometrics, A Practical Guide," Kenneth R. Beebe, Randy J. Pell and Mary Beth Seasholtz, ISBN 0-471-12451-6, Wiley-Interscience, 1998.

[5] P. Werle, F. Slemr, K. Maurer, R. Kormann, R. Mucke, and B. Janker, "Near- and mid-infrared laser-optical sensors for gas analysis," Optics and Lasers in Engineering, **37,** 101 (2002).

[6] "Optical Electronics," 3rd Ed., Amnon Yariv, p315-22. ISBN 0-03-070289-5, HRW, CBS, New York, 1985.

[7] "The Art of Electronics," Paul Horowitz and Winfield Hill, 2nd Ed., p184. ISBN 0-521-37095-7, Cambridge University Press, 1996.

[8] "Lasers," Anthony E. Siegman, p418. ISBN 0-935702-11-3, University Science Books, 1986.

[9] "Trace-gas analysis using diode lasers in the near-IR and long-path techniques," G. Gagliardi, L. Gianfrani, Optics and Lasers in Engineering, **37,** 509(2002).

[10] James M. Supplee, Edward A. Whittaker, and Wilfried Lenth, "Theoretical description of frequency modulation and wavelength modulation spectroscopy," Applied Optics **33**, 6294-6302 (1994).

[11] Joel A. Silver, "Frequency-modulated spectroscopy for trace species detection: theory and comparison among experimental methods," Applied Optics, **31**(6), 707(1992).

[12] R. L. Barger, M. S. Sorem, and J. L. Hall, "Frequency stabilization of a cw dye laser," Appl. Phys. Lett. **22,** 573(1973).

[13] R. W. P. Drever, J. L. Hall, F. V. Kowalski, J. Hough, G. M. Ford, A. J. Munley and H. Ward, "Laser phase and frequency stabilization using an optical resonator," Appl. Phys. B. **31,** 97(1983).

[14] Ch. Salomon, D. Hils, and J. L. Hall, "Laser stabilization at the millihertz level," J. Opt. Soc. Am. B. **5,** 1576(1988).

[15] J. L. Hall, Long-Sheng Ma, M. Taubman, B. Tiemann, F. Hong, O. Pfister, J. Ye, "Stabilization and frequency measurement of the I_2-stabilized Nd:YAG laser," Conference on Precision Electromagnetic Measurements, 6-10 July 1998, Washington, DC, USA.

[16] "Laser Spectroscopy, Basic Concepts and Instrumentation," 2nd Ed., Wolfgang Demtröder, p65. ISBN 3-540-57171-X, Springer-Verlag, Berlin, Heidelberg, New York.

[17] "A General Non-Complex Analytical Expression for the Nth Fourier Component of a Wavelength-Modulated Lorentzian Lineshape Function," O. Axner, P. Kluczynski, A.M. Lindberg, submitted to J.Q.S.R.T.

[18] "Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables," Ed. Milton Abramowitz and Irene A. Stegun, p 363. U.S. Government Printing Office, Washington, D.C. (1972).

[19] "Dynamic response of a Fabry-Perot interferometer," M.J. Lawrence, B. Willke, M.E. Husman, E.K. Gustafson and R.L. Byer, J. Opt. Soc. Am. B., **16**(4), 523 (1999).

[20] T. L. Myers, R. M Williams, M. S. Taubman, C. Gmachl, F. Capasso, D. L. Sivco, J. N. Baillargeon, A. Y. Cho, "Free-Running Frequency Stability of Mid-Infrared Quantum Cascade Lasers," Optics Letters, **27**(3), 170 (2002).

[21] Matthew S. Taubman, Tanya L. Myers, Bret D. Cannon, Richard M. Williams, Federico Capasso, Claire Gmachl, Deborah L. Sivco, and Alfred Y. Cho, "Frequency stabilization of quantum cascade lasers by use of optical cavities," Optics Letters **27** 2164-2166 (2002).

[22] Robert A. Toth, Applied Optics, **23,** 1825-1834 (1984). The HITRAN reported line strengths for N_2O are estimated as less than 2% for these lines.

[23] Clinton B. Carlisle, David E. Cooper, and Horst Preier, "Quantum noise-limited FM spectroscopy with a lead-salt diode laser," Applied Optics, **28**(13), 2567 (1989).

[24] David S. Bomse, Alan C. Stanton, and Joel A. Silver, "Frequency modulation and wavelength modulation spectroscopies: comparison of experimental methods using a lead-salt diode laser," Applied Optics, **31**(6), 718 (1992).

[25] "Laser Spectroscopy, Basic Concepts and Instrumentation," 2nd Ed., Wolfgang Demtröder, p436. ISBN 3-540-57171-X, Springer-Verlag, Berlin, Heidelberg, New York.

[26] "Laser Spectroscopy, Basic Concepts and Instrumentation," 2nd Ed., Wolfgang Demtröder, p253. ISBN 3-540-57171-X, Springer-Verlag, Berlin, Heidelberg, New York.

[27] "Laser Spectroscopy, Basic Concepts and Instrumentation," 2nd Ed., Wolfgang Demtröder, p441. ISBN 3-540-57171-X, Springer-Verlag, Berlin, Heidelberg, New York.