An atomic fountain clock based on metastable $^{3}\text{He}$

Tom Jeltes
Department of Physics and Astronomy
Vrije Universiteit
Amsterdam

April 9, 2002
Abstract

In atomic fountain clocks, atoms are cooled with lasers and captured in a magneto-optical trap. The trapped cloud of cold atoms is launched upwards and falls back due to gravity, passing twice through an rf magnetic field that is contained in a microwave cavity. In between the cavity passages, the atoms are subjected to a uniform magnetic field. Below the interaction region, the hyperfine state of the atoms is detected and an error signal is constructed to lock the frequency of the rf field that is coupled into the cavity to the atomic hyperfine transition.

The configuration and requirements regarding the various parts of a fountain clock based on metastable $^3$He were investigated. Furthermore, an extended cavity diode laser system at 1083 nm that can be used to cool and manipulate the metastable helium atoms was studied and locked to transitions in $^3$He using the Pound-Drever frequency modulation method.
## Contents

1 Introduction  
1.1 The concept of time .......................... 7  
1.2 Measuring time ................................ 8  
   1.2.1 Mechanical clocks .......................... 8  
   1.2.2 Quartz clocks ............................... 8  
   1.2.3 Atomic clocks ............................... 9  
   1.2.4 Applications of atomic clocks .......... 9  
1.3 This report ...................................... 10  

2 The $^3$He$^*$ atomic fountain clock  
2.1 Motivation for using $^3$He$^*$ in a clock .......... 11  
2.2 The clock configuration ........................... 12  
   2.2.1 Metastable helium beam source .............. 12  
   2.2.2 Zeeman slower ............................ 13  
   2.2.3 Magneto-optical trap ....................... 14  
   2.2.4 Moving molasses ........................... 15  
   2.2.5 Microwave cavity ........................... 15  
   2.2.6 Free flight region: C-field ................. 15  
   2.2.7 Detection region ............................ 16  
2.3 Magnetic dependence of the clock transition frequency ... 16  
2.4 Ramsey’s method of separated oscillatory fields ...... 17  
   2.4.1 Unperturbed Hamiltonian .................... 17  
   2.4.2 Perturbing rf magnetic field ............... 18  
   2.4.3 Time evolution in a rotating frame .......... 20  
   2.4.4 Evolution operators ......................... 21  
   2.4.5 $\pi/2$ pulse .............................. 21  
   2.4.6 Ramsey fringes ............................. 23  
2.5 Accuracy and stability ............................ 23  
   2.5.1 Accuracy ................................ 23  
   2.5.2 Stability ................................ 25
## CONTENTS

### 3 The microwave cavity
- 3.1 Introduction ........................................ 27
- 3.2 Boundary conditions ................................ 27
- 3.3 Circular cylindrical cavity ......................... 29
- 3.4 $\text{TE}_{011}$ field configuration ................. 31
- 3.5 Q-value and bandwidth ................................ 32
- 3.6 Cavity holes .......................................... 35
- 3.7 Coupling of fields into the cavity .................. 36
- 3.8 Amplitude of the magnetic field ................... 36

### 4 The C-field
- 4.1 Introduction .......................................... 39
- 4.2 Homogeneity requirements for high and low field .... 39
  - 4.2.1 Spatial homogeneity ............................. 40
  - 4.2.2 Temporal stability ................................ 41
  - 4.2.3 Magnetic shielding ............................... 41
- 4.3 The solenoid magnet .................................. 42
- 4.4 Inhomogeneities ...................................... 42
  - 4.4.1 State of the art .................................. 43
- 4.5 Calculation of the field .............................. 43
  - 4.5.1 Current density integration ..................... 43
  - 4.5.2 Single coil summation ............................ 47
- 4.6 Conclusions .......................................... 50

### 5 Stabilisation of an extended cavity diode laser at 1083 nm
- 5.1 Saturated absorption spectroscopy .................. 51
  - 5.1.1 Experimental setup ............................... 52
  - 5.1.2 Doppler broadening and Lamb dips ................ 53
  - 5.1.3 Frequency reference: Fabry-Perot interferometer .... 54
- 5.2 Laser frequency control ............................. 55
  - 5.2.1 Temperature control ............................. 55
  - 5.2.2 Scan control ..................................... 55
  - 5.2.3 Current control .................................. 56
  - 5.2.4 PID regulator .................................... 56
- 5.3 Pound-Drever detection .............................. 57
  - 5.3.1 Pound-Drever theory ............................. 58
  - 5.3.2 Measured spectra ................................ 61
- 5.4 Laser stability: a beatnote experiment ............... 62

### 6 Conclusions ............................................. 67
**CONTENTS**

<table>
<thead>
<tr>
<th>A</th>
<th>The quantum structure of $^3$He</th>
<th>69</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.1</td>
<td>Fine structure</td>
<td>69</td>
</tr>
<tr>
<td>A.2</td>
<td>Hyperfine structure</td>
<td>69</td>
</tr>
<tr>
<td>A.2.1</td>
<td>The hyperfine energy spectrum</td>
<td>70</td>
</tr>
<tr>
<td>B</td>
<td>Electrodynamics</td>
<td>73</td>
</tr>
<tr>
<td>B.1</td>
<td>Electromagnetic fields</td>
<td>73</td>
</tr>
<tr>
<td>B.2</td>
<td>Maxwell equations</td>
<td>74</td>
</tr>
<tr>
<td>B.3</td>
<td>Electromagnetic waves</td>
<td>74</td>
</tr>
<tr>
<td>C</td>
<td>Notches</td>
<td>77</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 The concept of time

Time is a weird physical quantity. Throughout history, many different views of time have existed and still the final word has not been said on this matter. In ancient Greece, long before the appearance of modern science, time was the domain of the famous philosophers. Plato, for example, stated that time is nothing more than the circular motion of the heavens [1]. Aristotle did not quite agree and said that time is not the motion itself, but the measure of the motion. In the seventeenth century Gottfried Leibniz argued that time did not exist independently of events (which was Isaac Newton’s opinion), but was just the ordering of events that are not simultaneous. Yet another view emerges from Einstein’s theory of relativity, in which time is treated almost on equal footing with the three spatial dimensions.

In quantum theory, like in any other physical theory, time plays an essential role. It is of course one of the parameters in the Schrödinger equation and can be seen as some kind of counterpart of another fundamental physical quantity: energy. For example, a quantum mechanical system that is in a superposition of two states oscillates with a characteristic period that is connected to the difference in energy between these two states. In general, repetition in time is connected to energy: this is most simply expressed in the formula for the energy of a photon: \( E = h\nu = h/T \). This connection between time and energy is used in atomic clocks. However, experts on the subject still do not agree about what the right conceptual picture of time would be, in spite of the well-defined character of time as a parameter in formulae that describe the physical world in great detail.
1.2 Measuring time

1.2.1 Mechanical clocks

Although it may be difficult to say sensible things about the right conceptual view of time, it was probably the first physical quantity that was measured by man and nowadays it is the one that can be measured with the greatest accuracy. Our natural environment provides time-measuring instruments of several kinds. The motion of the earth around the sun, the moon around the earth and the earth around its axis provided the basic structure of our time-keeping (years, months and days, respectively). Days were divided into smaller units using sundials and clocks based on the movement of water leaking from a basin with a small hole in it [2].

In the early 14th century the first weight-driven mechanical clocks appeared in Italian churches, followed by smaller spring-powered clocks at the beginning of the 16th century. In 1656, Christiaan Huygens developed the first pendulum clock, based on the natural period of oscillation of the pendulum. It had an error of less than one minute a day, which was unprecedented at the time. The same Dutch scientist invented the balance wheel and spring assembly, which was less accurate but suitable for watches. By 1721, George Graham reached an accuracy of one second a day with a pendulum clock by compensating for the changes in the length of the pendulum due to temperature variations and reducing the friction in the moving parts of the clock. Another interesting innovation was made by W.H. Shortt, who constructed a clock consisting of two pendulums. One is called the slave pendulum and drives the master pendulum as well as the hands of the clock. The master pendulum therefore does not have to fulfill mechanical tasks that can disturb its regularity.

1.2.2 Quartz clocks

The introduction of quartz clocks in the 1930s meant a giant leap forward in timekeeping. Quartz clocks are based on the piezoelectric properties of quartz crystals. When the crystal is compressed or extended along the crystalline axis, an electric potential gradient is generated and vice versa. When integrated in an electronic circuit, the crystal can be used to amplify electronic noise at the crystal eigenfrequency (typically 32 kHz for watches) [3]. The exact eigenfrequency of the quartz crystal depends on the actual shape of the crystal, which provides the main source of inaccuracy in quartz clocks. The typical accuracy of cheap quartz watches is at the present about ±4 minutes per year [4].
1.2 Measuring time

1.2.3 Atomic clocks

All types of clocks described above have something in common: they consist of two basic elements: a stable frequency source and a counting device that displays the time that passes, which is connected to that frequency source. In the case of atomic clocks the frequency source is split into two parts: the frequency source and a local oscillator. The oscillations of the local oscillator are counted and its frequency is locked to that of the frequency source, which is the frequency of the electromagnetic field that induces an atomic transition. In a sense the local oscillator connects the frequency source (atoms) to the counting device\(^1\).

It is easy to see that an increase of the counting rate of the clock amounts to dividing time into smaller segments, which allows more accurate determination of time intervals. And even when a counting rate is reached that is high enough for any practical application, one would like to create a situation in which each ‘tick’ of the clock proceeds from the counting of many cycles of the frequency source, thus increasing the stability of the clock. The ‘counting’ of cycles is done by electronic devices, which limits the range of useful clock frequencies since electronic equipment cannot handle frequencies above the microwave regime. This is why up until now magnetic dipole transitions are used as clock transitions. These transitions are induced by time-dependent magnetic fields with frequencies of several GHz. Recent developments show that it is possible to build an atomic clock based on an optical transition (ca. \(10^{14}\) Hz), by using a new frequency dividing scheme, based on optical frequency combs generated by mode-locked femtosecond lasers [5].

At the present, the best atomic clocks are so-called atomic fountain clocks and use the frequency of a hyperfine transition of cesium in the ground state (this is the frequency that defines the second as the time corresponding to 9,192,631,770 cycles of the electromagnetic field that is resonant with the transition). The NIST-F1, located in Boulder, Colorado, USA, for example, has an uncertainty of \(2 \times 10^{-15}\), which corresponds to less than a second in 20 million years [6]. Many atomic clocks together define Coordinated Universal Time (UTC), the official world time.

1.2.4 Applications of atomic clocks

Apart from defining the official world time to great accuracy, atomic clocks play an important role in several other fields. They are used for position determination systems, such as for example the Global Positioning System (GPS), that uses 24 satellites, all equipped with atomic frequency standards. The satellites transmit radio signals with a time stamp, such that the com-

\(^1\)Note that in the Shortt clock the role of the local oscillator is played by the slave pendulum. This clock is in that respect very similar to modern atomic clocks.
bination of four of those signals from different satellites by a receiver yields the position of the receiver. Since light travels a distance of 300 meters in a microsecond, the accuracy of the time stamp is critical for the accuracy of the position determination.

In communication systems frequency stability and time synchronisation are essential. Time synchronisation between transmitters and receivers is needed in order to identify strings of data. Atomic clocks are placed at the main nodes of communication networks to ensure the desired accuracy level of time-keeping. Power industry uses atomic clocks to determine the time of occurrence of faults in large networks and all over the world clocks and frequency generators are calibrated using atomic clocks, both for public and scientific use [7].

1.3 This report

In the following chapters, I will report on the research that I have done as a graduation project in the atomic physics group of the Vrije Universiteit in Amsterdam. The prospects of an atomic fountain clock that uses metastable $^3$He were investigated. In chapter 2, the experimental setup of the clock is described as well as the quantum physics that governs its operation. Chapter 3 focusses on the microwave cavity in which the microwave magnetic field that induces the hyperfine transition is generated. The high-homogeneity constant magnetic field that is needed in the helium clock is discussed in chapter 4. Finally, some experimental work performed with a 1083 nm diode laser system is described in chapter 5.
Chapter 2

The $^3\text{He}^*$ atomic fountain clock

2.1 Motivation for using $^3\text{He}^*$ in a clock

In this report an atomic fountain clock based on $^3\text{He}$ is described. The helium atoms that are used are in the metastable $2^3S_1$ state ($^3\text{He}^*$). A transition between the metastable state and a higher exited state is induced by light with a wavelength of 1083 nm. This closed transition is the cooling transition that is needed to cool the helium atoms. Light in this wavelength region can be generated by diode lasers. $^3\text{He}$ instead of the more common $^4\text{He}$ is used because the latter has no nuclear spin and therefore no hyperfine transitions that can be used as a frequency reference in a clock.

The fact that one uses the metastable species of helium provides the possibility of detection of the atoms with an efficiency of almost 100% by a microchannel plate detector, because of the internal energy of the metastables of about 20 eV.

The main advantage of $^3\text{He}^*$ has to do with the fact that it is a fermion. The accuracy of clocks based on bosonic atoms is limited by frequency shifts caused by so-called s-wave collisions between the bosons. For slow atoms the next term in the expansion of the scattering cross-section (the p-wave term) is much smaller. Identical fermions experience no s-wave collisions, because of the symmetry of the wavefunction [8].

Another reason to use helium has to do with the possible variation of the fine structure constant $\alpha$ with time [9]. The relativistic corrections of the hyperfine energy splitting are a function of $Z\alpha$ where $Z$ is the atomic number. Comparing clocks based on different atoms thus provides a way of studying the possible variations of the fine structure constant over time. Clocks based on Rb, Cs, H and Hg$^+$ are already available. Another clock based on an atom with a low atomic number (helium has $Z = 2$) would be a nice addition.
Figure 2.1: The fountain setup: it consists of three parts: 1) a metastable helium source, 2) a cooling section which is formed by a laser collimator, Zeeman slower and magneto-optical trap and 3) the fountain part, containing a microwave cavity and a very homogeneous constant magnetic field. (Figure taken from [14]).

2.2 The clock configuration

The operating principle of atomic fountain clocks is as follows: laser-cooled atoms are launched upwards and fall back due to gravity. The cloud of atoms passes twice through a microwave magnetic field. At the end of the trajectory, the hyperfine state of the atoms is detected. The frequency of the microwave field is locked to the frequency of the hyperfine transition that is probed. This frequency is fed to an readout device that uses the periodic signal for a time display.

2.2.1 Metastable helium beam source

The metastable \(^3\)He atoms are produced in a liquid nitrogen cooled dc discharge source. The atoms are transferred to the metastable \(^2\)\(^3\)S\(_1\) state (see figure 2.2) by a discharge current of a few mA, in a narrow tube that is called a nozzle. Only \(\sim 0.001\%\) of the atoms are transferred to the metastable state [10]. The metastable atoms are cooled by laser light; all other helium atoms are a source of unwanted background pressure.

The beam of metastable atoms enters the collimation section through the nozzle and a skimmer, which provides a first transverse velocity selection. Using focussed laser beams, the velocity component perpendicular to the
2.2.2 Zeeman slower

The average velocity of the metastable atoms in the beam is much too high to catch them in a magneto-optical trap\textsuperscript{1}. The atoms already have transverse velocities that are sufficiently low, because of the transverse cooling mentioned in the previous section. The longitudinal velocity can be decreased using the Zeeman slowing technique: a red-detuned\textsuperscript{2} laser beam travels in the direction opposite to the direction of the atoms; because of the Doppler shift, the atoms ‘see’ the light as resonant with the $^2\!\ ^3S_1$ to $^2\!\ ^3P_2$ transition at 1083 nm. The helium atoms absorb photons that have momenta directed along the laser beam and they emit photons in a random direction when falling back to the $^2\!\ ^3S_1$ state. This procedure generates a net force directed against the movement of the atoms.

When the atoms have absorbed enough photons to have been slowed down by a significant amount, the laser beam will become non-resonant for these atoms. The Zeeman slower is used to compensate for the decreasing Doppler shift experienced by the atoms. The magnetic field generated by the cone-shaped solenoid (see figure 2.1) shifts the resonance frequency of the atoms (the \textit{Zeeman shift}). The atomic resonance frequency decreases with the decreasing magnetic field, thereby compensating for the decreasing Doppler shift.

The maximum force exerted on the atoms in the Zeeman slower is limited by the amount of photons that are absorbed per unit time. This amount

\textsuperscript{1}The longitudinal velocity is about 1100 m/s when cooled by liquid nitrogen. The velocity of uncooled metastable helium atoms is about 2000 m/s.

\textsuperscript{2}This red-detuned laserlight has a frequency that is about 500 MHz below the atomic resonance frequency. The light is called blue-detuned if its frequency is higher than the resonance frequency.
depends on the laser intensity and the lifetime $\tau$ of the exited state of the cooling transition. Furthermore, the force depends on the momentum $p$ of the absorbed photons and therefore on the transition frequency $\nu$. The deceleration of the atoms also involves the mass of the atoms. At laser intensities for which the transition is saturated, the atoms spend half of the time in the upper state, so we have for the number of absorbed photons per unit time $N = \frac{1}{2\tau}$. The force exerted by the laser beam is then given by:

$$F = Np = \frac{1}{2\tau} \frac{h\nu}{c}$$

(2.1)

from which the deceleration $a$ of a mass $m$ is easily found to be:

$$a = \frac{h\nu}{2mc\tau}$$

(2.2)

The maximum deceleration can be calculated from equation (2.2) and is $4.7 \times 10^5 \text{ m/s}^2$ (!) [12]. In practice, about half of the maximum stopping force is chosen, thereby slowing down the atoms from 1100 m/s to 20 m/s in a distance of 2.4 m.

### 2.2.3 Magneto-optical trap

Once the atoms are slowed down to sufficiently small velocities they are caught in a magneto-optical trap (MOT). This device combines laser beams and a magnetic field to trap atoms both in momentum and real space. Six laser beams, arranged in counterpropagating couples that are orthogonal to each other (see figure 2.1), exert a velocity-dependent force on the atoms that is always directed against the direction of movement of the atoms (just like in the Zeeman slower). This effect is called optical molasses. The magnetic field, generated by a set of coils in anti-Helmholtz configuration yields a space-dependent force, directed towards the centre of the trap. The two counterpropagating beams of each couple are circularly polarised: one beam is $\sigma^-$-polarised and the other one is a beam of $\sigma^+$ light. Both are slightly red-detuned. Because the magnetic field shifts the energy of the magnetic sublevels of the atom\(^3\), and transitions to the different magnetic sublevels are polarisation-dependent\(^4\), either absorption from the $\sigma^+$ or the $\sigma^-$ beam is favoured. In this way the atoms can be pushed back towards the centre of the trap, even if they started out with velocities low enough not to be resonant with the red-detuned light, thus able to escape from the optical molasses.

\(^3\)In the $2^3S_1$ state of $^3\text{He}$ these are the $J = 1, m_J = -1, 0, 1$ levels that are split by the magnetic field with an amount proportional to $m_J$.

\(^4\) $\sigma^+$-polarised light only induces transitions with $\Delta m = +1$ and $\sigma^-$-polarised light only induces transitions with $\Delta m = -1$. 
2.2.4 Moving molasses

Once the metastable helium atoms are caught in the MOT, the actual operation of the clock can be started. The cold atoms are on average at rest in the ‘lab frame’. The temperature of the atoms in the MOT is about 1 mK \[13\]. Now the magnetic trapping field is switched off. The atoms are no longer trapped in real space, but compression in momentum space continues. When the minimal temperature of ca. 200 $\mu$K is reached, the frequency of the laser beam that enters the MOT-region from below is increased a little, and the frequency of the counterpropagating beam is decreased. The result is a cloud of atoms of very low temperature that gain a significant velocity in upward direction. Because of the huge accelerations that are induced by the laser light, the launching of the atoms takes place on a timescale that is much shorter than the time needed by the cloud to expand significantly by thermal diffusion. Since the definition of temperature is related to the width of the velocity distribution of the gas (and has nothing to do with ‘absolute’ velocities\(^5\), the temperature of the atoms does not change significantly, but the cloud of atoms is nevertheless launched upwards with a speed of about 3 m/s. This is called moving molasses. Before the cloud of atoms enters the interaction region, all atoms are optically pumped into the lower hyperfine state of the clock transition.

2.2.5 Microwave cavity

The internal structure of the $^3$He atoms is connected to the clock output via an rf magnetic field. This field is contained in a microwave cavity: a copper cylindrical cavity with a diameter and height of a few centimeters. The magnetic field in the cavity causes an excitation in the hyperfine spectrum of the helium atoms if it has the right frequency. Half of the atoms should be excited after one cavity passage for proper operation of the clock. This requires a certain amplitude of the magnetic field, in addition to the requirement that the field is resonant with the atomic transition\(^6\). The cavity operates in the so-called $TE_{011}$ mode.

2.2.6 Free flight region: C-field

The time that the atoms spend above the microwave cavity is crucial for the accuracy of the clock\(^7\). What is also very important is the fact that the microwave resonance frequency of the atoms (the clock frequency) should not differ between the atoms during the free flight. Since this resonance frequency depends on the magnetic fields present in the free flight region,

\(^{5}\)There is of course no such thing as an absolute velocity, it all depends on the frame of reference.

\(^{6}\)For more details on the microwave cavity see chapter 3.

\(^{7}\)The accuracy of the clock increases with the time between the two cavity passages.
high magnetic field homogeneity is required. This is why this region has to be shielded from magnetic influences from outside, such as the Earth’s magnetic field (see figure 2.1). The presence of a magnetic field is needed to break the degeneracy of the magnetic sublevels and for $^3$He the magnitude of this field has to be about 803 Gauss, in order to minimise the dependence of the clock frequency on the magnetic field. More about this constant field ($C$-field) can be read in chapter 4.

2.2.7 Detection region

After passing twice through the microwave cavity, the hyperfine state of the atoms that enter the region below the MOT is determined in the detection region below the MOT region. A state-selective mechanism is used to detect the atoms in one of the hyperfine states by a microchannel plate detector (MCP). The MCP signal is a measure of the number of atoms that end up in the upper hyperfine state, which depends on the frequency detuning of the microwave field with respect to the atomic resonance frequency (see section 2.4.6).

2.3 Magnetic dependence of the clock transition frequency

The clock frequency is locked to the transition between levels $|2\rangle$ and $|5\rangle$ of the $^2S_1$ state of $^3$He. The $|2\rangle$ and $|5\rangle$ states originate from $|F,m_F\rangle = |1/2,−1/2\rangle$ and $|3/2,−1/2\rangle$ respectively (see appendix A). The frequency of this hyperfine transition depends on the amplitude of the magnetic fields that are present, since the field interacts with the atoms. This Zeeman interaction is also described in appendix A.

The relation between the frequency of the transition between the clock states and the magnetic induction $B$ is given by\(^8\):

\[
\nu = E\left[1 - \frac{2}{3}(g_J - g_I)\mu_B B/E + ((g_J - g_I)\mu_B B/E)^2\right]^{1/2}
\]

(2.3)

in this equation $g_J$ and $g_I$ are the electronic angular momentum and nuclear spin $g$-factors, respectively. $\mu_B$ is the Bohr magneton and $E$ is the zero field hyperfine splitting energy corresponding to 6.7 GHz.

In most atomic frequency standards a hyperfine transition between two $m_F = 0$ levels is studied. The main motivation for this choice is that to first order the transition frequency is independent of the magnetic fields present (the small bias C-field and the rf transition-inducing field). But since $^3$He is a fermion (it has half-integer total angular momentum), it has no $m_F = 0$

\(^8\)If we express the magnetic induction $B$ in Gauss, we have $(g_J - g_I)\mu_B / E \approx 4.15 \times 10^{-4} \text{ G}^{-1}$.\
levels. However, if a C-field is applied such that the transition frequency is minimal, the first order dependence of the energy splitting on the magnetic induction vanishes. One can make a Taylor expansion around $B_0$ for the transition frequency:

$$\nu = \nu(B_0) + \frac{\partial \nu}{\partial B} |_{B_0} (B - B_0) + \frac{1}{2} \frac{\partial^2 \nu}{\partial B^2} |_{B_0} (B - B_0)^2 + ...$$

(2.4)

Taking the value for $B_0$ that corresponds to the minimum of $\nu$ ($B_0 \approx 802.6$ G), we can see that the second term vanishes and that the $B$-dependence of the energy splitting is determined by the third term, which is quadratic in $B$ and proportional to the second derivative of $\nu$ with respect to $B$. This value of the magnetic field is called the anticrossing field. Consulting equation (2.3) yields a transition frequency of 6.354251207 GHz. This is the configuration in which the helium fountain clock will operate.

2.4 Ramsey’s method of separated oscillatory fields

2.4.1 Unperturbed Hamiltonian

Let’s suppose that all helium atoms are prepared in the $2^3S_1$ state with $|F, m_F\rangle = |3/2, -1/2\rangle$ and all conditions are such that the only other state of importance is the $|1/2, -1/2\rangle$ state. Once the atoms are subjected to a strong magnetic field, the $|F, m_F\rangle$ labelling is no longer appropriate (see also appendix A). We will therefore say that all atoms are prepared in state $|5\rangle$ and the only other relevant level is $|2\rangle$. We will develop a theoretical description in which the effect of the rf magnetic field is treated as a perturbation.
of a system that is described by a Hamiltonian $H_0$, with known eigenvalues: $H_0 \ket{2} = E_2 \ket{2}$ and $H_0 \ket{5} = E_5 \ket{5}$. Using the definition $E_2 - E_5 = \hbar \omega_0$ and setting the zero of the energy scale to $\frac{1}{2}(E_2 + E_5)$, we get $H_0 \ket{2} = \frac{1}{2} \hbar \omega_0 \ket{2}$ and $H_0 \ket{5} = -\frac{1}{2} \hbar \omega_0 \ket{5}$. In the \{\ket{2}, \ket{5}\} representation in which $\ket{2}$ and $\ket{5}$ are represented by the vectors (1,0) and (0,1) respectively, one can write for $H_0$:

$$H_0 = \frac{\hbar \omega_0}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$  \tag{2.5}$$

### 2.4.2 Perturbing rf magnetic field

The probability of a transition caused by a magnetic field, oscillating at the resonance frequency, can be obtained using first order perturbation theory. One has to take into account that the quantum states in the presence of a magnetic field are not the $(F, m_F)$ eigenstates, nor eigenstates of the relevant quantum operators $\mathbf{S}$ and $\mathbf{I}$, but superpositions of those states.

The perturbing magnetic induction is directed along the $z$-direction and can be written:

$$\mathbf{B}_{rf} = B_{rf}(t) \hat{z} = B_{rf} \hat{z} \cos \omega t = \frac{1}{2} B_{rf} \hat{z} (e^{i\omega t} + e^{-i\omega t})$$  \tag{2.6}$$

Now the Hamiltonian for the interaction of the rf magnetic field with the total electronic and nuclear spins (the second term of equation (A.1)) can be expressed as:

$$H_{rf} = \mu_B (g_J J_z + g_I I_z) B_{rf}(t) \approx \mu_B g_J S_z B_{rf}(t)$$ \tag{2.7}$$

since $g_I \ll g_J$ and for the $2^3 S_1$ state $L = 0$. In order to do the calculation, it is sufficient to label the quantum states with the $m_S$-eigenvalues. We can express the states $\ket{2}$ and $\ket{5}$ as \[14\]:

$$\ket{2} = \frac{1}{N_2} \left\{ (3 - 2\sqrt{2}) \left[ \sqrt{\frac{7}{3}} \ket{0} + \sqrt{\frac{1}{3}} \ket{-1} \right] \right. + \left. \left[ \sqrt{\frac{1}{3}} \ket{0} - \sqrt{\frac{7}{3}} \ket{-1} \right] \right\}$$  \tag{2.8}$$

$$\ket{5} = \frac{1}{N_5} \left\{ (-3 - 2\sqrt{2}) \left[ \sqrt{\frac{7}{3}} \ket{0} + \sqrt{\frac{1}{3}} \ket{-1} \right] \right. + \left. \left[ \sqrt{\frac{1}{3}} \ket{0} - \sqrt{\frac{7}{3}} \ket{-1} \right] \right\}$$  \tag{2.9}$$

\[9\text{In the conventional atomic beam clocks, there usually is a phase difference between the magnetic fields of the two interaction regions. This results in a factor } e^{i\phi} \text{ in the expression for the field and eventually generates a complex Rabi frequency. Because in a fountain clock the atoms experience the same field twice, this phase difference is zero.}\]
where $N_2$ and $N_5$ are normalisation factors and are given by $N_2 = (18 - 12\sqrt{2})^{1/2}$ and $N_5 = (18 + 12\sqrt{2})^{1/2}$. The expressions in [] are the original $(F,m_F) = (3/2,-1/2)$ and $(1/2,-1/2)$ states respectively. Using these expressions to calculate the transition matrix element between these two states, we get:\n
$$\langle 5 | H_{\text{rf}} | 2 \rangle = \frac{13}{18} g_J \mu_B B_{\text{rf}}(t)$$ (2.11)\n
One might expect the rf magnetic field to cause a time-dependent change of the transition frequency due to the Zeeman effect. This appears not to be the case. In order to investigate this effect, the expressions $\langle 2 | H_{\text{rf}} | 2 \rangle$ and $\langle 5 | H_{\text{rf}} | 5 \rangle$ must be calculated. These expressions give a first order approximation of the change in energy of the two states, caused by the rf field, treated as a perturbation of the C-field Hamiltonian. Again, the states should be expressed in eigenstates of $S_z$ in order to do the calculation. Both expressions appear to be equal, namely $-\frac{1}{2} g_J \mu_B B_{\text{rf}}$. No net frequency change results. Similar calculations show that the nuclear spin part of the interaction energy is even zero for each level individually.

When applied to the set of states $\{ | 2 \rangle, | 5 \rangle \}$ $H_{\text{rf}}$ becomes:

$$H_{\text{rf}} = \begin{pmatrix} 0 & \frac{13}{18} g_J \mu_B B_{\text{rf}}(t) \\ \frac{13}{18} g_J \mu_B B_{\text{rf}}(t) & 0 \end{pmatrix}$$ (2.12)\n
The matrix above can be rewritten for practical purposes. The most subtle change is the fact that the time-dependence of the rf field can be expressed in the so-called rotating wave approximation. Since the elements of the Hamiltonian matrix that are not on the main diagonal will be used to calculate transition probabilities, only one of the two exponentials in equation (2.6) is relevant (because it yields a resonant term). The lower left element corresponds to the stimulated emission of a photon, the upper right one corresponds to absorption of a photon. When we also define the Rabi angular frequency $\omega_R = \frac{13}{9\hbar} \mu_B B_{\text{rf}}$, in which $B_{\text{rf}}$ is the amplitude of the oscillating magnetic induction, and furthermore use the fact that $g_J \approx 2$ we obtain:

$$H_{\text{rf}} = \begin{pmatrix} 0 & \frac{\hbar}{2} \omega_R e^{i\omega t} \\ \frac{\hbar}{2} \omega_R e^{-i\omega t} & 0 \end{pmatrix}$$ (2.13)\n
The total Hamiltonian thus becomes:

$$H_0 + H_{\text{rf}} = \begin{pmatrix} \frac{1}{2} \hbar \omega_0 & \frac{\hbar}{2} \omega_R e^{-i\omega t} \\ \frac{\hbar}{2} \omega_R e^{i\omega t} & -\frac{1}{2} \hbar \omega_0 \end{pmatrix}$$ (2.14)\n
\textsuperscript{10}The same calculation can be performed for the zero-field case, assuming pure $(F,m_F)$ states, yielding:

$$\langle 3/2, -1/2 | H_{\text{rf}} | 1/2, -1/2 \rangle = \frac{\sqrt{2}}{3} g_J \mu_B B_{\text{rf}}(t)$$ (2.10)
2.4.3 Time evolution in a rotating frame

In order to study the evolution of the wave functions under the influence of the Hamiltonian of equation (2.14) one can use the Schrödinger equation:

\[
i \hbar \frac{d}{dt} |\Psi\rangle = (H_0 + H_{\text{rf}}) |\Psi\rangle \quad (2.15)
\]

Solving the Schrödinger equation (2.15) would give us the time evolution of $|\Psi\rangle$. We can get rid of the time-dependence in the Hamiltonian if we make a transformation to a frame rotating with the frequency of the magnetic field $\omega$ [15].

\[
|\Psi\rangle \rightarrow R R^\dagger |\tilde{\Psi}\rangle \quad R = e^{i\omega\sigma_z t/2} \quad R^{-1} = R^\dagger \quad (2.16)
\]

The $\sigma_z$ in the exponential is the Pauli $z$ matrix. The exponential is a symbolic notation of the explicit expression given by the following equation:

\[
\exp(i \sigma \cdot A) = 1 \cos|A| + i \sigma \cdot \frac{A}{|A|} \sin|A| \quad (2.17)
\]

We now have:

\[
i \hbar \frac{d}{dt} (R R^\dagger |\tilde{\Psi}\rangle) = \frac{\hbar \omega}{2} \sigma_z R R^\dagger |\tilde{\Psi}\rangle + i \hbar R R^\dagger \frac{d}{dt} |\tilde{\Psi}\rangle = H R R^\dagger |\tilde{\Psi}\rangle
\]

\[
\Leftrightarrow i \hbar \frac{d}{dt} |\tilde{\Psi}\rangle = R HR^\dagger |\tilde{\Psi}\rangle - \frac{\hbar \omega}{2} \sigma_z |\tilde{\Psi}\rangle \equiv \tilde{H} |\tilde{\Psi}\rangle \quad (2.18)
\]

From equation (2.18) we can easily find the explicit expression of the Hamiltonian in the rotating frame $\tilde{H}$ in the \{2, 5\} representation:

\[
\tilde{H} = \frac{\hbar}{2} \begin{pmatrix} -\delta & \omega R \\ \omega R & -\delta \end{pmatrix} = -\frac{\hbar}{2} (\omega R \hat{x} + \delta \hat{z}) \cdot \sigma \quad (2.19)
\]

In equation (2.19), $\delta = \omega - \omega_0$ is the detuning of the applied rf-field with respect to the transition frequency and $\sigma$ is the vector of Pauli matrices. The Schrödinger equation becomes:

\[
\frac{d}{dt} |\tilde{\Psi}\rangle = i \frac{\hbar}{2} (\omega R \hat{x} + \delta \hat{z}) \cdot \sigma |\tilde{\Psi}\rangle \quad (2.20)
\]

This equation can be integrated directly with respect to time:

\[
|\tilde{\Psi}(t)\rangle = \exp \left( i \sigma \cdot \frac{(-\omega R \hat{x} + \delta \hat{z})}{2} (t - t_0) \right) |\tilde{\Psi}(t_0)\rangle \quad (2.21)
\]

Using equation (2.17), the time evolution of the coherent cloud of atoms can be written as follows:

\[
|\tilde{\Psi}(t)\rangle = \left\{ 1 \cos \left( \frac{\Omega}{2} t \right) + i \sigma \cdot \left( \frac{-\omega R \hat{x} + \delta \hat{z}}{\Omega} \right) \sin \left( \frac{\Omega}{2} t \right) \right\} |\tilde{\Psi}(t_0)\rangle \quad (2.22)
\]

in which $\Omega = \sqrt{\omega_R^2 + \delta^2}$ is called the Rabi flopping frequency.
2.4.4 Evolution operators

Equation (2.22) can be applied consecutively for each stage in the journey of the cloud of atoms (first cavity passage, free flight, second cavity passage). Each stage has its own *evolution operator* that can be derived from (2.22).

- First cavity passage \(0 < t < \tau\): \(U_1 = 1 \cos(\frac{\Omega}{2} \tau) + i \sigma \cdot \left(\frac{-\omega R \hat{x} + \delta}{\Omega} \right) \sin(\frac{\Omega}{2} \tau)\)
- Free flight \(\tau < t < \tau + T\): \(U_2 = 1 \cos(\frac{\delta}{2} T) + i \sigma \cdot \hat{z} \sin(\frac{\delta}{2} T)\)
- Second cavity passage \(\tau + T < t < 2\tau + T\): \(U_3 = U_1\)

The final probability that an atom of the cloud of atoms ends up in a final state \(|f\rangle\) is given by:

\[
|\langle f | \hat{\Psi}\rangle|^2 = |\langle f | U_3 U_2 U_1 \hat{\Psi}(t_0)\rangle|^2 = |\langle f | U_3 U_2 U_1 \hat{\Psi}(t_0)\rangle|^2
\]

(2.23)

The last step is allowed because the transformation matrix has only diagonal elements, so the rotation matrices can be pulled through the product of operators. Using \(R^\dagger = R^{-1}\) does the rest of the trick. If we start with all atoms in the lower state: \(|\Psi(t_0)\rangle = |5\rangle\), then the probability that the atoms are in the upper state \(|2\rangle\) after the Ramsey interrogation becomes:

\[
P_{\text{Ramsey}} = |\langle 2 | U_3 U_2 U_1 |5\rangle|^2
\]

(2.24)

2.4.5 \(\pi/2\) pulse

In the process of Ramsey interrogation, one wants to start with all atoms in one of the two clock hyperfine states. After the first cavity passage, the atoms in the cloud should be in an equal superposition of the lower and upper states. The following must thus apply:

\[
|\langle 2 | U_1 |5\rangle|^2 = \frac{1}{2}
\]

(2.25)
Figure 2.5: Ramsey fringes: a representation of equation (2.27), where the transition probability is plotted against the detuning of the rf magnetic field. The other parameters are chosen in such a way that a π/2-pulse is represented. The atoms spend 0.6 sec. above the cavity.

One has to adjust the parameters of the clock setup such that equation (2.25) is satisfied for probe frequencies in resonance with the atomic transition. The only way to bring ‘all’ atoms from the lower to the upper state, is to subject the atoms twice to the resonant rf magnetic field. If the magnetic field is exactly resonant with the clock transition we have δ = 0 and ω_R = Ω. Equation (2.25) is then satisfied for ω_Rτ = ±π/2. That is why the ‘rf-pulses’ experienced by the atoms are called π/2 pulses. We now have:

$$\omega_R \tau = \pm \pi/2 = \mp \frac{13}{9\hbar} \mu_B B_{rf} \tau$$

(2.26)

The physical solution of this equation is of course

$$B_{rf} \tau = \frac{9\pi\hbar}{20\mu_B} \approx 1.23 \times 10^{-7} \text{ Gauss-second.}$$

In practice, one has to take in account the fact that the atoms have a velocity distribution which results in a slightly different cavity passage time τ for each atom. Furthermore, the rf magnetic field amplitude is not entirely constant in space inside the cavity (see chapter 3). Therefore, not all atoms will make the transition. The frequency at which the maximum number of atoms ends up in the upper state is determined. This is the frequency at which the the central Ramsey fringe is located (see next section).
2.4.6 Ramsey fringes

Inserting the explicit expressions for the evolution operators in equation (2.23), and investing some time in simplifying the resulting expressions, gives the formula for the so-called Ramsey fringes:

\[ P_{5\rightarrow 2} = 4 \frac{\omega^2_R}{\Omega^2} \sin^2(\Omega \tau/2) \left( \cos(\delta T/2) \cos(\Omega \tau/2) \right. \]
\[ \left. - \frac{\delta}{\Omega} \sin(\delta T/2) \sin(\Omega \tau/2) \right)^2 \] (2.27)

For small \( \delta \) the transition probability is a very sensitive function of \( \delta \), consisting of narrow cosine-like fringes that are enveloped by a sine squared. Locking the local oscillator frequency to the central fringe of this pattern makes this oscillator suitable for use as a frequency standard (or clock).

2.5 Accuracy and stability

2.5.1 Accuracy

There are two quantities that are important in the characterisation of the quality of a frequency standard: accuracy and stability. The first term has already been used in the introduction and is defined as ‘the degree to which the output frequency agrees with the value corresponding to the definition of the SI-second. It is expressed as the cumulative fractional uncertainty of the output frequency relative to that given by the definition of the unit’ [7, p.846].

The accuracy budget of an atomic frequency standard consists of a list of all effects that shift the frequency with respect to the atomic resonance frequency, the measured or calculated values of these shifts, and the uncertainty in these values. This uncertainty or inaccuracy contributes to the cumulative uncertainty of the output frequency.

The Zeeman shift, the Doppler effect, black body radiation, cavity pulling and the cold collision frequency shift are a few of the main contributors to the accuracy budget of atomic fountain clocks.

Zeeman shift

The Zeeman shift is the most important effect for the helium fountain clock, mainly because of the high C-field that is used. The quadratic Zeeman shift is proportional to the square of the magnetic induction, integrated over the trajectory of each atom and averaged over all atoms: \( \langle B^2 \rangle \). If the magnetic field is not entirely constant over space, \( \langle B^2 \rangle \) is different for each atom in the cloud, as is the Zeeman shift. At high C-field, the same fractional accuracy of the magnetic field can be achieved as at low fields, which means that the absolute uncertainty of \( \langle B^2 \rangle \), and therefore of the Zeeman shift, is much
larger than at low fields. If the exact magnetic field profile along the atom’s trajectory is known, one could compensate for differences in the value of $\langle B^2 \rangle$.

The average value of the magnetic field $\langle B \rangle$ can be measured using a transition in the atom that depends linearly on the magnetic field\textsuperscript{11}. If the field is constant over space, one automatically knows the value of $\langle B^2 \rangle$ to the same accuracy. However, if the field is not constant, $\langle B^2 \rangle \neq (\langle B \rangle)^2$. If this is the case, one can construct a profile of the axial magnetic field by measuring the state of atoms that reach different heights above the cavity. In the actual operation of the clock one can again discriminate between atoms that have reached different heights, because their detection is separated in time. It is much harder to correct for frequency shifts due to radial field inhomogeneities, though. Anyhow, the frequency uncertainties associated with the lack of knowledge of the actual magnetic field profile are a few orders of magnitude higher (see section 4.2) than all other uncertainties, including those described below, that are of the order of $10^{-15}$ [16].

Other contributions to the accuracy budget

The Doppler effect (described in section 5.1.2) causes a broadening of the absorption profile of the atomic resonance. The velocity-dependent frequency shift experienced by a single atom of velocity $v$ is given by [7]:

$$\delta \omega = \omega - \omega_0 = k \cdot v - \frac{1}{2} \omega_0 \frac{v^2}{c^2}$$

(2.28)

where $k$ is the wave vector of the absorbed photon. The first term is the first-order Doppler shift and the second term is called the second-order Doppler shift. For the atomic fountain clock, equation (2.28) has to be applied to the interaction region inside the microwave cavity. In the ideal case, all terms of (2.28) become zero when averaged over all atoms and absorbed photons, due to the cylindrical symmetry of the microwave field and the atomic velocity distribution, as well as the fact that the trajectory of the atoms should be symmetric with respect to the turning point of the atoms. An asymmetry in either the atomic velocity distribution or the microwave field causes a Doppler shift in the clock frequency.

Another effect of interest in atomic clocks is black body radiation. The walls of the ‘housing’ of the clock will emit electromagnetic radiation. The power radiated in each frequency interval is given by Planck’s law and the electromagnetic energy is equally distributed between the electric and magnetic fields. These fields will produce ac Stark and Zeeman effects respectively.

\textsuperscript{11}Because this transition depends much stronger on the magnetic field, comparison of the clock running in this mode with another atomic clock (that may be less accurate than the helium fountain clock) will yield a very accurate measurement of $\langle B \rangle$. 
Two effects that depend on the number of atoms in the cloud are cavity pulling and cold collisions [16]. If the resonance frequency of the cavity is not equal to the atomic resonance frequency, the clock frequency is pulled towards or pushed away from the cavity resonance frequency if the atoms are in the upper or lower state respectively. This cavity pulling is a result of the interference between the field that is emitted by the atoms and the field present in the cavity. The cold collision frequency shift is caused by the collisional interactions between the atoms in the cloud. The collisions can be described by the sum of s-wave, p-wave and higher order collisions. Since $^3$He is a fermion, s-wave collisions are forbidden due to the symmetry of the wavefunction. The p-wave contribution is extremely small compared to the s-wave collisions in atomic clocks based on bosonic systems.

### 2.5.2 Stability

The stability of a clock is defined as the Allan standard deviation $\sigma_y$ of the fractional frequency difference $y(t)$ [17]:

$$y(t) = \frac{\nu(t) - \nu_0}{\nu_0}$$  \hspace{1cm} (2.29)

The fractional frequency difference is integrated over a time $\tau$ in order to obtain an average value $\bar{y}_k$. Subsequently, the root mean square value of the difference between each pair of average fractional frequency differences is taken:

$$\bar{y}_k = \frac{1}{\tau} \int_{t_k}^{t_k + \tau} y(t) dt$$  \hspace{1cm} (2.30)

$$\sigma_y(\tau) = \frac{1}{\sqrt{2(N-1)}} \left[ \sum_{k=1}^{N-1} (\bar{y}_k - \bar{y}_{k-1})^2 \right]^{1/2}$$  \hspace{1cm} (2.31)

For atomic fountain clocks, the Allan deviation is typically given by [16]:

$$\sigma_y(\tau) = \frac{1}{\pi Q_{at}} \sqrt{\frac{T_c}{\tau}} \left( \frac{1}{N_{at}} + \frac{1}{N_{at} n_{ph}} + \frac{2\sigma_N^2}{N_{at}^2} + \gamma \right)^{1/2}$$  \hspace{1cm} (2.32)

where $\tau$ is the measuring time, $T_c$ the frequency cycle duration, $Q_{at} = \nu_0 / \Delta \nu$ is the atomic resonance quality factor and $N_{at}$ is the number of detected atoms. The second term in parentheses stands for shot noise in the detection signal (in the case of detection by a microchannel plate detector, the photon noise must be replaced by an electron noise contribution), the third term indicates the rms fluctuations in the number of detected atoms and the last term is the frequency noise of the interrogation oscillator. All of these terms are expected to be negligible in the helium fountain clock compared to the first term, that is called quantum projection noise and is related to the
statistical nature of quantum mechanics. If an atom is in a superposition of two states \(|\Psi\rangle = \alpha |g\rangle + \beta |e\rangle\), the probability of finding the atom in state \(|e\rangle\) is \(|\beta|^2 = p\). If we define \(P_e\) as the projection operator onto state \(|e\rangle\), the standard deviation of the quantum fluctuations of the measurement of \(|e\rangle\) is given by [16]:

\[
\sigma = \left( \langle \Psi | P_e^2 |\Psi\rangle - (\langle \Psi | P_e |\Psi\rangle)^2 \right)^{1/2} = \sqrt{p(1-p)} \tag{2.33}
\]

This standard deviation scales as \(N_{\text{at}}^{-1/2}\) [18]. In the operation of the clock, the transition probability is alternately measured on each side of the central Ramsey fringe, where \(p = 1/2\). The quantum projection noise therefore contributes to the stability budget of the clock.
Chapter 3

The microwave cavity

3.1 Introduction

The microwave magnetic field is in a way the heart of the atomic clock. The field interacts with the atoms and the result of the interaction is monitored, as well as the exact frequency of the field. The field acts as an intermediate between the intrinsic structure of the atoms and the output device of the clock. It will be explained how this crucial electromagnetic field can be constructed inside a cavity, starting from basic electrodynamics and finally calculating the relevant parameters of the cavity.

3.2 Boundary conditions

The Maxwell equations can be used to derive boundary conditions for the electromagnetic field configuration of a system of finite dimensions, such as a conducting cylindrical cavity\(^1\). Using the integral form of equations (B.5–B.8) to integrate the fields over Gaussian pillboxes and Amperian loops the following conditions for the fields near the boundary between two different media 1 and 2 can be derived [19].

\[
\begin{align*}
D_1^\perp - D_2^\perp &= \Sigma_f \\
B_1^\perp - B_2^\perp &= 0 \\
E_1^\parallel - E_2^\parallel &= 0 \\
H_1^\parallel - H_2^\parallel &= K_f \times \hat{n}
\end{align*}
\]

The subscripts \(\parallel\) and \(\perp\) stand for parallel and perpendicular to the boundary surface respectively. \(\hat{n}\) is the unit vector perpendicular to the surface, \(\Sigma_f\) and \(K_f\) are the free surface charge density and surface current.

These equations apply to the case of time-independent fields. If we deal with time-varying fields the boundary conditions for the static case

\(^1\)The basics of classical electrodynamics are summarised in appendix B.
are valid only if the frequency of the variations is sufficiently low for the charge and current distributions to arrange themselves into the equilibrium configurations. For microwave frequency fields this approximation holds for the boundary between a dielectric and a perfect conductor \((\sigma = \infty)\). Inside a perfect conductor no electric and magnetic fields exist; the surface charge density \(\Sigma_f\) and the surface current density \(K_f\) are supposed to change their configurations instantly in response to changes in the fields in order to give zero fields inside the conductor [20].

From the boundary conditions (3.1 ff.), one can easily see that in a medium near the surface of a perfect conductor only normal \(E\) and tangential \(B\) can exist \((E_\parallel = 0, B_\perp = 0)\), and that \(H_\parallel\) and \(D_\perp\) are discontinuous by an amount that is given by the distribution of the surface charge and current densities.

If we are dealing with a good conductor (of finite conductivity) instead of a perfect one, no instantaneous rearrangement of charge and current can be accomplished to force the fields inside the conductor to be zero. A successive approximation scheme can be developed to derive the fields near the boundary surface of a good conductor [20]. The field configuration will differ slightly from the perfect case; inside a good conductor the fields are attenuated exponentially in a characteristic length \(\delta\), called the skin depth:

\[
\delta = \sqrt{\frac{2}{\mu_0 \mu \omega \sigma}}
\]  

(3.5)

In this approximation the fields inside the conductor are of the form:

\[
H_c = H_\parallel e^{-\xi/\delta} e^{i\xi/\delta}
\]  

(3.6)

in which \(\xi\) is the normal coordinate inward into the conductor. The associated electric field is given by [20, 21]:

\[
E_c = \sqrt{\frac{\mu_0 \mu \omega}{2\sigma}} (1 - i) (\hat{n} \times H_\parallel) e^{-\xi/\delta} e^{i\xi/\delta}
\]  

(3.7)

The normal unit vector \(\hat{n}\) is directed outward from the conductor. In the derivation of equations (3.6) and (3.7), the electric displacement current (the second term at the right-hand side of equation (B.6)) is assumed to be much smaller than the electric current density \(J_f\) that exists in a small layer near the surface of the conductor\(^2\). Furthermore, one uses the fact that the spatial variations of the fields normal to the surface are much larger than those parallel to the surface\(^3\). A schematic representation of the fields is shown in figure 3.1.

\(^2\)It is now a volume current density, as opposed to the surface current density in the ideal situation of a perfect conductor.

\(^3\)This means that all derivatives with respect to parallel coordinates can be neglected compared to those with respect to normal coordinates.
3.3 Circular cylindrical cavity

The most common shape of microwave cavities used for atomic fountain clocks is that of a circular cylinder of radius $R$ and height $d$ (see figure 3.2). The axial symmetry of this configuration is convenient, since the constant magnetic field and the trajectory of the atoms are subjected to the same symmetry. In this section the fields in a circular cylindrical cavity are derived, assuming (as an approximation) perfectly conducting walls (skin depth $\delta = 0$) and a perfect vacuum inside the cavity ($\epsilon = \mu = 1$).

We start from the wave equations for the electric and magnetic fields (equations B.11 and B.12). In cylindrical coordinates, the fields can be split into transverse and axial components: $E = E_t + E_z$ and $H = H_t + H_z$. The same is done with the differential operator: $\nabla^2 = \nabla^2_t + \partial^2 / \partial z^2$. If we now assume that the solutions are plane waves of the form:

$$E = E_0 e^{-i(\omega t \pm \beta z)} \quad (3.8)$$

the wave equation for the $z$-components of the fields in cylindrical coordinates $(\rho, \phi, z)$ becomes:

$$\left\{ \nabla^2_t + \frac{\partial^2}{\partial z^2} + \omega^2 \epsilon_0 \mu_0 \right\} \left\{ \begin{array}{c} E_z \\ H_z \end{array} \right\} = 0 \quad (3.9)$$

Using the fact that $\partial^2 / \partial z^2 = -\beta^2$ and putting $\omega^2 \epsilon_0 \mu_0 - \beta^2 = \gamma^2$ we can
write:
\[
\left\{ \nabla^2 + \gamma^2 \right\} \left\{ \begin{array}{c} E_z \\ H_z \end{array} \right\} = \left\{ \begin{array}{c} \frac{\partial^2}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial}{\partial \rho} + \frac{1}{\rho^2} \frac{\partial^2}{\partial \phi^2} + \gamma^2 \end{array} \right\} \left\{ \begin{array}{c} E_z \\ H_z \end{array} \right\} = 0
\]
(3.10)

These equations can be solved using the method of separation of variables. Following this procedure, the differential equations are transformed into Bessel equations. The fields are therefore expressed in Bessel functions. The solutions can be written as superpositions of different modes, for example transverse electric (TE) modes for which \( E_z = 0 \) and transverse magnetic (TM) modes (\( H_z = 0 \)). For each of these modes the transverse field components can be calculated from the z-components, using the Maxwell equations. The boundary condition on the component of the \( B \)-field normal to the surface implies that the z-components of the magnetic field vanishes at \( z = 0 \) and \( z = d \) (because no \( B \)-field can exist in the perfectly conducting wall). This requires \( \beta = p\pi/d \), with \( p = 1, 2, 3, \ldots \). The so-called characteristic equation becomes:
\[
\gamma^2 = \omega^2 \varepsilon_0 \mu_0 - \left( \frac{p\pi}{d} \right)^2
\]
(3.11)

For each value of \( p \) there is an infinite number of eigenvalues \( \gamma_{mn} \) that satisfy the characteristic equation. The labels \( m \) and \( n \) are integers, \( m = 0, 1, 2, \ldots \), \( n = 1, 2, 3, \ldots \). The \( \gamma_{mn} \) are roots of the Bessel functions \( J_m(x) \) or \( J'_m(x) \) for TM and TE modes respectively. The corresponding angular frequencies \( \omega_{mnp} \) are the resonance frequencies of the cavity:
\[
\omega_{mnp} = \frac{1}{\sqrt{\varepsilon_0 \mu_0}} \sqrt{\gamma_{mn}^2 + \left( \frac{p\pi}{d} \right)^2}
\]
(3.12)

The electromagnetic field configurations that correspond to the \( \omega_{mnp} \) are called TE\(_{mnp} \) and TM\(_{mnp} \) modes.
3.4 TE011 field configuration

An atomic fountain clock needs an rf magnetic field that is aligned along the bias magnetic field. All other field components are unwanted. Therefore, TM modes are of no use, since by definition \( H_z = 0 \) for transverse magnetic modes. The TE mode that is most suited for the fountain clock is the TE011 mode. The components of the magnetic and electric fields of the TE011 mode of a circular cylindrical cavity are given by:

\[
E_z = E_\rho = H_\phi = 0
\]

(3.13)

\[
H_z = H_0 J_0(\gamma \rho) \sin\left(\frac{\pi z}{d}\right)
\]

(3.14)

\[
H_\rho = H_0 \frac{\pi}{\gamma d} J'_0(\gamma \rho) \cos\left(\frac{\pi z}{d}\right)
\]

(3.15)

\[
E_\phi = \frac{\omega \mu}{\gamma} H_0 J'_0(\gamma \rho) \sin\left(\frac{\pi z}{d}\right)
\]

(3.16)

in which \( \gamma = x'_{01}/R \) is the first root of the Bessel function \( J'_0(x) \) (\( x'_{01} = 3.832 \)).

As can be seen from equations (3.13 ff.) and figure 3.3, the TE011 mode has the advantage that it has only 3 nonzero components, \( H_z \) has its maximum at the centre of the cavity and the unwanted \( H_\rho \) is zero at the centre. Furthermore, because it is the lowest mode, the TE011 fields have the lowest number of nodes, meaning that the spatial variation of the fields is less than for higher modes.

The resonance frequency of the cavity is of course determined by the atomic resonance frequency that is calculated in section 2.3: \( \omega_c = 3.9925 \times \)
$10^{10}$ rad/s. From equation (3.12) we can calculate the corresponding radius and height of the cavity. These parameters are both completely determined by this equation if we require that $R = d$. This requirement means that the cavity has a convenient shape and it also guarantees a quality factor (see next section) that is of the right magnitude. Both height and radius of the cavity have to be 3.7208 cm.

### 3.5 Q-value and bandwidth

The quality factor $Q$ of a resonant cavity is defined as $2\pi$ times the ratio of the time-averaged stored energy in the cavity and the energy-loss per cycle, which can also be expressed as [20]:

$$Q = \frac{\omega_c \text{ stored energy}}{\text{power loss}}$$  \hspace{1cm} (3.17)

The quality factor does not only determine the power loss of the cavity, but it is also connected to the width of the resonance and it is related to the strength of the cavity pulling effect. Typical $Q$-values for copper microwave cavities of a few centimeters in height are a few times $10^4$. The $Q$-value should not be too low, because the cavity pulling effect described in section 2.5.1 increases with decreasing $Q$ [22]. As can be seen from figure 3.4, much larger $Q$ cannot be reached with copper cavities of practical size. Extremely high $Q$, that could for instance be reached with superconductors, is of no use. This can be seen by considering the fact that the frequency of the field in the cavity has to be varied a little in the operation of the clock. This means that the fields should not remain in the cavity forever (which would be the case for a superconducting cavity).

Equation (3.17) can be used to calculate the quality factor of the cavity for each mode. The time-averaged energy density in the cavity is, for fields with harmonic time-dependence, given by:

$$U_{EM} = \frac{1}{4} \left( \varepsilon_0 E^2 + \frac{1}{\mu_0} B^2 \right) = \frac{1}{4} \left( \varepsilon_0 E^2 + \mu_0 H^2 \right)$$  \hspace{1cm} (3.18)

where it is assumed that we have no medium in the cavity, so that we can take $B = \mu_0 H$. If an electrical circuit (like a microwave cavity) is at resonance, the electric and magnetic stored energies are equal:

$$W_E = \frac{\varepsilon_0}{4} \int_{V_c} E^2 d\tau = W_M = \frac{\mu_0}{4} \int_{V_c} H^2 d\tau$$  \hspace{1cm} (3.19)

From equations (3.13 ff.) can be seen that the total electromagnetic energy stored in our TE$_{011}$ cavity is given by:

$$W_{EM} = \frac{\varepsilon_0}{2} \int_{V_c} E_{\phi}^2 d\tau$$  \hspace{1cm} (3.20)
Evaluation of this integral gives for the stored energy [23]:

\[
W_{EM} = \frac{\omega^2 \mu_0^2 \epsilon_0 R^2 \pi d H_0^2}{8 \gamma^2} (\gamma R)
\]  

(3.21)

The time-average power loss per unit volume in the conducting walls is given by the expression \( P_c = \frac{1}{2} J \cdot E^* \). Using equation (3.7) as a first order approximation of the electric field inside the cavity walls, the total power loss becomes: [23]:

\[
P_c = \frac{R_s}{2} \int_S |H_\parallel|^2 ds
\]

(3.22)

in which \( R_s = \sqrt{\omega \mu_0 / 2 \sigma} \) is the surface resistivity of the metallic walls (\( \sigma \) would be the conductivity of copper: \( \sigma = 5.8 \times 10^7 \ \Omega^{-1} m^{-1} \)). The integral is over the surface of the walls of the cavity. For the TE\(_{011}\) cavity this comes down to:

\[
P_c = \frac{R_s}{2} \left\{ \int_{z=0}^{d} \int_{\phi=0}^{2\pi} |H_\parallel (\rho = R)|^2 R d\phi dz + 2 \int_{\rho=0}^{R} \int_{\phi=0}^{2\pi} |H_\parallel (z = 0)|^2 \rho d\phi d\rho \right\}
\]

(3.23)

The second integral appears twice, because the top and bottom end plates of the cavity yield two identical terms. Integration of (3.23) yields:

\[
P_c = \frac{R_s}{2} \pi H_0^2 J_0^2 (\gamma R) \left\{ \frac{R d}{2} + \left( \frac{\pi R}{\gamma d} \right)^2 \right\}
\]

(3.24)

Combining equations (3.21), (3.24) and (3.11) gives the quality factor of the cavity:

\[
Q_c = \sqrt{\frac{\mu_0}{\epsilon_0}} \left\{ \frac{\gamma^2 + \left( \frac{\pi}{d} \right)^2}{\gamma R + \frac{2}{d} \left( \frac{\pi}{d} \right)^2} \right\}^{3/2}
\]

(3.25)

If we take a copper cavity, the quality factor of the unloaded cavity becomes \( 32 \times 10^3 \). The effective \( Q \)-value of the whole system, including coupling lines and cavity holes will be lower. In fact, if the impedances of the coupling line and the cavity itself are matched such that no coupling power is reflected by the cavity, the external quality factor due to coupling \( Q_e \) has the same value as the intrinsic cavity quality factor \( Q_c \) [7, p.186]. In

\[4\]The skin depth at the atomic resonance frequency is in this case \( \delta = 0.83 \ \mu m \). The thickness of the cavity walls should be much larger than this skin depth, a condition that will be easily fulfilled.
order to obtain the loaded quality factor of the system $Q_L$ we can use the general formula:

$$\frac{1}{Q_L} = \frac{1}{Q_e} + \frac{1}{Q_c}$$

(3.26)

Impedance matching thus yields a loaded quality factor of $Q_c/2 = 16 \times 10^3$.

The quality factor is related to the FWHM of the resonance modes (that have a Lorentzian shape, see figure 3.5):

$$\Gamma = \omega_c/Q$$

(3.27)

Furthermore, a finite quality factor leads to a shift in the resonance frequency of (for $Q \gg 1$) [20]:

$$\Delta \omega_c = -\omega_c/2Q$$

(3.28)

The width of the resonance of a cavity with a quality factor of $16 \times 10^3$ can now be calculated to be about 0.4 MHz. If the radius and the height of the cavity are increased with 10 µm, the resonance frequency changes with an amount of 1.7 MHz. This example shows that small mechanical imperfections will yield a cavity that is not resonant with the desired atomic transition frequency. Because of the finite width of the cavity resonance, it will still be possible to couple fields into the cavity. The cavity pulling effect, however, is proportional to the cavity detuning with respect to the atomic resonance frequency. Fortunately, the resonance frequency of the cavity can be adjusted by changing the temperature of the cavity. The typical temperature dependence of the cavity resonance frequency is about 100 kHz/K [22].
3.6 Cavity holes

All formulae in the preceding sections are based on a circular cylindrical cavity that is closed. For proper operation in a atomic fountain clock, however, the cavity needs to have holes in the upper and lower plate in order to let the atoms pass through. The holes will cause distortions in the calculated electromagnetic field configuration. The exact form of these distortions is hard to calculate\(^5\), but the basic structure of the fields is maintained. In order to minimise the perturbations, one would like to minimise the size of the holes. On the other hand, increase of the size of the holes provides an increase in the number of atoms that can be led through the cavity, thus enhancing the measurement statistics.

One should, however, also keep in mind that not only the properties of the microwave field get worse further away from the \(z\)-axis; we have the same problem with the magnetic \(C\)-field above the cavity. One has to select atoms that have been within a certain region around the \(z\)-axis during their flight, thereby diminishing the advantages of large cavity holes (and of large atom cloud diameters).

The time the atoms spend in the interaction region is important. One would therefore like to have a microwave field inside the cavity only. Leakage of fields through the holes in the upper and lower plates must be suppressed. This can be accomplished by connecting waveguides of circular cross-section to the holes. If the diameter of these so called cut-off waveguides is small enough, the fields cannot escape from the cavity, because the wavelength of these fields is too large to be guided through the waveguide. The radius of

---

\(^5\)A finite elements calculation has to be performed to calculate the effect of the cavity holes on the fields. It is possible, however, to use results from other groups, since the dimensions of cavities suited for cesium or rubidium clocks can be scaled to dimensions adapted for the helium clock.
the cut-off waveguides should be smaller than the cut-off radius $R_c$ in order to keep fields of frequency $\omega$ inside the cavity [24]:

$$R_c = \frac{c r_{mn}'}{\omega}$$

for transverse electric TE$_{mn}$ fields. This equation provides an upper limit for the size of the holes in the end plates. The cut-off radius for our fountain clock cavity can be calculated to be 2.87 cm, almost as large as the cavity radius of 3.72 cm. Since the magnetic field amplitude changes too much over such a large distance, the cavity hole radius should be well below the cut-off radius.

### 3.7 Coupling of fields into the cavity

Now that we know which fields can be contained in a microwave cavity, we need to know how to couple these fields into the cavity. Electromagnetic fields can be coupled into a resonator or wave guide in two distinct ways: *electrically* and *magnetically*. In both cases the central conductor of a coaxial transmission line is used as an antenna. In electric coupling, the central conductor has to be aligned parallel to one of the electric field components of the wanted mode. Magnetic coupling is accomplished by bending the conductor into a loop, which has to be positioned perpendicular to the direction of one of the magnetic field components of the mode that one wants to excite [24]. Sometimes a combination of both methods is used. Only modes that have a resonance frequency that is within a certain range of the frequency of the coupling signal can be excited (the power that is coupled in obeys a Lorentzian spectrum centered around the cavity resonance frequency).

The fields can be coupled directly into the cavity or using a waveguide as an intermediate stage. In the latter case, the waveguide is connected to the cavity via a hole (iris) in the cavity wall. Symmetric coupling from two sides of the cavity is recommended by some groups [25].

Excitation of the TE$_{011}$ mode yields an interesting problem: this mode has the same resonance frequency as the TM$_{111}$ mode and it has unwanted radial and azimuthal magnetic components. Calculating the surface currents associated to the both modes, however, shows that for the TM$_{111}$ mode current flows from the cylindrical wall to the end plates. This is not the case for the TE$_{011}$ mode. This provides the opportunity to exclude the TM$_{111}$ mode from the cavity by placing insulating rings between the cylindrical wall and the end plates [7, p.1194].

### 3.8 Amplitude of the magnetic field

The amplitude of the magnetic field component directed along the $z$-axis can be calculated from the height of the cavity, the velocity of the atoms
that are launched through the cavity (ca. 3 m/s) and the conditions for the \( \pi/2 \)-pulse as derived in section 2.4.5. These conditions can be united in a single formula for the rf magnetic field amplitude expressed in Gauss:

\[
B_{rf} = \frac{9\pi h v}{26\mu_B d} \approx 1.23 \times 10^{-7} \times \frac{v}{d} \tag{3.30}
\]

The amplitude of the magnetic induction that is found from this formula is \( B_{rf} = 10^{-5} \) Gauss. The corresponding magnetic field strength is \( H_{rf} = B_{rf}/\mu_0 = 8 \times 10^{-4} \) T, if we assume a vacuum in the cavity\(^6\).

The total amount of electromagnetic energy stored in the cavity can now be calculated from equation (3.21) and is only \( 4 \times 10^{-18} \) J (corresponding to \( \sim 10^6 \) photons). Because of the high frequency, the amount of energy that has to be coupled in per second is much larger than the amount of energy present in the cavity: using equation (3.21) and the value obtained for the quality factor of the cavity (\( 16 \times 10^3 \)), we get for the incoupling power a value of \( 10^{-11} \) W.

\(^6\)The cavity is in a vacuum chamber, but at the moment that the interaction takes place, the atoms will change the magnetic permeability a little. This effect is supposed to be negligible though.
The microwave cavity
Chapter 4

The C-field

4.1 Introduction

In atomic fountain clocks, a bias magnetic field is needed to break the degeneracy of the magnetic sublevels of the hyperfine states that are used. In this way a two-level system is created. This so-called C-field, that is generated with a solenoid magnet, also shifts the atomic resonance frequency (see section 2.3) and this shift has to be corrected for. The size of the correction has to be known with an accuracy of $\delta \nu/\nu = 10^{-12}$ in order to obtain a helium fountain clock that is useful for high-precision experiments. The desired accuracy of the clock imposes strict demands on the spatial homogeneity of the C-field as well as the temporal stability of the field.

In this chapter, the features of the C-field are discussed, including the possibilities for enhancement of the field homogeneity. Two models that are used to calculate the effects of adaptations of the solenoid shape are described at the end of the chapter.

4.2 Homogeneity requirements for high and low field

The Zeeman shift in resonance frequency of a single atom, as detected by the clock, can be calculated if the magnetic field profile is known to a certain accuracy. The position of the central Ramsey fringe, however, is determined by an ensemble of atoms that all experienced slightly different magnetic fields on their trip through the C-field. This results in inhomogeneous line broadening of the Ramsey fringes, an effect that is too complex to predict with high precision. Therefore, an estimate of the order of magnitude of the required C-field homogeneity is made by looking at the effect on the single atom resonance frequency of small deviations from the nominal value of the field amplitude.
The magnetic field dependence of the shift in resonance frequency can be calculated for each given magnetic field strength. Equation (2.4) yields the following relation between the fractional frequency shift and the C-field:

\[
\frac{\delta \nu}{\nu} = \frac{1}{\nu(B_0)} \frac{\partial \nu}{\partial B} |_{B_0} (B - B_0) + \frac{1}{\nu(B_0)} \frac{1}{2} \frac{\partial^2 \nu}{\partial^2 B} |_{B_0} (B - B_0)^2
\]

Equation (4.1) can be applied to two distinct situations: high and low C-field. In each case, the magnetic field homogeneity required for \( \delta \nu/\nu = 10^{-12} \) is calculated.

4.2.1 Spatial homogeneity

**Low C-field**

In conventional fountain clocks, based on bosonic atomic species, low C-fields are used, typically of the order of one Gauss. The atomic transition frequency is to first order independent of the magnetic field in those clocks. For the helium fountain clock, however, the first order term of the Taylor expansion (4.1) is dominant for small deviations from \( B_0 \) in this regime. For a magnetic induction of 0.8 Gauss we get \( \delta \nu = 9.3 \times 10^5 (B - B_0) \) Hz/G. At a transition frequency of \( \nu = 6.74 \) GHz requiring that \( \delta \nu/\nu = 10^{-12} \) yields \( (B - B_0) = \delta B = 7.2 \times 10^{-9} \) G. This comes down to a relative homogeneity\(^1\) of the field \( \delta B/B \) of \( 9 \times 10^{-9} \).

**High C-field**

For the situation of the (high) anticrossing field of 802.6 Gauss, the first order term of equation (4.1) vanishes\(^2\) and we get \( \delta \nu = 6.2 \times 10^2 (\delta B)^2 \) Hz/G\(^2\). We now have \( \nu = 6.35 \) GHz, yielding \( (\delta B)^2 = 1.0 \times 10^{-5} \) G\(^2\), and resulting in an absolute accuracy of \( 3.2 \times 10^{-3} \) G, or a relative homogeneity of \( \delta B/B = 4 \times 10^{-6} \).

It is clear that the homogeneity requirements imposed on the C-field are much less demanding for the anticrossing ‘high’ field than for low fields. Even though a given absolute accuracy can be much easier reached at low fields than at high fields, the fact that at anticrossing the field is to first order independent of the magnetic field fluctuations implies that the clock

\(^1\)Since the magnetic field is proportional to the electric current density through the solenoid, and increasing the magnetic field comes mainly down to increasing the total amount of current, one can see that inhomogeneities present at low fields are amplified with the increase of current, but without changing the relative homogeneity. This relative homogeneity is therefore the relevant parameter if the technical feasibility is concerned.

\(^2\)See section 2.3.
4.2 Homogeneity requirements for high and low field

Figure 4.1: Dependence of the Zeeman shift in the atomic resonance frequency on the deviations of the magnetic field around anticrossing field $\delta B$.

frequency accuracy requirements can be met more easily with anticrossing field.

In clocks based on bosons, e.g. cesium, the frequency dependence on the deviations of the magnetic field is also second-order (quadratic), and the absolute accuracy needed is of the same order of magnitude as in the helium clock. The required relative homogeneity, however, is much lower due to the smaller absolute magnitude of the fields.

In the remainder of this chapter, 803 G will be used as the value for the C-field and a relative homogeneity of $4 \times 10^{-6}$ is the goal that is aimed for.

4.2.2 Temporal stability

In addition to the spatial homogeneity of the C-field, also the temporal stability of the field is of importance. Because the field is proportional to the current through the windings of the solenoid, the stability of the current supply has to be also of the order of $10^{-6}$ during the operation of the clock. A high-stability current control that generates a 10 A dc current with a stability of $10^{-7}$ measured over a few hours has been designed and built in Torino [26].

4.2.3 Magnetic shielding

The C-field has to be shielded from external magnetic fields, such as the Earth’s magnetic field. The Earth’s magnetic field is ca. 0.3 G and will disturb both the spatial and temporal homogeneity of the C-field. Magnetic shielding is accomplished by a mumetal case that is placed around the interaction region of the clock. The part of the Earth’s magnetic field that penetrates into this region can easily be held below the mG range [27, 28].
4.3 The solenoid magnet

Constant magnetic fields that are uniform over a cylindrical region with a length of several tens of centimeters are usually generated with electromagnets. The design of the magnet is as follows: a solenoid is formed by several windings of copper wire around a cylinder with an inner diameter that is large enough to contain the microwave cavity. The cylinder (a copper pipe for example) also acts as a vacuum chamber.

4.4 Inhomogeneities

The main source of field inhomogeneity appears, in theory, to be the result of the finite length of the solenoid. The axial field has its maximum at the centre of the solenoid and the amplitude of this field component decreases towards the ends of the cylinder. This effect is of significant size. On the other hand, it is a very smooth effect that one might be able to account for when operating the clock. Still, an investigation was made on how

---

3The minimal height of the C-field region depends on the vertical velocity of the atoms. If this velocity is 3 m/s in the microwave cavity, the C-field region has to be at least 45 cm in height.

4It is possible to select only the atoms that have experienced a certain part of the C-field, by filtering out the atoms that are within a certain range of the velocity distribution.
to eliminate this problem. This effect is what will be called a systematic inhomogeneity. The radial dependence of the axial magnetic field and the unwanted radial components are effects of the same kind. Inhomogeneity effects that result from construction errors will be called random inhomogeneities. Bulges in the outer surface of the cylindrical pipe are an example of random inhomogeneities [28]. It is almost impossible to account for these effects in a sufficient manner, but attempts have been made (see section 4.4.1).

4.4.1 State of the art

Now that we know something about the requirements that are imposed on the uniformity of the magnetic field, it is of course instructive to have a look at the homogeneities that have been achieved throughout the world. The most common area to find high-homogeneity magnetic fields is in the Nuclear Magnetic Resonance (NMR) industry. Field homogeneities of well below $10^{-6}$ can be reached, for example using superconducting magnets, but always over regions of only a few centimeters.

Random inhomogeneities can be ‘counterbalanced’ by a technique called shimming, a local compensation technique. There are two ways of shimming: in passive shimming ferromagnetic rings are placed inside the solenoid and in active shimming current turns are placed on the outer surface of the main winding [28]. The latter is better suited for adaptation of the field after constructing the main coil. An NMR probe can be used to measure the magnetic field in the solenoid. Starting from the measured magnetic field distribution, a least-squares numerical calculation yields an optimal configuration for the shimming coils. In order to reach a 15 p.p.m. homogeneity the Costanzo group in Torino needed 7 separate power supplies and 18 shim coils [28]. In a more recent article they claim to observe a 60 p.p.m. disuniformity, in this case derived from the actual performance of a thermal cesium beam clock that incorporated this C-field [29].

4.5 Calculation of the field

There are at least two distinct ways of calculating the magnetic field generated solenoid magnets. Using these models one can try to estimate the effects on the magnetic field of inhomogeneities in the current distribution.

4.5.1 Current density integration

One possible approach of calculating the field of a coil is to integrate the expression for the field of a single loop along the axial and radial direction,

This limits the number of atoms that can be used, however, which implies worse statistics.
using the approximation of a smooth volume current density. In some specific cases this can be done analytically. One would of course like to have an analytical expression for the normalised root-mean-square (rms) value of the deviations of the magnetic field amplitude from the average value. One likes to solve the systematic problem of the finite coil length. Several adaptations of the geometry of the coil can be made in order to decrease the normalised rms value of the inhomogeneity, that is defined by:

\[ \text{rms} = \sqrt[3]{\int (H_z(r) - \overline{H}_z)^2 \, d^3r} \]  

(4.2)

\[ \overline{H}_z = \frac{\int H_z \, d^3r}{\int d^3r} \]  

(4.3)

If the homogeneity of the field on the \( z \)-axis is calculated, the integrations are over the \( z \)-coordinate only.

Differentiating this expression to the variables that represent the shape of the solenoid would lead to the optimal shape. Unfortunately, no analytical expression of the rms value of a solenoid of a shape that would partially compensate the finite length problem can be found. Numerical optimisation is the way of dealing with these kinds of problems, but for that procedure the approach described in section 4.5.2 is more suitable. Let’s first have a look at the expressions for the magnetic field in the current density integration approach.

The field on the axis of a single loop of radius \( R \), carrying a current \( I \) is given by:

\[ H_z(z) = \frac{\mu I}{2} \frac{R^2}{(R^2 + z^2)^{3/2}} \]  

(4.4)

If an uniform current density \( j \) is assumed, this expression can be integrated over the length and thickness of the coil:

\[ H_z(z) = \frac{\mu j}{2} \int_{a_1}^{a_2} \int_{-b}^{b} \frac{R^2}{(R^2 + z^2)^{3/2}} \, dz \, dR \]  

(4.5)

The axial magnetic field \( (H_z) \) on the axis at the centre of the solenoid can then be written as a function of the normalised length \( \beta = l/2a_1 \) and the parameter \( \alpha = a_2/a_1 \) where \( l = 2b \) is the length and \( a_1 \) and \( a_2 \) are the inner and outer radii of the solenoid respectively (see figure 4.2). The only other relevant parameters are the current density \( j \) and \( \lambda \) which is defined as the ratio of the active section of the winding (the part that carries the current) and the total section of the winding. The field at the centre of the solenoid, \( H_0 \), is now calculated by integrating equation (4.5) [30]:

\[ H_0 = j\lambda F(\alpha, \beta) \]  

(4.6)

\[ F(\alpha, \beta) = \frac{4\pi \beta}{10} \ln \frac{\alpha + (\alpha^2 + \beta^2)^{1/2}}{1 + (1 + \beta^2)^{1/2}} \]  

(4.7)
4.5 Calculation of the field

Figure 4.3: Notches: explanation of the parameters used in equations (4.9) and (C.7). The notches are cutaways in the primary coil that can be used to reduce the finite coil length effect within the radius of convergence of the expansions mentioned. (Taken from [30])

Now we use the argument that the end field of any coil is half the central field of a coil of twice the length [30]. We can derive the field $H_z$ as a function of the normalised $z$ variable:

$$H_z \left( \frac{z}{a_1} \right) = \frac{j \lambda a_1}{2} \left[ F(\alpha, \beta + z/a_1) + F(\alpha, \beta - z/a_1) \right]$$

(4.8)

The effects of the finite coil length can be compensated to some extent by superimposing secondary coils, carrying reversed currents, on the primary coil. This amounts of course to cutting away a part of the current distribution of the primary coil. To find the actual dimensions of these so-called compensating coils, an expansion can be made of equation (4.8) [30]:

$$H_z = j \lambda a_1 \left\{ F_0(\alpha, \beta) - F_c(\alpha_c, \beta_c) + \left[ F_0 E_2(\alpha, \beta) - F_c E_2(\alpha_c, \beta_c) \right] \left( \frac{z}{a_1} \right)^2 \
+ \left[ F_0 E_4(\alpha, \beta) - F_c E_4(\alpha_c, \beta_c) \right] \left( \frac{z}{a_1} \right)^4 + \ldots \right\}$$

(4.9)

This equation applies to the inside and the Helmholtz notches (see figure 4.3). A slightly different version exists for the outside notch (see appendix
The compensating coils can of course also be characterised by the normalised length $\beta$ and thickness $\alpha$. In equation (4.9) these parameters are given the subscript $c$. The error coefficients $E_n$ depend on yet another set of parameters $C_m$. The expressions for these parameters are given in appendix C. The inside notch has two parameters and is able to cancel the second and fourth order terms of (4.9). One can use the programme Mathematica to find a graphical solution of this problem.

Expansion (4.9) is of course only valid for points on the $z$-axis sufficiently close to the centre of the solenoid. The expansion does not converge for points further away from the centre than the closest point where the current density is nonzero. This comes down to the requirement that $z < a_1$. This means that calculating the optimal dimensions of an inside or Helmholtz notch is useful only for estimating the achievable homogeneity in a small region at the centre of the coil, which is much too small a region for the purpose of an atomic clock. Still, we can derive useful conclusions from this investigation.

In the following calculations, a main coil with a length of 700 mm, an inner radius of 45 mm and an outer radius of 69 mm is used, which are typical values for the solenoid C-field magnet [28].

**Inside notch**

Equation (4.9) can be used to calculate the exact shape of the inside notch that compensates the second and fourth order deviations from the central field. The prefactors of both the second and fourth order terms should be zero. The values of $\alpha_c$ and $\beta_c$ that satisfy both conditions can be found using a contour plot of the expressions given in appendix C.

As can be seen from figure 4.4, the inside notch should have $\alpha_c = 1.001\alpha_1$, which means that the depth of the notch should be 0.1% of the inner radius of the main coil (45 $\mu$m). The normalised length of the notch has...
to be $\beta_c = 0.9$, which comes down to a notch length of 1.8 times the length of the inner radius of the coil (81 mm). Should one be able to construct such an inside notch, this would mean a reduction of the normalised root mean square deviation from the average magnetic field value from $2 \times 10^{-4}$ to $8 \times 10^{-5}$ within the radius of convergence of the expansion. Over the range that is needed for the clock, ca. 45 cm, the homogeneity increases from $9.3 \times 10^{-3}$ to $8.9 \times 10^{-3}$.

**Outside notch**

The procedure for the outside notch is similar. One has to use a slightly adapted version of equation (4.9) (see appendix C). The results are also quite similar: the outside notch comes down to a cutaway with a depth of 0.17% of the outer radius of the main coil, extending over a length of 1.8 times the outer radius of the main coil. The outer radius of the main coil depends on the number of windings on it, but 69 mm was used in the simulation. Using this outer radius, the depth of the cutaway has to be $117 \, \mu m$. This would lead to an increase in magnetic field homogeneity to $2.6 \times 10^{-6}$ over a region of 90 mm and the outside notch would yield a homogeneity of $8.5 \times 10^{-3}$ over a region of 45 cm.

**4.5.2 Single coil summation**

In this picture one examines the magnetic field generated by an infinitely thin wire, carrying a current $I$. The field generated by the solenoid is then approximated as being the sum of the fields that correspond to a number of loops, each a certain distance apart. This view can be extended to solenoids.
of finite thickness by summing the fields from several layers of coaxial loops of different radius (see figure 4.2). This approach is mainly suitable for estimating the effect of small displacements of some of the current loops, corresponding to possible construction defects that one has to take in account when the clock is built. In principle, this approach could also be used for numerical optimisation of the number of coils at the different positions along the \(z\)-axis, aiming at the elimination of the finite coil length effect. Some investigations have been made towards this subject, but the observed sensitivity of the field with respect to mechanical imperfections suggests that the practical implementation of the results of the calculations would be useless.

As mentioned in the previous section, the field on the axis of a single loop is given by:

\[
H_z(z) = \frac{\mu I}{2} \frac{R^2}{(R^2 + z^2)^{3/2}}
\]  

(4.10)

If we want to investigate the effects of inhomogeneous current distributions, we are not allowed in this case to integrate this expression to obtain the resultant magnetic field. Instead we will sum over a large number of separate current loops, allowing the freedom to adapt each of the contributing terms to our liking. Starting out with \(N\) different sites along the \(z\)-axis, a distance \(c_1\) apart and allowing for \(N_k\) windings on top of each other, with a radius increasing with steps of \(c_2\) we can write:

\[
H_z(z) = \sum_{k=0}^{N} \sum_{m=0}^{N_k} \frac{\mu I}{2} \frac{(R + mc_2)^2}{((R + mc_2)^2 + (z - kc_1)^2)^{3/2}}
\]  

(4.11)

The number \(N_k\) is in principle a function of \(k\). Inserting equation (4.11) in Mathematica and using for example a random generator to vary the distances between the loops gives an estimate of the effects of nonuniform winding of the copper wires.

**Results**

Mathematica was used to calculate the rms value of the field homogeneity as defined in equation (4.2) for several perturbations of the positions of the current loops. A solenoid with 221 loop sites was chosen, each 3 mm apart, yielding a length of 66 cm. The default setting for the spacing between the loops in the radial direction was 1 mm. Formula (4.11) was used to calculate the magnetic field and the corresponding rms homogeneity.

Random displacements were added to the positions of the current loops along the \(z\)-axis and in the radial direction. The magnitude of the displacements were chosen by a random generator from several intervals, ranging from \((0, 0.01)\) mm to \((-3, 3)\) mm. The outcome of these calculations was
4.5 Calculation of the field

Table 4.1: Results from the single coil summation simulation, in which a bulge in the outer surface of the pipe on which the wire is wound was simulated. The bulge extended over a region of 1.5 cm at 10 cm from the centre of the solenoid. The rms homogeneities over 30 and 15 cm around the centre are given. A bulge of 0.5 mm in height already introduces an additional inhomogeneity of the order of $10^{-4}$.

<table>
<thead>
<tr>
<th>bulge height</th>
<th>length of C-field region</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 mm</td>
<td>$2 \times 10^{-4}$</td>
</tr>
<tr>
<td>0.5 mm</td>
<td>$4 \times 10^{-4}$</td>
</tr>
<tr>
<td>1 mm</td>
<td>$8 \times 10^{-4}$</td>
</tr>
<tr>
<td>3 mm</td>
<td>$2 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

Table 4.1: Results from the single coil summation simulation, in which a bulge in the outer surface of the pipe on which the wire is wound was simulated. The bulge extended over a region of 1.5 cm at 10 cm from the centre of the solenoid. The rms homogeneities over 30 and 15 cm around the centre are given. A bulge of 0.5 mm in height already introduces an additional inhomogeneity of the order of $10^{-4}$.

that, due to the randomness of the perturbations, no significant extra inhomogeneity is added to the inhomogeneity caused by the finite coil length effect (ca. $2 \times 10^{-4}$ over 30 cm). Moreover, random perturbations are not entirely realistic, since in practice disuniformities in the current distribution are correlated: the wires have a certain thickness and cannot be in the same place! Therefore, a simulation of a bulge in the outer surface of the pipe on which the copper wire is wound was made. The radius of all loops was increased with a certain amount in a region of 1.5 cm at 10 cm from the centre of the solenoid. In order to be able to compare the effects of random perturbations and finite coil length, the rms value was calculated over 30 cm and 15 cm (around the centre of the solenoid) for each interval. Some results are shown in table 4.1.

The largest influence of the finite coil length was taken away by choosing a larger number of loops in the radial direction at both ends of the solenoid. The configuration was as follows: (90, 90, 88, 88, 24, 24...24, 24, 88, 88, 90, 90), where the numbers indicate the number of loops at each site along the $z$-axis. An inner radius of the solenoid of 45 mm was used.

One might wonder whether the discrete nature of this approach causes unwanted effects. The model is based on infinitely thin wires, whereas in practice for example rectangular copper wire of $1 \times 3$ mm, coated with thin insulating material is used [28]. One could expect periodic dips in the magnetic field amplitude, of which the extrema are located right in between two loop sites. These effects would to some extend be artificial, but in reality something like this could also occur due to gaps between the wires (caused by the coating or air spaces). Plotting the magnetic field with Mathematica shows that no ripples due to the nature of the simulation occur in the magnetic field. To be more precise, they are not of a size that would influence the outcome of the simulations. In practice, the current distribution through the wires is much more uniform than that of the single
coil summation model, so no ‘ripple effect’ is to be expected.

4.6 Conclusions

In order to obtain a clock with a relative accuracy of \(10^{-12}\), a C-field with a homogeneity of \(4 \times 10^{-6}\) is needed. The temporal stability has to be of the same order of magnitude and is completely determined by the stability of the current through the solenoid, provided that the C-field region is magnetically shielded. Neither the shielding nor the current stability should be limiting factors. The spatial homogeneity was evaluated with two models. The current density integration approach shows that the desired theoretical homogeneity can be reached, in a much too small region (!), by applying rather small corrections to the basic shape of the solenoid. The outside notch is the best candidate for such a correction. The finite coil length effect is a problem if one tries to reach homogeneities below \(10^{-3}\). The single coil summation approach, however, yields the insight that even if that systematic problem is solved (which seems no trivial thing to do) ‘random’ construction errors of sizes that are not unimaginable cause inhomogeneities of the order of \(10^{-4}\). These inhomogeneities have to be corrected by shimming techniques. This implies a time-consuming procedure of which even the best results known would not satisfy our needs [28, 29].
Chapter 5

Stabilisation of an extended cavity diode laser at 1083 nm

Cooling and manipulating the helium atoms that are needed for the fountain clock is done by lasers emitting light at 1083 nm. An extended cavity diode laser system was acquired for this purpose. The system features an elaborate control system, including temperature stabilisation and diode-current control. Furthermore, the length of the laser cavity can be scanned, applying a voltage to a piezo crystal on which one of the cavity mirrors is mounted. The measured intensity of the laser beam is fed back into the system in order to control the laser frequency. In this way the laser can be frequency-stabilised on a spectral feature, such as a transmission or absorption peak of a Fabry-Perot interferometer or a Lamb dip in a gas cell. The main goal of the experimental work described in the following chapter was to lock the laser output frequency to an absorption line of metastable $^3\text{He}$.

5.1 Saturated absorption spectroscopy

Saturated absorption spectroscopy is a technique that is used to study the spectrum of lines corresponding to transitions in an atom or molecule. A beam of laser light is led through a gaseous medium containing the species that is to be examined. In our case, the medium is a gas of $^3\text{He}$, contained in a discharge cell, in which the atoms are excited from the ground state to the $2^3S_1$ metastable state (see appendix A for more details on the quantum structure of helium). The gas will absorb light of certain specific frequencies, corresponding to the energy difference between two quantum states of the atom. Within the frequency range of the diode laser are the 9 different transitions from the $2^3S_1$ to the $2^3P_J$ states (see figure 5.1).
Figure 5.1: Scheme in which the nomenclature of the spectral lines of $^3\text{He}$ at 1083 nm is defined.

Figure 5.2: The experimental setup used for saturated absorption spectroscopy, including the diode laser, polarising beam splitter, quarter wave plate, $^3\text{He}$ gas cell and the Fabry-Perot interferometer.

### 5.1.1 Experimental setup

The experimental setup used for saturated absorption spectroscopy is shown in figure 5.2. A positive lens is used to suppress the divergence of the beam. An uncoated glass plate (low reflectivity, LR) is used to send $\sim 4\%$ of the light through the polarising beam splitter. Since the laser beam is vertically polarised, the beam is deflected towards the gas cell. Behind the cell, the beam is reflected by a high reflectivity (HR) mirror. Before arriving at the beam splitter for the second time, the beam has passed through the quarter wave plate twice, so its polarisation has been rotated over $\pi/2$. Therefore, the beam is transmitted by the beam splitter and the light is sent onto a fast photodiode. Another part of the laser beam is sent through the FPI in order to have a frequency reference.
5.1 Saturated absorption spectroscopy

5.1.2 Doppler broadening and Lamb dips

The width of the velocity distribution of a gas at room temperature gives rise to Doppler broadening of the absorption profile. It is an example of an inhomogeneous line broadening process. As mentioned in chapter 2, the average speed of the helium atoms in the gas cell is of the order of thousands of meters per second. This means that the Doppler effect plays an important role in the absorption of photons by the fast-moving atoms. The absorption profile of a single atom, which displays the absorption intensity versus the laser light frequency, has a Lorentzian shape.[12]:

\[ L(\omega - \omega_0) = \frac{\Gamma/2\pi}{(\omega - \omega_0)^2 + (\Gamma/2)^2} \]  

(5.1)

with \( \Gamma/2\pi = 1/\tau \) the natural linewidth of the transition. Each atom with velocity \( \mathbf{v} \) sees the laser frequency \( \omega \) shifted to a frequency \( \omega' \):

\[ \omega' = \omega - \mathbf{k} \cdot \mathbf{v} \]  

(5.2)

where \( \mathbf{k} \) is the wave vector of the photon; \( |\mathbf{k}| = 2\pi/\lambda \). Assuming a Maxwell-Boltzmann velocity distribution, the probability for absorption of a photon in the frequency interval \((\nu, \nu + d\nu)\) by an atom with transition frequency \(\nu_0\) is given by \(g(\nu)d\nu\) and

\[ g(\nu) = \frac{c}{\nu_0} \sqrt{\frac{m}{2\pi k_B T}} \exp\left(-\frac{c^2 (\nu - \nu_0)^2}{2 k_B T/m}\right) \]  

(5.3)

This is approximately the shape of the Doppler broadened profile\(^1\).

The natural linewidth of the transitions in \(^3\)He at 1083 nm is about 1.6 MHz. The Doppler broadened profile has a width of \(~1\) GHz. The Doppler broadened profiles of lines that are close to each other tend to overlap to great extend, thus forming a broad absorption profile of a few GHz. It is impossible to stabilise a laser to a few MHz by locking to this profile. One likes to be able to detect the individual underlying transition lines, that are much narrower than the Doppler broadened profile.

Saturated absorption spectroscopy provides a means of recapturing the individual lines. A pump beam is used to saturate the transition. At the rear end of the gas cell, the beam is reflected and sent back along the path of the pump beam. The retroreflected beam is called the probe beam. The intensity of the probe beam is detected by a photodetector. Atoms that have a velocity component along the direction of the laser beams will experience a different Doppler shift for the pump and probe beams. Saturation of these

\(^1\)Actually, in order to calculate the Doppler broadened profile, the convolution of this expression with the Lorentzian associated with the natural linewidth should be taken, yielding a Voigt profile. And even the Lorentzian will be broadened by saturation broadening.
atoms by the pump beam will not influence the transmission of the probe beam. Only the atoms that have no velocity component along the laser beams, will be ‘transparent’ for the probe beam if they are saturated by the pump beam. This transparency causes dips in the Doppler profile of the transmission of the probe beam at the atomic transition frequencies. These dips are called Lamb dips. Around these Lamb dips, the probe beam is partially absorbed by the atoms that have the right velocity to be resonant with the laser frequency\(^2\).

Far from resonance, the light is not absorbed at all, because the velocity of the atoms is not high enough to compensate for the laser detuning by means of the Doppler shift. Additional to the Lamb dips, there are peaks that are not localised at transition frequencies, but are exactly halfway two transition lines that share a common ground (or excited) level: so-called cross-overs. These transmission lines are caused by atoms that have a velocity that makes them resonant with the pump beam for one transition and resonant with the probe beam for the other transition. The shared ground level is depleted by the pump beam, so no transitions from this level can be induced by the probe beam. States that share the same upper level generate also cross-overs; the upper level is populated by the pump beam and the probe beam induces stimulated emission.

5.1.3 Frequency reference: Fabry-Perot interferometer

In order to identify the peaks that are measured, one needs a frequency reference. Scanning of the laser frequency is done by applying a voltage to the piezo actuator. The length of the laser cavity is varied by an amount proportional to this scan voltage. Within a limited range (up to about 5 GHz for our laser system) the frequency of the laser is proportional to the length of the cavity. So-called mode hops occur when the length of the cavity is changed too much\(^3\). Besides the piezo actuator, also the current through the diode and the temperature of the laser can be varied to change the laser frequency.

The current can be varied much faster than the cavity length and is used to modulate the laser frequency when using the Pound-Drever locking method (see section 5.3). The current is also adapted during piezo scanning, in order get a ‘mode hop free scanning range’ that is as large as possible. This is accomplished via feed forward control. Mode hops can be recognised easily using an Fabry-Perot interferometer (FPI). An FPI is a cavity consisting of two confocal spherical mirrors of high reflectivity. Light is transmitted if

---

\(^2\)This is the Doppler profile region.

\(^3\)Of the band of frequencies generated by the diode, only those that are resonant with the laser cavity are amplified. At a mode hop, the laser suddenly starts lasing at a frequency that corresponds to a different mode: a different number of wavelengths fitting in the cavity.
the distance between the mirrors is equal to a multiple of a quarter of a wavelength (the light travels twice back and forth through the FPI, before overlapping with its own trajectory). The frequency distance between two transmission peaks is called the free spectral range of the interferometer. The free spectral range of the FPI that was used is 750 MHz. Mode hops are characterised by irregularities in the distance between the transmission peaks. The distance is in this case the distance in time (proportional to the distance in frequency, due to the linearity of the scan voltage and also proportional to the piezo deformation caused by this voltage).

5.2 Laser frequency control

As mentioned before, the laser frequency can be controlled by adjusting the laser temperature, diode current and piezo voltage. Diode lasers that consist only of a pn-junction have a linewidth of almost 100 MHz [32]. This linewidth can be reduced by forming a larger laser cavity, using the Littrow setup: the light emitted by the diode is collimated onto a reflection grating, from which the first order is sent back into the diode. Since the reflectivity of the grating is higher than the reflectivity of the diode front facet, the first order light experiences an extended cavity of a few centimeters, which has a smaller free spectral range and higher finesse than the diode cavity. This yields a linewidth of below 1 MHz\(^4\). A micrometer screw can be used to tilt the grating and coarsely tune the frequency that is reflected back to the diode. The laser can be tuned over a range of 7 nm (1800 GHz) in this way [32].

5.2.1 Temperature control

The temperature of the laser can be tuned with an accuracy of 0.1°C using a thermo-electric Peltier element. A thermistor temperature sensor is cemented into the base plate of the laser head to provide a feedback signal. The temperature of the laser head can be set in the range of 10°C to 50°C using a maximum current through the thermo-electric element of 2 A [32].

5.2.2 Scan control

The output frequency of the laser can be adjusted by the piezo scan control much faster than by adjusting the temperature. The high voltage version of the laser system provides possible output voltages between \(-75\) V and \(+150\) V. The scan control consists of three parts: the amplitude, offset and frequency tuning. The amplitude potentiometer controls the amplitude

\(^4\)Putting the laser on the slope of a FPI transmission peak (not locked), yielded a measured full range submillisecond variation of about 750 kHz, as measured on the oscilloscope.
of the triangular scan ramp that is sent to the piezo element (maximum peak-to-peak amplitude of 150 V). The offset of this periodic signal can be adjusted between 4 V and 150 V. The voltage is limited to a maximum of 150 V, which causes ‘clipped’ triangles when the offset is set above 75 V (at full amplitude). The frequency of the scan ramp can be varied over ten orders of magnitude.

The maximum mode hop free tuning range using the scan control was just above 4.5 GHz, measured with the transmission from the FPI. According to the manual, the piezo element extends linear with the applied voltage: 33 nm/V, and up to 5 μm [32]. Since the piezo actuator is integrated in the grating holder, the length of the cavity is not extended by the same amount: the effect of the piezo length consists of a combination of cavity extension and grating tilting.

5.2.3 Current control

Fastest adjustment of the laser frequency is accomplished by the current through the diode. The current can be changed by hand between the lasing threshold current of about 22 mA and the maximum current of 76 mA. The length of the pn-junction cavity depends on the current through it, and the frequency of the light that escapes from this minute cavity (ca. 100 μm) depends on the distance between the two ends of the diode. About 2 GHz can be scanned adjusting the current. The current can be phase-modulated with frequencies between 5 and 40 MHz, via a bias-T, located at the rear end of the laser head (see section 5.3). Another current input, the FET-current control, is able to adjust the diode current with a bandwidth of about 5 MHz. The FET-current control can be used as part of the feedback control loop that includes the PDD and PID regulator.

5.2.4 PID regulator

The PID regulator is an electronic circuit that creates a regulator output signal, using a photodiode or PDD error signal as input. First the set point reference voltage is subtracted from the input signal. If one tries to lock to a peak using the PDD, the setpoint has to be adjusted in such a way that the error signal goes through zero at the position of the peak.\(^5\)

PID stands for Proportional, Integral, Differential. The P-section generates a signal proportional to the input signal (often with a minus sign). The I-part integrates the input signal and the D-part gives a signal proportional to the rate of change of the input signal. The relative strength of all three components of the PID regulator can be adjusted using three trimming potentiometers. The resulting output can again be amplified before it is sent

\(^5\)This comes down to subtracting an offset that is caused by, for instance, the slope of the Doppler broadened profile.
5.3 Pound-Drever detection

The transfer function of the PID-circuit for the complex amplitude of the electronic signal $U$ can be written as [33]:

$$\frac{U_o}{U_i} = A_P \left[ 1 + i\omega\tau_D + \frac{1}{i\omega\tau_I} \right]$$

in this formula is $A_P$ the proportional gain and $\tau_D$ and $\tau_I$ are the time constants associated with the differential and integral part of the circuit respectively.

5.3 Pound-Drever detection

The Pound-Drever detector unit (PDD) can be used to lock to the top of spectral features. It generates an error signal using the probe beam intensity as input. If one tries to lock to the side of a fringe, the difference between the measured intensity of the probe beam and a reference voltage set by the user is fed to a PID regulator which converts this signal to an output voltage suitable for use as a feedback signal to the scan and current control. If the frequency of the laser changes in such a way that it becomes closer to the the top of the peak, the feedback signal has to be of opposite sign compared to when the laser frequency drifts away from the peak. The error signal is positive or negative respectively. But not only frequency changes cause a change in light intensity measured by a photodetector: laser intensity fluctuations have the same effect. In this case a feedback signal is generated where it is not needed and the laser might unlock. If one tries to lock to the top of the resonance, a frequency shift of the laser to either side of the peak gives the same sign for the error signal, which is obviously not what is needed to create the right feedback regulator signal. Frequency modulation of the laser output is used to overcome this problem.

The Pound-Drever unit generates an error signal that is independent of the absolute laser intensity. Instead, it uses a modulation of the diode current to construct a ‘derivative’ of the photodetector signal. This is done by a device that consists of two components: an internal oscillator and a phase detector. The internal oscillator provides a sinusoidal signal of (in our case) 20 MHz, which is added to the diode current signal via the bias-T. In the laser spectrum, a sideband structure is generated by the phase modulation of the diode current. Because of the high modulation frequency, an error signal can be constructed that is very sensitive of the peak that is to be locked to (see figure 5.5).

---

6One can compensate for intensity fluctuations by dividing the probe beam intensity by the intensity of the beam that is sent directly on a photodetector.
5.3.1 Pound-Drever theory

Modulated beam

The electric field of the phase modulated beam is given by [34]:

\[ E = E_0 e^{i(\omega t + \beta \sin \Omega t)} \] (5.5)

\( \Omega \) is the phase modulation frequency and \( \beta \) is called the modulation depth. The phase modulation generates a carrier beam which is accompanied symmetrically by intensity peaks at frequencies that are integer multiples of \( \Omega \) away from the carrier. The intensity of these sidebands decreases fast with increasing frequency distance to the carrier frequency for small \( \beta \). The electric field can be expanded in Bessel functions \( J_n \). Neglecting the higher order sidebands, we get:

\[ E = E_0 \left[ J_0(\beta) e^{i\omega t} + J_1(\beta) e^{i(\omega + \Omega)t} - J_1(\beta) e^{i(\omega - \Omega)t} \right] \] (5.6)

The asymmetric transmission\(^7\) of both first order sidebands when the carrier frequency is very close to the top of a resonance is used to determine an error signal. The medium (in our case the helium gas) has a transmission coefficient \( T(\omega) \), defined as the ratio of the transmitted and incoming electric field intensities: \( T(\omega) = E_t/E_i \). We can now easily write down the expression for the transmitted electric field, when the incoming field is that of equation (5.6).

\[ E_t = E_0 \left[ T(\omega) J_0(\beta) e^{i\omega t} + T(\omega + \Omega) J_1(\beta) e^{i(\omega + \Omega)t} - T(\omega - \Omega) J_1(\beta) e^{i(\omega - \Omega)t} \right] \] (5.7)

Transmitted beam

The power in the transmitted beam is what is measured by the photodetector. This power is simply \( P_t = |E_t|^2 \). If we define \( P_c \) and \( P_s \) as the power in the carrier beam and each of the sidebands respectively, we can write [34]

\[ P_t = P_c |T(\omega)|^2 + P_s \left\{ |T(\omega + \Omega)|^2 + |T(\omega - \Omega)|^2 \right\} \\
+ 2 \sqrt{P_c P_s} \left\{ \text{Re} [T(\omega)T^*(\omega + \Omega) - T^*(\omega)T(\omega - \Omega)] \cos \Omega t \right. \\
+ \left. \text{Im} [T(\omega)T^*(\omega + \Omega) - T^*(\omega)T(\omega - \Omega)] \sin \Omega t \right\} + (2\Omega \text{ terms}) \] (5.8)

Error signal

We are interested in the terms of frequency \( \Omega \); the value of the term

\[ T(\omega)T^*(\omega + \Omega) - T^*(\omega)T(\omega - \Omega) \] (5.9)

\(^7\)Or reflection, in the case of locking a laser to the top of a FPI resonance.
provides information about the frequency $\omega$ of the carrier beam. At high modulation frequencies (much larger than the width of the spectral feature), this term is purely imaginary and only the sine term in equation (5.8) survives. At modulation frequencies much lower than the width of the spectral feature, expression (5.9) becomes real and only the cosine term is important. The value of (5.9) is measured by mixing the photodiode signal with the modulation signal provided by the internal oscillator. The product of two sine waves is the following:

$$\sin(\Omega t)\sin(\Omega' t) = \frac{1}{2} \left[ \cos(\Omega - \Omega' t) - \cos(\Omega + \Omega' t) \right]$$  \hspace{1cm} (5.10)

Mixing of the photodiode signal and the internal oscillator yields a signal made up of a dc component and a $2\Omega$ component, since the frequency of the internal oscillator and the frequency difference between the carrier and the first-order sidebands are equal. The $2\Omega$ component is filtered out, using a low-pass filter. The relative phase of the two signals can be adjusted in such a way that the dc signal either becomes the prefactor of the cosine term in equation (5.8), or the prefactor of the sine term.

A theoretical error signal can be constructed if the frequency dependence of the transmission coefficient associated with the medium is known. In general, the transmission coefficient can be expressed as [35]:

$$T(\omega) = e^{-\delta(\omega) - i\phi(\omega)}$$  \hspace{1cm} (5.11)

where $\delta$ is the amplitude attenuation and $\phi$ is the optical phase shift of the electric field.

The shape of the spectral feature that is to be locked to is Lorentzian in the case of a Lamb dip [35]:

$$\delta(\omega) = \delta_{\text{peak}} \left( \frac{1}{R^2(\omega) + 1} \right)$$  \hspace{1cm} (5.12)

$$\phi(\omega) = \delta_{\text{peak}} \left( \frac{R(\omega)}{R^2(\omega) + 1} \right)$$  \hspace{1cm} (5.13)

where $\delta_{\text{peak}}$ is the peak attenuation at the line centre and $R(\omega)$ is a normalised frequency scale given by:

$$R(\omega) = \frac{\omega - \Omega}{\Delta\Omega/2}$$  \hspace{1cm} (5.14)

where $\Omega$ and $\Delta\Omega$ are the line centre frequency and FWHM of the Lorentzian, respectively.

The shape of the Pound-Drever error signal depends strongly on the ratio of the modulation frequency and the width of the spectral feature. In our case, the modulation frequency was 20 MHz and the FWHM of the Lamb dips was about 16 MHz. In order to obtain the characteristic Pound-Drever
Figure 5.3: Pound-Drever error signal: plot of the real part of equation (5.9), for a modulation frequency $\Omega = 5$ MHz and FWHM of the Lorentzian feature of 16 MHz.

Figure 5.4: Pound-Drever error signal: plot of the real part of equation (5.9), for a modulation frequency $\Omega = 20$ MHz and FWHM of the Lorentzian feature of 16 MHz. This is the situation of the experiment (compare figures 5.8 and 5.9).
Figure 5.5: Pound-Drever error signal: plot of the real part of equation (5.9), for a modulation frequency $\Omega = 400$ MHz and FWHM of the Lorentzian feature of 16 MHz. This is the characteristic Pound-Drever error signal.

error signal, the modulation frequency should be much larger than the width of the Lamb dips. In any case, either the real or imaginary part of expression (5.9) can be selected from the total photodetector signal, in order to serve as the error signal that is fed to the PID section of the system. The real part of (5.9) is plotted for a few ratios of the modulation frequency and the FWHM of the spectral feature.

A scheme of the setup that was used to lock to the Lamb dips is shown in figure 5.6.

In figures 5.3 to 5.5 theoretical error signals are plotted. Figure 5.3 represents the error signal that can be obtained with low frequency modulation. Figure 5.5 shows the characteristic Pound-Drever error signal for high modulation frequencies. This is the shape of the error signal one would like to achieve. Unfortunately, the modulation frequency of 20 MHz, provided by the local oscillator, was not high enough compared to the width of the Lamb dips to obtain an error signal like figure 5.5. The theoretical shape of the error signal shown in figure 5.4, calculated using the measured Lamb dip width of ca. 16 MHz, does very well resemble the measured error signal shown in figures 5.8 and 5.9.

5.3.2 Measured spectra

Figures 5.8 and 5.9 show measured signals plotted against time. As can be seen from figure 5.9b, the horizontal axis is also linearly proportional to the laser frequency (because the FPI transmission peaks are equidistant). A range of about 3 GHz is covered. The graphs show only four transition lines and two cross-overs. Other lines are beyond the mode hop free scanning range at these laser settings. One might think that in principle
the temperature of the laser could be used to scan towards the other lines, counting the mode hops that are encountered and subtracting from the cumulative frequency the frequency interval that is jumped back at a mode hop (∼3 GHz). When a wavelength meter was used to check the hopping behaviour of the diode laser, some highly irregular behaviour was observed. The laser frequency hopped over distances that were orders of magnitude larger than the simple mode hop distance of ca. 3 GHz. Moreover, the hops were not always opposite to the scanning direction (as is the case for regular mode hops). As a result of this behaviour, the grating of the laser had to be adjusted in order to reach higher frequencies and see for example the $C_8$ and $C_9$ peaks. The chaotic behaviour of the laser made it impossible to identify lines that have no neighbouring lines within the mode hop free scanning range, without a wavelength meter.

Figure 5.8b shows the Doppler broadened spectrum with Lamb dips. The Doppler profiles of all lines visible in this picture overlap, thus forming a profile of ca. 3 GHz. The error signal displayed above the Doppler profile is such that even the smallest cross-over can be locked to. Two features of these graphs deserve attention: the $C_5$ line is much stronger than could be expected from figure 5.7 and the Lamb dips are a factor of ten broader than the natural linewidth. This factor is too large to be explained by saturation broadening. It is probably due to the fact that the pump and probe beams do not overlap perfectly. A small angle between these beams causes a little Doppler broadening of the Lamb dips.

5.4 Laser stability: a beatnote experiment

In order to be able to say something about the stability of the laser output frequency when it is locked to an atomic resonance, a beat experiment was
5.4 Laser stability: a beatnote experiment

Figure 5.7: Scheme in which the relative positions and strength of the spectral lines of $^4$He and $^3$He at 1083 nm are shown. (Taken from [31]).

Figure 5.8: Graphs of a) regulator input signal against increasing frequency and b) corresponding probe beam transmission intensity, featuring Lamb dips. The Lamb dips are superimposed on the Doppler broadened profile formed by the Doppler profiles of the constituent lines.
performed. In such an experiment, the beams of two lasers are overlapped and sent onto a photodetector. The electric fields of both beams generate a temporal interference pattern if the beams have the same polarisation. The power measured by the photodetector can be written as:

\[ P(t) = \left| E_1 e^{i\omega_1 t} + E_2 e^{i\omega_2 t} \right|^2 = E_1^2 + E_2^2 + 2E_1E_2\cos(\omega_1 - \omega_2) \]  

(5.15)

The photodetector signal is used as input for a spectrum analyzer, which displays the intensity of the frequency components of the signal.

In our experiment a LNA laser with a known bandwidth of 150 kHz was detuned with an amount of 35 MHz in order to detect the cosine component of equation (5.15). The shape of the spectrum analyzer signal is the convolution of the shape of both individual beams. The intensity profile of both lasers is Lorentzian [37]. The full width at half maximum of the convolution of two Lorentz profiles is the sum of the FWHM’s of these Lorentz profiles.

Results

The output frequency of the laser changes significantly, compared to the width of the laser, over larger timescales within the time that the spectrum analyzer needs to scan over the appropriate frequency range (2.5 ms).
5.4 Laser stability: a beatnote experiment

Figure 5.10: Spectra of the beatnote signal of the extended cavity diode laser and the LNA laser. The spectrum analyzer performed a single sweep that takes 2.5 ms. The position of the beatnote changes during the sweep. The width of the signal depends on the direction in which the laser frequency changes during the sweep.

Therefore, the spectrum analyzer cannot be used to determine the width of the laser on millisecond timescale. The result of ‘single-sweep’ measurements are shown in figure 5.10. One can see that both width and position of each peak vary significantly between measurements. Averaging has to be performed over a series of measurements in order to be able to say something about the stability of the laser over longer timescales.

The spectra that are shown in figure 5.11 can be used to determine the stability of the laser on a timescale of a few minutes. It can clearly be seen that the FWHM of the averaged profile differs from one measurement to the other, as well as the exact position of the centre of the beatnote. Both the LNA and the diode laser had been out of lock in between measurements. The width of the diode laser seems to depend on the exact settings of the control system. Since the beat signal is displayed on a logarithmic scale, the half maximum of the signal is 3 dBm below the peak value. The FWHM of the diode laser that can be found from the displayed measurement are 480, 430, 410 and 370 kHz respectively, where the FWHM of the LNA laser is taken to be 150 kHz.
Figure 5.11: Spectra of the beatnote signal of the extended cavity diode laser and the LNA laser. The spectrum analyzer performed 5000 sweeps of 2.5 ms with a dead time of 50 ms in between. The measurements were averaged. The data was taken in a continuous series of 6 minute-measurements.
Chapter 6

Conclusions

The optimal shape of the microwave cavity in which the rf magnetic field is contained is circular cylindrical. The size of the cavity has to be such that the TE_{011} mode is resonant with the atomic hyperfine transition. If the radius is set equal to the height of the cavity (yielding a convenient shape) both radius and height should be 3.7208 cm. The cavity cannot be constructed to such a precision that it is resonant with the atomic transition, but the resonance frequency of the cavity can be tuned by changing the temperature. The amplitude of the magnetic field directed along the vertical has to be about $10^{-5}$ G. This amplitude depends on the velocity of the atoms that are detected. The power of the rf field that is coupled into the cavity is about $10^{-11}$ Watt.

A magnetic field in the free flight region above the cavity must be present to break the degeneracy of the magnetic substates of the hyperfine transition. At the most favourable field magnitude of 803 G, this C-field must have a relative homogeneity of $4 \times 10^{-6}$. The C-field region should be shielded from external magnetic fields. Furthermore, the temporal stability of the current supply that generates the current through the magnet solenoid has to be, over the operating time of the clock, of the same order of magnitude as the spatial homogeneity. The spatial homogeneity is the limiting factor though. The relative homogeneity that is needed to build a clock with a fractional frequency accuracy of $10^{-12}$ seems impossible to achieve without extreme effort, and maybe not possible to achieve at all.

The fact that helium is a fermion and therefore has no hyperfine transitions that are to first order independent of the magnetic field at low field amplitude, implies that the requirements for the relative field homogeneity are much more stringent than for fields used in clocks based on bosonic atoms. The requirements for the absolute field homogeneity are of the same order of magnitude as for cesium fountain clock. The helium clock, however, has to operate at an anticrossing C-field of 803 G to avoid the first order field dependence of the Zeeman shift. Therefore, the relative homogeneity of the
C-field has to be much better and, unfortunately, the relative homogeneity is the relevant parameter that determines the quality (and feasibility) of a magnetic field.

The severeness of the magnetic field homogeneity requirement is even increased by the fact that the width of the velocity distribution of the cooled helium atoms is larger than for heavier atoms. Therefore the cloud of atoms will expand during the free flight and the number of atoms that can be detected below the interaction region is only a small fraction of number of atoms captured in the MOT. Selection of atoms within a certain range of the velocity distribution, thereby selecting atoms that have to some extend experienced the same magnetic field, will decrease the number of detected atoms so much that statistical noise will limit the stability of the clock. If no velocity selection takes place, inhomogeneous line broadening due to the field inhomogeneity will destroy the Ramsey fringe pattern. In either case it seems impossible to reach the desired $10^{-12}$ accuracy of the clock [14].

An extended cavity diode laser can be locked to atomic transition lines of metastable helium in a discharge gas cell, using a 20 MHz modulation of the diode current and a Pound-Drever detector. The laser was locked to Lamb dips in the transmission spectrum of helium gas in a saturated absorption setup. The spectral width of the laser at a timescale of minutes was of the order of half a MHz, as measured in a beat experiment. The Lamb dips were too broad to observe the characteristic Pound-Drever error signal at the modulation frequency used.
Appendix A

The quantum structure of $^3$He

A.1 Fine structure

The $^3$He atom has a nucleus consisting of two protons and a neutron, with two electrons electrically bound to the nucleus. Each of the nucleons has spin $1/2$, which results, for a nucleus in the ground state, in a total nuclear spin quantum number $I = 1/2$. The electronic state of $^3$He that is used in the fountain clock is given in spectroscopic notation as $n^{2S+1}L_J = 2^3S_1$. This state is formed by two electrons in s-orbitals ($l = 0$) and with parallel spins, such that $S = s_1 + s_2 = 1$ and $L = l_1 + l_2 = 0$. The total electronic angular momentum quantum number can take only one value: $J = |J| = |L + S| = 1$. The principal quantum number associated with one electron (in the lowest hydrogenic state) is $n = 1$. Because of the Pauli principle the other electron must be in an excited state, the lowest of which has $n = 2$. The $n$ in $n^{2S+1}L_J$ stands for the principal quantum number of the excited electron. The $2^3S_1$ state is called a metastable state because of its long lifetime due to the fact that this state cannot decay to the $1^1S_0$ ground state via an electric dipole transition. Metastable $2^3S_1$ atoms are produced by electron bombardment of ground state atoms. The lifetime of this state is 8000 seconds.

A.2 Hyperfine structure

Each given state, characterised by its electronic configuration, is split into several hyperfine states, when the relative orientation of electronic and nuclear angular momentum is taken into account. The total angular momentum quantum number $F = |F| = |J + I|$ and its projection on the quantisation axis $m_F$ are used to label the hyperfine states. Combination of
The quantum structure of $^3\text{He}$

Figure A.1: Zeeman splitting of the magnetic sublevels of the $2^3S_1$ state of $^3\text{He}$. The transition between states $|2\rangle$ and $|5\rangle$ is the transition that provides the frequency standard for the helium fountain clock. (Taken from [36]).

$J = 1$ and $I = 1/2$ leads to $F = 3/2$ ($m_F = \pm 3/2, \pm 1/2$) and $F = 1/2$ ($m_F = \pm 1/2$).

A.2.1 The hyperfine energy spectrum

The energies of the hyperfine levels of the $2^3S_1$ state of $^3\text{He}$ in a magnetic induction $B$ are approximately given by a Hamiltonian that consists of a hyperfine interaction term and a Zeeman interaction term\(^1\):

$$H = H_{hf} + H_Z = A\mathbf{I} \cdot \mathbf{J} + \mu_B(g_J \mathbf{J} + g_I \mathbf{I}) \cdot \mathbf{B} \quad (A.1)$$

The energies of the six hyperfine levels are labelled $|1\rangle$ to $|6\rangle$. $|1\rangle$ is the state with the highest energy for small $B$ and $|6\rangle$ has the lowest energy. This labelling is useful to identify the states, because for nonzero magnetic induction the quantum number $F$ is no longer a good quantum number. Still it is common practice to write down expressions for the energies (or rather transition frequencies) of the six states, using their zero-field ($F, m_F$) labels [36]:

$$W(3/2, \pm 3/2) = \pm (g_J + g_I/2)\mu_B B \quad (A.2)$$

\(^1\)Note that this hamiltonian assumes that the eigenvales of the angular momentum quantum operators are numbers (e.g. $I_z |m_i\rangle = m_i |m_i\rangle$). Other conventions use eigenvalues that are multiples of $\hbar$ or $\hbar^2$; the Hamiltonian is adapted accordingly.
\[ W(F, \pm 1/2) = E/2 \pm (g_J/2)\mu_B B + \]
\[ \frac{1}{2} \left\{ (-1)^{F-1/2} \left[ E^2 \pm \frac{2}{3} E(g_J - g_I)\mu_B B + (g_J - g_I)^2(\mu_B B)^2 \right]^{1/2} \right\} \quad (A.3) \]

in which \( g_J \) is the electronic angular momentum \( g \)-factor \( g_J = 2.00223738(14) \) and \( g_I \) is the \(^3\)He nuclear spin \( g \)-factor \( g_I = 2.3174823(7) \times 10^{-3} \). \( E \) is the zero field hyperfine splitting energy of 6739.701177(16) MHz [36].
Appendix B

Electrodynamics

B.1 Electromagnetic fields

If we define an electric field-vector $\mathbf{E}$ and an electric polarisation $\mathbf{P}$ that is induced in a medium by the electric field:

$$\mathbf{P} = \epsilon_0 \chi_e \mathbf{E} \quad \text{(B.1)}$$

an associated auxiliary field $\mathbf{D}$ (the electric displacement) can be defined:

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} = \epsilon_0 (1 + \chi_e) \mathbf{E} = \epsilon_0 \epsilon_r \mathbf{E} \quad \text{(B.2)}$$

In the above equations $\epsilon_0$ is the electric permittivity of vacuum and $\chi_e$ is called the electric susceptibility of the medium. $\epsilon_r$ is the dielectric constant of the material and is often written without subscript$^1$.

Likewise we can define the magnetic field $\mathbf{H}$ and the magnetic polarisation or magnetisation $\mathbf{M}$ that is induced by it:

$$\mathbf{M} = \mu_0 \chi_m \mathbf{H} \quad \text{(B.3)}$$

Analogously to the electric situation we define the magnetic induction $\mathbf{B}$, although this field stands in some respects on the same footing as the electric field in stead of the electric displacement.$^2$

$$\mathbf{B} = \mu_0 \mathbf{H} + \mathbf{M} = \mu_0 (1 + \chi_m) \mathbf{H} = \mu_0 \mu_r \mathbf{H} \quad \text{(B.4)}$$

$^1$Many authors use the definition $\epsilon = \epsilon_0 \epsilon_r$, which can be confusing; I will only use $\epsilon$ (and $\mu$) as a dimensionless parameter.

$^2$e.g. $\mathbf{E}$ and $\mathbf{B}$ are both derived from the vector and scalar potentials $\mathbf{A}$ and $\Phi$:

$$\mathbf{E} = -\nabla \Phi - \frac{\partial \mathbf{A}}{\partial t}$$

$$\mathbf{B} = \nabla \times \mathbf{A}$$

Moreover, $\mathbf{E}$ and $\mathbf{B}$ appear on equal footing in the Lorentz Force Law:

$$\mathbf{F} = \mathbf{E} + \mathbf{v} \times \mathbf{B}$$
We now have a magnetic permeability of vacuum $\mu_0$ and relative permeability $\mu_\varepsilon$. The dimensionless parameter $\chi_m$ is called magnetic susceptibility.

Both the electric and magnetic susceptibilities are in principle dimensionless complex tensors [24]. The polarisation can thus have a different direction than the electric field. In the case of an isotropic medium, the tensor becomes a scalar (which is just a tensor of rank 1). If the medium is lossless, the imaginary parts of the susceptibilities vanish.\(^3\)

### B.2 Maxwell equations

The electromagnetic fields that exist in the medium inside the cavity are described by the Maxwell equations. In the MKSA system of units these equations read:

\[
\nabla \cdot \mathbf{D} = \rho_f \\
\nabla \times \mathbf{H} = \mathbf{J}_f + \frac{\partial \mathbf{D}}{\partial t} \\
\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \\
\n\nabla \cdot \mathbf{B} = 0
\]

in which $\rho_f$ is the free electric charge density in the medium and $\mathbf{J}_f$ is the free electric current density that is caused by the electric field if free charges are present in the medium:

\[
\mathbf{J}_f = \sigma \mathbf{E}
\]

in which $\sigma$ is called the conductivity of a medium.

### B.3 Electromagnetic waves

We can use the Maxwell equations to obtain equations for the electric and magnetic fields that relate the second derivatives of these fields with respect to space and time [24]. If we consider the case of an isotropic medium, such that the associated electric susceptibility is a scalar and assuming no net charge and current densities are present ($\rho_f = 0$, $\mathbf{J}_f = 0$), we can take the curl of equation (B.7) and get:

\[
\nabla \times \nabla \times \mathbf{E} = -\frac{\partial}{\partial t}(\nabla \times \mathbf{B})
\]  

\(^3\)The same arguments hold for the conductivity $\sigma$ of section B.2
Using an appropriate vector identity\(^4\) on the left-hand side of this equation and subsequently applying equations (B.5) and (B.2), we obtain \(-\nabla^2\mathbf{E}\) on the left-hand side. On the right-hand side we replace \(\mathbf{B}\) by \(\mu_0\mu\mathbf{H}\) and then use equation (B.6). In the resulting expression we write \(\mathbf{D} = \epsilon_0\epsilon\mathbf{E}\) in order to obtain a wave equation for \(\mathbf{E}\):

\[
\nabla^2\mathbf{E} = \epsilon_0\epsilon\mu_0\mu \frac{\partial^2\mathbf{E}}{\partial t^2}
\]

(B.11)

Similar manipulations yield the wave equation for the magnetic field:

\[
\nabla^2\mathbf{H} = \epsilon_0\epsilon\mu_0\mu \frac{\partial^2\mathbf{H}}{\partial t^2}
\]

(B.12)

Equations (B.11) and (B.12) are wave equations for waves travelling with velocity \(1/\sqrt{\epsilon_0\epsilon\mu_0\mu}\), if we consider fields in a vacuum we have \(\epsilon = 1\) and \(\mu = 1\) and the velocity becomes the speed of light in vacuum:

\[
c = 1/\sqrt{\epsilon_0\mu_0}
\]

(B.13)

Solutions of the wave equations are plane waves, travelling, for instance, in the \(z\)-direction in a cartesian coordinate system (this choice is of course arbitrary, we can simply define the coordinate system such that the \(z\)-direction coincides with the direction the wave is travelling). These solutions are then of the form:

\[
\mathbf{E} = \mathbf{E}_0 e^{-i(\omega t \pm \beta z)}
\]

(B.14)

\(\omega\) is the \textit{angular frequency} of the wave and \(\beta\) is called the \textit{phase constant}. If we substitute the electric field (B.14) into the wave equation (B.11) we obtain an expression for the phase constant:

\[
\beta^2 = \omega^2\epsilon_0\epsilon\mu_0\mu
\]

(B.15)

The phase constant \(\beta\) is related to the wavelength \(\lambda\) as:

\[
\lambda = 2\pi/\beta = \frac{2\pi}{\omega \sqrt{\epsilon_0\epsilon\mu_0\mu}} = \frac{c}{\nu \sqrt{\epsilon\mu}}
\]

(B.16)

\(^4\)For all vectors \(\mathbf{V}\) the following relation holds:

\[
\nabla \times \nabla \times \mathbf{V} = \nabla(\nabla \cdot \mathbf{V}) - \nabla^2\mathbf{V}
\]
Appendix C

Notches

In section 4.5.1, the Taylor expansion (4.9) for the inside and Helmholtz notches is given. The error coefficients that are used in equations (4.9) and (C.7) are given by the following equations [30]:

\[ C_1 = \frac{1}{1 + \beta^2} \quad C_2 = \frac{\beta^2}{1 + \beta^2} \quad C_3 = \frac{\alpha^2}{\alpha^2 + \beta^2} \quad C_4 = \frac{\beta^2}{\alpha^2 + \beta^2} \]  
(C.1)

\[ F = \frac{4\pi}{10} \beta \left( \sinh^{-1} \frac{\alpha}{\beta} - \sinh^{-1} \frac{1}{\beta} \right) \]  
(C.2)

\[ FE_2(\alpha, \beta) = \frac{4\pi}{10} \frac{1}{2\beta} (C_{3/2}^1 - C_{3/2}^3) \]  
(C.3)

\[ FE_4(\alpha, \beta) = \frac{4\pi}{10} \frac{1}{24\beta^4} [C_{1/2}^1 (2 + 3C_2^2 + 15C_2^4) - C_{3/2}^3 (2 + 3C_4^2 + 15C_4^4)] \]  
(C.4)

\[ FE_6(\alpha, \beta) = \frac{4\pi}{10} \frac{1}{240\beta^6} [C_{1/2}^1 (8 + 12C_2^2 + 15C_2^4 - 70C_2^6 + 315C_2^8) \quad (C.5) \]

\[ -C_{3/2}^3 (8 + 12C_4^2 + 15C_4^4 - 70C_4^6 + 315C_4^8)] \]  
(C.6)

Figure C.1: Contour plot that shows the parameters of the outside notch for which the second order term (dashed line) and the fourth order term (full line) in the Taylor expansion (C.7) vanishes. The intersection of both curves gives the desired values of \( \alpha_c \) and \( \beta_c \) for the 6th order outside notch.
The expansion that can be applied to the outside notch reads [30]:

\[
H_z = j \lambda a_2 \left\{ \frac{F_0(\alpha, \beta)}{\alpha} - \frac{F_c(\alpha_c, \beta_c)}{\alpha_c} + \left[ F_0 E_2(\alpha, \beta) \alpha - F_c E_2(\alpha_c, \beta_c) \alpha_c \right] \left( \frac{z}{a_2} \right)^2 \\
+ \left[ F_0 E_4(\alpha, \beta) \alpha^3 - F_c E_4(\alpha_c, \beta_c) \alpha_c^3 \right] \left( \frac{z}{a_2} \right)^4 + \ldots \right\} \tag{C.7}
\]
Bibliography


[14] Roland Stas (private communication).


