Effective Dual Gas Species Monitoring in Porous Scattering Media

Master's Thesis
By
Mikael Ingemansson

Lund reports on Atomic Physics LRAP 373
Department of Physics, Lund Institute of Technology
Lund, February 2007

Abstract

Absorption spectroscopy has a wide array of applications and new fields are constantly being investigated. The possibility of measuring gas inside a scattering material is fairly new and is therefore very much in its developing stage. A demand for measuring absolute values has arisen, this can be accomplished in a couple of different ways. The approach ventured here is to measure two gases simultaneously, where the concentration of one of the gases is known in advance. The measurement of the known gas will hence serve as a means of calibration.

A method for separating the signals by making the modulation phase of the two lasers perpendicular, removing signal crosstalk, has been developed. The practise of normalizing the second Fourier harmonic with the first has been investigated and implemented. Additionally, software has been constructed with the purpose of controlling the setup units.

Measurements have been performed over time on pieces of balsa wood and a sintered glass plate during drying processes. The acquired data have shown good resemblance to previous relevant measurements.

Contents

Chapter 1

Introduction

1.1 Background of thesis

Absorption spectroscopy is a commonly used technique for determining the concentration of gases. In the traditional variety of this spectroscopy, light of known spectral composition propagates over a known path length. By studying the light intensity decay under its propagation, the concentration of the gas in the light path can be determined. This is rather straight forward, as long as there is only one unknown, namely the concentration of the gas [1]. The reason for the change in light level is a consequence of light interacting with materia. Atoms and molecules, which make up materia, are associated with a probability to absorb certain energies. However, the size of the energy absorbed can not be chosen arbitrarily, but has to be changed in steps. That is, the energy of atoms and molecules is quantized. This is the background to why atoms and molecules only absorb certain wavelengths, the energy of the light (which corresponds to the frequency of the photon) has to fit the energy step in the material. It is this fact that is used when identifying different atomic/molecular species.
The basic principle described above has a wide field of applicability. Common uses include environmental monitoring, process control, etc. In an attempt to further utilize the technique, a method has been developed which enables absorption spectroscopy measurements to be performed on gas in a porus solid material. The technique, referred to as Gas in Scattering Media Absorption Spectroscopy (GASMAS), was developed at Lund University and has since been investigated in several steps [2, 3, 4]. Compared to conventional absorption spectroscopy, the path-length traveled by the light is not known. The GASMAS technique can be used in practically any solid material where gas-pockets are present, for instance organic materials (such as

wood, fruit) or to study human sinus cavities.

1.2 Motivation of thesis

The aim of this thesis is to investigate the possibility of detecting two gases simultaneously and to experimentally implement it. The reason for developing a method for dual measurements is the need for absolute measurements. If the concentration of one gas is previously known, the result of this known concentration can be used to calibrate the measurement of an unknown concentration.
There is no absolute demand on the measurements for the two species being done simultaneously, they must though be closely spaced in time, so that the measured object does not change between measurements.
There are some limitations set on the method, as a consequence of the final goal, which is to implement the method into a setup for making measurements on the human sinuses. This makes some approaches more suitable than others. The choice of light source, signal manipulation and lasers are all influenced by the final goal in mind.

1.3 Scope of thesis

This master's thesis is divided into more or less distinct parts. First, the light source used, a diode laser, is described. Here a couple of important different diode lasers are described in some detail. In the two subsequent sections, the process of light interacting with materia is treated from a practical standpoint. In these sections one can draw information on what can be expected when making a GASMAS measurement. The signals registered in absorption spectroscopy are mostly very weak and there is a constant battle to increase the signal and decrease the noise. Different techniques to this affect are described in Sect. 5. In this section, both the nature and origin of the noise is treated, as well as ways of reducing the impact of it. One of the problems arising when measuring at two wavelengths, namely how to normalize the signal, is also dealt with here. In Sect. 6 the GASMAS technique is investigated and the procedures used are studied. The following section is dedicated to shedding some light on the different ways of accomplishing two-wavelength measurements. In Sect. 9 the experimental setup is described. This includes the units used, as well as various problems which have arisen and a description of how they have been overcome. Besides this,

much of this section is dedicated to the techniques of deciding different parameters crucial for the system to work satisfactorily. Sect. 10 deals with the measurements made. Here the nature of the measurement objects used are described, as well as how the measured values should be interpreted. The measurements made and what conclusions can be drawn from them are presented in Sect. 11. This is followed by a section where future improvements to the system are suggested. The paper finishes with an Appendix, in which the computer-control programs are described in detail.

Chapter 2

Diode Lasers

Semiconductor lasers have the advantage over many other types of lasers, in that they have a compact design, are reliable (robust and stable over time) and can be tuned in wavelength [5]. Their applicability and widespread use have made them rather inexpensive as well. Modern diode lasers are made with simultaneous use of two different materials, from two different groups of the periodic table, for instance

Group	Wavelength
III–V	630-1600 nm
II–VI	"blue green"
IV–VI	4-29 μm

Here the first column refers to the groups in the periodic table which the semiconductors belong to . The use of a laser in gas absorption spectroscopy, rather than a conventional light source is absolutely imperative. The single-mode diode lasers used in high-resolution absorption spectroscopy have a linewidth smaller than what is otherwise achievable (Sect. 2.4.2).

2.1 Basic principles

Diode lasers are based upon the principle of stimulated emission in a semiconductor material. Electrons are, by adding energy, lifted form the valence band to the higher lying conduction band. Changing the electron distribution this way is referred to as pumping of the material. This can be accomplished in a number of ways (see Sect. 2.2). The exited electrons have a finite lifetime and will return spontaneously to their initial state. If a photon interacts with the electron when in its exited state, the relaxation energy given off will

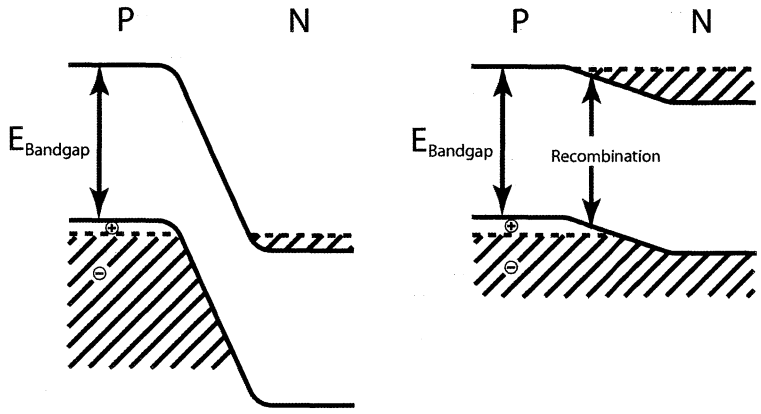

Figure 2.1: *Left: No voltage is applied to the junction and the holes in the valence band are separated from the electrons in the conduction band. Right: A voltage is applied and electrons will be injected into the conduction band and they will recombine with holes in the valence band. In both pictures dashed lines indicates the quasi-Fermi levels.*

be in the form of an identical photon. When the stimulated emission described above dominates over the absorption in the material, laser action is achieved. The difference in energy between the valence and conduction-band, the bandgap, will be inversely proportional to the wavelength of the emitted photon. The bandgap will be determined by the choice of semiconductor material as well as the present temperature. The material used must have a direct gap, which means that electrons are able to jump between bands without phonons (lattice vibrations) being added.

2.2 Pumping

In order to lift the electrons from the valence to the conduction band, energy has to be added to the semiconductor. This can be done in a number of fashions; by the use of another laser, an electron beam or by simply connecting the semiconductor like a diode and sending a current through it in the forward direction. A diode consists of two semiconductor slabs with different doping (adding positive or negative carriers), fused together. The area where the slabs are fused are called a junction and the vicinity of the junction makes up the active region of the laser. By applying a current over the junction, electrons will be added to the conduction band and holes (lack of electrons) will be added to the valence band.

2.3 Lasing condition

When stimulated emission described above dominates over absorption in the material, laser action is achieved. This "threshold" can be calculated by the use of quasi-Fermi levels. These levels indicate at what point inside the valence bands there are occupied levels below and no electrons above. The same is true if the Fermi level is situated in the conduction band. Doping a semiconductor will affect the position of its Fermi levels. By significantly doping a semiconductor slab with positive carriers, the Fermi level will be situated in the valence band. If negative carriers are introduced instead, the Fermi level will shift to the conduction band (see Fig. 2.1). The Fermi levels also dependend on the pump rate and the threshold condition is given by

$$E_g \leq hv \leq E_{F_c} - E_{F_v} \tag{2.1}$$

Here E_g is the energy corresponding to the bandgap, F_c is the quasi-Fermi level in the conduction band and F_v is the quasi-Fermi level in the valence band. hv denotes the energy lost when creating a photon. It can be deducted that the pump rate must be equal to the energy gap in the semiconductor, in order for the Fermi levels to be satisfactorily separated and for recombination to take place (see Fig. 2.1).

2.4 The composition of a diode laser

There are two different principles how to construct a diode laser; the homojunction (using only one material, p and n doped) and heterojunction approaches (using two different materials, p and n doped) [5]. The homojunction laser was the first semiconductor laser constructed. It has a major drawback, it has to be operated at low temperature (less than 77 K, since higher temperature will brake down the junction), which has made it obsolete, hence it will not be discussed here. The diode lasers used today are of the double-heterojunction type, discussed below.

2.4.1 Double-heterojunction lasers (DH)

The difference in construction for a double-heterojunction laser compared to a homojunction laser gives the former a threshold about two orders of magnitude lower. The reason behind this is that the active layer material is sandwiched between another material with a different band-gap and reflective index. These differences give the active layer a guiding quality, which

keeps the photons from being lost to the non-gain part of the structure. However, the basic principle of how laser action is achieved is the same for semi conductor lasers.

2.4.2 The cavity

In order to get a strong laser signal the light amplifying material has to be placed inside a cavity. For semiconductor lasers this is done by polishing the ends of the active region, and hence getting a cavity with reflective ends. A specific property of the diode lasers is that the outgoing beam has different divergence in the two directions perpendicular to the propagation direction. This is due to the fact that the active layer is not square, and thus giving different diffraction in the two directions. The lasers used for this project are single-mode lasers. This means that only one standing wave is experiencing gain in the laser cavity at a given time. This results in a very narrow bandwidth which is necessary in order to resolve the absorption peaks of free gas (Sect. 3.1).

2.4.3 Pigtailed laser

When high precision and a very low noise level is required a pigtailed laser (see Fig. 2.2) is used. The optical fiber is attached directly to the exit facet of the laser. By doing so one can maximize the amount of light entering the fiber as well as eliminating the path the light travels outside a controlled medium. If the light beam instead is coupled to the fiber via a lens the beam propagates through air, giving a faulty signal when measuring O_2 concentrations. Since the fiber will be glued on to the laser, the pigtail arrangement will have a higher degree of stability. Another approach to eliminate the spurious O_2 is to flush the path between the laser and the fiber tip with N_2. There are some problems associated with pigtailed lasers, such as inability to chop the light with a rotating disc and a significantly higher price.

2.5 Wavelength tuning

As mentioned earlier (Sect. 2.1) the laser wavelength depends on the temperature and the driving current applied. This is because of two different effects. The cavity length will depend on the temperature as well as the gain profile (dictates which mode will experience gain). Normally the laser is run so that the coarse tuning will be done by controlling the temperature, and the fine tuning with the drive current. The current flowing through the

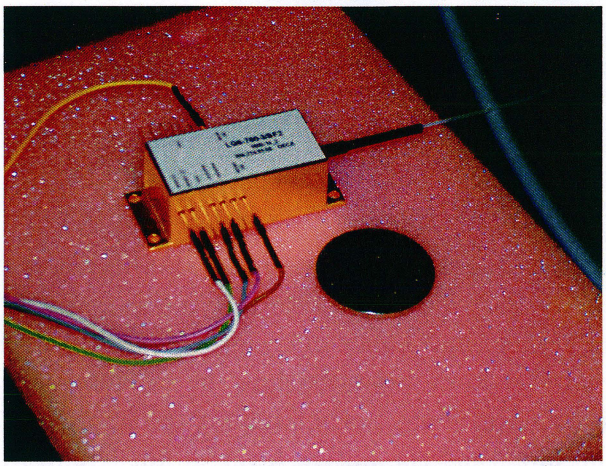

Figure 2.2: *A pigtailed laser for oxygen spectroscopy at 760 nm. A Swedish coin (five kronor with a diameter of 28 mm) is put next to the laser in order to appreciate the scale*

diode serves here not only as a means of adding energy, but also as a mean of heating it. Typical tuning speeds are 0.01 nm/mA and 0.06 nm/^{o}C, giving a tuning window in the order of a few nm (see Fig. 2.3). Due to the fact that the two tuning speeds are not identical, mode-jumps occur, that is the laser periodically jumps wavelength sections.

2.6 Distributed Feedback Lasers (DFB)

The distributed feedback laser is devised in much the same way as the DH laser. However, in the DFB laser the sandwiching material on one side of the active medium has a periodic change in its thickness. In addition, the periodicity is broken by a phase shift, halfway through. This will constitute two Bragg gratings which will impose a condition being set on the cavity. The condition will only make it possible for one mode to exist. The DFB laser has not got the problem with mode-hopping mentioned in Sect. 2.5. The main reason for using a DFB laser is to get a device with good single-mode operation. The single standing wave will, however, lead to hole burning in the gain medium, reducing the laser intensity. Therefore, a DFB laser will provide higher stability, though at the cost of lower intensity.

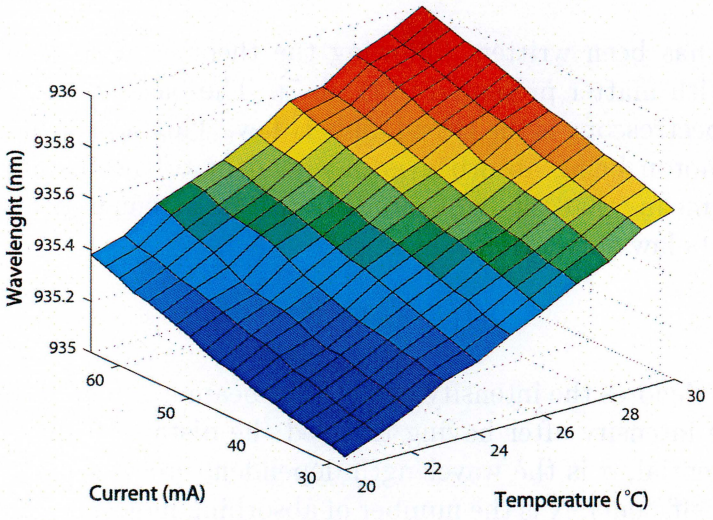

Figure 2.3: *A typical tuning diagram for a diode laser operating at 935 nm. The wavelength is given in nanometers for different temperatures (°C) and operating currents (mA).*

Chapter 3

Absorption spectroscopy

This section has been written, following the theoretical treatment of light interacting with matter presented in [1, 6, 7]. The basis for all quantitative absorption spectroscopy is the Beer-Lambert law. This states that the propability of a photon being absorbed is independent of the distance traveled. This in turn means that the light intensity will be attenuated exponentially. Beer-Lamberts law states that:

$$I(L) = I_0 e^{-\sigma NL} \tag{3.1}$$

In Eq. 3.1, I_0 denotes the intensity of the light beam as it enters the absorber and $I(L)$ the intensity after having traveled the distance L in the medium. In the exponential, σ is the wavelength dependent cross-section (that is the absorption coefficient), N is the number of absorbing molecules and L is again the distance traveled. The background to the wavelength dependent σ is that different wavelengths correspond to different molecular transitions, which all have their own transition probability.

3.1 Solid- vs. gas-phase absorption spectroscopy

The absorption spectroscopy performed on gases differs a lot from absorption spectroscopy done on liquid or solid materials. When molecules are in liquid or solid phase they are constantly interacting through collisions. This leads to an exited molecule being deexited very much faster, than it would have been in a gas. The fact that the sharpness of the absorption lines are inversely proportional to the lifetime of the exited state, results in a almost continuous absorption spectrum for non gaseous absorbers. The time spent in an exited

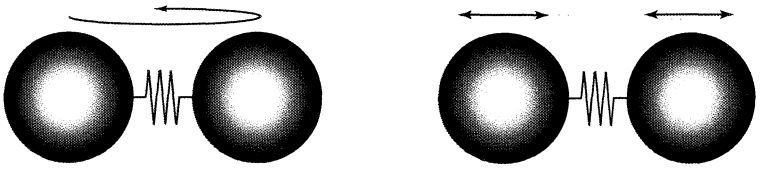

Figure 3.1: *The origin of the rotational (left) and vibrational (right) energies for molecules. The springs in the schematic picture are representing binding forces.*

state is not the sole reason for the smearing of the absorption lines, but perturbation of the energy levels (due to for instance the Stark effect (electric fields of molecules interacting)) will also play a role. Absorption peaks in gas is usually 10,000 times narrower than those corresponding to solids and liquids. There are several broadening mechanisms concerning gases as well, including the natural broadening mentioned above. Broadening mechanisms will be treated theoretically in Sect. 3.4.

3.2 Ro-vibrational transitions

There are some fundamental differences between atoms and molecules when performing spectroscopic measurements [1]. When the atoms only have electronic transitions, the molecules have apart from electronic, also rotational and vibrational transitions (see Figs 3.1 and 3.2). The different transitions can be explained according to the Born-Oppenheimer approximation, which states that they can be treated individually and then superpositioned. They are separable, due to the fact that the processes are taking place in very different timeframes, leaving each other unaffected.

3.2.1 Rotational transitions

When a molecule with two or more atoms rotates, an energy will be associated with the rotation. The system must be treated quantum mechanically and therefore the Schrödinger equation is being used. If the solid rotor approximation is made, meaning that the distance between the atoms is presumed constant, the rotational energies are:

$$E_r = hcBJ(J+1) \tag{3.2}$$

Here J is an integer given by $(J = 1, 2...)$; it can be changed through transitions in a way that $\Delta J = \pm 1$. B is the rotational constant given by

$$B = \frac{h}{8\pi^2 cI} \tag{3.3}$$

Here I is the moment of inertia. This is equal to the reduced mass μ (Eq. 3.4) times the distance between the atoms squared.

$$\mu = \frac{m_1 m_2}{m_1 + m_2} \tag{3.4}$$

In a more exact calculation of the rotational energy levels, the distance between atoms cannot be assumed to be constant. The centrifugal force will increase the distance as the rotational energy increases. A correction is made by the introduction of a small second-degree term:

$$E_r = hc[BJ(J+1) + DJ^2(J+1)^2] \tag{3.5}$$

Here D is a small (in comparison to B) correction constant.

3.2.2 Vibrational transitions

As stated in the previous section the inter-molecular distances are not rigid. This also manifests itself by the fact that atoms in a molecule vibrate in relation to each other. In molecules with more than two atoms there is bending and twisting, as well as the pure vibration. Commonly the vibration of atoms is approximated by a harmonic oscillator (that is a sinusoidal motion). In quantum mechanics the energies of a harmonic oscillator is given by

$$E_v = h\nu(v + \frac{1}{2}) \tag{3.6}$$

Here ν is the vibrational frequency. The vibrational quantum number (v) can change as $\Delta v = \pm 1$ in a transition. The energies based on a pure harmonic oscillator can be refined by adding higher order terms to Eq. 3.5. This correction will deal with the model discrepancies at atomic distances far from equilibrium. Since the vibrational motion will change the atomic distances it will have an effect on the rotational energies. This is included in the rotational energies by introducing corrected values of B and D.

$$\begin{aligned} B_v &= B - \alpha(v + \frac{1}{2}) \\ D_v &= D - \beta(v + \frac{1}{2}) \end{aligned} \tag{3.7}$$

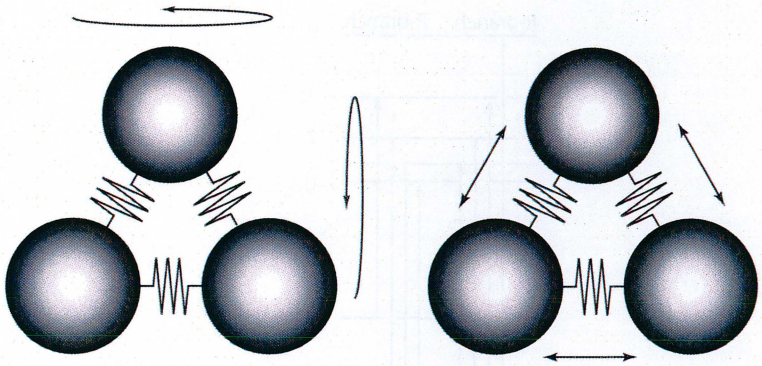

Figure 3.2: *The rotational (left) and vibrational (right) energies for three atomic molecules. This figure, in comparison to the one in 3.1, highlights the increased complexity when larger molecules are studied. The springs in the schematic picture are representing binding forces.*

Here B_v and D_v are the corrected values of B and D. The correction constants α and β are small compared to B and D. When molecules with more than two atoms involved are studied, there are more than one vibrational mode. Any molecule has $3N$ degrees of freedom, N being the number of atoms forming the molecule. Since three degrees of freedom is taken up by translational movement of the molecule as a whole and three by spinning around its own axes (only two for molecules with two atoms), there are $3N - 6$ vibrational modes. This gives three atomic molecules, such as H_2O, a more complex absorption spectra than two atomic ones, such as O_2.

3.2.3 Ro-vibrational spectra

Rotational-vibrational transitions have energies corresponding to wavelengths in the infrared. The frequencies for the two different types of transitions are given in inverse centimeters (cm^{-1}) by

$$\bar{\nu}_v = \frac{E(v')}{hc} - \frac{E(v'')}{hc} \tag{3.8}$$

$$\bar{\nu}_r = \frac{1}{hc}[E_{rv'}(J') - E_{rv''}(J'')] \tag{3.9}$$

Here v' denotes the vibrational quantum number of the upper state and v'' correspondingly for the lower one. If the frequencies in Eqs 3.8 and 3.9 are added and the expression for the rotational and vibrational energies are adopted from Eq. 3.2 and Eq. 3.6 respectively, the total transitional energy (see Fig. 3.3) can be written:

$$\bar{\nu} = \bar{\nu}_v + B'J'(J'+1) - B''J''(J''+1) \tag{3.10}$$

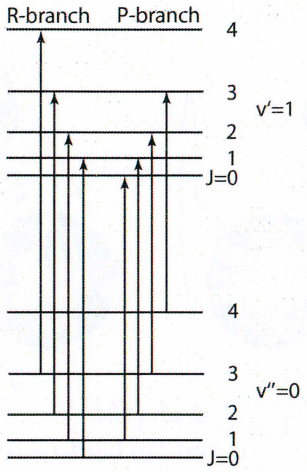

Figure 3.3: *Rotational and vibrational energy levels for the ground and first exited vibrational level.*

For a more general description of the molecular energies the electronic energy is added to the vibrational and rotational, to form the total energy of the molecules. This can be done by simply adding the different energies, as long as the Born-Oppenheimer approximation can be used.

3.3 Electronic transitions

Electronic transitions have much higher energies than the previously discussed rotational and vibrational transitions. The former will have energies in the visible and the latter ones in the microwave and infrared regions. The electronic transitions will normally be associated with ro-vibrational transitions as well, leading to a band-structure in the visible. The electrons are much smaller than the nucleus and electronic transitions are therefore taking place without the nucleus having time to react. This is the essence of the Franck-Condon principle. It manifests itself in a way that the transition probability for a certain electronic transition depends on the distance to the nucleus and the wave-functions for the electronic states involved. The transition will be dependent on several other properties, besides those stated above; for instance if the wave-function is even or uneven (see Fig. 3.4), the multiplicity of the state, etc.. In the oxygen transition used on the investigations of this paper, the multiplicity changes from triplet to singlet . The difference in multiplicity of the states makes the transition highly unlikely (forbidden) and the absorption imprint will therefore be small.

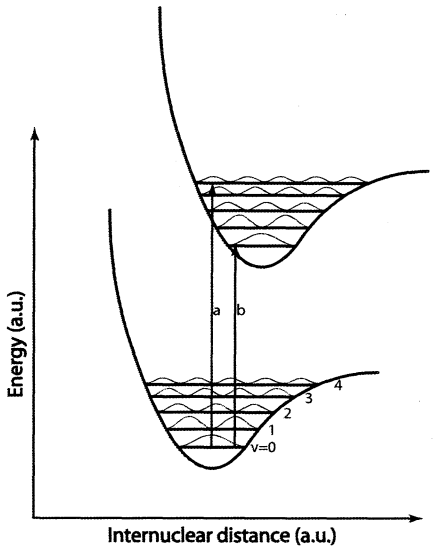

Figure 3.4: *The wavefunction associated with the vibrational states will not change as the electron jumps between states. The electron can thus only make transitions between states where the wavefunctions have a large overlap. This is the case for transition a), but not for the poor overlap of transition b).*

3.4 Line-broadening mechanisms

There are three different types of line-broadening mechanisms. The first is the natural linebroadening which only depends on the transition used. The second and third are the Doppler broadening and the pressure broadening, which each dominates under different conditions [2].

3.4.1 Natural broadening

The natural line broadening is a direct consequence of Heisenberg's uncertanty relation, which states that

$$\Delta E \Delta t \geq \frac{\hbar}{2} \tag{3.11}$$

The relation dictates that the combined uncertainty in time and energy cannot be made arbitrarily small. That is the precision in an ideal energy measurement is reversely proportional to the measurement time.
The time atoms and molecules spend in exited states are not infinite and the uncertainty in energy will depend on the excitation time of the transition. The natural line broadening will thus give an absolute restriction on the frequency resolution.

3.4.2 Doppler broadening

The Doppler broadening mechanism arises since the molecules are not at a standstill while absorbing light. Since different molecules in the gas travel at different speeds and directions, there will be a possibility for a range of frequencies to be absorbed. The broadening is due to individual effects and is therefore referred to as inhomogeneous. Doppler broadening is proportional to the square root of the temperature and will be the dominant effect when the pressure is low. It is also linearly dependent on frequency. The Doppler broadened peak will have a Gaussian lineshape and a half width at half maximum (HWHM) given by (3.12)

$$\Delta v_D = v_0 \sqrt{\frac{2kT}{Mc^2} ln2} \tag{3.12}$$

Here Δv_D is the frequency broadening, v_0 is the central frequency, T is the temperature and M is the molecular weight.

3.4.3 Pressure broadening

Pressure broadening is an effect which is connected to the natural line broadening. The latter one is a consequence of Heisenberg's uncertainty principle, given in Eq. 3.11. When the pressure increases the molecules will experience a growing interference from neighboring molecules. This interference leads to shorter excitation times and hence, by Heisenberg's uncertainty principle, also broader line profiles.

$$\Delta v_L = \sum_i \gamma_i p_i \tag{3.13}$$

The line profile is a Lorentzian and the line width is proportional to the partial pressure p_i with the constant of proportionality γ_i. The partial pressure is the pressure exerted from the individual gas species i. The pressure broadened line is, however, influenced by temperature as well.

$$\Delta v_L(T) = \Delta v_L(T_0)(\frac{T_0}{T})^n \tag{3.14}$$

Here T is the temperature and n is a transition- and species-dependent coefficient (equal to 0.7 for O_2). The pressure broadening is proportional to the pressure, reversely proportional to the temperature, but independent of frequency.

3.4.4 Voigt profile

In the previous two sections the two major contributions to line broadening have been stated. The Doppler broadening dominates when the pressure is low and at high frequencies and the opposite is true for the pressure broadening. In the intermediate case, that is when neither the Gaussian nor the Lorentzian will give a satisfactory solution, a convolution of the two is used. This convolution is called a Voigt line profile. The HWHM can be calculated from the linewidths given above (Eqs 3.12 and 3.14). However, the rather lengthy equation will not be given here [2].

Chapter 4

Light transport in scattering media

The interaction between light and a medium is characterized by three parameters, absorption events per unit length (μ_a), scattering events per unit length (μ_s) and the scattering angle distribution (g) [8]. The scattering medium considered here is primarily tissue. Tissue is forward scattering, that is a scattering event mostly contributes to a small angular change. The scattering medium in case of infrared radiation is characterized by a scattering coefficient several orders of magnitude larger than its absorption coefficient. Both the absorption and scattering are wavelength dependent, though there is a significantly larger wavelength dependence for the absorption coefficient. There are two reasons for light scattering in tissue, first, the quantum mechanical interaction of light with molecules and secondly, index of refraction variations. The later is due to the heterogeneousness of tissue, leading to variations in the same or larger scale as the wavelength of light being used. The scattering events considered here are elastic, that is the photon energy and hence wavelength remains unchanged by a scattering event. There are three different types of elastic scattering, resonant absorption, Mie and Rayleigh scattering [8]. When a photon has a wavelength corresponding to a resonance in a atom, the atom will upon impact stay in its resonant state until it sends out a identical photon. Mie scattering is based on the Maxwell equations and is due to the interaction of electromagnetic fields. It is used to describe scattering events when the wavelength is small in comparison to the particle size. For Rayleigh scattering, on the other hand, the photons are viewed as particles striking other particles. The Rayleigh model is used when the wavelength is large in comparison to the particles.

Many scattering events result in a complete randomization of phase, polarization, temporal and spatial properties. There are two commonly used ways to

describe photon travel through a diffuse (highly scattering, low-absorption) material; analytical calculations based on Maxwell's equations, or transport theory [2] The latter one uses a classical approach, treating photons as particles. In doing so, less demanding mathematics is used and it is possible to do calculations on complex structures. The transport equation is:

$$\frac{1}{v}\frac{\delta}{\delta t}L(r,s,t) = -s \cdot \nabla L(r,s,t) - (\mu_a + \mu_s)L(r,s,t) \qquad (4.1)$$
$$+\mu_s \int_{4\pi} L(r,s',t)p(s,s')d\Omega' + Q(r,s,t)$$

Here v is the photon velocity, $p(s,s')$ is the probability that a scattering event will change the direction from s to s'. Eq. 4.1 describes what happens to photons at depth r, traveling in direction s at time t. The first terms on the left hand side of the equation represents the photon loss outside the boundaries, the second is the loss due to scattering and absorption. The third term adds the photons scattered in from another direction, and, finally, the fourth term gives the photons added by the light source. This equation provides a statistical approach often used by computer-aided models. However, it may also be solved analytically.

Chapter 5

Noise and noise reducing schemes

When performing measurements where a small signal is to be registered, the elimination of noise is always important [9]. This is especially true when the interfering noise behaves similarly as the desired signal. There are several different types of noise and ways for it to affect the signal.

5.1 Sources of noise

There are two major groups into which different kinds of noise can be divided: external noise and interference feed to the setup from the environment, and intrinsic noise inherent to the experimental setup. Some of the common types of intrinsic noise are listed below:

- *Thermal noise*
  This is due to the thermal agitation of charge carriers in the detection system, inducing a "dark current". The thermal noise is independent of frequency and it has a white noise character.

- *Shot noise*
  Shot noise is an effect of the quantum nature of light, which is detectable when the detection rate is very low.

- *Flicker noise*
  Diode lasers yield an inherent noise the intensity of which is inversely proportional to the frequency. The (1/f) laser noise together with similar external noise makes up the flicker noise. It arises as a consequence

of fluctuations in carrier and photon density, refraction index, partition noise (noise due to mode hopping), laser injection current and laser temperature.

- *RAM noise*
  RAM noise is a consequence of the modulation techniques treated in Sect. 5.2.

The other type of noise, the external noise is to a high degree eliminated by isolating the measurement setup form the environment. There are a five ways in which the external noise can couple to the measurement signal, these include, capacitive coupling, inductive coupling, resistive coupling or ground loops, microphonics (vibrations) and thermocouple effects. By using proper cables and a correct laboratory environment these effects can more or less be avoided completely. There are, however, optical phenomena which are harder to get to grips with:

- *Interference fringes*
  Fringes constitute a considerable problem when making diode laser measurements. Whenever optical fibres or any type of optical device with parallel reflective surfaces is used, fringes will appear. This arises since reflection between parallel surfaces will lead to standing waves which will interfere, adding an oscillating disturbance to the signal.

- *Fibers*
  The measurements performed in this paper are done using multi-mode fibers. Such fibers will behave slightly different for various wavelengths and therefore corrupt the signal. This can be avoided by using a single mode fibre. The single-mode fibers do, however, have a smaller diameter, leading to a more difficult coupling situation, as well as diminished light transmittance. The difficulty of coupling light into the fibre is circumvented by using pigtailed lasers.

5.2 Noise elimination techniques

The effects of noise can be eliminated in two different ways, either by increasing the signal or decreasing the noise. Both methods will lead to a better signal-to-noise ratio (SNR). The way of increasing the signal when performing absorption measurements on a particular spectral line is to make the absorption path longer. However, this will not be possible in the situation focused on here. Another commonly used way of increasing the SNR is to shift the detection window towards higher frequency and thus avoid the

flicker noise. This is done by modulation techniques, discussed in Sect. 5.3. In some setups, the light from the laser is divided into two parts. The major part of the light is used to make the absorption measurement, while a smaller part is lead directly to a second detector. By subtracting the absorption signal with the second appropriately scaled signal, the common noise can be eliminated. Common noise is here referring to fringes and other effect which are identical in the two arms of the setup, therefore white noise such as shot-noise will not be eliminated, but rather amplified. The method is called balanced detection and will be treated separately in Sect. 5.5. Another way of increasing the signal, apart from lengthening the absorption path, is to use as strong an absorption line as possible. The absorption lines within a specific wavelength region may vary widely due to different transition probabilities for the different changes in energy. The strength of a specific absorption line is in turn reversely proportional to its width. In other words, the integrated absorption will be the same independent off the line broadening, but the absorption strength of the peak will go down as the line spreads out in frequency.

5.3 Lock-in amplifiers

A lock-in amplifier is basically a phase sensitive bandpass filter [9]. The combination of identifying a AC signal using both frequency and phase makes the apparatus very sensitive, and enables correct evaluation of signals drowned in noise. When using lock-in detection in diode-laser-based absorption experiment, a sinusoidal current imposes a variation in the laser wavelength. By feeding the same current as a reference to the lock-in amplifier, phase-sensitive detection can be accomplished. The signal to be detected and the reference can be written as:

$$V_{sig} = sin(\omega_{sig}t + \theta_{sig}) \tag{5.1}$$

$$V_{ref} = sin(\omega_{ref}t + \theta_{ref}) \tag{5.2}$$

Here ω denotes frequency and θ phase. The phase of the reference is not the same as the phase of the signal. The reference feed to the lock-in is not directly used but merely serves as a templet for the internally generated reference in the lock-in. In this process the reference loses its phase, which has to be set manually. The lock-in amplifier multiplies the two signals:

$$V_{psd} = V_{sig}V_{ref}sin(\omega_{ref}t + \theta_{ref})sin(\omega_{sig}t + \theta_{sig})$$

$$= \frac{1}{2}V_{sig}V_{ref}cos([\omega_{sig} - \omega_{ref}]t + \theta_{sig} - \theta_{ref}) - \frac{1}{2}V_{sig}V_{ref}cos([\omega_{sig} + \omega_{ref}]t + \theta_{sig} + \theta_{ref}) \quad (5.3)$$

The equation shows that when ω_{sig} is equal to ω_{ref} the V_{psd} will have a DC signal and an AC signal at the sum frequency. If V_{psd} is lowpass filtered the result is:

$$V_{psd} = \frac{1}{2}V_{sig}V_{ref}cos(\theta_{sig} - \theta_{ref}) \quad (5.4)$$

The signal arrived at is proportional to the intensity of the signal, V_{sig} and can be maximized by setting θ_{ref} equal to θ_{sig}.
The method of imposing a sinusoidal variation around a central value is referred to as modulation. There are a number of different ways in which a signal can be modulated, frequency (FM), wavelength (WM) and intensity (IM) modulation. The first two are closely related and differ only in the frequency to amplitude relation used on the modulation. If the wavelength of the modulation is large in comparison to the absorption feature, it is called FM. Subsequently, the lasers in this paper is wavelength modulated. This is, however, only semantics and the modulation will be referred to as FM.

5.3.1 The mathematical interpretation of the lock-in method

The signal processed by the lock-in is complex [10], but can using Fourier analysis be constructed by superpositioning pure sinus functions. The lock-in performs a convolution of the signal and a pure sinusoidal signal, as described in Eq. 5.3. By doing so using the first single Fourier component, the 1f signal is isolated. The Fourier components have similar appearance as the corresponding derivative of the signal, that is the n:th harmonic is similar to the n:th derivative (see Fig. 5.2). The reason for the derivative appearance of the Fourier components are that the frequency modulation will create a process closely resembling that of derivation (see Fig. 5.1). Often the second Fourier component is used instead of the first, causing undefined offset values to disappear, and thereby creating a pure AC signal. When registering the signal at the modulation frequency, the process as described above will be a pure derivation. However, a derivation is only exact if Δx approaches zero, this will not be the case here, since the modulation amplitude will be significant. This discrepancy leads to that the frequency modulation will not produce a clean sinusoidal amplitude variation (see Fig. 5.2). The residual will also be sinusoidal, but with the double frequency (plus to a lesser extent higher order Fourier components).

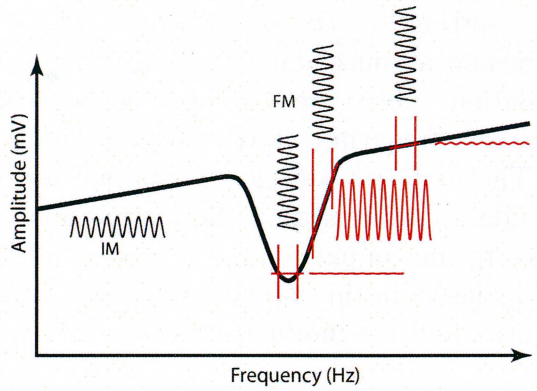

Figure 5.1: *The figure shows the ramped signal with an absorption feather in the middle, as it is received at the detector. The IM imposed on the laser current will be not generate a signal dependent on the absorption, but rather a constant sinusoidal intensity fluctuation. The FM of the laser will on the other hand not be seen where the signal is fairly independent to frequency changes. At the absorption dip, the signal will change significantly over a small frequency range and the FM modulation will generate a large amplitude fluctuation. This last statement is true in the beginning and end of the absorption peak. At the center of the absorption feature, no signal will be seen (see Fig. 5.2). Due to this the harmonics will have a derivative character.*

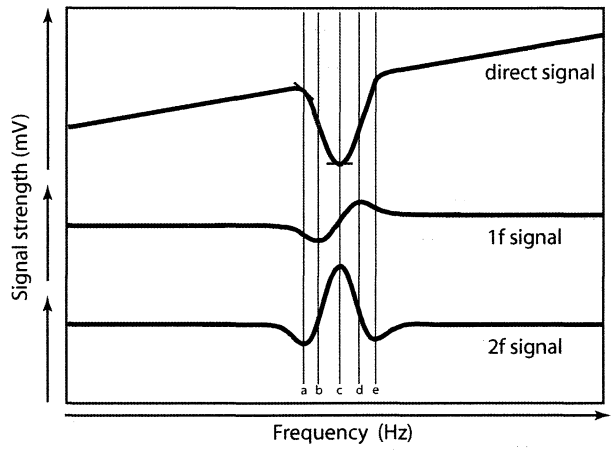

Figure 5.2: *The functions of the lock-in amplifier. The top curve is the direct signal, that is the unprocessed signal. The middle curve is the first harmonic, this is the gradient (first derivative) of the direct signal. The bottom curve is the second harmonic and it is corresponding to the deviation from the "straight line"-approximation made in creating the 1f signal. At b) and d) the gradient of the direct signal is maximal, which is reflected in the first harmonic. At a) and e) the "straight line" is a poor approximation, and even more so in c), hence the second harmonic has extreme points at a), c) and e).*

5.3.2 The nature of the signal

The full laser light intensity passed through the sample, taken for all frequencies, is referred to as the direct signal. The absorption signal can be seen as a dip in the otherwise constant direct signal. The 1f signal has the appearance of the first derivative and the 2f signal has the appearance of the second derivative of the absorption dip (see Fig. 5.2).

5.3.3 Calibrating the lock-in signal

In order to get a dimensionless and calibrated value for absorption the lock-in signal has to be normalized. It is of importance that the same signal is arrived at, no matter how much light passes through the sample. This is normally done by dividing the lock-in signal by the direct signal. By creating a dimensionless value, much of the change in different setups and so on can be canceled out. If the direct signal cannot be obtained, several studies [11] have shown it possible to normalize the second Fourier signal with the first Fourier signal (which has been found in this work as well). The 2f signal will consist of a pure AC signal, whereas the 1f signal will have an offset proportional to the total light intensity. The 1f signal will naturally contain the measured gas concentration as well, this information will, however, when normalizing, be disregarded in favor of the pure DC component.

5.4 Normalizing with the 1f signal

The reason for the 1f signal offset being proportional to the ramped light intensity is the modulation of the signal. The laser frequency is ramped by changing the drive current. By doing so the amplitude of the laser light will change as well (FM modulation cannot be made without getting intensity modulation (IM)). This means that a pure frequency modulation is not made but rather a mix between frequency and intensity modulation. (see Fig. 5.1) The amplitude change has two different constituents, a linear part and a non linear one, referred to as Residual Amplitude Modulation (RAM)(see Sect. 5.4.3). When, as is done here, the first harmonic is used to normalize the measurement, the linear part of the IM is used.
The intensity modulation will be a small wave on a big offset. If the offset is halved, the modulation amplitude is halved, leading to the linear IM signal being proportional to the total intensity. In turn, the 1f signal offset is proportional to the linear IM signal and a normalization can be made. The 1f signal contains all the information seen in the 2f signal. The reason for not using solely the 1f-signal, is the fact that the offset is large in comparison

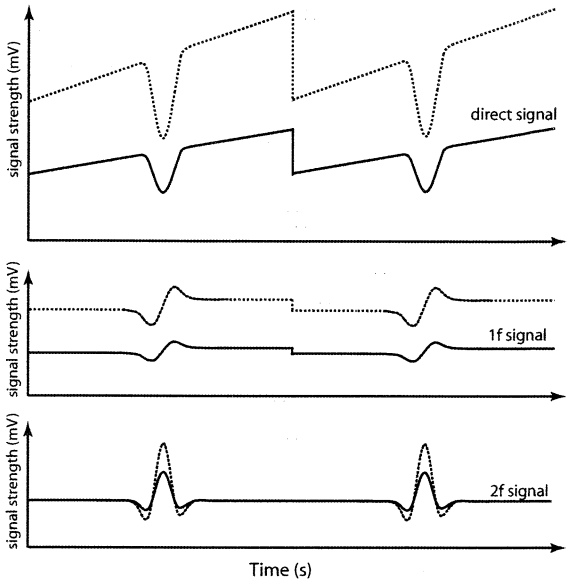

Figure 5.3: *Top picture: The direct signal. As the light hitting the detector is doubled, the absorption dip will also be doubled. It is in order to eliminate this effect a normalization is made. Middle picture: As the intensity is doubled the offset of the 1f signal will be doubled as well, making resemble the behavior of the direct signal. Bottom picture: The 2f signal will be linearly dependent on the amount of light hitting the detector, making some kind of normalization necessary.*

to the signal (AC part) leading to poor dynamic range when measuring the latter.
The FM and IM modulation will be manifested differently in the 1f signal. The signal due to the frequency modulation will be proportional to the absorption feature. The intensity modulation will, on the other hand, be the reason for an DC offset. The phase of the FM and IM signals will be shifted in relation to each other, typically 90 degrees [12].
An average taken over all frequencies for the 1f signal should be a good approximation when using it for normalization, since the base line is constant. Both aspects of the first harmonic will be linear in respect to frequency.

5.4.1 The 1f signal compared to the direct signal

The direct signal is proportional to the total laser output. The beam has been attenuated, as it has traveled through the setup. The 1f signal on the other hand is proportional to the modulation amplitude of the laser. This means

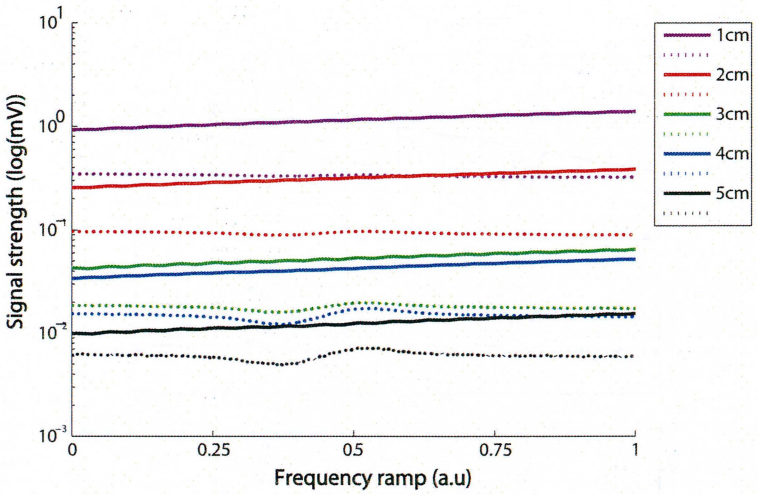

Figure 5.4: *The appearance of the 1f and direct signals, recorder with a transmission setup for various styrofoam thicknesses. The dashed lines are the 1f signals.*

that the two should only be separated by a factor. In Fig. 5.4 the direct and 1f signals have been recorded for a couple of different thicknesses of styrofoam. It can be seen that the average value for the two different signals seems to be linked. The averages as a function of thickness are presented in Fig. 5.5. This does show that the mean value of the two measurements are only separated by a constant factor.

5.4.2 Formulating

The transmission coefficient dictates the light intensity present at the detector. That is $I_{detector} = \tau \cdot I_0$, where I_0 is the laser output intensity and τ is the transmission coefficient.
When the laser frequency is being modulated according to

$$\bar{v} + \Delta v \cos(\omega t) \tag{5.5}$$

the transmission coefficient is according to [13]

$$\tau(v(t)) = \sum_{k=0}^{+\inf} H_k(\bar{v}, \Delta v) \cos(k\omega_m t), \tag{5.6}$$

where

$$H_0(\bar{v}, \delta v) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} \tau(\bar{v} + \Delta v \cos(u)) du \tag{5.7}$$

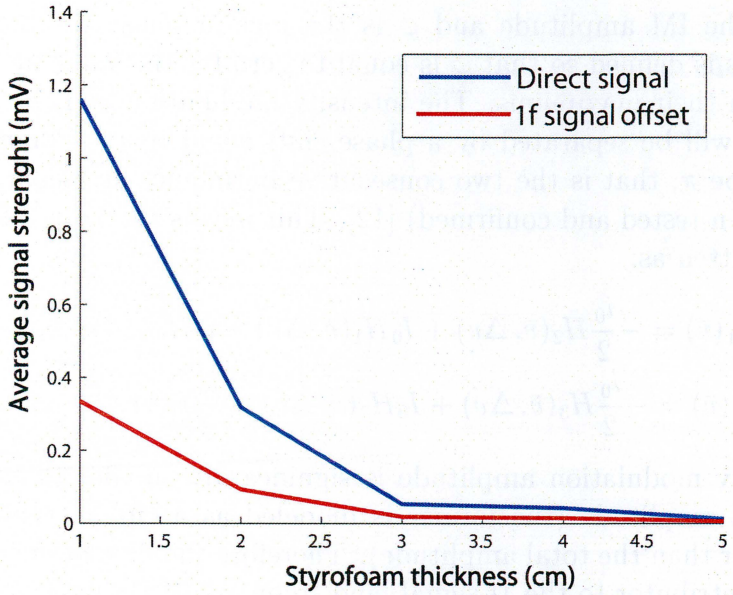

Figure 5.5: *Displaying the averaged taken over one frequency ramp (see Fig. 5.4) for the 1f and direct signals as a function of thickness. The to curves are only separated by a factor, hence confirming the possibility of using the offset of the 1f signal for normalization.*

$$H_k(\bar{v}, \Delta v) = \frac{1}{\pi} \int_{-\pi}^{+\pi} \tau(\bar{v} + \Delta v \cos(u)) \cos(ku) du, \ k > 0 \tag{5.8}$$

The index for H is referring to a harmonic, 1f, 2f and so on.

ω_m	Modulation frequency
$v(t)$	Frequency of the optical wave
$\bar{v}$	Average frequency of the optical wave
Δv	Frequency modulation amplitude

The first- and second-harmonic signals are low-pass filtered in the lock-in amplifier and the signals at the output appear as [13]:

$$S_1(\bar{v}) = \frac{i_0}{2} H_2(\bar{v}, \Delta v) cos(2\psi + \varphi) + I_0 H_1(\bar{v}, \Delta v) cos(\psi + \varphi) + i_0 H_0(\bar{v}, \Delta v) cos(\varphi) \tag{5.9}$$

$$S_2(\bar{v}) = \frac{i_0}{2} H_3(\bar{v}, \Delta v) cos(2\psi + \varphi) + I_0 H_2(\bar{v}, \Delta v) cos(\psi + \varphi) + \frac{i_0}{2} H_1(\bar{v}, \Delta v) cos(\varphi) \tag{5.10}$$

Here i_0 is the IM amplitude and φ is the lock-in phase for the 1f, 2f...-signals. This is defined so that φ is equal to zero for the value at which the signals reach their maximums. The intensity modulation and the frequency modulation will be separated by a phase shift equal to ψ. Commonly this is taken to be π, that is the two consecutive harmonics are separated by $\frac{\pi}{2}$ (this has been tested and confirmed) [12]. This means that Eqs 5.9 and 5.11 can be rewritten as:

$$S_1(\bar{v}) = -\frac{i_0}{2}H_2(\bar{v}, \Delta v) + I_0 H_1(\bar{v}, \Delta v) - i_0 H_0(\bar{v}, \Delta v) \tag{5.11}$$

$$S_2(\bar{v}) = -\frac{i_0}{2}H_3(\bar{v}, \Delta v) + I_0 H_2(\bar{v}, \Delta v) - \frac{i_0}{2}H_1(\bar{v}, \Delta v) \tag{5.12}$$

The intensity modulation amplitude is significantly smaller than the total laser output amplitude (the former is modeled as a ripple, typically 5-10 times smaller than the total amplitude). Therefore the first harmonic will be the main contributor to the 1f signal, and so on.
It can be seen from Eqs 5.7 and 5.8 that it is only the zeroth harmonic that has an offset. From Eq. 5.11 it can thus be deducted that the measured 1f signal does have an offset originating from the frequency modulation. However, this will be small, 20-30 times smaller than the offset due to intensity modulation. When measuring the DC signal originating from the IM, the small FM offset signal will be measured as well, resulting in an error in the order of 5%.

5.4.3 A detailed look at the 1f signal

As described in the prior sections, the 1f signal offset comes as a consequence of the intensity modulation. The laser intensity will have, due to nonlinear behavior, the appearance of

$$I_0(t) = \bar{I}_0[1 + i_0 \cos(\omega t + \psi_1) + i_2 \cos(2\omega t + \psi_2)] \tag{5.13}$$

where $I_0(t)$ is the laser intensity and i_0 and i_2 are the linear, and nonlinear IM amplitudes and ψ_1 and ψ_2 are the phase shifts between the FM and IM signals (linear respectively nonlinear). With correct notation only the i_2 term is corresponding to the RAM signal [14, 10], which is about 2% of the entire intensity modulation. It is also worth mentioning that it is because the nonlinear part is so small that it can be neglected, that it has not been accounted for in Eqs 5.11 and 5.12. If the nonlinear part is included, S_1 will also have a H_3 term and S_2 will have H_0 and H_4 terms. These added terms will be scaled by i_2.

It has been stated that the phase of the FM and IM signals will be shifted 90

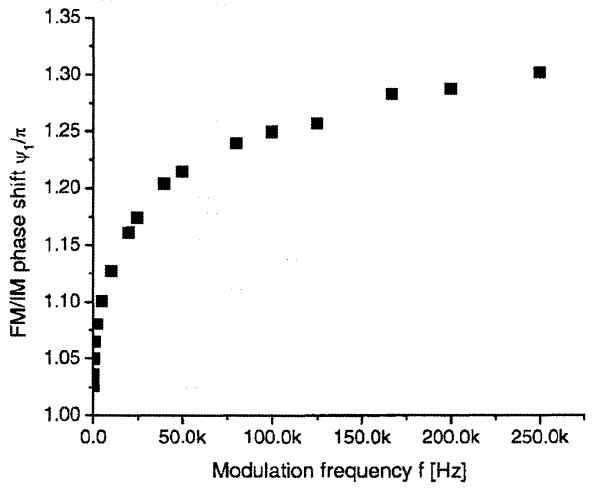

Figure 5.6: *The linear FM/IM phase shift as a function of modulation frequency. This curve is an example and the exact values may very well not coincide with the lasers used here [12].*

degrees in relation to each other. This is only true for the linear part of the IM generated signal. For the nonlinear, however, the situation is different. The nonlinear part is referred to as the RAM and is typically phase shifted 24 degrees in relation to the frequency modulation. Although the phase shifts are often taken to be largely universally constant, it is not the case. Both phases are actually depending on several experimental factors and should be determined through experiments [12]. It has been shown that the two phases vary as the center frequency, the modulation amplitude or the modulation frequency is changed. An example of the latter is illustrated in Fig. 5.6. For the case of using the 1f signal offset as a mean of normalizing, the linear part is used and hence it is ψ_1 which dictates the phase shift. The fact that it is not equal to 90 degrees poses some problems when one wants to make the signals perpendicular to each other (see Sect. 9.7).

5.5 Balanced detection

The purpose of balanced detection is to eliminate fringes, non-linear detector response and laser intensity noise [15]. The intensity of the fringe noise will decrease if anti-reflection coatings are applied or the surfaces are cut at an angle. However, this is not enough when performing trace gas detection. Another technique used to eliminate fringes is mechanical vibration followed by averaging (see Fig. 5.7). The vibrations will lead to a constantly changing

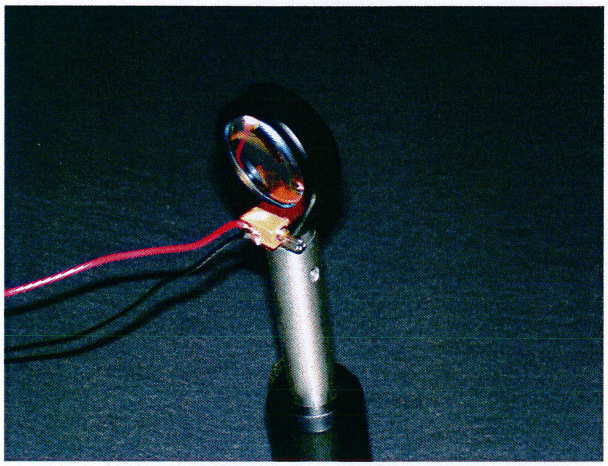

Figure 5.7: *A vibrator attached to a lens in order to eliminate fringes.*

path, eliminating the variations over time. However, this method cannot be used in junction with a pigtailed laser (see Sect. 2.4.3). The idea behind balanced detection is to split the laser beam into two parts, one part passing through the sample and one feed directly to the detector. The splitting is such that the reference beam will be significantly lower in intensity, since a high intensity is wanted through the sample. The reference part will contain the non-desired part of the signal (fringes, spurious signal and so on). By subtracting the signal with the properly weighted reference, a cleaner, fringeless signal is achieved. The result of the subtraction should be zero at wavelenghts where no oxygen absorption is present. However, balanced detection is not without its problems. The fluctuating signals make it hard to perform balanced detection in real time. This is therefore done as a part of the post measurement data processing.

5.6 Noise elimination by signal processing

When the signals have been collected and stored in the computer they have to be processed. In this process certain values are extracted from the signals (see Eq. 9.2). Even if the signal has the proper appearance, single values in the measurements can deviate from the general line, due to noise. To minimize the influence of noise, a "base function" is used. This is basically an ideal signal, taken for some arbitrarily large species concentration/distance traveled. By making a convolution between the base function and the noisy measurement, the base function is scaled to fit the appearance of the noisy signal. The base function now has (ideally) the properties of the noisy signal

minus the noise. The values are thus taken from this latter one.

Chapter 6

GASMAS

Gas in scattering media absorption spectroscopy (GASMAS) is a relatively new technique [16]. The scope of the method is to measure gas contents inside scattering materials. This is made possible by the different appearance associated with absorption by gas contra solid materials (see Sect. 3.4). GASMAS is not necessarily confined to studies in solid material. Gas dissolved in liquid would work equally well; this does, however, not have as apparent applications. The size of the gas pockets under investigation varies widely, from pores in wood [17, 3] and fruits [18, 19] to sinus cavities [20, 21]. The difficulty associated with making absorption measurements in scattering materials is that the Beer-Lambertian law is not directly applicable. This is because the length traveled by the light through the sample is not known. A demand on the experiment is that the absorption in the bulk material is rather small for the wavelength used. For materials such as tissue and other materials high in water content, wavelengths in the near infrared have to be used (below 1.4 μm). This imposes a second demand, the gas under investigation has to have strong, sharp absorption peaks in the wavelength region conventionally chosen. This is true for several gases of interest, for instance O_2, H_2O and CO_2.

6.1 Standard addition

A way of calibrating the GASMAS setup, in the sense of determining a absorption path length, is to use standard addition [1]. Standard addition is widely used both in chemistry and in absorption spectroscopy. Often the method is used in that a specimen with unknown concentration is measured, then a known concentration is added to the specimen, followed by a new measurement. By repeating the procedure and comparing the measurements

a concentration estimation can be made. Thereby a concentration scale is gained. The standard addition method is used in the same fashion in a GASMAS application, with the difference of not adding a known concentration, but rather a known distance of air. The distance traveled in the specimen is not actually measured and the standard addition value is referred to as equivalent distance in ambient air. This relation is expressed as

$$a = \sigma c_{air} N_{tot} L_{eq} = \sigma c_{sm} N_{tot} L_{sm} \tag{6.1}$$

Here a is the absorbance, c and L are the concentration and path length, respectively. The subscript sm indicate that the value is taken in the scattering medium, air is the ambient air and eq denotes equivalent distance in ambient air. When making a standard-addition measurement one would like the total change in path length to be large in comparison to the smallest laser-detector distance. In this way a calibration curve with high precision can be accieved. The standard addition measurement is of most importance, since it will be used to calibrate every subsequent measurement.

6.2 Transmission or backscattering detection geometry

When making absorption measurements in scattering media a transmission or backscattering detection geometry (see Fig. 6.1) is used. The difference lies in whether the detector is on the same side of the sample as the source or not. The transmission detection geometry is more straightforward, though not applicable to specimens of large thickness. The difference is especially prominent when a gas layer is extending over the entire plane. This means that all photons at the detector in the transmission detection geometry have passed through the layer. Meanwhile for the backscattering detection geometry, some photons will have passed through the gas pocket twice and some have not passed it at all. This makes it easier to evaluate the measurement situation in the transmission detection geometry. Another advantage for the transmission detection geometry is that no precautions have to be taken in order for the source not to block the detector. This is, however, not the case for the backscattering detection geometry. The source has to be located at the center of the detector and this has to be masked so that photons do not short cut to it from the source but rather have penetrated deeper. In the backscattering detection geometry, the photons registered can be somewhat controlled by varying the mask blocking the central part of the detector. Even with the use of a mask, the transmission detection geometry will result

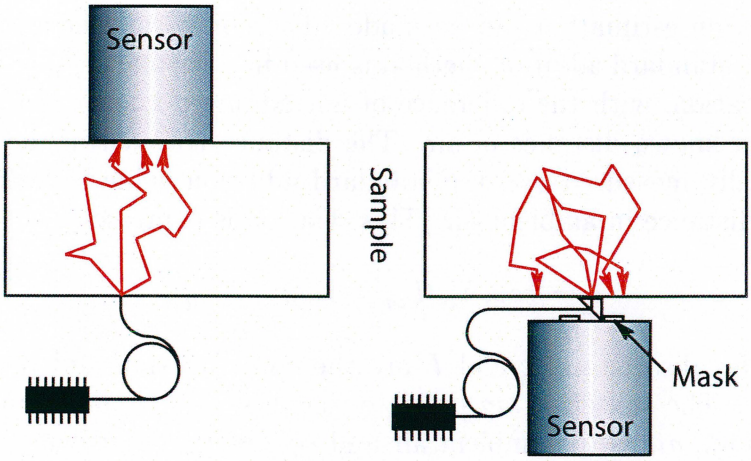

Figure 6.1: *Transmission (left) and backscattering (right) detection geometry.*

in a stronger gas imprint. This is foremost because of photons cannot reach the detector unless they have passed the gas layer.
The nature of the sample dictates how the setup should be arranged. The intended final measurement object for this setup (not to be reached in this specific project) is the human sinuses (see Fig. 6.2). A transmission detection geometry can be used for the maxillary sinuses, but reflection detection geometry has to be used on the frontal sinuses [20]. This situation arises since there is no way of getting on the other side of the frontal sinuses, whereas the maxillary sinuses are measured through the cheek. A factor that separates measurements on sinuses from other GASMAS applications is the nature of the gas pocket. In dealing with wood, fruits and so on the gas pockets are very small compared to the volume probed. The sinuses, however can be modeled as a non-scattering and absorbing slab, in an otherwise scattering and non-absorbing environment.

6.3 Etalon fringes

In Sect. 5.1 fringes were briefly discussed as part of a treatment of noise. The phenomenon can be used as a means of calibration, if the fringes are the result of a well-defined etalon. An etalon has two parallel surfaces with high reflection. The light entering the etalon will be multiply reflected and upon exiting, it will interfere with itself [2]. The interference is constructive if

$$2L = m\lambda, \tag{6.2}$$

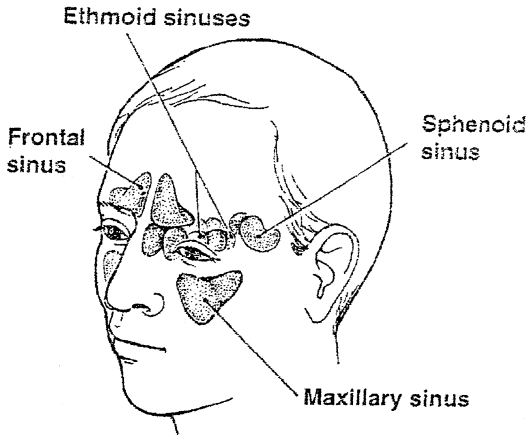

Figure 6.2: *A schematic picture of the different human sinuses [22].*

where the distance between the reflective surfaces is equal to L, λ is the wavelength and $m = 1, 2, ...$ By running a laser beam with the wavelength sweep mentioned in Sect. 9.1 through an etalon, a frequency scale can be obtained. This is possible by counting the number of fringes per second and using the known etalon free spectral range (fsr). The free spectral range denotes the distance given in frequency units between maxima (constructive interferences) and is given by

$$\Delta\nu_{fsr} = \frac{c}{2nL}, \tag{6.3}$$

where n is the index of refraction. With a calibrated frequency scale the half width at half maximum (HWHM) value can be calculated for the absorption signal. Here a foundation for receiving a high accuracy is to use ha etalon with high finesse, given by

$$F = \frac{\pi}{2arcsin(\frac{1}{\sqrt{f}})} \tag{6.4}$$

$$f = \frac{4R}{(1-R)^2}$$

The finesse is measuring the free spectral range compared to the HWHM value of the maxima peak, it is thus giving the contrast in the fringes received.

Chapter 7

Time-resolved measurements

Up till this point all measurements have been normalized and scaled in order for a comparison to be meaningful. To be able to determine absolute concentration values the average path length through the sample has to be measured (see Eq. 6.1). This can be done by a time-resolved measurement [2], where the time history of the photons is followed. The absorption mean path length (L_{sm}) can be calculated as

$$L_{sm} = \frac{c_0}{n} < t >, \tag{7.1}$$

where $< t >$ is the average time measured in the time-resolved method. A portion of the light is feed directly from the source to the detector, in order to get a start signal. The time relation between the start signal and the light distribution which has passed through the sample is measured. Making time resolved measurements complicates the setup considerably. Another method of getting absolute values by measuring two different concentrations, is discussed below, following the approach suggested in [3].

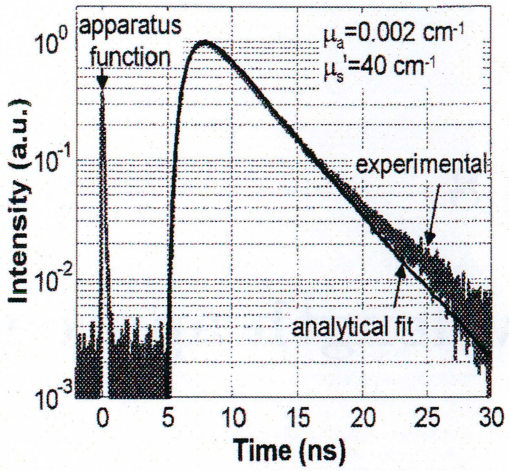

Figure 7.1: *A 39 mm thick polystyrene slab has been transiluminated by a pulsed light source. The figure shows the light pulse, as it is sent (apparatus function) and the appearance of the pulse as it has traveled through the polystyrene slab [4].*

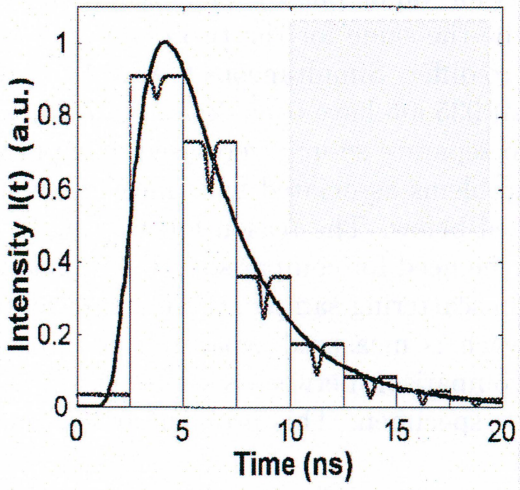

Figure 7.2: *A measurement taken in a transmission setup where short pulses have been used. The light intensity decays exponentially with time, the absorption feature does however show a different dependence. The light hitting the detector late will have traveled longer and will be more attenuated at the absorption frequency than the early light. This effect will compensate for the low light intensity for light arriving late [4].*

Chapter 8

Two-wavelength approach

The object of this paper is to investigate the possibility of making a GASMAS measurement with two different wavelengths at the same time in the same point. There is, however, not a total demand regarding there being no time separation at all; however, the time separation must be as short as possible. A time separation is acceptably short when system fluctuation and gas diffusion will not change the experimental situation between two measurements. The two gases to be measured is O_2 at 760 nm and $H_2O_{(v)}$ at 935 nm (v stands for vapor phase). The scattering coefficient in tissue will be assumed to be the same for the two closelying wavelengths, leaving only the absorption to differ. Simultaneous GASMAS measurement on O_2 at 760 nm and $H_2O_{(v)}$ at 935 nm have been done [1], however, not in exactly the same point, and two separate setups were used, complicating the situation. There are several problems associated with measurements of this kind and some will be discussed below. The reason for wanting to measure two gases "simultaneously" is the need for comparison. When measurements are being made on an unknown scattering sample, there is no way of knowing how the light travels. If oxygen is measured separately, no absolute values can be found. That is no comparison between samples can be made, only changes over time for a single specimen. This problem arises since the average path length is not known.

Considering water, the amount of water vapor can easily be determined. If there is liquid water in the sample and the temperature is know, the number of water molecules per volume unit can be calculated. Liquid water can be found for instance in the sinuses. Since the water vapor concentration is known, the effective average path length can be calculated by comparison with the GASMAS water vapor measurement. The water vapor measurements will hence serve as a calibration for the oxygen measurements. In other applications the opposite situation may occur, that is water vapor measure-

ments being calibrated by a measurement of a known oxygen concentration. There is of course no fundamental restriction to any specific gas. Any two gases could be used, by simply changing the wavelengths used. However, the concentration of one gas has to be known.

8.1 Signal separation

The most obvious difficulty in lock-in detection of signals due to two gases is how to distinguish one signal from the other. This can be done in either the frequency plane or the time plane. In using the latter option the signals are recorded intermittently and absorption measurements are made on a "every other" basis. In the frequency plane the signals can be separated by either using different modulation phase or modulation frequency (see Sect. 5.3). By having a 90^o phase difference between the two signals the lock-in amplifier will register one signal at once; see Eq. 5.4. If the signals are truly orthogonal, no crosstalk is theoretically possible. A different phase approach is restricted to the two wavelength case. By instead using different modulation frequencies any number of signals can be used in the system. This requires a frequency generator for every signal and a computer controlled switch to toggle between reference signals to the lock-in (multiplexing).
A critical part of the phase separation approach is the implementation of the 90 degree phase shift. A phase difference of 90 degrees is frequently used in telecom applications and there refereed to as IQ modulation. This can be done on the analog electric signal with the use of an all-pass filter. An all-pass filter will have a flat (constant) amplitude response and a sloped frequency response. This will give the possibility of tuning the phase shift for different frequencies with otherwise unchanged signal. In this paper the phase difference will be accomplished by simply generating two different signals, separated by an arbitrary phase. This is possible (with no increase in cost nor complexity) with the use of a waveform-generator computer card.

8.2 Time separation

When time separated measurements are performed, only one laser beam is passed through the sample at one time. The lasers cannot simply be turned on and off, since they require some time to reach a steady state. Due to this a mechanical chopper has to be used, which has to be synchronized with the reading of the signal. A mechanical chopper can, however, not be used in combination with a pigtailed-laser, since there are no gaps in which to put it.

8.3 Signal normalization

The signals used to monitor the absorption is the 2nd Fourier component provided by the lock-in amplifier. This is normalized by the direct signal in the one wavelength GASMAS case . When using two wavelengths the direct signal is a superposition of the two direct signals rendering it unusable. There have been several articles written on the subject of normalizing the 2nd Fourier signal with use of the 1st. This is unlike the direct signal not affected by superposition, since it is modulated. The lock-in amplifiers being used can measure both the 1:st and 2:nd Fourier signal, though not simultaneously. If in addition to this, one lock-in amplifier is used to register both signals, every measurement sequence will contain four measurements.

Chapter 9

The setup

The setup consists of a laser, a detector, a lock-in amplifier and an oscilloscope (see Figs 9.1 and 9.2). The different units are largely computer controlled. There is always a desire to rationalize a setup as far as possible, without loosing measurement quality. The reason for this is in part to create a small and manageable instrument, and in part to minimize production costs.

9.1 Lasers and laser control

In order to amplify and separate the signals originating from the two lasers, the lasers will have to be modulated.
The wavelength region for the laser is chosen by the temperature at which the laser is running. Inside this region the wavelength tuning is performed by controlling the laser diode drive current (see Sect. 2.1). While the temperature is kept constant during the measurement on an absorption peak, the drive current is repeatedly swept at approximately 4 Hz. This results in the laser wavelength constantly being swept across the absorption peak, rendering an updated signal feed to the oscilloscope. In order to modulate the frequency so that the lock-in amplifier can pick up the signal, the current ramp is superpositioned with a sinusoidal wave (see Fig. 9.3). This will have a frequency around 10 kHz, around 2500 times the ramp frequency. Due to the significant frequency separation, the two variations of the current will not influence each other.

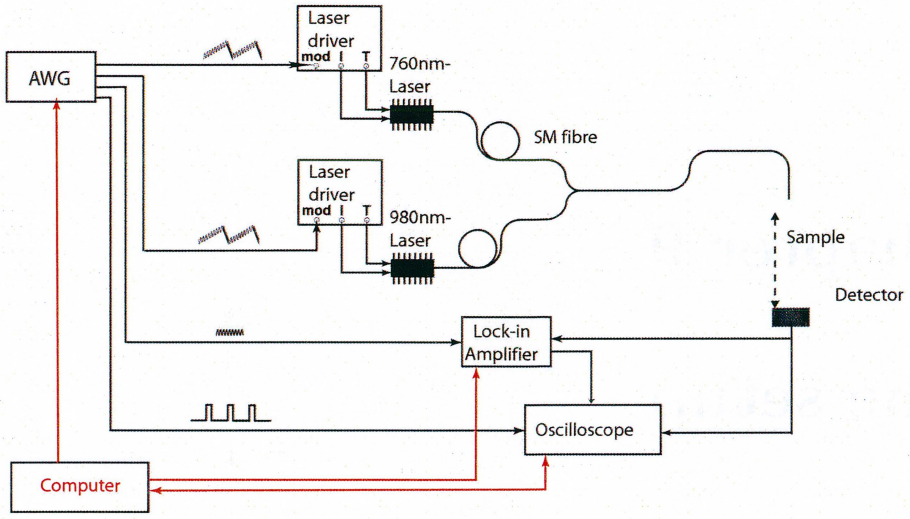

Figure 9.1: *Schematic diagram of the experimental setup.*

9.2 Achieving phase separation

The signals will be separated by a phase shift and the problem is then how to generate two current ramps and two sinusoidal signals, separated by a phase angle α. This could be done by the use of a function generator and a phase shifting device. Such a device is in its simplest form comprised of an all-pass filter. However, this can also be accomplished by using an Arbitrary waveform generator (AWG) card (see Sect. 9.2.2). With the AWG card two different signals are generated, though governed by a common clock.

9.2.1 Waveform generation

The ADW card is feed discrete values at a constant bit rate, from which a analog signal is created. The discrete values are in turn generated by LabVIEW program code. The code is written in such a way that all parameters which determine the appearance of the waveform can be adjusted by the user. There is a rather difficult relation between the card bit-rate and the frequency of the signal wanted. This will be discussed in App. A.

9.2.2 The arbitrary waveform generator card

The ADW card used is a DA4300 unit and has some specific characteristics. There is a relationship between the bit-rate, the clock-frequency and the

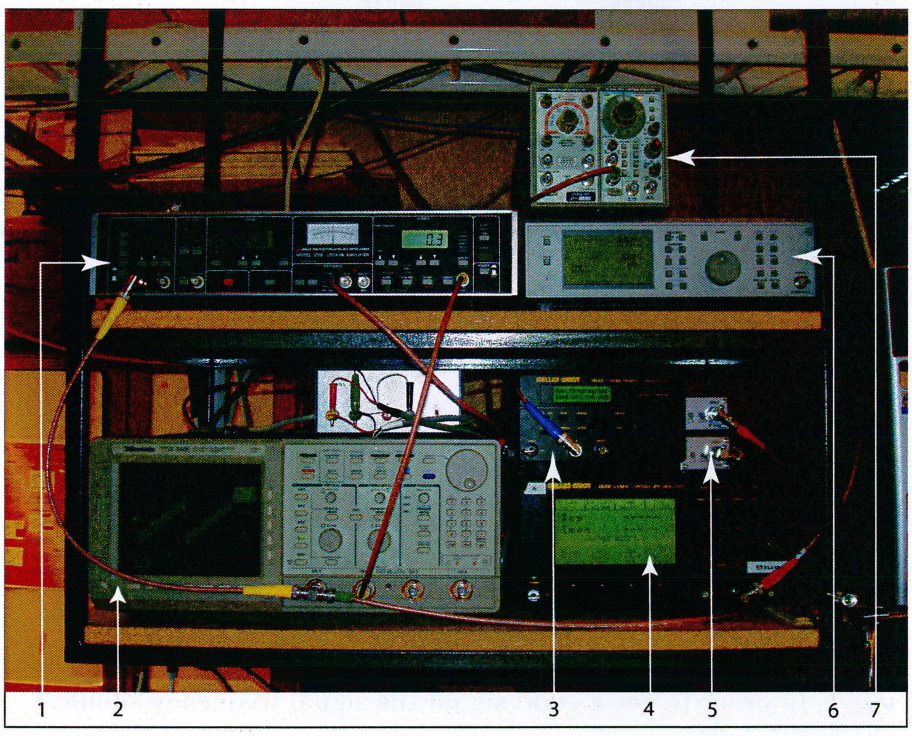

Figure 9.2: *The components used in the measurement setup. 1) Lock-in amplifier 2) Oscilloscope 3) Laser driver (Water vapor) 4) Laser driver (Oxygen) 5) Signal amplifiers 6) Ramp generator (later replaced by the AWG card) 7) Signal generator (later replaced by the AWG card).*

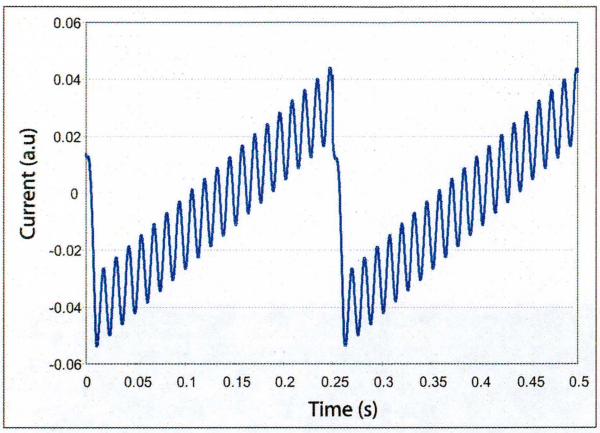

Figure 9.3: *The modulation signal sent to the laser (The frequency of the modulation is not to scale. (In reality the frequency of the modulation is about a factor 2500 higher than the frequency of the sawtooth.))*

resolution of the waveform. This can be stated as:

$$\frac{\nu_{clock}}{numpoints} = \nu_{limit} \tag{9.1}$$

where

- ν_{clock}: The clock frequency (should be set to $\frac{300\ MHz}{2^n}$, $n = 1, 2, 3...$)
- $numpoints$: Number of datapoints used
- ν_{limit}: In order to get a good signal the signal frequency should be kept above this value.

The slowest frequency component used is 4 Hz (the sawtooth) and hence the constraining factor in Eq. 9.1. A distortion-free signal is achieved with $\nu_{clock} = 300/2^8$ MHz, $numpoints = 292976$, resulting in $\nu_{limit} = 4$. There are several signal ports on the card and they have been utilized as follows:

- CLK IN: External clock frequency, this is used when clock-frequencies below 75 MHz are needed.
- TRIG IN: (Not used)
- AOUT 1: Modulation signal (nr 1) + Ramp (nr 1)
- AOUT 2: Modulation signal (nr 2) + Ramp (nr 2)
- AOUT 3: Modulation signal (nr 1)

- AOUT 4: Modulation signal (nr 2)
- MKR OUT: A signal used to trigger the oscilloscope.

9.2.3 User interface

When using the program to control the waveform, there is a balance between too much and to little parameter control. The waveform created and sent to the AWG card will for instance not be displayed, since screen space is limited. The measured absorption curves will neither be displayed in real-time, for the same reasons. The following will be available to determine for the user:

- The modulation frequencies
- The modulation amplitudes
- The ramp frequencies
- The ramp amplitudes
- The phase difference between the two modulation signals

Items one through four will be set for both signals (one for each laser being used). This is, however, somewhat unnecessary for the ramp frequencies, which will always be the same in our case.

9.3 Lock-in amplifier

As stated in Sect. 5.3 the lock-in amplifier is filtering out the information which is modulated with a specific frequency and phase. This leads to the fact that a separation between two signals can be made if they differ in either phase or frequency. The downside to changing the frequency is that the lock-in reference signal has to be changed simultaneously. By instead changing the phase of the modulation signals in comparison to the lock-in reference, a distinction can be made. This is done by setting the reference signal phase to zero and then generating two signals phase shifted 0 degrees and 90 degrees. Since the signals are perpendicular they will not interfere with each other (linearly independent).
The two different Fourier components (1f and 2f) both contain vital information. The 1f component has an offset which corresponds to the amount of light reaching the detector. The 2f component gives information on how much of the light was absorbed by the specific absorption line. This means that the lock-in amplifier output has to be read at the different settings depicted

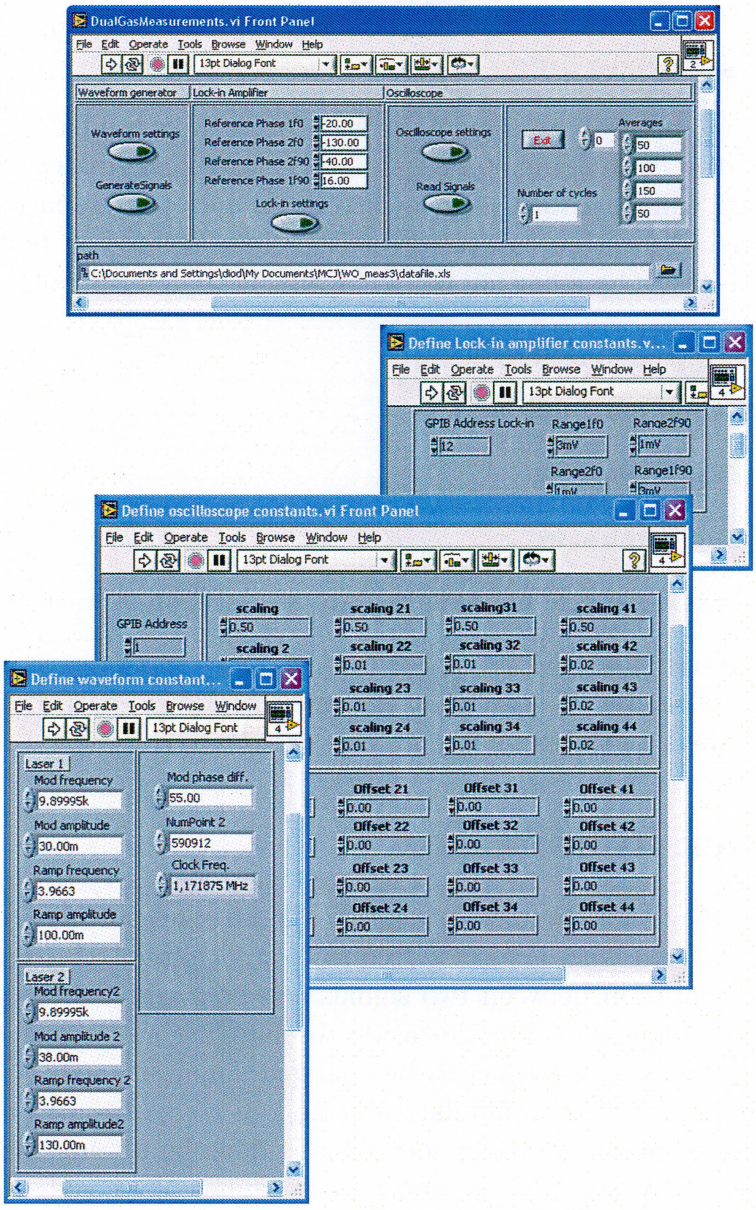

Figure 9.4: *User interface panels. The dialogue panel at the top is visible at all times and the three others can be called if modifications need to be made (using the "settings" buttons).*

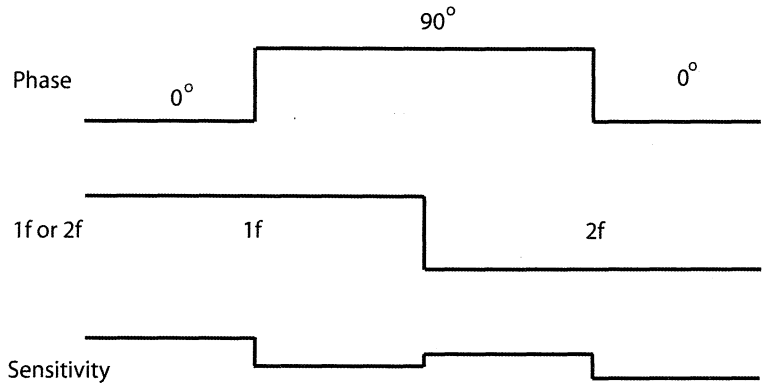

Figure 9.5: *Commands of change made to the lock-in amplifier*

in Fig. 9.5. The four different signal measurements have to be interlaced and therefore the phase of the lock-in, as well as the choice of Fourier component has to be made subsequently during a session. This is done through computer control of the lock-in amplifier. In addition to this the sensitivity (pre-amplification) has to be adjusted in-between every measurement in order to maximize the signal resolution.

9.4 Oscilloscope

An oscilloscope is used to register periodic signals. The signals displayed on the oscilloscope are feed from the lock-in amplifier output. In order to get a good signal, free of noise, many signals are registered and averaged. The information is then exported to the computer. The oscilloscope sensitivity and the number of averages made have to be controlled by the computer, so that full automation is achieved. It is important to be able to set the sensitivity and number of averages for every measurement, since the strength and quality of the signals vary widely. The option to set an offset for the oscilloscope has been included into the computer program as well. The AWG card mentioned above generates the trigger signal to the oscilloscope.
The number of averages that has to be made is dependent on the signal quality, the nature of the signal and if the setup is being subjected to vibrations. A common way of determining how many averages is suitable is to make an Allan plot. When making the plot, one starts by taking several single measurements. The first point in the plot is just one measurement, the second is the average of two measurements, and so on. A closely related method has been used for the measurements in this paper. In the Allan plot one single

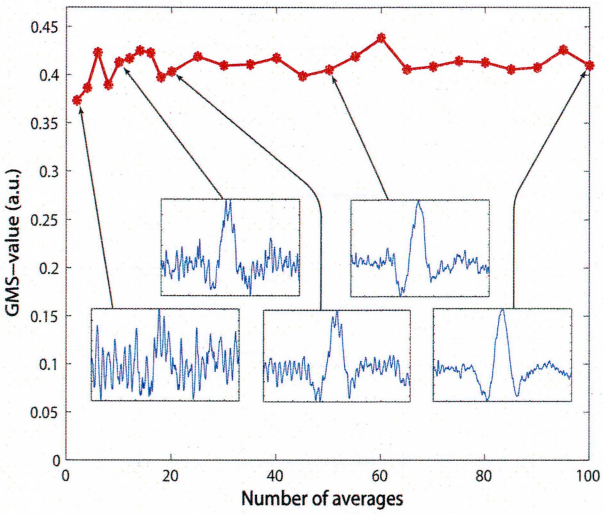

Figure 9.6: *Data received using the alternative Allan plot method. The number of averages is changed for the second harmonic and the appearance of the 2f signal can be seen at some strategic numbers.*

measurement is added to form the average for the next point. In the alternative method, entirely new sets of data is used for every new point. This means that the averaging can be made in the oscilloscope, as it will be in the subsequent measurements on samples. The Allan plot will relay a clearer picture of the influence of increasing averaging, the alternative method will, however, be closer to the actual measuring situation. Theoretically, both plots will display less variation between points for every increase in number of averages made. A set of measurements made are displayed in Fig. 9.7. The number of averages needed has to be established for the 2f signal as well as the 1f signal. When the number of averages for the 2f signal is plotted, the maximum number of averages is used for the 1f signal and vise versa.

9.5 Computer control

The entire experimental setup is controlled by a computer, as mentioned above. This is important for a couple of reasons:

- The ability to control all aspects of the experiment with one interface
- To make mid-experiment adjustments
- To gain the possibility of making unattended measurements

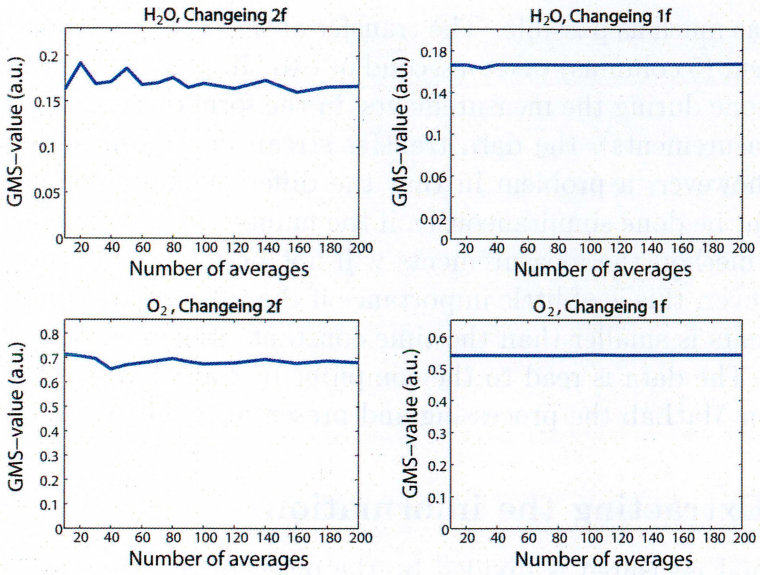

Figure 9.7: *The data received using the alternative Allan plot method. Common for the four plots is that they all show the 2f signal normalized by the 1f signal. The influence of the two is tested separately. Variations in the constituents will effect the normalized signal in the same way.*

- Being able to set parameters, which are not of concern to the user, in advance.
- Reading the result to the computer and thereby making data processing easy.

All transfers, apart from the communication with the AWG card, are done using GPIB. The prime advantage of GPIB is the fact that it was developed for laboratory instrument control and is therefore always compatible. It is, however, not the fastest communication, making it advantageous to use a large amount of oscilloscope averages before down-loading the data. The computer control is achieved using LabVIEW. The programs created are described in App. A.

9.6 Data analysis

There are two different measurement taken for each of the two lasers. For each of these, data are read from the four oscilloscope channels (plus time data). This yields 20 columns of data, each with thousands of value. The data transferred has to be kept to a minimum, in order to shorten measurement

downtime as much as possible. The transfer of the oscilloscope data after one measurement (5 columns) takes a second or two. If some of the processing had not been done during the measurements, in the form of averaging (averaging over n measurements), the data transfer stream had been n times greater. There is, however, a problem in that the different measurements will to a lesser extent be done simultaneously, if the number of averages increases. By using this method the measurements will not be made at exactly the same time. However, this is of little importance if the time separation between the measurements is smaller than the time constant associated with the process measured. The data is read to the computer by LabVIEW and exported to MatLab. In MatLab the processing and presentation of the data is made.

9.6.1 Extracting the information

The 2f signal measured is divided by the offset of the 1f signal in order to get a dimensionless entity (independent of the amount of light sent into the sample). This normalization is in most cases done using the direct signal [3, 4, 16]. Since two lasers are run simultaneously the direct signal cannot be utilized, due to that it cannot be separated into to two signals. However, the 1f signal can be separated into two independent signals, one for each laser. The non-DC characteristics of the 1f signal is ignored.
The 2f signal has two minima and one maximum. The value of interest is the difference between the signals maximum and minimum of the signal (see Fig. 9.8). We define the GMS signal as

$$GMS = max - \frac{min_1 + min_2}{2}, \tag{9.2}$$

The GMS signal will be proportional to the strength of the normalized 2f signal and therefore to the gas concentration. However, the GMS signal will be dependent on the scattering coefficient as well. This dependence can be removed by measuring the time-constant [21] or by, as is done here, calibration using dual gas measurements.

9.7 Settings

The modulation amplitude, lock-in phase and ramp amplitude have to be optimized. The optimization of the modulation amplitude is done according to Reid *et al* [9]. These steps are all performed in order to arrive at a maximal signal with the proper characteristics. The 2f signal is recorded

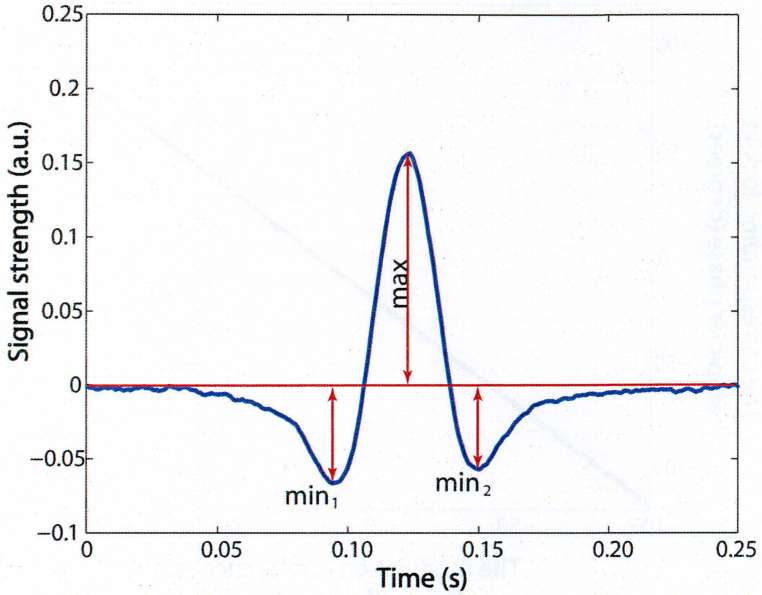

Figure 9.8: *Displaying the extraction of the GMS value.*

for a number of different modulation amplitudes. The modulation amplitude which generates a signal, where the maximum value is 2.2 times the minimum value (both to be taken as the absolute distance to zero) can be shown to be the optimal one (see Fig. 9.8). The lock-in phase which gives the maximum signal is found by tuning until the signal disappears, and then add 90 degrees. When setting the lock-in phase for the 1f signal, certain aspects must be considered. Maximizing the 1f signal (AC part) does not mean that the offset is maximized. This comes naturally since the two arise for different reasons (see Sect. 5.4). Finally, the ramp amplitude is chosen so that the absorption peak takes up a reasonable part of the frequency sweep. Too large a sweep reduces the resolution and too narrow a sweep increases the error in the position of the baseline (the horizontal part outside the signal).

The lock-in phase of the two laser 2f signals should be perpendicular. This is not simply achieved by setting the modulation signal of the two lasers perpendicular. This was found by plotting the lock-in phase that gave the maximal signal as a function of the phase shift between the reference and the laser modulation (see Fig. 9.10). The graphs make it apparent that the 2f signals being perpendicular does not result in the 1f signals being perpendicular to each other (see Fig. 9.9). In the 2f case, the phase dependence has a bistable quality; however, it does not seem to have an influence on the

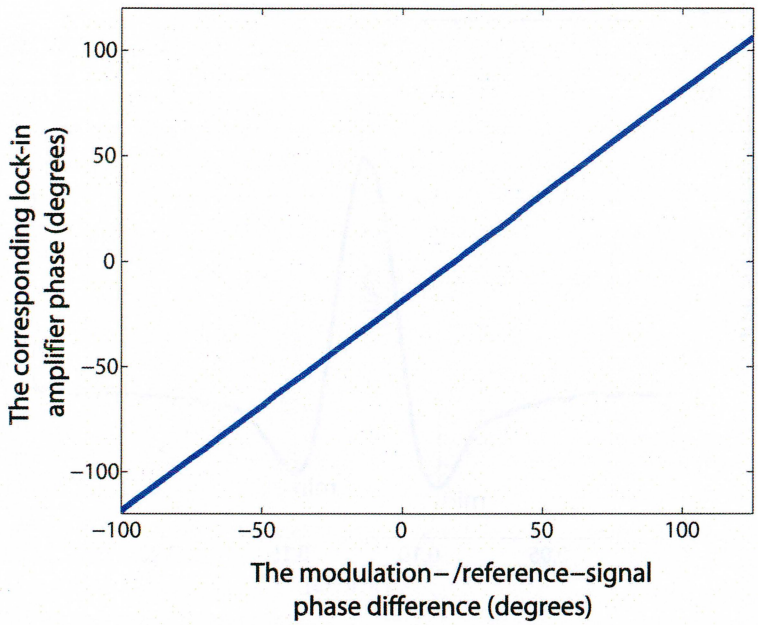

Figure 9.9: *The lock-in phase that gave the maximal signal, as a function of the phase shift for the 1f-signal. Zero on the horizontal axis will not correspond to zero on the vertical axis, since the phase is not preserved when the reference is mimicked internally by the lock-in amplifier (see Sect. 5.3).*

measurements. This is because of the fact that the phase does not seem to change when it once has been set. The bistability is seen if the wanted phase angle is approached from higher to lower contra lower to higher phase. The difference between the two branches is 45 degrees. This leads in turn to the assumption that the "problem" arises when the internal reference wave locks on to the reference signal feed to the lock-in amplifier.

If the 1f signal (or actually the offset of the 1f signal) is linearly dependent on sin(Lock-in phase) (see Fig. 9.12), they could still be separated by the use of

$$\begin{aligned} X \cdot \cos(0) + Y \cdot \cos(\alpha) &= A \\ Y \cdot \cos(0) + X \cdot \cos(\alpha) &= B \end{aligned} \tag{9.3}$$

Here A is the 1f signal offset when the X-signal is optimized and the Y-signal is off by α degrees and correspondingly for B (see Sect. 5.4.3). Since the requested unknowns are X and Y, the formulae are rearranged to form

$$X = \frac{A - B \cdot \cos(\alpha)}{\sin(\alpha)^2}$$

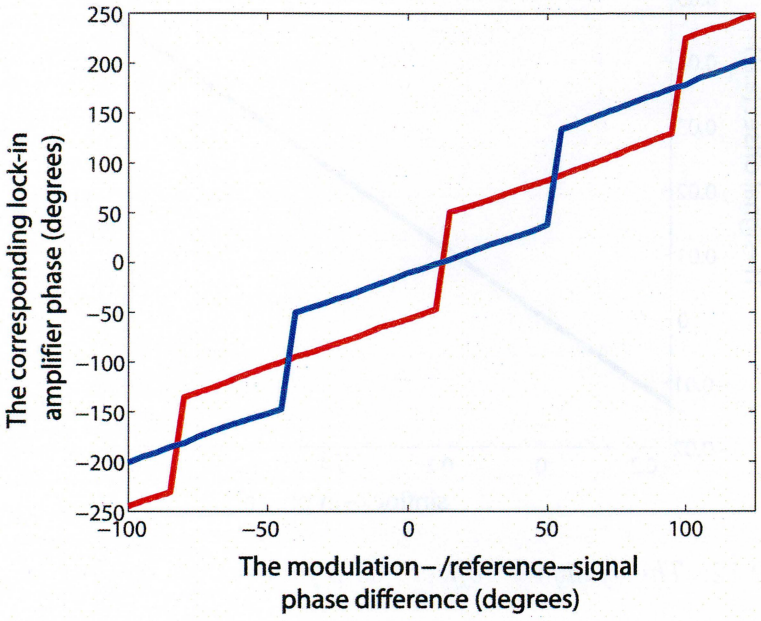

Figure 9.10: *The lock-in phase that gave the maximal signal, as a function of the phase shift for the 2f signal.*

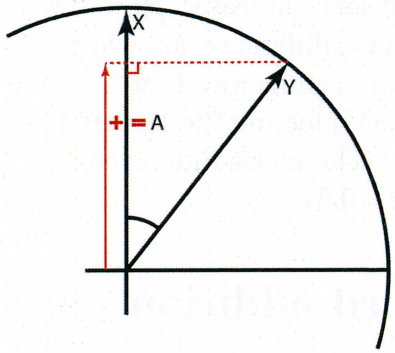

Figure 9.11: *The 1f signal "offset" measuring situation. The circle represents the phase plane (with zero defined as "up") and the two arrows represent the separated 1f signals.*

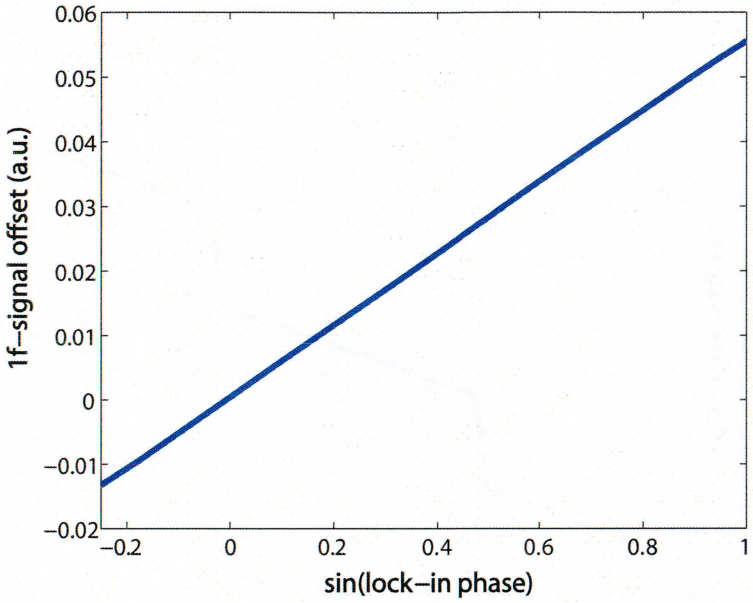

Figure 9.12: *The 1f signal "offset" as a function of sin(lock-in phase)*

$$Y = \frac{B - A \cdot \cos(\alpha)}{\sin(\alpha)^2} \tag{9.4}$$

We end up with two equations and two unknowns. In this manner the crosstalk due to non-perpendicular phase can be circumvented.
A demand set on the system by the method above is that the 1f signal is linearly dependent on sin(Lock-in phase). Measurements illustrated in Fig. 9.12 show that this is in fact the case. However, there is a problem associated with finding the phase difference for which the formula fully separates the 1f signals. The most certain way is to make a measurement with only one laser running and then plotting the 1f signal as a function of phase difference used in the formula. The phase difference for which one signal is zero, is the proper one (see Fig. 10.3).

9.8 Standard addition measurements

To evaluate the performance of the setup, the standard-addition method mentioned above (see Sect 6.1) was used for the two lasers. The measurements for the two different wavelengths were done separately, that is only one laser was running at a time. The laser beam first passed through a lens, making it collimated. The intensity is reduced with a neutral density filter in order to get a proper intensity (as large as possible, without saturating the

detectors and amplifiers). Before the beam reaches the detector it travels a variable distance. The distance is changed 60 mm in increments of 5 mm, added to an unavoidable offset. The offset (non-variable distance) should be kept to a minimum, making the impact of the distance variation maximally significant.
There are several reasons for performing a standard addition measurement.

- Estimating the system offset. Not including the STD offset, there is an inherent offset associated with every setup. This arises from couplings between fibres, lenses, filters and so on.

- Establishing the sensitivity, that is how much will the signal increase, for a known increase in the distance traveled by the beam (and the other way around).

- Getting a clear picture of the nature of using the 1f signal for normalizing. The STD measurements having the expected appearance is an indicator to whether it is an acceptable way of making the normalization.

- The measurement dependence on lock-in sensitivity. As discussed in Sect. 5.3, the lock-in sensitivity value corresponds to the pre-amplification of the signal in the lock-in amplifier. Therefore the signal should scale proportionally to the sensitivity setting for the 2f signal and inversely proportional to the sensitivity setting for the 1f signal. This is because the 1f signal is used to normalize the 2f signal. That is, the GMS value will be proportional to $\frac{2f}{1f}$ (see Figs 9.14 and 9.16).

The interesting information in the STD curves, are the gradients of the lines. It is the gradients that are used to calibrate the subsequent measurements (see Figs 9.13 and 9.15). The measurements later made on samples result in GMS values. In order to get the more interesting "Equivalent mean path" value, the GMS value has to be divided by the corresponding gradient. When calibrating the subsequent measurements, the gradient has to be recorded for the same sensitivities as the measurement to be calibrated. Since this will require about 100 STD measurements, some (five) have been done and the others have been calculated by the use of the fact that the GMS value (and therefore the gradient) is proportional to the lock-in sensitivities. By extrapolating the information received in the two groups of five, two entire matrixes can be derived. These are graphical displayed in Fig. 9.17.

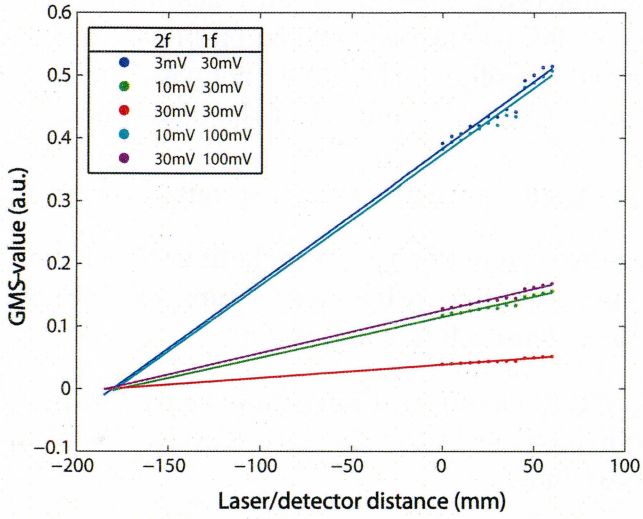

Figure 9.13: *The STD measurements for oxygen. Three different 2f sensitivities and two different 1f sensitivities have been used. It can be seen that the GMS values are in fact proportional to the sensitivities as stated above.*

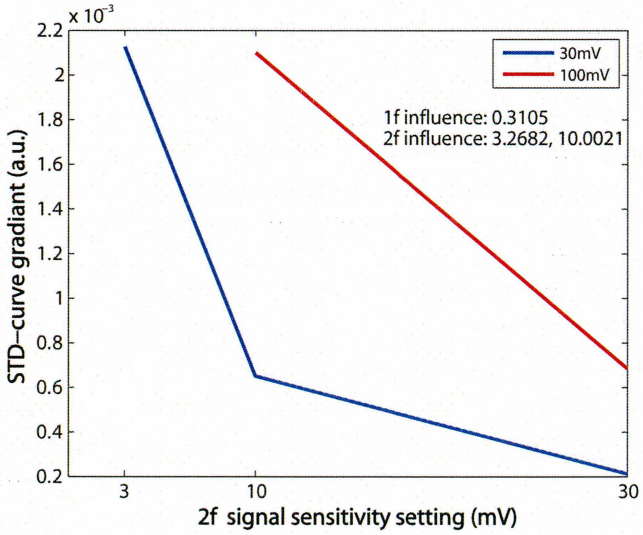

Figure 9.14: *The two lines represent different 1f sensitivity setting. It can be seen that the two lines differ in both points by a factor 0.3105. This should be* $\frac{30mV}{100mV} = 0.3000$. *The difference between the first and second point (from the left) on the horizontal axis, for the blue line is 3.2682. The expected value is* $\frac{10mV}{3mV} = 3.3333$. *Finally, the difference between the first and the third point along the horizontal axis is 10.0021. The expected value is* $\frac{30mV}{3mV} = 10.000$.

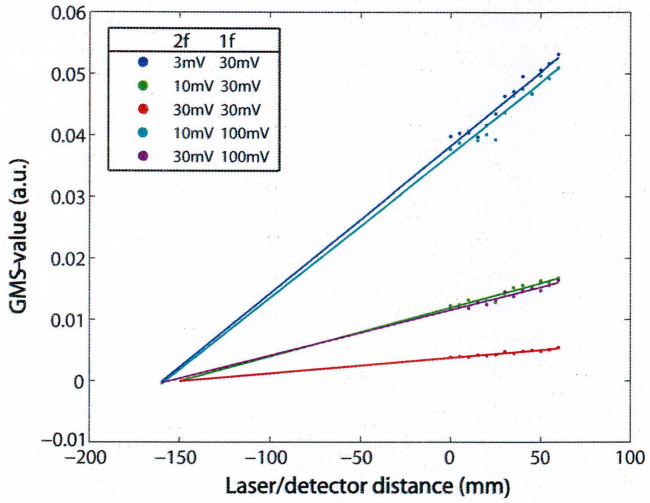

Figure 9.15: *The STD measurements for water vapor. Three different 2f sensitivities and two different 1f sensitivities have been used.*

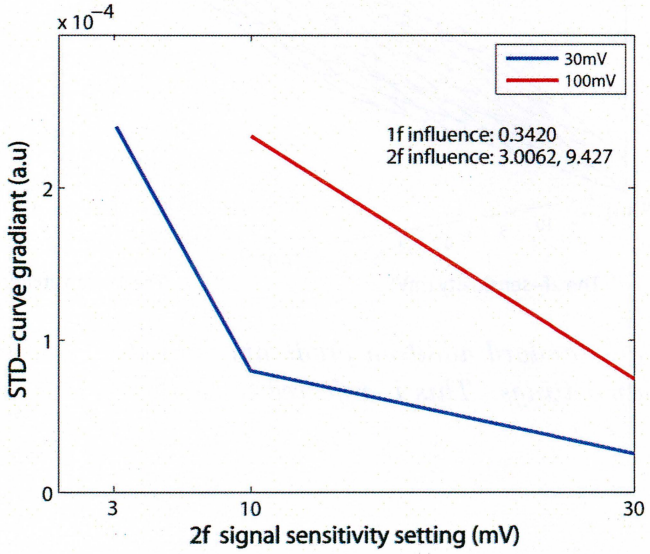

Figure 9.16: *Data as in Fig. 9.14, but for water vapor. The two lines differ in both points by a factor 0.3421 (should be* 0.3000*). The difference between the first and second point on the horizontal axis, for the blue line is 3.0062 (should be* 3.3333*) and 9.427 between the first and the third point (should be* 10.000*). The values arrived at here are not as good as those in Fig. 9.14; this is due to a much lower signal-to-noise ratio for the water vapor measurements in air. The difference in signal strength is dealt with in Sect. 10.2.4.*

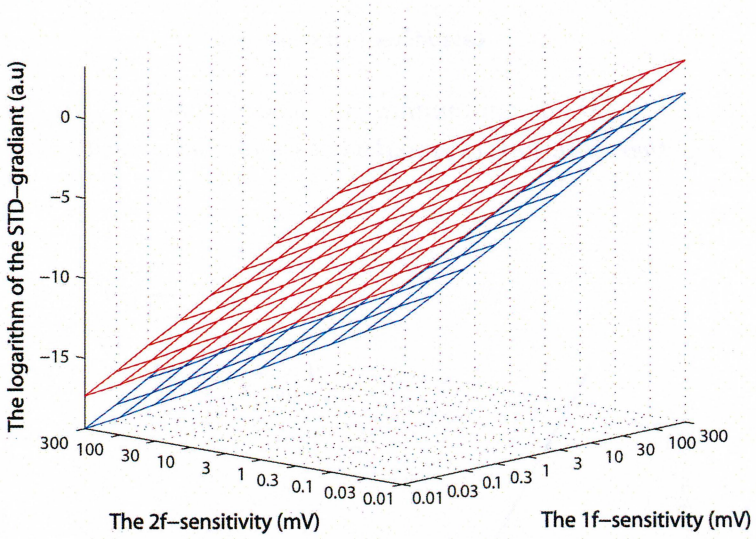

Figure 9.17: *The standard addition gradients plotted as a function of the 1f and 2f sensitivity settings. This is referred to as the STD matrix.*

Chapter 10

Measurements

The measurements made can be divided up into two main groups. The first group contains measurements which are designed to generate data that will serve the calculations performed on later measurements. This is for instance the case for the STD measurements, discussed in Sect. 6.1. The latter group is comprised of the measurements for which the system is designed, that is, two gases are measured at the same time. The purpose of the measurements in the latter group is to verify the accuracy and over-all performance of the method. The first group will be referred to as "Acquisition measurements" while the later will be "System measurements".

10.1 Acquisition measurements

The acquisition measurements constitute measurements, that are made only once every time the setup significantly changes. As stated above, the STD measurements fall into this category, as well as the recording of base functions. The method of measuring the STD curve has been thoroughly dealt with above. The base functions are measured in a situation where the signal is so strong that the noise is negligible in comparison. This has been accomplished by making a measurement over a considerable distance.

10.2 System measurements

In the system measurements category, all the method-verifying measurements have been included.

10.2.1 Measurements made over time

The data received in doing the GASMAS measurements are not absolute values of the concentration of the gas in question. The measurements are therefore used to make comparisons. If a comparison between different measurement situations is to be made, a comparison over time is close at hand. In the measurements testing the system, continuous registering of the signal is performed while the object is subjected to some kind of change. This change might be in the form of oxygen refilling a nitrogen flushed sinus cavity [21] or some kind of drying process [3].
In this specific setup, oxygen and water vapor are measured. For example, in the case of air refilling a nitrogen flushed cavity, both signals will increase as a function of time. When measuring a drying process, the situation is a bit more complex and will be dealt with in detail in Sect. 10.2.3.

10.2.2 Measurement object

The systems measurement have here been performed on drying balsa wood. The reason for using balsa is two-fold. First, balsa is a hard wood, see Fig. 10.1 and secondly, it is rather translucent in the wavelength region of interest. In comparison with soft wood, hard wood has a simpler construction. Softwood is comprised of four different kinds of cells, where as hard wood only incorporates two. Thus, we have a simpler measurement situation in the hardwood case. For a more in-depth discussion concerning wood drying processes, see Ref. [3]. When a piece of wood has been let to soak in water, its pores will ultimately be completely filled with liquid water. This means that theoretically no signal will come from either gas measurement. However, in practice there will always be a signal, due to gaps in the beam path. The drying of the wood is schematically described in Fig. 10.2. First, the liquid water will start to vaporize, followed by the cell walls retracting. The different effects this will have on the signal is treated in the subsequent chapter.

10.2.3 Interpretation of the signal in the case of drying wood

The value describing the gas concentration is, as stated above, dependent on the 2f and 1f signal. The 2f signal will be dependent on the gas concentration and on the amount of light arriving at the detector. The 1f signal will, on the other hand, only be dependent on the amount of light received by the detector. The first harmonic in the case of measuring gases for a drying piece

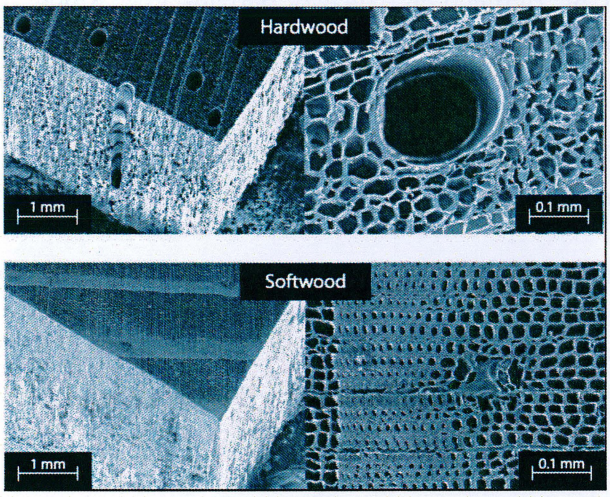

Figure 10.1: *Different types of wood. Top picture shows hard wood, for instance balsa. The lower picture shows soft wood, for instance pine (From [3]).*

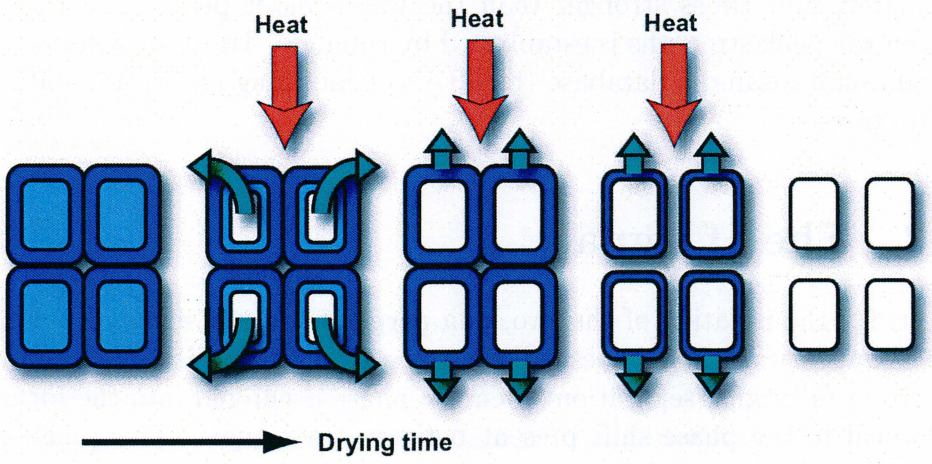

Figure 10.2: *The process of drying wood, displayed schematically (From [3]).*

of balsa wood, will thus have a simpler behavior than the 2f signal.
The 1f signal offset at the oxygen frequency will be high for water soaked balsa and it will be reduced as the wood dries [3]. Since the offset of the 1f signal is proportional to the amount of light at the detector, wet wood must be more translucent than dry. This comes as a consequence of the cell structure of the wood. When balsa is wet the cells will be filled with water and the light will be less scattered. When the wood dries, gas pockets will be introduced in the cells which, due to different refraction index will serve as scatterers.
In the case of measuring the 1f signal at the water vapor absorption wavelength, the situation is different. For a soaked piece of balsa, the signal will be small, increasing as the drying process goes along. This is not mainly a consequence of scattering, but rather due to the significant absorption of liquid water in the wavelength region used.

10.2.4 Relative intensity of the absorption peaks

The exact wavelengths for the absorption tops used for the measurements in this paper is:

Oxygen	761.003 nm
Water vapor	978.509 nm

These absorption peaks are related so that the oxygen peak absorbs approximately four times stronger than the water vapor peak. The relation between the peak strengths is established by running "Trans for windows", a program built around a database (HITRAN) containing absorption data for various gases.

10.3 The 1f signal

In Sect. 9.7 the isolation of the two, non-perpendicular 1f signals from each other is put forward. This does call for a practical affirmation. Eq. 9.3 will only give a full signal separation when the phase α entered into the formula is identical to the phase shift present between the signals. This phase can be found by measuring the DC part of the two 1f signals in the usual way, with the exception of only having one laser running. If a full separation is reached, one of the 1f signals will change over time, while the other will be be constantly zero (see Fig. 10.3).

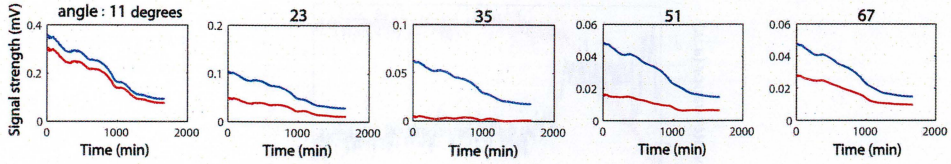

Figure 10.3: *The 1f signal offset as a function of time (Blue-Oxygen, Red-Water vapor) measured on drying balsa wood. Only one of the lasers are on (oxygen). The two curve are measurements of the parameters X and Y in Eq. 9.4. The different sub-plots indicate different phase entered into Eq. 9.4. When the phase in the equation is the same as the phase separation between the two lock-in readings, one signal should be constantly zero.*

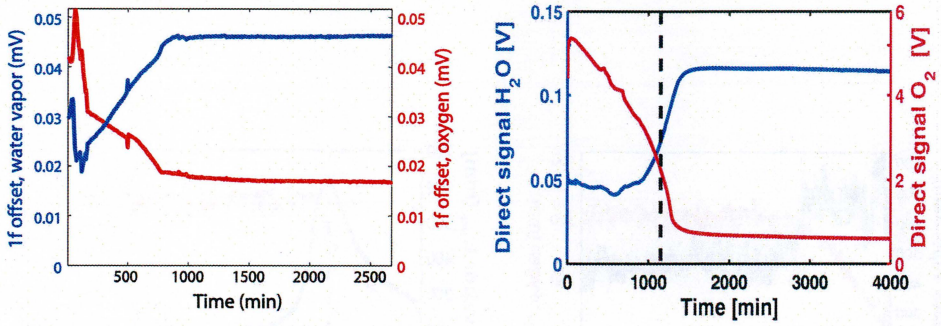

Figure 10.4: *Wood. The left picture is presenting the 1f signal as it is measured by the method described in this paper. The right picture is taken as a comparison [3]. The later shows the direct signals, registered separately for oxygen and water vapor.*

10.4 Measurements made on drying wood

The choice of measuring on wood came as a consequence of the fact that wood drying measurements has been done in the group. Previous to this, the oxygen and water vapor has not been measured in the same point. The preexisting data which are known to be correct, can be used for reference. These data have been published in [3]. The first aspect of the measurements which has to be investigated is, as stated above, the behavior of the first harmonic. It should depend on time in the same manner as the direct signal (see Fig. 10.4). The second parameters being measured is the GMS value. For one measurement session, the GMS value was recorded, normalized with the 1f offset and plotted in Fig. 10.7 as a function of time. This curve has also been compared to the corresponding one published in [3]. In order to get good, comparable results, the normalized GMS value has to be divided

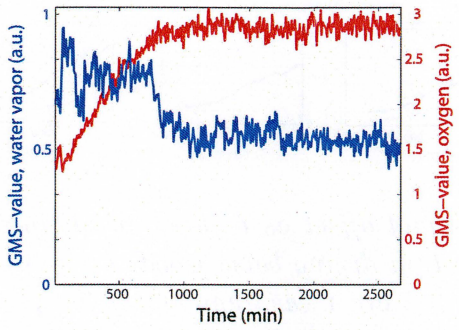

Figure 10.5: *Wood. The GMS values simultaneously registered by the two-wavelength setup. (No GMS curve was available in [3])*

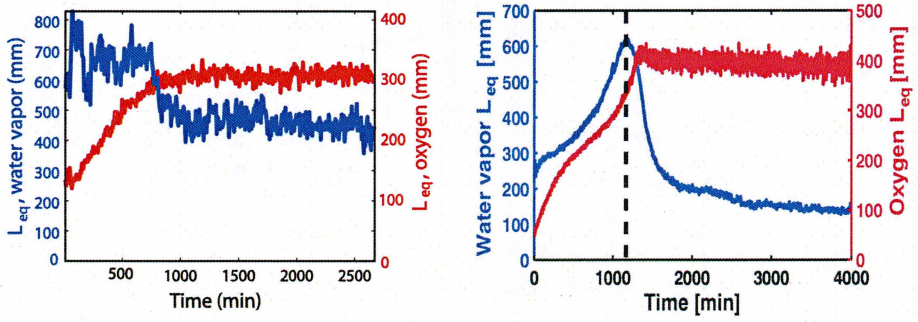

Figure 10.6: *Wood. The left picture is the EMP values simultaneously registered by the two-wavelength setup. The right picture is the EMP values from two separate measurements [3].*

by the corresponding STD value from 9.17. The value arrived at is called the equivalent mean path (L_{eq}) and is a quantity with can be compared between measurements and lasers. The equivalent mean path length has been plotted as a function of time for the measurement above in Fig. 10.6. In these examples, the comparisons are made between two measurements with are separated by different measuring situations, different samples and different environments, despite this clear similarities between the characteristics can be seen.

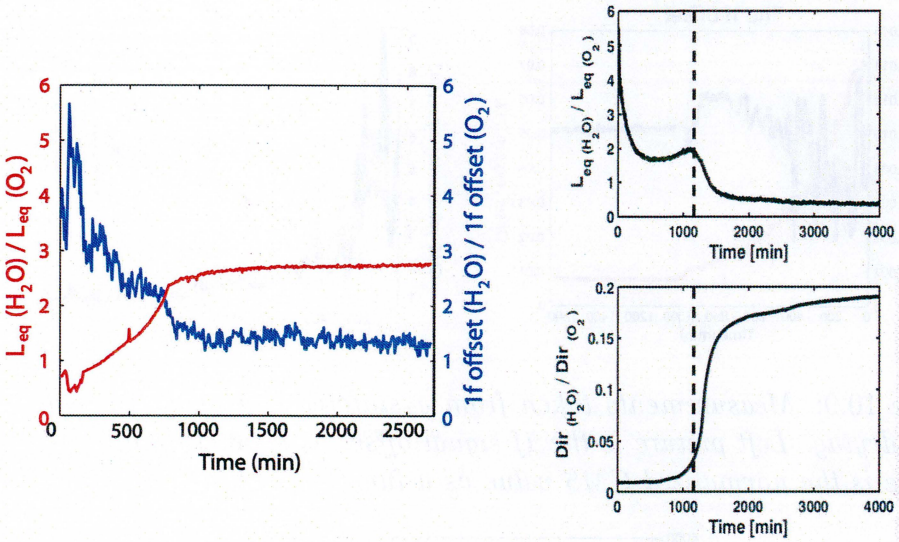

Figure 10.7: *Wood. The curves for water vapor divided by the curves for oxygen. Left figure: The red curve is the offset of the 1f signals divided by each other. The blue curve is the equivalent, taken for the equivalent mean path curves. Right figure: Corresponding curves, as they are found in [3].*

10.5 Measurements made on a drying slab of sintered glass

When testing and evaluating a new method it is desirable that the measurement object used has an as simple behavior as possible. This makes the wood measurements less than ideal. Therefore a sintered glass plate (see Fig. 10.8) has been measured in addition to the wood. The plate is 9 mm thick, 90 mm in diameter and designed to function as a filter in chemistry. The pores in the plate are, according to the factory specifications 16-40 μm. They are made

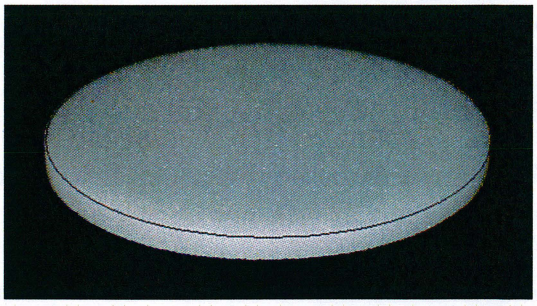

Figure 10.8: *The sintered glass plate used in the measurements.*

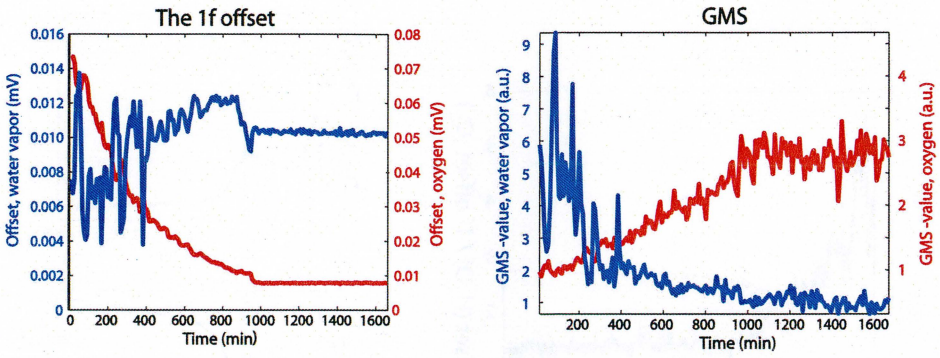

Figure 10.9: *Measurements taken from a sintered glass plate (9 mm thick) while drying. Left picture is the 1f signal offset as a function of time. Right picture is the normalized GMS value as a function of time.*

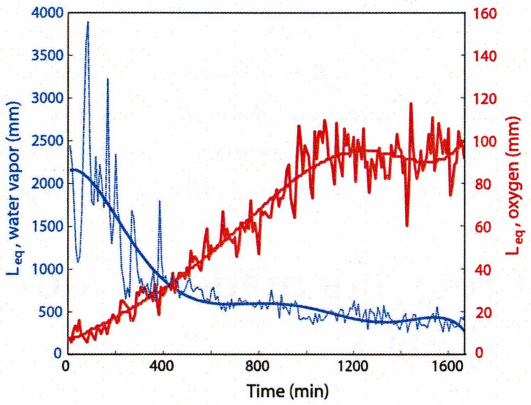

Figure 10.10: *Sintered glass. The equivalent mean path as a function of time. Calculated from the GMS curve in 10.9. A polynomial fit of the sixth order is included in the picture to indicate the overall tendencies.*

up of finely crushed glass, which is then heated and subjected to pressure, this makes the glass fuse together and form plates. The plates will, much like the wood, have pores which fills up with water when submerged in water. The structure of the plates will, however, not be affected in any other way by the water (the size of the pores is constant). It can be expected that the wood and sintered glass measurements will resemble each other, with some minor exceptions. The situation is, however, not clear, since the way oxygen diffuses into the sample when it dries is not known in either the wood nor the sintered glass case. The 1f signal offset and the normalized GMS value has been plotted in Fig. 10.9 as a function of time. The equivalent mean path is calculated and plotted in Fig. 10.10.

Chapter 11

Summary and conclusion

The purpose of this investigation has been to develop a method for making dual gas measurements for a common point in scattering materia. There are several aspects of the original GASMAS setup [4] which has had to be reworked. The most conspicuous one is the signal separation through phase shift (IQ modulation). The approach of changing the phase between lasers was implemented since it means that the reference signal can be kept unchanged. Phase separation is, however, as was seen as the work progressed, suffering from some difficulties. The overshadowing problem is the difficulty to achieve perpendicular signals. The first and second harmonic cannot simultaneously be made perpendicular. This problem can, however, be handled by the use of signal processing.
The second major problem is how to normalize the signal. Normally in GASMAS, this is done by dividing with the direct signal. The direct signal is, however, not attainable in this specific setup, since the lasers will be indistinguishable in the direct signal. The first harmonic will be proportional to the transmitted light as well, and since it is modulated, the signals from the two lasers can be separated.
It has been shown throughout this paper that this is not merely theoretical observations, but something which can be implemented in a functioning setup. Computer control programs have been made, with the object of performing the continuous changes necessary to the different devices during a measurement. The controls have been run from LabVIEW while the data evaluation has been performed in MatLab. In this manner the oxygen and water vapor content of different materials (balsa wood and sintered glass) have been measured while drying. In order to evaluate the results, a comparison has been made between the data received and the data presented in [3]. The first harmonic resemblance to the direct signal has been further treated in the comparisons mentioned. The measurements made are less than

ideal, due to time constraints, but never the less show the soundness of the method.

Chapter 12

Future improvements

This method, as well as most other methods can be improved in two ways, partly by increasing the performance of the system and in part by decreasing the complexity of it. The two are in some aspects linked, the former is, however, often significantly more difficult to accomplish. This is due to the fact that the performance restrictions are intrinsic to the setup. As an example it can be noted that the restrictions set on sensitivity when the concentration is measured is governed by the amount of fringes present. Much of the increase in performance will thus come down to a consideration of to which extent the fringes (and other noise) can be suppressed (see Sect. 5.1).
The other performance property which may be improved is the time a measurement cycle takes. This has not posed a problem yet, but it should not be excluded that measurements on more rapidly changing objects are desirable. In the setup as it is, the measuring time is rather rigid. This is because the restricting factor is due to the data transfer time (in the vicinity of 1 second) in the GPIB (oscilloscope to computer). This time cannot be controlled, and hence if shorter times are desirable an oscilloscope computer card would probably be the way to go.
The other important aspect of improving the system is to make it as simple as possible. This is desirable since it means easier handling, as well as cost reduction. The way of accomplishing this it to incorporate as much of the hardware (oscilloscope, lock-in amplifier, and so on) as possible as computer cards. This has already been done for the signal generators (replaced by the AWG card). When making this shift from separate boxes to computer cards, the system first has to be working properly, that is the setup must be significantly understood. The expected gains include lower cost, less complexity (the number of connected boxes decreases) and the communication between the different units become faster. However, the change is accompanied by some drawbacks. When a separate oscilloscope is being used, signals

are averaged in the oscilloscope, before sent to the computer. When using a oscilloscope computer card, this is not the case. The computer card lacks the ability to average and this leads to massive amounts of data that have to be processed, setting high demands on the computer being used.

Chapter 13

Acknowledgements

I would like to thank Linda Persson for support throughout this project. For helping me getting started with the project, as well as with the layout of my report. I thank Mats Andersson for technical and LabVIEW support, as well as incessantly introducing me to new aspects of the GASMAS method. I would like to thank Sune Svanberg for continues assistance and input. Finally I thank my parents for moral as well as financial support.

Bibliography

[1] S. Svanberg: *Atomic and Molecular Spectroscopy, fourth edition* Springer, Germany (2004)

[2] M. Sjöholm: *Laser Spectroscopy analysis of Atmospheric Gases in Scattering Media* Doctoral Thesis, Lund University (2006)

[3] M. Andersson, L. Persson, M. Sjöholm, S. Svanberg: Spectroscopic Studies of Wood-drying Processes. *Opt. Exp.* 14, 3641 (17 april 2006)

[4] G. Somesfalean, M. Sjöholm, J. Alanis, C. Klinteberg, S. Andersson-Engels, S. Svanberg: Concentration measurements of gas embedded in scattering media by employing absorption and time-resolved laser spectroscopy. *Appl. Opt.* 41, 3538 (20 June 2002)

[5] O. Svelto: *Principles of Lasers, Fourth edition* Plenum Press, New York (1998)

[6] A. Eckbreth: *Laser Diagnostics for Combustion Temperature and Species, Second edition* Gordon and Breach Publishers, Netherlands (1996)

[7] F. Predrotti, S. Leno, S. Pedrotti: *Introduction to Optics, Second edition* Prentice-Hall, New Jersy (1993)

[8] G. Boisd'e, A. Harmer: *Chemical and biochemical sensing with optical fibres and waveguides* Artech house, London (1996)

[9] J. Reid, D. Labrie: Second-harmonic Detection with Dunable Diode Lasers -Comparison of Experiment and Theory. In: *Appl. Phys. B* 26, 203, (1981)

[10] J. Liu, J. Jeffries, R. Hanson: Wavelenght Modulation Absorption Spectroscopy with 2f Detection using Multiplexed Diode Lasers for Rapid Temperature Measurements in Gaseous Flows. *Appl. Phys. B* 78, 503 (2004)

[11] M. Gharavi, G. Lehnasch, S. Buckley: Quantification of Near-IR Tunable Diode Laser Measurements in Flames *2nd Joint Meeting of the U.S. Sections of the Combustion Institute* Paper 167 (25-28 March 2001)

[12] H. Li, G. Rieker, X. Liu, J. Jeffries, R. Hanson: Extension of Wavelength-modulation Spectroscopy to Large Modulation Depth for Diode Laser Absorption Measurements in High-pressure Gases. *Appl. opt.* 45, 1052 (10 Feb. 2006)

[13] T. Fernholz, H. Teichert, V. Ebert: Digital, Phase-sensitive Detection for in situ Diode-laser Spectroscopy under Rapidly Changeing Transmission Conditions. *Appl. Phys. B* 75, 229 (2002)

[14] F. Burck, O Lopez: Correction of the Distortion in Frequency Modulation Spectroscopy. *Meas. Sci. Technol.* 15, 1327 (2004)

[15] L. Persson, F. Andersson, M.Andersson, S.Svanberg: Approach to Optical Interference Fringes Reduction in Diode Laser Absorption Spectroscopy. *Appl. Phys. B*, in press (2007)

[16] M. sjöholm, G. Somesfalean, J. Alnis, S. Andersson-Engels, S. Svanberg: Analysis of Gas Dispersed in Scattering Media. *Opt. Lett.* 26, 16 (2001)

[17] J. Alnis, B. Anderson, M. Sjöholm, G. Somesfalean, S. Svanberg: Laser Spectroscopy of Free Molecular Oxygen Dispersed in Wood Materials. *Appl. Phys. B* 77 691 (2003)

[18] L. Persson, H. Gao, M. Sjöholm, S. Svanberg: Diode Lasers Absorption Spectroscopy for Stidues of Gas in Fruits *Optics and Lasers in engineering* 44, 687 (2006)

[19] L. Persson, B. Andersson, M. Andersson, M. Sjöholm, S. Svanberg: Studies of Gas Exchange in Fruits using Laser Spectroscopic Techniques. *Proceedings of FRUTIC 05, Information and technology for sustainable fruit and vegetable pruduction* 543 (Sept 2005)

[20] L. Persson, K. Svanberg, S. Svanberg: On the Potential of Human Sinus Cavity Diagnostics Using Diode Laser Gas Spectroscopy *Appl. Phys. B* 82, 313 (2006)

[21] L. Persson, M. Andersson, T. Svensson, K. Svanberg and S. Svanberg: Non-intrusive Optical Study of Gas and its Exchange in Human Maxillary Sinuses. Manuscript in preparation (2006)

[22] Cystic Fibrosis Center At Stanford *http://cfcenter.stanford.edu/*

Appendix A

The LabVIEW programs

All the programs discussed below are stored in a file called *Programs*. It is important, in order for the programs to work, that this file is installed under *C: \ Program Files*. This is due to references used in the LabVIEW programs.
This section will mostly deal with the main programs written in LabVIEW. However, these main programs are made up of a considerable number of subprograms, for which the description will be somewhat more sketchy. The purpose of the programs is to, by use of the computer, control and run the experiments. The lock-in amplifier, the oscilloscope and the laser drivers will be controlled by the programs. In addition to this, the programs will read the information wanted, from the oscilloscope to the computer.

A.0.1 Performing dual gas measurements *DualGasMeasurements.vi*

The program front as shown in Fig. A.1 is what at first will be seen by the user. There are a few features which are readily available. These are:

- *Reference Phase* This controls the phase value feed to the lock-in amplifier. There are four different such values, for the 1f and 2f signals for the two measurements (gases A and B).

- *Path* Directs where the received data should be saved.

- *Number of Cycles* Lets the user type in the number of cycles the program should go through. One cycle includes four measurements (1f and 2f signals for the two gases). If continuous measurements are desirable, *Number of Cycles* is to be set to a large value (>1000).

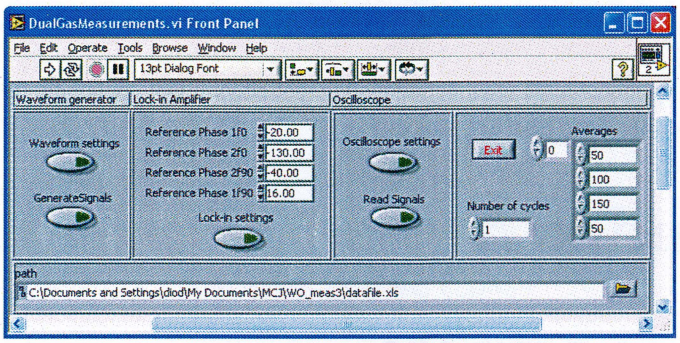

Figure A.1: *The user-interface for the Dual gas measurements program.*

- *Averages* Dictates the number of averages to be taken for the four measurements.

Besides these possible variable inputs, the program has several buttons. These are used as follows:

- *Waveform settings*
- *Lock-in settings*
- *Oscilloscope settings*

By pushing these buttons, additional windows (sub-windows) will open. They contain input boxes for parameters related to specific topic. For the *Waveform settings* button, the ramp amplitude, ramp frequency and so on can be set. The parameters set in this way are primarily data which will only be changed if the measurement situation has been altered significantly. When a change has been made in a sub-window, it has to be run in order for the changes to take affect.

There are three additional buttons:

- *Generate signals*
- *Read signals*
- *Exit*

While the exit button is rather self-explanatory, the other two demand a further description. When the *Waveform settings* sub-window has been run, the changes have been implemented. The changes have, however, not been transferred to the AWG card (Arbitrary Waveform Generator card). This is done by the *Generate signal* button. The parameters decided upon will be

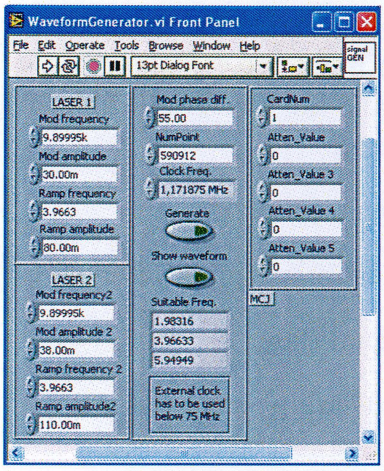

Figure A.2: *The user-interface for the Waveform generator program.*

uploaded to the card which in turn will generate the desired wave function. The *Read signals* button will start the measurements. It will feed the lock-in amplifier settings and the oscilloscope settings to the respective device. This will be done before every measurement. Four measurements will be made every cycle, the number of cycles can in turn be set arbitrarily. Every measurement will generate a data matrix which will be named in the following way:

datafile *"Measurement number (1-4)" "Cycle number (1-inf)"*.xls

For instance: datafile234.xls
(The name *datafile* can be chosen arbitrarily.)

A cycle takes about 4 minutes, this is of course governed by the amount of averaging used (set in the *Oscilloscope settings* sub-window).
Some of the sub-programs used above can be run separately from *Dual_Measurements.vi*; this is sometimes favorable (for instance during setup tests) and therefore these will be treated as well.

A.0.2 Generating the laser control *WaveformGenerator.vi*

When the *Generate signal* button in the main program is pushed the *WaveformGenerator.vi* program starts. This can be run separately, with some additional control over the parameters as a result. The program can be seen in Fig. A.2. Besides the functions in the main program, four different output

channel attenuations can be set. The attenuation sets the signal damping from 0 dB (no damping) to 63 dB. By pressing the *Show waveform* button before running the program, the waveform will be displayed on screen. There will be four graphs showing:

- The modulation signal (laser 1)
- The modulation plus ramp signal (laser 1)
- The modulation signal (laser 2)
- The modulation plus ramp signal (laser 2)

A.0.3 The waveform limitations

There are some limitations when it comes to generating functions with the AWG card. For instance:

> The values for: *Signal frequency, NumPoints* and *Clock frequency* cannot be set arbitrarily. This is a consequences of the digital nature of the signal.

The *Signal frequency* is the parameter you want to give a certain value, hence the *NumPoints* and the *Clock frequency* have to be fitted after wards. What do these parameters represent?

- *NumPoints* The card uses a set number of values which constitute the signal. This sequence of values is looped in infinity. The value of *NumPoints* is the number of values in a sequence. This value has to be $16 \cdot n$, where n is an integer.
- *Clock frequency* The values that constitutes the signal is feed from the card at a certain frequency. This "clock" can be generated by the card (300 MHz-75 MHz) or feed to the card from an external source. The demand on the clock is that it is equal to $\frac{300}{n}$ MHz, where n is an integer. (According to the manufacturer of the card, an external clock below 75 MHz should not be used, this does, however, not seem to cause any significant problems.)

The ramp-frequency used in this paper is 4 Hz. In order to lower the possible signal frequency either the *Numpoint* value has to be raised or the *Clock frequency* has to be lowered. Increasing the *Numpoint*-value above 1 M is not possible since the computer RAM memory runs out. This leaves the *Clock frequency* to be adjusted. The *Clock frequency* cannot, however, be

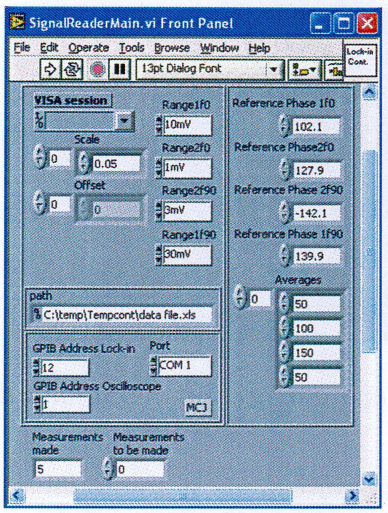

Figure A.3: *The user-interface for the Signal reader program.*

limitlessly lowered, since (in this paper) a 1 kHz modulation signal has to be represented as well. In the *WaveformGenerator* program suitable lower-limit frequencies are calculated and presented to the user.

A.0.4 Performing the measurements *SignalReaderMain.vi*

This is the program which deals with the steps performed when the *Read signals* button is pushed in the main program. The front shows parameters common to the ones set in the main program, though in a more compact way. The purpose of the *SignalReaderMain* as a separate program is to gain greater transparency in the code when running tests. The program can be seen in Fig. A.3

A.0.5 Performing standard addition measurements *StandardAddition_Measurements.vi*

This program is made to facilitate the standard addition measurements. As it is currently made it makes three 2f signal measurements and two 1f measurements at different settings of the lock-in amplifier sensitivity. The user interface can be seen in Fig. A.4. The lock-in amplifier phase of the five measurements can be set individually. The number of averages in a measurement is set globally and applies to all. The *Distance* is the position of the current measurement in millimeters (for instance 0 to 60 in increments

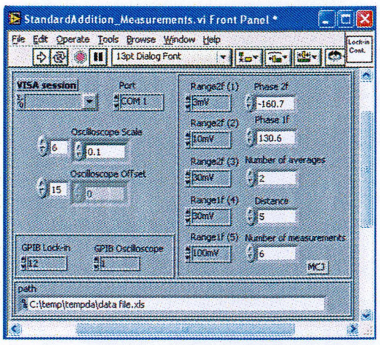

Figure A.4: *The user-interface for the Standard addition program.*

of 5). The *Number of measurements* is the number of measurements taken at every setting (to be averaged in MatLab). The program is semi-automatic in the sense that it stops every time the laser/detector distance has to be changed. Internally this program closely resembles the *SignalReaderMain* discussed above.

Appendix B

The MatLab programs

When the data have been acquired, they have to be processed in order to be appreciated. This is done using programs in MatLab. Different programs have been constructed for different applications and to be used in conjunction with the LabVIEW programs discussed above.

B.0.6 Dual gas measurements *DualMeasurement_evaluator.m*

This program is intimately connected to the LabVIEW program treated above, *Dual_Measurements.vi*. The MatLab program should be saved in the same folder as the measurement data. In addition to this the folder should also contain base functions for the four signals. Some parameters have to be set in the *DualMeasurement_evaluator.m* for it to work.

- *Angle* This denotes the angle between the lock-in phases of the two different lasers for the 1f signal.
- *Time* The total measurement time.
- *STD_O* This is the gradient of the standard addition curve for the lock-in sensitivity being used (Oxygen) (see Sect. B.0.10).
- *STD_W* This is the gradient of the standard addition curve for the lock-in sensitivity being used (Water vapor) (see Sect. B.0.10).
- *Ch* The channel being used on the oschilloscope
- *runs* The number of cycles performed by the LabVIEW program.

When the program has been run the following data is presented:

- During the running of the program the current absorption curve will be plotted. Since the calculations are rather fast, this information will appear as a video, displaying the changes over time.
- The GMS values will be plotted as a function of time.
- The Equivalent mean path will be plotted as a function of time.
- For single values, the respective matrixes is available.

B.0.7 Single measurement *SingleMeasurement_evaluator.m*

The single measurement program has a lot in common with the dual measurement one. The difference is that one gas is measured one time (that is a 1f and a 2f measurement). This means that the same parameters as above has to be set, with the exception of the parameter *runs*. The program returns the GMS value as well as a figure, displaying the 2f signal and the corresponding basefunction. This program is to be used in combination with single measurements made in LabVIEW. For this the program "Grab and store oscilloscope raw data.vi" (by Mats Andersson) can for instance be used.

B.0.8 Standard addition measurements *STDMeasurement_evaluator.m*

This program is to be used in association with *StandardAddition_Measurements.vi*. The data files read by the LabVIEW program is processed by the MatLab program. The processing results in the calculated GMS value as a function of the distance. There are a couple of parameters which has to be set.

- The lock-in sensitivity used. This is to be taken from the *StandardAddition_Measurements.vi* program.
- The distance moved, and the increments used.

It is important that the data files containing the acquired data have the proper names. The data files are to be named in the following manner:

- datafileXYZ.xls
- X: The current measurement

  - 1: Laser A, 1f signal
  - 2: Laser A, 2f signal
  - 3: Laser B, 2f signal
  - 4: Laser B, 1f signal

- Y: The distance between laser and detector (Only the variable part).
- Z: The sensitivity (the numbers corresponding to the different sensitivities are determined by the *StandardAddition_Measurements.vi*)

B.0.9 Calculating the standard addition matrix *CalculatingTheSTDMatrix.m*

When the standard addition gradients have been found by the use of the *STDMeasurement_evaluator.m* program, they have to be extrapolated into a plane. This is done with a MatLab program called *CalculatingTheSTDMatrix.m*. The measurements known are entered into the program together with which combination of sensitivity settings do they occur. This generates a a matrix, giving the gradient for ten different 2f sensitivity settings and ten different 1f sensitivity settings. The program makes simultaneous calculations for both measurement series (Laser A, B above) and plots the result.

B.0.10 Calculating the separation angle between the 1f-signals *OneFEvaluator.m*

The 1f signal is registered for the two different lock-in phase angles, that is when Laser A and Laser B respectively, have their maximum. However, for this measurement only Laser A will be running. The two data files (1f for A and B) is feed to the MatLab program. This program is designed to test the separation of the signals achieved by the used formula. The 1f offset is plotted as a function of the angle being used in the calculations. Since one knows that the 1f signal for laser B should be zero, a proper angle can be established. This is later feed to the other programs. The reason for the this program is the necessity for a good 1f signal separation and the difficulty of predicting a 1f separation angle.