Modulation transfer spectroscopy for fast, accurate laser stabilisation

Vladimir Negnevitsky
Student ID: 19849397

MONASH University

Supervised by
Dr. Lincoln Turner
School of Physics
Modulation transfer spectrometer constructed during this Honours project.
Abstract

This thesis presents a study of modulation transfer spectroscopy (MTS), a powerful technique used to create narrow spectral features suitable for the electronic frequency stabilisation of a diode laser. A spectrometer using an electronically demanding but inexpensive acousto-optic modulator was constructed. After optimising the electronic and optical elements, a control bandwidth of 3 MHz was achieved. This is comparable to more expensive electro-optical modulation methods. The spectrometer was applied to the rubidium $D_2$ transitions, and an unequal dependence on modulation frequency for different transitions was observed. Distortion of the MTS spectrum at high laser beam powers was noted, and a procedure for minimising residual amplitude modulation (RAM) was developed.
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Introduction</strong></td>
<td>1</td>
</tr>
<tr>
<td>1.1 Motivation for laser locking</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Techniques for producing an error signal</td>
<td>2</td>
</tr>
<tr>
<td><strong>Theory</strong></td>
<td>10</td>
</tr>
<tr>
<td>2.1 Modulated spectroscopic techniques</td>
<td>10</td>
</tr>
<tr>
<td>2.2 Modulation transfer spectroscopy</td>
<td>13</td>
</tr>
<tr>
<td>2.3 Simulation of MTS signals</td>
<td>15</td>
</tr>
<tr>
<td><strong>Optical setup</strong></td>
<td>21</td>
</tr>
<tr>
<td>3.1 External-cavity diode laser</td>
<td>21</td>
</tr>
<tr>
<td>3.2 Spectrometer layout</td>
<td>21</td>
</tr>
<tr>
<td>3.3 Photodetection</td>
<td>23</td>
</tr>
<tr>
<td><strong>Modulation</strong></td>
<td>24</td>
</tr>
<tr>
<td>4.1 Choice of modulation method</td>
<td>24</td>
</tr>
<tr>
<td>4.2 Generation of frequency-modulated signal</td>
<td>26</td>
</tr>
<tr>
<td>4.3 Characterisation of frequency-modulated signal</td>
<td>29</td>
</tr>
<tr>
<td>4.4 AOM double pass</td>
<td>31</td>
</tr>
<tr>
<td>4.5 Beat signals</td>
<td>33</td>
</tr>
<tr>
<td>4.6 Summary</td>
<td>36</td>
</tr>
<tr>
<td><strong>Demodulation</strong></td>
<td>37</td>
</tr>
<tr>
<td>5.1 Electronic demodulation</td>
<td>37</td>
</tr>
<tr>
<td><strong>Optimisation of modulation transfer...</strong></td>
<td>39</td>
</tr>
<tr>
<td>6.1 Modulation frequency</td>
<td>39</td>
</tr>
<tr>
<td>6.2 Minimising RAM distortion</td>
<td>42</td>
</tr>
<tr>
<td>6.3 Beam power</td>
<td>43</td>
</tr>
<tr>
<td>6.4 Optimising MTS for laser locking</td>
<td>46</td>
</tr>
<tr>
<td>6.5 Summary</td>
<td>47</td>
</tr>
<tr>
<td><strong>Conclusion</strong></td>
<td>48</td>
</tr>
</tbody>
</table>

## References


1 Introduction

1.1 Motivation for laser locking

Laser cooling and trapping is a burgeoning field that facilitates the cooling of atoms to millionths of a Kelvin. The theoretical study and experimental development of effective laser cooling and trapping techniques has been critical to the creation of Bose-Einstein condensates (BECs) in the last fifteen years [1], and a laser cooling setup is an important component of a BEC laboratory.

To illustrate laser cooling, consider an atom travelling through a laser beam towards the laser source. The atom has a Doppler shift $\delta(v)$ associated with its velocity $v$, and the laser light is blueshifted in the atom rest frame. If we wish the atom to absorb light at some atomic transition frequency $\omega_0$, the laser must be tuned to $\omega_L = \omega_0 - \delta(v)$; the atom has a significant scattering cross-section for photons of this frequency. For every photon absorbed, the atom momentum decreases by $\Delta p = \hbar \omega_L / c$; this can be thought of as ‘tennis balls being thrown at the front of a truck rolling forward’. The scattering occurs in random directions, thus the net effect of many such interactions is to slow the atom along the laser beam axis. Critically, atoms that have zero axial velocity or are moving away from the source will have a different Doppler shift, and will absorb much less light—thus we have a technique by which we can selectively slow atoms.

A commonly-used laser cooling technique is the magneto-optical trap (MOT). It relies on six such detuned laser beams\(^1\), arranged at the faces of a cube and focused on a narrow central volume of space, as shown in Figure 1.1. Within this volume, each laser lessens the component of an atom velocity vector directed towards it, and atoms gradually lose velocity (and thus kinetic energy) down to a quantum-mechanical threshold at tens of microkelvin [1]. The Doppler shift of a typical rubidium atom at 100 µK is 300 kHz, which places a lower limit on the laser precision needed to effectively slow atoms to this temperature. We wish to use an inexpensive laser diode, of the form used in consumer electronics such as CD players. The linewidth of a laser diode, however, is on the order of 50 MHz, and as we will see it is a nontrivial problem to reduce this down to a value usable for laser cooling. Despite the complications, this is still preferable to using a stable but very expensive titanium sapphire laser.

It turns out that the ‘linewidth’ of a laser diode is a result of rapid frequency shifts over timescales smaller than the sampling time of the detector we are using. Shifts much slower than the sampling time manifest as drifts in the line frequency. A multitude of effects cause frequency shifts, all with characteristic timescales; fast effects include laser current and voltage noise and acoustically-caused changes in the laser cavity, while slower drifts are caused by changes in temperature or air pressure.

The frequency of a laser diode can be adjusted very rapidly by varying the diode current, far more rapidly than the timescale of significant effects. We follow a strategy of rapidly applying such adjustments to compensate for these effects, which involves the design of a feedback system with a rapid response time. Applying such a system to stabilise a laser frequency is known as laser locking.

\(^1\)Beam polarisation is also important, but is beyond the scope of this introduction.
CHAPTER 1. INTRODUCTION

Laser frequency
Voltage
0
f₀ : desired frequency
a) b)
Adapted from 
http://quantumoptics.es/MorganMitchell/research/MOT.GIF , 5/6/09

Figure 1.1: a) The layout of a typical magneto-optical trap used in laser cooling, adapted from Reference 2. b) An ideal dispersive error signal, crossing zero at the exact frequency of the desired lock point.

Commercial feedback systems exist with excellent performance, such as the MOG DLC-202 laser diode controller used in this project, but all of these require information on the deviation of the laser frequency from some reference point. We supply this by generating an electronic error signal as shown in Figure 1.1 which provides information on the laser frequency. The signal crosses zero for the desired frequency, and the voltages obtained in the central slope region are proportional to frequency deviations. A good error signal has a steep zero crossing, to maximise the voltage deviation for a given frequency deviation; it is broad, to ensure locking over a wide range of frequencies; and it has a high signal-to-noise ratio (SNR) for locking stability. Such a signal is said to have a dispersive lineshape.

The remainder of this introduction will explore various techniques of generating error signals, and compare their advantages and disadvantages. Spectroscopic techniques will be focused on, for reasons outlined in the next section. It will be shown that modulation transfer spectroscopy (MTS) is the optimal technique for our purposes.

1.2 Techniques for producing an error signal

Non-spectroscopic techniques

Non-spectroscopic methods to produce an error signal are commonly used in scenarios where long-term frequency drifts are less important than short-term precision. A common and simple strategy is sending a laser through a carefully constructed optical cavity, or étalon, which consists of a pair of slightly transparent mirrors opposite one another. A beam is directed through the cavity, and the transmitted intensity is monitored. The cavity resonance condition is satisfied when an integral number of laser wavelengths fit within; constructive interference occurs for this cavity mode and a peak in transmitted intensity occurs. By using the Pound-Drever-Hall locking strategy, this peak can be converted to an error signal [3]; this process is discussed in §2.1.

It must be admitted that optical frequency references have been used to obtain the narrowest linewidths [4], but no macroscopic device can be perfectly shielded from acoustic vibrations, temperature changes and other physical disturbances; elaborate measures can be used to minimise these [4] however the advantages of convenience and inexpensiveness are lost. Atomic transitions, on the other hand, are absolute references independent of most macroscopic effects (except electromagnetic fields), and the spectrum of a pure atomic
1.2. TECHNIQUES FOR PRODUCING AN ERROR SIGNAL

sample such as rubidium is known to a high accuracy.

Most of the laser spectroscopic techniques discussed are usually used to investigate the unknown properties of a sample, by obtaining and analysing absorption and/or dispersion spectra across a spread of precisely-known frequencies. In laser locking the converse process is carried out: the sample spectrum is known to a high precision, and the laser frequency is deduced.

With these motivations in mind, we discuss several unmodulated (baseband) spectroscopic methods of producing an error signal; modulated techniques are discussed in the next chapter.

**Saturated absorption spectroscopy**

Ordinary absorption spectroscopy is straightforward: a laser beam of known frequency $\omega_c$ is sent through a sample, in our case rubidium vapour, and the intensity is measured by a photodetector on the other side. The vapour is transparent to light when $\omega_c$ is far from any transition frequencies of the atoms. When $\omega_c$ nears an atomic transition frequency $\omega_0$, atoms absorb photons from the beam and are excited. Each atom remains excited for a short time, then a photon is emitted in a random direction and the atom decays back to the ground state. The scattering rate defines how often this occurs. Because few photons are randomly emitted towards the photodetector, the intensity is reduced [5, p79].

The scattering rate of a transition also defines its natural linewidth $\gamma$, on the order of 6 MHz for the rubidium $D_2$ transitions [6]. We might imagine that a graph of intensity against $\omega_c$ would show a dip of similar width centred on $\omega_0$. This is not the case: the velocities of the vapour atoms along the beam axis are normally distributed (with an average magnitude of 300 m/s at 300K for rubidium), Doppler shifting their absorption frequencies in the rest frame. Now, even when $\omega_c$ is detuned from $\omega_0$, some atoms are suitably red- or blueshifted by the frequency needed to compensate for the detuning, and the vapour thus continues to absorb light. This is known as Doppler broadening, and is a strong effect: the rubidium $D_2$ linewidths are broadened by $\sim 400$ MHz at 300K, entirely obscuring finely-spaced spectroscopic features.

In a steady-state ensemble of simple two-level atoms, the excitation rate of the ground state population is equal to the decay rate of the excited state population. If a sufficiently intense laser beam is used, the excitation rate grows significantly, and atoms migrate from the ground to the excited state until equilibrium is reestablished. This is known as saturation\(^2\), and only occurs for those atoms whose velocity-shifted transition frequencies are near $\omega_c$. As shown in Figure 1.2, a hole forms in the ground state atom velocity distribution, at only the velocity where absorption can occur. This is called hole burning.

If a weak counter-propagating 'probe' beam is sent through the vapour at the same frequency $\omega_c$, the atoms interacting with this beam have the opposite Doppler shift to the atoms being excited by the saturating 'pump' beam. We can imagine varying $\omega_c$ as $\omega_c$ is swept from low to high frequency: point Q travels right, point P travels left, and the two meet at the centre when $\omega_c = \omega_0$. Here, the hole caused by the pump beam is detected by the probe beam as a decrease in absorption, as fewer atoms can absorb the probe beam. Equivalently, the two beams only mutually interact with a group of atoms whose velocity is zero along the beam axis. This condition is satisfied when $\omega_c = \omega_0$ within a natural linewidth.

This technique is known as saturated absorption spectroscopy (SAS), and provides sub-Doppler spectral resolution as shown in Figure 1.2b. Sharp peaks are obtained, and a laser

\(2\)In addition to absorption and random emission, stimulated emission becomes significant in this case, reducing the absorption of the saturating beam. The effect is tangential to the discussion. It is covered in Demtröder [5, p80-82].
CHAPTER 1. INTRODUCTION

Figure 1.2: a) Hole burning in a velocity distribution. Pump beam burns a hole at P, probe beam detects absorption at Q. The hole is detected by the probe beam when P is at Q, for atoms with zero Doppler shift at the centre of the distribution. Adapted from Reference 5, Figure 2.4. b) Saturated absorption spectrum of Rb D2 transitions. The lower line is the Doppler-broadened spectrum. Probe power: < 500 nW, pump power: 0.9 mW. Adapted from Reference 6, Figure 1.

Figure 1.3: The 85Rb D2 manifold. All spacings are given in megahertz unless otherwise marked. Allowed transitions are $F = 2 \rightarrow F' = \{1, 2, 3\}$ and $F = 3 \rightarrow F' = \{2, 3, 4\}$. Black lines show true atomic energy levels; red lines show crossovers that are seen in the saturated absorption spectrum. Adapted from Reference 8.

can be locked to the steep side of a peak by subtracting a DC offset from the spectroscopic signal. This basic technique is known as side locking, and is only as stable as the DC offset and associated noise. More advanced and stable techniques that apply the principles of SAS but lock to the top of a peak, such as FM spectroscopy and MTS, are discussed in later sections.

Before proceeding further, some points are mentioned about the structure of rubidium.

Hyperfine structure of rubidium

In rubidium, the 780 nm D2 line is due to the $5^2S_{1/2} \rightarrow 5^2P_{3/2}$ transition. The nuclear spin of 85Rb is $I = \frac{5}{2}$, and hyperfine coupling splits the fine structure energy levels into hyperfine states. We define a new quantum number $F = I + J$. The standard angular momentum selection rule for dipole transitions applies: $\Delta F = 0, \pm 1$. The ground states are now $F = \{2, 3\}$ and the excited states are $F' = \{1, 2, 3, 4\}$. The $F = 2 \rightarrow F' = 1$ and $F = 3 \rightarrow F' = 4$ transitions are closed, as once they are excited the atom can only return to one ground state.

Figure 1.3 shows an energy level diagram for the rubidium D2 transitions. Black lines show the energy levels of the transitions; red lines show crossover energies that are seen in
1.2. TECHNIQUES FOR PRODUCING AN ERROR SIGNAL

Figure 1.4: i shows a) the Doppler-broadened absorption of the vapour cell with no magnetic field; b) the absorption of right-handed light by the cell with a magnetic field applied; c) the absorption of left-handed light by the cell; d) the difference between spectra b) and c), which provides an error signal. The sign of the shift depends on the transition concerned. ii shows a) a standard Rb saturated absorption spectrum; b) an error signal from the same cell obtained using DAVLL. A 100-G magnetic field is applied. Note the ∼0.5 GHz recapture range. Adapted from Reference 9, Figures 2 and 3.

the saturated absorption spectrum for particular velocity classes of atoms whose Doppler shift brings the pump beam into resonance with one transition, and the probe into resonance with another. The $F = 3 \rightarrow F' = \{2, 4\}$ and $F = 3 \rightarrow F' = \{3, 4\}$ crossovers represent the large peaks in the saturated absorption spectrum in Figure 1.2b. The large size is due to hyperfine pumping, beyond the scope of this thesis. The physics of saturated absorption and hyperfine pumping in the rubidium $D_2$ transitions is discussed in Reference 7.

This structure is referred to in later chapters. Next we discuss several commonly used baseband laser locking techniques.

Dichroic atomic vapour laser lock (DAVLL)

DAVLL is a relatively simple, low-cost technique that, unlike FM and MTS, does not require frequency modulation of the laser beam. Additionally the frequency recapture range of a DAVLL system is broader than saturated absorption systems [9], thus a DAVLL system will remain locked despite jumps in the laser frequency that would throw other systems off. It has been applied to the rubidium $D_2$ transitions, recently achieving error signal amplitudes and gradients on the order of a conventional Doppler-free saturation spectrometer [10][11].

A linearly-polarised laser beam is sent through a vapour cell, with a magnetic field applied parallel to the beam wavevector. The beam can be treated as a superposition of right-handed and left-handed circularly polarised components $\sigma^+$ and $\sigma^-$, each of which interacts with some atomic resonances but not others. For instance if $\sigma^+$ interacts with a $(J, M_J) \rightarrow (J + 1, M_J + 1)$ transition, it will not interact with its degenerate $(J, M_J) \rightarrow (J + 1, M_J - 1)$ counterpart due to angular momentum conservation, while $\sigma^-$ will interact with the second transition but not the first. Since both transition energies are degenerate, the (Doppler-broadened) absorption profiles of $\sigma^+$ and $\sigma^-$ will be the same.

The degeneracy can be lifted, however, by applying a magnetic field along the beam axis, Zeeman-splitting degenerate energy levels. The absorption profiles of $\sigma^+$ and $\sigma^-$ are shifted away from the zero-field case, the shift being proportional to the magnetic field strength. The vapour absorbs the two polarisations unequally; thus it has become dichroic [11].

The transmitted beam is passed through a $\lambda/4$ plate, transforming the circularly-polarised components into orthogonal linearly-polarised components. These are split up into individual beams by a polarising beamsplitter, and sent to a pair of photodetectors. The difference in the beam intensities gives an error signal resembling a dispersion curve;
this is free from common-mode errors such as changes in laser intensity, although single-ended errors, such as the dimming of one optical path relative to the other, remain.

Additionally, the lock frequency can be directly tuned by rotating the χ/4 plate, altering the relative amplitudes of the components [11]— this is equivalent to adding an electronic DC offset to the DAVLL difference signal, but avoids that potential source of noise.

The components σ⁺ and σ⁻ encounter Doppler-broadened energy levels, leading to a lack of spectral resolution that is a feature of standard DAVLL spectra. Setups exist that eliminate the Doppler broadening [12], but the simplicity and large recapture range of standard DAVLL systems is lost. Unmodulated polarisation spectroscopy is an alternative technique based on similar principles which offers sub-Doppler resolution while requiring only a single detector, eliminating differential errors.

The key advantages of DAVLL systems are their simplicity and low cost. In our scenario, however, DAVLL is not ideal for several reasons. DAVLL requires careful alignment due to its fine dependence on polarisation angle, and though a DAVLL system has a large recapture range, long-term frequency drifts can occur due to thermal and other environmental effects [12]. Several interesting schemes have been applied to improve the amplitude and gradient of the error signal, such as using an electromagnetic coil as a heater to increase the vapour pressure in the cell [10] and sending a second, ‘pump’ beam through the cell to obtain DAVLL lineshapes with sub-Doppler features [12] (though the latter leads to a narrower recapture range); but DAVLL remains susceptible to frequency drift. While modulated spectroscopic techniques such as MTS lock to the peak of atomic transitions, this is not the case for DAVLL— this can be an advantage when a freely tunable lock is desired, but is a disadvantage in our case.

**Sagnac interferometry**

Sagnac interferometry consists of sending a laser beam through a beamsplitter, directing the components in opposite directions around an optical path, then recombining the beams. When applied to laser locking, a common strategy is to place a neutral density filter (NDF) and a vapour cell along the path [13, 14]. One beam interacts with the cell first, then is attenuated by the NDF; the other is attenuated by the NDF then enters the cell. Figure 1.5 shows the layout of a Sagnac interferometer.

The two beams play the same roles as a pump and probe beam in standard SAS. Both are equally attenuated **in total** by the NDF and the linear absorption of the vapour cell, but the pump beam is brighter when it reaches the cell. The intense pump beam saturates the vapour without being further affected, but the probe is further attenuated and phase shifted by the Doppler-free resonance it alone encounters [15].

A common method uses one detector for each interfered output path, as shown in Figure 1.5. This is the type analysed below. A linearly polarised laser beam, with frequency near an atomic transition of the vapour, is passed through a χ/4 plate, converting to circular polarisation. The pump beam is split off by the beamsplitter and passes around the triangle ACB anticlockwise, while the remaining probe beam travels clockwise. Both beams recombine and interfere at A. Part of the beam travels to photodetector P, while the rest travels downwards to a polarising beamsplitter (PBS) that deflects it to photodetector Q. A PBS must be used to avoid feedback into the laser [13]. Single-photodetector methods exist [15].

Without detailed analysis, we describe each beam by a plane wave. It can be shown [13]
1.2. TECHNIQUES FOR PRODUCING AN ERROR SIGNAL

Figure 1.5: Standard Sagnac setup, using two photodetectors. The probe and pump beam travel around paths ABCA and ACBA respectively, interfere at A and are measured at P and Q. Adapted from Reference [13] Figure 1.

that the sum and difference of the photodetector intensities $I_P$ and $I_Q$ are described by

$$I_P + I_Q = I_0^2 \frac{(e^{-\alpha_1 L} + e^{-\alpha_2 L})}{2}$$

$$I_P - I_Q = I_0^2 \left[ e^{-(\alpha_1 + \alpha_2)L/2} \sin \left( k \Delta n L \right) \right]$$

where $I_0$ is the initial laser intensity, $t$ is the ratio of pump to probe fields, $\alpha_{1,2}$ are the absorptions and refractive indices encountered in the cell by the pump and probe beams respectively, and $\Delta n = n_1 - n_2$.

When $k \Delta n L$ is small, which is true in our case, $\sin (k \Delta n L) \approx k \Delta n L$; thus $I_P - I_Q \propto \Delta n$.

For suitably low-power beams, $\exp(- (\alpha_1 + \alpha_2)L/2)$ alters the spectral profile very little, and the refractive index of the vapour cell is close to unity for the pump beam ($n_1 \simeq 1$). Thus, $\Delta n \approx 1 - n_2$.

The absorption spectrum can be modelled using a Lorentzian function [5, p62], and the dispersion spectrum is related to this by the Kramers-Kronig relations [5, p68]. The absorption and dispersion are described by

$$\alpha(\omega) = K_1 \omega_0 \frac{\gamma \omega}{(\omega^2 - \omega_0^2)^2 + \gamma^2 \omega^2}$$

$$\Delta n(\omega) = K_2 \frac{\omega_0^2 - \omega^2}{(\omega^2 - \omega_0^2)^2 + \gamma^2 \omega^2}$$

where $K_1$ and $K_2$ are physical constants not pertinent to the discussion, $\omega_0$ and $\gamma$ are the atomic transition frequency and natural linewidth, and $\omega$ is the laser frequency. Thus, the $I_P - I_Q$ signal follows a dispersion-like curve similar to (1.2), which crosses zero at $\omega = \omega_0$ and can thus be used as an error signal.

Sagnac interferometry is a scheme that is straightforward and inexpensive to implement. The apparatus is among the most compact of the commonly-used laser locking schemes [14], and unlike most DAVLL implementations it resolves sub-Doppler detail. Alignment is time-consuming, however, and the interferometer is misaligned more easily than other spectroscopic techniques. Being a baseband technique, it is susceptible to low-frequency laser and environmental noise. Thus, while it is an interesting and useful technique with many applications outside spectroscopy, such as verifying general relativity [16], it is not ideal for our needs.

Polarisation spectroscopy

Polarisation spectroscopy is related to SAS, but offers a significantly better signal to noise ratio [17]. Like DAVLL, it relies on inducing birefringence in a vapour cell, but this is done by circularly polarising the pump beam [5] instead of applying an external magnetic field.
A typical polarisation spectrometer is shown in Figure 1.6. A weak probe beam of frequency \( \omega \) passes through a linear polariser \( P_1 \), the vapour cell, and a second linear polariser \( P_2 \), orthogonal to the first, into a detector. Because the polarisers are crossed, the probe beam is almost entirely absorbed by the second polariser. An intense circularly-polarised pump beam of equal frequency passes in the opposite direction through the cell. We assume that our vapour cell consists of atoms with a P transition at frequency \( \omega_0 \), and the total angular momentum \( J \) shifts from \( J = 2 \) at the ground state, to \( J = 1 \) in the excited state. The quantum states are restricted to \((J, \{M_J\}) = (2, \{-2, -1, 0, +1, +2\}) \) and \((J, \{M_J\}) = (1, \{-1, 0, +1\}) \) for energy levels 1 and 2 respectively, where \( M_J \) is the projection of \( J \) along the pump beam axis. Assuming right-handed polarisation, to preserve conservation of angular momentum the pump beam can only stimulate transitions in which \( M_J \to M_J + 1 \), as shown in Figure 1.7. Thus the transitions from \((2, +1)\) and \((2, +2)\) are forbidden and the populations of these states are not depleted by the pump beam. Thus, while the energy of this transition is degenerate, the populations of its constituent states are not, and because these populations determine the absorption and refractive index of the vapour, the vapour becomes birefringent.

As in the DAVLL discussion, the probe beam can be decomposed into two circularly-polarised components, which interact unequally with the now-birefringent vapour. In a process related to Faraday rotation, one helicity is preferentially absorbed, breaking the symmetry of the components and leading to an elliptically-polarised beam. Now, rather than absorbing all of the probe beam, \( P_2 \) transmits a fraction of it, and the intensity at the detector is no longer zero.

As in standard SAS, the counter-propagating beams only mutually interact with those atoms whose axial Doppler shifts are less than the transition linewidth \( \gamma \); all other atoms are transparent to either the pump beam, the probe beam, or both. Thus, a sub-Doppler spectrum is obtained.

For a sweep of the laser frequency, the probe intensity will show Lorentzian lineshapes at transition frequencies, with a near-zero background. This is similar to an SAS spectrum, and can be used in the same way to create an error signal. A dispersive component can be added into the signal by slightly uncrossing the polarisers \cite{5, 18}, so that the
1.2. **TECHNIQUES FOR PRODUCING AN ERROR SIGNAL**

Figure 1.8: Polarisation spectrometer, with a polarisation beamsplitting cube used in place of crossed polarisers. Adapted from Reference 19, Figure 4

signal both rises and falls below the baseline as the probe beam’s orientation varies—the absorption and dispersion profiles are related by (1.1) and (1.2). A dispersive signal has a distinct zero crossing, and can be used as an error signal. The spectrometer can be modified by adjusting the beam intensities and polariser angles to produce a primarily dispersive signal—this is preferable in laser locking.

The polarisers can be replaced with a polarisation beamsplitting cube as shown in Figure 1.8 producing two beams of orthogonal linear polarisation that are sent to a pair of detectors. A good dispersion spectrum is obtained by subtracting the two signals; this method produces a signal whose dispersive component is an order of magnitude larger than that of the polariser technique discussed, and rejects common-mode laser noise [19].

Polarisation spectroscopy is a popular technique, owing to its improved signal-to-noise ratio over SAS with little additional complexity. However, relying on small changes in beam polarisation increases the susceptibility of polarisation spectroscopy to issues that do not affect other techniques, such as the stress in vapour cell walls and even waveplates. In our case, its disadvantages lie mainly in its DC nature and the superposition of both absorptive and dispersive components in the error signal, which shifts the zero crossing away from a transition line centre.

In the next chapter we discuss a class of techniques which have less noise still, and produce purely dispersive error signals.
2 Theory

We have discussed the motivations and basic principles of laser locking, and compared some commonly used baseband spectroscopic locking techniques. In this chapter we discuss the theory behind modulated spectroscopic techniques, of which modulation transfer spectroscopy is an example.

2.1 Modulated spectroscopic techniques

Frequency modulation

At a point in space, the electric field amplitude of a travelling monochromatic electromagnetic wave $E_1(t)$ is described by

$$E_1(t) = E_0 \sin (2\pi f_c t),$$

where $E_0$ is the peak field amplitude and we call $f_c$ the carrier frequency. If the carrier frequency is phase modulated sinusoidally at a modulation frequency $f_m$, to create a frequency modulated (FM) wave $E_2(t)$ (methods of achieving this are discussed in §4.1), then $E_2(t)$ can be treated spectrally as a carrier frequency $f_c$ with sidebands spaced at intervals of $f_m$, whose amplitudes are defined by Bessel functions [20],

$$E_2(t) = E_0 \sin \left[2\pi f_c t + \delta \sin(2\pi f_m t)\right],$$

where $\delta$ is the modulation index. Note that frequency and phase modulation are identical except in their mathematical description [20]; they are used interchangeably throughout this thesis. If the modulation index is less than 1, then $J_0(\delta) \simeq 1$, $J_{\pm 1}(\delta) \simeq \pm \delta/2$ and $J_{\pm n}(\delta) \simeq 0$, $|n| > 1$. Now all sidebands other than $J_{\pm 1}$ may be neglected. Thus (2.3) simplifies to

$$E_2(t) = E_0 \left\{\frac{\delta}{2} \sin[2\pi(f_c + f_m)t] + \sin(2\pi f_c t) - \frac{\delta}{2} \sin[2\pi(f_c - f_m)t]\right\}.$$  

The time derivative of the phase in (2.2) gives the instantaneous frequency $f_i$,

$$f_i(t) = f_c + \delta f_m \cos(2\pi f_m t),$$

from which the instantaneous frequency deviation $f_{Di}(t) = f_i(t) - f_c$ of the signal reaches a maximum value of

$$f_D = \delta f_m.$$  

There are times when either $f_D$ or $\delta$ is more convenient to use than the other, thus both are used in this thesis. Also, angular frequencies $\omega_c$, $\omega_m$ and $\omega_D$ are at times used in favour of frequencies $f_c$, $f_m$ and $f_D$ in some of the thesis, such as the next section; [2.1]–[2.6] remain valid for the substitutions $\omega = 2\pi f$. Optical frequencies are signified by $\nu$, while $f$ signifies rf frequencies: for instance, unmodulated 780 nm light has a frequency of $\nu_c \sim 380$ THz, while $f_c \sim 80$ MHz in much of the project electronics.
2.1. MODULATED SPECTROSCOPIC TECHNIQUES

Figure 2.1: Basic FM spectrometer. The beam is frequency modulated, sent through the sample where it acquires amplitude modulation (AM), and demodulated. The phase of the resultant signal is proportional to the frequency detuning. EOM: electro-optic modulator. PMPA: post-mixer pre-amplifier. Adapted from Reference 21

Frequency-modulated spectroscopy

While each baseband technique has clear advantages, they all rely on DC electronics to generate an error signal. To avoid noise entering the system due to ambient lighting shifts, the optical elements must be well shielded; this makes on-the-fly adjustment cumbersome. Modulated spectroscopic techniques such as Pound-Drever-Hall and MTS avoid this pitfall by modulating the laser frequency. The previously discussed techniques produce DC signals from the photodetectors, while FM produces a DC + AC signal. While the DC component is strongly affected by ambient lighting, laser intensity fluctuations and other low-frequency disturbances and noise sources, it is removed from the AC signal in demodulation. The radio-frequency (rf) electronics typically used in demodulation are generally insensitive to frequencies below 10 kHz, automatically removing very low-frequency noise, and a correctly tuned high-pass filter will remove any remainder. Thus, skews in the zero crossing and lock point of the system due to intensity fluctuations and low-frequency noise are avoided.

With these motivations in mind, a basic FM locking system modulates a laser beam using an electro-optic or acousto-optic modulator to create a beam whose electric field amplitude follows $E_2(t)$ in (2.4). Optical modulation is discussed in §4.1. The modulated beam is sent through a reference whose absorption changes as a function of frequency, such as a vapour cell for spectroscopic methods or an optical reference cavity for Pound-Drever-Hall and related techniques.

Figure 2.1 shows a simple FM spectrometer, with a sample FM spectrum. $E_2(t)$ is sent through the vapour cell. When $\omega_m$ is near a resonance, the sidebands experience different absorptions and dispersions relative to the carrier and to one another, and are shifted in amplitude and phase, resulting in $E_3(t)$. This process essentially adds amplitude-modulated (AM) spectral components to the beam. The square of $E_3(t)$ gives the intensity of the wave, which is picked up by a photodetector that produces a proportional output. Following the
derivation in \cite{21}, the intensity at the photodetector \(I_3(t)\) is given by
\[
I_3(t) = KE_0^2 \left[ 1 + (\alpha_j - \alpha_1)\delta \cos(\omega_m t) + (\phi_1 + \phi_1 - 2\phi_0)\delta \sin(\omega_m t) \right],
\] (2.7)
where \(K\) is a constant for our purposes, \(\alpha_j\) and \(\phi_j\) are the amplitude attenuation and phase shift experienced by a spectral band due to the vapour cell, and indices \(j = -1,0,1\) denote the \(\omega_c - \omega_m, \omega_c\) and \(\omega_c + \omega_m\) bands.

Thus the cosine term in (2.7) varies as the difference in sideband attenuation, and the sine term is proportional to the difference between the phase shift experienced by the carrier and the average of the phase shifts experienced by the sidebands. For the following discussion, we note that for small \(h\), \(f(\omega + h) - f(\omega - h) \simeq h \times df(\omega)/d\omega\). If \(\omega_m\) is small compared to the width of the spectral feature being analysed, the cosine and sine term coefficients \(\alpha_{-1} - \alpha_1\) and \(\phi_1 + \phi_1 - 2\phi_0\) approach:
\[
\alpha_{-1} - \alpha_1 = \alpha(\omega_c - \omega_m) - \alpha(\omega_c + \omega_m)
\]
\[
\simeq -\omega_m \frac{d\alpha(\omega)}{d\omega} \bigg|_{\omega = \omega_c};
\]
\[
\phi_1 + \phi_1 - 2\phi_0 = [\phi(\omega_c + \omega_m) - \phi(\omega_c)] - [\phi(\omega_c) - \phi(\omega_c - \omega_m)]
\]
\[
\simeq \omega_m \frac{d\phi(\omega)}{d\omega} \bigg|_{\omega = \omega_c - \omega_m/2} - \omega_m \frac{d\phi(\omega)}{d\omega} \bigg|_{\omega = \omega_c + \omega_m/2}
\]
\[
\simeq \omega_m^2 \frac{d^2\phi(\omega)}{d\omega^2} \bigg|_{\omega = \omega_c}. \tag{2.8}
\]

Thus the cosine and sine coefficients are proportional to the first derivative of the absorption and the second derivative of the dispersion of the vapour at \(\omega_c\) \cite{22}. The derivatives will cross zero precisely at the peak of a spectral feature, thus they can be used as error signals.

The coefficient amplitudes are obtained by electronically multiplying\(^1\) (2.7) by a signal proportional to \(\cos(\omega_m t + \phi)\). Let us rewrite (2.7) for clarity as
\[
I_3(t) = A + B \cos(\omega_m t) + C \cos(\omega_m t - \pi/2)
\]
where
\[
A = KE_0^2, \quad B = KE_0^2(\alpha_{-1} - \alpha_1)\delta, \quad \text{and} \quad C = KE_0^2(\phi_1 + \phi_1 - 2\phi_0)\delta.
\]

Multiplying \(I_3\) by \(\cos(\omega_m t + \phi)\) and using \(\cos a \cos b = [\cos(a + b)]/2 + [\cos(a - b)]/2\),
\[
I_3(t) \cos(\omega_m t + \phi) = A \cos(\omega_m t + \phi) + \frac{B}{2} \cos(2\omega_m t + \phi) + \cos(\phi)
\]
\[
+ \frac{C}{2} \left[ \cos(2\omega_m t - \pi/2 + \phi) + \cos(-\pi/2 - \phi) \right]. \tag{2.9}
\]

Next a low-pass filter is applied, removing all rapidly oscillating terms. Thus we are left with
\[
I_3(t) \cos(\omega_m t + \phi) = \frac{B}{2} \cos \phi - \frac{C}{2} \sin \phi \tag{2.10}
\]

Thus by altering \(\phi\), the phase of the mixing signal, we can electronically obtain \(B, C\), or a mixture of the two. After the phase is adjusted for the optimal signal amplitude and slope, the result can be used directly as a dispersive error signal.

To resolve sub-Doppler detail using FM spectroscopy, a counter-propagating unmodulated pump beam is sent through the sample to saturate it, analogous to SAS \cite{23,24}. Now

\(^1\)This process is also known as mixing, and is very widely used in communications and signal processing.
the modulated probe beam detects sub-Doppler features in the spectrum, which have steep slopes and are ideal for locking. The signal obtained is dispersive, as in the single-beam case, and can be directly used as an error signal. The Doppler-broadened background can be largely removed in such a system by rapidly cycling the pump beam on and off (‘chopping’) and comparing the signal in these two cases [21].

This setup is closely related to modulation transfer spectroscopy. While an in-depth treatment of Doppler-free FM spectroscopy is beyond the scope of this review, the model describing the error signal is similar to MTS [25] (although the physical process responsible is more complex for MTS).

Many modulated laser locking methods exist, and laser frequency is only one of the possible parameters that lend themselves to modulation. One alternative is a modulated form of Doppler-free DAVLL, which modulates the Zeeman splitting between rubidium energy levels [26]. This is simple to implement, but unlike with FM methods, the modulation frequency is limited by the vapour. A modulated variation of polarisation spectroscopy also exists [27].

Of the techniques discussed, FM methods are the most resilient to DC and environmental noise. Their key advantages lie in the high-frequency modulation of the laser beam, avoiding most noise, and the stability of their spectra when compared to techniques such as DAVLL, where disturbances such as room lighting can shift the lock point. Slow changes in, say, the laser intensity will vary the slope and amplitude of an FM error signal, but the zero crossing point will remain constant. While FM techniques such as MTS use costlier optics and electronics than other methods, their improved performance justifies the expense in our case. Next we discuss MTS, which offers significant improvements in error signal quality over standard FM spectroscopy with no major drawbacks for our purposes.

2.2 Modulation transfer spectroscopy

Doppler-free FM spectroscopy can be thought of as an intermediate technique between saturated absorption and MTS: it shares the electronic strategies of MTS to produce a stable signal and filter out noise, but nonlinear modulation transfer does not occur in an FM spectrometer. Ideally an FM error signal will be zero far from a transition, but this is rarely the case for real setups, as the absorption/dispersion of a sample is slightly frequency-dependent due to the Doppler-broadened background spectrum [21].

In contrast MTS is almost free of background, at the same time producing strong signals only for atomic transitions where the excited atom can return to only one ground state [28]: ‘closed’ transitions. This is desirable for two reasons. An MTS spectrum will have fewer closely spaced zero crossings than an FM spectrum in areas of dense spectral features such as the rubidium D2 transitions, reducing the chance of the lock point ‘skipping’ unpredictably from one crossing to another. Also, the ground state population of a closed transition is not depleted by saturation like an open transition, where atoms may relax into other ground states and be lost to the saturating pump beam, reducing the amplitude of the error signal. It is also particularly convenient in our case to lock a laser to a closed transition in rubidium, as the laser will be used for cooling and trapping rubidium atoms — these techniques are also most effective for closed transitions, again because a larger proportion of the atoms remains accessible to the cooling beam than for open transitions.

A recent MTS spectrometer is shown in Figure 2.2. A pump beam is modulated by an electro-optic modulator, discussed in §4.1 and passes through a vapour cell. Modulation is transferred to a probe beam passing through from the opposite side. In our project an entirely new layout was designed, discussed in Chapter 3. In the remainder of this chapter we summarise the theory of modulation transfer.
Figure 2.2: MTS apparatus. Modulated pump and unmodulated probe beams interact in the vapour cell, and modulation is transferred to the probe beam. PBS = polarising beamsplitter, PD = photodiode, EOM = electro-optic modulator. The dashed path to PD2 can be used to observe the pump beam and obtain an FM spectroscopy signal. Adapted from Reference 28, Figure 2.

Theory of modulation transfer

The electric field of the modulated pump beam is described by (2.3). If the laser frequency $\nu_c$ is near an atomic transition $\nu_0$, modulation is transferred the probe beam. This is discussed in the literature for several pump/probe beam intensities and layouts [25, 29] with several detailed theoretical models, in particular that discussed in [30] and extended to higher-order processes in [25]. Essentially, modulation is transferred through a process of nonlinear four-wave mixing through the $\chi^{(3)}$ susceptibility of the vapour; the pump and probe carrier waves at frequencies $\nu_c$ and a probe sideband at $\nu_c + nf_m$ ($n$ is a nonzero integer) beat nonlinearly to create a fourth wave: a probe beam sideband at $\nu_c + nf_m$. The sideband amplitudes and phases give information on the laser detuning $\nu_c - \nu_0$ as well as the properties of the vapour, similar to (2.7) for FM spectroscopy\(^2\).

The probe, having thus acquired sidebands, reaches the photodetector. The sidebands beat with the main band to generate an electronic signal at the modulation frequency $f_m$ [28], with amplitude

$$S(\Delta) = \frac{C}{\sqrt{\gamma^2 + f_m^2}} \sum_{n=-\infty}^{\infty} J_n(\delta)J_{n-1}(\delta) \left\{ \left( L_{(n+1)/2} + L_{(n-2)/2} \right) \cos(2\pi f_m t + \phi) \right\}$$

where

$$L_n = \frac{\gamma^2}{\gamma^2 + (\Delta - nf_m)^2} \quad \text{and} \quad D_n = \frac{\gamma(\Delta - nf_m)}{\gamma^2 + (\Delta - nf_m)^2}; \quad \Delta = \nu_c - \nu_0. \quad (2.12)$$

Here $\gamma$ is the natural linewidth of the $\nu_0$ transition, $\Delta$ is the laser detuning from the atomic transition, $\phi$ is the phase of the probe beam sidebands with respect to the modulating wave $\cos(2\pi f_m t)$, and $C$ is a function of the beam field amplitudes and physical constants. Note that $L_n$ is a Lorentzian curve and $D_n$ is related to it by the Kramers-Kronig relations, as in (1.1) and (1.2); this is because (2.11) is derived from an equation consisting of complex

\(^2\)The literature does not appear to provide a full analysis of the four-wave mixing process in MTS, though a similar scenario is covered in [31]. A model of four-wave mixing requires discussion of nonlinear optics and optical Bloch equations: an honours project in its own right!
exponentials \(^{32}\). The cosine and sine coefficients are orthogonal, and are referred to herein as the in-phase and quadrature components of the MTS signal.

When \(\delta\) is less than 1, high-order Bessel function coefficients are small and can be ignored. We need only consider \(n = 0, 1\), and \((2.11)\) becomes

\[
S(\Delta) = \frac{C}{\sqrt{\gamma^2 + f_m^2}} J_0(\delta) \left\{ \left[ J_1(\delta) \left( L_1 + L_{-1/2} \right) + J_{-1}(\delta) \left( L_{1/2} + L_{-1} \right) \right] \cos(2\pi f_m t + \phi) 
+ \left[ J_1(\delta) \left( D_1 - D_{-1/2} \right) + J_{-1}(\delta) \left( D_{1/2} - D_{-1} \right) \right] \sin(2\pi f_m t + \phi) \right\} 
\]

(2.13)

Since \(J_{-1} = -J_1\), this further simplifies to

\[
S(\Delta) = \frac{C}{\sqrt{\gamma^2 + f_m^2}} J_0(\delta) J_1(\delta) \left\{ \left( L_1 + L_{-1/2} - L_{1/2} - L_{-1} \right) \cos(2\pi f_m t + \phi) 
+ \left( D_1 - D_{1/2} - D_{-1/2} + D_{-1} \right) \sin(2\pi f_m t + \phi) \right\} 
\]

(2.14)

It can be shown \(^{25}\) that the in-phase and quadrature components are related to the absorption and dispersion of the vapour; intuitively this is clear from the similarity of \(L_n\) and \(D_n\) to absorptive and dispersive lineshapes as mentioned above. The relationship is more complex than the case of \((2.8)\), and produces dispersive lineshapes of better shape than in FM.

We can recover these coefficients by mixing with a sinusoid and low-pass filtering, as in the FM case, which provides a good-quality error signal free from the residual Doppler background of FM spectroscopy \(^{30}\). By adjusting the phase of the mixing signal with reference to the modulating sinusoid, the signs and relative contributions of the in-phase and quadrature components can be varied analogously to \((2.9)\) and \((2.10)\).

For reasons discussed previously, desirable properties of an error signal are a high signal-to-noise ratio, amplitude, slope and recapture range. While SNR levels were largely due to experimental parameters such as electronics and beam intensities, insight was gained into the other parameters by simulating MTS as a function of modulation frequency and phase, using \((2.11)\).

### 2.3 Simulation of MTS signals

The in-phase MTS signal is a sum of Lorentzian functions, whose centres are offset from the transition frequency by amounts proportional to the modulation frequency; Figure 2.3a shows this for the two-sideband approximation. At low modulation frequencies the Lorentzians largely cancel out one another, but as the modulation frequency is increased, their centre frequencies separate — increasing their sum near the centre of the signal (note that at the centre, their sum is always 0). As the modulation frequency is increased further, their overlap becomes smaller still, and the MTS signal develops a ‘kink’ due to the \(L_{\pm 1/2}\) Lorentzians. This increases with modulation frequency until the inner peaks reach the size of the outer peaks. The trend is shown in Figure 2.4a.

The quadrature signal is composed of four Lorentzian derivatives, as shown in Figure 2.3b. Its shape changes little except above 12 MHz, outside the range of our apparatus, however it is only 3/4 as wide as the in-phase signal; its properties are shown in Figure 2.4b.

The MTS signal amplitude, recapture range and zero crossing slope were simulated at various phases and modulation frequencies, with a constant modulation index of 0.6. The recapture range was obtained by numerically differentiating and obtaining the central range over which the MTS derivative remained a constant sign, in other words the range between the inner stationary points. Results are shown in Figure 2.5.
Figure 2.3: a) In-phase and b) quadrature Lorentzian components of MTS, for the two-sideband approximation; black lines show the resultant MTS signals. Modulation frequency is 6.066 MHz, such that \( f_m/\gamma = 1 \).

Figure 2.4: a) In-phase and b) quadrature MTS lineshapes at various modulation frequencies; \( \delta = 0.6 \). The modulation frequency runs from 500 kHz to 10 MHz in 500 kHz steps; illustrative signals are coloured. Note the different amplitude scales.

The simulated amplitude and slope were compared with Figure 1 in Reference 28; the results are in excellent agreement\(^3\). From the plots it is clear that only a small subset of MTS signals are useful for laser locking; any signal with a ‘kink’ such as Figure 2.5h2 is unusable. Signals without a true ‘kink’ but with a reduction in slope near their centre such as such as Figure 2.5h6 fall into a grey area; they are usable for locking, but the lock performance would be poor because of the low slope. Figure 2.5b clearly shows the peak slope at each phase; the slope reaches a maximum and then decreases due to the appearance of kinks. Note that the zero crossings in the slope are very close in frequency to the discontinuities in Figure 2.5c; the discontinuities reflect the emergence of the kinks, such as in Figure 2.5d3, when the algorithm detects new stationary points in the MTS signal and shows a sudden reduction in the recapture range.

We see that a tradeoff must be made between slope and recapture range when choosing a modulation frequency and phase. Amplitude and slope can be made relatively large, but this comes at the expense of recapture range; if the recapture range is increased, it

\(^3\)As far as I am aware, no recapture range simulation has been reported in the literature.
2.3. SIMULATION OF MTS SIGNALS

Figure 2.5: Simulations of MTS a) amplitude, b) slope and c) recapture range. Example signals are shown to scale in d), over a horizontal span of 40 MHz.

Inevitably leads to kinks and thus the slope is reduced. A completed laser lock system is required to decide which is more valuable; this stage was not reached in the project, and emphasis was placed on slope and amplitude, neglecting recapture range. This is typified by signals such as Figure 2.5d4, which arise at modulation frequencies between 3 and 5 MHz, and phases between 50 and 80°.

We have considered an idealised model of MTS, which predicts the expected MTS signals for various modulation frequencies and phases. Other interesting trends in the simulations are omitted for brevity, along with the properties of MTS at large modulation indices.

Residual amplitude modulation

So far we have only considered MTS produced with an ideal FM spectrum. This is seldom the case, however, as the FM sidebands are generally attenuated by different amounts due to both electronic and optical effects.

In (2.13), \( J_1(\delta) \) scales the amplitude of one pair of Lorentzians in both the in-phase and quadrature components, while \( J_{-1}(\delta) \) scales the second pair. If these coefficients are altered such that their amplitudes are no longer equal, the Lorentzians lose their odd symmetry around the zero crossing point. The MTS signal becomes skewed, as shown in Figure 2.6.

Such an asymmetry arises when the the pump beam encounters amplitude modulation (AM) as a function of its instantaneous frequency. In the two-sideband approximation, the spectrum of AM is similar to FM, but the AM sidebands are in phase rather than out of phase. When added to the FM sidebands, this results in an amplitude increase for one
sideband, and a similar decrease in the other. This result extends to multiple-sideband FM and AM spectra [20].

This process is an essential part of all modulated spectroscopic techniques, in which the amplitude modulation gives information on the laser detuning from resonance, but it is undesirable when caused by anything other than the vapour cell; thus we wish to minimise residual amplitude modulation (RAM) in the pump beam as much as possible. This is not always possible, but if an AM signal is added to the combined FM + AM spectrum that is opposite in phase to the existing RAM, the RAM can be cancelled out. A rigorous analysis of RAM in electro-optic modulators is carried out in [33], much of which applies to this project, and we discuss its minimisation and compensation in §6.2. RAM is a troublesome source of distortion in both FM and MTS signals [21, 34], and a significant body of literature exists on detecting and reducing its effects [35–37]. Strategies have been published to minimise RAM in FM [38] and MTS systems [37] with metrology rather than laser locking in mind; the techniques discussed in [37] are nonetheless directly applicable in this project. Care was taken to minimise sideband asymmetry, with reduction of RAM in mind.

Summary

Over the last two chapters we have discussed several spectroscopic locking techniques, each of which has advantages and disadvantages. These are summarised in Table 2.1 though this cannot adequately cover the multitude of variations of each technique.

MTS has three key advantages over other techniques. Due to modulation transfer only taking place when the laser frequency is near an atomic transition, the error signal is virtually zero far from transitions, unlike an FM signal. The nonlinear four-wave mixing process ensures that there is no residual component from linear absorption, as there is in FM. Finally, the $\chi^{(3)}$ susceptibility is large only near closed transitions [25], hence the error signal amplitude is highest for these; other techniques do not favour closed transitions.

A comparison between SAS, FM and MTS spectra made by other researchers is shown in Figure 2.7. The main weaknesses of MTS are complexity, a weak signal compared to FM, and distortion due to RAM [28], but the advantages of MTS heavily outweigh these caveats, which were overcome during the course of the project.

The first MTS experimental study was carried out in the early 1980s on iodine vapour...
2.3. SIMULATION OF MTS SIGNALS

<table>
<thead>
<tr>
<th>Technique</th>
<th>Stab.</th>
<th>Slope</th>
<th>Range</th>
<th>Backgr.</th>
<th>SNR</th>
<th>Tune.</th>
<th>Size</th>
<th>Compl.</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sat. Abs.</td>
<td>L</td>
<td>M</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L-M</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>DAVLL</td>
<td>M</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>M</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Sagnac</td>
<td>M-H</td>
<td>M</td>
<td>H</td>
<td>L-M</td>
<td>M</td>
<td>L-M</td>
<td>M-L</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>Polar.</td>
<td>M-H</td>
<td>H</td>
<td>L</td>
<td>L-M</td>
<td>M-H</td>
<td>L-M</td>
<td>M-H</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>FM</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>L-M</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>MTS</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
</tr>
</tbody>
</table>

Table 2.1: Comparison of alternative spectroscopic techniques. Stab.: frequency stability of lock point. Range: frequency recapture range. Backgr.: Doppler-broadened background. SNR: signal-to-noise ratio. Tune.: frequency tunability of the lock point. Compl.: complexity of method. \( \{L,M,H\} = \{\text{low, medium, high}\} \). Note that there are many variations on each technique, only the basic/original forms are compared above. The sidelocking technique is assumed in saturated absorption.

Figure 2.7: Spectra obtained using (a) MTS, (b) FM, (c) SAS. Note that while the MTS spectrum has smaller amplitude than the FM spectrum, there is zero background and peaks occur at closed transitions only. Adapted from Reference 28.

using a spectrometer similar to our own [30], in which the advantages of MTS over ordinary FM spectroscopy were noted. Since then, MTS has been applied to rubidium [39], iodine [37], caesium [34] and calcium [40] vapours, among others. Due to its long-term stability and low background, MTS has been proposed for use in achieving a new length standard [36].

Recently a laser was locked using the recommendations in [28], with several modifications to improve the MTS signal. An Allan deviation\(^4\) on the order of \(10^{-13}\) over 90 s was obtained [41]. In this Honour’s project a new spectrometer using an acousto-optic instead an electro-optic modulator was constructed; acousto-optic modulators (AOMs) are commonly used in existing laser locking setups, and offer several other advantages discussed in §4.1.

The remainder of this thesis summarises the experimental work and results of this project, interspersed with short theoretical sections. Optimisation of MTS signals was a particular focus, and is discussed in Chapter 6. Next we begin our discussion of the

\(^4\) Allan deviation is one measure of the frequency spread of a laser. It is defined as the standard deviation of differences between adjacent data samples.
spectrometer apparatus.
In previous chapters we have discussed the motivations for a MTS-based laser lock. This chapter briefly discusses the optical layout of the MTS system, and summarises the key optical elements.

### 3.1 External-cavity diode laser

The external-cavity diode laser (ECDL) was built at Monash University in 2008 [42] based on a Littrow layout [43]. The design consists of a 780 nm solid-state laser diode, a collimating lens, and an adjustable mount holding a diffraction grating and mirror, all mounted on an aluminium block. Two piezo-electric transducers (PZTs) are used for frequency control: a stack for sweeping and tuning over a range of several gigahertz, and a disc for more rapid tuning over a narrower range of several hundred megahertz. The frequency of a laser diode is highly dependent on temperature, and a servo loop was used to control a Peltier cooler, keeping the temperature within a narrow range.

A diode laser controller (Moglabs DLC-202) controlled the laser diode current, actuator voltages and temperature servo loop. It incorporated a control system for automatically adjusting the current and piezo actuators to lock the laser to a reference error signal. Some time was spent aligning the diffraction grating and tuning the laser; details are omitted. A diode current of 69-72 mA was used, with an optical output power of 50 mW. The beam was passed through a Faraday isolator (Electro-optic Technologies 04-780-00361) to avoid reflections entering the laser, and a 50% beamsplitter was used to divert 10 mW into the modulation transfer spectrometer. Thus, the laser source was set up for the spectrometer.

### 3.2 Spectrometer layout

After several design iterations, the layout in Figure 3.1 was constructed. The laser light was split into a pump and a probe beam by polarising beamsplitter PBS1, with half-wave plate HW1 used to adjust the ratio; the pump beam was created by modulating the transmitted component with the AOM double pass; this process and its evaluation is discussed in chapter 4. The pump beam was sent into PBS3, with a fraction being transmitted onwards to a beam expander composed of lenses L4 and L3 whose focal lengths were -50 mm and 180 mm respectively, widening the beam by a factor of 3.5. The component reflected by PBS3 was directed to the non-polarising beamsplitting cube BS2; the power ratio between the reflected and transmitted beams was adjusted using HW2.

Meanwhile, the component reflected by PBS1 was split again by BS1, a 50% beamsplitter, with the reflected portion passing into BS2. The transmitted portion was directed to another beam expander composed of L1 and L2, creating the probe beam, and the expanded pump and probe beams passed through one another inside the rubidium cell; modulation transfer from the pump to the probe beam occurred here. The resultant probe beam was narrowed by the other lens pair and completely reflected by PBS3 into the photodetector PD1 (Thorlabs PDA10A-EL). The pump beam power into the vapour cell was adjustable.
CHAPTER 3. OPTICAL SETUP

Figure 3.1: MTS spectrometer used in this project. A: apertures, AOM: acousto-optic modulator, BS: beamsplitters, HW: half-wave plates, L: lenses, M: mirrors, PBS: polarising beamsplitter cubes, PD: photodetectors, QW: quarter-wave plate. The main laser is split into a pump and a probe beam by PBS1; the pump beam is modulated by the AOM. The pump and probe beams are expanded to reduce their intensity, reducing power broadening \[28\], and sent through the rubidium vapour cell. Here modulation is transferred to the probe beam, which is measured at PD1. The two beams beat together at PD2 to provide the spectrum of the pump beam, independent of vapour cell interactions. Unimportant beam paths and optics are omitted for clarity.

from zero to 3 mW, and the probe beam from 200 µW to 5 mW (at 5 mW the laser was diverted entirely to the left by PBS1, leaving zero power for the pump beam).

The pump and probe beam components entering BS2 were carefully overlapped, then focused onto PD2 (Hamamatsu G4176-03). The two beams interfered and formed beat signals at several frequencies, from which the pump beam spectrum was obtained; this is discussed in §4.5.

Critical points were the correct alignment of the double-pass setup, the precise focusing of the beams onto PD1 and PD2, and the exact overlap of pump and probe beams within the vapour cell. The latter was critical for minimising RAM and obtaining dispersive MTS spectra, and is discussed in §6.2.
3.3 Photodetection

The photodetector PD1 was used to detect the probe beam. This required a constant gain from baseband to frequencies up to 10 MHz, to provide both MTS and saturated absorption signals. A typical photodiode produces a current proportional to the light incident upon its surface \[44\]. This current must be converted linearly to a voltage source (ideally with low source impedance) for use in electronics. A resistor is the simplest such conversion mechanism, however this forms an RC circuit with the photodiode capacitance that attenuates high signal frequencies. The resistor can be decreased for higher bandwidth, however this introduces thermal noise for low resistances.

A transimpedance amplifier is one solution for obtaining low-noise, high-bandwidth measurements, centred around a low-noise op-amp. It is a challenging design problem, however, involving compromises between amplifier and resistor noise, system bandwidth, and gain \[45\]. Significant time was spent on the design of a wideband transimpedance amplifier, based around a standalone photodiode, but it was decided that a commercial photodetector (Thorlabs PDA10A-EL) would be more time-efficient.

PD2 was used to measure AC beat signals (discussed in §4.5), a less demanding role that did not require DC amplification. A standard AC-coupled 30 dB rf amplifier (Mini-circuits ZFL-500LN) was used instead of a specialised transimpedance design. The photodetector (Hamamatsu G4176) had sufficient bandwidth even without deliberate biasing.

Next we will discuss the modulation of the pump beam, a key requirement in all modulated techniques.
In the previous chapters we established that modulated locking techniques are more stable and immune to noise than unmodulated techniques. To implement a modulated technique such as MTS, a precise and reliable means of frequency modulating a laser beam is critical. This chapter discusses the rationale for using an acousto-optic modulator, the modulating electronics used to drive it, and the optical techniques used to both implement and characterise the modulation.

4.1 Choice of modulation method

Electro-optic modulation

The simplest method of modulating a laser beam is an electro-optic modulator (EOM), which is a polished crystal of a material exhibiting the Pockels effect (such as LiNbO$_3$) between a pair of flat terminals. The refractive index of LiNbO$_3$ varies as a function of the electric field across it, thus a voltage applied across the terminals alters the refractive index of the EOM. The phase of a light wave of the form (2.1) passing through the crystal is modulated if a sinusoidal voltage is applied to the EOM; by tailoring the amplitude and frequency of the modulating voltage to induce a phase shift $\delta \sin(\omega_m t)$, a light wave of the form (2.3) can be created.

At modulation frequencies over a megahertz, several hundred volts must be applied across an EOM to induce sidebands with power over 10% of the carrier. An EOM crystal has a characteristic capacitance, and by connecting a suitable series inductor, a high-Q LC circuit is created that amplifies the modulating signal applied to the EOM by a factor of 10 or more [28,46]. This avoids both high-voltage amplifiers (such as the $\pm$200V Thorlabs HVA200, USD 82346 [47]) and undesired phase shift due to the EOM capacitance.

The inductor must be physically modified or replaced to alter the modulation frequency. This is especially inconvenient when the frequency dependence of a phenomenon such as MTS needs to be analysed in depth; for example, McCarron et al [28, Fig. 2] obtain MTS error signals at only six modulation frequencies, a lower resolution than desirable for a phenomenon as complex as MTS.

Acousto-optic modulation

An acousto-optic modulator (AOM) consists of a piezo-electric transducer (PZT) attached to a transparent crystal of a material such as TeO$_2$. The PZT changes its size as a function of voltage across it$^1$. The AOM is fed with an rf signal, which causes the PZT to vibrate at rf frequencies, creating travelling pressure waves of sound in the crystal. The refractive index at a point within the crystal varies as a function of the local pressure, hence the travelling waves modulate the refractive index along the crystal. This creates a moving diffraction grating whose spacing is equal to the sound wavelength; this Bragg diffracts a monochromatic laser beam and alters its direction and frequency [48].

$^1$Confusingly, some references use ‘PZT’ to refer to lead zirconate titanate, another piezo-electric material commonly used in AOMs.
4.1. CHOICE OF MODULATION METHOD

A quantised model sheds more light on AOM diffraction. The travelling vibrations are treated as phonons, which travel through the crystal perpendicular to the laser beam photons. A photon will sometimes ‘absorb’ a phonon, shifting its direction and frequency due to conservation of energy and momentum as shown in Figure 4.1. There are in fact many scenarios that occur, for example a photon can ‘absorb’ two phonons or ‘emit’ one, but an AOM is designed to maximise the rate of single-phonon ‘absorption’ by the laser beam photons within a narrow rf frequency range. In this case the photon frequency shift is precisely the rf frequency: $\nu_{c,\text{out}} = \nu_{c,\text{in}} + f_c$, where $\nu_{c,\text{in}}$ and $\nu_{c,\text{out}}$ are the input and output optical frequencies, and $f_c$ is the rf frequency. Altering the rf power varies the rate of phonons produced, affecting the probability of phonon-photon interactions and hence the intensity of the diffracted beam, while frequency adjustment varies the phonon energy, altering the frequency and deflected angle of the beam. Thus we can modulate the frequency and direction of the laser beam by frequency-modulating the rf signal into the AOM [49].

Differences between acousto-optic and electro-optic modulation

An AOM has several advantages over an EOM for our purposes. The primary benefit is the ease of tuning: unlike an RLC-fed EOM, the modulation can be easily varied by altering the modulation frequency and index of the rf signal up to 20 MHz, with no need for high-voltage amplifiers or tuned LC circuits.

Additionally an AOM offsets the frequency of the pump beam from the probe beam by the rf frequency $f_c$, which offsets the saturated-absorption spectrum by $f_c/2$. This is because the beams no longer mutually interact with the zero-velocity class of atoms, but with a class that travels towards the probe beam with a Doppler shift of $+f_c/2$; this class is Doppler-shifted by $-f_c/2$ from the pump beam (see §1.2). This spectrum is distinguishable from the unshifted spectrum created by undesirable same-beam saturation, and allows saturation due to stray reflections from the cell walls or other optics to be detected and minimised.

The frequency of a laser locked to a cooling transition $\nu_0$ (see §1.1) would also be shifted by $f_c/2$ ($f_c$ for a double pass, discussed in §4.4) from the lab frame frequency, which is useful if the main laser output is to be amplified and split into multiple beams; each beam can be shifted back to near the transition by a single AOM, but the exact frequencies are adjustable around the cooling transition by altering the rf into these AOMs. The is a commonly used layout, but usually a dedicated AOM is used to offset the pump beam; using an AOM for
both offset and modulation simplifies the apparatus. Previous Doppler-free FM and MTS experiments have used AOMs for offset but not modulation \cite{34,41}, using an AOM for both would simplify the experimental layout and save the cost of a modulating EOM ($5000).

AOMs also have disadvantages, however. Unlike an aligned EOM, an AOM deflects a beam by an angle roughly proportional to $f_c$, measured to be $2^\circ$ at $f_c = 80 \text{ MHz}$ for our AOM. To avoid spreading of the beam due to modulation, a double pass setup must be arranged; this is discussed in §4.4. A beam entering an AOM is never fully diffracted, and at least two beams are produced: the diffracted beam and the undiffracted remainder. The ratio of the two is known as the \textit{diffraction efficiency} of an AOM, and is nonlinearly dependent on the input RF power and frequency, the beam positioning, intensity, and incident angle, and temperature. For our AOM at $f_c = 80 \text{ MHz}$, an rf power$^2$ of 30 dBm, and beam diameter of 2 mm, the diffraction efficiency is 70%; a double pass reduces this to 50% at best. There was 10 mW of available light power during the course of the project, and a peak pump power on the order of 2 mW was obtained with the double pass setup, due to the rf power being lower than optimal and non-adjustable. AOM attenuation is a bigger drawback with setups where the available light power is 1 mW or less. The nonlinear dependence of diffraction efficiency on frequency is a drawback in itself, as it amplitude modulates the beam. This provided a method, however, of cancelling the RAM due to other components — §6.2 discusses RAM cancellation.

All of the above pitfalls can be overcome with careful alignment and input signal optimisation; the main drawback that has likely prevented more common use of AOMs in MTS is the difficulty of synthesising an FM signal at high enough modulation frequencies and deviations required for a low-noise and wide-bandwidth feedback system to be constructed. This challenge is discussed in the next section, and the AOM performance is discussed in §4.5.

\section*{4.2 Generation of frequency-modulated signal}

\subsection*{Frequency modulation of a voltage-controlled oscillator}

Using an FM-capable function generator is the obvious method of synthesising FM; however, the few models that are capable of modulation frequencies above 100 kHz cost thousands of dollars. The AOM used had a maximum diffraction efficiency at 80 MHz, hence a carrier frequency range of only 73 - 87 MHz was needed, assuming a maximum frequency deviation of 7 MHz (this assumption is justified later). A computer-controlled arbitrary waveform generator (Analog Devices AD9959) was an alternative, but difficulties with the computer interface prevented its use in this project.

A voltage-controlled oscillator (VCO, Minicircuits ZX95-78-S+ \cite{50}) provided an inexpensive alternative to the above options. A VCO generates a sinusoid with an instantaneous frequency roughly proportional to a controlling input voltage, plus an offset. If we ignore the slightly nonlinear response of the VCO \cite{50}, this is described by

$$f_i = f_{\text{min}} + \sigma V_c,$$

where $f_i$ is the instantaneous VCO frequency, $\sigma$ is the tuning sensitivity of the VCO (in MHz/V), $V_c$ is the control voltage, and $f_{\text{min}}$ is the base VCO frequency for zero control voltage. It was decided early in the project that a VCO would be used for FM synthesis by supplying a DC-offset sinusoidal voltage as the control signal:

$$V_c(t) = V_{\text{DC}} + V_p \cos(2\pi f_m t)$$  \hspace{1cm} (4.2)

\footnote{Units of dBm are defined with respect to one milliwatt: $\text{dBm} = 10 \log_{10} \frac{P_{\text{W}}}{1\text{mW}}$. Thus 1000 mW = 30 dBm.}
4.2. GENERATION OF FREQUENCY-MODULATED SIGNAL

A 10 MHz crystal oscillator (XO) is used as a stable reference by the phase-frequency detector (PFD) to precisely tune the frequency of the voltage-controlled oscillator (VCO) to 80 MHz. A modulating signal from a function generator (FG) is added to the PFD control signal using a bias tee, frequency-modulating the VCO output. 

Figure 4.2: a) Block diagram of the phase-locked loop (PLL) used to generate an 80 MHz FM signal for the AOM. A 10 MHz crystal oscillator (XO) is used as a stable reference by the phase-frequency detector (PFD) to precisely tune the frequency of the voltage-controlled oscillator (VCO) to 80 MHz. A modulating signal from a function generator (FG) is added to the PFD control signal using a bias tee, frequency-modulating the VCO output. b) Loop filter used in the PLL. c) Frequency response of the loop filter, note the 2.6 kHz shoulder.

where \( V_c(t) \) is the control signal, \( V_{DC} \) is the DC offset, \( V_p \) and \( f_m \) are the modulation amplitude and frequency. Now (4.1) becomes

\[
f_i(t) = \sigma V_{DC} + f_{\text{min}} + \sigma V_p \cos(2\pi f_m t).
\]

(4.3)

Compare this to (2.5). After equating terms between these equations,

\[
f_c = \sigma V_{DC} + f_{\text{min}} \quad \text{and} \quad \delta = \frac{\sigma V_p}{f_m},
\]

(4.4)

and it is clear that \( f_m \) is now the FM modulation frequency. Thus, by varying \( V_{DC}, V_p \) and \( f_m \), a wide range of FM signals may be synthesised.

The VCO chosen had a 3 dB modulation bandwidth\(^3\) of 5 MHz, to allow modulation at several megahertz. Universal values of \( f_{\text{min}} \) and \( \sigma \) were not obtained for the VCO, because the nonlinearity of its response reduced the accuracy of (4.1) at high or low control voltages; nonetheless, values measured in the linear region between 75 MHz and 85 MHz were \( f_{\text{min}} = 66.7(5) \) MHz and \( \sigma = 4.28(5) \) MHz/V. A second-order polynomial function of \( V_c \) was used instead of (4.1) to better model the nonlinear VCO response for §4.3, but the accuracy increase is small; it is left out for brevity.

Stable signal generation using a phase-locked loop

Even for a stable DC control voltage, the frequency of a VCO drifts due to factors such as temperature and power supply voltage [50]. To avoid drift and also allow direct, precise control of \( f_c \), a feedback system known as an integer phase-locked loop (PLL) was used (Analog Devices EVAL-ADF411XEBZ1 evaluation board [51]).

The components of an integer PLL are shown in Figure 4.2. The PFD chip contains digital counters denoted R and N. Let \( f_i \) be the reference signal frequency; \( f_c \), the VCO frequency; \( m_R \) and \( m_N \) the thresholds of the R and N counters. Each counter is incremented

\(^3\)Other desirable qualities were low frequency pulling and pushing, and low phase noise. A discussion of these is beyond the scope of this thesis.
in every period of its input signal, until it reaches a threshold and sends a pulse to the phase-frequency detector (PFD); thus the counters act as frequency dividers of their input signals. 

The PFD compares the relative timing of the counter pulses, and tunes the VCO frequency until the pulses are synchronised in frequency and phase ([52] §9.27, [53]). When the PLL is locked,

$$f_c = f_r \frac{m_N}{m_R},$$

(4.5)

and by adjusting the ratio $m_N/m_R$, a wide range of desired VCO frequencies can be obtained given merely a constant reference frequency. Suitable frequency-dependent gain and phase shift of the VCO control signal are required for the PLL to quickly attain this condition (‘lock’) without oscillating, which necessitates a loop filter.

On the PLL board, the reference is provided by a 10 MHz crystal oscillator, with a maximum frequency drift (due to temperature, aging and other effects) of below 30 Hz. The counters offer a programmable range of $1 \leq m_R \leq 16383$ and $1 \leq m_N \leq 8191$, allowing fine frequency control over a very broad range limited only by external factors. Owing to such flexibility, PLLs find application in a wide variety of electronics, such as mobile phones, where the transmitter/receiver frequencies must be rapidly and accurately altered in narrow steps. In this project, the stability rather than the flexibility of a PLL-locked signal was the most desirable attribute of PLL control, although adjusting the VCO frequency with ease and precision turned out to be a very useful benefit.

### PLL board setup

The PLL board was modified slightly to connect the standalone VCO. The layout is shown in Figure 4.2a. The PLL control output was passed through a surface-mount loop filter, whose output was connected to the ‘DC’ port of a bias tee (Mini-circuits JEBT-4R2G). This acted as a diplexer and added a high-frequency sinusoidal modulation signal to the low-frequency control signal from the ADF4001, producing a voltage of the form (4.2). The modulation signal was supplied to the ‘RF’ port of the bias tee by an inexpensive external function generator (Rigol DG1022).

A simple pole-zero loop filter was designed using the Analog Devices ADIsimPLL software, as shown in Figure 4.2a. A cutoff frequency of 2.6 kHz and a phase margin of 45° were chosen. The cutoff frequency choice determined the loop bandwidth of the PLL; noise and drift below it was efficiently cancelled by the loop, but the cutoff was chosen quite low so that the PLL would not ‘fight’ the modulation of the VCO for modulation frequencies below 500 kHz. The phase margin allowed for up to 45° of phase lag in the control signal at 2.6 kHz (potentially caused by the bias tee and other signal delays along the control path) before the servo loop would become unstable; too high a phase margin would have reduced the rolloff of the filter gain, while too low a value risks instability [54]. The filter frequency response is shown in Figure 4.2a; note the shoulder at 2.6 kHz.

In a more demanding application where improved noise cancellation was desired or the range between the cutoff frequency and the lowest unattenuated modulation frequency needed to be smaller, a higher-order loop filter could readily be designed using ADIsimPLL. This was unnecessary for our purposes.

The loop filter was implemented on the board, completing the PLL servo loop. The board internally passes the VCO output through an impedance matching network to the ADF4001 and an rf output; this output was connected to a 30 dB power amplifier (Delta RF LA2-1-525-30). The amplifier output was passed through a 20 dB coupler (Mini-circuits ZFDC-20-5-S+), which provided a sufficiently attenuated signal at its coupling port for
4.3. CHARACTERISATION OF FREQUENCY-MODULATED SIGNAL

viewing through a spectrum analyser (Anritsu MS2721A), while the bulk of the signal was sent directly to the AOM. Next we discuss the range of FM signals obtained with our electronic setup.

4.3 Characterisation of frequency-modulated signal

Measurements in this section and § were taken using the spectrum analyser, which was adjusted to obtain consistent results across all modulation frequencies and deviations. Key parameters were the resolution bandwidth (RBW), video bandwidth (VBW) and detection type: the RBW determines the width of the frequency ‘bins’ used to sample the spectrum, the VBW determines the level of smoothing applied the collected data, and the detection type determines what data processing is applied to the spectrum. Reference gives a classic summary of spectrum analyser use.

Referring to (4.5), our setup required a VCO frequency $f_c$ of 80 MHz, while the reference frequency $f_r$ was fixed at 10 MHz. The division ratios $m_N = 80$ and $m_R = 10$ were used, leading to a counter pulse frequency of 1 MHz. The spectral peak had a 3dB width of 100 Hz, with a frequency offset by 2 kHz. The offset varied, but remained within 2 kHz; it was most likely due to the uncalibrated difference between the reference crystals of the spectrum analyser and the PLL. Thus, the PLL performed very well at stabilising the frequency and narrowing the linewidth of the VCO.

A range of counter pulse frequencies was tried, but it was found that below around 200 kHz ($m_N = 400, m_R = 50$), noise sidebands appeared in the peak that were offset from the carrier by the pulse frequency (these are known as leakage spurs [56]). Such sidebands indicate distortion of the rf produced by the VCO, thus an effort was made to minimise them. Their power relative to the carrier decreased from -10 dB at 5 kHz to -80 dB at 200 kHz.

Next, the unmodulated VCO power output was measured from 66 to 88 MHz using the spectrum analyser to obtain a ‘power envelope’, an important characteristic of the VCO. Modulation was then applied at a frequency of 1 MHz and index of 0.6, and the upper and lower sideband power was measured for carrier frequencies between 70 to 84 MHz. At each carrier frequency, the envelope power 1 MHz below was subtracted from the corresponding power 1 MHz above, and this was compared with the difference between sideband powers at that frequency.

Finally, carrier and first-order sideband power were measured at 80 MHz across a broad range of modulation frequencies and indices (though widely-spaced, due to time constraints). The sideband-to-carrier ratios were predicted using

$$\frac{P_1}{P_0} = \left( \frac{J_1(\delta)}{J_0(\delta)} \right)^2 \tag{4.6}$$

which is derived from (2.3) (large modulation indices were used, hence (2.4) was inapplicable). Figure 4.4a shows the measured deviations from the theory. The independent variables were the modulation frequency and amplitude (i.e. $f_m$ and $V_p$ in (4.2)). From these, the modulation index $\delta$ was determined using a polynomial version of (4.4)\textsuperscript{4}.

The results deviated from theory by less than 0.3 dB at modulation frequencies below 5 MHz, and the setup performed very reliably throughout the project. Delays and nonidealities in the electronics attenuated the modulation at high modulation frequencies, weakening the sidebands with increasing modulation frequency. Below 5 MHz the oppo-

\textsuperscript{4}The parameters were obtained using regression from a plot of VCO frequency versus control voltage; the details are tangential to the discussion.
site was observed, which was unexpected: between 1 and 3 MHz the sidebands were more powerful than predicted!

It is possible that a nonlinearity in the VCO electronics, perhaps deliberate, strengthened the modulating signal in the range 1 - 5 MHz. This is suggested in the lines for $\delta = 0.3$ and higher. More data are required to precisely characterise the VCO response, but it is clear that despite slight nonlinearity, the 3 dB modulation bandwidth is significantly above the data-sheet value of 5 MHz [50].

The differences between the lower and upper sideband power in Figure 4.4 are a measure of the residual amplitude modulation (RAM) in the FM signal, as mentioned in §2.3. Their variation is due to the nonlinearity of the VCO envelope in Figure 4.3b; as the modulation frequency increases, the sidebands become more powerful than predicted. Figure 4.4 shows the deviation of average sideband power from predictions, for various modulation frequencies and indices. Figure 4.3: a) Spectrum of amplified VCO output at a modulation frequency of 1 MHz and modulation index of 0.6; the sidebands are individually attenuated by the VCO power envelope. b) Power output of VCO as a function of the carrier frequency with no modulation; this is the power envelope. c) Differences in sideband power (upper minus lower) as a function of the carrier frequency, at a modulation frequency of 1 MHz and modulation index of 0.6, with differences in neighbouring power envelope values overlaid.

Figure 4.4: a) Deviation of average sideband power from predictions, for various modulation frequencies and indices. b) Differences in sideband power at 80 MHz, for various modulation frequencies and indices.
4.4. AOM DOUBLE PASS

frequency increases and the sidebands move further from the carrier, the envelope attenuates
them unequally (due to the envelope asymmetry). This envelope asymmetry is the dotted
line in Figure 4.4b, and the trend of the sideband asymmetry agrees well with it.

Because the carrier frequency of the system could be freely tuned, it was straightforward
to tune out the RAM. Sideband difference as a function of carrier frequency is shown in
Figure 4.3, alongside the difference in envelope power at 1 MHz above and below the
carrier; clearly the two lines are strongly correlated, giving credence to the interpretation
above. Thus the RAM due to the electronics was minimised simply by tuning the carrier
frequency.

In summary, the VCO/PLL system was shown to be an accurate and stable FM source,
and RAM was minimised by tuning the carrier frequency. Even uncompensated, below
5 MHz the sideband power and unevenness were within 1 dB of predictions. The limiting
factor was not the VCO bandwidth, but the maximum allowable modulation amplitude — to attain a sideband-to-carrier power ratio of 10% (i.e. a modulation index of 0.6) at
10 MHz, a 1.5 V modulation voltage was required. If the DC voltage was over 3.5 V, to
balance the sidebands for instance, the added voltage would be over the maximum control
voltage of 5 V, saturating the electronics or worse. Thus a modulation amplitude of 1 V was
never exceeded throughout the project, corresponding to a maximum frequency deviation
(see (2.6)) of 4.28(2) MHz. For a 10% sideband-to-carrier ratio, this limited the maximum
modulation frequency to 7 MHz. It will be seen that this was not the chief constraint in
the system.

With the creation of a pure and well-characterised source of FM, a major challenge in
setting up AOM-based MTS was overcome. Next we will discuss the optical layout of the
AOM double pass, and the characterisation thereof.

4.4 AOM double pass

For this discussion we assume that the rf signal is of the form (2.3). A beam entering the
AOM is diffracted at an angle proportional to the instantaneous frequency of the rf supply,
as discussed in §4.1. For FM at a frequency deviation of 3 MHz (recall (2.6)) and a carrier
frequency of 80 MHz, the instantaneous frequency dithers between 77 and 83 MHz. Thus,
assuming the deflection is proportional to rf frequency and that an rf frequency of 80 MHz
causes a 2° deflection (obtained experimentally), the beam diverges over 2.5 mrad. The
diffracted beam frequency varies with deflection angle — as the rf instantaneous frequency
oscillates, the beam frequency and angle oscillate in turn. If we were to sample the diffracted
beam frequency at various points, we would find a ‘rainbow’, as shown in Figure 4.5a.

For an infinitely narrow incoming beam, each point in the diffracted beam would be
at a constant frequency due to the ‘rainbow’ effect, with no actual modulation occurring.
The incoming beam has a finite width, however, thus a point in the diffracted beam does
encounter a spread of frequencies, but the spread (deviation) is limited by the incoming
beam width and divergence. For a constant sideband power, higher modulation frequencies
require higher deviations; the beam divergence is proportional to the deviation, thus the
attainable deviation becomes limited by the physical parameters of the AOM and the
beam. This has limited the maximum modulation frequencies used in AOM-based MTS
spectrometers [37, 39].

This effect is cancelled by using the double pass layout shown in Figure 4.5c. The
diffracted beam is reflected back into the AOM using spherical mirror M5; a component
is diffracted twice and exits the AOM along the path on which it originally entered. Its

5Reference 37 describes this effect as ‘angular separation of sidebands exiting the AOM’. I find this
misleading, as the process is more easily understood and analysed in the time domain.
Figure 4.5: a) Angular spread of diffracted beam for an FM signal with a frequency deviation of 3 MHz. b) Infrared photographs of the diffracted beams after the second AOM pass, as the rf frequency is varied. Aperture A1 has been opened to show all orders, and darkness indicates intensity. c) Principle of the AOM double pass setup, with the same labelling as in Figure 3.1; a component of the input beam is diffracted twice, while other components are blocked. Aperture A1 is shown close to the AOM than in Figure 3.1 for clarity.

The double pass has a second effect on the beam, which can be treated as if it were diffracted by a pair of identical AOMs being supplied the same instantaneous rf frequency. The beam is thus frequency shifted twice: if the AOM is fed an instantaneous frequency \( f_i \), the optical frequency shift due to the double pass is \( \Delta \nu = 2f_i \). From (2.5) and (2.6) we know that the instantaneous rf frequency dithers in the range \( f_i = [f_c - f_D, f_c + f_D] \); thus the optical frequency shift falls in the range \( \Delta \nu = [2f_c - 2f_D, 2f_c + 2f_D] \). The beam carrier frequency is shifted by \( \Delta \nu = 2f_c \), and its frequency deviation is \( \nu_D = 2f_D \); the optical modulation frequency \( \nu_m \) remains equal to the FM modulation frequency \( f_m \), thus the modulation index \( \delta \) is doubled.

Thus after the double pass, the frequency-modulated light signal is

\[
E_2(t) = E_0 \sin \left[ 2\pi (\nu_c + 2f_c) t + 2\delta \sin(2\pi f_m t) \right],
\]

\[
= E_0 \sum_{n=-\infty}^{+\infty} J_n(2\delta) \sin \left\{ 2\pi [\nu_c + 2f_c + nf_m] t \right\}.
\]

Note that the sideband-to-carrier power ratio changes from the FM signal driving the AOM; this is because the modulation index is doubled. If we wish to obtain a modulation index of 0.6 in the pump beam, corresponding to a sideband-to-carrier ratio of -10 dB, we must supply rf with an index of 0.3 (sideband-to-carrier ratio of -16 dB).
4.5. BEAT SIGNALS

The AOM envelope was measured several times using a power meter, with results shown in Figure 4.6. The AOM could be deliberately misaligned using mirror M4 (see Figure 3.1) to alter the entry angle of the beam; this shifted the frequency of maximum diffraction efficiency, and allowed the AOM ‘envelope’ to be adjusted. This was used later for minimising RAM. The envelope peak could not be shifted down below 75 MHz without losing significant power, though it could be shifted up as high as 85 MHz before a similar loss was seen.

A deficiency of using a single spherical mirror is that after being focused at the centre of the AOM, the double-passed beam diverges, in contrast to a lens arrangement as used in [57](note that this divergence is more benign than that shown in Figure 4.5), as the spectrum is much more uniform across the beam). This was largely solved by adjusting the spacing between expansion lenses C3 and C4 to compensate for the divergence; this in turn led to a converging probe beam, which was not a problem since a separate lens (not shown in Figure 3.1) was used to focus the probe beam onto the photodetector surface. A more serious deficiency discovered late in the project was that the AOM modulation is optimal for beams narrower than those used; the effects of a lens arrangement to narrow the beam [57] will be tested after the conclusion of this project.

4.5 Beat signals

The experimenter was confident in the theory behind AOM-based modulation, but was less sure of the implementation. A method was needed to characterise the modulation of the pump beam; one common technique is a scanning étalon, consisting of an optical cavity similar to that used in Pound-Drever-Hall locking, but with a tunable length. The spectral resolution of the étalon used was too poor to detect sidebands at less than 10 MHz spacing, however. Instead the pump beam was beat with the unmodulated probe beam at PD2. We briefly discuss the theory before summarising the results.
Theoretical background

At a constant point in space, the electric field of an unmodulated coherent light beam oscillates according to (2.1),
\[ E_A(t) = A \cos(\omega_A t), \] (4.9)
where \( A \) is the electric field amplitude, \( \omega_A \) is the light frequency. Let us assume another beam \( E_B(t) \) with amplitude \( B \), frequency \( \omega_B \) and phase shift \( \phi \) overlaps the first. If the beams share the same polarisation, the resultant field \( E_s(t) \) is just a scalar sum of the two beam fields:
\[ E_s(t) = A \cos(\omega_A t) + B \cos(\omega_B t + \phi) \] (4.10)
The output current \( I_p \) of a photodetector is proportional to the light intensity falling upon it, hence proportional to the square of \( E_s(t) \). Thus, \( I_p = KE_s^2 \), where \( K \) is a factor we are not concerned with at present. Squaring (4.10) we obtain
\[ E_s^2(t) = A^2 \cos^2(\omega_A t) + B^2 \cos^2(\omega_B t + \phi) + AB \cos(\omega_A t) \cos(\omega_B t + \phi) \] (4.11)
Note that each term is a product of cosines, which allows us to apply the same identity used to obtain (2.9). We obtain
\[ E_s^2(t) = \frac{A^2}{2} + \frac{B^2}{2} + \frac{AB}{2} \cos[(\omega_B - \omega_A) t + \phi] \]
\[ + \frac{A^2}{2} \cos(2\omega_A t) + \frac{B^2}{2} \cos(2\omega_B t + 2\phi) + \frac{AB}{2} \cos[(\omega_B + \omega_A) t + 2\phi] \] (4.12)
The first two terms are the constant intensities of the two beams, while the terms in the second line oscillate at optical frequencies, undetectable with electronics. This leaves the third term, which oscillates at the difference in frequency of the beams, and whose amplitude is the product of the beam electric fields:
\[ V_B(t) = \frac{AB}{2} \cos[(\omega_B - \omega_A) t + \phi]. \] (4.13)
This term is the beat between the two beams. The reader may wish to compare the discussion with that in §2.1; here the photodetector acts as an ‘optical mixer’ due to its square-law response. When the difference in beam frequencies is on the order of gigahertz or lower, the beat is detectable using standard electronics.

Experimental verification

To test the principle, the pump and probe beams were focused on PD2 without any frequency modulation. The probe beam was at the laser frequency \( \nu_c \), while the pump was offset to \( \nu_c + 2f_c \) (by the double-pass AOM). A signal with frequency \( 2f_c \) was detected by connecting a spectrum analyser to PD2, exactly as predicted by (4.13). The power was -65(1) dBm, however, thus a 30-sample average was required to reduce the uncertainty to 0.2 dB.

The power envelope of the beat was obtained by sweeping the rf frequency, and a graph closely following Figure 4.6 was obtained, albeit with higher uncertainties. This demonstrated the linearity of the beat power as a function of pump beam power, as predicted by theory.

To test the FM in the pump beam, modulation was applied to the pump beam with a carrier frequency of 77 MHz (used to balance the rf sidebands), creating a signal of the form (4.8). The probe beam beat with each spectral band in the pump beam; the power of
4.5. BEAT SIGNALS

each beat was proportional to the power of the contributing band (since the probe power was equal for all the beats). Spectra were recorded at several modulation frequencies, all with a modulation index close to 0.6 except for the 10 MHz spectrum. To bypass the nonideal effects of the VCO/PLL, the modulation index was empirically calculated for each modulation frequency by measuring the sideband-to-carrier ratio of the rf spectrum, then numerically solving \((4.6)\) for the modulation index (recall that the two are related by Bessel functions). Table 4.1 shows the results; the departure of sideband power from predictions grows with increasing modulation frequency, but importantly the sideband power difference remains within 0.5 dB at and below 3 MHz. Uncertainties are large mainly due to the weakness of the beats.

<table>
<thead>
<tr>
<th>(f_m) (MHz)</th>
<th>0.1</th>
<th>0.5</th>
<th>3</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\delta)</td>
<td>0.61(2)</td>
<td>0.60(2)</td>
<td>0.60(2)</td>
<td>0.45(3)</td>
</tr>
<tr>
<td>Dep. (dB)</td>
<td>-0.3(2)</td>
<td>-0.4(2)</td>
<td>-4.2(2)</td>
<td>-13.1(3)</td>
</tr>
<tr>
<td>SB diff. (dB)</td>
<td>0.1(3)</td>
<td>0.0(3)</td>
<td>0.3(3)</td>
<td>1.6(4)</td>
</tr>
</tbody>
</table>

Table 4.1: Nonideality of beats as a function of modulation frequency. Dep.: departure of average sideband power from predictions. SB diff.: Power difference between upper and lower sidebands; a measure of RAM.

To further investigate the sideband attenuation, the rf deviation was varied at a constant modulation frequency of 3 MHz. The rf modulation indices were calculated as before. Figure 4.7 shows the results.

The \(\delta < 0.1\) points show that there was some attenuation due to the AOM envelope, and the remainder of the graph demonstrates a significant trend not present in Figure 4.4. The cause of this was unknown; possibilities include nonlinearity in the AOM or the double-pass setup, or some problem with the beat setup itself. Alignment and focusing of the beat setup was done repetitively and similar results were obtained, which suggests that the former was the culprit; the assumption was made that the beats accurately reflected the true spectrum of the pump beam. Data in Chapter 6 show that this assumption was probably inaccurate.

Sideband asymmetry at 3 MHz is also shown; the sidebands show attenuation by the AOM envelope just like the VCO envelope attenuated the VCO sidebands. By adjusting the mirror M4 and shifting the envelope peak, as shown in Figure 4.6, the sidebands were

Figure 4.7: a) Departure of average sideband power from predictions, and b) power difference between upper and lower sidebands, as a function of modulation index at a modulation frequency of 3 MHz.
equalised and RAM was minimised.

**Beating between the spectral bands of the pump beam**

Other than the pump-probe beats, the spectral bands in the pump beam also beat with one another. The carrier beats with the two first-order sidebands to generate two beats at the modulation frequency \( f_m \). For pure FM, the sidebands are phase shifted by \( \pi \) from one another \(^{[20]}\), thus the beats destructively interfere, resulting in zero power (recall that modulated techniques rely on the differential attenuation and phase shift of the sidebands due to a spectral feature to create a nonzero signal at the modulation frequency).

When the sidebands were balanced, however, this was not observed— the beat indeed grew weaker by a factor of 20 dB as the sidebands equalised, but did not fully disappear. A possible explanation is higher-order sidebands contributing to the power: any pair of neighbouring spectral bands beat with one another at the modulation frequency, and though their contributions are weak, they are enough to prevent the beat from disappearing.

In summary, the beats demonstrated the accuracy of the AOM modulation model at low modulation frequencies and indices, but as these were increased the strong nonlinearity in the AOM became apparent. The beats were used in the remainder of the experiment to accurately characterise the modulation index, however there were discrepancies between predicted and observed MTS parameters at modulation frequencies above 2 MHz that may have been caused by over-reliance on the beats to gauge the AOM modulation. This is discussed in §6.1.

### 4.6 Summary

We have discussed how modulation was applied to the pump beam, in preparation for obtaining MTS signals, and characterised the VCO and AOM nonidealities. The beats provided a direct way of measuring the modulation index and sideband asymmetry due to RAM, which was minimised by adjusting the rf frequency and/or AOM angle. The pump beam sidebands agreed well with theory at modulation frequencies below 500 kHz, but this worsened as the modulation frequency was raised; measurement of the beats allowed this to be accounted for. Next the electronic process of demodulation is summarised.
5 Demodulation

We have discussed the layout of the spectrometer, and the techniques used to modulate the pump beam. The modulated pump and unmodulated probe beams interact through the rubidium in the vapour cell, and modulation is transferred to the probe as discussed in §2.2. In this chapter we discuss the process of demodulating the photodetector output and recovering the MTS signal.

5.1 Electronic demodulation

After the vapour cell, the probe beam reaches the photodetector PD1. The probe beam equation is similar to \( (4.8) \), except the sideband powers are determined by the modulation transfer mediated by the vapour. The carrier at \( \nu_c \) beats with the sidebands at \( \nu_c \pm f_m \) at the photodetector, to create a signal of the form \( (2.11) \).

This oscillating signal is added onto the baseband saturated absorption profile. By amplifying the high frequencies of the combined signal, the MTS signal can be recovered; it is then multiplied by a phase-shifted version of the sinusoid used to modulate the VCO (see Figure 4.2). By adjusting this phase shift, it is possible to select purely the in-phase component, the quadrature component, or a mixture of both, for output as the final error signal; see §2.1 for a deeper discussion.

The completed modulation and demodulation setup is shown in Figure 5.1. The signal from PD1 was amplified by 60 dB using two rf amplifiers (ZFL-500LN+) whose minimum frequency was 100 kHz, avoiding any contamination from baseband and low-frequency noise. The dual-channel function generator used to modulate the VCO (Rigol DG1022) was set to output a phase-shifted version of the modulating signal from its second channel. This was mixed with the amplified photodiode signal using either a specialised phase detector (Mini-circuits ZRPD-1+) or an ordinary mixer (Mini-circuits ZLW-6+). The low -3 dB

![Figure 5.1: Simplified layout of the complete modulation and demodulation electronics.](image)
point of the mixer was 3 kHz rather than 1 MHz for the phase detector\(^1\), allowing the mixer to demodulate lower-frequency signals. Both will be referred to as phase detectors herein.

The phase detector is designed to operate in a nonlinear regime: when both of the input signals are above 0 dBm, the output signal is proportional to the phase difference between them, regardless of their power [58]. The photodetector signal power was kept below 0 dBm, to ensure that the output signal was proportional to both the amplitude and the phase of the MTS signal; the phase-shifted modulation signal was kept at 7 dBm to ensure the maximum phase detector amplitude without introducing higher-order nonlinearity. A phase detector has other sources of distortion such as offset voltage, but these were assumed to be negligible and the mixer and phase detector were not characterised aside from basic tests of functionality.

The phase detector output was amplified by 18 dB, and low-pass filtered with a cutoff at 20 kHz. A low-noise active filter was designed and built for this application; its circuit diagram is shown in Figure 5.2. A Sallen-Key topology was chosen for its simple characteristics, while a maximally smooth (Butterworth) transfer function was chosen to minimise distortion of signals below the cutoff frequency. The cutoff was chosen deliberately low to remove the frequency-doubled component of the phase detector output for modulation frequencies at and above 100 kHz. This made the active filter unusable for laser locking due to its high phase shift at and above 20 kHz, which would lower the bandwidth of a servo system in a similar way to the loop filter discussed in §4.2. As discussed shortly, this did not affect MTS data collection.

With the demodulation electronics in place, the MTS spectrometer was complete. Next the MTS signals were optimised for laser locking.

---

\(^1\) A lock-in amplifier was available, but it had a maximum -3 dB point of only 100 kHz.
6 Optimisation of modulation transfer spectroscopy

In the previous chapters, we have discussed the design and construction of an MTS spectrometer. After it was successfully tested, MTS signals were obtained under various conditions, and the effects of several parameters on MTS were investigated. In this chapter we discuss this process, and summarise the results obtained.

All the data except Figure 6.5 were taken in the $^{85}$Rb D$_2$ region, for the $F = 3 \rightarrow F' = \{2, 3, 4\}$ transitions and their crossovers (see §1.2). Figures 6.1 and 6.3 focus on the $F = 3 \rightarrow F' = 4$ closed transition in particular.

Data collection

A digital four-channel oscilloscope (Tektronix DPO 3014) was used as a simple data acquisition system. The outputs from photodiode PD1, the phase detector and the piezo voltage were simultaneously displayed. The frequency spacing between the peaks in the saturated absorption spectrum (from PD1) were known [8], and used to calibrate the horizontal axes of the MTS spectra.

The PZT stack inside the laser (see §3.1) was supplied with an external voltage between 5 and 75 Hz, which sinusoidally oscillated the laser frequency; the voltage levels and laser current were regularly adjusted to keep the laser frequency sweep in a constant region. As the frequency was increased, each ‘ripple’ due to acoustic noise covered a greater portion of the spectrum, and thus had less of a localised effect. Care was taken, however, to avoid distortion of the MTS signal due to the low-pass active filter. With this in mind, the sweep frequency was reduced to 35 Hz or lower for broadband sweeps.

With a reliable method of displaying and recording MTS spectra, data collection was begun. All spectra displayed below were obtained with 16-sample averaging and a 50-point moving average filter (each spectrum is 20 000 - 100 000 points), unless stated otherwise.

The mixer was trialled alongside the phase detector; despite having a minimum rating of 1 MHz, the phase detector output remained proportional to the mixer over the entire range. The mixer added a significant DC offset to the spectra near the zero crossings, especially for the quadrature component at low power, thus the data for this chapter were taken using the phase detector.

6.1 Modulation frequency

MTS for the $F = 3 \rightarrow F' = 4$ transition

Spectra were obtained of the $F = 3 \rightarrow F' = 4$ closed transition at modulation frequencies from 100 kHz to 5 MHz, at four phases per frequency. Beats were observed on the spectrum analyser, and the modulation index was adjusted at each frequency for a sideband-to-carrier ratio of -10(1) dB. Pump and probe powers were 150 and 200 µW. The peak-to-peak amplitude, recapture range and slope of the zero crossings were measured using the oscilloscope display. Uncertainties were obtained by observing the variation of the quantity
Figure 6.1: Measurements of a) amplitude, b) slope and c) recapture range, as a function of modulation frequency. Dotted lines show theoretical predictions (th. in the legend). In d), experimental MTS spectra (solid lines) are shown with the predicted spectra (dotted lines) for comparison. Note the different amplitude scales.

over several seconds; this method likely underestimated the uncertainties significantly, but due to time constraints more data were not collected. Figure 6.1 shows the results, alongside the predictions from \( (2.11) \).

Two experimental points need clarification. The recapture range was obtained between the ‘outer’ peaks of the MTS signal instead of the largest peaks; due to time constraints an automatic algorithm was not written to simulate this, and thus a discontinuity is predicted at 4.2 MHz that was not recorded. Between 4 and 5 MHz the MTS inner peaks did indeed ‘outgrow’ the outer peaks (see Figure 2.5d1 for an example of the spectrum when the peaks are equal), but this was not characterised in detail. Slopes were obtained by choosing two closely-spaced points near the zero crossing and dividing their voltage difference by their frequency difference; the slope is a derived quantity, thus its uncertainties are the largest of the three measurements.

The experimental recapture range is lower by a consistent amount for all phases. The cause of this is unknown, as \( (2.11) \) predicts that the recapture range is solely dependent on the modulation frequency and natural linewidth; variation in the former was ruled out by checking the spacing of the beat sidebands, while the latter was unlikely due to the low light intensities used (this is discussed further in §6.3). Despite the constant offset, the recapture range plot shows the best correspondence to theory; this is likely because the measurement was largely independent of the noise affecting amplitude and slope measurements.
The amplitude and slope predictions were uniformly scaled to match the experimental voltages (due to the unknown $C$ factor in (2.11)); due to time constraints a least-squares fit was not carried out, but the trend behaviour was clearly correct at modulation frequencies below 2.5 MHz. Despite the beats being used to adjust the sideband power, above 2.5 MHz the correspondence between predictions and measurements rapidly worsens, and no data were taken above 5 MHz due to the weak signals and high distortions.

The actual phase of the MTS spectra appeared to differ from predictions by $\Delta \phi \approx \pi \left( f_m - 3 \right)$. For instance, at a modulation frequency of 2.5 MHz and a displayed phase shift of 90°, the actual phase shift seemed to be 0° and the in-phase component was displayed instead of the quadrature. This effect was identified and corrected for during data collection. The likely cause was the propagation delay in the AOM crystal; this was identified afterwards. The speed of sound in TeO$_2$ (the AOM material) is 4202 m/s, and the distance in the crystal from the PZT to the laser beam centre is 1.5(5) mm. This causes a propagation delay of 350(120) ns, corresponding to a $\pi$ phase shift at 1.4(0.5) MHz, and a phase deviation of $\Delta \phi = \pi \left( f_m \right)$. The observed phase shift appeared very close to $\pi$/MHz, which was one of the reasons it was not noticed until data collection; this agrees with the lower predicted limit due to propagation delay. If the phase shift had been significantly more or less than $\pi$/MHz, the agreement with predictions of Figure 6.1 would have worsened significantly as the modulation frequency was increased.

The effect has no bearing on the results in later sections, as these were obtained at a constant modulation frequency of 3 MHz.

**MTS for the $F = 3 \rightarrow F' = \{2, 3, 4\}$ transitions**

Data were also taken at a constant modulation amplitude ($V_p$ in (4.1)) of 400 mV, keeping the frequency deviation nominally at 1.6(1) MHz. This led to modulation indices ($\delta$) of 16 at 100 kHz and 1.6 at 1 MHz, increasing the MTS amplitude (and SNR), and allowed low-noise comparison between MTS spectra for different transitions as a function of modulation frequency. According to simulations of (2.11) summed over sideband orders $-999 \leq n \leq 1000$, large modulation indices scale the MTS signal but do not affect its shape below 1 MHz. The scaling factor is highly nonlinear with modulation frequency, and its sign oscillates. This due to the Bessel function oscillation as a function of $\delta$; the effect was not characterised in depth.

The MTS spectra are shown in Figure 6.2. The 100 kHz in-phase spectrum is similar to the FM spectrum in Figure 2.7b, but the closed transition signal is smaller than the other features. The literature makes no mention of such an effect in rubidium to my knowledge; the large amplitudes of the other transitions are likely due to effects other than four-wave mixing, such as modulated hole burning [25]. Above 1 MHz, however, these effects are unable to follow the modulation and four-wave mixing dominates. This was reported for iodine at modulation frequencies below 100 kHz [30], but these results are new for rubidium. Several other unusual effects, such as variation between the zero crossing slopes at various modulation frequencies, were not characterised due to time constraints.

Several dispersive profiles are visible at detuned frequencies below -92 MHz for the 100 kHz and 500 kHz spectra; these correspond to weak hyperfine transitions that are difficult to distinguish in the saturated absorption spectrum.

In summary, data for the closed transition agreed with predictions below a modulation frequency of 2 MHz; above this there was a reduction in amplitude and slope, likely caused by experimental factors rather than the inherent physics. The MTS was further investigated at 3 MHz, the highest frequency at which the signal was still strong; the choice of modulation frequency is discussed further in §6.4. Next we discuss how distortion due to residual amplitude modulation (RAM) was minimised.
Figure 6.2: MTS spectra obtained at a variety of modulation frequencies, with the deviation nominally 1.6 MHz. Pump and probe powers are 140 and 200 µW. Dotted lines emphasise the zero crossings.

6.2 Minimising RAM distortion

As discussed in §2.2, RAM undesirably distorts MTS spectra. We have seen how RAM due to the VCO/PLL may be cancelled by adjusting the carrier frequency $f_c$ to tune out the sideband difference, as shown in Figure 4.3; similarly, RAM caused by the AOM can be minimised by altering the angle of the AOM mirror M4, shifting the AOM envelope, shown in Figure 4.6.

The largest source of RAM was the pump-probe beam overlap within the vapour cell, overshadowing RAM from both the VCO and AOM. For a pump beam with a spatially-uniform spectrum this would not have been a problem, as an incomplete overlap between the beams would merely have reduced the MTS amplitude. The shape of the signal also changed as the alignment was varied, however, demonstrating that the pump beam was not spectrally uniform. This is mentioned in the literature for single-pass AOMs [37], but a recent study of RAM noted this effect for an EOM-based spectrometer [33], which would not be expected to have a nonuniform pump beam like a single-pass AOM layout — this implies that the effect is more undisposed than I expected when characterising the spectrometer.

To ensure the MTS alignment was repeatable, I developed a routine for minimising the
RAM distortion. Initially the VCO was set to 77 MHz (the carrier frequency at which the VCO power output was maximised) and the modulation frequency was chosen and set. The AOM double pass was aligned for maximum diffraction efficiency using a power meter, and the modulation index was adjusted to the desired level using the beat spectrum as a guide. The pump and probe beams were aligned within the vapour cell by maximising the beam overlap before and after the cell at high beam powers to enhance visibility, then lowered to 100 and 200 µW respectively. Finally the beam overlap was adjusted to minimise distortion in the MTS signal, and the pump and probe powers were raised to the desired levels.

Any residual distortion thereafter (due to alteration of the modulation frequency or index) was cancelled by adjusting the AOM mirror M4. Figure 6.3 shows the MTS spectra at three different alignments of M4; ‘clockwise’ and ‘counterclockwise’ signify the direction of mirror rotation as viewed from Figure 3.1. Similar techniques have been used to optimise a single pass AOM-based MTS spectrometer [37], of which I was unaware until after much of the data had been collected. Adjusting the mirror altered the signal as expected, though the change in RAM was smaller for the clockwise rotation. Using only M4 minimally disturbed the optical alignment of the spectrometer, as the angle of the pump beam emerging from the double pass was independent of the mirror angle (this can be shown by geometric arguments). There was a slight reduction in the AOM diffraction efficiency as the angle was altered, thus the ‘clockwise’ and ‘counterclockwise’ spectra in Figure 6.3 have slightly smaller amplitudes than their ‘balanced’ counterparts. Note that after initial adjustment, the beam and AOM alignments were untouched for the duration of §6.1; this was to maintain consistency in the data, at the expense of some distortion in Figure 6.2.

Distortion caused by RAM changes as a function of phase, and a particular phase exists at which a signal is undistorted regardless of the RAM level [38]. Phases of 0° and 90° were used to alternately check the RAM distortion. A particularly convenient phase was 135°, at which both the in-phase and quadrature profiles contributed to the central ‘kink’ such that it was centred only at low RAM levels. This saved time during RAM cancellation; the signal is shown in Figure 6.3d.

Above 5 MHz, asymmetries entered the MTS signals that could not be completely cancelled using the above techniques. The signals were distorted in more complicated ways than Figure 6.3; it was suspected that nonuniform power and higher-order electromagnetic modes in the pump beam were partially responsible. Other researchers have removed similar effects by passing the pump beam through an optical fibre [41]; due to time constraints this was not carried out. Aperture A1 was instead narrowed to a diameter of 1 mm to function as a crude spatial filter, and the MTS signal did indeed grow more symmetric. The beat sideband power, however, fell dramatically as the aperture was narrowed; the increase in MTS amplitude implies an unexplained discrepancy between the beat spectrum and the spectrum within the vapour cell.

We have summarised the effects of RAM and other unexplained distortions on the MTS spectra, and outlined a protocol to reduce RAM distortion to a minimum. Next we investigate the effects of increasing beam power above the low levels used up to this point.

### 6.3 Beam power

Both the pump and probe beams had diameters of 10(5) mm within the vapour cell; intensities can be obtained from the powers listed below by multiplying by 8(5) × 10⁻⁷ m⁻², but due to the large uncertainty in beam diameter only beam power is discussed.

The pump beam power was adjusted by rotating half-waveplate HWP2, with a constant probe beam power of 200 µW. The effects are shown in Figure 6.4. Both the in-phase and quadrature spectra grow as the power is increased up to 1500 µW, but grow little beyond
CHAPTER 6. OPTIMISATION OF MODULATION TRANSFER SPECTROSCOPY

Figure 6.3: Distortion due to RAM in a 3 MHz MTS signal for three alignments of mirror M4 in Figure 3.1. Pump and probe powers are 140 and 200 µW; δ = 0.6. CW denotes clockwise rotation of the mirror, CCW = counterclockwise. The distortion is most visible at 135°, hence this phase was used to minimise RAM levels.

this. The MTS recapture range remains almost constant for the in-phase profile, and the central slope grows in proportion to the amplitude.

The in-phase profile remains dispersive and symmetric at all power levels, but distortion appears in the quadrature spectrum above 200 µW and grows stronger with increasing power. This distortion could be minimised by standard RAM techniques, however this added a corresponding distortion to the in-phase profile. The distortion was most likely due to a nonlinearity in the electronics or an unknown effect of the atomic physics, because it was directly dependent on power when no other parameters were changed. In (2.13), each sideband influences four Lorentzians, two absorptive and two dispersive; if one of these was influenced independently of the other by an unknown physical process, the standard RAM minimising techniques would fail in removing the effect — this was indeed observed.

Increasing the probe power caused a mirror effect, surprisingly lowering the quadrature centre instead of raising it. At any given pump power, the quadrature distortion was minimised at a certain probe power. Powers where the distortion was cancelled are listed in Table 6.1.

The complete absence of this distortion below 200 µW suggests an intensity threshold to some physical process. One possibility is power broadening of the spectral features. At large saturation intensities, the linewidths of the spectral features increase, broadening the saturated absorption peaks [5]; the intensity was reduced by using lenses to spread the beams over a larger area, but nonetheless some power broadening did occur, shown
6.3. BEAM POWER

Figure 6.4: Effect of pump power on a) in-phase and b) quadrature MTS spectra. The saturated absorption spectra are shown in c). The inset d) is a plot of the in-phase MTS amplitude as a function of pump power; the blue line is a guide to the eye. Probe power is 200 µW, δ = 0.6.

in Figure 6.4c: If the closed transition became sufficiently power broadened, its frequency range would begin to overlap one or more neighbouring transitions, and thus it would no longer be closed. This is speculative; exact intensity values must be obtained using a beam profiler for a more precise analysis.

According to [30], the MTS signal amplitude is proportional to both the pump and probe beam field amplitudes, thus the amplitude should be proportional to the square roots of the beam powers: \( S_{pp}(\Delta) = K \sqrt{P_s P_p} \), where \( S_{pp} \), \( P_s \) and \( P_p \) are the MTS peak-to-peak amplitude, pump power and probe power. This effect was investigated only for pump power, but good agreement was seen at power below 1 mW, above which the amplitude levelled out. The levelling was most likely caused by saturation of the transition due to high intensities, an effect not accounted for in [30]. Figure 6.4d shows the plot; a square-root fit for points below 1.5 mW (not shown) gives \( K = 1.11(2) \times 10^{-2} \) V/µW and an \( R^2 \) value of 0.99.

The SNR of the MTS spectrum fell considerably as the probe power was increased from the minimum (the half-waveplate HWP1 was low-order, thus the minimum was nonzero); this was a surprising result, as in most optical systems the SNR increases with power. This may however have been due to the electronics becoming saturated by the large MTS
Table 6.1: Pump and probe power settings at which there was no distortion in the MTS quadrature profile.

<table>
<thead>
<tr>
<th>Pump power (µW)</th>
<th>185(5)</th>
<th>318(5)</th>
<th>450(5)</th>
<th>605(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probe power (µW)</td>
<td>417(5)</td>
<td>881(5)</td>
<td>1235(5)</td>
<td>1405(5)</td>
</tr>
</tbody>
</table>

Figure 6.5: Unaveraged, unfiltered MTS and saturated absorption signals over the maximum scanning range of the laser. Modulation frequency is 3 MHz, δ = 0.6. Phase is 0°, RAM has been minimised and the pump and probe powers are 700 and 200 µW. MTS has thus been optimised for laser locking. The left Doppler-broadened profile is due to the 87Rb \( D_2 \) \( F = 2 \rightarrow F' = 3 \) transition; a laser will be locked to the closed transition therein.

6.4 Optimising MTS for laser locking

The final step in this project was to formulate recommendations for an optimal error signal, given the experimental conditions. As mentioned previously, desirable attributes are large amplitude, slope and recapture range. High modulation frequencies improve the SNR as there is less laser and electronic noise [59], however we have seen that there is a considerable reduction in amplitude as the modulation frequency is increased beyond 2 MHz, and distortion dominates the MTS spectrum above 5 MHz. Thus, 3 MHz was chosen, which was as high as possible without sacrificing signal strength.

A high pump power offered increased amplitude and slope, however there was little benefit in raising it above 1 mW. Tuning the phase allowed the amplitude and slope to be improved marginally, however the large distortion in the quadrature component for high pump powers negated this benefit. A phase of 0° was selected to allow high pump powers to be used without distortion of the MTS signal, with a pump power of between 500 and 1000 µW and a probe power of 200 µW.

An unmodified MTS spectrum for the above settings is shown in Figure 6.5. Visible ripple in the MTS spectrum remains, but this is due to electronic noise: there is no reason to suspect the MTS process itself. The SNR was not investigated in depth during this project, but a careful characterisation using a spectrum analyser will provide useful information,
and if necessary the photodiode design discussed in §3.3 will be completed — this should provide lower noise than the commercial photodiode used. High pump and probe powers can be used in this setup with hardly any broadening of the MTS signal, thanks to the lenses, and a large portion of the noise seen at high probe powers was probably electronic. Once the noise sources in the spectrometer are minimised, pump and probe powers at one of the settings in Table 6.1 in conjunction with a phase of between 50° and 80° will provide an optimal MTS signal.

Minimising the shift in the zero crossing from the closed transition frequency is also an important goal. The closed transition peak of the saturated absorption spectrum was used as a guide throughout this project, but the maximum precision obtainable with this method was ±2 MHz. Adjustment of RAM levels can be used to compensate small shifts at the expense of MTS distortion, as shown in Figure 6.3, and altering the AOM carrier frequency $f_c$ can shift the zero crossing over a range of ±5 MHz (though this alters RAM levels and requires realignment, and must be noted when shifting to other frequencies in a system similar to Figure 4.1). A combination of these techniques, in conjunction with an absolute frequency reference such as an optical comb, could be used to precisely tune the zero crossing to the desired point.

The DLC-202 laser controller (see §3.1) was used to lock the laser to an early MTS error signal, before the low-pass amplifier had been built. At input signal levels below -20 dBm the oscillatory components of the phase detector output were strongly weakened, negating the need for a low-pass filter; this was at the expense of considerably lower SNR. With current and PZT stack feedback the laser remained locked for over an hour, however the PZT disc was not used due to oscillation. Further investigation is needed, but the spectrometer can be used to lock a laser — how well is unknown, but the prospects are bright.

### 6.5 Summary

The effects of modulation frequency, RAM levels and beam power on the MTS spectrum have been investigated, with trends supporting the theory, and optimal values for locking have been selected. The advantages of MTS are clear from Figure 6.5: very low background, dispersive lineshapes with high slopes and high SNR, and large amplitudes only for closed transitions. Two unusual effects were observed that have not been reported in the literature; a strong quadrature distortion for high beam powers, and the unusual effects of modulation frequency on different transitions.
This thesis has demonstrated that an MTS spectrometer using acousto-optic modulation can be constructed with a bandwidth of 3 MHz, comparable to the bandwidth achieved by spectrometers using electro-optic modulation [28]. Of the various spectroscopic methods used for laser locking, modulation transfer spectroscopy offers several advantages for our application: zero background, a high signal-to-noise ratio, and large amplitudes over closed transitions. The theoretical properties of MTS were discussed and simulated, chiefly to investigate the effects of modulation frequency and phase on the error signal amplitude, slope and recapture range. The simulations agreed well with experimental data and published literature.

A new form of MTS spectrometer was set up, relying on an inexpensive acousto-optic modulator to frequency modulate the pump beam. A stable and flexible phase-locked loop was used to supply the AOM with a frequency-modulated signal. The frequency modulation of the pump beam was investigated by beating the pump and probe beams together, and the system matched predictions well up to a modulation frequency of 3 MHz — well above the bandwidth attained by previous AOM-based configurations. The limitations of the AOM were likely responsible for attenuating the modulation above this, rather than the electronics. A complete and well-justified protocol was established for removing RAM influence and distortion, extending similar work using a single-pass AOM layout [37].

In summary, it has been conclusively demonstrated that MTS using an AOM is a viable, inexpensive and convenient laser locking technique, with significant advantages over existing commonly-used methods.

Future work

Further characterisation of the spectrometer is required, in particular the SNR and the effects of narrowing the beam through the AOM. The laser will then be locked to the MTS error signal for the $^{87}\text{Rb}$ closed transition, and the lock will be evaluated, compared with existing alternatives such as FM spectroscopy, and optimised. An entirely digital implementation of the modulating and demodulating electronics may be built, further improving the SNR and flexibility of MTS. It would be satisfying and rewarding if the work of this Honours project will do a small part to encourage wider use of MTS for laser locking.

Acknowledgements

I would like to deeply thank my supervisor Dr. Lincoln Turner, not least for providing me with an empty laboratory and plenty of unopened boxes of equipment — no better environment can be imagined for a student interested in experimental physics! This project has been enjoyable and rewarding thanks chiefly to his unceasing enthusiasm, patience and insight. Thanks also to A/Profs David Paganin and Michael Morgan for illuminating discussion and invaluable advice, and all the professional staff that greatly helped me with this project’s experimental component, chiefly Antonio Benci and Stephen Downing.
References


REFERENCES


REFERENCES


