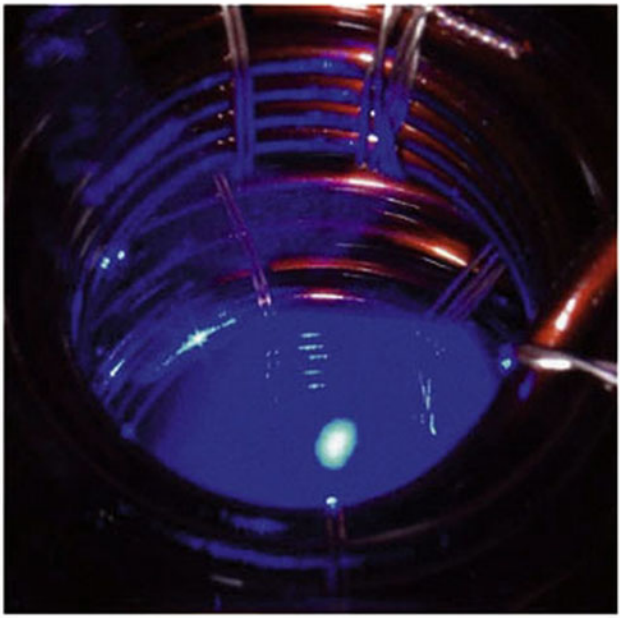
Fritz Riehle

WILEY-VCH

Frequency Standards

Basics and Applications



Fritz Riehle

Frequency Standards

Basics and Applications

WILEY-VCH Verlag GmbH Co. KGaA

Fritz Riehle

Frequency Standards

Basics and Applications

Fritz Riehle

Frequency Standards

Basics and Applications

WILEY-VCH Verlag GmbH Co. KGaA

Dr. Fritz Riehle
Physikalisch-Technische Bundesanstalt, Braunschweig
Germany
e-mail: fritz.riehle@ptb.de

Cover picture: Magneto-optical trap with 10 millions laser cooled calcium atoms in an optical frequency standard. Physikalisch-Technische Bundesanstalt, Braunschweig

Library of Congress Card No.: applied for

A catalogue record for this book is available from the British Library.

Bibliographic information published by Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at http://dnb.ddb.de

Printed in the Federal Republic of Germany.
Printed on acid-free paper.

Printing betz-druck GmbH, Darmstadt
Bookbinding J. Schäffer GmbH & Co. KG, Grünstadt

ISBN 3-527-40230-6

To the memory of my parents

and to Hildegard and Ruth

Contents

Preface

The contributions of accurate time and frequency measurements to global trade, traffic and most sub-fields of technology and science, can hardly be overestimated. The availability of stable sources with accurately known frequencies is prerequisite to the operation of world-wide digital data networks and to accurate satellite positioning, to name only two examples. Accurate frequency measurements currently give the strongest bounds on the validity of fundamental theories. Frequency standards are intimately connected with developments in all of these and many other fields as they allow one to build the most accurate clocks and to combine the measurements, taken at different times and in different locations, into a common system.

The rapid development in these fields produces new knowledge and insight with breathtaking speed. This book is devoted to the basics and applications of frequency standards. Most of the material relevant to frequency standards is scattered in excellent books, review articles, or in scientific journals for use in the fields of electrical engineering, physics, metrology, astronomy, or others. In most cases such a treatise focusses on the specific applications, needs, and notations of the particular sub-field and often it is written for specialists. The present book is meant to serve a broader community of readers. It addresses both graduate students and practising engineers or physicists interested in a general and introductory actual view of a rapidly evolving field. The volume evolved from courses for graduate students given by the author at the universities of Hannover and Konstanz. In particular, the monograph aims to serve several purposes.

First, the book reviews the basic concepts of frequency standards from the microwave to the optical regime in a unified picture to be applied to the different areas. It includes selected topics from mechanics, atomic and solid state physics, optics, and methods of servo control. If possible, the topics which are commonly regarded as complicated, e.g., the principles and consequences of the theory of relativity, start with a simple physical description. The subject is then developed to the required level for an adequate understanding within the scope of this book.

Second, the realisation of commonly used components like oscillators or macroscopic and atomic frequency references, is discussed. Emphasis is laid not only on the understanding of basic principles and their applications but also on practical examples. Some of the subjects treated here may be of interest primarily to the more specialised reader. In these cases, for the sake of conciseness, the reader is supplied with an evaluated list of references addressing the subject in necessary detail.

Third, the book should provide the reader with a sufficiently detailed description of the most important frequency standards such as, e.g., the rubidium clock, the hydrogen maser, the caesium atomic clock, ion traps or frequency-stabilised lasers. The criteria for the “impor-

tance" of a frequency standard include their previous, current, and future impact on science and technology. Apart from record-breaking primary clocks our interest also focusses on tiny, cheap, and easy-to-handle standards as well as on systems that utilise synchronised clocks, e.g., in Global Navigation Satellite Systems.

Fourth, the book presents various applications of frequency standards in contemporary high-technology areas, at the forefront of basic research, in metrology, or for the quest for most accurate clocks. Even though it is possible only to a limited extent to predict future technical evolution on larger time scales, some likely developments will be outlined. The principal limits set by fundamental principles will be explored to enable the reader to understand the concepts now discussed and to reach or circumvent these limitations. Finally, apart from the aspect of providing a reference for students, engineers, and researchers the book is also meant to allow the reader to have intellectual fun and enjoyment on this guided walk through physics and technology.

Chapter 1 reviews the basic glossary and gives a brief history of the development of clocks. Chapters 2 and 3 deal with the characterisation of ideal and real oscillators. In Chapter 4 the properties of macroscopic and in Chapter 5 that of microscopic, i.e., atomic and molecular frequency references, are investigated. The most important methods for preparation and interrogation of the latter are given in Chapter 6. Particular examples of frequency standards from the microwave to the optical domain are treated in Chapters 7 to 10, emphasising their peculiarities and different working areas together with their main applications. Chapter 11 addresses selected principles and methods of measuring optical frequencies relevant for the most evolved current and future frequency standards. The measurement of time as a particular application of frequency standards is treated in Chapter 12. The remainder of the book is devoted to special applications and to the basic limits.

I would like to thank all colleagues for continuous help with useful discussions and for supporting me with all kinds of information and figures. I am thankful to the team of Wiley–VCH for their patience and help and to Hildegard for her permanent encouragement and for helping me with the figures and references. I am particularly grateful to A. Bauch, T. Binnewies, C. Degenhardt, J. Helmcke, P. Hetzel, H. Knöckel, E. Peik, D. Piester, J. Stenger, U. Sterr, Ch. Tamm, H. Telle, S. Weyers, and R. Wynands for careful reading parts of the manuscript. These colleagues are, however, not responsible for any deficiencies or the fact that particular topics in this book may require more patience and labour as adequate in order to be understood. Furthermore, as in any frequency standard, feedback is necessary and highly welcome to eliminate errors or to suggest better approaches for the benefit of future readers.

Fritz Riehle
(fritz.riehle@ptb.de)
Braunschweig
June 2004

1 Introduction

1.1 Features of Frequency Standards and Clocks

Of all measurement quantities, frequency represents the one that can be determined with by far the highest degree of accuracy. The progress in frequency measurements achieved in the past allowed one to perform measurements of other physical and technical quantities with unprecedented precision, whenever they could be traced back to a frequency measurement. It is now possible to measure frequencies that are accurate to better than 1 part in 10^{15}. In order to compare and link the results to those that are obtained in different fields, at different locations, or at different times, a common base for the frequency measurements is necessary. Frequency standards are devices which are capable of producing stable and well known frequencies with a given accuracy and, hence, provide the necessary references over the huge range of frequencies (Fig. 1.1) of interest for science and technology. Frequency standards link the different areas by using a common unit, the hertz. As an example, consider two identical clocks whose

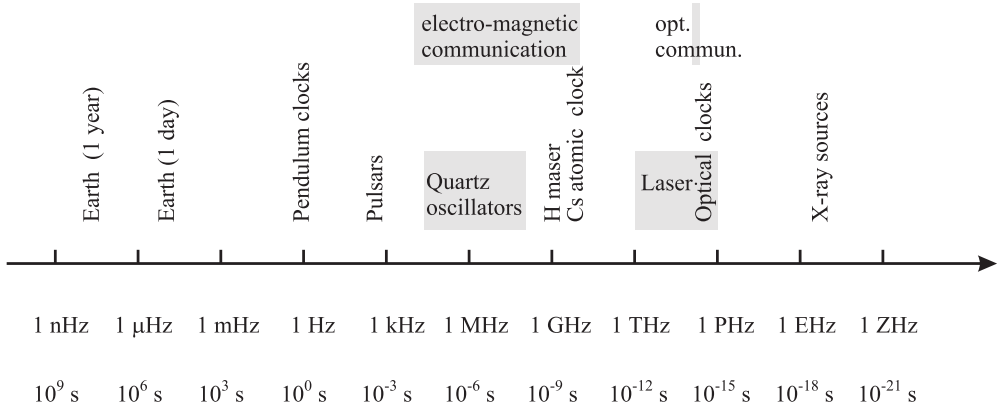


Figure 1.1: Frequency and corresponding time scale with clocks and relevant technical areas.

relative frequencies differ by 1×10^{-15}. Their readings would disagree by one second only after thirty million years. Apart from the important application to realise accurate clocks and time scales, frequency standards offer a wide range of applications due to the fact that numerous physical quantities can be determined very accurately from measurements of related frequencies. A prominent example of this is the measurement of the quantity *length*. Large distances are readily measured to a very high degree of accuracy by measurement of the time interval that a pulse of electromagnetic waves takes to traverse this distance. Radar guns used

by the police represent another example where the quantity of interest, i.e. the speed of a vehicle is determined by a time or frequency measurement. Other quantities like magnetic fields or electric voltages can be related directly to a frequency measurement using the field-dependent precession frequency of protons or using the Josephson effect, allowing for exceptionally high accuracies for the measurement of these quantities.

The progress in understanding and handling the results and inter-relationships of celestial mechanics, mechanics, solid-state physics and electronics, atomic physics, and optics has allowed one to master steadily increasing frequencies (Fig. 1.1) with correspondingly higher accuracy (Fig. 1.2). This evolution can be traced from the mechanical clocks (of resonant

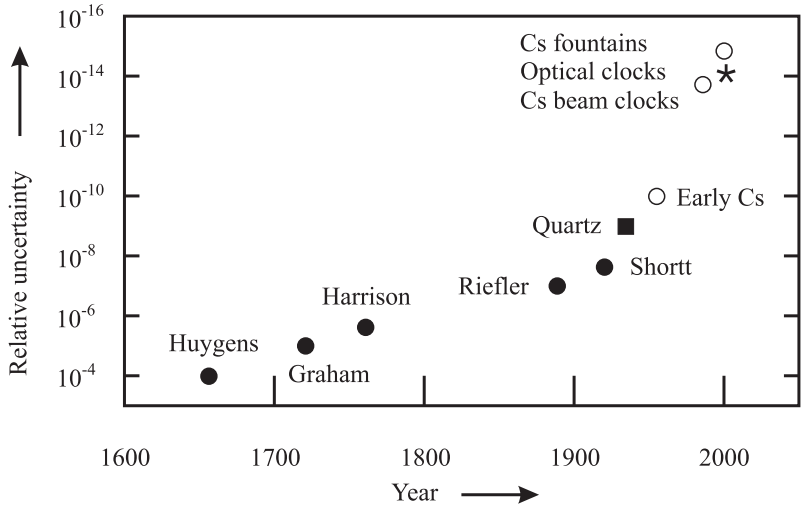


Figure 1.2: Relative uncertainty of different clocks. Mechanical pendulum clocks (full circles); quartz clock (full square); Cs atomic clocks (open circles); optical clocks (asterisk). For more details see Section 1.2.

frequencies $\nu_0 \approx 10^0\,\mathrm{Hz}$) via the quartz and radio transmitter technology ($10^3\,\mathrm{Hz} \leq \nu_0 \leq 10^8\,\mathrm{Hz}$), the microwave atomic clocks ($10^8\,\mathrm{Hz} \leq \nu_0 \leq 10^{10}\,\mathrm{Hz}$) to today's first optical clocks based on lasers ($\nu_0 \lesssim 10^{15}\,\mathrm{Hz}$). In parallel, present-day manufacturing technology with the development of smaller, more reliable, more powerful, and at the same time much cheaper electronic components, has extended the applications of frequency technology. The increasing use of quartz and radio controlled clocks, satellite based navigation for ships, aircraft and cars as well as the implementation of high-speed data networks would not have been possible without the parallel development of the corresponding oscillators, frequency standards, and synchronisation techniques.

Frequency standards are often characterised as active or passive devices. A "passive" frequency standard comprises a device or a material of particular sensitivity to a single frequency or a group of well defined frequencies (Fig. 1.3). Such a frequency reference may be based on macroscopic resonant devices like resonators (Section 4) or on microscopic quantum systems (Section 5) like an ensemble of atoms in an absorber cell. When interrogated by a suitable oscillator, the frequency dependence of the frequency reference may result in an absorption

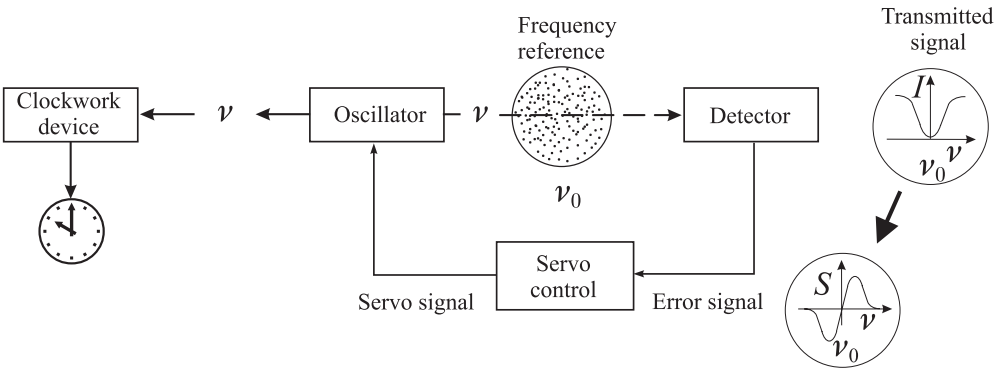


Figure 1.3: Schematics of frequency standard and clock.

line with a minimum of the transmission at the resonance frequency ν_0. From a symmetric absorption signal I an anti-symmetric error signal S may be derived that can be used in the servo-control system to generate a servo signal. The servo signal acting on the servo input of the oscillator is supposed to tune the frequency ν of the oscillator as close as possible to the frequency ν_0 of the reference. With a closed servo loop the frequency ν of the oscillator is "stabilised" or "locked" close to the reference frequency ν_0 and the device can be used as a frequency standard provided that ν is adequately known and stable.

In contrast to the passive standard an "active" standard is understood as a device where, e.g., an ensemble of excited atomic oscillators directly produces a signal with a given frequency determined by the properties of the atoms. The signal is highly coherent if a fraction of the emitted radiation is used to stimulate the emission of other excited atoms. Examples of active frequency standards include the active hydrogen maser (Section 8.1) or a gas laser like the He-Ne laser (Section 9.1).

A frequency standard can be used as a clock (Fig. 1.3) if the frequency is suitably divided in a clockwork device and displayed. As an example consider the case of a wrist watch where a quartz resonator (Section 4.1) defines the frequency of the oscillator at 32 768 Hz = 2^{15} Hz that is used with a divider to generate the pulses for a stepping motor that drives the second hand of the watch.

The specific requirements in different areas lead to a variety of different devices that are utilised as frequency standards. Despite the various different realisations of frequency standards for these different applications, two requirements are indispensable for any one of these devices. First, the frequency generated by the device has to be stable in time. The frequency, however, that is produced by a real device will in general vary to some extent. The variation may depend, e.g., on fluctuations of the ambient temperature, humidity, pressure, or on the operational conditions. We value a "good" standard by its capability to produce a stable frequency with only small variations.

A stable frequency source on its own, however, does not yet represent a frequency standard. It is furthermore necessary that the frequency ν is known in terms of absolute units. In the internationally adopted system of units (Systéme International: SI) the frequency is mea-

sured in units of Hertz representing the number of cycles in one second (1 Hz = 1/s). If the frequency of a particular stable device has been measured by comparing it to the frequency of another source that can be traced back to the frequency of a primary standard [1] used to *realise* the SI unit, our stable device then – and only then – represents a frequency standard.

After having fulfilled these two prerequisites, the device can be used to calibrate other stable oscillators as further secondary standards.

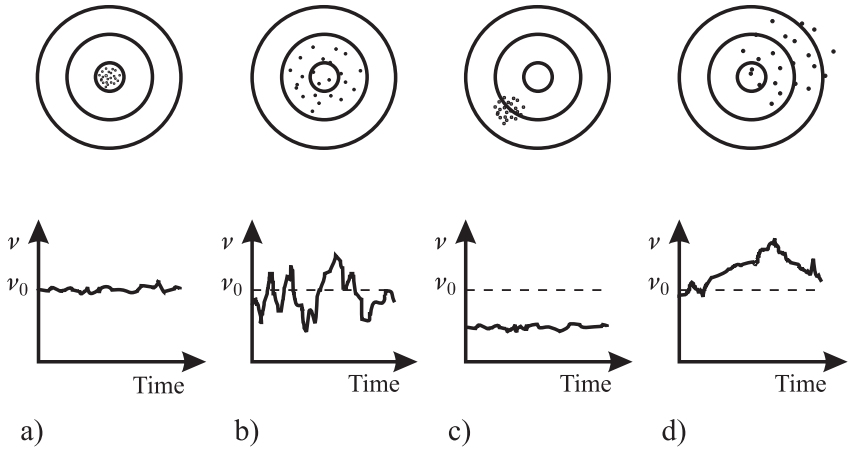


Figure 1.4: Bullet holes on a target (upper row) show four different patterns that are precise and accurate (a), not precise but accurate (b), precise but not accurate (c), not precise and not accurate (d). Correspondingly a frequency source (lower row) shows a frequency output that is stable and accurate (a), not stable but accurate (b), stable but not accurate (c), and not stable and not accurate (d).

There are certain terms like stability, precision, and accuracy that are often used to characterise the quality of a frequency standard. Some of those are nicely visualised in a picture used by Vig [2] who compared the temporal output of an oscillator with a marksman's sequence of bullet holes on a target (Fig. 1.4). The first figure from the left shows the results of a highly skilled marksman having a good gun at his disposal. All holes are positioned accurately in the centre with high precision from shot to shot. In a frequency source the sequence of firing bullets is replaced by consecutive measurements of the frequency ν, where the deviation of the frequency from the centre frequency ν_0 corresponds to the distance of each bullet hole from the centre of the target.[2] Such a stable and accurate frequency source may be used as a frequency standard. In the second picture of Fig. 1.4 the marks are scattered with lower precision but enclosing the centre accurately. The corresponding frequency source would suffer from reduced temporal stability but the mean frequency averaged over a longer period would be accurate. In the third picture all bullet holes are precisely located at a position off the centre. The corresponding frequency source would have a frequency offset from the desired

[1] A primary frequency standard is a frequency standard whose frequency correponds to the adopted definition of the second , with its specified accuracy achieved without external calibration of the device [1].

[2] The distances of bullet holes in the lower half plane are counted negative.

frequency ν_0. If this offset is stable in time the source can be used as a frequency standard provided that the offset is determined and subsequently corrected for. In the fourth picture most bullet holes are located to the right of the centre, maybe due to reduced mental concentration of the marksman. The corresponding oscillator produces a frequency being neither stable nor accurate and, hence, cannot be used as a frequency standard.

The accuracy and stability of the frequency source depicted in the third picture of Fig. 1.4 can be quantified by giving the deviation from the centre frequency and the scatter of the frequencies, respectively, in hertz. To compare completely different frequency standards the relative quantities "relative accuracy" ("relative stability", etc.) are used where the corresponding frequency deviation (frequency scatter) is divided by the centre frequency. As well as the terms accuracy, stability and precision the terms inaccuracy, instability and imprecision are also in current use and these allow one to characterise, e.g., a good standard with *low* inaccuracy by a *small* number corresponding to the *small* frequency deviation, whereas a *high* accuracy corresponds to a *small* frequency deviation.

The simple picture of a target of a marksman (Fig. 1.4) used to characterise the quality of a frequency standard is not adequate, however, in a number of very important cases. Consider, e.g., a standard which is believed to outperform all other available standards. Hence, there is no direct means to determine the accuracy with respect to a superior reference. This situation is equivalent to a plain target having neither a marked centre nor concentric rings. Shooting at the target, the precision of a gun or the marksman can still be determined but the accuracy cannot. It is, however, possible to "estimate" the uncertainty of a frequency standard similarly as it is done by measuring an *a priori* unknown measurand. There are now generally agreed procedures to determine the uncertainty in the Guide to the Expression of Uncertainty in Measurement (GUM) [3]. The specified uncertainty hence represents the "limits of the confidence interval of a measured or calculated quantity" [1] where the probability of the confidence limits should be specified. If the probability distribution is a Gaussian this is usually done by the standard deviation (1σ value) [3] corresponding to a confidence level of 68 %. For clarity we repeat here also the more exact definitions of accuracy as "the degree of conformity of a measured or calculated value to its definition" and precision as "the degree of mutual agreement among a series of individual measurements; often but not necessarily expressed by the standard deviation" [1].

1.2 Historical Perspective of Clocks and Frequency Standards

1.2.1 Nature's Clocks

The periodicity of the apparent movement of celestial bodies and the associated variations in daylight, seasons, or the tides at the seashore has governed all life on Earth from the very beginning. It seemed therefore obvious for mankind to group the relevant events and dates in chronological order by using the time intervals found in these periodicities as natural measures

[3] In cases where this confidence level is too low, expanded uncertainties with $k\,\sigma$ can be given, with, e.g., 95.5% ($k = 2$) or 99.7% ($k = 3$).

of time. Hence, the corresponding early calendars were based on days, months and years related to the standard frequencies of Earth's rotation around its polar axis (once a day), Earth's revolution around the Sun (once a year) and the monthly revolution of the moon around the Earth (once a month), respectively. The communication of a time interval between two or more parties had no ambiguity if all members referred to the same unit of time, e.g., the day, which then served as a natural standard of time. Similarly, a natural standard of frequency (one cycle per day) can be derived from such a natural clock. The calendar therefore allowed one to set up a time scale based on an agreed starting point and on the scale unit.[4] The establishment of a calendar was somewhat complicated by the fact that the ratios of the three above mentioned standard frequencies of revolution are not integers, as presently the tropical year [5] comprises 365.2422 days and the synodical month 29.5306 days.[6] Today's solar calendar with 365 days a year and a leap year with 366 days occurring every fourth year dates back to a Roman calendar introduced by Julius Caesar in the year 45 B.C.[7]

The use of Nature's clocks based on the movement of celestial bodies has two disadvantages. First, a good time scale requires that the scale unit must not vary with time. Arguments delivered by astronomy and geochronometry show that the ratio of Earth's orbital angular frequency around the sun and the angular frequency around its polar axis is not constant in time.[8] Second, as a result of the low revolution frequency of macroscopic celestial bodies the scale unit is in general too large for technical applications.[9]

1.2.2 Man-made Clocks and Frequency Standards

Consequently, during the time of the great civilisations of the Sumerians in the valley of Tigris and Euphrates and of the Egyptians, the time of the day was already divided into shorter sections and the calendars were supplemented by man-made clocks. A clock is a device that indicates equal increments of elapsed time. In the long time till the end of the Middle Ages the precursors of today's clocks included sundials, water clocks, or sand glasses with a variety of modifications. The latter clocks use water or sand flowing at a more or less constant rate and use the integrated quantity of moved substance to approximate a constant flow of time. Progress in clock making arose when oscillatory systems were employed that

[4] The set-up of a time scale, however, is by no means exclusively related to cyclic events. In particular, for larger periods of time, the exponential decay of some radioactive substances, e.g., of the carbon isotope ^{14}C allows one to infer the duration of an elapsed time interval from the determination of a continuously decreasing ratio ^{14}C/^{12}C.

[5] The tropical year is the time interval between two successive passages of the sun through the vernal equinox, i.e. the beginning of spring on the northern hemisphere.

[6] The synodical month is the time interval between two successive new moon events. The term "synode" meaning "gathering" refers to the new moon, when moon and sun gather together as viewed from the earth.

[7] The rule for the leap year was modified by Pope Gregor XIII in the year 1582 so that for year numbers being an integer multiple of 100, there is no leap year except for those years being an integer multiple of 400. According to this, the mean year in the Gregorian Calendar has 365.2425 days, close to its value given above.

[8] The growth of reef corals shows ridges comparable to the tree rings that have been interpreted as variations in the rate of carbonate secretion both with a daily and annual variation. The corresponding ratios of the ridges are explained by the fact that the year in the Jurassic (135 million years ago) had about 377 days [4].

[9] Rapidly spinning millisecond pulsars can represent "Nature's most stable clocks" [5], but their frequency is still too low for a number of today's requirements.

operate at a specific resonance frequency defined by the properties of the oscillatory system. If the oscillation frequency ν_0 of this system is known, its reciprocal defines a time increment $T = 1/\nu_0$. Hence, any time interval can be measured by counting the number of elapsed cycles and multiplying this number with the period of time T. Any device that produces a known frequency is called a frequency standard and, hence, can be used to set up a clock. To produce a good clock requires the design of a system where the oscillation frequency is not perturbed either by changes in the environment, by the operating conditions or by the clockwork.

1.2.2.1 Mechanical Clocks

In mechanical clocks, the clockwork fulfils two different tasks. Its first function is to measure and to display the frequency of the oscillator or the elapsed time. Secondly, it feeds back to the oscillator the energy that is required to sustain the oscillation. This energy from an external source is needed since any freely oscillating system is coupled to the environment and the dissipated energy will eventually cause the oscillating system to come to rest. In mechanical devices the energy flow is regulated by a so-called escapement whose function is to steer the clockwork with as little as possible back action onto the oscillator. From the early fourteenth century large mechanical clocks based on oscillating systems were used in the clock towers of Italian cathedrals. The energy for the clockwork was provided by weights that lose potential energy while descending in the gravitational potential of the Earth. These clocks were regulated by a so-called verge-and-foliot escapement which was based on a kind of torsion pendulum. Even though these clocks rested essentially on the same principles (later successfully used for much higher accuracies) their actual realisation made them very susceptible to friction in the clockwork and to the driving force. They are believed to have been accurate to about a quarter of an hour a day. The relative uncertainty of the frequency of the oscillator steering these clocks hence can be described by a fractional uncertainty of $\Delta T/T = \Delta\nu/\nu \approx 1\,\%$. The starting point of high-quality pendulum clocks is often traced back to an observation of the Italian researcher Galileo Galilei (1564 – 1642). Galilei found that the oscillation period of a pendulum for not too large excursions virtually does not depend on the excursion but rather is a function of the length of the pendulum. The first workable pendulum clock, however, was invented in 1656 by the Dutch physicist Christian Huygens. This clock is reported to have been accurate to a minute per day and later to better than ten seconds per day corresponding to $\Delta T/T \approx 10^{-4}$ (see Fig. 1.2). Huygens is also credited with the development of a balance-wheel-and-spring assembly. The pendulum clock was further improved by George Graham (1721) who used a compensation technique for the temperature dependent length of the pendulum arriving at an accuracy of one second per day ($\Delta T/T \approx 10^{-5}$).

The contribution of accurate clocks to the progress in traffic and traffic safety can be exemplified from the development of a marine chronometer by John Harrison in the year 1761. Based on a spring-and-balance-wheel escapement the clock was accurate to 0.2 seconds per day ($\Delta T/T \approx 2 - 3 \times 10^{-6}$) even in a rolling marine vessel. Harrison's chronometer for the first time solved the problem of how to accurately determine longitude during a journey [6]. Continuous improvements culminated in very stable pendulum clocks like the ones manufactured by Riefler in Germany at the end of the nineteenth century. Riefler clocks were stable to a hundredth of a second a day ($\Delta T/T \approx 10^{-7}$) and served as time-interval standards in

the newly established National Standards Institutes until about the twenties of the past century before being replaced by the Shortt clock. William H. Shortt in 1920 developed a clock with two synchronised pendulums. One pendulum, the master, swung as unperturbed as possible in an evacuated housing. The slave pendulum driving the clockwork device was synchronised via an electromagnetic linkage and in turn, every half a minute, initialised a gentle push to the master pendulum to compensate for the dissipated energy. The Shortt clocks kept time better than 2 milliseconds a day ($\Delta T/T \approx 2 \times 10^{-8}$) and to better than a second per year ($\Delta T/T \approx 3 \times 10^{-8}$).

1.2.2.2 Quartz Clocks

Around 1930 quartz oscillators (Section 4.1) oscillating at frequencies around 100 kHz, with auxiliary circuitry and temperature-control equipment, were used as standards of radio frequency and later replaced mechanical clocks for time measurement. The frequency of quartz clocks depends on the period of a suitable elastic oscillation of a carefully cut and prepared quartz crystal. The mechanical oscillation is coupled to electronically generated electric oscillation via the piezoelectric effect. Quartz oscillators drifted in frequency about 1 ms per day ($\Delta\nu/\nu \approx 10^{-8}$) [7] and, hence, did not represent a frequency standard unless calibrated. At this time, frequency calibration was derived from the difference in accurate measurements of mean solar time determined from astronomical observations.

The quartz oscillators (denoted as "Quartz" in Fig. 1.2) proved their superiority with respect to mechanical clocks and the rotating Earth at the latest when Scheibe and Adelsberger showed [7] that from the beginning of 1934 till mid 1935, the three quartz clocks of the Physikalisch-Technische Reichsanstalt, Germany all showed the same deviation from the siderial day. The researchers concluded that the apparent deviations resulted from a systematic error with the time determination of the astronomical institutes as a result of the variation of Earth's angular velocity.[10] Today, quartz oscillators are used in numerous applications and virtually all battery operated watches are based on quartz oscillators.

1.2.2.3 Microwave Atomic Clocks

Atomic clocks differ from mechanical clocks in such a way that they employ a quantum mechanical system as a "pendulum" where the oscillation frequency is related to the energy difference between two quantum states. These oscillators could be interrogated, i.e. coupled to a clockwork device only after coherent electromagnetic waves could be produced. Consequently, this development took place shortly after the development of the suitable radar and microwave technology in the 1940s. Detailed descriptions of the early history that led to the invention of atomic clocks are available from the researchers of that period (see e.g. [9–12]) and we can restrict ourselves here to briefly highlighting some of the breakthroughs. One of the earliest suggestions to build an atomic clock using magnetic resonance in an atomic beam was given by Isidor Rabi who received the Nobel prize in 1944 for the invention of this spectroscopic technique. The successful story of the Cs atomic clocks began between 1948 and

[10] T. Jones [8] points out that "The first indications of seasonal variations in the Earth's rotation were gleaned by the use of Shortt clocks."

1955 when several teams in the USA including the National Bureau of Standards (NBS, now National Institute of Standards and Technology, NIST) and in England at the National Physical Laboratory (NPL) developed atomic beam machines. They relied on Norman Ramsey's idea of using separated field excitation (Section 6.6) to achieve the desired small linewidth of the resonance. Essen and Parry at NPL (denoted as "Early Cs" in Fig. 1.2) operated the first laboratory Cs atomic frequency standard and measured the frequency of the Cs ground-state hyperfine transition [13, 14]. Soon after (1958) the first commercial Cs atomic clocks became available [15]. In the following decades a number of Cs laboratory frequency standards were developed all over the world with the accuracy of the best clocks improving roughly by an order of magnitude per decade. This development led to the re-definition of the second in 1967 when the 13^{th} General Conference on Weights and Measures (CGPM) defined the second as "the duration of 9 192 631 770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the caesium 133 atom". Two decades later the relative uncertainty of a caesium beam clock (e.g., CS2 in 1986 at the Physikalisch-Technische Bundesanstalt (PTB), Germany, denoted as "Cs beam cock" in Fig. 1.2) was already as low as 2.2×10^{-14} [16].

A new era of caesium clocks began when the prototype of an atomic Cs fountain was set up [17] at the Laboratoire Primaire du Temps and Fréquences (LPTF; now BNM–SYRTE) in Paris. In such clocks Cs atoms are laser cooled and follow a ballistic flight in the gravitational field for about one second. The long interaction time made possible by the methods of laser cooling (Section 6.3.1) leads to a reduced linewidth of the resonance curve. The low velocities of the caesium atoms allowed one to reduce several contributions that shift the frequency of the clock. Less than a decade after the first implementation, the relative uncertainty of fountain clocks was about 1×10^{-15} [18–20] (see "Cs fountain clock" in Fig. 1.2).

1.2.2.4 Optical Clocks and Outlook to the Future

As a conclusion of the historical overview one finds that the development of increasingly more accurate frequency standards was paralleled by an increased frequency of the employed oscillator. From the hertz regime of pendulum clocks via the megahertz regime of quartz oscillators to the gigahertz regime of microwave atomic clocks, the frequency of the oscillators has been increased by ten orders of magnitude. The higher frequency has several advantages. First, for a given linewidth $\Delta\nu$ of the absorption feature, the reciprocal of the relative linewidth, often referred to as the line quality factor

$$Q \equiv \nu_0/\Delta\nu, \tag{1.1}$$

increases. For a given capability to "split the line", i.e. to locate the centre of a resonance line, the frequency uncertainty is proportional to Q and hence, to the frequency of the interrogating oscillator. The second advantage of higher frequencies becomes clear if one considers two of the best pendulum clocks with the same frequency of about 1 Hz which differ by a second after a year ($\Delta\nu/\nu_0 \approx 3 \times 10^{-8}$). If both pendulums are swinging in phase it takes about half a year to detect the pendulums of the two clocks being out of phase by 180 °. With two clocks operating at a frequency near 10 GHz the same difference would show up after 1.6 ms. Hence, the investigation and the suppression of systematic effects that shift the frequency of a standard is greatly facilitated by the use of higher frequencies. One can therefore expect

further improvements by the use of optical frequency standards by as much as five orders of magnitude higher frequencies, compared with that of the microwave standards. The recent development of frequency dividers from the optical to the microwave domain (Section 11) also makes them available for optical clocks [21] which become competitive with the best microwave clocks (see "Optical clocks" in Fig. 1.2).

It can now be foreseen that several (mainly optical) frequency standards might be realised whose reproducibilities are superior to the best clocks based on Caesium. As long as the definition of the unit of time is based on the hyperfine transition in caesium, these standards will not be capable to realise the second or the hertz better than the best caesium clocks. However, they will serve as secondary standards and will allow more accurate frequency ratios and eventually may lead to a new definition of the unit of time.

2 Basics of Frequency Standards

2.1 Mathematical Description of Oscillations

A great variety of processes in nature and technology are each unique in the sense that the same event occurs periodically after a well defined time interval T. The height of the sea level shows a maximum roughly every twelve hours ($T \approx 12.4\,\mathrm{h}$). Similarly, the swing of a pendulum ($T \lesssim 1\,\mathrm{s}$), the electric voltage available at the wall socket ($T \approx 0.02\,\mathrm{s}$), the electric field strength of an FM radio transmitter ($T \approx 10^{-8}\,\mathrm{s}$) or of a light wave emitted by an atom ($T \approx 2 \times 10^{-15}\,\mathrm{s}$) represent periodic events. In each case a particular physical quantity $U(t)$, e.g., the height of the water above mean sea level or the voltage of the power line, performs oscillations.

2.1.1 Ideal and Real Harmonic Oscillators

Even though the time interval T and the corresponding frequency $\nu_0 \equiv 1/T$ differ markedly in the examples given, their oscillations are often described by an (ideal) harmonic oscillation

$$U(t) = U_0 \cos(\omega_0 t + \phi). \tag{2.1}$$

Given the amplitude U_0, the frequency

$$\nu_0 = \frac{\omega_0}{2\pi} \tag{2.2}$$

and the initial phase ϕ, the instantaneous value of the quantity of interest $U(t)$ of the oscillator is known at any time t. ω_0 is referred to as the angular frequency and $\varphi \equiv \omega_0 t + \phi$ as the instantaneous phase of the harmonic oscillator. The initial phase determines $U(t)$ for the (arbitrarily chosen) starting time at $t = 0$.

The harmonic oscillation (2.1) is the solution of a differential equation describing an ideal harmonic oscillator. As an example, consider the mechanical oscillator where a massive body is connected to a steel spring. If the spring is elongated by U from the equilibrium position there is a force trying to pull back the mass m. For a number of materials the restoring force $F(t)$ is to a good approximation proportional to the elongation

$$F(t) = -DU(t). \qquad \text{(Hooke's law)} \tag{2.3}$$

The constant D in Hooke's law (2.3) is determined by the stiffness of the spring which depends on the material and the dimensions of the spring. This force, on the other hand, accelerates

the mass with an acceleration $a(t) = d^2U(t)/dt^2 = F/m$. Equating both conditions for any instant of time t leads to the differential equation

$$\frac{d^2U(t)}{dt^2} + \omega_0^2\, U(t) = 0 \qquad \text{with} \qquad \omega_0 \equiv \sqrt{\frac{D}{m}}. \tag{2.4}$$

(2.1) is a solution of (2.4) as can be readily checked. The angular frequency ω_0 of the oscillator is determined by the material properties of the oscillator. In the case of the oscillating mass, the angular frequency ω_0 is given according to (2.4) by the mass m and the spring constant D.

If we had chosen the example of an electrical resonant circuit, comprising a capacitor of capacitance C and a coil of inductance L the frequency angular would be $\omega_0 = 1/(\sqrt{LC})$. In contrast, for an atomic oscillator the resonant frequency is determined by atomic properties. In the remainder of this chapter and in the next one we will not specify the properties of particular oscillators but rather deal with a more general description.

It is common to all oscillators that a certain amount of energy is needed to start the oscillation. In the case of a spring system potential energy is stored in the compressed spring elongated from equilibrium by U_0. When the system is left on its own, the spring will exert a force to the massive body and accelerate it. The velocity $v = dU(t)/dt$ of the body will increase and it will gain the kinetic energy

$$E_{\mathrm{kin}}(t) = \frac{1}{2}mv^2 = \frac{1}{2}m\left[\frac{d\,U(t)}{dt}\right]^2 = \frac{1}{2}m\omega_0^2U_0^2\sin^2(\omega_0 t + \phi) \tag{2.5}$$

where we have made use of (2.1). The kinetic energy of the oscillating system increases to a maximum value as long as there is a force acting on the body. This force vanishes at equilibrium, i.e. when $\sin^2(\omega_0 t + \phi) = 1$ and the total energy equals the maximum kinetic (or maximum potential) energy

$$E_{\mathrm{tot}} = \frac{1}{2}m\omega_0^2U_0^2 = \frac{1}{2}DU_0^2. \tag{2.6}$$

The proportionality between the energy [1] stored in the oscillatory motion and the square of the amplitude is a feature which is common to all oscillators.

Rather than using a cosine function to describe the harmonic oscillation of (2.1) we could also use a sine function. As is evident from $\cos\varphi = \sin(\varphi + \frac{\pi}{2})$ only the starting phase ϕ would change by $\pi/2$. More generally, each harmonic oscillation can be described as a

[1] The energy discussed here is the energy stored in the oscillation of an oscillator that has been switched on and that would be oscillating forever if no dissipative process would reduce this energy. It must not be mixed with the energy that can be extracted from a technical oscillator which uses another source of energy to sustain the oscillation. The voltage $U(t)$, for instance, present at the terminals of such an oscillator is capable of supplying a current $I(t)$ to a device of input resistance R. This current $I(t) = U(t)/R$ produces a temporally varying electrical power $P(t) = U(t)I(t) = U^2(t)/R = U_0^2/R\cos^2(\omega_0 t + \phi)$ at the external device. The mean power $\bar{P}$, i.e. the power integrated for one period $\int_0^T U_0^2/R\cos^2(\omega_0 t + \phi)dt = U_0^2/(2R)$ is also proportional to U_0^2, as well as the energy $E(t') = \int_0^{t'} P(t)dt = U_0^2/R\int_0^{t'}\cos^2(\omega_0 t + \phi)dt$ delivered by the oscillator within the time t'. In contrast to the energy stored in an undamped oscillator this energy $\int_0^{t'} P(t)dt$ increases linearly with time t'.

superposition of a sine function and a cosine function having the same frequency as follows

$$\begin{aligned} U(t) = U_0 \cos(\omega_0 t + \phi) &= U_0 \cos(\omega_0 t) \cos(\phi) - U_0 \sin(\omega_0 t) \sin(\phi) \\ &= U_{01} \cos(\omega_0 t) - U_{02} \sin(\omega_0 t) \end{aligned} \tag{2.7}$$

where we have used $\cos(\alpha+\beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta$. The two quantities $U_{01} = U_0 \cos\phi$ and $U_{02} = U_0 \sin\phi$ are termed quadrature amplitudes of the oscillation. As computations including sine and cosine functions can sometimes become awkward it is more convenient to describe the harmonic oscillation by a complex exponential using Euler's formula $\exp i\varphi = \cos\varphi + i \sin\varphi$. Then (2.1) can be replaced by

$$\begin{aligned} U(t) &= \Re\mathrm{e}\left\{U_0 e^{i(\omega_0 t+\phi)}\right\} = \Re\mathrm{e}\left\{\widetilde{U}_0 e^{i\omega_0 t}\right\} = \frac{\widetilde{U}_0 e^{i\omega_0 t} + \widetilde{U}_0^* e^{-i\omega_0 t}}{2} \\ &= \frac{1}{2}\left\{\widetilde{U}_0 e^{i\omega_0 t} + c.c.\right\}, \end{aligned} \tag{2.8}$$

with the complex phasor

$$\widetilde{U}_0 = U_0\, e^{i\phi} = U_{01} + i\, U_{02}. \tag{2.9}$$

The phasor $\widetilde{U}_0$ contains the modulus $U_0 = |U(t)|$ and the starting phase angle in a single complex number. Calculations using the complex representations of the oscillation take advantage of the simple rules for dealing with complex exponentials. Having obtained the final (complex) result one keeps only the real part.[2] Accordingly, there are different ways to represent the ideal harmonic oscillation of (2.8) graphically. To depict the oscillation in the time

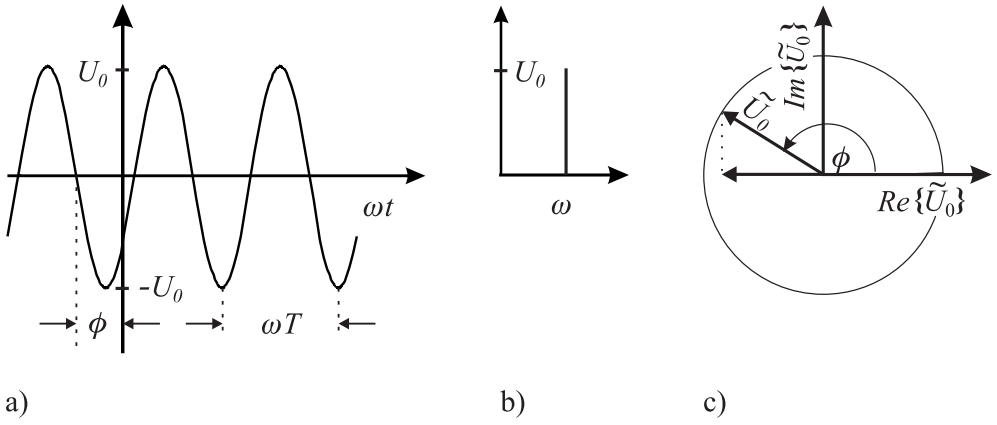


Figure 2.1: Ideal harmonic oscillator. a) Time-domain representation. b) Frequency-domain representation. c) Phasor representation.

[2] For simplicity, the operator $\Re\mathrm{e}$ is often not written in the course of the complex computations and the real part is taken only at the final result. Notice, however, that this procedure is only applicable in the case of linear operations as e.g. addition, multiplication with a number, integration or differentiation, but not in the case of non-linear operations. This can be seen in the case of the product of two complex numbers where obviously in general $\Re\mathrm{e}\,(A^2) \neq [\Re\mathrm{e}\,(A)]^2$.

domain (Fig. 2.1 a) it is necessary to know the initial phase ϕ, the amplitude U_0, and frequency $\nu_0 = 1/T$ of the oscillation. The oscillation in the frequency domain (Fig. 2.1 b) does not contain any information on the phase of the oscillator. If represented by a complex phasor, $\widetilde{U}_0 = U_0 \exp(i\phi)$ may be visualised in the complex plane (Argand diagram; Fig. 2.1 c) by a pointer of length U_0 that can be represented either in polar coordinates or in Cartesian coordinates. The initial phase is depicted as the angle ϕ between the real coordinate axis and the pointer. The phasor must not be mixed with the complex pointer $U_0 \exp[i(\omega_0 t + \phi)]$ that rotates counterclockwise [3] at constant angular velocity ω_0.

A specific property of the ideal harmonic oscillator is that we can predict its phase according to (2.1) starting from the initial conditions (phase, amplitude, and frequency) at any instant with any desired accuracy. For real oscillators used as examples above, these properties can be predicted only with an inherent uncertainty. For instance, the tidewaters do not always rise and fall to the same levels, but also show from time to time exceptionally high spring tides. In this case the amplitude of the oscillation resulting from the attraction of the moon is also "modulated" by the gravitational influence of the sun. In the example of the swinging pendulum the amplitude is constant only if the energy dissipated by friction is compensated for. Otherwise the amplitude of the swinging pendulum will die away similar to the amplitude of an oscillating atom emitting a wave train. In reality neither the amplitude nor the frequency of a real oscillator are truly constant. The long-term frequency variation may be very small as in the case of the ocean tides, where the angular velocity of the earth is decreasing gradually by friction processes induced by the tides of the waters and the solid earth but will become important after a large number of oscillations (see footnote 8 in Chapter 1). Apart from the natural modulations encountered in these two examples, the frequency of an oscillator may also be modulated on purpose. The frequency of the electromagnetic field produced by a FM (frequency modulated) transmitter is modulated to transmit speech and music. Basically, one refers to any temporal variation in the amplitude of an oscillator as amplitude modulation and the variation of its phase or frequency as phase modulation or frequency modulation, respectively. In the following we will investigate the processes of amplitude and phase modulation of an oscillator in more detail and we will develop the methods for the description.

For the oscillators relevant for frequency standards one may assume that the modulation represents only a small perturbation of the constant amplitude U_0 and of the phase $\omega_0 t$. An amplitude-modulated signal can then be written as

$$U(t) = U_0(t)\cos\varphi(t) = [U_0 + \Delta U_0(t)]\cos[\omega_0 t + \phi(t)]\,. \tag{2.10}$$

The instantaneous frequency

$$\nu(t) \equiv \frac{1}{2\pi}\frac{d\varphi(t)}{dt} = \frac{1}{2\pi}\frac{d}{dt}[2\pi\nu_0 t + \phi(t)] = \nu_0 + \frac{1}{2\pi}\frac{d\phi(t)}{dt} \tag{2.11}$$

differs from the frequency ν_0 of the ideal oscillator by

$$\Delta\nu(t) \equiv \frac{1}{2\pi}\frac{d\phi(t)}{dt}. \tag{2.12}$$

[3] Actually, there is an equivalent way to describe the oscillation mathematically by choosing a negative phase in (2.8), i.e. writing $U(t) = \Re\mathrm{e}\left\{\widetilde{U}_0 e^{-i\omega_0 t}\right\}$. From $e^{-i\phi} = \cos\phi - i\sin\phi$ it is clear that the pointer then rotates clockwise in the Argand diagram. As a consequence, one would have to change the sign of the imaginary part of several quantities. In the cases when this might lead to confusion we shall explicitly refer to this point again.

In the following we shall investigate amplitude modulation and phase modulation of an oscillator separately.

2.1.2 Amplitude Modulation

In general, the temporal variation of the amplitude $\Delta U_0(t)$ of a real oscillator may be very complicated and it will be not possible to describe the time dependence analytically. In this case the stochastic temporal behaviour may be described in terms of probability distributions. We will deal with these cases in Section 3. In this section we investigate deterministic modulations and consider two special cases of an amplitude modulation where one can give the explicit time dependence of $\Delta U_0(t)$. As examples we will consider a harmonic modulation and the case of an exponentially decreasing amplitude.

2.1.2.1 Spectrum of an Oscillator with Harmonically Modulated Amplitude

Let us assume that the amplitude varies by a pure sine or cosine function around the mean value U_0 with maximum deviation ΔU_0 and the modulation frequency $\nu_m = \omega_m/(2\pi)$ (Fig. 2.2 a). In technical applications the modulation frequency ν_m is usually much lower than

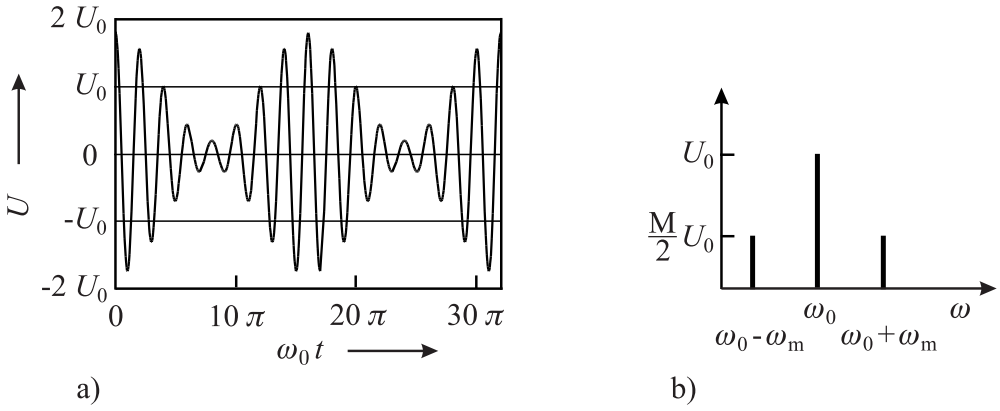


Figure 2.2: a) Time dependence of an amplitude modulated oscillation according to (2.13) using $\omega_m = \omega_0/8$ and $M = 0.8$. b) Frequency spectrum of the harmonically amplitude modulated oscillation according to (2.13) displayed in Fig. 2.2 a).

the frequency of the oscillation ν. The amplitude modulated (AM) oscillation is described by

$$\begin{aligned} U_{\mathrm{AM}}(t) &= (U_0 + \Delta U_0 \cos\ \omega_m t) \cos\ \omega_0 t \\ &= U_0(1 + M \cos \omega_m t) \cos \omega_0 t \end{aligned} \tag{2.13}$$

where

$$M \equiv \frac{\Delta U_0}{U_0} \tag{2.14}$$

is referred to as the modulation index of the amplitude modulation. Using the identity $\cos\alpha\ \cos\beta = \frac{1}{2}\cos(\alpha+\beta) + \frac{1}{2}\cos(\alpha-\beta)$ one writes (2.13) as

$$U_{\mathrm{AM}}(t) = U_0\left[\cos\omega_0 t + \frac{M}{2}\cos(\omega_0+\omega_m)t + \frac{M}{2}\cos(\omega_0-\omega_m)t\right]. \tag{2.15}$$

The frequency spectrum of the amplitude modulated oscillation (2.15) comprises three components (Fig. 2.2 b). The first term in square brackets of (2.15) represents the so called carrier, i.e. the component with the previously unmodulated angular frequency ω_0. The second and third terms are referred to as the high-frequency side band [4] and the low-frequency side band, respectively, representing the components generated by the modulation. The frequency of each side band is separated by the modulation frequency ω_m with respect to the frequency of the carrier. Both side bands are present with the same amplitude which is determined by the amplitude of the carrier and the modulation index.

One expects to obtain the same result as (2.15) if one uses the complex representation of the oscillation of (2.1)

$$\begin{aligned} U_{\mathrm{AM}}(t) &= U_0\Re\mathrm{e}\left\{[1+M\cos\ \omega_m t]\, e^{i\omega_0 t}\right\} \\ &= U_0\Re\mathrm{e}\left\{\left[1+\frac{M}{2}\left(e^{i\omega_m t}+e^{-i\omega_m t}\right)\right] e^{i\omega_0 t}\right\} \\ &= U_0\Re\mathrm{e}\left\{e^{i\omega_0 t}+\frac{M}{2}e^{i(\omega_0+\omega_m)t}+\frac{M}{2}e^{i(\omega_0-\omega_m)t}\right\}. \end{aligned} \tag{2.16}$$

In the phasor plot the harmonically amplitude modulated oscillation of the last line of (2.16) can be represented by three phasors in the complex plane. The phasor of the carrier is fixed on the real axis [5], whereas the phasor of the high-frequency side band is rotating counter-clockwise with the anguar frequency $\omega_m t$ with respect to the phasor of the carrier and the phasor of the low-frequency side band is rotating clockwise with $-\omega_m t$. The side band phasors have the length $M/2$. The influence of the amplitude modulation is defined by the sum of both side-band phasors. The phasor resulting from this sum is always parallel to the phasor of the carrier but changes its length and direction with the period of the modulation frequency. Consequently, the length of the phasor resulting from all three phasors and describing the harmonically modulated oscillation is also changed periodically (Fig. 2.3).

The power contained in the harmonically amplitude modulated oscillation is proportional to the amplitude squared (cf. footnote 1 on page 12) which in the case of a complex amplitude

[4] The phrase "band" refers to the more common case of amplitude modulation with a band of frequencies.

[5] In our case the phasor of the carrier is depicted on the real axis. Electrical engineers, however, prefer to represent the harmonic oscillation in contrast to (2.1) as a real sine function. The corresponding phasor is then rotated by $\exp(i\pi/2)$, i.e. by 90° and is pointing along the imaginary axis.

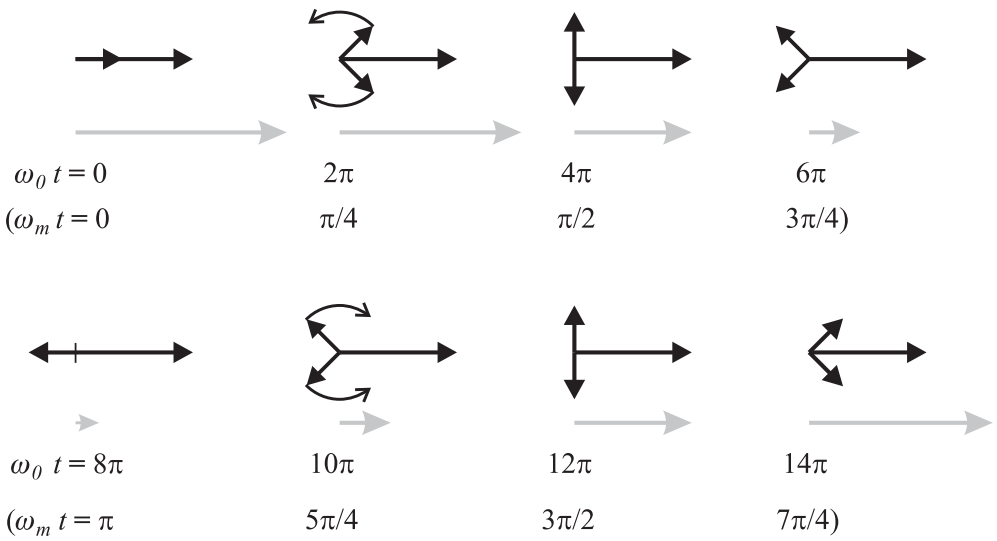


Figure 2.3: Phasor representation of an oscillation whose amplitude is modulated harmonically according to (2.16) and shown in Fig. 2.2 a) using $M = 0.8$ and $\omega_m = \omega_0/8$. The phasor of the high-frequency side band rotates with angular velocity ω_m and the phasor of the low-frequency side band rotates as -ω_m with respect to the carrier. The length of the phasor (grey arrow) resulting from the three single phasors varies periodically.

(2.16) has to be taken as the product of the amplitude of and its complex conjugate leading to

$$\begin{aligned} P_{\mathrm{AM}} &\propto U_0 \left[e^{i\omega_0 t} + \frac{M}{2} e^{i(\omega_0+\omega_m)t} + \frac{M}{2} e^{i(\omega_0-\omega_m)t} \right] \\ &\times U_0^* \left[e^{-i\omega_0 t} + \frac{M}{2} e^{-i(\omega_0+\omega_m)t} + \frac{M}{2} e^{-i(\omega_0-\omega_m)t} \right] \\ &= |U_0|^2 \left[1 + 2\frac{M}{2} e^{-i\omega_m t} + 2\frac{M}{2} e^{i\omega_m t} + 2\frac{M^2}{4} + 2\frac{M^2}{4} e^{2i\omega_m t} + 2\frac{M^2}{4} e^{-2i\omega_m t} \right] \\ &= |U_0|^2 \left[1 + \frac{M^2}{2} + 2M\cos\omega_m t + \frac{M^2}{2}\cos(2\omega_m t) \right]. \end{aligned} \tag{2.17}$$

Since the rapidly oscillating cosine terms average to zero for measurement times $t \gg 2\pi/\omega_m$ one obtains

$$P_{\mathrm{AM}} \propto |U_0|^2 \left[1 + \frac{M^2}{2} \right]. \tag{2.18}$$

Hence, the total power of the amplitude modulated oscillation is given by the power contained in the unmodulated carrier augmented by the power contained in both side bands.

Simple harmonic modulation of the amplitude of an oscillator results in two additional frequencies. One might therefore expect that more complicated forms of modulation will lead to a more complicated frequency spectrum comprising a larger number of side bands. In

the case of the harmonic AM the corresponding frequency spectrum was obtained by applying the simple rules for adding harmonic functions. The general procedure to derive the frequency spectrum, when the amplitude function is known in the time domain, is given by the Fourier transformation.

2.1.2.2 Fourier Transformation

The Fourier transformation makes use of a theorem of Jean Baptiste Joseph, Baron de Fourier (1768 – 1830). According to this theorem, any periodic function $f(t)$ characterised by the temporal period T can be represented unambiguously by a sum of harmonic functions defined by temporal periods T_i being integral sub-multiples of T, e.g. $T,\ T/2,\ T/3,\ T/4, \cdots$. This theorem can be stated in an equivalent way by saying that any periodic (time-) function $U(t)$ of angular frequency $\omega_g = 2\pi/T$ can be represented by a (finite or infinite) sum of sine and cosine terms having angular frequencies that are integral multiples of the basic frequency ω_g. The amplitudes of these so-called higher harmonics of the basic frequency represent the weights of the constituents necessary to synthesise the time function of interest. The time function of the harmonically amplitude modulated oscillator of Fig. 2.2 contains three single components of angular frequencies $\omega_0 - \omega_m$, ω_0 and $\omega_0 + \omega_m$ with the respective weights (amplitudes) $\frac{M}{2}U_0,\ U_0$ and $\frac{M}{2}U_0$. From a different point of view we can synthesise the temporal function depicted in Fig. 2.2 by these three purely harmonic functions. Generalisation of the Fourier series to non-periodic functions leads to a Fourier integral

$$U(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} A(\omega) e^{i\omega t} d\omega, \tag{2.19}$$

for any time function $U(t)$ represented by harmonic functions of angular frequencies ω. Similarly to the description of the harmonic oscillation, we use here the complex representation of the Fourier integral which is mathematically simpler to work with. The complex spectral function $A(\omega)$ gives the weights of all harmonic constituents (often called the Fourier components with the Fourier frequencies ω) contained in the time function $U(t)$. To determine the weight of a particular (complex) Fourier component we have to use the (complex) Fourier transformation

$$A(\omega) = \Re\mathrm{e}\, A(\omega) + i\Im\mathrm{m}\, A(\omega) = \mathcal{F}\{U(t)\} \equiv \int_{-\infty}^{\infty} U(t) e^{-i\omega t} dt. \tag{2.20}$$

Unfortunately, the Fourier transformation according to (2.20) and the inverse Fourier transformation according to (2.19) are not defined in a consistent way in the literature. Depending on the choice of the phase of the complex oscillation to be positive or negative (see footnote 3) the phase angles of the complex exponential functions of (2.20) and (2.19) interchange their positive and negative signs. Sometimes also the factor $1/(2\pi)$ in (2.20) is divided evenly as $1/\sqrt{2\pi}$ to (2.20) and (2.19).

Applying the complex Fourier transformation to the harmonic amplitude modulated oscillation of (2.13) leads to

$$\begin{aligned} A(\omega) &= U_0 \int_{-\infty}^{+\infty} \frac{\left(e^{i\omega_0 t} + e^{-i\omega_0 t}\right)}{2} e^{-i\omega t} dt \\ &+ MU_0 \int_{-\infty}^{+\infty} \frac{\left(e^{i\omega_m t} + e^{-i\omega_m t}\right)}{2} \frac{\left(e^{i\omega_0 t} + e^{-i\omega_0 t}\right)}{2} e^{-i\omega t} dt \\ &= \frac{U_0}{2} \int_{-\infty}^{+\infty} e^{-i(\omega-\omega_0)t} dt + \frac{U_0}{2} \int_{-\infty}^{+\infty} e^{-i(\omega+\omega_0)t} dt \\ &+ \frac{U_0 M}{4} \int_{-\infty}^{+\infty} e^{-i(\omega-\omega_0-\omega_m)t} dt + \frac{U_0 M}{4} \int_{-\infty}^{+\infty} e^{-i(\omega+\omega_0-\omega_m)t} dt \\ &+ \frac{U_0 M}{4} \int_{-\infty}^{+\infty} e^{-i(\omega-\omega_0+\omega_m)t} dt + \frac{U_0 M}{4} \int_{-\infty}^{+\infty} e^{-i(\omega+\omega_0+\omega_m)t} dt. \end{aligned} \tag{2.21}$$

The integrals of the kind $\int_{-\infty}^{+\infty} e^{-i(\omega-\omega')t} dt$ encountered in the previous equation are a special representation of the Dirac delta function [6]

$$\delta(\omega - \omega') = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-i(\omega-\omega')t} dt. \tag{2.23}$$

Hence,

$$\begin{aligned} A(\omega) &= 2\pi \frac{U_0}{2} \delta(\omega + \omega_0) + 2\pi \frac{U_0 M}{4} \delta(\omega + (\omega_0 + \omega_m)) + 2\pi \frac{U_0 M}{4} \delta(\omega + (\omega_0 - \omega_m)) \\ &+ 2\pi \frac{U_0}{2} \delta(\omega - \omega_0) + 2\pi \frac{U_0 M}{4} \delta(\omega - (\omega_0 + \omega_m)) + 2\pi \frac{U_0 M}{4} \delta(\omega - (\omega_0 - \omega_m)). \end{aligned} \tag{2.24}$$

In contrast to the amplitude spectrum of (2.16) the amplitude spectrum of (2.24) is not restricted to components at the angular frequencies $\omega = \omega_0 - \omega_m$, ω and $\omega_0 + \omega_m$, but also

[6] To be more specific the so-called Dirac delta function is not a function. It is defined by the so-called sifting property

$$\int_{-\infty}^{\infty} \delta(\omega - \omega') f(\omega) d\omega = f(\omega') \tag{2.22}$$

where $f(\omega)$ corresponds to any function which is continuous at ω'. If one substitutes $f(\omega) = 1$ one obtains $\int_{-\infty}^{\infty} \delta(\omega - \omega') d\omega = 1$. The delta function approaches ∞ for $\omega' \to \omega$.

includes negative angular frequencies $\omega = -(\omega_0 - \omega_m)$, $-\omega$ and $-(\omega_0 + \omega_m)$. Hence, the complex Fourier analysis leads to an amplitude spectrum which is symmetrical with respect to the angular frequency zero. Since the power contained in the oscillation must not depend on the description obtained by complex or real Fourier analysis, the spectrum of the latter is distributed evenly to the positive and negative angular (mirror) frequencies. Consequently, the amplitude spectra of (2.24) and (2.15) differ by a factor of 1/2.

2.1.2.3 Spectrum of a Damped Oscillator

A different type of amplitude modulation to be discussed now occurs when the amplitude of the oscillation dies out after some time. This situation is encountered, for instance, when an excited atom loses its energy by emitting electromagnetic radiation. To describe the rate of loss one often assumes that the portion of energy dW emitted during a time dt is at any instant t proportional to the energy $W(t)$ stored in the oscillator at this particular instant, i.e. $dW(t) = -\Gamma W(t)dt$. [7] Integrating $dW(t)/W(t) = -\Gamma t$ leads to $\ln W(t) - \ln W_0 = -\Gamma t$ and

$$W(t) = W(t=0)\exp(-\Gamma t). \tag{2.25}$$

As the energy stored in an oscillator is proportional to the square of the amplitude $U(t)$ the damped oscillation can be written as

$$U(t) = U_0 e^{-\frac{\Gamma}{2}t}\cos\omega_0 t. \tag{2.26}$$

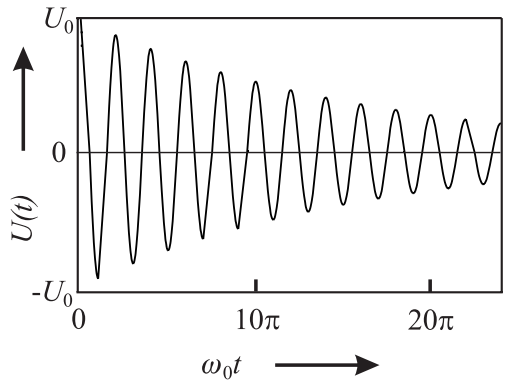


Figure 2.4: Damped harmonic oscillation according to (2.26) with $\Gamma = 0.04\,\omega_0$.

Equation (2.26) is an approximation for $\Gamma \ll \omega_0$ of the solution of a damped harmonic oscillator of mass m, spring constant D, damping constant α, described by the differential equation

$$\frac{d^2U(t)}{dt^2} + \Gamma\frac{dU(t)}{dt} + \omega_0^2 U(t) = 0 \tag{2.27}$$

$$\text{with} \qquad \omega_0 \equiv \sqrt{\frac{D}{m}} \qquad \text{and} \qquad \Gamma \equiv \frac{\alpha}{m} \qquad \text{for} \qquad \Gamma \ll \omega_0.$$

[7] As in this book the frequency ν and the angular frequency ω are equally used in the formulas, $\Gamma = 2\pi\gamma$ denotes the damping constant, linewidth, etc. in the angular frequency domain and γ the corresponding quantities in the frequency domain.

Solving the differential equation (2.27) by the ansatz $U(t) = U_0 \exp(i\omega t)$ leads to

$$-\omega^2 U(t) + i\omega\Gamma U(t) + \omega_0^2 U(t) = 0. \tag{2.28}$$

For arbitrary amplitudes $U(t) \neq 0$ this equation is solved by

$$\omega_{1,2} = \frac{i\Gamma}{2} \pm \sqrt{\frac{-\Gamma^2}{4} + \omega_0^2} \tag{2.29}$$

which leads to $\omega_{1,2} \approx \frac{i\Gamma}{2} \pm \omega_0$ for $\Gamma \ll \omega_0$. Inserting these two solutions into the ansatz $U(t) = U_0' \exp(i\omega t)$ and taking the sum of the two particular solutions leads to (2.25) with $U_0' = U_0/2$.

To obtain the frequency spectrum of the damped harmonic oscillator, (2.26) is Fourier transformed by use of (2.20) as

$$A(\omega) = \int_0^\infty U_0 e^{-\frac{\Gamma}{2}t} \cos(\omega_0 t) e^{-i\omega t} dt \tag{2.30}$$

where the lower integration limit has been changed from $-\infty$ to 0 since $U(t) = 0$ for $t < 0$.

$$\begin{aligned} A(\omega) &= \int_0^\infty U_0 e^{-\frac{\Gamma}{2}t} \left\{ \frac{e^{i\omega_0 t} + e^{-i\omega_0 t}}{2} \right\} e^{-i\omega t} dt \qquad (2.31) \\ &= \int_0^\infty \frac{U_0}{2} e^{[i(\omega_0-\omega)-\frac{\Gamma}{2}]t} dt + \int_0^\infty \frac{U_0}{2} e^{[i(-\omega_0-\omega)-\frac{\Gamma}{2}]t} dt \\ &= \frac{U_0}{2} \frac{1}{i(\omega_0-\omega)-\frac{\Gamma}{2}} \left[e^{\left[i(\omega_0-\omega)-\frac{\Gamma}{2}\right]t} \right]_0^\infty \\ &+ \frac{U_0}{2} \frac{1}{i(-\omega_0-\omega)-\frac{\Gamma}{2}} \left[e^{\left[i(-\omega_0-\omega)-\frac{\Gamma}{2}\right]t} \right]_0^\infty \\ &= \frac{U_0}{2} \left\{ \frac{1}{i(\omega-\omega_0)+\frac{\Gamma}{2}} + \frac{1}{i(\omega+\omega_0)+\frac{\Gamma}{2}} \right\}. \end{aligned}$$

If one is interested in the spectrum of frequencies ω close to ω_0, i.e. for $\omega - \omega_0 \ll \omega_0$, the second term is in general much smaller compared to the first one and can be neglected.[8] After multiplying numerator and denominator using the complex conjugate of the denominator one obtains

$$A(\omega) = \frac{U_0}{2} \frac{-i(\omega-\omega_0)+\frac{\Gamma}{2}}{\left[i(\omega-\omega_0)+\frac{\Gamma}{2}\right]\left[-i(\omega-\omega_0)+\frac{\Gamma}{2}\right]} = \frac{U_0}{2} \frac{-i(\omega-\omega_0)+\frac{\Gamma}{2}}{(\omega-\omega_0)^2+(\frac{\Gamma}{2})^2}. \tag{2.32}$$

[8] This approximation is often referred to as the "rotating wave approximation".

In contrast to the example of the harmonically modulated oscillator the spectral function $A(\omega) = \Re\mathrm{e}\, A(\omega) + i\Im\mathrm{m}\, A(\omega)$ comprises a real and an imaginary part

$$\begin{aligned} \Re\mathrm{e}\, A(\omega) &= \frac{U_0}{2}\frac{\frac{\Gamma}{2}}{(\omega-\omega_0)^2+(\frac{\Gamma}{2})^2} \qquad \text{and} \\ \Im\mathrm{m}\, A(\omega) &= -\frac{U_0}{2}\frac{\omega-\omega_0}{(\omega-\omega_0)^2+(\frac{\Gamma}{2})^2} \end{aligned} \tag{2.33}$$

which are displayed in Fig. 2.5 a) and b).[9]

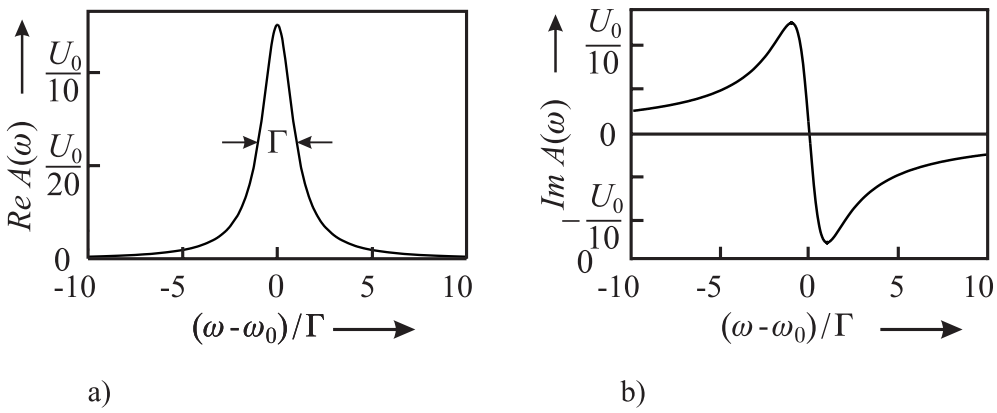


Figure 2.5: Spectral function of a damped oscillator according to (2.33) with $\Gamma = 0.04\,\omega_0$.a) Real part. b) Imaginary part.

Like $\Re\mathrm{e}\, A(\omega)$ the power spectrum contained in the Fourier components $P(\omega) \propto A(\omega)A^*(\omega) = [\Re\mathrm{e}\, A(\omega)]^2 + [\Im\mathrm{m}\, A(\omega)]^2$, i.e.

$$P(\omega) \propto \frac{U_0^2}{4}\frac{(\omega-\omega_0)^2+(\frac{\Gamma}{2})^2}{\left[(\omega-\omega_0)^2+(\frac{\Gamma}{2})^2\right]^2} = \frac{U_0^2}{4}\frac{1}{(\omega-\omega_0)^2+(\frac{\Gamma}{2})^2}, \tag{2.34}$$

has also a Lorentzian lineshape (Fig. 2.6). Hence, the exponential decay of the amplitude of the damped harmonic oscillation leads to a continuous band of frequencies of linewidth $\Delta\omega$. To determine the width $\Delta\omega$ of the band one first determines the maximum of the Lorentzian as $A(\omega = \omega_0)A^*(\omega = \omega_0) = U_0^2/\Gamma^2$. To determine the Full Width at Half Maximum (FWHM) one calculates the frequency $\omega_{1/2}$, where the corresponding value $A(\omega_{1/2})A^*(\omega_{1/2})$ has dropped to one-half of the maximal value $A(\omega_{1/2})A^*(\omega_{1/2}) = 1/2A(\omega_0)A^*(\omega_0)$ as

$$\frac{1}{2}\frac{U_0^2}{\Gamma^2} = \frac{U_0^2}{4}\frac{1}{(\omega_{1/2}-\omega_0)^2+(\frac{\Gamma}{2})^2}. \tag{2.35}$$

[9] The sign of the imaginary part of (2.33) and Fig. 2.5 b) is a consequence of our choice of the Fourier pair (2.19) and (2.20). Interchanging the signs in the complex exponential functions in (2.19) and (2.20) would lead to a sign change in (2.33) and Fig. 2.5 b).

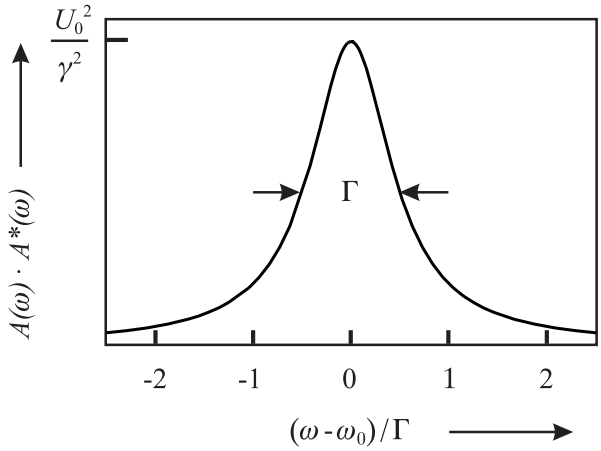


Figure 2.6: The square of the modulus of the Fourier transform of the damped harmonic oscillation is a Lorentzian.

From (2.35) one finds $(\omega_{1/2} - \omega_0)^2 = (\frac{\Gamma}{2})^2$. Hence, the full width at half maximum $\Delta\omega_{\mathrm{FWHM}} \equiv 2(\omega_{1/2} - \omega_0)$ of a Lorentzian is given by [10] the "damping constant" Γ as

$$\Delta\omega_{\mathrm{FWHM}} = \Gamma. \tag{2.36}$$

According to (2.25) $1/\Gamma \equiv \tau$ represents a characteristic time after which the energy stored in the oscillator has been reduced by $1/e$. In the case of a damped atom where the excitation energy of the atomic oscillator decays with the same time constant τ as the irradiated power, the linewidth can be related to the lifetime τ of the excited state as

$$\Delta\omega = \Gamma = \frac{1}{\tau}. \tag{2.37}$$

If the starting and terminating level both decay with the lifetimes τ_2 and τ_1, respectively, the bandwidth of the line is given as

$$\Delta\omega = \Gamma = \frac{1}{\tau_1} + \frac{1}{\tau_2}. \tag{2.38}$$

Optical transitions in atoms often start from excited states having lifetimes of a few nanoseconds. As an example, consider the calcium atom in the excited $^1\mathrm{P}_1$ state that decays to the ground state with a time constant $\tau \approx 4.6\,\mathrm{ns}$ emitting blue radiation with a wavelength $\lambda \approx 423\,\mathrm{nm}$ (frequency $\nu = c/\lambda \approx 7 \times 10^{14}\,\mathrm{Hz}$). Relating the decay time to the oscillation period of this transition one notices that the amplitude of the emitted wave train decays to $1/e \approx 0.37$ only after about three million oscillations. The change of the amplitude during one oscillation is therefore very small and would be hardly visible in a plot like Fig. 2.4.

[10] Here, and in the remainder of this book we use Γ to denote a full width at half maximum in the angular frequency (ω) domain and use γ for the full width at half maximum in the frequency ν domain.

To characterise the damped oscillation, a quality factor (Q factor) is used which is defined via the average stored energy W divided by the average dissipated energy as

$$Q \equiv \frac{\omega_0 W}{-dW/dt}. \tag{2.39}$$

Using (2.26) one derives $W \propto \overline{U(t)^2} = U_0^2/2 \exp(-\Gamma t)$ and $dW/dt \propto -\Gamma U_0^2/2 \exp(-\Gamma t)$ and, hence, using (2.36)

$$Q = \frac{\omega_0}{\Gamma} = \frac{\omega_0}{\Delta\omega}. \tag{2.40}$$

Thus the Q factor can be derived from the fractional linewidth $\Delta\omega/\omega_0$ of a measured line and the definitions (2.39) and (1.1) are equivalent.

As a typical example of optical transitions, the aforementioned spectral line of the Calcium atom has a line quality factor $Q \approx 2 \times 10^7$. The same atom, however, can be excited to a long-lived state designated by $^3P_1$ from which, after a lifetime of τ = 0.4 ms, the atom decays to the ground state thereby emitting red radiation (λ = 657 nm). The quality factor of the corresponding damped oscillation is $Q > 1.1 \times 10^{12}$.

The relation $\Delta\omega\,\tau = 1$ (2.37) between the decay time of an oscillation and the width of the corresponding frequency band has been derived here by the Fourier transformation. Multiplication with Planck's constant $\hbar = h/(2\pi)$ leads to $\hbar\Delta\omega\,\tau = \Delta E\,\tau = \hbar$ which shows an intimate connection of (2.37) to the Heisenberg uncertainty principle of quantum mechanics

$$\Delta E\,\Delta t \geq \frac{\hbar}{2}. \tag{2.41}$$

The time-domain and the complementary frequency-domain description of the oscillation are linked by the fact that the integrated power spectra in both domains have to be equal as will be shown in the following. The total energy contained in the damped oscillation is (cf. footnote 1 of Chapter 2) proportional to

$$\int_{-\infty}^{\infty} |U(t)|^2 dt = \int_{-\infty}^{\infty} U(t)U^*(t)dt = \int_{-\infty}^{\infty} U(t) \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} A^*(\omega) e^{-i\omega t} d\omega \right] dt. \tag{2.42}$$

Interchanging the sequence of integration one obtains

$$\int_{-\infty}^{\infty} |U(t)|^2 dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} A^*(\omega) \left[\int_{-\infty}^{\infty} U(t) e^{-i\omega t} dt \right] d\omega$$

and

$$\int_{-\infty}^{\infty} |U(t)|^2 dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |A(\omega)|^2 d\omega \qquad \text{(Parseval's formula)} \tag{2.43}$$

where we have used (2.19) and (2.20). As $|A(\omega)|^2 d\omega$ is the power within the angular frequency interval between ω and $\omega + d\omega$, it is found that $|A(\omega|^2$ represents a so-called power

spectral density in the angular frequency domain. Hence, Parseval's formula states that the total power integrated over the time domain equals the total power integrated in the Fourier-frequency domain.

To summarise the results of this section, one finds that any amplitude modulation of a harmonic oscillation leads to additional frequency components besides the carrier frequency. Hence, a small linewidth of an oscillator can be achieved only if the amplitude of the oscillator has a high temporal stability. In the case of a damped oscillation decaying with a time constant τ a band of frequencies occurs whose width is reciprocal to the decay time τ. Similarly, any oscillation that is limited to a finite period or interrogated during a finite observation time τ corresponds to a frequency band whose width is indirectly proportional to τ (see Section 5.4.1).

2.1.3 Phase Modulation

In this section the modulation of the phase of a harmonic oscillation and its effect on the frequency spectrum is investigated. To simplify the mathematics and to identify the related effects as purely as possible, the amplitude is kept constant and for simplicity a harmonic modulation of the phase is chosen as follows

$$U_{\mathrm{PM}}(t) = U_0 \cos\varphi = U_0 \ \cos(\omega_0 t + \delta \cos \ \omega_m t). \tag{2.44}$$

The modulation index δ (index of phase modulation) corresponds to the maximum difference between the phase of the modulated oscillator and that of an unmodulated one. The instantaneous angular frequency $\omega(t)$ of (2.44) can be derived by using (2.11) as

$$\omega(t) = \omega_0 - \omega_m \ \delta \ \sin\omega_m t \equiv \omega_0 - \Delta\omega \sin\omega_m t \tag{2.45}$$

where

$$\Delta\omega = \omega_m \ \delta. \tag{2.46}$$

represents the maximal deviation of the instantaneous angular frequency from the unperturbed angular frequency ω_0.

According to (2.46) phase modulation and frequency modulation are closely related and the terms are used in parallel. Radio engineers refer to phase modulation if the modulation index δ is kept constant independent of the modulation frequency ω_m. In this case the frequency deviation $\Delta\omega = \omega_m\delta$ increases linearly with the modulation frequency. If in the modulation process the frequency deviation $\Delta\omega$ is fixed and does not depend on the modulation frequency ω_m, one refers to frequency modulation. The modulation index $\delta = \Delta\omega/\omega_m$ is reciprocal to the modulation frequency ω_m.

Figure 2.7 shows an oscillation where the phase is harmonically modulated with $\omega_m = 0.1\,\omega_0$ using a modulation index of $\delta = 7.5$. From (2.46) one finds that this modulation index corresponds to a frequency deviation $\Delta\omega = 0.75\,\omega_0$ and that the instantaneous frequency swings between $0.25\,\omega_0$ and $1.75\,\omega_0$. Obviously, the phase modulated oscillation cannot be represented by a single frequency. To investigate the frequency spectrum of the phase modu-

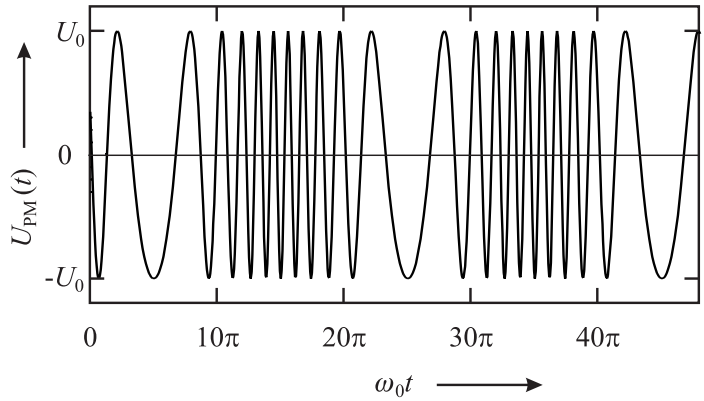


Figure 2.7: Time dependence of the amplitude of a phase-modulated oscillation according to (2.44) with $\omega_m = 0.1\,\omega_0$ and modulation index $\delta = 7.5$.

lated oscillation we write (2.44) as

$$\begin{aligned} U_{\mathrm{PM}}(t) &= U_0 \cos(\omega_0 t + \delta \cos \omega_m t) \\ &= U_0 \Re\mathrm{e}\,\{\exp(i\omega_0 t)\exp(i\delta \cos \omega_m t)\}. \end{aligned} \tag{2.47}$$

We expand the second complex exponential function into a power series and transform the higher powers of $\cos \omega_m t$ into cosine terms of higher harmonics $n\ \omega_m t$ by the appropriate trigonometric formulas obtaining

$$\begin{aligned} &\exp[i\delta \cos(\omega_m t)] \\ &= 1 \quad + \quad i\delta \cos(\omega_m t) \\ &\quad + \quad i^2 \frac{1}{2\,!}\delta^2 \frac{1}{2}[1 + \cos(2\ \omega_m t)] \\ &\quad + \quad i^3 \frac{1}{3\,!}\delta^3 \frac{1}{4}[3\ \cos(\omega_m t) + \cos(3\ \omega_m t)] \\ &\quad + \quad i^4 \frac{1}{4\,!}\delta^4 \frac{1}{8}[3 + 4\ \cos(2\ \omega_m t) + \cos(4\ \omega_m t)] \\ &\quad + \quad i^5 \frac{1}{5\,!}\delta^5 \frac{1}{16}[10\ \cos(\omega_m t) + 5\ \cos(3\ \omega_m t) + \cos(5\omega_m t)] \\ &\quad + \quad i^6 \frac{1}{6\,!}\delta^6 \frac{1}{32}[10 + 15\ \cos(2\ \omega_m t) + 6\ \cos(4\ \omega_m t) + \cos(6\ \omega_m t)] \\ &\quad + \quad i^7 \frac{1}{7\,!}\delta^7 \frac{1}{64}[35\ \cos(\omega_m t) + 21\ \cos(3\ \omega_m t) + 7\ \cos(5\ \omega_m t) + \cos(7\ \omega_m t)] \\ &\quad + \quad \cdots . \end{aligned}$$

After rearranging the terms one finds

$$\begin{aligned} \exp[i\delta \cos(\omega_m t)] = J_0(\delta) + 2\ i\ J_1(\delta)\cos(\omega_m t) + 2\ i^2 J_2(\delta)\cos(2\ \omega_m t) \\ + \cdots + 2\ i^n J_n(\delta)\cos(n\ \omega_m t)\cdots \end{aligned} \tag{2.48}$$

where the Bessel functions of the first kind J_n are given as follows

$$\begin{aligned} J_0(\delta) &= 1 - \left(\frac{\delta}{2}\right)^2 + \frac{1}{4}\left(\frac{\delta}{2}\right)^4 - \frac{1}{36}\left(\frac{\delta}{2}\right)^6 + \cdots \\ J_1(\delta) &= \left(\frac{\delta}{2}\right) - \frac{1}{2}\left(\frac{\delta}{2}\right)^3 + \frac{1}{12}\left(\frac{\delta}{2}\right)^5 - \cdots \\ J_2(\delta) &= \frac{1}{2}\left(\frac{\delta}{2}\right)^2 - \frac{1}{6}\left(\frac{\delta}{2}\right)^4 + \frac{1}{48}\left(\frac{\delta}{2}\right)^6 - \cdots \\ J_3(\delta) &= \frac{1}{6}\left(\frac{\delta}{2}\right)^3 + \frac{1}{24}\left(\frac{\delta}{2}\right)^5 + \frac{1}{240}\left(\frac{\delta}{2}\right)^7 - \cdots \\ &\vdots \end{aligned} \tag{2.49}$$

The Bessel functions J_0 to J_{10} are shown in Fig. 2.8.

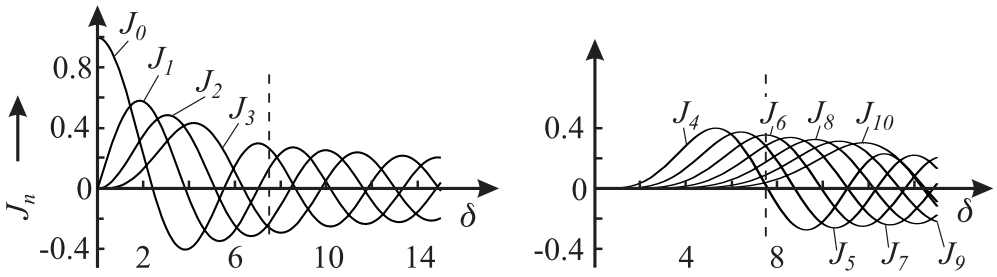


Figure 2.8: Bessel functions (of the first kind) $J_n(\delta)$ of order $0 \leq m \leq 10$ as a function of δ. The dashed line indicates the modulation index δ = 7.5 used in Fig. 2.7.

Hence, (2.47) can now be written as

$$U_{\mathrm{PM}}(t) = U_0 \sum_{n=-\infty}^{\infty} \Re\mathrm{e}\left\{(i)^n J_n(\delta) \exp\left[i(\omega_0 + n\omega_m)t\right]\right\}. \tag{2.50}$$

The Bessel functions of negative order can be computed from

$$J_{-n} = (-1)^n \, J_n. \tag{2.51}$$

By writing (2.50) explicitly as follows

$$
\begin{aligned}
U_{\mathrm{PM}}(t) = U_0 \Re\mathrm{e}\, \{ & J_0(\delta) \exp(i\omega_0 t) \\
& + \; iJ_1(\delta)\,[\exp i(\omega_0 t + \omega_m t) + \exp i(\omega_0 t - \omega_m t)] \\
& - \; J_2(\delta)\,[\exp i(\omega_0 t + 2\omega_m t) + \exp i(\omega_0 t - 2\omega_m t)] \\
& - \; iJ_3(\delta)\,[\exp i(\omega_0 t + 3\omega_m t) + \exp i(\omega_0 t - 3\omega_m t)] \\
& + \; J_4(\delta)\,[\exp i(\omega_0 t + 4\omega_m t) + \exp i(\omega_0 t - 4\omega_m t)] \\
& + \; i \cdots \} \\
= U_0 \{ & J_0(\delta) \cos\omega_0 t \\
& - \; J_1(\delta) \sin(\omega_0 t + \omega_m t) - J_1(\delta) \sin(\omega_0 t - \omega_m t) \\
& - \; J_2(\delta) \cos(\omega_0 t + 2\omega_m t) - J_2(\delta) \cos(\omega_0 t - 2\omega_m t) \\
& + \; J_3(\delta) \sin(\omega_0 t + 3\omega_m t) + J_3(\delta) \sin(\omega_0 t - 3\omega_m t) \\
& + \; J_4(\delta) \cos(\omega_0 t + 4\omega_m t) + J_4(\delta) \cos(\omega_0 t - 4\omega_m t) \\
& - \; \cdots \}
\end{aligned}
\tag{2.52}
$$

one finds in the frequency spectrum of the harmonic phase-modulated oscillation the carrier at angular frequency ω and an infinite number of components at the side-band angular frequencies $\omega \pm n\,\omega_m$ spaced at multiples of the angular modulation frequency ω_m. This is in contrast to the harmonically amplitude-modulated oscillation where there is only one first-order frequency component $\omega \pm \omega_m$ on either side of the carrier. From Fig. 2.8 we see that Bessel functions of higher orders become important only when the modulation index becomes larger than unity. To illustrate this in more detail we investigate the phase modulated oscillation shown in Fig. 2.7. Since the modulation index of this example was δ = 7.5 one has to compute $J_n(7.5)$ according to (2.49) or one has to look them up in Fig. 2.8 at the dashed line representing δ = 7.5. From the amplitudes $J_n(7.5)$ displayed in Fig. 2.9 a) one finds that about 10 components on either side of the carrier are necessary to represent the phase modulated oscillation with modulation index $\delta = 7.5$. The amplitudes of the higher-order components become rapidly less important. As a rule of thumb, the number of side bands contributing significantly to the spectrum is given by the modulation index. This is readily seen because the highest instantaneous frequency according to (2.46) is $\omega_{\max} = \omega_0 + \Delta\omega = \omega_0 + \omega_m\delta$. To synthesise the most rapid oscillations in Fig. 2.7 accurately we mainly need frequencies up to this maximum frequency.

The square of the amplitude components of Fig. 2.9 a) is a measure of the power contained in the different frequency components representing the phase modulated oscillation. The power spectrum is symmetrical with respect to the carrier frequency. In the example chosen only about 7 % of the total power of the oscillation remains in the carrier. Even though it is not obvious from (2.50) the total power contained in the carrier and all of the side bands has to equal that of the unmodulated oscillation as can be concluded from Parseval's formula (2.43). Hence, the phase modulation leaves the total power unchanged, but distributes it onto the side bands in dependence on the modulation index. This is in contrast to the amplitude modulation (see (2.18)) where the total power increases with increasing (amplitude) modulation index M.

Oscillators utilised in frequency standards in general are supposed to exhibit only little phase modulation so that the modulation index is small ($\delta < 1$). In this case it is sufficient

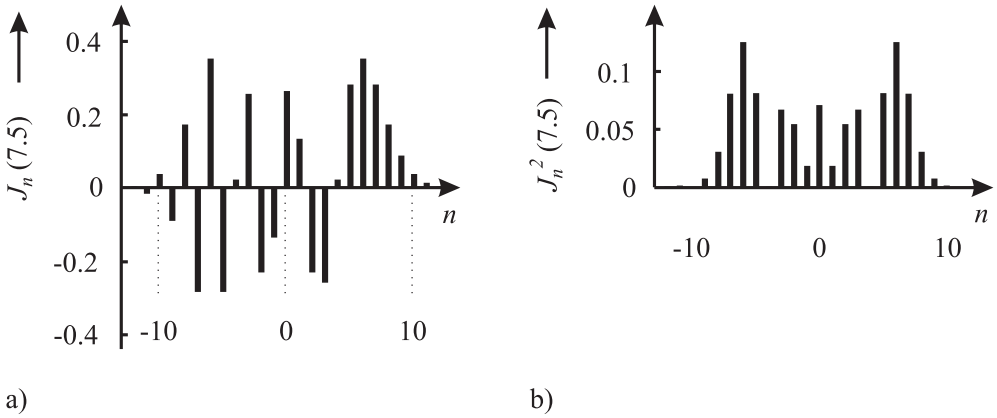


Figure 2.9: a) Bessel functions $J_n(\delta = 7.5)$ contributing to the spectrum of a phase modulated oscillation according to Fig. 2.7. b) Amplitudes of Fig. 2.9 a) are squared and represent the power contained in the side bands of the harmonically phase modulated oscillation of Fig. 2.7.

to take into account only the carrier and the first-order side bands determined by J_0 and J_1, respectively, because the Bessel functions of higher orders become very small (see Fig. 2.8 and (2.52)). Similar to the case of an amplitude modulated oscillation, there is a carrier at ω_0 and the two side bands at $\omega_0 + \omega_m$ and at $\omega_0 - \omega_m$. There is, however, a significant difference stemming from the phase differences between the cosine function of the carrier and the sine functions of the first-order side bands in (2.52). To see this more clearly we investigate a phase modulated oscillation in the phasor plot using a modulation index δ = 1 (Fig. 2.10). In this case $J_0 = 0.765$, $J_1 = 0.44$ and $J_2 = 0.115$ and it may be justified to keep only the carrier and the two first-order side bands. As usual, the phases $\omega_0 t$ are chosen as multiples of 2π such that the carrier phasor is on the real axis. At the starting time t = 0, i.e. $\omega_0 t$ = 0 the phasors representing both side bands are rotated by 90° with respect to the carrier phasor due to the phase shift of 90° between the cosine and sine function (or the phase factor $i = \exp(i\pi/2)$) in (2.52). The phase modulated oscillation at $\omega_0 t = 0$ is represented by the phasor resulting from the addition of the phasor of the carrier and the two side-band phasors both being perpendicular to the carrier phasor. The resultant phasor points at an angle of $\arctan \alpha = (2 \times 0.4)/1 \approx 38.7°$. When the carrier phase has gained 2π (second picture) the phasor representing the low-frequency side band fell back by 45° and the phasor representing the high-frequency one advanced by 45°. In this case and in all cases depicted in Fig. 2.10 both side-band phasors are symmetrical with respect to the imaginary axis and the resultant phasor of both first-order side bands is always perpendicular to the carrier phasor. From Fig. 2.10 one sees that in the case of a phase modulated oscillation the resultant phasor "swings" around the phasor representing the carrier. With respect to the fixed phasor of the carrier the resultant is sometimes ahead of and sometimes falls back behind the carrier.

At first glance it is surprising that the length of the resultant phasor is not constant in Fig. 2.10 which is equivalent to an additional amplitude modulation even though we have assumed a "pure" phase modulation with the amplitude being constant. The apparent amplitude

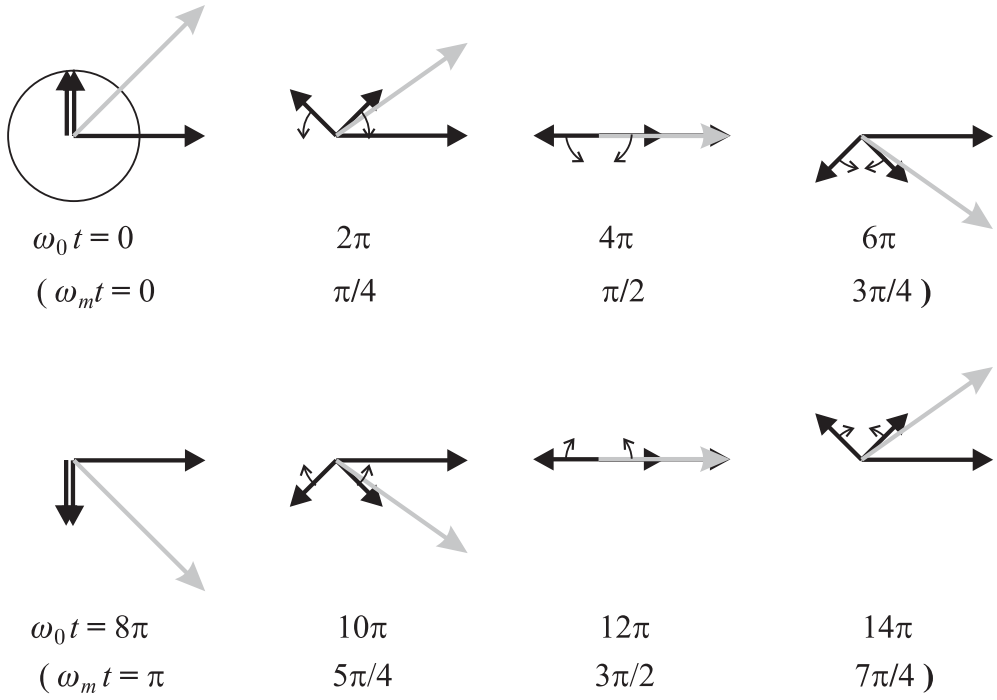


Figure 2.10: Phasor representation of a phase modulated oscillation according to (2.52) of modulation index $\delta = 1$ and a modulation frequency $\omega_m = \omega_0/8$ by the phasors of the carrier pointing to the right on the abscissa and the two rotating side-band phasors (see text).

modulation results because we took into account only the phasors of the first-order side bands. For an exact treatment one needs to include the phasors of all higher-order side bands with frequencies $\omega_0 + n\omega_m$ und $\omega - n\omega_m$.

Real oscillators are in general not modulated in the simple ways described so far. It has to be expected that real oscillators show at least to some degree modulations of phase and amplitude at the same time. The modulations in general cannot be represented by a single modulation frequency or by a harmonic time function. If one were to know the explicit time-dependent modulation function one could decompose it according to Fourier by a finite or infinite number of harmonic functions. In such a case the spectrum of the side bands could be very dense. In general, the modulation of a real oscillator cannot even be described by an analytical time function because the temporal evolution of the phase and the amplitude of the oscillator fluctuates in a non-deterministic way. The methods describing the fluctuations of the amplitude and frequency of such oscillators will be presented in Section 3.

2.2 Oscillator with Feedback

From the results of the previous section it follows that any amplitude and phase modulation has to be kept at a minimum in an oscillator to be used for a frequency standard. The keep the amplitude of an oscillator constant the power extracted from or dissipated by the oscillator has to be compensated for. The compensation can be achieved by splitting off a fraction of the emitted power which is amplified and fed back to the oscillator with the proper phase (Fig. 2.11 a). The feedback works as well if the arrangement of the power splitter and the amplifier is interchanged (Fig. 2.11 b).

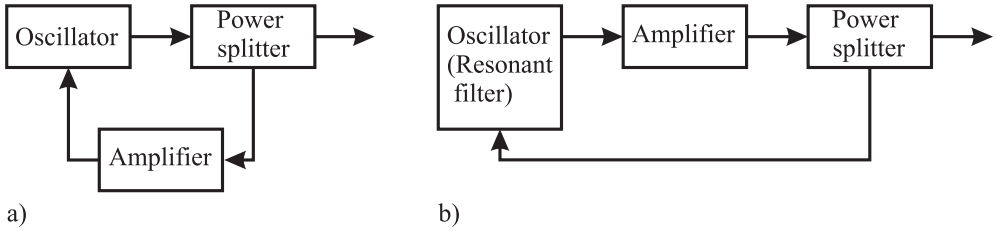


Figure 2.11: a) To compensate for the power extracted from the oscillator, part of this extracted power is split off in a power splitting device, amplified and fed back to the oscillator. b) The sequence of power splitter and amplifier in Fig. 2.11 a) can be interchanged.

To sustain continuous oscillation with constant amplitude certain conditions have to be fulfilled that can be derived from a balance when the system described by Fig. 2.11 a) is in equilibrium. Consider the oscillating signal $U_{\rm out}(t) = U_0 \exp(i\omega t)$ at the output terminals of the oscillator which can represent, e.g., a voltage, a microwave power or the field strength of a light wave. A fraction k of this signal amplitude is split off in the power splitter and fed back via the amplifier with amplitude gain factor A to the oscillator. In the steady state the power ($\propto U_{\rm out}^2$) extracted from the oscillator has to be supplied to the input of the oscillator ($\propto U_{\rm in}^2$) leading to the identity for the amplitudes [11]

$$U_{\rm out} = U_0 \exp(i\omega t) = U_{\rm in} = kAU_0 \exp i[\omega t - \alpha(\omega) - \beta(\omega)]. \tag{2.53}$$

In (2.53) α and β take into account the frequency-dependent phase shift resulting from the finite velocity of the signal with the corresponding transit time of the signal in the feedback path and the phase shift in the amplifier, respectively. From the oscillation condition (2.53) separate conditions for the amplitude and the phase can be derived as

$$kA = 1 \qquad \text{(amplitude condition)} \tag{2.54}$$

$$\text{and} \qquad \alpha + \beta = 0, 2\pi, \cdots \qquad \text{(phase condition).} \tag{2.55}$$

The phase condition $\alpha + \beta = 0, 2\pi, \cdots$ requires that the signal fed back to the oscillator is in phase with the oscillator. According to the amplitude condition harmonic oscillations occur only when the gain compensates all losses. For a smaller gain the amplitude of the oscillation decays exponentially, for a higher gain it increases exponentially with time. Since

[11] Here we have assumed for simplicity that the impedances for the input and the output are equal.

the output of the amplifier is limited, the amplitude of the output signal will saturate at a level determined by, e.g., the voltage supplied by the power supply. The non-linearity of the amplifier, however, will distort the signal of the oscillator and will no longer deliver a purely harmonic signal. Hence, besides the fundamental frequency the output signal will also include higher-harmonic frequency components. To avoid this harmonic distortion a control system for the amplifier has to ensure that the amplitude condition (2.54) holds.

In the feedback system of Fig. 2.11 a) part of the power of the oscillator is fed back to the oscillator via the amplifier and the oscillator is now no longer freely oscillating but acts as a driven oscillator. As the sequence of the power splitter and the amplifier can be interchanged (see Fig. 2.11 b) an alternative interpretation for this feedback loop assumes that part of the power delivered by the amplifier is coupled back to the input of the amplifier via the oscillator. In this case the oscillator acts as a resonant filter whose properties will be investigated in the following.

To calculate the frequency dependence of the driven oscillator (resonant filter) we modify the dynamic equation of the damped harmonic oscillator (2.27) by adding an additional periodic force $F(t)$ that acts on the oscillator. To simplify the calculation one uses $F(t) = u_0/m \exp(i\omega t)$ in complex representation and obtains

$$\frac{d^2U(t)}{dt^2} + \Gamma\frac{dU(t)}{dt} + \omega_0^2 U(t) = \frac{u_0}{m}e^{i\omega t}. \tag{2.56}$$

(2.56) can be solved using the ansatz $U(t) = \widetilde{U}_0 \exp(i\omega t)$ [12] yielding

$$\begin{aligned} \widetilde{U}_0 &= \frac{u_0}{m(\omega_0^2 - \omega^2 + i\Gamma\omega)} \\ &= \frac{u_0(\omega_0^2 - \omega^2)}{m(\omega_0^2 - \omega^2) + m\Gamma^2\omega^2)} - i\frac{u_0\Gamma\omega}{m(\omega_0^2 - \omega^2)^2 + m\Gamma^2\omega^2}. \end{aligned} \tag{2.57}$$

$\widetilde{U}_0$ is the frequency-dependent complex response of the driven oscillator (resonant filter). Often, a complex transfer function is defined as the ratio of the response and the driving force

$$\chi(\omega) \equiv \frac{\widetilde{U}_0}{u_0/m}. \tag{2.58}$$

The response of the resonant filter can be represented in the complex plane either (as in (2.58) as Cartesian coordinates) by its real and imaginary part $\Re\mathrm{e}\,\widetilde{U}_0$ and $\Im\mathrm{m}\,\widetilde{U}_0$ or (in polar coordinates) by amplitude and phase angle

$$\begin{aligned} \widetilde{U}_0 &= a(\omega)e^{i\varphi} \qquad \text{with} \\ a(\omega) = |\widetilde{U}_0| &= \sqrt{\widetilde{U}_0\widetilde{U}_0^*} = \sqrt{\Re\mathrm{e}\,\widetilde{U}_0^2 + \Im\mathrm{m}\,\widetilde{U}_0^2} \\ &= \sqrt{\frac{u_0^2\left[(\omega_0^2 - \omega^2)^2 + \Gamma^2\omega^2\right]}{m^2\left[(\omega_0^2 - \omega^2) + \Gamma^2\omega^2\right]^2}} = \frac{u_0}{m\sqrt{(\omega_0^2 - \omega^2)^2 + \Gamma^2\omega^2}} \end{aligned} \tag{2.59}$$

$$\text{and} \qquad \tan\varphi = \frac{\Im\mathrm{m}\,\widetilde{U}_0}{\Re\mathrm{e}\,\widetilde{U}_0} = \frac{\Gamma\omega}{\omega^2 - \omega_0^2}. \tag{2.60}$$

[12] This particular solution is not necessarily the general solution.

The amplitude gain of the resonant filter, given by the modulus of the complex transfer function (2.59), is almost constant for low frequencies $\omega \ll \omega_0$ (see Fig. 2.12a) and increases when it approaches ω_0 provided that $\Gamma < \omega_0$. After passing its maximum value at ω_0 it falls off roughly proportional to $1/\omega$ and for high values of $\omega \gg \omega_0$ it drops proportional to $1/\omega^2$. In the vicinity of the resonance frequency the phase of the resonant filter changes from $0°$ to

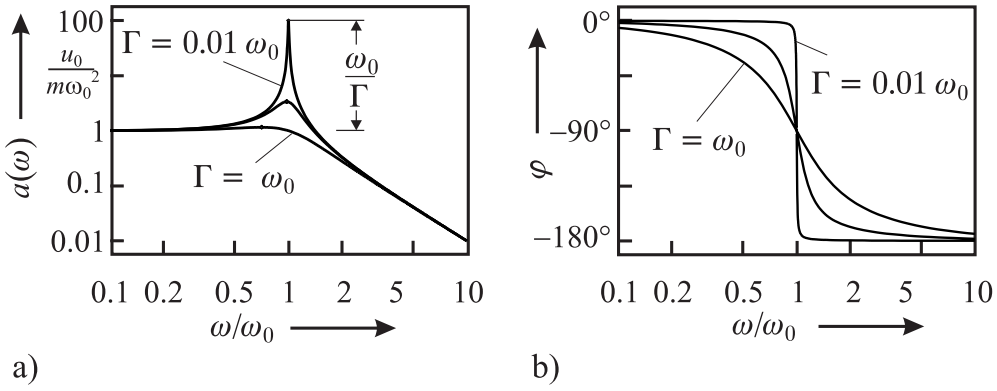


Figure 2.12: Bode plot of a resonant filter for $\Gamma = \omega_0, \Gamma = 0.1\omega_0$, and $\Gamma = 0.01\omega_0$. a) Amplitude. b) Phase.

$-180°$ as can be seen from Fig. 2.12 b).

The representation of a particular element in a feedback loop by a complex transfer function whose modulus describes the frequency response of the gain of the signal amplitude and whose phase represents the phase shift the signal suffers in the element is a common procedure. If more than one element is in the loop the total phase shift in the loop is calculated by adding the individual phase shifts whereas the resultant amplitude is given by the product of the individual components. Hence, the modulus $\alpha(\omega)$ and phase $\varphi(\omega)$ of the complex transfer function of an electronic element are often represented like in Fig. 2.12 in a so-called Bode plot where the (amplitude) gain is displayed logarithmically and the phase shift is displayed linearly. The Bode plots of the particular components are then added to yield the Bode plot of the combined system. As it is often useful to know the frequency response of the transfer functions in a frequency range of several orders of magnitude, a logarithmic scale is used for the frequency axis of the Bode plot.

The rapid phase change in the vicinity of the resonance frequency can be used to keep the frequency of an interrogating oscillator close to the resonance frequency. Thus it is interesting to investigate the phase variation, displayed in Fig. 2.12 b) close to resonance, in more detail. From (2.60) one finds

$$\varphi = \arctan \frac{\Gamma\omega}{\omega^2 - \omega_0^2} \approx -\frac{\pi}{2} - \frac{\omega^2 - \omega_0^2}{\Gamma\omega} \tag{2.61}$$

where we have made use of $\arctan x = \pm\pi/2 - 1/x + 1/(3x^2) - \cdots$. At $\omega = \omega_0$ one finds

the slope

$$\left.\frac{d\varphi}{d\omega}\right|_{\omega=\omega_0} = -\left.\frac{\omega^2+\omega_0^2}{\Gamma\omega^2}\right|_{\omega=\omega_0} = -\frac{2}{\Gamma} = -\frac{2Q}{\omega_0} \tag{2.62}$$

where (2.40) has been used.

2.3 Frequency Stabilisation

The steep variation of the response of a macroscopic or microscopic frequency reference in the vicinity of its resonance can be used to stabilise the frequency of an oscillator either making use of the amplitude or phase response of Fig. 2.12.

2.3.1 Model of a Servo Loop

Consider a servo system of a frequency standard depicted in Fig. 1.3. If the feedback loop [13] is open there is no fixed relationship between the frequency ν_i of the free-running oscillator and the reference frequency ν_0. The stabilisation scheme comprises a component called the discriminator which produces an error signal S that is a measure of the frequency deviation $\delta\nu \equiv \nu_s - \nu_0$ between the actual frequency of the oscillator ν_s and the reference frequency ν_0. Note that one has to distinguish the frequency ν_i of the free-running oscillator from the frequency of the oscillator ν_s even if the feedback loop is closed. There is a variety of techniques to produce an error signal, some of them will be discussed later. Here, for simplicity we assume that the error signal S is proportional to the frequency deviation

$$S \approx C\,(\nu_s - \nu_0) = C\,\delta\nu \tag{2.63}$$

for frequency deviations $\delta\nu$ not being too large. The error signal in general is processed further by suitable filtering and amplification in a servo amplifier which then produces the servo signal. The servo signal U_R acts on a servo element capable of changing the frequency of the oscillator in order to minimise its frequency deviation $\delta\nu$. Describing the servo amplifier by a frequency-dependent gain $g(f)$ one writes [14]

$$U_R = g(f)\,C\,\delta\nu. \tag{2.64}$$

Assuming that the servo element changes the frequency of the oscillator proportional to the servo signal U_R with a frequency-dependent response $D(f)$ the combined frequency-dependent transfer function of the servo loop is $D(f)\;g(f)\;C$. With the servo loop closed the frequency change of the oscillator initiated by the servo element has to counteract the frequency deviation $\delta\nu$ by so-called negative feedback and, hence,

$$\nu_s = \nu_i - D\;g(f)\;C\;\delta\nu. \tag{2.65}$$

[13] This is a different feedback loop from the one considered in Section 2.2 where it was assumed that the oscillator becomes oscillating as a result of the feedback. Here, we assume that the oscillator is already oscillating and the feedback is used for frequency stabilisation.

[14] Here, we have to distinguish between the frequency of the oscillator ν in the so-called carrier frequency domain and the Fourier frequency f used to describe the (Fourier) spectrum of the frequency deviations $\delta\nu(f)$.

Subtracting ν_0 on both sides of (2.65)

$$\nu_s - \nu_0 = \nu_i - \nu_0 - D\, g(f)\, C\, \delta\nu \tag{2.66}$$

one finds that the left-hand side of (2.66) represents the deviation $\delta\nu$ of the frequency ν_s of the frequency-stabilised oscillator from the reference frequency ν_0. Denoting the deviation of the frequency of the free-running oscillator from the reference frequency as $\nu_i - \nu_0 \equiv \Delta\nu$, (2.66) can be written as

$$\Delta\nu = \delta\nu + D\, g(f)\, C\, \delta\nu = \delta\nu[1 + D\, g(f)\, C] \tag{2.67}$$

from which one arrives at

$$\delta\nu = \frac{\Delta\nu}{1 + C\, D\, g(f)}. \tag{2.68}$$

In this simple model of a servo loop with proportional gain and negative feedback the frequency deviation $\Delta\nu$ of the free-running oscillator is reduced by the factor $1 + C\ D\ g(f)$. The somewhat surprising result is that there is a non-zero residual frequency deviation $\delta\nu$. To keep the frequency of the stabilised oscillator as close as possible to the reference frequency the overall gain $C\ D\ g(f)$ of the loop has to be as high as possible. In general the gain of the amplifier as well as the sensitivity of the servo element is frequency dependent. For optimum performance of the servo loop one needs to know the complex frequency response of each component including the amplitude gain as well as the phase. Dead times or transit times in servo elements, cables or in the amplifiers will lead to a frequency-dependent phase shift in the loop. Phase shifts are of utmost importance as they add up and a total phase shift of 180° produces positive rather than negative feedback, thereby increasing any frequency deviations. For characterising or designing an optimised servo loop it is therefore necessary to have profound knowledge of the frequency-dependent transfer functions of all components. Examples of the transfer functions of some often-used electronic components will be given in Section 2.4.

2.3.2 Generation of an Error Signal

There is a variety of different methods to generate an error signal. The error signal does not have to be a linear function of the frequency offset from the reference frequency as has been assumed in (2.63) but it should be monotonous with a zero crossing at resonance. This sign change at the resonance frequency allows the servo system to discriminate between the cases where the frequency of the oscillator is larger or lower than the reference frequency and to counteract the frequency deviations accordingly. Measuring the phase change near the resonance of a filter like that of Fig. 2.12 b) and subtracting a constant phase of $-\pi/2$ immediately leads to an error signal with the desired property. The transmission profile shown in Fig. 2.12 a), however, that is observed by monitoring the power in the loop when the frequency of the oscillator is scanned across the resonance, has an almost symmetric bell-shaped line profile. In the following we give two examples of how to generate an anti-symmetric dispersive signal from such a symmetric resonance feature. More examples will be given when describing particular frequency standards, e.g., in Section 9.

2.3.2.1 Side-lock Stabilisation

A particular simple method to stabilise the frequency of an oscillator is possible if the reference frequency is chosen at a suitable point at one side of the resonance line where the transmitted signal is at a level between maximum and minimum transmission (see Fig. 2.13).

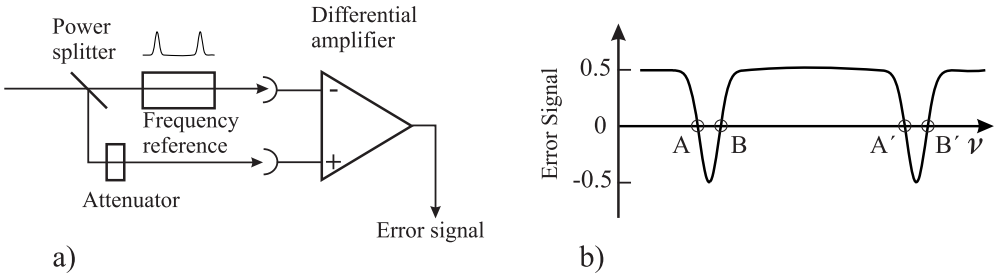


Figure 2.13: a) Schematics of a side lock. b) Error signal obtained from the difference of the signals from a detector measuring the power transmitted by the frequency reference and from a reference detector.

In this side-lock scheme, part of the power of the oscillator whose frequency is to be stabilised to a frequency reference, is divided roughly equally by a power splitter to different paths. In one path the signal interacts with the reference. To discuss the side-lock technique we need not specify the particular kind of oscillator and reference. Regardless whether we think of a combination of a laser and a Fabry–Pérot interferometer or of a microwave oscillator in connection with an absorption cell, the transmitted power is detected by a suitable photodiode or microwave detector and leads to a signal with the frequency dependence of the particular reference. The second path provides the reference signal which is obtained from a second detector exposed to a fraction of the power from the oscillator split off by the power splitter and which can be adjusted by a suitable attenuator. Subtracting both signals, e.g., in a differential amplifier, leads to a signal similar to the one shown in Fig. 2.13 b) which has two suitable lock points A and B for each resonance of the frequency reference. Near the lock points A and B the difference signal has the desired property of a monotonous curve that changes sign at the lock point. The side lock is simple to implement and it allows one to tune the frequency within a limited range by tuning the offset. The difference signal has the advantage of being largely independent of amplitude fluctuations of the power from the oscillator. The simplicity of the side lock has to be paid for by some disadvantages. Firstly, the chosen lock point does not coincide with the centre of the resonance, but is defined by the offset given by the setting of the attenuator. Secondly, as a consequence, the lock point is not very stable. Power fluctuations will only be compensated to the extent they lead to the same variations of the respective detector signals. A variation of the coupling to the frequency reference will also lead to a variation of the transmitted power and to a frequency shift. Such a variation is not unlikely, e.g., in the case of a Fabry–Pérot interferometer where the mode matching can be easily degraded by a limited beam pointing stability of the incoming laser beam. Thirdly, the capture range of the servo stabilisation is in general very asymmetric as can be seen from Fig. 2.13 b). Consider zero crossing A chosen as the lock point for a given polarity of the error signal. Any perturbation that reduces the frequency of the oscillator leads to a positive error

signal and is counteracted by the servo unit. The influence of a perturbation that increases the frequency of the oscillator will be reduced by the servo unit as long as the error signal is negative, i.e. if the frequency of the oscillator is within the range between ν(A) and ν(B). If, however, the perturbation leads to an increased frequency higher than ν(B) the error signal will again be positive and consequently the frequency of the oscillator will be repelled from the resonance and may jump to a corresponding lock point A$'$ of the next higher lying resonance. These disadvantages can be overcome by locking schemes employing modulation techniques.

2.3.2.2 Generation of an Error Signal Using Modulation Techniques

The stabilisation of the frequency ν_s of an oscillator to the centre frequency ν_0 of a resonance can also be achieved by modulating the frequency difference $\nu_s - \nu_0$. In general, either the frequency of the oscillator or the centre frequency of the reference can be modulated. The latter method is often used with macroscopic frequency references but is also available for frequency standards based on microscopic quantum absorbers if their clock transitions depend on external parameters that can be used for modulation, e.g., an absorption line with pronounced Zeeman effect. For non-tuneable resonance lines the frequency of the interrogating oscillator has to be modulated either by directly modulating the frequency of the oscillator itself or by use of an external modulator. In microwave and optical frequency standards two kinds of modulation are common, i.e. square wave modulation and harmonic modulation. Using square-wave modulation the frequency ν_s of the oscillator is periodically switched between by $\nu_s + \delta\nu$ and $\nu_s - \delta\nu$ for a time interval $\tau/2$ (Fig. 2.14 a).

Consider the case where the power transmitted through the absorber and measured by a detector is integrated for the durations of both half-periods $\tau/2$. The difference of the two integrated signals $\Delta I(\nu_s) \equiv I(\nu_s + \delta\nu) - I(\nu_s - \delta\nu)$ as a function of the frequency deviation $\nu_s - \nu_0$ of the oscillator from the centre frequency leads to an anti-symmetric discriminant curve (see Fig. 2.14 b).

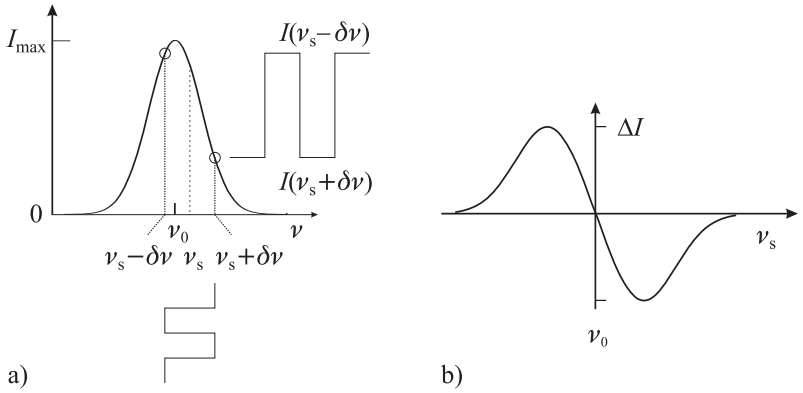


Figure 2.14: a) Power behind a bell shaped resonance line centred at ν_0. b) The difference signal $\Delta I(\nu_s) = I(\nu_s + \delta\nu) - I(\nu_s - \delta\nu)$ of Fig. 2.14 a) leads to a discriminant curve with a zero crossing at $\nu_s = \nu_0$.

Hence, the difference signal can be utilised as an unambiguous error signal for a servo unit. There is a positive error signal if the frequency of the oscillator is lower than the centre frequency of the reference and vice versa. The integrated difference signal can be obtained by using a so-called lock-in amplifier. This device integrates a signal synchronously to the frequency of a modulation signal thereby changing the polarity of the input signal after half the modulation period. This phase-sensitive detection technique is capable of detecting very weak periodic signals buried in a very strong background since the dc and all frequency components except the one at the modulation frequency are integrated to zero.

Square-wave modulation of the signal is often not appropriate for several reasons. Square-wave modulation introduces very high harmonics that can mix high-frequency noise components to the base band by so-called aliasing (Section 3.5.3). Furthermore, the rapid switching may introduce ringing effects in a narrow filter which may depend on the actual frequency spectrum of the modulation. Hence, often a harmonic modulation is used where the amplitude modulated signal is detected as a function of a frequency modulation either of the interrogating oscillator or the resonance of the reference (see Fig. 2.15).

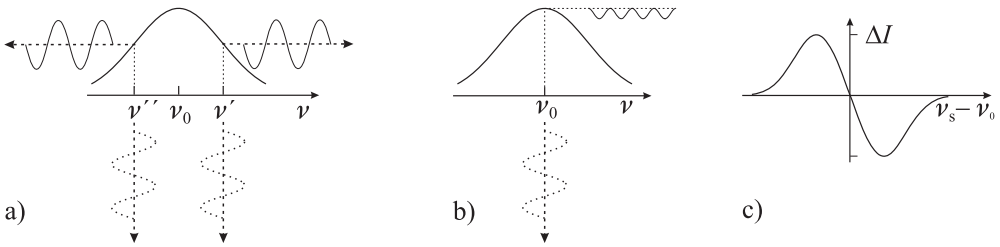


Figure 2.15: Harmonic modulation of the oscillator frequency (dots) tuned to different frequencies of the resonance line results in different modulations of the signal amplitude (solid lines). a) Modulating the frequency near the high-frequency halfwidth (ν') leads to a reduction of the signal for positive frequency excursions. In contrast, on the low-frequency slope (ν'') positive frequency excursions lead to an increased signal. b) Modulation of the frequency near the centre frequency (ν_0) leads to a modulation of the signal with twice the modulation frequency. c) Phase-sensitive detection of the amplitude modulated signal results in an anti-symmetric discriminant curve.

The phase-sensitively signal, detected, e.g., by means of a synchronous detector or lock-in amplifier, shows an anti-symmetric discriminant curve (see Fig. 2.15 c) as a function of the difference in the frequency of the oscillator and the centre frequency of the resonance. As can be seen from Fig. 2.15 a) the sign change of the error signal results from the phase shift by π of the modulated signals for frequencies above or below resonance.

2.4 Electronic Servo Systems

In the servo control unit of Fig. 1.3 the error signal generated by the discriminator is converted into a servo signal which is fed back to the oscillator when the servo loop is closed. In this section the function and frequency response of some widely used electronic elements in the servo control units are discussed.

2.4.1 Components

To determine the frequency response of the servo system the particular electronic components, e.g., electronic amplifiers or filters and also other mechanical and thermal components can be described by electronic equivalent circuits.

2.4.1.1 Low-pass Filter

As a first example, consider an electronic low pass (Fig. 2.16). As the low-pass behaviour is also found in mechanical components, amplifiers with a high-frequency cut-off, piezoelectric actuators or others, the electronic low pass often serves as an equivalent circuit for such devices.

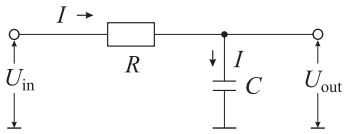


Figure 2.16: Low-pass filter.

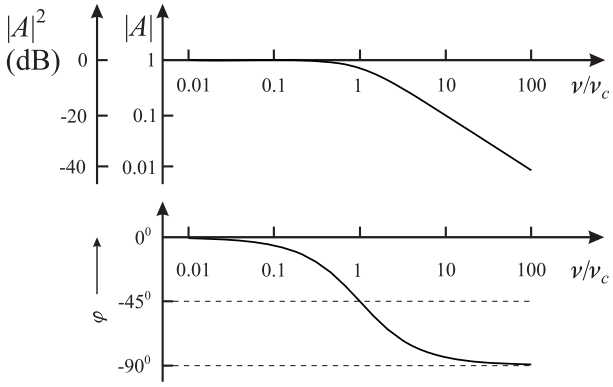


Figure 2.17: Bode plot of a low-pass filter.

To determine the frequency-dependent complex amplitude gain

$$A(\omega) \equiv \frac{U_{\rm out}}{U_{\rm in}} \tag{2.69}$$

one finds from Fig. 2.16 that the applied voltage $U_{\rm in}$ leads to a current I determined by the purely ohmic resistance R and the frequency-dependent complex impedance $R_C = 1/(i\,\omega C)$ of the capacitor. This current I at the same time leads to a voltage drop at the capacitor which represents the output voltage $U_{\rm out}$. Hence, it follows that

$$I = \frac{U_{\rm in}}{R + \frac{1}{i\omega C}} \qquad \text{and} \qquad I = \frac{U_{\rm out}}{\frac{1}{i\omega C}}. \tag{2.70}$$

Equating both results leads to the complex amplitude response (see (2.69))

$$A(\omega) = \frac{1}{1 + i\omega RC} = \frac{1}{1 + \omega^2\, R^2\, C^2} - i\frac{\omega RC}{1 + \omega^2\, R^2\, C^2}. \tag{2.71}$$

The behaviour of $A(\omega)$ is characterised by the dimensionless quantity ωRC where RC is a reciprocal characteristic angular "corner" frequency ω_c with

$$\omega_c = 2\pi\nu_c \equiv \frac{1}{R\,C}. \tag{2.72}$$

The frequency dependence of the modulus $|A|$ and the phase φ are calculated as

$$\begin{aligned} |A| &= \sqrt{\frac{1^2+\omega^2\,R^2\,C^2}{(1+\omega^2\,R^2\,C^2)^2}} = \frac{1}{\sqrt{1+\omega^2R^2C^2}} \qquad \text{and} \\ \tan\,\varphi &= \frac{\Im\mathrm{m}\,\{A(\omega)\}}{\Re\mathrm{e}\,\{A(\omega)\}} = -\omega RC \qquad \text{or} \\ \varphi &= -\arctan\omega RC. \end{aligned} \tag{2.73}$$

The modulus of the amplitude response is nearly constant for frequencies below the corner frequency ($\nu \ll \nu_c$ or $\omega RC \ll 1$) and rolls off with $1/\nu$ for frequencies $\omega RC \gg 1$ well above the corner frequency (Fig. 2.17).

Similarly as in the Bode plot of the resonant filter, the response of an electronic element in the servo loop is often characterised by a logarithmic quantity rather than by the ratio of the amplitudes (2.69). The unit of the decadic logarithm of the power ratio $P_{\mathrm{out}}/P_{\mathrm{in}}$ is 1 B (Bel) = 10 dB (deciBel). Hence, the power ratio in dB is

$$10\log\frac{P_{\mathrm{out}}}{P_{\mathrm{in}}} = 10\log\frac{U_{\mathrm{out}}^2}{U_{\mathrm{in}}^2} = 20\log\frac{U_{\mathrm{out}}}{U_{\mathrm{in}}}. \tag{2.74}$$

If a broad-band amplifier with constant gain of say 40 dB (amplitude gain A = 100) is put in series to the low-pass filter in the servo loop, the Bode plot of the amplitude gain in Fig. 2.17 is raised by 40 dB in the combined system. Above the characteristic frequency the gain of the low-pass filter is reduced by -20 dB per decade or about 6 dB per octave. The frequency where the gain is A = 1 (or 0 dB) is referred to as the unity-gain frequency. Increasing the (frequency independent) proportional gain leads to a higher unity-gain frequency. Adding several components each with a frequency-dependent phase shift may result in a total phase shift exceeding 180° for Fourier frequencies higher than a particular frequency. Hence, the negative feedback is converted into a positive feedback and the frequency fluctuations are amplified by the servo loop provided the gain is not $A \leq 1$.

2.4.1.2 Operational Amplifier

To increase the gain in the servo loop in general, circuits with operational amplifiers are used. The operational amplifier symbolised in Fig. 2.18 is an integrated electronic device with specific properties.

It is often powered by a symmetric voltage $\pm U_S$ of $12V \leq U_S \leq 15V$. The operational amplifier has two inputs realised as ground free differential inputs. The output voltage U_{out} with respect to ground can have values between $+U_{\max}$ and $-U_{\max}$ with $U_{\max}$ being slightly lower than the supply voltage $+U_S$. The open-loop gain A is large

$$10^5 \lesssim \frac{U_{\mathrm{out}}}{U_{\mathrm{in}}} \lesssim 10^6. \tag{2.75}$$

Hence, assuming $U_{\max} = 10\ V$ and $A = 10^5$ an input voltage of 100 μV is sufficient to saturate the output voltage with the gain curve of Fig. 2.19. With the negative input as reference the output voltage follows the polarity of the input voltage. Hence, this input is called the

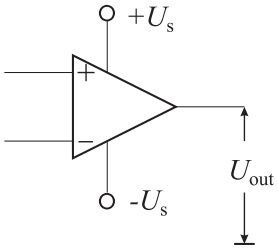


Figure 2.18: Symbol of an operational amplifier.

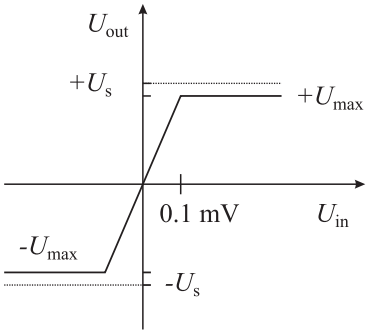


Figure 2.19: Gain curve of an operational amplifier.

non-inverting input in contrast to the second one (inverting input) where the output voltage changes polarity when the inverting input is referenced to the non-inverting input. As a consequence of the high gain, $+U_{\mathrm{in}} \approx -U_{\mathrm{in}}$ holds as long as the output voltage is in the linear range (Fig. 2.19).

Despite the peculiar gain curve (Fig. 2.19) of such an element where the output voltage is saturated at the smallest input voltage, the operational amplifier is a very useful device if used with external feedback circuitry. To describe the behaviour of an operational amplifier with feedback, two "golden rules" of the operational amplifier can be applied [22] as follows.

I. The output voltage of the operational amplifier is always such that the voltage difference between the two inputs is zero. (2.76)

II. The input currents are very low (zero). (2.77)

These rules, however, only apply if the operational amplifier does not saturate.

2.4.1.3 Non-inverting Proportional Amplifier

In the circuit of Fig. 2.20 the input voltage is applied to the non-inverting input and a small fraction of the output voltage is fed back to the inverting input via the resistor R_2. As a consequence of the negative feedback the output voltage will be such that the partial voltage fed back compensates the voltage difference between the input terminals, i.e. golden rule I (2.76) applies.

Since the inverting and the non-inverting inputs are at the same potential ($U_{R_2} \approx U_{\mathrm{in}}$) according to golden rule II (2.77) there is almost no input current and hence, the voltages are divided as follows

$$U_{R_2} = U_{\mathrm{out}} \frac{R_1}{R_1 + R_2} \approx U_{\mathrm{in}}. \tag{2.78}$$

As a consequence, the gain

$$A \equiv \frac{U_{\mathrm{out}}}{U_{\mathrm{in}}} = \frac{R_1 + R_2}{R_2} \tag{2.79}$$

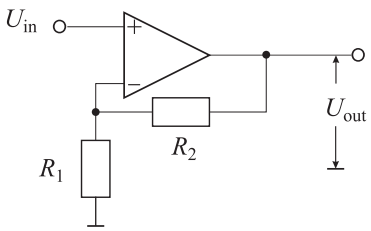


Figure 2.20: Non-inverting amplifier.

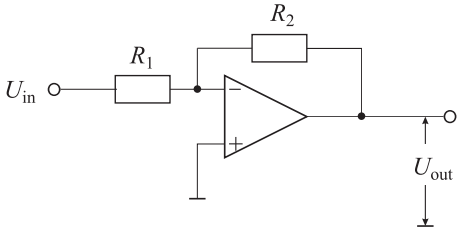


Figure 2.21: Inverting amplifier.

no longer depends on the open-loop gain but rather on the values of the external resistors R_1 and R_2 and the circuit represents a proportional amplifier. For high frequencies, however, depending on the particular type of operational amplifier the gain is reduced and the frequency is that of a low-pass filter including the associated phase shift.

2.4.1.4 Inverting Amplifier

In Fig. 2.21 where the non-inverting input is taken as the reference, a positive signal at the inverting input will lead to a negative output voltage from where a small part is fed back to the inverting input via the resistor R_2. According to golden rule II (2.77) no current flows into the input of the operational amplifier and, consequently, the input current I_{in} equals the output current I_{out}

$$I_{\text{in}} = \frac{U_{\text{in}}}{R_1} = \frac{U_{\text{out}}}{R_2} = I_{\text{out}}. \tag{2.80}$$

The proportional gain is

$$A = \frac{U_{\text{out}}}{U_{\text{in}}} = -\frac{R_2}{R_1} \tag{2.81}$$

where the minus sign reflects the fact that the output voltage has the opposite polarity from the input signal. In comparison to the non-inverting amplifier which according to golden rule II (2.77) has an infinite input impedance, this amplifier draws a disadvantageously high input current when at high gain R_2/R_1 the input resistance R_1 is small.

2.4.1.5 Integrator

The inverting amplifier is converted into an integrating amplifier (Fig. 2.22) if the ohmic resistor R in the feedback network is replaced by a capacitor with the capacitance C.

As a result of the frequency-dependent impedance $1/\omega C$ of the capacitor the amplitude gain decreases like $1/\nu$ (Fig. 2.23). To understand the way the network acts let us assume that a voltage U_{in} is applied to the inverting input. Applying golden rule I (2.76) one finds both input terminals at the same potential. Hence, the current $I = U_{\text{in}}/R$ flows via the feedback loop and thus C is charged. As a consequence the voltage $U = Q/C$ across the capacitor C increases. This corresponds to a temporally increasing resistance and, hence, the output

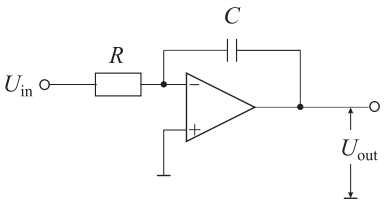


Figure 2.22: Operational amplifier as integrator.

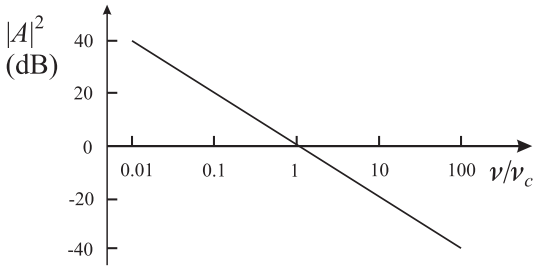


Figure 2.23: Frequency-dependent gain of an integrator.

voltage $U_{\rm out}$ increases linearly with time. If at a later instant the input voltage is set to zero ($U_{\rm in} = 0$) and the output $U_{\rm out}$ is not saturated, the voltage between the input terminals is zero (see (2.76)). Hence, no current flows via R leaving the charge on the capacitor unchanged and the output voltage remains constant. If, however, the input voltage $U_{\rm in}$ is applied again the capacitor will be charged further. Applying golden rule II (2.77) the input current $I_{\rm in}$ flows via the feedback loop to the output $I_{\rm out}$

$$I_{\rm in} = \frac{U_{\rm in}}{R} = -\frac{d\,Q}{d\,t} = -\frac{d}{d\,t}(C\,U_{\rm out}) = -C\,\frac{d\,U_{\rm out}}{d\,t}. \tag{2.82}$$

Integration of (2.82) leads to

$$U_{\rm out} = \frac{1}{RC}\int U_{\rm in}\,dt + const \tag{2.83}$$

and it is seen from (2.83) that the output voltage is proportional to the time-integrated input voltage.[15] In reality, however, the output voltage of an integrator may increase until saturation even if the inputs are shorted as a result of unavoidable offsets in the operational amplifier. External terminals of the operational amplifiers can be used to adjust these offsets. To properly take into account its temperature dependence, the adjustment of the offset has to be performed under the actual working conditions.

2.4.1.6 P-I Amplifier

Consider the case of Fig. 2.24 where the feedback loop comprises a capacitor and an ohmic resistor in series. For high frequencies the gain of the device is determined by the resistance and for small frequencies by the reactance $1/(\omega C)$ of the capacitor (Fig. 2.25). Similarly, to (2.74) one derives

$$|A| = \sqrt{1 + \frac{1}{(\omega^2 R^2 C^2)^2}} \tag{2.84}$$

and

$$\varphi = -\arctan\frac{1}{\omega RC}. \tag{2.85}$$

[15] Without the use of an input resistor R, likewise the input current can be integrated.

The corner frequency ν_c where the integrating gain (dotted line in Fig. 2.25 top) intersects the proportional gain (dashed line in Fig. 2.25 top) of the "Proportional-Integrating" (P-I) amplifier is determined by $\omega_c RC = 1$ (see (2.84)).

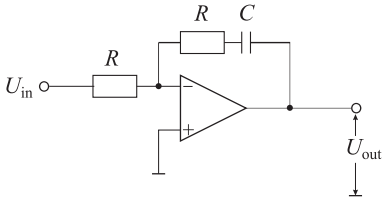


Figure 2.24: Operational amplifier with a feedback network as a P-I amplifier.

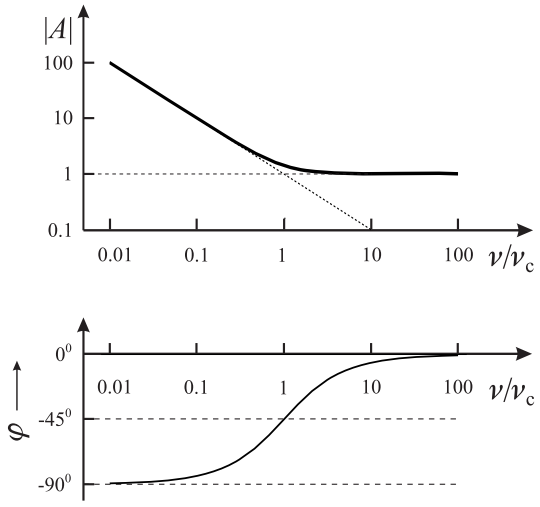


Figure 2.25: Frequency-dependent amplitude gain of a P-I amplifier.

2.4.2 Example of an Electronic Servo System

The combination of various circuits discussed so far and similar ones, allows one to tailor a servo controller with the desired frequency response. As an example, consider a servo system (Fig. 2.26) that has been used to stabilise the frequency of a diode laser system to a Fabry–Pérot interferometer [23]. Two independent servo elements were used to vary the frequency of the laser namely the current input of the laser diode (fast input) and a piezo element for variation of the length of the laser cavity (slow output). Hence, the servo controller of Fig. 2.26 has a fast output for the current input of the laser diode and a slow output for the piezo actuator. Since the driver providing the high voltage for the piezo actuator had a frequency response of a low pass with a 3 dB frequency of 32 Hz the operational amplifier D is used as an integrator for $\nu < 32$ Hz and has a constant gain of 0 dB for higher frequencies leading to an integrating behaviour of the combination. The corner frequency of $\nu_c \approx 32$ Hz where the integrating characteristics changes to the proportional gain is given by $1/(2\pi\nu_c C) = R$ of the capacitor C = 470 nF and the resistor R = 10 kΩ. The combined frequency response of the fast output of the servo controller together with the fast input of the laser diode was chosen to show an overall integrating behaviour for frequencies above about 4.5 kHz. As the measured amplitude transfer function of the laser showed an integrating $(1/\nu)$ behaviour for frequencies 40 kHz $\leq \nu \leq$ 350 kHz and a doubly integrating $(1/\nu^2)$ behaviour for frequencies above about 350 kHz, the frequency-dependent gain of the operational amplifier B and C was chosen as shown in

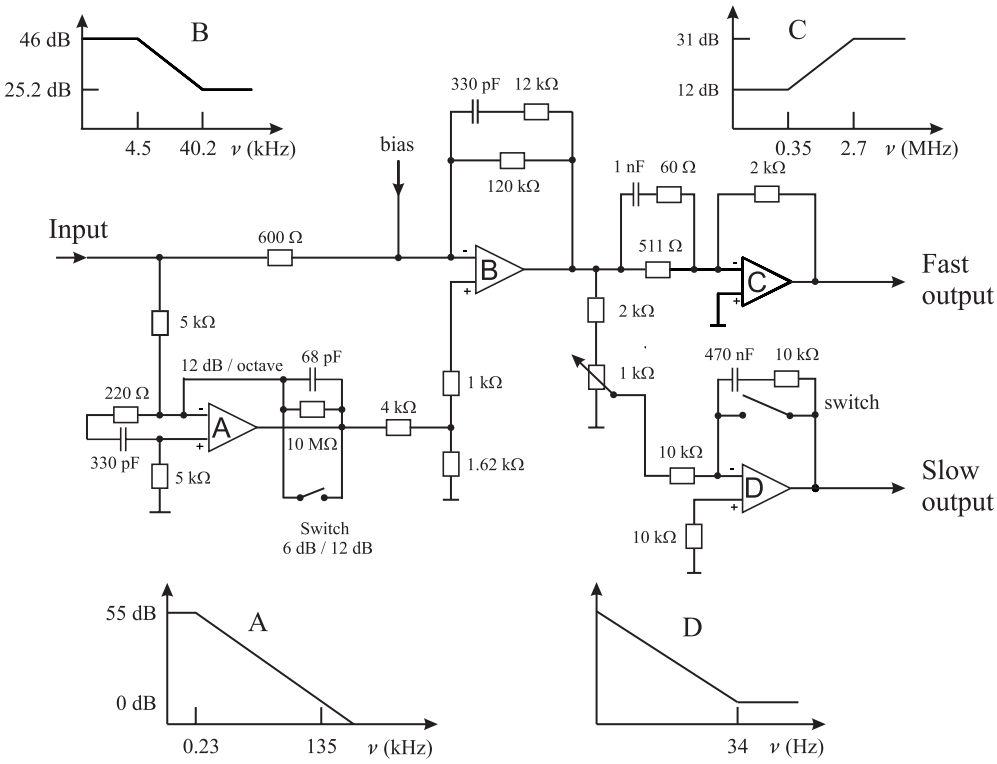


Figure 2.26: Simplified example of a servo controller for a frequency stabilised laser system [23] with the Bode plots for the particular operational amplifiers.

the insets B and C of Fig. 2.26. Consider operational amplifier B. For very low frequencies the capacitor (330 pF) in the feedback loop has a very high impedance and the amplitude gain is given by $120\,\mathrm{k\Omega}/600\,\Omega = 200$ or 46 dB. For very high frequencies the impedance of the capacitor can be neglected and the feedback resistance of 10.9 kΩ is given by the two parallel resistors of 120 kΩ and 12 kΩ leading to a gain of 25.2 dB. The lower corner frequency ($\nu_c = 4.5\,\mathrm{kHz}$) is determined as the frequency where the combined impedance $1/(2\,\pi\nu_c\ 330\,\mathrm{pF})$ + 12 kΩ equals 120 kΩ whereas the upper corner frequency is given for $1/(2\,\pi\nu_c\ 330\,\mathrm{pF}) = 12\,\mathrm{k\Omega}$.

Operational amplifier A acts as a second integrator for frequencies above 230 Hz if the switch in the feedback loop is open.[16] The frequency response in the doubly integrating regime rolls off with $1/\nu^2$. The amplitude gain of 55.2 dB results from the gain of the amplifier (10 MΩ/5 kΩ) and the subsequent reduction at the voltage divider (1.62 kΩ/5.62 kΩ). The gain is reduced by a factor of 4 per octave, i.e. for a factor of two increase in frequency, corresponding to a power reduction of about 12 dB per octave. This second integrator pro-

[16] The R-C network at the input of operational amplifier A is used to prevent oscillation of the operational amplifier stage.

vides the high gain at low Fourier frequencies to allow one to suppress frequency fluctuations (see (2.68)) in the regime where technical noise contributes the most. For a more detailed discussion of electronic laser stabilisation schemes the reader is referred to [24].

3 Characterisation of Amplitude and Frequency Noise

The frequency and amplitude of even the most advanced oscillators are not really constant in time, but fluctuate. In the previous chapter we have analysed the situation where these quantities were modulated in a strictly deterministic way. The harmonic modulation of the amplitude of an oscillation was found to lead to discrete sideband frequencies whereas the exponential temporal decay of the amplitude resulted in a continuous band of frequencies. In both cases, for any instant in the past or in the future one was able to predict the exact value of the instantaneous amplitude, frequency and the phase. For real oscillators, however, a large variety of physical processes that are not under control can affect these quantities in a complicated way. As a result, the amplitude, phase or frequency of any oscillator will fluctuate in an irregular way that in general can no longer be represented by an analytic function of time. These unwanted fluctuations are often referred to as noise or jitter. To describe these fluctuations, statistical measures have to be applied. The characterisation of frequency standards in terms of statistical quantities nevertheless allows one to select the most suitable standard or to infer information about possible sources that degrade the performance of the standard.

For frequency standards one deals in general with the best available oscillators where often the statistical "modulations" of amplitude and phase are small. Consequently, one uses a model of the oscillator where the instantaneous output signal of the oscillator is written as in (2.10)

$$U(t) = [U_0 + \Delta U_0(t)] \cos\left(2\pi\nu_0 t + \phi(t)\right). \tag{3.1}$$

The quantity $U(t)$ may represent, e.g., the signal from a quartz oscillator or the electric field of an oscillator in the microwave or optical domain. In contrast to (2.10) $\Delta U_0(t)$ now represents random rather than deterministic amplitude fluctuations around U_0. Similarly, the fluctuations $\phi(t)$ of the phase result from a random process. In (3.1) it is furthermore assumed that the fluctuations of the phase and amplitude are orthogonal meaning that no amplitude fluctuations are transferred to phase fluctuations and vice versa. To compare frequency standards operating at different frequencies ν_0 it is helpful to define the normalised phase fluctuations

$$x(t) \equiv \frac{\phi(t)}{2\pi\nu_0} \tag{3.2}$$

which are sometimes referred to as the phase time. Similarly, rather than using the fluctuations of the instantaneous frequency (see (2.11)) itself, the instantaneous fractional (or normalised) frequency deviation

$$y(t) \equiv \frac{\Delta\nu(t)}{\nu_0} = \frac{dx(t)}{dt} \tag{3.3}$$

is defined where (2.12) has been utilised to derive the latter equation.

3.1 Time-domain Description of Frequency Fluctuations

Consider the time sequence of a fluctuating quantity measured as a continuous function $y(t)$ (Fig. 3.1 a) or as series of discrete readings y_i (Fig. 3.1 b). The latter may be obtained, e.g.,

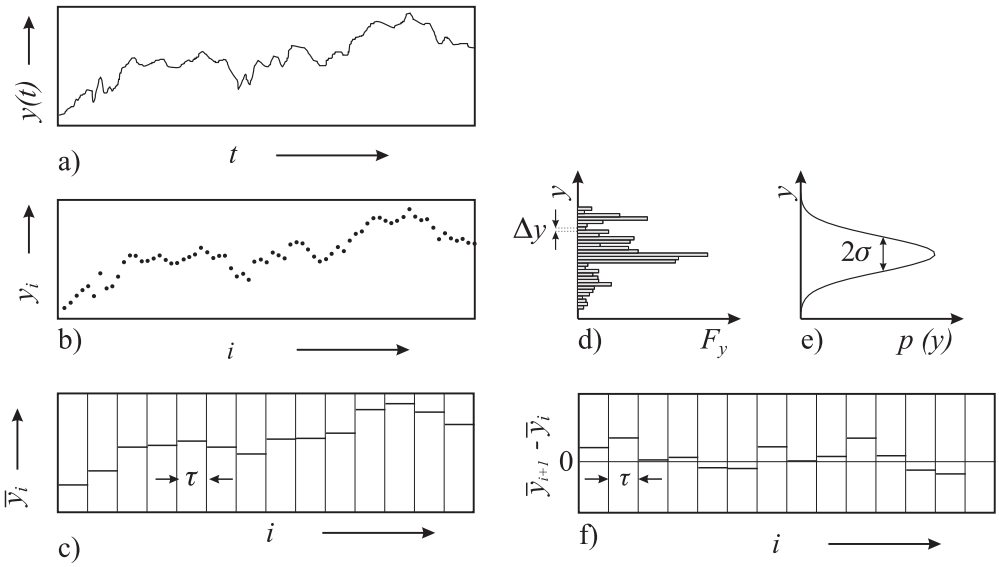


Figure 3.1: a) Continuous time sequence $y(t)$. b) Discrete time series y_i of a fluctuating quantity. c) Consecutive mean values of $y(t)$ (see Fig. 3.1 a) where the values $\overline{y}_i$ are taken during a duration τ. d) Histogram F_y corresponding to the distribution of y in the bin size Δy. e) Corresponding Gaussian probability density $p(y)$. f) $\overline{y}_{i+1} - \overline{y}_i$ used to compute the Allan variance according to (3.13).

if the measurement of $y(t)$ was performed by using a frequency counter. As a result the continuous function $y(t)$ is reduced to a discrete series of consecutive measurements averaged over the measurement time τ

$$\overline{y}_i = \frac{1}{\tau} \int_{t_i}^{t_i+\tau} y(t)dt \tag{3.4}$$

(Fig. 3.1 c), referred to as the normalised frequency deviation averaged over the duration τ. The experimental determination of these quantities will be discussed in Section 3.5. As repeated measurements of y_i in general differ from each other (see Fig. 3.1 b), we recall in the following the statistical means usually employed to characterise such a data set. It is well known that the mean value and the square of the experimental standard deviation are

$$\overline{y} = \frac{1}{N} \sum_{i=1}^{N} y_i \tag{3.5}$$

and

$$s_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \overline{y})^2 = \frac{1}{N-1}\left[\sum_{i=1}^{N} y_i^2 - \frac{1}{N}\left(\sum_{i=1}^{N} y_i\right)^2\right], \tag{3.6}$$

respectively. The standard deviation of the mean is

$$s_{\overline{y}} = \frac{s_y}{\sqrt{N}}. \tag{3.7}$$

s_y is a measure of the width of the histogram F_y (see Fig. 3.1 d) where the values of $y(t)$ (or y_i) have been grouped into bins of width Δy as a function of $y(t)$.

Often the fluctuations of $y(t)$ are thought to result from a statistical process. If the process causing the fluctuations of $y(t)$ is stationary,[1] according to the central-limit theorem of probability theory one expects that for $T \to \infty$, F_y evolves into a Gaussian probability density

$$p(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y-\overline{y})^2}{2\sigma^2}\right) \tag{3.8}$$

(Fig. 3.1 d) with variance σ^2. The statistical process is characterised by the expectation value

$$\langle y\rangle \equiv \int_{-\infty}^{\infty} y p(y) dy \tag{3.9}$$

and the variance

$$\sigma^2 = \int_{-\infty}^{\infty} (y - \langle y\rangle)^2 p(y) dy. \tag{3.10}$$

Using the notation of (3.9), (3.10) can be written as

$$\sigma^2 = \langle (y - \langle y\rangle)^2\rangle = \langle y^2 - 2y\langle y\rangle + \langle y\rangle^2\rangle = \langle y^2\rangle - \langle y\rangle^2. \tag{3.11}$$

The expectation value (3.9) and variance (3.10) of a statistical process can be only *estimated* from the measured finite sequence of the fluctuating quantity in such a way that the mean value (3.5) is an estimate for the expectation value $\langle y\rangle$ of the Gaussian process and the square of the standard deviation (3.6) is an estimate of its variance σ^2.

Besides defining mean value and standard deviation from consecutive measurements of, e.g., the frequency of a single oscillator, analogously, the mean value and standard deviation can be defined as a statistical average for a sample of identical oscillators. For a stationary process such a sample average is independent of the chosen time of the measurement. For an ergodic process,[2] σ^2 can be estimated either from the time average or from the sample average.[3]

[1] A statistical process is called stationary if the statistical measures describing the process, e.g., the mean value or the variance, are time independent.

[2] A process where the average over an infinite number of samples is identical to the infinite time average ($\langle y\rangle = \overline{y}$) is referred to as an ergodic process.

The use of the statistical tools of mean value and standard deviation meets with difficulties if applied to fluctuating quantities with correlations. This can be seen if one divides the times series of Fig. 3.1 a) into equidistant intervals like in Fig. 3.1 c). A quick look reveals that the data of Fig. 3.1 a) or b) within each subset scatter much less than the data within the total interval. The corresponding experimental standard deviations (3.6) calculated within each sub-interval in general are much smaller than the one calculated from the entire data set. This indicates that the adjacent data points are not independent of each other but are somehow correlated. Consequently, the standard deviation of the mean is not reduced by $1/\sqrt{N}$ (see (3.7)) for N as would be the case for uncorrelated data. Thus, the determination of standard deviations from different subsets of the data can be used to get information about the existence of correlations. It has to be pointed out that the statistics of a fluctuating quantity with correlations can sometimes be well described by a Gaussian distribution and, hence, the lack of this property cannot be used to identify correlations.

3.1.1 Allan Variance

To make a meaningful estimate of the statistical process in the presence of correlations one has to specify the number N of measurements (samples), the measuring time τ of a single sample and the time T between consecutive measurements which may differ from τ by the dead time $T - \tau$ (see Fig. 3.2). After having done this one can readily define a so-called

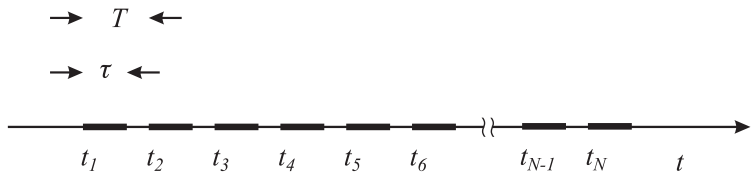


Figure 3.2: Measurement cycle.

N-sample variance for this data set in analogy to (3.6) as [4]

$$\sigma^2(N,T,\tau) = \frac{1}{N-1}\sum_{i=1}^{N}\left(\overline{y}_i - \frac{1}{N}\sum_{j=1}^{N}\overline{y}_j\right)^2 \tag{3.12}$$

for a given number N of samples and given values of T and τ (see Fig. 3.2). It is now generally agreed [25] to follow a proposition made by Dave Allan [26,27] and to select from all possible sample variances the expectation value of the so-called two-sample variance with $N = 2$ and

[3] Stationarity and ergodicity are mathematical properties that are often assigned to the statistical processes used to model the fluctuations of real frequency standards. As a result of the limited time available for any measurements and the limited number of identical frequency standards at hand these properties cannot be proven but merely represent reasonable assumptions. Care has to be taken when the results derived on these assumptions are applied to practical cases. During their lifetimes, e.g., frequency standards may become more "noisy" and stationarity may not be granted over this time.

[4] To be more specific, there is more than one possible definition for the N sample variance. The various definitions differ by the pre-factor and each have their advantages for a particular type of noise [25].

$T = \tau$. Hence, this so-called Allan variance $\sigma_y^2(2,\tau,\tau)$ which is alternatively referred to as short-hand notation $\sigma_y^2(2,\tau)$ or $\sigma_y^2(\tau)$, is defined using (3.12) as

$$\sigma_y^2(\tau) = \left\langle \sum_{i=1}^{2} \left(\overline{y}_i - \frac{1}{2} \sum_{j=1}^{2} \overline{y}_j \right)^2 \right\rangle = \frac{1}{2} \left\langle \left(\overline{y}_2 - \overline{y}_1 \right)^2 \right\rangle. \tag{3.13}$$

The Allan variance and its square root sometimes termed the Allan (standard) deviation is based on differences of adjacent frequency values rather than on frequency differences from the mean value, as is the "true" standard deviation.

Alternatively, the Allan variance can be determined from the phase deviation $\phi(t)$ or the normalised phase deviation $x(t)$. For a given measuring interval τ it follows from (3.3) that

$$\overline{y}_i = \frac{\overline{x}_{i+1} - \overline{x}_i}{\tau} \tag{3.14}$$

which after insertion into (3.13) gives

$$\sigma_y^2(\tau) = \frac{1}{2\tau^2} \left\langle \left(\overline{x}_{i+2} - 2\overline{x}_{i+1} + \overline{x}_i \right)^2 \right\rangle. \tag{3.15}$$

3.1.1.1 Practical Determination of the Allan Variance

In the experiment the Allan variance of a particular oscillator "1" may be determined, e.g., from a beat note (3.87), i.e., the frequency difference with respect to a second oscillator "2" (reference oscillator) using a counter gated with the measuring time τ. According to the definition it has to be ensured that there is no dead time between two adjacent measurements. From the squared normalised frequency differences between two adjacent pairs ν_i and ν_{i+1} the mean value is computed and divided by 2 to give the Allan variance $\sigma_{\mathrm{y,tot}}^2$ for the particular measuring time τ. To make a good approximation of the expectation value ($\langle\rangle$) of (3.13) a sufficiently large number of frequency differences has to be used. The procedure has to be repeated for the different times τ and may lead to Allan deviations such as the ones displayed in Fig. 3.3. In Fig. 3.3 the Allan deviations $\sigma_y(\tau)$ of various frequency standards and oscillators are compared with frequencies ranging from the microwave region to the optical regime.

In practice the Allan variance is determined in slightly different ways in order to allow for the minimum measurement time necessary to retrieve the full information required. The counter is set to the shortest gate time τ_0 where the Allan variance is to be determined and the frequency difference $\overline{y}_{i,\tau_0}$ between the oscillators is measured repeatedly and the data are stored making sure that no deadtime occurs during the data aquisition (see Fig. 3.4 a). To derive the data for longer times, e.g., $\tau = 3\tau_0$ the consecutive values of $\overline{y}_{1,\tau} = (\overline{y}_{1,\tau_0} + \overline{y}_{2,\tau_0} + \overline{y}_{3,\tau_0})/3$, $\overline{y}_{2,\tau} = (\overline{y}_{4,\tau_0} + \overline{y}_{5,\tau_0} + \overline{y}_{6,\tau_0})/3$, $\overline{y}_{3,\tau} = \cdots$ are determined in a post processing (Fig. 3.4 b) to estimate the Allan variance for the time $\tau = 3\tau_0$ and accordingly for all other times τ.

To make even better use of the stored data, roughly n times more values of $\overline{y}_{i,\tau=n\tau_0}$ can be obtained if the data processing is done in the way depicted in Fig. 3.4 c) where $\overline{y}_{1,\tau} = (\overline{y}_{1,\tau_0} + \overline{y}_{2,\tau_0} + \overline{y}_{3,\tau_0})/3$, $\overline{y}_{2,\tau} = (\overline{y}_{2,\tau_0} + \overline{y}_{3,\tau_0} + \overline{y}_{4,\tau_0})/3$, $\overline{y}_{3,\tau} = \cdots$ are taken.

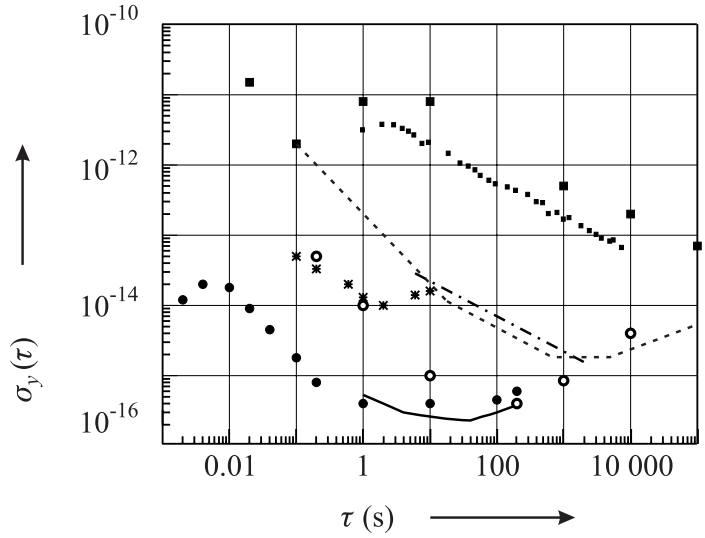


Figure 3.3: Allan deviation $\sigma_y(\tau)$ as a function of the measuring time τ for various highly stable oscillators used as frequency standards and discussed in this book: commercial caesium atomic clock (big squares: [28], small squares: [29]), hydrogen maser (typical, dashed line; see also Fig. 8.5), caesium fountain (dashed dotted line) [18], sapphire loaded cavity microwave-oscillator (thick line) [30], superconducting-cavity stabilised microwave oscillator (open circles [30], laser stabilised to a Fabry–Pérot cavity (full circles) [31], Ca stabilised laser (asterisks) [32].

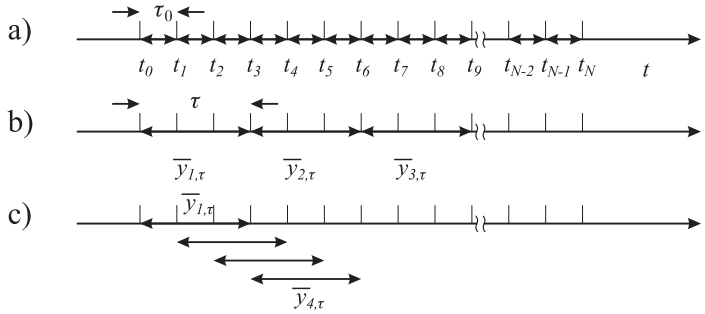


Figure 3.4: Alternative methods of calculating the Allan variance.

If the reference oscillator is known to be of superior stability with respect to the oscillator under test, the Allan variance is a measure of the instability of the latter one. If the Allan variance of two identical oscillators "1" and "2" is taken one is led to assume that both oscillators contribute equally to the instability and the measured Allan variance $\sigma_{\mathrm{y,tot}}$ is attributed evenly

to both oscillators as follows

$$\begin{aligned} \sigma^2_{\mathrm{y,tot}}(\tau) &= \sigma^2_{y,1}(\tau) + \sigma^2_{y,2}(\tau) \qquad \text{and} \\ \sigma_{y,1}(\tau) &= \sigma_{y,2}(\tau) = \frac{1}{\sqrt{2}}\sigma_{\mathrm{y,tot}}(\tau). \end{aligned} \tag{3.16}$$

The Allan variance $\sigma^2_y(\tau)$ is a useful time-domain measure of the frequency instability of an oscillator. It allows one to select the ideal oscillator for a particular application. As an example, consider the Allan variance of a typical hydrogen maser and one of the best frequency stabilised lasers shown in Fig. 3.3. The latter one has an optimum stability of $\sigma_y \leq 5 \times 10^{-16}$ for measurement times τ between 1 s and 100 s whereas the former one reaches its optimum frequency stability at one to several hours. In the plot of $\sigma_y(\tau)$ one often can identify regions where the frequency instability of a particular frequency standard follows a well defined power law. We will see in Section 3.1.1.2 that a linear drift leads to an Allan deviation proportional to τ. The relationship between the τ^{-1} and $\tau^{-1/2}$ dependencies recognised, e.g., in the plot of the hydrogen maser (Fig. 3.3) and the underlying noise processes will be discussed in Section 3.3.

Besides stochastic fluctuations, deterministic variations of the frequency of a given oscillator have a profound impact on the measured Allan variance. In the following we will investigate two important cases, a linear frequency drift and an harmonic frequency modulation.

3.1.1.2 Influence of a Linear Frequency Drift

Consider an oscillator whose normalised frequency shows a linear drift $y(t) = at$ where a is the slope of the drift. With $\overline{y}_1 = [at_0 + a(t_0 + \tau)]/2$ and $\overline{y}_2 = [a(t_0 + \tau) + a(t_0 + 2\tau)]/2$ one calculates from (3.13)

$$\sigma_y(\tau) = \langle a\tau/\sqrt{2}\rangle = \frac{a}{\sqrt{2}}\tau \qquad \text{for linear frequency drift.} \tag{3.17}$$

Hence, a linear frequency drift leads to an Allan deviation that linearly increases with measuring time τ.

3.1.1.3 Influence of an Harmonic Modulation

Next we consider an oscillator whose frequency is modulated with a sinusoidal modulation frequency f_m as [5]

$$y(t) = \frac{\Delta\nu_0}{\nu_0}\sin(2\pi f_m t). \tag{3.18}$$

Calculating (3.13) by use of (3.18) leads to [25]

$$\sigma_y(\tau) = \frac{\Delta\nu_0}{\nu_0}\frac{\sin^2(\pi f_m \tau)}{\pi f_m \tau} \qquad \text{for modulation with sinusoidal signal.} \tag{3.19}$$

[5] In this chapter modulation and Fourier frequencies are denoted by f rather than by ν to allow for better distinction with respect to the carrier frequency.

From (3.19) one finds that the influence of the frequency modulation on the Allan deviation becomes zero for $\tau = 1/f_m$, i.e., when τ equals the modulation period $1/f_m$ or one of its harmonics, where the influence of the modulation is averaged to zero. It is maximal for $\tau \approx n/(2f_m)$ with n an odd integer.

3.1.2 Correlated Fluctuations

A simple method of identifying correlations between measured data is to plot each measured value as function of the preceding one (see Fig. 3.5). As an example of a fluctuating quantity with correlations consider the simple model

$$y_{k+1} = \alpha y_k + \epsilon \tag{3.20}$$

where each value of the fluctuating quantity y has a purely statistical contribution ϵ but also is influenced by the previous value depending on a correlation factor $0 \leq \alpha \leq 1$. In Fig. 3.5 a)

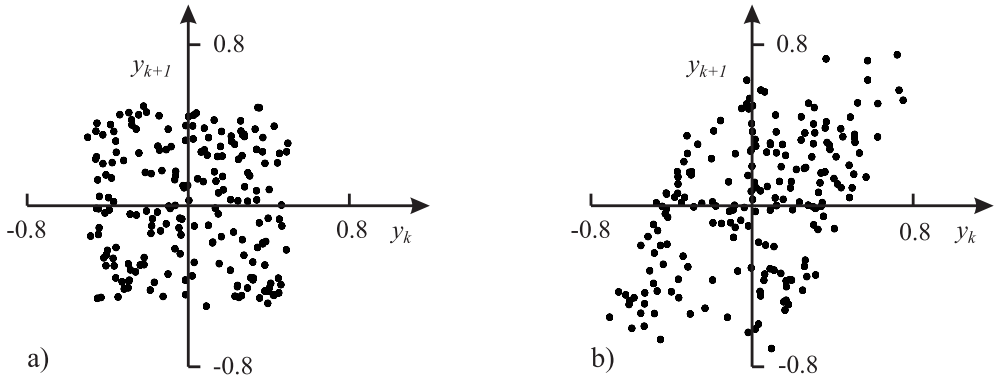


Figure 3.5: Series of 200 pseudo-random data calculated according to (3.20). a) Uncorrelated data ($\alpha = 0$). b) Correlated values using $\alpha = 0.5$.

with $\alpha = 0$ the values of $y_{k+1}(y_k)$ are evenly distributed in the four quadrants and no correlation between the data of adjacent values can be seen. This is in contrast to Fig. 3.5 b) where the correlations identify themselves by the fact that the data points cluster in the first and third quadrant. In the following we discuss methods that are better suited as the common statistical ones to cope with the correlated time series that may be encountered when dealing with frequency standards.

Commonly, any fluctuating signal $B(t)$, e.g., $y(t), U(t)$ or $\Phi(t)$ is decomposed into a purely fluctuating contribution $b(t)$ and a mean value $\overline{B(t)}$ as follows

$$B(t) = b(t) + \overline{B(t)}. \tag{3.21}$$

Consider the autocorrelation function of the signal fluctuations defined by

$$R_b(\tau) \equiv \overline{b(t+\tau)b(t)} = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} b(t+\tau)b(t)dt \tag{3.22}$$

which multiplies the signal fluctuations $b(t)$ at the instant t with the signal fluctuation $b(t+\tau)$ at the instant $t+\tau$ and takes the mean over all epochs. If the fluctuations were totally uncorrrelated, the time-averaged product $\overline{b(t+\tau)b(t)}$ would cancel for any τ. For stationary processes the autocorrelation function must be an even function since $R_b(-\tau) = R_b(\tau)$ holds. Comparing the definition of the autocorrelation function (3.22) for $\tau = 0$ and the right-hand side of (3.11) for a purely fluctuating quantity (i.e. for $\langle B\rangle^2 = 0$) the value of the autocorrelation function for τ = 0 represents the variance of the signal fluctuations

$$R_b(\tau = 0) = \sigma_b^2. \tag{3.23}$$

For very large times τ one may assume that the power fluctuations are not correlated and the autocorrelation function approaches zero for $\tau \to \infty$. It has been shown in the previous chapter that the Fourier transform of a temporal varying amplitude function represents the amplitude spectrum in the Fourier frequency domain. In the case of the statistically fluctuating power of the oscillator the time function $U(t)$ is not known but the autocorrelation function $R_b(\tau)$ might have been determined. To perform the integration in (3.22) we consider $b(t)$ as the Fourier transform $b(t) = \mathcal{F}(a(\omega))$ (see (2.19)) of a quantity $\alpha(\omega)$ whose relevance will become clear later and obtain

$$\begin{aligned} R_b(\tau) &= \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} a(\omega) e^{i(\omega t+\tau)} d\omega \int_{-\infty}^{\infty} a(\omega') e^{i\omega t'} d\omega' dt \\ &= \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} e^{i(\omega t+t')} dt \right] a(\omega) a(\omega') e^{i\omega\tau} d\omega' d\omega, \end{aligned} \tag{3.24}$$

after we have interchanged the orders of integration. In the limit $T \to \infty$ the term in square brackets can be expressed by the Dirac delta function (see (2.23)) and, hence,

$$\begin{aligned} R_b(\tau) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} a(\omega)\, a(\omega')\, e^{i\omega\tau}\, \delta(\omega+\omega')\, d\omega\, d\omega' \\ &= \int_{-\infty}^{\infty} \frac{\mid a(\omega) a(\omega') \mid}{2\pi}\, e^{i\omega\tau}\, d\omega \\ &\equiv \int_{-\infty}^{\infty} S_b(f)\, e^{i2\pi f\tau} df. \end{aligned} \tag{3.25}$$

To find the significance of $S_b(f)$ we set $\tau = 0$ in (3.25) and obtain

$$R_b(0) = \int_{-\infty}^{\infty} S_b(f) df. \tag{3.26}$$

Recalling that the left-hand side of (3.26) is the averaged square of the fluctuating quantity $b(t)$ (see (3.22)), S_b represents a power spectral density. In the case of a fluctuating voltage the spectral density is given in units of V^2/Hz.

The autocorrelation function $R_b(\tau)$ and the spectral density function $S_b(f)$ form a Fourier transform pair

$$S_b^{2-\mathrm{sided}}(f) \equiv \mathcal{F}^*\{R_b(\tau)\} = \int_{-\infty}^{\infty} R_b(\tau)\exp(-i\,2\pi f\tau)d\tau \tag{3.27}$$

$$R_b(\tau) \equiv \mathcal{F}\{S_b^{2-\mathrm{sided}}(f)\} = \int_{-\infty}^{\infty} S_b(f)\exp(i\,2\pi f\tau)df \tag{3.28}$$

where the meaning of the index {2-sided} will be discussed below. (3.27) is one form of the so-called Wiener–Khintchine theorem and allows one to determine the spectral density function from the autocorrelation function of the time-dependent signal amplitude.

If one chooses the power fluctuations $\delta P(t)$ of the oscillator rather than the amplitude fluctuations $b(t)$, the Fourier transformation of the corresponding autocorrelation function $R_{\delta P}(\tau)$ leads to a spectral density of the square of the power fluctuations (in units of W^2/Hz).[6]

Similarly, the fluctuations of the phase $\phi(t)$ with time [7] result in a power spectral density of phase fluctuations in units of rad^2/Hz. Caution is necessary as sometimes in the literature also the square root of $S_b(f) \propto a(\omega)$ (see (3.25)) is used.

The power spectral density of the frequency fluctuations in the Fourier domain represented by (3.27) is defined for Fourier frequencies $-\infty < f < \infty$ thereby extending to both the positive and negative side of the frequency spectrum. Consequently, $S_b(f)$ is referred to as the two-sided power spectral density $S_b^{2-\mathrm{sided}}(f)$. From $R_b(\tau) = R_b(-\tau)$ it follows that $S_b(f)$ is a real, non-negative and even function, i.e. $S_b(-f) = S_b(f)$. In experimental work, however, only positive frequencies are of interest. Hence, a one-sided power spectral density is often introduced for Fourier frequencies $0 \le f < \infty$ (see Fig. 3.6) with

$$S_b^{1-\mathrm{sided}}(f) = 2\,S_b^{2-\mathrm{sided}}(f). \tag{3.30}$$

As the power spectral density is a real quantity, it suffices to use a real Fourier transform pair rather than (3.27) and (3.28). Changing also the limits of the integrals the Wiener–

[6] This quantity is closely related to the so-called "Relative Intensity Noise" (RIN)

$$\mathrm{RIN}(f) \equiv \frac{S_{\delta P}}{P_0^2} \tag{3.29}$$

often used to describe the power fluctuations of lasers oscillators.

[7] Power spectral densities are used not only to describe fluctuations of a physical quantity with time but also, e.g., to characterise the roughness of a technical surface [33].

Khintchine relations for a one-sided spectral density function $S_b^{1-\mathrm{sided}}(f)$ are written as

$$S_b^{1-\mathrm{sided}}(f) = 4\int_0^\infty R_b(\tau)\cos(2\pi f\tau)d\tau \tag{3.31}$$

$$R_b(\tau) = \int_0^\infty S_b^{1-\mathrm{sided}}(f)\cos(2\pi f\tau)df. \tag{3.32}$$

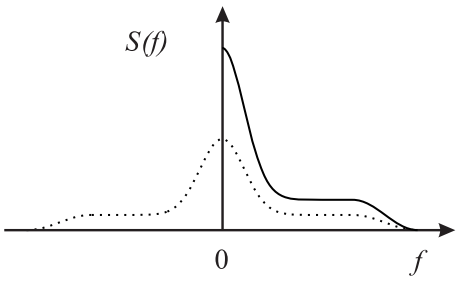


Figure 3.6: Two-sided (dots) and one-sided (line) power spectral densities.

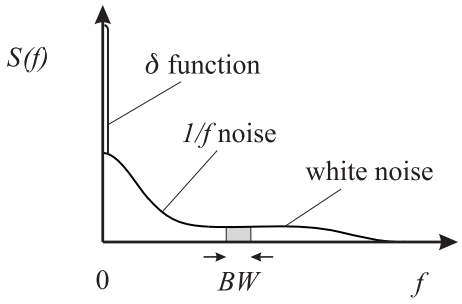


Figure 3.7: Different regimes in a power spectral density. BW: band width.

3.2 Fourier-domain Description of Frequency Fluctuations

For a reasonable frequency-stable oscillator, the instantaneous frequency $\nu(t)$ as function of time can be expected to deviate only slightly from the temporal mean $\overline{\nu}$ and

$$\Delta\nu(t) \equiv \nu(t) - \overline{\nu} \ll \overline{\nu} \tag{3.33}$$

holds.

We assume that the frequency excursions $\Delta\nu(t)$ are stationary distributed, i.e., that their distribution is time independent. Similarly as in (3.22), we define the autocorrelation function of the frequency deviations

$$R_\nu(\tau) \equiv \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\Delta\nu(t+\tau)\Delta\nu(t)dt \tag{3.34}$$

as a measure of this distribution and use the Wiener–Khintchine relationship to obtain the power spectral density of the frequency deviations from the autocorrelation of the frequency deviations

$$S_\nu^{2-\mathrm{sided}}(f) = \int_{-\infty}^{\infty} R_\nu(\tau)\exp(-i2\pi f\tau)d\tau. \tag{3.35}$$

Besides the power spectral density of the frequency fluctuations $S_\nu(f)$ the power spectral density $S_y(f)$ of the fractional frequency fluctuations $y(t)$ (see (3.3), (3.34) and (3.35)) can be found as

$$S_y(f) = \frac{1}{\nu_0^2} S_\nu(f). \tag{3.36}$$

Similarly one defines a power spectral density of phase fluctuations $S_\phi(f)$ and, by taking into account that the frequency fluctuations are essentially the time derivative of the phase fluctuations ($2\pi\Delta\nu(t) = d/dt\Delta\phi(t)$), one obtains by comparison with (3.34) and (3.35)

$$S_\nu(f) = f^2 S_\phi(f). \tag{3.37}$$

From the last two equations it follows that

$$S_y(f) = \left(\frac{f}{\nu_0}\right)^2 S_\phi(f). \tag{3.38}$$

Each of the defined three power spectral densities contains the same information.

In the typical power spectral density of Fig. 3.7 one can identify different regimes. The delta function at $f = 0$ occurs if $B(t)$ has a non-vanishing mean value $\overline{B}(t)$ and does not show up for a purely fluctuating quantity $b(t)$. The contributions at low Fourier frequencies decreasing with increasing frequency are termed $1/f$-noise. In an intermediate regime the power spectral density of the frequency fluctuations is often independent of the frequency referred to as white frequency noise. The total power contained in the frequency fluctuations is obtained from

$$\int_0^\infty S_\nu^{1-\text{sided}}(f)df = \int_{-\infty}^\infty S_\nu^{2-\text{sided}}(f)df = \langle[\Delta\nu(t)]^2\rangle = \sigma_\nu^2 \tag{3.39}$$

where we have made use of (3.23) and (3.26). From reasons of energy conservation this total power must be finite and one therefore expects that for higher frequencies the power spectral density of the frequency fluctuations decreases again (Fig. 3.7).

The determinations of spectral densities of different frequency sources reaching from quartz oscillators to atomic frequency standards has shown that the observed spectral density $S_y(f)$ can be reasonably well modelled by a superposition of five independent noise processes obeying power laws with integer exponents $-2 \leq \alpha \leq 2$

$$S_y(f) = \sum_{\alpha=-2}^{2} h_\alpha f^\alpha \tag{3.40}$$

(see Table 3.1).

The particular contributions also have characteristic appearances in the time domain (Fig. 3.8).

In a doubly logarithmic plot the particular contributions to (3.40) can be identified readily by their slope, thereby allowing identification of the causes of the noise mechanisms in the

Table 3.1: Model of a power law of the power spectral density of fractional frequency fluctuations $S_y(f) = h_\alpha f^\alpha$ and the corresponding power spectral density of phase fluctuations $S_\phi(f)$. The corresponding Allan variance $\sigma_y^2(\tau)$ derived in Section 3.3 holds for a low-pass filter with cut-off frequency f_h when $2\pi f_h \tau \gg 1$.

$S_y(f)$	$S_\phi(f)$	Type of noise	$\sigma_y^2(\tau)$
$h_{-2}f^{-2}$	$\nu_0^2 h_{-2} f^{-4}$	Random walk of frequency noise	$(2\pi^2 h_{-2}/3)\tau^{+1}$
$h_{-1}f^{-1}$	$\nu_0^2 h_{-1} f^{-3}$	Flicker frequency noise	$2h_{-1}\ln 2\tau^0$
$h_0 f^0$	$\nu_0^2 h_0 f^{-2}$	White frequency noise (Random walk of phase noise)	$(h_0/2)\tau^{-1}$
$h_1 f$	$\nu_0^2 h_1 f^{-1}$	Flicker phase noise	$h_1[1.038 + 3\ln(2\pi f_h \tau)]/(4\pi^2)\tau^{-2}$
$h_2 f^2$	$\nu_0^2 h_2 f^0$	White phase noise	$[3h_2 f_h/(4\pi^2)]\tau^{-2}$

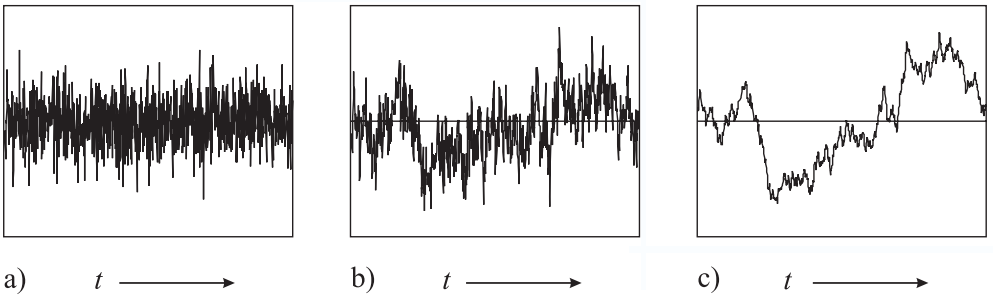


Figure 3.8: Time-domain signal with a) white frequency noise. b) $1/f$ noise. c) $1/f^2$ noise.

oscillators. The particular contributions listed in Table 3.1 can sometimes be identified in frequency standards [25]. The random walk of frequency noise ($\alpha = -2$) is often caused by the influences of environmental parameters, e.g., temperature, vibrations, etc.. Flicker frequency noise ($\alpha = -1$) is observed in active devices such as quartz crystal oscillators, hydrogen masers or laser diodes, but also in passive frequency standards like the Cs clock. White frequency noise ($\alpha = 0$) can result from thermal noise in the oscillator loop of active standards. It is also present in passive standards and may result, e.g., from the shot noise of the photons or atoms where it represents the quantum limit. Flicker phase noise ($\alpha = 1$) often results from contributions of noisy electronics whose level can be reduced by selected components. White phase noise ($\alpha = 2$) becomes important for high Fourier frequencies and can be reduced by band-pass filtering the output of a frequency standard.

One has to keep in mind that the pure power laws of (3.40) represent a theoretical model, which is not always observed in this form. The low frequency contributions to the noise

sometimes referred to as $1/f$ noise often follow a $f^{-\beta}$ dependence with $0.5 \leq \beta \leq 2$ (see, e.g., Fig. 3.10) where the observed power law may also be due to a superposition of several noise processes.

3.3 Conversion from Fourier-frequency Domain to Time Domain

So far we have described the frequency instability of an oscillator either in the Fourier frequency domain by power spectral densities or in the time domain by the Allan variance. In the following we develop the procedure that allows one to calculate the Allan variance from a given power spectral density.

The Allan variance defined by (3.13) and (3.4) can be written as

$$\sigma_y^2(\tau) = \frac{1}{2}\langle(\overline{y}_2 - \overline{y}_1)^2\rangle = \left\langle \frac{1}{2}\left(\frac{1}{\tau}\int_{t_{k+1}}^{t_{k+2}} y(t')dt' - \frac{1}{\tau}\int_{t_k}^{t_{k+1}} y(t')dt'\right)^2 \right\rangle \tag{3.41}$$

with $t_{k+i} - t_k = \tau$ for all i. In (3.41) a single sample is obtained by one-half of the squared difference of the mean values of the function $y(t)$ derived from two adjacent intervals of duration τ and the Allan variance is then the expectation value of this quantity. To obtain many samples of the Allan variance it is not necessary to divide the function $y(t')$ into discrete time intervals but rather we derive a sample for each instant t as follows

$$\sigma_y^2(\tau) = \frac{1}{2}\left\langle \left(\frac{1}{\tau}\int_{t}^{t+\tau} y(t')dt' - \frac{1}{\tau}\int_{t-\tau}^{t} y(t')dt'\right)^2 \right\rangle. \tag{3.42}$$

(3.42) can be written as

$$\sigma_y^2(\tau) = \left\langle \left(\int_{-\infty}^{\infty} y(t')h_\tau(t-t')dt'\right)^2 \right\rangle \tag{3.43}$$

by introducing the function $h_\tau(t)$ with

$$h_\tau(t) = \begin{cases} -\dfrac{1}{\sqrt{2}\tau} & \text{for} \quad -\tau < t < 0 \\ +\dfrac{1}{\sqrt{2}\tau} & \text{for} \quad 0 \leq t < \tau \\ 0 & \text{otherwise} \end{cases} \tag{3.44}$$

(see Fig. 3.9 a). The integral in (3.43) represents a convolution of the time series $y(t)$ with the function $h_\tau(t)$. One recognises the meaning of the function $h_\tau(t)$ by applying, e.g., a

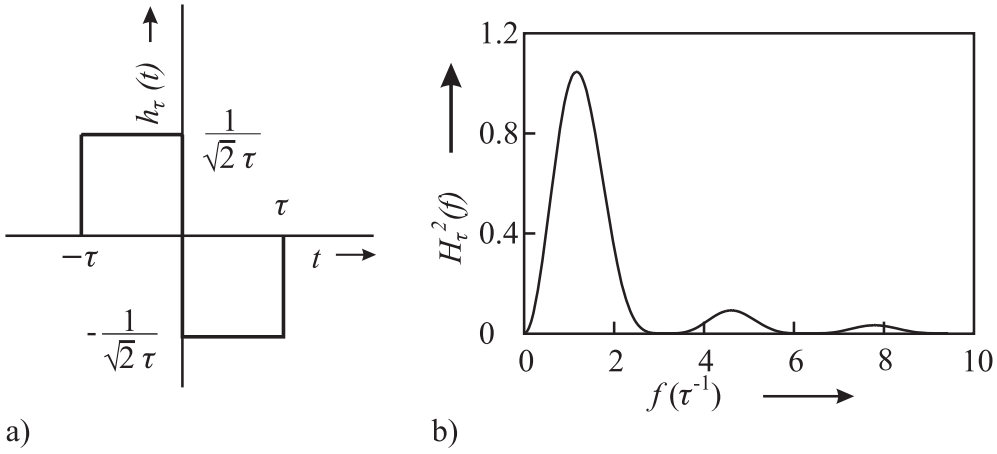


Figure 3.9: a) Filter function $h_\tau(t)$ according to (3.44). b) Transfer function $|H_\tau(f)|^2$ corresponding to the filter function of Fig. 3.9 a).

sharp pulse (Dirac's δ function) for $y(t)$ which reproduces $h_\tau(t)$ (see (2.22)). Hence, the convolution integral in (3.43) can be interpreted as the temporal response of a hypothetical linear "filter" with impulse response $h_\tau(t)$ to an input signal $y(t)$. Consequently, the Allan variance is the mean square of the accordingly filtered temporal fluctuations at the output of this filter. On the other hand the (true) variance of an unfiltered signal $y(t)$ that has a zero mean is given by (3.39) as an integral over the corresponding power spectral density.

To take into account the influence of the filter $h_\tau(t)$ on the power spectral density we recall that, according to the convolution theorem, a convolution of $y(t)$ and $h_\tau(t)$ in the time domain corresponds in the Fourier frequency domain to a multiplication of the Fourier transformed $\mathcal{F}(y(t))$ and $\mathcal{F}(h_\tau(t))$. Similarly, in the Fourier frequency domain the filtered power spectral density is given as a product of the unfiltered power spectral density and the appropriate weighting function, i.e., the square of the filter function.[8] Hence,

$$\sigma_y^2(\tau) = \int_0^\infty |H_\tau(f)|^2 S_y^{1-\text{sided}}(f) df \tag{3.45}$$

where the transfer function

$$H_\tau(f) = \mathcal{F}\{h_\tau(t)\} \tag{3.46}$$

represents the Fourier transform of the filter function $h(t)$.

[8] Note that this only holds if the corresponding time functions are uncorrelated.

We are now in a position to calculate the transfer function for the filter function $h_\tau(t)$ of (3.44) as

$$\begin{aligned} H(f) &= -\int_\tau^0 \frac{1}{\sqrt{2}\tau}\exp(i2\pi ft)dt + \int_0^\tau \frac{1}{\sqrt{2}\tau}\exp(i2\pi ft)dt \\ &= \frac{1}{\sqrt{2}\tau}\left\{-\frac{1}{i2\pi f}\left[\exp(i2\pi ft)\right]_{-\tau}^0 + \frac{1}{i2\pi f}\left[\exp(i2\pi ft)\right]_0^\tau\right\} \\ &= \frac{1}{\sqrt{2}i2\pi f\tau}\left[-1+\exp(-i2\pi f\tau)+\exp(i2\pi f\tau)-1\right] \\ &= \frac{1}{\sqrt{2}i2\pi f\tau}2[\cos(2\pi f\tau)-1] = \frac{1}{\sqrt{2}i\pi f\tau}2\sin^2(\pi f\tau). \end{aligned} \tag{3.47}$$

Hence,

$$|H(f)|^2 = 2\frac{\sin^4(\pi\tau f)}{(\pi\tau f)^2} \tag{3.48}$$

and

$$\sigma_y^2(\tau) = 2\int_0^\infty S_y(f)\frac{\sin^4(\pi\tau f)}{(\pi\tau f)^2}df \tag{3.49}$$

allows one to compute the Allan variance directly from the (one-sided) power spectral density.

As an example, we calculate the Allan variance for white phase noise ($S_y = h_2 f^2$) using (3.49) as follows

$$\sigma_y^2(\tau) = 2\int_0^\infty h_2 f^2\frac{\sin^4(\pi\tau f)}{(\pi\tau f)^2}df = \frac{2h_2}{\pi^2\tau^2}\int_0^\infty \sin^4(\pi\tau f)df. \tag{3.50}$$

The integral (3.50) diverges for $f \to \infty$. In the experiment this does not present a problem since the frequency bandwidth of each measuring device is limited towards high frequencies. Modelling this bandwidth by a low-pass filter with sharp cutoff frequency f_h, (3.50) can be solved with the help of $\int \sin^4 axdx = 3/8x - 1/(4a)\sin 2ax + 1/(32a)\sin 4ax$ as

$$\sigma_y^2(\tau) = \frac{2h_2}{\pi^2\tau^2}\int_0^{f_h} \sin^4(\pi\tau f)df = \frac{3h_2 f_h}{4\pi^2\tau^2} + \mathcal{O}(\tau^{-3}). \tag{3.51}$$

Since the term $\mathcal{O}(\tau^{-3})$ in general can be neglected for $f_h \gg 1/(2\pi\tau)$ the Allan variance for white phase noise shows a power law $\propto \tau^{-2}$. Similarly, $\sigma_y(\tau)$ is calculated for other power spectral densities where each one also shows a well-defined power law dependence of the Allan variance (see Table 3.1).

The integral (3.49) also diverges for flicker phase noise ($S_y(f) = h_1 f$) as can be seen from the fact that $|H(f)|^2$ with its infinite number of side lobes (see Fig. 3.9 b) decreases

as $1/f^2$. Exactly, as in the case of white phase noise discussed above, the low-pass filtering $S_y(f)$ results in an Allan variance that depends on the cut-off frequency of the low-pass filter. In general, the integral (3.49) diverges for $f \to \infty$ for all power laws of (3.40) with $\alpha \geq -1$. The integrals over the model spectral densities with $\alpha = -1$ and $\alpha = -2$ are also diverging for $f \to 0$. In reality, however, an infinite variance will not be observed as both cases cannot be realised experimentally. The case $f \to 0$ would require an infinite measurement time and $f \to \infty$ would require an infinite bandwidth of the measurement equipment. Nevertheless, the situation is not very satisfactory as for these cases $\sigma_y^2(\tau)$ depends on the maximum measurement time or on the bandwidth of the equipment.

The integral (3.49) converges for all $S_y(f) \propto f^\alpha$ with $\alpha \geq -2$ at the low-frequency cut-off. If $S_y(f)$ can be represented by a power law according to (3.40), the calculation of the respective two-sample standard deviation by means of (3.49) in general leads also to a power law for the dependence of $\sigma_y^2(\tau)$ as a function of the measurement time τ. In the cases when $\alpha = -2, -1, 0$ the exponent in the power law and in the Allan variance can be related unambiguously. The dependence of the Allan variance, however, in practice does not allow one to distinguish between flicker phase noise ($S_y(f) \propto f^{-2}$; $\sigma_y^2(\tau) \propto \tau^{-2}$) and white phase noise ($S_y(f) \propto f^{-1}$; $\sigma_y^2(\tau) \propto \tau^{-2}[1.038 + 3\ln(2\pi f_h \tau]$) (see Table 3.1). To overcome this deficiency a so-called modified Allan variance

$$\text{Mod}\ \sigma_y^2(\tau) = \frac{1}{2}\Big\langle \left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{n}\sum_{k=1}^{n}\overline{y}_{i+k+n,\tau_0} - \frac{1}{n}\sum_{k=1}^{n}\overline{y}_{i+k,\tau_0}\right)\right]^2 \Big\rangle \tag{3.52}$$

has been introduced [1, 34]. This type of variance does not influence the variance of the first four types of power spectral densities given in Table 3.1 but increases the sensitivity for white phase noise. Hence, for $S_y(f) = h_2 f^2$ the modified Allan variance Mod $\sigma_y^2(\tau) = 3h_2 f_h \tau_0/(4\pi^2)/\tau^3$ is proportional to τ^{-3} in contrast to the regular Allan variance that goes as τ^{-2}.

The time-domain description of the instability of oscillators by the Allan variance is often chosen as it is easily calculated from the time series measured with simple counters. The description of fluctuations by power spectral densities in the Fourier frequency domain, however, contains the full information about the noise process if properly determined. It furthermore allows one to calculate the Allan variance according to (3.49). In contrast, the calculation of the power spectral density from the measured Allan variance requires the solution of an integral equation which is possible only in simple cases, e.g., those where the power spectral density is given by a simple power law as discussed. It is, however, sufficient if the power spectral density follows a power law within limited frequency ranges which can then be interpreted as resulting from a superposition of different noise processes. As an example, consider the Allan deviation of the hydrogen maser (Fig. 3.3) which at low measuring times is dominated by white phase noise ($\propto \tau^{-1}$) or perhaps flicker phase noise (which also is roughly proportional to τ^{-1}), at higher measuring times by white frequency noise ($\propto \tau^{-1/2}$). The Allan deviation then reaches the so-called flicker floor, before it may increase again, e.g., due to a frequency drift ($\propto \tau^{1/2}$). The underlying physical processes will be described in more detail in Section 8.1.

3.4 From Fourier-frequency to Carrier-frequency Domain

Often, when dealing with laser or microwave frequency standards, one is interested in the power spectrum of the oscillator in the carrier frequency domain. An ideal oscillator operating at the frequency ν_0 would consist of a delta function at ν_0 in the carrier frequency domain. For a real oscillator perturbed by noise processes the power is spread over a frequency range around the centre frequency ν_0. The power spectrum can be measured by different methods. As a first method consider a bandpass filter whose centre frequency is tuned over a frequency range in the vicinity of the centre frequency of the oscillator. The power spectrum of the oscillator is directly related to the power transmitted through the filter measured as a function of the frequency setting of the filter. In the optical domain, a tuneable Fabry–Pérot interferometer (Section 4.3.1) is often chosen as a filter to sweep across laser lines. Another possibility of measuring the power spectrum in the carrier frequency domain is to feed the signal from the oscillator simultaneously to a parallel filter bank. The parallel filter bank can also be simulated by a fast Fourier transform of a digitised and numerically filtered signal. It has to be pointed out, however, that the concept of a power spectrum with a well defined form and linewidth is in general not applicable to all noise processes. As an example, consider a power spectral density with large $1/f$ contribution. For long observation times corresponding to low Fourier frequencies the central frequency may drift away and, hence, there is no unique "linewidth" as the measured width of the power spectrum will depend on the observation time.

With this note of caution in mind, we show in this section how the shape of the emission line in the carrier frequency domain can be determined from a particular noise spectral density, e.g., $S_\nu(\nu)$ determined in the Fourier domain. The power spectrum of the electric field $S_E(\nu)$ can be evaluated by following [35–37]. In analogy to (3.27) and (3.28) one defines the two-sided power spectral density as the Fourier transform

$$S_E(\nu) = \int_{-\infty}^{\infty} \exp\left(-i2\pi\nu t\right) R_E(\tau) d\tau \tag{3.53}$$

of the autocorrelation function

$$R_E(\tau) = \left\langle E(t+\tau)E^*(t)\right\rangle \tag{3.54}$$

of the electric field $E(t)$. For a complex representation of the electric field of the electromagnetic wave with negligible amplitude fluctuations and real amplitude E_0

$$E(t) = E_0 \exp i[2\pi\nu_0 t + \phi(t)] \tag{3.55}$$

the autocorrelation function becomes

$$R_E(\tau) = E_0^2 \exp[i2\pi\nu_0\tau]\left\langle \exp i[\phi(t+\tau) - \phi(\tau)]\right\rangle. \tag{3.56}$$

Now, $\langle \exp i[\phi(t+\tau) - \phi(\tau)]\rangle$ has to be expressed in terms of the spectral density of phase fluctuations $S_\phi(f)$. To begin with, one assumes that the noise process is ergodic, i.e., that the temporal average is identical to the corresponding ensemble average

$$\overline{\exp[i\Phi(t,\tau)]} = \left\langle \exp[i\Phi(t,\tau)]\right\rangle = \int_{-\infty}^{\infty} p(\Phi)\exp(i\Phi) d\Phi \tag{3.57}$$

where

$$\Phi(t,\tau) \equiv \phi(t+\tau) - \phi(\tau) \tag{3.58}$$

is the phase accumulated during the interval τ. The right-hand side of (3.57) uses the usual definition of the expectation value of the quantity $\exp[i\Phi(t,\tau)]$ if the probability density $p(\Phi)$ is known. For a large number of uncorrelated phase-shifting events the central limit theorem allows one to use the Gaussian probability density

$$p(\Phi) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\Phi^2}{2\sigma^2}\right) \tag{3.59}$$

with the classical variance σ^2. As $p(\Phi)$ is an even function, only the real (cosine) part of the complex exponential of (3.57) survives. (3.57) is evaluated by using (3.59) and $\int_{-\infty}^{\infty} \exp(-a^2x^2)\cos x dx = \sqrt{\pi}/a \exp(1/4a^2)$ leading to

$$\left\langle \exp[i\Phi(t,\tau)] \right\rangle = \exp\left(-\frac{\sigma^2}{2}\right). \tag{3.60}$$

According to (3.11) with vanishing mean value $\langle\Phi\rangle = 0$ and (3.58)

$$\begin{aligned} \sigma^2(\Phi) &= \left\langle \Phi^2 \right\rangle = \left\langle \left[\phi(t+\tau) - \phi(\tau)\right]^2 \right\rangle \\ &= \left\langle \left[\phi(t+\tau)\right]^2 \right\rangle - 2\left\langle \left[\phi(t+\tau)\phi(\tau)\right] \right\rangle + \left\langle \left[\phi(\tau)\right]^2 \right\rangle. \end{aligned} \tag{3.61}$$

Using (3.54) and (3.32) one finds

$$\left\langle \left[\phi(t+\tau)\phi(\tau)\right] \right\rangle = \int_0^\infty S_\phi(f)\cos(2\pi f\tau) df = R_\phi(\tau) \tag{3.62}$$

$$\left\langle \left[\phi(t+\tau)\right]^2 \right\rangle = \left\langle \left[\phi(\tau)\right]^2 \right\rangle = \int_0^\infty S_\phi(f) df = R_\phi(0). \tag{3.63}$$

Insertion of (3.62) and (3.63) into (3.61) leads to

$$\sigma^2 = 2\int_0^\infty S_\phi(f)\left[1 - \cos 2\pi f\tau\right] df \tag{3.64}$$

which can be used to derive the autocorrelation function from (3.56)

$$R_E(\tau) = E_0^2 \exp(i2\pi\nu_0\tau) \exp\left(-\int_0^\infty S_\phi(f)\left[1 - \cos 2\pi f\tau\right] df\right). \tag{3.65}$$

From (3.53) and (3.65) the power spectral density in the carrier frequency domain

$$S_E(\nu - \nu_0) = E_0^2 \int_{-\infty}^{\infty} \exp -[i2\pi(\nu - \nu_0)\tau] \exp\left(-\int_0^\infty S_\phi(f)\left[1 - \cos 2\pi f\tau\right] df\right) d\tau \tag{3.66}$$

can be derived for a given phase noise spectral density $S_\phi(f)$ (see (3.37)) provided that the integral in brackets in (3.66) converges.

3.4.1 Power Spectrum of a Source with White Frequency Noise

We now consider a source whose power spectral density in the Fourier-frequency domain can be represented as white (frequency independent) frequency noise S_ν^0 (see Table 3.1). Consequently,

$$S_\phi(f) = \frac{S_\nu^0}{f^2} = \frac{\nu_0^2 h_0}{f^2} \tag{3.67}$$

holds and the integral in the exponential of (3.66) can be solved analytically using $\int_0^\infty [1 - \cos(bx)]/x^2 dx = \pi|b|/2$ leading to

$$\begin{aligned} S_E(\nu - \nu_0) &= E_0^2 \int_{-\infty}^{\infty} \exp -[i2\pi(\nu - \nu_0)\tau] \exp\left(-\pi^2 h_0 \nu_0^2 |\tau|\right) d\tau \\ &= 2E_0^2 \int_0^{\infty} \exp -\tau \left[i2\pi(\nu - \nu_0) + \pi^2 h_0 \nu_0^2\right] d\tau. \end{aligned} \tag{3.68}$$

Solving the integral (3.68) and keeping the real part leads to the power spectral density of

$$S_E(\nu - \nu_0) = 2E_0^2 \frac{h_0 \pi^2 \nu_0^2}{h_0^2 \pi^4 \nu_0^4 + 4\pi^2(\nu - \nu_0)^2} = 2E_0^2 \frac{\gamma/2}{(\gamma/2)^2 + 4\pi^2(\nu - \nu_0)^2} \tag{3.69}$$

with $\gamma \equiv 2h_0\pi^2\nu_0^2 = 2\pi(\pi h_0 \nu_0^2) = 2\pi(\pi S_\nu^0)$. Hence, the power spectral density of frequency fluctuations in the carrier-frequency domain of an oscillator with white frequency noise S_ν^0 in the Fourier-frequency domain, is a Lorentzian whose full width at half maximum is given by

$$\Delta\nu_{\mathrm{FWHM}} = \pi S_\nu^0. \tag{3.70}$$

Similarly, other types of phase noise spectral densities can be calculated accordingly. Godone and Levi have furthermore treated the case of white phase noise and flicker phase noise [38].

3.4.2 Spectrum of a Diode Laser

As an example of white frequency noise, consider the frequency fluctuations in a laser resulting from the spontaneous emission of photons [39]. They lead to the so-called Schawlow–Townes linewidth

$$\Delta\nu_{\mathrm{QNL}} = \frac{2\pi h\nu_0 (\Delta\nu_{1/2})^2 \mu}{P}. \tag{3.71}$$

where $h\nu_0$ is the photon energy, $\Delta\nu_{1/2}$ is the full width at half maximum of the passive laser resonator, $\mu \equiv N_2/(N_2 - N_1)$ is a parameter describing the population inversion in the laser medium, and P is the output power of the laser. This quantum-noise limited power spectral density (which is enhanced for laser diodes by Henry's linewidth enhancement factor; see (9.37)) can be found in the measured spectral noise of a solitary diode laser (Fig. 3.10) at Fourier frequencies above a corner frequency of about 80 kHz. At frequencies below the

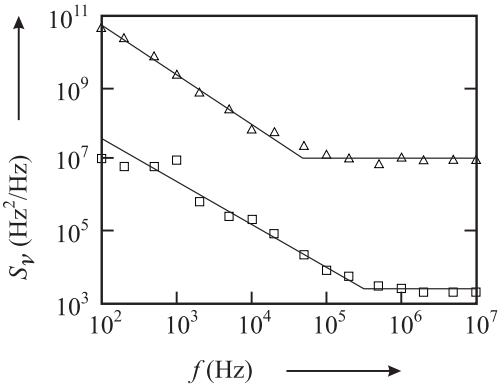


Figure 3.10: Measured power spectral densities of frequency fluctuations versus Fourier frequency f of a diode laser without optical feedback (triangles) and with optical feedback from a grating (squares) after [40] with permission.

corner frequency the power spectral density increases with a power law of roughly $1/f$. The white frequency noise regime is also visible above the corner frequency f_c of about 200 kHz if the cavity of the diode laser is extended (Section 9.3.2.5) but $S_\nu(f)$ is reduced by about 33 dB according to the reduced linewidth $\Delta\nu_{1/2}$ (see (3.71)).

As the $1/f$-like behaviour often results from technical noise which is present in any oscillator to some degree it is interesting to investigate the validity of (3.69). O'Mahony and Henning [41] have investigated the effect of low frequency (1/f) carrier noise on the linewidth of a semiconductor laser. From their findings Koch [40] gives a criterion that allows one to obtain information about the lineshape from the positions of the corner frequencies f_c as follows

$$S_\nu(f_c)/f_c \gg 1 :\Rightarrow \quad \text{Lorentzian lineshape} \tag{3.72}$$

$$S_\nu(f_c)/f_c \ll 1 :\Rightarrow \quad \text{Gaussian lineshape.} \tag{3.73}$$

We apply these criteria to the power spectral density of frequency noise displayed in Fig. 3.10 where one finds, for the solitary laser diode (triangles), $S_\nu(f_c)/f_c > 100$ and, hence, criterion (3.72) applies. With (3.70) one expects a Lorentzian profile of about 5 MHz linewidth. From the power spectral density of frequency fluctuations (squares in Fig. 3.10) of another diode laser with extended cavity (Section 9.3.2.5) one finds $S_\nu(f_c)/f_c \approx 10^{-2}$ and hence expects a Gaussian lineshape according to criterion (3.73). The origin of the Gaussian lineshape can be thought of as resulting from a small Lorentzian line whose width is given by (3.70) which statistically wanders around a central frequency. The width of the Gaussian depends on the time T of averaging, as the measurement time T also defines the lowest measurable Fourier frequency $1/T$. For a true $1/f$ behaviour of S_ν the linewidth would be infinite as $\int_{1/T}^{\infty} S_\nu(f)df = \infty$ holds (see (3.66)). Experimentally, however, one always finds a finite linewidth resulting from the finite measurement time T with the low-frequency cut off $1/T$.

The mean frequency excursion $\Delta\nu_{\mathrm{rms}}$ (linewidth) can be computed as

$$\Delta\nu_{\mathrm{rms}} = \sqrt{\int_{1/T}^{f_c} S_\nu(f)df} \tag{3.74}$$

from (3.39). In the case of the laser with optical feedback in an extended cavity arrangement (squares in Fig. 3.10) one derives a FWHM of the Gaussian of about 120 kHz for a measurement time of 10 ms.

3.4.3 Low-noise Spectrum of a Source with White Phase Noise

With the help of (3.62) and (3.63) we can write (3.66) as

$$S_E(\nu-\nu_0) = E_0^2 \int_{-\infty}^{\infty} \exp\left[-R_\phi(0)\right] \exp\left[R_\phi(\tau)\right] \exp[-i2\pi(\nu-\nu_0)\tau]d\tau. \tag{3.75}$$

For very low phase fluctuations, i.e., for $\int_0^\infty S_\phi(f)df \ll 1$ it is justified to expand the first two exponential functions in (3.75) and to keep only the first terms as

$$S_E(\nu-\nu_0) \approx E_0^2 \int_{-\infty}^{\infty} \left[1 - R_\phi(0) + R_\phi(\tau)\right] \exp[-i2\pi(\nu-\nu_0)\tau]d\tau. \tag{3.76}$$

Using the definition of Dirac's delta function $\delta(\nu-\nu_0)$ (see (2.23)) and the Wiener–Khintchine relation (3.28) one finds

$$S_E(\nu-\nu_0) \approx E_0^2 \left[1 - R_\phi(0)\right] \delta(\nu-\nu_0) + E_0^2 S_\phi^{2-\mathrm{sided}}(\nu-\nu_0). \tag{3.77}$$

Hence, the spectrum in the carrier frequency domain comprises a carrier (delta function) at $\nu = \nu_0$ and two symmetric sidebands with the level of the phase noise spectral density S_ϕ at $f = |\nu - \nu_0|$.

Often commercial oscillators are specified by the measure of the so-called spectral purity $\mathcal{L}(f)$, i.e., the noise found on each side of the carrier when the signal of an oscillator is measured directly with a spectrum analyser [1]

$$\mathcal{L}(f) \equiv \frac{S_\Phi^{2-\mathrm{sided}}(\nu-\nu_0)}{1/2E_0^2}. \tag{3.78}$$

Here it is assumed that the amplitude noise is negligible as compared to the phase noise. Then the spectral purity represents all phase noise for all Fourier frequencies except for the origin, i.e., the delta function of (3.77).

3.5 Measurement Techniques

In practical applications, the power spectral density of the frequency (phase) fluctuation of a signal is determined from measured time sequences of $\Delta\nu(t)$ ($\Delta\phi(t)$) by different methods. One may measure $S_\nu(f)$ by separating the spectrum into frequency classes by a number of filters of different centre frequencies and then measuring the (ac) power transmitted by each one of these filters. After division of the respective bandwidths of the filters these discrete power values represent $S_\nu(f)$ at the Fourier frequencies of the centre frequencies of the filters. Another method makes use of digital spectrum analysers, using the Fast Fourier Transformation (FFT) algorithm giving, e.g.,

$$\Delta\phi(f) = \mathcal{F}(\Delta\phi(t)). \tag{3.79}$$

From this, the power spectral density of phase fluctuations can be determined from

$$S_\phi(f) = \frac{[\Delta\phi(f)]^2}{BW} \tag{3.80}$$

where the measurement bandwith BW in Hertz has to be chosen such that BW $\ll f$. This procedure is equivalent to the definition of the power spectral density via the autocorrelation function for reasonably well behaved noise processes. To allow the reader to trace possible implications, the derivation of (3.80) is given in the following. We rewrite (3.31) and (3.32) in complex notation [42] as

$$S_b(f) = 4 \lim_{T\to\infty} \frac{1}{T} \int_0^T \left[\int_0^T b(t)b(t+\tau)dt \right] \exp(2\pi i f\tau)d\tau. \tag{3.81}$$

Substitution of

$$\tau \to z - t \tag{3.82}$$

leads to

$$S_b(f) = 4 \lim_{T\to\infty} \frac{1}{T} \int_0^T \left[\int_0^T b(z)\exp(2\pi i f z)dz \right] b(t)\exp(-2\pi i f t)dt. \tag{3.83}$$

As the expression in squared brackets is a complex number it can be moved out of the integral

$$S_b(f) = 4 \lim_{T\to\infty} \frac{1}{T} \int_0^T b(z)\exp(2\pi i f z)dz \int_0^T b(t)\exp(-2\pi i f t)dt. \tag{3.84}$$

If the value of the integral does not change after the shift of the integration interval by τ (see (3.82)), each one of these integrals is the complex conjugate of the other one and, hence,

we finally end up with the relation

$$S_b(f) = 4 \lim_{T\to\infty} \frac{1}{T} \left| \int_0^T b(t) \exp(-2\pi i f t) dt \right|^2 \tag{3.85}$$

for the one-sided power spectral density $S_b(f)$.

Frequency fluctuations can be converted into amplitude or power fluctuations by means of a discriminator. The slopes of an electronic filter, an optical Fabry–Pérot Interferometer (FPI) or an absorption line can be employed as such a discriminator. If the oscillator or the filter is tuned such that the carrier frequency of the oscillator is at the slope, preferably near the inflection point (Fig. 3.11), the power transmitted by the filter varies, to first order, linearly with the frequency of the signal as

$$V(\nu - \nu_S) = (\nu - \nu_S)\, k_d + V(\nu_S) \tag{3.86}$$

where k_d is the slope of the filter at ν_S. The detector behind the filter of Fig. 3.11 a) converts

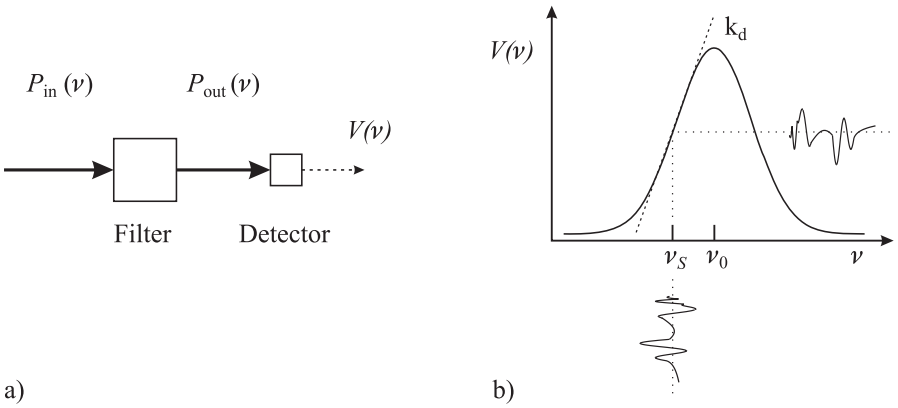


Figure 3.11: a) The transmission of a filter can be used to convert the frequency fluctuations of a signal $P(\nu)$ into voltage fluctuations $V(\nu)$. b) At a suitable working point ν_S the filter acts as a frequency discriminator where the voltage fluctuations are approximately proportional to the frequency fluctuations (see (3.86)).

the power fluctuations into the fluctuations of a voltage that can be analysed by means of an electronic spectrum analyser. Modern spectrum analysers show a quantity directly related to the spectral density of the fluctuations of the signal. To obtain the power spectral density of the frequency fluctuations in Hz^2/Hz the slope k_d of the frequency discriminator has to be determined. The application of this method requires that the contributions of other noise sources do not affect the measurement. Fluctuations of the centre frequency of the filter or fluctuations of the signal amplitude may mimic a higher spectral density of frequency fluctuations. The latter contribution has to be eliminated by stabilising the input or by using the amplitude fluctuations measured directly with a second detector of frequency independent response to

normalise the signal transmitted by the filter. The influence of amplitude fluctuations can easily be detected by tuning the centre frequency of the filter or the frequency of the signal such that both coincide. For not too high frequency excursions, the flat portion of the filter at the centre frequency is less susceptible to frequency fluctuations and the observed fluctuations of the transmitted signal can be related to amplitude fluctuations of the incoming signal. The frequency of electric signals up to a few gigahertz can be measured directly with an electronic counter. In its simplest form the counter measures the number of cycles during a given time interval τ by the number of detected zero crossings with positive slope. The time interval is provided by a well defined number of cycles of a reference frequency. In this simple case the resolution of the frequency to be measured is limited to ± 1 cycle. Often, electronic counters employ interpolation techniques in order to estimate the fractions of a cycle. In any case, the uncertainty of a frequency measurement with a counter decreases with $1/\tau$.

3.5.1 Heterodyne Measurements of Frequency

Higher frequencies can be measured by the heterodyne technique. The heterodyne technique produces the difference signal between the signal of the Device Under Test (DUT) of frequency ν and a reference signal of frequency ν_0, by mixing them. Consider two harmonic

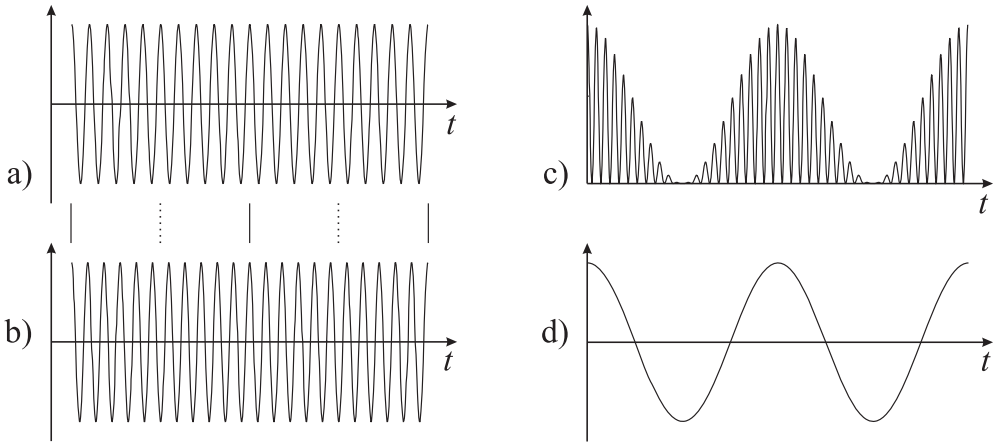


Figure 3.12: a, b) Signals differing in frequency by 10%. c) Squared sum of signals a) and b). d) Beat signal.

signals of very high frequencies ν and ν_0 (see Fig. 3.12 a, b), e.g., two laser beams that are superimposed on a photodetector. The frequencies of the particular fields are far too high to allow the electronics to trace them directly. However, the power, being proportional to the amplitude squared of the resulting field, shows an amplitude-modulated signal (Fig. 3.12 c). After filtering with a low pass, the so-called beat frequency ν_{beat} of this signal (Fig. 3.12 d) is the difference frequency

$$\nu_{\mathrm{beat}} = |\nu - \nu_0|. \tag{3.87}$$

Since the photodetector delivers a signal which is proportional to the product of the electric field, it responds in a highly non-linear way to the input field. Similarly in the radio-frequency regime non-linear devices are used that form the product of two inputs. We briefly recall the essential features of such a product mixer. The mixer multiplies two input signals often termed the Radio Frequency (RF) and the Local Oscillator (LO) leading to an output signal called the Intermediate Frequency (IF). If the two input signals are harmonic signals the output

$$\cos(\omega_{\mathrm{RF}}t)\cos(\omega_{\mathrm{LO}}t) = \tfrac{1}{2}\cos[(\omega_{\mathrm{RF}}+\omega_{\mathrm{LO}})t] + \tfrac{1}{2}\cos[(\omega_{\mathrm{RF}}-\omega_{\mathrm{LO}})t] \tag{3.88}$$

contains the sum and the difference of the two input signals but not the input signals or their harmonics. The Double Balanced Mixer (DBM) (Fig. 3.13) often used for frequency mixing is an electronic device based on four diodes and two transformers, in order to facilitate the generation of only the sum and the difference signals, but not the input signals nor their harmonics.

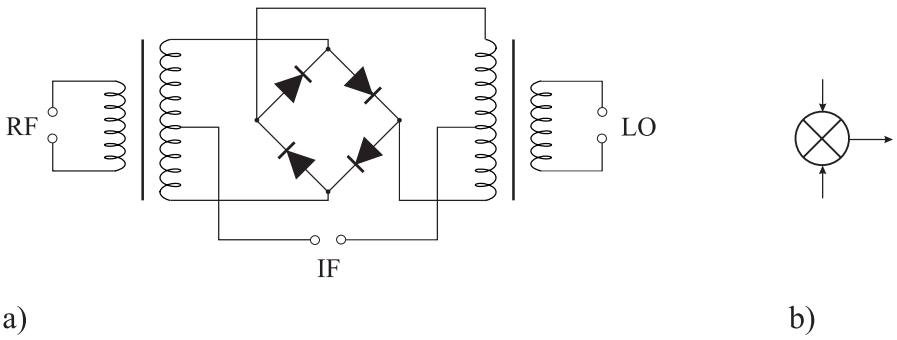


Figure 3.13: Double-Balanced Mixer. a) Set up. b) Symbol.

Important applications are found for $\omega_{\mathrm{LO}} = \omega_{\mathrm{RF}} \equiv \omega$ where two signals with the same frequency ω but a phase difference are delivered to the inputs of the mixer. Using (3.88) one finds

$$\cos(\omega t+\phi)\cos(\omega t) = \frac{1}{2}\left[\cos(2\omega t+\phi)+\cos\phi\right]. \tag{3.89}$$

The output signal contains an ac component $1/2\cos(2\omega t+\phi)$ of twice the input frequency superimposed on a dc signal $1/2\cos\phi$ that depends on the phase difference between the two input signals.

Consequently, the mixer can be used as a discriminator to detect phase fluctuations as in Fig. 3.14 [1] where the phase of the signal from an oscillator under test is compared to the phase of the signal from a reference oscillator. If the voltage fluctuations from the mixer are measured by a spectrum analyser, they are a measure of the fluctuations of the phase difference provided that amplitude modulation is negligible. These in turn can be attributed to the phase fluctuations of the oscillator under test only if the phase of the reference oscillator can be regarded as more stable. To have a constant slope for the phase discriminator it has to be operated close to 90° during the measurement, where the cosine function (3.89) near

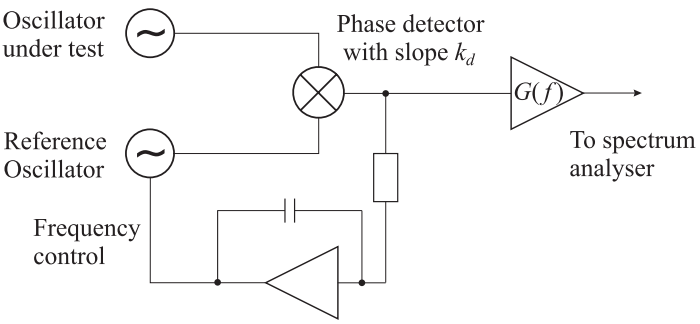


Figure 3.14: System for measurement of phase noise of an oscillator under test.

zero can be approximated by a linear discriminant curve. This is the case only if the mean frequencies of both oscillators are kept the same during that time, i.e., if their mean phases are locked. This condition can be fulfilled by employing a Phase-Locked Loop (PLL) that uses the signal from the mixer representing the phase difference between the oscillators as an error signal. The error signal is integrated to give a servo signal which is used to control the frequency of the reference oscillator. Provided that the slope of the phase discriminator k_d and the frequency-dependent gain $G(f)$ of the amplifier are known, the power spectral density of phase fluctuations at Fourier frequency f from the carrier are measured from the mean-squared phase fluctuations in a measurement bandwidth of 1 Hz according to (3.80).

Care has to be taken in determining the power spectral density of phase fluctuations if the frequency is translated by summation, multiplication or division. In a beat-note measurement (see (3.87)) the frequency is translated and the stochastic phase modulation (PM) noise of the reference oscillator and the PM noise in the non-linear device $S_\phi^{\mathrm{trans}}(f)$ performing the translation are added to the PM noise of the oscillator under test (DUT) as

$$S_\phi(\nu, f) = S_\phi^{\mathrm{DUT}}(\nu_0, f) + S_\phi^{\mathrm{RefOsc}}(\nu_1, f) + S_\phi^{\mathrm{trans}}(f). \tag{3.90}$$

Frequency multiplication in a non-linear device by a factor N at the same time multiplies the phase by the same factor N and, hence, the phase fluctuations. From (3.80) one finds that in this case the PM noise is increased by N^2

$$S_\phi(N\nu_0, f) = N^2 S_\phi(\nu_0, f) + S_\phi^{\mathrm{mult}}(f). \tag{3.91}$$

where $S_\phi^{\mathrm{mult}}(f)$ is the PM noise added by the multiplication process in the non-linear device. Similarly, frequency division by $1/N$ reduces the PM noise by $1/N^2$. Hence, the residual noise is reduced if a high-frequency signal is divided rather than mixed down [43].

3.5.2 Self-heterodyning

The phase noise of a single oscillator can be measured also by using a self-heterodyning technique where the phase fluctuations of an oscillator are determined by comparing the signal of the oscillator with the same signal at a previous epoch (Fig. 3.15). The signal is split

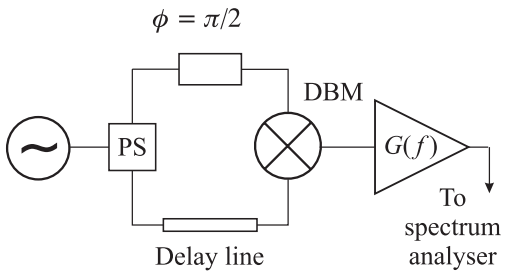


Figure 3.15: Set-up for measurement of phase noise by use of a delay line discriminator

in a power splitter PS and the particular signal in one path is delayed with respect to the other one before being mixed again and amplified. The splitting in the power splitter and the recombination in the mixer makes the arrangement essentially to an interferometer plus square-wave detector.

3.5.2.1 Sensitive Measurement of AM and PM Noise in the Microwave Regime

For low-noise microwave devices the sensitive measurement of phase and amplitude noise close to the carrier meets with difficulties since the power in the carrier is typically several orders of magnitude higher than the noise power to be measured. A highly sensitive method based on interferometric techniques has been used as early as 1968 [44] and has been applied later to achieve ultra-high sensitivity [45]. In such a microwave interferometer (Fig. 3.16) the

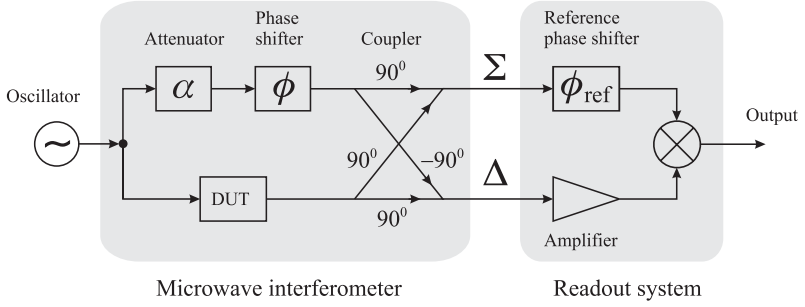


Figure 3.16: Microwave interferometer for the measurement of AM and PM noise (DUT: Device Under Test).

carrier is cancelled in the branch referred to as Δ by adding the signal from the device under test phase-shifted by +90° with the suitably adjusted reference phase-shifted by –90°. Hence, the Δ branch is operated on a dark fringe of the interferometer. The signal with suppressed carrier is amplified in the readout system and mixed with the reference branch Σ. Depending on the phase shift of the latter signal adjusted in the reference phase shifter either phase noise or the amplitude noise can be measured.

3.5.2.2 Fibre Optical Interferometer

Interferometric systems are also often employed in the optical telecommunication bands to investigate the linewidth of diode lasers where the interferometer comprises an optical fibre as the delay line and an acousto-optic modulator (AOM) (see Section 11.2.1), that shifts the frequency in one arm of the interferometer. For a delay time small compared to the coherence time of the laser the power spectral density comprises of a delta function peak at the frequency of the AOM, superimposed on a pedestal [46,47] (Fig. 3.17). As the delay time increases until the phase of the optical field becomes uncorrelated the peak becomes a self-convolution of the lineshape, which is the case when the delay time becomes approximately six times larger than the coherence time [46]. For highly coherent laser oscillators with a linewidth of 10 kHz, a fibre optic delay of several tens of kilometres would be necessary. In contrast to the several kilometre long fibres, a 100 m long fibre leads to an optical path difference which is small compared to the coherence length of the laser. In this case, each frequency component of the interference signal corresponds to a Fourier component of the frequency noise of the laser. In this regime a comparison between the calculated and measured spectrum (Fig. 3.17) can be used to derive the linewidth of the laser. It furthermore gives hints for additional influences that degrade the phase stability [46].

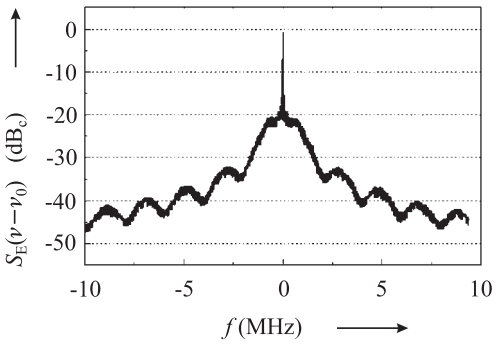


Figure 3.17: Power spectral density $S_E(\nu - \nu_0)$ of a diode laser at 1.5 µm measured at the exit of a heterodyne fibre interferometer with a laser linewidth of 0.4 MHz. Courtesy of U. Sterr.

3.5.3 Aliasing

Care has to be taken if a continuous signal is sampled digitally to obtain spectral densities. Such a situation is encountered, e.g., if a digital spectrum analyser is used in Fig. 3.15. It is well known that a harmonic signal can be sampled digitally unambiguously only if at least two samples are taken per period T (see, e.g., open circles in Fig. 3.18 a). The corresponding minimal sampling frequency

$$\nu_N \equiv \frac{2}{T} \tag{3.92}$$

is referred to as the Nyquist frequency. Consider a harmonic signal that is sampled with less than two points per period T (dots in Fig. 3.18 a). The sampling period Δt is higher than

half of the signal period T and, hence, the frequency of the signal is higher than the Nyquist frequency. In Fig. 3.18 a) eleven samples are taken in ten periods of the component with

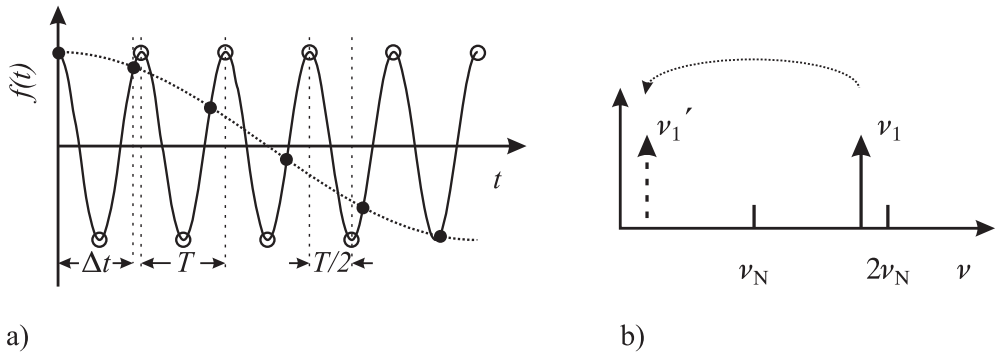


Figure 3.18: Aliasing. a) A cosine signal (full line) can be sampled unambiguously only with at least two samples per period (open circles). If the sampling is done with less than two samples per period (dots) the reconstructed signal shows an apparent lower frequency (dotted line). b) The true frequency component $\nu_1 = 1/T$ is higher than the Nyquist frequency ν_N (3.92) and appears as a low-frequency component ν_1'.

frequency ν. Hence, the frequency $\nu = 20/11\nu_N$ of the sampled signal is 82 % larger than the Nyquist frequency. The sampled signal appears to have a lower frequency compared to the original signal (dotted curve of Fig. 3.18 a). In the frequency domain (Fig. 3.18 b) this signal appears at $\nu' = \nu/10 = 0.18\,\nu$, i.e., 82 % below the Nyquist frequency. The frequency component above the Nyquist frequency seems to be "reflected" at the latter one. In a case where a complete spectrum has been sampled in too crude a manner and the power spectral density has non-zero components above the Nyquist frequency, the spectral density between ν_N and $2\nu_N$ is reflected into the interval $0 \leq \nu \leq \nu_N$. As a result, the power spectral density is corrupted by contributions from outside this interval. This corruption of the power spectral density resulting from insufficient sampling is referred to as "aliasing".

3.6 Frequency Stabilization with a Noisy Signal

In a scheme of the frequency standard (Fig. 1.3) where the signal from a passive frequency reference is used to stabilise the frequency of the standard, fluctuations of the signal I from the passive reference lead to fluctuations of the stabilised frequency. As an example consider, e.g., the noisy absorption signal of a passive reference (Fig. 3.19) where the deviation of the frequency of the oscillator from the centre frequency ν_0 is detected by modulating the frequency by $\pm\Delta\nu/2$ and comparing the signals at both sides of the absorption line. The fluctuations of the signal at a given working point ν_{wp} described by $\sigma^{(I)}$ and the corresponding uncertainty in the frequency ν_{wp} described by $\sigma^{(\nu)}$ are linked by the slope of the absorption

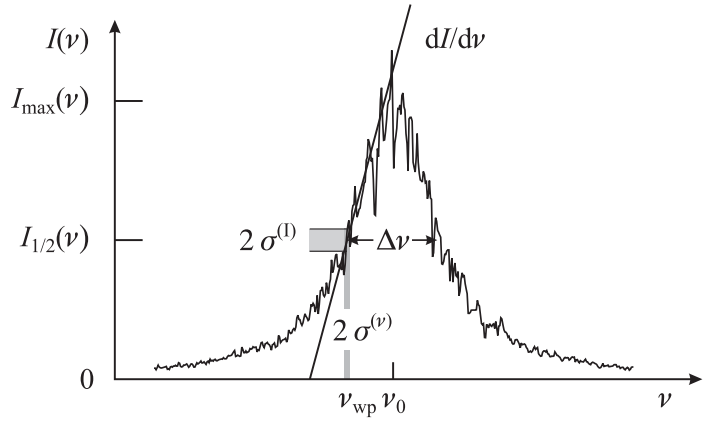


Figure 3.19: Effect of fluctuations of the signal from a passive frequency reference on the frequency stability.

line at $\nu_{\rm wp}$ as follows

$$\sigma^{(\nu)} = \sigma^{(I)} \frac{1}{\left.\dfrac{dI(\nu)}{d\nu}\right|_{\nu_{\rm wp}}}. \tag{3.93}$$

In general, the slope can be written as

$$\frac{dI(\nu)}{d\nu} = K\frac{I_{\max}}{\Delta\nu} \tag{3.94}$$

where K is a constant of the order unity that depends on $I(\nu)$ and $\Delta\nu$ is the full width at half maximum. As an example, we approximate $I(\nu)$ of Fig. 3.19 by a symmetric triangle of hight $I_{\max}$ and baseline $2\Delta\nu$. In this simplified case $K = 1$ holds. A second important example is $I(\nu) = I_{\max}\left\{1 + \cos[2\pi(\nu - \nu_0)t]\right\}/2$, which is encountered in the case of background-free Ramsey excitation (see (6.44)). In this case one finds $K = \pi$. Often the fluctuations of the signal are limited to white frequency noise, resulting, e.g., from the shot noise of the detected photons or atoms and

$$\sigma^I(\tau) = \sigma^I(\tau = 1\,\mathrm{s})\frac{1}{\sqrt{\tau/\mathrm{s}}} \tag{3.95}$$

holds. From (3.93) a useful formula can be derived to estimate the ultimate frequency stability of a frequency standard by

$$\sigma_y(\tau) = \frac{1}{K}\frac{1}{Q}\frac{1}{S/N}\frac{1}{\sqrt{\tau/\mathrm{s}}} \tag{3.96}$$

where we have used $Q = \nu_0/\Delta\nu$, $\sigma^{(\nu)} = \nu_0\sigma_y(\tau)$ and the signal-to-noise ratio S/N is determined by the signal $S = I_{\max}$ and the noise N by $\sigma^I(\tau = 1\,\mathrm{s})$. (3.96) has to be

modified if the atoms are interrogated only for a fraction τ/T_c of the total cycle time T_c used for preparation and interrogation of the atoms leading to

$$\sigma_y(\tau) = \frac{1}{K}\frac{1}{Q}\frac{1}{S/N}\sqrt{\frac{T_c}{\tau}}. \tag{3.97}$$

(3.96) and (3.97) are widely used to determine the achievable instability given by the Allan deviation if the signal-to-noise ratio is known.

3.6.1 Degradation of the Frequency Stability Due to Aliasing

In frequency standards where the frequency of a local oscillator is stabilised to a passive reference by modulation techniques the frequency stability expected from (3.97) is often degraded by different aliasing effects. These effects map high-frequency noise of the free-running local oscillator onto the output frequency of the stabilised oscillator.

Such an effect [48, 49] referred to as "intermodulation effect" results from the nonlinear slope of the discriminant curve used to derive the error signal (Fig. 3.19). A nonlinear element in the servo control loop acts as an element which generates difference frequencies between the harmonics of the modulation frequency and the nearby high-frequency noise components of the oscillator which are thereby mixed into the base band. These low-frequency fluctuations are interpreted by the servo control unit as frequency fluctuations which are counteracted, thereby increasing the fluctuations of the frequency of the stabilised oscillator.

This effect is in particular pronounced in passive frequency standards that are operated or interrogated in a pulsed mode, e.g., atomic fountain clocks (Section 7.3) or single-ion frequency standards (Section 10) where the atoms are prepared and interrogated in a sequence. The periodic activation of the control loop only during the interrogation again can mix down frequency noise of the oscillator near the harmonics of the interrogation frequency into the base band and, hence, into the bandwidth of the control loop. Increased noise due to this so-called "Dick effect" has been predicted by Dick [50] and observed in atomic frequency standards [51].

The degradation depends on the particular mode of interrogation and has been calculated for a number of different operational schemes and parameters [51–53]. Often, use is made of the so-called "sensitivity function" $g(t)$ which is the response of the atomic system to a phase step of the interrogating oscillator or the impulse response with respect to a frequency change occuring at time t [51]. Hence the sensitivity function $g(t)$ takes into account that during the interrogation of the passive resonance the effect of frequency fluctuations of the local oscillator on the error signal can be very different during the cycle time T_c. As an example, consider the case of a sinusoidally modulated local oscillator where frequency fluctuations have almost no effect when the instantaneous frequency is near the maximum of the line, but contribute maximally at instantaneous frequencies near the maximum slope (Fig. 2.15). The sensitivity function is represented as a function periodic in T_c with Fourier series

$$g(t) = \sum_{m=-\infty}^{\infty} g_m e^{i2\pi m f_c t} \tag{3.98}$$

with $f_c = 1/T_c$. The contributions of the power spectral density of fractional frequency fluctuations $S_y^{\mathrm{LO}}(f)$ of the free-running local oscillator weighted by the corresponding Fourier components of the sensitivity function can be used to determine [51, 53] the resulting Allan variation as follows

$$\sigma_y^2(\tau) = \frac{1}{\tau} \sum_{m=1}^{\infty} \frac{|g_m|^2}{g_0^2} S_y^{\mathrm{LO}}(m f_c) \tag{3.99}$$

where g_0 is the mean value of $g(t)$ over the cycle time T_c. As the Dick effect and the intermodulation effect can seriously affect the achievable stability of the most advanced frequency standards local oscillators have to be selected whose noise properties match the chosen interrogation scheme [51].

4 Macroscopic Frequency References

Macroscopic structures of specifically designed geometries can be excited to oscillate at distinct resonance frequencies. The eigenfrequencies of such resonators are uniquely determined by their dimensions and the properties of the medium from which they are made. Resonators find applications in stabilising the frequency of oscillators which in turn can serve as secondary frequency references as in the case of the quartz oscillator (Section 4.1) where a vibrating quartz plate defines the frequency of the oscillator. In the microwave (Section 4.2) and optical domain (Section 4.3) dielectric resonators or empty cavities serve as stable fly wheels for frequency standards. Furthermore, cavity resonators are also applied to interrogate quantum absorbers at well defined frequencies.

4.1 Piezoelectric Crystal Frequency References

In solid bodies, mechanical deformations can be excited that oscillate with resonance frequencies which depend on the properties and dimensions of the chosen material. Certain materials like quartz (SiO_2) or langasite ($La_3Ga_5SiO_{14}$) and its isomorphs furthermore have piezoelectric properties that allow one to couple such a mechanical resonator to an electric circuit and to generate high-frequency electric oscillations.

4.1.1 Basic Properties of Piezoelectric Materials

Consider a plate of a piezoelectric material where two opposite surfaces are coated with a thin metallic layer. When stress is applied to this plate, the centres of the positive and negative charges in each unit cell are displaced with respect to each other. Inside the materials the charges are compensated. On the metallised surfaces, however, the uncompensated charges give rise to a voltage between these surfaces depending on the capacitance C of the plate with the two conducting surfaces. As a result, any mechanical oscillation of the piezoelectric plate is directly connected to an oscillating electric voltage. Conversely, a voltage applied to the surfaces generates a strain and an electric ac voltage can be used to excite mechanical oscillations in the plate.

From the several modifications of SiO_2, α-quartz is the thermodynamically stable form at temperatures up to 573 °C. There are distinct properties that make α-quartz unique for building oscillators. First, the high mechanical stiffness and the high elasticity of the material allows one to excite mechanical oscillations of a suitably cut plate with high frequencies and high Q. Second, quartz is a material that can be grown with high purity and in large quantities at low cost. Third, the material can be easily machined.

Langasite ($La_3Ga_5SiO_{14}$) and its isomorphs langanite ($La_3Ga_{5.5}Nb_{0.5}O_{14}$) and langatate ($La_3Ga_{5.5}Ta_{0.5}O_{14}$) have higher Q than the best suitable cuts of quartz and higher piezoelectric coupling. Furthermore these materials have no phase transition up to their melting point around 1400 ° C. Hence, in some applications these and other piezoelectric materials [54] may replace the more familiar quartz.

4.1.2 Mechanical Resonances

The mechanical vibrations that can be excited in quartz specimens can be very complex. In order to derive the frequencies associated with particular modes of motions we restrict ourselves to the simple cases of rectangular plates where we consider the extensional, flexural and shear modes (Fig. 4.1). Consider a rectangular slab of a homogeneous and isotropic medium of length l, width a and height b. Applying a force F to both ends of the slab will stretch its length (Fig. 4.1 a) by Δl. In the elastic regime where Hooke's law holds, the fractional length

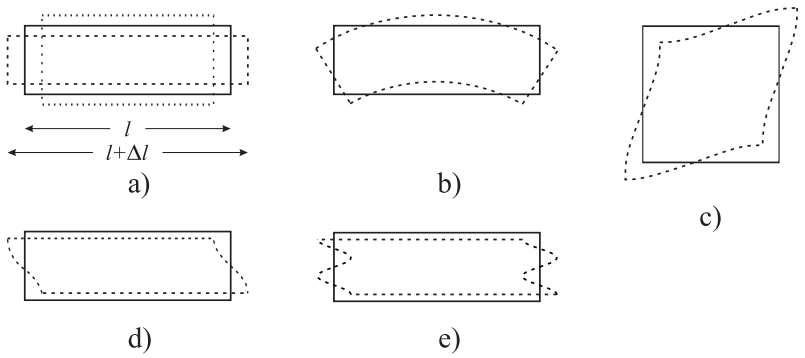


Figure 4.1: Different modes of strain in a quartz crystal plate. Extensional mode (a), flexure mode (b), face shear mode (c), thickness shear mode (fundamental mode, d), and third-overtone thickness shear mode (e).

change $\Delta l/l$ is proportional to the force which can be written as $F = EA\Delta l/l$ with $A = a \cdot b$ and E being Young's modulus of elasticity. Alternatively one writes

$$S = Es \tag{4.1}$$

for the stress $S \equiv F/A$ and the ensuing strain $s \equiv \Delta l/l$. Similarly, as in the case of a spring obeying Hooke's law (2.3), the slab can vibrate where the elastic medium is alternatively stretched and compressed. When the length of the slab is increased, at the same time its width and height are reduced (Fig. 4.1 a). In an isotropic medium the reduction of the thickness occurs in the same way for the width and the height. The respective fractional thickness variations are commonly described by $\Delta a/a = \Delta b/b = -\sigma\Delta l/l$ where σ is referred to as Poisson's ratio. However, in general, (4.1) has to be replaced by a tensor equation since a stress along any direction leads not only to strain along this direction but also in all other directions.

In the following we briefly recall the derivation of the eigenfrequencies of a long thin bar having all surfaces free. Here the thickness variation is small and may be neglected so that the only stress that needs to be considered is along the x direction of the slender slab. The eigenfrequencies of the (longitudinal) vibrations can be derived from the fundamental relationships of Newton's axiom and Hooke's law. In contrast to the isolated mass used to derive (2.4) one has now to treat a large number of mass elements Δm_i coupled by forces $\sum F_{i,j}$ to all other elements. We model the thin slab by a linear chain of mass elements Δm_i where each element at the i^{th} position q_i interacts only with its direct neighbours. Newton's second axiom then reads

$$\frac{d}{dt}\left(\Delta m_i \frac{d}{dt} q_i\right) = F_{i,i+1} - F_{i-1,i} \tag{4.2}$$

for each mass element. Introducing the density $\rho \equiv \Delta m_i/(A\Delta x)$ and the continuous function $u(x,t)$ describing the displacement of a mass element, the transition from the single masses Δm_i to an elastic continuum can be performed by letting

$$\frac{d}{dt}\left(\Delta m_i \frac{d}{dt} q_i\right) \quad \rightarrow \quad A\Delta x \rho \frac{\partial}{\partial t}\frac{\partial u(x,t)}{\partial t} \tag{4.3}$$

and

$$F_{i,i+1} - F_{i-1,i} \quad \rightarrow \quad A\frac{\partial S(x,t)}{\partial z}\Delta x. \tag{4.4}$$

Equating (4.3) and (4.4) according to (4.2) one finds

$$\rho \frac{\partial}{\partial t}\frac{\partial u(x,t)}{\partial t} = \frac{\partial S(x,t)}{\partial x}. \tag{4.5}$$

Relating the strains $S(x,t)$ at any position of the slab and at any time to the stresses, we insert Hooke's law

$$S(x,t) = Es(x,t) = E\frac{\partial u(x,t)}{\partial x} \tag{4.6}$$

into (4.5) and end up with a wave equation for $u(x,t)$

$$\frac{\partial^2}{\partial x^2}u(x,t) - \frac{\rho}{E}\frac{\partial^2}{\partial t^2}u(x,t) = 0. \tag{4.7}$$

The wave equation (4.7) is solved, e.g., by the ansatz

$$u(x,t) = u(x)\exp(i\omega t) \tag{4.8}$$

which leads to

$$\frac{\partial^2}{\partial x^2}u(x) - k^2 u(x) = 0 \tag{4.9}$$

with the wavenumber

$$k = \omega\sqrt{\frac{\rho}{E}}. \tag{4.10}$$

(4.9) is solved either by a cosine or sine function or a linear combination of both. There are, however, so-called "boundary conditions" which restrict the possible solutions. If the ends of the bar are free the stress has to be zero at $x = 0$ and at $x = l$. From (4.6) we find $\partial u(x)/\partial x = 0$ and, hence, it follows that

$$u(x) = A \cos kx \tag{4.11}$$

with $k = m\pi/l$. Combining this result with (4.10) we find equidistant angular eigenfrequencies of the oscillating slab as

$$\omega_m = \frac{m\pi}{l}\sqrt{E/\rho} \tag{4.12}$$

where $m = 1, 2, 3, \cdots$.

For a different limiting case of a thin plate whose length l' (thickness) is small with respect to the transverse dimensions ($l' < a$ and $l' < b$) the other stress components can no longer be neglected. Similar, but more lengthy calculations [55] lead to the eigenfrequencies

$$\omega_m = \frac{m\pi}{l'}\sqrt{\frac{E}{\rho}\frac{1-\sigma}{(1+\sigma)(1-2\sigma)}}. \tag{4.13}$$

Comparison of (4.12) and (4.13) shows that $\omega_m l'$ of (4.13) is always smaller than $\omega_m l$ of (4.12) by the factor $\sqrt{(1-\sigma)/(1-\sigma-2\sigma^2}$ depending on Poisson's ratio. As (4.12) and (4.13) describe the limiting cases of a rectangular block, one expects that the frequency of the longitudinal mode becomes gradually lower as the width of the bar is increased. Consequently, for such a block there are specific eigenfrequencies corresponding to the extensional vibrations along the different directions.

This simple model does not account for the variety of effects that can occur in real plates, e.g., in a quartz plate. First, apart from the extensional vibrations, other modes can be excited (see Fig. 4.1) depending on the direction of the external force applied to the plate. Second, as a result of the mechanical coupling between the different directions described by Poisson's ratio, the respective eigenfrequencies also exhibit coupling effects. Third, a real quartz plate is a highly anisotropic medium as can be concluded already from the complicated shape of the macroscopic quartz crystal (see Fig. 4.2 a) which gives a clue to anisotropic growth along the different directions. Similarly the elastic constants and the coefficients of the thermal expansion differ along the different directions. The latter properties lead to dependencies of the eigenfrequencies on external parameters that degrade the frequency stability of quartz oscillators. These influences could be dramatically reduced since it was discovered that cutting the crystals along selected crystallographic directions leads to devices where the temperature coefficients of the electrical properties can be made exceedingly small as in the so-called AT-cut. The SC-cut furthermore has an excellent stress compensation. On the other hand, LC-cut quartz crystals show an almost linear temperature coefficient and can be used as a quartz thermometer where the variation of the eigenfrequency is used to monitor temperature variations. The locations of some common cuts and their angles with respect to the crystallographic axes, are shown in Fig. 4.2 b) and c).

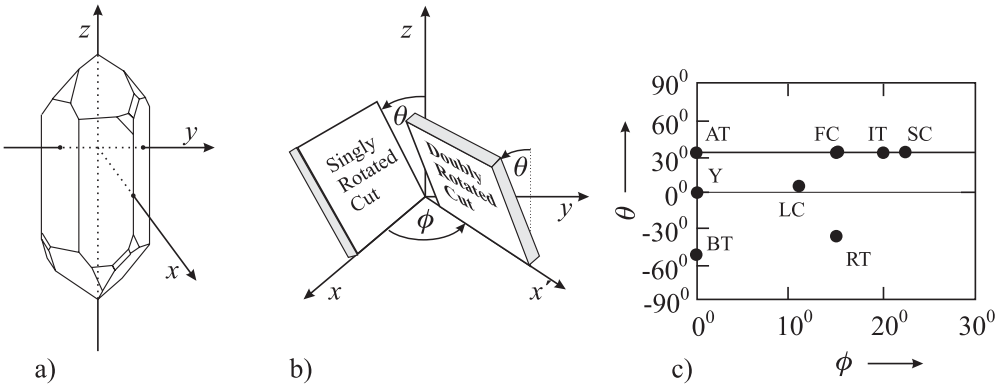


Figure 4.2: a) Natural quartz crystal. b) Orientations of singly and doubly rotated quartz crystal cuts. c) Angles of some important cuts after [2]. For the definition of the angles θ and ϕ see b).

4.1.3 Equivalent Circuit

In the quartz oscillator the energy is swapped between the electric energy stored in the capacitor and the mechanic energy stored in the elastic deformation of the crystal. This energy exchange between its electric and mechanical form is analogous to the case of an oscillating circuit. There, the energy of the system is stored alternatively in the magnetic field of a coil with inductance L and the electric energy of the capacitor having a capacitance C. Hence, the quartz crystal unit is readily represented by an electric equivalent circuit (Fig. 4.3). The resistance R takes into account the dissipation of the energy from the oscillation to mostly thermal energy in the crystal itself and in the mounts. The $C - L - R$ branch is called the motional arm representing the electric equivalent of the mechanical vibrating body of the resonator coupled into the circuit by means of the piezoelectric effect. C_0 is the static capacitance of the electrodes and the leads. To determine the resonance frequency of the crystal oscillator

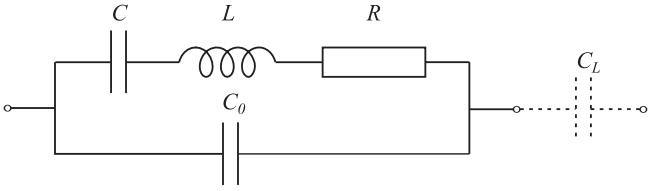


Figure 4.3: Equivalent electrical circuit of a quartz crystal unit.

one calculates the impedance Z of the equivalent circuit of Fig. 4.3 from

$$\frac{1}{Z} = \frac{1}{Z_1} + \frac{1}{Z_2} = \frac{1}{i\omega L + \frac{1}{i\omega C} + R} + \frac{1}{\frac{1}{i\omega C_0}}. \tag{4.14}$$

The impedance Z of the device in general is a complex quantity that determines its current response to an oscillating voltage. The real part is the conventional resistance to electric

current and the imaginary part, referred to as the reactance, is positive when the phase of the current lags behind the voltage. From the typical values of a 4 MHz quartz ($L \approx 100$ mH, $C \approx 0.015$ pF, $C_0 \approx 5$ pF, and $R \approx 100\,\Omega$) one finds that, close to the resonance, the resistance R is small compared to other contributions to the impedance. Hence, neglecting R in the following, one obtains the reactance

$$Z = \frac{\frac{L}{C_0} - \frac{1}{\omega^2 C C_0}}{\frac{1}{i\omega C_0} + i\omega L + \frac{1}{i\omega C}} = \frac{\omega^2 LC - 1}{\frac{\omega C}{i} + i\omega^3 C C_0 L + \frac{\omega C_0}{i}} = \frac{i}{\omega} \frac{\omega^2 LC - 1}{C_0 + C - \omega^2 L C C_0}. \quad (4.15)$$

There are two specific frequencies of the crystal oscillator defined by $Z = 0$ and $Z = \infty$ called the series resonance ν_s and the parallel resonance ν_p, respectively. They can be calculated by finding the zeros of the nominator and the denominator of (4.15) as

$$\nu_s = \frac{1}{2\pi\sqrt{LC}} \qquad \text{series resonance} \quad (4.16)$$

$$\nu_p = \frac{1}{2\pi\sqrt{LC}} \sqrt{1 + \frac{C}{C_0}} \qquad \text{parallel resonance.} \quad (4.17)$$

The frequencies of the series and parallel resonances are closely spaced (Fig. 4.4). In the

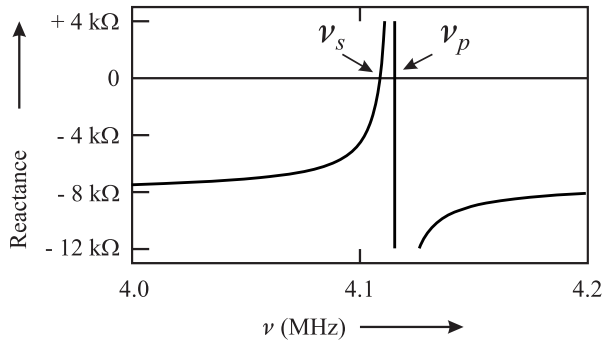


Figure 4.4: Reactance of a typical 4 MHz quartz calculated according to (4.15) using L = 100 mH, C = 0.015 pF and C_0 = 5 pF.

example chosen they differ by 0.15 %. The resonance frequency of the series resonance depends on the well defined quantity LC, whereas the parallel resonance also includes C_0, resulting from the less well defined capacities of the electrodes and leads.

Crystal oscillators make use of the steep slope of the reactance between ν_s and ν_p (Fig. 4.4). In principle the oscillator comprises an amplifier as an active element and the quartz inserted into a feedback loop (Fig. 4.5) which can be realised by a variety of different circuits (see e.g. [22]). The oscillator sustains steady oscillation if the gain in the feedback loop is unity (2.54) and the phase shift is an integer multiple of 2π (2.55). The steep frequency dependence of the reactance allows the oscillator to maintain these conditions with very small frequency deviations even if the parameters of the remaining elements in the oscillating loop vary, e.g., in dependence on the temperature. Consider any phase fluctuation $\Delta\phi_{\mathrm{amp}}$ in the amplifier of Fig. 4.5. To sustain the phase condition of $n2\pi$ for oscillation,

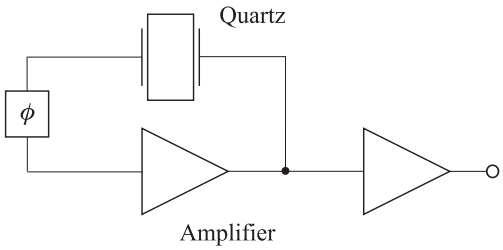


Figure 4.5: Schematics of a quartz oscillator.

this phase fluctuation has to be compensated by an opposite phase fluctuation in the resonator $\Delta\phi_{\rm res} = -\Delta\phi_{\rm amp}$ which causes fluctuations in the frequency of the resonator that can be estimated from the phase shift in the vicinity of the resonance (Fig. 4.4). From (2.62) it follows that

$$\frac{\Delta\nu}{\nu_0} = \frac{1}{2Q_L}\Delta\phi_{\rm amp} \tag{4.18}$$

where Q_L is the loaded Q of the resonator. Optimum frequency stability, hence, requires a high-Q resonator and an amplifier with low phase fluctuations $\Delta\phi_{\rm amp}$.

Often it is necessary to tune the frequency of a quartz oscillator. This can be achieved, e.g., by a load capacitor C_L (Fig. 4.3) that affects the series resonance frequency. It can be calculated from the series impedance of the corresponding equivalent circuit given by (4.15) and adding $1/(i\omega C_L)$ which leads to

$$Z' = \frac{1}{i\omega C_L}\frac{C + C_0 + C_L - \omega^2 LC(C_0 + C_L)}{C_0 + C - \omega^2 LCC_0}. \tag{4.19}$$

From the zero of the nominator one obtains the shifted series resonance frequency as

$$\nu_s' = \frac{1}{2\pi\sqrt{LC}}\sqrt{\frac{C + C_0 + C_L}{C_0 + C_L}} = \nu_s\sqrt{1 + \frac{C}{C_0 + C_L}}. \tag{4.20}$$

For $C \ll C_0 + C_L$ the square root can be expanded leading to

$$\omega_s' = \omega_s\left[1 + \frac{C}{2(C_0 + C_L)}\right]. \tag{4.21}$$

Hence, the load capacitance C_L changes the frequency of the series resonance by

$$\frac{\Delta\nu_s}{\nu_s} = \frac{C}{2(C_0 + C_L)}. \tag{4.22}$$

Often the load capacitor comprises a fixed capacitor in series with a varactor, i.e., a capacitor whose capacitance can be varied by an applied voltage in order to fine-tune the frequency of the quartz-crystal oscillator. Such a Voltage Controlled Crystal Oscillator (VCXO) compromises between the good intrinsic stability of a crystal oscillator and the tunability

($10^{-5} \lesssim \Delta\nu/\nu \lesssim 10^{-4}$). The frequency stability of a VCXO is typically a few parts per million within the normal operational temperature range.

The production of quartz plates thin enough to produce fundamental frequencies higher than 30 MHz meets with difficulties. To realise quartz oscillators with higher frequencies, electrical circuits are used that, e.g., selectively excite higher modes of the quartz oscillator or lock an LC oscillator via a phase-locked loop (PLL) to higher harmonics of a low-frequency quartz.

4.1.4 Stability and Accuracy of Quartz Oscillators

The noise and instability of the phase and frequency of a quartz oscillator results from various contributions such as ageing, the sensitivity to external influences such as temperature, acceleration, magnetic fields, or noise in the electronic circuits, to name only a few. Ageing can result from relief of stress in the crystal or the bonds which settles down after some months, from the absorption or desorption of molecules, from dc bias on the crystal, from load reactance change (see (4.22)) but also for a number of other reasons. The mounting of the quartz plate is of utmost importance to achieve high Q, low cross coupling between the different modes (see Fig. 4.1) and minimum external stress. Coating the quartz plate with metallic electrodes, in general leads to stress that is partially reduced during the lifetime of the resonator and, hence, leads to a variation of the resonance frequency which contributes to the "ageing". This source of ageing can be avoided in a so-called "BVA" electrodeless resonator structure [56] (Fig. 4.6). There the electrodes are on auxiliary plates with a small gap of a

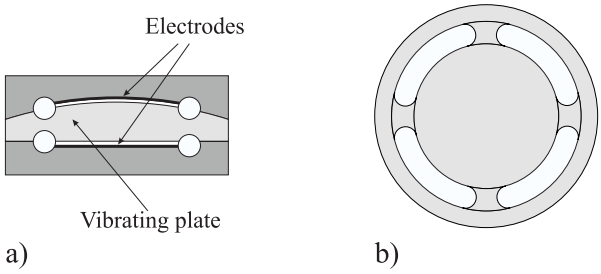


Figure 4.6: BVA quartz resonator structure. a) Side view of vibrating quartz plate clamped by two quartz holders with deposited electrodes. b) Top view of vibrating plate with four quartz bridges connecting the vibrating part to the outer ring being clamped.

few micrometres to the vibrating plate rather than on the vibrating plate itself. The absence of the electrodes from the vibrating plate furthermore allows the achievement of a high Q which is not degraded by the damping of the electrode material.

From the external influences, temperature fluctuations affect the quartz crystal's frequency despite of the use of temperature compensated cuts. Higher immunity to temperature fluctuations can be obtained by use of a Temperature Compensated Crystal Oscillator (TCXO) where the temperature sensitivity of the quartz crystal is compensated by temperature sensitive reactances. Better performance (see Table 4.1) is obtained with the Microcomputer Compensated

Crystal Oscillator (MCXO) which makes use of a dual mode oscillator that operates simultaneously on the fundamental frequency ν_1 and on the third overtone $(\nu_3 \approx 3\nu_1)$.

Table 4.1: Properties of quartz crystal oscillators after [2].

	TCXO	MCXO	OCXO
Fractional inaccuracy / year	2×10^{-6}	5×10^{-8}	1×10^{-8}
Ageing / year	5×10^{-7}	2×10^{-8}	5×10^{-9}
Temp. stability / year (-55° C to +85° C)	5×10^{-7}	3×10^{-8}	1×10^{-9}
$\sigma_y(\tau = 1\,\mathrm{s})$	1×10^{-9}	3×10^{-10}	1×10^{-12}

Since the frequency difference $3\nu_1 - \nu_3$ exhibits a monotonous almost linear variation with the temperature of about -14 Hz/K, in the MCXO this frequency difference is monitored by a microcomputer that corrects the output frequency accordingly. In the Oven Controlled Crystal Oscillator (OCXO) a crystal with a zero temperature-coefficient cut is operated in an oven at a temperature of $\gtrsim 80°$ C kept constant by a thermostat. A typical commercial high performance oscillator operating at 10 MHz can have an instability of a few parts in 10^{11} over periods of seconds to hours.

The frequency stability of a quartz oscillator depends on a variety of ambient parameters. Even though quartz is diamagnetic, quartz crystal resonators are slightly sensitive to external magnetic fields that can influence the circuitry with fractional frequency changes of $2 \times 10^{-8}\,\mathrm{T}^{-1}$ [57]. The quartz resonator is particularly sensitive to vibrations and shocks. The associated frequency excursions result from the sensitivity of the resonator to stress. The best available SC-cut BVA resonators show a fractional frequency change with acceleration between $10^{-10}\,\mathrm{g}^{-1}$ and a few times $10^{-12}\,\mathrm{g}^{-1}$ [57]. Fluctuations of the ambient humidity and pressure affect the frequency by deforming the oscillator package and the oscillator circuitry. Fluctuations of the supply voltage can change the resonator drive level and the load reactance, which in turn can change the amplitude or phase of the signal in the oscillator loop. Noise in the electronic circuit represents another source of frequency instability of quartz oscillators.

Hyperstable quartz oscillators at 10 MHz employing a 3^{rd} overtone, SC-cut BVA quartz have been shown [57] to have instabilities below $\sigma_y(\tau) < 10^{-13}$ for $0.3\,\mathrm{s} \leq \tau \leq 500\,\mathrm{s}$. Ageing ratios of the best specimens between 2×10^{-11} per day and 5×10^{-13} per day have been observed.

4.2 Microwave Cavity Resonators

In microwave frequency standards cavities are often used where electromagnetic fields are contained in structures bounded by electrically conducting surfaces. The boundaries affect the properties of the electromagnetic fields inside the cavity. For a quantitative analysis of such microwave cavity-resonators, Maxwell's equations have to be solved with the boundary conditions resulting from the particular shape of the cavity. Here, we restrict ourselves to the right circular cylindrical resonator which is wide-spread in frequency standards, e.g., in the hydrogen maser (Section 8.1), the rubidium clock (Section 8.2) or the caesium fountain clock (Section 7.3). Right cylindrical resonators can be thought of as resulting from a finite cylindrical wave guide which is sealed at its end with metal caps (Fig. 4.7).

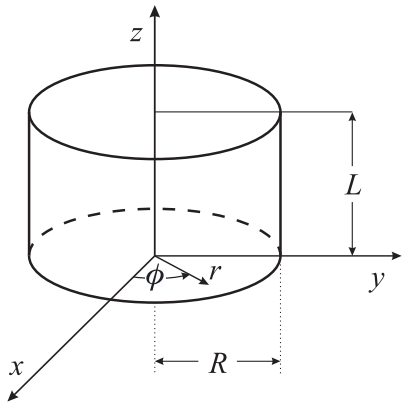


Figure 4.7: Right circular cylinder with coordinate system as used in a microwave cavity.

We take the z axis of the coordinate system to coincide with the axis of the cylindrical cavity of length L and radius R with circular end faces. The radial distance from the axis is denoted as r, the axial distance z is conveniently chosen having a zero position $z = 0$ at the surface at one end plane, and the angle ϕ is the angle around the axis with respect to a defined radial direction. In the mathematical description of the properties of those resonators we will apply a two-step procedure and begin by using Maxwell's equations to find the electromagnetic waves that can propagate in the circular wave guide. Afterwards we look for the standing waves and resonance frequencies that result from the introduction of the metal caps. This type of empty microwave resonator will be treated in the following section. Later we will also discuss resonators filled with a dielectric medium finding applications as fly wheels in frequency standards.

4.2.1 Electromagnetic Wave Equations

The treatment often starts [58] with Maxwell's equations in differential form for the electric field $\vec{E}$ and the magnetic induction $\vec{B}$ in a volume of space which is free of currents and charge and is filled with a uniform, non-dissipative medium.

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \tag{4.23}$$

$$\vec{\nabla} \times \vec{B} = \mu\mu_0\epsilon\epsilon_0 \frac{\partial \vec{E}}{\partial t} \tag{4.24}$$

$$\vec{\nabla} \cdot \vec{E} = 0 \tag{4.25}$$

$$\vec{\nabla} \cdot \vec{B} = 0. \tag{4.26}$$

Here, $\mu_0 = 4\pi \times 10^{-7}\,\mathrm{V\,s\,/\,(A\,m)}$ and $\epsilon_0 = 8.854 \times 10^{-12}\,\mathrm{A\,s\,/\,(V\,m)}$ are the permeability and the permittivity of the vacuum, respectively. From these equations a wave equation for the electric field can be derived by taking the curl of (4.23), i.e., $\vec{\nabla} \times (\vec{\nabla} \times \vec{E}) = \vec{\nabla} \times (-\partial\vec{B}/\partial t)$. Using the well kown relationship $\vec{\nabla} \times \vec{\nabla} \times \vec{E} = \vec{\nabla}(\vec{\nabla} \cdot \vec{E}) - \vec{\nabla} \cdot (\vec{\nabla}\vec{E})$ one ends up with

$0 - \nabla^2 \vec{E} = \vec{\nabla} \times (-\partial \vec{B}/\partial t) = -\partial/\partial t \vec{\nabla} \times \vec{B} = -\mu\mu_0\epsilon\epsilon_0 \partial^2 \vec{E}/\partial t^2$ leading to wave equations for the electric field and similarly for the magnetic field as follows

$$\nabla^2 \vec{E} - \frac{\mu\epsilon}{c^2} \frac{\partial^2 \vec{E}}{\partial t^2} = 0 \tag{4.27}$$

$$\nabla^2 \vec{B} - \frac{\mu\epsilon}{c^2} \frac{\partial^2 \vec{B}}{\partial t^2} = 0. \tag{4.28}$$

Here, we have made use of the speed of light in the vacuum $c = 1/\sqrt{\mu_0\epsilon_0}$. Assuming a harmonic time dependence, e.g., $\exp(i\omega t)$ of the fields, one derives a differential equation for the spatial part of the electromagnetic wave

$$\nabla^2 \vec{E} + k_0^2 \vec{E} = 0 \tag{4.29}$$

$$\nabla^2 \vec{B} + k_0^2 \vec{B} = 0 \tag{4.30}$$

with

$$k_0 = \frac{\sqrt{\mu\epsilon}}{c}\omega \tag{4.31}$$

being the wave number. To solve these wave equations one has to take into account boundary conditions similarly as in Section 4.1.2. For simplicity, we assume that the walls have infinite conductivity and no losses. For an ideal conducting surface the tangential electric field on the surface must vanish since the charges inside a perfect conductor would move under the influence of this field, thereby creating a field distribution of the charges that completely counteracts the applied field. At the boundary between two media "1" and "2" the tangential component of the electric field $\vec{E}$ is continuous as well as the normal component of the magnetic induction $\vec{B}$, i.e.,

$$\vec{n} \times \vec{E}_1 = \vec{n} \times \vec{E}_2 \tag{4.32}$$

$$\vec{n} \cdot \vec{B}_1 = \vec{n} \cdot \vec{B}_2 \tag{4.33}$$

where $\vec{n}$ denotes the normal vector of the surface. Hence, for $\vec{E}_2 = 0$ and $\vec{B}_2 = 0$ in the metal the boundary conditions take the simple forms

$$\vec{n} \times \vec{E}\,|_S = 0 \tag{4.34}$$

$$\vec{n} \cdot \vec{B}\,|_S = 0. \tag{4.35}$$

Hence, the electric field $\vec{E}$ at the surface S of the cylinder has to be perpendicular to the surface leading to (4.34). Similarly, the normal component of the magnetic field vanishes at the surface of a conductor of infinite conductivity (see (4.35)) since any magnetic field reaching into the conductor would induce internal currents with opposite magnetic field. As the wave equations (4.27) and (4.28) only deal with $\vec{E}$ and $\vec{B}$ their solutions only require the boundary conditions (4.34) and (4.35), respectively. The solutions to the wave equations are the types of waves that are possible in the wave guide. To this end one only needs to know the z component of $\vec{E}$ for (4.29) or $\vec{B}$ for (4.30) in order to derive all other components, as will become obvious in the following.

4.2.2 Electromagnetic Fields in Cylindrical Wave Guides

For the electric field we utilise a cylindrical coordinate system with the coordinates r, ϕ, and z and the corresponding unit vectors $\hat{e}_r, \hat{e}_\phi, \hat{e}_z$, respectively. The fields of the wave travelling in the wave guide along the $+z$ direction are given by

$$\vec{E}(\vec{r}) = \vec{E}(r,\phi)e^{-ikz} \tag{4.36}$$

$$\vec{B}(\vec{r}) = \vec{B}(r,\phi)e^{-ikz} \tag{4.37}$$

with k the wave number in the wave guide. If the axial components are known, the radial and azimuthal components can be determined from Maxwell's equations (4.23) and (4.24) as follows. For example, consider (4.24) with the curl of $\vec{B}$ in cylindrical coordinates given as

$$\vec{\nabla} \times \vec{B} = \left(\frac{1}{r}\frac{\partial B_z}{\partial \phi} - \frac{\partial B_\phi}{\partial z}\right)\hat{e}_r + \left(\frac{\partial B_r}{\partial z} - \frac{\partial B_z}{\partial r}\right)\hat{e}_\phi + \left(\frac{1}{r}\frac{\partial (rB_\phi)}{\partial r} - \frac{1}{r}\frac{\partial B_r}{\partial \phi}\right)\hat{e}_z. \tag{4.38}$$

Applying (4.24) and (4.38) by components, one obtains for the z dependent fields of (4.36) and (4.37)

$$i\omega\mu\mu_0\epsilon\epsilon_0 E_r = \frac{1}{r}\frac{\partial B_z}{\partial \phi} + ikB_\phi, \tag{4.39}$$

$$i\omega\mu\mu_0\epsilon\epsilon_0 E_\phi = -ikB_r - \frac{\partial B_z}{\partial r} \tag{4.40}$$

$$i\omega\mu\mu_0\epsilon\epsilon_0 E_z = \frac{1}{r}\frac{\partial (rB_\phi)}{\partial r} - \frac{1}{r}\frac{\partial B_r}{\partial \phi}, \tag{4.41}$$

and similarly from (4.23)

$$-i\omega B_r = \frac{1}{r}\frac{\partial E_z}{\partial \phi} + ikE_\phi, \tag{4.42}$$

$$-i\omega B_\phi = -ikE_r - \frac{\partial E_z}{\partial r}, \tag{4.43}$$

$$-i\omega B_z = \frac{1}{r}\frac{\partial (rE_\phi)}{\partial r} - \frac{1}{r}\frac{\partial E_r}{\partial \phi}. \tag{4.44}$$

These equations can be combined (e.g., (4.39) and (4.43) lead to (4.45)) to express the transverse components E_r, E_ϕ, B_r, and B_ϕ by the axial components E_z and B_z as follows

$$E_r = -\frac{ik}{\omega^2\mu\mu_0\epsilon\epsilon_0 - k^2}\left(\frac{\partial E_z}{\partial r} + \frac{\omega}{k}\frac{1}{r}\frac{\partial B_z}{\partial \phi}\right), \tag{4.45}$$

$$E_\phi = -\frac{ik}{\omega^2\mu\mu_0\epsilon\epsilon_0 - k^2}\left(\frac{1}{r}\frac{\partial E_z}{\partial \phi} - \frac{\omega}{k}\frac{\partial B_z}{\partial r}\right), \tag{4.46}$$

$$B_r = -\frac{ik}{\omega^2\mu\mu_0\epsilon\epsilon_0 - k^2}\left(\frac{\partial B_z}{\partial r} - \mu\mu_0\epsilon\epsilon_0\frac{\omega}{k}\frac{1}{r}\frac{\partial E_z}{\partial \phi}\right), \tag{4.47}$$

$$B_\phi = -\frac{ik}{\omega^2\mu\mu_0\epsilon\epsilon_0 - k^2}\left(\frac{1}{r}\frac{\partial B_z}{\partial \phi} + \mu\mu_0\epsilon\epsilon_0\frac{\omega}{k}\frac{\partial E_z}{\partial r}\right). \tag{4.48}$$

Hence, the wave equations have to be solved with the appropriate boundary conditions only for the z components, e.g., E_z and B_z for a cylindrical resonator (Fig. 4.7) from which later

the other components can be derived by use of (4.45) to (4.48). From (4.29) and (4.36) one finds

$$\nabla^2 \left[E_z(r,\phi)e^{-ikz}\right] + k_0^2 E_z(r,\phi)e^{-ikz} = 0 \tag{4.49}$$

$$\nabla^2 \left[E_z(r,\phi)\right] e^{-ikz} + E_z(r,\phi)\nabla^2\left(e^{-ikz}\right) + k_0^2 E_z(r,\phi)e^{-ikz} = 0 \tag{4.50}$$

or

$$\nabla^2 \left[E_z(r,\phi)\right] + (k_0^2 - k^2)E_z(r,\phi) = 0 \tag{4.51}$$

$$\nabla^2 \left[B_z(r,\phi)\right] + (k_0^2 - k^2)B_z(r,\phi) = 0. \tag{4.52}$$

The electromagnetic field inside the cavity reflects the cylindrical symmetry and the internal boundary of the electrically conducting surface leads to different boundary conditions on the magnetic induction $\vec{B}$ and on the electric field $\vec{E}$ (4.34) and (4.35). Since the boundary conditions in general cannot be satisfied simultaneously there are two distinct categories of field configurations. They are referred to as the transverse magnetic (TM) case with $B_z = 0$ everywhere and the transverse electric (TE) case with $E_z = 0$ everywhere [58]. Alternatively, the TM waves and TE waves are referred to as H waves and E waves, respectively.

To determine the r and ϕ dependence $\psi(r,\phi)$ of E_z and B_z the wave equations (4.27) and (4.28) have to be solved with the Laplacean operator in cylindrical coordinates as follows

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \phi^2} + \frac{\partial^2}{\partial z^2}. \tag{4.53}$$

Insertion of (4.53) into (4.51) yields

$$\left[\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \phi^2} + \gamma^2\right] E_z(r,\phi) = 0 \tag{4.54}$$

with

$$\gamma^2 \equiv k_0^2 - k^2. \tag{4.55}$$

We look for solutions to (4.54) of the form

$$E_z(r,\phi) = A(r)\Phi(\phi) \tag{4.56}$$

and derive

$$r^2 \frac{\frac{\partial^2 A(r)}{\partial r^2}}{A(r)} + r\frac{\frac{\partial A(r)}{\partial r}}{A(r)} + r^2\gamma^2 = -\frac{\frac{\partial^2 \Phi(\phi)}{\partial \phi^2}}{\Phi(\phi)}. \tag{4.57}$$

Since the left-hand side and the right-hand side of (4.57) depend independently on r and ϕ, respectively, both sides are equal to the same real constant which we call m^2. We obtain two regular differential equations

$$\frac{\partial^2}{\partial r^2}A(r) + \frac{1}{r}\frac{\partial}{\partial r}A(r) + \left(\gamma^2 - \frac{m^2}{r^2}\right)A(r) = 0 \tag{4.58}$$

$$\frac{\partial^2}{\partial \phi^2}\Phi(\phi) + m^2\Phi(\phi) = 0. \tag{4.59}$$

Particular solutions to (4.59) are $\sin m\phi$ and $\cos m\phi$. In order to ensure that E_z and B_z are single-valued functions of ϕ the azimuthal solution has to fulfil $m = 0, 1, 2, \cdots$. To each $m \neq 0$ there is a pair of degenerate azimuthal eigenfunctions, one with $\sin\phi$ and one with $\cos\phi$. There is, however, only a single solution with rotational symmetry $(n = 0)$.

(4.58) is known as Bessel's differential equation [59]. From the particular solutions to (4.59) only the Bessel functions of the first kind $J_{\pm m}(\gamma r)$ (see (2.49)) are relevant which stay finite for $r = 0$. Since the boundary condition (4.34) requires that $E_z(r = R) = 0$, the Bessel function $J_m(\gamma r)$ of order m is a solution to the problem only if it is zero at $r = R$. Thus,

$$x_{mn} = \gamma_{mn} R \tag{4.60}$$

is the n^{th} root of the equation $J_m(\gamma R) = 0$ with the first few values given in Table 4.2.

Table 4.2: Roots of the Bessel function of the first kind of order $m (J_m(x) = 0)$.

m	x_{m1}	x_{m2}	x_{m3}	x_{m4}
0	2.405	5.520	8.654	11.792
1	3.832	7.016	10.173	13.324
2	5.136	8.417	11.620	14.796
3	6.380	9.761	13.015	16.223

Table 4.3: Maxima or minima of the Bessel function of the first kind of order m $(J'_m(x) = 0)$.

m	x'_{m1}	x'_{m2}	x'_{m3}	x'_{m4}
0	3.832	7.016	10.173	13.324
1	1.841	5.331	8.536	11.706
2	3.054	6.706	9.969	13.170
3	4.201	8.015	11.346	14.586

Hence, the solution for the z component of the electric field is

$$E_z(r, \phi, z) = E_0 J_m\left(x_{mn}\frac{r}{R}\right)\left\{\begin{array}{l}\sin m\phi \\ \cos m\phi\end{array}\right. \exp(-ikz). \tag{4.61}$$

Using (4.61) and $\vec{\nabla} \times \vec{E} = -i\omega\vec{B}$ (which one derives from (4.23) and a harmonic time dependence of the fields) one immediately finds that B_z vanishes. Hence, (4.61) corresponds to a solution where the magnetic field has only transverse components. This solution is therefore the transverse magnetic (TM) wave sometimes called the E wave.

We now start with (4.30) and use the boundary condition (4.35). At the curved surface of the cavity the fields E_ϕ and B_r must vanish. This is equivalent to the condition that the derivative of the Bessel function $J'_0(\gamma r)$ is zero at $r = R$. Similar to the Bessel function itself, the resonance frequencies of a cylindrical resonator are given by the zeros of $J'_0(\gamma r) = 0$ labelled as $x'_{mn} = \gamma_{mn} R$ and shown in Table 4.3 for the first few values. Hence,

$$B_z(r, \phi, z) = B_0 J_m\left(x'_{mn}\frac{r}{R}\right)\left\{\begin{array}{l}\sin m\phi \\ \cos m\phi\end{array}\right. \exp(-ikz) \tag{4.62}$$

holds for the transverse electric (TE) waves (H waves).

4.2.3 Cylindrical Cavity Resonators

We now consider the fields in cylindrical resonators. There, the wave travelling along the $+z$ direction described by the e^{-ikz} factor is reflected by the end cap. Neglecting losses, the

reflected wave has the same amplitude but an e^{+ikz} factor which leads to standing waves with the z dependence $A\sin kz + B\cos kz$. We furthermore have to take into account the boundary conditions at the end caps

$$k = q\frac{\pi}{L}, \qquad q = 0, 1, 2, 3\cdots . \tag{4.63}$$

Due to the boundary conditions (4.34) the transverse components of the electric field E_r and E_ϕ must vanish at the bottom ($z = 0$) and at the top ($z = L$) of the cylinder. On the other hand, the z component of the magnetic field has to vanish at the bottom and the top of the cylinder (see (4.35)). Hence, eigenoscillations being TM with respect to z are given by

$$E_z(r,\phi,z) = E_0 J_m\left(x_{mn}\frac{r}{R}\right)\left\{\begin{array}{c}\sin m\phi\\ \cos m\phi\end{array}\right.\cos\left(\frac{q\pi z}{L}\right) \quad \text{with} \begin{array}{rcl} m &=& 0,1,2,\cdots\\ n &=& 1,2,3,\cdots\\ q &=& 0,1,2,\cdots\end{array} \tag{4.64}$$

and eigenoscillations being TE with respect to z are given by

$$B_z(r,\phi,z) = B_0 J_m\left(x'_{mn}\frac{r}{R}\right)\left\{\begin{array}{c}\sin m\phi\\ \cos m\phi\end{array}\right.\sin\left(\frac{q\pi z}{L}\right) \quad \text{with} \begin{array}{rcl} m &=& 0,1,2,\cdots\\ n &=& 1,2,3,\cdots\\ q &=& 1,2,3,\cdots\end{array} \tag{4.65}$$

The three integers $m = 0, 1, 2, \cdots$, $n = 1, 2, 3, \cdots$ and q are related to the numbers of zeros of the fields along the ϕ, r and z coordinates, respectively. Furthermore, m denotes the order of the Bessel function of the first kind. The field configurations determined by these integers are called the modes of the cavity. Instead of (4.55) the eigenvalue problem now leads to

$$\gamma_{mn}^2 = \mu\epsilon\frac{\omega^2}{c^2} - \left(\frac{q\pi}{L}\right)^2 \tag{4.66}$$

from where the resonant angular frequencies ω_{mnq} associated with these modes can be calculated using the corresponding zeros $x_{mn} = \gamma_{mn}R$ from Table 4.2 as

$$\nu_{mnq}^{(TM)} = \frac{c}{2\pi\sqrt{\mu\epsilon}}\sqrt{\frac{x_{mn}^2}{R^2} + \frac{q^2\pi^2}{L^2}}. \tag{4.67}$$

Using $D = 2R$, (4.67) can be given as a linear relation for $(D/\lambda)^2$ as function of the dimension $(D/L)^2$ (see Fig. 4.8 a).

Similarly, the resonance frequencies for the TE waves are

$$\nu_{mnq}^{(TE)} = \frac{c}{2\pi\sqrt{\mu\epsilon}}\sqrt{\frac{(x')_{mn}^2}{R^2} + \frac{q^2\pi^2}{L^2}} \tag{4.68}$$

(see Fig. 4.8 b). In general, for each $m \neq 0$ according to (4.56) there are two degenerate oscillations with $\sin m\phi$ and $\cos m\phi$ (Fig. 4.8). The oscillation with the lowest eigenfrequency is TE_{111} for $2R/L < 0.985$ and TM_{010} for $2R/L > 0.985$.

As an example, consider the TE_{011} resonance in a right cylindrical resonator (Fig. 4.9). In frequency standards like the Cs fountain or the hydrogen maser this mode is often used to

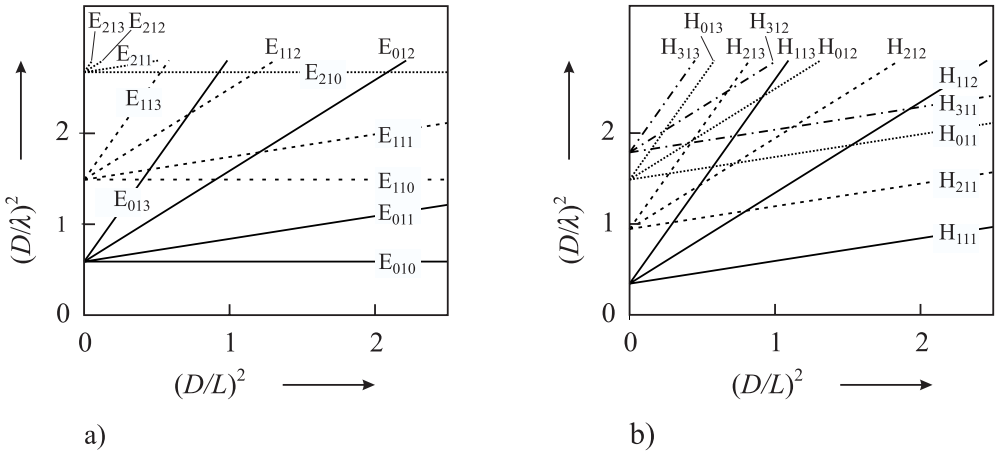


Figure 4.8: Resonance wavelengths in a cylindrical resonator with height L and radius $R = D/2$. a) TM waves (E waves) calculated according to (4.67) and b) TE waves (H waves).

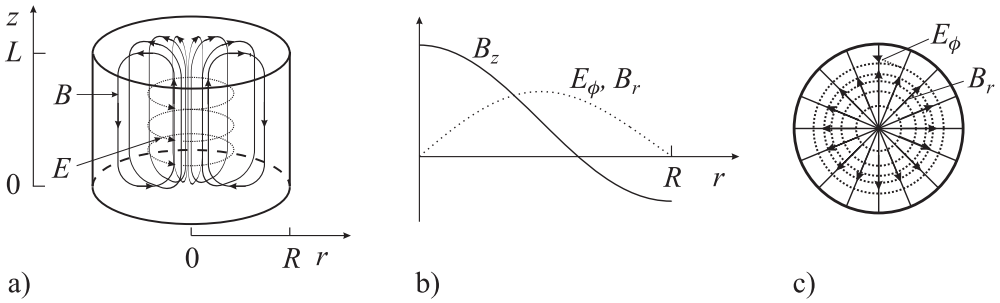


Figure 4.9: a) Magnetic (full lines) and electric field lines (dashed lines) of a resonator mode TE_{011} in a right circular microwave cavity. b) Radial field components according to (4.69), (4.72) and (4.73) c) Top view.

interrogate or to excite magnetic hyperfine transitions. Atoms supposed to interact with the magnetic field enter and leave the resonator through small holes in the centre of the bottom and top of the right circular cylinder.

From $m = 0$ there is no azimuthal dependence of the magnetic field and the z dependence is taken from (4.65) using $q = 1$. Hence, the magnetic field has a maximum along the z axis of the cylinder with

$$B_z = B_0 J_0(3.832\frac{r}{R}) \sin\left(\frac{\pi z}{L}\right) \exp(i\omega t) \tag{4.69}$$

$$E_z = 0. \tag{4.70}$$

The remaining components of the TM_{011} mode can be calculated by applying (4.45) to (4.48). By insertion of (4.69) and (4.70) into (4.45) one finds

$$E_r = 0. \tag{4.71}$$

Similarly, we obtain

$$E_\phi = B_0 \frac{\omega}{\omega^2 \mu\mu_0 \epsilon\epsilon_0 - k^2} J_1(3.832\frac{r}{R}) \cos\left(\frac{\pi z}{L}\right) \quad \text{and} \tag{4.72}$$

$$B_r = B_0 \frac{k}{\omega^2 \mu\mu_0 \epsilon\epsilon_0 - k^2} J_1(3.832\frac{r}{R}) \cos\left(\frac{\pi z}{L}\right), \tag{4.73}$$

where we have used [59]

$$J'_m(x) \equiv \frac{dJ_m(x)}{dx} = \frac{m}{x} J_m(x) - J_{m+1}(x). \tag{4.74}$$

The factor i in (4.45) – (4.48) leads to the phase shift of $\pi/2$ between E_ϕ (see (4.72)), B_r (see (4.73)) and B_z (see (4.70)). From insertion of (4.69) and (4.70) into (4.48) one finds

$$B_\phi = 0. \tag{4.75}$$

The radial dependencies of B_z, E_ϕ and B_r are shown in Fig. 4.9.

4.2.4 Losses due to Finite Conductivity

For finite conductivity of the walls of the resonator the high-frequency electromagnetic field penetrates into the metallic walls of the cavity. At the same time the electric currents in the walls suffer from ohmic losses and the eigenoscillations are damped. The field distribution inside cylindrical cavities, however, does not differ considerably from that of an ideal resonator with infinite conductivity. In particular, for eigenoscillations (m n q) with one index equal to zero these modifications of the field are small. A fraction of the energy of the electromagnetic wave is dissipated continuously to heat by the ohmic losses of the wall currents and hence the energy flux in the wall decreases exponentially with a characteristic length. In a regular conducting material such as, e.g., copper this characteristic length is given by the skin depth δ_S of a few micrometres. In a superconducting material the penetration depth is given by the much smaller London depth of, e.g., $\lambda_L \approx 30\,\text{nm}$ in niobium.

To determine the quality factor (see (2.39)) of a cavity mode, the ratio of the stored electromagnetic energy and the power dissipated into the walls

$$dW/dt = R_s \oint |H_t|^2 dA \tag{4.76}$$

has to be calculated [58] where the integration has to be extended over all walls of the cavity. H_t are the components of the magnetic field transverse to the surfaces of the cavity and R_s is the surface resistivity. From such calculations, in general, the quality factor can be written as

$$Q \equiv \omega \frac{W}{-dW/dt} = \frac{\Gamma}{R_s} \tag{4.77}$$

where $\Gamma = \mu_0 c G(R/L)$ is of the order of the resistivity of the vacuum $\mu_0 c = 376.73\,\Omega$ for low-order modes and the geometrical factor $G(R/L)$ is of order unity. The surface resistivity of copper and superconducting niobium (at 1.8 K) of $R_s \approx 5\,\text{m}\Omega$ and $R_s \approx 7\,\text{n}\Omega$ leads to

a quality factor of 50 000 and 4 $\times 10^{10}$, respectively. In general the TE_{0nq} modes have the highest Q factors as a result of the associated low wall losses. If microwave resonators with highest Q factors are required superconducting materials are employed. Ultrastable superconducting cavity resonators [30] are used in basic research [60, 61], in accelerator physics and for space applications [62].

4.2.5 Dielectric Resonators

Near 10 GHz, monocrystalline sapphire shows a high permittivity of $\epsilon = 11.5$ along the crystal's c axis allowing one to build compact cavities (see (4.67) and (4.68)). Consider a sapphire cavity resonator exhibiting cylindrical symmetry with the cavity axis coinciding with the sapphire crystal's c axis and the sapphire surface coated with a conducting material. Such a resonator can be thought of as a vacuum cavity filled with the dielectric sapphire material with the transverse electric (TE_{mnq}) or transverse magnetic (TM_{mnq}) modes described in Section 4.2.2.

The unloaded quality factor of such a cavity is given [30] as

$$Q = \frac{1}{R_s \Gamma^{-1} + p_\epsilon \tan\delta + p_\mu \chi''}. \tag{4.78}$$

where p_ϵ and p_μ are the electric and magnetic filling factors, respectively. Apart from the losses in the metallic shield described by R_S/Γ (see (4.77)) the losses of the medium reduce the Q factor. The loss tangent $\tan\delta \equiv \epsilon''/\epsilon'$ of the dielectric material is determined by the real and imaginary part ϵ' and ϵ'', respectively, of the relative dielectric constant. χ'' is the imaginary part of the ac susceptibility resulting from paramagnetic impurities. For an ideally conducting shield and no paramagnetic impurities, the Q factor is determined by the dielectric losses that can be calculated from (2.39) and from the maximal electric energy density leading to $W = \epsilon'\epsilon_0 \int |E|^2 dV$ and from $-dW/dt = \omega_0 \epsilon'' \epsilon_0 \int |E|^2 dV$ as $Q = 1/\tan\delta$. When the shield is a direct metallic coating of the sapphire body the surface resistance has maximum influence on the quality factor Q. Hence, in typical resonators, the metallic shield is placed at some distance from the surface and a mode configuration is chosen which is optimally confined in the dielectric medium, e.g., "whispering gallery modes" [1] (WG modes; see Fig. 4.10). In general the modes in such an arrangement are hybrid, but whispering gallery modes can have dominant axial electric fields referred to as E modes or quasi-TM or WGH modes. In the case of a dominant axial magnetic field dependence they are called H modes or quasi-TE or WGE modes.

For a well designed sapphire resonator the Q factor of the unloaded cavity (4.78) is governed by the loss tangent of the dielectric material. At cryogenic temperatures, unloaded quality factors can be between 10^8 at 50 K and 10^{10} at 2 K for frequencies near 12 GHz [30]. The frequency stability of such a resonator is limited by temperature stability since, e.g., the mechanical dimensions as well as the permittivity ϵ have significant temperature dependencies. Different techniques such as, e.g., paramagnetic compensation, dielectric compensation or mechanical compensation have been used to achieve compensation of the different

[1] These modes are named after the "whispering gallery" in St. Paul's cathedral in London where the acoustics allow one to hear a whispered word at the opposite end of the gallery.

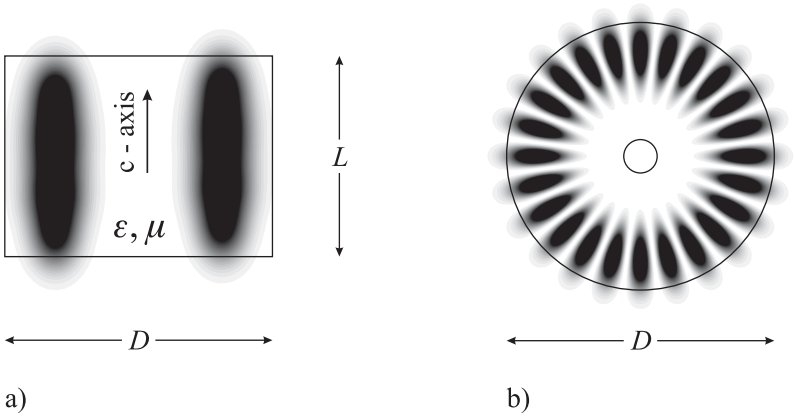


Figure 4.10: Example of the electric field configuration (dark area) in a dielectric resonator with azimuthal mode number $m = 12$ (metallic shield not shown). a) Side view. b) Top view.

temperature-dependent effects [63]. The first technique uses the contribution of a suitably chosen portion of paramagnetic ions in the dielectric to compensate the temperature dependence of the permittivity by the temperature dependence of the magnetic susceptibility. The second method uses compensation with two different dielectric materials, e.g., rutile and sapphire showing opposite slopes of the temperature dependence of the dielectric constant and having low microwave losses at the same time. Composite sapphire-rutile resonators with a turning point in the temperature dependence of the resonance frequency show excellent frequency stability and very high Q values. The third method uses two different materials with different thermal expansion to compensate for the change in ϵ. Various concepts have been realised that led to exceptional short and medium-term stabilities [63]. For integration times τ between a few seconds and about one hundred seconds, the best sapphire oscillators can reach a flicker floor in the Allan deviation at $\sigma(\tau) \approx 3 \times 10^{-16}$ [30]. Ultra low-noise microwave resonators based on dielectric materials like single-crystal sapphire, have proven to be excellent flywheel oscillators, e.g., for Cs fountain clocks [18, 64] or for deep-space applications (see Section 13.1.2.2, [65]).

4.3 Optical Resonators

Optical resonators differ from the microwave resonators in the sense that the wavelength of about one micrometre is typically very small compared to the dimensions of the resonator. Diffraction effects are often therefore not very relevant and the resonator structures need not be confined in all three dimensions but can be set up using discrete mirrors. The most simple arrangement consists of two reflecting mirrors facing each other separated by the distance L (Fig. 4.11 a). More than two mirrors can be arranged in a ring configuration (Fig. 4.11 b) or even in a three-dimensional arrangement.

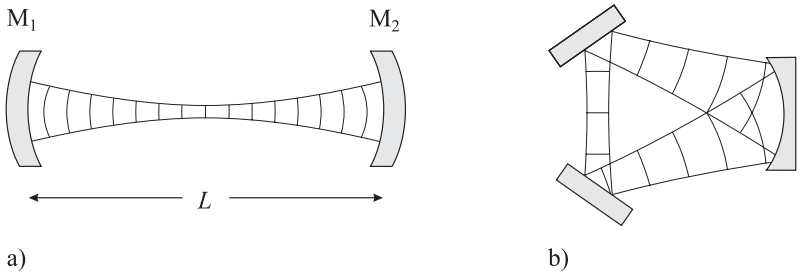


Figure 4.11: a) Linear optical resonator. b) Optical ring resonator.

4.3.1 Reflection and Transmission at the Fabry–Pérot Interferometer

For simplicity, we investigate the case of a linear resonator with plane mirrors often called a plane Fabry–Pérot interferometer (FPI). E_0, E_r, and E_t denote the complex amplitudes of the electromagnetic wave incident on the first mirror M_1 (Fig. 4.12) representing the input coupler of the resonator, the amplitude of the wave reflected by the input coupler, and the amplitude transmitted by the resonator, respectively. We assume that the reflection takes place at the surfaces of the mirrors pointing towards the inner side of the resonator. Hence, the reflected wave will suffer from a phase shift of π at this interface when travelling from the medium with the lower index of refraction to the medium with the higher index of refraction. The mirrors are characterised by their (amplitude) reflection coefficients r_1 and r_2 and by their (amplitude) transmission coefficients t_1 and t_2. The phase factor of the incident wave $\exp i(\omega t - \vec{k} \cdot \vec{r})$ is chosen such that it is unity at the surface of the entrance mirror.

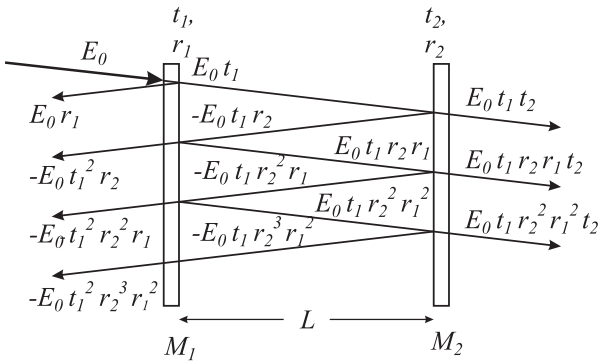


Figure 4.12: Reflection and transmission from an optical resonator (plane Fabry–Pérot interferometer) with the amplitudes of the partial waves.

In the following we calculate the amplitudes of the waves that are transmitted and reflected by the resonator (see Fig. 4.12). To visualise the particular reflected or transmitted partial waves we slightly tilt the impinging beam but do not take into account the phase shift associated with this tilt in the following calculation. The complex amplitude of the transmitted wave is a superposition of all partial amplitudes resulting from the contribution of the directly

transmitted part and the ones coupled out after circulating once, twice, three times, and so on inside the resonator and each time acquiring a phase factor $\exp(-i\vec{k}\cdot\vec{r}) = \exp(-i\omega/c \cdot 2L)$ leading to

$$\begin{aligned} E_T &= E_0 t_1 t_2 e^{-i\omega L/c} + E_0 t_1 t_2 r_1 r_2 e^{-i\omega 3L/c} + E_0 t_1 t_2 r_1^2 r_2^2 e^{-i\omega 5L/c} \cdots \\ &= E_0 t_1 t_2 e^{-i\omega L/c} \left[1 + r_1 r_2 e^{-i\omega 2L/c} + r_1^2 r_2^2 e^{-i\omega 4L/c} \cdots\right]. \end{aligned} \tag{4.79}$$

The geometrical series of the terms inside the brackets of (4.79) can be evaluated by using

$$\sum_{n=0}^{\infty} q^n = \frac{1}{1-q} \qquad \text{and} \qquad q = r_1 r_2 e^{-i\omega 2L/c} \tag{4.80}$$

as $1/[1 - r_1 r_2 \exp(-i\omega 2L/c)]$ and one obtains

$$E_T = E_0 \frac{t_1 t_2 \exp(-i\omega L/c)}{1 - r_1 r_2 \exp(-i\omega 2L/c)}. \tag{4.81}$$

From

$$\begin{aligned} E_T &= E_0 \frac{t_1 t_2 \exp(-i\omega L/c)[1 - r_1 r_2 \exp(i\omega 2L/c)]}{[1 - r_1 r_2 \exp(-i\omega 2L/c)][1 - r_1 r_2 \exp(i\omega 2L/c)]} \\ &= E_0 \frac{t_1 t_2 \left[\exp(-i\omega L/c) - r_1 r_2 \exp(i\omega L/c)\right]}{1 + r_1^2 r_2^2 - 2 r_1 r_2 \cos(2\omega L/c)}. \end{aligned} \tag{4.82}$$

one calculates

$$E_T E_T^* = E_0^2 \frac{t_1^2 t_2^2}{1 + r_1^2 r_2^2 - 2 r_1 r_2 \cos(\omega 2L/c)} \tag{4.83}$$

being proportional to the power transmitted through the FPI representing the so-called Airy function (Fig. 4.13) which depends on the phase shift

$$\Delta\phi = \omega\ 2\ L/c. \tag{4.84}$$

between adjacent partial waves.

If this phase shift corresponds to an integral (q) multiple of 2π, all partial waves interfere constructively in contrast to all other cases where the partial waves interfere more or less destructively. Obviously, the phase difference between the partial waves, and consequently the transmitted power, varies with the angular frequency ω of the incident radiation. The frequency difference that leads to a phase shift of 2π between two consecutive round trips of the radiation inside the resonating cavity is the so-called free spectral range (FSR) of the Fabry–Pérot interferometer

$$\text{FSR} = \frac{c}{2\,L}. \tag{4.85}$$

The linewidth $2\pi\delta\nu$ (FWHM) of the interference structure becomes sharper if more partial waves contribute to the transmitted amplitude, i.e., if the reflectivities r_1 and r_2 of the mirrors become higher. The quantitative relation between $\delta\nu$ and the reflectivities can be derived

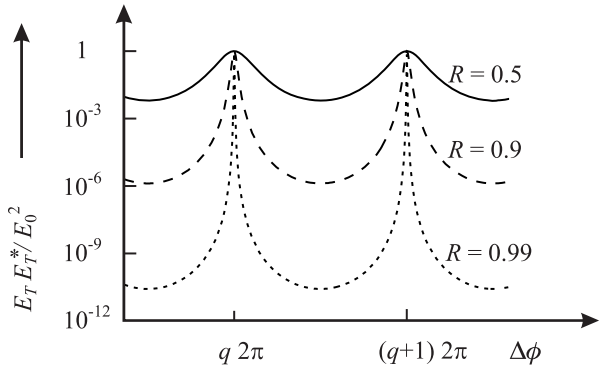


Figure 4.13: The fraction of the incident power transmitted through a Fabry–Pérot interferometer shows an Airy function according to (4.83) for $R = r_1^2 = r_2^2 = 0.5, R = 0.9, R = 0.99$ and $T = t_1^2 = t_2^2 = 1 - R$.

readily for small phase shifts, i.e., for $\Delta\phi = 2\omega L/c \ll 2\pi$. Expanding the cosine function in the denominator of (4.83) leads to

$$E_T E_T^* = E_0^2 \frac{t_1^2 t_2^2}{1 + r_1^2 r_2^2 - 2r_1 r_2 (1 - \frac{4\omega^2 L^2}{2c^2} + \cdots)} \approx E_0^2 \frac{t_1^2 t_2^2}{(1 - r_1 r_2)^2 + 4 r_1 r_2 \frac{\omega^2 L^2}{c^2}}. \tag{4.86}$$

In this approximation the resonance curve becomes Lorentzian (see (2.34)) where the transmitted power is reduced to 50 % at the angular frequency $\omega_{\frac{1}{2}}$. The full width at half maximum (FWHM) $2\pi\delta\nu = 2\omega_{\frac{1}{2}}$ can be calculated from the condition $I_T(\omega = \omega_{1/2}) = 1/2 I_T(\omega = 0)$ as $\omega_{1/2}^2 = c^2(1-R)^2/(L^2 4R)$ as

$$\delta\nu = \frac{2\omega_{\frac{1}{2}}}{2\pi} = \frac{(1 - r_1 r_2)}{\pi\sqrt{r_1 r_2}} \frac{c}{2L}. \tag{4.87}$$

The linewidth normalised to the free spectral range (4.85) is referred to as the finesse F^* of the Fabry–Pérot resonator and using (4.87) it is given as

$$F^* \equiv \frac{\text{FSR}}{\delta\nu} = \frac{\pi\sqrt{r_1 r_2}}{1 - r_1 r_2}. \tag{4.88}$$

As an example, a Fabry–Pérot interferometer of length L = 30 cm and a reflectivity of the mirrors, $R = r_1 r_2$ = 99 %, has a finesse $F^* \approx 314$, a free spectral range FSR = 500 MHz and a linewidth $\delta\nu \approx 1.6$ MHz.

In the photon picture, an optical high-finesse resonator stores the photons for a mean time τ before they eventually escape through the output mirror. As the finesse and the time τ are correlated, the former can be determined from the decay time of the power stored in the resonator by measuring the decreasing power behind the output coupler after abruptly blocking the input power to the resonator (Fig. 4.14). The linewidth of the spectrum of the

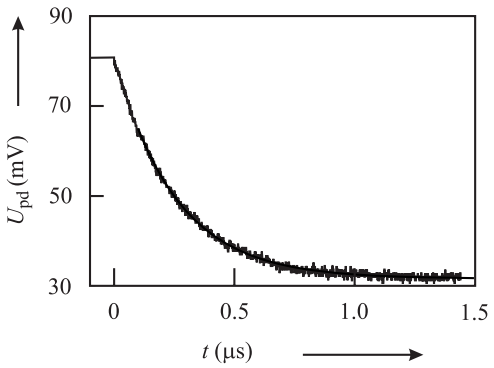


Figure 4.14: Exponential decay of the power stored in an optical resonator as measured by the voltage $U_{\rm pd}$ of a fast photo diode after abruptly switching off the power incident onto the resonator. The experimental data can be fitted by a decay time τ = 252 μs corresponding to a linewidth $\delta\nu$ = 630 kHz and a $Q = \nu/\delta\nu \approx 7.5 \times 10^8$.

exponentially decreasing electromagnetic wave is related to the decay time as $\delta\nu = \delta\omega/2\pi = 1/(2\pi\tau)$ (2.37). Hence, the finesse is

$$F^* = \frac{c/2L}{\delta\nu} = \frac{c}{2L} 2\pi\tau. \tag{4.89}$$

The measurement of the storage time τ is a suitable way of determining the linewidth of an optical high-finesse resonator and of characterising super mirrors with very high reflectivities with the help of (4.89) [66,67]. The relationship between the linewidth of the resonator and the combined reflectivities of the mirrors can be derived from (4.88) and (4.89) for $1 - r_1 r_2 << 1$ as

$$r_1 r_2 = 1 - \frac{L}{c\tau}. \tag{4.90}$$

The determination of the reflectivity of the mirrors has to be performed with an "empty" resonator. Any absorbing medium inside the resonator reduces the number of round trips and the storage time and consequently, the actual linewidth will increase.

In the same way as for the transmitted amplitude, the amplitude of the electromagnetic wave reflected by the optical resonator is calculated (see Fig. 4.12) as

$$\begin{aligned} E_R &= E_0 r_1 - E_0 t_1 r_2 t_1 e^{-i\omega 2L/c} - E_0 t_1^2 r_1 r_2^2 e^{-i\omega 4L/c} - E_0 t_1^2 r_1^2 r_2^3 e^{-i\omega 6L/c} - \cdots \\ &= E_0 \left[r_1 - t_1^2 r_2 e^{-i\omega 2L/c} - t_1^2 r_1 r_2^2 e^{-i\omega 4L/c} - t_1^2 r_1^2 r_2^3 e^{-i\omega 6L/c} - \cdots \right] \\ &= E_0 r_1 - E_0 t_1^2 r_2 e^{-i\omega 2L/c} \left[1 + r_1 r_2 e^{-i\omega 2L/c} + r_1^2 r_2^2 e^{-i\omega 4L/c} + \cdots \right] \end{aligned} \tag{4.91}$$

Evaluating the geometrical series in the square brackets of the last line of (4.91) one arrives at

$$E_R = E_0 r_1 - E_0 \frac{t_1^2 r_2 \exp(-i\omega 2L/c)}{1 - r_1 r_2 \exp(-i\omega 2L/c)}. \tag{4.92}$$

The (amplitude) reflection coefficient of the Fabry–Pérot interferometer

$$r_{\mathrm{FP}}(\omega) \equiv \frac{E_R}{E_0} = \frac{r_1 - r_2(r_1^2 + t_1^2)\exp(-i\omega 2L/c)}{1 - r_1 r_2 \exp(-i\omega 2L/c)} \tag{4.93}$$

does not depend on the (amplitude) transmission coefficient t_2 of the second mirror used as output coupler. From (4.92) and from Fig. 4.11 one finds that the amplitude of the wave reflected from the resonator is composed of two contributions. The first one results from the part of the incident wave that is directly reflected by the input mirror. Its negative sign results from the phase shift of π radians that occurs at the optical interface when the wave comes from and is reflected into the material with the higher index of refraction. The second term is the part of the wave circulating in the resonator that is transmitted through the same mirror. Exactly at resonance ($\omega 2L/c = 2\pi$) the reflected wave exhibits a minimum (see (4.93)). Hence, these two contributions are out of phase by π which is provided for by the minus sign of the second term [2] in (4.93).

The minimum goes to zero for a loss-less symmetrical cavity with $r_1 = r_2 = r$ and $t_1 = t_2 = t$ with $r^2 + t^2 = R + T = 1$. For the symmetrical loss-less cavity the complex reflection coefficient (4.93) becomes

$$r_{\mathrm{FP}}(\omega) = r\frac{1 - \exp(-i\omega 2L/c)}{1 - r^2 \exp(-i\omega 2L/c)}. \tag{4.94}$$

The separation of the real and imaginary part of (4.93)

$$\begin{aligned} \Re\mathrm{e}\, r_{\mathrm{FP}} &= \frac{-r_1 - t_1^2 r_1 r_2^2 + 2(t_1^2 + 2r_1^2)\cos(\omega 2L/c)}{1 - 2r_1 r_2 \cos(2\omega L/c) + r_1^2 r_2^2} \\ \Im\mathrm{m}\, r_{\mathrm{FP}} &= -\frac{r_2 t_1^2 \sin(2\omega L/c)}{1 - 2r_1 r_2 \cos(2\omega L/c) + r_1^2 r_2^2} \end{aligned} \tag{4.95}$$

can be used to derive the power reflection factor

$$r_{\mathrm{FP}} r_{\mathrm{FP}}^* = (\Re\mathrm{e}\, r_{\mathrm{FP}})^2 + (\Im\mathrm{m}\, r_{\mathrm{FP}})^2 = \frac{r_1^2 + r_2^2 - r_1^2 r_2^2 \cos(\omega 2L/c)}{1 + r_1^2 r_2^2 - 2r_1 r_2 \cos(\omega 2L/c)} \tag{4.96}$$

and the relation for the phase ϕ_R of the reflected wave as

$$\tan\phi_R = \frac{\Im\mathrm{m}\, r_{\mathrm{FP}}}{\Re\mathrm{e}\, r_{\mathrm{FP}}} = \frac{r_2 t_1^2 \sin(2\omega L/c)}{r_1\left[1 + r_2^2(r_1^2 + t_1^2)\right] - r_2(2r_1^2 + t_1^2)\cos(2\omega L/c)}. \tag{4.97}$$

In the last equation, we have made use of $\cos\delta = 1 - 2\sin^2(\delta/2)$. In the vicinity of a resonance frequency $\omega_q/2\pi = qc/(2L)$ of the Fabry–Pérot interferometer, the reflected and the transmitted wave change by π (Fig. 4.16). The slope of the phase change increases with increasing reflectivity of the mirrors.

The complex reflection coefficient (4.94) in the vicinity of the q^{th} resonance frequency ω_q can be simplified by introducing the detuning $\Delta\omega \equiv \omega - \omega_q$. For a detuning $\Delta\omega$ smaller

[2] The $+$ and $-$ signs in the nominator of (4.93) are sometimes interchanged in the literature.

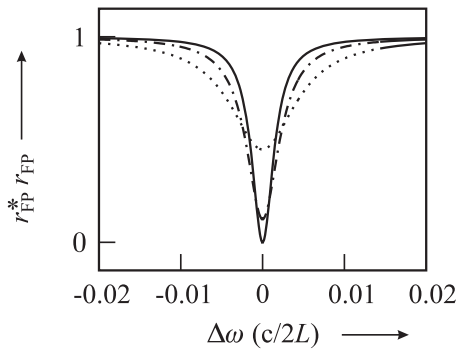


Figure 4.15: Power reflection factor of the Fabry–Pérot interferometer versus frequency detuning according to (4.96) for $R_1 = r_1^2 = 0.99$ and $R_2 = r_2^2 = 0.99$ (full line), $R_2 = r_2^2 = 0.98$ (dashed-dotted line) and $R_2 = r_2^2 = 0.95$ (dots).

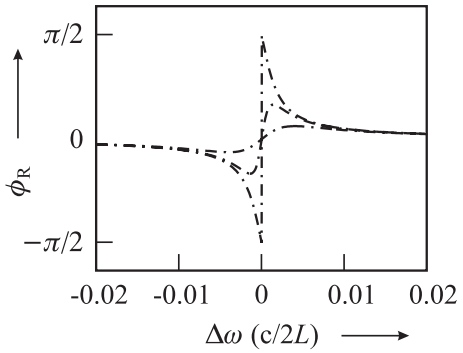


Figure 4.16: Phase shift of the wave reflected by the Fabry–Pérot interferometer versus angular frequency detuning according to (4.97) for $R_1 = r_1^2 = 0.99$ and $R_2 = r_2^2 = 0.99$ (full line), $R_2 = r_2^2 = 0.98$ (dashed-dotted line) and $R_2 = r_2^2 = 0.95$ (dots).

than the free spectral range $c/(2L)$ we approximate the phase factor $\exp(-i\omega 2L/c)$ by $1 - i(\Delta\omega)2L/c + \cdots$ and obtain

$$r_{\mathrm{FP}}(\Delta\omega) \approx r \frac{i\Delta\omega 2L/c}{1 - r^2 + r^2 i\Delta\omega 2L/c}. \tag{4.98}$$

For a cavity with high finesse ($R - 1 \ll 1$) we approximate $(1 - r^2)/r^2 \approx (1 - r^2)/r$ and use (4.87) together with $\Gamma \equiv 2\pi\delta\nu$ to derive

$$r_{\mathrm{FP}}(\Delta\omega) \approx \frac{1}{r} \frac{i\Delta\omega}{\Gamma/2 + i\Delta\omega} \tag{4.99}$$

from which the amplitude and phase variation in the vicinity of the resonance is readily derived.

4.3.2 Radial Modes

In the Fabry–Pérot resonator the maxima of the Airy function (Fig. 4.13) occur at the eigenfrequencies

$$\omega_q = q\, 2\pi \frac{c}{2L} \qquad \text{with } q \in \mathcal{N}. \tag{4.100}$$

The electromagnetic field components with ω_q are called the longitudinal or axial modes of the linear optical resonator. The corresponding field components are plane waves along the optical axis of the resonator. As a plane wave would extent to infinity in the transverse direction, the energy contained in such a wave would be unlimited. A more realistic treatment of the electromagnetic field in the resonator has to include an amplitude dependence from the transverse coordinates and a transverse confinement of the wave inside the resonator. On

the other hand, any transverse confinement will result in diffraction effects which in turn will modify the z dependence of the $\exp(\pm ikz)$ factor in the plane wave.

To determine the transverse beam profile we follow [68, 69] and make use of the wave equation (4.27) for the electric field $E(x, y, z, t)$ of a laser beam. Factorising $E(x, y, z, t)$ into a spatial and a temporal part $E(x, y, z, t) = E(x, y, z) \exp(i\omega t)$ and separating the time dependence $\exp(i\omega t)$ leads to a differential equation for the spatial part of the electromagnetic wave. For the latter one we use the ansatz

$$E(x, y, z) = \widetilde{u}(x, y, z)e^{-ikz} \tag{4.101}$$

where $\widetilde{u}(x, y, z)$ represents a complex scalar wave amplitude describing the transverse profile of the laser beam. For simplicity, we use the scalar form

$$\left[\nabla^2 + k^2\right] E(x, y, z) = 0. \tag{4.102}$$

Substituting (4.101) into (4.102) yields a differential equation for the complex scalar amplitude of the wave

$$\frac{\partial^2 \widetilde{u}}{\partial x^2} + \frac{\partial^2 \widetilde{u}}{\partial y^2} + \frac{\partial^2 \widetilde{u}}{\partial z^2} - 2ik\frac{\partial \widetilde{u}}{\partial z} = 0. \tag{4.103}$$

If it is justified to neglect the second derivative of $\widetilde{u}$ with respect to z as compared to the first-order derivative with respect to z and compared to the second derivatives with respect to x and y (paraxial approximation) one obtains

$$\frac{\partial^2 \widetilde{u}}{\partial x^2} + \frac{\partial^2 \widetilde{u}}{\partial y^2} - 2ik\frac{\partial \widetilde{u}}{\partial z} = 0. \tag{4.104}$$

This is the paraxial wave equation which can be solved by the ansatz

$$\widetilde{u}(x, y, z) = A(z)\ \exp\left(-ik\frac{x^2 + y^2}{2\widetilde{q}(z)}\right). \tag{4.105}$$

leading to the differential equation

$$\left[\left(\frac{k}{\widetilde{q}}\right)^2 \left\{\frac{d\widetilde{q}}{dz} - 1\right\}(x^2 + y^2) - \frac{2ik}{\widetilde{q}}\left\{\frac{\widetilde{q}}{A}\frac{dA}{dz} + 1\right\}\right] A(z) = 0. \tag{4.106}$$

(4.106) can be solved for all x and y only if both terms in the braces are identically zero

$$\frac{d\widetilde{q}}{dz} = 1 \qquad \text{and} \qquad \frac{dA(z)}{dz} = -\frac{A(z)}{\widetilde{q}(z)} \tag{4.107}$$

After integration we obtain

$$\widetilde{q}(z) = \widetilde{q}_0 + z \qquad \text{and} \qquad \frac{A(z)}{A_0} = \frac{\widetilde{q}_0}{\widetilde{q}}(z). \tag{4.108}$$

For simplicity, we have chosen the integration constant $z_0 = 0$ in (4.108). The first of these equations describes the evolution of the complex beam parameter $\widetilde{q}$ from the value $\widetilde{q}_0$ in a plane through z_0 to the value $\widetilde{q}(z)$ in a plane through z.

To identify the meaning of the beam parameter $\widetilde{q}$ we write it as the sum of a real and an imaginary part as

$$\frac{1}{\widetilde{q}(z)} = \frac{1}{R(z)} - i\frac{\lambda}{\pi w^2(z)} \qquad \text{with} \qquad k = \frac{2\pi}{\lambda} \tag{4.109}$$

and substitute it into (4.105). Thus

$$\widetilde{u}(x,y,z) = A_0 \frac{\widetilde{q}_0}{\widetilde{q}(z)} \exp\left[-ik\frac{x^2+y^2}{2R(z)} - \frac{x^2+y^2}{w^2(z)}\right] \tag{4.110}$$

holds. The distribution $\widetilde{u}(x,y,z)$ has a purely real part

$$\exp\left[-\frac{x^2+y^2}{w^2(z)}\right] \equiv \exp\left[-\frac{r^2}{w^2(z)}\right] \tag{4.111}$$

which yields a two-dimensional Gaussian amplitude distribution $\widetilde{u}$ along the transverse coordinates x and y. $w(z)$ is called the beam radius at the location z describing the transverse distance where the amplitude of the beam is reduced to $1/e$ of the maximum.

The complex part

$$\exp\left[ik\frac{x^2+y^2}{2q_{\mathrm{re}}}\right] \equiv \exp\left[ik\frac{x^2+y^2}{2R(z)}\right] \tag{4.112}$$

represents a phase factor of a spherical wave with $R(z)$ being the real radius of curvature of the wave front intersecting the axis the point z. The representation of the beam parameter $q(z)$ by a real and an imaginary part contains all the physics of a Gaussian wave. Starting at the location $z = 0$ where the wave front is a plane wave, i.e., where the radius of curvature $R(z=0) = \infty$ leads to

$$\begin{aligned} \frac{1}{\widetilde{q}(z=0)} &\equiv \frac{1}{\widetilde{q}_0} = -i\frac{1}{q_{\mathrm{im}}(z=0)} = -i\frac{\lambda}{\pi w_0^2} \qquad \text{or} \\ \widetilde{q}_0 &= i\frac{\pi w_0^2}{\lambda}. \end{aligned} \tag{4.113}$$

w_0 is called the waist of the Gaussian wave. $\widetilde{q}(z)$ can be derived by using of (4.113) as follows

$$\widetilde{q}(z) = \widetilde{q}_0 + z = i\frac{\pi w_0^2}{\lambda} + z \equiv iz_R + z \tag{4.114}$$

where $z_R = \pi w_0^2/\lambda$ is known as the Rayleigh range. By inserting (4.114) into (4.109) and by equating the imaginary parts we obtain

$$w^2(z) = w_0^2\left[1 + \left(\frac{\lambda z}{\pi w_0^2}\right)^2\right] = w_0^2\left[1 + \left(\frac{z}{z_R}\right)^2\right]. \tag{4.115}$$

After equating the real parts it follows that

$$R(z) = z\left[1 + \left(\frac{\pi w_0^2}{\lambda z}\right)^2\right] = z\left[1 + \left(\frac{z_R}{z}\right)^2\right]. \tag{4.116}$$

A wave with a Gaussian profile will remain a Gaussian with the beam diameter $2w(z)$ (Fig. 4.17). The radius of curvature and the beam radius evolve as given by (4.115) and by (4.116), respectively.

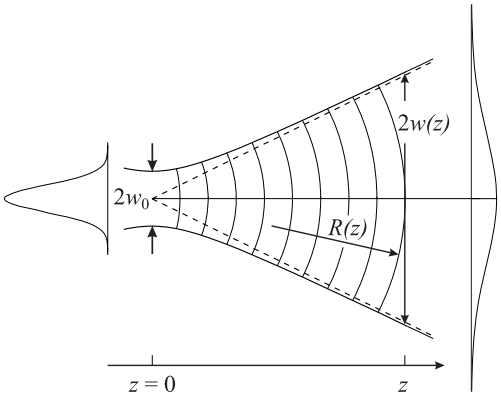


Figure 4.17: Evolution of a Gaussian beam.

Finally, from (4.110), (4.114), (4.115) and (4.116) one finds

$$\widetilde{u}(x,y,z) = A_0 \frac{i}{i + \frac{z}{z_R}} \exp\left[-ik\frac{x^2+y^2}{2z(1+\frac{z_R^2}{z^2})} - \frac{x^2+y^2}{w_0^2(1+\frac{z_R^2}{z^2})}\right]. \tag{4.117}$$

The evolution of the Gaussian wave in the way depicted in Fig. 4.17 is a result of the diffraction to which any wave is subjected. Even though at the location of the waist w_0 the wave fronts are described by a plane wave the finite transverse extension leads to a transverse expansion. This is a result of the diffraction described by Huyghens' principle stating that in an isotropic medium any point of a wave front by itself is the origin of a spherical wavelet. The characteristic distance where the diffractive expansion becomes more and more noticeable is determined by the Rayleigh range $\pi w_0^2/\lambda$. The principle of the diffractive expansion of a wave front whose transverse extension is restricted to a small area is a general one. Even the so-called diffraction-free beams [70] obey this principle, however, in this case the apparent diffractional expansion of particular small transverse structures (on a wide pedestal) are considerably reduced.

The beam radius $w(z)$ in Fig. 4.17 evolves like a hyperbola and for large distances $w \gg w_0$ holds and the second term in square brackets of (4.115) becomes the dominant one

$$w(z) \approx \frac{\lambda z}{\pi w_0}. \tag{4.118}$$

The angle between the asymptotes and the axis is $\theta \approx \tan\theta \approx w/z$ and hence

$$\theta = \frac{\lambda}{\pi w_0}. \tag{4.119}$$

The smaller the waist w_0 the larger is the expansion of the beam as a consequence of the diffraction.

The ansatz (4.105) representing the fundamental Gaussian mode is by no means the only possible solution of the paraxial wave equation. Suppose we had rather chosen an ansatz of the form

$$\widetilde{u}(x,y,z) = g\left(\frac{x}{w}\right) h\left(\frac{y}{w}\right) \exp\left[-ik\frac{x^2+y^2}{2\widetilde{q}(z)}\right] \tag{4.120}$$

then $\widetilde{u}(x,y,z)$ would lead to higher transverse modes represented by the product of two Hermite polynomials $H_m(\sqrt{2}\frac{x}{w})$, $H_n(\sqrt{2}\frac{y}{w})$ and a Gaussian function. The four lowest Hermite polynomials are

$$\begin{aligned} H_0(\sqrt{2}\frac{x}{w}) &= 1, \\ H_1(\sqrt{2}\frac{x}{w}) &= 2(\sqrt{2}\frac{x}{w}), \\ H_2(\sqrt{2}\frac{x}{w}) &= 4(\sqrt{2}\frac{x}{w})^2 - 2, \\ H_3(\sqrt{2}\frac{x}{w}) &= 8(\sqrt{2}\frac{x}{w})^3 - 12(\sqrt{2}\frac{x}{w}). \end{aligned} \tag{4.121}$$

$H_n(\sqrt{2}x/w)$ has $n-1$ zeros along the x direction leading to $n-1$ dark regions in the transverse power profile. Consequently, the higher order transverse modes are characterised by the number of zeros m and n (in Cartesian coordinates). They are called Transverse Electro-Magnetic waves of order m and n, i.e., TEM_{mn}.

Hence, the TEM_{00} mode has no zeros along the x and y direction and represents the Gaussian profile. Owing to the contributions of the Hermite polynomials the area taken by the higher transverse modes becomes larger with increasing m and n. This behaviour can be utilised to suppress higher transverse modes by putting a diaphragm into the resonator with a diameter large enough to allow the fundamental mode to pass but small enough to cut off considerable power from the outer rim of the higher-order modes. If the resonator has a truly cylindrical symmetry it is more appropriate to use polar coordinates r and ϕ rather than Cartesian coordinates. In this case, the modes are described by a product of Laguerre and Gaussian functions.

Both, the Hermite-Gauss polynomials as well as the Laguerre–Gauss polynomials form a complete system of eigenfunctions. Hence, the modes can be represented in both systems and a mode described in one system of eigenfunctions can be represented in general as a superposition of modes in the other one [71]. Fig. 4.19 depicts a Laguerre–Gauss mode (in polar coordinates) made up by three Hermite–Gauss modes. The Laguerre–Gauss modes are characterised similarly by two independent integers denoting the zeros of the field along the radial (r) coordinate and along the azimuthal (ϕ) coordinate. The three Hermite–Gauss modes of Fig. 4.19 necessary to make up the TEM_{31} with three zeros along the ϕ coordinate and one

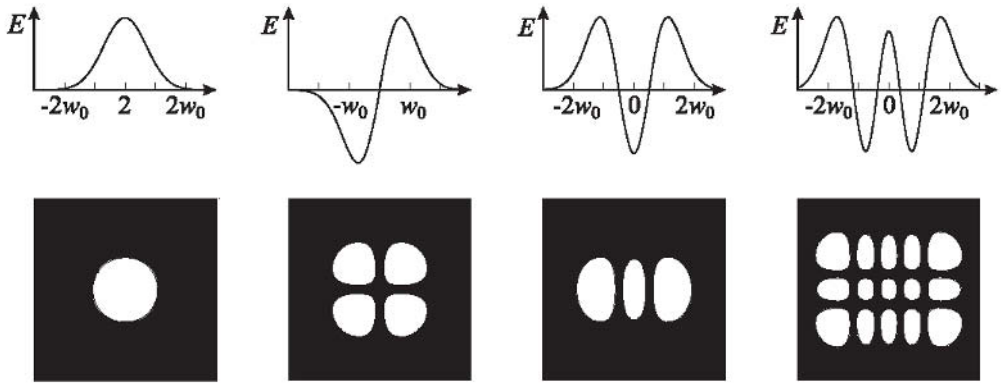


Figure 4.18: Distribution of the electric field and mode pattern.

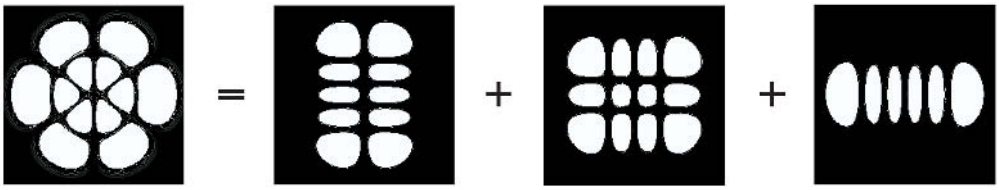


Figure 4.19: Representation of a TEM_{31} Laguerre–Gauss mode by three Hermite–Gauss modes TEM_{14}, TEM_{32} and TEM_{50}.

zero in the radial direction r exhibit, 1 and 4, 3 and 2, 5 and 0 zeros along the x and y coordinates, respectively.

Practical optical resonators use curved mirrors rather than planar ones. In fact, resonators with flat mirrors have in general larger diffraction losses and hence are not ideally suited for high-finesse resonators. In general they are not even truly Gaussian in profile. For a linear resonator comprising two concave mirrors with the radii of curvature R_1 and R_2 separated by the distance L, the eigenfrequencies of the modes also depend on the radii of curvature. The calculation of the eigenfrequencies is somewhat lengthy and can be found, e.g., in [68] to lead to

$$\nu_{mnq} = \frac{c}{2L}\left[q + \frac{1}{\pi}(m+n+1)\arccos\sqrt{(1-\frac{L}{R_1})(1-\frac{L}{R_2})}\right]. \tag{4.122}$$

For $m = n = 0$ the frequencies of (4.122) correspond to the ν_q of the axial modes defined only by the optical length of the resonator.

A particular simple but rather important case is the one where the radii of curvature of both mirrors equal the distance between both mirrors, i.e., $L = R_1 = R_2$. As a result, the square-root term in (4.122) disappears and the *arccos* term becomes $\pi/2$ and the eigenfrequencies depend on ν, m and n as follows

$$\nu_{mnq} = \frac{c}{2L}\left[q + \frac{1}{2}(m+n+1)\right] \tag{4.123}$$

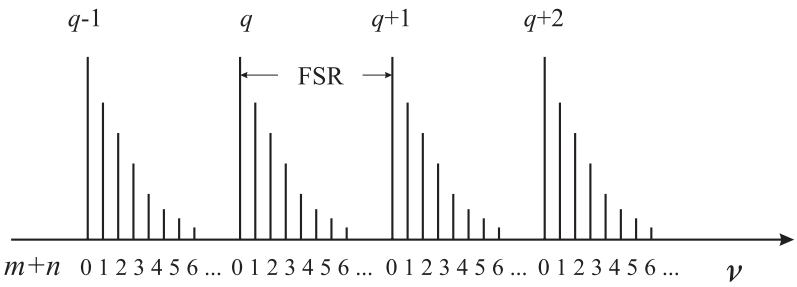


Figure 4.20: Schematic mode spectrum of a Fabry–Pérot interferometer.

For an even number of $m + n$ the eigenfrequencies of the transverse modes coincide with the frequencies of the fundamental modes, i.e., they are degenerate. The eigenfrequencies of the transversal modes with odd $m + n$, however, are shifted by $c/(4L)$ with respect to the axial modes separated by $c/(2L)$. Consequently, in the confocal Fabry–Pérot resonator modes with frequencies can be excited that differ by

$$\delta\nu = \frac{c}{4\,L}. \tag{4.124}$$

A linear optical resonator can be comprised of two mirrors with different radii of curvature R_1 and R_2. If the length $L \neq R_1, R_2$ the degeneracy of the eigenfrequencies of the transverse modes is lifted (Fig. 4.20).

An electromagnetic wave impinging onto one of the mirrors of the optical resonator can only excite those modes whose frequencies coincide with that of the wave. In a confocal resonator an infinite number of axial modes have the same eigenfrequencies but have different distribution of the transverse field distribution on the surface of the mirror. Consequently, a wave of a given field distribution will predominantly excite that particular mode inside the resonator whose field distribution coincides with the one of the impinging wave. Mathematically speaking, the incident wave will be decomposed into a linear combination of the modes, i.e., of the eigenfunctions representing the field inside the resonator. The coupling of the incident wave to the particular modes is determined by the coupling coefficients which are determined by the overlap integrals between the modes in the resonator and the incident wave. If only one mode is to be excited, the field distributions of the incident wave and of the resonator mode have to coincide exactly at the surface of the resonator mirror.

For optical frequency standards, Fabry–Pérot interferometers (FPIs) are particularly useful for analysing the frequency spectrum of laser radiation or to pre-stabilise the frequency of such a laser to a suitable eigenfrequeny of the FPI. For the first purpose, the FPI can be used as a tuneable filter since a particular eigenfrequency of the comb of equidistant resonances can be adjusted to match the desired frequency of the laser. Applying a variable high voltage to a piezoelectric element between the spacer and a resonator mirror allows one to vary the length of the resonator by a few wavelengths. Here, a confocal FPI is often preferred since the frequencies of all modes with odd $m + n$ are grouped together at the frequencies of the axial modes and all modes with even $m + n$ are located at frequencies shifted by half a free spectral range. Consequently, accurate mode matching is not necessary. For achieving very narrow

linewidths of the FPI or for use of a frequency discriminator, however, the confocal FPI is not well suited, since in general it is not possible, to realise the condition $L = R$ with sufficient accuracy.

4.3.3 Microsphere Resonators

In the optical domain, whispering gallery modes in dielectric microspheres of fused silica can exhibit exceedingly high Q values and, hence, represent an alternative to resonators of the Fabry–Pérot type. They also have the potential to be used in ultra-compact optical frequency standards. Microspheres are readily prepared by a fusion technology from a high-quality fused silica rod in an oxygen-hydrogen torch. The surface tension of the melting material leads to a spheroid with a diameter $D = 2R$ of a few tens to a few hundreds of micrometres. The quality factor Q of a whispering gallery mode is determined by the radiative loss as a result of the curvature, the scattering on residual surface inhomogeneities, surface contaminants and intrinsic material losses [72, 73]. From the latter, a principal limit of $Q = 9 \times 10^9$ has been derived at $\lambda = 633\,\text{nm}$ and $Q = 1.5 \times 10^{11}$ at $\lambda = 1.55\,\mu\text{m}$. Quality factors $Q = 8 \times 10^9$ have been measured in three resonators with diameters between 0.6 mm and 0.9 mm [73] at 633 nm immediately after preparation of the microspheres. The large Q, however, was shown to deteriorate rapidly by adsorption of atmospheric water and thus it might be necessary to keep the sphere in a hermetically sealed chamber [74].

The TE and TM eigenmodes of an electromagnetic field in a dielectric sphere with a refractive index n are characterised by three integers l, m, q. The number of field maxima along the radius of the sphere is given by $q \geq 1$, the number of field maxima in the equatorial plane equals m and l is the mode number [75, 76]. The latter can be thought of as roughly the number of wavelengths λ on the circumference of the sphere ($l \approx 2\pi Rn/\lambda$). For $l \gg q$ one refers to the modes as whispering gallery modes. As a result of the production process the final shape of the microspheres differs from a perfect sphere and is better approximated by an ellipsoid. Denoting the axes of the ellipsoid by a and b, typically $10^{-2} < \epsilon^2 < 10^{-1}$ is achieved [74] where the eccentricity ϵ of the ellipsoid is given by $\epsilon^2 \equiv 1 - b^2/a^2$.

The eigenfrequencies of the eccentric microsphere are given by [74]

$$\gamma_{qlm}^{E,H} = \Delta_0 \left[l + \frac{1}{2} - A_q \sqrt[3]{(l+1/2)/2} - \Delta_{E,H} \pm \epsilon^2 (l - |m|)/2 \right] \tag{4.125}$$

where $A_q = 2.338, 4.088, 5.521, 6.787, \cdots$ are the qth zeros of the Airy function. $\Delta_{E,H}$ takes into account that the two waves of orthogonal polarisations are differently confined near the surface and, hence, experience a different index of refraction. The positive and negative sign has to be chosen for an oblate and a stretched spheroid, respectively. The free spectral range of the microresonator is $\Delta_0 = c/(\pi Dn)$ which is about 180 GHz for a $D = 370\,\mu\text{m}$ diameter sphere and $n = 1.45$ at $\lambda = 852\,\text{nm}$.

Tuning of the eigenfrequencies can be achieved by temperature variation or by strain. The resonance frequency is shifted by $\Delta\nu/\nu = -\Delta a/a - \Delta n/n$ where a is the radius and n is the index of refraction of the sphere. The temperature dependence reduces the mode frequency by a few gigahertz per degree. Larger tuning over several hundred gigahertz has been achieved in the near infrared [77] by applying a compressive force near the "polar" regions in a "micro

vice" or by stretching a microsphere with two attached stems [78]. The dominant effect on the detuning is due to the associated equatorial expansion and, to a lesser degree, to the change of the index of refraction. Relative variations in the circumference of about 10^{-3} have been achieved without noticeably degrading the quality factor of $Q \approx 10^9$.

The high Q of whispering gallery modes indicates that these modes have extremely low losses and hence are very weakly coupled to free space which means that, on the other hand, they can hardly be excited by free-space beams. Hence, light has to be coupled into the microsphere by other methods [79], e.g., by placing side-polished optical fibres in close contact with the microspheres or via a coupling prism (Fig. 4.21). With the former method a more than 99.8 % optical power transfer to the sphere has been achieved [80]. In the latter method

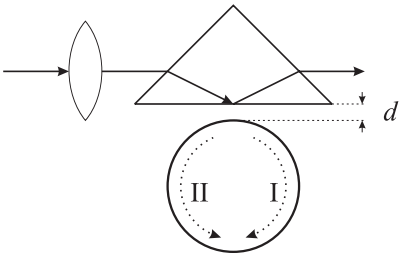


Figure 4.21: Coupling of laser radiation to a whispering-gallery microresonator by a prism using frustrated internal reflection.

a laser beam is focused by a lens onto the inner surface of a coupling prism. If the microsphere is placed at a distance $d \approx \lambda/(2\pi)$ from the surface of the microsphere despite the frustrated internal total reflection light from the near-field evanescent wave can be coupled into the whispering gallery mode (I). Resonances excited in the microsphere resonator show up as dips in the reflected light leaving the coupling prism. The coupling to the resonator is varied by adjusting the gap width d. Rayleigh backscattering inside the microsphere, together with the high Q, in general leads to the build-up of a backward reflected wave (II) that can be used for frequency locking of a diode laser [74]. A diode laser whose frequency is stabilised to a microsphere might lead to a sub-kHz-linewidth laser [74] for applications in very compact frequency standards.

4.4 Stability of Resonators

The eigenfrequencies of the resonators treated in this chapter depend on the macroscopic dimensions of the resonators (see (4.100)). Consequently, there is a linear variation of the frequency associated with any variation of the length relevant to the frequency

$$\frac{d\nu}{dL} = -\frac{qc}{2L^2} = -\frac{\nu}{L} \tag{4.126}$$

or

$$\frac{\Delta\nu}{\nu} = -\frac{\Delta L}{L}. \tag{4.127}$$

Here, we have replaced the differential quotient by the ratio of the differences. One of the most significant environmental parameters that affect the stability of the mechanical dimensions of

macroscopic resonators is the temperature. Temperature fluctuations ΔT around the working temperature T_0 of the resonator result in a variation of the length L_0 to $L(T)$ which can be described as a Taylor series with linear (α), quadratic (β), cubic (γ), ... coefficients of thermal expansion (CTE)

$$L(T) = L(T_0) + L(T_0)\,\alpha\,\Delta T + L(T_0)\,\beta\,(\Delta T)^2 + L(T_0)\,\gamma\,(\Delta T)^3 + \cdots . \quad (4.128)$$

In most cases it suffices to refer to the linear coefficient of thermal expansion α and consequently the fractional frequency shift of a particular mode is given by

$$\frac{\Delta\nu}{\nu} \approx -\alpha\Delta T. \quad (4.129)$$

High frequency stability thus asks for minimisation of the temperature fluctuations and the employment of materials with a low thermal expansion coefficient (see Table 4.4). At room temperature the linear coefficient of thermal expansion of copper is $\alpha_{\mathrm{Cu}} \approx 1.65 \times 10^{-5}\,\mathrm{K}^{-1}$ and that of the temperature compensated nickel iron steel Invar [3] is about an order of magnitude lower and comparable to that of fused silica (see Table 4.4). Much lower values of the thermal expansion are provided by a mixture of glass and ceramic materials such as Zerodur or temperature compensated glasses called Ultra-Low-Expansion glass (Corning ULE 7971); comprising about 80 % SiO_2 and 20 % TiO_2). These materials are specifically tailored by

Table 4.4: Mechanical properties of materials suitable for macroscopic resonators. α: linear coefficient of thermal expansion; E: Young's modulus of elasticity; ρ: density; c_p: specific heat; λ: heat conductivity.

Symbol	Units	Copper	Invar	Fused silica	ULE	Zerodur M	Sapphire (4.2 K)
α	10^{-8}/K	1650	150	55	0.3	< 1	5×10^{-4}
E	10^9 N/m^2	130	145	73	68	89	435
ρ	10^3 kg/m^3	8.92	8.13	2.2	2.21	2.52	4.0
c_p	J/(kg K)	385	500	703	0.77	0.81	5.9×10^{-6}
λ	W/(m K)	400	10.5	1.38	1.31	1.63	280

a suitable composition to show a zero crossing or an extremum of the coefficient of thermal expansion $\alpha(T)$ at a given temperature, e.g., near 25° C. Very low coefficients of thermal expansion can be achieved at this temperature.

In contrast to crystalline materials, glass or glass-ceramic materials suffer from long-term length variations. Such a behaviour is expected, e.g., from the thermal diffusion of the molecules in a glass leading to the formation of crystalline domains with the associated reduction in volume. The length variation due to this ageing effect or "creep" in general slows down with an exponential decrease in the length [81–83]. As an example consider a Fabry–Pérot interferometer made of Zerodur M where a particular eigenfrequency was monitored over more

[3] Super Invar (31% Ni 5% Co 64% Fe) has an even smaller coefficient of linear thermal expansion of $\alpha \approx -19 \times 10^{-8}$/ K at 20° C.

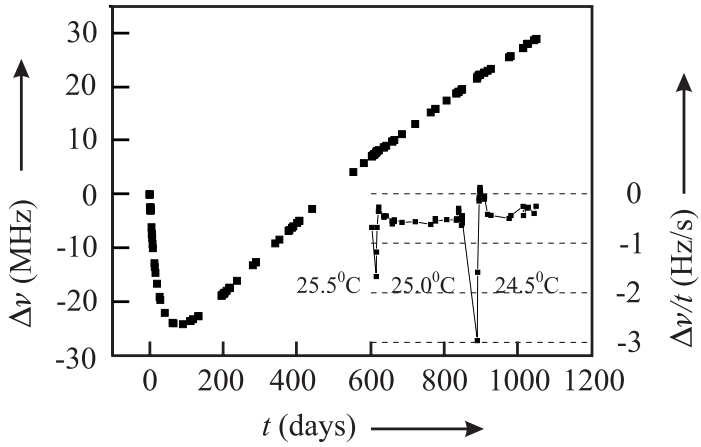


Figure 4.22: Temporal length variation of a Fabry–Pérot interferometer made of Zerodur M, as measured by the frequency of a suitable resonance, shows a drift due to ageing. The inset shows the effect when the temperature was reduced by 0.5 K.

than three years (Fig. 4.22). After about one hundred days a monotonous drift of the frequency was observed that gradually slowed down, which corresponds to a length variation. The earlier behaviour was probably dominated by a heat treatment after the Fabry–Pérot resonator had been put into vacuum. A variation in the temperature (inset of Fig. 4.22) by twice 0.5 K did not show a big effect. For $\alpha \neq 0$ one would expect a discontinuous change in the eigenfrequency at the respective instants. The measured drift of about $\Delta\nu \approx 0.4$ Hz/s at 456 THz translates into a relative length variation $\Delta L/L = -\Delta\nu/\nu < 10^{-15}$/s which corresponds to a relative length change of about 8×10^{-11} per day. Marmet *et al.* [84] reported on an optical resonator of ULE with a differential coefficient of thermal expansion of $2 \times 10^{-9}/\mathrm{K}^2$. With a temperature stability of 50 µK achieved with a two-stage temperature stabilisation the residual relative length variation was no longer limited by temperature fluctuations but by the creep of about 1×10^{-11} per day.

The thermal expansion of a solid crystalline material, e.g., quartz can be explained by the anharmonicity of the lattice oscillations. For purely harmonic oscillations of the atoms around their equilibrium, the mean length of a crystal would not change with temperature. The thermal expansion is therefore intimately connected to the modes of the lattice oscillations in a real crystal. Since all the modes contribute to the specific heat of the crystal the temperature dependence of the specific heat determines the thermal expansion of the crystal. In general according to Debye's model, the specific heat and consequently the linear thermal expansion coefficient for low temperatures decreases with the third power of the temperature. Ultra-stable cryogenic optical resonators with sapphire spacers have been set up [85, 86] and operated at 1.9 K. The low coefficient of thermal expansion of sapphire (see Table 4.4) leads to a reduced sensitivity to ambient temperature changes. It has been pointed out [86] that at cryogenic temperatures the thermal diffusivity is also strongly enhanced in comparison to room temperature and leads to a greater precision in active temperature stabilisation. The ther-

mal diffusivity, i.e., $\lambda/(\rho c_p)$ is a measure of how fast the material can react to a variation in the heat transport. When selecting a suitable material for a resonator, either for the microwave or the optical domain, attention has to be given also to Young's modulus of elasticity E of the material (see Table 4.4). The higher E the smaller the deformation and, hence, the smaller the variation of the eigenfrequencies due to tilt or acceleration.

5 Atomic and Molecular Frequency References

The resonance frequencies of the resonators described in Chapter 4 depend on dimensional quantities. These in turn are influenced in a delicate way by environmental parameters as, e.g., temperature, air pressure, vibrations, gravity. Consequently, the stability of a resonance frequency of a *macroscopic oscillator* can be kept constant only to that extent to which these parameters can be precisely controlled. If, on the other hand, electromagnetic transitions in free atoms, ions, or molecules, i.e., *microscopic oscillators*, are utilised to stabilise the frequency of an oscillator, the influence of the external parameters on its frequency is usually very small. The use of these quantum oscillators relies on the fact that the emission and absorption of electromagnetic radiation by the atomic particles occurs at well defined frequencies which are characteristic for each species. According to Bohr such a frequency ν occurs when the absorbing particle undergoes a transition between two discrete states with energies E_1 and E_2. Energy conservation immediately leads to the well known relationship between the energy of the photon and the energy difference between these states

$$\Delta E = E_2 - E_1 = h\nu \equiv \hbar\omega \tag{5.1}$$

where h is Planck's constant. Another advantage of the microscopic quantum systems over macroscopic oscillators is based on the fact that all atomic systems of a given species are the same and consequently have the same transition frequencies. Thus, after having determined the frequency of a particular microscopic oscillator, an unlimited number of identical copies of this frequency standard can be cast to realise a fixed frequency in the electromagnetic spectrum. In contrast to the particular group of macroscopic resonators treated in Chapter 4, with their combs of nearly equidistant resonant frequencies, atoms and ions exhibit only a few suitable absorption lines in a given frequency range. In contrast, molecular quantum systems have a large number of transitions that can serve as frequency references in a much wider regime.

In this chapter we first recall the basic properties of quantum transitions in atoms (Section 5.1) and molecules (Section 5.2) with particular emphasis to examples relevant for frequency standards. Later, the quantitative description of the interaction of radiation with a two-level system is introduced (Section 5.3) for current use in further chapters. Last, we consider effects that are able to shift and broaden the observed quantum transitions (Section 5.4) and thus ultimately limit the accuracy of frequency standards.

5.1 Energy Levels of Atoms

The possible energy states of an isolated quantum system such as a single atom, ion or molecule are readily determined by quantum theory where the quantum state is described by a wavefunction ψ or by a state vector often denoted using Dirac's bra ($\langle\psi|$) and ket ($|\psi\rangle$) notation. The temporal evolution of the state ψ is described by the time-dependent Schrödinger equation

$$\mathcal{H}\psi = i\hbar\frac{\partial\psi(t)}{\partial t} \tag{5.2}$$

which reduces for stationary cases to the time independent Schrödinger equation $\mathcal{H}\psi_n = E_n\psi_n$. The Hamiltonian operator $\mathcal{H}$ has to be chosen such that E_n contains all relevant contributions to the energy of the system. Important contributions include the electrostatic Coulomb attraction each electron experiences in the central field of the nucleus, the mutual Coulomb repulsion between the electrons, the magnetic interactions between the magnetic moments associated with the angular momenta and spins of the electrons and the nucleus as well as the interaction of the system with external fields.

5.1.1 Single-electron Atoms

To start with, we recall the simplest system where a single electron moves in the central field of a positively charged nucleus of charge $Z\ e$. This situation is realised in the hydrogen atom (Z = 1) or in hydrogenic atoms or ions. In the so-called "central-field approximation" only the kinetic energies of the nucleus and the electron and their respective Coulomb interaction are taken into account. As a solution of the time-independent Schrödinger equation, one obtains the wavefunction $\psi_n(\vec{r})$ describing the probability amplitude to find the electron at the position $\vec{r}$ together with the eigenvalues E_n of the energy. The wavefunctions of the hydrogenic atoms can be separated into a radial function $R(r)$ and a spherical harmonic function $Y_{l,m}(\theta,\phi)$

$$\psi_{n,l,m}(\vec{r}) = R_{n,l}(r)Y_{l,m}(\theta,\phi). \tag{5.3}$$

The wavefunction depends on the principal quantum number $n = 1, 2, 3\ldots$, designated as K, L, M ..., the orbital angular momentum quantum number $l = 0, 1, 2, 3\ldots$, $(n-1)$ designated as s, p, d, f ..., and the magnetic quantum number $m = -l, (-l+1), \ldots, (l-1), l$. The quantum number m defines the projection of the orbital angular momentum of an electron onto a chosen (z) axis.

In the central-field approximation, the energies of the discrete states

$$E_n = -hcR\frac{Z^2}{n^2} \equiv -\frac{m_r c^2}{2}\frac{Z^2\alpha^2}{n^2} \tag{5.4}$$

depend only on the principal quantum number n being an integer. The fine structure constant α and the Rydberg constant R are given as follows

$$\alpha \equiv \frac{e^2}{4\pi\epsilon_0\hbar c} \tag{5.5}$$

$$R \equiv \frac{m_r e^4}{8\epsilon_0^2 h^3 c} \equiv \frac{m_r}{m_e}R_\infty. \tag{5.6}$$

Here, e and m_e are the elementary charge and the rest mass of the electron, respectively. c is the speed of light and ϵ_0 is the dielectric permeability of free space. The reduced mass is given as

$$m_r = \frac{m_e m_n}{m_e + m_n} \tag{5.7}$$

where m_n is the mass of the nucleus. R_∞ is the Rydberg constant for a nucleus of infinite mass where the reduced mass coincides with the rest mass of the electron m_e. The energies of the hydrogenic atoms in the central-field approximation are represented in the energy level diagram shown in the left-hand part of Fig. 5.1. Since the energy levels of hydrogen-like

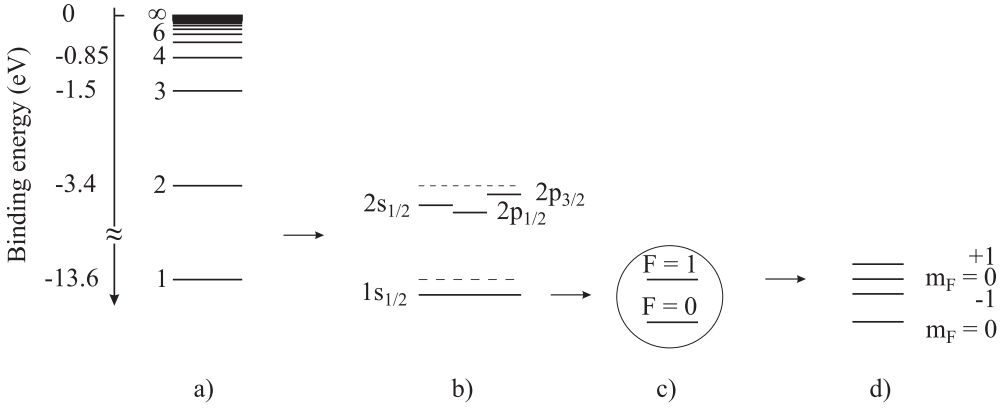


Figure 5.1: Schematic energy diagram of atomic hydrogen. a) Central-field approximation. b) Inclusion of the spin-orbit interaction and QED effects. c) Interaction with the nuclear spin. d) Interaction with a magnetic field (Zeeman effect).

atoms depend on the mass of the nucleus (see (5.4)) the energy levels of different isotopes of the same species are shifted by the so-called isotopic shift.[1]

Besides the Coulomb interaction leading to the central field approximation, the magnetic interactions of the magnetic moments $\vec{\mu}$ associated with the orbital angular momentum of the electron, the spin of the electron and of the nucleus contribute to the energy of the atomic system

$$E_{\mathrm{mag}} = -\vec{\mu} \cdot \vec{B}. \tag{5.8}$$

Like in classical physics, in quantum mechanics the magnetic moment of a rotating charge q is proportional to the angular momentum $\vec{J}$. For an atom the magnetic moment is determined by the electrons with charge $e = -q = -1.602 \times 10^{-19}$ A s and hence the magnetic moment of an atom

$$\vec{\mu} = -g_J \frac{e}{2m_e}\vec{J} = -g_J \frac{e\hbar}{2m_e}\frac{\vec{J}}{\hbar} \equiv -g_J \mu_B \frac{\vec{J}}{\hbar} \tag{5.9}$$

[1] If necessary, one explicitly indicates the total number of protons and neutrons in the nucleus for atoms, ions, or molecules, e.g., ^{40}Ca, ^{6}Be^{+}, or ^{127}I$_2$, respectively.

is always antiparallel to the angular momentum. Here, the Landé factor g_J is a dimensionless constant of order unity that can be calculated from quantum mechanics and $\mu_B = e\hbar/(2m_e) = 9.274 \times 10^{-24}$ J/T is referred to as the Bohr magneton with m_e the mass of the electron. The g factor is $g = 1$ for a pure orbital angular momentum $\vec{l}$ of the electron and $g \approx 2$ for a pure spin angular momentum $\vec{s}$. Conventionally the magnetic moment of the nucleus is written as

$$\vec{\mu} = g_I \frac{e}{2m_p} \vec{I} = g_I \frac{e\hbar}{2m_p} \frac{\vec{I}}{\hbar} \equiv g_I \mu_n \frac{\vec{I}}{\hbar} \tag{5.10}$$

where $\vec{I}$ is the nuclear spin and $\mu_n = e\hbar/(2m_p) = 5.051 \times 10^{-27}$ J/T is the nuclear magneton using the mass m_p of the proton.

The magnetic interactions of the magnetic moments associated with the spin and the orbital angular momentum also contribute to the energy of the atomic system. Hence, the simple energy level structure given by the central-field approximation (5.4) is modified by the spin-orbit coupling, resulting in a so-called fine structure. The corrections are calculated either by taking into account the spin-orbit interaction and the relativistic corrections due to the high velocity ($v/c > 10^{-2}$) of the electron or by the Dirac equation and lead to a modification of (5.4) as (see [87, 88])

$$E_{n,j} = -hcR\frac{Z^2}{n^2}\left[1 + \frac{(Z\alpha)^2}{n^2}\left(\frac{n}{j+1/2} - \frac{3}{4}\right) + \cdots\right]. \tag{5.11}$$

The spin-orbit interaction therefore reduces the energies of the possible levels of the electron in the hydrogenic atom in dependence of the total angular momentum quantum number j of the electron. According to (5.11) the separation between the $p_{3/2}$ and $p_{1/2}$, i.e., the fine structure of the n = 2 state, is about $3 \times 10^{-6} hcR$. As in the case of (5.4) for each principal quantum number n there are n^2 possible energy states which belong to the same energy. This degeneracy is only partly removed with respect to the total angular momentum j, but not with respect to l. This degeneracy is a particular property of the Coulomb potential where the energy levels do not depend on the quantum number l of the orbital angular momentum $\vec{l}$ (see (5.11)).[2]

Similar to the electronic shell the nucleus can have a total nuclear angular momentum $\vec{I}$ resulting from the spins and the particular angular momenta of the protons and neutrons that make up the nucleus. For the case of $\vec{I} \neq 0$ the coupling of $\vec{I}$ with the total angular momentum $\vec{J}$ of the shell to the total angular momentum $\vec{F}$ has to be taken into account. According to the rules of quantum mechanics the total angular momentum quantum number can be $F = J + I, J + I - 1, \ldots, |J - I|$. For the ground state of hydrogen $J = j = 1/2$ and $I = 1/2$ are coupled to F = 1 and F = 0 which leads to an associated splitting of the energy levels called the hyperfine structure (see Fig. 5.1 c). In a magnetic field the magnetic moment of the $F = 1$ state can be oriented in three different ways with respect to the direction of the magnetic induction $\vec{B}$. The three directions with their z component being parallel, perpendicular, and anti-parallel to $\vec{B}$ are designated by the quantum numbers $m_F = 1, 0, -1$,

[2] In fact, there is a small energy difference between the $s_{1/2}$ and $p_{1/2}$ states resulting from the Lamb shift.

respectively. The corresponding three energy states belong to three different energies in the magnetic field (see Fig. 5.1 d).

The discrete energy levels of hydrogenic atoms result in discrete absorption lines for electromagnetic radiation according to (5.1). In the hydrogen atom, the well-known Lyman, Balmer, or Paschen series can be excited for transitions from the states with n_1 = 1, 2, or 3, respectively, to states $n_2 = n_1 + 1, n_1 + 2, n_1 + 3, \ldots$ with higher energies. Many of the transitions allowed by Bohr's principle are not observed as a consequence of so-called selection rules reflecting conservation laws for particular physical quantities. As an example, consider the selection rule applicable for the interaction with electric dipole radiation. The photon carries a spin angular momentum of $\hbar$ and angular momentum conservation requires that the angular momentum of the atom changes by the same amount when a photon is absorbed or emitted. Hence,

$$\Delta J = 0, \pm 1, \qquad \text{except for} \quad J = 0 \leftrightarrow J = 0 \tag{5.12}$$

holds.

For frequency standards where the frequency of an oscillator is stabilised to the transition frequency of an atomic absorber, narrow-linewidth transitions are preferred that connect long-lived states. The hyperfine-split ground states represent such long-lived states,[3] e.g, the transition between the F = 1 and F = 0 ground states of the hydrogen atom. The convenient frequency separation of $\Delta\nu \approx 1.4\,\text{GHz}$ made this transition very suitable for a frequency standard which is exploited in the hydrogen maser (Section 8.1). Similarly, other highly accurate frequency standards are based on the hyperfine separation of the ground states of atoms and ions (see Table 5.1). With these magnetic dipole transitions, the selection rule $\Delta F = 0, \pm 1$ (but without $F = 0 \leftrightarrow F = 0$) has to be fulfilled exactly and $\Delta J = 0, \pm 1$ (again without $J = 0 \leftrightarrow J = 0$) approximately. Since the spontaneous decay rates for allowed dipole

Table 5.1: Ground-state splitting due to the hyperfine structure in neutral atoms. For ions see Table 10.1.

Atom	Frequency (Hz)	Standard	Reference
^{1}H	1 420 405 751.770(3)	H-maser (Section 8.1)	[1,90]
^{87}Rb	6 834 682 610.904 29(9)	Rb-clock (Section 8.2)	[91]
^{133}Cs	9 192 631 770.0 (exact)	Cs-clock (Section 7)	[1,92]

transitions are proportional to the third power of the transition frequency (see (5.133)) their linewidths rapidly increase from the microwave to the optical regime. For a dipole-allowed optical transition the lifetime of the excited state is typically of the order of a few nanoseconds and below and the associated linewidth resulting from (2.37) is several ten megahertz or

[3] The higher sublevels of the hyperfine Zeeman sub-states in the ground state decay predominantly by magnetic dipole radiation with a rate given by (5.133). Itano *et al.* [89] derive a spontaneous decay rate for a magnetic dipole transition at $\omega = 2\pi \times 30\,\text{GHz}$ of $2.7 \times 10^{-11}\,\text{s}^{-1}$ corresponding to a lifetime of the energetically higher state of about 1200 years.

more. Hence, mostly so-called (dipole) forbidden transitions are relevant for optical frequency standards.

5.1.2 Multi-electron Systems

Apart from the hydrogenic atoms discussed so far, all other atoms have many electrons and the mutual interaction of the electrons, in general, makes the energy structure of the respective atoms or ions more complicated. As an example of multi-electron atoms we consider alkaline earth atoms like magnesium or calcium (Fig. 5.2) where there are two electrons in the outer 3s or 4s shell, respectively.[4] In calcium, the most abundant isotope ^{40}Ca has no nuclear spin and

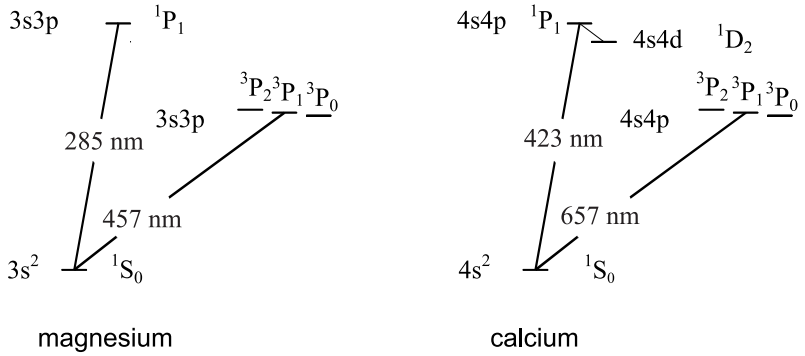


Figure 5.2: Partial energy diagrams of the alkaline earth atoms magnesium and calcium.

consequently, there is no hyperfine structure. The eighteen inner electrons fill the first shells $1s^2$, $2s^2$, $2p^6$, $3s^2$, and $3p^6$. The angular momenta of the two outer electrons can be described by the so-called LS coupling scheme. Besides the total angular momentum $\vec{J}$, in this scheme the total spin $\vec{S} = \sum \vec{s}_i$ resulting from a coupling of the spins of the particular electrons and the total orbital angular momentum $\vec{L} = \sum \vec{l}_i$ are conserved to a good approximation. As a result the additional selection rules for electric dipole radiation

$$\Delta L = 0, \pm 1 \tag{5.13}$$

and

$$\Delta S = 0 \tag{5.14}$$

are approximately valid. In the ground state of the calcium atom the spins of the two outer electrons of the 4s shell are anti-parallel, leading to $L = 0$, $S = 0$, and $J = 0$. We describe the states by using the nomenclature $n\,^{2S+1}L_J$ as $4\,^1S_0$ state, as 4s4s $^1S_0$ state, or as $4s^2\,^1S_0$ state. Owing to the multiplicity $2S + 1 = 1$ the ground state is a singlet state. The lowest excited states result from a 4s and a 4p single electron state. The coupling to a singlet state of

[4] A similar situation is encountered with the singly charged ions of the third group of the periodic system of the elements, e.g., in indium or thallium. These absorbers will be discussed in more detail in Section 9.4.4 and Section 10 for their use as optical frequency standards.

total spin $S = 0$ leads to the so-called $^1P_1$ state which decays with a decay time τ = 4.6 ns to the ground state $^1S_0$ by emission of blue light ($\lambda = 423$ nm). Besides the singlet state, there is also a triplet of states at lower energy resulting from a combination of the 4s and 4p states. In this state the two spins are parallel (S = 1) and with L = 1 the three states $^3P_2$,$^3P_1$ and $^3P_0$ result from the three combinations of $J = L + S = 2, L + S - 1 = 1$ and $L - S = 0$, respectively.

5.1.2.1 Forbidden Atomic Transitions for Optical Frequency Standards

In the simplified picture of Fig. 5.2 the transition of the atom from the $^3P_1$ state to the ground state ($\lambda = 657$ nm) requires a spin flip of one of the two electrons, which cannot be accomplished by an electric field (see (5.14)). Thus this transition is forbidden for electric dipole radiation if the LS coupling would be exactly fulfilled, i.e., if S and L were conserved exactly. Experimentally one finds that the natural lifetime of the excited $^3P_1$ state is $\tau \approx 0.5$ ms. This lifetime is about five orders of magnitude higher than that of the excited $^1P_1$ state in the singlet configuration. The intercombination transition between the singlet and triplet systems is therefore forbidden in such a sense that it is about five orders of magnitude less probable compared to the corresponding transition in the singlet system, $^1P_1 \rightarrow ^1S_0$ at $\lambda = 423$ nm.

The lighter the atoms are, the better the LS coupling scheme applies. Correspondingly, the intercombination selection rule is weakened for atoms with a higher number of protons (Z) and electrons. As a result, the lifetime of the $^3P_1$ state (Fig. 5.2) decreases from Mg to Ba. The

Table 5.2: Selected narrow transitions in atoms suitable for optical frequency standards. The frequencies are from [93], [94], [95], [96], [97], [98] for H, Mg, Ca, Sr, Ag, Xe, respectively. Other candidates can be found, e.g., in reference [99]. For ions see Table 10.2. Wavelengths denoted by an asterisk refer to that of one of the two photons required to excite the two-photon transition.

Atom	Transition	Frequency THz	Wavelength nm	Width Hz
^{1}H	1S - 2S	2 466.061 413 187 103(46)	243.13 *	1
^{24}Mg	3^1S_0 - 3^3P_1	655.658 9	457.24	40
^{40}Ca	4^1S_0 - 4^3P_1	455.986 240 494 15	657.46	370
^{88}Sr	5^1S_0 - 5^3P_1	434.829 121 311(10)	689.45	6 900
^{109}Ag	$5s\ ^2S_{1/2}$ - $4d^9\,5s^2\ ^2D_{5/2}$	453.320 4	661.33 *	0.8
^{132}Xe	$6s'[1/2]_0$ - $6s[3/2]_2$	136.844	2190.76 *	1.2 [5]

transition $^3P_2 \rightarrow ^1S_0$ with $\Delta J = 2$ is only allowed for electric quadrupole radiation and hence the excited state lifetime is more than 5000 seconds in magnesium. Excited states connected to lower energetic states only by higher-order multipole transitions can have extraordinarily long lifetimes of, e.g., ten years in the $^2F_{7/2}$ state of the $^{171}Yb^+$ ion that is connected to the

[5] At room temperature, however, a much larger linewidth of 12 Hz is observed as a result of transitions induced by black-body radiation to energetically higher states which subsequently can decay [100].

ground state via an octupole transition [101] (Section 10.3.2.2). The transition $^3P_0 \rightarrow {}^1S_0$ (Fig. 5.2) as an example of a $J = 0 \rightarrow J = 0$ transition is totally forbidden in ^{40}Ca. This is a consequence of angular momentum conservation where the photon has to carry away an angular momentum of at least $\hbar$ but the atom has no net angular momentum in both states.

The combination of the selection rules (5.13) and (5.14) also makes the 1S – 2S transition (Fig. 5.1) in atomic hydrogen a forbidden transition with a natural linewidth of about 1 Hz. This transition can be excited by intense laser fields where two photons with a wavelength of 243.1 nm (see Table 5.2) are absorbed at the same time [102]. The 1S – 2S [93] and the 2S – 8S/D transitions [103] and other transitions [88] have been utilised recently to set up optical frequency standards or to perform precision measurements. Atomic silver also has a forbidden transition (see [99, 104, 105], Table 5.2) that can be excited by two photons. The realisation of standards based on these transitions is described in Section 9.

5.2 Energy States of Molecules

In contrast to atoms and ions, the spectrum of molecules contains many more spectral lines as a result of their complex energy structure. In the following we begin with the description of molecules built up from two identical atoms like the $^{127}I_2$ molecule that is used as an absorber in a variety of optical frequency standards. The two atoms in the molecule are separated by the distance R. For a large separation $R \rightarrow \infty$, the total energy of the molecule composed of two atoms is given by the sum of the energies of the two isolated atoms. If the two atoms come closer they may attract or repel each other and the energy states of the molecule depend on the distance R and split into so-called bonding and anti-bonding states (Fig. 5.3). For very small distances R between the nuclei the two atoms repel each other due to their Coulomb interaction. Consequently, when the distance R is reduced, the energy of a bonding state initially decreases to a minimum at the equilibrium distance R_0 and increases again towards even smaller distances. In contrast to the spherical symmetry of the central potential of an

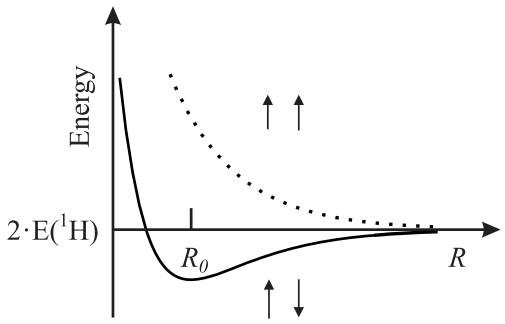


Figure 5.3: The energy states of a molecule split into bonding (full line) and anti-bonding (dotted line) states, both depending on the distance R between the two atoms. In the case of an H_2 molecule the bonding state and the anti-bonding state correspond to the cases where the electron spins are anti-parallel and parallel, respectively. R_0 represents the equilibrium distance between the atoms in the bonding state of the molecule.

atom the potential of a two-dimensional molecule has a cylindrical symmetry. The axis of the cylinder coincides with the inter-nuclear axis of the molecule, i.e., the line of sight between the nuclei. In the case of a small interaction between the spin and the orbital angular momentum, i.e., for small multiplet splitting, the total spin quantum number S of the molecule results from the two spins $\vec{S}_1$ and $\vec{S}_2$ of the individual atoms. The total spin is a vector combination leading to the values $S = S_1 + S_2, S_1 + S_2 - 1, \ldots |S_1 - S_2|$. The orbital angular momenta $\vec{L}_1$ and $\vec{L}_2$ of the individual atoms, however, are quantised with respect to the symmetry line of the molecule. The quantum number of the component of the resulting electronic orbital angular momentum is defined as Λ with the values Λ = 0, 1, 2, ... In an analogous way to the atom, the molecular states are labelled as $\Sigma, \Pi, \Delta, \Phi, \ldots$ for Λ = 0, 1, 2, 3, ..., respectively. The total angular momentum $\vec{\Omega}$ of the electrons projected onto the inter-nuclear axis is composed of the orbital angular momentum $\vec{\Lambda}$ and the spin $\vec{S}$ in an analogous way as the total electronic angular momentum $\vec{J}$ results from $\vec{L}$ and $\vec{S}$ for atoms. The symmetry of the potential of the molecule moreover requires that the spatial density of the electronic charge distribution of the molecule is symmetric with respect to any mirror plane containing the centres of both nuclei. Consequently, the symmetry of the electronic wavefunction can be either even or odd which is indicated by the plus or minus sign as $\Omega = 0^+, 1^+, \ldots$ and $\Omega = 0^-, 1^-, \ldots$ in the symmetric and anti-symmetric case, respectively. In a homonuclear molecule comprising two identical atoms the centre of symmetry is located halfway between the two atoms. As the electronic charge density will reflect this symmetry the charge density will not change if all coordinates of the electrons are inverted with respect to this centre of symmetry and the corresponding wavefunction will have an even or an odd symmetry. Wavefunctions representing an even function or an odd function with respect to this symmetry operation are denoted by "g" and "u", respectively.[6] To identify the different electronic states, traditionally, the ground state is referred to as the X state and the higher lying electronic states are labelled by A, B, C, ... in a somewhat arbitrary manner depending on the order of their first identification.

5.2.1 Ro-vibronic Structure

A quantum mechanical treatment of the molecule is often based on the Born–Oppenheimer approximation.[7] The wavefunctions of the electrons and the nuclei are separated and as a result the energy of the molecule is given as the kinetic energy of the centre-of-mass system and a contribution that depends only on the distance R between the two nuclei. The latter part is described by the Schrödinger equation [106, 107]

$$\left\{ -\frac{\hbar^2}{2m_r}\frac{\partial^2}{\partial R^2} - \frac{\hbar^2}{m_r}\frac{1}{R}\frac{\partial}{\partial R} + \frac{\hbar^2 J(J+1)}{2m_r R^2} + V(R) \right\} \chi(R) = E_{v,J} \cdot \chi(R) \qquad (5.15)$$

where m_r is the reduced mass of the molecule and $\chi(R)$ is the wavefunction describing the relative motion of the two nuclei.

In the vicinity of the equilibrium distance R_0, the potential energy $V(R)$ (Fig. 5.3) can be approximated by a parabola $V(R) = V(R_0) + (R - R_0)^2/2$, which is the potential corre-

[6] g and u are the abbreviations for the German words *gerade* (even) and *ungerade* (odd).

[7] In the Born–Oppenheimer approximation the vibration and rotation of the nuclei are treated separately from the motion of the electrons since the electrons can follow the movement of the nuclei almost instantaneously.

sponding to an harmonic oscillator. For $J = 0$ (5.15) then represents the Schrödinger equation of a quantised harmonic oscillator whose equilibrium position is shifted by R_0 from the origin of the coordinate system with the vibrational eigenvalues

$$E_{\mathrm{vib}} = \hbar\omega_{\mathrm{vib}} \left(v + \frac{1}{2} \right). \tag{5.16}$$

$\omega_{\mathrm{vib}}/(2\pi)$ is the vibrational frequency and v is the vibrational quantum number. In this approximation the nuclei perform harmonic oscillations along the line connecting the two nuclei.

Since the amplitude of the oscillation is small with respect to the equilibrium distance, $R \approx R_0$ holds, and the third term in curled brackets of (5.15) is approximately constant. Hence the energy eigenvalues to (5.15) are given by

$$E = V(R_0) + \left(v + \frac{1}{2} \right) \hbar\omega_{\mathrm{vib}} + \frac{\hbar^2 J(J+1)}{2\Theta} \tag{5.17}$$

with $\Theta = m_r R_0^2$. The latter term corresponds to the rotational energy $E_{\mathrm{rot}} = J^2/(2\Theta)$ of a classical rotating dumb-bell with two masses $m_r/2$ at a distance $2R_0$. The rotational energy of the quantum mechanical rotator depends on the moment of inertia Θ and the angular momentum $\vec{J}$ with the quantum numbers $J = 0, 1, \ldots$. Hence, each vibrational state in the potential energy of the electronic states is accompanied by a ladder of bound rotational states (Fig. 5.4). Compared to an atom, a molecule has additional degrees of freedom, i.e., the

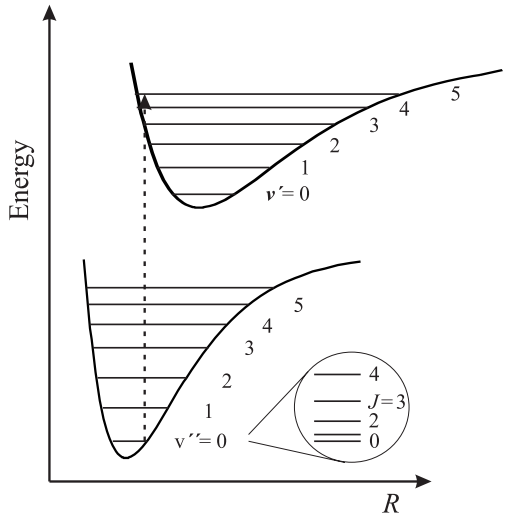


Figure 5.4: Energy level scheme of a molecule (not to scale).

vibrations of the atoms with respect to the centre of gravity of the molecule, and the rotation of the molecule.

For larger excursions from the equilibrium distance R_0 the asymmetry of the potential curve has to be taken into account and the vibration becomes anharmonic. In comparison to the purely harmonic oscillator the anharmonicity leads to a modification of the vibrational energy levels E_{vib} that are no longer equidistantly spaced.

Similarly, the model of a rigid rotator with a constant moment of inertia Θ has its deficits when applied to the description of a real molecule. Owing to the anharmonicity of the potential curve and to the stretching due to the centrifugal force, the energy levels depend on the vibrational quantum number and on the rotational quantum number, respectively. In general, the energies of the rotational and vibrational energy levels are no longer de-coupled and, hence, they are often referred to as ro-vibronic levels. To describe the energy states and to predict the experimentally observed transitions, more realistic potentials $V(R)$ have to be used which are usually derived from a comparison with experimentally determined transition frequencies.

5.2.2 Optical Transitions in Molecular Iodine

For optical frequency standards the iodine molecule is one of the most prominent absorbers. By absorbing a photon the molecule makes a transition from an energetically lower electronic state referred to as $E''_{\rm el}$ with the vibrational state $E''_{\rm vib}$ and the rotational state $E''_{\rm rot}$ to an energetically higher electronic state with $E'_{\rm el}, E'_{\rm vib}$ and $E'_{\rm rot}$ (Fig. 5.4). For low-lying vibrational levels in the I_2 molecule the vibrational frequencies $\omega_{\rm vib}/(2\pi)$ are separated by about 6 THz in the X state and by about 4 THz in the B state. For low-lying rotational levels of the iodine molecule the rotational frequencies $\omega_{\rm rot}/(2\pi)$ are separated by about 3 GHz. At room temperature, the thermal energy $k_BT \approx hc/\lambda$ corresponds to $1/\lambda \approx 200\,{\rm cm}^{-1} = 6\,{\rm THz}$. Hence, only the low-lying $v'' = 0, 1, 2$ vibrational states are populated with up to a hundred rotational states. Together with the roughly 80 accessible vibrational states v' there are about 60 000 fine structure lines resulting from transitions in the B – X system in the spectral region between about 500 nm and roughly 900 nm. Due to the symmetries and the resulting conservation laws not all transitions are allowed. The allowed ones obey specific selection rules (see Table 5.3). In molecules, the transition probabilities (see (5.133)) can vary largely, even

Table 5.3: Selection rules for electric dipole radiation in molecules.

$\Delta\Lambda$	=	$0, \pm 1$	
$\pm$	$\leftrightarrow$	$\pm$	
g	$\leftrightarrow$	u	
$\Delta J \equiv J' - J''$	=	-1	(P - branch)
	=	0	(Q - branch; but not $J' = 0 \leftrightarrow J'' = 0$)
	=	+1	(R - branch)

for transitions allowed by the selection rules. The observed absorption strength of a particular transition is determined by the Franck–Condon principle [8] reflecting the structures of the relevant wavefunctions. Consider the transition indicated in Fig. 5.4 connecting regions of R near the boundaries of the potential wells. In these regions the wavefunctions go to zero and, hence, the matrix elements corresponding to the transition probabilities will become small. On the other hand, transitions occurring at distances where the wavefunctions of both states

[8] The quantum mechanical transition probability is given by the Franck–Condon integral $\int \chi_{v'}(\vec{R})\chi_{v''}(\vec{R})dR$. $\chi_{v''}(R)$ and $\chi_{v'}(R)$ are the wavefunctions describing the nuclear oscillations in the ground state with energy E'' and the excited state with energy E', respectively.

have large values, i.e., where the charge density shows anti-nodes, can be expected to exhibit large transition probabilities.

The particular absorption lines are classified by comparison with calculations similar to (5.17) by reference to the quantum numbers. As an example consider the transition from the ground state X $^1\Sigma_g^+$ to the excited state B Π_u^+ (11 - 5), R(127) which coincides with the Doppler broadened emission line of the He-Ne laser and is used as an optical frequency standard (see Section 9.1.3 and Table 5.4). Σ and Π denote the orbital angular momentum around the line connecting both nuclei, $\Lambda = 0$ and $\Lambda = 1$, respectively. The parity of the wavefunctions of the ground state (excited state) is even (odd) denoted by g (u), i.e., the wavefunction keeps (changes) its sign by inversion at a centre of symmetry of the molecule. The ground state (excited state) wavefunction of the molecule is symmetric (anti-symmetric) referred to as $+$ ($-$) with respect to a mirror plane intersecting the line of sight between the nuclei. The vibrational quantum number of the electronic ground state (excited state) is $v'' = 5$ ($v' = 11$).[9] The rotational angular momentum quantum number is $J'' = 127$ and since this transition belongs to the so-called R-branch ($J' = J''$ + 1; see Table 5.3) $J' = 128$ holds.

The Doppler-broadened absorption spectrum of iodine has been measured using Fourier transform spectroscopy by Gerstenkorn and coworkers between 11 000 cm^{-1} (905 nm) and 20 000 cm^{-1} (500 nm) [108, 109]. Kato presented an atlas with Doppler-free lines in the range between 15 000 cm^{-1} (905 nm) and 19 000 cm^{-1} (500 nm) [110]. A large number of accurate determinations of transition frequencies are available (see [111, 112] and references therein).

5.2.2.1 Determination of Molecular Potentials

Since the potential $V(R)$ is intimately connected with the energy levels in a molecule (see (5.15)), the experimentally determined transitions can be used to identify the energy levels and to determine more realistic potential curves. These potential curves, on the other hand, allow one to extrapolate to hitherto unknown energies and line spectra. A particular approach according to Dunham [113] uses a series expansion of the potential $V(R)$ of a vibrating rotor which leads to energy levels

$$E(v, J) = \sum_{k,l} Y_{k,l} \left(v + \frac{1}{2} \right)^k [J(J+1)]^l , \qquad k, l = 0, 1, 2, \ldots \tag{5.18}$$

The $Y_{k,l}$ are referred to as Dunham coefficients. The Dunham coefficients can be determined from a fit of a measured ro-vibronic spectrum allowing one to construct potentials $V(R)$. Gerstenkorn and Luc [114] were able to represent their measured 17 800 iodine lines by 46 molecular parameters with a fractional uncertainty of about 10^{-7}. More recently, Knöckel *et al.* [112] described selected bands in the B – X spectrum of iodine between 778 nm and 815 nm using Dunham coefficients with an uncertainty of less than 200 kHz. A different model based on analytical molecular potentials allowed a prediction of the iodine lines between 515 nm and 815 nm with an uncertainty of less than 12 MHz.

[9] As in atoms, in molecules the quantum numbers of the energetically higher state are denoted by a prime. The energetically lower state to which this state is connected by a transition is usually denoted by a double prime in molecular spectroscopy in contrast to atomic spectroscopy where it is not primed.

Alternatively, a fully quantum mechanical description of the ro-vibrational structure can be performed by using analytical potentials $V(R)$ and integrating the Schrödinger equation (5.15) numerically (see e.g. [115]).

5.2.2.2 Influence of the Hyperfine Structure

The hyperfine interaction of the magnetic moment associated with the electrons in the shell and the magnetic moment of the nuclei leads to a splitting of the lines (Fig. 5.5). The two

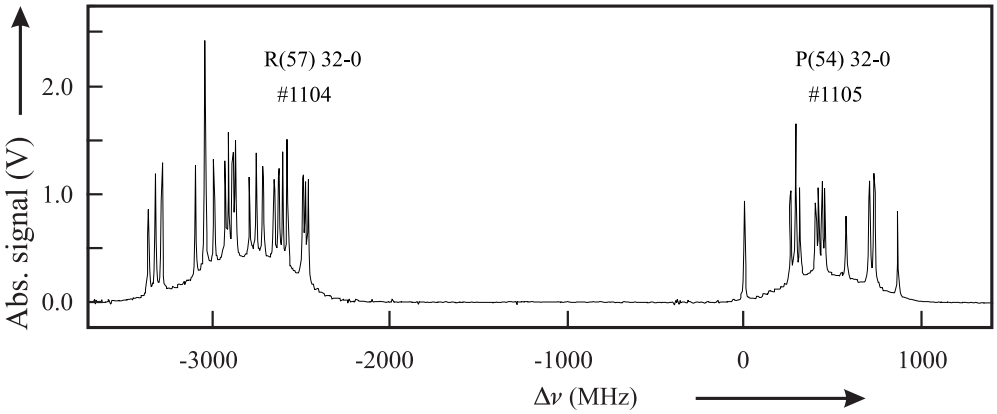


Figure 5.5: Observed hyperfine multiplet of the R(57) 32-0 and P(54) 32-0 lines of $^{127}I_2$ molecules also known as lines 1104 and 1105 according to the iodine atlas of Gerstenkorn *et al.* [108]. Courtesy of H. Schnatz.

multiplets in Fig. 5.5 of the so-called 1104 and 1105 lines comprise 15 or 21 hyperfine components, respectively, which results from the spin statistics and the selection rules for electronic dipole radiation. The nuclear spin quantum number of an iodine atom in the $^{127}I_2$ molecule is $I_1 = I_2 = 5/2$ and $I = 7/2$ for $^{129}I_2$, resulting in a total nuclear spin quantum number of $I = |I_1 - I_2|, |I_1 - I_2 + 1|, \ldots, |I_1 + I_2| = 0, 1, \ldots, 5$ for $^{127}I_2$ and $|I_1 + I_2| = 0, \ldots 7$ for $^{129}I_2$. Since the nuclei are fermions the wavefunction which is the product of the spin wavefunction and the spatial wavefunction including the rotational function has to be anti-symmetric with respect to an exchange of the nuclei. Consequently, for any state with even parity (like the X ground state) a symmetric wavefunction of the nuclear spins requires an odd rotational wavefunction with odd values of J and vice versa for an anti-symmetric nuclear spin wavefunction a symmetric rotational wavefunction is necessary. In the iodine molecule the anti-symmetric nuclear spin wavefunction requires $I = I_1 + I_2 = 0, 2, 4$ and the symmetric one $I = 1, 3, 5$. Hence, anti-symmetric wavefunctions of the X state are obtained by coupling even J'' and even I states or odd J'' and odd I states. Consequently, the ground state in $^{127}I_2$ splits into

$$\begin{aligned} \sum_{I=0,2,4} (2I+1) &= 15 \qquad \text{for even } J'' \\ \sum_{I=1,3,5} (2I+1) &= 21 \qquad \text{for odd } J'' \end{aligned} \tag{5.19}$$

hyperfine components and in the case of $^{129}I_2$ into 28 for even J'' and 36 for odd J''. Electric dipole transitions obey the selection rules $\Delta J = \pm 1, \Delta I = 0$ and $\Delta F = 0, \pm 1$. For higher values of J only $\Delta F = \Delta J$ are relevant [106]. An anti-symmetric wavefunction of the excited B state results from a coupling of states with even J' and odd I (21 hyperfine levels $^{127}I_2$) or odd J' and even I states (15 hyperfine levels for $^{127}I_2$). Hence, the optical transitions for high J'' of the ground state also show 15 (21) components for even (odd) J''. This can be seen from Fig. 5.5 where the line R(57) of $^{127}I_2$ with an odd $J'' = 57$ in the X state exhibits 21 hyperfine transitions in contrast to the P(54) line with even J'' which has 15 hyperfine absorption lines. Arguments like these allow one to characterise the observed transitions and to compare the measured frequencies with the calculated ones.

The hyperfine structure results from the contributions of the electric and magnetic interactions of the nuclear moments with the electrons and the moments of the other nuclei in the molecule. Despite the complexity of the problem, often an effective hyperfine Hamiltonian can be introduced [116] that allows one to describe the hyperfine energies of a particular ro-vibronic state as follows

$$\mathcal{H}_{\mathrm{hfs,eff}} = \mathcal{H}_{\mathrm{EQ}} + \mathcal{H}_{\mathrm{SR}} + \mathcal{H}_{\mathrm{SSS}} + \mathcal{H}_{\mathrm{TSS}}. \tag{5.20}$$

In (5.20) $\mathcal{H}_{\mathrm{EQ}}$ is the electric quadrupole interaction, $\mathcal{H}_{\mathrm{SR}}$ is the spin-rotation interaction, $\mathcal{H}_{\mathrm{SSS}}$ is the scalar spin–spin interaction, and $\mathcal{H}_{\mathrm{TSS}}$ is the tensor spin–spin interaction. The matrix elements of these contributions are often separated into a product of different geometrical factors g_i and four hyperfine parameters eQq, C, A, D [111,117] leading to the hyperfine energy splitting

$$\langle (J'I'), F|\mathcal{H}_{\mathrm{hfs,eff}}|(J,I),F\rangle = eQq \cdot g_{eQq} + C \cdot g_{\mathrm{SR}} + A \cdot g_{\mathrm{SSS}} + D \cdot g_{\mathrm{TSS}}. \tag{5.21}$$

Based on physical models, interpolation formulae for the hyperfine splittings have been derived and fitted to measured frequency separations (see, e.g., [111, 117] and references therein). Bodermann *et al.* give an uncertainty for the calculation of the hyperfine splittings of < 30 kHz for the wavelength range between 514 nm and 820 nm [111]. Within the hyperfine multiplet of a particular transition the four parameters of (5.21) can be used to fit the observed hyperfine splittings with residuals below 1 kHz [118].

5.2.3 Optical Transitions in Acetylene

In the infrared and in particular in the region of the optical telecommunication bands around 1.3 µm and 1.5 µm there is a need for transitions that can be used as reference lines. A prominent example of such a molecular absorber is the acetylene molecule (C_2H_2; H−C≡C−H) [47, 119–123]. The molecule has a strong three-fold bond between the central C atoms and a weaker single bond between the respective C and H atoms. The acetylene molecule has a linear symmetry and can vibrate in different modes (Fig. 5.6). The vibrational mode ν_2 results when the C atoms vibrate towards each other and each H atom vibrates more or less in phase with its neighbouring C atom corresponding predominantly to a stretching of the C ≡ C bond. In contrast, the ν_3 vibrational mode results mainly from the stretching of the C – H bond. ν_4 and ν_5 correspond to a bending of the C ≡ C and C – H bonds, respectively. The frequencies (wavenumbers; wavelengths) corresponding to the vibrational modes $\nu_1, \nu_2, \nu_3, \nu_4$, and ν_5 are

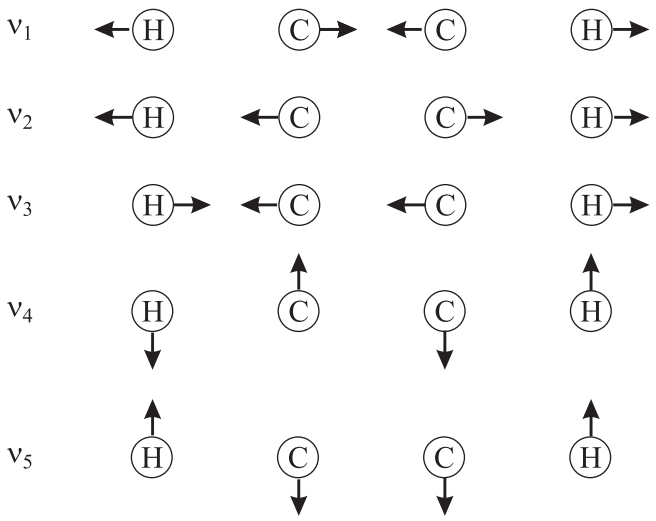


Figure 5.6: Five normal vibrations of acetylene (H – C ≡ C – H). The arrows indicate the motions of the respective atoms at a particular instant.

101.1 THz (3373 cm^{-1}; 2.965 µm), 59.2 THz (1974 cm^{-1}; 5.066 µm), 98.4 THz (3282 cm^{-1}; 3.047 µm), 18.4 THz (613 cm^{-1}; 16.31 µm), and 21.9 THz (730 cm^{-1}; 13.70 µm), respectively [47]. The anharmonicity of the potential (Fig. 5.4) corresponds to a non-linearity between a driving force and the excursion of the atoms. Hence, so-called overtone spectra with $2\nu, 3\nu, \ldots$ can be observed, however, with strongly decreasing intensities. The non-linear interaction furthermore leads to combinations, e.g., $\nu_1 + \nu_2, \nu_1 - \nu_2, 2\nu_1 - \nu_2, \ldots$.[10] The transitions resulting from a combination of the ν_1 and ν_3 oscillations (Figs. 5.7 b) and 13.5) are frequently used as frequency and wavelength references in the telecommunication band (see Table 13.1, [124]).

The origin of the R (P) branch results from the almost linearly increasing (decreasing) energy differences (see Fig. 5.7 a) from $J'' \rightarrow J'$ for increasing J'' and $\Delta J = +1$ ($\Delta J = -1$). The intensity variations in Fig. 5.7 b) between even and odd J transitions are a result of the nuclear spin which affects the number of possible molecular states similarly as discussed in the case of iodine (Section 5.2.2.2). The shape of the envelope of the spectra results from the interplay between the increasing number of accessible states with increasing J and the decreasing thermal occupation of these states. The absorption lines of the $^{13}C_2H_2$ are shifted by about 8 nm towards higher wavelengths with respect to the ones of the $^{12}C_2H_2$ molecules. This shift can be explained by the different masses of the two carbon isotopes, therefore leading to a variation of the vibrational frequency of the C – H bonds.

[10] Note that the frequencies of the combined oscillations agree roughly, but not exactly, with the combined frequencies of the particular oscillations.

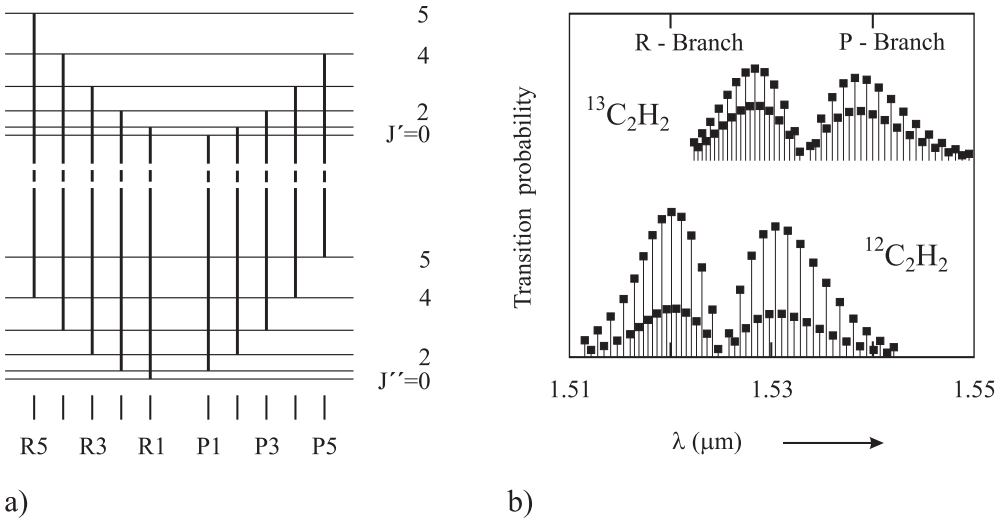


Figure 5.7: a) Energy levels and rotational transitions in acetylene. b) Corresponding calculated rotational bands of the acetylene spectrum resulting from the $\nu_1 + \nu_3$ modes of Fig. 5.6.

5.2.4 Other Molecular Absorbers

The transitions in molecular iodine are restricted to the green, red, and near infrared part of the electromagnetic spectrum. For the infrared spectral range there are other suitable absorbers, e.g., H_2O, NH_3, HCN, HI, Cs_2, O_2, or others [47]. In the blue and green spectral range tellurium Te_2 is often used [125–128].

In the methane molecule CH_4 used for stabilisation of the He-Ne laser (see Table 5.4) near 3.39 µm, the four H atoms form a regular tetrahedron around a C atom in the centre. The high symmetry leads to four fundamental vibrational modes, a two-fold degenerate vibration, and two three-fold degenerate vibrations referred to as A_1, E, and F_2, respectively. Methane stabilised lasers used as high-accuracy optical frequency standards are treated in Section 9.1.4. The osmium tetraoxide (OsO_4) molecule has a three-fold degenerate ν_3 mode with the absorption band near 10.42 µm that coincides with the emission of the CO_2 laser (Section 9.1.5). A more detailed compilation of molecular transitions for frequency stabilisation of lasers can be found, e.g., in reference [47].

5.3 Interaction of Simple Quantum Systems with Electromagnetic Radiation

5.3.1 The Two-level System

To model the interaction of an atomic quantum system with a monochromatic electromagnetic field it often suffices to treat the system as a two-level system having only two states with energies E_1 and E_2 with $E_2 > E_1$. These states are referred to in many ways, such as $|1\rangle$ and

Table 5.4: Selected optical transitions in molecules used for frequency standards. Others may be found in [47] and Table 9.1. DL: Diode Laser; Dye: Dye Laser.

Molecule	Transition	Frequency THz	Wavelength μm	Lasers	Ref.
OsO_4	R(12)	29.096 274 952 34	10.303	CO_2	[95]
CH_4		88.376 181 600 18	3.392	He-Ne	[95]
$^{12}C_2H_2$	P(21) $\nu 1 + \nu 3$	194.916 199 55(15)	1.538	DL	[124]
$^{13}C_2H_2$	R(23) $\nu 1 + \nu 3$	196.929 745 92(15)	1.522		[124]
$^{12}C_2H_2$	R(18) $\nu 1 + \nu 3$	197.750 466 56(15)	1.516	DL	[124]
HCN	P(27)	192.622 446 9(1)	1.556	DL	[129]
$^{127}I_2$	R(42) 0-17 b_1	367.615 127 628(14)	0.816	DL	[130]
	R(127) 11-5 a_{13}	473.612 214 705	0.633	He-Ne	[95]
$^{130}Te_2$	d_4	613.881 149 1(5)	0.488	Dye	[127]
		642.116 513 6(6)	0.467	Dye	[128], [131]

$|2\rangle$, as the ground state and the excited state, as $|g\rangle$ and $|e\rangle$, or $|\downarrow\rangle$ and $|\uparrow\rangle$, respectively. We briefly go through this description following the approach given in textbooks [11, 132, 133] in order to derive the relevant tools and formulas that will be used in the remainder of this book. We start with the time-dependent Schrödinger equation of (5.2), writing the Hamiltonian as

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_{\rm int}. \tag{5.22}$$

$\mathcal{H}_0$ describes the system without interaction with the radiation using the time-independent Schrödinger equation neglecting spontaneous emission

$$\mathcal{H}_0 \phi_k(\vec{r}) = E_k \phi_k(\vec{r}) \tag{5.23}$$

with $k = 1, 2$ and where $\vec{r}$ denotes all internal degrees of freedom, such as positions of electrons, spins, etc. The operator $\mathcal{H}_{int}$ represents the contribution of the perturbation due to interaction with the radiation field. We assume that $\mathcal{H}_{\rm int}$ has no diagonal elements.[11] Since the eigenfunctions $\phi_k(\vec{r})$ form a complete set, the general solution $\psi(\vec{r}, t)$ of (5.2) can be expressed in the interaction picture as a linear combination of this basis set as follows

$$\psi(\vec{r}, t) = c_1(t) e^{-iE_1 t/\hbar} \phi_1(\vec{r}) + c_2(t) e^{-iE_2 t/\hbar} \phi_2(\vec{r}). \tag{5.24}$$

[11] One can easily remove any constant energy by going to new basis states which shift the energies of both states, which is then included in $\mathcal{H}_0$.

In general, the $c_k(t)$ now include the time dependence introduced by the perturbation operator $\mathcal{H}_{\rm int}$. To find the explicit time dependence (5.22) and (5.24) are inserted into (5.2) which leads to

$$(\mathcal{H}_0+\mathcal{H}_{\rm int})\left[c_1(t)e^{-iE_1t/\hbar}\phi_1(\vec{r})+c_2(t)e^{-iE_2t/\hbar}\phi_2(\vec{r})\right] = i\hbar\frac{\partial}{\partial t}\left[c_1(t)e^{-iE_1t/\hbar}\phi_1(\vec{r})+c_2(t)e^{-iE_2t/\hbar}\phi_2(\vec{r})\right] \tag{5.25}$$

or

$$\begin{aligned} & E_1c_1(t)e^{-iE_1t/\hbar}\phi_1(\vec{r})+E_2c_2(t)e^{-iE_2t/\hbar}\phi_2(\vec{r}) \\ & +\mathcal{H}_{\rm int}c_1(t)e^{-iE_1t/\hbar}\phi_1(\vec{r})+\mathcal{H}_{\rm int}c_2(t)e^{-iE_2t/\hbar}\phi_2(\vec{r}) \\ = \;& i\hbar\frac{dc_1(t)}{dt}e^{-iE_1t/\hbar}\phi_1(\vec{r})+E_1c_1(t)e^{-iE_1t/\hbar}\phi_1(\vec{r}) \\ & +i\hbar\frac{dc_2(t)}{dt}e^{-iE_2t/\hbar}\phi_2(\vec{r})+E_2c_2(t)e^{-iE_2t/\hbar}\phi_2(\vec{r}) \end{aligned} \tag{5.26}$$

where we have used (5.23) on the left-hand side. After subsequent multiplication from the left by $\phi_1^*(\vec{r})$ and $\phi_2^*(\vec{r})$ and by integrating over the spatial coordinates $\vec{r}$ one obtains

$$i\hbar\frac{dc_1}{dt} = c_2(t)H_{12}(t)e^{-i\omega_0t} \tag{5.27}$$

$$i\hbar\frac{dc_2}{dt} = c_1(t)H_{21}(t)e^{+i\omega_0t}. \tag{5.28}$$

with $E_2-E_1\equiv\hbar\omega_0$ and the time-dependent matrix elements

$$H_{21}(t) \equiv \int\phi_2^*(\vec{r})\mathcal{H}_{\rm int}(t)\phi_1(\vec{r})d^3r\equiv\langle 2|\mathcal{H}_{\rm int}(t)|1\rangle \tag{5.29}$$

and

$$H_{12}(t) \equiv \int\phi_1^*(\vec{r})\mathcal{H}_{\rm int}(t)\phi_2(\vec{r})d^3r\equiv\langle 1|\mathcal{H}_{\rm int}(t)|2\rangle. \tag{5.30}$$

Since $\mathcal{H}_{\rm int}$ is Hermitian it follows that

$$H_{21}=H_{12}^*. \tag{5.31}$$

For the probabilities $|c_1|^2$ and $|c_2|^2$ to find the two-level atom in state $|1\rangle$ and state $|2\rangle$, respectively, the relation holds

$$|c_1(t)|^2+|c_2(t)|^2=1. \tag{5.32}$$

The solutions of the coupled differential equations (5.27) and (5.28) for $c_1(t)$ and $c_2(t)$ can be obtained if the wavefunctions or state vectors for the two levels are known together with the Hamiltonian that couples the two states.

The Hamiltonian describing the particular interaction of the atom with the electromagnetic field can be derived, e.g., from the minimal coupling between a charged particle of mass m and charge q at the position $\vec{r}$ with the vector potential $A(\vec{r},t)$ of the electromagnetic field [134, 135]. If the wavelength of the radiation is large compared to the extension of the

atom the field can be expanded into multipoles at the centre–of–mass position $\vec{r}_0$ of the atom. For frequency standards some of the most relevant interaction Hamiltonians are given [11] by

$$\mathcal{H}_{\rm int} = -\vec{d}\cdot\vec{E}(\vec{r}_0,t) = +q\vec{r}\cdot\vec{E} \quad \text{(electric dipole interaction),}^{12} \tag{5.33}$$

$$\mathcal{H}_{\rm int} = -\vec{\mu}\cdot\vec{B}(\vec{r}_0,t) \quad \text{(magnetic dipole interaction),} \tag{5.34}$$

$$\mathcal{H}_{\rm int} = +\frac{q}{2}\vec{r}\cdot\vec{r}\cdot\vec{\nabla}_{r_0}\vec{E}(\vec{r}_0,t) \quad \text{(electric quadrupole interaction)} \tag{5.35}$$

where $\vec{d}$, $\vec{\mu}$, etc have to be interpreted as quantum mechanical operators.

The electric dipole interaction is relevant for optical frequency standards like the ones based on the alkaline earth atoms (Section 9.4.4). In the approximation of (5.33) the electric dipole moment $\vec{d} = q\vec{r} = e\vec{r}$ (where the elementary charge $e = 1.602 \times 10^{-19}$ A s must be taken positive) of the atom interacts with the mean electric field at the position $\vec{r}_0$ of the atom.

The Hamiltonian (5.34) refers to the interaction of the magnetic dipole moment $\vec{\mu}$ of the atom with the magnetic $\vec{B}(\vec{r}_0,t)$ field of the electromagnetic radiation. Magnetic dipole transitions between hyperfine transitions are employed in microwave frequency standards as, e.g., the caesium atomic clock (Section 7), the hydrogen maser (Section 8.1), or a large number of trapped ion standards. In optical frequency standards also electric quadrupole transitions (Hg^+ and Yb^+) or even an octupole transition (Yb^+) (Section 10.3.2) are used.

Let us consider in more detail the electric dipole interaction (5.33). The presence of an electric field $\vec{E}$ first of all tends to separate the positive and negative charges in an atom, molecule or ion thus polarising the microscopic particle and changing its energy. The polarisation of the atom is usually described by the expectation value of the (induced) electric dipole operator $\vec{d} = -q\sum_{i=1}^{N}\vec{r}_i$ with the position coordinates $\vec{r}_i$ of the N electrons of the atom taken with the perturbed field-depending state of the atom. In analogy to the classical case, the dipole moment $\vec{d}$ can be thought of as the distance which is now the electronic $\vec{r}$ operator between the charges. For simplicity we assume that the electric dipole is parallel to the electric field direction $\hat{\epsilon}$ of a linearly polarised electromagnetic wave $\vec{E}(\vec{r}_0,t) = E_0\hat{\epsilon}\cos(\omega t)$. Applying the dipole approximation, i.e., neglecting the spatial variation of the electric field across the atom, (5.27) can be re-written as

$$\begin{aligned} i\hbar\frac{dc_1(t)}{dt} &= c_2(t)\left(\int\phi_1^*(\vec{r})\,\vec{d}\cdot\vec{E}\,\phi_2(\vec{r})d^3r\right)e^{-i\omega_0 t}\frac{1}{2}\left[e^{i\omega t}+e^{-i\omega t}\right] \\ &\equiv c_2(t)\frac{\hbar\Omega_R}{2}\left[e^{i(\omega-\omega_0)t}+e^{-i(\omega+\omega_0)t}\right] \end{aligned} \tag{5.36}$$

where

$$\Omega_R = \frac{eE_0}{\hbar}\int\phi_1^*(\vec{r})\,\vec{r}\cdot\vec{\epsilon}\,\phi_2(\vec{r})d^3r \tag{5.37}$$

is called the Rabi frequency. It is always possible to adjust the relative phases of the states ϕ_1 and ϕ_2 such that the matrix elements (5.29) and (5.30), and hence the Rabi frequency, are

[12] The dipole moment is a vector in the direction from $-q$ to $+q$. In an atom the electric field of the electromagnetic wave leads to an induced dipole moment where the electron follows the electric field but the position of the positive cloud of charge, determined mainly by the nucleus, is virtually fixed. If $\vec{r}$ is the vector between the nucleus and the electron it is antiparallel to the dipole moment which leads to the positive sign on the right-hand side of (5.33).

real [136]. Then we obtain

$$i\hbar\frac{dc_2(t)}{dt} = c_1(t)\frac{\hbar\Omega_R}{2}\left[e^{-i(\omega-\omega_0)t} + e^{i(\omega+\omega_0)t}\right]. \tag{5.38}$$

Near resonance ($\omega \approx \omega_0$) the second terms in the square brackets of (5.36) and (5.38) are rapidly oscillating with about twice the frequency of the electromagnetic field, in contrast to the terms containing the detuning

$$\Delta\omega \equiv \omega - \omega_0. \tag{5.39}$$

It can be shown that the rapid oscillation gives rise to only a small frequency shift, referred to as the Bloch–Siegert shift [11,137,138]. Hence, often the terms rapidly oscillating with $\omega+\omega_0$ are neglected, a procedure which is known as the rotating-wave approximation. According to Vanier and Audoin [11] one easily finds that this rapidly oscillating term has an negligible influence if one integrates (5.36) for times short enough that $c_2(t)$ can be regarded as constant, e.g., for $c_2(t) \approx 1$. In this case, $c_1(t)$ is the sum of a term with the resonance denominator $1/\Delta\omega$ and a term with $1/(\omega+\omega_0)$, where the first is much larger than the second one. In the "rotating-wave approximation" we replace (5.36) and (5.38) by

$$\frac{dc_1(t)}{dt} = -ic_2(t)\frac{\Omega_R}{2}e^{i\Delta\omega t} \qquad \text{and} \tag{5.40}$$

$$\frac{dc_2(t)}{dt} = -ic_1(t)\frac{\Omega_R}{2}e^{-i\Delta\omega t}. \tag{5.41}$$

To solve (5.40) and (5.41) we try the ansatz

$$c_1(t) = e^{i\alpha t} \tag{5.42}$$

with $dc_1(t)/dt = i\alpha\exp(i\alpha t)$ and substitute it into (5.40). The result

$$c_2(t) = -\alpha\frac{2}{\Omega_R}e^{i(\alpha-\Delta\omega)t} \tag{5.43}$$

when inserted into (5.41) leads to a quadratic equation $\alpha^2 - \alpha\Delta\omega - \Omega_R^2/4 = 0$ with the two solutions

$$\alpha_{1,2} = \frac{\Delta\omega}{2} \pm \frac{1}{2}\sqrt{\Delta\omega^2 + \Omega_R^2}. \tag{5.44}$$

By denoting

$$\Omega'_R \equiv \sqrt{\Omega_R^2 + \Delta\omega^2} \tag{5.45}$$

equations (5.42) and (5.43) can now be written as

$$c_1(t) = e^{i\frac{\Delta\omega t}{2}}\left[Ae^{i\frac{\Omega'_R t}{2}} + Be^{-i\frac{\Omega'_R t}{2}}\right] \tag{5.46}$$

and

$$c_2(t) = e^{-i\frac{\Delta\omega t}{2}}\left[-A\frac{\Delta\omega+\Omega'_R}{\Omega_R}e^{i\frac{\Omega'_R t}{2}} - B\frac{\Delta\omega-\Omega'_R}{\Omega_R}e^{-i\frac{\Omega'_R t}{2}}\right]. \tag{5.47}$$

The coefficients A and B can be determined from the initial conditions $c_1(t=0)=1$ and $c_2(t=0)=0$ as $1=A+B$ and $A(\Delta\omega+\sqrt{\Delta\omega^2+\Omega_R^2})=-B(\Delta\omega-\sqrt{\Delta\omega^2+\Omega_R^2})$ or

$$A = -\frac{\Delta\omega-\Omega'_R}{2\Omega'_R} \tag{5.48}$$

$$B = \frac{\Delta\omega+\Omega'_R}{2\Omega'_R}. \tag{5.49}$$

Substituting (5.48) and (5.49) into (5.46) and (5.47) we end up with

$$c_1(t) = \left[\cos\frac{\Omega'_R t}{2} - i\frac{\Delta\omega}{\Omega'_R}\sin\frac{\Omega'_R t}{2}\right]\exp\left[i\frac{\Delta\omega}{2}t\right] \tag{5.50}$$

$$c_2(t) = -i\frac{\Omega_R}{\Omega'_R}\sin\frac{\Omega'_R t}{2}\exp\left[-i\frac{\Delta\omega}{2}t\right]. \tag{5.51}$$

The probability of finding the two-level atom in either one of the two states given by $|c_1(t)|^2$ or $|c_2(t)|^2$ oscillates at frequency Ω'_R (see Fig. 5.8). These oscillations are referred to as Rabi oscillations. This angular frequency depends on the detuning of the frequency of the

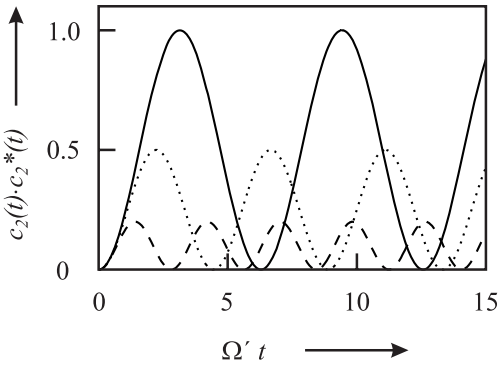


Figure 5.8: The probability $|c_2(t)|^2$ of finding a two-level atom in the excited state calculated from (5.51) for $\Delta\omega=0$ (solid line), $\Delta\omega=\Omega_R$ (dotted line), $\Delta\omega=2\Omega_R$ (dashed line) versus time, showing Rabi oscillations.

radiation from the transition frequency and on the Rabi frequency Ω_R (5.37) defined by the particular interaction such as (5.33) – (5.35). For zero detuning the probability of finding the atom in the excited state increases to unity after some time. A pulse of radiation applied for the corresponding time t is called a π pulse since it changes the phases of the Rabi oscillations (full line in Fig. 5.8) by π. In general, the angle

$$\theta_R = \Omega'_R t \tag{5.52}$$

is referred to as the Rabi angle. As can be seen from Fig. 5.8, with increasing detuning, this oscillation frequency increases. However, at the same time the amplitude of the oscillation decreases.

5.3.2 Optical Bloch Equations

The evolution of the two-level system near resonance described by the Schrödinger equation [13] (5.2) can be visualised by a geometrical picture that has been given by Feynman, Vernon and Hellwarth [139]. To avoid confusion with the convention which led to (5.24) we do not follow their notation but write

$$\psi(\vec{r},t) = C_1(t)\phi_1(\vec{r}) + C_2(t)\phi_2(\vec{r}). \tag{5.53}$$

The equations of motion are then derived from (5.2) and by using (5.25) similarly as has been done in Section 5.3.1 as follows

$$\frac{dC_1(t)}{dt} = +i\frac{\omega_0}{2}C_1(t) - \frac{i}{\hbar}C_2(t)H_{12}(t) \tag{5.54}$$

$$\frac{dC_2(t)}{dt} = -i\frac{\omega_0}{2}C_2(t) - \frac{i}{\hbar}C_1(t)H_{21}(t). \tag{5.55}$$

Here again, $H_{21} \equiv \int \phi_2^* \mathcal{H}_{\rm int}\phi_1 d^3r$, $H_{12} \equiv \int \phi_1^* \mathcal{H}_{\rm int}\phi_2 d^3r = H_{21}^*$ and $\hbar\omega_0 = E_2 - E_1$ holds. The zero of the energy of the two-level system is chosen as $(E_2 + E_1)/2$ leading to $E_2 = \hbar\omega_0/2$ and $E_1 = -\hbar\omega_0/2$. The difference between the different time-dependent coefficients is such that $c_1(t)$ and $c_2(t)$ of (5.27) and (5.28) in the interaction picture show a slowly varying time dependence determined by the interaction energy $\mathcal{H}_{\rm int}$ alone, whereas the more rapidly oscillating $C_1(t)$ and $C_2(t)$ of (5.54) and (5.55) in the Schrödinger picture are determined by the total Hamiltonian. Nevertheless, $C_1C_1^* = c_1c_1^*$ and $C_2C_2^* = c_2c_2^*$ holds.

Feynman *et al.* used the three real functions depending on $C_1(t)$ and $C_2(t)$

$$R_1'(t) \equiv C_2(t)C_1^*(t) + C_2^*(t)C_1(t) \tag{5.56}$$

$$R_2'(t) \equiv i\left[C_2(t)C_1^*(t) - C_2^*(t)C_1(t)\right] \tag{5.57}$$

$$R_3'(t) \equiv C_2(t)C_2^*(t) - C_1(t)C_1^*(t) \tag{5.58}$$

to define a vector $\vec{R}'(t) = (R_1'(t), R_2'(t), R_3'(t))$. $\vec{R}'(t)$ is commonly referred to as a fictitious spin vector or pseudo-spin vector whose name will be justified below. Equating (5.56) to (5.58) and (5.32) one finds

$$R'^2_1(t) + R'^2_2(t) + R'^2_3(t) = \left[C_2(t)C_2^*(t) + C_1(t)C_1^*(t)\right]^2 = \left(|c_2(t)|^2 + |c_1(t)|^2\right)^2 = 1 \tag{5.59}$$

which means that the length of the vector $\vec{R}'(t)$ is constant, i.e., the tip of $R'(t)$ traces an orbit on a unit sphere referred to as a "Bloch sphere". To find out the significance of the three components of the pseudo-spin vector $\vec{R}'(t)$ we determine the dynamics of the pseudo-spin vector by deriving the motional equations from $d\vec{R}'(t)/dt$.

As an example we calculate the vector component $dR_1'(t)/dt$ from (5.56) as

$$\frac{dR_1'(t)}{dt} = \frac{dC_2(t)}{dt}C_1^*(t) + C_2\frac{dC_1^*(t)}{dt} + \frac{dC_2^*(t)}{dt}C_1(t) + C_2^*(t)\frac{dC_1(t)}{dt} \tag{5.60}$$

[13] Equivalently, the Heisenberg picture can be used [138].

and find by using (5.54) and (5.55) together with (5.57) and (5.58)

$$\begin{aligned}\frac{dR_1'(t)}{dt} &= \frac{1}{i\hbar}\frac{\hbar\omega_0}{2}C_2C_1^* + \frac{1}{i\hbar}C_1C_1^*H_{21} + \frac{1}{i\hbar}\frac{\hbar\omega_0}{2}C_2C_1^* - \frac{1}{i\hbar}C_2C_2^*H_{12}^* \\ &- \frac{1}{i\hbar}\frac{\hbar\omega_0}{2}C_2^*C_1 - \frac{1}{i\hbar}C_1C_1^*H_{21}^* - \frac{1}{i\hbar}\frac{\hbar\omega_0}{2}C_2^*C_1 + \frac{1}{i\hbar}C_2C_2^*H_{12} \\ &= \frac{2\omega_0}{2i}\left(C_2C_1^* - C_2^*C_1\right) + \frac{1}{i\hbar}C_1C_1^*\left(H_{21} - H_{21}^*\right) + \frac{1}{i\hbar}C_2C_2^*\left(H_{12} - H_{12}^*\right) \\ &= -\omega_0 R_2'(t) - \frac{2}{\hbar}\Im\mathrm{m}\left(H_{21}\right)R_3'(t). \qquad (5.61)\end{aligned}$$

Similarly, we find the equations for the components $\dot{R}_2(t)$ and $\dot{R}_3(t)$. The complete set of equations given in the following is referred to as the optical Bloch equations

$$\frac{dR_1'(t)}{dt} = -\omega_0 R_2'(t) - \frac{2}{\hbar}\Im\mathrm{m}\left(H_{21}\right)R_3'(t) \qquad (5.62)$$

$$\frac{dR_2'(t)}{dt} = +\omega_0 R_1'(t) - \frac{2}{\hbar}\Re\mathrm{e}\left(H_{21}\right)R_3'(t) \qquad (5.63)$$

$$\frac{dR_3'(t)}{dt} = +\frac{2}{\hbar}\Re\mathrm{e}\left(H_{21}\right)R_2'(t) + \frac{2}{\hbar}\Im\mathrm{m}\left(H_{21}\right)R_1'(t). \qquad (5.64)$$

This set of equations can be written in a more compact form as

$$\frac{d\vec{R}'(t)}{dt} = \vec{\Omega}' \times \vec{R}'(t) \qquad (5.65)$$

where $\vec{\Omega}'$ represents a "torque" vector [14] with the three real components [15]

$$\vec{\Omega}' \equiv \left(\frac{2}{\hbar}\Re\mathrm{e}\left(H_{21}\right), -\frac{2}{\hbar}\Im\mathrm{m}\left(H_{21}\right), \omega_0\right). \qquad (5.66)$$

Equation (5.65) resembles the equation used to describe the precession of a spinning top under the action of a torque (Fig. 5.9) or the precession of a spin-1/2 particle with a magnetic moment gyrating in a magnetic field which led one to adopt the name "pseudo-spin vector". The optical Bloch equations describe the coupling of a radiation field to a two-level system. The R_1' and R_2' components of the Bloch vector correspond to the real and imaginary part of the polarisation of the atom, whereas the R_3' component gives the difference in the probabilities to find the two-level system in the upper state ϕ_2 ($|e\rangle$, $|2\rangle$) and in the lower state ϕ_1 ($|g\rangle$, $|1\rangle$), i.e., it represents the inversion of the two-level system. For an atom in the ground state ($|1\rangle$) or in the excited state ($|2\rangle$) $\vec{R}'$ points downward or upward, respectively.

The simple picture of $\vec{R}'$ precessing around $\vec{\Omega}'$ is adequate only in cases where $\vec{\Omega}'$ varies slowly during the time needed for $\vec{R}'$ to precess around $\vec{\Omega}'$. Near resonance $\omega \approx \omega_0$ holds and $\vec{\Omega}'(t)$ has components where the time dependence (of the interaction operator) is determined by the frequency of the electromagnetic field ω. As an example consider the case of a π pulse

[14] The vector $\vec{\Omega}'$ must not be confused with the Rabi frequency Ω_R'.

[15] If we had chosen H_{12} rather than H_{21} in (5.62) – (5.64) then the second component of the "torque" vector would read $+2/\hbar\Im\mathrm{m}\,H_{12}$.

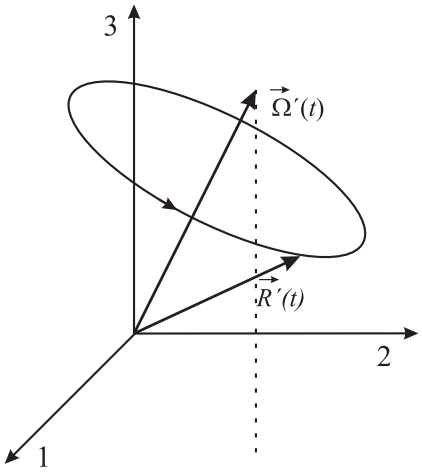


Figure 5.9: The pseudo-spin vector $\vec{R}'$ precesses around the $\vec{\Omega}'$ vector.

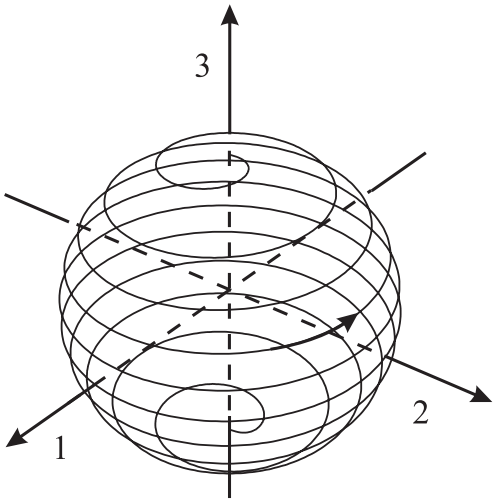


Figure 5.10: Evolution of the R' vector on the Bloch sphere for a resonant π pulse applied to an atom in the ground state.

where the interaction is constant for a time τ such that $\Omega_R \tau = \pi$ (Fig. 5.10). In microwave (optical) frequency standards the time τ is typically of the order of microseconds or longer and hence the pseudo-spin vector spirals at least a thousand (thousand millions) times around the Bloch sphere. Thus often a transformation is applied where a reference frame with the new coordinates u, v, w is used, that rotates at the frequency of the electromagnetic wave ω around the 3 coordinate, with the w coordinate being parallel to the 3 coordinate [133, 138–140]. Before performing this transformation we again specify the particular interaction as the electric dipole interaction where $\mathcal{H}_{12} = \mathcal{H}_{21}^* = -\vec{d} \cdot \vec{E}$ (see (5.33)) and the dipole moment $d_{12} = d_r + i d_i$ in general is a complex vector. With the same argument as before the phases of the states ψ_1 and ψ_2 can be adjusted in such a way that $d_i = 0$. Thus we can write (5.62) – (5.64) as

$$\frac{dR_1'(t)}{dt} = -\omega_0 R_2'(t) \tag{5.67}$$

$$\frac{dR_2'(t)}{dt} = +\omega_0 R_1'(t) + \frac{2d_r}{\hbar} E_0 \cos\omega t \, R_3'(t) \tag{5.68}$$

$$\frac{dR_3'(t)}{dt} = -\frac{2d_r}{\hbar} E_0 \cos\omega t \, R_2'(t). \tag{5.69}$$

which we now transform into the rotating system by means of the transformations

$$R_1'(t) = u\cos\omega t - v\sin\omega t \tag{5.70}$$

$$R_2'(t) = u\sin\omega t + v\cos\omega t \tag{5.71}$$

$$R_3'(t) = w. \tag{5.72}$$

Insertion of, e.g., (5.71) into (5.67) and equating the result with the derivative of (5.70) leads to

$$\dot{u}\cos\omega t - \dot{v}\sin\omega t = (\omega-\omega_0)\,u\sin\omega t + (\omega-\omega_0)\,v\cos\omega t \tag{5.73}$$

$$\dot{u}\sin\omega t + \dot{v}\cos\omega t = -(\omega-\omega_0)\,u\cos\omega t + (\omega-\omega_0)\,v\sin\omega t + \frac{2d_r}{\hbar}E_0\cos\omega t\,w \tag{5.74}$$

$$\dot{w} = -\frac{2d_r}{\hbar}E_0\cos\omega t\sin\omega t\,u - \frac{2d_r}{\hbar}E_0\cos^2\omega t\,v. \tag{5.75}$$

Adding or subtracting (5.73) and (5.74) after suitable multiplication by $\cos\omega t$ and $\sin\omega t$, leads to the optical Bloch equations in the rotating frame

$$\dot{u} = (\omega-\omega_0)v + \frac{d_r}{\hbar}E_0\sin 2\omega t\,w \tag{5.76}$$

$$\dot{v} = -(\omega-\omega_0)u + \frac{d_r}{\hbar}E_0(1+\cos 2\omega t)\,w \tag{5.77}$$

$$\dot{w} = -\frac{d_r}{\hbar}E_0\sin 2\omega t\,u - \frac{d_r}{\hbar}E_0(1+\cos 2\omega t)\,v. \tag{5.78}$$

Hence the transformation to the rotating frame leads to terms slowly varying with the detuning as $(\omega-\omega_0)t$ and rapidly oscillating terms oscillating with $2\omega t$. In microwave and optical frequency standards where the pulse duration τ is large compared to $1/\omega$ these terms in general have only a small influence (the so-called Bloch–Siegert shift, [11, 137, 138]). Applying the rotating-wave approximation where the terms oscillating with $2\omega t$ are neglected and using the Rabi frequency Ω_R (see (5.37)) we find the optical Bloch equations in the rotating frame and in the rotating-wave approximation as follows

$$\dot{u} = (\omega-\omega_0)v \tag{5.79}$$

$$\dot{v} = -(\omega-\omega_0)u + \Omega_R\,w \tag{5.80}$$

$$\dot{w} = -\Omega_R\,v. \tag{5.81}$$

The transformation to the rotating frame again leaves the length of the vector constant and again the pseudo-spin vector $\vec{R} = (u, v, w)$ traces out an orbit on the Bloch sphere. Similarly as before (see (5.65)) these equations can be written as a single vector equation

$$\frac{d\vec{R}(t)}{dt} = \vec{\Omega}\times\vec{R}(t) \tag{5.82}$$

with the "torque" vector

$$\vec{\Omega} = (-\Omega_R, 0, \omega_0-\omega) \tag{5.83}$$

where the third component of $\vec{\Omega}$ now is the negative detuning $-(\omega-\omega_0)$.

The picture of the pseudo-spin vector on the Bloch sphere is particularly useful in the context of pulsed excitations, e.g., in the Ramsey technique (Section 6.6) and it will be used in more detail in the next chapters. In Fig. 5.11 we therefore consider two particular cases which represent "building blocks" for such sequences.

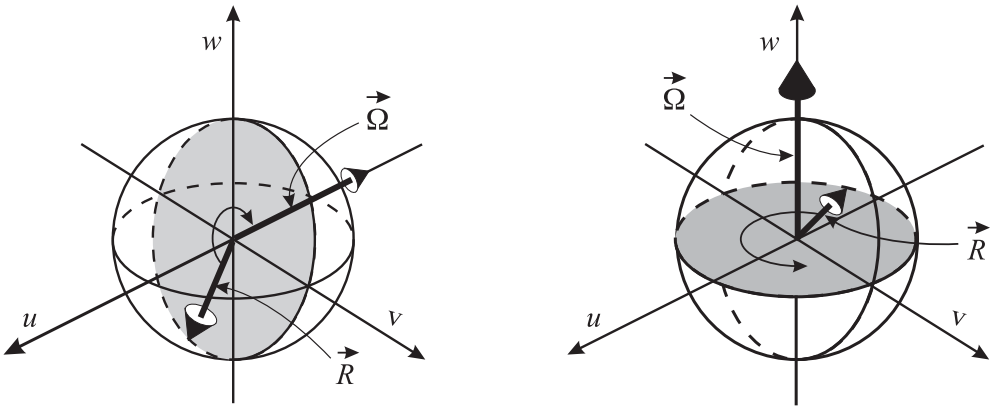


Figure 5.11: a) For zero detuning, the pseudo-spin vector precesses in the $v - w$ plane. b) Evolution of the pseudo-spin vector representing a pure coherence of the two-level system.

As a first case (Fig. 5.11 a) consider the interaction of a monochromatic field with a frequency that coincides with the atomic transition frequency. As the detuning is zero, the "torque vector" $\vec{\Omega}$ points along the $-u$ axis (see (5.83)) and the pseudo-spin vector $\vec{R}$ precesses around this axis. To find the evolution of the pseudo-spin vector in general one uses (5.79) to (5.81). Since at the beginning the two-level system is in the ground state the pseudo-spin vector points to the south pole of the Bloch sphere with $u = v = 0$ and $w = -1$. Hence, $\dot{u} = 0$ (see (5.79)), $\dot{w} = 0$ (see (5.81)), $\dot{v} = -\Omega_R$ (see (5.81)) and consequently the pseudo-spin vector acquires a $-v$ component. On its path on the Bloch sphere the pseudo-spin vector passes the north pole and later the south pole again and again as long as the interaction is switched on. As the third component of $\vec{R}$ represents the inversion of the two-level system we recover here the on-resonance Rabi oscillations of Fig. 5.8.

Next, we consider the case (Fig. 5.11 b) where the "torque vector" has only a w component, i.e., there is a detuning $\omega - \omega_0 \neq 0$ but no field ($\Omega_1 = 0$). Such a case can be encountered after a $\pi/2$ pulse resulting from an infinitely short interaction of the atom with a detuned field. The Bloch vector precesses now in a the $u - v$ plane and its third component does not change with time. The positive or negative detuning defines if the pseudo-spin vector rotates clockwise or counter-clockwise.

In the more general case where the "torque vector" has a u and a w component, i.e., during interaction of the two-level system with a detuned radiation field, the Bloch vector acquires a non-zero u component. Consequently, the Bloch vector never can reach the north pole of the Bloch sphere again and hence the Rabi oscillations with finite detuning never lead to full inversion (Fig. 5.8).

5.3.2.1 Density Operator

Up to now we have described a single two-level atom by a state vector or a wave function (5.24) that contains all possible information about the system. Often, however, this information is not available, e.g., if spontaneous decay, described by a decay rate γ from the upper level to the lower level, has to be taken into account. In such a case one does not know the state of the system but rather knows the probabilities of finding the system in the particular states and the system is described by the density operator. The density operator is defined as the sum of the projectors onto the possible state vectors $|\psi_i\rangle$ each properly weighted by a classical probability P_i

$$\rho \equiv \sum_i P_i |\psi_i\rangle\langle\psi_i| \tag{5.84}$$

(see e.g. [133, 135]). For a two-level system where the state of the system is given by $|\psi\rangle = C_1|1\rangle + C_2|2\rangle$ the matrix elements of the density operator

$$\rho = \begin{pmatrix} C_1C_1^* & C_1C_2^* \\ C_2C_1^* & C_2C_2^* \end{pmatrix} = \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix} \tag{5.85}$$

can be written as $\rho_{ij} = \langle j|\rho|i\rangle$ where

$\rho_{11} = C_1C_1^*$ is the probability to find the system in the lower level,
$\rho_{22} = C_2C_2^*$ is the probability to find the system in the upper level,
$\rho_{12} = C_1C_2^* = \rho_{21}^*$ is the so-called coherence.

The diagonal elements of the density matrix represent the population of the atomic states. Similar to the R_1 and R_2 components of the Bloch vector, the non-diagonal elements represent the induced polarisation which is responsible for the absorption and dispersion. The relaxation of terms resulting from spontaneous emission, collisions, and other damping mechanisms are included in a phenomenological way by decay rates γ. A complex non-diagonal element describes the coherent superposition of two states and is therefore referred to as a coherence.

The pseudo-spin vector on the Bloch sphere can be constructed by using the elements of the density matrix ρ in an equivalent way to (5.56) – (5.58) [133].

5.3.3 Three-level Systems

The simple two-level picture used so far is capable of describing a large body of the effects associated with the absorption and emission of photons by quantum mechanical systems and hence the physics of frequency standards. There are, however, a number of effects, such as optical pumping [141], the occurrence of dark states, or coherent population trapping [142] that are based on the interaction of radiation fields with absorbers having more than two levels. Most of these effects can be visualised in a three-level system in a simplified way (Fig. 5.12).

5.3.3.1 Optical Pumping

Optical pumping, first described by Kastler [141], can occur, e.g., in multi-level systems. Consider an ensemble of atoms described by the simplified energy level scheme of Fig. 5.12 a)

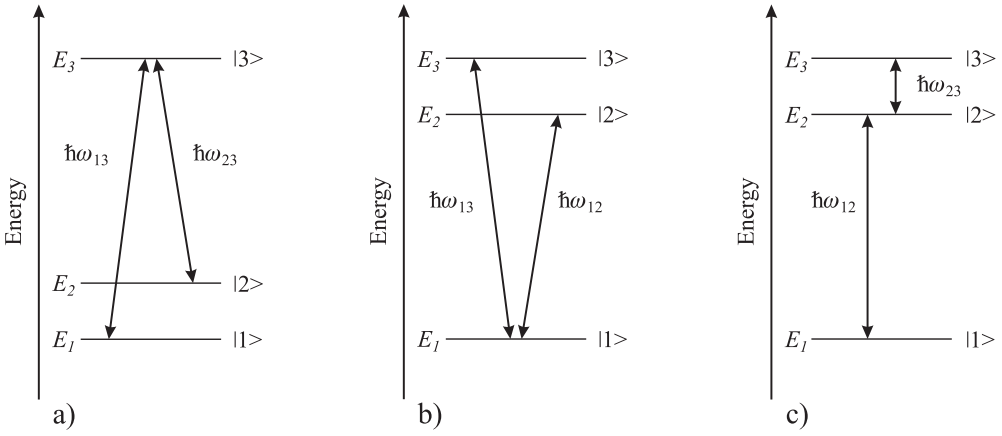


Figure 5.12: Three-level systems in a) Λ configuration, b) V configuration and c) cascade configuration.

where the atom, if excited in the state $|3\rangle$, can decay by allowed transitions with energies $\hbar\omega_{13}$ and $\hbar\omega_{23}$ to the two states $|1\rangle$ and $|2\rangle$. If a monochromatic light source is tuned, e.g., to the transition $\hbar\omega_{23}$, the atoms spontaneously decaying into the state $|2\rangle$ will be re-excited, whereas the atoms in the state $|1\rangle$ remain unaffected. After a number of excitation–emission cycles, virtually all atoms will be found in the $|1\rangle$ state. For frequency standards this optical pumping is important in several aspects. First, it is used to prepare multi-level atoms in a well-defined state, e.g., in the optically pumped Cs clock, where the $F = 4$ to $F' = 3$ transition ($\lambda = 852\,\mathrm{nm}$) (see Fig. 7.8) can be used to redistribute the atomic population in the ground states. Second, optical pumping may terminate laser cooling, where, e.g., in the case of the alkaline atoms its influence has to be counteracted by an additional laser that re-pumps the atoms.

5.3.3.2 Coherent Population Trapping

Very peculiar features can be observed when two coherent (laser) fields interact with a three-level system, e.g., the Λ system of Fig. 5.12 a), whose angular frequencies ω_{L1} and ω_{L2} are near the two-photon Raman resonance condition

$$\hbar\omega_{13} - \hbar\omega_{23} = E_2 - E_1. \tag{5.86}$$

The detunings

$$\delta\omega_{L1} = \omega_{L1} - \omega_{13}, \tag{5.87}$$

$$\delta\omega_{L2} = \omega_{L2} - \omega_{23} \tag{5.88}$$

differ from zero. The features include the occurrence of so-called coherent population trapping (CPT) or dark resonances. They can be calculated, e.g., by formulating the optical Bloch equations for the density matrix elements of a three-level system [142–144] and considering the stationary solutions $\dot{\rho} = 0$. As an example consider the three elements of the density matrix ρ_{11}, ρ_{22}, and ρ_{33} (Fig. 5.13) representing the populations of the three states $|1\rangle$, $|2\rangle$, and $|3\rangle$

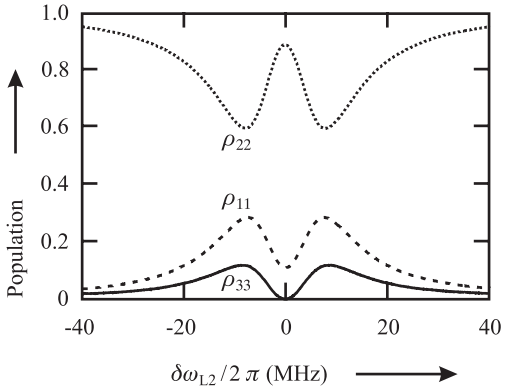


Figure 5.13: Diagonal elements of the density matrix describing the steady-state populations in a closed three-level Λ system according to Fig. 5.12 a), as a function of the detuning of the second laser $\delta\omega_{L1}$ and the first laser on resonance ($\delta\omega_{L1} = 0$). The transition rates and the Rabi frequencies were chosen as $\Gamma_{13} = \Omega_{\mathrm{R}\,13} = 2\pi \times 15\,\mathrm{MHz}$, $\Gamma_{23} = \Omega_{\mathrm{R}\,23} = 2\pi \times 5\,\mathrm{MHz}$, and $\Gamma_{23} = 0$ (similarly to [145]).

of Fig. 5.12 a), respectively. Away from the two-photon resonance, the population ρ_{33} in the excited state $|3\rangle$ depends Lorentzian-like on the detuning of the laser frequency. However, in a narrow region close to resonance the population is decreased and vanishes exactly when both lasers are tuned to resonance. As a result the population is trapped in the two lower states by this coherence effect, and hence it is termed "coherent population trapping" (CPT). The two dipoles associated with the transitions $|1\rangle \rightarrow |3\rangle$ and $|2\rangle \rightarrow |3\rangle$ are coupled via the non-diagonal elements of the density matrix in the optical Bloch equations. It can be shown [142] that exactly at resonance the coherent superposition of both dipoles leads to destructive interference for the transition to the state $|3\rangle$. Since due to the missing population in the state $|3\rangle$ at resonance the fluorescence is suppressed, the term "dark resonance" was coined. Far from resonance the distribution of the population distribution in the states $|1\rangle$ and $|2\rangle$ is determined by the effect of optical pumping and consequently depends on the effective Rabi frequencies of the transitions driven by the respective lasers (see Fig. 5.13). The steady-state populations in the states $|1\rangle$ and $|2\rangle$ depend on the ratio of the associated Rabi frequencies $\Omega_{\mathrm{R}\,13}/\Omega_{\mathrm{R}\,23}$ and on ratio of the decay rates Γ_{13} and γ_{23}. If in Fig. 5.13 the first laser were also detuned from the transition ($\delta\omega_{L1} \neq 0$) the dark resonance would be shifted by the amount given by the two-photon Raman resonance condition.

Dark resonances and coherent population trapping find increasing application in frequency standards since they allow one to use a very small linewidth in compact devices (Section 8.2). The width of the dark resonance is ultimately limited by the lifetime of the coherence between the two lowest states $|1\rangle$ and $|2\rangle$. If these states result, e.g., from the hyperfine split ground state of an alkali atom such as caesium or rubidium, these states are coupled by weak magnetic dipole transitions that would have an extremely long spontaneous lifetime. Consequently, the observed linewidth is mainly determined by the phase stability of the laser fields and other experimental parameters.

5.4 Line Shifts and Line Broadening

The transitions in atoms, ions, or molecules considered so far are widely used to stabilise the frequencies of external oscillators. Any effects that shift the transition frequencies in an uncontrolled way may limit the performance of such frequency standards. Hence, the knowledge of the origins, magnitudes, and properties of these shifts is essential to reduce and control them in order to design and to operate frequency standards at optimum performance. In this sense all microscopic quantum systems of a particular species behave in the same way only to the extent that one is able to create the same conditions for all of these systems. Frequency shifts of a transition line immediately lead to line broadening if either a particular perturbation to the quantum system is not stable during the interaction or if the interaction includes many systems subject to different perturbations. If the interaction takes place with an ensemble of particles it is useful to distinguish two different categories of line broadening. One refers to a homogeneous broadening if all particles show the same individually broadened line. If, on the other hand, the particles of an ensemble that is interrogated experience different perturbations of their transition frequencies, one refers to an inhomogeneous broadening of the line.

5.4.1 Interaction Time Broadening

The natural linewidth resulting from the finite lifetime of the excited state considered in 2.1.2.3 represents an example of a homogeneous broadening since it is common to all atoms. Similarly, the finite interaction time of an interrogating field with the absorber leads to a finite homogeneous width of the corresponding spectrum in the frequency domain. In the following we consider two special cases where the interaction is switched on and off either abruptly or smoothly and calculate the corresponding frequency spectrum.

If the interaction of the field with the atom is switched on for a finite duration τ_p the field spectrum can be described by

$$f(t) = \begin{cases} A_0 \cos\omega_0 t & \text{for} \quad -\tau_P/2 < t < \tau_p/2 \\ 0 & \text{elsewhere.} \end{cases}$$

The frequency-domain spectrum can be calculated as the inverse Fourier transformation of $f(t)$ according to (2.19) as

$$\begin{aligned} F(\omega) &= \frac{A_0}{2} \int_{-\infty}^{\infty} \{\exp(i\omega_0 t) + \exp(-i\omega_0 t)\} \exp(-i\omega t) dt \qquad (5.89) \\ &= \frac{A_0}{2} \int_{-\tau_p/2}^{\tau_p/2} \exp[i(\omega_0 - \omega)t] dt + \frac{A_0}{2} \int_{-\tau_p/2}^{\tau_p/2} \exp[-i(\omega_0 + \omega)t] dt. \end{aligned}$$

By writing the complex exponential as cosine and sine terms, one finds

$$F(\omega) = \frac{A_0}{2} \int_{-\tau_p/2}^{\tau_p/2} \cos[(\omega - \omega_0)t] dt + \frac{A_0}{2} \int_{-\tau_p/2}^{\tau_p/2} \cos[(\omega + \omega_0)t] dt \qquad (5.90)$$

where the integration of the odd (sine) function cancels. From the two terms in (5.90) $F(\omega) \equiv F_+(\omega-\omega_0) + F_-(\omega+\omega_0)$ the second one $F_-(\omega+\omega_0)$ is the one that takes into account the negative "mirror frequencies". Evaluation of the first integral leads to

$$\begin{aligned} F_+(\omega-\omega_0) &= \frac{A_0}{2}\frac{1}{\omega-\omega_0}\left[\sin\frac{(\omega-\omega_0)\tau_p}{2} - \sin\frac{-(\omega-\omega_0)\tau_p}{2}\right] \\ &= \frac{A_0\tau_p}{2}\frac{\sin[(\omega-\omega_0)\tau_p/2]}{(\omega-\omega_0)\tau_p/2}. \end{aligned} \tag{5.91}$$

The power spectrum [16] $|F_+(\omega-\omega_0)|^2$ (Fig. 5.14) shows side lobes besides the main peak at ω_0. These side lobes result from the steep slopes when the interaction is switched on and off and from the corresponding high-frequency Fourier components necessary to reproduce these steep slopes. To determine the full width at half maximum $\Delta\omega_p$ of the main peak of Fig. 5.14

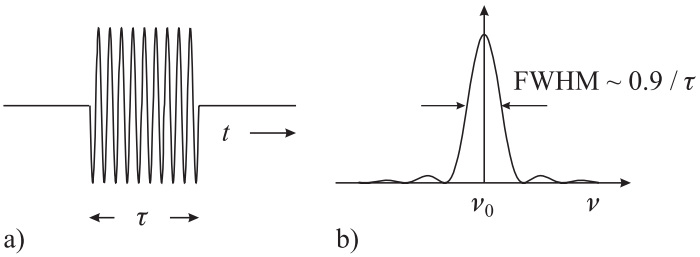


Figure 5.14: a) Square pulse of finite length τ. b) Corresponding power spectrum.

we set

$$\frac{1}{2} = \frac{\sin^2[(\omega_{1/2}-\omega_0)\tau_p/2]}{[(\omega_{1/2}-\omega_0)\tau_p/2]^2}. \tag{5.92}$$

Solving this equation is equivalent to solving $1/\sqrt{2}x = \sin x$ with $x \equiv (\omega_{1/2}-\omega_0)\tau_p/2 = 1.3916$. From the full width at half maximum $\Delta\omega_p = 2(\omega_{1/2}-\omega_0) = 4\times 1.3916/\tau_p$ of the pulse we find $\Delta\nu_p \times \tau_p = 0.8859$. This is a special case of the time–bandwidth product relating the temporal width τ_p of a pulse [17] to the width $\Delta\nu_p$ of the corresponding power spectrum $F(\nu)$.

Often, the interaction is "switched on and off" smoothly. This situation occurs, e.g., when a moving atom crosses a laser beam with a Gaussian profile. In this case the pulse can be described by

$$f(t) = A_0 \exp\left(-\frac{t^2}{2\sigma^2}\right)\cos\omega_0 t \tag{5.93}$$

[16] Note that the total power is given by $|F_+(\omega-\omega_0) + F_-(\omega+\omega_0)|^2$.

[17] For short laser pulses in general the time–bandwidth product is given for the square of the field amplitude rather than the amplitude itself.

where A_0 denotes the amplitude at $t = 0$ (Fig. 5.15 a). The spectrum in the frequency domain is obtained from the inverse Fourier transformation

$$\begin{aligned} F(\omega) &= \frac{A_0}{2}\int_{-\infty}^{\infty}\exp\left(\frac{-t^2}{2\sigma^2}\right)[\exp(i\omega_0 t)+\exp(-i\omega_0 t)]\exp(-i\omega t)dt \qquad (5.94)\\ &= \frac{A_0}{2}\int_{-\infty}^{\infty}\exp\left(\frac{-t^2}{2\sigma^2}\right)\exp[i(\omega_0-\omega)t]dt \\ &\quad +\frac{A_0}{2}\int_{-\infty}^{\infty}\exp\left(\frac{-t^2}{2\sigma^2}\right)\exp[-i(\omega_0+\omega)t]dt. \end{aligned}$$

To evaluate the left-hand term of (5.94) $F(\omega) \equiv F_+(\omega-\omega_0)+F_-(\omega+\omega_0)$ we complete the square of the exponent

$$\begin{aligned} F_+(\omega) &= \frac{A_0}{2}\int_{-\infty}^{\infty}\exp\left\{-\left[\frac{t^2}{2\sigma^2}-i(\omega_0-\omega)t\right]\right\}dt \\ &= \frac{A_0}{2}\int_{-\infty}^{\infty}\exp\left[-\left(\frac{t}{\sqrt{2}\sigma}-i\frac{(\omega_0-\omega)\sqrt{2}\sigma}{2}\right)^2-\frac{(\omega_0-\omega)^2\sigma^2}{2}\right]dt \\ &= \frac{A_0}{2}\exp\left[-\frac{(\omega_0-\omega)^2\sigma^2}{2}\right]\int_{-\infty}^{\infty}\exp\left[-\left(\frac{t}{\sqrt{2}\sigma}-i\frac{(\omega_0-\omega)\sigma}{\sqrt{2}}\right)^2\right]dt. \qquad (5.95) \end{aligned}$$

The value of the definite integral can be looked up in tables after simplifying it by taking $\frac{t}{\sqrt{2}\sigma}-i\frac{(\omega_0-\omega)\sigma}{\sqrt{2}} \equiv \beta$ as

$$\int_{-\infty}^{\infty}\exp\left[-\left(\frac{t}{\sqrt{2}\sigma}-i\frac{(\omega_0-\omega)\sigma}{\sqrt{2}}\right)^2\right]dt \equiv \sqrt{2}\sigma\int_{-\infty}^{\infty}\exp(-\beta^2)d\beta = \sqrt{2}\sigma\sqrt{\pi}. \qquad (5.96)$$

Hence, the Fourier transform of the Gaussian (5.93)

$$F_+(\omega) = \frac{A_0\sqrt{2\pi}\sigma}{2}\exp\left[-\frac{(\omega_0-\omega)^2\sigma^2}{2}\right] \qquad (5.97)$$

is again a Gaussian function of width $\sigma' = 1/\sigma$ (Fig. 5.15). The $(1/e)$ width of the squared amplitude of both pulses (in the time domain and in the frequency domain) is related to the respective halfwidth by $2\sqrt{\ln 2}$. Hence, the corresponding time–bandwidth product is given by $4\ln 2/2\pi \times \Delta\nu_p \times \tau_p = 0.4413$. The time–bandwidth product for the most common pulses (see e.g. Table 5.5) are of the order of unity and, hence,

$$\Delta\nu_p \times \tau_p \lesssim 1 \qquad (5.98)$$

holds.

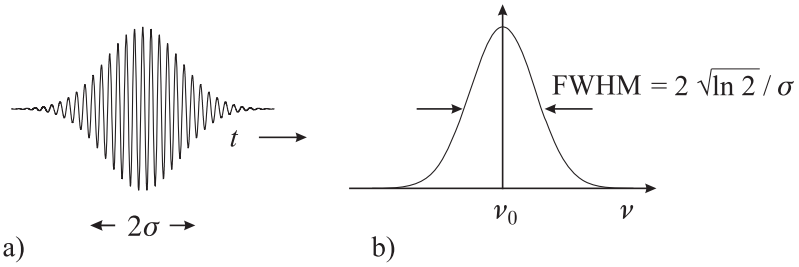


Figure 5.15: a) A Gaussian pulse of length σ. b) The corresponding power spectrum (b)

Table 5.5: Time–bandwidth product of different pulse shapes. τ_p: pulse length (FWHM) of $[f(t)]^2$; $\Delta\nu_p$: bandwidth (FWHM) of $[\mathcal{F}(f(t))]^2$.

Shape	$I(t)$	$\Delta\nu_P \times \tau_p$
Rectangular pulse	$\begin{cases} 1, \lvert t\rvert \le \tau_p/2 \\ 0, \lvert t\rvert > \tau_p/2 \end{cases}$	0.8859
Gaussian	$e^{-(4\ln 2)t^2/\tau_p^2}$	0.4413
Secans Hyperbolicus	$\mathrm{sech}^2\left(1.7627\dfrac{t}{\tau_p}\right)$	0.3148
Lorentzian	$\dfrac{1}{1+\left(\frac{2t}{\tau_p}\right)^2}$	0.2206

5.4.2 Doppler Effect and Recoil Effect

In contrast to a single, ideally isolated two-level system at rest, real quantum systems such as, e.g., caesium atoms in an atomic clock or iodine molecules in an absorption cell used to stabilise a laser are moving during their interaction with the radiation field. As a consequence of this movement the angular frequency of the absorbed or emitted photon ω differs from Bohr's angular frequency $\omega_0 = (E_2 - E_1)/\hbar$ of the unperturbed absorber. We are going to calculate this frequency difference by assuming that the energy and momentum of the system will be conserved during the absorption or emission of a photon. Energy conservation requires

$$\frac{p_1^2}{2m} + E_1 + \hbar\omega = \frac{p_2^2}{2m} + E_2. \tag{5.99}$$

Momentum conservation for absorption and emission of a photon requires

$$\vec{p}_1 + \hbar\vec{k} = \vec{p}_2 \tag{5.100}$$

$\hbar\omega$ and $\hbar\vec{k}$ are the energy and momentum, respectively, of a photon corresponding to the wave number $k = \omega/c$ being absorbed or emitted by a two-level quantum system having the energies E_1 and E_2 with $E_2 > E_1$. The corresponding momenta of the two-level system are $\vec{p}_1$ and $\vec{p}_2$. Rearranging (5.99) and inserting (5.100) into (5.99) leads to

$$\begin{aligned}\hbar\omega &= E_2 - E_1 + \frac{p_2^2}{2m} - \frac{p_1^2}{2m} \\ &= E_2 - E_1 + \frac{p_1^2}{2m} + \frac{\vec{p}_1 \cdot \hbar\vec{k}}{m} + \frac{\hbar^2 k^2}{2m} - \frac{p_1^2}{2m} \\ &= \hbar\omega_0 + \vec{v}_1 \cdot \hbar\vec{k} + \frac{\hbar^2 k^2}{2m}. \qquad \text{(absorption)} \end{aligned} \tag{5.101}$$

The energy $\hbar\omega$ of the absorbed photon differs from the Bohr energy $\hbar\omega_0$ by the terms $\vec{v}_1 \cdot \hbar\vec{k}$ and $\frac{\hbar^2 k^2}{2m} = \frac{(\hbar\omega)^2}{2mc^2}$. The first term varies linearly with the velocity of the absorber. It is referred to as the (linear) Doppler effect. The second term results from the recoil of the absorber owing to the photon momentum transferred to the atom (Eq. (5.100)).

In the case of the emission of a photon the atom is initially in the state $|2\rangle$ with momentum $\vec{p}_2$. Hence we derive in a similar way

$$\hbar\omega = \hbar\omega_0 + \vec{v}_2 \cdot \hbar\vec{k} - \frac{\hbar^2 k^2}{2m} \qquad \text{(emission)} \tag{5.102}$$

where the sign of the recoil term is changed. From (5.101) and (5.102) one finds that the energy difference between the photon absorbed by an atom which is initially at rest ($v_1 = 0$) and the photon emitted by an atom initially at rest ($v_2 = 0$) is

$$\hbar\Delta\omega = \frac{(\hbar\omega)^2}{mc^2}. \tag{5.103}$$

As the recoil term depends on ω^2 its influence in general can be neglected in microwave frequency standards but becomes very important in the frequency standards in the optical regime where its magnitude is about ten orders of magnitude higher.

In the following we use the relativistic relationship $E = \sqrt{p^2c^2 + m_0^2c^4}$ rather than the classical one $E = p^2/(2m)$ between the energy and the momentum of a free particle and we will find additional terms occurring in (5.101) and (5.102) as seen in (5.108). Relativistic energy conservation requires that

$$\begin{aligned}\hbar\omega &= \sqrt{p_2^2c^2 + (m_0c^2 + \hbar\omega_0)^2} - \sqrt{p_1^2c^2 + m_0^2c^4} \\ &= (m_0c^2 + \hbar\omega_0)\sqrt{1 + \frac{p_2^2c^2}{(m_0c^2 + \hbar\omega_0)^2}} - m_0c^2\sqrt{1 + \frac{p_1^2c^2}{m_0^2c^4}}. \end{aligned} \tag{5.104}$$

By expanding the square roots and neglecting all terms in powers higher than the square of v/c we find

$$\begin{aligned}\hbar\omega &= (m_0c^2+\hbar\omega_0)\left(1+\frac{1}{2}\frac{p_2^2c^2}{(m_0c^2+\hbar\omega_0)^2}-\cdots\right)-m_0c^2\left(1+\frac{1}{2}\frac{p_1^2c^2}{m_0^2c^4}-\cdots\right)\\ &\approx m_0c^2+\hbar\omega_0+\frac{p_2^2c^2}{2(m_0c^2+\hbar\omega_0)}-m_0c^2-\frac{p_1^2c^2}{2m_0c^2}\\ &= \hbar\omega_0+\frac{p_2^2c^2}{2m_0c^2(1+\frac{\hbar\omega_0}{m_0c^2})}-\frac{p_1^2c^2}{2m_0c^2}. \end{aligned} \tag{5.105}$$

Expanding the term in brackets in the denominator results in

$$\begin{aligned}\hbar\omega &= \hbar\omega_0+\frac{p_2^2c^2}{2m_0c^2}\left(1-\frac{\hbar\omega_0}{m_0c^2}+\cdots\right)-\frac{p_1^2c^2}{2m_0c^2}\\ &\approx \hbar\omega_0+\frac{p_2^2c^2}{2m_0c^2}-\frac{\hbar\omega_0p_2^2c^2}{2m_0^2c^4}-\frac{p_1^2c^2}{2m_0c^2}. \end{aligned} \tag{5.106}$$

Considering the case of absorption of a photon we use the conservation of momentum

$$p_2^2=p_1^2+\vec{p}_1\cdot\hbar\vec{k}+\hbar^2k^2 \tag{5.107}$$

which we insert into (5.106) to find

$$\hbar\omega=\hbar\omega_0+\hbar\vec{v}_1\cdot\vec{k}+\frac{(\hbar\omega)^2}{2m_0c^2}-\hbar\omega_0\frac{v_1^2}{2c^2}+\ldots \qquad \text{(absorption)} \tag{5.108}$$

Like in (5.101) the first three terms are the Bohr frequency $\hbar\omega_0$, the linear Doppler shift and the recoil shift. The fourth term is sometimes called the quadratic Doppler shift as it depends on v^2/c^2. Since it is due to the time dilation in the moving frame when observed by an external observer the resulting shift is often called the "time dilation shift". In the case of the emission of a photon we obtain

$$\hbar\omega=\hbar\omega_0+\hbar\vec{v}_2\cdot\vec{k}-\frac{(\hbar\omega)^2}{2m_0c^2}-\hbar\omega_0\frac{v_2^2}{2c^2}+\ldots. \qquad \text{(emission)} \tag{5.109}$$

The correction terms accounting for the energy difference between the Bohr energy $\hbar\omega_0 = E_2 - E_1$ and the energy absorbed (or emitted) by the particle, i.e., the first-order Doppler shift, the second-order Doppler shift, and the recoil shift result from the fact that the energy–momentum relationship of a photon is linear in contrast to the non-linear relationship of a particle with rest mass (Fig. 5.16).

5.4.2.1 First-order Doppler Shift and Broadening

Consider an absorber moving with velocity $\vec{v}$ in the field of an electromagnetic wave described by a plane wave with wavevector $\vec{k}$ with $|\vec{k}| = 2\pi/\lambda = 2\pi\nu/c$ and frequency ν_0 which corresponds to the transition frequency of an atom at rest. According to (5.108) the frequency ν experienced by the absorber in its frame is

$$\nu=\nu_0+\frac{\vec{k}\cdot\vec{v}}{2\pi}-\frac{v^2}{2c^2}\nu_0 \tag{5.110}$$

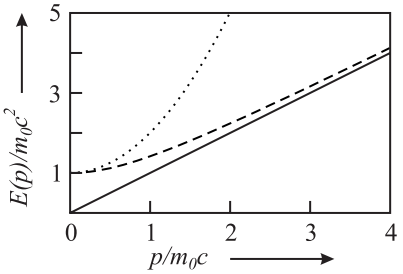


Figure 5.16: The recoil effect and the first- and higher-order Doppler effects result from the mismatch between the energy momentum relationship of the photon ($E = pc$, solid line) and a particle with rest mass m_0c^2 either with the classical energy ($E = m_0c^2 + p^2/(2m_0)$, dots) or the relativistic energy ($\sqrt{p^2c^2 + m_0^2c^4}$, dashed line), respectively.

which differs from the transition frequency by the first-order Doppler shift and the second-order Doppler shift. We first look to the frequency shift due to the first-order Doppler shift

$$\nu - \nu_0 = \frac{\vec{k} \cdot \vec{v}}{2\pi} = \frac{|\vec{v}|}{\lambda} \cos\alpha = \nu \frac{|\vec{v}|}{c} \cos\alpha \tag{5.111}$$

which depends on the angle α between $\vec{k}$ and $\vec{v}$. Consider an ensemble of quantum absorbers, e.g., atoms kept in an absorption cell. We assume that the velocities of atoms in thermal equilibrium with a reservoir of temperature T are isotropically distributed. For a given direction chosen as the v_z-direction the probability of finding an atom within the velocity interval v and $v_z + dv_z$ is described by a Maxwellian distribution (Fig. 5.17)

$$p(v_z)dv_z = \frac{1}{\sqrt{\pi}u} \exp\left[-\left(\frac{v_z}{u}\right)^2\right] dv_z \tag{5.112}$$

with a most probable velocity u that can be derived from

$$\frac{mu^2}{2} = k_B\, T \qquad \text{as} \qquad u = \sqrt{\frac{2\, k_B\, T}{m}}. \tag{5.113}$$

$k_B = 1.38 \times 10^{-23}$ Ws/K is the Boltzmann constant. As a result of (5.111) each atom experiences a (linear) Doppler shift that depends on its velocity.

The influence of the first-order Doppler effect on the profile of an absorption line of atoms obeying a Maxwellian velocity distribution is easily calculated from (5.112) by inserting $v_z = (\nu - \nu_0)\lambda = c(\nu - \nu_0)/\nu_0$ from (5.111) as

$$p\left(\frac{c(\nu - \nu_0)}{\nu_0}\right) \propto \exp\left[-\frac{mc^2}{2k_BT}\left(\frac{\nu - \nu_o}{\nu_0}\right)^2\right]. \tag{5.114}$$

Hence, as a result of the first-order Doppler effect, a transition line is broadened according to (5.114) where the Doppler-broadened profile of the spectral line (Fig. 5.18) represents a

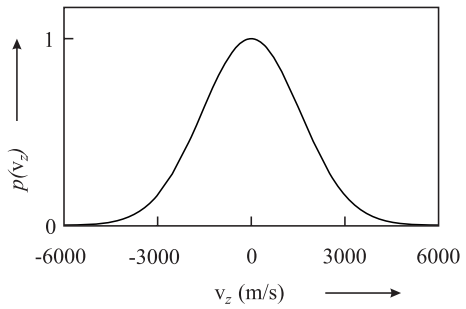


Figure 5.17: Maxwellian velocity distribution (5.112) for atomic hydrogen at $T = 300\,\mathrm{K}$.

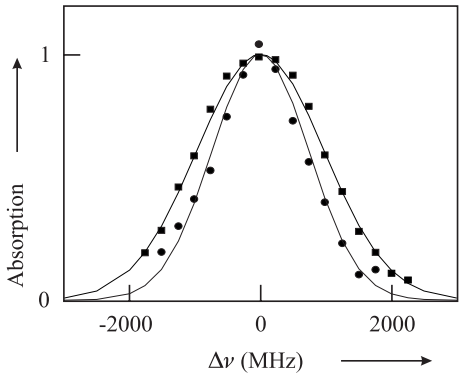


Figure 5.18: Doppler broadened spectral line of the Ca intercombination line $\lambda = 657\,\mathrm{nm}$ for two temperatures $765\,^{\circ}\mathrm{C}$ (squares) and $625\,^{\circ}\mathrm{C}$ (dots) and fits according to (5.114).

Gaussian whose full width at half maximum (FWHM) is

$$\mathrm{FWHM} = \nu_0 \sqrt{2 \ln 2 \frac{k_B T}{mc^2}}. \tag{5.115}$$

As a consequence of the Doppler broadening (Fig. 5.18), for example, at room temperature (300 K) the gain profile of the neon atoms in the HeNe laser at λ = 633 nm has a width of about 1.5 GHz.

5.4.2.2 Second-order Doppler Effect

For larger velocities of the absorbers the second-order Doppler effect becomes more and more important with respect to the first-order Doppler effect. The dramatic influence of the second-order Doppler effect (second term of (5.110)) on the shape and centre of a spectral line can be seen in particular in the spectra of absorbers having low masses and high transition frequencies, as for instance, in the 1S – 2S transition of atomic hydrogen (see Fig. 5.19). These spectra were taken at different temperatures of the absorbers corresponding to different mean velocities. At room temperature the centre of gravity of this absorption line is shifted by about 40 kHz to lower frequencies. The quadratic dependence on the velocity furthermore leads to a strong asymmetry of the line.

5.4.3 Saturation Broadening

The observed shape and width of an absorption line in an ensemble of absorbers furthermore depends on the irradiance of the radiation field. Consider a number N of two-level atoms with populations N_1 and N_2 in the ground state and excited state, respectively, and $N = N_1 + N_2$. The interaction of the two-level system with a radiation field of spectral energy density $\rho(\nu)$ is often described phenomenologically by the Einstein coefficients A_{21}, B_{21} and

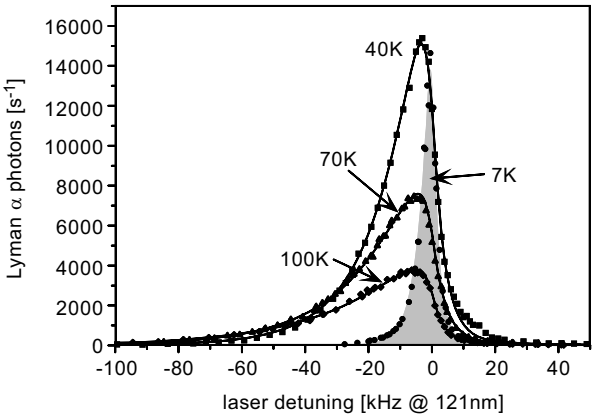


Figure 5.19: Influence of the second-order Doppler effect on the 1S – 2S two-photon absorption line in atomic hydrogen. Courtesy of Th. Hänsch.

B_{12} of spontaneous emission, stimulated emission and stimulated absorption, respectively. In the stationary case, the total absorption rate $N_1 B_{12}\rho(\nu)$ has to be equal to the sum of the rate due to spontaneous emission ($A_{21}N_2$) and that due to induced emission ($B_{21}\rho(\nu)N_2$)

$$(B_{21}\rho(\nu) + A_{21})N_2 = B_{12}N_1\rho(\nu). \tag{5.116}$$

Hence, stimulated emission and absorption strongly modify the population distribution. By defining a dimensionless saturation parameter [18]

$$S \equiv 2B_{12}\frac{\rho(\nu)}{A_{21}} \tag{5.117}$$

and using $B_{21} = B_{12}$, the population ratio in the excited state can be expressed as

$$\frac{N_2}{N} = \frac{S}{2(1+S)}. \tag{5.118}$$

In the vicinity of the atomic transition the frequency dependence of the absorption rate, and hence that of B_{12} and S, is given by a Lorentzian

$$S = S_0\frac{\left(\frac{\gamma}{2}\right)^2}{\left(\frac{\gamma}{2}\right)^2 + \delta\nu^2}. \tag{5.119}$$

The frequency detuning $\delta\nu = \nu - \nu_0$ represents the difference between the frequency of the electromagnetic radiation and the transition frequency which in general may be shifted by the Doppler effect and external fields. The on-resonance saturation parameter

$$S_0 \equiv \frac{I}{I_{\mathrm{sat}}} \tag{5.120}$$

[18] Note that in the literature a different definition of the saturation parameter is also used ($S = B_{12}\frac{\rho(\nu)}{A_{21}}$) which then leads to saturation intensities being twice those shown in Table 5.6.

includes the so-called saturation intensity [19]

$$I_{\rm sat} = \frac{2\pi^2 hc\gamma}{3\lambda^3}. \tag{5.122}$$

The saturation intensity is the irradiance (energy flow density) where $S_0 = 1$ and, hence, where according to (5.118) the steady-state population difference is $(N_1 - N_2)/N = 0.5$. For $S \gg 1$ the population difference tends to zero. Saturation intensities for some relevant atomic transitions are shown in Table 5.6.

Table 5.6: Saturation intensities of some resonance lines in atoms relevant for frequency standards. Others may be found, e.g., in [132].

Atom	Transition	Wavelength nm	$\gamma = 1/(2\pi\tau)$ MHz	$I_{\rm sat}$ mW/cm^2
^{1}H	$1^2S_{1/2}$ - $2^2P_{3/2}$	121.57	99.58	7244
^{24}Mg	3^1S_0 - 3^1P_1	285.30	81	455
^{40}Ca	4^1S_0 - 4^1P_1	422.79	34	59
^{85}Rb	$5^2S_{1/2}$ - $5^2P_{3/2}$	780.24	6	1.6
^{88}Sr	5^1S_0 - 5^1P_1	460.86	32	43
^{133}Cs	$6^2S_{1/2}$ - $6^2P_{3/2}$	852.35	5.2	1.1

By inserting (5.119) into (5.118) one finds

$$\frac{N_2}{N} = \frac{S_0}{2}\frac{(\gamma/2)^2}{(1+S_0)(\gamma/2)^2 + \delta\nu^2} = \frac{S_0}{2(1+S_0)}\frac{1}{1+((2\delta\nu)/\gamma')^2} \tag{5.123}$$

which again is a Lorentzian, however, with an increased linewidth

$$\gamma' = \gamma\sqrt{1+S_0}. \tag{5.124}$$

In the centre of the spectral line the number of scattered photons saturates faster than in the wings and the line is broadened (Fig. 5.20). This effect is referred to as "power broadening" or "saturation broadening". It can become very important in the case of frequency standards where narrow lines with low decay rates γ easily lead to large saturation ratios.

[19] The word "intensity" is often used for different radiation quantities which can lead to confusions when comparing the results from different sources as has been illustrated by Hilborn [146]. Often, the word intensity is used for the physical quantity "irradiance" (units: W/m^2) as in the case of the saturation intensity (see (5.122)). The relationship between the irradiance I, time-averaged over the fast oscillation, and the electric field amplitude E_0 of a plane electromagnetic wave $E(t,z) = E_0\cos(\omega t - kz)$ is given by

$$I = \frac{\epsilon_0 c}{2}E_0^2. \tag{5.121}$$

Except for the saturation intensity as a well known quantity, whenever possible, here the word intensity in conjunction with radiation quantities is avoided.

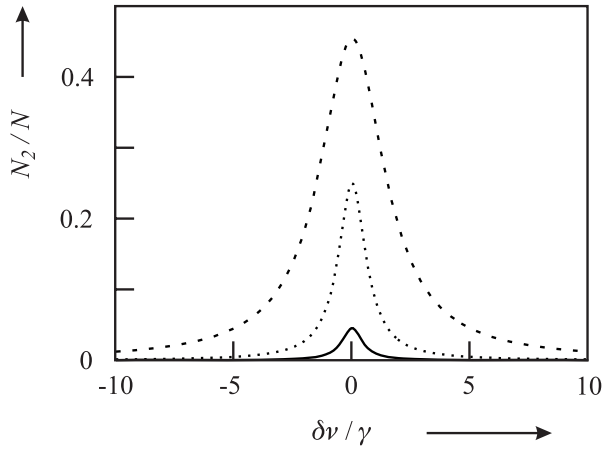


Figure 5.20: Power broadening of a spectral line for on-resonance saturation parameters $S_0 = 0.1$ (solid line), $S_0 = 1$ (dotted line) and $S_0 = 10$ (dashed line) according to (5.123).

5.4.4 Collisional Shift and Collisional Broadening

In an ensemble of atomic or molecular particles where the particles move with a velocity defined by their kinetic energy (5.112), close encounters between particles may occur. One refers to such processes as collisions if energy and momentum are exchanged between these particles. In so-called elastic collisions the total kinetic energy is conserved, whereas in inelastic collisions an energy exchange takes place between the external and the internal degrees of of freedom of the partners involved in the collision. During the collision process the energy levels of the particles involved in the collision are modified depending on their relative distance, e.g., as depicted in Fig. 5.3. The modification of the energy level structure during the collision in general may also lead to a change of the transition frequency between two levels. The time-integrated effect on the position and the width of the absorption line are referred to as collisional shift and collisional broadening, respectively.

The magnitude of these effects depends on a variety of conditions and no universal theory exists that is capable of describing all effects. Most of the microscopic absorbers, however, are used in the form of dilute gases and, hence, the temperature of the gas and the associated velocities of the particles lead to very different aspects in the different regimes. As we are interested in the shift and broadenings occurring for particular frequency standards, in the following we briefly sketch those effects in the relevant regimes.

5.4.4.1 Collisions at Thermal Energies

For thermal energies at room temperature and above, the velocities of the atoms are of the order of a few hundred metres per second (see (5.113)). The collision diameter, i.e., the range where the particles interact with each other, corresponds typically to a few diameters of the particles. Hence, for atoms with a velocity of a few hundred metres per second and a collision diameter of about a nanometre, the time of the collision T_{col} is a few picoseconds. This

time is still large enough that the electronic energy levels of the collision partners are able to follow "adiabatically" the perturbation by the other partners as depicted in the potential curves of Fig. 5.3. If one of the particles involved in the collision emits electromagnetic radiation during the collision the frequency of the emitted radiation is temporarily shifted. If T_{col} is small compared to one oscillation period $T = 1/\nu$ of the electromagnetic radiation, an instantaneous phase shift occurs as is visualised in Fig. 5.21. This approximation is valid in

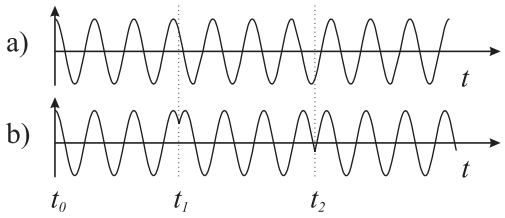


Figure 5.21: The influence of collisions on the phase of the emitted wave. a) Unperturbed emission. b) Phase shifting collisions at time t_1 and t_2.

the case of microwave clocks operating at a few gigahertz and corresponding times of several nanoseconds but not in the case of optical transitions where the corresponding oscillation times T are of the order of femtoseconds. Consider the case where the collision rate is high enough that multiple collisions occur during the emission of radiation. Hence, the emitted radiation consists of finite wave trains of a mean duration $\tau_c = \overline{t_2 - t_1}$ of unperturbed phase (see Fig. 5.21) with random phase shifts due to the collisions. A Fourier analysis similar to that performed in Section 2.1.2.2 leads to a Lorentzian lineshape.[20] The halfwidth $\Delta\nu_c$ is calculated, e.g., by analogy to (2.38) as

$$\Delta\nu_c = \frac{1}{\pi\tau_c} \tag{5.125}$$

by assuming that the lifetimes of both states are limited by the mean time τ_c between the collisions. The lifetime can be calculated from kinetic gas theory, leading to

$$\Delta\nu_c = \sqrt{\frac{3}{4mk_BT}}d^2p \tag{5.126}$$

where m and d are the mass and diameter of the absorbers, respectively. T and p are the temperature and the pressure of the gas. Since the broadening according to (5.126) depends on the pressure it is sometimes referred to as pressure broadening. This model (with $T_{\mathrm{col}} \ll T$) is capable of explaining the occurrence of collision broadening. Collisional shifts then occur, e.g., as a result of asymmetric lineshapes where the broadening affects the maximum or the centre of gravity of the line.

In the case of $T \ll T_{\mathrm{col}}$, e.g., for optical transitions, the frequency of the electromagnetic radiation absorbed or emitted during the collision can be severely altered for a considerable

[20] If, however, a finite duration of the collision process is taken into account small but detectable deviations from the Lorentzian lineshape have been observed [147].

fraction of the total wave train, which immediately explains the occurrence of collisional shifts and broadening. As a rule of thumb for optical frequency standards, the pressure shift is lower by at least an order of magnitude compared to the pressure broadening. In optical frequency standards such as the iodine-stabilised lasers (Section 9) the pressure shift is of utmost importance for the achievable accuracy. For iodine near room temperature the pressure shift is roughly proportional to the pressure, with the proportionality constant being of the order of 10 kHz/Pa.

5.4.4.2 Collisions of Cold Quantum Systems

For absorbers at very low temperatures, e.g., for laser-cooled atoms, the very quantum nature of the systems becomes more and more important for a meaningful description of the collision process. To begin with, at these temperatures the velocity of an atom is so low that the de Broglie wavelength

$$\lambda_{\mathrm{dB}} = \frac{h}{mv} \tag{5.127}$$

becomes much larger than the diameter of the atom. Hence, the picture of interacting wave packets is more appropriate as the concept of particles and the collision process has to be treated as a scattering process. The wavefunction describing both incoming particles in the centre-of-mass system is expanded in partial waves with a well-defined angular momentum with the quantum number l. It is appropriate to solve the time-dependent Schrödinger equation (5.15) in spherical coordinates with V the interaction potential.

$$V(R) = -\frac{C_n}{R^n}. \tag{5.128}$$

If two identical atoms collide, the long-range potential can be given in some special cases. If both atoms are in an S state, the interaction is of the van der Waals type and $n = 6$ holds. This case is realised, e.g., in the collision of two alkaline-earth atoms (^{20}Mg, ^{40}Ca,or ^{88}Sr) being in their ground states. If one of the two atoms were in the excited state the atoms are coupled by an electric dipole transition and the corresponding dipole interaction has an R^{-3} dependence. Hence, the dipole potential is relevant for the collisions of alkaline earth atoms used in optical frequency standards (Section 9.4.4), e.g., in a magneto-optical trap operated on the $^1\mathrm{P}_1$ – $^1\mathrm{S}_0$ transition which might then limit the achievable density. The S – P interaction between excited-state atoms and ground-state atoms also may lead to a collisional shift of the optical clock transition $^3\mathrm{P}_1 - ^1\mathrm{S}_0$. As the C_3 coefficients are proportional to the transition strength of the relevant transition, the interaction potential is much weaker in the latter case. If both colliding atoms are in the P state $n = 5$ holds as a result of the quadrupole interaction.

Apart from the collisional aspects discussed so far quantum mechanical scattering resonances may play an important role as, e.g., in caesium (Section 7.3.2.1) where they may lead to inelastic collisions and the occurrence of so-called Feshbach resonances.[21]

[21] A Feshbach resonance occurs when two colliding atoms are resonantly coupled to a bound state of the molecular pair.

5.4.5 Influence of External Fields

External fields, e.g., dc or ac electric and magnetic fields, in general, shift the energy levels of the atomic, ionic, or molecular absorbers employed in frequency standards and hence have a profound influence on the accuracy and stability of the standards.

5.4.5.1 Electric Fields

The interaction of an atom with an external electric field resulting in frequency shifts of atomic levels and the associated line shifts and splittings is referred to as Stark effect. Since the interaction energy with the external field in general is small with respect to the internal energy of the atom the Stark effect can be calculated by perturbation theory. The induced dipole moment $\vec{d}$ is proportional to the electric field $\vec{E}$ and the ratio

$$\alpha \equiv \frac{d}{E} \tag{5.129}$$

of the electric dipole moment d and the electric field E is called the "polarisability". For an atom without permanent dipole moment the Stark shift of energy E_m of the unperturbed level m is calculated as a second-order perturbation as

$$\Delta E_m = \sum_n \frac{|\langle m|\vec{d}|n\rangle|^2}{E_m - E_n} E^2 \tag{5.130}$$

where the summation includes all discrete and continuum states n that are coupled with the state m by the dipole operator but excludes $m = n$.

Apart from static electric fields the quantum system can interact with alternating (ac) electric fields whose frequencies can vary from dc to the optical ones if, e.g., the interaction with the electric field of an electromagnetic wave is considered. The polarisability α can be calculated [148] classically in the so-called Lorentz model where the oscillation of an electron with charge $-e$ and mass m driven by the electric field $E(t)$ is described by the equation of motion as $\ddot{x} + \Gamma_\omega \dot{x} + \omega_0^2 x = -eE(t)/m$ (Section 2.2). By integrating this equation of motion and by using $\alpha = -ex/E_0$, where E_0 is the amplitude of the electric field, it follows from $x = x_0 \exp(i\omega t)$ that

$$\alpha = \frac{e^2}{m} \frac{1}{\omega_0^2 - \omega^2 + i\omega\Gamma_\omega}. \tag{5.131}$$

The classical damping rate due to the radiative loss is given [58] from Larmor's formula by

$$\Gamma_\omega = \frac{e^2\omega^2}{6\pi\epsilon_0 m c^3}. \tag{5.132}$$

Following Grimm *et al.* [148] we introduce an on-resonance damping rate $\Gamma \equiv \Gamma_{\omega_0} = (\omega_0/\omega)^2\Gamma_\omega$. By insertion of (5.132) into (5.131) one obtains [22]

$$\alpha = 6\pi\epsilon_0 c^3 \frac{\Gamma/\omega_0^2}{\omega_0^2 - \omega^2 + i(\omega^3/\omega_0^2)\Gamma}. \tag{5.134}$$

By writing $\vec{E} = \vec{E}_0 \cos\omega t = \frac{1}{2}\vec{E}_0 \exp(i\omega t) + c.c. = \frac{1}{2}\vec{E}_0 [\exp(i\omega t) + \exp(-i\omega t)]$ and $\vec{d} = \alpha\vec{E} = \frac{1}{2}\alpha\vec{E}_0 \exp(i\omega t) + \frac{1}{2}\alpha^*\vec{E}_0 \exp(-i\omega t)$ we find the interaction energy of the induced dipole moment $\vec{d}$ in the electric field $\vec{E}$

$$W_{\rm dip} = -\frac{1}{2}\langle\vec{d}\vec{E}\rangle = -\Re\mathrm{e}\{\alpha\}\left(\tfrac{1}{2}E_0\right)^2 \tag{5.135}$$

where $\langle\,\rangle$ indicates that the oscillating terms have to be time averaged. The oscillating dipole absorbs power from the driving field that is calculated as

$$P_{\rm abs} = \langle\left(\frac{d}{dt}\vec{d}\right)\vec{E}\rangle = -2\omega\Im\mathrm{m}\{\alpha\}\left(\tfrac{1}{2}E_0\right)^2. \tag{5.136}$$

Hence, the real part (5.135) and imaginary part (5.136) of the polarisability α describes the in-phase and out-of-phase component of the dipole oscillation with the dispersive and absorptive properties, respectively. We will use these properties later to describe optical traps for neutral atoms (Section 6.4)

In the following we discuss in more detail the polarisability depending on the driving frequency where we have to distinguish four different cases [149, 150]. Consider a state $|m\rangle$ whose energy E_m is shifted by ΔE_m under the influence of radiation of angular frequency ω that couples the level $|m\rangle$ to other levels $|n\rangle$ with the matrix elements $\langle m|\vec{d}\,|n\rangle$.

First, if the electric field oscillates with an angular frequency ω that is slow with respect to the inverse lifetime τ ($\omega \ll 1/\tau$) of the atomic transition, the atom follows adiabatically the perturbation by the electric field and the Stark effect can be calculated as if the field were static with an amplitude varying with the angular frequency ω.

A second important case is encountered if the frequency of the electric field is larger than the spectral width of the transition but much smaller than the transition frequency of a dipole transition between the states $|m\rangle$ and $|n\rangle$ with the energies E_m and E_n, respectively, i.e., $1/\tau \ll \hbar\omega \ll |E_m - E_n|$. In this case the atom can no longer follow the oscillation of the electric field, but responds to the average (rms) electric field. Hence the ac-Stark shift is calculated by using the static response of the atom and the square averaged electric field strength.

---

[22] If the atomic polarisability is calculated by using a two-level quantum system and a classical radiation field one derives for low saturation an expression similar to (5.134) where the classical damping rate is replaced by

$$\Gamma = \frac{\omega_0^3}{3\pi\epsilon_0\hbar c^3}|\langle 2|\hat{d}\,|1\rangle|^2. \tag{5.133}$$

Nevertheless, the classical expression (5.132) is a good approximation for strong dipole-allowed transitions from the ground state [146, 148].

In the third regime the angular frequency ω of the alternating electric field becomes comparable to a transition angular frequency $|E_m - E_n|/\hbar$ between two levels that can be connected by an allowed dipole transition with matrix element $\langle m|\vec{d}|n\rangle$ but is not resonant, the m^{th} energy level is shifted by

$$\begin{aligned}\Delta E_m &= \frac{1}{4}\sum_n |\langle m|\vec{d}|n\rangle|^2 E_0^2 \left(\frac{1}{E_m - E_n - \hbar\omega} + \frac{1}{E_m - E_n + \hbar\omega}\right) \\ &= \frac{1}{2}\sum_n |\langle m|\vec{d}|n\rangle|^2 E_0^2 \frac{E_m - E_n}{(E_m - E_n)^2 - (\hbar\omega)^2}. \end{aligned} \tag{5.137}$$

The classical analogue of this formula is readily derived from (5.135) for large detunings ($\hbar\omega_0 - \hbar\omega \equiv E_m - E_n - \hbar\omega \gg \Gamma$).

A fourth important case occurs if the electric field is resonant with an allowed transition ($\hbar\omega \approx E_m - E_n$) where transitions between the states can be induced. The coupling between these states leads to a splitting of the states. Exactly at resonance the splitting is given by $\hbar\Omega_R$ where Ω_R is the Rabi frequency. Resonant transitions between the states may be induced also by thermal radiation thereby reducing the lifetime of the excited state of the atoms. Such reduction of the lifetime was, e.g., held responsible for the unexpectedly low excited-state lifetime of a xenon transition [100] proposed as an optical frequency standard ([151], Section 9.4.6).

The frequencies of clocks and frequency standards operating on a transition between the hyperfine-split ground states $^2S_{1/2}$ as in the Cs atomic clock are affected by the averaged electric field associated with the radiation field at ambient temperature which will be considered in more detail in Section 5.4.5.2 and in Section 7.1.3.4.

Frequency shifts due to near-resonant optical radiation referred to as "ac-Stark shift" or as "light shift" can ultimately limit the accuracy of frequency standards, e.g., in optically pumped standards (Section 8.2.2 or Section 7.2). Spatially dependent light shifts are furthermore important for the realisation of optical traps for neutral atoms (Section 6.4).

5.4.5.2 Black-body Radiation

In frequency standards operated at a temperature $T \neq 0\,\text{K}$ the atoms are exposed to the ambient electromagnetic temperature-radiation field which perturbs the energy levels and hence leads to frequency shifts of the clock transition. To estimate the magnitude of the temperature field, one considers the radiation field of an ideal black-body radiator whose spectral energy density $\rho_2(\nu, T)$ emitted into a solid angle of 4π is given by Planck's formula as

$$\rho_2(\nu, T)d\nu = \frac{8\pi h\nu^3}{c^3}\frac{1}{e^{\frac{h\nu}{k_B T}} - 1}d\nu. \tag{5.138}$$

Hence, the ac-Stark shift resulting from the thermal radiation field is sometimes also referred to as the "black-body shift".

For a monochromatic field of amplitude E_0 the spectral energy density is given as

$$\rho_2 = \epsilon_0\langle E^2\rangle = \tfrac{1}{2}\epsilon_0 E_0^2. \tag{5.139}$$

Since the field of the temperature radiation represents a continuous spectrum of frequencies the electric field, e.g., in (5.137) has to be substituted by an integral over its spectral density which can be connected to the spectral density $\rho_2(\nu, T)$ of the black-body radiator as follows

$$\langle E^2(t)\rangle = \frac{1}{\epsilon_0}\int_0^\infty \rho_2(\nu, T)d\nu. \tag{5.140}$$

(5.140) can be solved by using (5.138) and with the help of $\int_0^\infty \frac{x^3}{e^x-1} = \frac{\pi^4}{15}$. The averaged squared amplitude of the electric field integrated over all frequencies then becomes

$$\langle E^2(t)\rangle = \frac{1}{\epsilon_0}\frac{8\pi k^4 T^4}{c^3h^3}\int_0^\infty \frac{(\frac{h\nu}{kT})^3}{e^{\frac{h\nu}{kT}}-1}d(\frac{h\nu}{kT}) = \frac{8\pi^5 k^4}{15c^3h^3\epsilon_0}T^4 = \frac{4\sigma}{c\epsilon_0}T^4 \tag{5.141}$$

where we have introduced the Stefan–Boltzmann constant $\sigma \equiv (2\pi^5k^4)/(15c^2h^3) = 5.6705\times10^{-8}\,\mathrm{W\,m^{-2}\,K^{-4}}$. Hence, one finds from (5.141) the time-averaged quadratic electric field is

$$\langle E^2(t)\rangle = (831.9\,\mathrm{V/m})^2 \qquad \text{for } T = 300\,\mathrm{K}. \tag{5.142}$$

There is also a time-averaged quadratic magnetic field that can be calculated by using $\langle B^2\rangle = \langle E^2\rangle/c$ as $\langle B^2(t)\rangle = (2.775\,\mu\mathrm{T})^2$, which corresponds to about 5 % of Earth's magnetic field. In general the frequency shift due to the electric field is much larger. The influence of the black-body radiation on the frequency of the caesium clock will be discussed in Section 7.1.3.4.

5.4.5.3 Magnetic Fields

The Hamiltonian of an atom in an external magnetic field can be written conveniently as

$$\mathcal{H} = \mathcal{H}_{LS} + \mathcal{H}_{\mathrm{hfs}} + \mathcal{H}_B \tag{5.143}$$

where $\mathcal{H}_{LS}$, $\mathcal{H}_{\mathrm{hfs}}$ and $\mathcal{H}_B$ take into account the spin-orbit interaction in LS coupling, the hyperfine interaction, and the interaction of the electronic shell with the magnetic field, respectively. In a small magnetic field where $\mathcal{H}_B$ can be considered as a perturbation to $\mathcal{H}_{LS}$ the energy levels of the atoms show a Zeeman splitting according to (5.8) and (5.9) where the energy of the level depending on the magnetic quantum number m_J and on the Landé factor g_J as follows

$$\Delta E_{\mathrm{Zeeman}} = g_J\mu_B m_J B = \frac{J(J+1)+S(S+1)-L(L+1)}{2J(J+1)}\mu_B m_J B. \tag{5.144}$$

For large magnetic fields where the interaction energy of the electronic shell with the magnetic field is larger than the interaction energy between the spin and the orbital angular momenta and the associated magnetic moments both are decoupled and precess independently around the magnetic field. In this so-called Paschen–Back regime

$$\Delta E_{\mathrm{P-B}} = \mu_B(m_L + 2m_S)B \tag{5.145}$$

holds. We now consider the hyperfine interaction which is much smaller than the spin-orbit interaction. In the regime of a small magnetic field the magnetic moment m_F precesses around the field and the energy levels depend on m_F. For a small magnetic field, i.e., if the interaction energy corresponding to $\mathcal{H}_B$ is much smaller than the hyperfine coupling between I and J leading to F, one observes the Zeeman effect of the hyperfine structure (see left part of Fig. 5.22). In the strong-field regime I and J precess independently around the magnetic field and one observes the Paschen-Back effect of the hyperfine structure (see right part of Fig. 5.22). In the intermediate region the interaction Hamiltonian with the magnetic field can be written as

$$\mathcal{H}_B = \frac{g_J \mu_B B}{\hbar} J_z + \frac{g_I \mu_n B}{\hbar} I_z \approx \omega_J J_z \tag{5.146}$$

where the second term can in general be neglected since $g_I \mu_n \ll g_J \mu_B$. The basis states $|IJFm_F\rangle$, however, are no longer eigenstates of the operator J_z and hence the operator $H_{\mathrm{hfs}} + \omega_J(B) J_z$ has to be diagonalised as a function of the external field B. In the intermediate regime the states with the same m_F but different F are mixed without coupling between the states with different m_F, leading to $(2m_F + 1)$-dimensional sub-matrices that have to be diagonalised independently.

In the ground state of hydrogen and the alkaline metals where $J = 1/2$ leads to only two hyperfine levels the sub-matrices are of rank 2. The solution is known as the Breit–Rabi formula

$$E_{\mathrm{hfs}}(B, m_F) = -\frac{E_{\mathrm{hfs}}}{2(2I+1)} + m_F g_I \mu_n B \pm \frac{E_{\mathrm{hfs}}}{2} \sqrt{1 + \frac{4m_F}{2I+1} x + x^2} \tag{5.147}$$

$$x = \frac{g_J \mu_B - g_I \mu_n}{E_{\mathrm{hfs}}} B \approx \frac{2\mu_B}{E_{\mathrm{hfs}}} B \tag{5.148}$$

where in (5.148) we have used $g_I \mu_n \ll g_J \mu_B$ and $g_J = 2$ for $S = J = 1/2$.

As an example we consider the case of hydrogen whose nucleus is composed of a single proton orbited by a single electron. The angular momentum of the electronic shell described by the quantum number $J = L + S = 1/2$ results entirely from the spin $S = 1/2$ of the electron since its angular momentum in the ground state is $L = 0$. The angular momentum $I = 1/2$ of the nucleus is made up by the proton. Since both angular momenta I and J interact with each other via their associated magnetic moments only the total angular momentum $F = I + J$ is a "good" quantum number. Its possible values are $F = |I+J| = 1; m_F = +1, m_F = 0, m_F = -1$ forming a triplet and $F = |I - J| = 0; m_F = 0$ (singlet). The energies of the $F = 1$ and $F = 0$ states differ by $\Delta W = h\Delta\nu = 6.6 \times 10^{-34}\,\mathrm{W\,s^2} \times 1.42\,\mathrm{GHz}$ and transitions from one state to the other can be stimulated by electromagnetic radiation of 1.42 GHz (see Fig. 5.22) calculated by using (5.147)). Since a magnetic moment of $M(m_F) = \mu_B g_F m_F$ is associated with the atom in the three triplet states $F = 1$ the energies of the three states of the hydrogen atom in a magnetic field of magnetic induction B differ. The $F = 1, m_F = 1$ and $F = 1, m_F = -1$ show a linear Zeeman effect whereas the $m_F = 0$ states in the singlet and in the triplet at low magnetic fields show a small quadratic dependence of the energy on the magnetic field.

The quadratic dependence of the energy of the $F = 1, m_F = 0$ and the $F = 0, m_F = 0$ states on the magnetic field is small at low fields and fluctuations of magnetic fields do not

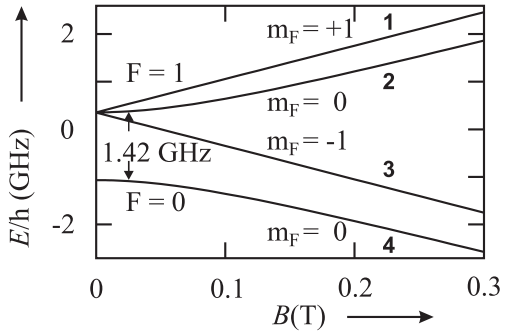


Figure 5.22: Energy levels of the ground states of hydrogen in a magnetic field showing the $F = 1$ triplet and the $F = 0$ singlet separated by 1.42 GHz. The labels **1, 2, 3, 4** for the energy levels are used for the description of the hydrogen maser in Section 8.1.

much affect the energy separation between these states. Consequently, the hydrogen maser as a frequency standard (Section 8.1) is usually operated on the $\Delta F = 1, \Delta m_F = 0$ hyperfine transition (Fig. 5.22) with a frequency separation of 1.42 GHz. The particular dependencies of the energies of the different states on the magnetic field can also be used to manipulate the atoms. From Fig. 5.22 it can be seen that the energies of the $F = 0, m_F = 0$ and the $F = 1, m_F = -1$ states are reduced in a magnetic field. Consequently, in an inhomogeneous magnetic field, these atoms will be accelerated into the regions of high magnetic fields. Atoms in such states are therefore referred to as high-field seekers in contrast to the $F = 1, m_F = 1$ and $F = 1, m_F = 0$ states which are called low-field seekers since they are attracted into regions of a smaller magnetic field.

5.4.6 Line Shifts and Uncertainty of a Frequency Standard

The particular physical effects that have been discussed in this section are examples that affect the stability and accuracy of any frequency standard. The design of good frequency standards thus requires that the associated lineshifts are kept small and constant and that they can be determined accurately. During operation of the standard it has to be ensured that the corresponding offsets of the frequency are known in order to correct for them. The operation of a frequency standard, hence, asks for a careful evaluation of all the effects that are capable of shifting the frequency from the frequency of the unperturbed transition.

The standard procedure is to determine the sensitivity of the frequency of the standard to all relevant parameters, e.g., the temperature of the absorbers (to account for the Doppler effect), the density and temperature of the absorbers (to account for the collisional shift), or the magnitude of the external magnetic or electric dc and ac fields. When these sensitivities to the particular effects have been determined the frequency of the standard is corrected for the combined offset at the chosen working point.

This correction, however, can be performed only with a limited degree of accuracy since, e.g., the sensitivity to a particular parameter is determined with a given uncertainty. Furthermore, the device can be operated at the chosen nominal working point only with a given

uncertainty, e.g., at $T = (20 \pm 1)\,°\mathrm{C}$, at $(60 \pm 10)\,\mu\mathrm{T}$, etc. The corresponding contributions to the uncertainty of the frequency measurement have to be estimated and are usually presented in an uncertainty budget. The procedure of evaluating and stating the uncertainty of a measurement is now standardised [3]. All individual contributions are added in quadrature, leading to a single figure that gives the estimated standard uncertainty and, hence reflects, to the best of the operator's knowledge, the validity of the measurement. In a, not strictly exact, way this figure is often interpreted as the "estimated standard uncertainty of the frequency standard" even though it is merely the estimated standard uncertainty of the particular frequency measurement for which the evaluation has been performed.

6 Preparation and Interrogation of Atoms and Molecules

As has been emphasised in the previous chapter, the linewidth and the centre frequency of a transition is affected by the external degrees of freedom, i.e., the positions and the velocities of the microscopic oscillators interacting with the interrogating electromagnetic radiation. Hence, the progress achieved in the past with frequency standards was intimately related to the development of novel methods for interrogating and manipulating the absorbers under well defined conditions. In this chapter, methods are discussed for overcoming Doppler shifts and broadenings, for increasing the interaction time and for confining the microscopic absorbers to spatial locations of well defined properties. Depending on the methods used to prepare and interrogate the absorbers the observed linewidth can vary over several orders of magnitude (Fig. 6.1).

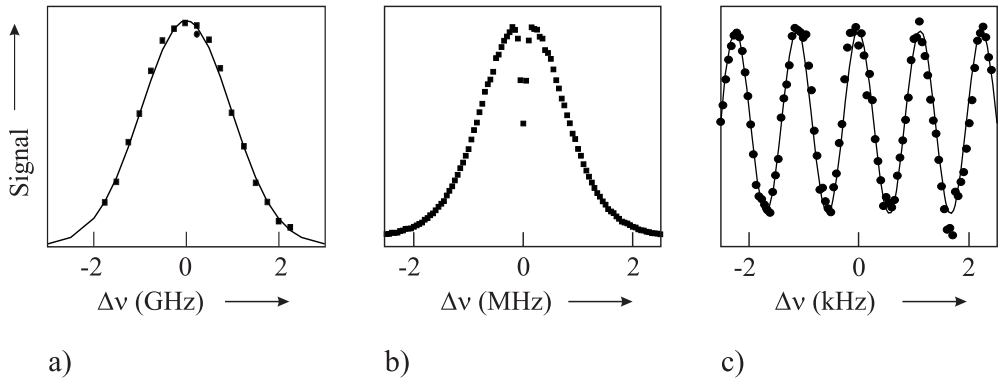


Figure 6.1: Optical transition ($\lambda = 657\,\text{nm}$) in Ca atoms of a natural linewidth of $\Delta\nu \approx 0.37\,\text{kHz}$ obtained under different conditions. a) The absorption in a heated cell (Section 6.1) shows a Doppler broadened linewidth of $\Delta\nu \approx 2\,\text{GHz}$. b) Transverse excitation in a collimated effusive atomic beam (Section 6.2) shows a reduced Doppler width of $\Delta\nu \approx 2\,\text{MHz}$ and the saturated absorption (Section 6.5.1) leads to a transit-time limited absorption dip ($\Delta\nu \approx 150\,\text{kHz}$). c) In a laser cooled (Section 6.3.1) atomic cloud, separated field excitation (Section 6.6) allows one to resolve the line close to the natural linewidth.

6.1 Storage of Atoms and Molecules in a Cell

The confinement of absorbers in a suitable container for interrogation has a number of distinct advantages. Besides allowing for a compact set-up, the absorbing material can be prepared and kept at well defined conditions, e.g., at constant temperature or pressure. It furthermore allows one to make economic use of expensive material with defined purity or isotopic composition.

The advantage of increasing the interaction time with the radiation by containing the absorbing gases or vapours in a suitable cell but is limited to the extent that collisions with the wall during the interrogation perturb the coherent response of the absorbers to the electromagnetic field. Consequently, much work has been devoted to finding or preparing surfaces where the collisions with the surface have as little influence as possible on the clock transition. In the case of absorbers based on microwave transitions between magnetic sub-states, metallic surfaces and surfaces or surface layers with magnetic interactions have to be avoided. Coating the wall of the bulb in the hydrogen maser with PTFE (Teflon) or helium, or coating the absorption cells in the rubidium standard with paraffin, has proved to be an efficient method of reducing the influence of collisions. To increase the time the absorber spends in the cell before reaching the wall, sometimes the cell contains additionally buffer gases of inert atoms or molecules, besides the absorbers.

Confining the absorbers in cells whose relevant dimensions are smaller than half the wavelength of the interrogating electromagnetic radiation, allows one to probe the transition without perturbation by the first-order Doppler effect. This so-called Lamb-Dicke regime (see Section 10.1.4) is readily obtained in microwave standards where the radiation has a wavelength of a few centimetres. In this regime, collisions can narrow the linewidth rather than broadening it [152, 153]. Line narrowing in a very thin cell has also been observed in the optical domain [154].

In a variety of frequency standards the microscopic particles are contained in a cell with walls that can let through the electromagnetic radiation used for interrogation of the quantum systems. Cells containing atoms or molecules are applied likewise in microwave standards, e.g., the hydrogen maser (Section 8.1) and the rubidium standard (Section 8.2) or in optical frequency standards (see, e.g., Section 9.1 and Section 9.4). The particular properties of the cells and their influence will be discussed in more detail later in conjunction with the relevant frequency standards.

6.2 Collimated Atomic and Molecular Beams

Atomic and molecular absorbers are often prepared in beams rather than in absorption cells. Beams are applied, e.g., in cases where gases like hydrogen or vapours from metals with high melting points are not stable as atomic species, but are readily available in molecular form or as solid material. After the preparation of atomic particles from those materials, collisions and the formation of molecules or solids prior to the interrogation can be avoided in a beam.

The specific properties of atomic and molecular beams and the relevant techniques for their preparation can be found in [155, 156] and hence it suffices here to recall only the basics that are relevant to frequency standards. Often atoms effuse from an oven kept at temperature T. Inside the oven, the velocities of the atoms in any direction are given by the Maxwellian

velocity distribution (see (5.112)). The probability of an atom emerging from the oven is proportional to v. Hence, the intensity $I(v)dv$ of a beam, i.e., the number of atoms per unit area and unit time in the velocity interval between v and $v + dv$, effusing from an oven kept at a temperature T, through a nozzle, is given by

$$I(v) = \frac{2I_0}{u^4} v^3 e^{-(v/u)^2} \tag{6.1}$$

with the most probable velocity u given by $u = \sqrt{2k_B T/m}$ (5.113). Here, I_0 is the full beam intensity and v is the velocity of the atom or molecule. The transverse extension of the beam and the velocities of particles effusing from the orifice of an oven result from the position and diameter of the chosen diaphragms. For limiting apertures, the maximum full opening angle is given by

$$2\alpha = (d + D)/L \tag{6.2}$$

where d and D are the diameters of the orifice and the diaphragm, respectively, and L is the distance that separates them along the z direction. For a narrow collimation angle α equivalent to $d \ll L, D \ll L$, only particles with transverse velocities $v_\perp = \sqrt{v_x^2 + v_y^2}$ are selected, where $v_\perp/v_z < \alpha$. For typical collimation angles $0.01\,\text{rad} \lesssim \alpha \lesssim 0.001\,\text{rad}$, $v = \sqrt{v_x^2 + v_y^2 + v_z^2} \approx v_z$ holds.

In optical frequency standards a laser beam is often used that crosses the atomic beam perpendicularly. In the case of flat wave fronts the residual Doppler broadening can be calculated in a similar way to the case of a vapour in an absorber cell (Section 5.4.2.1) except for the fact that the width is now reduced by $\sin\alpha \approx \alpha$, i.e.,

$$\Delta\nu_{\text{beam},\perp} \approx \alpha \Delta\nu_{\text{vapour}} \tag{6.3}$$

in an atomic beam as compared to a cell. The associated striking reduction of the broadening can be seen by comparing the Doppler curves of Fig. 6.1 a) and Fig. 6.1 b).

The Maxwellian velocity distribution is modified if collisions take place inside the nozzle of the oven. Collisions become relevant if the Knudsen number $K = \bar{\lambda}/l$ approaches unity where $\bar{\lambda}$ is the mean free path of the atoms and l is the length of the nozzle. The mean number of collisions of the atoms in the nozzle described by the Knudsen number gives rise to a reduction of the slow atoms in the beam.[1] For high-precision experiments with hydrogen atoms this effect had to be taken into account to describe the lineshapes of the measured two-photon absorption spectrum properly [157].

For frequency standards, the use of atomic or molecular beams with a defined velocity direction meets with both advantages and disadvantages. Well defined trajectories can be used to spatially separate the particles with different internal states as is done in Cs atomic clocks or the hydrogen maser with magnetic state selection (Section 7.1, Section 8.1.3.2). In optical frequency standards, the use of collimated atomic beams allows the excitation of the absorbing particles by transversal laser beams, thereby reducing the first-order Doppler

[1] This effect was considered to be the reason for the unsuccessful attempt to operate the first atomic fountain employing a thermal beam by Zacharias [15, 17, 155].

shift and broadening by orders of magnitude depending on the collimation ratio. On the other hand, if the laser beam intersects the atomic beam at an angle deviating from 90°, then immediately a residual first-order Doppler shift occurs. The excitation of long-lived states that are particularly useful in frequency standards allows one to spatially separate the excitation and detection zones with obvious advantages for the signal-to-noise ratio.

In an ensemble of particles of finite temperature their velocities are generally spread over a large velocity range. Hence, the perturbations by the first and second-order Doppler effect and the interaction-time broadening can be reduced if absorbers with lower than the mean velocity are selected. Selection of the signal from the slow atoms in a thermal beam of hydrogen atoms has been achieved, e.g., by gating the excitation and detection of the excited atoms in a beam [157] and selecting the velocity class by the time delay between both processes (Section 9.4.5). Generally, as in the case of the Maxwell-Boltzmann distribution, the number of particles in a given velocity range is dramatically reduced for selected velocities which differ considerably from the mean velocity. Thus, the suppression of possible perturbations has to be paid for by a reduced number of contributing absorbers and hence by a reduced signal.

6.3 Cooling

The most rigorous approach for reducing the velocity of the interrogated particles and, hence, the Doppler effect in all orders and for increasing the interaction time, is to cool the particles. Cooling rather than selection of the slowest particles furthermore has the advantages that in general more particles contribute to the signal.

6.3.1 Laser Cooling

The advances in laser cooling and the development of cheap and convenient tuneable lasers has made laser cooling [132] one of the most efficient methods of cooling atoms for use in optical frequency standards. The description of laser cooling and the applied methods differ somewhat if one considers free atoms or particles bound to a trap [158]. The common underlying principle is that on the average the energy of the radiation absorbed by the particles is smaller than the energy irradiated by the particles during the subsequent emission. The energy difference is supplied by the kinetic energy of the particles leading to a reduction of the mean value and width of the velocity distribution. Soon after the invention of tuneable lasers, this effect was proposed for reducing the velocity of atoms in a gas by Hänsch and Schawlow [159] or of trapped ions by Wineland and Dehmelt [160]. Here, we consider laser cooling of free atoms and leave the treatment of laser cooling of bound particles to Section 10.2.2.3.

6.3.1.1 Optical Molasses

Consider processes where a two-level atom with the two states E_g and E_e absorbs a photon having a momentum $\hbar\vec{k}$ from a laser beam with wave vector $\vec{k}$ and subsequently emits a photon by spontaneous emission. The laser is red detuned with respect to the transition frequency $(E_e - E_g)/h$. The momentum of the photon is transferred to the atom thereby changing

the momentum $\vec{p} = m\vec{v}$ of the atom. The Doppler shift $\Delta\nu = p/(m\lambda)$ associated with the absorption is assumed to be small compared to the natural width $\gamma = 1/(2\pi\tau)$ of the absorption line where τ is the lifetime of the excited state, i.e., $\Delta\nu = h/(m\lambda^2) \ll \gamma$. In this approximation, the momentum transfer $\Delta\vec{p} = \hbar\vec{k}$ can be averaged over many processes of absorption and re-emission, leading to a classical force $\vec{F}$ onto the atom. The photons from the spontaneous decay are emitted isotropically and, hence, do not contribute to the force, whereas the averaged force due to the absorption of the photons is

$$\vec{F} = \frac{N_e}{N}\frac{\hbar\vec{k}}{\tau}. \tag{6.4}$$

Here, N_e is the averaged number of atoms in the excited state, $N = N_e + N_g$ is the total number of atoms including the number N_g of atoms in the ground state. The ratio N_e/N can be expressed by the saturation parameter (5.119) leading to

$$\vec{F} = \frac{\hbar\vec{k}}{2\tau}\frac{S_0}{1 + S_0 + \left(\frac{\delta\nu}{\gamma/2}\right)^2}. \tag{6.5}$$

At low enough laser irradiance, i.e. $S_0 \ll 1$, the spontaneous force (see (6.5)) shows a Lorentzian profile as a function of the detuning $\delta\nu$ determined by the natural linewidth of the atomic transition. For a particular atom with velocity $\vec{v}$, where the detuning depends on the velocity of the atom, the laser frequency in the reference frame of the atom is shifted by the Doppler effect, and the detuning is $\delta\nu = \nu - \nu_0 - \vec{k}\cdot\vec{v}/(2\pi)$. Next we consider the movement of an atom with a velocity $\vec{v}$ in the field of two counter-propagating laser beams of equal intensities realised, e.g., in a retro-reflected laser beam. In the low-intensity limit ($S_0 \ll 1$) the forces of the two counter-propagating light waves acting on the atom can simply be added to yield

$$\begin{aligned}\vec{F}_{\rm om} &= \frac{\hbar\vec{k}}{2\tau}\left(\frac{S_0}{1 + S_0 + 4(\nu - \nu_0 - \frac{\vec{k}\cdot\vec{v}}{2\pi})^2/\gamma^2} - \frac{S_0}{1 + S_0 + 4(\nu - \nu_0 + \frac{\vec{k}\cdot\vec{v}}{2\pi})^2/\gamma^2}\right)\\ &= \frac{\hbar\vec{k}}{2\tau}S_0\frac{16(\nu - \nu_0)\frac{\vec{k}\cdot\vec{v}}{2\pi\gamma^2}}{\left[1 + S_0 + \frac{4(\nu-\nu_0)^2}{\gamma^2} + \left(\frac{k^2v^2}{\pi^2\gamma^2}\right)\right]^2 - \left[8(\nu - \nu_0)\frac{\vec{k}\cdot\vec{v}}{2\pi\gamma^2}\right]^2}.\end{aligned} \tag{6.6}$$

Fig. 6.2 shows the force for a saturation parameter of $S_0 = 0.3$ and a frequency of the laser ν that is red detuned from the resonance ν_0 by a linewidth γ, i.e. $\nu - \nu_0 = -\gamma$.

For very low velocities ($v < \gamma\lambda$) in (6.6) the terms $(\vec{k}\cdot\vec{v}/\gamma^2)^2$ and higher can be neglected and one finds

$$\vec{F}_{\rm om} \approx \frac{8\hbar k^2 S_0(\nu - \nu_0)}{\gamma\left(1 + S_0 + \frac{4(\nu-\nu_0)^2}{\gamma^2}\right)^2}\vec{v} \equiv \alpha\vec{v}, \tag{6.7}$$

i.e., near zero velocity the combined force varies monotonically with the velocity of the atoms. For red detuning ($\nu - \nu_0 < 0$) one finds $\vec{F}_{\rm om} = -\alpha\vec{v}$, which represents a friction force. Hence, as a result of the Doppler effect, atoms moving with a velocity $\vec{v}$ are more resonant with the

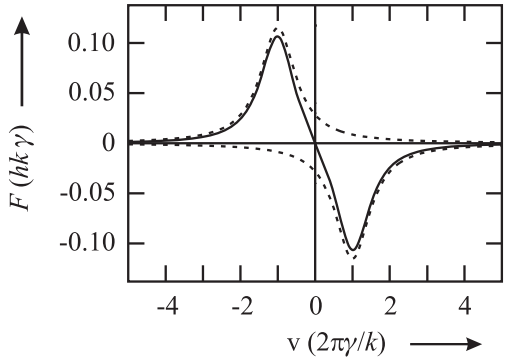


Figure 6.2: The velocity-dependent force exerted on an atom due to the near-resonant absorption of photons according to (6.6) with the parameters $S_0 = 0.3$ and $\nu - \nu_0 = -\gamma$.

counter-propagating beam as compared to the co-propagating beam. As a consequence, the atom is decelerated by the viscous damping in the red detuned near-resonant light field [161, 162] and the name "optical molasses" has been coined to describe the damping interaction of the light field on the moving atom.

6.3.1.2 Doppler Limit

One might expect that the residual motion of the atoms gradually decreases and the atoms come to rest and reach a temperature $T = 0$. This clearly unphysical result does not take into account that even an atom at rest would absorb and emit photons. The recoil energy transferred to each atom during the absorption process of $+\hbar^2k^2/(2m)$ and the emission process of $-\hbar^2k^2/(2m)$ leads to heating which in total corresponds to a mean increase in the kinetic energy per particle by $2\hbar^2k^2/(2m)$ (see (5.103)).[2] In equilibrium

$$\dot{E}_{\mathrm{heat}} = -\dot{E}_{\mathrm{cool}} \tag{6.8}$$

holds. The heating rate $\dot{E}_{\mathrm{heat}}$ for each of the two beams due to the rate of the recoil transfers is proportional to the fraction of atoms in the excited state N_e/N (see (5.123)) and to the decay rate $1/\tau = 2\pi\gamma$ (see (2.37)). Hence, the heating rate is

$$\dot{E}_{\mathrm{heat}} = 2\frac{(\hbar k)^2}{2m}\frac{2\pi\gamma}{2}\frac{2S_0}{1 + 2S_0 + 4(\nu - \nu_0)^2/\gamma^2} \tag{6.9}$$

where we have assumed that the on-resonance saturation parameter in the two counter-propagating beams is $2S_0$. The cooling rate due to the loss of kinetic energy resulting from the damping is calculated from

$$\dot{E}_{\mathrm{cool}} = \frac{\partial}{\partial t}\frac{p^2}{2m} = \dot{p}\frac{p}{m} = F(v)v = -\alpha v^2. \tag{6.10}$$

[2] Averaging is required as the atom undergoes a random walk in momentum space as a result of the spontaneous re-emission processes [158] where the value given holds for isotropic emission.

Insertion of (6.9) and (6.10) together with (6.7) into (6.8) and replacing v^2 with its mean value $\langle v^2 \rangle$ leads to

$$m\langle v^2 \rangle = \frac{h\gamma}{4} \frac{\left[1 + 2S_0 + (2(\nu - \nu_0)/\gamma)^2\right]}{2(\nu - \nu_0)/\gamma}. \tag{6.11}$$

(6.11) has a minimum for $\nu - \nu_0 = \gamma/2$ from which one derives by use of $m \left\langle v^2 \right\rangle /2 = k_B T/2$ the minimal temperature in the limit $S_0 \to 0$ as

$$T_D = \frac{h\gamma}{2k_B} = \frac{\hbar\Gamma}{2k_B}. \qquad \text{(Doppler limit)} \tag{6.12}$$

The Doppler temperature T_D represents the minimal temperature that can be reached by means of Doppler cooling and hence is often referred to as the Doppler limit.

The Doppler limit in three dimensions is derived similarly [162]. It suffices to note that the cooling rate is the same as in the one-dimensional case but the heating rate is three times higher if six beams are used rather than two. At the same time, however, the three degrees of freedom result in $m \left\langle v^2 \right\rangle_{3\mathrm{D}} /2 = 3k_B T/2$ and the Doppler limit in three dimensions is the same as in (6.12). For typical cases like the Cs $6\,^2\mathrm{S}_{1/2} - 6\,^2\mathrm{P}_{3/2}$ transition ($\lambda = 852$ nm, $\gamma = 5.18$ MHz) and the Ca $4\,^1\mathrm{S}_0 - 4\,^1\mathrm{P}_1$ transition ($\lambda = 423$ nm, $\gamma = 34.6$ MHz), the Doppler temperature (6.12) is 0.12 mK and 0.83 mK, respectively. The velocity corresponding to the Doppler limit is calculated from $1/2mv_D^2 = k_B T_D/2$ as

$$v_D = \sqrt{\frac{h\gamma}{2m}}. \tag{6.13}$$

For the cases given above this velocity corresponds to $v_{D,\,\mathrm{Cs}} = 8.82$ cm/s and $v_{D,\,\mathrm{Ca}} = 41.5$ cm/s.

6.3.1.3 Sub-Doppler Cooling

The velocities, determined by the Doppler limit of suitable resonance lines for atoms used in frequency standards, range from several cm/s to several tens of cm/s which in cases where a long interaction time is required as, e.g., in an atomic fountain (Section 7.3) would lead to a high loss of atoms. Fortunately, in the case of atoms with magnetic or hyperfine split ground states, e.g., the alkali atoms Cs and Rb used in atomic clocks (Section 7) and (Section 8.2), there are a number of mechanisms (see e.g. [132]) that conveniently allow one to reach lower temperatures. As an example we consider briefly the case of "Sisyphus cooling" where the atoms move in a laser field whose polarisation has strong gradients (see Fig. 6.3 a) as it changes completely over half a wavelength. Such a polarisation gradient is achieved, e.g., in two counter-propagating beams of the same frequency and amplitude, but whose linear polarisations are perpendicular in a so-called "lin $\perp$ lin" arrangement

$$\begin{aligned} \vec{E} &= E_0 \hat{x} \cos(\omega t - kz) + E_0 \hat{y} \cos(\omega t + kz) \\ &= E_0 \left[(\hat{x} + \hat{y}) \cos\omega t \cos kz + (\hat{x} - \hat{y}) \sin\omega t \sin kz \right]. \end{aligned} \tag{6.14}$$

From (6.14) one finds (see Fig. 6.3 a) that for $kz = 0$ there is a linear polarisation under 45° between $\hat{x}$ and $\hat{y}$. It changes to a perpendicular polarisation for a quarter of a wavelength $kz =$

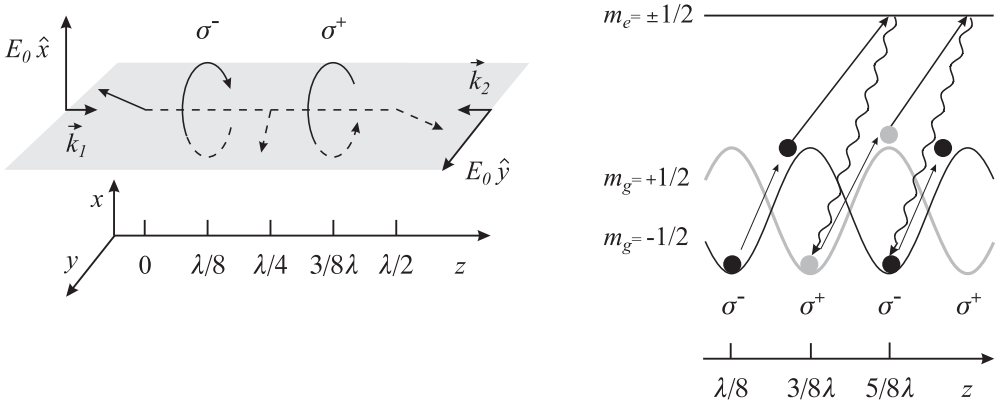


Figure 6.3: Sisyphus cooling. a) The polarisation in a "lin ⊥ lin" standing wave. b) The light shift in the standing wave of Fig. 6.3 a) modulates the energies of the ground states ($m_g = +1/2$ and $m_g = -1/2$) and leads to a spatially periodic modulation of the optical pumping.

$\pi/2$ with circular polarisation at $kz = \pi/4$, i.e., at $z = \lambda/8$. In such a field the $m_{\pm 1/2}$ ground states of an atomic system with a $J_g = 1/2 \rightarrow J_e = 3/2$ transition experience different light shifts which change with position, due to the position-dependent polarisation (see Fig. 6.3 b). Consider an $m_g = -1/2$ ground state atom which has the lower potential energy in the valley at $z = \lambda/8$. If this atom moves to the $+z$ direction it has to climb up the potential hill thereby lowering its kinetic energy. Near the top of the hill the radiation becomes σ^+ and the atom is predominantly optically pumped into the $m_g = +1/2$ state via the $m_e = +1/2$ state. Moving on again, the atom loses further kinetic energy while climbing up the next hill from where it is optically pumped by the σ^- radiation via the $m_e = -1/2$ state to the $m_g = -1/2$ state. Referring to the ancient Greek fairy tale where Sisyphys had to continuously move a stone uphill which escaped there and rolled down again, the process is referred to as Sisyphus cooling. The process works best if the average time needed for the pumping process corresponds to the time an atom takes to move over the distance of $\lambda/2$. Besides the "lin ⊥ lin" standing wave, other combinations like the $\sigma^+ - \sigma^-$ counter-propagating waves exhibit polarisation gradients, too. Furthermore, constant polarisation in combination with a magnetic field can lead to magnetically induced sub-Doppler cooling [132].

The minimum achievable temperature in a three-dimensional optical molasses for Cs, where sub-Doppler cooling mechanisms are present, is around 2.5 µK which is considerably lower than the Doppler limit of 0.12 mK, but slightly higher than the temperature corresponding to the recoil energy

$$k_B T > E_r = (\hbar k)^2/2m. \qquad \text{(recoil limit)} \qquad (6.15)$$

In cases where there is no ground-state splitting as, e.g., in the even alkaline earth isotopes like ^{20}Mg, ^{40}Ca, ^{88}Sr used as optical frequency standards, the velocity can be reduced below the Doppler limit of the strong resonance line by a second-stage Doppler cooling on a narrow

"forbidden" line [163] (Sr). Even in cases where the corresponding linewidth is so small that the cooling force is too weak to efficiently counteract gravity, schemes have been found to achieve temperatures in the microkelvin regime [164, 165] (Ca), however, at the expense of additional lasers.

Thus the recoil limit (6.15) constitutes a much more severe limit to the achievable temperature than the Doppler limit. Nevertheless, several clever schemes have been devised that allow one to achieve temperatures even below the recoil limit by using methods of coherent population trapping [166, 167] or Raman transitions with tailored detuning sequences [168]. Up to now, however, these latter techniques have not found wide applications in frequency standards.

6.3.2 Cooling and Deceleration of Molecules

In contrast to atoms, the cooling of molecules by conventional laser-cooling techniques meets with difficulties since after laser excitation to an excited state a molecule can radiatively decay into myriads of ro-vibrational states of the deeper lying electronic state and hence, no cycling transitions exist. Consequently, for the most advanced frequency standards molecular references have become of less importance. This situation might change if suitable cooling mechanisms will be applied, some of which will be discussed in the following.

Cooling with Buffer Gas Paramagnetic atoms and ions can be trapped in a magnetic trap with a minimum of the magnetic field. Doyle and co-workers at Harvard used such a trap to trap various atoms and molecules and cool them by collisions with cryogenically cooled ^{3}He at a temperature around 0.3 K [169]. In their apparatus, two superconducting magnetic coils generate a spherical quadrupole field trap. The atoms or molecules to be investigated are set free by laser ablation from a solid target. The particles lose kinetic energy by collisions with the helium atoms. Inside the trap, atoms or molecules with their magnetic moment anti-parallel to the trapping field are attracted into the region of the low field in the centre of the trap. The "high-field seeking particles" (Section 6.4) with their magnetic moment parallel to the magnetic field are lost. Large numbers of atoms of up to 10^{12} and as many as 10^8 molecules (CaH) have been trapped. For frequency standards with neutral atoms or molecules, buffer gas cooling has not been used yet in contrast to ions (see Section 10.2.2.2).

Electrostatic Deceleration The interaction of the dipole moment of molecules with a time-varying electric field has been used to decelerate these particles [170, 171]. In this method a bunch of molecules to be slowed down pass through an array of electrode pairs which are oriented perpendicularly to the trajectory of the molecular beam. Low-field seeking molecules, with their electric dipole moments anti-parallel to the gradient of the electric field of the first electrode pair, are decelerated when penetrating into the high electric field between the electrodes. When the pulse of molecules arrives at the region of maximum field between the electrodes, the electric field of this electrode pair is rapidly switched off and the molecules have to climb up the potential hill again which is created by the next pair of electrodes that are connected to a high voltage of several kilovolts. Hence, the low-field seeking molecules lose their kinetic energies and can later be stored in a trap.

Photoassociation of Molecules Molecules which are with much lower kinetic energies than those achieved by buffer gas cooling or electrostatic deceleration can be produced from ultracold atoms by means of photoassociation. As shown in Fig. 5.4, the potential energies of the ground state and a particular excited electronic state of two atoms that collide with each other, are functions of the internuclear distance of the two atoms. A properly tuned laser beam which is red detuned with respect to the transition frequency of the free atom can put the molecule into one of the bound excited states. The molecules formed by this photoassociation process can radiatively decay into a ro-vibrational bound state of the ground state for close distances of the atoms, when the Franck–Condon factors are high enough. This method is mainly restricted to the formation of dimers from cold atoms as, e.g., alkali atoms [172, 173]. In general this method produces molecules in high vibrational states but the use of two lasers can access very low vibrational states [174]. Extremely cold diatomic molecules can be produced by starting from a Bose Einstein condensate [175].

Presently the particular type of cold molecules accessible by these techniques are not the best candidates for frequency standards and it remains to be shown to what extent these techniques can be applied to molecules more suitable for the latter purpose.

6.4 Trapping of Atoms

For frequency standards it is often highly desirable to keep the absorbers in a well defined location during interrogation. Electric, magnetic, gravitational and light forces can be used to manipulate the external degrees of freedom of ions, atoms, or molecules in order to confine them to the desired region of space. Several restrictions, however, exclude the design of stable traps of particular design. In a volume free of electric charge $\Delta\Phi = 0$ holds [3] and consequently no configuration of static electric fields can be designed that exhibits maxima or minima of the electrostatic potential Φ. This fact is sometimes referred to as Earnshaw's theorem. As a consequence of Earnshaw's theorem no electrostatic ion trap can be constructed. Wing [176] has shown that in a space region, free of currents and charges, there are no maxima of the moduli of electric or magnetic fields. Consequently, no electro- or magneto-static trap can be set up for neutral atoms in the lowest energy state. Wing's theorem has been generalised by Ketterle and Pritchard [177] to any combination of static electrical, magnetic and gravitational fields. Hence, inclusion of the graviational field does not change these findings. Ashkin and Gordon [178] have derived a so-called "optical Earnshaw theorem" stating that no stable trap can be constructed based on time-independent optical fields where the force on an atom is proportional to the irradiance.

Atomic or molecular ions are easily confined in ion traps where Earnshaw's theorem does not exclude the use of an electric potential having saddle points where positive and negative gradients are generated with high frequency. The electric force $\vec{F} = q\vec{E}$ is strong enough to create a deep trapping potential of a few electron volts (1 eV $\hat{=}$ 11 600 K) that easily confines ions of thermal energies. We leave the discussion of such ion traps and the frequency standards based thereon to Section 10.1.1.

[3] If the total charge density ρ in the trapping volume is zero, Maxwell's equation (4.25) $div\vec{E} = \vec{\nabla}\cdot\vec{E} = \vec{\nabla}\cdot\vec{\nabla}\Phi = \Delta\Phi = \rho/\epsilon_0 = 0$ holds.

The forces acting on neutral atoms and molecules are much weaker. They are based on the interactions (5.33) and (5.34) of external electric or magnetic field gradients with permanent or induced electric or magnetic moments. Due to their inversion symmetry, unperturbed atoms cannot have a permanent electric dipole moment and hence atoms can be trapped in electric fields only via induced dipole moments. On the other hand, atoms can be easily prepared in states where they have a magnetic moment. The applied field shifts the energy levels of the atoms. Any spatial gradient of this energy level shift, i.e., of the potential, leads to a force on the centre-of-mass motion of the atom. As a result of the perturbation by the external field the energy of atoms in the ground state is lowered and those atoms are accelerated towards regions of higher field and are sometimes referred to as "high-field seekers". Atoms in excited states, however, can be "low-field seekers" which are attracted to regions of lower fields. As a consequence of Wing's theorem in static magnetic [179] or electric traps [180] only low-field seekers can be stored [177].

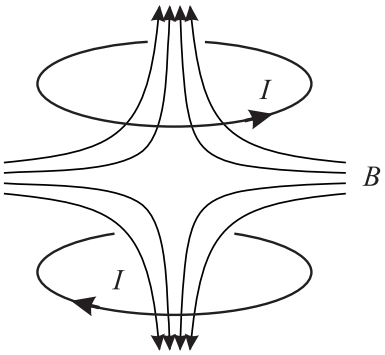


Figure 6.4: Magnetic quadrupole trap formed by a pair of coils in an anti-Helmholtz configuration.

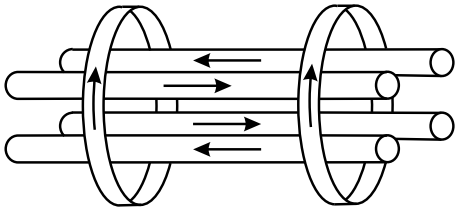


Figure 6.5: Magnetic Ioffe trap realised with four linear rods carrying a current and two end coils. The arrows indicate the direction of the currents.

The most simple magnetic trap, among other possibilities [181], can be realised by an anti-Helmholtz coil configuration (Fig. 6.4). These coils generate a magnetic field with radial symmetry in the x and y plane and a zero field at the centre. Near the centre the magnetic induction varies linearly ($B_x = \{\partial B_x/\partial x\} \cdot x, B_y = \{\partial B_y/\partial y\} \cdot y, B_z = \{\partial B_z/\partial x\} \cdot z$). From $div\vec{B} = \vec{\nabla} \cdot \vec{B}(\vec{r})$ (see (4.26)) it follows that $2\partial B_x/\partial x = 2\partial B_y/\partial y = -\partial B_z/\partial z$. Hence, the field gradient is twice as large along the z direction as compared to the x and y directions and in the opposite direction. Neutral atoms were trapped for the first time in such a trap [179]. A quadrupole magnetic trap with its zero-field at the centre has the disadvantage that atomic spins have a non-vanishing probability to not follow adiabatically when passing through this zero field region and may suffer from so-called Majorana spin flips [182]. By this effect low-field seekers are converted into high-field seekers and are expelled from the trap. A solution to this problem represents the Ioffe trap with an additional bias field. It can be realised, e.g., by a two-dimensional magnetic quadrupole field generated by four current carrying wires (Fig. 6.5) [181]. Confinement along the longitudinal direction is achieved by two current loops such that there is no longer a zero field in the centre line. These traps are

shallow and, hence, can only be used to confine neutral particles with very low temperature. Larger potential depth can be achieved with dynamical traps, e.g., the time-averaged orbiting potential trap (TOP trap) [183] which also is a trap for low-field seekers.

Atoms without a magnetic dipole moment can be trapped using the electric dipole moment induced in an electric field [148]. Very high electric fields, and hence sufficiently strong gradients, can be generated in a focussed laser beam. For a laser beam whose frequency is red detuned from resonance (Fig. 6.6) the energy of the ground state in a two-level system is lowered whereas the energy of the excited state is increased. For red and blue detuning the induced dipole oscillates in-phase and out of phase and the atom experiences a force pulling it into and out of the maxima of the irradiance, respectively.

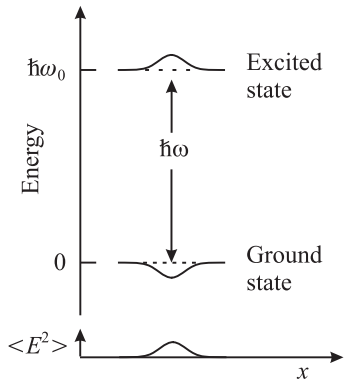


Figure 6.6: The interaction of a two-level atom with the spatially dependent field distribution of a near-resonant laser beam, leads to a spatially dependent "light shift" of the atomic levels.

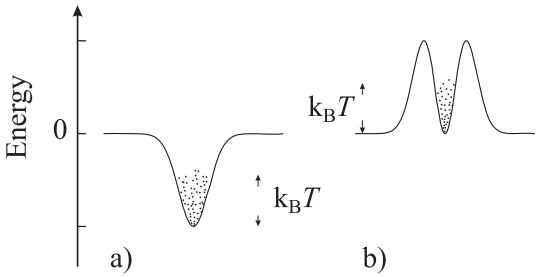


Figure 6.7: Schematic illustration of dipole traps with red (a) and blue (b) detuning. The red detuned dipole trap can be realised by a simple Gaussian laser beam. The blue detuned dipole trap can be realised, e.g., by a Laguerre–Gaussian LG_{01} "doughnut" mode.

The potential energy of an atom in the laser beam with the electric field amplitude E_0 is derived from (5.135) and (5.134) as

$$\begin{aligned} W_{\rm dip}(r,z) &= -\frac{6\pi\epsilon_0 c^3}{\omega_0^2}\frac{\Gamma\left(\omega_0^2-\omega^2\right)}{\left(\omega_0^2-\omega^2\right)^2+\omega^6\Gamma^2/\omega_0^4}\frac{E_0^2}{4} \\ &\approx -\frac{3\pi\epsilon_0 c^3}{4\omega_0^3}\left[\frac{\Gamma}{\omega_0-\omega}+\frac{\Gamma}{\omega_0+\omega}\right]E_0^2 \approx \frac{\hbar}{8}\frac{\Gamma^2}{\omega-\omega_0}\frac{I(r,z)}{I_{\rm sat}}. \end{aligned} \tag{6.16}$$

In (6.16) we have assumed first that the detuning is much larger than the line width ($\omega-\omega_0 \gg \Gamma$). Furthermore we have applied the rotating wave approximation by neglecting the second term in square brackets and introduced the irradiance of the laser beam ($I(r,z)=(\epsilon_0 c/2)E_0^2$) and the saturation intensity $I_{\rm sat}$ given by (5.122). The most simple optical trap for red detuning (see Fig. 6.7 a) is realised by a focussed Gaussian laser beam (4.110) which generates a three-

dimensional spatial maximum of the irradiance at the waist. From (4.117) we find

$$I(r,z) = \frac{2P}{\pi w_0^2(1+\frac{z^2}{z_R^2})} \exp\left[-\frac{2r^2}{w_0(1+\frac{z^2}{z_R^2})}\right] \approx \frac{2P}{\pi w_0^2}\left(1-\frac{2r^2}{w_0^2}-\frac{z^2}{z_R^2}\right), \qquad (6.17)$$

where P defines the power in the Gaussian beam with waist w_0 and $z_R = \pi w_0/\lambda$ is the Rayleigh range. The approximation in (6.17) holds for small distances from the centre of the waist, i.e., $z < z_R$ and $r < w_0$ and shows that there is an harmonic potential along the z and r directions.

In contrast to the spontaneous force, the maximum dipole force does not saturate. On the other hand, spontaneous emissions in the dipole trap lead to radiative heating which is proportional to the rate of scattered photons. The scattering rate $\Gamma_{\rm sc}$, i.e. the number of photons that are scattered per second by the atoms, can be calculated from (5.136) and (5.134) with the same approximations as before leading to

$$\Gamma_{\rm sc} = \frac{P_{\rm abs}}{\hbar\omega} = -\frac{2}{\hbar}\Im\mathrm{m}\{\alpha\}\frac{E_0^2}{4} \approx \frac{\Gamma^3}{8(\omega-\omega_0)^2}\left(\frac{\omega}{\omega_0}\right)^3\frac{I}{I_{\rm sat}}. \qquad (6.18)$$

The scattering rate and hence the heating becomes less important for a large detuning $\omega-\omega_0$ since the scattering rate decreases as $(\omega-\omega_0)^{-2}$ (see (6.18)). Thus red-detuned traps are most often operated as Far Off-Resonant Traps (FORT) [184].

Blue detuned traps do not suffer from these disadvantages. The blue detuned dipole trap realised, e.g., by a Laguerre–Gaussian LG_{01} "doughnut" mode (Fig. 6.7 b) can provide the same potential depth and the same curvature in the trap centre as the red detuned dipole trap. However, for the same detuning of both trapping laser fields, the laser power has to be increased for the blue detuned trap by a factor of e^2 [148]. Three-dimensional blue-detuned traps have also been generated by light sheets [185] or as a pyramid trap [186]. The blue detuned trap, suitable for low-field seekers, is more advantageous for frequency standards as the ac Stark shift is several orders of magnitude smaller than in a red detuned trap. In a blue detuned trap Davidson *et al.* [185], e.g., were able to observe the hyperfine clock transition in sodium atoms by means of Ramsey excitation (Section 6.6).

6.4.1 Magneto-optical Trap

In an optical molasses atoms are decelerated to very low velocities. However, there is only a damping force but no force that binds the atoms to a particular point in space. Such a force can be generated, e.g., in an inhomogeneous magnetic field. Consider an atom with a ground state E_g with total angular momentum $J=0$ and an excited state E_e with $J=1$ (Fig. 6.8). Such conditions are readily encountered in a number of cases (see Table 5.6) as, e.g., the alkaline earth metals. In a magnetic field the energy of the ground state is not affected to a good approximation whereas the energy level of the excited state is split into three magnetic sub-states. In contrast to the $m_J=0$ sub-state whose energy is almost independent of the magnetic field, the energies of the two other sub-states ($m_J=\pm1$) vary linearly but with opposite signs, with an applied magnetic field. Suppose that the magnetic induction B varies

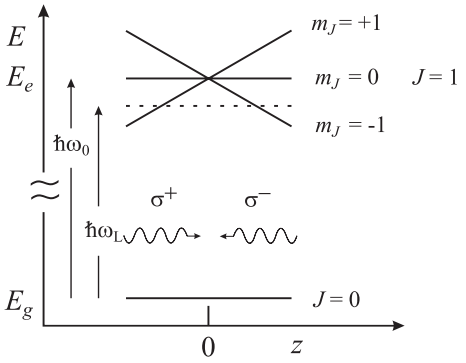


Figure 6.8: Energy levels of an atom in a magneto-optical trap.

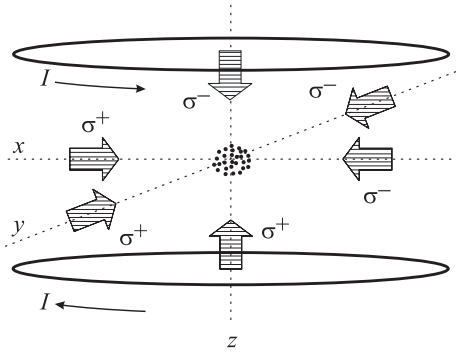


Figure 6.9: Schematics of a magneto-optical trap.

along the z direction linearly with distance z from a centre $z = 0$ as

$$B_z(z) = bz. \tag{6.19}$$

The corresponding Zeeman shift of the energy of the $m_J \neq 0$ excited states

$$\Delta E(z) = \pm g_J \mu_B bz \tag{6.20}$$

introduces a spatially dependent term into the detuning

$$\delta\nu = \nu - \nu_0 \mp \frac{v}{\lambda} \mp \frac{g_J \mu_B}{h} bz \tag{6.21}$$

where g_J is the Landé factor of the excited state and μ_B is the Bohr magneton ($\mu_B/h = 1.4 \times 10^{10}$Hz/T). Along the z axis, transitions to the $m_J = 1$ and $m_J = -1$ can be excited selectively by means of circularly polarised σ^+ and σ^- radiation, respectively. If one performs the calculation of (6.7) with the spatially dependent term rather than with the velocity dependent detuning term of (6.21), one ends up with a force that is linear in z

$$F_z(z) = -Dz \tag{6.22}$$

where the constant D is given as

$$D \approx \frac{8\mu_B bk S_0(\nu - \nu_0)}{\gamma\left(1 + S_0 + \frac{4(\nu-\nu_0)^2}{\gamma^2}\right)^2}. \tag{6.23}$$

As a result of this force, resembling Hooke's law for a small distance z from the centre, there is a harmonic potential $V(z) = Dz^2/2$ which is capable of trapping the atoms. Given that the two laser beams have the same irradiance, the centre of the trap coincides with the zero of the magnetic field. The combined force including the damping resulting from the optical molasses and the harmonic potential resulting from the spatially varying magnetic field, is given by

$$F_z(z, v) = -Dz - \alpha v. \tag{6.24}$$

The corresponding one-dimensional equation of motion for an atom of mass m is described by a damped linear harmonic oscillator with the angular frequency of the undamped oscillation with $\omega_0 = \sqrt{\frac{D}{m}}$ and the damping constant $\Gamma = \frac{\alpha}{m}$ (see (2.27)).

The extension of this scheme to a three-dimensional magneto-optical trap (MOT) is straightforward [187] when for each spatial dimension a pair of laser beams with the proper circular polarisation is employed (Fig. 6.9). A magnetic field that vanishes at the centre of the trap and increases approximately linearly with the distance from the centre can be generated by a pair of anti-Helmholtz coils (Fig. 6.4) with gradients of the magnetic induction varying between about 0.05 T/m and 0.5 T/m. To calculate typical values for the angular frequency ω_0 and the damping constant Γ we use (6.23) and (6.7) and a trap for ^{40}Ca atoms with $b = 0.1$T/m, $\omega - \omega_0 = \Gamma/2$, $k = 2\pi/423\,\text{nm}$ and $m = 40 \times 1.66 \times 10^{-27}\,\text{kg}$. With $\omega_0 \approx 2\pi \times 2.4\,\text{kHz}$ and $\Gamma \approx 1.56 \times 10^5$/s one finds that the motion of the atoms in a MOT is strongly overdamped.

Loading of a Magneto-optical Trap The maximal velocity of the atoms which can be captured in a MOT is $v_c \approx (2F_{\max}r/m)^{1/2} = (\hbar k\gamma r/m)^{1/2}$ [188] where r is the radius of the MOT. Hence atoms with velocities up to $v_c \lesssim 30\,\text{m/s}$ can be loaded directly from an uncooled vapour [189] or from an uncooled thermal atomic beam [104, 187, 190, 191]. However, in such set-ups only atoms from the low velocity tail of the Maxwell–Boltzmann distribution are captured. In order to obtain a good duty cycle between preparing and interrogating the atoms, in optical frequency standards it is desirable to load a large number of atoms in a short time. The number N of atoms trapped in a magneto-optical trap can be derived from a rate equation

$$\frac{dN}{dt} = R_c - \frac{N}{\tau_{\text{MOT}}} - \beta N^2 \tag{6.25}$$

where R_c is the capture rate and τ_{MOT} is the average time of an atom spent in the MOT. The second term represents the collisions of the trapped atoms with the background gas and the third term describes collisions between the trapped atoms themselves. (6.25) is solved if the last term which becomes important only at high densities can be neglected by

$$N(t) = (N(0) - R_c\tau_{\text{MOT}})e^{-t/\tau_{\text{MOT}}} + R_c\tau_{\text{MOT}}. \tag{6.26}$$

The loading curve (see Fig. 6.10) approaches the equilibrium value $N(t \to \infty) = R_c\tau_{\text{MOT}}$ with the time constant τ_{MOT}. Starting with an empty trap $(N(0) = 0)$ the loading curve is

$$N(t) = R_c\tau_{\text{MOT}}\left(1 - e^{-t/\tau_{\text{MOT}}}\right). \tag{6.27}$$

The capture range and hence the flux of atoms loaded into the trap can be increased considerably by employing two or more optical frequencies separated from each other by about a natural linewidth of the cooling transition [190, 192].

Alternatively, a Zeeman cooling technique [193, 194] is often employed to load a MOT [195–198] where a circularly polarised cooling laser beam counter-propagates to the atoms in a beam whose changing Doppler shift is compensated by a longitudinal magnetic field. In frequency standards, sometimes the Zeeman slower is used in conjunction with a tilted molasses to deflect only the slowest atoms into the MOT [199]. Dense samples ($\rho \geq 10^{10}$ atoms/cm^3 of

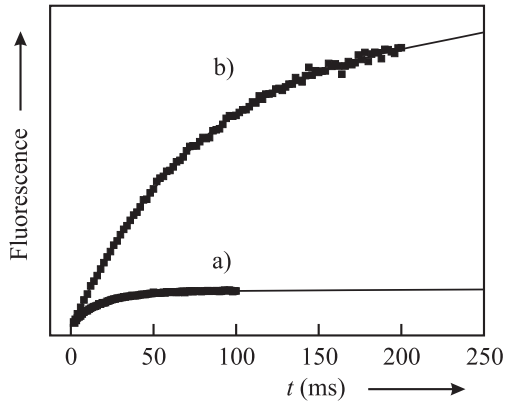


Figure 6.10: Measured loading curves (squares) of a Ca magneto-optical trap and fits (lines) according to (6.27). When a particular trap loss channel is closed (where the atoms lost into the $^1D_2$ state (Fig. 5.2) are re-pumped) the lifetime increases from $\tau_{\rm MOT} = 19\,$ms (a) to $\tau_{\rm MOT} = 83\,$ms (b).

$N \gg 10^7$ atoms are easily cooled in a MOT which after switching off the MOT fields (lasers and magnetic field) can serve as a cold ballistic ensemble for interrogation in microwave and optical frequency standards.

6.4.2 Optical lattices

Atoms can also be trapped in an optical lattice. An optical lattice is generated by the interference between two or more light fields, that leads to a stationary field pattern in space. Consider a standing wave created by the interference between two counter-propagating waves of the same frequency $\nu = c/\lambda$, polarisation and irradiance

$$\begin{aligned} \vec{E} &= E_0\hat{\epsilon}\cos(\omega t - kz) + E_0\hat{\epsilon}\cos(\omega t + kz) \\ &= 2E_0\hat{\epsilon}\cos kz\cos\omega t. \end{aligned} \tag{6.28}$$

The light-shift (ac Stark shift) potential ($\propto E^2$) experienced by atoms in the region of the interference also forms a one-dimensional periodic potential structure with the nodes and anti-nodes separated by $\lambda/4$. Atoms with low enough kinetic energy may become trapped in the corresponding potential wells and may be localised to a region significantly smaller than the wavelength of the light creating the potential wells. A two-dimensional optical lattice can be realised in the intersection of two perpendicular laser beams (Fig. 6.11) [200]. The interference pattern depends on the orientation of the polarisation of the beams and shows a regular arrangement of potential wells (Fig. 6.12).

It has been found that lattices in n dimensions may be formed by $n+1$ laser beams [201] with the appropriate choice of polarisations. In three dimensions a simple and convenient realisation is shown in Fig. 6.13. The two penetrating beams from the left in the $x-y$ plane and from the right in the $y-z$ plane form a standing wave in the x and z direction, respectively. Along the y direction a standing wave is formed by the interference between all four beams.

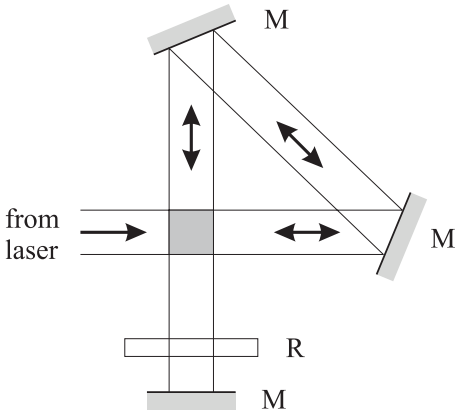


Figure 6.11: Generation of a two-dimensional optical lattice. M: mirror, R: optional retarding plate.

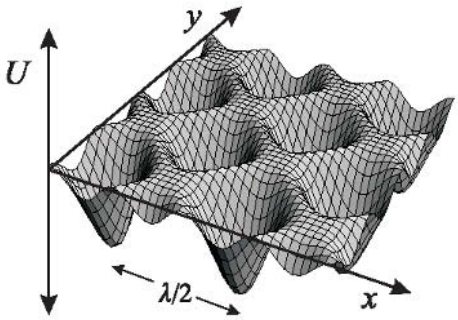


Figure 6.12: Potential of a two-dimensional optical lattice.

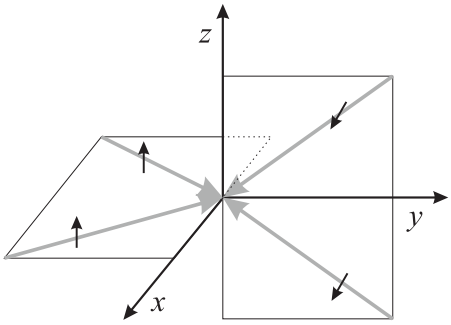


Figure 6.13: Four-beam geometry used to produce a three-dimensional optical lattice. The two beams from the left and the two beams from the right are propagating within two orthogonal planes. All beams are linearly polarised with the polarisation perpendicular to the respective plane as indicated by the arrows.

By filling an optical lattice from a cold source of atoms, in general, a very dilute occupation of the lattice sites is achieved. There are, however, methods that allow the achievement of densities near unity occupation of the lattice sites [202]. For use in an optical frequency standard, optical lattices could allow very long interaction times, even in the presence of the gravitational field, also with neutrals confined to the Lamb–Dicke regime where the first-order Doppler effect can be eliminated. A promising method has been devised [203] where the perturbations of the clock transition by the lattice laser beams can be avoided (Section 14.2.2).

6.4.3 Characterisation of Cold Atomic Samples

The performance of a frequency standard usually depends in various ways on particular properties of the ensemble (see Section 5.4) to which its frequency is stabilised. A variety of methods and techniques are routinely employed [204] to measure the number of absorbers, their density or their temperature, some of which will be discussed in the following.

6.4.3.1 Measurement of the Number and Density of the Particles

The total number of particles forming the sample is usually determined by one of three methods, namely from measurements of the power of the fluorescent light emitted by the sample, from the light absorbed by the sample or from the phase shift that light suffers when transmitted through the sample. The spatial variation of the corresponding signals also allows one to determine the size and distribution of the atoms in the cloud. A weak probe laser beam passing through a medium along the z axis experiences the influence of a complex index of refraction $n = n' + in''$ as

$$E(z,t) = E_0 e^{-i(\omega t - n'kz)} e^{-n''kz} \equiv t E_0 e^{-i(\omega t - \phi)} \tag{6.29}$$

and thus suffers from a phase shift $\phi = n'kz$ and from absorption. In a medium comprising two-level systems the absorption characterised by the (amplitude) transmission coefficient $t = \exp(-n''kz)$ is given as

$$t \equiv e^{-\widetilde{D}/2} = \exp\left(-\frac{\widetilde{\rho}\sigma_0}{2}\frac{1}{1+\frac{(\omega-\omega_0)^2}{(\Gamma/2)^2}}\right) \tag{6.30}$$

and the phase shift as

$$\phi = \exp\left(-\frac{\widetilde{\rho}\sigma_0}{2}\frac{\frac{(\omega-\omega_0)^2}{(\Gamma/2)^2}}{1+\frac{(\omega-\omega_0)^2}{(\Gamma/2)^2}}\right) \tag{6.31}$$

as can be derived, e.g., from (5.131). $\widetilde{\rho} = \int \rho dz$ is the density of absorbers integrated along a column through the sample and

$$\sigma_0 = \frac{3\lambda^2}{2\pi} \tag{6.32}$$

is the (on-resonance) cross-section for scattering [146, 204, 205]. Consider a set-up where a probe beam with diameter larger than the atomic cloud is directed onto a charge coupled device (CCD) camera where it produces a shadow image of the cloud. The density of the atomic cloud can be determined as follows

$$D(x,y) = -\ln\left(\frac{I_{\text{with cloud}}(x,y) - I_0(x,y)}{I_{\text{w/o cloud}}(x,y) - I_0(x,y)}\right) \tag{6.33}$$

where $I_{\text{with cloud}}(x,y), I_{\text{w/o cloud}}(x,y), \text{and} I_0(x,y)$ are the images taken with the probe beam illuminating the cloud, the probe beam alone, and the dark image without cloud and without probe beam, respectively. Often the two-dimensional density distribution can be described by a Gaussian

$$D(x,y) = D_{\max} e^{-\frac{x^2+y^2}{2r_0^2}}. \tag{6.34}$$

To determine $D_{\max}$ one has to know the atomic distribution along the z axis. For a Gaussian with the same radius r_0 in the z direction it follows that

$$D_{\max} = \int \sigma_0 \rho_{\max} \exp\left(-\frac{z^2}{2r_0^2}\right) dz \tag{6.35}$$

which after evaluation of the integral yields

$$\rho_{\max} = \frac{D_{\max}}{\sqrt{2\pi}\sigma_0 r_0}. \tag{6.36}$$

Dark-ground imaging, phase-contrast imaging, or polarisation-contrast imaging are other methods that probe the atomic dispersion for non-destructive imaging of atomic clouds [204].

6.4.3.2 Temperature

Several methods can be employed to measure the temperature of cold atomic clouds either in the trap or after the atoms have been released from the trap.

Time-of-flight Measurements These techniques measure in one way or another [206] the ballistic expansion of the atomic cloud and relate the derived distribution in space and time to the kinetic energy of the particles. Given that the atoms are in thermal equilibrium, which is not always the case in a MOT [207], this distribution is related to the temperature of the ensemble. One of the simplest and earliest methods now referred to as the "release and recapture" technique was employed to determine the temperature of atoms in an optical molasses [161]. In this method the number of atoms (or a quantity being proportional to this number) is measured before and after the confining fields, e.g., magnetic field or laser fields are turned off abruptly for a variable time, τ_{off}. During this time the fastest atoms in the ballistically expanding cloud leave the capture range of the trap which contains less atoms after being switched on again. From the fraction of atoms remaining as a function of τ_{off} the temperature can be determined to the extent that the shape and size of the capture region can be determined. As the trapping fields have to be switched fast enough, this method is particularly suited for optical traps where the fluorescence of the atoms in the trapping field can be measured directly without extra equipment.

Methods that allow higher accuracy employ an extra laser beam and image the cloud sequentially by one of the methods described in Section 6.4.3.1. From a sequence of absorption or phase contrast images taken after various times t, the evolution of the size of the atomic cloud can be determined. As an example, consider the case of a spherical cloud of Gaussian distribution with radius r_0 immediately after switching off the trap. Since the position of each atom results from the initial position and the path travelled after shutting off the trap, both quantities are uncorrelated and hence can be added quadratically as follows

$$r(t) = \sqrt{r_0^2 + \langle v^2 \rangle t^2}. \tag{6.37}$$

As a third method, consider a sheet of light with low vertical extension and high horizontal extension placed a few centimetres below the trapped atoms. To determine the vertical

velocity distribution of the atoms released from the trap, the time-dependent fluorescence of the atoms falling through the light sheet is measured. Imaging the fluorescence radiation from the horizontal plane gives the velocities in the two horizontal directions [208].

Another method determines the height of the atoms released from the trap in the gravitational potential [162, 206].

Trap Centre Oscillations The temperature of atoms confined in a trap with a harmonic potential, e.g., a MOT, dipole trap, or a magnetic trap, can be measured using their response to an external force. In thermal equilibrium the thermal energy equals the average potential energy or the average kinetic energy as follows

$$k_B T = D\langle x^2 \rangle = m\langle v^2 \rangle. \tag{6.38}$$

Determination of the spring constant D (see (2.6)) together with a measurement of the extension of the trapped cloud, e.g., by a CCD camera, allows one to determine the temperature using (6.38) [209]. The spring constant can be determined by measuring the response of the cloud as a function, e.g., of the light pressure of a pushing laser beam or of the movement of the centre of the magnetic trapping fields. If the atomic cloud is driven by an external harmonic force to perform forced oscillations, the averaged damping constant and the spring constant can be determined by measuring the frequency-dependent amplitude response and phase using (2.59) and (2.60) [210].

Measurement of the Doppler Broadening Quantum systems used as frequency references in frequency standards often posess a narrow transition. Given that its homogeneous linewidth is small enough to allow the inhomogeneous Doppler broadening to be measured, the broadening can be utilised to derive the velocity distribution in the sample. In optical frequency standards the broadening of the clock transition is often used directly (Fig. 6.1). The method requires a laser of high short-term and medium-term stability which, in any case, is required as a local oscillator for the frequency standard. The velocity distribution in microwave standards, e.g., the Cs atomic fountain (Section 7.3), can be probed using Raman transitions (Fig. 5.12 a) with two phase-coherent laser fields [211]. For Raman transitions the relative phase of the two laser fields has to be stable rather than the phases of both laser fields themselves. Hence, the two necessary phase-coherent fields can be obtained either from a single laser of moderate frequency stability by the use of acousto-optic (Section 11.2.1) or electro-optic frequency shifters (Section 11.2.2) or for larger frequency separations by two phase-locked lasers.

6.5 Doppler-free Non-linear Spectroscopy

6.5.1 Saturation Spectroscopy

In optical frequency standards where the Doppler broadening becomes of utmost importance an easy-to-use method of obtaining narrow lines is based on non-linear spectroscopy or saturation spectroscopy [212]. Consider the case of a strong laser beam of frequency ν slightly blue detuned ($\nu > \nu_0$) with respect to the transition frequency ν_0 interacting with an ensemble of two-level systems with velocities obeying a Maxwellian distribution. Atoms whose velocities

$\vec{v}'$ satisfy the Doppler condition $\nu - \nu_0 = \vec{k} \cdot \vec{v}'$ are transferred from the lower energy level to the higher one and a hole, sometimes called a Bennett hole [213], is "burned" into the ground-state velocity distribution (Fig. 6.14 a). In the case of large saturation $S_0 \gg 1$ approximately

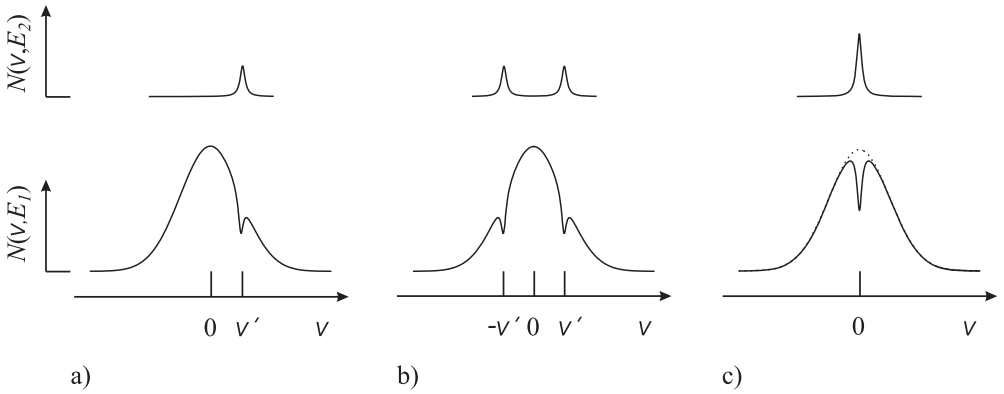


Figure 6.14: Population in the ground state E_1 and in the excited state E_2. a) A travelling wave with a positive detuning with respect to the transition centre transfers population from the state with lower energy E_1 to the state with higher energy E_2 thereby burning a hole in the lower state velocity distribution. b) Two counter-propagating waves interact with different velocity groups in the Doppler broadened profile if slightly detuned from resonance. c) Two counter-propagating waves tuned to resonance interact with the same velocity group.

one half of the atoms is transferred to the excited state. By tuning the frequency of the laser across the line and measuring the absorbed power or the fluorescence light from the decay of the excited atoms, e.g., by using a set-up like the one shown in Fig. 6.15 a), the Doppler broadened absorption line is recorded. Consider now the case where a second laser beam with the same frequency ν crosses the absorbing ensemble counter-propagating to the first laser beam, e.g., by re-directing the laser beam by means of a retro-reflector Fig. 6.15 a). For an "optically thin" atomic beam, each one of the two laser beams, if tuned across the absorption line independently, would show the absorption line and the total absorption suffered by both beams would be twice that of a single beam. For a fixed frequency the second beam interacting with the same ensemble burns a second hole into the ground-state velocity distribution at $v = -v'$ (see Fig. 6.14 b). For $\nu \neq \nu_0$ the two waves interact with different velocity groups, however, if the frequency ν of the laser is tuned to the transition frequency ν_0 (see Fig. 6.14 c) both waves interact with the same velocity class. The velocity of these atoms is characterised by a zero Doppler shift along the axis of the laser beam (say the z axis), i.e., by the velocity ($v_z = 0$). If the first laser beam has already saturated the transition ($S_0 \gg 1$), the second laser beam virtually does not experinece absorption, and hence, at resonance the total absorption is reduced by almost a factor of two as compared to the off-resonant absorption of the two single beams (see Fig. 6.16, open circles).

Such non-linear absorption occurs also for weaker saturation since atoms of this velocity group are saturated by twice the saturation parameter S_0 compared to atoms with a velocity far from resonance ($\nu \neq \nu_0$). The dip due to the saturated absorption at the centre of the absorption line is often called "Lamb dip" [212]. The width of the Lamb dip can be as narrow

as the homogeneous power broadened (Fig. 5.20) linewidth. The Lamb dip thus in general provides a more narrow spectroscopic feature than the absorption profile itself and can be utilised to stabilise the frequency of a laser with higher stability, due to the higher slope of the corresponding discriminator curve.

The intuitive picture given above for the saturated absorption is valid only in the weak saturation approximation [212]. Strong-field effects in coherent saturation spectroscopy can modify the absorption feature dramatically. To correctly describe the absorption spectrum in the strong-field case, the influence of the photon recoil, the actual laser beam profile and multiple momentum exchanges between atoms and light, have to be taken into account [214–216]. The latter becomes important if a standing wave is used for excitation. As an example, compare the absorption lines observed by exciting the transition $^1S_0 - ^3P_1$ ($\lambda = 657\,\mathrm{nm}$) in a Ca atomic beam with two spatially separated travelling waves and a standing wave (Fig. 6.16) [216].

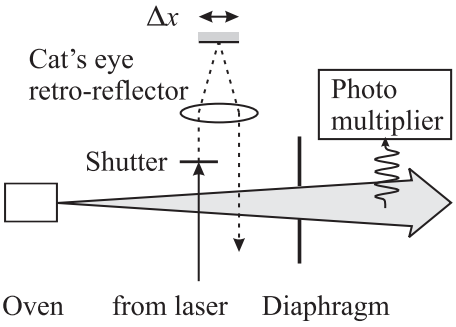


Figure 6.15: Saturated absorption in a beam. By shifting the cat's eye retroflector by $\Delta x/2$ the incoming beam and the reflected beam overlap leading to a standing wave.

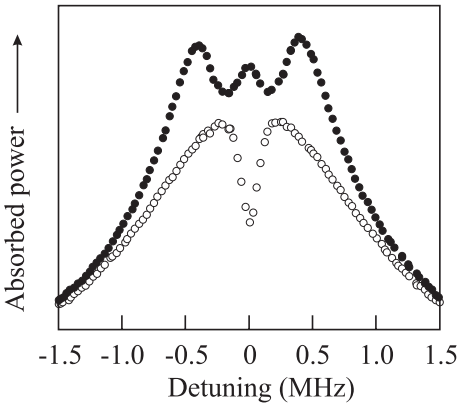


Figure 6.16: Power absorbed in a Ca atomic beam measured by the fluorescence of the excited atoms [216] in two counter-propagating travelling waves (open circles) and in a standing wave (dots) using the set-up of Fig. 6.15.

Due to the recoil effect the saturated absorption line (Fig. 6.16) is splitted into two components [217]. The origin of such a doublet can be inferred from Fig. 6.17 where the energy momentum parabola is shown for atoms in the ground state E_1 and the excited state E_2. Consider an atom at rest in the ground state (see Fig. 6.17 a). After the absorption of a photon the momentum of the atom is $p^2/(2m) = (\hbar k)^2/(2m)$. Hence the two counter-propagating laser waves interact with the same (zero-) velocity group in the ground state and $\hbar\omega = E_2 - E_1 + (\hbar k)^2/(2m)$, i.e. the saturated absorption is blue shifted in angular frequency $\Delta\omega = \hbar k^2/(2m)$. A saturated absorption feature appears also if the two laser beams interact with the same (zero -) velocity group of atoms in the excited state. In this case conservation of energy and momentum requires that the line is red shifted by the same amount. Hence the

frequency separation (recoil splitting) between both absorption lines is (see (5.103))

$$\Delta\nu = \frac{h}{m\lambda^2} \tag{6.39}$$

which in the case of the Ca intercombination line ($\lambda = 657.46\,\text{nm}$) is 23.1 kHz (Fig. 6.18).

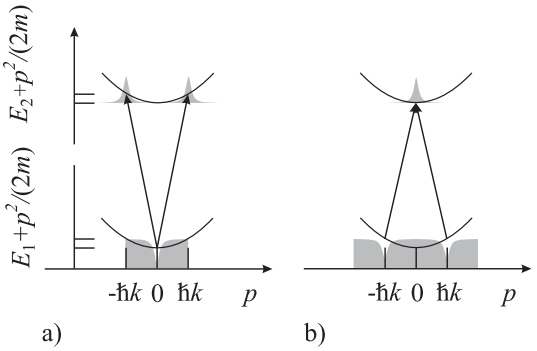


Figure 6.17: Energy–momentum diagrams for saturated absorption. a) The two counter-propagating laser waves interact with the same (zero-) velocity group in the ground state or b) in the excited state.

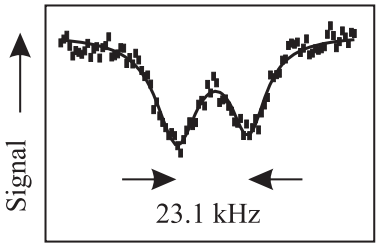


Figure 6.18: Recoil splitted saturated absorption line in a laser cooled cloud of Ca atoms.

Since the high-frequency recoil component results when both laser beams interact with the same velocity group in the excited state this component can be affected by spontaneous decay of the excited atoms during the desired and required long interaction time. Hence the high-frequency component in general shows a smaller saturated absorption feature than the low-frequency component. Thus if both recoil components are not resolved the centre of gravity of both lines is not necessarily at the frequency of the unperturbed line. For an optical frequency standard it might therefore be necessary to suppress one the components. Methods have been developed to suppress either the low-frequency component [218] or the high-frequency one [219, 220].

In saturated absorption spectroscopy only the absorbers from the zero velocity group contribute to the signal and hence give a first-order Doppler free signal. Nevertheless the second-order Doppler effect might still be large, e.g., in an atomic beam.

6.5.2 Power-dependent Selection of Low-velocity Absorbers

The first-order and second-order Doppler effect can be reduced significantly if only the absorption signal from the slowest particles, e.g., in a gas are selected. Such an effect occurs if the irradiance of the laser beam is so low that it is too small to markedly excite the faster particles since for optimum excitation of a two-level system a π pulse is required. Consider a particle with velocity v that crosses a coherent interacting radiation field of diameter $2w_0$ perpendicularly. The Rabi angle (5.52) during transit is given as $\theta_R = \Omega_R 2w_0/v$ (see (5.52)). If the Rabi frequency Ω_R (5.37) is kept deliberately low by choosing a small enough field amplitude, only those particles are excited whose velocities are small enough to lead to a significant Rabi angle.

The selection of slow molecules or atoms requires low pressure and low saturation intensity. Consequently, the saturation signal becomes very low. Despite of these difficulties the method has been used in frequency standards. Bagayev *et al.* [221] in their methane stabilised He-Ne laser, placed the methane absorption cell inside the laser cavity to benefit from the amplification of the signal with the laser operating just above threshold. The laser comprised an 8 m long internal absorption cell which was cooled to a temperature of 77 K. The laser beam had a waist of $2w_0 = 15\,\text{cm}$ to allow for sufficient interaction time. The group at the Université Paris-Nord [222] applied the same method by using a multiple-pass CO_2 laser beam ($\lambda \approx 10\,\mu\text{m}$ of a radius $w_0 = 3.5\,\text{cm}$ in an 18 m long OsO_4 absorption cell with an effective path length of 108 m. The cell was kept at room temperature and the pressure was below 3×10^{-4} Pa. With a power of 30 nW, molecules with an effective temperature of $T_{\text{eff}} = 0.6\,\text{K}$ could be selected. To utilise the signal contrast of only 10^{-6}, heterodyne detection and double modulation techniques had to be employed.

6.5.3 Two-photon Spectroscopy

First-order Doppler free absorption lines can be recorded if the energy energy required for a transition in a quantum absorber is provided by two photons (Fig. 6.19) from two counter-propagating beams of the same frequency [223, 224]. If an absorber moving with a velocity

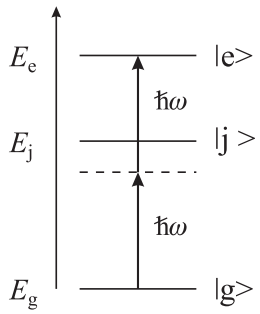


Figure 6.19: Two-photon transition. The dashed line indicates a virtual level.

$\vec{v}$ is to be excited at the same time by a photon from each beam according to the first-order Doppler effect the energy of one photon ($\hbar\omega_1$) is red-shifted and that of the other one ($\hbar\omega_2$) is blue shifted in the rest frame of the absorber. The combined energy of both photons

$$\hbar\omega_1 + \hbar\omega_2 = \hbar\omega_0 \left(1 - \frac{\vec{v}\cdot\vec{k}}{\omega_0}\right) + \hbar\omega_0 \left(1 + \frac{\vec{v}\cdot\vec{k}}{\omega_0}\right) = 2\hbar\omega_0 \tag{6.40}$$

is not affected by the first-order Doppler effect. Two-photon transitions can induce transitions between states that are otherwise not coupled by single-photon dipole radiation, e.g., if they have the same parity.

The transition amplitude for a two-photon transition has to be calculated in second-order time-dependent perturbation theory since the first-order amplitude for a transition such as 1S – 2S in hydrogen would be zero. Consider an atom moving in the field of two counter-propagating waves of the same angular frequency ω, amplitudes E_1 and E_2 and polarisations

described by the unit vectors $\vec{\epsilon}_1$ and $\vec{\epsilon}_2$. It has been shown [223–225] that the probability for two-photon absorption is

$$
\begin{aligned}
c^{(2)}(\omega) &= \sum_j \frac{e^2E_1^2}{4\hbar^2}\frac{\langle g|e\vec{r}\cdot\vec{\epsilon}_1|j\rangle\langle j|e\vec{r}\cdot\vec{\epsilon}_1|e\rangle}{\omega_{j\,g}-\omega}\times\frac{\exp\left[i(\omega_{e\,g}-2\omega+2\vec{k}\cdot\vec{v})t\right]}{(\omega_{g\,e}-2\omega+2\vec{k}\cdot\vec{v})-i\Gamma_e/2} \\
&+\sum_j \frac{e^2E_2^2}{4\hbar^2}\frac{\langle g|e\vec{r}\cdot\vec{\epsilon}_2|j\rangle\langle j|e\vec{r}\cdot\vec{\epsilon}_2|e\rangle}{\omega_{j\,g}-\omega}\times\frac{\exp\left[i(\omega_{e\,g}-2\omega-2\vec{k}\cdot\vec{v})t\right]}{(\omega_{g\,e}-2\omega-2\vec{k}\cdot\vec{v})-i\Gamma_e/2} \\
&+\sum_j \frac{e^2E_1E_2}{4\hbar^2}\left[\frac{\langle g|e\vec{r}\cdot\vec{\epsilon}_1|j\rangle\langle j|e\vec{r}\cdot\vec{\epsilon}_2|e\rangle}{\omega_{j\,g}-\omega}\right. \\
&\qquad\left.+\frac{\langle g|e\vec{r}\cdot\vec{\epsilon}_2|j\rangle\langle j|e\vec{r}\cdot\vec{\epsilon}_1|e\rangle}{\omega_{j\,g}-\omega}\right]\times\frac{\exp\left[i(\omega_{e\,g}-2\omega)t\right]}{(\omega_{g\,e}-2\omega)-i\Gamma_e/2}
\end{aligned}
\tag{6.41}
$$

where all atomic eigenstates have to be included in the sum and $\Gamma_e = 2\pi\gamma_e$ is the inverse lifetime of the state $|e\rangle$, provided that the ground state has an infinite lifetime. As a result of the resonance denominator $\omega_{jg}-\omega$ the largest contributions come from any intermediate level j close to the virtual level (Fig. 6.19). The first two terms in (6.41) describe the absorption of two photons from any one of the running waves. Consequently, both terms include the first-order Doppler shift. The third term describes the absorption of a photon from each of the two counter-propagating waves where the first-order Doppler shift cancels according to (6.40). Squaring (6.41) and averaging over all velocities of the atoms, the first two terms lead to a Gaussian Doppler background [223]. In general, for frequency standards, the width of the Doppler background is large compared to the linewidth of the two-photon transition and the Doppler pedestal contributes a very weak, almost constant background. For a linearly polarised standing wave ($E \equiv E_1 = E_2$ and $\vec{\epsilon} \equiv \vec{\epsilon}_1 = \vec{\epsilon}_2$) the third term of (6.41) leads to a Doppler-free Lorentzian profile

$$
P^{(2)}(\omega) = \frac{e^4E^4}{4\hbar^4}\left|\sum_j \frac{\langle g|e\vec{r}\cdot\vec{\epsilon}|j\rangle\langle j|e\vec{r}\cdot\vec{\epsilon}|e\rangle}{\omega_{j\,g}-\omega}\right|^2 \frac{\Gamma_e}{(\omega_{g\,e}-2\omega)^2+\Gamma_e^2/4} \tag{6.42}
$$

whose linewidth is the width of the final state. The transition probability depends on the fourth power of the field amplitude, i.e., on the square of the laser irradiance. Hence, the laser beam often has to be focussed to a small diameter to drive the weak second-order transitions. The high irradiance often leads to a considerable ac Stark shift which has to be measured and corrected for in frequency standards. For fast atoms the small diameter of the laser beam may result in a transit time broadening that changes the lineshape of (6.42) [226]. In contrast to saturation spectroscopy, two-photon spectroscopy has the big advantage that all atoms contribute to the signal despite their velocity.

Two-photon transitions are used in frequency standards based on the 1S – 2S transition in atomic hydrogen (Section 9.4.5), the $5S_{1/2}$ – $5D_{5/2}$ transition in Rb (Section 9.4.3), or the $4d^{10}\,5s\,{}^2S_{1/2}$ – $4d^9\,5s^2\,{}^2D_{5/2}$ transition in Ag (Section 9.4.6).

6.6 Interrogation by Multiple Coherent Interactions

In order to make the best use of the narrow linewidth achievable with quantum absorbers, also the interaction time broadening has to be reduced by allowing a coherent interaction with the applied electromagnetic field over the necessary long time. An extremely fruitful method has been developed for microwave standards by Norman Ramsey [155, 227, 228] where, rather than applying a coherent field over the total time T, the quantum absorbers are exposed to the field for short durations τ that are separated by a time T where no field is present. This "Ramsey excitation" is equally well applied to atomic or molecular beams interacting with spatially separated fields or to devices where the absorbers interact at the same location with a field that is switched on and off in a time sequence. Ramsey excitation using two separated interactions is applied likewise in caesium atomic clocks and other microwave frequency standards. In the optical regime the method has to be modified where in general more than two coherent interactions are used.

As will be shown, a main advantage of the Ramsey excitation with multiple coherent interactions in contrast to, e.g., saturation spectroscopic methods, results from the fact that the interaction time broadening and the resolution can be adjusted independently. The former can be increased by choosing short interaction times τ thereby also allowing absorbers with Doppler shifts $v\nu/c \lesssim 1/(2\pi\tau)$ to contribute to the signal. The resolution, however, is determined mainly by the much longer time T between the interactions.

6.6.1 Ramsey Excitation in Microwave Frequency Standards

To describe the interaction of two short pulses of an electromagnetic field with a suitable quantum absorber, e.g., the 9.2 GHz radiation with the hyperfine split ground states of Cs atoms, we first outline the calculations of the resonance features using the methods developed in Section 5.3.1 for two-level systems. To complement the description we also visualise the relevant processes using the Bloch vector picture developed in Section 5.3.2 (Fig. 6.21).

Consider two subsequent interactions of duration τ of a two-level atom with an interrogating field where these two interaction times are separated by a time T where the field is switched off. Such a situation is realised, e.g., in a caesium beam machine (Section 7.1) where the caesium atoms on their way pass two spatially separated interaction zones or in the caesium fountain clock (Section 7.3) where the caesium atoms during their vertical ballistic flight cross the same microwave field on their way up and down.

In the calculation of the probability amplitudes $c_1(t)$ and $c_2(t)$ of finding the two-level system in the ground state and in the excited state given by (5.50) and (5.51), respectively, (Section 5.3.1) we have assumed that the atom immediately before the interaction was in the ground state ($c_1(t = 0) = 1, c_2(t = 0) = 0$). In order to investigate the combined influence of the field during the first and second interaction and the evolution of the atomic states between both, Ramsey [155, 227] considered a more general expression. He calculated the evolution of the amplitudes $c_1(t_1 + t)$ and $c_2(t_1 + t)$ for an interaction applied between the time t_1 and $t_1 + t$, where at the beginning of the interaction, the two-level atom is described by given probability amplitudes $c_1(t_1)$ and $c_2(t_1)$ as a result of any previous interaction. The probability amplitude to find the two-level atom in the excited state after the second interaction is the sum of two contributions. One term describes the probability amplitude that the atom

has been excited during the first interaction and enters the second interaction in this state. The other term describes the probability amplitude for the atom to leave the first interaction in the ground state and being excited during the second interaction. Hence, by calculating the total probability to find the atom in the excited state after the second interaction one expects an interference between both amplitudes that depends on the phase difference between the two amplitudes.

Ramsey [155, 227] calculated the probability to find the two-level system in the excited state after the second interaction as

$$\begin{aligned} p(\tau+T+\tau) &\equiv |c_2(\tau+T+\tau)|^2 \qquad (6.43)\\ &= 4\frac{\Omega_R^2}{{\Omega'_R}^2}\sin^2\frac{\Omega'_R\tau}{2}\left(\cos\frac{\Omega'_R\tau}{2}\cos\frac{\Delta\omega T}{2}-\frac{\Delta\omega}{\Omega'_R}\sin\frac{\Omega'_R\tau}{2}\sin\frac{\Delta\omega T}{2}\right)^2, \end{aligned}$$

where the Rabi frequency Ω_R (for electric dipole interaction (5.37) or equivalently for magnetic dipole interaction), Ω'_R (see (5.45)), and the detuning $\Delta\omega$ (see (5.39)) are defined as before.

In the immediate vicinity of the resonance, i.e., for $\Delta\omega \ll \Omega_R$, $\Omega_R \approx \Omega'_R$ holds and (6.43) leads to

$$p(\tau+T+\tau) \approx \tfrac{1}{2}\sin^2\Omega_R\tau\,[1+\cos 2\pi(\nu-\nu_0)T]\,. \qquad (6.44)$$

From (6.44) one finds that optimal excitation of the atom is achieved for two interactions with $\Omega_R\tau = \pi/2$, i.e. $\pi/2$ pulses. The full width at half maximum of the resonance curve is given by

$$\Delta\nu = \frac{1}{2T}. \qquad (6.45)$$

Hence, the achievable resolution with Ramsey excitation is about twice that of Rabi excitation for the same interrogation time (Fig. 6.20).

In practical frequency standards, that make use of the Ramsey scheme, there are important effects that require a modification of (6.44). As an example we consider the Cs atomic clock described in more detail in Section 7. There, the Ramsey scheme is applied to a beam of caesium atoms that successively interact with two phase coherent electromagnetic fields derived from the same microwave source. If there is a phase difference $\Delta\Phi = \Phi_2 - \Phi_1$ between the phases Φ_1 and Φ_2 in the first and second zone, respectively, then (6.44) has to be modified to

$$p(\tau+T+\tau) \approx \tfrac{1}{2}\sin^2\Omega_R\tau\,[1+\cos 2\pi(\nu-\nu_0)T+\Delta\Phi]\,. \qquad (6.46)$$

Phase differences $\Delta\Phi$ in general shift the centre of the Ramsey structure to $\nu \neq \nu_0$ by

$$\frac{\Delta\nu_\Phi}{\nu_0} = -\frac{\Phi}{2\pi\nu_0 T}. \qquad (6.47)$$

Consequently, for the operation of precise frequency standards methods have been devised for keeping such phase shifts as low and as constant as possible as will be discussed in more detail later. Compared to a continuous interaction the application of Ramsey's technique in atomic clocks has distinct advantages if frequency shifts are considered that result from a

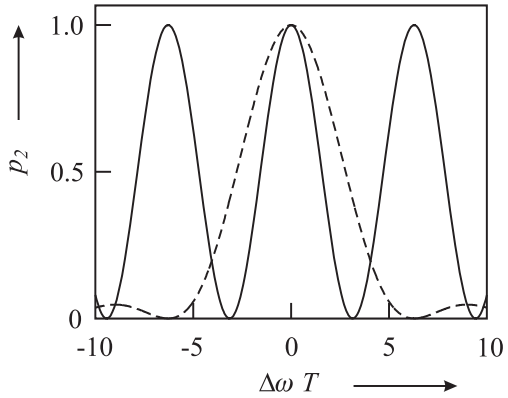


Figure 6.20: Comparison of the probability to find an atom in the excited state after excitation by two short Ramsey pulses of duration $\tau \ll T$ separated by a field-free time T (full curve; (6.44)) and a constant interaction (dashed curve; (5.51), so-called Rabi pulse) with the same interaction time.

varying phase during the coherent interaction of the atomic systems with the electromagnetic field. This is due to the fact that, in general, the phase in the limited interaction regions can be controlled more accurately as in the extended region.

In a set-up as the Cs atomic clock (Section 7) furthermore the influence of the velocity distribution of the atoms has to be taken into account. The atoms in the beam having different velocities v enter the second interaction zone after a time $T = L/v$ and hence (6.44) leads to different resolutions for the different atomic velocity classes. Consequently, the Ramsey fringes for a thermal ensemble are washed out for larger detunings and, mainly, the central fringes persist (Fig. 7.4). The velocity furthermore affects the time τ that the atom spends in the interaction zones and hence affects the Rabi angle that the atoms acquire. These effects will be discussed in more detail in connection with the particular frequency standards.

We now use the pseudo-spin picture to visualise the Ramsey method with two phase-coherent pulses of very short duration τ and the interaction, whose frequency is in resonance with the transition, adjusted such that each interaction corresponds to a $\pi/2$ pulse. As before, the time between the pulses is T. A two-level system in the ground state before the first interaction is represented by a pseudo-spin vector pointing towards the south pole of the Bloch sphere (Fig. 6.21 a). As described in Fig. 5.11 the first $\pi/2$ pulse rotates the pseudo spin by the angle $\theta = \pi/2$ around the $-u$ axis (Fig. 6.21 b). Afterwards, during the time T the pseudo spin rotates in the $u - v$ plane by the angle $2\pi(\nu - \nu_0)T$ (Fig. 6.21 c_1). If the time is such that this angle is $2n\pi$, a second $\pi/2$ pulse rotates the pseudo spin by $\pi/2$ around the $-u$ axis leaving the atom with unity probability in the excited state (Fig. 6.21 d_1). If this time is chosen such that, during this time, the pseudo spin rotates around the w axis by $3/4\pi$ as depicted in (Fig. 6.21 c_2) the second pulse would not change the orientation of the pseudo spin as the rotation axis is now the v axis. A subsequent measurement projecting the wave function onto the eigenstate $|e\rangle$ would show only 50 % of the atoms in the excited state. If the time T, on the contrary, was chosen to give an angle $\pi(\nu - \nu_0)T$ in the equatorial plane (Fig. 6.21 c_3)

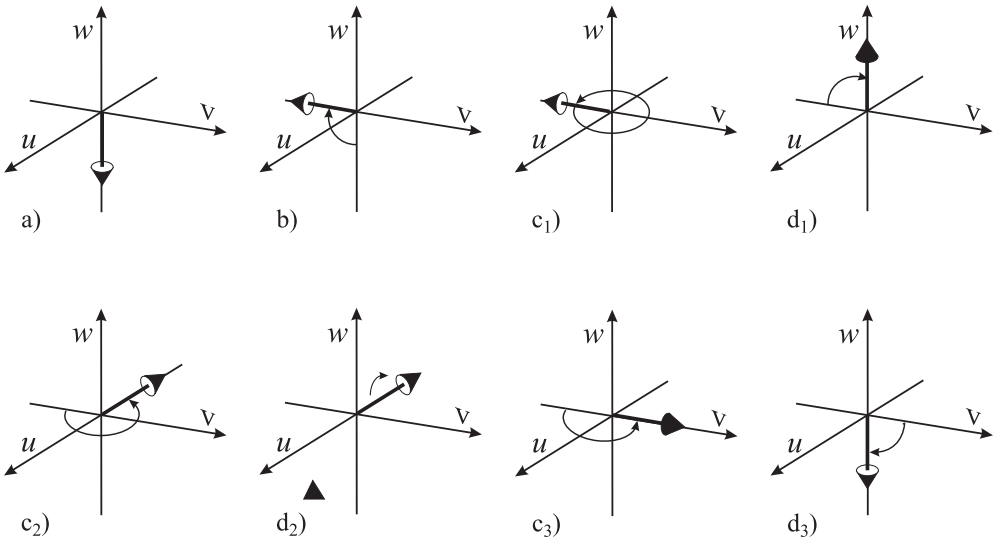


Figure 6.21: Evolution of the fictitious spin (Bloch vector) after excitation by two very short $\pi/2$ pulses $\Omega_R\tau \ll \Delta\omega T$ in a Ramsey excitation separated by different times T. a) – d_1) $\Delta\omega T = 2\pi$. a), b), c_2, d_2) $\Delta\omega T = 3/4\pi$. a), b), c_3, d_3) $\Delta\omega T = \pi$.

the precession of the pseudo spin under the "torque" of the second pulse (around the $-u$ axis) would flip the pseudo spin pointing to the south pole (Fig. 6.21 d_3) and the atom would be found in the ground state. It follows from such arguments that the probability of finding the atom in the excited state after the two Ramsey pulses shows a sinusoidal variation as a function of $2\pi(\nu - \nu_0)T$, i.e., either for fixed T as a function of the detuning or for fixed detuning as function of the time T. In other words, the Ramsey technique measures the difference in the phase $2\pi\nu T$ of the external oscillating field and the internal phase $2\pi\nu_0 T$ of the quantum system by, e.g., (6.44).

6.6.2 Multiple Coherent Interactions in Optical Frequency Standards

In the optical regime the wavelength of the radiation is so small that, in general, freely moving atoms are not confined to the Lamb-Dicke regime. The associated phase shifts do not allow one to observe a stationary Ramsey interference pattern. Thus the generation of optical Ramsey resonances requires either extra elements affecting the trajectories of the absorbers [229] or non-linear optical Doppler-free excitation schemes such as Doppler-free two-photon excitation or three [230] and more [215] suitable separated excitation zones.

6.6.2.1 Linear Optical Ramsey Resonances

As a first example consider Fig. 6.22 a) where the absorbers in a beam interact consecutively with two standing optical waves forming a Ramsey interrogation scheme. The electric field of the linearly polarised standing wave comprised of two counter-propagating linearly polarised

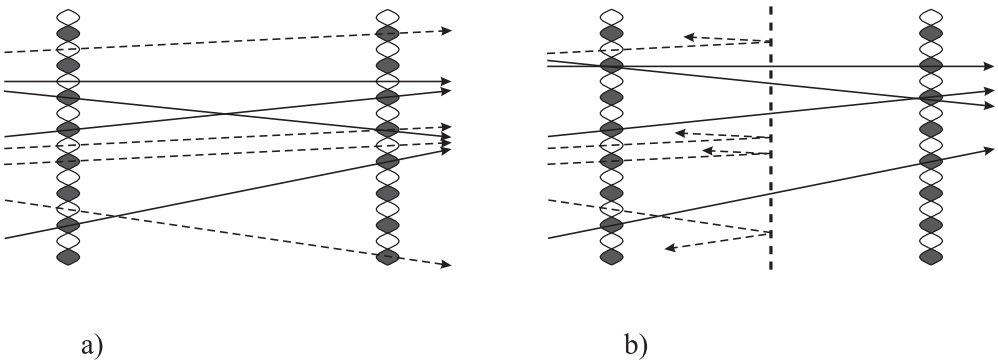


Figure 6.22: Linear Ramsey excitation of a molecular beam with two spatially separated standing waves of wavelength λ according to Kramer [229]. a) Trajectories of the absorbers passing anti-nodes in the two standing waves with an even phase difference of $2n\pi$ (solid lines) or with phase difference $(2n+1)\pi$ (dashed lines). (b) Insertion of a grating (grating constant $\lambda/2$) blocks the trajectories that lead to inverted Ramsey fringes.

travelling waves used to excite a transition in the atoms or molecules in the beam changes its phase by π every $\lambda/2$ (see (6.28)). Trajectories of the molecules passing anti-nodes in the two standing waves with an odd phase difference $(2n+1)\pi$ (dashed lines) give rise to Ramsey fringes that are phase shifted by π, i.e., inverted with respect to the Ramsey fringes generated by molecules that experience an even phase difference of $2n\pi$ (solid lines). Blocking the trajectories of either group by means of a transmission grating for the molecules with a period of $\lambda/2$ allowed Kramer [229,231] to use optical Ramsey fringes in optical CH_4 (Section 9.1.4) or OsO_4 frequency standards.

6.6.2.2 Optical Ramsey Resonances Using Non-linear Spectroscopy

Ramsey's method of two spatially separated fields can be transferred to the optical regime if first-order Doppler-free methods are employed.

Two-photon Optical Ramsey Resonances After Baklanov *et al.* [232] suggested two-photon excitation to apply Ramsey's method to the optical domain Salour and Cohen–Tannoudji [233] used two time-delayed short pulses to excite so-called "optical Ramsey resonances" in a sodium cell. Two spatially separated standing waves allowed Lee *et al.* [234] to observe two-photon Ramsey resonances in rubidium Rydberg atoms. The method has been employed recently to the hydrogen 1S – 2S transition [235] to obtain narrow resonances and is suitable for all two-photon clock transitions.

Saturated Absorption with Spatially Separated Laser Fields To overcome the Doppler shift in the optical domain where, in general, the atoms are no longer in the Lamb-Dicke regime, Ramsey excitation is often performed with three or more interaction zones. Shortly after Baklanov *et al.* [230] suggested the use of three equidistantly separated standing waves Bergquist *et al.* [236] used this technique to observe optical Ramsey resonances in

a metastable neon beam. Later Barger *et al.* [237] applied the method to the Ca intercombination transition ($\lambda = 657$ nm) used today as optical frequency standard (Section 9.4.4). With laser beam separations up to 21 cm Barger was able to resolve the Ca line in a thermal atomic beam with a width as narrow as 1 kHz [238] thereby resolving the recoil dublet and demonstrating the shifts and broadening due to the second-order Doppler effect. Kisters *et al.* [196] applied the method in the time domain to a laser-cooled sample of Ca atoms and used three pulses of a standing wave.

Baba and Shimoda [239] used three separated beams from a He-Ne laser to excite Ramsey resonances at 3.39 µm in a CH_4 cell. The two outermost beams were counterpropagating travelling waves with a central standing wave. Bordé *et al.* [240] and Helmcke *et al.* [241] showed that optical Ramsey resonances with much higher contrast could be obtained by using a scheme of two counter-propagating pairs of co-propagating laser beams which will be discussed in more detail below.

6.6.2.3 Optical Ramsey Resonances as Bordé Atom Interferences

We consider in the following the interaction of an atom in the ground state that subsequently interacts with two counter-propagating pairs of parallel laser beams. Rather than describing the interaction scheme in the pseudo-spin picture which has been done in [215] we use, in the following, a more illustrating and conceptionally simple picture that has been introduced by Bordé [242, 243] to explain the origin of the Ramsey interferences in terms of atom interferences.

Consider a two-level atom in the ground state $|g\rangle$ with momentum $\vec{p} = m\vec{v}$ interacting with a laser beam with wave vector $\vec{k}$ (Fig. 6.23 a). If a photon is absorbed, the atom is excited

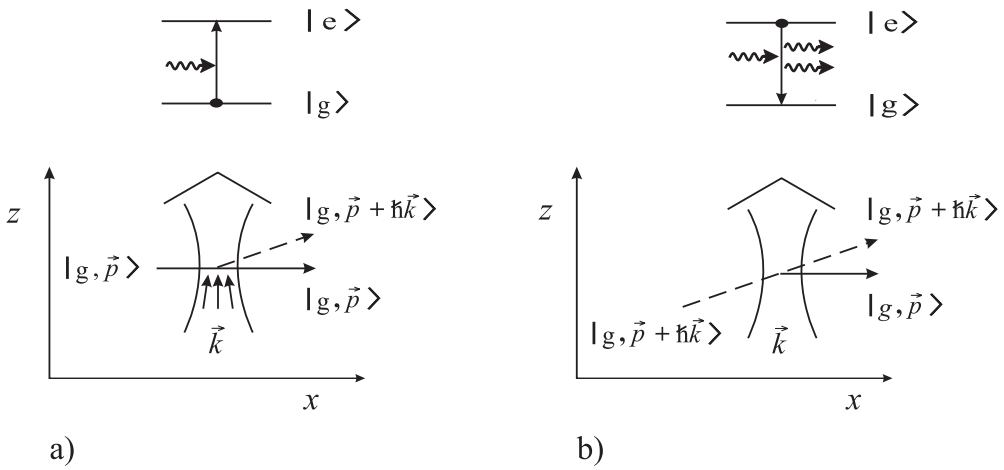


Figure 6.23: Near-resonant interaction of a photon with a two-level atom. a) Stimulated absorption. b) Stimulated emission.

to the state $|e\rangle$ and the momentum $\hbar\vec{k}$ with $\hbar k = h\nu/c$ of the photon is transferred to the atom. As a consequence of this momentum transfer, the momentum and the trajectory of the

atom are changed (dashed line in Fig. 6.23 a). If no photon is absorbed, however, the atom leaves the interaction zone without changing its direction. The probability of finding the atom in the excited state after the interaction with the field depends on the Rabi angle (Fig. 5.8) which can be adjusted by the amplitude and the interaction time depending on the particular matrix element of the interaction. If the Rabi angle is chosen to be $\theta = \pi/2$, the atom is put into a coherent superposition of both states $|e\rangle$ and $|g\rangle$ with equal amplitudes. Both sub-states have different trajectories which in the picture of an atomic "particle" would mean that the atom "breaks up" into two parts that separate. It is therefore more appropriate to describe the atomic particle as an atomic wave packet that is split into two partial wave packets travelling into two different directions. The wavelength associated with each of the wave packets is the well known de Broglie wavelength (5.127) $\lambda_{\mathrm{dB}} = h/mv$ depending on the momentum mv of the particle. Similar arguments hold for the case of the stimulated emission of a photon (Fig. 6.23 b).

In this picture the interaction of a photon field with a two-level atom can be thought of as an atomic beam splitter for the atomic wave packet. When two of those partial wave packets are recombined they interfere where the resulting amplitude is a function of the phase shift between the two partial waves. In the following, we use this picture to calculate the phase shift quantitatively, following Sterr *et al.* [244, 245].

The (non-relativistic) energy conservation

$$\frac{\vec{p}^{\,2}}{2m} + \hbar\omega = \frac{(\vec{p} + \hbar\vec{k})^2}{2m} + \hbar\omega_0 \tag{6.48}$$

leads to

$$\frac{\vec{k}\cdot\vec{p}}{m} = \omega - \omega_0 - \frac{\hbar k^2}{2m}. \tag{6.49}$$

From (6.49) it follows that, in general, for a detuning not equal to the recoil term ($\omega - \omega_0 \neq \hbar k^2/(2m)$) also a momentum component $\hbar k_x$ is transferred to the atom which is parallel to the momentum of the atom and perpendicular to the z direction. This somewhat surprising result can be explained by the fact that the interacting electromagnetic field is localised to a spatial area and consequently is composed of a bundle of wave vectors with different directions (Fig. 6.23 a) in contrast to the single wave vector of a plane wave of infinite transverse extension. As a result of the momentum transfer in the x and z direction, the two partial wave packets are displaced in space by

$$\Delta z = T\frac{\hbar k_z}{m} \tag{6.50}$$

and

$$\Delta x = T\hbar\frac{(\omega - \omega_0) - \hbar k^2/(2m) - k_z p_z/m}{p_x}. \tag{6.51}$$

The three terms in the displacement Δx (6.51) are due to the detuning, the recoil, and the Doppler shift. If another interaction zone is added (Fig. 6.24) to provide a second 50 % beam splitter, there are two partial wave packets in the ground state $|g\rangle$ and two wave packets in the excited state $|e\rangle$. The two partial wave packets in the state $|e\rangle$ (and likewise in the state $|g\rangle$) are identical with respect to their quantum numbers. If the spatial shifts (6.50) and (6.51)

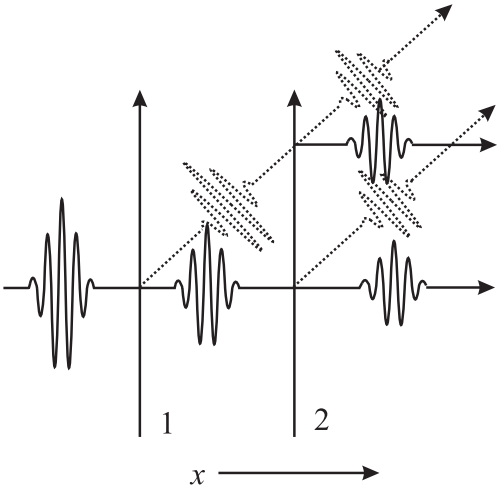


Figure 6.24: Atom interferometer comprising two laser beam splitters 1 and 2. Solid and dotted wave packets represent atoms in the ground state and excited state, respectively. Interferences between the respective wave packets leaving the interferometer after the second interaction zone occur if the two wave packets overlap in the x and z directions.

are smaller than the corresponding widths of the wave packets, the amplitudes of both wave particles have to be added, where the total amplitude depends on the phase shift between the two interfering partial wave packets. In the x direction, one partial wave is shifted with respect to the other by the phase shift $k\Delta x = (2\pi/\lambda_{\mathrm{dB}})\Delta x$ and additionally by the phase ϕ_i of the electromagnetic wave in the i^{th} interaction process.[4]

Consequently, the periodic variation of the number of atoms in the excited state as a function of the detuning (see (6.51)), previously referred to as Ramsey resonances, is now interpreted as an interference term between the amplitudes of both (atomic) partial waves in the excited state $|e\rangle$.

From (6.50) and (6.51) and Fig. 6.24, the difference between the Ramsey resonances in the microwave and in the optical domain becomes obvious. In the microwave domain the modulus of the wave vector $k = 2\pi/\lambda$ is about five orders of magnitude smaller than in the optical domain, and hence only the detuning term is relevant in the x direction and practically no transverse separation (Δz) between the wave packets occurs. Thus two interaction zones are sufficient for observing atom interferences. In the optical domain, however, the separation is large and the wave packets have to be re-directed, e.g., by additional beam splitters (Fig. 6.25, Fig. 6.26).

In the case of an atom interferometer with four travelling laser beams as beam splitters (Fig. 6.25) there are two different interferometers (indicated by the two grey trapezoids) with

[4] This latter phase shift is a result of the fact that the interaction operator, e.g., the magnetic or electric dipole operator that initiates the transition, explicitly depends on the phase of the electromagnetic wave. Solving the Schrödinger equation with this interaction operator leads to an atomic wave function with the same time dependence, and hence the same phase, as the electromagnetic field [246].

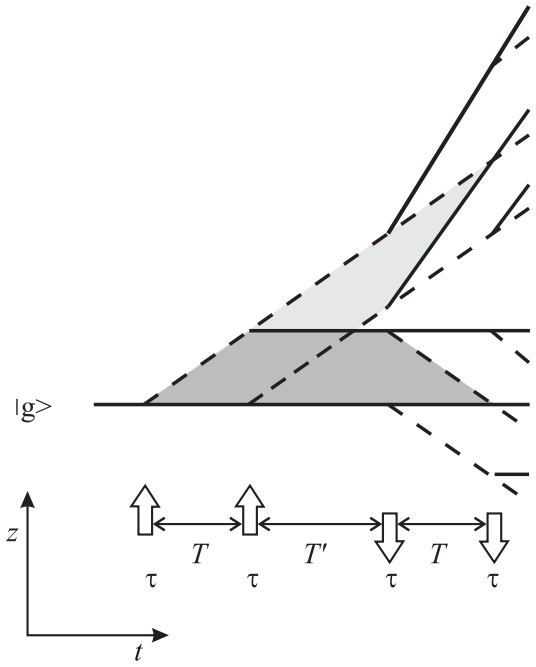


Figure 6.25: Time-domain atom interferometer comprising four coherent interactions with four travelling waves. Trajectories of wave packets in the ground state $|g\rangle$ are shown as full lines, trajectories depicting excited state wave packets $|e\rangle$ are dashed lines.

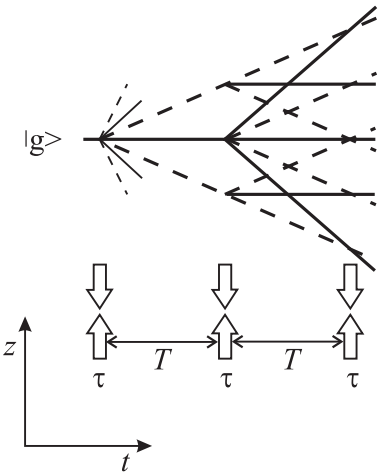


Figure 6.26: Atom interferometer with three standing waves. Trajectories of wave packets in the ground state $|g\rangle$ are shown as full lines, trajectories depicting excited state wave packets $|e\rangle$ are dashed lines.

different directions of the recoil shift. The changed direction of the second pair of laser beams with respect to the first one, leads to a final displacement at the exit ports of the interferometers of $\Delta z = 0$ and $\Delta x = 2T\hbar[(\omega - \omega_0) \pm \hbar k^2/(2m)]/p_x$ where the $+$ and $-$ signs hold for the red and blue recoil component, respectively.

The contrast of the atom interferences near zero detuning for four travelling waves can be calculated from Fig. 6.25 by counting the interfering and the non-interfering paths that lead to excited states. Consider an atomic wave packet entering the interferometer. Assuming an ideal beam splitter in each interaction zone, an incoming wave packet is split into two wave packets, one being in the ground state $|g\rangle$ and one in the excited state $|e\rangle$, with the amplitudes reduced by $1/\sqrt{2}$. Hence, from the sixteen partial waves with amplitudes $1/4$ leaving the fourth interaction zone, there are four wave packets in the excited state that lead to an incoherent background with probability $4 \times 1/16 = 1/4$. The same reasoning leads to a probability of $1/4$ for the atom leaving the interferometer in the ground state. At each exit port of each of the two interferometers, two partial waves with amplitude $1/4$ have to be added coherently. Hence the probability of finding the atom in the excited state behind the

fourth interaction is given by

$$p_{|e\rangle} = \frac{1}{4} + \frac{1}{8}\left(1+\cos\left[2T\left(\omega-\omega_0+\frac{\hbar k^2}{2m}\right)+\phi_2-\phi_1+\phi_4-\phi_3\right]\right) + \frac{1}{8}\left(1+\cos\left[2T\left(\omega-\omega_0-\frac{\hbar k^2}{2m}\right)+\phi_2-\phi_1+\phi_4-\phi_3\right]\right). \tag{6.52}$$

Up to now we have considered only the interference pattern generated by a single atom. Consider the interaction of the atoms in a beam with two field zones separated by the distance D. For an ensemble of atoms the sinusoidal oscillation (see (6.52)) can be observed only in cases where the time $T = d/v$ between the two interactions is the same for all atoms. This can be achieved if all atoms have the same velocity v crossing the spatially separated interaction zones or if all atoms interact with a sequence of phase coherent pulses. The former case represents an atom interferometer in the space domain, the latter an atom interferometer in the time domain. In a space-domain atom interferometer the beam of atoms in general has a velocity spread Δv corresponding to a coherence length $x_{\mathrm{coh}} = \hbar/(2m\Delta v)$ and the fringes can be observed only as far as the displacement is smaller than the coherence length $\Delta x < x_{\mathrm{coh}}$. Hence, for a thermal beam, only a few fringes can be observed (Fig. 7.4).

Separated field excitation with three standing waves (Fig. 6.26) was first proposed by Baklanov *et al.* [230]. A standing wave acting as a beam splitter, in general, produces more than one deflected partial wave. Furthermore, there are six interferometers, two of which are symmetric and hence are not sensitive to the detuning. Because of this and due to the additional diffraction orders in the beam splitters, the achievable contrast is lower than in the set-up with four travelling waves. Each interferometer (triangle) has a mirror image. As a result of the symmetry, this type of atom interferometer is less susceptible to phase errors introduced by misalignments of the laser beams which, to first order, result in reduced contrast but not in a phase shift.

The Ramsey resonances interpreted as atom interferences, where a photon splits and recombines an atomic wave packet, meets with difficulties if we apply the photon picture in a naive way. We have started our consideration with the interaction of a photon with the wave packet in each interaction zone. On the other hand, we know that an excited atom leaving the second interaction zone has absorbed in total only a single quantum $\hbar\omega_0$ from both interaction regions. We therefore have to regard both interaction zones as a single field and the absorbed photon as the quantum of the total field. As a consequence, we cannot distinguish in which interaction zone the photon has been absorbed and which path the "atom" has taken, a situation that is encountered in each interferometry experiment.

The principal limit of accuracy that can be achieved with such an interferometer and hence with an atomic clock is set by the fluctuations of the phase to be detected. Jacobson *et al.* [247] have shown that the smallest detectable phase is given by

$$\delta\phi_{\mathrm{min}} = \frac{1}{2}\sqrt{\frac{N_{\mathrm{at}} + 4N_{\mathrm{phot}}}{N_{\mathrm{at}}N_{\mathrm{phot}}}} \tag{6.53}$$

where N_{at} and N_{phot} are the number of atoms and photons, respectively. In a typical atom interferometer, the number of photons in the beam splitting fields is much larger than the

number of atoms ($N_{\rm phot} \gg N_{\rm at}$) and the minimum detectable phase of (6.53) reduces to $\delta\phi_{\rm min} \approx 1/\sqrt{N_{\rm at}}$, i.e., the shot noise of the atoms.

These excitation schemes can also be applied to an isotropically expanding sample of laser cooled atoms by using three pulses from a standing wave or two pulses separated by a time T from each one of two counter-propagating running waves. These time-domain atom interferometers differ from the separated field set-ups in such a way that energy conservation according to (6.48) has not to be fulfilled. In contrast, the energy spread associated with a short pulse of duration τ, in general, can provide the energy difference $\hbar(\omega - \omega_0)$. Since in the pulsed excitation all atoms are subject to the same duration T between the pulses, in general many more fringes can be observed despite the velocity spread of the atomic ensemble.

Sequences of more than four pulses have also been used to obtain narrow fringes [248]. For a given total interrogation time, however, the resolution is the same if a sequence of $2n$ pulses with a dark time T in between ($\Delta\nu = 1/(2nT)$) or if a pulse sequence with two pulses of a dark time nT is used.

7 Caesium Atomic Clocks

Amongst all the atomic clocks, the Cs clock has a special place, since the unit of time is at present based on a microwave transition in Cs. The 13$^{\text{th}}$ General Conference for Weights and Measures (CGPM) has defined in 1967 [92] that

> the second is the duration of 9 192 631 770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the caesium 133 atom.

The only stable isotope ^{133}Cs has a nuclear spin quantum number of $I = 7/2$ which together with the total spin $J = 1/2$ of the electron shell leads to the two hyperfine states $F = I + J = 4$ and $F = I - J = 3$ which split, in the magnetic field, into 16 components (Fig. 7.1). The Cs clock uses the transition with the smallest sensitivity to magnetic fields, i.e., between the $|F = 4, m_F = 0\rangle \rightarrow |F = 3, m_F = 0\rangle$ transition (Fig. 7.2).

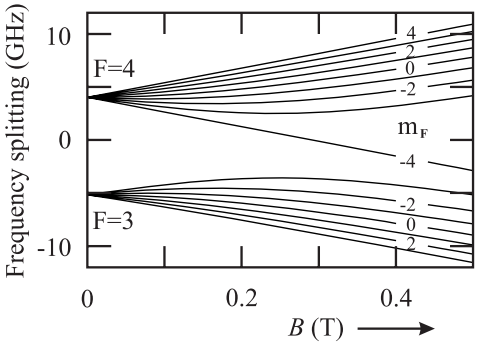


Figure 7.1: The energies of the hyperfine components $F = 3$ and $F = 4$ of the $6\,^2S_{1/2}$ state in ^{133}Cs, calculated according to (5.147), are split in a magnetic field into 16 components ($F = 4, +4 \leq m_F \leq -4$) and ($F = 3, +3 \leq m_F \leq -3$).

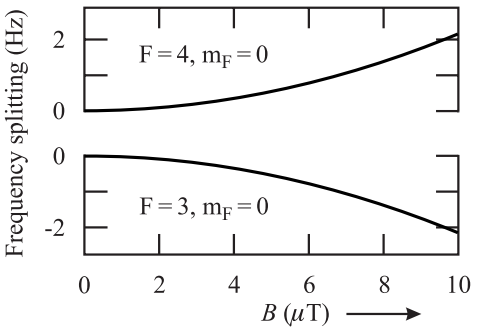


Figure 7.2: The transition frequency between the $F = 4, m_F = 0$ and the $F = 3, m_F = 0$ states of 9 192 631 770 Hz at zero magnetic field is used to define the unit of time, the second.

To excite this $\Delta m_F = 0$ magnetic dipole transition the magnetic quantisation field and the magnetic component of the oscillating field levels are chosen to be parallel. For a weak magnetic field in the μT range used in atomic clocks the other transitions between the levels with $m_F \neq 0$ show a linear Zeeman effect

$$\Delta\nu_B = (g_{F=4} - g_{F=3}) g_J m_F \mu_B B \approx 6.998 \times 10^3\,\text{Hz}\; m_F \frac{B}{\mu\text{T}} \tag{7.1}$$

(see Fig. 7.1). For weak magnetic fields the $m_F = 0$ levels have a small quadratic dependence on the field (see Fig. 7.2) leading to a frequency shift of

$$\Delta\nu_{B^2} \approx 4.274\,5 \times 10^{-2}\,\mathrm{Hz} \left(\frac{B}{\mu\mathrm{T}}\right)^2 . \tag{7.2}$$

In a caesium atomic clock the atoms are prepared such that they are either in the $F = 4, m_F = 0$ or in the $F = 3, m_F = 0$ state. Afterwards the atoms interact with an electromagnetic field that induces transitions into the former unoccupied state. The atoms in this state are detected and allow one to determine the frequency of the interrogating field where the transition probability has a maximum. The observed transition frequency is corrected for all known frequency offsets that would shift the transition frequency from the unperturbed transition and is used to produce a standard frequency or a Pulse Per Second (PPS) every 9 192 631 770 cycles.

In the following we describe first how this preparation and interrogation is implemented in commercial caesium clocks where a compromise must be found between the achievable accuracy and stability and the corresponding weight, power consumption and costs of the devices. The perturbing effects and the methods used to suppress or avoid the associated frequency shifts are discussed later in the examples of primary laboratory standards.

7.1 Caesium Atomic Beam Clocks with Magnetic State Selection

Most Cs atomic clocks operated today use an atomic Cs beam in a high vacuum chamber. The principal design is similar to the one used in the predecessor of all Cs atomic clocks developed in the National Physical Laboratory, England by Essen and Parry [14]. The atoms effuse through a nozzle or a system of channels from an oven, heated to a temperature of about 100° C or higher, which is filled with a few grams of caesium. Due to the small energy separation between the $F = 3$ and $F = 4$ states both levels are almost equally populated in the thermal beam. Consequently, the atoms must be prepared in one of these states in order to allow the detection of transitions between the levels induced by an external oscillating field. In conventional Cs atomic beam machines, like most of the commercial Cs atomic clocks and older primary clocks, atoms in a particular state are selected by means of their magnetic moments. As can be seen from Fig. 7.1, in a strong magnetic field of more than about 0.4 T the energy of a Cs atom in any one of the $F = 3$ states is reduced when the magnetic field is increased. The same is true for the $F = 4, m_F = -4$ state. The energies of atoms in any one of the other $F = 4$ states, however, increase when the magnetic field increases. Consequently, a magnetic induction $B(z)$ with spatial variation along the z direction leads to a force acting on the Cs atoms

$$F_{\mathrm{mag}} = -\frac{\partial W}{\partial z} = -\frac{\partial W}{\partial B}\frac{\partial B}{\partial z} \equiv -\mu_{\mathrm{eff}}\frac{\partial B}{\partial z} \tag{7.3}$$

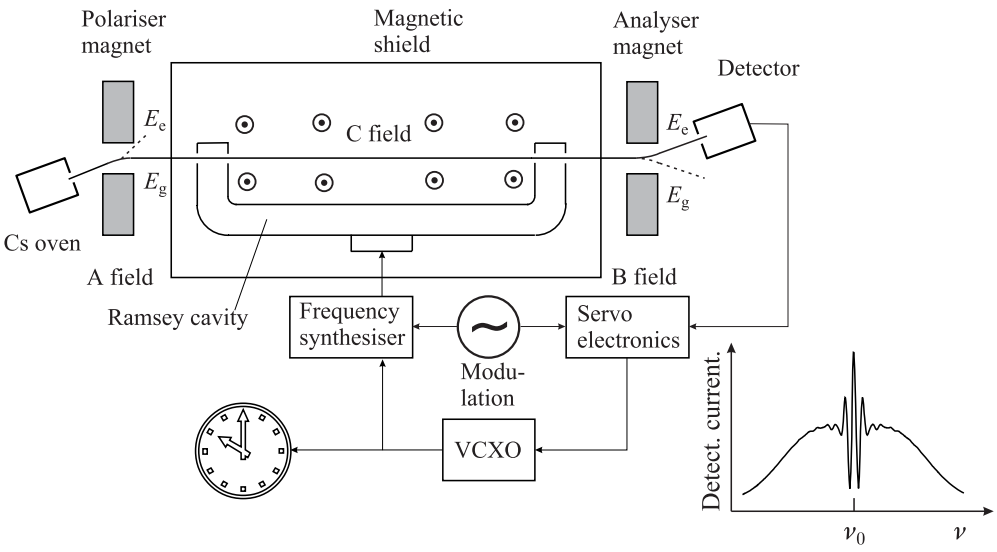


Figure 7.3: Schematic layout of a commercial Cs atomic clock. The magnetic quantisation field (C field) is perpendicular to the paper plain. The inset shows the detector current when the frequency of the synthesiser is tuned across the atomic resonance displaying the Ramsey resonance on the Rabi pedestal.

depending on their potential energy W and is given by the effective magnetic moment μ_{eff} [1] which is proportional to the gradient of the curves in Fig. 7.1.

7.1.1 Commercial Caesium Clocks

Caesium atomic clocks have been produced since the 1950s beginning with the Atomichron [15] the earliest commercially available caesium-beam atomic frequency standard. Fifty years later most commercial products [29] use a basic layout shown in Fig. 7.3. A so-called magnetic polariser with an inhomogeneous magnetic field deflects the atoms according to their magnetic moment and can be used to select the atoms in the desired state. Consider the $m_F = 0$ states of the clock transition. As the $F = 3, m_F = 0$ state is a high-field seeker in contrast to the $F = 4, m_F = 0$ state which is a low-field seeker (Fig. 7.2) one of these can be eliminated from the beam. In Fig. 7.3 it is assumed that the latter one, with energy E_g, enters the interaction regions provided by the U-shaped microwave resonator tuned to 9.192 GHz.

The resonator is usually a standard waveguide of rectangular cross section with end plates that short circuit the wave guide. The resonator may be bent in a way shown in Fig. 7.3 and the atomic beam enters and leaves the two end sections of the resonator through small holes near the end plates. In Fig. 7.3 the transverse dimensions of the waveguide resonator are chosen such that the magnetic field lines of the standing electromagnetic field inside the resonator

[1] Only at very weak magnetic fields (Zeeman regime) and at very high magnetic fields (Paschen-Back regime) is the magnetic moment $\frac{\partial W}{\partial B}$ given by the slopes of the curves in Fig. 7.1 constant. In the intermediate regime the effective magnetic moment μ_{eff} varies with the field.

are perpendicular to the paper plane. In this case the atoms crossing the terminal parts of the resonator experience a constant field inside the resonator. The resonator is fed by the radio frequency derived from a voltage-controlled quartz crystal oscillator (VCXO), typically an OCXO as characterised in Table 4.1. The atoms in the E_g state pass the first and later the second end segment of the U-shaped microwave resonator where transitions from the $F = 3$ state to the $F = 4$ state of the Cs atoms can be induced in a Ramsey excitation (Section 6.6) depending on the detuning between the frequency of the external oscillator and the atomic transition. A second magnet combination, i.e., the analyser, deflects the atoms in the upper state into the detector.[2]

A constant magnetic field is used to separate energetically the otherwise degenerate magnetic sub-levels (Fig. 7.1) in order to allow the excitation of the clock transition $|F = 3, m_F = 0\rangle \rightarrow |F = 4, m_F = 0\rangle$ isolated from the other transitions. By convention such a field is referred to as the C field as it is applied between the fields of the polariser and the analyser which historically were called the A field and the B field, respectively. The magnitude of the C field is chosen as a compromise between two conflicting requirements. First, it has to be large enough to separate the otherwise overlapping resonances. Second, the C field shifts the resonance frequency quadratically according to (7.2) which has to be corrected, as will be discussed below. However, in a larger field the frequency of the clock is influenced to a larger extent by fluctuations of the magnetic field. In the scheme of a commercial Cs clock depicted in Fig. 7.3 the C field is often generated by a coil with windings in the paper plane wound around the Ramsey resonator and hence, points perpendicularly to that plane. Owing to the dependence of the frequency of the clock transition from the magnetic field, efficient magnetic shielding has to be provided in order to attenuate the ambient magnetic field and the magnitude of the associated fluctuations.

The detector for the Cs atoms can consist of a hot-wire detector (Langmuir–Taylor detector) consisting of a transition-metal ribbon made from tungsten or iridium-platinum which is heated to prevent surface layers due to adsorbed gases. Due to the large difference in the work functions of metallic Cs (1.7 eV) and, e.g., metallic tungsten (4.5 eV) a Cs atom is easily ionised at the hot wire, giving its outer electron to the tungsten metal. An applied voltage allows detection of the positively charged Cs atoms, e.g., directly in a Faraday cup as often employed in primary caesium clocks. In commercial caesium clocks a mass filter is used to select the caesium ions from other ions that are produced near the detector. The caesium ions are then directed onto the first stage of a photoelectron multiplier where the current of the electrons emitted by the impinging ions is amplified. In contrast to collecting the caesium ions directly in a Faraday cup, this detection method is much faster and allows one to utilise higher modulation frequencies for the stabilisation onto the resonance.

Scanning the frequency ν of the synthesiser around the frequency ν_0 of the atomic resonance leads to a detector current like the one shown in the inset of Fig. 7.3. The signal shows the Ramsey resonance structure on a broader, so-called, Rabi pedestal. As has been shown in Section 6.6 the Ramsey resonance results from atoms that have been coherently excited in the two interaction zones of the Ramsey resonator. The interaction with the rf field

[2] Depending on the position of the detector, the Cs atoms that are excited in the resonator are either deflected into the detector or are guided to miss it. These two "flop-in" and "flop-out" techniques therefore lead to a peak or to a dip of the signal at resonance, respectively.

in the first interaction zone puts the caesium atoms into a coherent superposition between the $F = 4, m_F = 0$ and the $F = 3, m_F = 0$ state. The temporal evolution of the quantum mechanical state of the Cs atom occurs with a frequency corresponding to the energy difference of these states. After the interaction of the atom in the second zone of the Ramsey cavity the probability of finding the atom in the $F = 4$ state or in the $F = 3$ state depends whether the external rf field is in phase or out of phase with the atomic oscillator. Hence, the number of atoms in either the $F = 4$ or in the $F = 3$ state oscillates as a function of the frequency of the external oscillator thereby leading to the Ramsey interference structure. Due to the large width of the atomic velocity distribution, only the central Ramsey fringes survive the associated velocity averaging. By contrast, the Rabi pedestal reflects the Doppler broadened line resulting from the interaction of the atoms in a single zone.

The central feature with its maximum at the transition frequency ν_0 is used to stabilise the frequency of the VCXO to the atomic transition frequency. To this end, the frequency from the synthesiser is modulated across the central peak. The signal from the detector is phase-sensitively detected in the servo electronics (Section 2.3.2), integrated and the servo signal is used for stabilising the frequency of the VCXO. From this suitable output frequencies are derived, such as 5 MHz or a 1 PPS signal.

As a result of the second-order Zeeman effect the atoms experience in the C field region, the centre frequency of the Ramsey resonance is shifted from the frequency of the frequency of the unperturbed transition defined as 9 192 631 770 Hz. To take into account the corresponding frequency offset, the value of the chosen C field is usually determined by using (7.1) and the associated frequency shift is determined from (7.2) and added in the synthesiser to make sure that the output frequency from the VCXO represents the exact SI value.

Caesium atomic beam clocks are commercially available from several manufacturers [29]. They fit into a 19 inch rack and have a mass of less than 25 kg and a power consumption of less than 50 W. The specified fractional inaccuracies range from 2×10^{-12} to 5×10^{-13}. The measured instability of commercial clocks is shown in Fig. 3.3. The flicker floor of 5×10^{-15} can be reached after about 10 days of averaging [28, 29].

Commercial caesium atomic clocks find applications in several fields. To begin with, they are used in time-keeping laboratories. About two hundred of them are used to give a stable atomic time scale (TAI) (Section 12.1.2). Commercial caesium atomic clocks are furthermore used in Global Navigation Satellite Systems (GNSS) (Section 12.5) such as GPS, GLONASS, or GALILEO. They are used there in the ground segments and sometimes also in the satellites. A third application for commercial caesium clocks is in the telecommunications sector covering all forms of distance communications, including radio, telegraphy, television, telephony, data communication and computer networking. There, atomic clocks are used to synchronise the different networks. Caesium atomic clocks are furthermore used, e.g., to synchronise radio-controlled clocks by a time code transmitted by various radio transmitters (Section 12.4).

7.1.2 Primary Laboratory Standards

Better accuracy than obtained with commercial caesium clocks is achieved with primary laboratory standards where the scheme of Fig. 7.3 is modified in a number of aspects. In the following we discuss the effects that contribute to the achievable accuracy using the examples

of primary laboratory standards even though the effects that limit the accuracy by and large are the same as in commercial devices. For the discussion we often refer to the examples of PTB's primary clocks CS1 and CS2 [29, 249, 250] as those standards allowed already in the 1980s to achieve fractional uncertainties as low as 10^{-14}. First, in order to achieve higher resolution, the length of the Ramsey in primary clocks is about five times larger than in commercial clocks, e.g., 76 cm in the case of PTB's primary clocks CS1 and CS2. Second, in order to use a sufficient flux of atoms, the one-dimensional deflection in a two-pole magnet (Fig. 7.3) is replaced by magnetic lenses comprising four- or six-pole magnets (Section 8.1.3.2) for the polariser and analyser. In contrast to the scheme of Fig. 7.3 the oven and the detector are arranged on a line of sight with suitable means to block the trajectories of atoms with the undesired magnetic moments. Magnetic selectors of this kind focus the atoms in a focal point that depends on the velocity of the atoms. Even though for a typical oven temperature of $T \approx 450\,\mathrm{K}$ the most probable velocity is about 250 m/s, the velocity selection in the magnetic selectors of CS1 leads to a mean velocity of about 95 m/s and a much narrower velocity distribution contributes to the Ramsey signal than in a thermal beam of Cs atoms. Consequently, in the former, a larger number of Ramsey fringes is visible (Fig. 7.4).[3] In primary laboratory

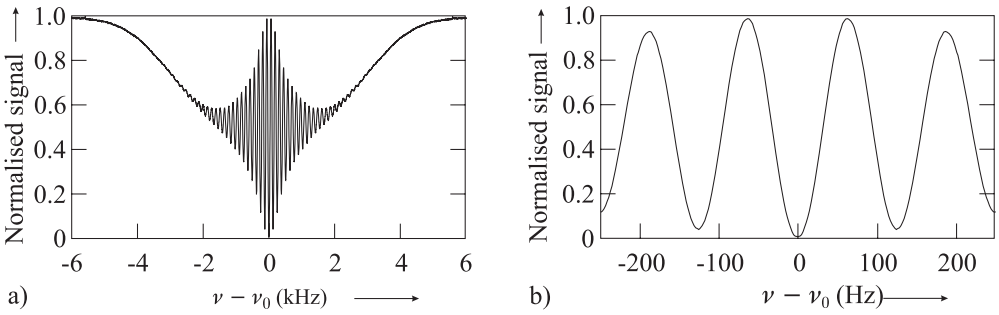


Figure 7.4: a) Ramsey resonance structure of the $F = 4, m_F = 0 \rightarrow F = 3, m_F = 0$ transition recorded using PTB's primary caesium atomic clock CS1. b) Central Ramsey fringes.

clocks, the homogeneity of the C field obtained by the method described in Section 7.1.1 is not sufficient. Hence, in general a configuration is chosen where the field is produced by a solenoid whose axis coincides with the trajectory of the Cs beam.

7.1.3 Frequency Shifts in Caesium Beam-Clocks

As a representative of the state-of-the-art primary beam-machine clocks, PTB's CS1 has a stability of $5 \times 10^{-12}\sqrt{\tau/\mathrm{s}}$ and an accuracy of 7×10^{-15} [249]. To derive a meaningful figure for such an uncertainty a careful analysis of all the effects has to be performed that can shift the frequency of the clock transition. Once the influences of these effects have been quantitatively established for a particular device, the frequency of the clock is corrected for the associated shifts in order to derive the frequency of the unperturbed transition. The

[3] In CS1 the detector signal is reduced if an atom undergoes a transition and hence at resonance there is a minimum in Fig. 7.4 in contrast to the resonance curve shown in Fig. 7.3, where there is a maximum.

correction, however, can be performed only with limited accuracy due to the uncertainties associated with the correction procedure. The overall uncertainty of the clock is then derived from a budget of all relevant contributions. In the following we discuss the most important sources for frequency shifts in a primary Cs atomic clock.

7.1.3.1 Influence of the Magnetic Field

The largest frequency offset results from the Zeeman effect that shifts the energy levels in the presence of a magnetic field (Fig. 7.1 and Fig. 7.2). The seven transition lines in Fig. 7.5 originate from the seven magnetic sub-states ($F = 3, -3 \leq m_F \leq 3$) with the selection rule $\Delta m_F = 0$. The highly asymmetrical population of the different Zeeman states of Fig. 7.5 results from the state selection which is performed by deflecting the Cs atoms in an inhomogeneous magnetic field.

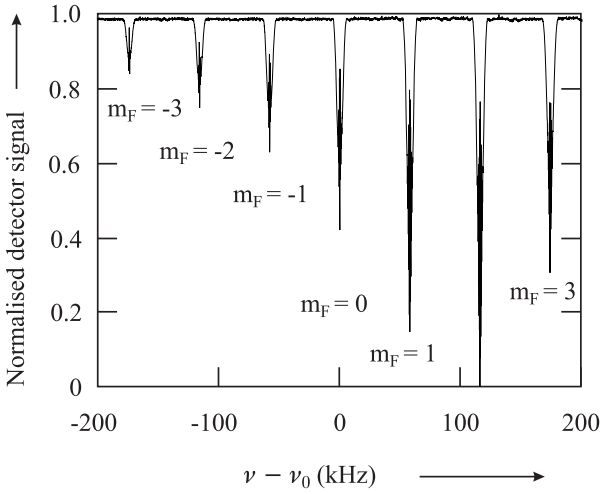


Figure 7.5: Zeeman splitting of the rf resonance transition $F = 4 \rightarrow F = 3, \Delta m_F = 0$ in a weak magnetic field of about 8 µT in PTB's caesium atomic clock CS1.

For a typical value of the C field near 8 µT the frequency shift is 2.7 Hz corresponding to a relative frequency shift of 3×10^{-10}. The average magnetic field $\langle B \rangle$, to which the Cs atoms are exposed during their passage between the two Ramsey zones, can be determined from the frequency separation of two rf resonances shown in Fig. 7.5. To determine the frequency shift of the clock transition one has to keep in mind that the correction for the quadratic Zeeman shift asks for the determination of $\langle B^2 \rangle$ (see (7.2)) which equals $\langle B \rangle^2$ only in the case of a constant C field. Consequently, an excellent homogeneity of the B field is required within the C field region of the rf cavity. A fractional (rms) deviation of $\Delta B/B$ of a few 10^{-4} has been reported, which allows one to correct for the quadratic Zeeman shift with a relative uncertainty of 1×10^{-15} [249]. From $\Delta\nu_{B^2}/\nu \approx 2\Delta B/B$ a relative field fluctuation of 5×10^{-5} leads to a relative frequency shift of about 3×10^{-14} at 8 µT which requires efficient shielding of any external magnetic fields and stable field generation.

7.1.3.2 Cavity Phase Shifts

The Ramsey resonance feature (Fig. 7.4) essentially relies on a comparison between the temporally evolving internal phase $\omega_0 T$ of the Cs atom and the phase difference of the rf field ωT at the locations of the two interactions of the atom with the field (see (6.44)). These interactions take place in the two end zones of a U-shaped rf resonator of rectangular cross-section. The orientation of the magnetic C field, which defines the quantisation axis, and the resonator structure is chosen such that the magnetic field lines are parallel to the C field in the interaction zones in order to allow only excitation of the transitions with $\Delta m_F = 0$. In the arrangement shown in Fig. 7.3 the magnetic field lines of the electromagnetic field inside the resonator are perpendicular to the paper plane as does the magnetic component of the standing-wave rf field in the resonator. In primary laboratory standards, the resonator often has a shape similar to the one shown in Fig. 7.6 a, b). Consequently the C field, generated by a solenoid, is parallel to the atomic beam. Even though the electromagnetic energy of the field fed into the resonator is

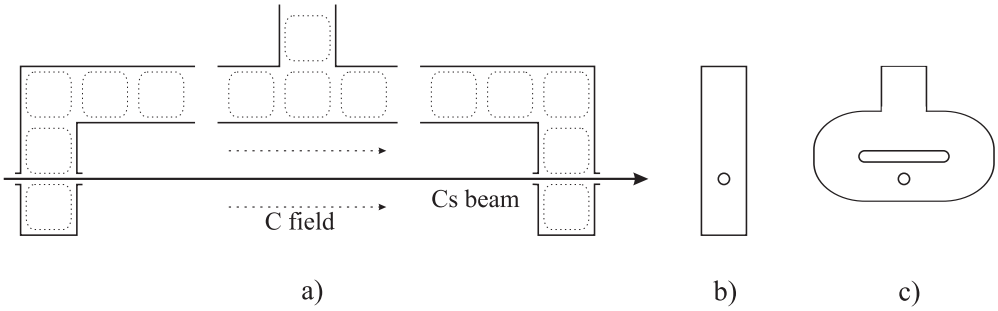


Figure 7.6: a) Cross-section of the Ramsey resonator with the magnetic field lines of the standing rf wave. b) Side view of the end segment with the hole for the trajectories of the Cs atoms. c) Side view of the race-track end section according to [251].

equally divided into both arms, allowing two interactions of each atom with exactly the same frequency ν, phase differences $\Delta\Phi$ of the electromagnetic field in the two interaction zones, in general, lead to a frequency offset according to (6.46). In the following we discuss the origin of such phases and the associated frequency shifts together methods with methods to reduce their influence.

Phase shifts occur as a result of the ohmic losses in the microwave resonator where in particular the electric losses at the end parts of the resonator lead to a reduction in the amplitude of the reflected wave with respect to the impinging wave. Hence, the resulting field inside the resonator is not completely described by a standing wave but also includes a contribution from a travelling wave towards the end caps of the resonator. If we assume that the reflecting surface has an amplitude reflection coefficient $1-\delta$ for a wave travelling along the z axis with $k = 2\pi/\lambda$, the field is

$$\begin{aligned} B &= B_0 \cos(\omega t - k_g z) + B_0(1-\delta)\cos(\omega t + k_g z) \\ &= 2B_0 \left[\cos(\omega t)\cos(k_g z) - \frac{\delta}{2}\cos(\omega t + k_g z)\right], \end{aligned} \tag{7.4}$$

where we have made use of $\cos(\alpha \pm \beta) = \cos\alpha\cos\beta \mp \sin\alpha\sin\beta$. In (7.4) k_g is the wave vector depending on the group velocity inside the wave guide. For a wavelength $\lambda_g = 2\pi/k_g \approx 4.65\,\text{cm}$, corresponding to the transition frequency of 9.192 GHz in a standard X-band waveguide as compared to the vacuum wavelength $\lambda \approx 3.26\,\text{cm}$. The first term in the brackets of (7.4) represents a standing wave whose spatial dependence is given by the $\cos k_g z$ factor thereby representing an spatial modulation of the amplitude of the magnetic field along the z coordinate. However, the sign change of this factor at odd multiples of $k_g z = \pi/2$ corresponds to a phase shift of π at these positions leading to a phase variation as depicted in Fig. 7.7 a) along the two halves of the resonator fed symmetrically at the centre. In the Cs

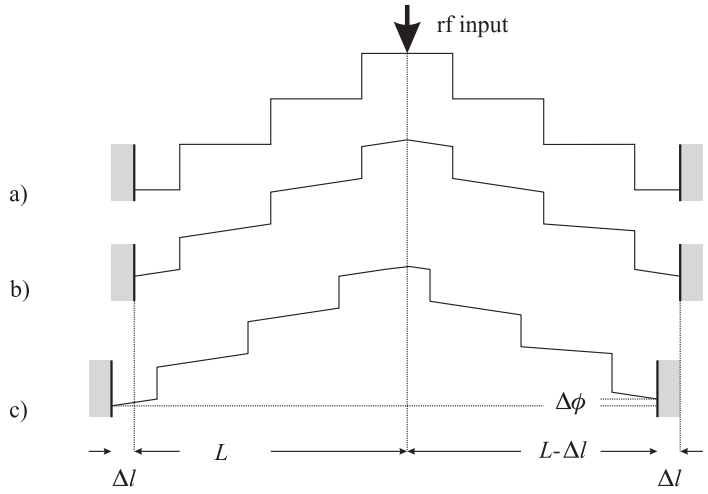


Figure 7.7: a) Phase of a standing wave in an symmetric resonator (Fig. 7.6) with infinite conductivity. b) Influence of the contribution to the phase by a travelling wave component. c) Influence of asymmetric resonator.

atomic clocks the interaction regions with the atomic beam are located near the anti-nodes of the magnetic field component so that all atoms even in a beam with finite transverse extension are subject to the same spatial phase.

For finite conductivity, i.e., for $\delta \neq 0$ there is a travelling wave contribution in (7.4) whose amplitude is proportional to the non-ideal reflection and whose phase increases linearly with $k_g z$ (see Fig. 7.7 b). This contribution of the travelling wave gives rise to a spatial variation in the phase of the microwave field in the resonator which is often termed the distributed phase shift. As a consequence, the atoms in the beam experience a position-dependent phase and the frequency of the observed resonance maximum depends on the trajectory of the beam. In order to reduce the phase gradient, de Marchi *et al.* [251] devised a special resonator (see Fig. 7.6 c) where the standing wave field splits into two parts that travel in opposite directions in the annular end part. The superposition of the two waves in the region of the atomic beam should result in a negligible energy flux at the interaction point with the atoms. Thus, the position dependence of the phase is expected to vary only quadratically in the vertical and horizontal direction [251] in contrast to the conventional resonator design where the phase

varies linearly with position. The use of such annular end parts in primary clocks has reduced the distributed cavity phase shift considerably [249, 252, 253].

Next, we consider the case where the total length of the resonator is still correct but the two arms of the resonator are of unequal lengths (Fig. 7.7 c). Since both arms are fed by the same source at the junction, at this point the phase has the same value for both arms. Since the lengths of both arms are different the phases at the terminating end parts are different leading to a so-called end-to-end cavity shift. The end-to-end phase shift changes sign if the direction of the atomic beam is reversed (see (6.46)). Consequently, in the primary caesium beam standards, such as PTB's CS1 and CS2, the fractional beam-reversal frequency shift is determined on a regular basis [250] and is about 6×10^{-13} and 5×10^{-13}, respectively. To this end in CS1 the beam reversal requires breaking the vacuum and interchanging the detector and oven assemblies with the respective magnetic lenses. In CS2, which is equipped with an oven assembly and a detector assembly at each end, the oven and detector at each end can be interchanged without breaking the vacuum. The residual fractional uncertainties after applying the beam reversal were estimated to be 0.6×10^{-14} and 1×10^{-14} for CS1 and CS2, respectively.

7.1.3.3 Influence of Neighbouring Transitions

The multilevel structure of the ^{133}Cs atom with its 16 ground states sub-levels allows the coupling of several levels apart from the $m_F = 0 \rightarrow m_F = 0$ clock transition that can shift the experimentally determined transition frequency in different ways. If one keeps in mind that the linewidth of about 60 Hz, corresponding to a $Q \approx 1.5 \times 10^8$, has to be split to about 10^{-6} to reach a relative uncertainty below 10^{-14} one appreciates the importance of even the slightest contributions to any asymmetry of the lineshape.

Rabi Pulling A particular effect that may shift the apparent central minimum (Fig. 7.4) from the frequency of the unperturbed atom is called Rabi pulling [254]. It results from an overlap of the central $F = 3, m_F = 0 \leftrightarrow F = 4, m_F = 0$ line, with the wings of the adjacent $F = 3, m_F = 1 \leftrightarrow F = 4, m_F = 1$ and $F = 3, m_F = -1 \leftrightarrow F = 4, m_F = -1$ lines (Fig. 7.5). If the two latter contributions are not symmetric, as is the case in the signal shown in Fig. 7.5, their influence leads to an asymmetric Rabi pedestal of the $m_F = 0 \rightarrow m_F = 0$ transition. A quantitative treatment of the Rabi pulling [11, 254, 255] leads to a frequency shift that is proportional to the population difference of the ± 1 Zeeman components. Hence this shift can be largely reduced by generating a more symmetric spectrum or by depleting the $m_F \neq 0$ states, e.g., by optical pumping. The shift is smaller in clocks with smaller linewidth or with a larger C field as the latter determines the separation between the lines (Fig. 7.5). Since the amplitudes in the wings of the transitions is proportional to the microwave power the frequency shift is also proportional to this power. Hence a power dependence of the frequency of the caesium clock may indicate the presence of Rabi pulling.

Ramsey Pulling Another type of frequency shift referred to as Ramsey pulling [255, 256] results from the contributions of transitions with $\Delta m_F = \pm 1$. Such transitions, barely visible in Fig. 7.5, e.g., between the components labelled $m_F = 2$ and $m_F = 3$, are excited by components of the rf field which are orthogonal to the desired direction of the quantisation

field (C field). Since the $\Delta m_F = \pm 1$ transitions $F = 3, m_F = 0 \leftrightarrow F = 4, m_F = \pm 1$ and $F = 3, m_F = \pm 1 \leftrightarrow F = 3, m_F = 0$ couple to the states of the clock transition $F = 3, m_F = 0 \leftrightarrow F = 3, m_F = 0$ these states can be perturbed when the above transitions occur giving rise to a frequency shift of the clock transition. Due to the finite extension of the atomic beam and the distribution of the field inside the cavity, the transition amplitude into a particular sub-state differs for atoms following different trajectories. As the different paths cannot be identified the associated transition amplitudes give rise to an interference structure that can shift the apparent minimum or maximum used to determine the clock transition. The quantitative theoretical treatment given in the literature is inconsistent [255, 256]. Experimentally, the frequency shift can be identified by its oscillation with a variation of the C field strength. Its influence can result in a frequency shift of a few parts in 10^{13} for short (commercial) clocks but has been shown to be two orders of magnitude smaller in a primary standard with a homogeneous C field [249].

Majorana Pulling Frequency shifts can also result from transitions induced between different magnetic sub-states of the same hyperfine transition ($\Delta F = 0, \Delta m_F \neq 0$; Majorana transitions; [182, 257]). Majorana transitions occur if the magnetic moment of a state-selected atom cannot follow adiabatically the inhomogeneous static magnetic field along its path between the polariser and the analyser. The influence of Majorana pulling is affected when the magnitude of the C field is altered and when the rf power is varied allowing one to identify these contributions [249, 257].

7.1.3.4 Frequency Shift due to Black-body Radiation

In neutral-atom atomic clocks such as the Cs clock, a dc electric field in general does not represent a major source of uncertainty as electric fields can be shielded effectively. However, the Cs atoms are exposed inevitably to the electromagnetic temperature-radiation field which is emitted from each body with a temperature $T \neq 0\,\mathrm{K}$.

As was discussed in Section 5.4.5.2 the ac hyperfine Stark shift due to an electric field is approximately equal to the shift due to a static field with the same rms value. Itano *et al.* [258] calculated the shift for several alkali metals and have shown that the above approximation is justified. In the case of the Cs atomic clock they calculate the fractional ac Stark shift of the hyperfine transition to be $-1.69(4) \times 10^{-14}\,[T/300\,\mathrm{K}]^4$. Pal'chikov *et al.* [259] calculated the fractional frequency shift as

$$\frac{\delta\nu}{\nu} = -17.2 \times 10^{-15} \left(\frac{T}{300\,\mathrm{K}}\right)^4 \times \left[1 + 0.014 \left(\frac{T}{300\,\mathrm{K}}\right)^2 - 3.18 \times 10^{-5} \left(\frac{T}{300\,\mathrm{K}}\right)^2\right] \tag{7.5}$$

where the first term corresponds to the polarisability and the second term inside the square brackets is a correction due to the separation in the frequencies of the D1 ($\lambda = 894\,\mathrm{nm}$) and D2 ($\lambda = 852\,\mathrm{nm}$, Fig. 7.8 a) lines in caesium. The third term in the square brackets represents the third-order term in the polarisability of the atom (hyperpolarisability) giving rise to a higher-order Stark effect. Bauch and Schröder [260] have measured the black-body shift in a thermal beam apparatus to be $-(16.6 \pm 2) \times 10^{-15}$ in good agreement with the expected values.

Simon *et al.* [261] have measured directly the dc-Stark shift in a Cs atomic fountain clock to be $\delta\nu = -2.271(4)\times 10^{-10}\, E^2\,\mathrm{Hz\,(V/m)^{-2}}$. From this value and the time averaged electric field of the black-body radiation at 300 K (see (5.142)) one calculates $\delta\nu/\nu = 17.09(3)\times 10^{-15}$ with an order-of-magnitude reduced uncertainty. For the best Cs clocks the black-body shift at room temperature is more than an order of magnitude larger than the uncertainty of realising the unperturbed line centre of the clock. Hence, it is now generally agreed to correct the Cs atomic clocks for the black-body shift, for realisation of the unit of time. In the long run the influence of the time-dependent electric field of the black-body radiation might be the largest source contributing to the uncertainty of the most advanced frequency standards. However, this influence can be largely avoided by keeping the absorbers at cryogenic temperature as is done, e.g., in the case of the optical mercury frequency standard (Section 10.3.2.4).

7.1.3.5 Gravitational Frequency Shift

Since the definition of the second is based on proper time, a direct comparison of the frequencies of two non-local clocks has to take into account the influence of the gravitational potential Φ at the location of each clock (see Section 12.2). According to general relativity the frequency is shifted by $\Delta\nu = -\nu\,\Phi/c^2$ where the potential includes the contributions due to gravitational acceleration as well as those due to the acceleration, e.g., on the surface of the rotating Earth. If the potential is referenced to the rotating geoid the gravitational potential can be expressed by $\Phi = gh/c^2$ where $g \approx 9.81$ m/s^2 is the local acceleration due to Earth's gravitation and h is the height above the surface of the geoid. For small heights the fractional frequency shift is therefore 1.09×10^{-16} m^{-1}. The corresponding corrections are: -8.7×10^{-15} in the case of PTB's clocks at an altitude of 79.5 m; or -180.54×10^{-15} [20,262] for NIST's clocks in Boulder, at an altitude of 1.6 km. For such a large difference in altitudes and the accuracies obtained with present days best clocks it suffices no longer to approximate the potential by $\Phi = gh/c^2$ [263].

7.1.3.6 Second-order Doppler Shift

Time dilation leads to a frequency shift $\delta\nu = -\nu v^2/(2c^2)$ with respect to the proper frequency if the Cs atoms moving with velocity v are observed in the laboratory frame (Section 5.4.2 and Section 12.2). For a mean velocity of 95 m/s in CS1 the fractional shift is about 5×10^{-14}. The observed Ramsey pattern results from the contributions of all atoms with their different velocities and hence the information about the relevant velocity distribution is contained in this spectrum. In general, the velocity distribution can be calculated from the Fourier transforms of the Ramsey lineshapes (see e.g., [264,265] and references therein). The fractional contribution to the uncertainty after correction of the second-order Doppler shift leads to 0.5×10^{-15} and 1×10^{-15} for PTB's CS1 and CS2, respectively.

7.1.3.7 Frequency Shifts due to Non-perfect Instrumentation

Cavity Pulling Cavity pulling occurs if the eigenfrequency of the microwave cavity is not tuned exactly to the atomic resonance. Consider a stabilisation scheme where the frequency

of the interrogating oscillator is modulated by $\pm\gamma/2$ with γ the full width of the central Ramsey fringe. If the cavity is detuned from the resonance, the two frequencies are located on different positions of the resonance curve of the resonator which leads to different excitation probabilities of the atoms. An approximation for the corresponding fractional frequency shift in Cs clocks is given as follows [11, 266]

$$\frac{\Delta\nu_c}{\nu_0} \approx \frac{\nu_c - \nu_0}{\nu_0} K_c \frac{Q_c^2}{Q_{\mathrm{at}}^2}. \tag{7.6}$$

In (7.6) ν_c and ν_0 are the resonance frequencies of the atomic absorber and the resonator, respectively. Q_{at} and Q_c are the line quality factors of the atomic resonance as measured in the given device, respectively. K_c is a constant depending on the modulation width, the velocity distribution, and the dipole moment of the atoms. Thus, the cavity pulling effect can be reduced by employing a resonant cavity with low Q_c. If the clock is operated at optimum power, i.e., with $\pi/2$ excitation in each zone, the transition probability is largely independent of the power. Varying the power allows one to identify the cavity pulling effect.

Since the cavity pulling effect is more pronounced in the case of the Rabi pedestal, as in the Ramsey fringes, a power-dependent shift can be used to determine and to correct for this influence as Shirley *et al.* [267] have pointed out. There are several other effects that affect the Rabi pedestal more than the Ramsey fringes including shifts due to the inhomogeneity of the magnetic field, Rabi pulling, an asymmetric microwave spectrum, or light shift in optically pumped standards.

Frequency Shifts Related to the Electronics Even carefully devised electronics and selected components may introduce shifts that limit the performance of the clocks. Examples include integrator offsets in the servo amplifier or insufficient purity of the signals. As the servo control is supposed to bring the error signal to zero, any offset will be counteracted by a frequency offset that compensates the electronic offset. Spurious frequency components in the microwave field used to interrogate the atoms can also induce frequency shifts. The different microwave signals present at various stages of a frequency standard are later processed in very different applications, e.g., to derive timing signals. Any non-linear element in the corresponding electronic devices may result in phase or frequency shifts if the spectral purity of the primary signal is not sufficient. Mixtures of amplitude and phase modulation may occur that shift the output frequency. Great care has to be taken and the art of clock-making depends on the ability to reduce the associated deficiencies to a level where they are no longer significant.

Microwave Leakage Spurious microwave leakage in the resonator can induce transitions outside the well-defined interaction zones, leading to asymmetric resonance curves and can shift the centre frequency considerably [268]. In general, these effects can be identified by varying the microwave power applied to the clock from the optimum $\pi/2$ Rabi angle to higher odd integer multiples $(3\pi/2, 5\pi/2, \dots)$.

Uncertainty budget For primary standards, careful examination of all relevant shifts is usually performed on a regular basis and published where the offsets are corrected and reported. The uncertainties of the relevant contributions are evaluated in a standardised way [3] leading to a so-called "uncertainty budget". It is now generally agreed to add all contributions to the

uncertainty in quadrature and to give a single value for the uncertainty of the measurement corresponding to the evaluation.

7.2 Optically-pumped Caesium Beam Clocks

The magnetic state selection, previously described, suffers from the fact that at thermal energies all 16 states (Fig. 7.1) are almost equally populated and only 1/16 of all atoms, i.e., the ones in the desired state, e.g., the $F = 3, m_F = 0$ state are selected. If all atoms could be prepared in the $|F = 3, m_F = 0\rangle$ state, the signal could be greatly increased. Preparation rather than selection can be performed by optical pumping (Section 5.3.3.1) as has been first demonstrated by Picqué [269] using an atomic beam. In the simplest approach (see Fig. 7.8 a), light

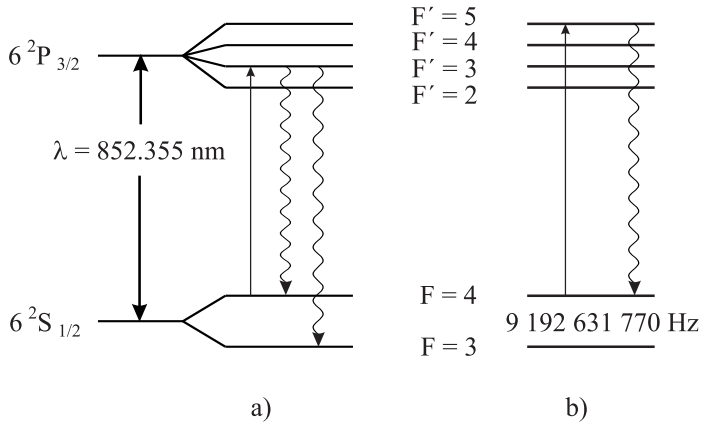


Figure 7.8: A simple version of optical pumping of a Cs atomic clock for state preparation.

resonant with the D_2 line at $\lambda = 852.355$ nm between the $F = 4$ and $F' = 3$ states is absorbed by the atoms in the $F = 4$ state. The atoms excited into the $F' = 3$ states decay after about 30 ns by spontaneous emission into the $F = 4$ and $F = 3$ hyperfine-split ground states. Since only the atoms in the $F = 4$ state are excited again and again, after a few cycles all population from the $F = 4$ state is optically pumped into the $F = 3$ state. Other pumping schemes and the resulting distributions in the population of the states have been described by Avila *et al.* [270]. There is another important advantage of optical state preparation as compared to magnetic selection. If the optical pumping beam is perpendicular to the atomic beam, the excitation is not velocity selective. In contrast to the limited angle of acceptance in the magnetic selector, the optical pumping produces a beam which can have a higher number of atoms and can be spatially more homogeneous. As a third advantage the optical state preparation avoids the strong magnetic gradients used in magnetic selectors. Therefore, Majorana transitions that can lead to additional frequency shifts (Section 7.1.3.3) can be mainly avoided.

After excitation in the Ramsey zones there are atoms in the $F = 4$ state. These atoms can be detected by exciting them by a laser beam tuned to the $F = 4$ to $F' = 5$ transition and by monitoring the fluorescence photons from this decay (see Fig. 7.8 b). The $F = 4$ to $F' = 5$

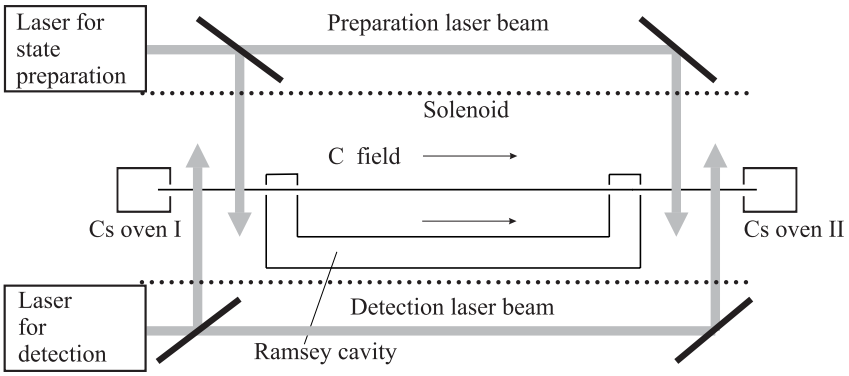


Figure 7.9: Simplified set-up of a Cs atomic clock with optical state selection and detection.

transition is called a cycling transition because, due to the quantum mechanical selection rule $\Delta F = 0, 1$, atoms excited to the $F' = 5$ level can decay only back to the $F = 4$ state and the excitation and emission process can take place many times. As a consequence each excited atom can emit a large number N of photons. Even for a small detection probability $p \ll 1$ the number of detected photons $N \times p$ easily exceeds unity, thereby allowing one to detect virtually each excited atom.

In general, a Cs beam clock using optical excitation for state preparation and detection (Fig. 7.9) has two laser systems for this purpose. Depending on the chosen scheme, more lasers are used to deliver the necessary frequencies. The central part with the Ramsey excitation is similar to the one employed in beam machines and hence the uncertainty budget has much in common with the clocks using magnetic state selection. A particular error source that is connected with optically pumped frequency standards is due to spurious light from the pumping and detection region that is present along the atomic beam path between the interaction regions. This radiation can introduce a light shift due to the ac Stark effect. It has been modelled for particular cases [11, 271] but it can also be studied experimentally by changing the laser power. Atomic beam reversal for determining the cavity shift is much more easily accomplished since the ovens do not have to be moved.

Several primary Cs clocks using optical state preparation and detection have been built and operated, e.g., in Japan at the former NRLM (now NMIJ) [272, 273] and at the former CRL (now National Institute of Information and Telecommunications Technology, NICT [253]), in France at the former LPTF (now BNM–SYRTE) [271, 274], and in the USA at the NIST [253, 275]. Fractional instabilities of about $1 \times 10^{-12}\sqrt{\tau/\mathrm{s}}$ [253, 276] and $3.5 \times 10^{-13}\sqrt{\tau/\mathrm{s}}$ [271] have been achieved. Relative uncertainties of between 10^{-14} and 10^{-15} have been reported [253, 271, 273] for these types of clocks.

7.3 Fountain Clocks

The line quality factor achieved with typical laboratory atomic clocks based on thermal Cs beams is $Q \approx 10^8$. To attain fractional uncertainties around 10^{-14} requires that the centre

of the atomic transition is located to 10^{-6} of the linewidth, which can hardly be improved further. The development of more accurate clocks thus requires increasing the interaction time and thereby reducing the linewidth of the atomic transition. Laser cooled samples with velocities of a few centimetres per second allow interaction times of the order of seconds. Horizontal beam machines are no longer appropriate for these low-velocity atoms as they would fall several metres during this time, as a result of the gravitational acceleration.

To interrogate low-velocity atoms in the gravitational field the concept of an atomic "fountain" sometimes referred to as Zacharias fountain [4] is now widely used for clocks with greatly increased accuracy.

7.3.1 Schematics of a Fountain Clock

In an atomic fountain (Fig. 7.10) a cloud of cold atoms is launched vertically through an interaction region with a velocity of a few metres per second. The gravitational acceleration g forces the atoms to slow down and to fall back, thereby passing the same interaction zone with the electromagnetic field a second time. As in the Cs atomic clocks with beam machines, the resulting resonance feature displays a Ramsey interference structure with a resolution that is determined by the time T between the two interactions. This time is calculated from the time necessary for the atoms to climb up to the apogee and the same time to fall down

$$T = 2\sqrt{\frac{2H}{g}}. \tag{7.7}$$

For an apparatus with a typical height between the Ramsey resonator and the apogee of H= 1 m, this time is T = 0.9 s and the necessary starting velocity of the atoms is $v = \sqrt{2gH} =$ 4.5 m/s. The low velocities of the atoms in the atomic cloud are the prerequisite for an efficient operation of such an atomic fountain inasmuch as the number of atoms returning through the opening in the Ramsey cavity on their way down depends on the transverse velocity. An initially very small cloud of caesium atoms with a temperature $T = 2\,\mu$K will spread to 1.1 cm in 1 s thereby allowing about 40 % of the atoms to enter the 1 cm interaction region, in contrast to about 0.7 % that would enter the same region if the sample was cooled only to the Doppler limit of 125 µK [277]. The first successful predecessor of a fountain clock experiment used laser cooled sodium atoms in a pulsed radio frequency cavity and demonstrated a linewidth of 2 Hz [278]. The first atomic fountain clock with Cs atoms was realised in the former LPTF [17, 279]. Later a variety of atomic fountains using Cs or Rb atoms were operated or investigated for atomic clocks in several institutions [280–289]. Even though these realisations differ somewhat in their design, each set-up essentially contains three sections (Fig. 7.10): a preparation zone where the atoms are collected and cooled, an excitation zone containing the cavity resonator and the zone for the ballistical flight and a third zone for the detection. The cloud of cold atoms is often prepared from a thermal vapour fed by a Cs reservoir at a base pressure of about typically 10^{-6} Pa. In a magneto-optical trap about 10^7 Cs atoms are collected and further cooled in an optical molasses to typically about 2 µK. If necessary the Cs atoms are sometimes further cooled by means of special cooling techniques [168, 290].

[4] Zacharias at the Massachussetts Institute of Technology is reported to have undertaken the first attempt to use the low-velocity atoms from an effusive source in a fountain apparatus [15].

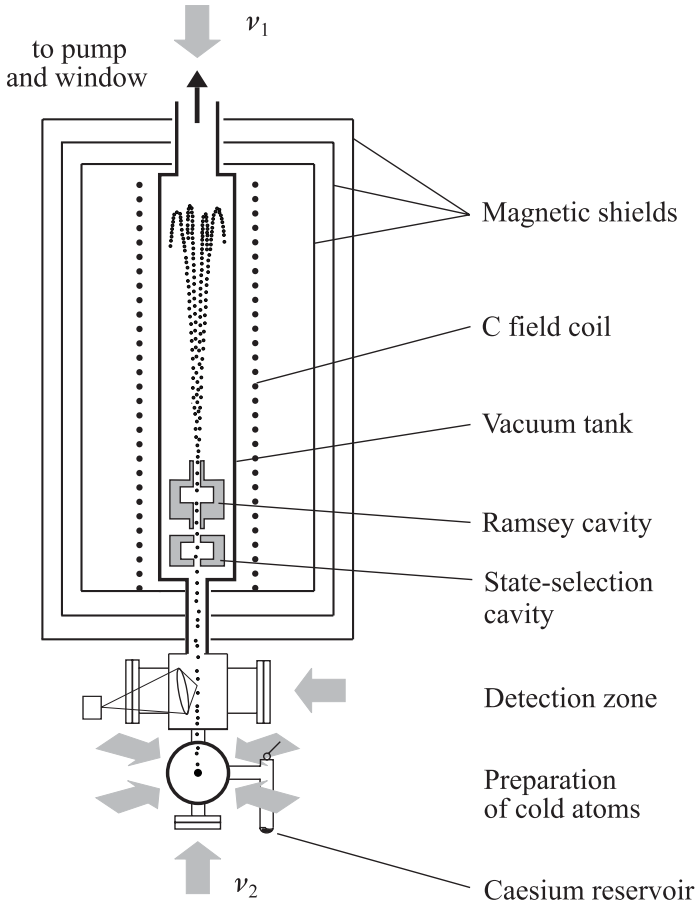


Figure 7.10: Set-up of an atomic fountain clock.

In the next step the atoms have to be launched upward without increasing their temperature. This is conveniently done in a moving molasses where the frequency ν_1 of the downward pointing vertical laser beam is red-detuned by $\delta\nu$ and the frequency ν_2 of the upward pointing beam is blue-detuned by $\delta\nu$ with respect to the frequency ν used to cool the atoms in the molasses. Choosing z in the upward direction, the superposition of two counter-propagating waves of equal amplitudes can be written as

$$\begin{aligned} E(z,t) &= E_0 \exp i[(\omega+\delta\omega)t - kz] + E_0 \exp i[(\omega t - \delta\omega) + kz] \\ &= 2E_0 \exp i\omega t \cos(\delta\omega t - kz). \end{aligned} \tag{7.8}$$

Consider the phase fronts with $\delta\omega t - kz = 2\pi\delta\nu t - 2\pi z/\lambda = 0$. From this relation one finds that they move upward with a velocity $v = z/t$ given by

$$v = \lambda\delta\nu. \tag{7.9}$$

As the atoms are laser cooled in this "walking wave" they also move upward with the velocity of the phase fronts.

Besides the arrangement of the fountain shown in Fig. 7.10, where four of the six beams for the MOT are arranged horizontally and two vertically, a so-called "1 1 1" arrangement is often used. In this arrangement the hypothetical "cube" whose six planes, defined by the six MOT beams, is oriented such that two opposite corners are aligned along the vertical direction. Each of the mutually perpendicular three pairs of counter-propagating beams intersects the tube axis at the same angle ($\approx 54.7°$). A detuning $\delta\nu$ between the counter-propagating beams creates a moving molasses [17] that leads to a velocity $v = \lambda\delta\nu\sqrt{3}/2$ of the atoms passing through the microwave cavity for interrogation. This arrangement may lead to less stray light in the cavities and the drift region. To launch the atoms, all three beams pointing downward have to be red-detuned and all upward beams blue-detuned.

When the detuned beams are switched off, the atoms follow the trajectories of a ballistic flight. The atoms pass the microwave resonator, e.g., a TE_{011} cavity on their way up and down, and hence experience a Ramsey excitation, where transitions between the ground states $|F = 4, m_F = 0\rangle \leftrightarrow |F = 3, m_F = 0\rangle$ are induced. In PTB's clock CSF1, typically 5×10^5 atoms come back through the resonator to the detection zone. Atoms that have undergone a transition in the microwave cavity can be detected in different ways. A scheme which is often used [281] detects the atoms in the $|F = 3\rangle$ and $|F = 4\rangle$ states separately. First, the atoms cross a standing-wave laser field tuned to the $|F = 4\rangle \rightarrow |F' = 5\rangle$ transition (Fig. 7.8). Cycling the atoms between these states, a large number of fluorescence photons can be detected by a photodetector whose signal is proportional to the number of atoms $N(F = 4)$ in the $|F = 4\rangle$ state. The standing wave is employed in order not to accelerate the atoms by a unidirectional light pressure. Behind the first detection zone, these atoms in the $|F = 4\rangle$ state are pushed away by a single transverse beam in order to prevent them from reaching the second detection zone. There, the atoms in the $|F = 3\rangle$ state are pumped by a second standing-wave laser field into the $|F = 4\rangle$ state from where they are detected in the same way, by a second photodetector, as the atoms in the first detection zone. Hence, the laser field in the second detection zone comprises the two frequencies necessary to excite the $|F = 3\rangle \rightarrow |F' = 4\rangle$ (pumping) transition and the $|F = 4\rangle \rightarrow |F' = 5\rangle$ (cycling) transition. The signals from the two photodetectors are combined to give the rate of atoms excited by the microwave field, normalised to the total number of atoms

$$p \propto \frac{N(F = 3)}{N(F = 3) + N(F = 4)}. \tag{7.10}$$

Excitation spectra for a particularly low apogee and a typical height of the atoms in the fountain, i.e., for low and high resolution, are shown in Fig. 7.11 and Fig. 7.12, respectively.

State-selection Cavity Several effects that may lead to frequency shifts of a Cs atomic clock are directly related to population in states others than $|F = 3, m_F = 0\rangle$ and $|F = 4, m_F = 0\rangle$ like the shifts due to Rabi pulling, Ramsey pulling, Majorana pulling, cavity pulling or cold collisions (Section 7.3.2.1). Hence, most of the fountain clocks comprise a second microwave resonant cavity that can be used to prepare the atomic population in the desired states. In PTB's fountain, atoms in the state $|F = 4, m_F = 0\rangle$ are transferred in the state-selection cavity by a π pulse to the state $|F = 3, m_F = 0\rangle$. Atoms remaining in the state $|F = 4\rangle$ that

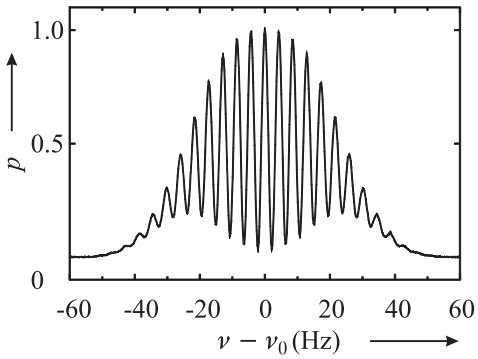


Figure 7.11: Measured probability p of exciting Cs atoms in the fountain clock of PTB with the apogee of the atomic trajectories about 5 cm above the resonator.

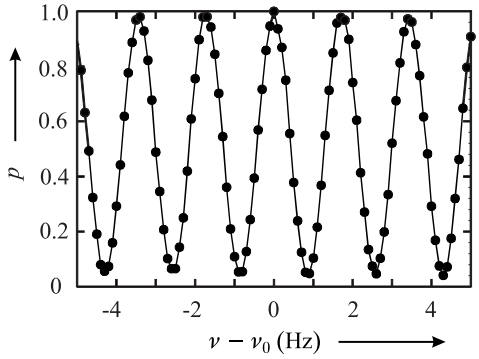


Figure 7.12: Central part of a high-resolution Ramsey spectrum similar to the one shown in Fig. 7.11 with the apogee of the atomic trajectories being 0.4 m above the resonator.

have not experienced an exact π pulse as a result of inhomogeneities in the field, or due to their different velocities, are pushed away by radiation from a short pulse of the downward laser beam.

7.3.2 Uncertainty of Measurements Using Fountain Clocks

Most of the effects that limit the uncertainty of the measurements performed with fountain clocks, and the methods, to explore and correct for them, are similar to the ones described with Cs beam clocks, as discussed in Section 7.1.3. However, most of these contributions are smaller in a fountain as in the clocks using thermal beams. The much reduced velocity v of the atoms in the former one reduces all contributions that scale with v or v^2. Moreover, the use of a single resonator for interrogation, in general, leads to a smaller influence of the cavity phase shifts. Furthermore, the special arrangement of the fountain allows for a convenient method to map the magnetic field in the region of free ballistic flight. Variation in the height of the apogee by changing the initial vertical velocity of the atomic cloud allows the atoms to spend most of their time near this region. The magnetic field is then easily determined as a function of the height from the shifted transition lines affected by the linear Zeeman effect.

Fractional uncertainties of primary Cs fountain clocks have been reported around 1×10^{-15} from several institutions [18–20] where one of the largest contributions results from the collisional shift.

7.3.2.1 Cold Collision Shifts

With Cs atoms, cold collisions have been found to be one of the biggest sources of systematic shifts since they can produce a frequency shift of $\Delta\nu/\nu = -1.7 \times 10^{-12}$ at an atomic density of $10^9\,\mathrm{cm}^{-3}$ [291, 292]. At low temperatures of a few millikelvin, the collision process between the Cs atoms is essentially described by s-wave scattering. This can be seen by calculating the maximum angular momentum $L_{\max} = b_{\max} \times m_r v$ where $b_{\max}$ is the

maximal impact parameter, i.e., the distance of closest approach of two caesium atoms for a straight-line trajectory. The reduced mass of the temporary Cs_2 molecule is $m_r = m_{Cs}/2$ and $v \approx 1$ cm/s is the atomic velocity at a temperature of 1 µK. Assuming that the maximal impact parameter b_{max} is within the range of the molecular potential, i.e., $b_{max} = 30$ nm, the corresponding maximal angular momentum is $L_{max} = \sqrt{l(l+1)}\hbar \approx 0.3\hbar$. Since $l < 1$, i.e., $l =$ s, only s-wave scattering is possible. At the same time the de Broglie wavelength $\lambda_{dB} = h/(m_r v) \approx 150$ nm is larger than b_{max} which again indicates that there is no p-wave scattering. Hence at low temperature, collisions can be described by a single parameter, i.e., the s-wave scattering length. The large scattering cross-section of Cs at low temperature has two causes. First, for cold atoms the collisional cross-section $\lambda_{dB}^2/(2\pi)$ becomes huge as a result of the increased de Broglie wavelength λ_{dB} in the microkelvin temperature range. Second, in Cs there are molecular bound states of the Cs_2 molecule very close to the dissociation energy, leading to resonances in the scattering amplitude. Hence, in Cs fountains used for atomic clocks the density, in general, is kept at a lower level and the shift is measured for different densities and an extrapolation to zero density has to be performed. If one refers to measured frequency shift coefficients [292] an accurate determination of the absolute density has to be performed. The resulting correction is one of the largest contributions to the uncertainty of the Cs fountain clocks. As a consequence of the large scattering lengths in Cs the cold-collision shift is expected to be strongly temperature dependent. A reversal of the cold-collision shift in Cs has been predicted to occur at very low temperatures below 100 nK [293]. It remains to be shown whether this dependence can be used to improve the accuracy of Cs fountain clocks.

The measurement of the collisional frequency shift is usually performed by alternating sequences of frequency measurements with a high and low atomic density, performed with respect to a reference frequency of sufficient short-term stability, e.g., a hydrogen maser (Section 8.1). In a fountain, the frequency shift is proportional to the "effective" density, which takes into account the density change during the ballistic flight and which depends on the initial spatial and velocity distribution [286]. The density for the two conditions is often changed by loading different numbers of atoms in the MOT. This is readily achieved by varying the loading parameters such as the power of the trapping laser beams or the time the beams are switched on.

Another method keeps these parameters constant but changes the microwave power in the selection cavity used to prepare the atoms in the $|F = 3, m_F = 0\rangle$ state. In general, a π pulse transfers all atoms from the $|F = 4, m_F = 0\rangle$ to the $|F = 3, m_F = 0\rangle$ state and a laser beam tuned to the $|F = 4\rangle \rightarrow |F = 5'\rangle$ transition is subsequently applied to remove the remaining atoms. Using a $\pi/2$ pulse rather than a π pulse leaves one-half of the atoms in the lower ground state thereby reducing the number of atoms by one half.

An elegant method that allows one to vary the number of atoms by an exact factor of two, but does not depend, e.g., on inhomogeneities in the microwave field in the selection cavity, has been devised by Pereira Dos Santos *et al.* [294]. Adiabatic fast passage allows one to transfer all atoms from a state $|g\rangle$ to a state $|e\rangle$ by continuously tuning the frequency of the interacting field from very large negative detunings to very far positive detunings. The pseudo-spin representing the ensemble in the rotating frame then spirals from the south pole of the Bloch sphere toward the north pole in a similar way to that depicted in Fig. 5.10. If the rapid

tuning of the frequency of the driving field is stopped exactly at resonance, the pseudo-spin vector of the ensemble is in the equatorial plane and the population is evenly distributed on the two ground states. Provided that no other losses occur during this process, the number of atoms, and hence the density, can be reduced to one-half of the former density. With the two latter methods it is not necessary to measure the absolute density for the determination of the collisional shift and to rely on the knowledge of the exact value of the coefficient of the collisional shift [291, 292].

The influence of the density-dependent shift is regularly measured in fountains used as frequency standards or clocks [19,20,289]. Fountain clocks are usually operated using atomic densities such that the associated uncertainty contributes not more than 50 % to the total uncertainty which typically limits the number below 10^6 atoms per bunch.

7.3.3 Stability

In a fountain clock the Allan deviation can be expressed as [64]

$$\sigma_y(\tau) = \frac{1}{\pi Q_{\mathrm{at}}} \sqrt{\frac{T_c}{\tau}} \sqrt{\left(\frac{1}{N_{\mathrm{at}}} + \frac{1}{N_{\mathrm{at}} n_{\mathrm{ph}}} + \frac{2\sigma_{\delta N}^2}{N_{\mathrm{at}}^2} + \gamma \right)} \tag{7.11}$$

which is a straightforward extension of (3.97). In (7.11) τ is the measurement time in seconds, T_c is the duration of a cycle of the measurement and $Q_{\mathrm{at}} = \nu_0/\Delta\nu$ is the line quality factor. The first term in (7.11) is due to the quantum projection noise [89] resulting from the quantum fluctuations of the population, after a measurement that projects the quantum mechanical superposition of the two states on either state. This term is not the only one in (7.11) that depends on the number of detected atoms N_{at}. The second term results from the photon shot noise due to the fluorescence detection of the atoms when n_{ph} photons are detected. As a large number of photons is scattered by every atom in this process, this term is in general much smaller than the first one. The third term results from the noise of the detection system where the number of atoms in the $|F = 4, m_F = 0\rangle$ and the $|F = 3, m_F = 0\rangle$ states are measured separately. $\sigma_{\delta N}$ represents the uncorrelated rms fluctuations of the atom number in each measurement. γ is the contribution of noise added from the interrogation oscillator. For example, non-linearities in the control loop or in the interrogation process can occur as the pulsed interrogation can mix down high-frequency noise of the oscillator into the base band. There it contributes to fluctuations of the signal via aliasing effects such as the Dick effect or intermodulation effects (Section 3.5.3). By using a low-noise cryogenic sapphire oscillator [18, 64] and by varying the atomic number N_{at} in the fountain between 10^5 and 6×10^5, the principal limit set by the quantum projection noise for $10^5 \leq N_{\mathrm{at}} \leq 5 \times 10^5$ atoms could be reached. An instability of $4 \times 10^{-14} (\tau/\mathrm{s})^{-1/2}$ could be obtained using 6×10^5 atoms [64].

7.3.4 Alternative Clocks

7.3.4.1 Rubidium Fountain

The collisional shift in ^{87}Rb has been found to be at least a factor of thirty smaller than in ^{133}Cs [286, 295]. Consequently, in comparison to a Cs fountain, an atomic fountain based on ^{87}Rb

atoms has the advantage of a good short-term stability since larger densities of atomic clouds can be employed without degrading the accuracy. Different ^{87}Rb fountains have been operated [91,295] and the ground state splitting of ^{87}Rb has been measured (see Table 5.1). In their Rb fountain, Sortais *et al.* [286] obtained a fractional instability of $\sigma_y(\tau) = 1.5 \times 10^{-13}(\tau/\mathrm{s})^{1/2}$. An instability of $\sigma_y(\tau) = 1 \times 10^{-14}(\tau/\mathrm{s})^{1/2}$ and a potential relative uncertainty in the 10^{-17} was envisaged [91].

7.3.4.2 Fountains with Increased Duty Cycle

The high accuracy achievable with caesium fountain clocks is hampered by their instability which allows one to make use of their accuracy only after a long measurement time of about a day. For a given line Q, the instability can be reduced by increasing the rate of detected atoms (see (7.11)) which is limited, on the other hand, by the pulsed operation of the fountain. To increase this rate, methods have been devised to multiply launching clouds of atoms or to employ quasi-continuous beams. The concepts of multiple launching or continuous beams in a fountain, require specific provisions to be made in order not to perturb the energy levels of the atoms, already in free flight, by the ac-Stark shift coming from stray light.

Juggling Fountain The successive launch of more than ten laser-cooled clouds of atoms simultaneously in the atomic fountain clock, has been proposed [296, 297] and demonstrated [287, 298]. Juggling many clouds at the same time has two advantages. First, for a given density corresponding to a given collisional frequency shift, the number of atoms detected and hence the signal-to-noise ratio can be increased. Second, juggling mainly eliminates the dead times between the measurements thereby imposing less stringent requirements for the stability of the local oscillator. However, shutters have to be used to block stray light from the interrogation region of the clock. Furthermore, the launching rate is limited by the fact that the frequency shift, due to collisions between the clouds, increases.

Fountain Using a Continuous Beam of Atoms A completely different concept that combines the advantages of a continuous beam and the fountain, has been jointly developed in the Observatoire de Neuchatel and the Swiss National Metrology Institute METAS [288,299–301] (Fig. 7.13). The continuous beam fountain has two advantageous properties. First, the atomic density at a given atomic flux is reduced by up to two orders of magnitude in a continuous beam as compared to a pulsed fountain. This property can be used to reduce the collisional shift by the same magnitude. On the other hand, for a given allowable shift, the increased number of atoms can reduce the instability (see (7.11)). Second, the Dick effect, as a result of an intermittent interrogation scheme, is expected to be much reduced, also promising higher short-term stability [52].

Due to the peculiarity of a fountain based on a continuous beam the powerful method which uses the same microwave field twice to interrogate the atoms has to be be given up. However, employing parabolic trajectories with a small separation of a few centimetres between the rising and the descending atoms at the interrogation region, a single coaxial cavity TE_{021} can be used [52] (Fig. 7.14).

The atoms pass through the coaxial cavity at opposite sides but at the same radial distance from the axis. At the positions of the atomic beam the magnetic field configuration of

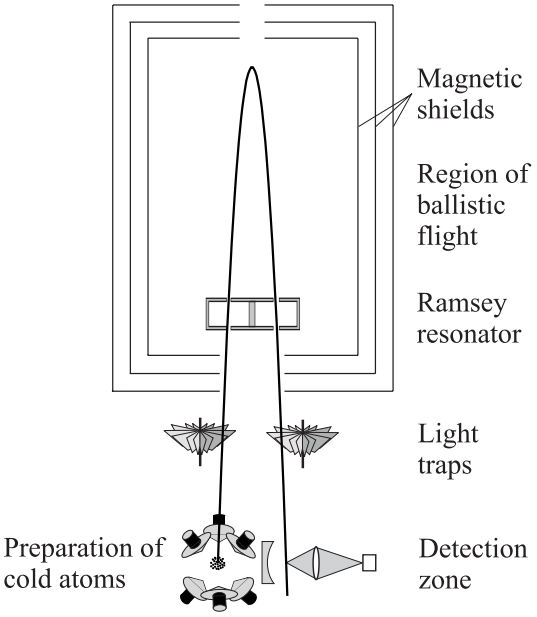


Figure 7.13: Schematics of a fountain with a continuous atomic beam.

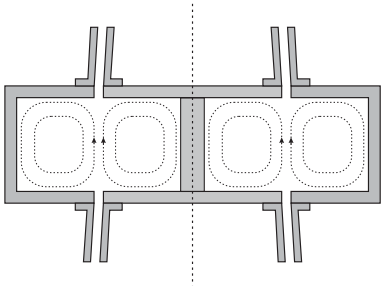


Figure 7.14: Vertical section of the TE_{021} cavity with schematic B field lines.

the TE_{021} cavity mode is similar to the field in the TE_{011} cavity used in typical fountains (Section 7.3.1) in a sense that there is a maximum of the B_z component with a sinusoidal dependence along the axis of the cylindrical resonator and no azimuthal dependence on the B field. It is expected that the relative frequency offset, related to the end-to-end phase shift, can be below 10^{-15} [288]. In order to allow one to investigate the phase shift in the microwave resonator, despite the fact that reversal of the atomic beam is not possible, the TE_{021} cavity (Fig. 7.14) can be rotated by 180° around the vertical axis. Thus the interaction of the caesium atoms with the interaction zones in time-reversed order also reverses the phase shift which can be determined and corrected for.

To prevent stray light, produced by the preparation of the cold beam, from perturbing the atoms which could lead to a relative frequency shift of up to 10^{-12}, light traps are necessary. Light traps based on turbine wheels with black blades mounted under 45°, rotating with an angular velocity matched to the velocity of the caesium beam, have been used. Such light traps have been shown to suppress the stray light to better than 10^{-5} [299] thereby attenuating the flux of atoms only by about 10 %. A short-term instability of $2.5 \times 10^{-13}\sqrt{\tau/\mathrm{s}}$ has been measured and a fractional uncertainty around 10^{-15} is anticipated [288].

7.4 Clocks in Microgravitation

The interaction time in a fountain is limited to about 1 s for the typical height of a laboratory clock. To go beyond is practically not possible as the height of the fountain increases quadratically with the desired interrogation time (see (7.7)). A longer interrogation time of up to 10 s corresponding to a linewidth as low as 50 mHz might be achieved in a microgravity environment such as is provided in a satellite orbiting Earth.

Several clocks are planned for use on board the International Space Station (ISS) orbiting Earth at a height of up to 450 km [302–306]. In the microgravity environment the use of a fountain is no longer possible or necessary and the current designs are based on a low-velocity beam.

A prototype of such a clock (PHARAO: Projet d'Horloge Atomique par Refroidissement d'Atomes en Orbite) already has been flown in a special aircraft to allow the simulation of microgravity in parabolic flights for about a minute [302]. The source is a sample of about 10^7 atoms laser cooled to about 2 µK in a six-beam optical molasses (Fig. 7.15). The six beams in

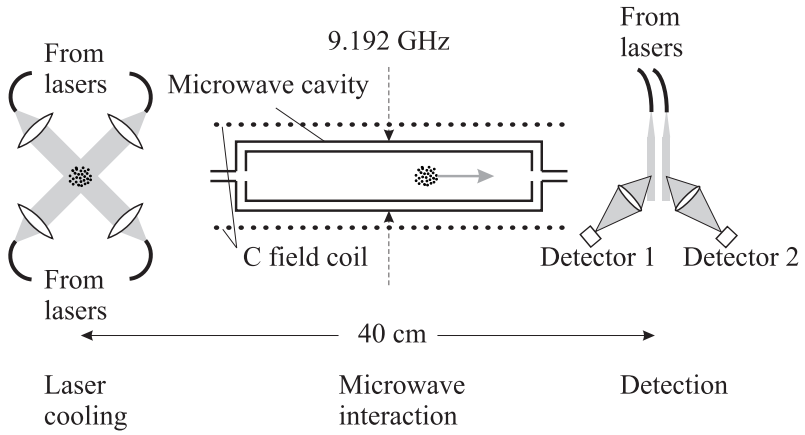


Figure 7.15: Schematic of a clock for use in space [302] with vacuum system, magnetic shields and lasers omitted for clarity.

the "1 1 1" arrangement are delivered to the vacuum apparatus by polarising optical fibres and the power of each beam can be adjusted individually. After leaving the cavity the atoms in the two atomic states $|F = 4\rangle$ and $|F = 3\rangle$ are detected separately using two separated detection laser beams in order to get a signal largely free of fluctuations in the atomic flux. In an earlier version [302] the microwave cavity was a cylindrical TE_{013} of length about 19.5 cm with three half-wavelengths of the standing wave fitting along the cylindrical axis. In the vicinity of the two nodes of the standing wave along the axis of this cavity, the residual spurious travelling waves led to phase inhomogeneities with a resulting large first-order Doppler shift. Hence, it was concluded [303] that cavities $TE_{01n}(n > 1)$ have to be given up to equip a space cold atom clock and, as an alternative, a ring resonator was used. In this resonator one coupling system feeds two symmetrical lateral waveguides which meet at the two interaction

zones [304]. The expected relative frequency instability and the fractional inaccuracy are $\sigma_y(\tau) = 10^{-13}(\tau/\mathrm{s})^{1/2}$ and 10^{-16}, respectively [304].

Another clock for the ISS is the PARCS (Primary Atomic Reference Clock in Space) which is designed to perform tests of relativity and fundamental physics and to serve as a primary clock [305, 307]. The set-up differs from that of Fig. 7.15 in a way that the microwave interrogation will take place in two independent cavities ($Q \approx 20\,000$) operated in the TE_{011} mode. The end-to-end cavity phase shift is expected to be measured with a fractional uncertainty of 2×10^{-17} by varying the atomic velocity and extrapolating to zero velocity [307]. Most of the contributions to the relative uncertainty are expected to be below 10^{-16} [305]. One of the biggest contributions to the uncertainty results from the frequency shift induced by the black-body radiation which, at the operational temperature of 37° C, is about 2×10^{-14} and cannot be corrected with a relative uncertainty in the 10^{-17} range. The PARCS clock will be compared to a superconducting microwave oscillator during flight.

Other clocks for space operation have been devised both for Cs or Rb [306]. An anticipated relative uncertainty of 10^{-16} or below requires a low instability of the microwave clocks in microgravity. With 10^6 detected atoms, a shot-noise limited signal-to-noise S/N = 10^3 can be achieved. Using an interrogation time of 10 s and a corresponding $\Delta\nu \approx 50$ mHz one calculates from (7.11) $\sigma_y(\tau = 10\,\mathrm{s}) \approx 5 \times 10^{-15}$ for Cs and $\sigma_y(\tau = 10\,\mathrm{s}) \approx 8 \times 10^{-15}$ for Rb allowing one to reach $\sigma_y(\tau) \approx 10^{-16}$ within a few hours.

8 Microwave Frequency Standards

In addition to the Cs atomic clock, which holds the special position of direct realisation of the unit of time and frequency, according to the definition, there are other microwave standards based on neutral atoms, which are used for different purposes. Amongst them are hydrogen masers with a short-and medium-term stability which surpasses the best caesium clocks, or the much cheaper and more compact clocks, such as the rubidium clock, used for applications with less demanding accuracy.

8.1 Masers

The first microwave amplification by stimulated emission of radiation (maser) was proposed in 1954 [308, 309] and the acronym maser is now used for all devices based on this type of process. Masers have been built using various atomic or molecular species in order to perform high-resolution microwave spectroscopic investigations. From the various kinds of masers developed for frequency standards and based on, e.g., ammonia, hydrogen, rubidium, or caesium, the hydrogen maser (H maser) has achieved the greatest widespread use and will be discussed in the following.

8.1.1 Principle of the Hydrogen Maser

The H maser utilises a transition between two ground state levels $|F = 1, m_F = 0\rangle$ and $|F = 0, m_F = 0\rangle$ of atomic hydrogen, with a frequency separation of 1.42 GHz (Fig. 5.22). Hydrogen masers used today do not differ much from the first realisation in the group of Norman Ramsey [310–312].

Hydrogen atoms from the source (Fig. 8.1) effuse in a vacuum of about 10^{-4} Pa maintained with an ion pump. Low-field seeking atoms are focussed into a storage bulb by means of a state selecting magnet which, at the same time, deflects the high-field seekers in the lowest states such that they do not enter the bulb. Hence, the storage cell, which is placed in a microwave resonator, contains more atoms in the upper state which can emit radiation by stimulated emission. The radiation is detected by an antenna and used to tune the frequency of a voltage-controlled crystal oscillator (VCXO) to the transition of the hydrogen atoms. The maser design allows the detection of the atomic transition with a small linewidth as a result of the long possible interaction time of about a second which is the time the atoms typically spend in the storage bulb. During this time the atoms are kept in a volume whose extension of about 15 cm is smaller than the wavelength of the transition. This confinement to the Lamb–

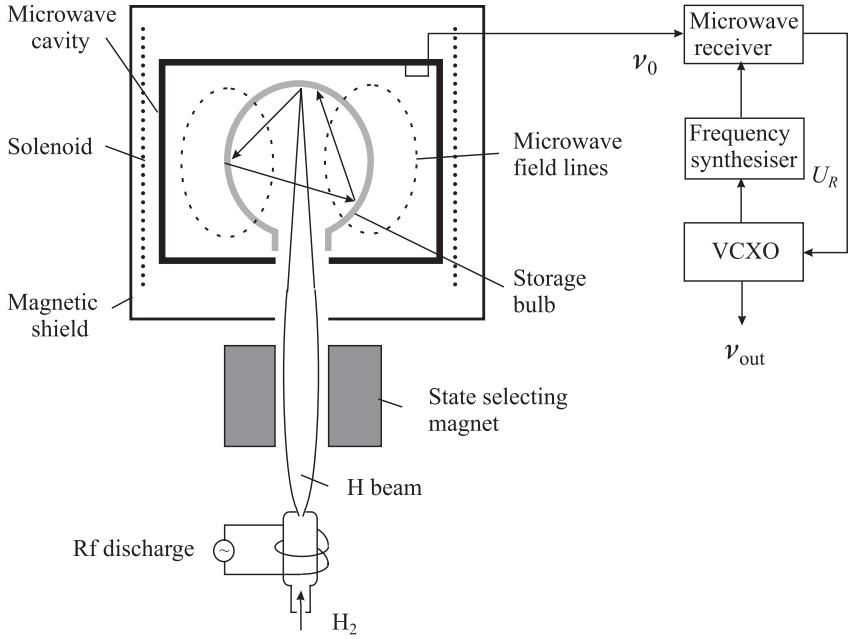


Figure 8.1: Schematic of an active hydrogen maser.

Dicke regime leads to a suppression of the first-order Doppler effect of the atoms interacting with the standing-wave field allowed in the microwave resonator.

In the here described "active H maser" conditions are such that continuous oscillation is sustained. However, the acronym maser is also used for a "passive H maser" (Section 8.1.4), where self-sustained oscillation is not achieved, but where the hydrogen atoms in the upper state resonantly amplify externally injected microwave radiation of the proper frequency.

8.1.2 Theoretical Description of the Hydrogen Maser

The theory of the hydrogen maser is presented in several publications [11,311–313] and hence here we only give a comprehensive sketch following Bender [314] and Vanier [313]. There is a substantial difference in the interaction of the hydrogen atoms in a H maser and the Cs atoms in a Cs clock. In the latter, discussed in Section 7, the atoms were prepared in a pure state before the first interaction. Moreover, the resolution was limited by the interaction time which was much smaller than the lifetime of the states involved in the transitions. Hence we described the Cs atom by a state vector and applied the Bloch vector picture to this state. This treatment is no longer appropriate in the case of the hydrogen maser where the collisions of the hydrogen atoms with the wall, and with each other, in general lead to a mixture of states. Hence one can only give statistical estimates about the probability of finding the ensemble of atoms in the excited state. The theoretical framework makes use of the density operator

(Section 5.3.2.1) which can be written in the equilibrium as the matrix

$$\rho = \begin{pmatrix} \rho_{11} & 0 & 0 & 0 \\ 0 & \rho_{22} & 0 & \rho_{42} \\ 0 & 0 & \rho_{33} & 0 \\ 0 & \rho_{24} & 0 & \rho_{44} \end{pmatrix} \tag{8.1}$$

where the four magnetic sub-states of the $S_{1/2}$ state are numbered according to their energy in the magnetic field (Fig. 5.22) beginning with $|F = 1, m_F = 1\rangle$ as "1" and ending with $|F = 0, m_F = 0\rangle$ as "4" [313]. The clock transition connects states "2" und "4" and consequently only the coherence between these states, given by the off-diagonal elements, is considered. The different elements of the density matrix are affected by four relevant processes occurring in the H maser. In the following we discuss these processes and the assumptions used to model them based on the rate equation

$$\frac{d\rho}{dt} = \left(\frac{d\rho}{dt}\right)_{\text{flow}} + \left(\frac{d\rho}{dt}\right)_{\text{wall}} + \left(\frac{d\rho}{dt}\right)_{\text{spin exchange}} + \left(\frac{d\rho}{dt}\right)_{\text{radiation}}. \tag{8.2}$$

First, there is a positive change in the number of atoms in the bulb, affecting ρ_{11} and ρ_{22}, as only atoms in the low-field seeking states are deflected into the bulb. At the same time there are losses by atoms escaping through the entrance hole of the bulb that depend on the bulb geometry and, thus, change all elements of the density matrix in the same way by $-\Gamma_b \rho_{ij}$, where Γ_b represents the relaxation rate due to the escape. Hence,

$$\left(\frac{d\rho}{dt}\right)_{\text{flow}} = \begin{pmatrix} I_1/N & 0 & 0 & 0 \\ 0 & I_2/N & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} - \Gamma_b \rho \tag{8.3}$$

where I_1 and I_2 is the flux of atoms entering the storage bulb in the states "1" and "2", respectively, and N is the total number of atoms.

Second, there are losses when atoms hit the wall. A fraction of these atoms are lost since they are absorbed on the surface or form a molecule with another atom thereby giving the binding energy to the wall. If these losses affect the atoms in all states in the same way, the rate of change is $d\rho_{ii}/dt = -\Gamma_w \rho_{ii}$ where Γ_w is the wall relaxation rate. Besides this effect, in less hard collisions with the wall, the coherence may also suffer from a phase change similar of to that described in Section 5.4.4. Hence $d\rho_{24}/dt = -\Gamma_w \rho_{24} + i\Omega_w \rho_{24}$ where it is often assumed that the coherence decays with the same rate Γ_w as the populations and where the imaginary term includes the phase change of the coherence resulting from the associated frequency shift Ω_w.

Third, there are collisions between two hydrogen atoms that can exchange their spins mutually thereby changing ρ_{22} and ρ_{44} by a rate Γ_{se}. Furthermore, the spin-exchange collisions also lead to a phase shift of the coherence. It has been shown both theoretically and experimentally by Berg, [315] that in the case of spin-exchange collisions between two hydrogen atoms, the relaxation time (dephasing time) of the coherence ρ_{24} is $\frac{1}{2}\Gamma_{\text{se}}$.

Fourth, besides these relaxation processes the dynamic change due to the interaction with the magnetic component of the electromagnetic field is obtained from the equation of motion

of ρ in the Schrödinger picture

$$\frac{d}{dt}\rho_{ij} = \frac{1}{i\hbar}\sum_k [\mathcal{H}_{i\,k}\rho_{kj} - \rho_{ik}\mathcal{H}_{k\,j}] \tag{8.4}$$

which is the quantum mechanical equivalent of Liouville's equation. The calculation can be performed using the Hamiltonian

$$\mathcal{H}_{24} = -\tfrac{1}{2}\mu_B g_J B_z(r)\cos(\omega t + \phi) \tag{8.5}$$

describing the magnetic interaction (5.34) for the hyperfine transition between two states with $m_F = 0$ and where μ_B is the Bohr magneton and $g_J = 2$. B_z is the only component of the magnetic field that is capable to induce transitions between the $|F = 1, m_F = 0\rangle$ and $|F = 0, m_F = 0\rangle$ states since it is parallel to the static quantisation field provided by the solenoid.

In equilibrium, the total rate of change for all four contributions is $d\rho_{\rm tot}/dt = 0$ which allows one to calculate the response of the atomic ensemble in the microwave field including the relaxation processes. In the following we give some results relevant to the practical use of the maser.

Let ΔI denote the difference in the flux of atoms entering the storage bulb in the $|F = 1, m_F = 0\rangle$ and $|F = 0, m_F = 0\rangle$ states. Then the averaged power radiated by this beam is $P = \Delta I\, h\nu\, p$ where p is the averaged probability of finding the atoms in the excited state "2". Neglecting for the time being any losses of the hydrogen atoms in state "2" by the processes discussed above, we could simply use (5.51) and (5.45) to derive $p = |c_2|^2$ and

$$P_{\text{without losses}} = \frac{1}{2}\Delta I h\nu \frac{b^2}{b^2 + [2\pi(\nu - \nu_0)]^2}. \tag{8.6}$$

In (8.6) the Rabi frequency b is defined as

$$b = \mu_B \frac{\langle B_z\rangle_b}{\hbar} \tag{8.7}$$

of the atoms induced by $\langle B_z\rangle_b$, i.e., the B_z component of the rf field, averaged over the bulb. Here, and in the following, the indices b and c represent bulb and cavity, respectively. According to (8.6) the power at resonance would linearly increase with the inversion ΔI independent of the Rabi frequency and hence of the power of the microwave field. It turns out that this behaviour is not observed in the H maser.

To correctly describe the hydrogen maser, the relaxation of the atoms in the upper state has to be taken into account. Using the processes discussed above, one can define decay times T_1 and T_2 given by

$$\frac{1}{T_1} \equiv \Gamma_b + \Gamma_w + \Gamma_{\rm se} \qquad \text{and} \tag{8.8}$$

$$\frac{1}{T_2} \equiv \Gamma_b + \Gamma_w + \Gamma_{\rm se}/2 \tag{8.9}$$

that are related to the total decay of population and the total loss of coherence, respectively. By including the decay times T_1 and T_2, Bender [314] and Kleppner *et al.* [312] obtained for the power

$$P = \frac{1}{2}\Delta I h\nu \frac{b^2}{\dfrac{1}{T_1 T_2} + b^2 + \left(\dfrac{T_2}{T_1}\right)\left[2\pi(\nu - \nu_0)\right]^2} \tag{8.10}$$

instead of (8.6).

For a large amplitude of the rf field, the b^2 term in the denominator becomes dominant and the power saturates with the on-resonance saturation parameter $S_0 = T_1 T_2 b^2$. The width (FWHM) of the resonance curve (8.10) is calculated as

$$\Delta\nu = \frac{1}{\pi}\sqrt{\frac{1}{T_2^2} + \left(\frac{T_1}{T_2}\right) b^2}. \tag{8.11}$$

A necessary condition for self-sustained oscillation of the maser is that the power delivered by the beam is equal to the power dissipated in the microwave resonator dW/dt. We recall that the quality factor of the rf cavity is $Q_c = \omega_c W/(-dW/dT)$ (see (2.39)) and that the energy W stored in the magnetic field in the cavity is

$$W = \frac{1}{2\mu_0}\int_{V_c} B^2 dV \equiv \frac{V_c}{2\mu_0}\langle B^2\rangle_c \tag{8.12}$$

where $\langle B^2\rangle_c$ is the mean squared amplitude of the magnetic field amplitude over the volume V_c of the cavity. We thus calculate the power dissipated by the cavity as

$$\frac{dW}{dt} = \frac{\omega_c V_c \langle B^2\rangle_c}{2\mu_0 Q_c}. \tag{8.13}$$

Defining a "filling factor"

$$\eta \equiv \frac{\langle B\rangle_b^2}{\langle B^2\rangle_c}, \tag{8.14}$$

the power dissipated by the cavity field (see (8.13)) can be expressed by a corresponding Rabi frequency b (see (8.7)) as follows

$$\frac{dW}{dt} = \frac{\omega_c V_c}{2\mu_0 Q_c}\frac{\hbar^2}{\eta\mu_B^2} b^2. \tag{8.15}$$

Following [312] we determine the power in the maser as function of the atomic flux ΔI. Taking (8.10) at resonance ($\nu = \nu_0$) one obtains

$$\begin{aligned} P &= \frac{1}{2}\Delta I h\nu - \frac{1}{T_1 T_2}\frac{P}{b^2} \qquad \text{which leads to} \\ \frac{P}{P_c} &= \frac{\Delta I h\nu}{2P_c} - \left(1 + \frac{3\Gamma_{\rm se}}{2(\Gamma_w + \Gamma_b)} + \frac{\Gamma_{\rm se}^2}{2(\Gamma_w + \Gamma_b)^2}\right) \end{aligned} \tag{8.16}$$

after inserting (8.8), (8.9), and (8.15) and defining P_c as

$$P_c \equiv (\Gamma_w + \Gamma_b)^2 \frac{\omega_c V_c}{2\mu_0 Q_c} \frac{\hbar^2}{\eta \mu_B^2}. \tag{8.17}$$

Defining a threshold flux

$$I_{\rm thr} \equiv \frac{2P_c}{h\nu}, \tag{8.18}$$

i.e., the minimum flux that is required to sustain oscillation in the cavity, then above threshold one would expect the power to increase linearly with the flux of hydrogen atoms if spin-exchange collisions can be neglected in (8.16).

However, the probability for spin-exchange collisions to occur is proportional to the density n of hydrogen atoms, which in turn depends on the total flux $I_{\rm tot}$ of hydrogen atoms as follows [312]

$$\Gamma_{\rm se} = n\sigma\bar{v}_r \qquad \text{with} \qquad n = \frac{I_{\rm tot}}{V_b \Gamma_b}. \tag{8.19}$$

In (8.19) σ is the cross-section for spin-exchange collisions and $\bar{v}_r$ is the average relative velocity of the hydrogen atoms. The density $n = N/V_b$, with V_b the volume of the bulb, is calculated from $dN/dt = I_{\rm tot} - N\Gamma_b = 0$.

By inserting (8.19) and (8.18) into (8.16) one obtains

$$\frac{P}{P_c} = \frac{\Delta I}{I_{\rm thr}} - \left[1 + 3q\frac{\Delta I}{I_{\rm thr}} + 2q^2\left(\frac{\Delta I}{I_{\rm thr}}\right)^2\right] \tag{8.20}$$

with

$$q = \frac{\sigma\bar{v}_r\hbar}{2\mu_B^2\mu_0}\frac{V_c}{V_b}\frac{1}{Q\eta}\frac{I_{\rm tot}}{\Delta I}\frac{\Gamma_w + \Gamma_b}{\Gamma_b}. \tag{8.21}$$

Fig. 8.2 displays the normalised power of the maser calculated according to (8.20) as a function of $\Delta I/I_{\rm thr}$ for different parameters $q \geq 0$. Hence, self-sustained oscillation ($P/P_c > 0$) of the H maser occurs only if the flux of hydrogen atoms is limited to a range between a minimal flux $\Delta I_{\rm min}$ and a maximal flux $\Delta I_{\rm max}$. These limits can be derived by solving the quadratic equation (8.20) for $P/P_c = 0$ leading to

$$\frac{\Delta I_{\rm max}}{I_{\rm thr}} = \frac{1 - 3q + \sqrt{1 - 6q + q^2}}{4q^2} \qquad \text{and} \tag{8.22}$$

$$\frac{\Delta I_{\rm min}}{I_{\rm thr}} = \frac{1 - 3q - \sqrt{1 - 6q + q^2}}{4q^2}. \tag{8.23}$$

The maser cannot oscillate for $q > 3 - \sqrt{8} \approx 0.171$ as can be derived by equating (8.22) and (8.23), i.e., $\Delta I_{\rm max} = \Delta I_{\rm min}$. Typical values of the power delivered by a hydrogen maser are $P \approx 1\,\text{pW}$. The maser quality parameter q in general is less than 0.1. The ratio $\Delta I/I_{\rm tot}$ is a measure of how effectively the state-selective magnet works.

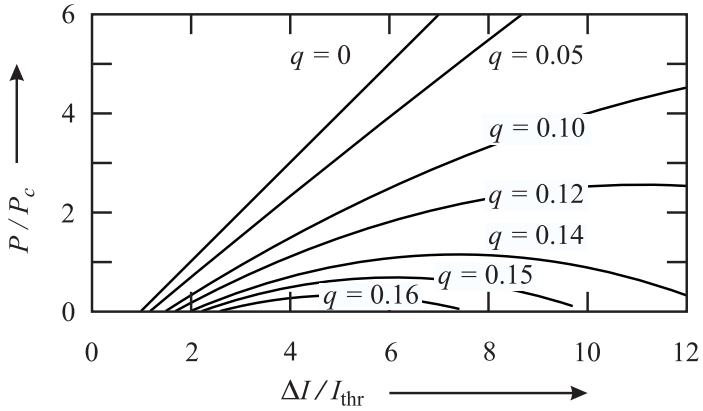


Figure 8.2: Normalised power of the H maser versus normalised flux of hydrogen atoms for different values of the parameter q (see (8.20) and (8.21)).

8.1.3 Design of the Hydrogen Maser

8.1.3.1 Hydrogen Source

In the set-up of Fig. 8.1 the atomic hydrogen (H) is produced in an intense electrical discharge from the chemically more stable molecular hydrogen gas (H_2) which is supplied by a reservoir bottle or from a heated metal hydride, e.g., $LaNiH_x$. The molecules from either source pass a purifier stage and diffuse to the discharge source through a sealed tube with thin walls consisting often of a silver palladium alloy. This so-called palladium-silver "leak" acts as a filter whose temperature-dependent mass flow can be varied by means of a heater coil. Hence, the pressure in the discharge region can be adjusted rapidly within typically 10 Pa and 100 Pa. In the discharge with a characteristic red colour a high-frequency electric field of about 200 MHz leads to a dissociation of the molecular hydrogen via molecular impact ionisation in the electric field.

8.1.3.2 State Selecting Magnet

In the beam of hydrogen atoms, all of the four states (Fig. 5.22) are populated with almost the same probabilities p as can be calculated from the Boltzmann distribution as $p(F = 1)/p(F = 0) = \exp[-h \times 1.42\,\text{GHz}/(k_B T)] \approx 0.99976$. The, almost equal, thermal population of both states requires the selection of the atoms in the high-energy states which is often performed in an inhomogeneous magnetic field.

Assuming that the magnitude of the magnetic moment of the moving particle (see (7.3)) is constant ($\vec{\mu}_{\text{eff}} = \mu$) and that the direction of the magnetic field, as viewed from the frame of the moving particle, changes slowly compared to the Larmor frequency, i.e. the precession of the magnetic moment in the magnetic field, then the energy of the system is proportional to the modulus of the magnetic induction $|\vec{B}|$ and does not change with the direction of $\vec{B}$.

Different inhomogeneous magnetic field configurations can be used for state selection, however, higher-order magnets like hexapoles also focus the beam and lead to a higher usable flux of the desired atoms.

Consider a hexapole field whose magnetic induction has the components

$$B_x = \frac{D}{2\mu}(x^2 - y^2), \qquad B_y = -\frac{D}{\mu}xy, \qquad B_z = 0 \tag{8.24}$$

where D is a constant. Such a field (Fig. 8.3) can be generated by six magnets facing each other under 60° with alternating north and south poles or by six wires arranged on a regular hexagon and alternatively carrying the same current in different directions [316–318].

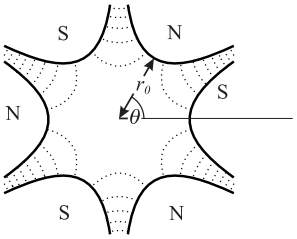


Figure 8.3: Hexapole magnet with alternating north (N) and south (S) poles. Dotted curves sketch the magnetic induction of (8.24).

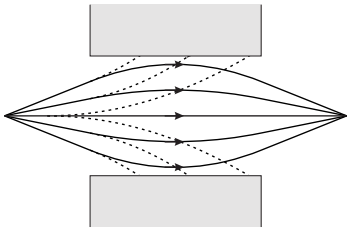


Figure 8.4: Trajectories of low-field seeking (full lines) and high-field seeking (dashed lines) paramagnetic atoms in a hexapole focussing magnet.

The modulus of the field

$$\sqrt{B_x^2 + B_y^2} = \sqrt{\frac{D^2}{4\mu^2}(x^2 - y^2)^2 + \frac{D^2}{4\mu^2}4x^2y^2} = \frac{D}{2\mu}(x^2 + y^2) \tag{8.25}$$

varies quadratically with the radius $r = \sqrt{x^2 + y^2}$.

In a real magnetic hexapole lens the magnetic field depends on the details of the realisation, e.g., the spacings of the gaps between the poles and the dimensions of the poles. The magnetic field (Fig. 8.3) can be calculated from Laplace's equation for the magnetic potential for a given set of boundary conditions. For a chosen geometry where the pole tips were arranged on a cylinder of radius r_0 with equal pole and air gap spacings (Fig. 8.3) the magnetic induction was calculated [132, 319] (or equivalently for the six current-carrying wires [317]) as follows

$$B(r,\theta) = B_0 \left(\frac{r}{r_0}\right)^2 \sqrt{1 - 2\left(\frac{r}{r_0}\right)^6 \cos 6\theta + \mathcal{O}\left(\frac{r}{r_0}\right)^{12}}. \tag{8.26}$$

The leading term of (8.26) is again quadratic in r. Hence from (8.25) and (8.26), a particle with magnetic moment μ and mass m experiences a force (see (7.3)) $\vec{F}_{\mathrm{mag}} = -\vec{\nabla}\mu|B| = -D\vec{r}$ that increases linearly with the distance from the origin $x = y = 0$. Consequently, an atom with a negative magnetic moment and a sufficiently small excursion and radial velocity, would undergo an oscillation with angular frequency

$$\omega = \sqrt{2\mu B_0/(mr_0^2)}. \tag{8.27}$$

As for any harmonic oscillation the angular frequency (8.27) does not depend on the transverse velocity v_t. Hence, all atoms starting with the longitudinal velocity v_l from a source point on the axis of the hexapole will cross the axis again after performing half an oscillatory cycle corresponding to the time $T = \pi/\omega$. The distance from the source point $L = v_l T$ depends on the longitudinal velocity and, hence, all atoms of the same kind and with the same longitudinal velocity are focussed into the same point (Fig. 8.4).

The hexapole magnet acting as a magnetic lens is often used in the hydrogen maser to select the low-field seeking atoms by focussing them into the entrance of the storage bulb. Since the focal length depends on the velocity of the atoms, the magnetic lens shows a "chromatic aberration" and the focussing properties for a thermal ensemble of atoms are limited. If permanent magnets are used, B_0 is limited to less than about 1 T and from the, typically, 10^{16} atoms/s of a collimated source, a fraction of, typically 5×10^{-4} atoms in the upper state can be focussed. The use of rare-earth magnetic materials allows one to design iron-free magnets where the field pattern can be predicted on the 1 % level if the geometry of the individual segments forming the multipole is known [320].

8.1.3.3 Storage Bulb

The storage bulb containing the atoms has to fulfil several requirements. As it is located inside the microwave cavity with a desired high Q_c it has to be made of a dielectric material with low dielectric loss. It is often made of a fused silica sphere with a typical diameter of 15 cm with a collimator at the inlet of the atomic beam. The wall is usually coated with a material that minimises the wall shift and what is even more important, that produces a temporally stable wall shift. Typically, the wall is coated with a fluorocarbon polymer such as Teflon using special procedures [311, 321]. The wall shift depends on the particular type of fluorocarbon [90].

8.1.3.4 Microwave Resonator

In the hydrogen maser, the TE_{011} mode of a right cylindrical microwave resonator is used (Section 4.2.3) where the magnetic field in the centre points along the axis of the cylinder (Fig. 4.9). Often, a cavity is employed where the length and diameter of the cavity are the same, in order to minimise the losses in the walls of the cavity and hence to obtain a large quality factor Q_c. For the resonance frequency of $\nu_{011}^{(\mathrm{TE})} = 1.42$ GHz, the length is calculated from (4.68) as $L \lesssim 27$ cm. The spherical quartz bulb is located symmetrically in the cavity with respect to the cylinder axis. Its diameter along the radial direction ($D \lesssim 15$ cm) has to be small enough to confine the atoms in a region where the axial B_z field does not reverse its direction (see Fig. 4.9 b). The quartz bulb does not much affect the Q_c factor of the microwave resonator, but for the given frequency the resonant length is reduced by about 5 cm for a quartz bulb with 1 mm thick walls due to the increased dielectric constant $\epsilon > 1$. The body of the resonator has to be very rigid in order to achieve ultimate stability of the length and diameter to avoid frequency shifts due to cavity pulling, i.e., an offset of the maser frequency from the frequency of the atomic transition resulting from a detuning of the eigenfrequency of the resonator. In order to avoid furthermore fluctuations of the eigenfrequency of the resonator, resulting from temperature-induced variations of the size, the resonator is kept in a

temperature-stabilised vacuum chamber. The cavity is often made from a material with low thermal expansion coefficient such as, e.g., a glass ceramic (see Section 4.4) whose surface is internally coated with silver thus forming resonator walls with a low ohmic resistance.

To obtain a larger filling factor η (see (8.14)), sometimes the length of the resonator is chosen to be larger than its diameter by about a factor of 1.5 [322,323] with a similar form for the bulb.

The size of the resonator, together with the necessary vacuum housing and magnetic shieldings, determines the dimensions and the weight of the hydrogen maser which in stationary devices can be up to 0.5 m^3 and more than 200 kg. In order to reduce the size of the cavity, sometimes a dielectric material such as a quartz tube used to hold the bulb [322] or a sapphire ring, is introduced which allows one to reduce the volume of the cavity as a result of the high dielectric constant of these materials.

A hair-pin loop antenna, located at a position near the end plates of the cavity where the magnetic field is strong enough, is used to couple out the power delivered by the hydrogen atoms to a 50 Ω coaxial cable leading to an amplifier. Often a second loop antenna is used in combination with a varactor to fine-tune the cavity.

8.1.3.5 Frequency Shifts

The actual frequency of a hydrogen maser in general differs from the frequency of the ground-state splitting of the unperturbed hydrogen atom as a result of several effects. In the following these effects are addressed in the order of magnitude in which they contribute to the frequency shift in a hydrogen maser.

Second-order Doppler Effect Due to their small mass, hydrogen atoms in thermal equilibrium near room temperature have a rather large average velocity and hence suffer from a considerable second-order Doppler shift. Equating the mean thermal energy $3k_BT/2$ to the kinetic energy $mv^2/2$ of a hydrogen atom with mass m and velocity v, the second-order Doppler shift or time dilation shift (5.110) is

$$\frac{\Delta\nu_{\text{time dil.}}}{\nu} = -\frac{3k_BT}{2mc^2}. \tag{8.28}$$

Hence, the transition frequency varies linearly with the temperature. At a typical operational temperature for masers of $t = 40°\,\mathrm{C}$ ($T = 313\,\mathrm{K}$) the frequency is shifted by -4.3×10^{-11}. Keeping the temperature of the bulb constant to 0.1 K leads to an associated fractional uncertainty of the corresponding frequency of 1.4×10^{-14}. Consequently, the hydrogen masers in general have a single- or double-stage temperature stabilisation set-up.

Wall Shift Hydrogen atoms, interacting with the radiation field, and colliding with the walls of the storage bulb, suffer from a mean phase shift of the radiating atom per collision which depends on the coating of the walls and the temperature of the atoms. For a given interaction time the number of collisions is inversely proportional to the diameter D of the bulb. Thus the offset between the frequency of the maser and the frequency of the unperturbed hydrogen atoms referred to as the "wall shift" is sometimes [90,266] expressed as

$$\Delta\nu_w = \frac{K(t)}{D} \tag{8.29}$$

where $K(t)$ is a temperature-dependent constant. At a temperature of $t = 40°$ C the constant has been determined for different Teflon coatings [90] by using bulbs of different size and extrapolating the measured frequency offsets as function of $1/D$ to infinite diameter D. The value of $K(t = 40°\text{C}) \approx -0.5\,\text{Hz}\,\text{cm}$ (for Teflon FEP 120) leads to a fractional shift of 2.3×10^{-11} for a spherical bulb of $D = 15\,\text{cm}$. The value of this shift is estimated to be accurate to 10% [266] which is relevant only if the frequency of the unperturbed ground state hyperfine splitting (ν_H = 1 420 405 751.768(2) Hz [90, 324], see also Table 5.1) is determined. In most other cases it is more important that the wall shift is stable in time. Around 40 °C, $K(t)$ varies by about -10^{-2}/K [90] which requires that the temperature of the bulb is kept stable to at least 0.1 K to restrict the fluctuations to the lower 10^{-14} regime.

Magnetic Field The $|F = 1, m_F = 0\rangle \rightarrow |F = 0, m_F = 0\rangle$ clock transition does not show a linear Zeeman effect and for weak magnetic fields the transition frequency varies quadratically with the magnetic field as

$$\Delta\nu_{B^2} \approx 2.7730 \times 10^{-1}\,\text{Hz} \left(\frac{B}{\mu\text{T}}\right)^2, \tag{8.30}$$

which is more than six times higher than in the Cs clock (see (7.2)). For a typical value of the magnetic field of $0.1\,\mu\text{T}$ the fractional frequency shift is $\Delta\nu_{B^2}/\nu \approx 2 \times 10^{-12}$. It can be determined by measuring the linear Zeeman shift of the $|F = 1, m_F = \pm 1\rangle$ states. Frequency shifts due to a variation in the magnetic field inside the storage bulb are usually controlled by using four to five layers of soft magnetic material like permalloy. For higher requirements or harsher environments, active field compensation [325] is used, where leakage of the external magnetic field through the outer shielding layer is sensed by a magnetometer and largely reduced by using a compensating coil wound around the next inner layer. By this means the internal field fluctuation could be reduced by 2×10^6 when the external field was varied by $\pm 50\,\mu\text{T}$ [325].

Cavity Pulling If the resonance frequency ν_c of the microwave resonator is detuned from the frequency of the atomic transition ν_0, the maser frequency ν is shifted by [326]

$$\Delta\nu_c = \nu - \nu_0 = \frac{Q_c}{Q_{\text{at}}}\,(\nu_c - \nu_0) \tag{8.31}$$

where Q_c and Q_{at} are the quality factors of the cavity and the atomic resonance line, respectively. The shift can be understood in a simple way if one considers the hydrogen ensemble as an amplifying oscillator and the microwave cavity as a resonant filter in the feedback loop (Section 2.2). The radiation fed back from the detuned cavity suffers from a phase shift. According to the phase condition (see (2.55)) this phase shift is compensated by a slightly changed frequency of the oscillator. In contrast to the cavity-pulling shift in Cs clocks, which varies proportionally to $(Q_c/Q_{\text{at}})^2$ (see (7.6)), the linear ratio is as large as $Q_c/Q_{\text{at}} \approx 3.5 \times 10^{-5}$ for $Q_c \approx 5 \times 10^4$ and $Q_{\text{at}} \approx 1.4 \times 10^9$. In order to reduce the associated frequency shift, hydrogen masers in general are equipped with one or another method of "auto-tuning".

A first method keeps the eigenfrequency of the cavity constant by electronically stabilising it to a chosen frequency close to the resonance frequency. To this end the resonance

frequency of the cavity is square-wave modulated by means of a tuning varactor attached to the microwave resonator, driven with a modulation frequency of a few ten hertz [322] to a few hundred hertz. This frequency is much higher than the inverse lifetime of the radiating atoms in the cavity whose emission is not much affected by the modulation. Detuning of the microwave cavity changes the gain and, hence, the output power of the maser. If the frequency of the cavity is modulated symmetrically with respect to the eigenfrequency, the modulation of the output power vanishes and the centre frequency of the resonant cavity can be locked to this position. This method works only if the fluctuations of the frequency used to interrogate the cavity are smaller than the fluctuations or drift of the cavity frequency. As the modulation frequency is sufficiently fast the requirements for medium-term stability of the frequency reference can be met by an ultra-stable quartz oscillator used as a reference for for the synthesiser.

A second method of locking the eigenfrequency of the microwave resonator to a given value [327] involves directly injecting a frequency-modulated microwave signal into the cavity and then measuring the reflected signal to determine the offset of the eigenfrequency from the frequency of the signal serving as a reference. Depending on the obtained error signal the cavity can be tuned to the desired resonance frequency. In order not to perturb the atoms by a resonant injected microwave signal, a square-wave frequency-modulated signal is chosen where the modulation frequency is an even sub-multiple of the modulation depth, since in this case the carrier is totally suppressed. This method has the advantage that the error signal can have a better signal-to-noise ratio as compared to the previously described technique since higher power can be employed as delivered by the atoms themselves.

Spin-exchange Collisions Collisions between hydrogen atoms in the "clock states" $|F = 0, m_F = 0\rangle$ and $|F = 1, m_F = 0\rangle$, where the spin of both atoms is exchanged, lead directly to a frequency shift as a consequence of the exchange interaction. In addition to this direct shift, there is a broadening of the atomic resonance as a consequence of the associated relaxation. The additional broadening may lead to a frequency shift via the cavity pulling effect. This combination then becomes a major source for frequency shifts as it couples the oscillation frequency to parameters like the density of the hydrogen atoms that can hardly be controlled with the desired uncertainty. The spin-exchange frequency shift has been calculated by Koelman *et al.* [328] from zero temperature to 1000 K. However, measurements at room temperature [329] and cryogenic temperatures [330] revealed significant discrepancies with theory [331].

The combined effects of the spin-exchange shift and the cavity pulling are often used for auto-tuning by the method of "spin-exchange tuning". It is based on tuning the cavity such that the cavity pulling and the spin-exchange collision shift cancel each other to a high degree [332]. As described by (8.19) the spin relaxation rate and hence the quality factor Q_{at} of the atomic resonance can be easily changed by changing the flux of atoms. In general, the variation of Q_{at} affects the frequency pulling (see (8.31)). At the same time, variation of the atom number affects the number of spin-exchange collisions and, hence, the associated frequency shift. The working point where both effects cancel each other can be found by slowly switching the atomic flux, with a modulation time of a few minutes, between two different values $I_{\mathrm{tot},1}$ and $I_{\mathrm{tot},2}$, with the associated quality factors $Q_{\mathrm{at}}(I_{\mathrm{tot},1})$ and $Q_{\mathrm{at}}(I_{\mathrm{tot},2})$.

The measurement of the associated frequency shift requires a very stable reference frequency which, in cases where the highest precision is required, is often supplied by a second maser.

Both effects cannot be separated but the influence of the spin exchange collisions is believed to be reduced to the 10^{-13} level [266] by this method.

8.1.3.6 Frequency Stability of Active Hydrogen Masers

The frequency instability of a hydrogen maser represented by the Allan deviation shows distinct regimes for different averaging times τ (Fig. 8.5).

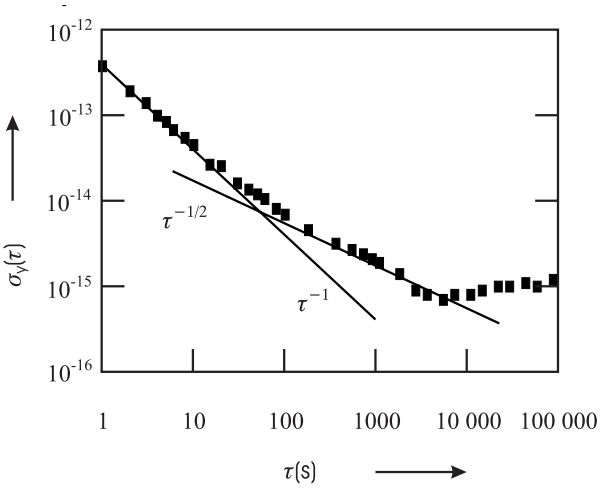


Figure 8.5: Combined instability (Allan deviation) of two commercial hydrogen masers at PTB measured in 1998. For the measurements shown, both masers were operated without auto-tuning.

The short-term instability can be represented by a $1/\tau$ dependency, whereas the medium-term instability shows a $1/\sqrt{\tau}$ behaviour. For longer times the instability increases again after a flicker floor.

The long-term instability of Fig. 8.5 is determined mainly by the cavity drift. This contribution to the instability can be improved if auto-tuning is employed leading also to a reduction of the flicker floor. The medium-term instability with the $1/\sqrt{\tau}$ decrease results from two principal processes ultimately limiting the frequency stability of any maser activity. The phase of the electromagnetic field in the cavity resulting from the stimulated emission is perturbed by the presence of the thermal radiation field with random phase contributing to the excited mode of the resonator. Hence, the power spectral density of fractional frequency fluctuations of the hydrogen maser [11, 266, 333]

$$S_y(f) = \frac{k_B T}{P Q_{\mathrm{at}}^2} + \frac{4 k_B T}{P} \frac{f^2}{\nu_0^2} \left(1 + F \frac{P}{P_r}\right) \tag{8.32}$$

comprises a first term with the corresponding white frequency noise. The associated frequency instability described by the Allan deviation is calculated from Table 3.1 as

$$\sigma_y(\tau) = \frac{1}{Q_{\mathrm{at}}}\sqrt{\frac{k_B T}{2P}}\frac{1}{\sqrt{\tau}}. \tag{8.33}$$

The stability due to this process is limited by the power P delivered by the atoms and the atomic Q_{at} factor, typically to $3 \times 10^{-14}\,\tau^{-1/2}$. A similar value is derived from Fig. 8.5 if the Allan deviation is evenly attributed to each maser using (3.16).

The short-term instability is limited by additional white phase noise added to the signal represented by the second term in (8.32). Such phase fluctuations can occur by length fluctuations of the microwave resonator and by phase fluctuations in the electronic circuit, e.g., the amplifier that depends on the amplifier noise factor F and the power P_r that is received by the amplifier. The corresponding contribution to the Allan deviation is calculated (see Table 3.1) as

$$\sigma_y(\tau) = \sqrt{\frac{3k_B T f_h}{\pi^2 \nu_0^2 P}\left(1 + F\frac{P}{P_r}\right)}\frac{1}{\tau}. \tag{8.34}$$

The factor f_h is the cut-off frequency defining the bandwidth of the equipment used to measure the frequency fluctuations (Section 3.3). As both noise contributions (8.33) and (8.34) are independent they can be added in the Allan deviation plot limiting the stability at different averaging times. The flicker floor of commercial masers (Fig. 8.5) corresponding to a frequency instability of below 10^{-15} is reached for periods from 1 000 to 10 000 seconds. For higher measuring times τ frequency drifts cause the Allan deviation to increase again. Drift rates of a few times 10^{-14} per year have been reported for five different masers kept in a very stable environment [334].

8.1.4 Passive Hydrogen Maser

Below threshold, where active self-sustained oscillation is not achieved, radiation whose frequency coincides with the frequency of the transition of the hydrogen atoms can still be amplified. In this regime the maser acts as an amplifier within a narrow frequency range set by the linewidth of the atomic transition. In such a "passive maser" an external generator provides electromagnetic radiation of 1.42 GHz which is coupled into the resonator. The resonance line is interrogated by monitoring the amplified signal and the frequency of the external oscillator is tuned to the maximum output signal. Passive hydrogen masers combine high short-term stability with smaller size and weight compared to active masers.

To achieve smaller size sometimes a so-called magnetron microwave resonator is employed [266] which contains an internal concentric capacitive-inductive structure of annular metal electrodes resembling the electrode structure in a magnetron. With a magnetron cavity a mass of a few ten kilograms can be achieved [335]. The contributions that limit the frequency stability of the passive maser are similar to that of the active maser. The medium-term instability also follows $\tau^{-1/2}$ but in contrast to (8.33) it is increased by a factor of about 10 which depends on the experimental conditions. A small-size passive maser has a fractional instability of $\sigma_y(\tau) \leq 10^{-12}/\sqrt{\tau/\mathrm{s}}$, up to about 10^4 seconds, which then levels to a flicker floor of a few times 10^{-15} [336, 337].

8.1.5 Cryogenic Masers

The exceptionally low fractional frequency instability of a maser, operated near room temperature, of below $\sigma(\tau) \cong 10^{-15}$ for about 10^4 seconds, was expected to be even reduced to the 10^{-18} [338] domain for cryogenic temperatures. Temperature affects the instability of the hydrogen maser via various effects. First, as can be seen from (8.19) the spin-exchange linewidth is proportional to the velocity of the hydrogen atoms which can be reduced by an order of magnitude or more by using hydrogen atoms at a temperature of a few Kelvin. Reduction of the linewidth leads to an increased line $Q_{\rm at}$ of the maser with the associated reduced instability (8.33). Second, the temperature enters (8.33) directly as temperature radiation induces incoherent emission of photons and hence leads to fluctuations in the phase of the radiation field in the maser. Third, at cryogenic temperatures the walls can be coated with a helium film which leads to a better controlled and a temporally more stable interaction with the hydrogen atoms than the Teflon coating. A temperature can be found where the wall shift is, to first order, independent of the temperature (Fig. 8.6). Fourth, as the maximal flux of atoms in the maser depends on the relaxation rates, a reduction of the spin-exchange rate at lower temperatures allows one to operate the maser with a larger number of atoms (see (8.20) and (8.21)) thereby increasing the usable output power and reducing the instability.

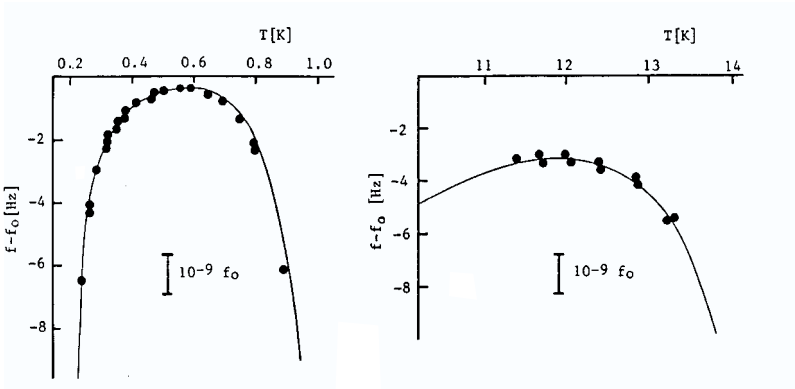


Figure 8.6: Frequency shifts resulting from collisions for liquid helium surfaces and solid neon surfaces. After [339] with permission.

Cryogenic masers have been realised [330, 338, 340–342] and it was expected that a fractional instability in the 10^{-18} range could be obtained [338]. However, later it turned out that serious discrepancies existed between measured and calculated hydrogen spin-exchange cross-sections [331, 343]. These difficulties, together with the necessary additional effort to realise cryogenic masers, have restricted their use mainly to basic research.

8.1.6 Applications

Masers are used in a variety of challenging applications like time transfer, navigation, tracking of space crafts or performing leading-edge experiments in basic research.

8.1.6.1 Hydrogen Masers for Time Scales

With respect to their excellent stability for periods from about 10 s to a day, masers are superior to Cs clocks. Timing laboratories usually have several masers as so-called "fly wheels" to increase the short-term stability of their time scale (Section 12.1.2). Sometimes an ensemble of masers augmented by commercial caesium clocks is used to provide a time scale [334] in order to have a convenient frequency reference for evaluating the stability of frequency standards. The long-term stability of such an ensemble at the National Institute for Standards and Technology (NIST, USA) was given as $\sigma_y(\tau = 0.1\,\text{day}) \approx 1 \times 10^{-15}$, $\sigma_y(\tau = 1\,\text{day}) \approx 4 \times 10^{-16}$, $\sigma_y(\tau = 10\,\text{days}) \approx 2.5 \times 10^{-16}$, $\sigma_y(\tau = 100\,\text{days}) \approx 8 \times 10^{-16}$ with a frequency drift of less than 3×10^{-15}/year.

8.1.6.2 Gravitation and Relativity Experiments

The influence of gravitational and acceleration potentials on clocks and frequency standards is described by Einstein's Theory of General Relativity (Section 12.2). Consequently, the comparison of remote clocks represents a suitable means of checking the predictions of theory in order to determine the validity of the basic ingredients and assumptions.

An early precision experiment, now referred to as "Gravity Probe A", [1] used to test Einstein's general theory of relativity, measured the frequency difference between a hydrogen maser onboard a rocket and masers on Earth [344]. To this end a hydrogen maser was launched in a nearly vertical trajectory in a two-hour sub-orbital flight by a Scout rocket to an altitude of about 10 000 km. During flight, microwave links allowed comparison of the frequencies of the masers which were subjected to different gravitational potentials, different time-dilation shifts (second-order Doppler shift; Section 5.4.2) and different first-order Doppler shift due to acceleration. When analysed, the observed frequency shift was consistent with Einstein's predicted gravitational red shift within a fractional uncertainty of 7×10^{-5} [344].

The clocks onboard the "space-ship Earth" also probe the variation of the gravitational potential $U(t)$ of the sun during the annual elliptical orbital motion of Earth around the sun. According to Einstein's equivalence principle, clocks on Earth experience an associated fractional frequency shift $\Delta U(t)/c^2$ of $\pm 3.3 \times 10^{-10}$. The so-called "principle of local position invariance" furthermore requires that this gravitational shift is independent of the atomic species used as reference in the atomic clock. Bauch and Weyers [345] have tested this basic assumption by comparing the frequencies of a hydrogen maser and a caesium atomic clock for about one year and did not find a related variation of the frequency ratio within a fractional uncertainty of 2.1×10^{-5}.

Measurements of the Zeeman frequency splitting in a hydrogen maser performed by Phillips *et al.* [346] and the unobserved sidereal variations were used to constrain possible Lorentz and CPT violations.[2] Nowadays the standard model of particles is believed to represent the low-energy limit of a more general theory that might include general relativity. Rea-

[1] In April 2004 "Gravity Probe B" a relativity gyroscope experiment was launched to check tiny changes in the direction of the spins of four gyroscopes, contained in an Earth satellite orbiting at 650 km altitude directly over the poles, resulting from space-time dragging effects by the rotating Earth.

[2] The CPT theorem states that physical laws are invariant under simultaneous application of Charge conjugation, Parity inversion, and Time reversal.

sonable, but only guessed, extensions of the standard model often lead to spontaneous Lorentz symmetry violation [347] whose bounds have been determined by these measurements in the proton.

8.1.6.3 Other Applications

Passive masers are employed in global navigation satellite systems (Section 12.5) such as the future European satellite system GALILEO with space-qualified versions for use onboard the satellites. There, the primary function of the masers is to provide clocks with high short-term stability. Long-term stability and accuracy is of less importance in this application since the clocks can be synchronised with clocks on Earth.

Active masers find applications in astronomy and geodesy, e.g., to correlate the signals from different antennas in Very Large Baseline Interferometry (VLBI) (Section 12.6.1). The maser at each antenna site provides the clock signal that is recorded with the radio signal from the telescope, so that the signals from each pair of telescopes in the VLBI network can be later properly correlated.

8.1.6.4 Cosmic Masers

Masers are also "operated" by nature. The first cosmic maser was discovered accidentally in 1965 [348] when radio astronomers measured the absorption of OH molecules against thermal background sources in order to map the distribution of gas in molecular clouds. In the meantime, a number of natural masers have been found, based on transitions in OH, H_2O, SiO, CH_3OH, or NH_3 molecules, to name only a few [349]. The evidence for the maser action results from the combination of spectral features, namely narrow linewidths, high polarisation and high brightness. The measurement of these properties can rule out a black-body source since the measured linewidths require low temperatures, whereas the brightness would correspond to a black-body radiator with a temperature as high as 10^{15} K. In contrast to man-made masers, cosmic masers are single-pass amplifiers and rely on the large path lengths in molecular clouds to provide the required gain. The spatial coherence is assumed to be very small.

Cosmic masers have been used to derive information about a large variety of astro-physical phenomena which are otherwise not accessible [349]. They have been used to probe the velocities of winds of matter flowing from newly formed stars or the envelopes of red giant stars. In some fortunate cases, the appearance of maser action in SiO, H_2O and OH has been observed from the same object in different shells with diameters of more than 10^{10} km, allowing one to derive useful information about the composition of stellar objects. Similarly, supernova remnants or nuclei of active galaxies have been investigated. Other investigations include the measurements of distances via the Doppler effect or of magnetic fields via the line splitting by the Zeeman effect.

8.2 Rubidium-cell Frequency Standards

In a rubidium-cell standard the frequency of an oscillator is locked to the 6.83 GHz transition between the hyperfine-split ground state of the isotope ^{87}Rb (Fig. 8.7). This isotope makes up about 28 % of the natural abundance of rubidium, with the remaining 72 % composed of the isotope ^{85}Rb which has a lower hyperfine splitting of 3.04 GHz (Fig. 8.7). The quantum

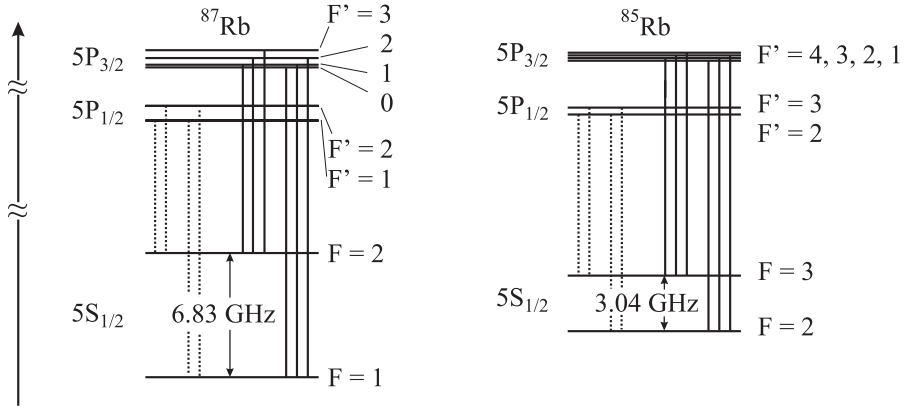


Figure 8.7: Hyperfine structure of the ground state and the first excited electronic states of ^{87}Rb and ^{85}Rb including the D1 transitions ($\lambda = 794.7$ nm; dotted lines) and D2 transitions ($\lambda = 780.0$ nm; full lines).

numbers of the nuclear spins of ^{87}Rb and ^{85}Rb are $I = 3/2$ and $I = 5/2$, respectively, and the angular momentum in the ground state is given for both isotopes by the spin of the single electron in the outer shell, i.e., $J = 1/2$. The coupling of both angular momenta leads to the two hyperfine states with the total angular momenta $F = 2$ and $F = 1$ for ^{87}Rb, and $F = 3$ and $F = 2$ for ^{85}Rb.

In ^{87}Rb the sub-states associated with the quantum numbers $F = 1, -1 \leq m_F \leq +1$ and $F = 2, -2 \leq m_F \leq +2$, split in a magnetic field (Fig. 8.8). Again, the transition between the states with only a weak quadratic dependence on the magnetic field

$$\Delta\nu_{B^2} \approx 5.74 \times 10^{-2}\,\mathrm{Hz} \left(\frac{B}{\mu\mathrm{T}}\right)^2 , \tag{8.35}$$

i.e., the $|F = 2, m_F = 0\rangle$ and $|F = 1, m_F = 0\rangle$ states, is chosen as the clock transition.

8.2.1 Principle and Set-up

Rubidium microwave standards are in current use as frequency references that are very compact, require low-power and are transportable. The heart of such a standard (Fig. 8.9) is a glass cell containing isotope-enriched ^{87}Rb vapour whose ground-state transition is interrogated by a radiation field in a microwave resonator.

As in the H maser and the Cs clock, near room temperature, both ground states are almost equally populated. In the rubidium standard the population is re-distributed by optical

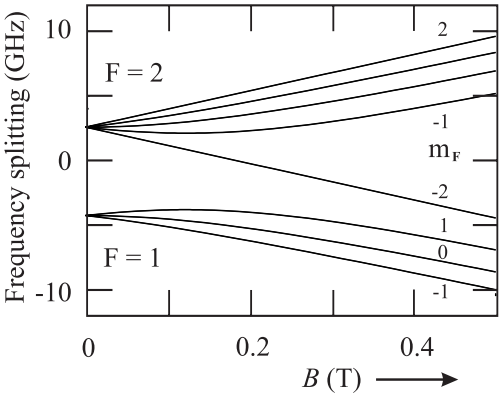


Figure 8.8: The energies of the hyperfine states of ^{87}Rb in a magnetic field.

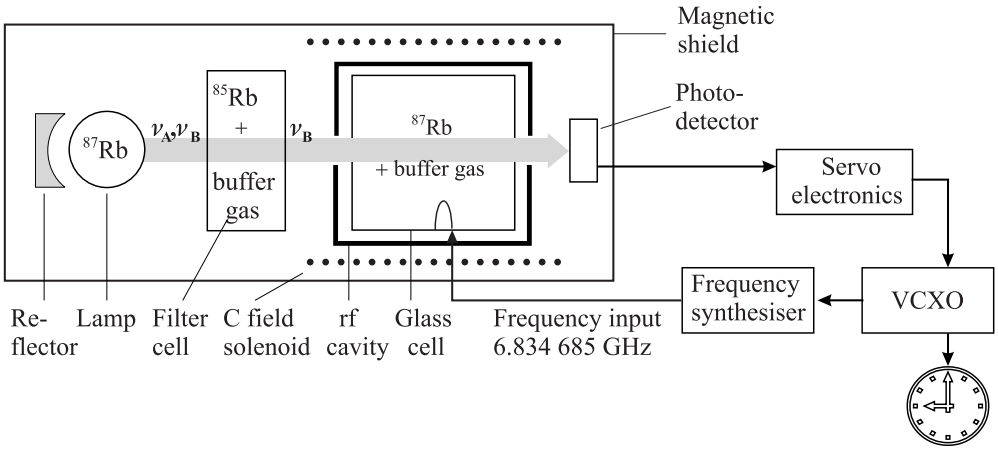


Figure 8.9: Layout of a rubidium frequency standard.

pumping to allow absorption of the microwave at 6.83 GHz. Due to a fortuitous coincidence between the spectra of ^{87}Rb and ^{85}Rb, optical pumping can be achieved by using a simple ^{87}Rb discharge lamp and a ^{85}Rb absorption filter. A discharge lamp on its own would not allow one to achieve optical pumping, since there are emission lines that always couple both ground states to the same excited state.

The effect of the ^{87}Rb lamp – ^{85}Rb filter combination can be understood from Fig. 8.7 and Fig. 8.10. The excited $5\,P_{1/2}$ and $5\,P_{3/2}$ states are connected to the hyperfine-split $5\,S_{1/2}$ ground states by 8 and 12 allowed optical transitions, referred to as the D1 ($\lambda = 795$ nm) and D2 ($\lambda = 780$ nm) resonance lines and indicated in Fig. 8.7 by dotted and full vertical lines, respectively. The natural linewidths of these lines shown in Fig. 8.10 are about 6 MHz. As a result of the Doppler and collisional broadening in the lamp, the two D1 lines starting from the $F' = 1$ and the $F' = 2$ states (separated by 816 MHz) in general overlap and two lines (ν_A and ν_B) are emitted which are separated by 6.83 GHz. Due to a fortuitous coincidence

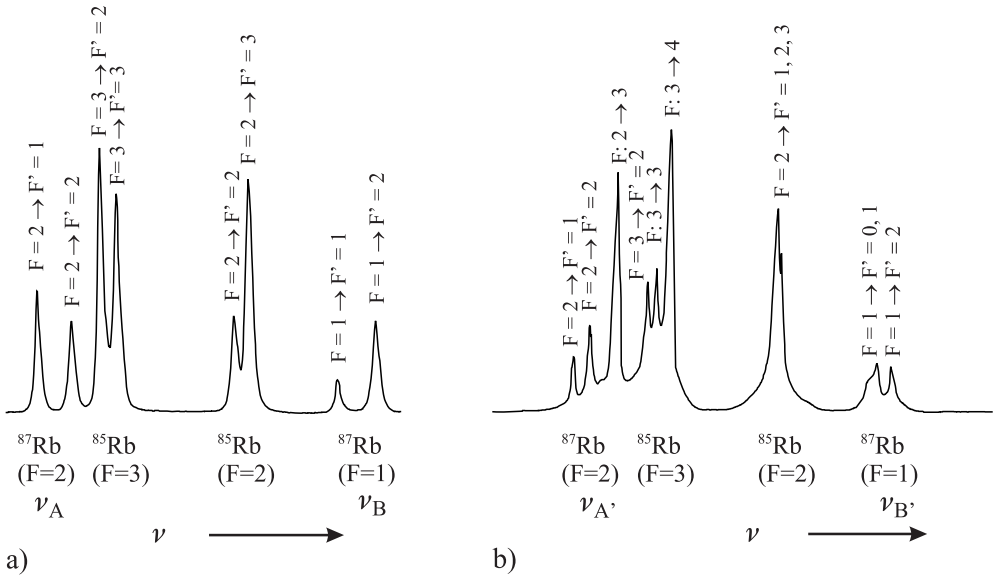


Figure 8.10: Measured Doppler-free absorption spectra of a mixture of ^{85}Rb and ^{87}Rb. a) D1 line ($\lambda = 795\,\text{nm}$). b) D2 line ($\lambda = 780\,\text{nm}$).

(Fig. 8.10 a), the frequency of the transition $F = 2 \rightarrow F' = 1, 2$ is only about 1.3 GHz lower than the centre of the Doppler broadened D1 transition from the $F = 3$ state to the excited $F' = 3$ and the $F' = 2$ states in ^{85}Rb. In contrast, the separation between the transition of ^{87}Rb designated as $\nu_B(F = 1)$ and the transition in ^{85}Rb designated as $(F = 2)$ is larger than the Doppler widths and hence, the radiation ν_B from the ^{87}Rb lamp is not absorbed in the ^{85}Rb filter. The situation is analogous for the D2 lines (Fig. 8.10 b). Consequently, behind the filter cell the radiation from the lamp (Fig. 8.9) contains only the components $\nu_B, \nu_{B'}$ that can excite the transition from the $F = 1$ ground state.

After a few absorption-emission cycles in the absorption cell, virtually all ^{87}Rb atoms are optically pumped into the $F = 2$ state and the absorption cell becomes transparent to the filtered light, whose power is monitored with a photodetector. When the microwave radiation is tuned into resonance with the ^{87}Rb atoms in the absorption cell, transitions from $F = 2$ to $F = 1$ take place. Hence, the absorption of the $\lambda = 795\,\text{nm}$ radiation is increased with an associated decrease of the signal from the photodetector. A servo electronics uses this absorption feature at the microwave resonance to tune a crystal oscillator (VCXO) such that the microwave frequency from the synthesiser is kept at the atomic resonance.

Discharge Lamp The discharge lamp contains a noble gas like krypton and about a milligram of rubidium which is either isotope enriched ^{87}Rb or a mixture of ^{87}Rb and ^{85}Rb given by its natural abundance. The rubidium atoms fluoresce in a radio frequency excited discharge at an operational temperature of up to 140° C. Rubidium lamps are supposed to operate for as long as twenty years. Even though the bulb of the lamp is made of alkali-resistant glass

ageing during this time results, e.g., from the diffusion of the rubidium into the glass which can be as high as 100 micrograms after one year of operation.

Absorption Cell In order to keep the Rb atoms in the Lamb–Dicke regime (see Sections 6.1, 10.1.4) the length L of the cell has to be $L < \lambda/2 = c/(2\nu) \approx 2\,\text{cm}$ where the wavelength corresponding to the transition in free space is $\lambda = 4.4\,\text{cm}$. For optimum signal-to-noise from the resonance light transmitted through the absorption cell, it is important that the rubidium vapour be neither too optically thin nor too dense. An appropriate density for the length of a cell of about $1\,\text{cm} \leq L \leq 2\,\text{cm}$ is obtained by operating the cell at a temperature between 70 °C and 80 °C.

Rubidium absorption cells also contain a buffer gas for different reasons. The most important one is that the collisions with the cell walls lead to a large spin-relaxation rate and hence limit the interaction time. Adding an inert gas like nitrogen or a rare gas like neon leads to frequent collisions of the rubidium atoms with the particles of the buffer gas, thereby increasing the time before the atoms reach the walls of the cell. Furthermore, the use of a buffer gas avoids blackening of the glass cell due to the high chemical activity of alkali metals. The pressure of the buffer gas is chosen to be of the order of 1 kPa (680 Pa N_2; [350] or 4 kPa Ne; [351]) to obtain diffusion velocities of the rubidium atoms of about 1 cm/s or below, depending on the size of the cell.

The collisions of the rubidium atoms with the constituents of the buffer gas, in general leads to a frequency shift of the resonance line. If the pressure is such that the rubidium atoms predominantly suffer from two-body collisions with the atoms or molecules of the buffer gas, the frequency shift at a given temperature is proportional to the pressure. Light buffer gases like helium, neon or nitrogen lead to a positive pressure shift whereas rare gases of heavier atoms like argon, krypton and xenon, reduce the frequency of the resonance. The collisional shift under typical conditions for a single-component buffer gas can be as large as one kilohertz, corresponding to a fractional shift of about 1.5×10^{-7}. In order to keep the frequency shifts that result from temperature fluctuations with the corresponding pressure fluctuations small, a suitable mixture of two buffer gases is often chosen. For example, a mixture of 12% neon and 88% argon at a pressure of 5.3 kPa leads to a fractional temperature shift of about -1.5×10^{-9}/K.

Buffer gases can be avoided if the walls of the cells can be coated with organic materials like paraffin wax [352] that can reduce the spin-relaxation rate due to wall collisions by four orders of magnitude. Other types of wall coatings, coating procedures, associated processes for alkalis have been described by Stephens *et al.* [353].

Microwave Resonator Often microwave cavities sustaining the TE_{111} or TE_{011} modes are used [354] with a loaded $Q_c \lesssim 400$ of the microwave resonator including the resonance cell [355]. To reduce the size of the rubidium frequency standard, which is limited by the size of the microwave resonator, sometimes a magnetron-type microwave resonator is employed [266, 356]. Couplet *et al.* [356] report on a design where the magnetic field is concentrated at the end region of the cell. In order to obtain an even more compact design, the functions of the filter and the absorption cell were combined in a single cell. In this "integrated filter technique", the irradiance of the unwanted hyperfine components, e.g., ν_A, is steadily decreasing along the light path in the cell and optical pumping becomes spatially dependent. The associated processes are in general more complicated than discussed above.

Electronics The electronics package has to serve several purposes. It contains the controllers to keep the lamp, the filter and the absorption cell at their respective optimal temperatures. The electronics package has also to provide the 6.83 GHz signal for interrogation which is derived from a high-quality quartz oscillator in a phase-coherent way by means of a synthesiser. The signal is frequency modulated and the associated variation of the photocurrent is phase-sensitively detected and used to stabilise the oscillator to the atomic transition. Standard frequencies, e.g., 10 MHz, are provided directly which are derived from the quartz together with a PPS signal. The power consumption of such rubidium standards is typically below 10 watts for continuous operation but somewhat higher for the warm-up phase.

8.2.2 Performance of Lamp-pumped Rubidium Standards

As in any passive standard, the frequency instability of the rubidium standard is ultimately limited by the fluctuations of the signal derived from the interrogation process. Provided that the detected photocurrent exhibits white frequency noise the stability is determined by the signal-to-noise ratio S/N of the photocurrent and the $Q_{\rm at}$ factor of the atomic resonance the ultimate stability of the lamp-pumped rubidium standard can be calculated using (3.96). Couplet *et al.* [356] give for their rubidium standard an amplitude of the absorption dip of 1 μA and a photocell current of 150 μA corresponding to 9×10^{14} electrons per second from which one calculates a shot-noise of 5 pA/$\sqrt{\rm Hz}$. From the $S/N \approx 2 \times 10^5 \sqrt{\rm Hz}$ and $Q_{\rm at} \approx 3.6 \times 10^6$ one calculates $\sigma_y(\tau) \approx 1.4 \times 10^{-12}/\sqrt{\tau/{\rm s}}$. Due to other noise contributions and the dead time in the interrogation scheme the observed instability is higher. In the best cases, lamp-pumped devices show an instability of $\sigma_y(\tau) \gtrsim 4 \times 10^{-12}\sqrt{\tau/{\rm s}}$ for $1\,{\rm s} < \tau < 1000\,{\rm s}$ [355–358]. Depending on the environmental conditions and the particular device, the relative instability may reach a flicker floor between 10^{-12} and 10^{-13} after about 1000 s. For times longer than about 10^4 s the Allan deviation increases again mainly due to fluctuations of the frequency shifts, resulting from the collisions of the rubidium atoms with the atoms or molecules of the buffer gas and from light shifts.

Considering that different manufacturers use different compositions of the buffer gases and that the filling pressure of different cells can be controlled only to a limited extent, it is obvious that rubidium cell standards are not accurate frequency standards on their own. Furthermore, in the long run, the composition and the pressure of the buffer gas and the Rb vapour in the absorption cell can change, e.g., due to diffusion into the cell walls. As a result, the frequencies of rubidium clocks are largely affected by ambient conditions, in particular by the temperature. Typically, the sensitivity is $\Delta\nu/\nu \approx 10^{-10}$/K. For applications requiring less sensitivity to ambient temperature fluctuations, optimised design is expected to lead to a sensitivity of 10^{-13}/K [358].

Another effect that influences the frequency of rubidium clocks results from light shifts that occur if the frequency of the light that is used for optical pumping is detuned from the optical resonance frequency. The complicated processes involving temperature-dependent Doppler shifts and broadenings, collisional shifts, isotope dependencies or processes associated with the filtering, may easily shift the centre of gravity of the lines and lead to a frequency shift of the rubidium clock of a few hertz.

As a result of these influences rubidium clocks drift by a few times 10^{-11} per month [358]. To be used as frequency standards they have to be calibrated against more accurate standards.

Consequently, the rubidium clocks are often "disciplined" [359] by GPS (see Section 12.5) where the signals from the GPS satellites are used to long-term stabilise the frequency of the oscillator in the rubidium standard. Depending on the instability of the rubidium clock and on the conditions of the GPS signal, the disciplining algorithm is such that GPS takes over for times above 1000 s or 10 000 s to ensure the long-term stability and accuracy.

8.2.3 Applications of Rubidium Standards

Rubidium frequency standards have their market since they can be compact with a volume of a litre or below and are low priced. They are best used when instabilities in the 10^{-11} regime are needed as this is where quartz crystals become very costly.

Rubidium clocks are used when timing is critical but the host device may need to perform autonomously. As an example we consider a satellite where an internal clock is synchronised via microwave links with clocks on Earth. Occasionally the connection may be interrupted and, by re-establishing the contact, both the internal and Earth clocks must be synchronised to maintain communication. As an example we consider the performance of the rubidium atomic clock carried onboard the United States Milstar FLT-2 satellite in geo-synchronous orbit used to provide secure military communications. After its activation in November 1995, the clock's linear frequency ageing rate was determined as $+7 \times 10^{-14}$/day and $1 \times 10^{-15}(\tau/\mathrm{s})^{-1/2}$ [360].

Similarly, rubidium clocks are used in base stations of cellular telephones where incoming and outgoing signals received from and transmitted to the cellular phones must be synchronised. The accuracy of the signal synchronisation is critical since often a large number of cellular phones access the same site at the same time. Usually, GPS-disciplined rubidium clocks are used where the rubidium clock acts as a fly-wheel providing the short-term stability and the internal synchronisation in the case where there is an insufficient GPS signal.

Other areas where rubidium clocks are employed are audio broadcasting, analogue and digital television transmissions, navigation, military communications and tracking, and guidance control.

8.3 Alternative Microwave Standards

8.3.1 Laser-based Rubidium Cell Standards

Depite the simplicity of rubidium standards based on lamps the use of lasers for state preparation and detection of the microwave-induced hyperfine transition has distinct advantages. The broad-band light from the lamp that does not directly contribute to optical pumping leads to an increased background signal on the photodetector, with the associated degradation of the achievable signal-to-noise ratio. Furthermore, near-resonant components of the light can limit the number of excited atoms by de-excitation.

Consequently there have been many attempts to establish rubidium cell standards where the lamp and filter is replaced by a laser for optical pumping. It has been shown that the use of diode lasers in rubidium cell standards can reduce the instability to a few parts in 10^{13} [354, 361, 362] and improve the short-term stability between 1 and 10 seconds by an

order of magnitude. However, the medium-term stability for $\tau > 100\,\mathrm{s}$ is still limited by the effects on the buffer gas cell and by the light shift and hence, the flicker floor of around 7×10^{-13} [362] achieved with laser pumping has been reached also with lamp-filter designs. The light shift can be largely reduced with a narrow-band laser whose frequency is kept exactly at resonance at the expense of an additional stabilisation scheme thereby further reducing the medium-term stability.

8.3.2 All-optical Interrogation of Hyperfine Transitions

Rubidium or caesium cell standards that can be even more compact and have much less power consumption may be based on an interrogation scheme that avoids the microwave resonator altogether. Interrogation of the hyperfine transition in an alkali vapour cell is accomplished by using optical rather than microwave radiation [350]. Consider two coherent laser radiation fields tuned such that they connect the two hyperfine ground states to either the $5P_{1/2}$ (D1 line) or to the $5P_{3/2}$ states (D2 line) in the Λ configuration of Fig. 5.12 a). The interaction of the two fields creates a coherence in the ground states by the coherent population trapping (CPT) mechanism (Section 5.3.3.2). If the frequency difference of the two laser fields equals the frequency difference of the two ground states, minimal absorption occurs, sometimes referred to as electromagnetically induced transparency.

The resonance can be detected by several methods. One may use the small variation of the absorption of the exciting beam present in the signal from a photodetector behind the absorption cell. Alternatively, one may detect the fluorescence from the excited atoms. A much more sensitive detection of the resonance signal superimposed on a weak background signal can be obtained if the transmission of an additional weak probe beam is detected [350].

The two optical frequencies necessary for the optical interrogation can be supplied either by two phase-locked diode lasers or by a single-frequency single-mode diode laser whose injection current is modulated to produce phase-coherent side bands, besides the laser carrier frequency. For the latter use, Vertical-Cavity Surface Emitting Lasers (VCSELs) with a high-quality single-spatial and spectral mode are well suited as they exhibit extraordinarily high modulation bandwidths of up to 10 GHz.

We discuss the principal scheme of such a clock (Fig. 8.11) according to Kitching *et al.* [363] where the first-order optical sidebands of the laser, used to interrogate the ground-state coherence of the caesium atoms, were generated by the 4.6 GHz (3.2 GHz in case of ^{87}Rb) modulation frequency synthesiser which was driven by a 5 MHz VCXO.

The 10 kHz servo system was used to stabilise the optical frequency of the laser close to the Doppler-broadened absorption line, having a width of about 1.4 GHz. The 0.5 kHz servo system stabilised the side-band frequency to the dark-line resonance of a width of about 100 Hz. A small magnetic field was applied to select the $\Delta m = 0$ transition. The fractional frequency instability of such a low-power, small scale frequency reference was demonstrated to be $\sigma_y(\tau) < 3 \times 10^{-11}(\tau/\mathrm{s})^{-1/2}$ for $1 < \tau < 10^5$ s.

The optical interrogation of the hyperfine transition is seriously affected by light shifts. Zhu and Cutler [364] have investigated the light shift in systems comprising either two phase-locked lasers or a single frequency-modulated laser. They found that the total light shift could be controlled by adding extra frequency components which can be produced, e.g., by changing the index of the current modulation. The short-term stability of such a CPT-based rubidium

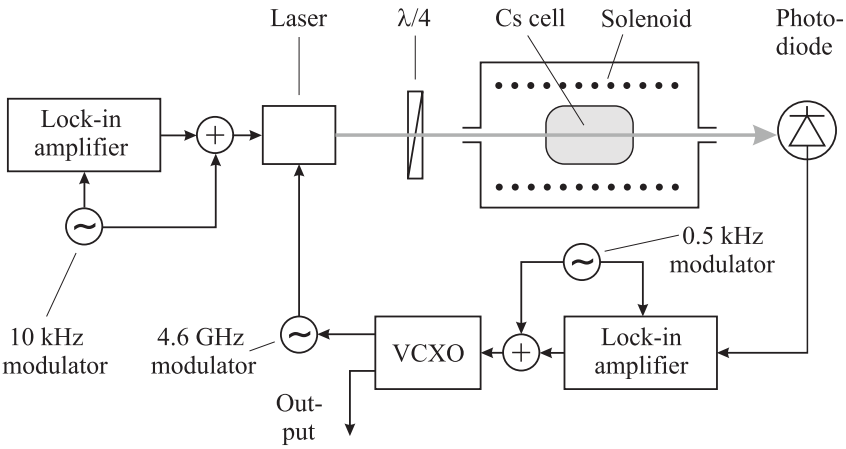


Figure 8.11: Schematic of a compact cell-based frequency reference.

vapour cell standard has been reported as $\sigma_y(\tau) = 1.3 \times 10^{-12}(\tau/\mathrm{s})^{-1/2}$ with a flicker floor reached below 2×10^{-13} for 100 s to 10 000 s [364].

A different method used stimulated Raman scattering in Rb vapour [351] where a laser beam was tuned about 1 GHz from the Rb D1 transition at 795 nm. Due to the stimulated Raman scattering related to dark lines in the Rb cell a second optical field is generated, co-propagating with the first one, but frequency shifted by the ground-state hyperfine splitting. Both radiation fields produce a beat note signal on a fast photodetector with the frequency coinciding with the one of the hyperfine transitions of Rb or Cs that can be used to lock an oscillator.

The ongoing quest for more compact, cheaper and less power-consuming atomic clocks requires the reduction of the size of the absorption cell. Conventional glass-blowing techniques can lead to cell volumes of a few mm^3 but the production and filling procedures for each cell place severe limitations on a low price and therefore on widespread use for, e.g., replacing quartz oscillators. Chip-scale atomic clocks have been suggested [365, 366] based on small-scale vapour cells. A concept has been proposed where a silicon spacer with a large number of regularly arranged holes each of 1.5 mm diameter is anodically bonded in the desired buffer gas environment to two pyrex wafers, forming the windows. The composite wafer can later be cut, leading to a large number of single buffer gas cells [366, 367]. The buffer gas pressure in miniature vapour cells has to be increased as the wall collisions for a given pressure increase with reduced cell dimensions. In a miniature caesium vapour cell containing temperature-compensated N_2/Ar buffer gas at 65 °C at a pressure of 12 kPa, a linewidth of 0.44 kHz ($Q_{\mathrm{at}} \approx 2 \times 10^7$) has been observed leading to an Allan deviation of $\sigma_y(\tau) = 1.5 \times 10^{-10}(\tau/\mathrm{s})^{-1/2}$ falling below 1×10^{-11} for 1000 s [366].

9 Laser Frequency Standards

Lasers represent oscillators with frequencies from the terahertz to the petahertz regime ($10^{12} - 10^{15}$ Hz). Lasers operating in the visible part of the electromagnetic spectrum exceed the frequencies of the oscillators in the microwave domain by about five orders of magnitude and often show frequency stabilities that are highly competitive to the best microwave oscillators. Hence, when the same measurement uncertainty has to be achieved, frequency comparisons between two oscillators in the optical regime can be performed within a much shorter time than between microwave oscillators. In the optical regime there are highly forbidden transitions that can be used for stabilising the frequency of a laser oscillator. The reduction of the wavelength associated with the increase in frequency, however, leads to particular difficulties as it is much harder to eliminate the influence of the Doppler effect with the desired accuracy. Consequently, much work has been devoted in the past years to developing Doppler-free methods that can be applied to optical frequency standards.

Depending on the application, the design of laser frequency standards has followed two different routes. One approach relies on easy–to–operate and simple lasers and uses preferably molecular references with transitions that coincide fortuitously with the frequencies of those lasers. The iodine stabilised He-Ne laser at 633 nm (Section 9.1.3), which is used as a frequency and wavelength standard for the realisation of the length unit, represents a prominent example of this group discussed in Section 9.1. The second strategy first identifies an "ideal" atomic, ionic, or molecular reference and makes use of a tuneable laser, e.g., a dye laser or a diode laser to access the corresponding transition frequency. This approach is often chosen if the highest accuracy is desired.

The particular laser oscillators used for both groups of standards furthermore differ with respect to their inherent noise properties. As an example, consider the power spectral noise densities of a He-Ne laser and a dye laser (Fig. 9.1). At low Fourier frequencies, in both cases technical noise sources are dominating and the resulting laser linewidth is approximately Gaussian. Without proper reduction of these noise contributions the resulting laser linewidth is much larger than that of the Lorentzian expected for quantum noise [369]. For higher Fourier frequencies the noise spectra differ substantially and, hence, in this chapter both groups will be discussed separately beginning with the gas laser frequency standards (Section 9.1). Afterwards (Section 9.2) suitable methods will be discussed that allow one to narrow the linewidths and to stabilise the frequencies of the lasers used as oscillators for optical frequency standards. In Section 9.3 we will treat tuneable lasers where the spectral width of the gain medium of the lasers can extend over a considerable fraction of the spectrum. The larger this width the more provisions have to be made to suppress all but one of the possible laser modes.

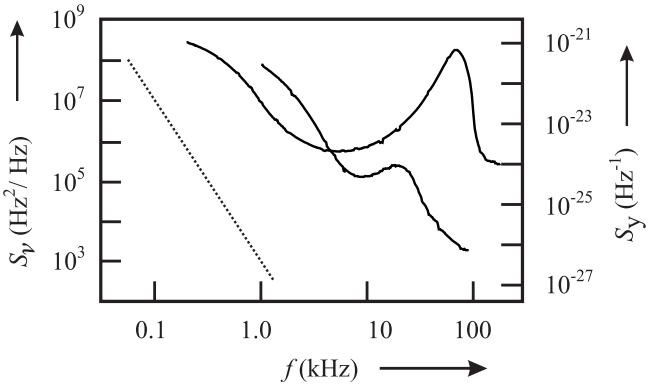


Figure 9.1: Power spectral densities of frequency fluctuations of a free-running dye laser with two different nozzles (lines; courtesy of J. Helmcke) and a He-Ne laser (dots; from [368]).

In the last section of this chapter several optical frequency standards based on neutral absorbers will be treated in more detail. Optical frequency standards based on ionic absorbers will be described separately in Chapter 10 together with the microwave ion standards.

9.1 Gas Laser Standards

The He-Ne laser was one of the very first examples of an optical frequency standard. Owing to its simplicity it is still widely employed in different set-ups and wavelength regions for use as optical wavelength standards of moderate to high accuracy. Other gas lasers used for these purposes are CO_2 lasers and to a lesser extent Ar^+ lasers.

9.1.1 He-Ne Laser

In the He-Ne laser the amplifying medium is provided by neon atoms that are efficiently pumped into the excited states predominantly by excited helium atoms. In a glass capillary a mixture of helium and neon at a pressure $p_{\mathrm{He}} \lesssim 200\,\mathrm{Pa}$ and $p_{\mathrm{Ne}} \lesssim 10\,\mathrm{Pa}$ is excited in an electrical discharge. The discharge is started by an electrical breakdown when an electrical voltage of several thousand volts is applied and it is sustained by a voltage of about 1.5 kV or higher.

Helium atoms are excited by collisions with electrons in the discharge to the $2\ ^1\mathrm{S}_0$- and to the $2\ ^3\mathrm{S}_1$-states. Collisions between the excited helium atoms and the neon atoms lead to a transfer of the energy from the excited He atoms to the Ne atoms according to

$$\mathrm{He}^* + \mathrm{Ne} \rightarrow \mathrm{He} + \mathrm{Ne}^* + \Delta E. \tag{9.1}$$

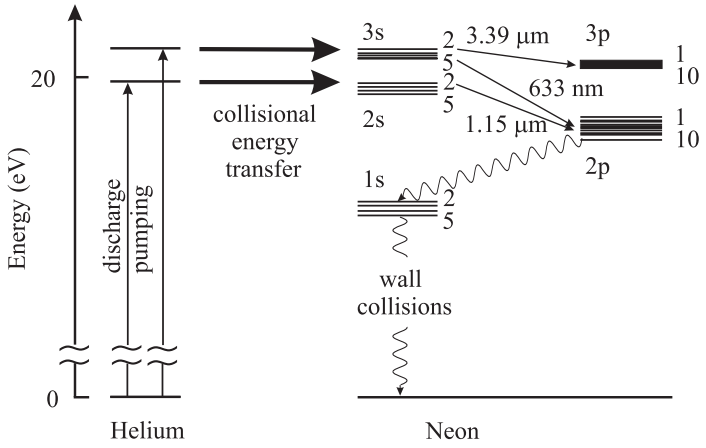


Figure 9.2: Energy levels of helium and neon relevant for the He-Ne laser.

These inelastic collisions are nearly resonant, particularly since the energies of the $3s_2$ and $2s_2$ levels of the Neon atoms are close to the energy levels 2^3S_1 and 2^1S_0 of the Helium atoms (Fig. 9.2).[1]

As a consequence of the population of the higher energy levels, population inversion of the neon atoms can be achieved. From these excited states a variety of different radiating transitions into energetically lower states occur which can be utilised for stimulated emission (e.g. the well known transition $3s_2 \rightarrow 2p_4$ at 633 nm; see Table 9.1).

Table 9.1: Selected transitions of the He-Ne laser used for optical frequency standards. Detailed descriptions of the operational procedures can be found in [95, 370].

Transition		Wavelength	Stab. frequency	Absorber
LS coupling [1]	Paschen	nm	THz	
$5s\ ^1P_1^o \rightarrow 4p\ ^3P_2$	$3s_2 \rightarrow 3p_4$	3.391 µm	88.376 181 600 18	CH_4
$4s\ ^1P_1^o \rightarrow 3p\ ^3P_2$	$2s_2 \rightarrow 2p_4$	1.153 µm	260.103 404 2	$^{127}I_2$
$5s\ ^1P_1^o \rightarrow 3p\ ^3P_1$	$3s_2 \rightarrow 2p_2$	0.640 µm	468.218 332 4	$^{127}I_2$
$5s\ ^1P_1^o \rightarrow 3p\ ^3P_2$	$3s_2 \rightarrow 2p_4$	0.633 µm	473.612 353 604	$^{127}I_2$
$5s\ ^1P_1^o \rightarrow 3p\ ^1D_2$	$3s_2 \rightarrow 2p_6$	0.612 µm	489.880 354 9	$^{127}I_2$
$5s\ ^1P_1^o \rightarrow 3p\ ^3S_1$	$3s_2 \rightarrow 2p_{10}$	0.543 µm	551.579 482 97	$^{127}I_2$

[1] The electron states of the helium atoms are given in the LS coupling scheme. The excited states of neon with the ground state configuration of $1s^22s^22p^6$ in Fig. 9.2 result from a single electron excited to the 3s, 4s, $\cdots$, or 3p, 4p, $\cdots$ states and the remaining shell $1s^22s^22p^5$. The departure from the LS coupling scheme in neon allows the LS designations to be used only in a few cases [94] and often a different, purely phenomenological, notation according to Paschen is used, where the sub-levels of the excited electrons are counted from 2 to 5 for s states and from 1 to 10 for the p states (see Table 9.1).

The lower $2p_4$ laser state of the 633 nm line radiatively decays further to the 1s levels. At higher pressure this state is re-populated by radiative capture and at higher discharge current by collisions with electrons thus increasing the effective lifetime $\tau_2 \geq 20$ ns of the lower laser level. The 1s states are effectively de-populated by collisions with the wall. The design

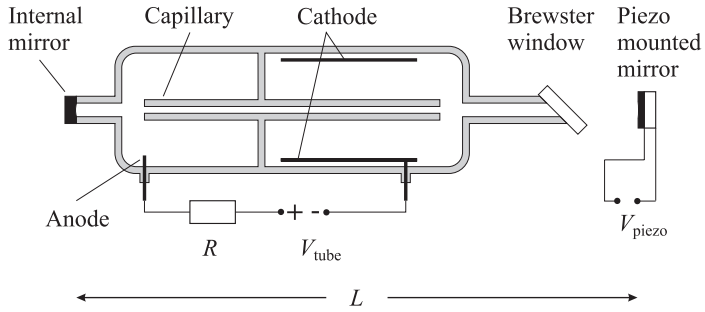


Figure 9.3: Schematics of a Helium-Neon-laser

of the He-Ne laser (Fig. 9.3) takes these peculiarities into account. The discharge current in a capillary is limited to a current of 5 mA $\lesssim$ I $\lesssim$ 20 mA by a series resistor ($R \approx 70$ kΩ). The discharge tube is sealed either with laser mirrors forming the optical laser resonator or with Brewster angled windows which allow reduction of the Fresnel reflection losses in the window, at least for one polarisation component. In the latter case external mirrors attached to a rigid framework make up the optical resonator.

The natural linewidth of any transition is determined by the lifetimes of the initial and final states (see (2.38)). Using the lifetimes $\tau_1 \approx 10$ ns and $\tau_2 \approx 20$ ns of the 3s state and the 4p state, respectively, one calculates the natural linewidth of the 633 nm-laser transition as $\Delta\nu \approx$ 20 MHz. In the laser this line is homogeneously broadened, i.e. broadened by mechanisms that affect all absorbers homogeneously. These mechanisms include pressure broadening by collisions (FWHM ≈ 20 MHz) and saturation broadening (FWHM < 100 MHz). The largest (inhomogeneous) line broadening effect results from the Doppler shift which is different for all atoms. From (5.115) one calculates for neon at λ = 633 nm a width of the Doppler broadened line of about 1.5 GHz.

For a typical length of the laser resonator (Fig. 9.3) of L = 30 cm, the frequency separation of the longitudinal modes is FSR $= c/2L \approx 500$ MHz. In general, more than one mode can overcome the threshold within the gain profile (Fig. 9.4) where the particular laser modes interact with different velocity classes. If the free spectral range is smaller than the homogeneous linewidth, different modes interact with the atoms of the same velocity class leading to a coupling of these modes. In such a case the mode with the highest number of photons uses up the gain and the modes with less photons die out. Strong mode coupling is observed in the argon ion laser and in diode lasers. The distribution of the modes need not be temporally stable since a particular mode can start oscillating as a consequence of, e.g., the length fluctuation of the laser resonator. Like any other amplitude modulation (see Section 2.1.2) the fluctuations of the amplitudes of the particular modes necessarily lead to a broadening of the linewidth of the laser, commonly referred to as mode-partition noise.

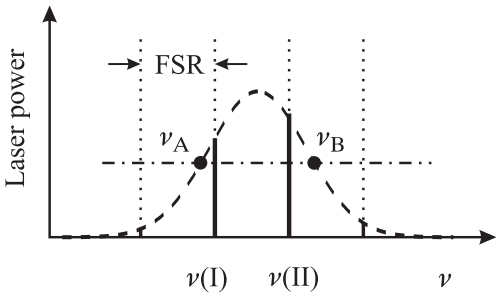


Figure 9.4: Two-mode operation of a laser with a free spectral range (FSR) smaller than the gain profile. Dashed-dotted line: laser threshold.

The noise sources of a gas laser like the He-Ne laser are mainly of technical origin resulting from the fluctuations of the discharge current and of the length of the resonator. The latter ones are affected, e.g., by fluctuations of the temperature, by mechanical vibrations of the set-up or by acoustical perturbations of the ambient air. The corresponding contribution of the technical noise to the power noise spectral density (Fig. 9.1) is dominating for Fourier frequencies below about 10 kHz. For higher Fourier frequencies where this contribution rapidly rolls off, the white noise resulting from the spontaneous emission becomes dominant. For a laser power of P = 1 mW, a length of the resonator L = 30 cm and a reflectivity R = 0.98 we expect a Schawlow–Townes linewidth (3.71) of about 18 mHz that translates to a power spectral density (3.70) of $S_\nu \approx 6 \times 10^{-3}\,\mathrm{Hz^2/Hz}$. Hence, a servo bandwidth of a few kilohertz is sufficient for effectively stabilising the frequency of a He-Ne laser.

9.1.2 Frequency Stabilisation to the Gain Profile

The gain curve of a He-Ne laser resulting from the Doppler broadening of the atomic laser transition has a typical linewidth of about 1.5 GHz. For a resonator whose length is short enough that the free spectral range is larger than the gain profile single-frequency operation is achieved, however, this frequency can be located at any frequency between ν_A and ν_B (Fig. 9.4). The corresponding relative frequency uncertainty of a free-running He-Ne laser of $\Delta\nu/\nu = 1.5\ \mathrm{GHz}/473.6\,\mathrm{THz} \approx 3 \times 10^{-6}$ can be reduced if the laser frequency is kept at a well defined position of the gain profile. The frequency dependence of the output power on the gain profile of the He-Ne laser has been utilised in a variety of convenient and cheap means to stabilise the laser frequency.

9.1.2.1 Two-mode Stabilisation

Consider a gas laser where the length of the resonator and, hence, the axial mode separation is chosen such that over a large tuning range only two adjacent axial modes oscillate. Such a situation is typically encountered if the length of the resonator is about 30 cm corresponding to a mode separation of 500 MHz. Using laser tubes without apparent polarisation-dependent losses i.e. with internal mirrors rather than Brewster angled mirrors, these two modes in general are polarised orthogonally since in this case the two modes experience reduced mode competition losses. The two orthogonally polarised modes can be easily separated behind the rear mirror in a polarising beam splitter, e.g., a Wollaston prism (Fig. 9.5). When the length of the resonator is tuned both modes move across the gain profile and produce different

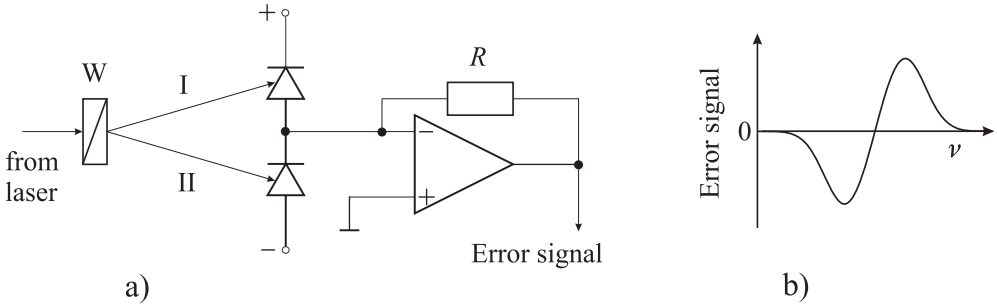


Figure 9.5: Two-mode polarisation scheme. The laser operates simultaneously with two orthogonally polarised modes (I and II), a) that are spatially separated by means of a Wollaston prism W, and whose powers are measured with two photodiodes. The difference of the signals from the two photodiodes can be used to generate a non-zero error signal, b) when the two modes have different intensities.

signals at the respective photodetectors behind the polarising beam splitter. Provided that both photodetectors have the same sensitivities and that the gain curve is symmetric the difference of both signals shows an anti-symmetric discriminant curve (Fig. 9.5 b) with a zero crossing at the line centre, i.e. at the atomic resonance. The difference signal can be used to stabilise the frequency of the laser where the frequencies of both modes are symmetrical with respect to the gain curve. The anti-symmetric shape of the error signal allows the servo amplifier to discriminate whether the length of the laser resonator has to be increased or reduced after a perturbation to set the laser frequency back to the reference frequency.

The two-mode stabilisation technique [371] is frequently used for He-Ne lasers operating in the red (λ = 633 nm) or green (λ = 543 nm) spectral range. The simplicity of the two-mode stabilisation with two orthogonally polarised resonator modes makes it useful for operation in interferometers, in particular in combination with heterodyne interferometers. Caution has to be exercised if one mode is suppressed, e.g., by a polariser. Since the remaining mode is not in the centre of the gain profile the mode frequency can be stabilised to either side of the gain profile depending on the polarity of the difference signal. When such a laser is used as a frequency or wavelength standard this frequency shift has to be corrected accordingly. Difficulties of the two-mode stabilisation technique sometimes arise with lasers as, e.g., the He-Ne laser at 543 nm where the modes change their polarisations when going through the centre of the gain profile. These polarisation jumps can be suppressed by placing a magnet near the gain tube [372]. Moreover, the electronic lock point may be shifted by unbalanced gain of the photodiodes and by electronic offsets.

Investigations of the frequencies of polarisation stabilised He-Ne lasers at $\lambda = 633\,\mathrm{nm}$ during a period of more than two years showed a drift of about 5 MHz [373] corresponding to a fractional frequency variation of 10^{-8}. These lasers often show frequency changes of similar magnitude which can be attributed to fluctuations of external magnetic fields, temperature fluctuations and ageing due to pressure loss in the gain tube.

9.1.2.2 Zeeman Stabilisation

If the laser tube is placed in an axial magnetic field the energy levels of the neon atoms in the amplifying medium are shifted due to the Zeeman effect with the shift being proportional to the applied magnetic field. As a consequence, the laser line splits into two oppositely circularly polarised waves, the frequencies of which differ depending on the magnetic field by typically 300 kHz to 2 MHz. In a Zeeman stabilised laser the two circularly polarised waves are converted into two orthogonally linearly polarised waves by means of a quarter-wave plate. Similarly to the case of the two-mode stabilised laser (Section 9.1.2.1) the difference in the powers of the two waves detected by two detectors can be used to stabilise the laser frequency. An alternative method [374] uses the fact that, due to the strong dispersion, the index of refraction in the centre of the laser line varies with the frequency of the laser line. The difference frequency of both Zeeman modes therefore shows a minimum at the line centre that can be utilised to stabilise the laser frequency. When the Zeeman splitting is used for frequency stabilisation the frequency difference of both modes is much smaller than the separation of the two resonator modes in the two-mode stabilised laser. The higher slope of the discriminant curve leads to a higher gain in the servo system. On the other hand this advantage is paid for by a reduced locking range.

9.1.2.3 Lamb-dip Stabilisation

Consider a laser where only a single resonator mode is excited with the eigenfrequency ν_L. When the laser frequency ν_L is tuned as a function of the length of the resonator the output power exhibits a sharp minimum in the centre of the Doppler-broadened absorption line. This minimum has been predicted by Willis Lamb Jr. [375,376] and, hence is often referred to as the Lamb dip. Its origin is understood by considering a standing wave in the linear laser resonator. The standing wave can be thought to result from two counter-propagating running waves with wave vectors $\vec{k}$ and $-\vec{k}$. These waves are in resonance with atoms whose velocities $\vec{v}\,'$ satisfy the Doppler condition $\nu_L - \nu_0 = \vec{k} \cdot \vec{v}\,'$. At $\nu_L \neq \nu_0$ the two waves interact with different velocity groups. At the corresponding frequencies spectral holes are burnt into the velocity distribution (see Fig. 6.14 b). If the frequency ν_L of the laser is tuned to the frequency ν_0 (see Fig. 6.14 c) of the laser transition by changing the length of the laser resonator both waves interact with the same velocity class. The velocity of these atoms is characterised by a zero Doppler shift along the axis of the laser beam (say the z axis), i.e. by the velocity ($v_z = 0$). In general, the transition of this velocity group is more saturated as in atoms with a velocity far from resonance ($\nu_L \neq \nu_0$) and, hence, the absorption is reduced. As a result, a smaller laser power is sufficient to reach the equilibrium condition where the saturation is increased to lower the gain to the extent needed to compensate for the losses. The Lamb dip thus provides a more narrow spectroscopic feature than the gain profile itself and can be utilised to stabilise the frequency of the laser.

The different types of gain stabilised lasers discussed so far have in common that the active laser medium itself is used for the frequency stabilisation, allowing one to build compact and simple devices. For optical frequency standards these methods suffer from the disadvantage that any variation associated with the laser medium, in general affects the frequency of the laser. Examples include the fluctuation of the discharge current with the associated effects on

the temperature, the electromagnetic field in the plasma or the index of refraction in the laser medium. To overcome these and other deficiencies in the iodine-stabilised He-Ne laser, the absorber (iodine) and the lasing medium (neon) consist of different species which are spatially separated.

9.1.3 Iodine Stabilised He-Ne Laser

The absorption spectrum of the iodine molecule (see Section 5.2.2) has myriads of hyperfine transitions in the green and red part of the visible spectrum and fortuitous coincidences occur with the emission lines of the He-Ne laser (see Table 9.1). Following the early work of Hanes and Dahlstrom [377] the emission frequencies of many gas lasers have been stabilised to iodine absorption lines. Most widely used is the He-Ne laser at λ = 633 nm where the coincidence between the Doppler broadened emission line of the isotope ^{22}Ne and the vibrational transition 11-5 of the R(127) line of the isotope $^{127}I_2$ is utilised. As a matter of fact, these absorption lines are weak and the irradiance required to detect the absorption signals with good signal-to-noise is in general not available in the output beam from a He-Ne laser. The irradiance can be increased by about two orders of magnitude if the absorbing medium is placed inside the laser resonator (Fig. 9.6).

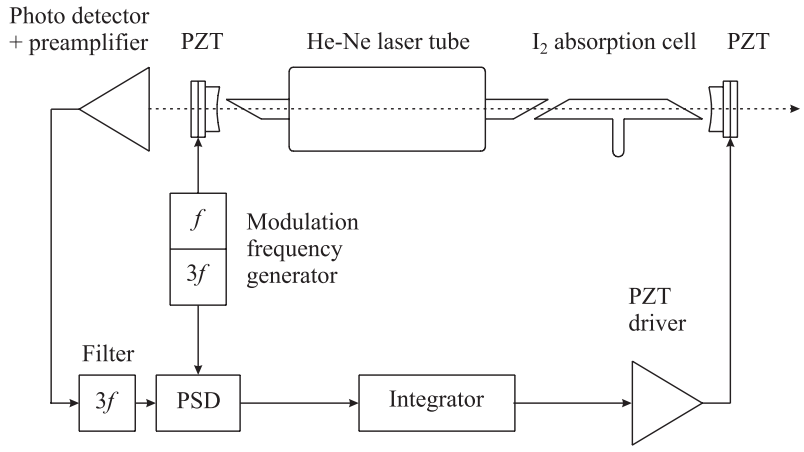


Figure 9.6: Schematics of an iodine stabilised He-Ne laser with the absorption cell inside the laser resonator: PZT, piezo actuator; PSD, Phase-sensitive detector.

Owing to the thermal velocity distribution of the iodine molecules in the absorption cell the hyperfine lines of iodine are Doppler broadened. For arbitrary laser frequency ν_L both counter-propagating laser beams inside the resonator in general interact resonantly with different Doppler-shifted velocity groups. If, however, the frequency of the laser coincides with an unperturbed transition frequency of the molecules, both laser beams are interacting with the same velocity group of molecules having zero velocity in the direction of the laser beams. Correspondingly, the absorption of these molecules is reduced due to the saturation of the corresponding non-linear absorption. Hence, the absorption losses inside the laser resonator

decrease and the output power of the laser increases. For the conditions of the typical iodine stabilised He-Ne laser at λ = 633 nm the output power increases by only about 0.1 %. Often, this variation is not directly detectable as a result of the amplitude noise associated with the discharge. Furthermore, the output power varies by about 10 % when the laser frequency is tuned across the absorption profile. In order to stabilise the frequency to the centre of the weak feature on a large and varying background, first- and higher-order modulation techniques are applied.

9.1.3.1 First-, Third- and Higher-harmonic Detection

To detect the absorption signal buried in the noise the frequency of the laser is modulated across the line and the corresponding synchronous variation of the laser power is phase-sensitively detected (Fig. 9.6). To modulate the laser frequency the length L of the resonator is periodically changed with a few kilohertz by modulating the high voltage applied to the piezoelectric transducer (lead zirconium titanate; PZT) supporting one of the laser mirrors. The phase-sensitive detector (PSD) consists of a lock-in amplifier that changes the polarity each half-cycle of the modulation signal which is later integrated. Due to the change of the polarity with the modulation frequency all frequency components in the detected signal integrate to zero except for the one occurring with the modulation frequency.

Consider a given variation of the output power $P_L(\omega)$ of the laser with the laser angular frequency ω representing, e.g., a narrow absorption feature on a frequency dependent Doppler background. As a consequence of the harmonic modulation of the angular frequency of the laser around the laser frequency ω_0, the output power of the laser varies as

$$P_L(\omega) = P_L(\omega_0 + \Delta\omega \sin \omega_m t) \tag{9.2}$$

where $\Delta\omega$ is the amplitude of the frequency modulation. One finds from Taylor's expansion

$$\begin{aligned} P_L(\omega) &= P_L(\omega_0) + \Delta\omega \sin\omega_m t \left.\frac{dP_L(\omega)}{d\omega}\right|_{\omega_0} + \frac{\Delta\omega^2}{2!}\sin^2\omega_m t \left.\frac{d^2P_L(\omega)}{d\omega^2}\right|_{\omega_0} \\ &+ \frac{\Delta\omega^3}{3!}\sin^3\omega_m t \left.\frac{d^3P_L(\omega)}{d\omega^3}\right|_{\omega_0} + \cdots \end{aligned} \tag{9.3}$$

that the output power $P_L(\omega)$ contains higher-order components of $\sin^n \omega_m$. According to the trigonometric rules for the harmonic functions, $\sin^n \omega_m$ contains terms $\propto \sin\ n\,\omega_m$. Hence, $P_L(\omega)$ includes contributions with higher harmonics $n\omega_m$ of the modulation frequency. According to (9.3) the amplitude of this contribution is proportional to the n^{th} derivative of the output power of the laser.

In the typical set-up of the iodine stabilised He-Ne laser stabilised with the third-harmonic technique (Fig. 9.6) the angular frequency components around $3\omega_m$ are filtered from the detected signal and fed into the lock-in amplifier which is triggered by the angular frequency of $3\omega_m$. This signal of the third harmonic (Fig. 9.7) does not contain the constant, linear and quadratic contributions of the background since it is proportional to the third derivative of the signal rather than to the signal itself. Hence, the central zero crossings of the signal (Fig. 9.7) are not shifted by these contributions and correspond, to a good approximation, to the centre

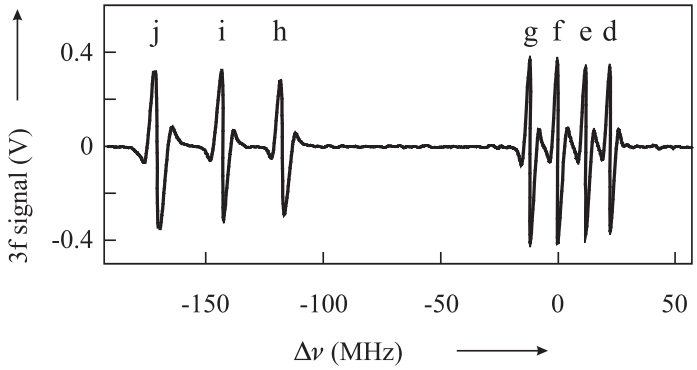


Figure 9.7: Signal of the third harmonic of the iodine $^{127}I_2$ hyperfine components in the R127 (11-5) rotational line. The frequency value of the central zero crossing of the "f" line as recommended by the International Committee of Weights and Measures (CIPM) is 473 612 353 604 kHz [370].

frequencies of the respective absorption lines and can be utilised for frequency stabilisation of the laser.

Even-order derivatives showing a maximum or minimum at the position of the maximal absorption are not suited for frequency stabilisation, in contrast to the odd-order derivatives exhibiting a zero crossing at the frequency where a maximum or minimum of the absorption occurs.

As a result of intercomparisons performed between the lasers operated in the different national metrology institutes over more than thirty years, the reproducibility of the iodine stabilised He-Ne laser is well documented (see e.g. [378, 379] and references therein). The frequency of the stabilised laser is known to depend on a number of operational parameters like the amplitude $\Delta\omega$ of the frequency modulation, the vapour pressure in the iodine absorption cell, or the laser power in the cavity. The variation of the frequency of the laser with the modulation amplitude can be explained by the effect of the residual Doppler background together with an asymmetry of the absorption line. The temperature of the cold finger of the iodine cell changes the vapour pressure and, hence, affects the rate and duration of the collisions of the iodine molecules leading to a pressure broadening and shift of the absorption line. The dependence of the frequency on the laser power results from the modification of the saturation parameter in the iodine cell and by a modification of the index of refraction in the gas discharge. The latter may lead to a gas lens and hence to a distortion of the wave front of the laser radiation with an associated first-order Doppler shift.

Typically, the frequency varies by 6 kHz/Pa near 15 °C and by -10 kHz/MHz as a consequence of the pressure and modulation dependency, respectively. International comparison showed that the frequencies of the majority of the iodine stabilised He-Ne lasers at $\lambda = 633$ nm coincide to about 10 kHz provided all lasers are operated under the same conditions. These standard conditions have been laid down in a recommendation by the International Committee of Weights and Measures (CIPM) [370]. The wall temperature of the iodine cell should be kept at 25 °C ± 5 °C with a cold finger temperature of 15 °C ± 0.2 °C to keep

the vapour pressure fixed. The full width of the modulation of the laser frequency is 6 MHz ± 0.3 MHz and the internal power of one of the counter-propagating waves should be 10 mW ± 5 mW. If these conditions are met and "good practice" is provided [370] a relative uncertainty of the frequency of the laser of 2.5×10^{-11} can be expected. The magnitude of the frequency dependence of the stabilised laser varies for lasers of different design. In order not to exceed the estimated relative uncertainty of 2.5×10^{-11} this contribution must be smaller or equal to 1.4 kHz/mW.

With the possibility at hand to measure the frequencies of a particular standard more regularly and at different locations by using femtosecond laser frequency combs (Section 11.5) it has been shown [380] that a particular laser was reproducible to 1×10^{-12} after transportation.

Even though the third-harmonic technique is capable of largely reducing the contributions of a non-linear background, there are higher-order contributions. Hence, a fifth-harmonic modulation has been used sometimes. Frequency offsets between the third- and fifth-harmonic techniques have been observed for different hyperfine components (d to g) in the range from 26 kHz to 35 kHz [379]. This indicates that the actual frequency of the iodine stabilised He-Ne laser differs considerably from the frequency of the respective transition in the unperturbed iodine molecules. The frequency reproducibility attributed to this type of standard critically depends on the extent of similarity of the design and operation.

The frequency modulation of the output of the iodine stabilised He-Ne laser is disadvantageous in cases where another laser or an interferometer is to be locked precisely to the standard. It has been demonstrated, however, that the dither modulation can be removed almost completely by use of an external acousto-optical modulator (AOM) [381]. In this work the AOM was operated in a double-pass configuration to avoid beam-displacement induced amplitude modulation. When the AOM was driven by the same frequency used to dither the iodine-stabilised laser with a suitable phase shift and a properly adjusted amplitude, the linewidth of the dither modulated laser could be reduced from 6 MHz to a few kilohertz. In principle, negative feedback alone could be used to suppress a measured frequency modulation. Taubmann and Hall [381] found, however, that this technique gave inferior results compared to a feed-forward technique, as a result of the added noise in the servo system with the required large bandwidth.

9.1.4 Methane Stabilised He-Ne Laser

The coincidence of the 3.39 µm radiation of the He-Ne laser (Fig. 9.2) with the methane transition ν_3, P(7) component $F^{(2)}$ has led to the development of a high-accuracy optical frequency standard which is also recommended for the realisation of the metre (see Table 13.1; [370]). Since the CH_4 molecule is highly symmetrical its energy levels are not easily perturbed and, hence, the frequency shifts associated with external perturbations are small. Compared to the iodine molecule the methane molecule has a low mass which leads to a large velocity at room temperature (see Table 9.2). To reduce the associated transit-time broadening, stationary lasers and portable lasers have been set up [382–384] with beam diameters in the absorption cell of up to 20 cm and more. These lasers allow one to resolve the hyperfine-structure triplet with a frequency separation of about 11 kHz (Fig. 9.8) or even the recoil doublet in each line separated by 2.15 kHz [217]. In the Lebedev Physical Institute a laser system was operated [383] comprising three He-Ne lasers. They were used as a reference laser with

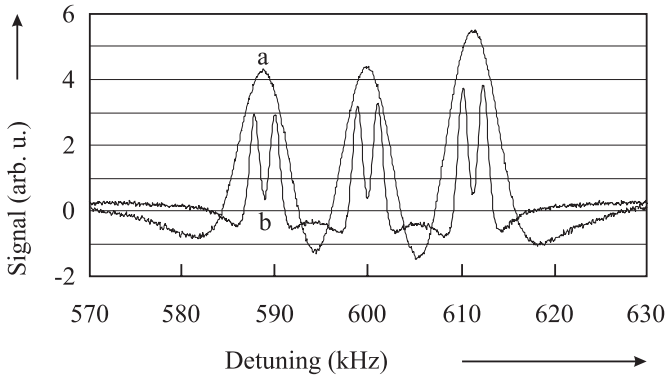


Figure 9.8: Double recordings of the first derivatives of saturated dispersion signals obtained with two different CH_4 stabilised He-Ne lasers. a) transportable laser, beam diameter 60 mm. b) Stationary laser, beam diameter 200 mm. Courtesy of M. Gubin.

narrow emission spectrum, a single-mode heterodyne laser and a main laser with a telescopic beam expander inside the cavity to resolve the hyperfine structure of methane. The first and third laser were operated in a double-mode regime where saturated absorption and saturated dispersion resonances were used to stabilise the laser to the methane transition. Saturation dispersion resonances were recorded by monitoring the beat frequency between both modes which varies owing to the frequency pulling effect near the centre of the absorption line. A frequency-to-voltage converter delivered the voltage as the fast error signal that was used to lock the laser. Transportable standards of this kind have been shown to have a frequency reproducibility for different devices of 1×10^{-12} and a frequency repeatability for a single device of 2×10^{-13} during several months [385]. Bagayev and co-workers [384] reported on a similar transportable CH_4/He-Ne laser system employing three lasers. The measured Allan deviation showed a minimum of 5×10^{-15} at $\tau \approx 10$ s. The reproducibility over three years was given as 30 Hz ($\Delta\nu/\nu \approx 10^{-12}$).

With stationary systems, much higher resolutions of the absorption feature can be obtained when only the slowest molecules are selected. The laser operated at the Institute of Laser Physics in Novosibirsk [221] comprised an 8 m long internal absorption cell which was cooled to 77 K. The slowest molecules were selected optically (Section 6.5.2) i.e., by applying low laser power to saturate the transition only in the slowest molecules. At a pressure of 6×10^{-4} Pa, Bagayev *et al.* [221] obtained a full width at half maximum of about 100 Hz. The use of the cold absorbers at the same time reduces the second-order Doppler effect considerably.

Compared to the F line the E line of the P(7) transition does not show a hyperfine splitting and, hence, has the potentially higher accuracy. The E line with its frequency about 3 GHz lower than the F line can be accessed by a Zeeman shifted He-Ne laser or by a tuneable source like an optical parametric oscillator. The frequency of the E line has been determined with respect to the Cs atomic clock as ν_E = 88 373 149 028 553 $\pm$ 200 Hz [386].

Owing to its simplicity and high frequency accuracy based on early frequency measurements [370,388–393] the methane stabilised HeNe laser has been used often as a starting point

Table 9.2: Comparison of properties of I_2 and CH_4 relevant to optical frequency standards at 532 nm and 3.4 µm, respectively, after [387].

	I_2	CH_4
Transition	a_{10} (line # 1110 [108])	F_2 - P(7) in the ν_3 band
Natural linewidth	380 kHz	10 Hz
Doppler broadening	430 MHz (300 K)	275 MHz (300 K)
2^{nd} order Doppler shift	5×10^{-12}	10^{-12}
Pressure broadening	0.11 MHz/Pa	0.11 kHz/Pa
Pressure shift	4 kHz/Pa	1 Hz/Pa
Transit-time broadening	5 kHz	170 kHz
	(2 w_0 = 2 mm, 300 K)	(2 w_0 = 2 mm, 77 K)
Saturation broadening	660 kHz $\sqrt{1 + I/I_{\text{sat}}}$	500 kHz $\sqrt{1 + I/I_{\text{sat}}}$
ac Stark shift	25 kHz/mW	
	(linearised at 1 mW)	
dc Stark shift		1 kHz/(V/cm)
1^{st} order Zeeman shift		2 kHz/mT
2^{nd} order Zeeman shift		0.1 kHz/(mT)^2
Hyperfine splitting	$\approx$ 10 MHz	11 kHz
Recoil splitting	5.55 kHz	2.2 kHz

for frequency chains to the visible and UV region to perform measurements of the highest accuracy [394, 395].

9.1.5 OsO_4 Stabilised CO_2 Laser

The two three-fold degenerate vibrational oscillations of the symmetry type F_2 in the OsO_4 molecule occurring at $\nu_3 = 28.9$ THz [396] coincide with the emission spectrum of the CO_2 laser and thus can be used to set up a frequency standard near 9.6 µm. The natural abundance of the isotopes ^{192}Os, ^{190}Os, and ^{189}Os is 41.0 %, 26.4 %, and 16.1 %, respectively. Like the CH_4 molecule, the OsO_4 molecule belongs to the spherical top molecules. In each of these molecules the principal moments of inertia are the same for rotations around the principal axes of the molecule. There are three different types of rotational levels referred to as A, E and F. In the molecule $^{192}Os^{16}O_4$ the nuclear spin of the four identical nuclei is zero and only the A rotational levels occur. Systematic frequency shifting effects are small in OsO_4 stabilised lasers since the energy levels are not easily perturbed by external fields. In the even isotopes of osmium there is no hyperfine structure. Owing to the high mass of the OsO_4 molecule the effects resulting from the second-order Doppler effect and the recoil splitting ($\approx$ 15 Hz) are small.

CO_2/OsO_4 frequency standards have been set up in several laboratories [222, 397–403]. A typical standard [402] comprised a 1 m long sealed CO_2 laser whose frequency was locked on the third derivative of the OsO_4 saturation peak. In this laser the OsO_4 absorbers were kept

in a cell of 1.5 m length kept inside a Fabry–Pérot high-finesse build-up cavity that allowed the transitions to be saturated with less than 1 μW of external laser power. The Fabry–Pérot cavity was modulated and a first-harmonic detection technique was used to keep the centre frequency on the molecular resonance. By comparing two independent systems over six months a fractional reproducibility of 2×10^{-13} has been found [403]. A short-term instability of $\sigma_y(\tau) = 6.6 \times 10^{-14}(\tau/\mathrm{s})^{-1/2}$ up to $\tau = 300\,\mathrm{s}$ with a minimum of the Allan deviation of 4×10^{-15} near about 500 s has been reported [404].

Several absolute frequency measurements have been performed (see e.g. [397, 398, 403, 405, 406] and references therein) with a fractional uncertainty as low as 7×10^{-13}. Some of the frequencies of the OsO_4 stabilised laser have been included into the list of radiations recommended by the CIPM for the realisation of the metre [370]. $^{189}OsO_4$ and $^{197}OsO_4$ with odd osmium isotopes and, hence, with hyperfine structure have been utilised to determine the spin-rotation constants and to set up a different grid of frequencies [407]. By selecting slow molecules in an 18 m long gas cell, a linewidth as low as 160 Hz at a pressure of 2×10^{-4} Pa has been reported [222].

9.2 Laser-frequency Stabilisation Techniques

The frequency of lasers can be stabilised by using the dispersive or absorptive feature of a suitable resonance frequency supplied by either microscopic references such as atoms, molecules, and ions or macroscopic ones provided, e.g., by a Fabry–Pérot resonator. Polarisation spectroscopic or rf phase-modulation spectroscopic techniques are characterised by a very high sensitivity. As an example of the former one we discuss the Hänsch–Couillaud technique in Section 9.2.1. A particular phase-modulation spectroscopic method, referred to as Pound–Drever–Hall technique, is often used to pre-stabilise a laser to a Fabry–Pérot resonator and thereby narrowing its linewidth (Section 9.2.2). Two often-used different phase modulation spectroscopic methods will be discussed later that allow one to long-term stabilise the frequency of a laser to reference frequencies of quantum systems.

9.2.1 Method of Hänsch and Couillaud

Hänsch and Couillaud [408] have devised a frequency stabilisation scheme that uses a polarisation spectroscopic method in combination with an optical resonator to derive a dispersion-like signal and consequently the technique is referred to as Hänsch–Couillaud technique. Consider a light field with amplitude $E^{(0)}$ impinging on a Fabry–Pérot interferometer (FPI). The FPI (Fig. 9.9) comprises an internal element such as a Brewster plate, a polariser or a birefringent crystal exhibiting polarisation-dependent losses.

As a result of the internal polarising element the polarisation components $E_{\parallel}^{(r)} = E^{(0)} \cos\theta$ and $E_{\perp}^{(r)} = E^{(0)} \sin\theta$ of the reflected light field being parallel and perpendicular to the direction of minimum loss will experience minimal and maximal losses, respectively. θ is defined as the angle between the polarisation of the incoming beam and the direction of minimum loss. The complex amplitudes of the reflected fields $E_{\perp}^{(r)}$ and $E_{\parallel}^{(r)}$ are given (see (4.92))

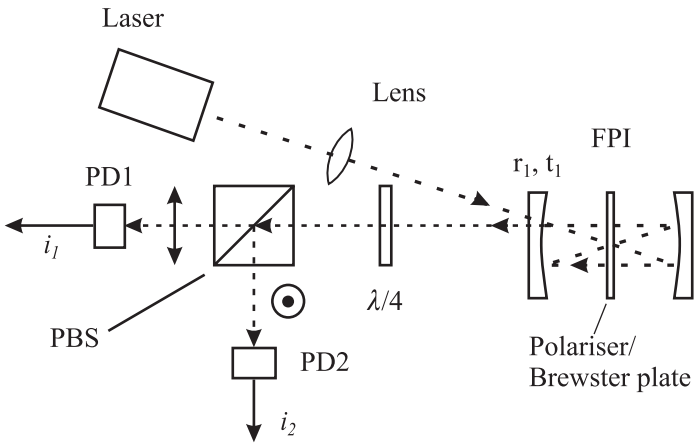


Figure 9.9: Set-up for the stabilisation of the frequency of a laser to a Fabry–Pérot Interferometer (FPI) by means of polarisation spectroscopy according to Hänsch and Couillaud [408]. PBS: Polarising Beam Splitter. $\updownarrow$ and $\odot$ indicate the in-plane and out-of-plane polarisations of the laser beam. PD: Photodiode.

as

$$E_{\perp}^{(r)} = E_{\perp}^{(0)} r_1 \tag{9.4}$$

and

$$\begin{aligned} E_{\parallel}^{(r)} &= E_{\parallel}^{(0)} \left(r_1 - \frac{t_1^2}{r_1} \frac{r e^{-i\delta}}{1 - r e^{-i\delta}} \right) \\ &= E_{\parallel}^{(0)} \left(r_1 - \frac{t_1^2 r^2}{r_1} \frac{-r^2 + \cos\delta - i \sin\delta}{(1-r^2)^2 + 4r^2 \sin^2(\delta/2)} \right) \end{aligned} \tag{9.5}$$

with $\delta \equiv 2\Delta\omega L/c$. Here, r_1 and t_1 are the amplitude reflection and transmission factors of the entrance mirror, respectively. r is the amplitude reduction factor that includes not only the reflection factor of the output mirror but also the internal losses including that due the two additional reflections in the confocal cavity of Fig. 9.9 used off-axis. As a consequence of the different finesse for the two polarisation directions the two components $E_{\parallel}^{(r)}$ and $E_{\perp}^{(r)}$ suffer from phase shifts that are in general different if the frequency of the laser light does not coincide with the eigenfrequency of the optical resonator (see Fig. 4.16). Due to its high loss the perpendicular component is essentially reflected at the entrance mirror with low frequency-dependent phase shift and, hence, can serve as a phase reference for the larger detuning-dependent phase shift of the parallel beam. The phase shift will lead to an ellipticity of the combined beam which is detected by an analyser for the polarisation state of the beam. The combination of the $\lambda/4$ plate and the polarising beam splitter serves as such an analyser. With a suitable alignment of the $\lambda/4$ plate the power of the linearly polarised light is divided equally between the two ports. Elliptical polarised light can be thought of as composed of left-handed and right-handed circularly polarised components with different amplitudes. The $\lambda/4$ plate

generates linearly polarised light from these two components with two orthogonal components that are detected by the two photodiodes. The photocurrents i_1 and i_2 are proportional to the squared amplitude $|E_1|^2$ and $|E_2|^2$, respectively of the waves behind the polarising beam splitter. $|E_1|^2$ and $|E_2|^2$ can be derived [408] by using the Jones matrices describing the evolution of polarised light (see e.g. [409])

$$E_{1,2} = \frac{1}{2}\begin{pmatrix} 1 & \pm 1 \\ \pm 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}\begin{pmatrix} E_{\parallel}^{(r)} \\ \\ E_{\perp}^{(r)} \end{pmatrix} \tag{9.6}$$

where the first Jones matrix describes a linear polariser set at $45\,^\circ$ and the second one a quarter-wave plate with the fast axis horizontal. Hence

$$|E_{1,2}|^2 = \left|\frac{1}{2}\left(E_{\parallel}^{(r)} \pm iE_{\perp}^{(r)}\right)\right|^2. \tag{9.7}$$

The difference signal of the photocurrents $i_1 - i_2 \propto |E_1|^2 - |E_2|^2$ is readily calculated by using (9.7) and (9.5) as follows

$$i_1 - i_2 \propto |E^{(0)}|^2\, 2\cos\theta \sin\theta \frac{t_1^2 r^2 \sin\delta}{(1-r^2)^2 + 4r^2 \sin^2 \delta/2}. \tag{9.8}$$

The corresponding signal (Fig. 9.10) can be used as an error signal for the stabilisation with a steep slope at resonance and a capture range extending halfway to the next resonance.

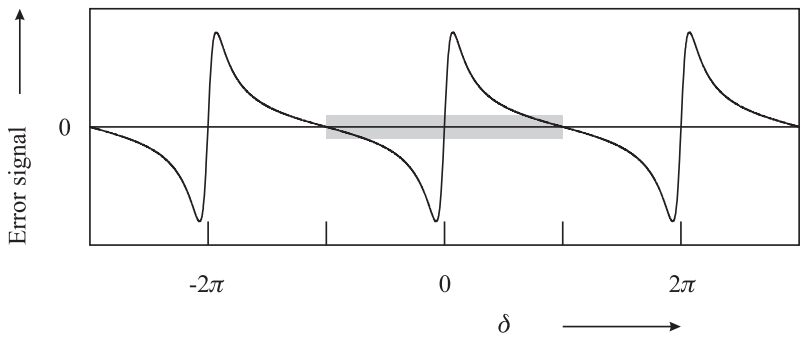


Figure 9.10: Error signal of the Hänsch-Couillaud method calculated according to (9.8) with cavity finesse $F^* = \pi r/(1-r^2) = 14$. The grey area indicates the capture range when locked to the central resonance.

The Hänsch–Couillaud technique is very versatile owing to its simple and inexpensive set-up and is often used for pre-stabilisation of a laser. However, as with any dc technique the locking point is sensitive to baseline drifts of the error signal and it is furthermore affected by the technical noise of the laser at low Fourier frequencies.

9.2.2 Pound–Drever–Hall Technique

The Pound–Drever–Hall technique named after its inventors [410] and R. V. Pound who used a corresponding technique in the microwave regime [411] is a phase modulation spectroscopic method which is applied to stabilise the frequency of a laser to an optical resonator. In the Pound–Drever–Hall technique (Fig. 9.11) the phase of the laser beam of angular frequency ω is modulated by an electro-optical modulator (Section 11.2.2) with the angular frequency ω_m.

In the case of small modulation index $\delta \ll 1$ it suffices to take into account the carrier of

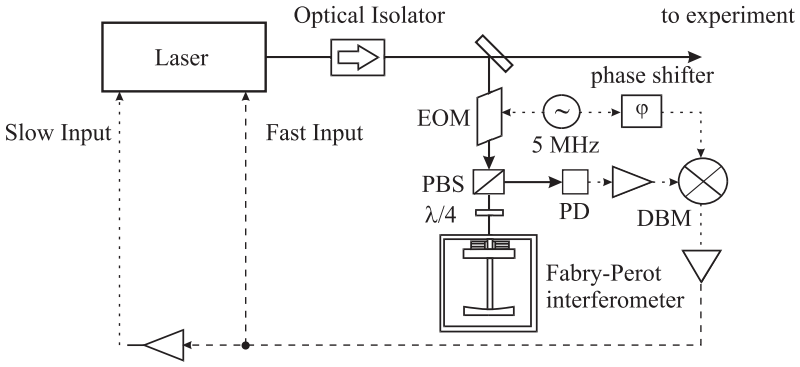


Figure 9.11: Pound-Drever-Hall stabilisation scheme. Optical and electrical paths are depicted as solid and dashed lines, respectively. EOM: Electro-Optical Modulator. PBS: polarising beam splitter. DBM: double - balanced mixer.

angular frequency ω (see (2.52)) and the two nearest sidebands at $\pm\omega_m$

$$E_{\mathrm{FM}}(\omega) = E_0 \left[J_0(\delta) e^{i\omega t} + J_1(\delta) e^{i(\omega+\omega_m)t} - J_1(\delta) e^{i(\omega-\omega_m)t} \right] + c.c.. \tag{9.9}$$

When the carrier and the sidebands are reflected from the Fabry–Pérot interferometer their amplitudes and phases are changed by the complex reflectivity coefficient $r_{\mathrm{FP}}(\omega)$ (see (4.92)) to

$$\begin{aligned} E_r(\omega) = \frac{E_0}{2} \Big[& r_{\mathrm{FP}}(\omega) J_0(\delta) e^{i\omega t} \quad + \quad r_{\mathrm{FP}}(\omega+\omega_m) J_1(\delta) e^{i(\omega+\omega_m)t} \\ & - \quad r_{\mathrm{FP}}(\omega-\omega_m) J_1(\delta) e^{i(\omega-\omega_m)t} \Big] + c.c.. \end{aligned} \tag{9.10}$$

To separate the beam reflected back by the Fabry–Pérot interferometer from the impinging beam, a quarter-wavelength plate or a Faraday rotator is employed (see Fig. 9.11). The application of either device leads to a rotation of the direction of the polarisation of the laser beam and the reflected beam is directed by a polarising beam splitter to a photodetector. The current i_{PD} of the photodiode with detection efficiency η_{PD} depends on the reflected power P_r

$$i_{\mathrm{PD}} \approx \eta_{\mathrm{PD}} P_r \propto E_r E_r^* \tag{9.11}$$

and, hence,

$$\begin{aligned} i_{\rm PD} \;\propto\;& \left[J_0^2(\delta)|r_{\rm FP}(\omega)|^2 + J_1^2(\delta)\left\{|r_{\rm FP}(\omega+\omega_m)|^2 + |r_{\rm FP}(\omega-\omega_m)|^2\right\}\right. \\ +\;& J_0 J_1 r_{\rm FP}(\omega) r_{\rm FP}^*(\omega+\omega_m) e^{-i\omega_m t} \\ -\;& J_0 J_1 r_{\rm FP}(\omega) r_{\rm FP}^*(\omega-\omega_m) e^{i\omega_m t} \\ +\;& J_0 J_1 r_{\rm FP}^*(\omega) r_{\rm FP}(\omega+\omega_m) e^{i\omega_m t} \\ -\;& J_0 J_1 r_{\rm FP}^*(\omega) r_{\rm FP}(\omega-\omega_m) e^{-i\omega_m t} \\ -\;& J_1^2\left\{r_{\rm FP}(\omega+\omega_m) r_{\rm FP}^*(\omega-\omega_m) e^{i2\omega_m t}\right. \\ & \left.\left. -r_{\rm FP}^*(\omega+\omega_m) r_{\rm FP}(\omega-\omega_m) e^{-i2\omega_m t}\right\}\right] \end{aligned} \tag{9.12}$$

The photocurrent (9.12) is comprised of the three dc components resulting from the power of the carrier and the two sidebands and of the components modulated with the beat frequencies between the three fields. Consider a detector which is constructed such that it is sensitive only to frequencies near the modulation frequency ω_m, i. e. for the beat note between the carrier and the side bands. Then the beat note between the high-frequency and the low-frequency side bands occurring at $2\omega_m$ is usually suppressed, e.g., by a notch filter in the detector. Hence, in this case, we need to keep only the terms from (9.12) oscillating with the modulation frequency ω_m and obtain

$$\begin{aligned} i_{\rm PD}^{(\omega_m)} \;\propto\;& J_0 J_1 \left\{[r_{\rm FP}(\omega) r_{\rm FP}^*(\omega+\omega_m) - r_{\rm FP}^*(\omega) r_{\rm FP}(\omega-\omega_m)]\exp[-i(\omega_m t)]\right. \\ +\;& \left.[r_{\rm FP}^*(\omega) r_{\rm FP}(\omega+\omega_m) - r_{\rm FP}(\omega) r_{\rm FP}^*(\omega-\omega_m)]\exp[i(\omega_m t)]\right\} \end{aligned} \tag{9.13}$$

which is equivalent to [2]

$$\begin{aligned} =\;& 2J_0 J_1 \Re\mathrm{e}\left\{r_{\rm FP}(\omega) r_{\rm FP}^*(\omega+\omega_m) - r_{\rm FP}^*(\omega) r_{\rm FP}(\omega-\omega_m)\right\}\cos\omega_m t \\ +\;& 2J_0 J_1 \Im\mathrm{m}\left\{r_{\rm FP}(\omega) r_{\rm FP}^*(\omega+\omega_m) - r_{\rm FP}^*(\omega) r_{\rm FP}(\omega-\omega_m)\right\}\sin\omega_m t. \end{aligned} \tag{9.14}$$

From (9.14) we find that the photocurrent

$$i_{\rm PD}^{(\omega_m)} \propto J_0(\delta) J_1(\delta)\left[A(\Delta\omega)\cos(\omega_m t) + D(\Delta\omega)\sin(\omega_m t)\right] \tag{9.15}$$

contains a sine term and a cosine term with the components $D(\Delta\omega)$ and $A(\Delta\omega)$. To calculate the coefficients A and D we use (4.98)

$$r_{\rm FP} = -\frac{\Delta\omega(\Delta\omega + i\Gamma/2)}{(\Gamma/2)^2 + \Delta\omega^2} \tag{9.16}$$

rather than the exact Airy function (4.92). From

$$\begin{aligned} & r_{\rm FP}(\omega) r_{\rm FP}^*(\omega+\omega_m) - r_{\rm FP}^*(\omega) r_{\rm FP}(\omega-\omega_m) \\ =\;& \frac{\Delta\omega\left[\Delta\omega + i(\Gamma/2)\right](\Delta\omega+\omega_m)\left[\Delta\omega+\omega_m - i(\Gamma/2)\right]}{\left[(\Gamma/2)^2+\Delta\omega^2\right]\left[(\Gamma/2)^2+(\Delta\omega+\omega_m)^2\right]} \\ -\;& \frac{\Delta\omega\left[\Delta\omega - i(\Gamma/2)\right](\Delta\omega-\omega_m)\left[\Delta\omega-\omega_m + i(\Gamma/2)\right]}{\left[(\Gamma/2)^2+\Delta\omega^2\right]\left[(\Gamma/2)^2+(\Delta\omega-\omega_m)^2\right]} \end{aligned} \tag{9.17}$$

[2] Using $A = a + ib$ one finds $A\exp(-i\omega t) + A^*\exp(i\omega t) = 2a\cos\omega t + 2b\sin\omega t$.

we can separate the real and imaginary part by comparison of (9.14) with (9.15) to determine the coefficients A and D. After some straightforward but tedious algebra we obtain

$$D(\Delta\omega) = -4\frac{\omega_m^2(\Gamma/2)\Delta\omega\left[(\Gamma/2)^2 - \Delta\omega^2 + \omega_m^2\right]}{\left[\Delta\omega^2 + (\Gamma/2)^2\right]\left[(\Delta\omega+\omega_m)^2 + (\Gamma/2)^2\right]\left[(\Delta\omega-\omega_m)^2 + (\Gamma/2)^2\right]} \quad (9.18)$$

and

$$A(\Delta\omega) = 4\frac{\omega_m(\Gamma/2)^2\Delta\omega\left[(\Gamma/2)^2 + \Delta\omega^2 + \omega_m^2\right]}{\left[\Delta\omega^2 + (\Gamma/2)^2\right]\left[(\Delta\omega+\omega_m)^2 + (\Gamma/2)^2\right]\left[(\Delta\omega-\omega_m)^2 + (\Gamma/2)^2\right]}. \quad (9.19)$$

The photocurrent (9.15) contains contributions depending on $\cos(\omega_m t)$ and on $\sin(\omega_m t)$. By comparison with the phase ϕ of the modulation frequency ω_m one can select the former (by multiplying with $\cos(\omega_m t)$) or the latter (by multiplying with $\sin(\omega_m t) = \cos(\omega_m t - 90°)$). In the experiment (Fig. 9.11) the phase comparison is readily done by employing a double-balanced mixer (DBM; Fig. 3.13). If a low pass filter is used only the difference frequency remains and the frequencies $\omega_{\mathrm{RF}} \approx \omega_{\mathrm{LO}}$ remain. Consider the case $\omega_{\mathrm{RF}} t = \omega_{\mathrm{LO}} t + \phi$ where the phase difference ϕ can be adjusted by a phase shifter (Fig. 9.11) and the IF signal at the output is $1/2\cos\phi$ (see (3.88)). Hence, the error signal applied to the RF input of the DBM is phase sensitively detected, i.e. rectified and integrated and the DBM acts like a lock-in amplifier.

Fig. 9.12 and Fig. 9.13 show the signals calculated from (9.18) and (9.19), i.e. the signal out of phase with the modulation frequency and the in–phase–signal, respectively, versus the detuning from the resonance $\Delta\omega = \omega - \omega_0$ of the resonance of a FPI for $\omega_m = 10\,\Gamma$.

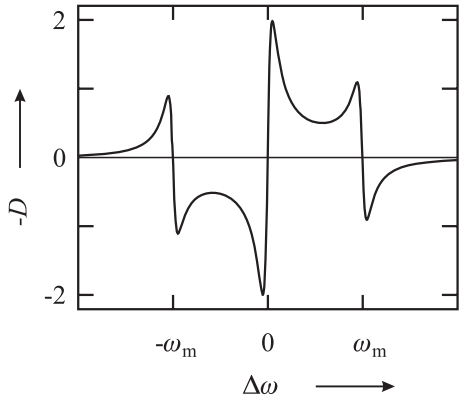


Figure 9.12: Dispersion term $-D(\Delta\omega)$ calculated from (9.18) for $\omega_m = 10\Gamma$.

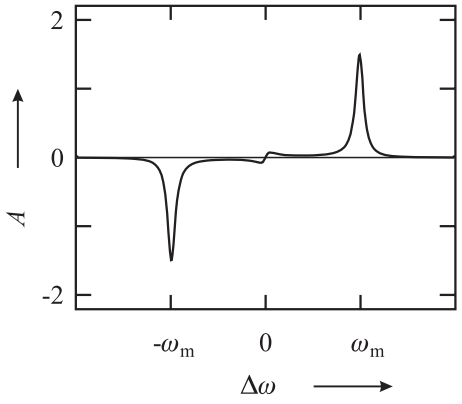


Figure 9.13: Absorption term $A(\Delta\omega)$ calculated from (9.19) for $\omega_m = 10\Gamma$.

The out–of–phase signal (Fig. 9.12) can be thought of as a superposition of three dispersion-like signals at the resonance frequency ω_0 and at $\omega_0 + \omega_m$ and $\omega_0 - \omega_m$ corresponding to the three contributions of the phase modulated reflected wave. As the sidebands are out of phase with the carrier the corresponding signals change sign. If the modulation frequency is considerably larger compared to the half width of the resonance of the interferometer the three structures are separated. Close to the resonance frequency of the FPI only a

small fraction of the power of the carrier is reflected but the phase of this wave varies strongly in this region (Fig. 2.5). Similarly to the Hänsch-Couillaud method the phase comparison between the three reflected fields contains the information about the detuning of the laser frequency from the resonance frequency of the FPI. The steep slope at the resonance frequency of the FPI can be used as an error signal for the frequency stabilisation of the laser. In this case the output signal of the mixer $\propto \omega_{\mathrm{RF}} - \omega_{\mathrm{LO}}$ is a measure of the fluctuations of the laser frequency.

By setting the phase difference between the modulation frequency and the signal from the photodiode to 0° the DBM averages over all phases (Fig. 9.13) and becomes particularly sensitive to variations of the amplitude, i.e. the absorption by the FPI. Hence, there are two absorptive structures at the sideband frequencies. At resonance the absorptive signal vanishes as both beat notes with the sideband frequencies are out of phase by π.

As the error signal is proportional to $J_0 \times J_1$ (see (9.15)) the highest signals are obtained for a modulation index of $\delta_{\max} \approx 1.08$ (Fig. 9.14).

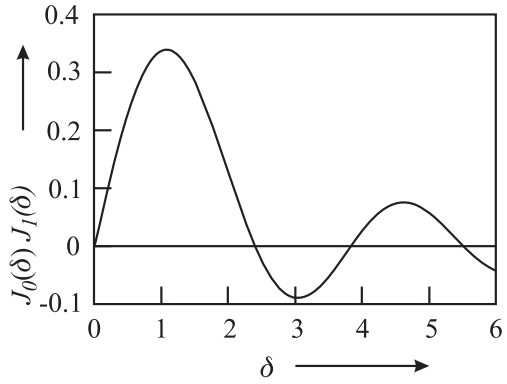


Figure 9.14: $J_0(\delta)J_1(\delta)$ versus modulation index δ. The curve is proportional to the slope of the error signal in the Pound-Drever-Hall technique.

The ultimate sensitivity that can be achieved by this technique is limited by the shot noise of the detected error signal. In a properly designed servo loop any deviation of the error signal from zero is counteracted by a feedback signal to the servo element.

We consider the case where the resonator is impedance-matched and the modulation frequency is large compared to the linewidth of the resonator. In this case only the optical power of the two sidebands is reflected from the resonator and is incident onto the detector. If the electrical noise of the photodetector with efficiency η is due to the shot-noise of the laser amplitude the spectral density of frequency fluctuations of the locked laser is given by

$$S_\nu = \frac{\delta\nu_c}{\nu}\sqrt{\frac{h\nu}{8\eta P_d}} \tag{9.20}$$

where P_d is the laser power impinging on the detector. The Allan deviation calculated from (9.20) is

$$\sigma_y(\tau) = \frac{1}{4Q}\sqrt{\frac{h\nu}{\eta P_d \tau}}. \tag{9.21}$$

Both, the Hänsch-Couillaud technique and the Pound-Drever-Hall technique have the advantage with respect to the side lock technique (Section 2.3.2.1, [412]) that the frequency of the oscillator is stabilised to the centre of the resonance. A particular advantage of the Hänsch-Couillaud technique is based on the fact that the oscillator frequency need not be modulated, whereas the Pound-Drever-Hall technique often (except in the case of diode lasers) needs an additional electro-optical modulator. In the latter case, the error signal of the detected frequency fluctuations is located in a band around the modulation frequency and has to be mixed down to the base band, whereas the error signal in the Hänsch-Couillaud technique is already in the base band. This additional complication of the Pound-Drever-Hall technique is often compensated by the fact that the modulation frequency can be chosen high enough for the technical noise of the laser to be no longer relevant. Particular emphasis has to be given to any spurious amplitude modulation with the same frequency as the modulation frequency, since this residual amplitude modulation is detected in the Pound-Drever-Hall technique and shifts the locking point. This is of particular importance for residual amplitude modulation produced in the electro-optical modulator as a result of any piezoelectric effects. Much attention has been devoted to the suppression of such effects [413, 414].

9.2.3 Phase-modulation Saturation Spectroscopy

Phase-modulation spectroscopy [415, 416] is often used to stabilise the frequency of a laser to the first-order Doppler free saturated absorption lines in an external absorber cell [118, 417–420]. The method is closely related to the Pound-Drever-Hall technique. Consider the experimental set-up (Fig. 9.15 a) of an iodine stabilised laser where an adjustable portion of the output of a laser is split off by use of a $\lambda/2$ plate and a polarising beam splitter PBS1. A similar combination (at PBS 2) allows one to generate a probe beam and a counter-propagating pump beam for interaction with the iodine molecules in the absorber cell. The probe beam is phase modulated with angular frequency ω_m and modulation index δ. To consider the interaction with the molecular absorbers we first assume that $\delta < 1$ where the electric field can be represented by a carrier together with a low-frequency and a high-frequency sideband (Section 2.1.3).

When passing the iodine absorption cell the phases and amplitudes of these three electric field components with the respective frequencies are affected differently by the interaction with the absorbers. Following Bjorklund [415] we take into account the absorption and the optical phase shift (dispersion) of each one of the three waves $l = -1, 0, +1$ by a factor $T_l = \exp[-\alpha_l - i\phi_l]$ where α_l describes the amplitude attenuation due to absorption and ϕ_l represents the phase shift experienced by the l^{th} field component. Hence, the probe field after the interaction is given by

$$E_{\text{probe}}(t) = E_{0,\text{probe}}/2 \left[T_0 e^{i\omega t} + T_1 \frac{\delta}{2} e^{i(\omega+\omega_m)t} - T_{-1} \frac{\delta}{2} e^{i(\omega-\omega_m)t} \right] + c.c. \quad (9.22)$$

Near resonance where the balance between the three waves is perturbed an amplitude modulation of the power of the laser beam occurs that can be detected with a photodetector

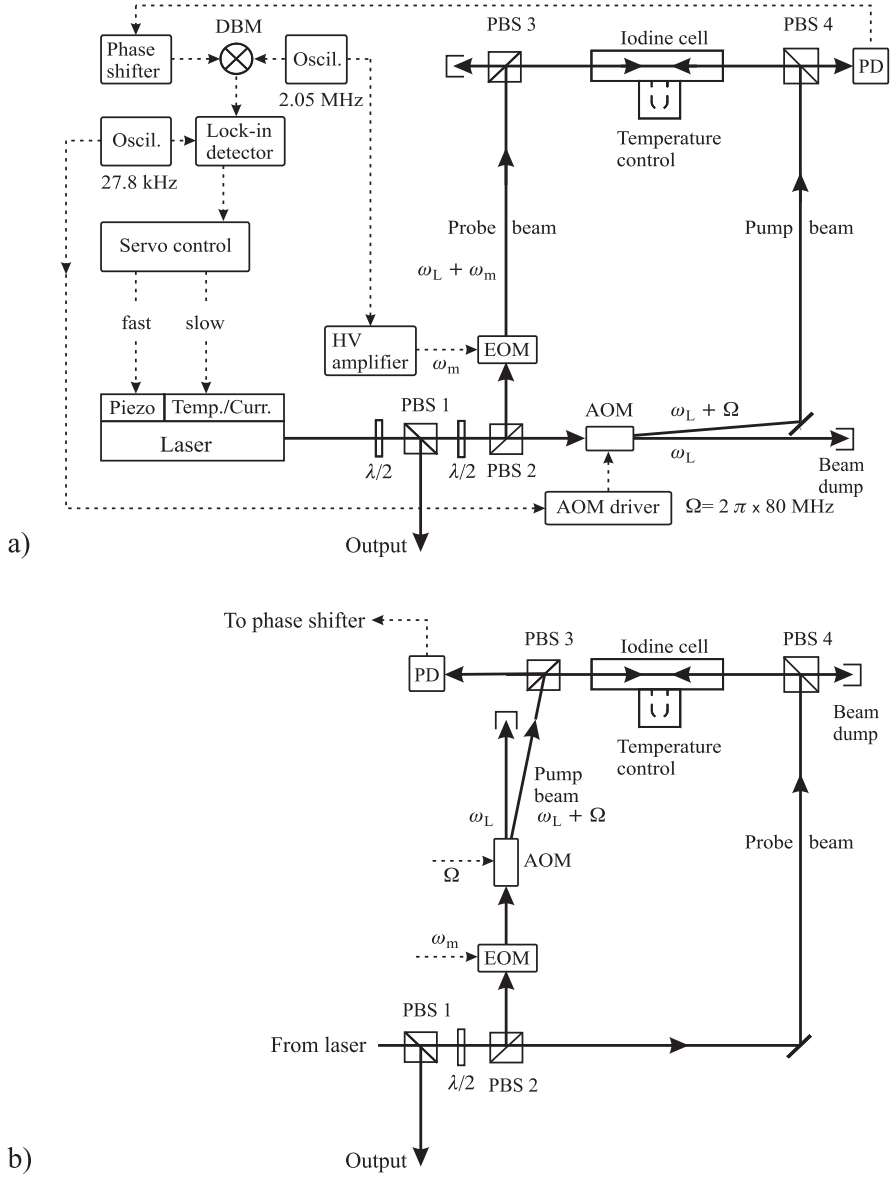


Figure 9.15: Set-up of an iodine stabilised frequency-doubled Nd:YAG laser at 532 nm using the technique of a): phase modulation and b): modulation-transfer. EOM: Electro-Optical Modulator; AOM: Acousto-Optical Modulator; PBS: Polarising Beam Splitter; PD: Photodetector. Optical and electrical paths are shown as solid and dashed lines, respectively.

PD in Fig. 9.15 a). The signal is calculated from

$$P_{\text{probe}} \propto |E_{0,\text{probe}}|^2 e^{-2\alpha} \Big| e^{-i\phi_0} e^{i\omega t} \quad + \quad \frac{\delta}{2} e^{-i(\alpha_0-\alpha_1)} e^{-i\phi_1} e^{i(\omega+\omega_m)t} \\ - \quad \frac{\delta}{2} e^{-i(\alpha_0-\alpha_{-1})} e^{-i\phi_{-1}} e^{i(\omega-\omega_m)t} \Big|^2. \quad (9.23)$$

Dropping terms with δ^2 one obtains

$$P_{\text{probe}} \propto e^{-2\alpha} \Big\{ 1 \quad + \quad \left[e^{-i(\alpha_0-\alpha_1)} \cos(\phi-\phi_0) - e^{-i(\alpha_0-\alpha_{-1})} \cos(\phi_0-\phi_{-1}) \right] \delta \cos\omega t \\ + \quad \left[e^{-i(\alpha_0-\alpha_1)} \sin(\phi-\phi_0) - e^{-i(\alpha_0-\alpha_{-1})} \sin(\phi_0-\phi_{-1}) \right] \delta \sin\omega t \Big\}. \quad (9.24)$$

For $|\alpha_0 - \alpha_1| \ll 1$, $|\alpha_0 - \alpha_{-1}| \ll 1$, $|\phi_0 - \phi_1| \ll 1$, and $|\phi_0 - \phi_{-1}| \ll 1$ the power on the detector is calculated [415] from (9.24) as

$$P_{\text{probe}} \propto e^{-2\alpha} \left[1 + (\alpha_{-1} - \alpha_1)\delta \cos(\omega_m t) + (\phi_1 - 2\phi_0 + \phi_{-1})\delta \sin(\omega_m t) \right]. \quad (9.25)$$

The cosine term in (9.25) is in phase with the modulation frequency and is proportional to the difference of the absorptions of the low and high frequency sidebands. The sine term is out of phase compared to the modulation frequency by $\pi/2$ being proportional to the phase differences.

The error signal can be computed if the spectral absorption and dispersion feature is known. Assuming a Lorentzian lineshape Bjorklund *et al.* [421] determined the absorption signal, the dispersion signal and the modulus of the total beat signal for a large range of parameters. In this case the absorptive component in (9.23) is given by two symmetric lines centred at $\omega - \omega_0$ and at $\omega + \omega_0$ representing the real part of the Lorentzian (Fig. 2.5 a). Hence one expects maximum signal of the in-phase component at $\omega = \omega_0 \pm \omega_m$. The dispersive component is a superposition of three dispersion-shaped features (see Fig. 2.5 b) at the angular frequencies $\omega - \omega_0$, ω_0, and $\omega + \omega_0$.

Hall *et al.* [416] and Shirley [422] have similarly calculated the phase and amplitude changes of the several spectral components including the non-linear absorption and dispersion resonances together with the second-order side bands. They use the notations

$$x_j = \frac{\omega - \omega_0 - j\omega_m}{\Gamma/2}, \qquad L_j = \frac{1}{1+x_j^2} \quad \text{and} \quad D_j = L_j x_j \quad (9.26)$$

where ω_m is the centre frequency of the absorption line and Γ is the homogeneous width of a velocity group with the power broadening by the pump and probe beam taken into account. The variable j can take the values [3] $-1, -1/2, 0, 1/2, 1$. The calculated error signal V_{PMS} is given [416, 422] as

$$V_{\text{PMS}} \quad \propto J_1(\delta) \times \quad (9.27)$$
$$\left\{ \left[(J_0(\delta) + J_2(\delta)) \left(L_{1/2} - L_{-1/2} \right) - J_2(\delta) \left(L_1 - L_{-1} \right) \right] \cos(\Phi) - \right.$$
$$\left. \left[(J_0(\delta) - J_2(\delta)) \left(D_{1/2} - 2D_0 + D_{-1/2} \right) + J_2(\delta) \left(D_1 - 2D_0 + D_{-1} \right) \right] \sin(\Phi) \right\}.$$

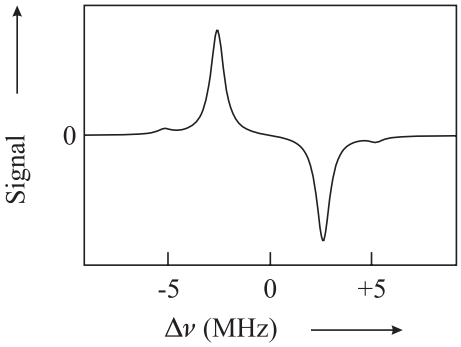


Figure 9.16: Phase-modulation spectroscopy signal calculated according to (9.27) using $\omega_m = 2\pi\times$ 5.185 MHz, $\delta = 0.6, \Gamma = 2\pi \times 0.4$ MHz and absorptive phase ($\Phi = 0^\circ$).

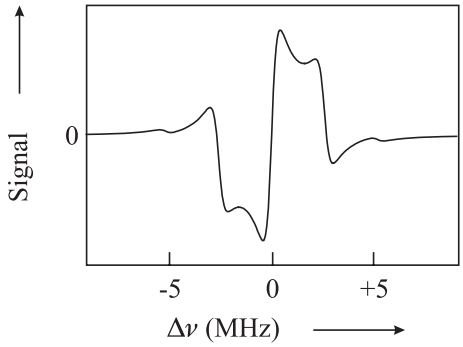


Figure 9.17: Phase–modulation spectroscopy signal calculated according to (9.27) using $\omega_m = 2\pi\times$ 5.185 MHz, $\delta = 0.6, \Gamma = 2\pi \times 0.4$ MHz and dispersive phase ($\Phi = 90^\circ$).

Fig. 9.16 and Fig. 9.17 show the absorptive and dispersive contributions of the different Lorentzians (Fig. 2.5). The steep slope of the experimental dispersive signal (Fig. 9.18) at the centre frequency is used for frequency stabilisation.

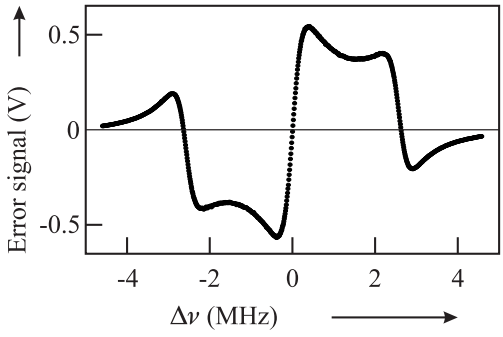


Figure 9.18: Error signal measured with the set-up of Fig. 9.15 a) based on phase-modulation spectroscopy. Compare also Fig. 9.17. Courtesy of H. Schnatz.

In order to allow first-order Doppler-free detection of the absorption line a laser beam counter-propagating to the phase modulated beam (see Fig. 9.15 a) is used to prepare the absorbers. The strong unmodulated pump beam is switched on and off by means of a acousto-optical modulator facilitating phase-sensitive detection of the probe signal when used in combination with a lock-in amplifier. The phase Φ of the photosignal can be adjusted by means of a phase shifter before the photosignal is multiplied in the double-balanced mixer with the signal from the EOM driver to yield an error signal V_{PMS}. The phase shifter (Fig. 9.15) allows one to select the purely absorptive ($\Phi = 0$; Fig. 9.16) or dispersive component ($\Phi = \pi/2$; Fig. 9.17) or any superposition.

[3] Since only the probe beam in the set-up of Fig. 9.15 a) is phase modulated but not the probe beam the first-order sidebands resonances are offset from the resonance frequency by $\omega_m/2$ rather than by ω_m (see Fig. 9.16 and Fig. 9.17).

The method is particularly useful since the modulation frequency can be chosen to be large enough to allow for shot-noise limited detection not affected by the (mainly low-frequency) technical noise of the laser. The accuracy to which a laser can be locked to the centre of an absorption line, however, is affected by any offset in the detected error signal. A particular problem arises from residual amplitude modulation (AM) of the probe beam which adds in-phase sidebands to the out-of-phase sidebands from the phase modulation. Spurious AM is always present to some degree and hence limits the achievable accuracy. This problem can be largely avoided if an unmodulated probe beam is employed as it is done in the spectroscopic method referred to as modulation transfer spectroscopy.

9.2.4 Modulation Transfer Spectrocopy

Near resonance, the interaction of the pump beam and the probe beam with the absorbers is sufficiently non-linear to transfer the modulation from a amplitude-modulated or phase-modulated beam to an unmodulated counter-propagating beam. The modulation transfer [422–427] represents an example of four-wave mixing where the probe-beam carrier, one of its sidebands and the counter-propagating unmodulated beam generate a fourth wave producing a sideband to the latter one. As the modulation transfer is essentially a non-linear process which requires non-linearity near the absorption feature there is virtually no non-resonant background absorption. Consequently, stabilisation techniques based on modulation transfer are not very susceptible to fluctuations of the baseline that can shift the locking point and, hence, the frequency of stabilised laser. The physical mechanisms have been elucidated by Shirley [422] resulting predominantly from modulated hole burning where the amplitude or frequency modulated beam burns holes of modulated depths into the velocity distribution of the absorbers. When the unmodulated beam interacts with the holes whose depths vary with the modulation frequency it experiences modulated absorption and dispersion. With the notations of (9.26) the error signal is given [422, 428] as

$$V_{\mathrm{MTS}} \propto J_0(\delta)J_1(\delta)\times\Big\{ \big(L_1 - L_{1/2} + L_{-1/2} + L_{-1}\big)\cos(\Phi) + \big(-D_1 + D_{1/2} + D_{-1/2} - D_{-1}\big)\sin(\Phi)\Big\}. \tag{9.28}$$

Besides the effect of the modulated hole burning taken into account to derive (9.28) (see Fig. 9.19 and Fig. 9.20) Shirley [422] has pointed out that in general additional weak resonances occur as a result of higher-order interactions with the saturating beam. Furthermore, portions of the carrier and the sidebands of the modulated wave could be Bragg reflected into the counter-propagating, previously unmodulated, beam. The Bragg grating results from a spatial modulation of the population in the standing wave [4] generated by the two counter-propagating pump and probe beams. This contribution changes the height of the inner absorption peaks at $\pm\omega_m/2$ with respect to the outer ones [422].

Due to their non-linear origins the modulation transfer signals in general are weak and, hence, it is important to obtain a large slope of the error signal. For frequencies higher than

[4] In the case of Fig. 9.15 where the unmodulated beam is frequency shifted by Ω by the acousto-optical modulator it is rather a “walking wave”.

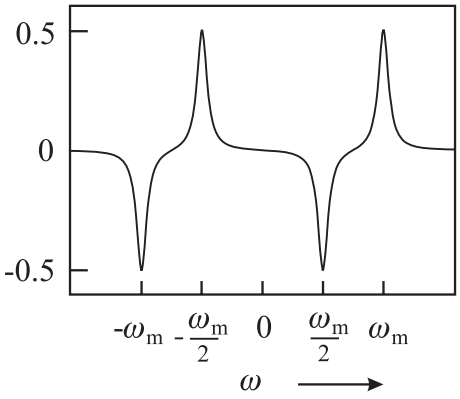


Figure 9.19: Modulation-transfer spectroscopy signal calculated according to (9.28) using $\omega_m = 10\Gamma$ and an absorptive phase setting ($\Phi = 0°$).

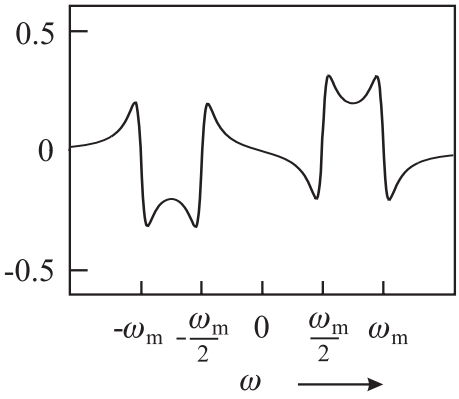


Figure 9.20: Modulation-transfer spectroscopy signal calculated according to (9.28) using $\omega_m = 10\Gamma$ and a dispersive phase setting ($\Phi = 90°$).

the Doppler broadened absorption the non-linearity of the medium rapidly decreases. Consequently, the modulation frequency is favourably chosen $\omega_m \lesssim \Gamma/2$ (Fig. 9.21) to give an error signal that can be used to stabilise the laser. From (9.28) and from experiments [428] it is found that an optimal slope of the discriminant curve (Fig. 9.22) is obtained for $\Phi \approx 50°$ and $\omega_m \approx 0.35\Gamma$.

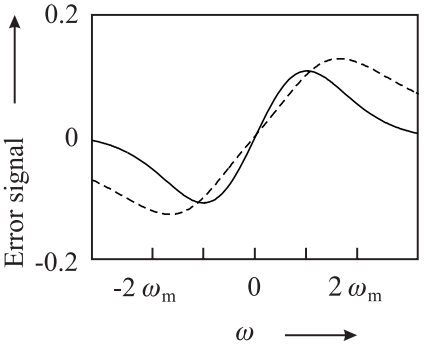


Figure 9.21: Error signal calculated according to (9.28) using $\omega_m = \Gamma/4$ with dispersive phase setting ($\Phi = 90°$; full line) and absorptive phase setting ($\Phi = 0°$; dashed line).

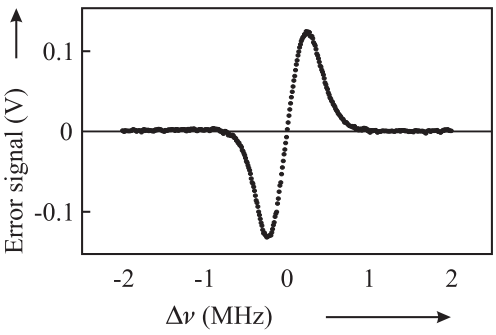


Figure 9.22: Experimental error signal from the iodine stabilisation set-up using modulation transfer spectroscopy (Fig. 9.15 b). Courtesy H. Schnatz.

Jaatinen [429] has calculated the parameters required to produce a maximal signal according (9.28). He finds that an amplitude modulated pump beam provides a higher slope than can be achieved with a frequency modulated beam. Even though the overall lineshape can be described well by simple theory, Eickhoff and Hall [118] find systematic residuals that

may eventually limit the achievable accuracy of frequency standards based on this type of spectroscopy.

Modulation transfer spectroscopy has been used to stabilise, e.g., the frequency of frequency-doubled Nd:YAG laser to iodine (Section 9.4.1).

9.3 Widely Tuneable Lasers

Tuneable lasers are employed to reach a particular absorption line at a given frequency. The cheapest, smallest lasers with lowest energy consumption and the ability to be tuned across several tens of nanometres are diode lasers. Even though they can be produced to operate as single mode over a large wavelength region (Fig. 9.23) they are often not available at the desired wavelength. Moreover, their output powers are often limited to the range of a few milliwatts to a few tens of milliwatts.

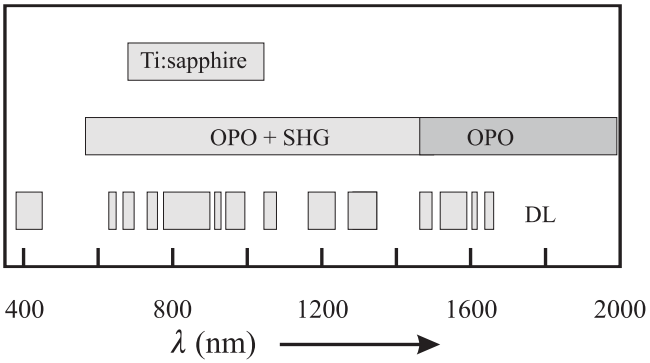


Figure 9.23: Widely tuneable solid-state lasers. Wavelength range of diode lasers (from [430]), tuning range of a Ti:sapphire laser. OPO: signal beam from an optical parametric oscillator based on a Nd:YAG pumped periodically poled $LiNbO_3$ crystal. OPO + SHG: emission range of an OPO whose idler beam is frequency doubled in an external resonator [431]. DL: diode lasers.

In the near infrared the titanium sapphire laser (see Section 11.5.1) allows one to reach any wavelength between about 0.7 µm and 1 µm with high power of up to a few watts. Other wavelength regions can be accessed by frequency conversion in non-linear crystals. In the longer wavelength region, e.g., a commercially available optical parametric oscillator based on a Nd:YAG laser at 1.06 µm and periodically poled $LiNbO_3$ crystals allows one to reach the region between 1.45 µm and 2 µm with the signal beam and between 2.4 µm and about 4 µm with the idler beam. To obtain tuneable radiation with large power in the visible in particular in the yellow and green spectral range where diode lasers are missing (Fig. 9.23) dye lasers are sometimes the only choice (Fig. 9.24). In the following the properties of these different lasers will be discussed to the extent relevant to optical frequency standards.

9.3.1 Dye Lasers

Continuous-wave dye lasers are versatile tools that allow one to produce coherent electromagnetic radiation from the ultraviolet to the infrared region without any gap (Fig. 9.24) and with large power. Their use is restricted to laboratory standards because of the required expensive pump lasers, their large technical noise and the required dye circulators. In the following we concentrate on the properties of dye lasers relevant to optical frequency standards; a more detailed description can be found in [432].

Dye lasers make use of organic molecules solved in organic solvents, e.g., ethylene glycole. The simplified energy level scheme of a dye molecule (Fig. 9.25) comprises a singlet electronic ground state, a singlet excited state and a triplet excited state, each one with a broad range of ro-vibrational levels. Owing to the strong interaction of the dye molecules with the molecules of the solvent the ro-vibrational states are strongly broadened by collisions. Hence, the fluorescence lines overlap and lead to a homogeneously broadened emission continuum. The excitation energy is supplied by optical pumping predominantly using high-power ion lasers in the UV or green spectral range or by frequency-doubled solid state lasers around 0.53 µm similarly to the Nd:YAG laser. The dye molecule is optically pumped from the $^1S_0$ ground state to the $^1S_1$ state (Fig. 9.25) from where it decays in less than 10^{-12} seconds to the lowest vibrational level.

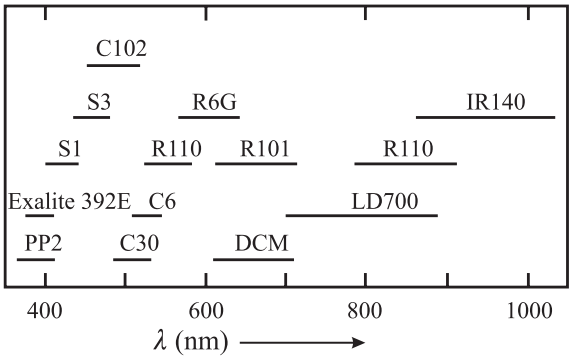


Figure 9.24: Tuning range of laser dyes covering the visible and near infrared part of the electromagnetic spectrum.

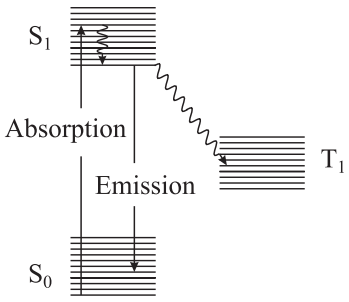


Figure 9.25: Simplified energy level scheme of dyes.

This level represents the starting point for laser emission to the ro-vibrational states of the electronic $^1S_0$ state. As a result of collisions with the molecules of the solvent radiation-less transitions into the triplet system can occur. The molecules in the long-lived triplet system do not contribute to the laser emission. To keep the number of these molecules small the dye solution is pumped with 0.4 MPa to 1.5 MPa through a nozzle to produce a dye jet with rectangular cross-section of about 0.2 mm – 1 mm width and 3 mm – 5 mm length. The dye molecules spend about 1 µs within the waist (≈ 10 µm) of the beam of the pumping laser.

The thickness fluctuations of the dye jet inside the laser resonator give rise to frequency fluctuations (Fig. 9.1). The mechanical resonances of the dye jet dominate the power spectral density of frequency fluctuations for Fourier frequencies below a few megahertz before rolling

off towards the white noise due to spontaneous emission. As a result, a servo bandwidth of a few megahertz is necessary to suppress the frequency fluctuations effectively.

Owing to the broad gain profile in a dye laser a large number of modes can be excited. Dye lasers can be easily tuned across $\Delta\lambda \approx 30\,\text{nm}$ or more (Fig. 9.24). The corresponding frequency width of $\Delta\nu = |c/\lambda^2|\Delta\lambda = 36\,\text{THz}$ at 500 nm is broad enough to sustain 120 000 modes of a laser with a perimeter of 1 m. To achieve single-mode operation wavelength selective elements are required which exhibit broad tuneability and low losses and which do not change the light path when the laser frequency is changed. In particular, interferometric optical devices fulfil these requirements. Coarse tuning is often achieved with a birefringent filter (Lyot filter) [5] with a free spectral range of each plate

$$\Delta\nu_{\text{Lyot,i}} = \frac{c}{(n_o - n_e)D_i} \tag{9.29}$$

of a few gigahertz depending on the thickness D_i and the index of refraction n_o and n_e of the ordinary and extraordinary beam. A further frequency selective element often used in a tuneable laser is the etalon where a thin glass plate with partially reflecting surfaces acts as a multiple beam interferometer. The free spectral range of the etalon

$$\Delta\nu_{\text{etalon}} = \frac{mc}{2D\sqrt{n^2 - \sin^2\alpha}} \tag{9.30}$$

is calculated [409, 433] similarly to that of the Fabry–Pérot interferometer. In (9.30) α is the angle of incidence, m denotes the interference order, D the thickness of the plate and n the index of refraction. Frequency tuning is achieved by tilting the etalon. The selectivity of a thin etalon $D = 1\,\text{mm}$ of glass corresponds to a free spectral range (FSR) of about 100 GHz which is not sufficient yet to achieve single-mode operation of a laser with a FSR $\approx 300\,\text{MHz}$. Hence, another interferometer with FSR $\approx 10\,\text{GHz}$ is necessary. A second etalon with $D \approx 1\,\text{cm}$ can be chosen. However, at a larger angle α, the interfering beams show a shear. This shear leads to an incomplete overlap of the interfering beams thereby reducing the contrast in the interferometer and to a considerable walk-off. Hence, the thicker etalon is sometimes replaced by a Mach-Zehnder interferometer [434] which has lower insertion loss as it can be stabilised to a dark fringe of the interferometer. If all frequency selective elements are properly tuned with respect to each other, single mode operation can be achieved as a consequence of the multiplicative transmissions of the single elements (Fig. 9.26). To counteract the technical frequency fluctuations of the dye laser effectively (see Fig. 9.1) one of the end mirrors of the laser resonator is mounted on a piezo actuator. Piezo elements in general are slow and, hence, are utilised to eliminate frequency fluctuations with large excursions but low Fourier frequencies. As fast servo elements internal-cavity electro-optical modulators are often used that allow variation of the optical path length by applying a voltage (see Section 11.2.2). In order to incorporate these different elements into the resonator of a dye laser a ring laser design with a folded beam path can be used (Fig. 9.27) where only one of the two possible counter-propagating running waves is allowed to oscillate. The other one is suppressed by

[5] The Lyot filter comprises of three birefringent plates of different thickness D_i. It is essentially a two-beam interferometer where the impinging beam is split into two different polarisation components which are subsequently recombined. Rotating the plates allows one to vary the relative path lengths of the two beams and thereby to achieve the typical cosine interference pattern of a two-beam interferometer (see e.g. [433]).

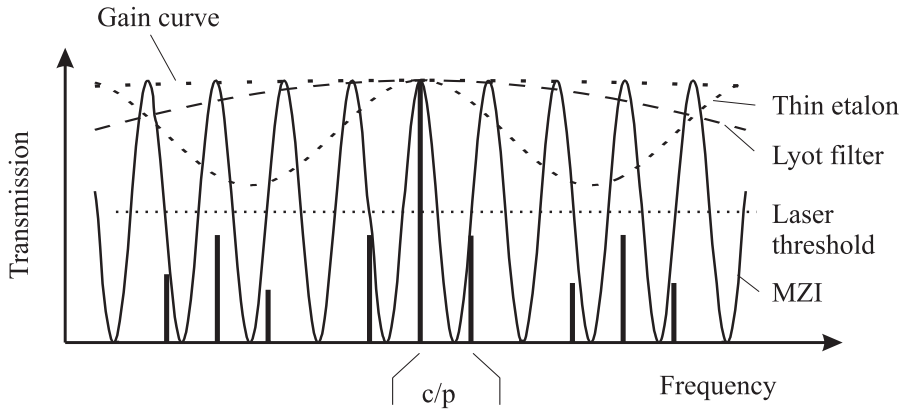


Figure 9.26: Mode selection in a tuneable laser (free spectral ranges of the particular elements to scale). MZI: Mach-Zehnder Interferometer. p: perimeter of the laser. c/p: free spectral range of the laser.

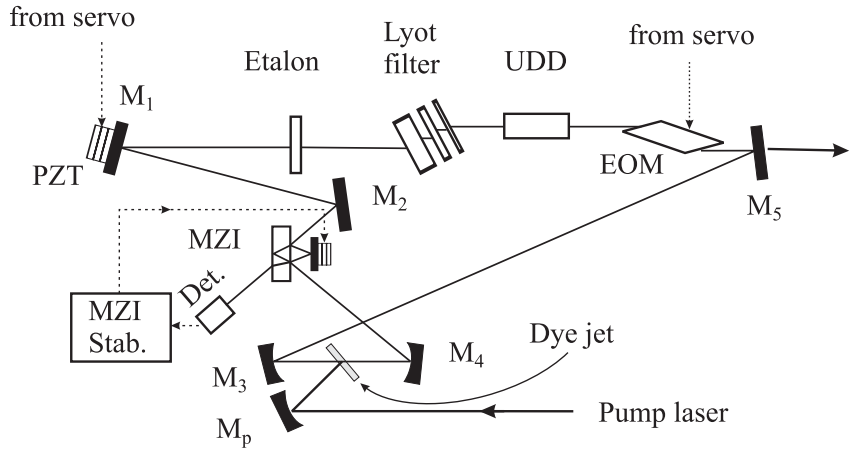


Figure 9.27: Ring dye laser. M_1 to M_5: Mirrors. M_p: Pump mirror. MZI: Mach-Zehnder interferometer. PZT: Piezo actuator. UDD: Unidirectional device. EOM: Electro-optic modulator.

use of a unidirectional device ([433], see also Section 9.4.1.1) that rotates the polarisation only for this direction. The wave travelling along this direction suffers from high losses at any polarisation discriminating element in the laser resonator, e.g., the Brewster angled dye jet. With only one travelling wave present in the laser resonator spatial hole burning in the active medium is avoided. The reflection angles at the folding spherical mirrors M_3 and M_4 are chosen such that the associated astigmatism compensates for the astigmatism introduced by the Brewster angled dye jet.

To narrow the linewidth of a free-running dye laser and to pre-stabilise its frequency, fast and efficient stabilisation schemes to a reference frequency of a Fabry–Pérot interferometer are

in current use. Among them are the side-lock technique (Section 2.3.2.1, [412]), polarisation spectroscopic methods (Section 9.2.1, [408]) and phase modulation spectroscopic methods (Section 9.2.2 [82, 410]). Despite the large free-running linewidth of about a megahertz or higher, reduction of the linewidth of a dye laser to below 1 Hz have been achieved [31].

9.3.2 Diode Lasers

9.3.2.1 Basics of Diode Lasers

In a diode laser the active medium is a semiconductor where the p-n junction emits electromagnetic radiation if excited by an electric current. Semiconductors are solid-state materials where the valence band is filled with electrons and the conduction band is empty at zero temperature. In contrast to isolators, the energy width of the gap between these bands is about 1 eV and, hence, at finite temperatures some electrons are thermally activated into the valence band. The Fermi energy, i.e. the energy level separating the filled energy levels from the empty ones, is located in a semiconducting material in the gap half-way between the conduction and the valence band. If the semiconductor is heavily doped with positive donor ions (p-type material) or negative acceptor ions (n-type) there are holes in the conduction band or electrons in the valence band and the Fermi energy is shifted into the valence band or in the conduction band, respectively.

In a p-n junction, i.e. in the contact region of a p-type and an n-type material, the Fermi energies level out. In a forward-biased diode where a voltage U is applied, the Fermi energies of both materials are shifted by the energy $e\,U$ and the electrons in the conduction band as well as the holes in the valence band are swept into a spatially confined area (Fig. 9.28). In this region of the p-n junction the population of the electrons is inverted and consequently the electrons may recombine with the holes by emission of photons.

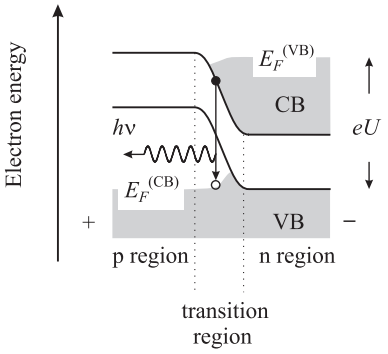


Figure 9.28: In a forward-biased p-n junction the electrons and holes are swept into the transition region leading to inversion.

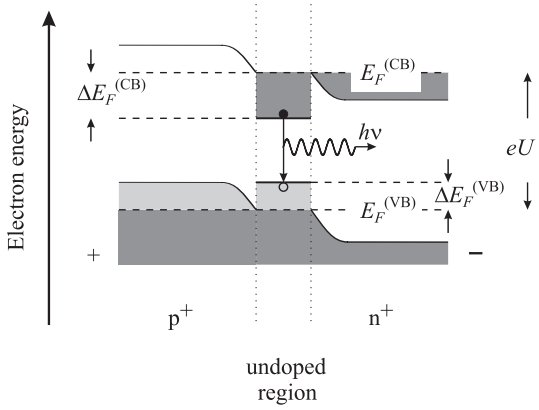


Figure 9.29: Energy bands of a forward-biased double-heterostructure laser diode. The p^+ and n^+ material could be e.g. $Ga_{1-y}Al_yAs$ and $Ga_{1-x}Al_xAs$, respectively, whereas the active region of a width of about 0.1 µm consists of undoped GaAs.

The energy gap and hence the energy of the emitted photons depends largely on the crystal structure of the semiconductor material and can be tailored over a large range by suitably choosing the composition as, e.g. in $In_xGa_{1-x}As$ or in $InAs_{1-x}\ P_x$. If the gain medium, i.e. the p-n junction, is put into an optical resonator and if the charge carrier density is high enough, laser action can occur. The simplest arrangement referred to as a Fabry–Pérot-type laser makes use of the cleaved surfaces of the semiconductor crystal as mirrors. Due to the high index of refraction $3.5 \leq n \leq 4$ the reflectance R due to the Fresnel equations [409] is

$$R = \left(\frac{n-1}{n+1}\right)^2 = \frac{(3.5-1)^2}{(3.5+1)^2} \approx 30\%, \tag{9.31}$$

which is sufficient to achieve laser oscillation thanks to the large gain in the active medium.

To achieve the carrier density necessary to allow for laser activity at not too high injection currents it is necessary to keep the range where the carriers recombine as small as possible. Thus the laser diode is designed in such a way that the laser beam is emitted by a wave guide where only the lowest transverse mode is sustained. The active layer thickness (height) of a p-n junction results from the diffusion zone of about $1\,\mu\text{m} \leq d \leq 2\,\mu\text{m}$. Its width is adjusted by suitable means from about $1\,\mu\text{m}$ to about $100\,\mu\text{m}$ (broad area diode lasers). As a result of the diffraction the laser beam originating from such a small area has a large divergence (see (4.119)) with opening angles θ of several tens of degrees. The large divergence of the laser field can be reduced by a lens of short focal length close to the front facet of the diode laser. Depending on the means taken to confine the wave-guide, laser diodes are grouped into gain-guided or index-guided laser diodes. In both types the electromagnetic wave is guided in the vertical direction by total reflection as a result of the index of refraction of the thin active layer, which is higher than the one in the adjacent material. In the gain-guided diode laser the horizontal guiding of the wave results from the geometry of the contact electrode leading to a confinement of the current flow with the resulting thermally induced gradient of the index of refraction. In contrast, index guided laser diodes contain a horizontal profile of the index of refraction that confines the electromagnetic wave inside the laser. The length L of the diode laser is typically 0.3 mm $\leq L \leq$ 0.5 mm. After one round trip inside the laser resonator the phase of the electromagnetic wave is changed by

$$\phi_a = \omega t = 2\pi\nu_a\, n(\nu)\frac{2L}{c}, \tag{9.32}$$

which depends on the frequency dependent index of refraction $n(\nu)$. Using

$$\frac{d\phi_a}{d\nu} = 2\pi\, n(\nu)\frac{2L}{c} + 2\pi\nu_a \frac{dn(\nu)}{d\nu_a}\,\frac{2L}{c} = 2\pi\frac{2nL}{c}\left(1 + \frac{\nu_a}{n}\frac{dn}{d\nu}\right) \approx \frac{\Delta\phi_a}{\Delta\nu} \tag{9.33}$$

the free spectral range (FSR) of the laser diode can be calculated as the frequency difference $\Delta\nu$ where $\Delta\phi = 2\pi$ as follows

$$\text{FSR} = \Delta\nu(2\pi) = \frac{c}{2\, nL(1 + \frac{\nu}{n}\,\frac{dn}{d\nu})}. \tag{9.34}$$

With typical values of L = 0.3 mm, $(\nu/n)(\ dn/d\nu) \approx 1.5$ and n = 3.5 for GaAs, the free spectral range becomes FSR $\approx$ 57 GHz corresponding to a wavelength separation of

$\Delta\lambda = \lambda\Delta\nu/\nu \approx 0.2$ nm of two adjacent longitudinal modes. Depending on the widths of the conduction band and the valence band the emission spectrum of a laser diode can be as wide as a few tens of nanometres. In general, as a consequence of their design, gain-guided diode lasers operate in a multi-longitudinal mode (Fig. 9.30) whereas index guided diode lasers show a single mode behaviour at high currents (Fig. 9.31).

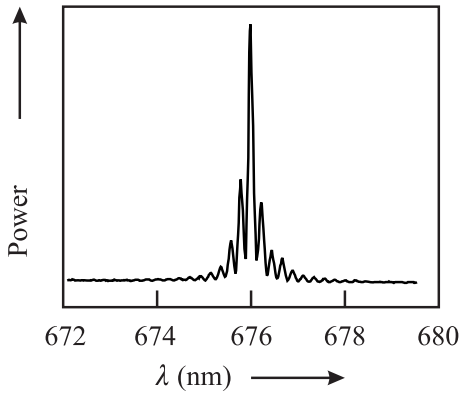


Figure 9.30: Spectrum of a gain-guided solitary laser diode.

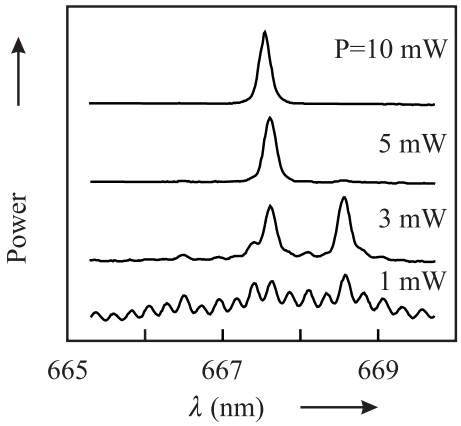


Figure 9.31: Spectrum of an index-guided solitary laser diode.

The typical transverse multi-mode structure of gain-guided laser diodes leads to an asymmetrical beam profile in the far field of the laser. For laser frequency standards, index-guided laser diodes are preferred. If they are not available, other means have to be applied to allow for single mode operation.

9.3.2.2 Noise in Diode Lasers

Frequency Noise In contrast to the majority of gas lasers and solid-state lasers the linewidth of a solitary diode laser is determined by the quantum process of spontaneous emission. Each photon spontaneously emitted into the laser mode can be multiplied by stimulated emission and the resulting field amplitude adds to the field in the diode laser. On the other hand, the small line-quality factor Q of the laser resonator does not sustain a very well defined phase of the internal field. As a result, the phase of the combined field shows considerable fluctuations owing to the statistically fluctuating contributions of the spontaneously emitted photons leading to a large Schawlow–Townes linewidth (see (3.71)). Hence, in a Fourier frequency regime where technical noise is not dominant the noise floor in diode lasers is large compared with other lasers as can be seen by comparing Fig. 9.1 and Fig. 9.32.

In diode lasers the spontaneous emission leads to an additional fluctuation of the index of refraction

$$\Delta n = \Delta n' + i\Delta n'' \tag{9.35}$$

within the duration of the relaxation oscillation, i.e. during ≈ 1 ns. Here, n' describes the dispersion and n'' the absorption. The fluctuation of $\Delta n''$ also leads to a fluctuating gain since

$\Delta n''$ is a result of the fluctuation of the density of the charge carriers ΔN. A fluctuation of the charge carrier density ΔN also leads to a fluctuation of the dispersive part of the index of refraction n' thereby leading to a fluctuation of the phase of the laser wave. In essence, there is a coupling between the phase and the amplitude of the light wave expressed by Henry's coupling parameter [435]

$$\alpha \equiv \frac{\Delta n'}{\Delta n''} = 2k \frac{dn/dN}{dg/dN} \tag{9.36}$$

which leads to an excess line broadening with respect to the Schawlow–Townes linewidth. In (9.36) k is the modulus of the wave vector, n the index of refraction, g the gain factor, and N the density of carriers. α depends on the material of the diode laser. For GaAs and $\lambda \approx 850\,\text{nm}$, $\alpha \approx 4$ has been determined.

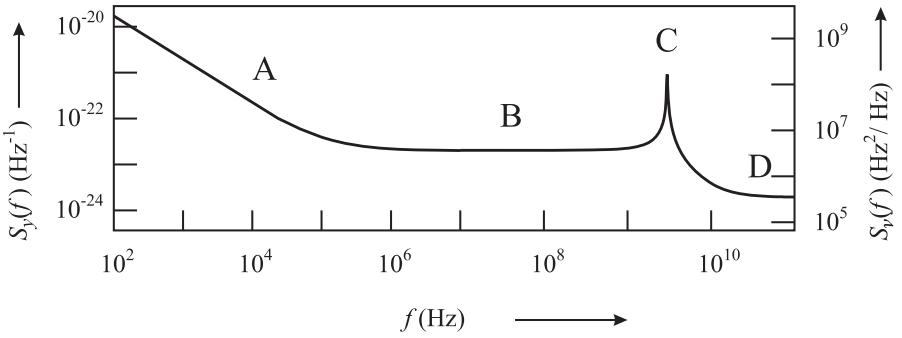


Figure 9.32: Typical power spectral densities $S_y(f)$ and $S_\nu(f)$ of a diode laser near 850 nm as function of the Fourier frequency f according to [436] with different regimes. A: $1/f$ fluctuations. B: fluctuations of the carrier density. C: relaxation oscillations. D: fluctuation of the spontaneous emission.

Starting with a thermodynamic model of the "phase diffusion" the emission line is found to be Lorentzian with a width (FWHM) given by the so-called "modified Schawlow–Townes linewidth" as follows [39]

$$\Delta\nu_{\text{LD}} = \frac{h\nu_0\mu}{2\pi\tau_P^2 P}\left(1+\alpha^2\right) = \frac{2\pi h\nu_0(\Delta\nu_{1/2})^2\mu}{P}\left(1+\alpha^2\right). \tag{9.37}$$

Here, P is the output power, $\mu \equiv N_2/(N_2 - N_1)$ is the parameter describing the inversion, τ_p is the lifetime of the photons in the passive resonator and $\Delta\nu_{1/2}$ is the half-width of the passive resonator.

In a semiconductor laser the coupling between amplitude and phase of the electromagnetic field is much more pronounced compared with other lasers since here the spectral profile of the gain, and hence that of the index of refraction, are asymmetrical with respect to the frequency of the laser. Fluctuations of the gain induced by the emission of spontaneous photons lead to fluctuations of the index of refraction and, hence, of the laser phase. The resulting excess phase noise accounts for the modification of the Schawlow–Townes linewidth (3.71) by the factor $(1+\alpha^2)$ in (9.37).

Intensity Noise In contrast to the phase, the amplitude of the electromagnetic field in a laser diode is much less subject to noise as the latter one is stabilised by the gain saturation. Hence, the intensity noise of a laser diode is extremely small for a wide range of Fourier frequencies with the exception of the regime near the relaxation frequency. A typical value of the relative spectral intensity noise density is about $10^{-6}\,\mathrm{Hz}^{-1}$ [436].

9.3.2.3 Frequency Stability and Tuning of Diode Lasers

To find the frequency of the m^{th} longitudinal mode the phase shifts occurring at the laser facets have to be taken into account. Often, the facet at the rear of the diode laser is coated as a mirror of high reflectivity and the laser field can be thought of as a standing wave with a node at the facet. This is obviously not true for the other facet that serves as an output coupler with a typical coefficient of reflection of $R \approx 35\ \%$. Consequently, there is a phase shift φ of the wave internally reflected from the output facet that is equivalent to an additional optical path length

$$m\,\lambda = 2n(\nu)\,L + \varphi\,\frac{\lambda}{2\pi} = m\frac{c}{\nu} \tag{9.38}$$

or

$$\nu_m = \frac{m\,c}{2\,n(\nu)L + \frac{\varphi \cdot c}{2\pi\nu}}. \tag{9.39}$$

Hence, the frequency ν of the emitted wave depends on the number m of the longitudinal modes, on the phase shift φ due to reflection, on the length $L = L(T)$ of the laser crystal and on the index of refraction n. In the case of small variations of these parameters one can assume [37]

$$\frac{\Delta\nu}{\nu} = \frac{\Delta m\,\mathrm{FSR}}{\nu} - \frac{\Delta\varphi\,\mathrm{FSR}}{2\pi\nu} - \frac{\Delta L}{L} - \frac{\Delta n}{n} \tag{9.40}$$

with the free spectral range FSR determined according to (9.34). The first contribution of (9.40) can lead to mode jumps of about 100 GHz (Fig. 9.33). The phase φ in the second term of (9.40) can be varied in particular by coupling back a part of the light emitted by the diode laser. This effect can be used for frequency stabilisation of diode lasers. On the other hand, spurious radiation reflected back, e.g., by the window of the housing of the diode laser, by the collimating lens or from other optical components, can alter the frequency of the laser, in particular if the phase of the back-reflected light fluctuates.

The last two contributions of (9.40) are influenced by the temperature of the laser. For small temperature variations the length $L(T)$ of the laser diode is expected to vary linearly with the temperature according to (4.128). The index of refraction n influences the laser frequency in a complicated way since it varies with the frequency ν, temperature T, the injection current I and the laser power P ($n = n(\nu, T, I, P)$). Temperature fluctuations affect the index of refraction via different effects [37, 39, 430, 436]. In general, raising the temperature of a laser diode increases its wavelength where the monotonic variation is interrupted by discontinuous jumps (Fig. 9.33).

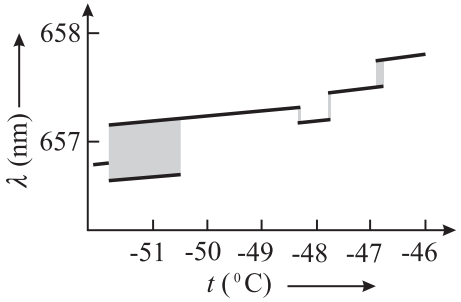


Figure 9.33: Tuning of the wavelength of a InGaAlP laser diode with injection current. The grey areas indicate regions of multi-mode operation.

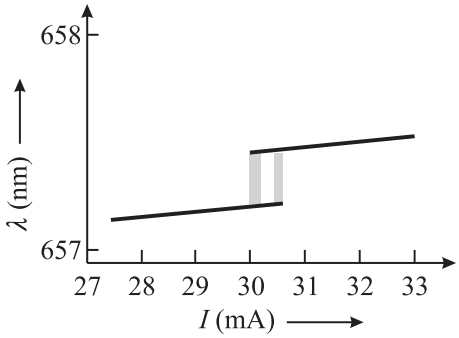


Figure 9.34: Tuning of the wavelength of a InGaAlP laser diode with temperature. The grey areas indicate regions of multi-mode operation.

The monotonous variation with a typical value of about -30 GHz/K is due to the length variation and the associated shift of the mode frequency. At the same time the temperature-dependent lattice constants of crystal and the associated variation of the band structure result in a shift of the gain profile of the laser. As a result, mode jumps of about 50 GHz – 100 GHz or more occur and the mean wavelength variation with temperature over a wide temperature range amounts to about -100 GHz/K. The wavelength, however, is not an unambiguous function of the temperature but rather shows hysteresis effects dependence whether the temperature is raised or lowered.

Besides the ambient temperature, the temperature of the diode laser is also affected by the injection current. With a nearly constant voltage drop across the p-n junction the dissipated power, and hence the temperature increase, is proportional to the injection current. Consequently, a smooth increase of the injection current results in a red detuning of the frequency of the laser due to the associated temperature variation. The shift of the frequency of a solitary diode laser as a function of the injection current varies from -5 GHz/mA to 1 GHz/mA for $(Al_xGa_{1-x})_yIn_{1-y}P$ to $Ga_xIn_{1-x}P_yAs_{1-y}$ for wavelengths between 635 nm and 1.5 µm, respectively [430]. To achieve good long-term frequency stability these values ask for a low-noise power supply with low ripple. A list of practical precautions and design criteria has been given by Fox *et al.* [430].

Variation in the current, in general, also affects the index of refraction by the changed number of the free charge carriers. For larger variations of the injection current and higher modulation frequencies this influence of the current prevails, which is of particular importance when the injection current is to be used as a fast input for frequency stabilisation [37, 437]. The current-to-frequency transfer function rapidly goes down for modulation frequencies high enough that the thermal effect dies out. For frequencies above 1 GHz the current to frequency transfer function often displays a resonance-like feature resulting from the relaxation oscillations.

The mechanism leading to the relaxation oscillations can be understood as follows. Consider a temporal fluctuation of the number of photons above the equilibrium ($n(t) > n_0$)

induced by spontaneous emission. As a consequence of the stimulated emission the inversion will be reduced below the equilibrium value ($N(t) < N_0$), provided there is a constant pump rate. The reduced number of emitted photons ($n(t) < n_0$) associated with the reduced inversion will eventually tend to increase the inversion above the equilibrium again ($N(t) > N_0$). Any delay between $N(t)$ and $n(t)$ then may lead to so-called relaxation oscillations.

For frequencies close to the relaxation oscillation small perturbations in $N(t)$ or $n(t)$ are amplified and consequently their influences can be seen as well in the spectral density of the power fluctuations as in the spectral density of the frequency fluctuations (Fig. 9.32).

There are methods for fast frequency modulation or frequency control of diode lasers that avoid the phase delays and amplitude modulation associated with modulation of the bias current such as the use of an intra-cavity electro-optic modulator [438] or the injection of "control light" [439]. If the wavelength of the control laser diode is tuned close to the transparency region of the laser diode to be controlled, i.e. to a wavelength between the two regions of absorption and stimulated emission, the index of refraction of the laser under control can be modulated without amplitude modulation.

9.3.2.4 Linewidth Reduction by Optical Feedback

The linewidth of a solitary diode laser of the Fabry–Pérot type can be as high as tens to hundreds of megahertz. Such a linewidth cannot be reduced by negative electronic feedback alone since the required large servo bandwidth can hardly be achieved. The limited frequency range and delay times of currently available electronic components and the phase shifts of the signals in the electronic feedback loop make it necessary to reduce the linewidth first by other than electronic means. These difficulties can be overcome by employing an optical feedback loop. It has been shown by Velichansky *et al.* [440] and by Fleming and Mooradian [441] that feedback from an external mirror can be used to change the linewidth of a laser diode. Consider a solitary diode laser where part of the emitted radiation is coupled back into the diode laser (Fig. 9.35). In this case, the electromagnetic field in the laser is composed of a

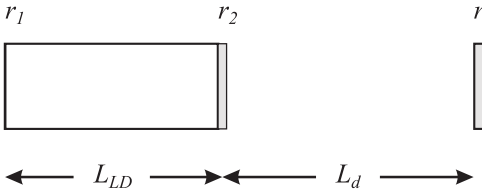


Figure 9.35: Diode laser with feedback from an external reflector referred to as an external cavity laser. For $r_2 \ll 1$ the device becomes an extended cavity diode laser.

superposition of the internal field and the back-coupled field. Depending on the phase between both fields the amplitude of the resulting field is increased or reduced. Due to the strong coupling between the phase and the amplitude of the diode laser a sudden power fluctuation in the active medium will lead to a fluctuation of the phase of the field in the diode laser. Consequently, the field coupled back after an external round trip time $\tau_d = 2L_d/c$ also suffers from a phase shift. Depending on the phase difference between the internal and the back-coupled field, the initial perturbation will be increased or reduced.

In the latter case, the influence of the internal field at an instant $t - \tau_L$, on the field at an instant t, leads to reduced frequency fluctuations by the so-called self-injection locking. As a consequence, the frequency of the laser is very sensitive to the phase of the back-reflected radiation. A small fraction

$$\beta \equiv \frac{P_R}{P_l} \tag{9.41}$$

of, say, $\beta \leq 10^{-6}$ of the power P_l of the laser coupled back into the laser is sufficient to influence the frequency of the diode laser considerably. For a more specific characterisation of the feedback process not only the fraction β of the power of the laser diode that is coupled back, but also the ratio of the round-trip times $\tau_d/\tau_{\mathrm{LD}}$ need to be taken into account where τ_{LD} is the round-trip time inside the laser diode.

The feedback is characterised by the parameter

$$C = \frac{\tau_d}{\tau_{\mathrm{LD}}} \frac{1 - r_2^2}{r_2} r_3 \sqrt{1 + \alpha^2} \tag{9.42}$$

where β is approximated by the (amplitude) reflection coefficients r_1 and r_2 (see Fig. 9.35). Different regimes have been identified [442, 443]. In *Regime I* ($C \ll 1$) the spectrum of the diode laser is stable. The attainable linewidth reduction or broadening is not very high and depends critically on the phase of the reflected radiation field. The radiation reflected by the window of the diode laser or from a glass plate close to the front facet of the diode laser (0.1 mm $\leq L \leq$ 0.5 mm) can be used for mode selection. In *Regime II* ($C \approx 1$) the laser likes to jump between several modes of the external resonator depending on the phase. In *Regime III* ($C > 1$) with weak feedback the laser operates stably on a single mode with reduced linewidth and independent of the external feedback phase. Increasing C, e.g., by increasing the distance between the external reflector or by increasing its reflectivity (*Regime IV*; $C \gg 1$) leads to a "coherence collapse" initiated by frequent mode jumps in the nanosecond regime. The associated linewidth of several gigahertz is a result of the mode partition noise. For even stronger feedback ($C \gg 1$) *Regime V* is again a stable one. Here, the linewidth is determined by the external resonator and is small. Often, the portion of the light fed back into the laser diode is frequency narrowed by, e.g., a diffraction grating or a Fabry–Pérot interferometer.

Depending on the elements and configurations used to accomplish the feedback, several terms have been coined which, however, are used differently in the literature. The laser diode by itself without any extra elements is referred to as "solitary diode laser". If an external reflective element is added to the solitary diode laser the combined arrangement is termed "External Cavity Laser" (ECL) since an external cavity is formed by the reflective element and the output coupler of the solitary laser. If the output coupler has a low reflectivity, e.g., from an anti-reflective coating, the laser cavity is formed by the rear mirror of the laser diode and the external reflector and, hence, acts as an "Extended Cavity Diode Laser" (ECDL).[6]

The various aspects of feedback from external elements with different feedback levels have been treated in the literature based on different theoretical methods [444–447]. One approach [447] considers the influence of the external elements, e.g., mirror, grating, interferometer or

[6] Sometimes, however, the acronym ECDL is also used for an external cavity laser and sometimes the extended cavity diode laser is abbreviated as XCDL.

atomic ensemble as a modification of the complex reflectivity of the front facet of the diode laser

$$r_{\text{eff}} = r(\omega)e^{i\Phi_r(t)}. \tag{9.43}$$

The modulus $r(\omega)$ takes into account the modified fraction of the internal field that is reflected at the front facet owing to the contribution reflected from the external components. The phase shift $\Phi_r(t)$ depends on the roundtrip time of the backcoupled light in the external extension. The ansatz (9.43) allows one to modify the rate equations for the solitary diode laser. Kazarinov and Henry [445] derived the linewidth reduction of a Lorentzian line in the presence of optical feedback as

$$\Delta\nu = \frac{\Delta\nu_0}{(1+A+B)^2} \tag{9.44}$$

where $\Delta\nu_0$ is the Lorentzian linewidth without feedback. The factors A and B

$$A - i\frac{B}{\alpha} = \frac{1}{i\tau_{\text{LD}}}\frac{d(\ln r_{\text{eff}})}{d\omega} = \frac{1}{\tau_{\text{LD}}}\frac{d\Phi_r}{d\omega} - \frac{i}{\tau_{\text{LD}}}\frac{d(\ln r(\omega))}{d\omega} \tag{9.45}$$

are determined by the increase in roundtrip phase in the laser diode with optical angular frequency ω and by the increase in reflectivity of the mirror of the laser diode cavity with ω, respectively.

As an example of how to determine the effective reflectivity (9.43) of a particular arrangement, consider the case of the effective reflectivity of the front facet of the solitary laser diode with a plain external mirror which is given [444, 445] [7] as

$$r_{\text{eff}} \quad = \quad \frac{r_2 + r_3(\omega)e^{i\omega\tau_d}}{1 + r_2 r_3(\omega)e^{i\omega\tau_d}}. \tag{9.46}$$

In (9.46) τ_d is the external roundtrip time. (9.46) can be calculated in a similar way to (4.92). The difference in the signs in (9.46) and (4.92) reflects the fact that the reflectivity of the input mirror of the Fabry–Pérot resonator is reduced at resonance, whereas in the case of the external mirror, the effective reflectivity of the output facet of the laser diode is increased.

Rather than using a plain mirror commercially available "External Cavity Diode Lasers" often comprise a diffraction grating placed at a distance of a few centimetres from the laser diode. The frequency stability of this type of laser is determined by the complicated interplay of the three different resonators formed by the front facet and the rear facet of the laser diode, the rear facet and the external reflector, and the front facet and the external reflector. Higher frequency and mode stability is achieved when the spectrally purified reflected light from an external Fabry–Pérot interferometer is coupled back into the solitary laser diode with low $\beta \ll 0.01$. The "Extended Cavity Diode Laser" comprises a laser diode with good antireflection coating on the front facet. The front facet reflector is replaced by an external grating or other reflector which couples back a large fraction $0.1 < \beta \lesssim 0.8$ of the power. The last two schemes will be discussed in more detail in the following.

[7] The authors use different phase conventions. Here we follow the one used in [445] and modify the signs of the reflection coefficients accordingly.

9.3.2.5 Diode Laser with Extended Cavity

Diode lasers used in optical frequency standards are often operated in *Regime V* where $R_2 \ll \beta$ holds, e.g., by placing a mirror at a distance L behind the output facet of the laser diode of length L_{LD}. In general, only a fraction of the power reflected by the external mirror is fed back into the active zone of the laser diode and the strong feedback *Regime V* can be accessed only when the reflectivity of the front facet is reduced by an anti-reflective (AR) coating. State-of-the-art AR coatings can achieve amplitude reflection coefficients $r_2 < 10^{-2}$. In this case the influence of the residual reflection from the front facet can often be neglected and the device is referred to as an Extended Cavity Diode Laser (ECDL).

The cavity length of the extended resonator can be as large as about 30 cm. At such a distance the longitudinal mode separation of about 500 MHz becomes sufficiently small to allow stable single-mode operation in the broad gain profile of the diode laser only at the expense of additional frequency selective elements such as, e.g., etalons, gratings or prisms. The wavelength selected by a diffraction grating can be calculated from the grating equation

$$m\lambda = a(\sin\theta_i + \sin\theta_d) \tag{9.47}$$

depending on the angle of incidence θ_i, the diffraction angle θ_d, the grating constant a and the diffraction order m.

In diode lasers, two particular arrangements are often used. The first one, called the Littrow configuration, employs the reflection grating as the output coupler of the extended cavity. The grating angle is set such that the first-order reflection coincides with the incident beam from the diode laser (Fig. 9.36). In this case $\theta_i = \theta_d = \theta$ holds. The zero-order reflection is used to couple out the output beam and the wavelength of the diode laser is adjusted by a rotation of the grating.

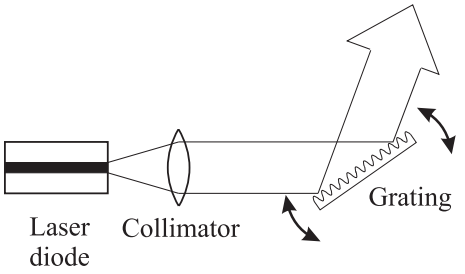


Figure 9.36: Extended-cavity diode laser using a grating as output coupler in Littrow configuration. Wavelength selection is achieved by rotating the grating.

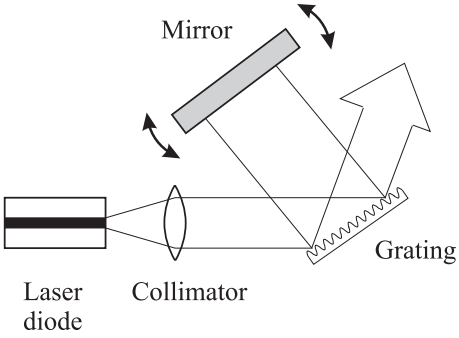


Figure 9.37: Extended-cavity diode laser with an intra-cavity grating in Littman configuration. Wavelength selection is achieved by rotating the mirror.

The Littman configuration [448] (Fig. 9.37) uses a folded laser cavity. In contrast to the Littrow configuration the incident beam and the diffracted beam are no longer collinear. The diffracted beam is reflected back from a mirror and is directed into the laser diode after a second diffraction at the grating. Tuning of the wavelength is achieved by rotating the mirror.

Again, the zero-order beam is used to couple out a fraction of the power circulating in the cavity. By comparing both configurations [449], the Littman configuration has the advantage that the tuning does not change the direction of the output beam. The double diffraction, however, leads to increased intra-cavity losses and requires gratings of high reflectivity. The double pass, on the other hand, leads to an increased selectivity. Another advantageous feature of the Littman configuration is the free choice of the angle of incidence independent of the wavelength. Consequently, this configuration allows one to use a large angle of incidence independent thereby illuminating a large number of grooves of the grating with the associated better resolution.

For an extended cavity diode laser with grating, the effective reflectivity (9.46) is modified by the frequency dependent reflectivity r_3 of the grating which has been given [444] as

$$r_3 = r_0 \exp\left[-\left(\frac{N_{\text{eff}}}{4}\right)^2 (\omega\rho - 2\pi m)^2\right]. \tag{9.48}$$

In (9.48) r_0 is the reflectivity of the grating with grating constant a which is illuminated by a Gaussian beam of a full spot size $2b$ (1/e irradiance) under the angle θ_i and

$$\rho = \frac{2a}{c} \sin\theta_i \tag{9.49}$$

and the number of illuminated lines of the grating

$$N_{\text{eff}} = \frac{2b}{a\cos\theta_i}. \tag{9.50}$$

Combining (9.46) and (9.48) leads to the effective reflectivity

$$r_{\text{eff}} = \frac{r_2 + r_0 \exp\left[-\left(\frac{N_{\text{eff}}}{4}\right)^2 (\omega\rho - 2\pi m)^2\right] e^{i\omega\tau_d}}{1 + r_2 r_0 \exp\left[-\left(\frac{N_{\text{eff}}}{4}\right)^2 (\omega\rho - 2\pi m)^2\right] e^{i\omega\tau_d}} \tag{9.51}$$

which can be used to calculate the minimum achievable linewidth with the help of (9.44) and (9.45) as

$$\Delta\nu = \frac{\Delta\nu_{\text{LD}}}{\left[1 + (\tau_d/\tau_{\text{LD}})\right]^2} = \frac{\Delta\nu_{\text{LD}}}{\left[1 + (L_d/nL_{\text{LD}})\right]^2} \tag{9.52}$$

where n is the index of refraction of the laser diode. The reduced noise in a laser with extended cavity can be seen, e.g., from Fig. 3.10. The power spectral density of frequency fluctuations S_ν of a solitary diode laser decreases for Fourier frequencies f below about 80 kHz faster than $1/f$ and shows white noise for higher frequencies. With grating the corner frequency is above 200 kHz and $S_\nu(f)$ is reduced by about 33 dB. Note the increase of the spectral noise density near $f \approx 1$ kHz which was attributed [40] to acoustical vibrations changing the length of the extended cavity. In general, laser with extended cavities are more susceptible to external perturbations. The reduced spectral density of frequency fluctuations also shows up as a considerably reduced linewidth of the ECDL (Fig. 9.38) as compared to the solitary laser. On the

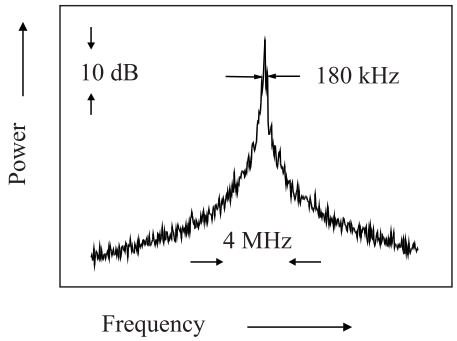


Figure 9.38: The beat note between a free-running extended cavity diode laser and a frequency stabilised dye laser with linewidth <1 kHz essentially shows the linewidth of the extended cavity diode laser. Measuring time: 2 s

other hand, in the ECDL set-up the desired reduction of frequency fluctuations is accompanied by a reduced sensitivity of the injection current when used as a servo actuator. Continuously tuning without mode hops over extended portions of the spectrum can be achieved when the cavity length and the grating or mirror are tuned synchronously [448, 450, 451].

Owing to the losses due to the various elements in the laser cavity, the power of frequency-narrowed diode lasers is often too low to be used directly in optical frequency standards. To achieve higher power often a second laser is used which is either injection locked [452] or used as an amplifier. Broad area laser diodes [438] or tapered amplifiers [191] are readily used for this purpose.

9.3.2.6 Optical Feedback from a Fabry–Pérot Interferometer

Consider a weak field of high spectral purity which is coupled back into the laser diode. This field can be produced by filtering the laser field by an optical resonator of reasonably high finesse. Conveniently, a confocal Fabry–Pérot interferometer (FPI) is used [446, 453] which allows one to mode match the field of the diode laser to the confocal FPI. In Fig. 9.39 a confocal FPI is used whose optical axis is tilted with respect to the axis of the FPI. The FPI acts as a V-shaped degenerate three-mirror cavity. With the tilted FPI, the light reflected in a frequency independent manner at the front mirror of the FPI (direction A) is no longer coupled back into the diode laser. The field fed back into the diode laser has been circulated inside the FPI and, hence, is spectrally filtered. To stay in the regime of weak coupling it is sufficient to couple back $10^{-8} < \beta < 10^{-4}$ of the laser power into the diode laser.

The effective reflectivity of the front facet of the laser again can be calculated from (9.46) using the modulus $r_3(\omega)$ and the phase of the complex reflection coefficient of the Fabry–Pérot interferometer (4.93) [447]. For the case when the frequency of the laser corresponds to an eigenfrequency of the confocal FPI ($\omega = \omega_q$) and the distance from the laser diode to the confocal FPI is approximately an integer times half the wavelength, the Lorentzian linewidth is reduced maximally [447] to

$$\Delta\nu = \frac{\Delta\nu_0}{(1+\alpha^2)\beta\left(\frac{F^*_{\mathrm{FPI}}L_{\mathrm{FPI}}}{F^*_{\mathrm{LD}}nl_{\mathrm{LD}}}\right)^2} \tag{9.53}$$

where F^*_{FPI} and F^*_{LD} is the finesse (see (4.89)) of the FPI and the solitary laser diode, respectively. From (9.53) one finds that apart from the coupling coefficient β and Henry's parameter

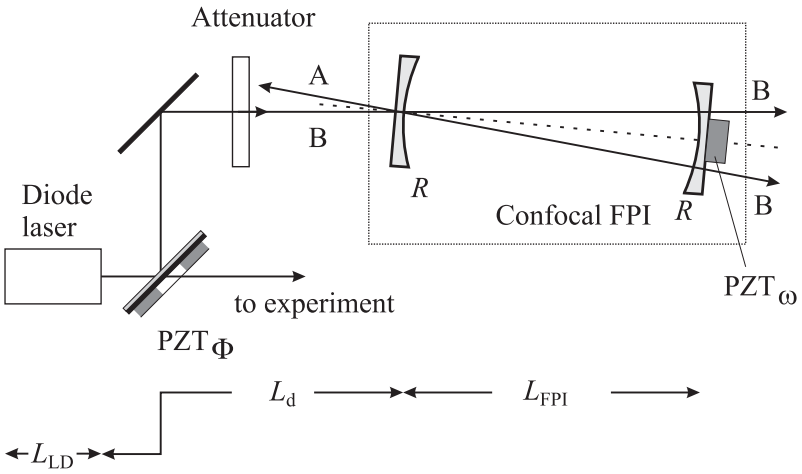


Figure 9.39: Optical feedback from a high-finesse confocal Fabry–Pérot interferometer (FPI) according to reference [453].

α the linewidth reduction is given by the the ratio of the product of the finesse and the optical length of the FPI and the laser diode. For typical values of $F^*_{\mathrm{FPI}} = 100$, $F_{\mathrm{LD}} = 2$, $L_{\mathrm{FPI}} = 20\,\mathrm{cm}$, $nL_{\mathrm{LD}} = 1\,\mathrm{mm}$ and $\beta = 10^{-3}$ the linewidth reduction is calculated from (9.53) as 10^{-8}, i.e., one would expect a linewidth of a few hertz. In practice, however, the minimal achievable linewidth was measured around a few kilohertz [446] which can be attributed to the influence of technical $1/f$ noise at Fourier frequencies below 1 MHz.

Laurent *et al.* [446] gave the steady-state solution of the rate equations describing the angular frequency ω_N of the laser with feedback from the confocal FPI as a function of the angular frequency of the laser ω without feedback as

$$\omega_N = \omega + K\frac{\sin[\omega(\tau_d + \tau_{\mathrm{FPI}}) + \Theta] - R^2\sin[\omega(\tau_d - \tau_{\mathrm{FPI}}) + \Theta]}{1 + F^2\sin^2\omega\tau_{\mathrm{FPI}}}. \tag{9.54}$$

The different quantities used in (9.54) are

$$\tau_d = \frac{2L_d}{c} \quad \text{and} \quad \tau_{\mathrm{FPI}} = \frac{2L_{\mathrm{FPI}}}{c}, \tag{9.55}$$

$$F = \frac{2R}{1-R^2}, \tag{9.56}$$

$$\Theta = \arctan(\alpha), \tag{9.57}$$

$$K = \sqrt{1+\alpha^2}\frac{c}{2nL_{\mathrm{LD}}}\sqrt{\beta}\,\frac{1-r_0^2}{r_0}\,r\,\frac{1-r^2}{(1-r^4)^2} \tag{9.58}$$

with τ_d and τ_{FPI} the roundtrip times, $r_0 = r_1 = r_2$ the amplitude reflection factors of the laser facets, r is the amplitude reflection factor of the mirrors of the confocal FPI. $\mathrm{FSR}_{\mathrm{LD}} = c/(2nL_{\mathrm{LD}})$ is the free spectral range, i.e., the mode separation of the solitary laser diode and β the power feedback coupling factor. Fig. 9.40 shows by using (9.54) how under the

influence of the feedback from the confocal FPI the frequency of the laser is "locked" to the eigenfrequencies of the FPI over large detunings $\omega - \omega_N$. This scheme has been used to set

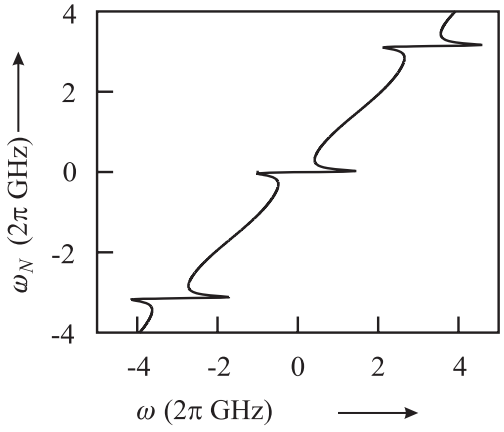


Figure 9.40: Angular frequency ω_N of the laser with feedback versus angular frequency ω of the free-running laser according to (9.54) with $\alpha = 5$, $\mathrm{FSR_{LD}}$ = 90 GHz, $r_0^2 = 0.4$, $\beta = -40$ dB, $r^2 = 0.97$, $c/(4L_{\mathrm{FPI}})$ = 0.5 GHz and $L_d/L_{\mathrm{FPI}} = 3$.

up pre-stabilised lasers with large tuning ranges as narrow band sources for high-resolution spectroscopy [454] which in turn could be frequency stabilised to molecular transitions [455].

9.3.3 Optical Parametric Oscillators

In cases where suitable absorption lines are not accessible by tuneable lasers as, e.g., in the infrared or ultraviolet, continuous-wave coherent radiation can also be generated by utilising non-linear interactions in suitable materials (Section 11.1.3). The development of powerful and easy-to operate solid state lasers like the Nd:YAG laser and its frequency-doubled version together with quasi-phase matched materials (see Table 9.3) makes optical parametric oscillators [456] available for tuneable sources of coherent radiation. The optical parametric process between the strong electric field of the pump laser and the non-linear material converts a photon from the pump field of energy $\hbar\omega_p$ into two photons referred to as the signal photon $\hbar\omega_s$ and the idler $\hbar\omega_i$. In the process energy and momentum are conserved

$$\hbar\omega_p = \hbar\omega_s + \hbar\omega_i \tag{9.59}$$

$$\vec{k}_p = \vec{k}_s + \vec{k}_i. \tag{9.60}$$

The latter condition (9.60) requires matching the phase of the respective waves which can be achieved by different means (Section 11.1.3). In order to generate the high fields necessary to make efficient use of the non-linear process, optical parametric oscillators employ a single, double or triple resonant cavity for one or more of the pump, signal, or idler beams. Tuning of an OPO can be achieved by rotating the crystal or by temperature tuning. In the case of periodically poled materials, the former procedure is not possible. In this case gratings with several periods are often implemented into the same substrate where each sub-device can

Table 9.3: Properties of phase-matching materials for optical parametric oscillators for the fundamental and frequency doubled wavelength of a Nd:YAG laser

Material	Transmission range (μm)	Phase matching range 0.532 μm	1.064 μm
BBO	0.19 – 2.56	0.67 – 2.5	
LBO	0.16 – 2.6	0.67 – 2.5	
$KNbO_3$	0.35 – 4.2	0.61 – 4.2	1.43 – 4.2
KTP	0.35 – 4.0	0.61 – 4.0	1.45 – 4.0
$LiNbO_3$	0.35 – 4.3	0.61 – 4.3	1.42 – 4.3
$AgGaS_2$	0.8 – 9	1.2 – 9.0	2.6 – 9.0

be temperature tuned over a few ten gigahertz and the tuning ranges overlap. An extended OPO cavity with an intra-cavity etalon and periodically poled $LiNbO_3$ with 33 gratings has been used [457] to derive Doppler-free absorption lines of the $F_2^{(2)}$ component of the P7 rovibrational transition in CH_4 with a linewidth of 100 kHz.

9.4 Optical Standards Based on Neutral Absorbers

9.4.1 Frequency Stabilised Nd:YAG Laser

Favourable properties such as high power, compact size and intrinsically high stability make the Nd:YAG laser a prominent oscillator for use in optical frequency standards. The laser action results from Nd^{3+} ions present in the cubic host crystal of Yttrium-Aluminium Garnet ($Y_3Al_5O_{12}$ known as YAG). In the Nd:YAG laser about 1% of the Y^{3+} ions are replaced by Nd^{3+} ions. The most versatile laser transition at 1.064 μm (Fig. 9.41) is part of a four-level laser system. The radiation of a diode laser at, e.g., 0.81 μm pumps the Nd^{3+} ion from the $^4I_{9/2}$ ground state to the $^4F_{5/2}$,$^2H_{9/2}$ pump bands from where the ion rapidly decays by radiationless transitions to the $^4F_{3/2}$ state. Since electric dipole transitions to the lower states are forbidden this upper laser state is long lived with a lifetime of 0.24 μs. The lower laser state ($^4I_{11/2}$) is rapidly emptied by radiationless transitions to the ground state. Near room temperature the lasing lines in the YAG are homogeneously broadened by lattice oscillations. The linewidth of about 100 GHz near room temperature is small in comparison with other solid state lasers, resulting in high gain for small pump power. This property together with the low threshold of the four-level system allows one to build lasers of considerable power but moderate size. The large numbers of host crystals and doping ions available give some flexibility for accessing (after frequency doubling) the transitions of atoms, ions (see e.g. [458]) and molecules which is highly interesting for optical frequency standards.

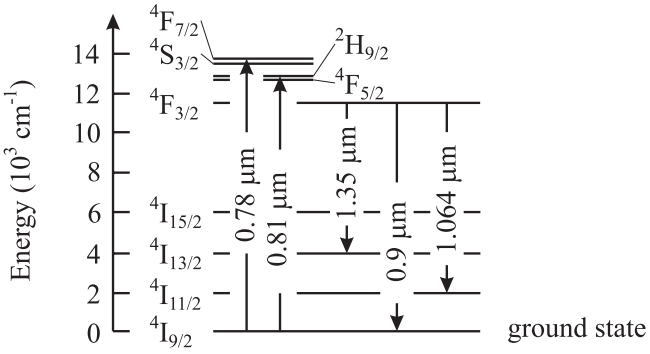


Figure 9.41: Simplified energy level scheme of Nd^{3+} ions in a Nd:YAG laser crystal with some absorption and laser wavelengths.

9.4.1.1 Monolithic Ring Laser

Solid-state laser media allow the construction of monolithic lasers where the reflecting surfaces of the resonator are directly attached to the lasing material with exceptional immunity to acoustic perturbations. The simplest linear resonator would make use of two polished and coated reflecting end surfaces. However, in a standing-wave resonator the stability is often degraded by mode competition and the output power is limited by spatial hole burning. Spatial hole burning occurs in a standing wave where stimulated emission is suppressed at the nodes of the standing wave with vanishing field amplitude. A ring laser with unidirectional and single-frequency operation can avoid these disadvantages. Monolithic ring lasers have been developed by Kane and Byer [459] in a non-planar ring configuration (Fig. 9.42) or as quasi-planar ring design [460]. To obtain unidirectional operation in a ring laser a combi-

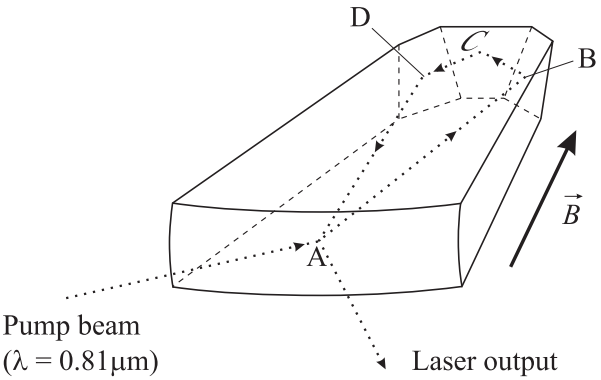


Figure 9.42: Monolithic non-planar Nd:YAG ring laser.

nation of elements is used that rotate the polarisation of the two counter-propagating waves differently in a non-reciprocal way. Consequently, the two counter-propagating waves suffer from different losses at any polarising element in the ring. Such a non-reciprocal device can

be constructed by a combination of a Faraday rotator [8] and a plate exhibiting natural optical activity, e.g., a quartz plate. The combination of both elements is chosen such that for the desired direction of the light wave the rotation in the first element is reversed by the second one. In the opposite direction the two polarisation rotations add up which leads to an increased round-trip loss at a suitable polarising element. In the non-planar ring laser of Fig. 9.42 the crystal itself has a positive Verdet constant. The internal reflections B and D at the inclined faces rotate the polarisation [459]. The reflection at the output coupler A is different for the horizontal and vertical polarisation direction and, hence, has polarising properties.

9.4.1.2 Atomic and Molecular Absorbers

The advantages of this kind of Nd:YAG laser such as low noise, compact size and high laser efficiency, make it an excellent oscillator for an optical frequency standard. Unfortunately, at the prominent wavelength of 1.064 µm there are only a few absorption lines among which are molecular $^{133}Cs_2$ [461–463], acetylene C_2H_2 [124], deuterated acetylene C_2HD [418], and carbon dioxide CO_2 [464]. The transition in molecular Cs is an electronic one whereas the others are due to ro-vibrational overtones and, hence, their line strengths are about nine orders of magnitude weaker as compared to the one in the Cs molecule. The Cs_2 stabilised Nd:YAG laser is not widely used, mainly since the Cs_2 molecular transition is subject to a large temperature induced frequency shift [462]. To utilise the weak overtone transitions in optical frequency standards the absorbers are often exposed to the increased irradiance in a resonant cavity (see Section 9.4.2).

9.4.1.3 Iodine Stabilised Frequency-doubled Nd:YAG Laser

In contrast to the fundamental frequency, the frequency-doubled radiation of the Nd:YAG laser at 532 nm falls into a regime where the molecular iodine spectrum has suitable absorption lines (Fig. 5.5). The frequencies of a variety of transitions have been determined accurately and are recommended for the realisation of the metre (see [370]).

Several methods have been used to stabilise the frequency of a frequency-doubled Nd:YAG laser to a suitable absorption line. A design from the Novosibirsk group [420] employed a method where the modulation of the fluorescence from the iodine cell was detected which was induced by a frequency modulated laser beam. A modulation frequency of 455 Hz with an amplitude of about 500 kHz was used and a third-harmonic detection scheme was employed. The laser captivates by its simple set-up, a good short-term stability with a minimum of the Allan deviation between two identical systems of $\sigma_y(\tau = 300\,\mathrm{s}) \approx 5 \times 10^{-14}$ and a relative uncertainty of 2×10^{-12}.

Using the method of modulation transfer an Allan deviation has been reported as low as $\sigma_y(\tau) = 5 \times 10^{-15}$ at τ = 1000 s [465]. Extensive comparisons of the frequencies of the iodine-stabilised frequency-doubled Nd:YAG lasers have been performed (see e.g. [419, 420, 466]). Hong *et al.* [466] reported a frequency reproducibility of this type of laser of about 1×10^{-13}. However, the frequency difference between different lasers could be as large

[8] A Faraday rotator comprises a material of length L in a magnetic field of induction B whose direction is parallel to the direction of the light. The polarisation vector is rotated by an angle $\alpha = VBL$ where the Verdet constant V is characteristic of the material.

as $2 - 5$ kHz. Nevsky *et al.* [420] stated uncertainties of their lasers of 2 kHz and 1.1 kHz, corresponding to 3.5×10^{-12} and 2×10^{-12}, respectively. Ye *et al.* [467] even recommend such a laser as an optical clock with a measured instability of 4.6×10^{-13} over one year.

9.4.2 Molecular Overtone Stabilised Lasers

Molecular ro-vibronic transitions are located predominantly in the infrared part of the electromagnetic spectrum. Hence, their richness of transitions is not directly accessible for optical frequency standards in the visible. Radiation in the visible or near infrared, however, is capable of exciting so-called molecular overtones where two or more different vibrational or rotational quanta are transferred to the molecule (Section 5.2.3). These transitions exhibit the same small linewidths in the kilohertz regime as their fundamentals but suffer from low dipole moments and the associated weak absorption. Hence, highly sensitive spectroscopic methods have to be applied to utilise these transitions. One method increases the absorption length by placing the absorber material into an optical cavity with high finesse [468]. Furthermore, highly sensitive phase modulation techniques could be applied where the phase shift of the carrier in the vicinity of the molecular resonance is probed, by comparison with the less affected side bands (Section 9.2.3). The combination of both techniques, however, meets with difficulties. To allow a sufficient side band power through the cavity, the modulation frequency has to be comparable to the linewidth of the cavity. The low modulation frequency leads to a low bandwidth of the servo system for the frequency stabilisation. At the same time the error signal is detected at a low frequency where considerable technical noise of the laser may be present. Furthermore, even though the carrier is kept near the maximum of the transmission curve of the resonator the sidebands are located in its wings. Any frequency fluctuations will affect the amplitudes of the sidebands and frequency noise will be converted into amplitude noise, hence limiting the achievable signal-to-noise of the molecular signal. This noise can be effectively reduced if the laser beam is modulated with a frequency that matches the free spectral range of the resonator. In this case the carrier as well as the sidebands can be kept at the centres of the respective transmission maxima leading to a largely reduced frequency-to-noise conversion. This so-called Noise-Immune Cavity-Enhanced Optical Heterodyne Molecular Spectroscopy (NICEOHMS, [469]) has been applied to weak overtones of, e.g., C_2H_2 [470] (see Fig. 9.43), C_2HD [469], CH_4 [471] or to O_2 [472]. The instability of an acetylene stabilised laser was determined to be $\sigma_y(\tau) = 4.5 \times 10^{-11}\sqrt{\mathrm{s}/\tau}$ for measurement times $\tau < 1000$ s [418]. Such standards find applications as wavelength references for optical communications [473] (see Section 13.1.4.1).

9.4.3 Two-photon Stabilised Rb Standard

The frequency of a laser stabilised to the two-photon transition $5S_{1/2} - 5D_{5/2}$ in rubidium (Fig. 9.44) has recently been recommended by the CIPM for the realisation of the length unit [370]. This reference frequency in the near infrared at 778 nm wavelength [474] makes use of easy-to-handle laser diodes of low frequency noise allowing the development of a transportable optical frequency standard of high precision. Different realisations of this optical frequency standard have been investigated. In a simple set-up, the collimated beam of an extended cavity diode laser passed through an absorption cell filled with rubidium vapour.

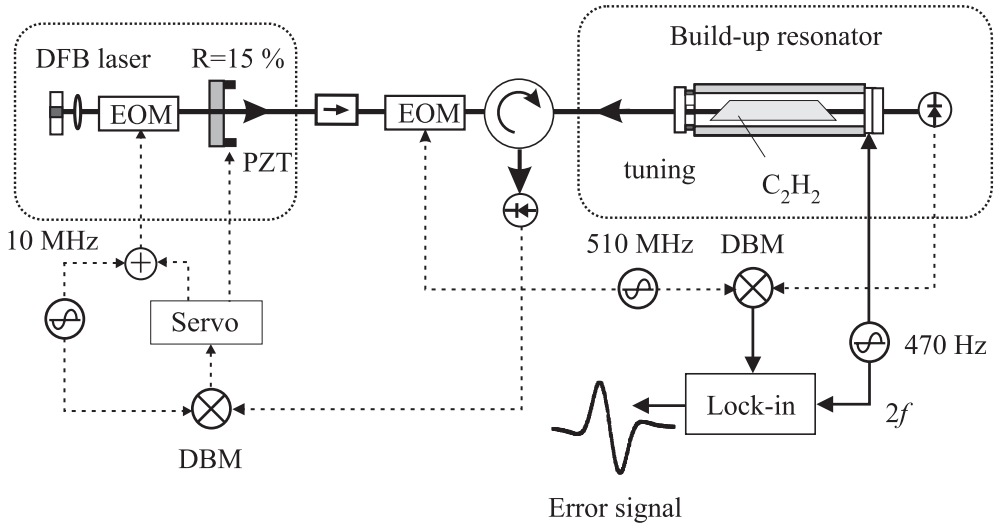


Figure 9.43: Experimental set-up of an acetylene stabilised laser using overtone spectroscopy [470]. Courtesy of U. Sterr.

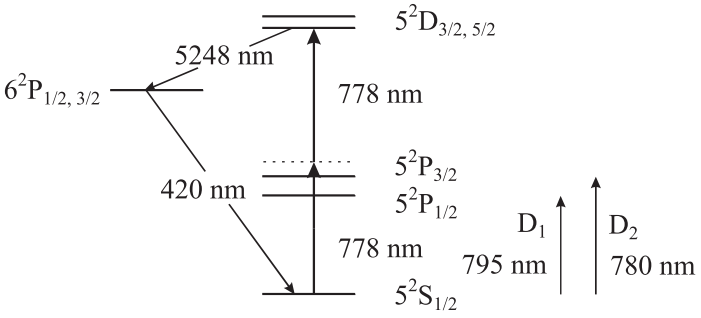


Figure 9.44: Partial atomic energy diagram of rubidium including the two-photon transition at 778 nm.

The cell was sealed with Brewster windows at its ends and filled with rubidium of natural abundance (73% ^{85}Rb and 27% ^{87}Rb). The beam was retro-reflected by a mirror or a cat's eye in order to achieve Doppler-free two-photon excitation. Optical feedback into the diode laser was avoided by Faraday isolators. When the laser frequency was scanned through resonance, the two-photon transition was observed via the blue fluorescence (420 nm) of the 6P – 5S transition in the cascade of the spontaneous decay 5D $\rightarrow$ 6P $\rightarrow$ 5S. With this set-up, a frequency stability of $\sigma_y(\tau = 2000\,\mathrm{s}) = 2 \times 10^{-14}$ has been achieved. Another design [475] operated the absorption cell inside a non-degenerate optical resonator which allowed a power build-up of the radiation and therefore an increase in the two-photon signal. Furthermore, this arrangement can provide exact retro-reflection of the laser beam which is necessary to suppress residual first-order Doppler shifts. It is known that the stabilised frequencies of

standards based on two-photon transitions suffer from light shifts whose magnitudes depend linearly on the irradiance (see (5.137)). Therefore, it is important to control the laser power and to prepare a well defined laser beam geometry for the excitation. In the set-up of [475] the laser frequency was pre-stabilised in a first step to a resonance of the cavity. Such a laser can then be tuned through the two-photon resonance if the length of the resonator is changed by applying voltage to the PZT transducer to which one of the mirrors is mounted. Compared to the results of the simple set-up described before, the observed frequency stability was approximately the same. However, the light shift could be better controlled and consequently the frequency of the transition could be extrapolated more precisely to zero laser power. The two-photon Rb standard provides a precise frequency reference for optical communication systems since the sub-harmonic coincides with the transmission band at 1.55 µm. Moreover, this laser has played an important role as an intermediate standard for the synthesis of optical frequencies [476] and as a reference for the frequency determination of the 2S – 8S/8D two-photon transition in atomic hydrogen and the determination of the Rydberg constant [103]. Further improvements may be achieved by laser cooling the rubidium atoms [474] and the same two-photon transition might then be used as an optical frequency standard of high but not ultra-high accuracy.

A two-photon 6S – 8S transition in caesium [477] has also been measured, although the natural linewidth of 1.5 MHz is a factor of three larger as compared to the corresponding one in rubidium.

9.4.4 Optical Frequency Standards Using Alkaline Earth Atoms

The intercombination transitions of the alkaline earth atoms have been long recognised to represent excellent references for optical frequency standards (see, e.g., [99] and references therein). For example, Mg, Ca and Sr (see Table 5.2 and Fig. 5.2) exhibit narrow natural linewidths of about 0.035 kHz, 0.37 kHz, and 6 kHz, respectively. Furthermore, the frequencies of the $\Delta m_J = 0$ transitions in these elements are very little perturbed by magnetic and electric fields. In the case of Ca the dependencies are $\Delta\nu/\nu = 1.3 \times 10^{-13}\,(\mathrm{mT})^{-2} \times B^2$ [191, 478] and $\Delta\nu/\nu = 5.4 \times 10^{-17}\,(\mathrm{V/cm})^{-2} \times E^2$ [479], respectively. The intercombination lines in an effusive beam have been investigated in magnesium [244], calcium, strontium [96, 480, 481], and barium [482].

9.4.4.1 Ca Beam Standard

Most work for optical frequency standards using alkaline earth elements has been performed with Ca. Already around 1980 Barger *et al.* obtained a resolution as low as 1 kHz [237, 238]. Frequency standards based on an effusive beam have been set up, e.g., at the Physikalisch-Technische Bundesanstalt (PTB) [483,484], at the National Research Laboratory of Metrology (now NMIJ) [485, 486] and at the National Institute of Standards and Technology (NIST) [191, 487].

A transportable standard based on an effusive beam (Fig. 9.45) has been developed at the PTB [484] that has been used to compare standards at the PTB in Braunschweig and NIST in Boulder.

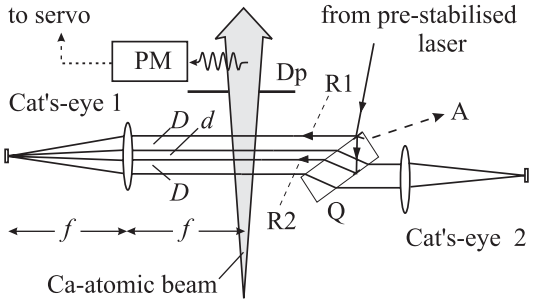


Figure 9.45: Optical frequency standard based on an effusive Ca beam [484]. Dp: Diaphragm; Q: Quartz plate; PM: Photomultiplier.

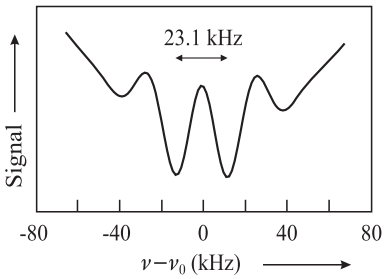


Figure 9.46: Optical Ramsey resonances (atom interferences) obtained with the set-up of Fig. 9.45 [484].

The set-up (Fig. 9.45) comprised an Extended Cavity Diode Laser (ECDL) system whose frequency was pre-stabilised to a Fabry–Pérot interferometer (FPI) by means of the Pound–Drever–Hall technique (Section 9.2.2). A fraction of the power from the laser of about two milliwatts was sent to a beam splitter / mirror configuration and was split into two beams R1 and R2 of equal power. From each one of the beams (R1 or R2) crossing the atomic beam perpendicularly, an excitation geometry with two pairs of counter-propagating laser beams was obtained by the help of two cat's-eye retro-reflectors. With one of the two directions (R1 or R2) blocked, the atoms excited in the four-beam geometry were detected by a photomultiplier that measured the fluorescence light associated with the decay of the excited atoms in the $^3P_1$ state. The detected fluorescence of the $\delta m_j = 0$ transition versus angular frequency detuning $\Delta\omega = \omega - \omega_0$ of the laser from the Ca intercombination transition (Fig. 9.46) can be described [215, 483] by

$$I(\Delta\omega) \propto \int_0^\infty A(P,v,\Delta\omega) f(v) \left\{ \cos\left[2T\left(\Delta\omega + \delta_{\rm rec} + \frac{\omega_0 v^2}{2}\right) + \Delta\Phi_L\right] + \cos\left[2T\left(\Delta\omega - \delta_{\rm rec} + \frac{\omega_0 v^2}{2}\right) + \Delta\Phi_L\right]\right\} dv + B(P,v,\Delta\omega). \quad (9.61)$$

(9.61) can be derived from the evolution matrices of the pseudo-spin vectors [215] (Section 5.3.1) or in the frame work of atom interferometry (Section 6.6.1). In (9.61), $A(P,v,\delta\omega)$ represents the contribution of a particular atom with velocity v to the signal, and $B(P,v,\delta\omega)$ takes into account the amplitude of the background of the Doppler broadened line including the saturation dip, both depending on the laser power P and weakly on the detuning $\Delta\omega$. The factor $f(v)$ represents the velocity distribution. According to (9.61) each velocity group v contributes to the signal with two cosine functions. The phases of the cosine functions given in square brackets in (9.61) depend on the detuning and the flight time $T = D/v$ of the atoms between two co-propagating beams. The phases furthermore include three contributions. There is a term $\delta_{\rm rec} = \hbar\vec{k}^2/(2m_{\rm Ca}c^2) = 2\pi \times 11.5\,$kHz resulting from the photon recoil where $\vec{k}$ is the wave vector of the laser field and $m_{\rm Ca}$ is the mass of a Ca atom. The term $\omega_0 v^2/(2c^2)$ is

due to the second-order Doppler shift which is a consequence of the relativistic time dilation.

$$\Delta\Phi_L = \Phi_4 - \Phi_3 + \Phi_2 - \Phi_1 \tag{9.62}$$

(see (6.52)) gives the residual phase transferred to the atomic de Broglie wave by the four exciting laser beams, having the individual phases $\Phi_i, i = 1, \cdots, 4$ in each interaction zone.

The two cosine functions are shifted symmetrically with respect to the resonance frequency and separated by the recoil splitting $2\delta_{\mathrm{rec}} = 2\pi \times 23.1\,\mathrm{kHz}$ with a period $1/(2T)$ depending on the time-of-flight through the distance D. Optimum visibility is expected if the superposition of the two cosines is such that the period is an integer fraction of the recoil splitting. The linewidth of the signal is given by FWHM $\approx 1/(4T)$. The measured fluorescence signal versus detuning (Fig. 9.46) shows the two central minima originating from the two cosine terms separated by the recoil splitting of 23.1 kHz in the saturation dip at the centre of the Doppler broadened line. The Doppler broadening of the signal displayed in Fig. 9.46 of 7.5 MHz was a result of the chosen collimation of the atomic beam. With increasing detuning the cosine structure is rapidly washed out, since all velocity groups v of the atomic beam contribute with a slightly different period (see (9.61)). The FWHM = 16 kHz of the resolved structure was determined by the separation of the beams $D = 10\,\mathrm{mm}$ and a most probable atomic velocity in the beam of $v_{\mathrm{prob}} = 620\,\mathrm{m/s}$.

The pre-stabilised laser with a linewidth of about 2 kHz was long-term stabilised to the central maximum of Fig. 9.46 by a third-harmonic technique using a sinusoidal modulation frequency of 325 Hz and a full width of the frequency modulation of 32 kHz. The stabilised laser had a short term instability of $\sigma_y(\tau) < 10^{-12}$ at an integration time τ = 1 s. The characterisation of the frequency standard showed that the major contributions to the frequency uncertainty were due to a residual first-order Doppler and second-order Doppler effect caused by the high atomic velocities. The influence of the first-order Doppler effect essentially results from misalignments of the four interrogating laser beams and curvatures of the wave fronts at the interaction regions.

From (9.61) it is evident that the corresponding phase shift $\Delta\Phi_L$ (see (9.62)) changes sign when the direction of the laser beams is reversed [483]. In the set-up of Fig. 9.45 the laser beams could be reversed by blocking alternatively either R1 or R2. The corresponding frequency shift allowed one to determine the residual phase shift and to correct its influence.

The uncertainty due to the residual first-order Doppler effect was estimated to be 500 Hz [23]. The large velocity of the atoms led to a considerable second-order Doppler effect whose influence could be corrected by calculating the influence of the velocity distribution $f(v)$ (see (9.61)) of the atoms on the resonance pattern. Rather than using a Maxwellian velocity distribution an effective velocity distribution had to be used that included modifications from several effects, e.g., the velocity dependent excitation and detection efficiency [216] affecting $A(P, v, \omega)$. The effective velocity distribution was determined from a Fourier analysis of the measured signal (Fig. 9.46) similarly to how it was done with Cs atomic clocks (see e.g. [264, 265] and references therein). The relative uncertainty of the transportable standard was estimated to be 1.3×10^{-12} [23]. With stationary standards, a relative uncertainty of 5×10^{-13} was estimated [486] mainly limited by the first-order and second-order Doppler effects associated with the high velocities of the atoms in the beam.

9.4.4.2 Optical Frequency Standard Using Cold Ballistic Alkaline Earth Atoms

Much work has been done to cool and trap alkaline earth atoms for use in frequency standards in order to minimise velocity-induced contributions to the frequency uncertainty. The alkaline earth atoms can be conveniently laser cooled and trapped in a magneto-optical trap (MOT) (Section 6.4.1) via fast transitions connecting the $^1P_1$ states with the $^1S_0$ ground states (Fig. 5.2). The missing splitting of the ground state makes the alkaline earth elements well suited for optical frequency standards but does not easily allow to reach temperatures below the Doppler limit (6.12) of the $^1S_0 \rightarrow ^3P_1$ transition. Magnesium has been cooled [488] and trapped [489] and the optical spectroscopy necessary for an optical frequency standard has been successfully demonstrated [490]. Laser cooled strontium atoms captured and stored [491] in a magneto-optical trap [163, 492] are of particular interest since the intercombination line can be used to further cool the atoms down to the recoil limit [163,493]. The straightforward application of the same method to other alkaline earth elements is hampered by the fact that the narrow linewidth leads to a cooling force that is close to the gravitational force (Ca) or even smaller (Mg). This problem can be overcome by quenching the intercombination transition and Ca has also been "quench-cooled" to temperatures well below the limit imposed by the width of the $^1P_1 - ^1S_0$ transition [164,165,494]. In the following, the progress made with laser-cooled optical frequency standards will be discussed by using the example of ^{40}Ca.

Optical frequency standards with laser-cooled ballistic Ca atoms have been realised with different set-ups [191, 196, 495]. One method used atoms effusing from an oven heated to about 600 °C which were subsequently laser cooled by a counter-propagating laser beam (423 nm) in a Zeeman slower, deflected by a standing laser beam, before being trapped in a magneto-optical trap [196,199]. A more compact apparatus can be set up if the Zeeman slower is omitted (Fig. 9.47, [191,495]). Hence, only the low velocity atoms of the atomic beam can

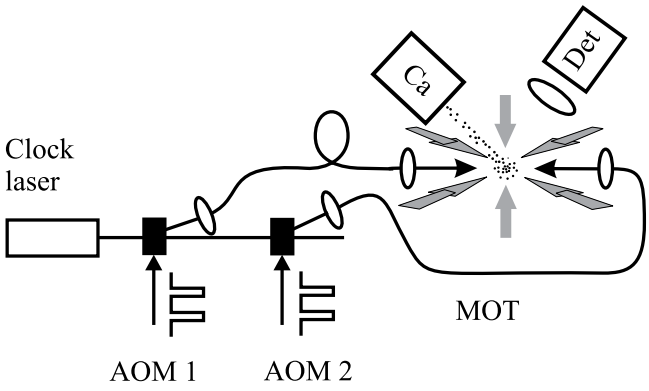


Figure 9.47: Magneto-optical trap (MOT) loaded by the low-velocity tail of the atomic distribution effusing from a calcium oven (Ca) [191, 496]. The acousto-optical modulators (AOM) cut the pulses of about 1 µs used for the excitation from the clock laser beam. Det: Detector.

be captured by the trap. To increase the loading rate the oven has to be placed close to the trap centre. The capture velocity can be increased by using two laser frequencies for the horizontal

trapping beams [190] or by an additional slowing beam and the deceleration within the fringe field of the magneto-optical trap [191]. The radiation at $\lambda = 423\,\text{nm}$ was either produced by a dye laser pumped by an argon UV-laser or later by a solid state laser system consisting of a frequency-doubled diode laser system or by a frequency doubled Ti:Sapphire laser. The traps are loaded typically for about five to twenty milliseconds leading up to 10^8 stored atoms. After this time the light of the trapping laser and (when higher accuracy is needed) also the magnetic field of the MOT, is shut off and the atoms are allowed to expand ballisticly. The intercombination transition at 657 nm of the freely falling atoms is then excited by the radiation of a high-resolution dye laser spectrometer [82] or a diode laser spectrometer [191, 452]. From the width ($\approx 3\,\text{MHz}$) and shape of the Doppler broadened intercombination line at $\lambda =$ 657 nm the temperature of the atomic sample can be determined indicating a temperature of the atoms of 2 mK – 3 mK. The simplest method to detect the percentage of excited atoms is to detect the photons ($\lambda = 657\,\text{nm}$) from the fluorescence decay. In this case, however, each excited atom emits only one photon which is detected in general with a probability around 0.1 % and, hence, the signal-to-noise ratio is limited by the shot noise of these photons. To detect each excited atom with almost unity probability different shelving techniques have been applied to magnesium [489] and calcium [191, 496, 497]. These methods monitor the fluorescence on the strong transition $^1\text{P}_1 - {}^1\text{S}_0$ which is coupled to the $^3\text{P}_1 - {}^1\text{S}_0$ via the common ground state (Fig. 5.2). The fluorescence on the strong transition ($\lambda = 423\,\text{nm}$) decreases at a rate proportional to the number of atoms excited to the $^3\text{P}_1$ state where they are shelved for some time. These methods represent modifications of the electron shelving technique used with single ions (see Section 10.2.3.3, [498]).

The root mean square velocity of the atoms in an ensemble cooled close to the Doppler limit of the $^1\text{P}_1 - {}^1\text{S}_0$ transition is about 1 m/s as can be derived from the Doppler broadening of about 3 MHz (Fig. 9.48) and (first-order) Doppler-free methods have to be applied for the interrogation. Excitation by two counter-propagating pulses leads to a saturation dip in the

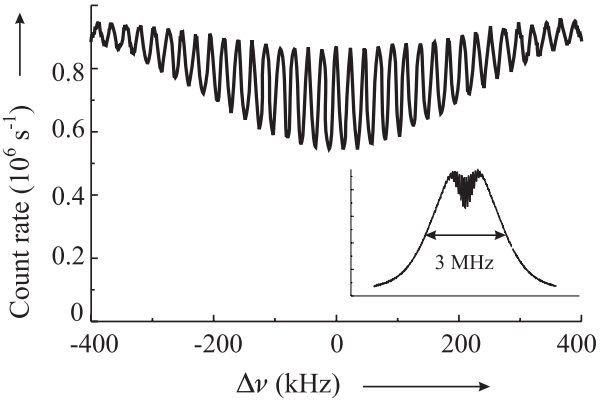


Figure 9.48: Atom interferences in the centre of the Doppler-broadened intercombination line $^1\text{S}_0 - {}^3\text{P}_1$ (see inset) observed with laser-cooled Ca atoms.

centre of the Doppler broadened line whose width depends on the duration of the pulse (Section 5.4.1). In calcium, an interaction time of 0.5 ms corresponding to the natural lifetime of

the excited state, however, results in a linewidth of about 0.37 kHz and consequently only a few times 10^{-4} of the atoms in the Doppler broadened velocity distribution are excited. High spectral resolution combined with a good signal-to-noise ratio (S/N) can be achieved by using a sequence of short pulses to realise the time-domain equivalent of the separated field excitation. The resulting optical Ramsey resonances or Bordé atom interferences (Section 6.6.1) of the number of excited atoms vary proportional to $\cos(2\pi\nu T - 2\pi\nu_0 T)$ (Fig. 9.48). Even during the short interaction time of 1 µs the pulses excite only a fraction of the cold ensemble of atoms (see inset of Fig. 9.48). The necessary high spectral resolution is achieved by a sufficiently large time separation T between the exciting pulses. If the length of the pulse is small compared to their separation, the width of the interference fringes $\Delta\nu = 1/(4T)$ is inversely proportional to T. Depending on the time T the resolution can be increased without changing the interaction time broadening and the signal does not degrade much with increasing resolution. Fringe widths below 300 Hz, i.e. close to the natural linewidth of the clock transition were obtained (see Fig. 6.1 c) and [191]). The interference pattern is comprised of the two contributions of the two recoil components with a frequency difference $2\delta_{\mathrm{rec}} = h/(m_{\mathrm{Ca}}\lambda^2)$ = 23.1 kHz. To obtain maximum contrast the resolution is usually adjusted in such a way that both patterns overlap. The contrast also depends on the kind of excitation, e.g., whether three pulses of a standing wave or two pulses of the laser beam in each direction are employed.

The central fringe is usually chosen to stabilise the frequency of the laser, i.e. to keep the frequency at a minimum or maximum of the fluorescence signal by means of an electronic servo-control system. From the periodicity of the fringes it might seem ambiguous to find the proper frequency of the intercombination transition. There are, however, several means of determining the central fringe. One method relies on the fact that for arbitrary variations of the fringe period only the positions of the minima corresponding to the frequencies at both recoil components are constant.

The performance of optical Calcium frequency standards has been investigated in the groups of the NIST [499] and the PTB [495, 500]. Uncertainty budgets have been given between 1×10^{-13} [500], 6×10^{-14} [501] and 2×10^{-14} [502, 503] with prospects to reach even smaller fractional uncertainties [497]. The frequency of the intercombination transition has been measured over the years with a phase-coherent frequency chain [504] and with a mode-locked femtosecond laser [501, 505] (Section 11). The uncertainty of the frequency of ν_{Ca} = 455 986 240 494 150 (9) Hz [502] makes the Ca stabilised laser one of the most accurate optical frequency standards. The uncertainty eventually achieved with the optical Ca standard might be limited by errors in the phase fronts of the laser pulses exciting the clock transition. Atoms moving perpendicular to the laser beams in curved phase fronts will experience a phase shift between consecutive laser pulses equivalent to a first-order Doppler shift. Since the atoms are accelerated by gravitation between the first and last laser pulse a frequency shift can occur if the laser beams are not exactly levelled horizontally [506]. The application of different types of excitation schemes equivalent to different kinds of atom interferometers has been used to reduce these effects [503, 506]. Frequency shifting radiation fields resulting from incomplete suppression of the cooling lasers or temperature radiation from the Ca oven lead to a so-called ac-Stark shift or black-body shift, respectively. Smaller contributions to the uncertainty include influences of the superposition of the two recoil components, the contribution of Doppler background, and the stabilisation scheme.

The uncertainty of these types of optical frequency standards based on a ballisticly expanding cloud of neutral atoms is ultimately limited by the velocity of the atoms and the associated phase shifts in tilted or curved wave fronts and can be efficiently reduced further by employing atoms of lower velocities. The measured Doppler broadened intercombination transition (Fig. 9.49 a) with ^{40}Ca atoms quench-cooled [507, 508] to the microkelvin range is much less affected by the velocity as, e.g., the one obtained with atoms in the millikelvin range (Fig. 9.48). Hence, the lineshape can be modelled more accurately (Fig. 9.49 b) which is prerequisite to determining the true line centre to a precision much lower than the linewidth.

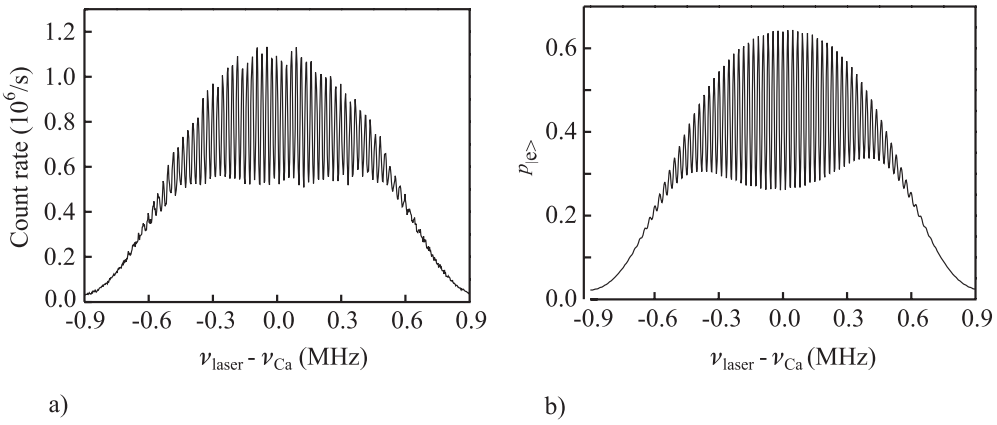


Figure 9.49: a) Atom interferences with four travelling waves using quench-cooled atoms [507]. b) Calculated spectrum according to [215].

The fractional instability of a Ca optical frequency standard was measured to be $\sigma_y(\tau = 1\,\mathrm{s}) = 4 \times 10^{-15}$ [499]. The minimal instability that can be obtained with such an optical neutral-atom frequency standard is given by the quantum projection noise (Section 14.1.3.1). For $N_0 = 1 \times 10^7$ atoms and a time $T = 0.6$ ms corresponding to a linewidth of about 0.6 kHz, one computes from (3.97) $\sigma_y(\tau) < 10^{-16}/\sqrt{\tau/\mathrm{s}}$ [497, 499].

9.4.5 Optical Hydrogen Standard

The natural linewidth around one hertz and the corresponding exceptionally high $Q \approx 10^{15}$ of the hydrogen 1S – 2S two-photon transition (Fig. 5.1, Table 5.2) makes it ideally suited for precision spectroscopy and an optical frequency standard. Based on the results obtained in the group of Th. Hänsch [93, 157, 235, 509] this transition has been recommended for the realisation of the metre [370]. In the set-up of this group (Fig. 9.50) the Doppler-free two-photon transition was excited by frequency doubled radiation from a dye laser operated at 486 nm [510] in an atomic beam. Owing to their small mass the hydrogen atoms in a room-temperature gas have a mean velocity near 2 km/s and the associated large second-order Doppler shift (Fig. 5.19) will eventually limit the achievable accuracy. Hence, a cold beam of atomic hydrogen has been produced in a liquid helium cooled nozzle kept at a temperature

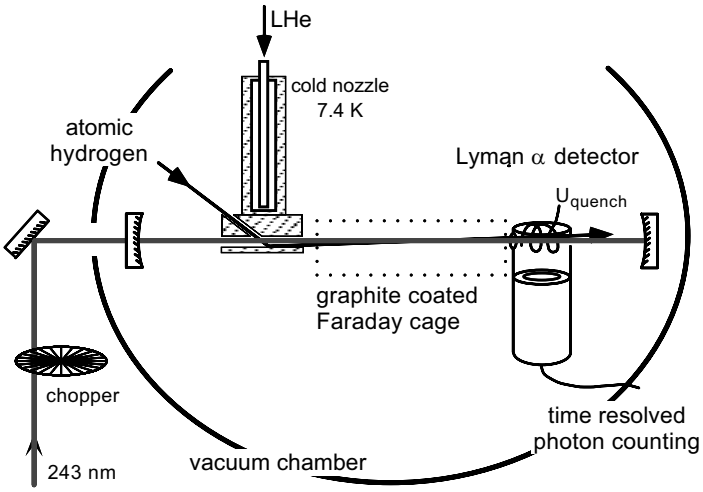


Figure 9.50: Set-up used to measure the frequency of the 1S – 2S hydrogen transition after [157]. LHe: Liquid Helium. Courtesy of Th. Hänsch.

of about 7 K. The atoms in the beam were excited by the collinearly aligned standing wave light of 243 nm resonantly enhanced in a linear cavity in the vacuum. To detect the excited atoms an electric field was applied that mixed the 2S and the 2P state. This field quenched the 121 nm transition whose photons could be detected by a solar-blind photomultiplier during the time the 243 nm radiation was blocked.

An estimation of the uncertainty of the optical hydrogen beam standard has been given with a fractional uncertainty of a few parts in 10^{14} and the frequency has been measured [93] with the same accuracy. Frequency shifting contributions result from the second-order Doppler effect, electric and magnetic fields, light shift, and a pressure shift. To find the line centre despite the asymmetric lineshape resulting from the second-order Doppler effect the experimental lineshape was modelled and compared to the experimental ones [157]. Furthermore, when the slowest atoms in the beam were selected, the second-order Doppler shift could be reduced further. The fluorescence of the excited atoms was monitored after a delay time τ, after blocking the laser light. Hence, the maximum velocity of the atoms is given by $v_{\max} = d/\tau$ where $d \approx 13$ cm is the distance between the nozzle and the detector.

Static electric fields shift the line by $\Delta\nu_{\mathrm{dcStark}} = 3.6\,\mathrm{kHz} \times E^2/(\mathrm{V/m})^2$ as a result of the mixing of the almost degenerate 2P and 2S levels. To reduce the electric field produced by static charges, and hence the dc Stark shift, to a minimum a graphite-coated Faraday cage was employed (Fig. 9.50). The 243 nm radiation is red detuned with respect to the 1S – 2S level separation thereby repelling these levels. Consequently, the ac Stark shift leads to a blue detuning of the resonance frequency depending on the one-way irradiance I as $\Delta\nu_{\mathrm{acStark}} = 2I \times 1.667 \times 10^{-4}\mathrm{Hz/(W/m^2)}$ [157]. The frequency shift was estimated to be several tens of hertz. A residual first-order Doppler effect may be present as a result of the

limited parallelism of the wave vectors in the resonator and the contribution of higher-order modes in the resonator. The ultimate uncertainty to be achieved with a frequency standard based on a two-photon transition seems to be limited to the extent that the ac-Stark shift can be determined and corrected for.

The resolution of the two-photon line might be further increased by means of optical Ramsey two-photon excitation which also has been demonstrated for the 1S – 2S transition [235]. Since no first-order Doppler shift occurs in the two-photon excitation two interaction zones are sufficient even in the optical domain where the wavelength of the interacting electromagnetic radiation is much smaller than the width of the atomic beam. The necessity of only two interaction zones in two-photon optical Ramsey excitation allows one to make efficient use of the concept of an atomic fountain as has been proposed also for ultra-cold hydrogen atoms [511]. The necessary cooling of hydrogen to lower temperatures has been achieved in a magnetic trap by using resonant light from a hydrogen lamp. The temperature achieved of 8 mK [512] is not far above the Doppler limit of about 3 mK. For a fountain, however, sub-Doppler cooling schemes have to be applied that require the use of coherent radiation. Generation of continuous coherent radiation near the Lyman-α line at 121.56 nm is very tedious. However, up to 200 nW have been obtained by four-wave-mixing of the radiation from three lasers [513] which might be sufficient for laser cooling.

9.4.6 Other Candidates for Neutral-absorber Optical Frequency Standards

Two-Photon Transition in Atomic Silver A promising two-photon clock transition with a natural linewidth of about 1 Hz can be excited by two photons with a wavelength of 661.2 nm in ^{109}Ag (Fig. 9.51; [514]). Silver transitions have been investigated in a thermal beam

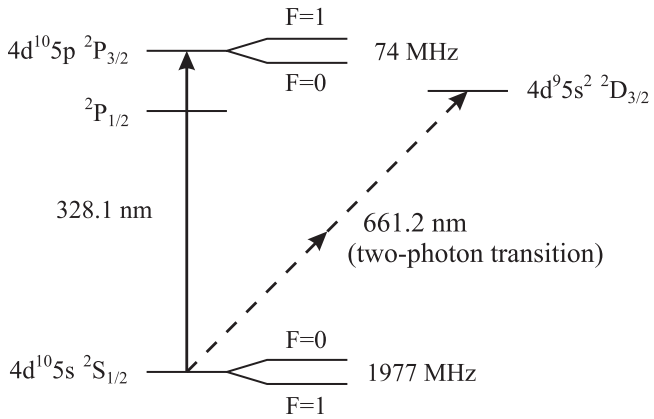


Figure 9.51: Partial energy level diagram of ^{119}Ag

[105, 515] and with laser-cooled atoms [104, 516]. Technically, the generation of the UV cooling radiation at 328.1 nm meets with difficulties. However, frequency doubling of the radiation of a dye laser using a LBO crystal or a diode laser with tapered amplifier and a

$LiNbO_3$ crystal have led to 50 mW and 5 mW, respectively [104]. A minimum temperature of 0.3 mK which is below the Doppler limit of 0.56 mK has been achieved and 3×10^6 atoms have been stored. The corresponding velocities of about 20 cm/s are already low enough to allow for a frequency standard with an expanding cloud of ballistic atoms. With further reduction of the velocity the setting up of a fountain seems possible. As in the case of the hydrogen standard the ultimate attainable uncertainty of a future silver standard may be limited by the ac-Stark shift.

Two-photon Transition in Xenon An optical frequency standard has been proposed [100, 151] based on a two-photon transition in xenon connecting a $^3P_2$ state with a $^3P_0$ state with two 2.19 µm photons. Xenon has been laser cooled in a magneto-optical trap to microkelvin temperatures. The promotion of this candidate was somewhat hampered by the observation that the clock state is de-excited by the ambient room-temperature black-body radiation [100] thereby increasing the linewidth considerably from the expected 2 Hz. It was pointed out, however, that by cooling the apparatus to liquid nitrogen temperature, the black-body induced broadening could be virtually eliminated and the associated frequency shifts could be reduced by three orders of magnitude [100].

10 Ion-trap Frequency Standards

The ideal reference in a frequency standard would consist of an absorber at rest in an environment free of perturbations by other particles or fields, with a high line quality factor Q and a strong signal when probed by an external oscillator. To keep microscopic particles at rest at a fixed position in space a strong binding force pointing to this point is required. Due to the weak interaction of neutral atoms or molecules with electric and magnetic fields, relatively strong fields or field gradients are required which may perturb the atomic energy levels. For ionised particles, however, with an electron added or one or more electrons removed from the shell, much smaller fields are needed for confinement in a so-called ion trap. There are several advantages of ion traps when used for frequency standards applications which will be briefly addressed here and considered in more detail below. First, storage times of days and more allow one to probe ultra-narrow lines with virtually no broadening due to the limited interaction time. Second, the concentration of the absorbers into a small volume allows effective application of the methods of laser cooling and detection of the induced signals originating from a single spot. The reduction of the velocity and the confinement to regions smaller than the wavelength of the probing radiation (Lamb–Dicke regime; see Section 10.1.4) allows the reduction of the Doppler effect in all orders. Third, collisions with other atomic species can be largely suppressed by the use of ultra-high vacuum environment which at the same time reduces the coupling to the outside world. Fourth, the strong interaction with other ions can be avoided by the use of a "mono-ion oscillator as potential ultimate laser frequency standard" as was proposed by Dehmelt in 1982 [517]. The use of ion traps in frequency standards has been the subject of a number of reviews (see e.g. [277, 518–521]). An account of the history of ion traps with important developments and references has been given by Thompson [519].

10.1 Basics of Ion Traps

Electrically charged particles can be confined almost indefinitely to a well defined region in space by suitably shaped electric and magnetic field configurations.

Since it is not possible to keep an electrically charged particle fixed in space using solely static electric fields (Earnshaw's theorem), a combination of static magnetic and electric fields (Penning trap) or a time-dependent inhomogeneous electric field (radio frequency (rf) trap or Paul trap) [1] must be used to trap ions.

[1] The pioneering developments of W. Paul and co-workers [522] to store electrically charged particles in the ion trap have been acknowledged in 1989 by the Nobel prize. The Penning trap is named after Penning [523] who investigated the effects of magnetic fields on low pressure discharges.

10.1.1 Radio-frequency Ion Traps

Consider an electric field configuration $\vec{E}(\vec{r})$ defined by a potential $\Phi(\vec{r})$ leading to a force on an ion with charge q at any space point inside the trapping volume pointing to the centre of this volume. In the following we consider singly charged ions, where the charge is $q = +e = 1.602 \times 10^{-19}$ A s. The force on the ion is

$$\vec{F}(\vec{r}) = e\vec{E}(\vec{r}) = -e \cdot \nabla\Phi(\vec{r}). \tag{10.1}$$

Here, $\nabla = (\partial/\partial x, \partial/\partial y, \partial/\partial z)$ denotes the Nabla operator, used to determine the gradient of a vector field in Cartesian coordinates. Preferentially, this force should increase linearly with the distance $\vec{r}$ from the centre $\vec{F}(\vec{r}) \propto \vec{r}$ since in this case the particles are expected to perform simple harmonic oscillations. The corresponding scalar potential $\Phi(x, y, z)$ has a parabolic shape and may be represented by

$$\Phi = const \cdot \left(ax^2 + by^2 + cz^2\right) \tag{10.2}$$

with the constant to be determined later. From Laplace's equation $\Delta\Phi \equiv \nabla^2\Phi = 0$ for regions of space where there is no charge density we derive the condition

$$a + b + c = 0 \tag{10.3}$$

for the constants determining the potential (10.2).

In the following, we will have a closer look to two particular solutions fulfilling (10.3), namely

$$a = 1, b = -1, c = 0 \qquad \text{(linear quadrupole configuration)} \tag{10.4}$$

and

$$a = b = 1, c = -2 \qquad \text{(three-dimensional quadrupole configuration).} \tag{10.5}$$

10.1.1.1 Linear Quadrupole Trap

The first solution (10.4) represents a trap configuration where the potential is independent of z

$$\Phi = const \cdot \left(x^2 - y^2\right) \tag{10.6}$$

leading to the configuration of a two-dimensional linear quadrupole as sketched in Fig. 10.1.

We will discuss this configuration first before treating the second case (10.5) leading to a trap in three dimensions.

This two-dimensional quadrupole potential can be generated by a set of four hyperbolic electrodes (Fig. 10.1) where the upper and lower electrodes are kept, e.g., at a negative potential and the two remaining ones are set to a positive potential. Consider a potential difference of Φ_0 between both sets of electrodes generated by an applied voltage. The constant of (10.2) and (10.6) is readily determined from $\Phi(r_0) = \Phi_0/2 = const \cdot r_0^2$ as $const = \Phi_0/(2r_0^2)$

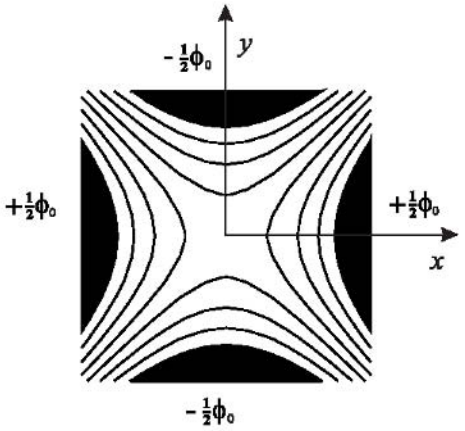


Figure 10.1: The two-dimensional quadrupole potential in the $x - y$ plane can be generated by four electrodes of hyperbolic shape (dark areas).

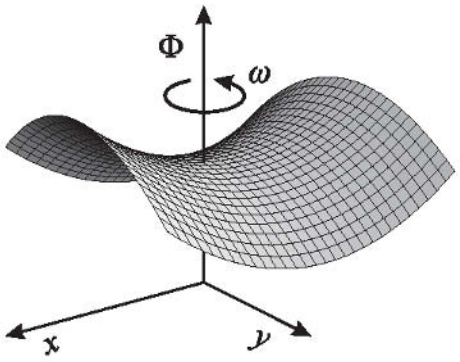


Figure 10.2: The potential surface of a quadrupole electrode (Fig. 10.1) system of hyperbolic shape exhibits a saddle point.

where $2r_0$ is the distance between two opposite electrodes. From this potential the electric field is calculated by use of (10.1) as

$$E_x = \frac{\Phi_0}{r_0^2}x, \qquad E_y = -\frac{\Phi_0}{r_0^2}y, \qquad E_z = 0. \tag{10.7}$$

Within this static field in the $x - y$ plane a particle of electric charge $+e$ will be repelled from the positive charged electrodes experiencing a repulsive force along the x direction towards the centre at $x = 0$. According to the linear dependence of the field strength E on the coordinate x (Hooke's law) we expect the ion to perform a harmonic oscillation in the x direction. In contrast, along the y direction the ion is accelerated towards the nearest negative electrode. The potential surface of (10.6) shows a saddle point at the centre with a minimum along the x direction but a maximum along the y direction (Fig. 10.2). Changing the polarity of the static field will lead to a confinement in the y direction and a corresponding repulsive force pointing away from the centre along the x direction.

To confine the ion in both directions, the potential of both pairs of electrodes is alternated periodically. The alternating rf voltage V_{ac} with a driving angular frequency ω may be added in general to a constant voltage U_{dc} so that

$$\Phi_0 = U_{\mathrm{dc}} - V_{\mathrm{ac}} \cos \omega t. \tag{10.8}$$

Hence, the potential surface of Fig. 10.2 rotates around the vertical axis through the saddle point with angular frequency ω. At first sight, it is not obvious that the alternating focussing and de-focussing along the x and y direction leads to trapping since we might expect that the time-dependent terms cancel on average and no net force acts onto the ion. This is, however, not true in a periodic inhomogeneous field where there is always a small mean force pointing towards the centre. Before we discuss the origin of this force we consider the motion of an ion

in the trapping potential in more detail following the treatment given by Dehmelt [524] and Paul ([525] and references therein).

Consider an ion near the centre of the trap experiencing a time-dependent force according to (10.8). The time-dependent position and velocity components can be obtained from the resulting equations of motion of the ion

$$\begin{aligned} F_x(t) = m\ddot{x}(t) = eE(x)\cos\omega t &= \frac{e}{r_0^2}\left(U_{\rm dc} - V_{\rm ac}\cos\omega t\right)x \\ F_y(t) = m\ddot{y}(t) = eE(y)\cos\omega t &= -\frac{e}{r_0^2}\left(U_{\rm dc} - V_{\rm ac}\cos\omega t\right)y \end{aligned} \tag{10.9}$$

where, as usual, the shorthand $\ddot{x}(t)$ represents d^2x/dt^2. Using (10.7) and the dimensionless parameters

$$\tau \equiv \frac{\omega}{2}t, \qquad a \equiv \frac{4eU_{\rm dc}}{m\omega^2 r_0^2}, \qquad q \equiv \frac{2eV_{\rm ac}}{m\omega^2 r_0^2} \tag{10.10}$$

the equations of (10.9) lead to Mathieu's differential equation [2]

$$\frac{d^2x(\tau)}{d\tau^2} + (a - 2q\cos 2\tau)\,x = 0 \tag{10.11}$$

and

$$\frac{d^2y(\tau)}{d\tau^2} - (a - 2q\cos 2\tau)\,y = 0. \tag{10.12}$$

Since the coefficients of Mathieu's equation (10.11) are periodic functions of τ there exists a so-called Floquet-type solution [59, 526] of the form

$$F_\mu(\tau) = e^{i\mu\tau}P(\tau). \tag{10.13}$$

$P(\tau)$ is a periodic function of the same period as that of the coefficients in (10.11), i.e. π. Every non-periodic solution of (10.11) is known to be a linear combination of the two linearly independent Floquet-type solutions $F_\mu(\tau)$ and $F_\mu(-\tau)$. The so-called characteristic exponent $\mu \equiv \alpha + i\beta$ depends uniquely on the parameters a and q. A complex characteristic exponent μ in general leads to an exponential growth of the amplitude (see (10.13)) and, hence, is termed the unstable solution. A characteristic exponent being real $\mu = \beta, \beta \in R$ leads to oscillations of the ions with uniformly bounded amplitude (stable solution). Obviously, for stable trapping in the experiment the amplitude has to be smaller than the distance from the centre to the electrodes. Since the characteristic exponent is a function of a and q the corresponding dependence $a(q)$ has to be computed for a given $\beta = f(a, q)$, e.g., by the method of the continued fraction [59, 526].

Calculations of $a(q)$ [526] displayed in Fig. 10.3 show stable (shaded) regions for $0 \le \beta \le 1$, $1 \le \beta \le 2$, $2 \le \beta \le 3$ beginning from bottom to top [3].

[2] This type of equation was treated in 1868 by the French mathematician E. Mathieu when he studied the vibration of an elliptical membrane.

[3] Mathematically, there are distinct differences depending on whether β is an integer or not as can be seen, for instance, by the fact that $\beta = 1$ represents the boundary of the first and the second stable region. In practice, however, we need not be concerned with these integer solutions since a trap has to be operated well inside a stable region to allow for stable trapping also in the presence of unavoidable fluctuations of the voltages defining a and q.

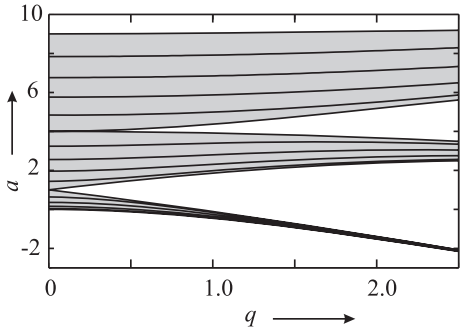


Figure 10.3: The computed $a(q)$ for $\beta = f(a,q)$ with $0 \leq \beta \leq 3$ in increments of 0.2 (lines) [526] shows three stable regions (shaded areas). The figure is symmetric, i.e. $a(q) = a(-q)$.

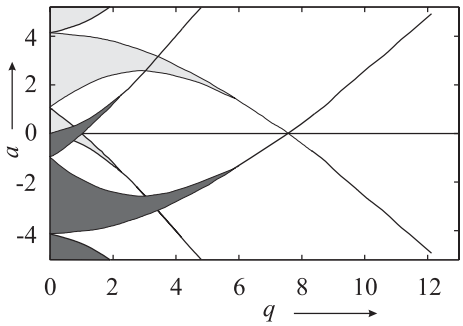


Figure 10.4: A composite diagram similar to Fig. 10.3 for both x and y shows overlapping regions of stability in both directions.

Since stable trapping depends exclusively on the choice of the parameters a and q it does not depend (within some reasonable restrictions) on the initial conditions, e.g., on the velocity of the ions.

To allow stable trapping in a two-dimensional trapping field like the one of Fig. 10.1 or in three dimensions the parameters a_i and q_i ($i = x, y$) have to be kept in a stable region independently. For the trajectory along the x and y directions a compound stability diagram can be constructed (Fig. 10.4) simply by plotting both diagrams into a single one using $+a$ and $+q$ for the x motion and $-a$ and $-q$ for the y motion together with the fact that $a(q) = a(-q)$. We identify stable regimes only where x stable regions and y stable regions overlap in the $a-q$ diagram. The first stability region which is almost exclusively used in the experiments, is plotted together with the curves representing constant stability parameters β for both separate coordinates (Fig. 10.5).

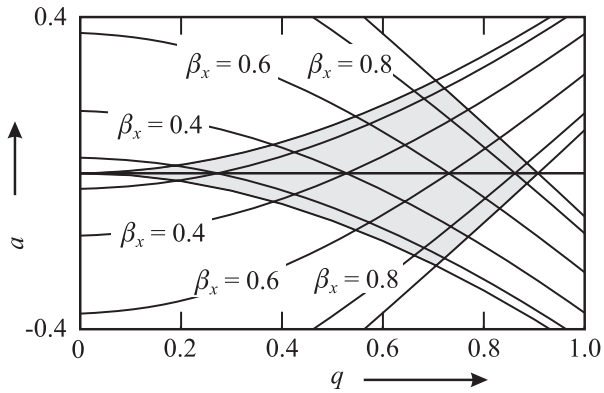


Figure 10.5: First stability region (shaded area) in a two-dimensional trapping configuration like the linear Paul trap or in the mass filter.

The stable solutions to the Mathieu equation (10.11) can be found from the linear combinations of the Floquet solutions $F_\mu(\tau)$ and $F_\mu(-\tau)$ (10.13) as an infinite series of harmonics of τ [59]

$$x(\tau) = A \sum_{n=-\infty}^{\infty} c_n \cos(2n+\beta)\tau + B \sum_{n=-\infty}^{\infty} c_n \sin(2n+\beta)\tau, \tag{10.14}$$

where A and B are constants depending on the initial conditions and c_n and β are functions of a and q. Hence, using (10.10) the motional angular frequency spectrum of the trapped ions is given by the values

$$\omega_n = (n \pm \frac{\beta}{2})\,\omega \tag{10.15}$$

defined by the driving rf angular frequency ω applied to the trap.

Rather than investigating (10.14) in more detail we now turn back to the question of why there is a residual trapping force on the ion in a periodic inhomogeneous electric field. To understand the origin of this force we summarise the treatment given by Dehmelt [524] and Paul ([525] and references therein). We first investigate the motion of an ion which has an excursion $\hat{x}$ from the centre of the trap. Even though in the vicinity of the particle's location the electric field varies locally we neglect this spatial variation for the time being and replace it by the mean constant field $\hat{E}$. Integrating the first of equations (10.9) $m\ddot{x}(t) = e\hat{E}\cos\omega t$ twice and assuming for simplicity that the ion was at rest at $t = 0$, we find the time-dependent position of the particle to be

$$x(t) = \hat{x} - \frac{e\hat{E}}{m\omega^2}\cos\omega t. \tag{10.16}$$

Hence, the ion oscillates with the frequency of the driving field but (because of the minus sign) it is out of phase by π with respect to the driving field. This oscillation is called micromotion and its phase lag is the reason that, in the spatially dependent potential of Fig. 10.2, the ion is accelerated towards the centre of the trap. When the ion is displaced from the mean position $\hat{x}$ by the micromotion towards the centre of the trap, the force on the ion points away from the centre and vice versa. We now lift the restriction of a constant field $\hat{E}$. We see that in the case when the ion is accelerated away from the centre of the trap, it is at a position closer to the centre of the trap where the field is smaller than the mean field $\hat{E}$. Consequently, the electric force acting on the ion is smaller than the force the ion experiences in the higher field at a distance $x > \hat{x}$. As a result, there is a net force acting in the direction of decreasing amplitude of the oscillating electric field. If the driving frequency is sufficiently high and the field amplitude is sufficiently low the resulting force can be thought of as being due to a potential, sometimes called the pseudopotential. In this case the excursion $x(t) - \hat{x}$ during one cycle of the oscillating field is small enough to allow one to use only the first terms of the Taylor series of the electric field

$$\begin{aligned} F(t) &= eE(\hat{x})\cos\omega t + e\frac{dE(\hat{x})}{dx}(x-\hat{x})\cos\omega t + \cdots \\ &\approx eE(\hat{x})\cos\omega t - \frac{e^2E(\hat{x})}{m\omega^2}\frac{dE(\hat{x})}{dx}\cos^2\omega t. \end{aligned} \tag{10.17}$$

Here, we have made use of (10.16) to determine $x(t) - \hat{x}$. Averaging the force over one cycle of the driving field, the first term of (10.17) cancels and the second one averages to

$$F_{\mathrm{av}}(\hat{x}) = -\frac{e^2 E(\hat{x})}{2m\omega^2}\frac{dE(\hat{x})}{dx}. \tag{10.18}$$

The pseudopotential Ψ_{pseudo} corresponding to this force can be obtained from (10.18) by use of (10.1) and extended to two dimensions as

$$\Psi_{\mathrm{pseudo}}(\hat{x},\hat{y}) = \frac{eE^2(\hat{x},\hat{y})}{4m\omega^2}. \tag{10.19}$$

In this so-called adiabatic approximation [524] the ion rapidly oscillates with the frequency applied to the trap (micromotion) and performs a much slower movement (macromotion or secular motion) in the pseudopotential. The secular frequency of this radial slow oscillation can be obtained by equating the potential energy derived from (10.19) and the kinetic energy as follows

$$e\Psi_{\mathrm{pseudo}} = \tfrac{1}{2}m\omega_r^2(x^2+y^2). \tag{10.20}$$

For simplicity, we assume $U_{\mathrm{dc}} = 0$ and insert $E^2(\hat{x},\hat{y}) = E_x^2 + E_y^2$ (see (10.7)) into (10.19) to obtain $\omega_r \approx eV_{\mathrm{ac}}/(\sqrt{2}m\omega r_0^2)$. This angular frequency corresponds to the resonance of lowest order of the solution of (10.14).

Up to now, we have looked for a two-dimensional potential which is capable of confining the ions only radially. For axial confinement, additional means have to be used. The trapping electrodes shown in Fig. 10.1 can be bent to form a ring structure [527, 528]. For frequency standards such a "race track" is not well suited since the position of a particular group of ions is not necessarily fixed and may therefore limit the interaction time with the interrogating field. Several methods have been used to make axial confinement in a so-called linear trap. Examples include additional ring electrodes [529] or segmented rods with a dc potential coupled to the outer rods [530] (Fig. 10.6) or separate dc end electrodes [531, 532].

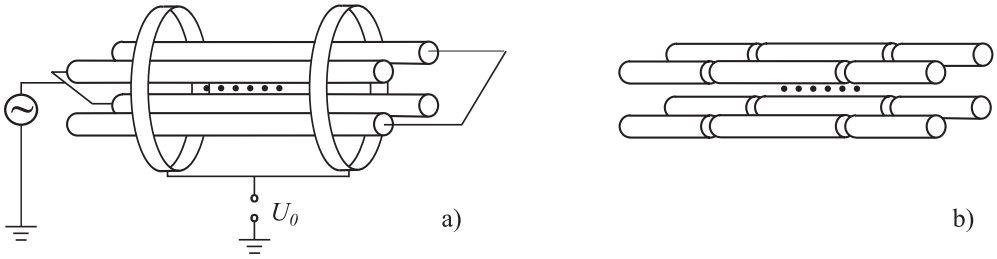


Figure 10.6: Realisations of linear traps generating a radial potential similar to the one of Fig. 10.1 with two additional ring electrodes a) or with additional rods b) for axial confinement.

10.1.1.2 Three-dimensional Paul Trap

The second particular solution (10.5) leads to a three-dimensional potential [533]

$$\Phi = \frac{\Phi_0}{x^2+y^2+2z_0^2}\cdot\left(x^2+y^2-2z^2\right) \tag{10.21}$$

which can be generated by the potential surfaces

$$x^2 + y^2 - 2z^2 \equiv r^2 - 2z^2 = \pm r_0^2. \tag{10.22}$$

The positive sign leads to a hyperbolic surface with rotational symmetry around the z axis generated by a ring electrode of an inner radius r_0 (Fig. 10.7). The negative sign results in two branches of the hyperbola of revolution separated by the distance $2z_0 = \sqrt{2}r_0$ also exhibiting rotational symmetry with respect to the z axis.

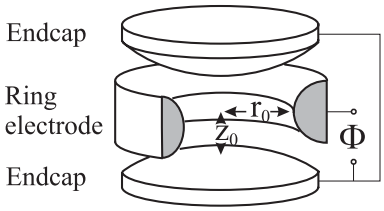


Figure 10.7: Three-dimensional Paul trap.

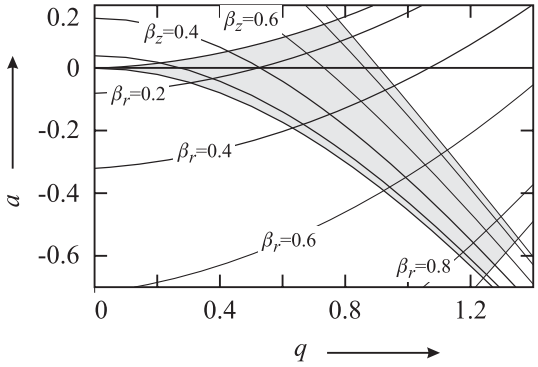


Figure 10.8: First stability region (shaded area) in a three-dimensional Paul trap.

The electric field in the radial direction (E_r) and in the axial z direction (E_z) differ by a factor of -2. The potential in cylindrical coordinates now reads

$$\Phi(r, z) = \frac{U_{\mathrm{dc}} + V_{\mathrm{ac}} \cos \omega t}{r_0^2 + 2z_0^2}(r^2 - 2z^2) \tag{10.23}$$

where r_0 and z_0 are defined as shown in Fig. 10.7. Consequently, we have to introduce individual constants a and q (10.10) for the radial direction (a_r, q_r) and for the axial direction (a_z, q_z) also differing by the same factor of -2

$$a_z = -2a_r \equiv a, \qquad q_z = -2q_r \equiv q. \tag{10.24}$$

Again, we plot a composite stability diagram $a(q)$ for the axial $a_z(q_z)$ and radial $a_r(q_r)$ diagrams (Fig. 10.8). We first plot the diagram for $a_z = a$ and $q_z = q$. In a second step we include into the same diagram $a_r(q_r)$, scaling the latter with the factor of two (see (10.24)). The resulting first stability region is somewhat deformed (Fig. 10.8) as compared to the two-dimensional case (Fig. 10.5).

Even though there is more than one region where stable confinement is possible the first stability region is used almost exclusively in the experiments. A particular trap used at PTB to store $^{171}\mathrm{Yb}^+$ ions had a radius of r_0 = 0.7 mm and was operated with a voltage of V_{ac} = 500 V at an angular frequency $\omega = 2\pi \times 16$ MHz and with a few volts for U_{dc}. We calculate the corresponding parameters $q_z = 0.11$ and $a_z \approx 2 \times 10^{-3}$ from (10.10) and (10.24) and find that the trap is operated well within the first stability region of Fig. 10.8.

The pseudopotential $\Psi_{\mathrm{pseudo}}(\hat{r},\hat{z})$ for the Paul trap can be obtained in a similar way as in the two-dimensional case (see (10.19)) as follows

$$\begin{aligned}\Psi_{\mathrm{pseudo}}(\hat{r},\hat{z}) &= \frac{U_{\mathrm{dc}}}{2r_0^2}(\hat{r}^2-2\hat{z}^2)+\frac{eV_{\mathrm{ac}}^2}{4m\omega^2 r_0^4}(\hat{r}^2+4\hat{z}^2)\\ &= \frac{m\omega^2}{16e}\left[(q_r^2+2a_r)\hat{r}^2+4(q_r^2-a_r)\hat{z}^2\right]. \end{aligned} \tag{10.25}$$

We refer to $e\Psi_{\mathrm{pseudo}}(r_0,0)$ and $e\Psi_{\mathrm{pseudo}}(r_0/\sqrt{2},0)$ as the radial and axial depth of the potential well, respectively. For an alternating potential without constant term (U_{dc} = 0) the depth of the axial well is twice that of the radial one. The potential can be made more symmetrical if a dc voltage (having the same sign as the charge of the ions to be stored) is applied to the ring electrode. For $a_r = q_r^2/2$ the potential becomes spherically symmetrical. The frequencies of the secular motion can be calculated from (10.25) as

$$\omega_r = \frac{\omega}{\sqrt{8}}\sqrt{q_r^2+2a_r} \qquad \text{and} \qquad \omega_z = \frac{\omega}{\sqrt{2}}\sqrt{q_r^2+2a_r}. \tag{10.26}$$

10.1.2 Penning Trap

The Penning trap makes use of the same electrode arrangement as in the radio frequency trap (Paul trap) but no high-frequency field is applied (V_{ac} = 0). The ions experience a repulsive potential along the z axis. In the $x-y$ plane, however, this potential would kick the ion out of the centre. To trap the ion, an additional magnetic field is applied along the z axis. The (classical) equations of motion of the ion in this case read

$$\begin{aligned} m\ddot{\vec{r}} &= e\vec{E}(\vec{r})+e\dot{\vec{r}}\times\vec{B}, \qquad \text{which is equivalent to}\\ m\ddot{x} &= e(E_r+\dot{y}B_z)\\ m\ddot{y} &= e(E_r-\dot{x}B_z)\\ m\ddot{z} &= eE_z. \end{aligned} \tag{10.27}$$

The components of the electric field can be obtained using the potential Φ of (10.23). Solving the last equation leads to a harmonic oscillation where the angular frequency

$$\omega_z^2 = \frac{4eU_{\mathrm{dc}}}{m(r_0^2+2z_0^2)} \tag{10.28}$$

is independent of B_z. If there was only a magnetic field of induction B_z, the charged particle would circle in the plane orthogonal to the magnetic field with its angular frequency

$$\omega_c = \frac{e}{m}B_z \qquad \text{(cyclotron angular frequency).} \tag{10.29}$$

The cyclotron frequency [4] of (10.29) can be derived from a balance of the Lorentz force and the centrifugal force $evB = mv^2/r$ or $eB = m\omega_c$. There is, however, an radial electric field

[4] The name cyclotron frequency refers to the cyclotron where this condition is essential to accelerate charged particles.

$\vec{E}_r$ which is orthogonal to the axial $\vec{B}$ field resulting in a $\vec{E} \times \vec{B}$ drift leading to a circular orbit in the $x - y$ plane around the z axis. The balance of the electric force and the Lorentz force $qE_r = qvB$ leads to the magnetron frequency [5]

$$\omega_m = \frac{E_r}{Br} \qquad \text{(magnetron angular frequency).} \tag{10.30}$$

For a typical magnetic induction of a few tesla and a voltage of a few tens of volts the magnetron angular frequency ω_m, the angular frequency of the axial oscillation ω_z, and the cyclotron angular frequency ω_c are a few tens of kilohertz, a few hundred kilohertz, and a few megahertz, respectively. Hence, for the three frequencies the following relations hold

$$\omega_c \gg \omega_z \gg \omega_m. \tag{10.31}$$

In this regime the trajectory of an ion in the Penning trap is a superposition of the three largely independent motions (Fig. 10.9). There is a fast cyclotron motion around a field line of the magnetic field (10.29), an oscillation along this field line described by (10.28) and a slow ($\vec{E} \times \vec{B}$) drift that can be calculated from (10.30). The trajectory can be described by an orbit with epicircles in the $x - y$ plane. In addition, the ion oscillates harmonically along the perpendicular z axis corresponding to the direction of the magnetic field. If, however, the

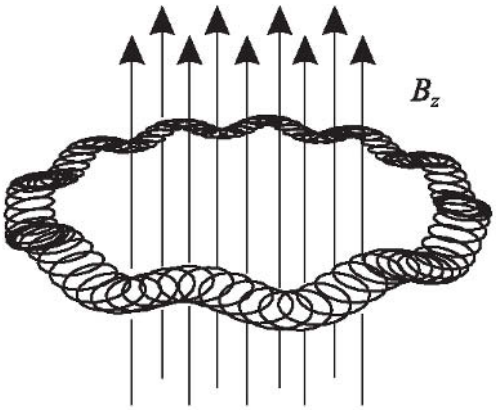


Figure 10.9: The trajectory of an ion in a Penning trap represents an orbit in the $x - y$ plane with epicircles and an additional perpendicular axial oscillation in the z direction, the axis of the magnetic field. The trajectory is drawn for $\omega_c = 10\,\omega_z = 100\,\omega_m$.

cyclotron frequency ω_c (10.29) is not higher than the magnetron frequency ω_m (10.30), the orbits in the $x - y$ plane can no longer be described by a epicircles. Several methods can be

[5] The magnetron frequency does not depend on the properties of the particle as, e.g., charge, mass, or velocity but on the parameters of the electric and magnetic fields. Its name results from the devices called magnetrons used to generate high-power microwave radiation.

applied to solve the two coupled differential equations of (10.27) for the $x-y$ components

$$\ddot{x} = \frac{e}{m}\left(\frac{2U_{\rm dc}}{r_0^2+2z_0^2}x+\dot{y}B_z\right)=\frac{\omega_z^2}{2}x+\omega_c\dot{y} \tag{10.32}$$

$$\ddot{y} = \frac{e}{m}\left(\frac{2U_{\rm dc}}{r_0^2+2z_0^2}y-\dot{x}B_z\right)=\frac{\omega_z^2}{2}x-\omega_c\dot{x}. \tag{10.33}$$

Probably the easiest approach is to add (10.32) and i times (10.33) and to use the complex quantity $r=x+iy$ [534] which leads to $\ddot{r}=\omega_z^2 r/2-i\omega_c\dot{r}$. The latter equation can be solved by $r=r_0\exp(i\omega t)$ leading to the quadratic equation $\omega^2-\omega\omega_c-\omega_z^2/2=0$ for ω. The two solutions of this equation for the angular frequencies are

$$\omega_c' = \frac{\omega_c}{2}+\sqrt{\frac{\omega_c^2}{4}-\frac{\omega_z^2}{2}} \qquad \text{(modified cyclotron frequency)} \tag{10.34}$$

$$\omega_m = \frac{\omega_c}{2}-\sqrt{\frac{\omega_c^2}{4}-\frac{\omega_z^2}{2}} \qquad \text{(magnetron frequency).} \tag{10.35}$$

If the argument of the square root in (10.34) and (10.35) is not negative, i.e. if $\omega_c \geq \sqrt{2}\omega_z$ this solution leads to two new frequencies called the modified cyclotron angular frequency ω_c' and the magnetron angular frequency ω_m. The modification to the "true" cyclotron angular frequency ω_c results from the repulsive term in the electrostatic potential (10.23).

Adding the two equations (10.34) and (10.35) and adding the same equations after squaring them leads to the following relations

$$\omega_c = \omega_c'+\omega_m \tag{10.36}$$

$$\omega_c^2 = {\omega_c'}^2+\omega_m^2+\omega_z^2. \tag{10.37}$$

Both equations can be used to calculate the cyclotron frequency (10.29) from which very accurate comparisons of ionic masses can be derived as will be discussed later. The last equation (10.37) is the more accurate one since it also holds if the magnetic field is misaligned with respect to the axis of the electrode configuration as has been shown by Brown and Gabrielse [535]. Another useful relation is

$$\omega_m=\frac{\omega_z^2}{2\omega_c'} \tag{10.38}$$

which can be derived by subtracting (10.34) from (10.35), squaring the result and inserting (10.37).

For a cloud of ions the $(\vec{E}\times\vec{B})$ drift leads to a rotation around the field lines of the magnetic field. The associated time dilation results in a second-order Doppler shift and represents a systematic shift of the frequency of microwave standards based on ions in a Penning trap. Since the diameter of the cloud grows with the number of ions, this shift increases with the number of trapped ions for otherwise fixed trapping conditions [536]. The radial trapping of ions in the Penning trap is a consequence of the orbits resulting from the balance between an electric force pointing away from the axis of the trap and a magnetic $(\vec{v}\times\vec{B})$ force pointing toward the axis. Consequently, in contrast to the Paul trap there is no restoring force acting on an ion that suffered, e.g., from a collision with a neutral particle. Hence, the ions may

diffuse out of the trap if collisions with the molecules of the background gas occur. There are distinct differences between the magnetron motion and the cyclotron and the axial motion. The latter represents a harmonic oscillation and consequently the energy is swapped between its kinetic and potential part. Due to the high velocity and the small radius, the energy of the cyclotron motion is mainly a kinetic one, whereas the nature of the magnetron motion is essentially potential energy. We can see this immediately by comparing the kinetic and potential energies of an ion performing a magnetron motion close to the centre of the trap and close to the radial electrodes. The potential energy of a singly charged ion decreases from zero to about $E_{\mathrm{pot}} = 5\,\mathrm{eV} = 8 \times 10^{-19}\,\mathrm{J}$ if we assume a voltage of 10 V between the radial and the axial electrodes. From (10.30) we calculate the magnetron velocity of the ion in a trap (magnetic induction $B = 5\,\mathrm{T}$; radius $r = 1\,\mathrm{mm}$) as $v \approx 1000\,\mathrm{m/s}$. For an ion having an intermediate mass of about 100 nucleons ($1.6 \times 10^{-25}\,\mathrm{kg}$) the corresponding kinetic energy of $E_{\mathrm{kin}} = 1/2mv^2 \approx 8 \times 10^{-20}\,\mathrm{J}$ is an order of magnitude smaller compared to the potential energy. Hence, the total energy decreases with increasing magnetron radius and collisions preferably will increase the magnetron radius leading eventually to a loss of the ions.

For frequency standards, the Penning traps suffer furthermore from the fact that the large magnetic field in general causes a large Zeeman shift which is undesirable when applied to frequency standards purposes. Nevertheless, in some exceptional cases frequency standards based on the ions trapped in a Penning trap have been realised or proposed (see e.g. [537–539]). An example will be given below.

10.1.3 Interactions of Trapped Ions

The results derived so far for the motion of the ions hold only for single ions since we have neglected the strong interaction between the ions due to their mutual Coulomb repulsion. If there are several ions in a trap with kinetic energies that are small compared to the energy of their mutual Coulomb interaction, the ions will be arranged in quasi-crystalline structures. In the field-free nodal line of a linear quadrupole trap a small number of ions may be aligned like the pearls on a string [528, 530] (Fig. 10.10). For higher numbers of ions more complicated structures like helices can occur [540]. After the first crystal-like structures were observed in a three-dimensional Paul trap [541, 542] large crystals containing up to 10^5 ions have been observed in a Penning trap [543] or in a linear quadrupole trap [544].

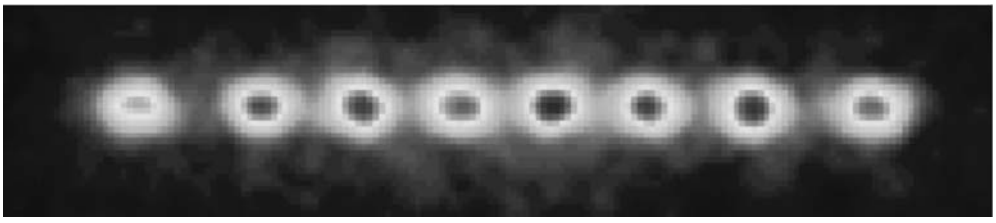


Figure 10.10: Fluorescence from eight ions trapped in a linear Paul quadrupole trap. Photograph courtesy of R. Blatt, University of Innsbruck.

Due to the collective oscillations of the ions new motional frequencies of the ions occur. For intermediate temperatures the non-linearity of the Coulomb interaction will lead to chaotic

behaviour of the ions, and the ions absorb energy from the trapping field, a phenomenon known as radio frequency heating. For even higher temperatures and a sufficiently small density the ions can be thought of as independent particles again and their motion can be described by Mathieu's equation.

Heating results from non-linear forces acting on the stored ions. It has been pointed out by Walther [540] that the non-linear Coulomb force between the ions can lead to substantial heating if the density of the ions is sufficiently high. Strong heating in clouds of ions has been observed which depended on the operating point of the Paul trap as well as on the number of ions in the cloud [545, 546]. This heating has been explained by deviations from the ideal quadrupole potential, resulting, e.g., from misalignment of the electrodes or holes in the electrodes. In this case coupling between the different degrees of freedom of the particles in the trap can occur and energy can be exchanged between the otherwise uncoupled oscillations. Hence, instabilities of the trap due to this heating occur at the resonances (see (10.15) and Fig. 10.11)

$$\frac{n_r \beta_r}{2} + \frac{n_z \beta_z}{2} = 1. \tag{10.39}$$

Here, n_r and n_z are integers with $n_r + n_z = N$, N is the multipole order of the potential, and β_r and β_z are the stability parameters.

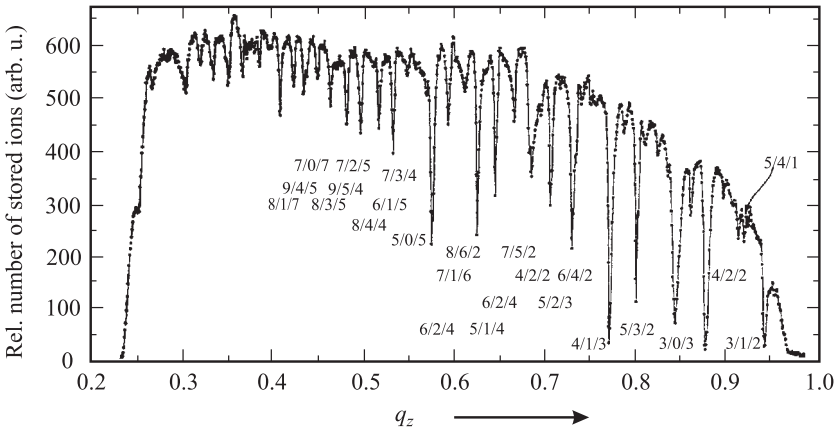


Figure 10.11: Scan revealing instabilities due to heating in a Paul trap [547]. The numbers assigned to the resonances are N, n_r, n_z. Courtesy of G. Werth.

10.1.4 Confinement to the Lamb–Dicke Regime

Besides its many advantages for interrogation of the ions discussed so far the high frequency of optical laser radiation also leads to a predominant broadening $\Delta\nu$ due to the first-order Doppler effect since the broadening is proportional to the frequency (see e.g. (5.111)). For an ion temperature of 1 mK the width of an optical transition is Doppler broadened to a few megahertz. Since, e.g., tuneable laser oscillators can be stabilised to below one hertz [31]

it is important to reduce this broadening. It has been realised by R. H. Dicke [152] that the radiation emitted by a particle confined in a box of dimensions much smaller than its wavelength does not suffer from the first-order Doppler effect. In the same way the Doppler broadening of a spectral line absorbed by an ion vanishes if the oscillation amplitude of the ion in the trap is restricted to much less than the wavelength of the absorbed radiation. To understand this, we recall that an ion oscillating harmonically in the trap with the angular frequency ω_m experiences a phase modulated radiation field

$$E(t) = E_0 \sin(\omega t + \delta \sin \omega_m t) \tag{10.40}$$

even though the field in the laboratory frame $E(t) = E_0 \sin \omega t$ is unmodulated. From (2.50) we know that such a phase modulated radiation in the time domain is represented in the frequency domain by distinct frequencies composed of the carrier with angular frequency ω and an infinite series of sidebands with angular frequencies $\omega \pm n\omega_m$ with $1 \leq n \leq \infty$. For weak phase modulation ($\delta \equiv \Delta\omega/\omega_m \ll 1$) only the carrier remains (see (2.52) and (2.49)). We rewrite this condition as

$$\delta \equiv \frac{\Delta\omega}{\omega_m} = \frac{\omega v_{\max}}{\omega_m c} = \frac{\omega x_{\max}}{c} = \frac{2\pi x_{\max}}{\lambda} < 1 \tag{10.41}$$

where we have made use of the Doppler shift $\Delta\omega = v_{\max}\omega/c$ and the energy balance for an harmonic oscillator $mv_{\max}^2/2 = Dx_{\max}^2/2$ or $v_{\max}^2 = \omega_m^2 x_{\max}^2$. Consider an ion whose vibration is confined to a spatial region $d = 2x_{\max}$. From (10.41) we find that if the diameter of this region is limited to

$$d < \frac{\lambda}{\pi}, \qquad \text{(Lamb–Dicke criterion)} \tag{10.42}$$

$\delta < 1$ holds. Hence predominatly only radiation at the carrier frequency but not at the frequencies of the sidebands is absorbed and the Doppler broadening of the absorption line becomes less significant the better (10.42) is fulfilled. Dicke [152] has derived a similar condition $d < \lambda/(2\pi)$ for a particle in a box of dimension d and consequently this regime is called the Dicke regime or the Lamb–Dicke regime.

10.2 Techniques for the Realisation of Ion Traps

10.2.1 Loading the Ion Trap

Ions can be confined in a trap only if the kinetic energy of the ion is smaller than the barriers surrounding the potential well. It is therefore not possible to inject ions with a defined energy into an ion trap with fixed height of the barrier. There are several ways to load ion traps. The most common method creates the ions inside the trap from neutral atoms, e.g., by ionising an atomic beam by electron impact. This method is not always suitable for isotopes or antiparticles of low abundance that have to be generated or selected for instance in accelerator facilities. Consequently, it may be necessary to raise the potential barrier after injection of the ions fast enough during the time the ions take to traverse the trap [548, 549]. Another method

that has been used is to rapidly reduce the kinetic energy of the ions by an effective cooling method within the time of their passage through the trap region [550]. Although easy to apply, the first method of loading an ion trap is not without deficiencies for use in frequency standards. A common problem often encountered with this technique is that the atoms from the atomic beam may be deposited on the electrodes and on insulating materials. A thin layer is sufficient to lead to a surface potential which can disturb the potential distribution in rf traps. As a result the ions may be shifted into a region where the rf field is no longer zero resulting in an increased micromotion. Modern traps often include additional heating elements to allow heating the trap after loading [551] or additional electrodes to compensate the fields. Ionising the atoms in the trapping region by means of, e.g. UV radiation, avoids these additional provisions.

10.2.2 Methods for Cooling Trapped Ions

The deep potential well in ion traps allows the ions to have high energies without being lost from the trap. The well depth can be as large as 20 eV and if we assume that the kinetic energy of the ions under high vacuum conditions is about 10 % of the well depth [552] the kinetic energy corresponds to a temperature of the ions which is about 80 times higher as compared to room temperature where $k_B T \approx 1/40$ eV. The fractional frequency shift due to the second-order Doppler effect for an ion of mass number 200 of $\Delta\nu/\nu = -v^2/(2c^2) = -(mv^2/2)/(mc^2) \approx 2\,\text{eV}/(200 \times 0.94\text{GeV}) \approx -10^{-11}$ is in general not acceptable for use in a frequency standard. Since the ions in the trap are well isolated from the environment they do not effectively thermalise to the temperature of the apparatus and they have to be cooled by an appropriate method. On the other hand, if the ions have been cooled to low temperatures, e.g., by laser cooling methods, the ions may be kept at temperatures below 1 K for extended times without further cooling. The phrase "cooling" of ions has to be utilised with care. Here, we use it to describe the reduction of the velocity of the ions rather than to reduce their temperature, since the concept of temperature is difficult to use for a single ion or for ions that are far from thermal equilibrium. The prominent methods for cooling the ions will be described in more detail now following reviews given by Holzscheiter [553] or Itano *et al.* [554].

10.2.2.1 Energy Dissipation in an Electric Circuit

Ions oscillating in the trap induce electric currents in the trap electrodes. If these currents are damped by an external resistor energy is removed from the ion and the motion of the ion is damped. If no heating takes place the ions will come to equilibrium at the temperature of the external circuit. A simple model has been given by Dehmelt [555] to describe the time dependence of the corresponding cooling process. Consider a single ion of mass m and charge q oscillating along the z direction between the electrodes of the trap separated by a distance $2z_0$ (Fig. 10.7). The current induced by the ion moving in the electric field E between the electrodes along the distance ds with the velocity v can be calculated from the energy $dW_z = qEds$, necessary to move the ion. If the corresponding power $dW_z/dt = qEds/dt \approx qUv/(2z_0)$ is provided by an external power supply connecting the electrodes, the current I flowing from one electrode to the other can be calculated from

$IU = qUv/(2z_0)$ as $I = qv/(2z_0)$. The approximation is equivalent to replacing the field of the electrode configuration of the trap by that of a parallel plate capacitor. The moving ion inside the capacitor therefore can be thought of as an ideal current source shunted by a capacitor dissipating the time averaged power $\langle I^2 R\rangle$ in an external resistor of resistance R which connects the electrodes. Here, we have assumed that the capacitance C of the electrodes is small enough that $R \ll 1/(\omega_z C)$ holds. The average power dissipated by the ion in the external circuit is calculated as

$$-\frac{dW_z}{dt} = \langle I^2 R\rangle = \frac{q^2 R W_z}{4mz_0^2} \tag{10.43}$$

where we have used $W_z = m\langle v_z^2\rangle$ for the energy of the ion. The solution to this differential equation represents an exponentially decaying energy of the ion with a damping time constant

$$t_0 = \frac{4mz_0^2}{q^2 R}. \tag{10.44}$$

This method can be applied to cool all ions. However, it is most effective for highly charged ions of small mass, as can be seen from (10.44). For cooling the axial motion in a Paul trap (Fig. 10.7) the end caps can be directly connected by the external resistance circuit. To damp the radial (x or y) direction the ring electrode has to be split and the external circuit must be connected between the opposite segments. This cooling method is not applicable for cooling the magnetron motion in Penning traps, since there the reduction of energy is paralleled by an increase of the diameter of the magnetron orbit and by an increase of the magnetron velocity.

In principle, the cooling time can be reduced by use of negative electrical feedback. In this case the electric signal induced in one of the end caps by the displacement of an ion could be used as an error signal to generate an amplified servo signal of opposite phase which is fed to the other endcap. For a cloud of ions only the centre-of-mass motion can be damped. Due to the spread of energies of the individual ions, after some time the centre-of-mass is displaced again and this so-called stochastic cooling can be applied again. The method has been demonstrated [556] but is not in current use in frequency standards.

10.2.2.2 Buffer Gas Cooling

In early trapping experiments with dust particles in a rf trap, Wuerker *et al.* [557] observed that the particles lost kinetic energy and their motion was damped when the pressure of the background was raised to a few hectopascals. Similarly, light buffer gases can be used to cool heavy ions in the same way as kinetic energy is transferred from a coin which is tossed to a second one having a smaller mass. It has been shown, e.g., for mercury ions in a helium pressure of about 10^{-3} Pa [558] that the fractional energy loss per collision would be

$$\frac{\Delta E_{\mathrm{kin}}}{E_{\mathrm{kin}}} = \frac{m_{\mathrm{He}}}{m_{\mathrm{Hg}}}. \tag{10.45}$$

Cutler *et al.* [552] have used this collisional cooling with neutral helium gas atoms to cool $^{199}Hg^+$ ions in a Paul-trap frequency standard. The secular (or macro) motion was cooled

in their experiments to room temperature but the micromotion corresponded to a higher temperature. The disadvantage of the buffer gas cooling results from the occurrence of pressure shifts [559] and from the loss of ions from the trap by collisions if the mass of the stored ions is not much larger than the mass of the buffer gas ions.

10.2.2.3 Laser Cooling

The idea of using lasers for cooling ions in a trap was proposed as early as 1975 by Wineland and Dehmelt [160]. Laser cooling of ions was first observed with barium ions by Neuhauser *et al.* [560] and with magnesium ions by Wineland *et al.* [561]. As discussed in Section 6.3.1 for free atoms, the method is based on the fact that, in general, the energy of the photon absorbed by the ion is smaller than the energy subsequently emitted by the ion. In contrast to free atoms, however, the ion is bound to the trap and the states of allowed energies are discrete (see (10.14)) separated by the energies corresponding to the motional frequencies of the ions in the trap. In cases, however, where these energy differences are small compared to $h\gamma$, with γ the natural linewidth of the cooling transition, the cooling process can be treated classically. Strong resonance lines used for cooling often have a linewidth of γ of a few tens of megahertz, whereas the frequencies in the spectrum of an ion bound in a trap are separated typically by a few megahertz or less. The corresponding cooling process is then very similar to the Doppler cooling of free atoms. The frequency of the cooling laser beam is red detuned, i. e. it is smaller compared to the transition frequency of the ion at rest. Consequently, mostly ions moving towards the direction of the cooling laser beam of wave vector $\vec{k}$ absorb photons from this beam. In the course of each absorption process the momentum of the ion is reduced by $\vec{p} = \hbar k$. The subsequently emitted spontaneous photons in general are randomly distributed into all directions and the average momentum transfer of the emitted photons is zero.

The lowest temperature that can be reached by this kind of laser cooling is given by the so-called Doppler limit $kT_D \equiv h\gamma/2$ (see (6.12)). This minimum temperature results from an equilibrium between the cooling and heating processes and has been treated in more detail in Section 6.3.1. The minimum temperature T_D is obtained, in general, when the detuning is chosen to be half a linewidth below resonance. The exact factor in the relation between temperature and linewidth depends on the particular situation. For a strong cooling transition, having a linewidth of a few ten megahertz, the minimum temperature is $T_D \approx 1\,\mathrm{mK}$. Laser Doppler cooling is possible in both Paul and Penning traps. However, large clouds of ions can be cooled by this technique only in Penning traps. In three-dimensional Paul traps, the radio frequency heating increases rapidly with the number of stored atoms since, due to the Coulomb repulsion, the ions experience the electric rf fields that increase with distance from the trap centre. Consequently, even in small clouds of about one hundred ions, the rf heating cannot be compensated any longer by laser cooling. In Penning traps the heating is much less pronounced and is mostly due to anharmonic trapping fields. Bollinger *et al.* [537] investigated heating effects in a Penning trap and found an increase of the kinetic energy of the ions to about 20 eV only after about 20 s. Hence, it is not surprising that Brewer *et al.* [562] observed laser cooling in clouds of more than 10 000 ions.

Heating effects experienced by laser-cooled single or few ions in radio-frequency traps are much weaker in general. However, they are of particular importance in frequency standards using very long interrogation times. There are different mechanisms that lead to motional

heating in Paul traps [563] such as collisions with background gases and fluctuating fields that exert a fluctuating force on the ions. The latter may be due to Johnson noise, fluctuating patch potentials on the electrodes or others. The Johnson noise, i.e. thermal electronic noise, may result from the resistance in the trap electrodes or in the external circuit. It has been found [563] that the largest contributions are more likely due to fluctuating patch potentials resulting from inhomogeneous surface conditions, e.g., randomly oriented domains or adsorbed material on the electrodes.

Motional Sideband Cooling For ions strongly bound to the trap corresponding to high angular frequencies ω or for weaker cooling transitions, the regime $2\pi\gamma \ll \omega$ is reached. If the laser linewidth as well as the recoil energy is smaller than the separation of the energy levels in this regime of resolved sidebands, a cooling method is possible that has been described by Wineland and Dehmelt [160]. Consider an ion oscillating in a trap with oscillation frequency ν_a. In a harmonic potential well the energies of this oscillating ion will be equidistant and separated by ν_a (Fig. 10.12). Due to the harmonic motion the emission spectrum as well as the absorption spectrum of the ion absorbing at a frequency ν_b in the laboratory system is phase modulated and exhibits resolved sidebands around the carrier $\nu_b \pm m\nu_a$ where m is a positive integer. If the atom is irradiated with a frequency tuned to $\nu_b - \nu_a$ the ion absorbs this frequency. The frequency of the emitted radiation, however, on the average is ν_b. Multiple absorption–emission cycles can bring the ion to the ground state in the trapping potential. In this case the absorption spectrum is dramatically modified as can be seen from an experiment performed by Diedrich *et al.* [564] (Fig. 10.13) who cooled a single $^{198}Hg^+$ ion by this method close to the zero point of motion.

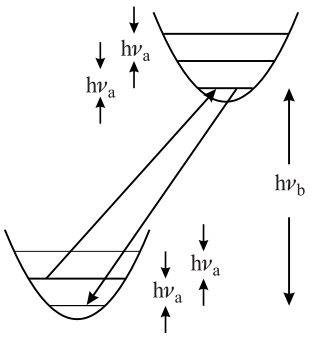


Figure 10.12: Principle of sideband cooling.

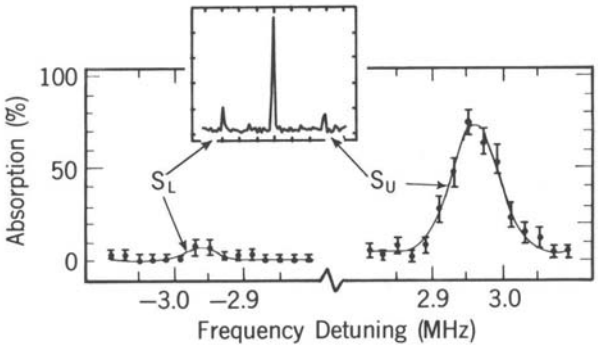


Figure 10.13: Absorption spectrum of the 281.5 nm transition in a single $^{198}Hg^+$ ion before (inset) and after side-band cooling using 194 nm radiation. Courtsey of D. Wineland, with permission from [564].

Consider an ion confined in the Lamb–Dicke regime (Section 10.1.4) where there are only the first-order sidebands in the spectrum due to the motion of the ion. If the ion is in the lowest vibrational level of the electronic ground state only laser radiation with the transition frequency ν_b or with the frequencies $\nu_b + m\nu_a$ can be absorbed. Consequently, the low frequency sideband disappears. From a comparison of the heights of the lower and

upper motional sidebands the population probabilities of the harmonic oscillator levels in the potential well can be derived. The different heights of the sidebands in Fig. 10.13 correspond to a situation where the ion is in the lowest level of the trapping potential for about 95 % of the time corresponding to a temperature of the ion of below 50 μK.

In a Paul trap the Coulomb repulsion between the ions, stray fields and surface potentials can lead to a micromotion of the particles with the frequency of the rf field. Consequently, in contrast to the secular motion which can be reduced for a cloud of ions by laser cooling the kinetic energy in the micromotion cannot be reduced. This problem can be reduced in a linear Paul trap where the rf field is zero on the nodal line.

Sympathetic Cooling In cases where laser cooling of an ionic species meets with difficulties due to, e.g., an unfavourable energy level structure, one species may be cooled by another one that is accessible to laser cooling. This so-called sympathetic cooling has been applied first to cool different isotopes of the same ionic species ($^{25}Mg^+$ and $^{26}Mg^+$ isotopes) by laser cooled $^{24}Mg^+$ ions [565]. Sympathetic cooling has also been applied to different species as in the case of $^{198}Hg^+$ ions that have been cooled sympathetically in a Penning trap by the use of laser cooled $^{9}Be^+$ ions [566]. The $^{198}Hg^+$ ions were cooled and maintained cold by the Coulomb interaction with the cold $^{9}Be^+$ ions. The dynamics of the ions in this kind of trap leads to some kind of spatial separation where the ions with the higher mass-to-charge ratio move radially to the outside of the species with the lower mass-to-charge ratio. In the latter experiment the Hg^+ ions were cooled to a temperature near 1 K by the Be^+ ions whose temperature was about an order of magnitude lower. In general, since laser cooling of large clouds is restricted to Penning traps, so also is sympathetic cooling.

10.2.3 Detection of Trapped and Excited Ions

Charged particles like ions leaving the trap can be easily detected by, e.g., a channel electron multiplier (channeltron) where the accelerated ion ejects electrons. The electrons are accelerated in the electric field inside the tubular channel coated with a high-resistance material and emit secondary electrons each time they hit the surface. The ion loss current can therefore be monitored with high gain. This method relies on the loss of the ions from the trap and therefore is preferentially used in mass spectrometry. For frequency standards, however, methods are preferred which detect the ions but still keep them trapped.

10.2.3.1 Electronic Detection

The motion of the ions in the trap can be detected by highly sophisticated electronic techniques. One so-called "bolometric" technique [567] was devised to detect rf transitions between suitable energy levels of ions by monitoring the translational temperature of the ion gas. There, the noise voltage in a resistor connected between the end caps induced by the moving ions is amplified by suitable electronics. Other more sensitive methods use active circuits where the ionic motion is driven by a weak voltage applied to the electrodes. When the driving circuit has a high Q, the absorption of energy by the ion from the circuit leads to a reduction of the Q value and to an associated voltage drop in the drive when the motional

frequency of the ion is resonant with the frequency of the drive. Very sensitive superconducting detection systems have been used [568,569] preferably in combination with Penning traps where low temperatures are required in any case in order to generate the magnetic field. The heating of the trapped ions during the detection process and the small signal-to-noise ratio makes electronic detection of minor importance for frequency standards.

10.2.3.2 Optical Detection

Trapped ions in a particular internal quantum state can be state-selectively excited to a higher electronic state by a laser tuned to the corresponding transition. The occurrence of this excitation can be monitored either by measuring the absorbed power or the fluorescence radiation from the subsequent spontaneous decay. If a cyclic transition is chosen, i.e. if the excited state always decays to the same initial state, a large number of photons can be scattered and detected from the same ion. This is of particular importance in ion traps where the size of the electrodes often places severe limits on the solid angle for the detection of the fluorescence radiation. Despite the complications of an additional laser for optical excitation, optical fluorescence detection was used even for the determination of microwave ground state splittings [570]. In cases where laser cooling is applied to reduce the velocity of the trapped ions the cooling laser is conveniently used to implement the optical detection scheme.

10.2.3.3 Electron Shelving and Quantum Jumps

A widely used technique relies on a double-resonance method referred to as "electron shelving" [498]. It is often applied to ions with a so-called V system where a strong (cooling) transition and the weak (clock) transition are connected at the ground state (see e.g., Figs. 10.18, 10.19, 10.20). Consider an ion which can be irradiated either with radiation which is resonant with the strong transition, or with radiation which is capable of exciting the clock transition. When the ion is irradiated by the light resonant with the strong transition the excited state will decay after only a few nanoseconds thereby emitting more than 10^8 fluorescence photons per second. Monitoring this fluorescence with a moderate detection efficiency of about 10^{-3} as many as 10^5 photons per second can be detected. If, however, a "quantum jump" to the long-lived excited state occurs, e.g., by absorbing a photon from the clock laser, the strong transition of the ion can no longer be excited as long as the electron is "shelved" in the long-lived state. The fluorescence of the strong transition will be terminated until eventually the ion returns to the ground state and the detected power of the fluorescence light will exhibit dark intervals (Fig. 10.14) during the time the electron spends in the long-lived state. After the first experimental demonstrations [498, 571] the detection of the dark intervals in the fluorescence of an ion due to quantum jumps to a long-lived state is now a routinely applied technique.

From the time of the dark phases the lifetime of the long-lived excited state can be determined directly by evaluating the number of observed dark phases as function of their duration (Fig. 10.15).

In frequency standards it is not necessary, in general, to wait for the spontaneous decay to occur which in special cases could be as high as years [131]. In contrast, after unambiguous identification of the quantum jump by a dark phase the ion might be brought back to the ground

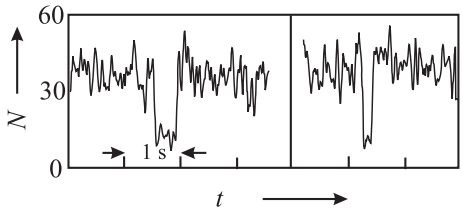


Figure 10.14: The dark intervals in the fluorescence spectrum of a single trapped In^+ ion [572] indicate the transitions to the long-lived state. Courtesy of E. Peik.

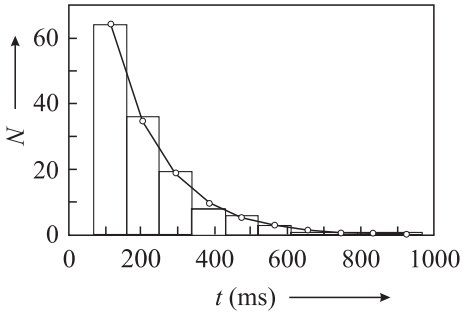


Figure 10.15: The number of observed dark phases taken from data like that of Fig. 10.14 [572] as a function of their duration allows one to fit an exponential decay and to determine the lifetime of the long-lived state.

state by use of a suitable excitation from the long-lived state and subsequent fast spontaneous decay.

The method of electron shelving is equivalent to a quantum amplification since the excitation of an ion can be monitored with detection probability that can approach unity despite the very low efficiency of detection of the photons from the direct excitation.

10.2.4 Other Trapping Configurations

There is, in general, an infinite number of possible trapping configurations aside from the ones described so far. Many of them as, e.g., the combined trap [533, 573] where radio frequency and magnetic fields are used at the same time are not particularly suited for frequency standards and will not be discussed here. Precise hyperbolic electrodes where the potential in the vicinity of a saddle point is parabolic to first order together with a suitable choice of the dimensions of the trapping electrodes allows one to minimise the higher-order contributions to the potential [574] which helps to avoid instability resonances as shown in Fig. 10.11. Despite the associated higher-order contributions that may lead to an increased radio-frequency heating rate, semi-spherical [575], spherical [529] or conical [576, 577] electrodes have been utilised that create a predominantly quadrupole field by simpler electrode arrangements. These configurations are often used for frequency standards.

10.2.4.1 Miniaturised Traps

In large traps, the laser cooling is not very effective since a single ion initially may perform oscillations with large amplitude and is only a short time within the laser field. Hence, at the beginning of the cooling cycle the cooling rate is weak. Good localisation of the ion can be expected from a miniaturised trap. The Paul trap, however, is not the best candidate for such a trap since the optical access for the laser beams used for cooling and excitation is only possible in the space between the electrodes. Consequently, the solid angle for collection of the detected photons is small and the stray light scattered by the electrodes increases.

In cases where good optical access to the stored ions is mandatory, often the so-called Paul–Straubel trap is used. Its roots go back to the early work of Straubel [578] who trapped oil drops in the 50 Hz alternating field of a single ring electrode. Yu *et al.* [579] used such a small ring electrode with a diameter of 100 μm to trap a single barium ion. This type of trap can be thought of as a Paul trap with the end caps being far away. Other modifications have been implemented also as, e.g., three thin ring electrodes [580] or a trap comprising only two end caps [581]. In comparison to a Paul trap with end caps and a ring electrode (Fig. 10.7) the Paul–Straubel trap exhibiting the same ring size requires a larger rf voltage to obtain the same potential depth. The reason for this is that the potential drop occurs in the former case close to the ring electrode whereas in the latter case it is extended over the distance to the grounded environment. Miniaturised traps offer particular advantages and disadvantages for frequency standards. The much smaller electrode size allows the use of small rf voltages. At the same time, however, the field resulting from contact potentials is typically larger so that its compensation is more critical. Since the region of a harmonic potential is small in a miniaturised trap, the rf heating due to the anharmonicity may also be large.

10.2.4.2 Higher-multipole Traps

Besides the traps discussed so far, i.e. with parabolic trapping potentials, there are other possible configurations. Due to their steeper confining potential, rf traps using higher-order electric multipole fields might be used to store a larger number of ions and to confine them in a smaller volume than for Paul traps of comparable size. An rf octupole trap has been analysed and used [582] with clouds of Ba^+ ions. In the Paul trap the motions of charged particles are described by linear, uncoupled equations of motion (Mathieu equations) which can be solved analytically. In contrast, the motion of an ion in a higher-order rf trap is described by non-linear, coupled, and explicitly time-dependent equations of motion which had to be solved by numerical integration. Experimentally, it was observed that the ion cloud radially showed two distinct maxima with a separation that was much larger than the width of the Gaussian spatial distribution in a Paul trap.

In a linear ion trap comprising four rods (Fig. 10.6) the degree of harmonicity depends on the diameters and the spacings of the rods. The use of a segmented cylinder with eight sectors, four at 60 degrees and four at 30 degrees angular width, leads to the desired quadrupole potential depending on the square of the distance ρ from the central nodal line. From the higher order terms the first one which is proportional to ρ^6 vanishes and the next remaining one shows a ρ^{10} dependence. This configuration can be approximated by an arrangement of twelve circular rods [583]. It was shown [584] that fluctuations of the clock frequency due to a fluctuating number of ions can be much smaller in ion clocks based on multipole traps than comparable clocks based on quadrupole linear traps.

10.3 Microwave and Optical Ion Standards

Trapped ions can provide reference frequencies in the microwave or in the optical domain. In the former case mostly magnetic dipole transitions between the ground states split by the hyperfine interaction are used. In the optical domain mostly forbidden electric dipole or

higher-order multipole transitions between different electronic states are utilised. Since the frequencies of both classes typically differ by four to five orders of magnitude, both have their particular advantages. Microwave transitions can be conveniently used to lock radio frequency oscillators and the generated signal frequencies can be easily processed and counted using conventional electronics. By contrast, the high frequencies of optical transitions allow frequency comparisons with the same uncertainty in a much shorter time. This advantage, however, has to be paid for by the additional equipment required to connect the optical frequencies to the microwave regime.

10.3.1 Microwave Frequency Standards Based on Trapped Ions

Several candidates for microwave standards based on trapped ions have been investigated (see Table 10.1). Here, we discuss the standards based on $^{9}Be^{+}$, $^{171}Yb^{+}$, and $^{199}Hg^{+}$ ions

Table 10.1: Selected microwave clock transitions in ions. Ground state hyperfine splitting in other ions may be found, e.g., in [585].

Ion	Frequency / Hz	References
$^{9}Be^{+}$	303 016 377.265 070(57)	[536, 537, 586]
$^{43}Ca^{+}$	3 255 608 286.4(3)	[587]
$^{137}Ba^{+}$	8 037 741 667.694(360)	[588, 589]
$^{113}Cd^{+}$	15 199 862 858.(2)	[590]
$^{171}Yb^{+}$	12 642 812 118.468 5(10)	[591–593]
$^{199}Hg^{+}$	40 507 347 996.841 59(44)	[529]

in more detail since they may serve as examples in the following to recall the state of the art with ion-trap frequency standards. Comprehensive compilations can be found also, e.g., in [277, 520].

10.3.1.1 $^{9}Be^{+}$ Ions in a Penning Trap

The $^{9}Be^{+}$ ion has an angular momentum of the nucleus of $I = 3/2$ and of the electron shell of $J = 1/2$. The ground states $F = 2$ and $F = 1$ split in the magnetic field according to Fig. 10.16. At a magnetic induction of $B = 0.8194$ T, which is conveniently obtained in a Penning trap, the transition frequency ν_1 between the $F = 1$, $M_I = -3/2$, $M_J = 1/2$ sub-state and the $F = 1$, $M_I = -1/2$, $M_J = 1/2$ sub-state becomes independent of the magnetic field to first order. The frequency $\nu_1 \approx 303$ MHz between these states depends only quadratically on the magnetic induction, ΔB, as $\Delta\nu_1/\nu_1 = -0.017(\Delta B/B)^2$. A frequency standard based on this clock transition has been operated at the United States National Institute of Standards and Technology in Boulder [537, 586, 594]. Different schemes have been used to obtain the desired population difference between the two sub-states. In the first realisation of the standard [537] the ions were laser cooled by the radiation of a frequency-doubled dye laser ($\lambda \approx 313$ nm) tuned to the transition from the $2s\ ^{2}S_{1/2}$ ($M_I = -3/2$, $M_J = -1/2$) state to the $2p\ ^{2}P_{3/2}$ ($M_I = -3/2$, $M_J = -1/2$) state, which leads to optical pumping to the $M_I = -3/2$, $M_J = -1/2$ state (see Fig. 10.16). Microwave (mixing) radiation of a frequency

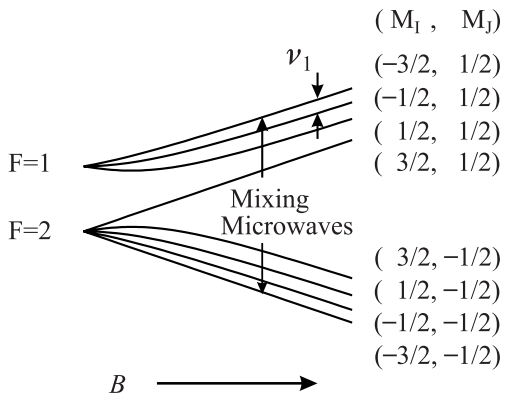


Figure 10.16: Ground state hyperfine energy levels of the $^9Be^+$ ion in a magnetic field. For a magnetic induction of B = 0.8194 T the frequency of the clock transition ν_1 from the $F = 1$, $M_I = -3/2$, $M_J = 1/2$ state to the F=1, $M_I = -1/2$, $M_J = 1/2$ state becomes independent from the magnetic field to first order.

of about 23.9 GHz was used to repump half of the ion population from the $M_I = -3/2$, $M_J = -1/2$ state to the $M_I = -3/2$, $M_J = -1/2$ state. Applying a microwave radiation near the clock transition at 303 MHz reduces the population in the higher state and allows the microwave mixing field at 29.5 GHz to further reduce the population in the lowest state ($M_I = -3/2$, $M_J = -1/2$). The depletion of this state therefore indicates that the clock transition is resonantly excited. This was monitored by the reduction of the fluorescence induced by the 313 nm laser. The interrogation of the clock transition by the 303 MHz radiation was performed using two pulses of duration t separated by a time T defining the resolution. This pulsed scheme represents the time-domain equivalent of the Ramsey excitation. The observed linewidth of 25 mHz corresponding to T = 19 s resulted in a line quality factor $Q = 1.2 \times 10^{-10}$. When a passive hydrogen maser was used as reference, the frequency was measured to be ν_1 = 303 016 377.265 070(57) Hz and a fractional frequency stability of $\sigma_y(\tau)$ ranging from $1.3 \times 10^{-11}(\tau/\mathrm{s})^{-1/2}$ to $4 \times 10^{-11}(\tau/\mathrm{s})^{-1/2}$. The fractional uncertainty of about 1.8 $\times 10^{-13}$ was dominated by the second-order Doppler shift. During the interrogation time the cooling laser and the microwave mixing radiation had to be shut off to avoid light shifts and ac Zeeman shifts. Within the measurement the temperature of the ion cloud consisting of a few hundred to about 2000 ions, increased from less than 1 K to about 35 K. To cool the ions also during the time necessary to interrogate the clock transition with high resolution, sympathetic laser cooling was used in the most recent version of the $^9Be^+$ frequency standard [586,594]. $^{26}Mg^+$ ions were loaded into the same trap together with the $^9Be^+$ ions. Since the frequency of the laser beam of a wavelength of 280 nm used to cool the magnesium ions is far off resonance, with any transitions of the $^9Be^+$ ion, continuous radiation can be used to cool the magnesium ions, and consequently the beryllium ions continuously. In contrast to the methods described before, the $^9Be^+$ ions were optically pumped into the $M_I = +3/2$, $M_J = +1/2$ state by the 313 nm radiation. After turning off this laser the ions were successively transferred to the $M_I = +1/2$, $M_J = +1/2$ state and from there to the $M_I = -1/2$, $M_J = +1/2$ state

by means of two π pulses of 321 MHz and 311 MHz, respectively. After interrogating the transition from the lower to the higher state by two Ramsey pulses the number of ions left in the lower state was probed by bringing them back to the $M_I = +3/2$, $M_J = +1/2$ using the π pulses in reversed order and monitoring the fluorescence when excited with the 313 nm laser. The time between the two Ramsey pulses could be made as long as 550 s corresponding to a linewidth of 0.9 mHz. The fractional instability was better than $3 \times 10^{-12}(\tau/\mathrm{s})^{-1/2}$ for $10^3\,\mathrm{s} < \tau < 10^4\,\mathrm{s}$. An unexpectedly large pressure shift from collisions with CH_4 molecules was observed which limited the stability to about 3×10^{-14}. A cryogenic environment was suggested as a remedy [536]. The fractional second-order Doppler shift corresponding to about 5×10^{-15} [594] seems to represent a limit for the fractional uncertainty of this standard.

10.3.1.2 $^{171}Yb^+$ Microwave Frequency Standards

The $^{171}Yb^+$ ion has caused many researchers to establish a microwave frequency standard. The advantages of this ion result from the low Doppler shifts associated with the high mass of the ytterbium ion, the simple hyperfine spectrum and the large transition frequency of 12.6 GHz between the hyperfine structure sub-levels of the ground state (Fig. 10.17). The

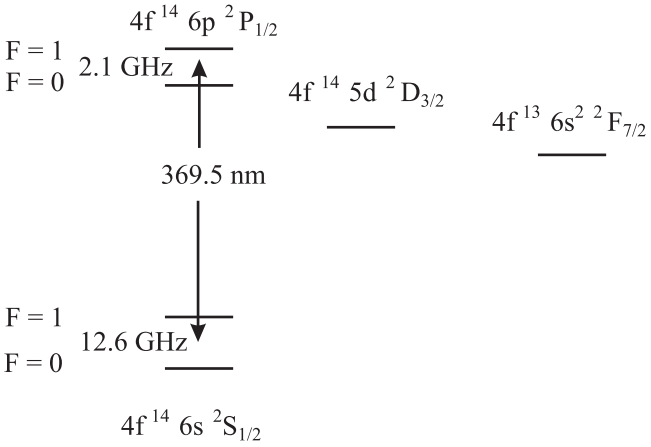


Figure 10.17: Partial energy level diagram of $^{171}Yb^+$ ions. The microwave clock is operated on the 12.6 GHz transition betwen the $F = 0$ and $F = 1$ ground states.

optical transition at $S_{1/2} \rightarrow P_{1/2}$, λ = 369.5 nm for cooling and detection can be reasonably well accessed by dye lasers or frequency-doubled solid-state lasers. Probably the first measurement of the frequency of the 12.6 GHz transition using trapped ions has been performed in the group of G. Werth at the University of Mainz [595]. These authors obtained a line quality factor $Q = 2 \times 10^{-11}$ using a He buffer gas cooled ($p \approx 10^{-4}$ Pa) cloud of about 10^5 ions in a Paul trap. Rather than to detect the tiny absorbed power directly and to determine the centre of the absorption curve an optical pumping scheme is in general employed to obtain a background-free signal [596]. The ions were optically pumped by a weak pulsed dye laser ($\lambda = 369.5$ nm) from the $S_{1/2}, F = 1$ state via the $P_{1/2}$ state to the $F = 0$ ground state, where the fluorescent decay served to monitor the population in the $F = 1$ state. Transitions

between the $F = 0$ and $F = 1, m_F = 0$ ground state sub-levels, induced by the applied microwave radiation, were monitored by the fluorescence detected perpendicular to the laser beam through one of the end caps of the trap.

The possibility of a microwave standard based on about 10^6 Yb^+ ions in a Paul trap was demonstrated later at the University of Hamburg [597] with a reduced uncertainty of the frequency measurement and a measured instability $\sigma_y(\tau) = 2\times10^{-11}(\tau/\mathrm{s})^{-1/2}$ for averaging times τ up to a few hundred seconds.

Several institutes including the Physikalisch-Technische Bundesanstalt (PTB) in Braunschweig [559, 592], the National Research Laboratory of Metrology (NRLM) in Tsukuba [598–600], the National Measurement Laboratory (NML) of the Commonwealth Scientific and Industrial Research Organisation of Australia (CSIRO) in Sydney [591, 593], or the California Institute of Technology's Jet Propulsion Laboratory (JPL) [601] used rf traps of hyperbolic shaped electrodes (PTB, NRLM) or linear rf traps (NMI, JPL) to investigate Yb^+ microwave standards. As in most other microwave ion standards, at PTB also a double-resonance technique using radio-frequency excitation and laser preparation and detection (λ = 370 nm) was applied to a cloud of up to 50 000 collision cooled ^{171}Yb ions [559]. The frequency of the ground state hyperfine transition was determined by comparison with the frequency of the primary Cs standards to be 12 642 812 118.471(9) Hz [592]. The uncertainty of 9 mHz was mainly limited by the high temperature in the region of 2000 K. Other sources of uncertainty included the quadratic ac-Stark effect with a coefficient of the fractional frequency shift of 2×10^{-17} $(\mathrm{V/cm^2})^{-1}$ and the pressure shift in helium. Pressure shifts relevant for microwave Yb^+ standards have been determined [559,593] for helium, nitrogen, neon and hydrogen, with relative frequency shifts ranging between 10^{-8}/Pa and 10^{-10}/Pa. In CSIRO, Fisk *et al.* [593] operated Yb^+ microwave frequency standards using linear trap configurations where rf heating and Stark shifts were substantially reduced since the node of the confining rf field is a line. In their standard IT-2 Fisk *et al.* used a cloud of a 2×10^4 ions of length 24 mm, radius about 2 mm and a temperature of about 400 K. The ions were excited by microwave $\pi/2$ pulses of 0.4 s duration separated by 25 s yielding Ramsey fringes with a period of 40 mHz. The frequency of the 12.6 GHz radiation was synthesised from the frequency supplied by a cryogenic sapphire resonator. The frequencies corrected to the frequency of the unperturbed Yb^+ ion are listed in Table 10.1 where the maximum correction was due to the second-order Zeeman effect of about 0.8 Hz. Earlier values for the transition frequencies have been compiled in [520].

Using laser-cooled $^{171}Yb^+$ ions the fractional uncertainty of the measured frequency of 12 642 812 118.468 5 Hz could be reduced to below 8×10^{-14} with a projected uncertainty of 4×10^{-15} [591].

Several researchers have reported that the fluorescence of a cloud of Yb^+ ions excited by the 369 nm radiation gradually becomes weaker with a corresponding degradation of the signal-to-noise ratio. It has been found that there is a possibility for the ions to get trapped in low-lying long-lived meta-stable D and F states (Fig. 10.17). It has been shown that the signal could be increased again by the use of additional laser radiation which repumped the ions from the meta-stable D states before they further decayed to the extremely long-lived F state [598, 602–605]. Another approach [601] investigated different buffer gases to quench the trapped population in excited Yb^+ and found nitrogen as the most suitable one.

The use of a single Yb^+ ion in a frequency standard has been investigated by the Hamburg group [606] and a potential fractional uncertainty of 10^{-16} has been envisaged.

10.3.1.3 $^{199}Hg^{+}$ Microwave Frequency Standards

The higher mass and the higher ground-state hyperfine splitting of mercury ions, as compared to ytterbium ions, led to the development of excellent microwave standards based on $^{199}Hg^{+}$ ions.

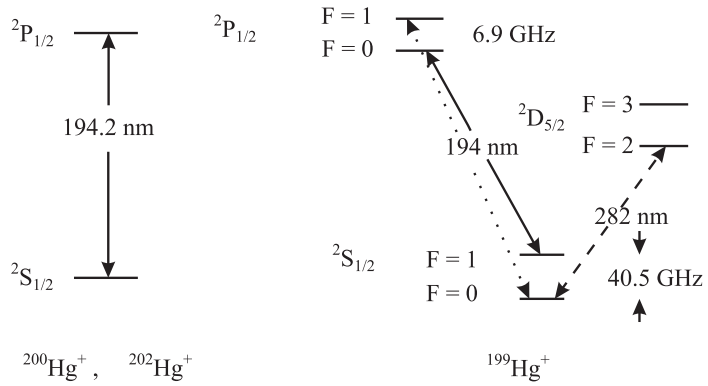


Figure 10.18: Partial energy-level diagram of $^{200}Hg^{+}$ and $^{199}Hg^{+}$ relevant for laser cooling and detection including the electric quadrupole transition at 282 nm used for an optical frequency standard.

Historically, after the observation of the 40.5 GHz ground-state hyperfine splitting [596] with $Q \approx 10^{10}$ the first prototypes of frequency standards were realised in the Laboratoire de l'Horloge Atomique (LHA) [607, 608] and at Hewlett–Packard [552, 609–611]. Both groups operated a Paul trap with hyperbolically shaped electrodes with typically 10^6 stored ions. The latter used helium buffer gas at a pressure of 1.3×10^{-3} Pa to cool the ions. Electromagnetic radiation of 40.5 GHz produced by a microwave source locked to a synthesiser was used to excite the atoms from the $F = 0$ to the $F = 1$ ground state (Fig. 10.18). Similarly, as in the rubidium clock (Section 8.2) there is a natural coincidence between the wavelengths of two isotopes that can be utilised. The $^2S_{1/2}-^2P_{1/2}$ transition (λ = 194.2 nm) of the $^{202}Hg^{+}$ isotope, that has no hyperfine splitting owing to its nuclear spin quantum number $I = 0$, coincides with the $^2S_{1/2}(F = 1) - {}^2P_{1/2}$ transition of the $^{199}Hg^{+}$ isotope. When a discharge lamp filled with $^{202}Hg^{+}$ irradiates the trapped $^{199}Hg^{+}$ ions only those ions previously excited by the microwave radiation to the $F = 1$ state will re-emit ultraviolet radiation at λ = 194.2 nm that can be detected with a photomultiplier. Three of those systems have been operated in the US Naval Observatory for several years. The group at the Jet Propulsion Laboratory has developed ultra-stable frequency standards [583, 612] based on linear traps [531] with typically 10^6 to 10^7 $^{199}Hg^{+}$ ions pumped by a ^{202}Hg rf discharge lamp and cooled with He buffer gas to near room temperature. In an extended linear trap ions are transferred between two confinement regions one for production and detection of the ions and one for Ramsey excitation. The contribution of the fractional second-order Doppler shift was estimated to be -4×10^{-13} [612]. This buffer gas cooled mercury trap has a very low instability expressed by the relative Allan deviation $\sigma_y(\tau) = 7 \times 10^{-14}(\tau/\mathrm{s})^{-1/2}$.

The uncertainties associated with a large cloud of Hg^+ ions can be avoided by storing a single ion or only a few ions in a trap. Very low numbers of laser cooled $^{198}Hg^+$ ions [542] in a Paul trap have been shown to exhibit clusters or crystals. The crystallisation of the ions takes place when the kinetic is reduced below the energy associated with the Coulomb repulsion between the ions. The group at NIST has operated a linear ion trap with only a few $^{199}Hg^+$ ions [613] that can be located near to the field-free nodal line in a regular "string of pearls" to investigate its use for frequency standards. From one to up to more than thirty $^{199}Hg^+$ ions could be stored in these experiments. The ions were loaded by leaking atomic mercury into the system at a pressure of about 10^{-6} Pa that was reduced by two orders of magnitude after loading the trap. Crystals have been observed showing "defects" [542, 613] resulting from impurities of other isotopes or molecular ions. At this pressure also background neutral Hg atoms cause losses of Hg^+ ions presumably by forming dimers with ions excited by the cooling laser. Consequently, the group at NIST has operated a cryogenic linear ion trap for a $^{199}Hg^+$ frequency standard [529, 551].

Due to the hyperfine separation of 6.9 GHz of the excited state in $^{199}Hg^+$ the optical pumping in the laser cooling scheme requires a second laser at 194 nm that is offset by 47.4 GHz with respect to the first one. The transition between the sub-states $^2S_{1/2}(F = 1)$ of the ground state and $^2P_{1/2}(F = 0)$ is a cycling transition since the excited ions can decay only to the same ground sub-state due to dipole selection rules. There is, however, a small probability that the strong cooling laser will also excite ions from the $^2S_{1/2}(F = 1)$ state to the $^2P_{1/2}(F = 1)$ state from where the ions can decay to the other ground state $^2S_{1/2}(F = 0)$ and are lost for cooling. A second weak laser is therefore necessary to repump those ions to the $^2P_{1/2}(F = 1)$ state from where they can decay to both ground states.

Further complications result from the magnetic sub-structure of the $F = 1$ sub-state which requires a strong magnetic field during laser cooling that has to be switched off during microwave Ramsey excitation [530] or modulated in the polarisation of two cooling laser beams [529]. Despite the technical difficulties associated with the hyperfine structure, the ultraviolet lasers, and the cryogenic environment at 4 K, impressive results have been obtained [529]. With seven ions and a Ramsey interrogation time of 100 s, a fractional frequency instability of $3.2 \times 10^{-13}(\tau/\mathrm{s})^{-1/2}$ was measured for measurement times of $\tau < 2$ hours. These authors derive a fractional uncertainty of 1.1×10^{-14} for the system where the uncertainty of the frequency (see Table 10.1) was dominated by the uncertainty with which the frequency could be referenced to the atomic time scale TAI (Section 12.1.2) at the time of the measurement.

10.3.2 Optical Frequency Standards with Trapped Ions

There are several candidates with trapped ions for a frequency standard in the optical domain. A short and by no means exhaustive list on this work is given in Table 10.2. The most important criterion for selecting an ion is the existence of a suitable clock transition and a convenient transition for laser cooling and detection. The realisation of several promising alternative schemes is currently hampered by the fact that the relevant wavelengths are in the deep ultraviolet where no convenient laser sources are available. The rapid progress encountered with laser techniques during the past years and the tremendous achievements currently achieved in material technology, however, make it foreseeable that this argument will lose importance

Table 10.2: Selected optical clock transitions in ions. The data are from Ba^+: [614–616], Sr^+: [617, 618], Ca^+: [619–621], Yb^+: [101, 131, 622–626], In^+: [627–629], Hg^+: [21, 499, 501]. In some cases other isotopes of these ions may be more advantageous. Other candidates can be found, e.g. in [630].

Ion	Transition	Frequency wavelength / μm	Natural line-width / Hz
$^{138}Ba^+$	$5d\ ^2D_{3/2}$ - $5d\ ^2D_{5/2}$	24 012 048 317 170 Hz 12.5	0.02
	$6s\ ^2S_{1/2}$ - $5d\ ^2D_{5/2}$	170.1 THz 1.762	0.005
$^{88}Sr^+$	$5s\ ^2S_{1/2}$ - $4d\ ^2D_{5/2}$	444 779 044 095 510(50) Hz 0.674	0.4
$^{43}Ca^+$	$4s\ S_{1/2}$ - $3d\ D_{5/2}$	411 THz 0.729	0.13
$^{171}Yb^+$	$6s^2S_{1/2}$ - $5d^2F_{7/2}$	642 121 496 772.6(1.2) kHz 0.467	5×10^{-10}
$^{171}Yb^+$	$6s^2S_{1/2}$ - $5d^2D_{3/2}$	688 358 979 309 312(6) Hz 0.435	3.2
$^{171}Yb^+$	$6s^2S_{1/2}$ - $5d^2D_{5/2}$	729 487 779 566(153) kHz 0.411	22
$^{115}In^+$	$5s^2\ ^1S_0$ - $5s5p\ ^3P_0$	1 267 402 452 899.92(23) kHz 0.2365	1.1
$^{199}Hg^+$	$6s\ ^2S_{1/2}$ - $5d^96s^2\ ^2D_{5/2}$	1 064 721 609 899 143(10) Hz 0.282	1.8

in the near future. Blue diode lasers, more efficient frequency doubling crystals or optical parametric oscillators, to name only a few, may pave the road to this development. We will therefore also discuss in the following, some of the technologically challenging candidates.

10.3.2.1 $^{88}Sr^+$ Optical Frequency Standards

The $^2S_{1/2}-{}^2D_{5/2}$ 674 nm transition in single trapped Sr^+ ions has been investigated for use in an optical frequency standard [84, 618, 631]. The advantage of this ion stems from the fact that the wavelengths of the clock transition at 674 nm and the transition needed for laser cooling and detection at 422 nm (Fig. 10.19) is conveniently generated by diode lasers and by frequency doubling of 844 nm diode laser light, respectively. The cooling cycle, however, is not closed and subsidiary light at 1092 nm is necessary to repump the ions lost from the cooling cycle by a decay channel to the $^2D_{3/2}$ meta-stable level. This light can also be obtained from a Nd^{3+} doped silica fibre laser [632]. The upper level of the clock transition has a lifetime of (347 ± 33) ms.

A laser with its frequency referenced to this transition has been recently recommended by the International Committee of Weights and Measures (CIPM) for the realisation of the

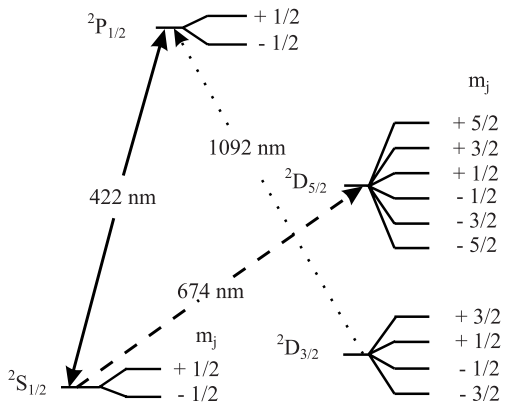


Figure 10.19: Partial energy diagram of Sr^+.

length unit [370]. The optical clock transition at 674 nm was thoroughly investigated at the National Physical Laboratory of England and the National Research Council of Canada and the frequencies of the trapped ^{88}Sr ions measured in both institutes yielded an agreement of about 10^{-13} [617,618,631]. Madej *et al.* [618] find it likely that the fractional uncertainty can be reduced to the 10^{-17} level for the ^{88}Sr standard.

A particular complication in this frequency standard results from the fact that the ^{88}Sr isotope has zero nuclear spin and the ten Zeeman components (Fig. 10.19) all have a linear linear dependence on the magnetic field. An alternative may be the $^2S_{1/2}(F = 5, m_F = 0)$–$^2D_{5/2}(F' = 7, m_{F'} = 0)$ transition in ^{87}Sr with a quadratic Zeeman shift of $6.4\,\mathrm{Hz}/(\mu\mathrm{T})^2$ [633].

10.3.2.2 $^{171}Yb^+$ Optical Frequency Standards

The single Yb^+ ion received a lot of attention recently as a candidate for an optical frequency standard. First, $^{171}Yb^+$ has a nuclear spin of $I = 1/2$ leading to reference transitions without linear Zeeman effect in a level system with relatively simple hyperfine and magnetic sub-level structure. Second, there are three different optical transitions with high frequencies in the blue spectral region (Table 10.2 and Fig. 10.20) that together with the other optical transitions needed for cooling or detection can be accessed by frequency-doubled near-infrared semiconductor laser sources. Fig. 10.21 shows a measurement of the 435 nm transition that has been recorded with a linewidth of less than 80 Hz; where the carrier together with the sidebands due to the oscillation frequencies of the ion in the trap in the radial (r_1, r_2) and the vertical (z) direction can be identified.

Tamm *et al.* [623] have compared two largely independent $^{171}Yb^+$ standards and found no significant deviation on the 1×10^{-15} level. The fractional instability was $\sigma_y(\tau = 1000\,\mathrm{s}) = 1 \times 10^{-15}$. The absolute frequency has been measured with a femtosecond comb (see Table 10.2). The ultimate accuracy of the ^{171}Yb ion may be limited by the quadrupole shift similarly as in the $^{201}Hg^+$ ion [634]. A transition without quadrupole shift, however, can be found in the isotope $^{173}Yb^+$ isotope to the $I = 5/2, F = 0$ state.

A particularly interesting transition is the highly forbidden electric octupole (E3) transition at 467 nm with an estimated lifetime of the $^2F_{7/2}$ excited state of ten years which is

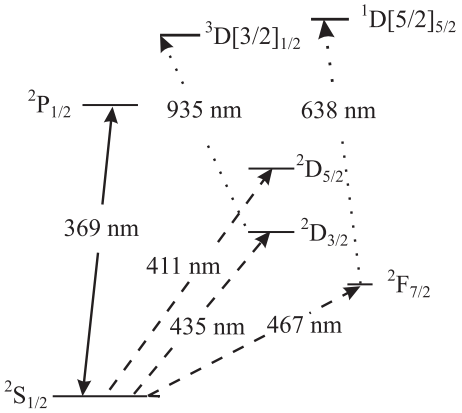


Figure 10.20: Partial energy diagram of Yb^+. Dashed lines (411 nm, 435 nm and 467 nm) represent the optical transitions proposed for optical frequency standards. The 369 nm line is used for cooling and detection.

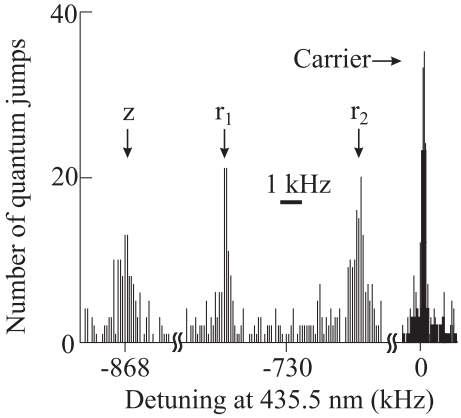


Figure 10.21: The spectrum of the transition of a single Yb^+ ion in a Paul trap [622] shows besides the carrier the motional sidebands along the radial (r_1 and r_2) and axial (z) directions. Courtesy of Chr. Tamm.

investigated for an optical frequency standard in the NPL (UK) [131]. The octupole transition is weak and, hence, high irradiance of the probe laser is required to interrogate the transition. The transition has been recorded in a trapped single $^{171}Yb^+$ ion with a linewidth of the clock laser of 4.5 kHz [101] and an irradiance of 10^7 W/m^2. In this experiment the high irradiance led to a considerable ac Stark shift (Section 6.6) of about 500 Hz. As the predicted ac Stark shift of 47 μHz W^{-1} m^2 scales with the linewidth of the laser for a clock laser having 0.5 Hz linewidth, a fractional shift of 1 part in 10^{16} is expected [101]. The transition is free from first-order Zeeman shift and the second-order Zeeman shift has been determined to be 2.1 mHz/mT^2.

10.3.2.3 $^{113}In^+$ and $^{115}In^+$ Optical Frequency Standards

On the way to the ideal frequency standard Dehmelt [517, 635] has suggested utilising the $J = 0 \rightarrow J = 0$ transitions in particular intercombination transitions in group III ions from the $^1S_0$ ground state to the $^3P_0$ state. Conservation of angular momentum does not allow this transition for any order of multipoles of the radiation field in a pure LS coupling scheme. Due to hyperfine interaction, however, other states of $J \neq 0$ are usually mixed to the $^3P_0$ state leading to a small transition probability of decay by dipole radiation. Owing to the vanishing angular momentum of the electron in the ground and the excited state, the perturbations of the frequencies of these transitions by external fields are expected to be extremely small. In contrast to the S $\rightarrow$ D transitions in Yb^+, Ca^+, Sr^+, and Hg^+ ions discussed also here, the static intrinsic quadrupole moment of the ion vanishes for $J \leq 1$. The interaction of the quadrupole moment with the gradient of the electric field leading to a fractional frequency shift that can be as large as 10^{-15} in other ions is missing in the In^+ ion. The single ground state represents a further advantage of the ions in the third group of the periodic system since

no pumping to other hyperfine states can occur during the cooling process and consequently no repumping laser is needed. From the singly ionised group III elements of the periodic system B^+, Al^+, Ga^+, In^+, and Tl^+ with the $^1S_1 \rightarrow ^3P_0$ transition the lighter elements are up to now not much suited due to a missing suitable cooling transition. The heavy elements In and Tl are much better candidates and consequently Dehmelt in his proposal has concentrated on Tl^+ [517].

The indium ion has been investigated as a frequency standard in the group at the Max Planck Institute at Garching [395, 627, 629]. The clock transition $^1S_0 \rightarrow ^1P_0$ (Fig. 10.22; $\lambda = 236.5$ nm) has an achievable line quality factor $Q = 1.2 \times 10^{15}$ due to its natural linewidth of 1.1 Hz. There is a strong transition in the singlet system ($^1S_0 \rightarrow ^1P_1$; λ = 158.1 nm;

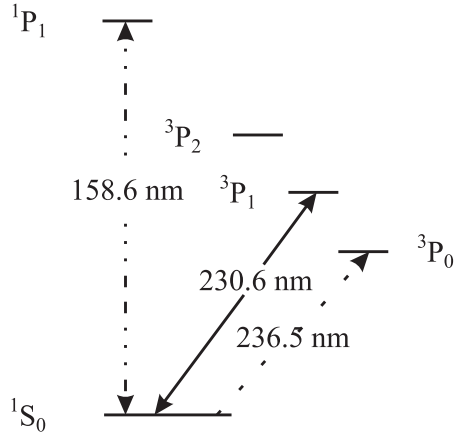


Figure 10.22: Partial energy diagram of In^+.

Fig. 10.22) that could be used to cool and detect the ion. However, the generation of the ultraviolet transition meets with technical difficulties. Peik *et al.* [572] have therefore used the intercombination transition $^1S_0 \rightarrow ^3P_1$ (λ = 230.6 nm) for this purpose. The lifetime of $\tau(^3P_1)$ = 0.44 µs corresponds to a linewidth of 360 kHz, which is about two orders of magnitude smaller compared with the cooling transitions in "alkali like" group II ions or other ions like Ba^+ or Hg^+. Since this linewidth is in general smaller than the motional frequencies of the ion, optical sideband cooling is possible and very low temperatures of 20 µK can be reached. On the other hand, the small linewidth of this transition leads to a reduced cooling rate and to a reduced fluorescence scattering rate when used for detection with the electron shelving technique. The radiation for the cooling transition has been obtained by frequency doubling the radiation of a dye laser at 461.2 nm using stilben 3 (Fig. 9.24) or solid state lasers. The clock transition fortuitously coincides with the fourth harmonic of the 946 nm line of the Nd:YAG laser and thus permitted the set-up of a narrow-linewidth solid state laser [636]. For the ultimate accuracy [637] of a standard based on the In^+ transition, the linear Zeeman effect of the clock transition may be disadvantageous. The corresponding frequency shift of about 2.4 kHz/mT for the transitions $^1S_0 \rightarrow ^3P_1$ with $m_F = \pm 1/2 \rightarrow m_F = \pm$ 1/2 due to the high nuclear spin of the stable isotopes $^{113}In^+$ and $^{115}In^+$ of I=9/2 requires control of the magnetic field below the nanotesla region. A narrow linewidth of the clock transition of 170 Hz has been observed, and its frequency has been measured by comparison with the

frequency of an iodine stabilised optical frequency standard at 532 nm [395] or by means of a femtosecond comb [629].

10.3.2.4 $^{199}Hg^{+}$ Optical Frequency Standard

The electric quadrupole transition in $^{199}Hg^{+}$ (λ = 282 nm; Fig. 10.18) has been investigated in the group at the NIST [21, 499, 638, 639] as a reference for an optical frequency standard. A standard based on this transition has only be made possible after overcoming a number of challenging technical difficulties. For example, the high vapour pressure of elemental mercury leads to a reduced storage time of the ions as a consequence of the recombination of the trapped ions with the neutral atoms. To reduce this effect, the trap is operated at liquid helium temperature with further challenges for the experimenters [640]. To use the narrow linewidth of the clock transition of 1.8 Hz, a laser has been developed with sub-hertz linewidth [31]. A Fourier limited linewidth as narrow as 6.7 Hz at 282 nm has been observed [641] corresponding to an experimentally realised $Q \approx 1.6 \times 10^{14}$. The projected accuracy of a single ion Hg^{+} standard has been estimated as low as 10^{-18} [641]. As in Yb^{+} the uncertainty finally achieved may be limited by the non-vanishing electronic quadrupole moment of the $^2D_{5/2}$ state and its interaction with the electric field of the trap. The quadrupole shift of the ^{199}Hg ion has been calculated [634] to be of the order of 1 Hz for 10^3 V/m^2. The frequency has been measured with a femtosecond comb [501] with a fractional uncertainty of less than 10^{-14}. By comparison with an optical Ca standard an upper limit of the instability with a $^{199}Hg^{+}$ clock of $\sigma_y(\tau) = 7 \times 10^{-15}(\tau/\mathrm{s})^{-1/2}$ has been derived [21]. These impressive results clearly show that suitable single-ion optical clocks can compete favourably with the best microwave clocks.

10.3.2.5 Other Candidate Ions

There is a large number of suitable candidate ions to be used in frequency standards of the future. Ba^{+} and Ca^{+} ions having energy level schemes similar to that of the Sr^{+} ion have been investigated [521]. The proposed clock transition (4S – 3D) of the Ca^{+} ion (see Table 10.2, Fig. 10.23) is an electric quadrupole transition with a lifetime of about 1 s [642]. The isotope $^{43}Ca^{+}$ with the odd nuclear spin $I = 7/2$ has a transition that is to first order independent of residual magnetic fields. Furthermore, the Ca^{+} ion is a hydrogen-like atom and, hence, is

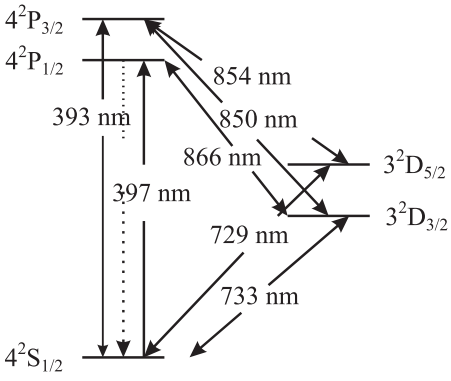


Figure 10.23: Partial energy diagram of Ca^{+}.

interesting because its wave functions can be calculated readily. Technically, this ion and its transitions become more interesting since all relevant optical transitions can be accessed by diode lasers.

The Tl^+ ion already proposed by Dehmelt [635] has the same outer electronic configuration as the indium ion discussed above. A completely different class of candidates with an even number of both the neutrons and protons (so-called gg nuclei) [6] and vanishing nuclear spin is found in the doubly ionised elements of the fourth group in the periodic system. The technical difficulties associated with the use of these elements ranging from C^{2+} to Pb^{2+} are enormous, since the relevant clock transitions $^1S_0 \rightarrow {}^3P_1$ are located in the deep ultraviolet between 199 nm and 166 nm. These difficulties may be overcome, however, thanks to a proposal [643] in which the clock ion is trapped simultaneously with an auxiliary ion species which could provide cooling and state detection of the clock ion (Section 14.2.1). It has been suggested that besides the Tl^+ ion, $^{10}B^+$ and $^{27}Al^+$ [644] might then also become suitable candidates.

10.4 Precision Measurements in Ion Traps

Apart from their use as frequency standards where the internal oscillations of the ionic quantum absorbers are utilised, the measurement of the external oscillation frequencies of ions in the trap allows a variety of measurements with unprecedented accuracy. As the external oscillation frequencies of trapped ions depend on the properties of the ions themselves, valuable information about these particles can be extracted from the frequency measurements. On the other hand these particles can be monitored over a long time and the same system can be studied again and again allowing one to use the stored particles as high sensitivity probes of tiny time-dependent and environment-dependent effects. From the numerous excellent experiments performed with stored ions in the past we restrict ourselves to a few examples closely related to applications in fundamental research, metrology and technology.

10.4.1 Mass Spectrometry

The knowledge of accurate masses of microscopic particles [645] is extremely important if they are used as probes in high precision experiments where the masses are the input data for other measurements. An example of this kind is presented in Section 13.6.1 where the measurement of the Rydberg constant with optical frequency standards allows one to derive fundamental constants only after correcting for the finite mass of the particle. Accurate masses are furthermore needed to build a consistent mass scale from the lightest to the heaviest atomic particles. Moreover, the atomic and nuclear binding forces contribute to the mass of an atomic particle and weighing atomic particles with low uncertainty allows one to test nuclear models.

The most accurate mass measurements are performed in Penning traps where the true cyclotron frequency (10.29) of an ion in a strong magnetic field relates the mass-to-charge ratio m/e to the magnetic induction B. The true cyclotron frequency can be derived by use of (10.37) from the measured trap-dependent modified cyclotron frequency ν_c' and the magnetron

[6] See footnote 6 on page 125.

frequency ν_m derived from the measured axial frequency according to (10.38). Alternatively, (10.37) can be used to determine the true cyclotron frequency from the three normal mode frequencies for the trapped ions, namely the observable trap-dependent modified cyclotron frequency [646], the axial frequency ν_z and the magnetron frequency ν_m.

DiFilippo *et al.* [646] have measured atomic masses of several atomic and molecular ions by relating these masses to the atomic mass unit $m_u = m(^{12}\mathrm{C})/12$. To obtain the high accuracy, perturbations by the electric field had to be kept small by applying only a small electric field. The resulting low axial frequency (160 kHz) was measured using a superconducting resonance circuit. To locate ions with different mass numbers at the same place in the trap the trap voltages were kept the same for the different ions. In order to derive the ratio of the masses of the ions by the ratio of their respective cyclotron frequencies, the magnetic field has to be kept constant during the measurement. To reduce the influence of temporal drifts of the magnetic field of several 10^{-9} per hour the cyclotron frequencies of two ion species were measured alternating between the two ions. As a result, a table of fundamental atomic masses could be derived with relative uncertainties of 10^{-10} and below [646]. To illustrate the impressive accuracy achieved so far it was almost possible to detect the mass difference of molecules due to their binding energies and to allow to "weigh molecular bonds".

Similar techniques have been applied to determine masses of unstable isotopes [647] produced in high-energy reactions at accelerator facilities. The accurate measurement of isotope sequences allows one to study the nuclear binding energy as a function of a broad range of proton and neutron numbers. As a consequence, nuclear properties, e.g., shell closures, pairing, or deformation effects of the nucleons can be inferred and used to test models of the nuclei. The short lifetime of the unstable isotopes, their different frequencies and the small number of ions available requires special techniques to detect the cyclotron resonances of the ions. One method uses the interaction of the magnetic moment of the ion's orbit with a gradient of the magnetic field, where energy associated with the cyclotron motion is converted into axial energy. After excitation of their motion and ejection from the trap, the ions excited with the cyclotron frequency drift faster towards the detector. The resolution for the masses of unstable isotopes can exceed 10^6 and the relative uncertainty of the mass determination can be lower than 10^{-7}.

The masses and mass ratios of fundamental particles represent another important field for precision frequency measurements in Penning traps. Examples include the determination of the mass of the proton, electron, positron, neutron or antiproton. The atomic mass scale is based on the mass of the ^{12}C isotope and, hence, the masses of all atomic particles have to be referenced to this isotope. Van Dyck *et al.* [648] have measured the proton's rest mass by comparing the mass ratio of a proton and a C^{4+} ion in a compensated Penning trap. In these types of traps apart from the hyperbolic end caps and ring electrode, guard ring electrodes were used to compensate for the non-quadratic even-order terms in the trapping potential. To efficiently drive and cool the cyclotron motion and to detect the cyclotron frequency directly the ring electrode was split into four equal parts. The mass ratio of a proton and a C^{4+} ion was determined from the free-space cyclotron frequencies $\nu_c(\mathrm{p}^+)$ and $\nu_c(\mathrm{C}^{4+})$ as $m_p = M(\mathrm{C}^{4+}) \times \nu_c(\mathrm{C}^{4+})/(4\nu_c(p^+))$. The latter were derived from the respective modified cyclotron resonance frequencies ν_c' using (10.37). To refer the mass of the proton to the mass of the neutral ^{12}C atom, rather than to that of the C^{4+} ion a correction for the binding energies $E_B = 148.019\,\mathrm{eV}$ and the masses m_e of all the four liberated electrons had to be

performed, leading to $M(^{12}\mathrm{C}) = M(^{12}\mathrm{C}^{4+}) - E_B + 4m_e$. The corrected value gave a mass of the proton with a relative uncertainty of about 3×10^{-9}. The mass ratios of positrons $m(e^+)/m(e^-)$ [649] and protons $m(p^+)/m(e^-)$ [650] relative to the electron had been determined also in Penning traps. The electron's mass has been determined alternatively [651] from a measurement of the g factor of the electron in $^{12}\mathrm{C}^{5+}$ [652] in combination with calculated quantum electrodynamical corrections to $0.000\,548\,579\,909\,2(4)m_u$ with a relative uncertainty of 7.3×10^{-10}.

10.4.2 Precision Measurements

The investigation of the hyperfine structure and the effects that lead to frequency shifts is of vital importance since a number of frequency standards are based on hyperfine transitions. At present the accuracy of measurement of the hyperfine splitting exceeds the capability of calculating these splittings from first principles. Ab initio calculations agree with the experiment on the 10^{-3} level [585]. Differential effects, however, e.g., by comparing the hyperfine structure and the g_I factors in a chain of isotopes are expected to clarify details of the magnetic structure of nuclei [585].

The measurements of magnetic moments or g factors of ions require large magnetic fields and, hence, are performed preferably in Penning traps. The accuracy of the measurement of g_J factors of ionic ground states is possible with a fractional uncertainty of a few times 10^{-7} with similar uncertainties of the calculations [585] where relativistic corrections become important. The g factors allow therefore sensitive tests for relativistic wave functions. The uncertainty is mainly limited by the uncertainty of the determination of the magnetic field strength at the position of the ions. In the group at the university of Mainz, therefore, a double trap configuration was used [652] where the transitions are induced and detected spatially separated in an analysis trap and a precision trap, respectively. In the precision trap the magnetic field is as homogeneous as possible to allow for narrow linewidths of the cyclotron resonance line when irradiated by the microwave field. After the ions are transported to the analysis trap by moving the electric potential minimum between the two traps spin flips are detected via the Larmor frequency in the inhomogeneous magnetic field that couples the spin orientation to the axial frequency of the ion. Accurate measurements of the g factor of the electron in $^{12}\mathrm{C}^{5+}$ were used together with quantum electrodynamical predictions to derive an independent determination of the electron's rest mass [651].

10.4.3 Tests of Fundamental Theories

The measurement of the anomalous magnetic moment of the electron (and similarly that of the positron) [326] challenged the theory since the deviation $g-2$ is believed to result mostly from quantum electrodynamic (QED) corrections.[7] The QED contribution can be written as a series of powers in the fine-structure constant α. Kinoshita [653, 654] has given a value for $g-2$ using α derived from measurements of the quantum Hall effect [655]. The comparison of the theoretical $(g-2)_{\mathrm{theor}} = (1\,159\,652\,156.4 \pm 23.8) \times 10^{-12}$ and experimental $(g-2)_{\mathrm{exp}} =$

[7] There are also hadronic and electro-weak contributions of a few parts in 10^{12} and 10^{14}, respectively.

$(1\,159\,652\,188.25 \pm 4.24) \times 10^{-12}$ [326] values of $g-2$ is at the same time a striking proof of the accuracy that can be achieved with experiments based on frequency measurements and the theoretical framework connecting very different fields of physics.

The test of the equivalence of the mass and charge of a particle and its antiparticle is particularly interesting since it is required from the invariance of physical laws under the so-called CPT transformation. The CPT theorem states that physical laws must remain the same under the combined transformation when the charge C is conjugated, and parity P and time T are reversed. Experiments have shown violations of P in the weak interaction, CP, or T, separately, but the combined transformation CPT is believed to leave the physical laws invariant. A comparison of the charge-to-mass ratios of the proton and the antiproton was performed by Gabrielse *et al.* [656, 657]. The Penning trap is not capable of trapping both positive and negative particles simultaneously, and one has to alternate either the voltage of the trapping electrodes or that of the magnetic field to trap either particle or antiparticle. To keep the magnetic field constant during the experiment the latter procedure is not practical. To allow an efficient transfer of the antiprotons into the trap, a large opening was required and stacked cylinders were used rather than hyperbolically shaped electrodes to produce the quadrupole potential of the trap. The high quality of the potential necessary to generate harmonic motions independent of the energy of the trapped ions was obtained by a careful choice of the length of the cylindrical electrodes and the voltages. Gabrielse *et al.* [656] obtained the same charge-to-mass ratio for the proton and the antiproton within a relative uncertainty of 9×10^{11} [657]. An even tighter bound on a possible CPT violation was deduced [658] from earlier $g-2$ measurements of the electron and positron.

Frequency standards based on ion traps have been utilised to perform several fundamental tests devised as null experiments with the aim either of setting new limits for the currently accepted theories or of finding deviations that require a refinement of our current understanding. The researchers at the National Institute of Standards and Technology have used their frequency standard based on Be^+ ions in a Penning trap (Section 10.3.1.1) to look for a possible dependence of the frequency of the 303 MHz ground-state hyperfine transition on various parameters [586]. In a first experiment, the frequency of the Be^+ transition was compared to the frequency of a hydrogen maser. Since the $^9Be^+$ nucleus has a quadrupole moment and the hydrogen atom has not, a possible anisotropy of space would lead to a diurnal variation of the frequency ratio between both standards. Within the experimental uncertainty no dependence of this ratio was observed that could be interpreted as resulting from a 24 hour period associated with a possible anisotropy of space [659]. Another experiment looked for the dependence of the frequency of the Be^+ ion on the way the transition was excited. For excitation, a radio-frequency pulse with a well defined Rabi angle θ was applied. The probability of finding the Be^+ ion in the upper of the two states $(-3/2, 1/2$; see Fig. 10.16) varies sinusoidally with the Rabi angle (5.52). In particular if a π pulse is applied the ion is excited with 100 % probability whereas a $\pi/2$ puts the ion into a coherent superposition with equal probabilities to find the ion in the two states $(-3/2, 1/2)$ and $(-1/2, 1/2)$; (Fig. 10.16)). It has been proposed by Weinberg [660] that a non-linear correction to quantum mechanics would result in a small shift of the transition frequency shift on the Rabi angle. Bollinger *et al.* [661] looked for frequency differences when the Rabi angle was near to 60 degrees or 120 degrees, but did not find such a term within their very low fractional uncertainty of 1.3×10^{-14}.

In another series of experiments Wineland *et al.* [662] compared the frequencies of the

303 MHz transition of Be^+ in devices where the magnetic field was generated either by a superconducting magnet or by a conventional electromagnet. From the null results, upper bounds were derived for possible but hitherto unknown spin-dependent interactions between the spins of the Be^+ ions and the spins of the electrons in the iron pole faces of the electromagnet and the nucleons in the Earth.

11 Synthesis and Division of Optical Frequencies

The advantages of the high frequencies of laser oscillators in the optical domain can be utilised to their full extent only if these frequencies can be generated or measured with the same ease of operation as their counterparts in the microwave domain. In order to connect these regimes, where the frequencies can differ by up to five orders of magnitude, special devices have been developed. They allow one to multiply the low frequencies or to divide the high frequencies in a similar way as does a gear for rotational frequencies of mechanical shafts. These "gears" comprise non-linear mixing stages (Section 11.1) or frequency shifting elements (Section 11.2) and allow the frequency synthesis by harmonic generation (Section 11.3), by frequency division (Section 11.4) or by frequency combs from ultra-short pulse lasers (Section 11.5).

11.1 Non-linear Elements

The response P of a medium perturbed by a small excitation U is often considered to be linear in the same way as a mechanical spring shows an elongation proportional to the applied force in the regime of Hooke's law. If the system is excited with a harmonic perturbation of angular frequency ω

$$U(t) = U_0 \cos \omega t \tag{11.1}$$

it responds with the same frequency ω. For larger perturbations, however, the response becomes non-linear

$$\begin{aligned} P(U) &= \alpha_1 U + \alpha_2 U^2 + \alpha_3 U^3 \cdots \\ &= \alpha_1 U_0 \cos \omega t + \frac{\alpha_2}{2} U_0^2 (1 + \cos 2\omega t) + \frac{\alpha_3}{4} U_0^3 (3 \cos \omega t + \cos 3\omega t) + \cdots \end{aligned} \tag{11.2}$$

and higher harmonics $2\omega, 3\omega, \cdots$, appear. The coefficients α_i can be considered as the expansion coefficients of order i of a Taylor series. Hence, the generation of higher harmonics 2ω, $3\omega, \cdots, n\omega$ is intimately related to the occurrence of that particular term in the Taylor expansion. In other words, a strong curvature of the transfer function (Fig. 11.1) of the non-linear device is essential for the generation of higher harmonics.

As can be seen from (11.2) the non-linear characteristic produces higher-order powers of a harmonic signal. If the input signal U is composed of two components with two different angular frequencies ω_1 and ω_2 the resulting signal (response) contains new frequencies with

$$\omega = m\omega_1 \pm n\omega_2 \tag{11.3}$$

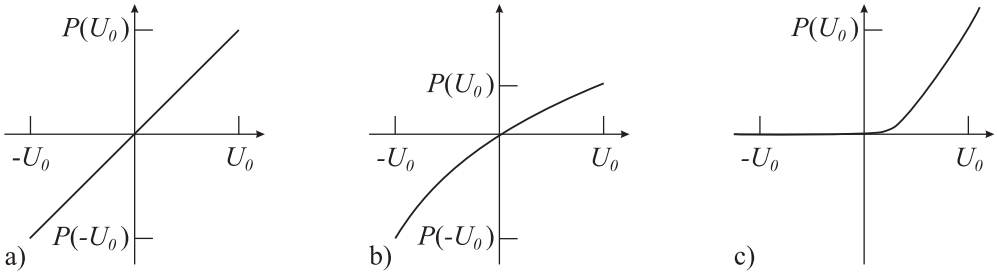


Figure 11.1: Linear (a) and non-linear (b, c) characteristic responses $P(U)$ of a medium excited by a perturbation U.

where m and n are integers and $m + n$ is the highest power exponent of the non-linear characteristic. In the following several non-linear effects will be discussed that lead to devices that can be used to produce and detect suitable combinations of frequencies of two or more signals.

11.1.1 Point-contact Diodes

A rectifying diode exhibits a particularly strong non-linear characteristic (see Fig. 11.1 c) and, hence, is well suited for generating high harmonics. To achieve suitable response to high frequencies the time constant of the device $\tau = RC$ has to be as low as possible. The series resistance R can be minimised by using materials of high conductivity made from metal or highly doped semiconductors. The minimal stray capacity C is obtained by using the smallest possible point contacts.

11.1.1.1 MIM Diodes

Point-contact metal-insulator-metal (MIM) diodes have long been used for absolute frequency measurements in the far infrared [663]. The point contact diode results from the MIM interface when a tip of, e.g., a tungsten wire is pressed onto an oxidised surface of a nickel or cobalt post. The tip of the tungsten wire of about 8 µm diameter is etched to a radius of about 30 nm. The radiations of a microwave or a far infrared laser coupled into the MIM interface can give rise to an electric signal to be further processed with frequencies at the difference or sum frequencies or higher harmonics of the impinging radiations. To use the contact wire as an antenna [664] for the electromagnetic radiation of a far infrared laser the laser beam has to be coupled properly to the antenna pattern with good focussing, e.g., by use of a high quality microscope objective with long working distance. In the infrared the coupling of the radiation into the MIM diode also critically depends on the polarisation. When visible radiation is used the polarisation and the orientation of the laser beam with respect to the antenna are much less critical [665].

The characteristic of a MIM diode in general is not as simple as sketched in Fig. 11.1 c) but can show either a more symmetric or a completely anti-symmetric shape [666, 667]. The corresponding signal critically depends on the obtained characteristic, the wavelength of the

radiation and the electrical polarisation of the MIM diode by a forward or reverse bias. The preparation of a good MIM diode is tricky. Furthermore, the fragility of the MIM diode set-up comprising a thin wire which is exposed to high laser power coupled to the point contact, often allows one to use a particular MIM diode for only a few hours.

11.1.2 Schottky Diodes

More stable mixers and detectors can be obtained by using Schottky barrier diodes that are based on a metal-semiconductor transition. The small resistance necessary to allow for high-frequency applications is achieved by using a thin semiconductor layer (thickness about 0.1 µm) which is on top of a highly doped substrate. Commercial Schottky diodes typically comprise several hundred AuPt metal anodes of a diameter $1\,\mu\text{m} \lesssim d \lesssim 2\mu\text{m}$ within a SiO_2 mask on the active n-GaAs layer. A tungsten whisker is used to contact any one of the PtAu-GaAs diodes and serves as the antenna for coupling the microwave radiation. Such a diode structure is in general more stable than a MIM diode. More detailed comparisons between the Schottky and MIM diodes have been performed, e.g., in [665, 667].

11.1.3 Optical Second Harmonic Generation

In the optical regime non-linearities are generally weak. However, with the strong laser fields available, they give rise to a number of important effects. An electromagnetic wave interacting with the atomic or molecular systems in a dielectric medium can drive or induce electric microscopic dipoles. In a harmonic approximation the polarisation $P(E)$, i.e. the sum of all microscopic dipoles varies linearly with the field E (Fig. 11.1 a) and the induced electric dipole moments oscillate with the frequency of the electromagnetic wave. These oscillating dipoles are sources of electromagnetic radiation. However, in general, the polarisation $P(E)$ is non-linear (Fig. 11.1 b) and (11.2) is written as

$$P(E) = \epsilon_0 \left[\chi^{(1)} E + \chi^{(2)} E^2 + \chi^{(3)} E^3 + \cdots \right]. \tag{11.4}$$

Here, the $\chi^{(i)}$ are the susceptibilities describing the relevant processes of order i and usually become smaller with increasing i. In (11.4) it has been assumed implicitly that the polarisation depends on the instantaneous value of the field and that the medium does not have any "memory". Consider the quadratic term with the second-order susceptibility $\chi^{(2)} E^2$ which represents the special case of a tensor equation

$$P_i = \epsilon \sum_{j,k=1}^{3} \chi^{(2)}_{i,j,k} E_j E_k, \qquad i,j,k = 1,2,3. \tag{11.5}$$

The general case of the superposition of any two waves E_1 and E_2 produces the product terms $E_j E_k$ occurring in (11.5)

$$\begin{aligned} (E_1 + E_2)^2 &= E_{01}^2 \cos^2 \omega_1 t + 2E_{01}E_{02} \cos\omega_1 t \cos\omega_2 t + E_{02}^2 \cos^2 \omega_2 t \qquad (11.6) \\ &= E_{01}^2/2\,(1 - \cos 2\omega_1 t) + E_{02}^2/2\,(1 - \cos 2\omega_2 t) \\ &+ E_{01}E_{02}\left[\cos(\omega_1 - \omega_2)t - \cos(\omega_1 + \omega_2)t\right] \qquad (11.7) \end{aligned}$$

and includes terms with the doubled frequencies $2\omega_1$ and $2\omega_2$ as well as the sum and difference frequencies. In general, frequency conversion experiments in non-linear optical materials can be thought of as resulting from the interactions between three different light fields with frequencies ν_1, ν_2, ν_3 and the corresponding vacuum wavelengths $\lambda_i = c/\nu_i$. The three frequencies and wavelengths are constrained by energy conservation as

$$\nu_1 + \nu_2 = \nu_3 \qquad \text{or} \qquad \frac{c}{\lambda_1} + \frac{c}{\lambda_2} = \frac{c}{\lambda_3}. \tag{11.8}$$

Several processes can be identified:

Reading (11.8) from left to right describes the process of sum frequency generation (SFG) in which two photons of ν_1 and ν_2 annihilate thereby producing a new photon with the frequency of the sum of the two previous photons. In the case of $\nu_1 = \nu_2$ this process is called second harmonic generation (SHG). Reading (11.8) from right to left allows the generation of any "signal" photon of desired frequency ν_1 together with a so-called "idler" photon of frequency ν_2 from a "pump" photon of the higher frequency ν_3. This second process is realised in an optical parametric oscillator (OPO) (Section 9.3.3). A third process is the difference frequency generation (DFG) where the frequency of the resulting photon equals the difference of the frequencies of the two initial photons. Energy conservation requires that the resulting photon with difference frequency $\nu_1 - \nu_2$ is accompanied by two other photons as can be seen by writing $\nu_1 + \nu_2 = \nu_1 + (-\nu_2 + \nu_2) + \nu_2 = (\nu_1 - \nu_2) + 2\nu_2$.

11.1.3.1 Phase Matching

As an example consider the frequency doubling in a non-linear crystal. In a medium with a second-order non-linear susceptibility $\chi^{(2)}$ the electric field with a fundamental frequency ω_1 produces a polarisation wave at the second harmonic frequency $\omega_2 = 2\omega_1$. This polarisation wave travels with the same velocity as its generating fundamental wave. Hence, this velocity is determined by the index of refraction n_1 at the fundamental wavelength. The polarisation wave, however, gives rise to a second harmonic wave travelling with a velocity which is determined by the index of refraction n_2 at the wavelength of the second harmonic. In general, these indices of refraction are not the same and vary monotonically with the wavelength as is shown in Fig. 11.2 for the case of $LiNbO_3$. This material may serve as a representative example for the following discussion as it is commonly used for efficient frequency doubling. $LiNbO_3$ is an uniaxially birefringent crystal, where the two eigenmodes of polarisation, referred to as the ordinary and the extraordinary waves, in general have different phase velocities c/n_o and c/n_e (Fig. 11.2). The ordinary beam whose electric field vector is perpendicular to the optical axis has the same velocity, i.e. the same index of refraction in each direction (Fig. 11.3) thereby behaving as if in an isotropic medium. The phase velocity and consequently the index of refraction of the extraordinary beam varies monotonically between the extreme velocities of the two eigenpolarisations. In contrast, the extraordinary beam with its polarisation perpendicular to the ordinary beam travels with a velocity that depends on the angle between the optical axis and the direction of the beam. If the direction of the extraordinary beam coincides with the optical axis its polarisation is also perpendicular to the axis. Hence, for this case the velocities of the extraordinary and the ordinary beam are the same.

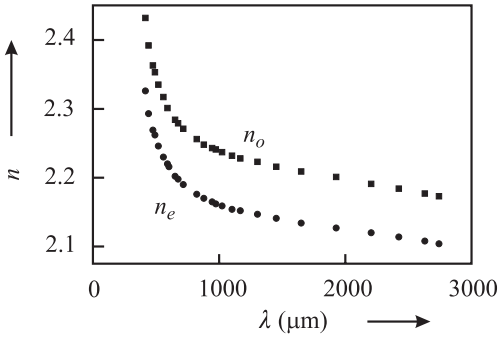


Figure 11.2: Ordinary refractive index n_o and extraordinary refractive index n_e of $LiNbO_3$.

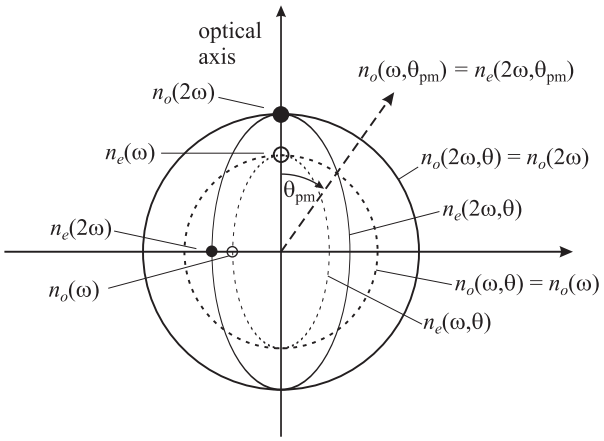


Figure 11.3: Cut through the surface of the refractive index in a negative uniaxial crystal ($n_e < n_o$) as, e.g., $LiNbO_3$. For the indicated crystal direction θ_{pm} the index of refraction of the extraordinary beam matches the index of refraction of the ordinary beam.

As a result of the, in general, different velocities of the fundamental wave and the second harmonic wave, the incremental second harmonic field generated in an incremental crystal element δz is out of phase by $\delta\phi$ with respect to the contribution to the field from the previous length element. The total field of the second harmonic contribution $E^{2\omega}(z)$ is obtained by adding all phasors of the incremental elements (see Fig. 11.4 a). After a length l_c called the coherence length both waves are out of phase by π and the power of the second harmonic wave falls to zero at 2 l_c (Fig. 11.5 a). The coherence length is calculated as follows

$$l_c = \frac{\pi}{\Delta k} = \frac{\lambda}{4(n_{2\omega} - n_\omega)}, \tag{11.9}$$

where λ is the wavelength in vacuum of the fundamental beam and where we have made use

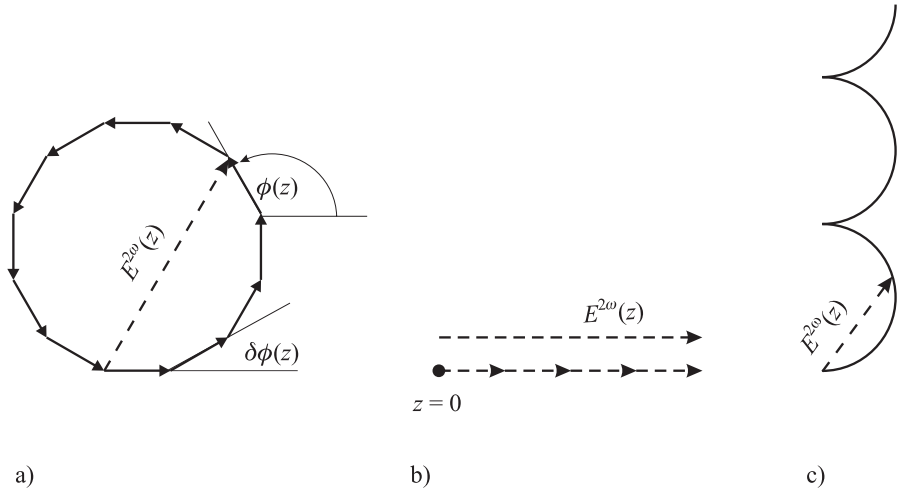


Figure 11.4: The phasor describing the evolution of the second harmonic wave $E^{2\omega}(z)$. a) In a non-phase-matched medium. b) For perfect phase matching. c) For quasi phase matching. In the latter case the polygon of a) has been replaced by a continuous curve of infinitesimal phasors.

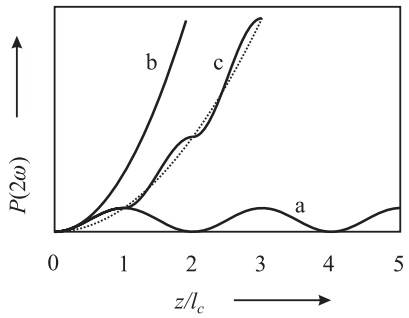


Figure 11.5: Increase of the second harmonic amplitude for phase mismatch between the fundamental and the second harmonic wave (a), for perfect phase matching (b) and for quasi phase matching (c).

of $\Delta k = k_{2\omega} - 2k_{\omega}$. The coherence length l_c is the optimum crystal length that is useful in producing the second harmonic power.[1] In order to achieve frequency doubled light over a long distance in the non-linear crystal the phases of the fundamental wave and the second harmonic wave have to be matched. Phase matching can be obtained in birefringent crystals that are characterised by two different indices of refraction depending on the orientation of the electric field vector with respect to the optical axis.

The so-called phase matching condition where $n(\omega) = n(2\omega)$ can be achieved for the direction in the crystal where the index of refraction of the ordinary beam matches the index of refraction of the extraordinary beam, i.e. $n_o(\omega) = n_e(\omega)$ (Fig. 11.3). In this case also the phase velocities of both waves of ω and 2ω are the same and consequently the energy flows continuously from the fundamental wave to the second harmonic wave. The electric field strength of the latter increases linearly with the distance travelled in the crystal (see Fig. 11.4 b)

[1] Sometimes [39] the coherence length l_c is defined as twice the value given in (11.9).

and the power increases quadratically (Fig. 11.5, curve b). For optimum phase matching the angle $\theta_{\rm pm}$ needs to be chosen accurately and consequently this type of phase matching is often referred to as critical phase matching. Besides this angular matching, in favourable cases at a certain temperature, the ellipsoidal surface of the extraordinary beam nestles against the sphere of the ordinary beam for the direction of beam propagation perpendicular to the optical axis. This "90° phase matching" or "non-critical phase matching" allows a particularly large conversion efficiency for two reasons. Firstly, there is no dependence of the conversion efficiency on beam divergence. Secondly, since no birefringence occurs the extraordinary beam will not walk off the ordinary beam and longer crystals can be used to increase the efficiency of the second harmonic generation.

11.1.3.2 Quasi Phase Matching

An elegant method referred to as quasi phase matching has been developed for efficient second harmonic generation [668, 669] which does not need to rely on the methods of critical or non-critical phase matching. The underlying idea can be understood from Fig. 11.4 c) and Fig. 11.5 c). Consider a crystal whose non-linear coefficient is periodically modulated such that the direction of the polarisation of the medium is changed at each integer multiple of l_c. This leads to a change of the phase of π for the second harmonic polarisation and, hence, for $E_{2\omega}$. The sum of all phasors representing the resultant contribution of all wavelets of the second harmonic generated in the quasi phase-matched crystal, increases continuously. Such a structure can be implemented, e.g., by periodically poling the crystal thereby creating a sequence of oppositely polarised optical domains. Stable domain inversion is achieved [669], e.g., by the tailored application of strong electric pulsed fields to the properly cut material where the spatial field pattern is determined by metallic masks with the desired periods. For a given length of the crystal and for the same non-linear coefficients, the second harmonic generation per unit length is not as efficient in quasi phase-matched devices as in phase-matched materials (Fig. 11.5). However, the method allows one to use non-linear coefficients with higher values from the diagonal elements of (11.5) that otherwise are not accessible for phase matching when relying on the natural birefringence. Second harmonic generation in suitable periodically poled materials has also been used to generate efficient coherent radiation in the UV [670].

11.1.4 Laser Diodes as Non-linear Elements

In laser diodes the processes including $\chi^{(2)}$ terms are of limited importance owing to the strong absorption bands in semiconductors where at least one of the contributing fields E_i is absorbed. As has been shown [40, 671] the relevant $\chi^{(3)}$ process is due to four-wave mixing. The three pump fields E_i with the three frequencies ν_i which are different in general can give rise to a fourth field E_4 with frequency ν_4. The allowed frequencies are dictated by energy conservation as follows

$$\pm h\nu_1 \pm h\nu_2 \pm h\nu_3 = h\nu_4. \tag{11.10}$$

Typically only one of the pump frequencies in (11.10) has a negative sign which leads to a fourth signal having a frequency close to that of the pump waves and, hence, lies in the trans-

parent spectral range of the laser diode. The origin of four-wave mixing in laser diodes for frequency differences of up to 30 GHz between the pump waves, results from interband modulation, i.e. by a modification of the carrier concentration. For higher frequency differences, spectral intraband modulation by spectral hole burning has been identified as the relevant mechanism [40]. Four-wave mixing in laser diodes has been demonstrated for frequency differences of the pump waves up to 3.1 THz and such laser diodes have been used for optical bisection of a frequency interval (Section 11.4.1).

11.2 Frequency Shifting Elements

For optical frequency measurements often elements are required that allow one to shift the frequency of an optical wave by an accurately determined frequency increment. The most important of such devices make use of acousto-optic or electro-optic effects.

11.2.1 Acousto-optic Modulator

Acousto-optic modulators based on materials such as $PbMoO_4$ or TeO_2 are characterised by a particularly high velocity v of the sound waves in the material, which are excited, e.g., by a piezoelectric transducer (Fig. 11.6). The sound wave of wavelength $\Lambda = 2\pi v/\omega_{\mathrm{sound}}$ modulates the density and, hence, the index of refraction n of the material. A laser beam

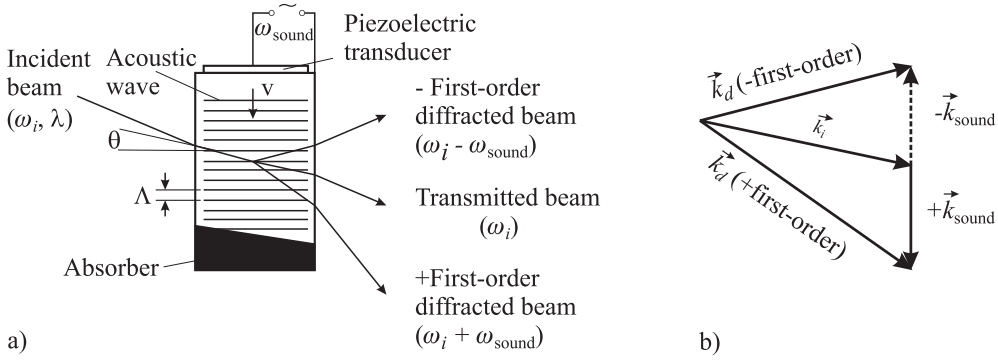


Figure 11.6: a) In an acousto-optic modulator the light wave is diffracted by a modulation of the index of refraction induced by a ultrasonic sound wave of angular frequency ω_{sound}. b) Momentum-conservation relation.

passing through the medium is diffracted by the periodic modulation of the index of refraction similarly to the case of an optical grating. In general, there is a big difference whether the light wave is deflected by a thin grating or by a thick one ($\lambda \times l > \Lambda^2$) where l is the thickness of the grating. The latter condition is referred to as the Bragg case. In this regime there is an exchange of energy and momentum between light waves and sound waves. In both waves the energy is quantised and depends linearly on the angular frequency ω of the wave. We refer to the quanta of the light wave and of the sound wave as photons $\hbar\omega_{\mathrm{photon}}$ and phonons $\hbar\omega_{\mathrm{sound}}$, respectively. The photons in the diffracted beam are deflected and consequently the

momentum $\hbar\vec{k}_d$ of each photon in the deflected beam differs from the momentum of that of the photons $\hbar\vec{k}_i$ in the incident beam. Conservation of momentum requires that this difference is supplied by the momentum of the phonons absorbed from the sound wave

$$\hbar\vec{k}_i + \hbar\vec{k}_{\text{sound}} = \hbar\vec{k}_d \tag{11.11}$$

(see Fig. 11.6 b)). Accordingly, energy conservation requires

$$\begin{aligned} \hbar\omega_i + \hbar\omega_{\text{sound}} &= \hbar\omega_d \qquad \text{or} \\ \omega_d - \omega_i &= \omega_{\text{sound}}. \end{aligned} \tag{11.12}$$

From (11.11) and Fig. 11.6 one finds that the angle between the undeflected laser beam (zero$^{\text{th}}$ order)and the diffracted beam depends on the frequency of the ultrasonic wave used for excitation since the transferred momentum increases with increasing angular frequency ω_{sound}. Commercial acousto-optic modulators in general are operated with frequencies $40\,\text{MHz} \lesssim \omega_{\text{sound}}/(2\pi) \lesssim 0.5\,\text{GHz}$. To obtain high diffraction efficiency the modulators are in general operated at low orders of diffraction resulting in an efficiency of up to about 85 % depending on the power of the ultrasonic wave. Consequently, this power dependence is readily used to modulate or adjust the amplitude of the transmitted and the diffracted laser beam by the power used to drive the piezo actuator. Moreover, according to (11.12) the frequency of the diffracted laser beam can be varied by the frequency of the oscillator driving the piezoelectric transducer. The use of positive or negative diffraction order allows an increase or decrease in the frequency of the light wave (Fig. 11.6). The frequency shifted and deflected beams can be easily separated from the undeflected laser beam even though the variation of the angle of the deflected beam with the variation of the frequency is sometimes disadvantageous in applications. The deflection angle $\alpha = 2\theta$ is calculated from the vector equation (11.11) as

$$\sin\theta = \frac{k_{\text{sound}}}{2k_i} = \frac{\lambda\nu_{\text{sound}}}{2v} \tag{11.13}$$

where $k_i \approx k_d$ has been used. From (11.13) the deflection angle of a laser beam ($\lambda = 633\,\text{nm}$) in an acousto-optic modulator made of $PbMoO_4$ (velocity of sound $v = 3650$ m/s) operated at 80 MHz is calculated as $\alpha \approx 2\sin\theta$ = 13.9 mrad.

From (11.12) we find (by multiplying with the time t) an equation that relates the phase of the diffracted light beam to the phase of the sound wave. Hence, the acousto-optic modulator can also be used as a phase shifter for the diffracted light wave.

11.2.2 Electro-optic Modulator

In certain crystals it is possible to modify the propagation properties of electromagnetic radiation by an applied electric field in a number of ways (see e.g. [39]). Consider a birefringent crystal like $(NH_4)H_2PO_4$ (ADP), $LiTaO_3$, $LiNbO_3$, etc., cut and used in such a way that the optical axis of the crystal is perpendicular to the direction of incidence of the laser beam (Fig. 11.7). A linearly polarised laser beam is split into two orthogonally polarised beams called the ordinary beam (which obeys Snell's law of refraction) and the extraordinary beam. The ordinary and the extraordinary beams travel with different velocities according to their different indices of refraction n_o and n_e, respectively. When an electric field is applied to

the electro-optic medium the centre of charge of the binding electrons may be displaced with respect to the ionic cores. The resulting polarisation of the material leads to a modification of the index of refraction. The anisotropic response of the crystal to an applied electric field in general has to be described by a tensor. For simplicity, in Fig. 11.7 the direction of the applied field E_z coincides with the direction of the polarisation of the extraordinary beam. The index

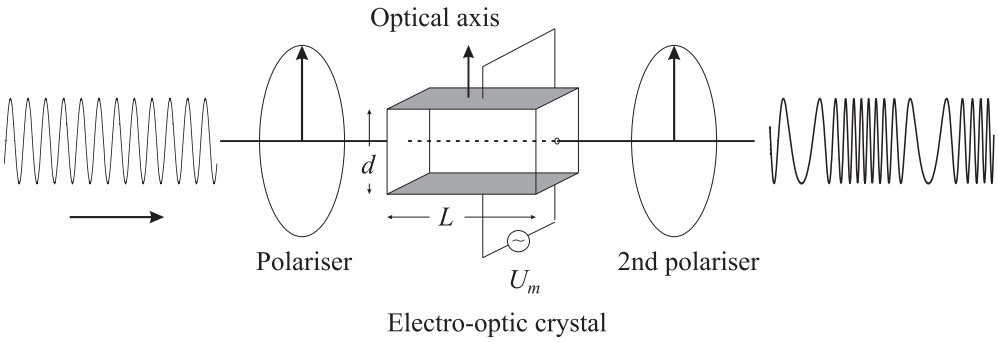


Figure 11.7: Electro-optic crystal operated as a transverse phase modulator.

of refraction n along this direction is modified by the applied electric field E_z as

$$n = (n_e - \tfrac{1}{2} n_e^3 \, r_{zz} \, E_z). \tag{11.14}$$

Here, r_{zz} is a diagonal element of the tensor describing the response of the crystal along the z direction if the field is applied along this z direction. The phase of the light wave is shifted in the crystal by

$$\delta\phi = \frac{2\pi n}{\lambda} L = \frac{n_e^3 \, r_{zz}}{\lambda} \, \frac{L}{d} \, \pi \, U_m \equiv \pi \, \frac{U_m}{V_\pi}, \tag{11.15}$$

when the electric field is realised by applying a voltage $U_m = dE_z$. The so-called "half-wave voltage"

$$V_\pi = \frac{\lambda}{n_e^3 \, r_{zz}} \, \frac{d}{L}, \tag{11.16}$$

refers to the voltage that changes the phase of the light wave after travelling through the crystal by $\delta\phi = \pi$. Such an electro-optic device can serve different purposes. It is readily used as a fast servo element for frequency stabilisation of lasers (Section 9) which allows one to vary the optical path length $n\,L$, and hence to counteract technical path length fluctuations. The device is furthermore used as an electro-optic (phase) modulator by applying a sinusoidal modulation voltage $U_m = U_0 \sin \omega_m t$ to modulate the phase of the light wave. In this case, $\pi U_0 / V_\pi$ (11.15) represents the modulation index and besides the carrier, the phase-modulated laser beam comprises sidebands whose amplitudes depend on the modulation index according to (2.52).

In a different arrangement the electro-optic modulator can be used for fast amplitude modulation of the light field, similarly to a Pockels cell. Consider a linearly polarised laser beam

directed to a crystal whose fast (slow) axis is now tilted by 45° with respect to the polarisation direction of the laser beam (Fig. 11.8). In this arrangement the different velocities of the two partial waves of different polarisations result in a phase shift between the two waves behind the crystal depending on the voltage applied to the crystal. With respect to the polarisation of the input beam a phase difference of 0° and 180° leads to linearly polarised light with the polarisation axes being parallel and perpendicular, respectively. A phase difference of 90° and 270° leads to circularly polarised light and all other phase differences result in elliptically polarised light. Placing a polariser behind the exit of the electro-optic modulator (Fig. 11.8) allows one to adjust the laser power by an applied voltage. Both, electro-optic

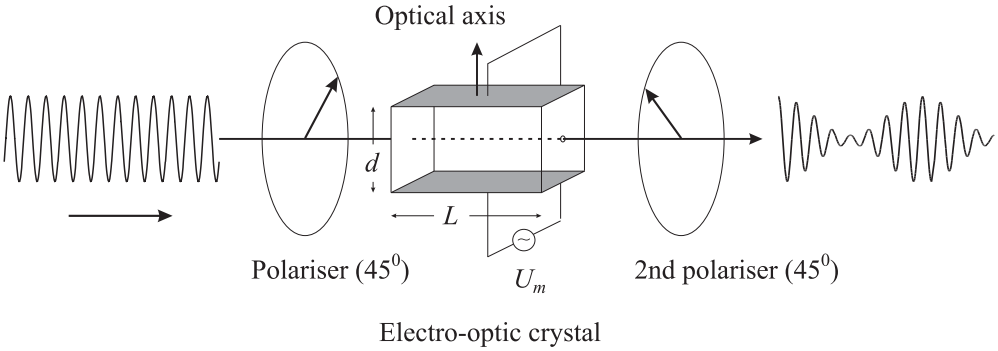


Figure 11.8: Electro-optic crystal operated as a transverse amplitude modulator.

modulators operated as amplitude modulators, and acousto-optic modulators are often used as a "noise eater" to eliminate fluctuations of the laser power. In such an arrangement part of the transmitted laser power is split off and directed onto a photodiode. The difference between the photo current and the reference current of a constant current source can be utilised as an error signal for an electronic servo system which uses the electro-optic modulator as a servo element to keep the laser power constant.

In optical frequency standards electro-optic modulators are often used as phase modulators with the requirement that no amplitude modulation occurs (Section 9.2.2). In practical devices, however, phase modulation is often accompanied by a certain contribution of amplitude modulation. Even for well aligned polarisation, amplitude modulation hardly can be avoided since all materials with non-vanishing electro-optic coefficients are also to some extent piezoelectric. As a consequence, the applied electric ac field used to modulate the phase of the light wave at the same time excites vibrations of the material. These vibrations in general also modulate the indices of refraction via the elasto-optic effect. Any resulting modulation in the direction of the beam in turn may be converted into amplitude modulation occurring, e.g., at a limiting aperture.

11.2.3 Electro-optic Frequency Comb Generator

The sidebands produced by an electro-optic modulator can be used to bridge frequency differences in the optical regime considerably larger than the difference frequencies that can be measured directly with photodiodes. The maximal frequency separation can be increased by

either using a high radio frequency or by producing a large number of sidebands. The first route has been followed by Kallenbach *et al.* [672] who placed an electro-optic modulator into a resonant microwave cavity to produce sidebands at 72 GHz. The second approach uses resonant microwave cavity in a resonant optical cavity where each sideband can again be the origin of new sidebands thereby increasing the total number of sidebands considerably (Fig. 11.9). In such an optical frequency comb generator any two adjacent frequencies are

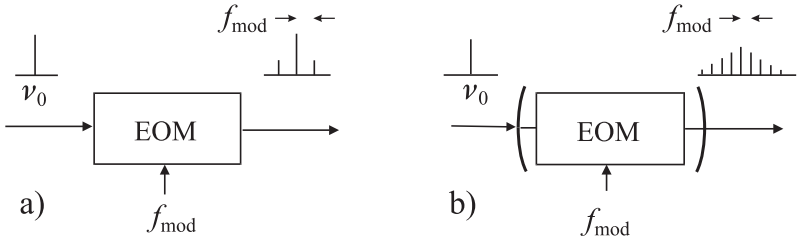


Figure 11.9: The sidebands created in an electro-optic modulator a) generate additional sidebands in a resonant optical cavity b), where the modulation frequency corresponds to an integer value of the free spectral range.

separated by the modulation frequency and there is a well defined phase relationship between all waves. Electro-optic frequency comb generators have been developed [673, 674] that span a frequency range of several terahertz. Telle and Sterr [675] have calculated the relative power of the k^{th} sideband with respect to the power of the carrier P_c as

$$\frac{P_k}{P_c} = \exp\left(-\frac{\pi|k|}{\delta \times F^*}\right) \tag{11.17}$$

where δ is the single-pass phase modulation index and F^* is the finesse of the optical resonator corresponding to -13.64 dB/($\delta \times F^*$) per sideband order. Assuming $\delta = 0.5$, $F^* = 100$ and $f_{\mathrm{mod}} = 9.2$ GHz the power of the comb varies like -30 dB/THz. With high enough power in the carrier and a sensitive detector, several terahertz can be bridged. At a frequency of several terahertz, however, the sideband power abruptly drops as a consequence of the group velocity dispersion [676, 677] where the frequencies of the sidebands no longer coincide with an integer multiple of the free spectral range. As the material dispersion typically decreases for longer wavelengths the span of 3 THz at 1.064 µm [677] increases to 7.7 THz at 1.54 µm [678] and may be more than 20 THz above 1.8 µm [678]. These values can be further increased if the cavity frequency is locked slightly off resonance from the laser frequency [677, 678]. Brothers and Wong [677] have utilised an intracavity prism pair to partially compensate the material dispersion of a lithium niobate modulator in order to increase the span of the comb at 1064 nm from 3.0 THz to 4.3 THz. The width of the comb can be furthermore expanded by non-linear interactions, e.g., by self-phase modulation in a fibre. Second harmonic generation in a periodically poled quasi-phase matched $LiNbO_3$ (Section 11.1.3.2) material has yielded a comb of 22 THz width [678]. Electro-optic comb generators have been used to measure frequency gaps for optical frequency measurements [679]. For frequency standard applications, typically only a particular sideband is desired. Ye *et al.* [680] used a three-mirror cavity with

two tuneable mirrors. With this set-up they could couple out the desired sideband resonantly, while keeping all other sidebands confined in the resonator for continued comb generation.

11.3 Frequency Synthesis by Multiplication

The comparison of the frequency of an optical laser oscillator with the frequency of an oscillator in the microwave region was for a long time not an easy task due to the large frequency ratio of about 10^5 between both domains. In the visible, one of the very first measurements of an optical frequency led to the frequency of a Lamb-dip stabilised He-Ne laser where the difference and the ratio of two optical frequencies were determined resulting in a relative uncertainty of 6×10^{-8} [681]. The measurement of the frequency of the iodine-stabilised He-Ne laser (λ = 633 nm; Section 9.1.3) was reported in 1983 [682]. In the following years several so-called "frequency measurement chains" have been developed connecting microwave clocks to oscillators in the infrared [391, 392, 397], from the infrared to the optical regime [683] or directly from the Cs clock to the optical regime [103, 504, 684]. These chains were based on a large number of intermediate oscillators with different frequencies that were compared and phase locked by harmonic mixing in non-linear elements and beat frequency measurements, respectively.

Optical frequencies can be measured by a technique well known from the measurement of microwave frequencies. There, the known frequency ν_1 of an oscillator is used to produce higher harmonics in a non-linear element. If a suitable harmonic $n\nu_1$ is close enough to the unknown frequency ν_2 of another oscillator the beat note $\delta\nu$ between the frequency of the latter and the n^{th} harmonic of ν_1 can be used to determine the previously unknown frequency by

$$\nu_2 = n\nu_1 \pm \delta\nu. \tag{11.18}$$

The plus or minus sign in (11.18) can easily be determined by monitoring the change of the beat frequency resulting from an increase of the frequency ν_1. Several stages of frequency multiplication can be combined to a so-called frequency multiplication chain to link a microwave frequency to a frequency in the optical domain. The non-linear devices used in the different steps of such a chain are Schottky diodes ($\nu \leq 5$ THz), metal-insulator-metal (MIM) diodes ($\nu \leq 120$ THz), and non-linear crystals ($\nu > 120$ THz). Fig. 11.10 shows an example of such a frequency multiplication chain that was used to reference the optical frequency (456 THz) of a Ca-stabilised laser to the frequency (9.2 GHz) of the Cs atomic clock [504, 667]. At the low-frequency end of the chain, the frequency of a hydrogen maser is locked to the frequency of a primary Cs atomic clock. This combination joins the long-term accuracy of the Cs atomic clock with the short-term stability of the hydrogen maser thereby allowing one to perform precise measurements at short integration times. The quartz oscillator at 100 MHz serves as a "flywheel" with a phase noise of $S_\Phi \approx -170\,\text{dB}_c$ at 10 kHz from the carrier. Its frequency is multiplied, driving a step-recovery diode which produces harmonics near the 22.7 GHz of the Gunn oscillator. The 17^{th} harmonic of the 22.7 GHz signal at 386 GHz derived from an F-band harmonic mixer is mixed in a Schottky diode with the radiation from a backward wave oscillator to lock the phase of the latter to the signal from the Gunn oscillator. Part of the radiation from a methanol laser at 4.25 THz is coupled to a Schottky barrier diode together with the radiation from the backward wave oscillator using

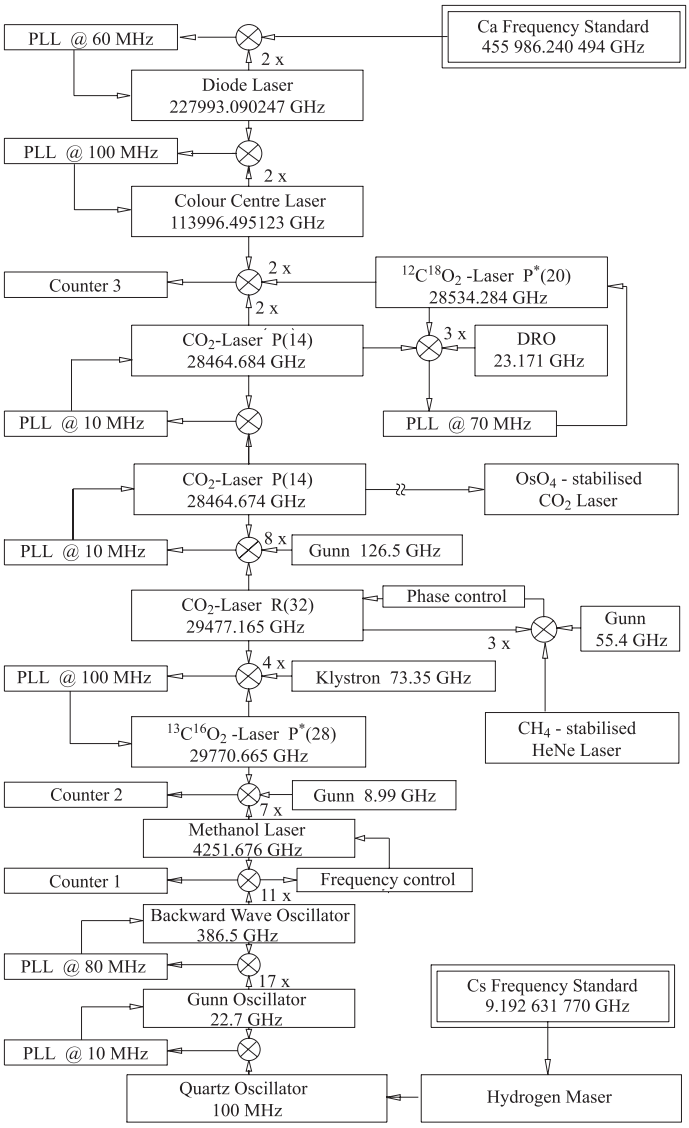


Figure 11.10: Frequency multiplication measurement chain based on the synthesis of frequencies according to Schnatz *et al.* [504]. PLL: Phase Lock Loop; DRO: Dielectric Resonator Oscillator.

the beat between its eleventh harmonic and the radiation from the methanol laser to keep the frequency of the latter loosely locked. In the next step, the beat note between the seventh harmonic signal of the methanol laser and a CO_2 laser is measured. The frequency of this laser is again locked via a series of other CO_2 lasers to two CO_2 lasers (CO_2 P(14) laser and

$^{12}C^{18}O_2$ P*(20) laser) at 28.5 THz, using MIM diodes. The sum of the frequencies of two photons from each of these lasers is very close to the frequency of a photon from a colour centre laser (KCL:Li) operated at 2.6 µm. The mixing signal at about 1.4 GHz from a MIM diode is counted. The next multiplication step uses second harmonic generation in an angle matched $AgGaS_2$ crystal. The beat note between this radiation and the radiation of a diode laser operating at about 1314 nm and pre-stabilised to a Fabry-Pérot interferometer is measured with a photodetector and used to phase lock the frequencies of the colour centre laser and the diode laser. Eventually, the radiation of the diode laser is frequency-doubled in a temperature-tuned Lithium-Beta-Borate (LBO) crystal. LBO allows non-critical phase matching at 1314 nm near 12 °C where the fundamental wave propagates as an ordinary beam and its polarisation is perpendicular to the second harmonic extraordinary beam (Type I phase matching). The doubled frequency is very close to the frequency of the Ca stabilised laser and the beat frequency between both radiations is measured with a photodetector. The frequency of the diode laser is controlled by a phase lock loop in such a way that their beat note is kept constant. Hence, the frequencies of the two oscillators, i.e. the Ca stabilised laser and the diode laser at 1 314 nm, are phase coherently related to each other.

The upper part of the frequency chain of Fig. 11.10 is phase locked to the Ca stabilised laser, the lower part to the Cs atomic clock. Thus if both sub-chains are phase coherently locked, the beat frequency measured by use of counter 3 together with all offset frequencies and multiplication factors can be used to obtain a truly phase coherent frequency ratio of the Ca stabilised laser and the Cs clock.

The far infrared methanol laser in the frequency chain of Fig. 11.10, however, is not directly locked to the adjacent stages by phase lock loops. This laser in general is tricky to operate [667] and some peculiarities do not allow to phase lock it directly. Nevertheless, the phase coherence in the chain can be preserved by the "transfer oscillator concept" [667, 685, 686]. To understand this concept, consider the upper and lower part of the chain to produce stable frequencies at the mixers above and below the methanol laser in Fig. 11.10. If the beat note measured by counter 2 between the seventh harmonic of the methanol laser and the CO_2 laser is electronically divided by a factor of seven the frequency fluctuations of the resulting signal are those of the methanol laser at counter 1. The same fluctuations with opposite sign show up in the beat note between the methanol laser and the eleventh harmonic from the backward wave oscillator. In the sum of both frequencies the frequency fluctuations of the methanol laser are eliminated. This elimination can be performed either afterwards by an electronic computation of the two readings from the counters, or by a direct electronic implementation using mixers and dividers [686]. Experimentally, the frequency of the transfer laser (methanol laser) is loosely locked to keep the beat frequencies within the filter bands of the employed electronic circuits. Using purely electronic means to track the phase of a pre-stabilised laser rather than to phase lock a laser in general seems the better choice, since the tracking range provided by an inertia-free electronic system can hardly be surpassed.

Measurements of different Ca frequency standards have been performed over several years [495, 504]. The mean value of all frequency measurements of ν_{Ca} = 455 986 240 494.13 kHz with a total relative uncertainty of 2.5×10^{-13} agrees well with more recent frequency measurements with femtosecond comb generators [501, 505]. The frequency measurement chain of Fig. 11.10 could also be used to measure simultaneously the frequencies of the CH_4 stabilised He-Ne laser at 3.39 µm (Section 9.1.4; [392, 667]) and the OsO_4 stabilised CO_2 laser at 10.6 µm (Section 9.1.5).

Frequency measurements based on multiplication in non-linear elements have been also used to connect novel frequency standards in the visible or near infrared to known frequency standards (see e.g. [475, 687]).

11.4 Optical Frequency Division

Apart from the conventional frequency multiplication techniques discussed so far, there is a completely different approach for the measurement of large optical frequency differences. This method divides an optical interval, with an accurately known factor, into a frequency interval that can be measured directly. If the frequency interval spans a complete octave, i.e. if it extends from ν to 2ν, the frequency ν is determined directly as $\nu = 2\nu - \nu$.

11.4.1 Frequency Interval Division

An optical frequency interval divider, that is capable of bisecting an optical frequency interval, has been invented by Telle, Meschede and Hänsch [688]. Consider the frequency interval $\nu_1 - \nu_2$ generated by two lasers of frequencies ν_1 and ν_2 (Fig. 11.11).

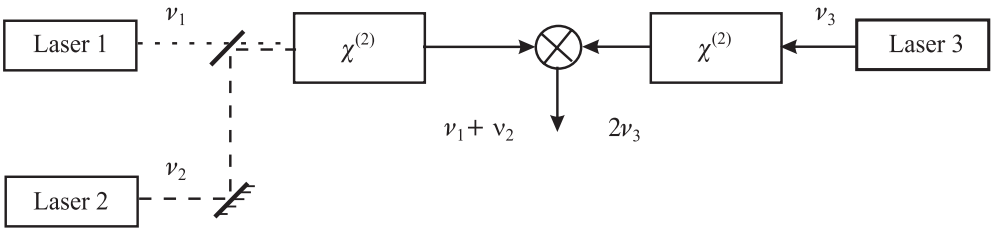


Figure 11.11: Division of an optical frequency interval according to Telle *et al.* [688].

In a suitable non-linear material with a second-order non-linear susceptibility $\chi^{(2)}$ the beams of these two lasers produce a wave with the sum frequency $\nu_1 + \nu_2$. This new wave is mixed with the frequency $2\nu_3$ obtained by second harmonic generation in a second non-linear device, from the beam of a third laser with frequency ν_3. If this frequency is adjusted in such a way that the beat note from the photodetector is zero, it follows that

$$\nu_1 + \nu_2 = 2\nu_3 \qquad \text{or} \qquad \nu_3 = \frac{\nu_1 + \nu_2}{2}. \tag{11.19}$$

In this case ν_3 is located exactly halfway between the frequencies ν_1 and ν_2. Any one of the two new frequency intervals between ν_1 and ν_3 or between ν_3 and ν_2 again can be bisected in the same way. This technique allows one to divide the initial interval to any desired frequency difference with the help of the required number of cascaded interval divider stages. A divider chain comprising n stages reduces the initial frequency difference by 2^n. It has been proposed to measure optical frequencies directly by comparison with the frequency of a microwave frequency standard using a divider chain where the initial frequency interval corresponds to an optical octave, i.e. $\nu_2 = 2\nu_1$. Such a situation is depicted in Fig. 11.12 for the frequency of the sub-harmonic of the Ca stabilised laser where the frequency ν_2 is the second harmonic

of ν_1. The number of necessary divider stages depends on the ability to measure the final interval.

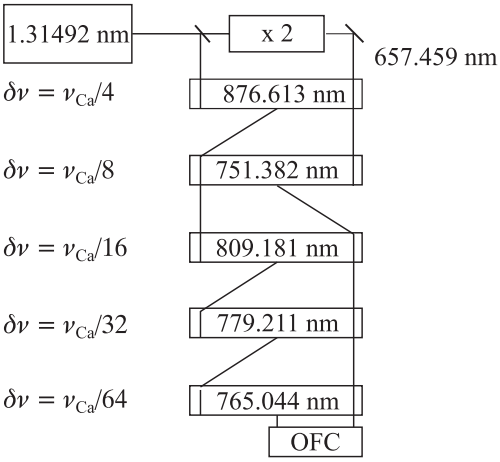


Figure 11.12: Frequency division chain based on frequency interval division. OFC: Optical Frequency Comb generator.

With the advent of optical frequency comb generators (Section 11.2.3) large optical frequency differences could be measured with a small number of divider stages. The wavelength range at the final frequency difference measurement is chosen depending on the availability of suitable laser diodes and non-linear crystals. The same starting points as in Fig. 11.12 have been used in a proposal [689] to end in a wavelength region around 875 nm. Similar schemes have been used, e.g., to measure optical frequencies by comparison with other optical or infrared standards [394, 395, 690].

11.4.2 Optical Parametric Oscillators as Frequency Dividers

A different divider scheme (Fig. 11.13) has been devised by Wong [691]. The scheme uses a laser with frequency ν_1 that is divided by an optical parametric oscillator (2:1) to yield a frequency $\nu_1/2$ and by a second one (3:1) to yield $2\nu_1/3$ (and at the same time $\nu_1/3$). The frequency of a second auxiliary laser (laser 2) is adjusted such that its frequency ν_2 is near to the frequency $2\nu_1/3$ derived from the first laser allowing the difference frequency x to be measured as a beat note. The frequency $2\nu_2/3$ is derived from the auxiliary laser by means of a third optical parametric oscillator (3:1) to allow measurement of the difference y. The frequencies are related by

$$\nu_2 = \tfrac{2}{3}\nu_1 + x \qquad \text{or} \qquad \tfrac{2}{3}\nu_2 = \tfrac{4}{9}\nu_1 + \tfrac{2}{3}x, \tag{11.20}$$

and by

$$\tfrac{2}{3}\nu_2 = \tfrac{1}{2}\nu_1 - y. \tag{11.21}$$

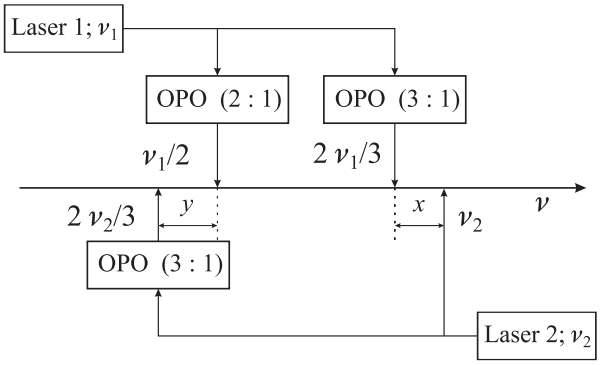


Figure 11.13: Frequency division scheme using three optical parametric oscillators (OPO) according to Wong [691].

Equating (11.20) and (11.21) one derives

$$\begin{aligned} \tfrac{4}{9}\nu_1 + \tfrac{2}{3}x &= \tfrac{1}{2}\nu_1 - y \qquad \text{or} \\ \nu_1 &= 12x + 18y \end{aligned} \tag{11.22}$$

which allows one to compute the frequency of laser 1 directly from the two measured beat notes x and y. Schemes to measure optical frequencies with optical parametric oscillators by phase coherent links to the microwave region have been worked out [692] and the ability to perform cycle-slip free measurements has been demonstrated to an accuracy of 5×10^{-18} [693].

11.5 Ultra-short Pulse Lasers and Frequency Combs

Frequency combs much wider than the ones produced by electro optic generators Section 11.2.3 can be obtained from ultra-short pulse lasers. A periodic train of ultra-short light pulses emitted by a continuous wave mode-locked laser with repetition frequency $f_{\rm rep}$ gives rise to a comb of equidistant frequencies. This is easily understood from the corresponding statement that the coherent superposition of harmonic signals with frequencies differing by the same constant frequency offset $\Delta\omega$ results in a periodic signal with pulses in the time domain (Fig. 11.14) separated by $T = 2\pi/\Delta\omega = 1/f_{\rm rep}$. An example is shown in Fig. 11.14 where the periodic pulsed signal resulting from 21 equidistant frequencies has been calculated. To find the envelope of N pulses one calculates

$$\begin{aligned} E(t) = \sum_{n=0}^{N-1} e^{i(\omega_0 + n\Delta\omega)t} = e^{i\omega_0 t} \sum_{n=0}^{N-1} e^{in\Delta\omega t} = e^{i\omega_0 t} \left[\sum_{n=0}^{\infty} e^{in\Delta\omega t} - \sum_{n=N}^{\infty} e^{in\Delta\omega t} \right] \\ = e^{i\omega_0 t} \left[\frac{1}{1 - e^{i\Delta\omega t}} - e^{iN\Delta\omega t} \frac{1}{1 - e^{i\Delta\omega t}} \right] = \frac{1 - e^{iN\Delta\omega t}}{1 - e^{i\Delta\omega t}} e^{i\omega_0 t}, \end{aligned} \tag{11.23}$$

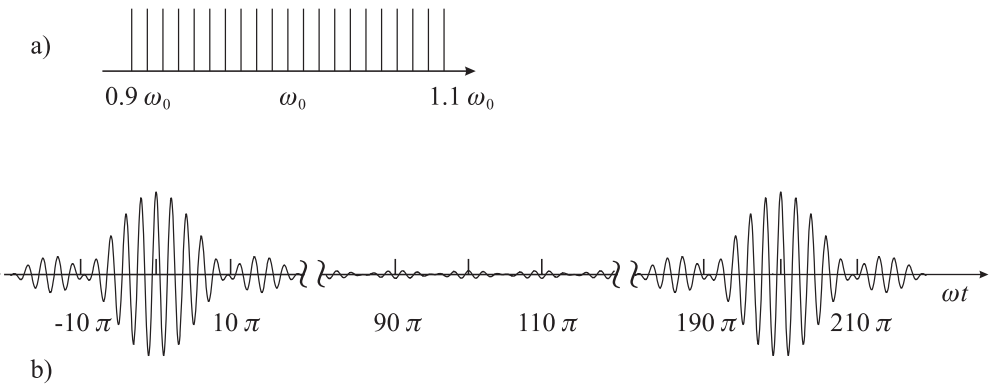


Figure 11.14: Comb of 21 frequencies separated by $\Delta\omega = 2\pi f_{\mathrm{rep}} = 0.1\omega$, a) and corresponding pulse trains in the time domain, b) calculated from the phase-coherent superposition of the 21 equidistant amplitudes of a).

where use has been made of $\sum_{n=0}^{\infty} q^n = 1/(1-q)$ for $|q| < 1$. The corresponding irradiance is given by

$$I(t) \propto |E(t)|^2 = \frac{1-\cos N\Delta\omega t}{1-\cos \Delta\omega t} = \frac{\sin^2 N\Delta\omega t/2}{\sin^2 \Delta\omega t/2}. \tag{11.24}$$

From (11.24) one can estimate the width of the pulse for $N \gg 1$ by finding the first zero of the nominator that occurs at $\pm N\Delta\omega t_0/2 = \pi$ corresponding to a full base width of $2t_0 = 4\pi/(N\Delta\omega)$. The full width at half maximum of the pulse τ_p is approximately half of this value

$$\tau_p \approx \frac{2\pi}{\Delta\omega N}. \tag{11.25}$$

From (11.25) one finds that the pulses become shorter the more frequencies (N) that contribute to the signal. Equivalently (11.25) can be interpreted in such a way that the spectral width of the comb scales inversely with the pulse duration. With the advent of femtosecond mode-locked Ti:sapphire lasers, frequency combs can be generated that cover a considerable fraction of the optical spectrum.

11.5.1 Titanium Sapphire Laser

The Ti:sapphire laser shows an exceptionally large gain between about 670 nm and about 1100 nm. The active medium of this laser is formed by Ti^{3+} ions in a sapphire crystal (Al_2O_3) where they replace a significant fraction of the Al^{3+} ions. The crystal is doped with as much as a few times 0.1 weight percent titanium. The energy levels of the Ti^{3+} ion are strongly shifted by the Stark effect due to their bonding to the neighbouring ions in the lattice. The $3d^1$ configuration of an ion resulting from the single electron in the outer shell is split by the cubic part of the crystal field into the ^2E and $^2T_2$ states (Fig. 11.15).

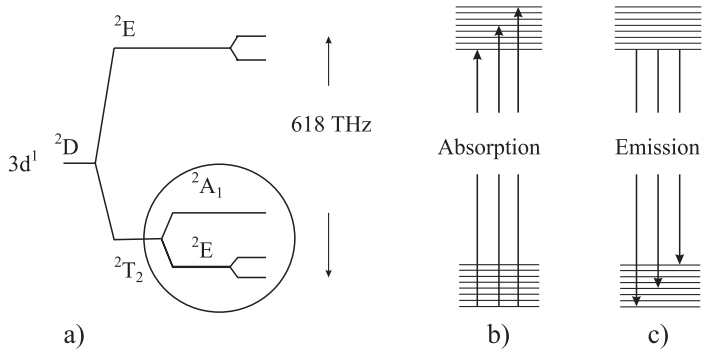


Figure 11.15: The splitting of the energy levels of the Ti^{3+} ion in the sapphire crystal leads to broad absorption and emission bands.

The trigonal part of the crystal field, together with the spin-orbit coupling leads to further splitting. As a result of the interaction with lattice oscillations these levels are strongly broadened and overlap. In essence, there are two well separated bands. Absorption of photons can take place from the bottom of the lower band to any energy level in the upper band. The resulting broad absorption band of Ti:sapphire has a maximum near 500 nm allowing these levels to be pumped with about 50 % efficiency by an argon ion laser or a frequency doubled Nd:YVO_4 laser. The excited Ti^{3+} ions rapidly decay to the lowest levels of the upper band by non-radiative processes. There, the lifetime is about 3 µs at 20 °C and radiative emission can occur to any state in the lower band. The large width of this emission band is prerequisite to the generation of ultra-short pulses provided that there is an efficient mode coupling process.

11.5.2 Mode Locking

If a broad-band frequency comb is to be generated in a laser, all laser modes contributing to the comb have to have a well defined and constant phase relationship. This locking of the modes with different frequencies separated by the free spectral range of the laser resonator can be achieved by active or passive means. The more modes that are locked, the shorter are the pulses emitted by the laser. In fact, the shortest pulses of around 5 fs are obtained by passive mode locking [694].

11.5.2.1 Active Mode Locking

The active mode locking schemes often employ acousto-optic (AOM) or electro-optic modulators (EOM) to modulate the losses in the laser resonator. Consider an intracavity AOM driven by a sinusoidal voltage of frequency f where the first-order diffracted beam is coupled out of the resonator. Photons passing the AOM in phase with zero applied voltage will suffer from minimal losses if their round trip time in the laser resonator of perimeter z corresponds to the inverse repetition rate $f_{\mathrm{rep}} = c/z$. In each pass the power in the leading and trailing tails of the pulse is reduced with respect to the central part. The repetitive shaping and amplification of the pulse circulating in the resonator leads to a shortening of the pulse and

a pulse train is emitted by the laser. With acousto-optic modulators, continuous pulse trains with pulse times below 100 ps can be achieved. Shorter pulses can be obtained with passive mode locking schemes using saturable absorbers or Kerr-lens mode locking [69, 695].

11.5.2.2 Saturable Absorbers

Efficient passive mode locking in lasers is achieved when saturable absorbers, e.g., organic dyes or semiconductors are placed in the laser resonator. The transmission of a saturable absorber depends on the irradiance of the laser beam and the absorber becomes transparent at high power. A laser with a saturable absorber begins to oscillate with the different modes unlocked. Fluctuations with higher intensities corresponding to an in-phase superposition of different modes lead to higher transmission of the saturable absorber and, hence, suffer from lower losses, thereby being amplified predominantly in the subsequent circulations. Any other fluctuating mode acquiring the proper phase with respect to the modes already phase-locked will contribute to even higher pulse power thereby further reducing the absorption in the saturated absorber. Semiconductor saturable absorber mirrors (SESAMs) [696, 697] are sometimes used to achieve self-starting of mode locked lasers.

11.5.2.3 Kerr Lens Mode Locking

Kerr lens mode locking is based on the optical Kerr effect which represents a non-linear change in the index of refraction with increasing optical irradiance. It is based on the third-order term $\chi^{(3)}E^3$ of (11.4). Unlike the second-order susceptibility term which is identically zero in centrosymmetric materials, the third-order term is non-vanishing in all optical materials. Keeping the linear and third-order term of (11.4) one obtains for the displacement resulting from the influence of the electric field

$$D = \epsilon_0 E + P = \epsilon_0 \left(1 + \chi^{(1)}\right) E + \chi^{(3)} E^3 = \epsilon_0 \left[1 + \chi^{(1)} + \epsilon_0^{-1} \chi^{(3)} E^2\right] E. \quad (11.26)$$

The quantity in square brackets in (11.26) can be regarded as a non-linear dielectric "constant"

$$\epsilon' = \epsilon_1 + \epsilon_2 E^2 \quad (11.27)$$

with a linear dielectric constant $\epsilon_1 \equiv 1 + \chi^{(1)}$ and a non-linear contribution $\epsilon_2 \equiv \chi^{(3)}/\epsilon_0$. Using $n = \sqrt{\epsilon'}$ one derives from (11.27) the index of refraction as $n \approx n_0 + n_2' E^2$ with $n_2' = \epsilon_2/2$ or

$$n \approx n_0 + n_2 I. \quad (11.28)$$

Hence, the index change is proportional to the irradiance I of the laser beam. The non-linear index of refraction, e.g., for glass used in optical fibre material and for sapphire near 800 nm is $n_2 \approx 10^{-16}\text{cm}^2/\text{W}$ [69] and $n_2 \approx 3.2 \times 10^{-16}\text{cm}^2/\text{W}$ [698], respectively. The propagation of a light pulse in a Kerr medium is affected in two different ways referred to as the transverse Kerr effect and the longitudinal Kerr effect. The transverse spatial variation of the irradiance in a Gaussian beam leads to a spatial variation of the phase shift in the Kerr medium that acts in the same way as does a lens. This so-called Kerr lens leads to a self-focussing of the laser

beam, focussing the high-irradiance part of the beam more strongly than the low irradiance part. The Kerr lens can be used to achieve passive mode locking since an aperture behind the Kerr medium produces less loss for the high irradiance. Similarly to the case of the saturable absorber, mode-locking occurs since it leads to higher power and less loss at the aperture. In the Ti:sapphire laser, the crystal, a few millimetres in length, serves as both the laser medium and the Kerr medium. Mode locking can also be achieved by optimising the overlap of the pump beam with the mode of the resonator, taking into account the Kerr lens acting as a soft aperture. Practical hints for optimisation of the mode locking can be found in [699].

The longitudinal Kerr effect is based on the temporal irradiance dependence in an optical pulse. The temporal envelope of the pulses emitted by a femtosecond laser can often be approximated by a hyperbolic secant pulse

$$E(t) = \frac{1}{\pi\tau_p}\,\mathrm{sech}\left(\frac{t}{\tau_p}\right) e^{i\omega_0 t} = \frac{1}{\pi\tau_p}\,\cosh^{-1}\left(\frac{t}{\tau_p}\right) e^{i\omega_0 t}. \tag{11.29}$$

In comparison with the Gaussian pulse

$$E(t) = \frac{1}{\sqrt{2\pi}\tau_p}\exp\left(-\frac{t^2}{2\tau_p^2}\right) e^{i\omega_0 t} \tag{11.30}$$

the wings of the sech pulse are more pronounced (Fig. 11.16). Since both pulse shapes,

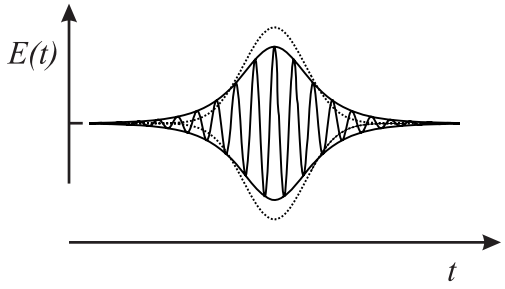


Figure 11.16: The hyperbolic secant short pulse with its envelope (lines; (11.29)) in comparison with the envelope of a Gaussian pulse (dots, (11.30)). The amplitudes of both pulse envelopes are normalised to the same area.

however, are not so different we prefer to work with the Gaussian pulse in the following since it is easier to handle mathematically.

Consider a pulse with the irradiance distribution $I(t) = I_0 \exp[-(t/\tau_p)^2]$ corresponding to the Gaussian pulse of (11.30) and travelling through a Kerr medium of length L thereby acquiring a phase factor $\exp(i\Phi) = \exp(i\omega_0 L n/c)$. Using (11.28) and expanding the Gaussian pulse near the centre as $I(t) = I_0\left[1 - (t/\tau_p)^2 + \cdots\right]$ the field of the pulse behind the Kerr medium of length L is given as

$$E(t) \propto \exp[-(t/\tau_p)^2]\exp(i\omega_0 t)\exp(i\omega_0 L/c\{n_0 + n_2 I_0[1 - (t/\tau_p)^2]\}). \tag{11.31}$$

From the phase

$$\Phi(t) = \omega_0 t + \omega_0 L/c\{n_0 + n_2 I_0[1 - (t/\tau_p)^2]\} \tag{11.32}$$

the instantaneous frequency ω is calculated as

$$\omega(t) \equiv \frac{d}{dt}\Phi(t) = \omega_0 - 2\omega_0 \frac{n_2 I_0 L}{c\tau_p^2} t. \tag{11.33}$$

Hence, near the centre of the pulse the frequency varies linearly with t. This frequency chirp means that for $n_2 > 0$ (i.e. positive dispersion) the frequency decreases with increasing time t. The central part of the pulse with the highest irradiance is retarded thereby producing a red shift on the leading edge of the pulse and a blue shift on the trailing part. This effect is sometimes referred to as self-phase modulation.

11.5.3 Propagation of Ultra-short Pulses

The propagation of ultra-short pulses in a medium like the Ti:sapphire crystal has several peculiarities usually not encountered when the propagation of almost monochromatic radiation is considered. The spread of frequencies corresponding to ultra-short pulses over a considerable part of the spectrum according to (5.5) requires the inclusion of the dispersion, i.e. the dependence of the index of refraction on the wavelength. As we will see in the following, assuming normal dispersion ($dn/d\omega > 0$), the high-frequency components are delayed with respect to the low-frequency components. This group velocity dispersion (GVD) results in a temporal broadening of the pulse and a variable instant frequency called a chirp (Fig. 11.17). Since the

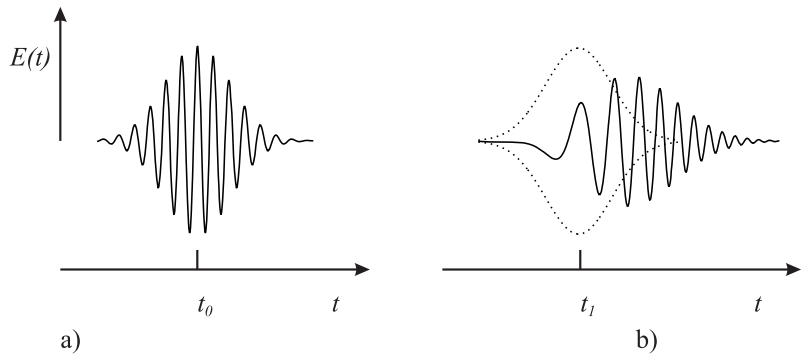


Figure 11.17: Amplitude of an ultra-short Gaussian pulse before, a) and after passing a material with normal dispersion, b). The pulse suffers from a delay with respect to the original pulse (dotted line), it is stretched and it acquires a chirp with the blue shifted frequency components at the trailing edge of the pulse.

pulse shape is dramatically affected by the dispersion one has to include its effect on the propagation of the pulse in a medium described by the propagation constant $k = 2\pi/\lambda = \omega n(\omega)/c$. We expand the propagation constant in a Taylor series about its value at ω_0 as

$$k(\omega) = k(\omega_0) + (\omega - \omega_0) \left.\frac{dk}{d\omega}\right|_{\omega=\omega_0} + \frac{1}{2}(\omega - \omega_0)^2 \left.\frac{d^2k}{d\omega^2}\right|_{\omega=\omega_0} + \cdots \tag{11.34}$$

The different coefficients in the power series affect the propagation of the pulse in a dispersive medium in different ways.

$$k(\omega_0) \equiv \frac{\omega_0}{v_\phi} \tag{11.35}$$

is a measure of how the angular frequency ω_0 of the sinusoidal carrier propagates within the pulse envelope. The corresponding phase delay after a distance z is $k(\omega_0)z$ and the corresponding time delay is $t_\phi = k(\omega_0)z/\omega_0 = z/v_\phi$. The phase velocity v_ϕ is related to the index of refraction by $n \equiv c/v_\phi$.

$$\left.\frac{dk}{d\omega}\right|_{\omega=\omega_0} = \frac{1}{v_g} \tag{11.36}$$

is a measure of how fast the envelope moves with the group velocity v_g. Hence, there is also an index of refraction of the wave group

$$n_g(\lambda) \equiv \frac{c}{v_g} = c\frac{dk}{d\omega} = c\frac{d}{d\omega}\frac{\omega \cdot n}{c} = n + \omega\frac{dn}{d\omega} = n(\lambda) - \lambda\frac{d}{d\lambda}n(\lambda) \tag{11.37}$$

where we have made use of $d\omega/d\lambda = -\omega/\lambda$. If the group velocity is smaller than the phase velocity ($v_g < v_\phi$) the cycles of the carrier frequency move from the trailing tail through the pulse envelope to the leading edge as depicted in Fig. 11.24 a) for consecutive pulses in a mode-locked laser. The second derivative in the third term of (11.34)

$$\left.\frac{d^2k}{d\omega^2}\right|_{\omega=\omega_0} = \frac{d}{d\omega_0}\left(\frac{1}{v_g(\omega)}\right) \tag{11.38}$$

represents the group velocity dispersion. This term will lead to a distortion of the pulse when travelling through a medium with a non-vanishing group velocity dispersion. The group velocity dispersion in optical materials, e.g., in optical fibres is often characterised by the dispersion

$$D \equiv \frac{1}{L}\frac{dT}{d\lambda} \tag{11.39}$$

where λ is the vacuum wavelength and T is the time a pulse takes to travel through the material of length L. This time is $T = L/v_g$ and hence

$$D = \frac{d\frac{1}{v_g}}{d\lambda} = -\frac{\omega}{\lambda}\frac{d\frac{1}{v_g}}{d\lambda} = -\frac{2\pi c}{\lambda^2}\frac{d^2k}{d\omega^2} \tag{11.40}$$

where we have used (11.36). For quartz glass, which is a typical material for optical fibres, the index of refraction $n(\lambda)$ and the group index of refraction $n_g(\lambda)$ are shown in Fig. 11.18 and the corresponding dispersion is displayed in Fig. 11.19.

In a wave guide or in an optical fibre the group velocity dispersion results not only from the material dispersion, but also from waveguide dispersion, modal dispersion, or polarisation mode dispersion. The latter three result from the confinement of the mode in a waveguide

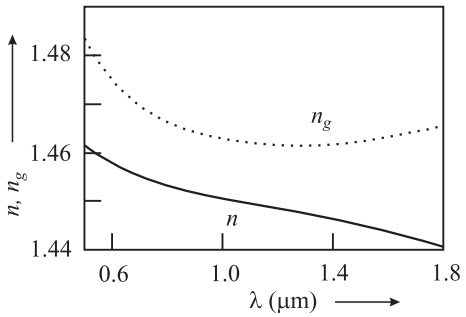


Figure 11.18: Index of refraction $n(\lambda)$ and group index of refraction $n_g(\lambda)$ (see (11.39)) of quartz glass.

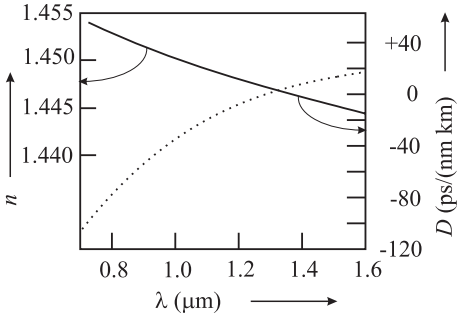


Figure 11.19: Index of refraction $n(\lambda)$ and dispersion (see (11.37)) of quartz glass.

where the propagation constant k and hence v_g depend on the angular frequency ω. If present, both material and waveguide dispersion contribute to the group velocity dispersion [39]

$$D = -\frac{\lambda}{c}\left[\left(\frac{\partial^2 n}{\partial \lambda^2}\right)_m + \left(\frac{\partial^2 n}{\partial \lambda^2}\right)_w\right] \tag{11.41}$$

where the indices m and w refer to the material and waveguide, respectively. The pure material dispersion of a quartz glass fibre shows zero dispersion near a wavelength of 1.3 µm. At this wavelength very short pulses can be transmitted over long distances without dispersive spreading. Taking into account the modal dispersion, this zero of dispersion can be shifted by suitable preparation of the waveguide where the material dispersion is compensated by the waveguide dispersion. The use of special fibres (see Section 11.5.5) with the zero of dispersion shifted to about 800 nm where short pulses can be generated by a Ti:sapphire laser allows one to generate extremely broad frequency combs.

11.5.4 Mode-locked Ti:sapphire Femtosecond Laser

A Ti:saphire femtosecond laser based on Kerr-lens mode locking often has a linear resonator formed by the output coupler and the end mirror similar to the one shown in Fig. 11.20. The Ti:sapphire crystal with Brewster angled facets is typically pumped, e.g., by a 10 W single-frequency 532 nm frequency-doubled Nd:YVO_4 laser. To produce short pulses the normal group velocity dispersion in the Ti:Sa crystal has to be compensated, e.g., by a pair of fused silica prisms [700] inside the laser resonator. Adjustment of the prisms allows one to correct the round-trip velocity dispersion.

Self-starting of a femtosecond Ti:sapphire laser with a repetition rate of about 100 MHz is not always easily obtained since the peak power changes by about six orders of magnitude when the laser switches from cw to pulsed operation. As a consequence, the Kerr-lens mode locking process is very weak in the cw regime. To accomplish self-starting of the pulsed laser sometimes a semiconductor saturated absorber (SESAM) is included [505, 696]. The broadband saturable absorber is a composite mirror consisting of a 5 µm silver mirror on a silicon

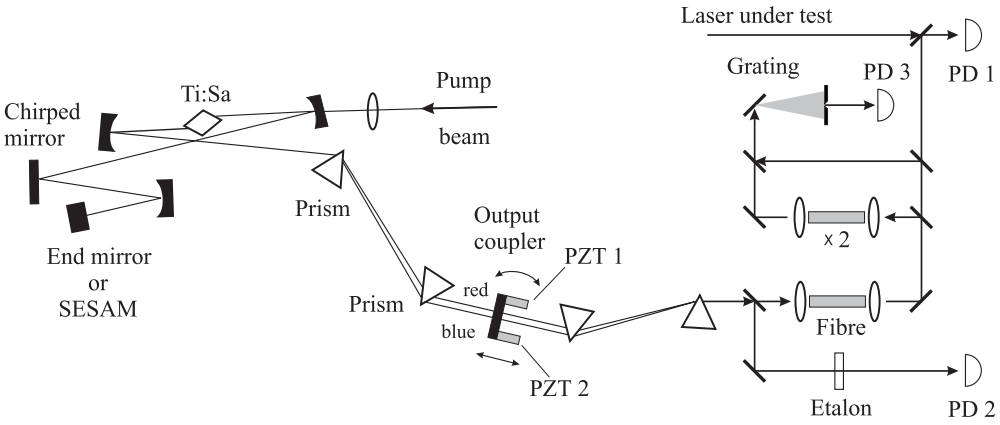


Figure 11.20: Layout of a Ti:sapphire laser for generation of femtosecond pulses [505]. Ti:Sa: Ti:sapphire crystal; SESAM: semiconductor saturable absorber; PD 1 – PD 3: photodiodes.

substrate and a complex sequence of films including a low temperature GaAs semiconductor absorber layer of 15 nm thickness [696]. With such a mirror, laser pulses of 6.5 fs were obtained [697] covering a wavelength range between about 690 nm and 900 nm.

11.5.4.1 Chirped Mirrors

Even though the prism pair compensates the group-velocity dispersion, the higher-order dispersion of the fused silica prisms is the most important limitation to the generation of ultrashort pulses. To achieve simultaneously a high reflectivity and compensation of the group-delay over an extended bandwidth the use of chirped mirrors [701] is imperative. They consist of multilayers of TiO_2 and SiO_2 ion beam sputtered onto a SiO_2 substrate. The alternating layers exhibiting a high index of reflection $n(TiO_2) \approx 2.3$ and $n(SiO_2) \approx 1.45$ (near 800 nm) can be regarded as a Bragg reflector. Each step in the index of refraction leads to Fresnel refraction with a Fresnel reflectivity for the amplitude of $r = [n(TiO_2) - n(SiO_2)]/[n(TiO_2)+n(SiO_2)] \approx 0.23$. With a sequence of alternating layers of the same thickness a Bragg reflection occurs for an angle [2] θ with $n\lambda = 2a \sin\theta$. To compensate for the group-delay dispersion in other elements the chirped mirror requires a group delay that varies approximately linearly with wavelength. Such a variation can be obtained for mirrors where the multilayer period is not constant (Fig. 11.21) sometimes referred to as chirped mirrors. In a chirped mirror a wave packet is reflected by the multilayer structure in a depth where the period matches the centre wavelength of the wave packet. In Fig. 11.21 longer wavelengths are reflected in a deeper region of the mirror as compared to shorter wavelengths thereby leading to a larger group delay of longer wavelengths. However, chirped multilayer coatings with monotonic variations of the layer thickness are not suitable due to Fabry-Pérot-like resonances that strongly perturb

[2] Note that the Bragg angle θ is defined differently to the angle of incidence α of ordinary optics. α is the angle between the incident ray and the normal of the surface whereas θ represents the angle between the wave vector of the incident beam and the surface.

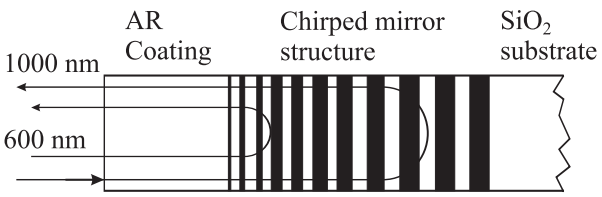


Figure 11.21: Schematic representation of the universal structure of a double-chirped mirror according to [702].

the group-delay dispersion. These interferences for longer wavelengths result from partial reflection from the front section of the mirror and from the Bragg reflection at the back of the mirror. The interferences can be avoided by an appropriate adjustment of the layer thickness [701] which then can result in double chirped mirrors [702] (Fig. 11.21). The reflectivities of these mirrors can be considerably higher than 99 % for very broad ranges [701, 702].

With these techniques Kerr-lens mode-locked lasers with pulse lengths down to about 10 fs can be routinely operated. Such a pulse has a length in space of only $\Delta l = c \times 10\,\text{fs} \approx 3\,\mu\text{m}$ corresponding to $\Delta\lambda/\lambda = \Delta\nu/\nu \approx 10\,\%$ (see (5.98) and Table 5.5). For application of a frequency comb to the measurement of optical frequencies it is advantageous if a whole octave or even more is covered (Section 11.5.6).

11.5.5 Extending the Frequency Comb

An even broader spectrum can be generated when the pulse from a femtosecond laser undergoes self-phase modulation in an optical fibre. With conventional fibres, however, the short pulses of a few tens of picoseconds produced by a Kerr-lens mode-locked laser are rapidly broadened already in the first few hundred micrometres in the fibre. At the same time the peak power of the pulse rapidly decreases and the self-phase modulation becomes inefficient. The development of microstructured fibres allowed one to shift the zero dispersion wavelength close to the midband wavelength $\lambda \approx 0.8\,\mu\text{m}$ of femtosecond pulses from a Kerr-lens mode-locked Ti:sapphire laser. In such a fibre these pulses can travel several centimetres without serious pulse spreading.

Self-phase modulation in the fibre can be regarded as a four-wave mixing process. Consider two adjacent angular frequencies ω_1 and $\omega_2 = \omega_1 + \delta$ of the comb representing the short pulse in the frequency domain. In the optical fibre non-linear processes may lead to the frequencies $2\omega_1$ and $2\omega_2 = 2\omega_1 + 2\delta$. The difference frequencies of these two new frequencies together with the original frequencies $2\omega_1 - \omega_2$ and $2\omega_2 - \omega_1$ immediately lead to new frequencies at $\omega_1 - \delta$ and $\omega_2 + \delta = \omega_1 + 2\delta$. As a result, new frequencies are generated that broaden the comb. As a consequence of the group velocity dispersion, ultra-short pulses are rapidly broadened and the peak power necessary to produce an efficient self-phase modulation cannot be sustained over a certain distance in the fibre. The group velocity dispersion can be tailored in a specific class of optical fibres. Fibres can be made to form a two-dimensional periodic array of closely packed hollow silica fibres (so-called holey fibres) to exhibit a two-dimensional photonic crystal [703]. If the core of such a fibre consists of a fibre with no hole,

radiation with frequencies inside the photonic band gap cannot penetrate into the cladding of the fibre. Holey fibres can be fabricated [703, 704] by a procedure where thin glass capillaries are stacked into a periodic array and the stack is fused and drawn at high temperature. When the process is repeated it results in a desired structure as depicted in Fig. 11.22. The authors of reference [705] used a microstructure fibre consisting of an 1.7 µm-diameter silica core surrounded by an array of 1.3 µm diameter air holes in a hexagonal close-packed arrangement. In such a fibre the waveguide dispersion can be tailored to compensate the material dispersion as a function of the core diameter between wavelengths of about 0.7 µm, 0.8 µm, and above 0.9 µm for core diameters of 1.4 µm, 1.7 µm, and 4 µm, respectively [705].

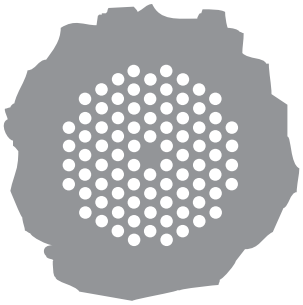


Figure 11.22: Cross-section of a fibre with a periodic array of air holes running down its length. The central hole is absent and this region with a high index of refraction and surrounded by the air gaps guides the light.

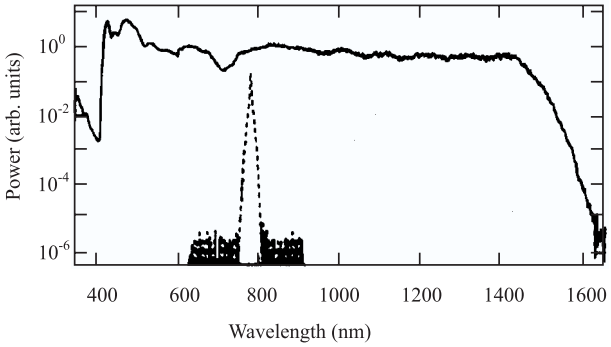


Figure 11.23: Optical spectrum of the continuum generated in a 75 cm section of a microstructure fibre. The dashed curve shows the spectrum of the initial 100 fs pulse. With permission from [705].

The use of a microstructured fibre does not require any longer to generate pulses below 10 fs and, hence, most of the femtosecond lasers used for the measurement of optical frequencies do not make use of a SESAM any longer.

In more detail, the "supercontinuum generation" in a microstructured fibre includes several processes among them self-phase modulation, soliton fission, four-wave mixing and Raman scattering [706]. For use of these combs it is essential that these processes conserve phase coherence. Fundamental limits to the noise result, e.g., from spontaneous Raman scattering down the fibre [707]. Nevertheless, with microstructured fibres frequency combs are readily produced that extend over more than an octave and where the phase coherence between each line of the comb is preserved.

11.5.6 Measurement of Optical Frequencies with fs Lasers

The mode-locked laser (Fig. 11.20) emits pulses of a few femtoseconds with a repetition rate given by the free spectral range of the laser resonator. Typical repetition frequencies are $100\,\mathrm{MHz} \leq f_{\mathrm{rep}} \leq 1\,\mathrm{GHz}$.

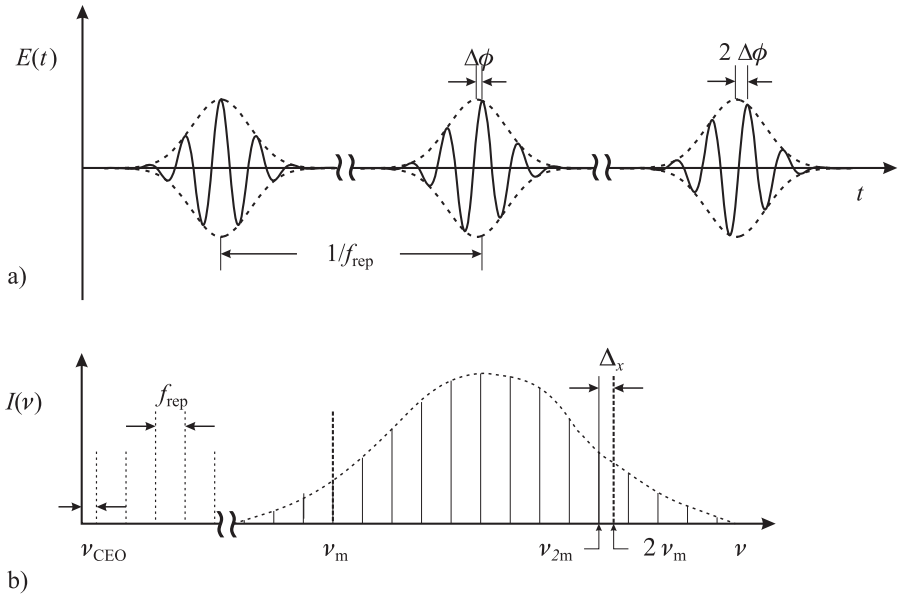


Figure 11.24: Time domain a), and frequency domain b), spectra of a mode locked femtosecond laser [708].

Since in general the group velocity of the pulse differs from the phase velocity of the wave, the phase of the carrier is shifted with respect to the envelope by $\Delta\Phi$ after each round trip of the pulse in the laser resonator (Fig. 11.24). Hence, the amplitude spectrum in the time domain is not periodic and consequently each line of the corresponding frequency spectrum is not an exact multiple of the repetition frequency. If this shift occurs with a constant rate the whole comb is shifted by a so-called carrier envelope offset frequency shift

$$\nu_{\mathrm{CEO}} = \frac{d\Phi}{dt} = \frac{\Delta\Phi}{2\pi T}. \tag{11.42}$$

Hence, each frequency ν_m of the comb can be determined by

$$\nu_m = m f_{\mathrm{rep}} + \nu_{\mathrm{CEO}} \tag{11.43}$$

provided that the repetition frequency f_{rep} and ν_{CEO} are measured and provided that m is known. With repetition frequencies of the order of several hundreds of megahertz a wavemeter is suited to determine the latter without any ambiguity. The repetition frequency can be easily measured by means of a photodiode (PD 2 of Fig. 11.20) and compared to a frequency standard in the microwave range. In general it is desirable to measure a higher harmonic (say, near 10 GHz) rather than the repetition frequency itself since in this case a fast InGaAs PIN photodiode allows one to achieve a higher signal-to-noise ratio [43]. To this end an etalon with a free spectral range of 10 GHz in front of the photodiode PD 2 (Fig. 11.20) is used.

The shift of the frequency comb ν_{CEO} can be determined if the comb spans more than an frequency octave, i.e. if the comb includes the frequency $\nu_m = m f_{\mathrm{rep}} + \nu_{\mathrm{CEO}}$ together with

a frequency $\nu_{2m} = 2mf_{\mathrm{rep}} + \nu_{\mathrm{CEO}}$, both with sufficient power. Consider radiation from a laser whose frequency is phase locked to the frequency ν_m of the frequency comb and which is frequency doubled in a non-linear crystal. The frequency of the frequency doubled light is simply given by $2\nu_m = 2mf_{\mathrm{rep}} + 2\nu_{\mathrm{CEO}}$. The frequency of the beat note Δx of this line with the line of the comb ν_{2m} is

$$\Delta x = 2\nu_m - \nu_{2m} = 2(mf_{\mathrm{rep}} + \nu_{\mathrm{CEO}}) - (2mf_{\mathrm{rep}} + \nu_{\mathrm{CEO}}) = \nu_{\mathrm{CEO}} \tag{11.44}$$

and exactly gives the carrier offset frequency ν_{CEO}. To implement this scheme the frequency doubled comb (Fig. 11.20) is superimposed with the fundamental comb and directed to a grating. The number of modes contributing to the doubling process is determined by the spectral width of the phase matching in the non-linear crystal. The resolution of the grating is chosen such that only those modes are selected behind a suitable aperture which achieve an optimal signal-to-noise ratio with the photodetector PD 3. If a complete octave is not available from the frequency comb, alternative schemes to determine ν_{CEO} are available [709].

Fine tuning of the frequency comb is necessary if the repetition frequency f_{rep} and the optical frequency ν_m have to be stabilised or kept in the optimum range of filters or phase lock loops. An end mirror placed on a piezoelectric transducer (PZT; Fig. 11.20) allows one to vary the length of the laser cavity thus shifting all modes of the frequency comb at the same time [710, 711]. The end mirror behind the prism pair can also be tilted by means of a second PZT thereby affecting predominantly the intracavity dispersion. As a consequence, the mode spacing which equals the repetition rate, can be controlled. These two actuators, tilt and length change, are not completely orthogonal as can be seen from the following equations. Taking into account the influence of the Ti:sapphire crystal the repetition frequency in a femtosecond laser is given by

$$f_{\mathrm{rep}} = \frac{c}{z + l_{\mathrm{Ti:Sa}}(n_g + n_{2g}I - 1)} \tag{11.45}$$

where z is the perimeter of the laser, $l_{\mathrm{Ti:Sa}}$ is the length of the Ti:sapphire crystal, n_g and n_{2g} are the (group) indices of refraction in analogy to (11.28). The carrier frequency of the m^{th} mode is given as

$$\nu_m = \frac{mc}{z + l_{\mathrm{Ti:Sa}}(n_p + n_{2p}I - 1)} \tag{11.46}$$

with n_p and n_{2p} the phase indices of refraction. The combined effects of the influences leading to fluctuations of the repetition frequency and the carrier frequency in a femtosecond laser show up as noise (see e.g. Fig. 11.25). Their influences can be found from inspection of (11.45) and (11.46). The path length z includes an increase due to the path inside the prisms and to the air inside the resonator leading to a variation of the carrier frequency by a few gigahertz and a few terahertz, respectively. Path length fluctuations are predominantly due to acoustical perturbations. Thermal fluctuations from the heat dissipated in the Ti:sapphire crystal and from n_2 induced fluctuations are other sources of noise. Apart from tilting the output coupler, there are also other techniques that can be used to provide a fast servo element such as a variation of the irradiance, e.g., by varying the pump power [709]. A very fast modulator input can also be generated by injecting an additional laser field into the laser

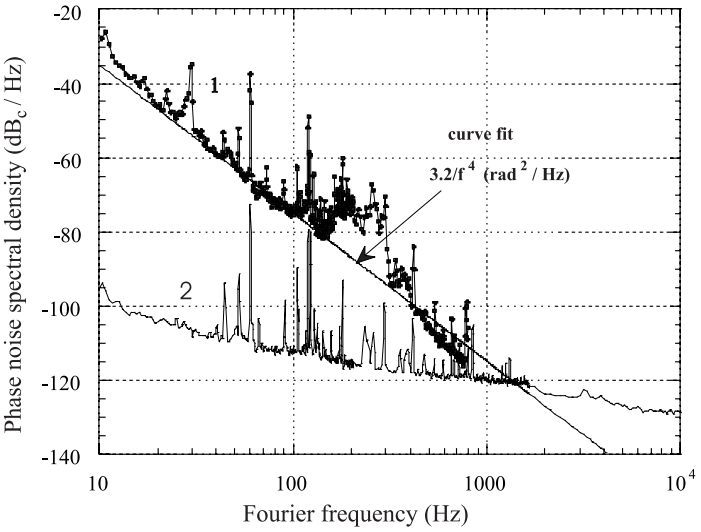


Figure 11.25: Measured phase noise spectral density of the repetition frequency $f_{\rm rep}$ for a free-running Kerr-lens mode-locked laser (1) compared to the phase noise of the frequency synthesiser (2). Courtesy of L. Hollberg.

crystal thereby modulating the Kerr lens [712]. Technically, the fluctuations in general can be controlled to the extent that true phase coherent measurements can be performed.

By stabilising the repetition rate $f_{\rm rep}$ and the carrier offset frequency $\nu_{\rm CEO}$ to a microwave frequency standard, e.g., a Cs atomic clock, the frequency comb represents a self-referenced "frequency ruler" with well known optical frequencies.

Several comparisons have shown the equivalence between frequency measurements performed by frequency chains using frequency division by a femtosecond comb and harmonic generation synthesis [501, 505, 713] or with a frequency interval divider chain [679] or with two independent combs [714].

Telle *et al.* [686] have applied the transfer oscillator concept to optical frequency measurements that are not degraded by the noise properties of the Kerr-lens mode-locked laser. Consider the signal processing scheme of Fig. 11.26. x is the frequency difference between the unknown frequency ν_x of the laser under test and the nearest comb line. The measured repetition frequency $f_{\rm rep}$ is mixed in mixer M1 with the frequency of the rf source as local oscillator providing the microwave frequency $f_{\rm LO}$. The difference signal is selected and multiplied by m_2, e.g., using a harmonic phase lock loop, leading to

$$\nu_A = m_2(f_{\rm LO} - m_1 f_{\rm rep}). \tag{11.47}$$

In mixer M2 the sum of $\nu_{\rm CEO}$ and x is generated and after the divider the frequency is

$$\nu_B = \frac{\nu_{\rm CEO} + x}{m_3}. \tag{11.48}$$

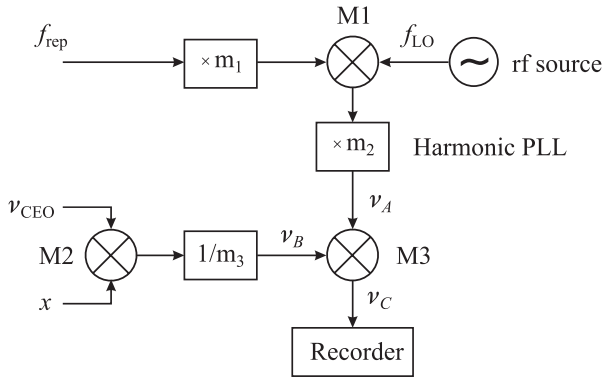


Figure 11.26: Signal processing scheme for linking optical and radio frequencies after [686]. M1, M2, M3: mixers.

Behind the mixer M3 the difference $\nu_A - \nu_B$ is generated as follows

$$\nu_C = \nu_A + \nu_B = m_2 f_{\rm LO} - \left(m_1 m_2 f_{\rm rep} + \frac{\nu_{\rm CEO} + x}{m_3} \right). \tag{11.49}$$

On the other hand, from (11.43) it follows that

$$\frac{\nu_x}{m_3} = m_1 m_2 f_{\rm rep} + \frac{\nu_{\rm CEO} + x}{m_3}. \tag{11.50}$$

Putting (11.50) into (11.49) one ends up with

$$\nu_x = m_2 m_3 f_{\rm LO} - m_3 \nu_C. \tag{11.51}$$

Hence, the scheme of Fig. 11.26 allows one to measure the optical frequency ν_x without degradation by the noise of the Kerr-lens mode-locked laser, provided that the phases of all signals in the scheme of Fig. 11.26 can be tracked unambiguously.

Similarly, the measurement of optical frequency ratios need not be degraded by the instability of the microwave oscillators. Consider the case of two different optical frequencies ν_1 and ν_2 from two independent optical frequency standards, locked to suitable frequencies of the same frequency comb

$$\nu_1 = \nu_{\rm CEO} + m f_{\rm rep} + \Delta x \tag{11.52}$$

$$\nu_2 = \nu_{\rm CEO} + n f_{\rm rep} + \Delta y. \tag{11.53}$$

Provided the values of ν_1 and ν_2 are known much better than the pulse repetition frequency $f_{\rm rep}$, then m and n are known without ambiguity. Since $\nu_{\rm CEO}, \Delta x$ and Δy are measured there are only two unknowns ν_1/ν_2 and $f_{\rm rep}$ which can be determined by the use of (11.52) and (11.53). This method has been used to show that a frequency ratio can be measured with an inaccuracy of below 6×10^{-19} [715].

Besides the Ti:sapphire laser there are other solid-state femtosecond systems [698] that might be used to produce a frequency comb. Cr:LiSAF (Cr:$LiSrAlF_6$) is another candidate

capable of producing pulses of a width below 60 fs [499, 716]. Such a laser can be directly pumped by 670 nm diode lasers thereby allowing one to set up a transportable battery powered system. The repetition frequency of such a laser comprising a set-up similar to the one shown in Fig. 11.20 with an intracavity SESAM and a prism pair for adjusting the group velocity dispersion has been locked to a 100 MHz quartz oscillator [716]. Very promising candidates are Er-doped amplifier lasers [717]. They are much less expensive as they can set up using commercial telecommunications lasers and components. They can be built very compact and may be even suitable for space applications.

Immediately after its invention the femtosecond comb generator was used for frequency measurements of the Cs D_1 line [718] and of a number of optical frequency standards [501, 505, 626, 713, 719–721]. In a rapid succession of experiments with increasing accuracy, it has been shown [715, 718] that the relative inaccuracy of a femtosecond optical frequency generator can be smaller than 6×10^{-19}. In contrast to frequency chains based on multiplication, the division of optical frequencies with a comb generator has a number of advantages. First, the frequency grid of well known frequencies extending over the visible, infrared and near UV spectral range makes the measurement of any optical frequency an easier task, in contrast to the use of frequency chains that were dedicated to particular optical transitions. Second, the all solid-state system sustains a high reliability and this, together with compact size and affordable prize, makes it a convenient clockwork for optical clocks. Third, the division of frequencies avoids the growth of phase noise associated with each multiplication step. Consequently, the long standing problem of sufficiently simple measurements of optical frequencies is solved by the femtosecond comb generator allowing one to utilise the full potential of optical frequency standards and clocks.

12 Time Scales and Time Dissemination

Accurate time and frequency is of fundamental importance to technology and science. Techniques taken for granted in our daily life, e.g., navigation for ships, aircraft and vehicles, geodetic positioning, wide area networks or high-speed digital telecommunication are based on accurate time and frequency. Less obviously to the public, deep space navigation, very long base line interferometry, the measurement of fundamental constants and the realisation of units in basic metrology present other examples at the frontier of science.

In one way or another all these applications rely on the dissemination of frequency or time signals. At distant locations the received time and frequency information allows one to compare, generate or synchronise local time scales, to discipline oscillators or to measure propagation delay times between transmitters and receivers. The measured path delays and the fact that time signals are travelling along the transmission path with the speed of light, permit the determination of geometric distances or accurate positions. The transfer techniques, however, have to meet different requirements depending on whether time information or frequency is transmitted. For time transfer the various contributions to the delay time, as in the cables, equipment and propagation path have to be taken into account properly. Hence, the uncertainty achieved in the time transfer is ultimately affected by the signal structure, the accuracy with which the delays in the transmitting and receiving equipment can be determined, the stability of the signal delays in the equipment and on the propagation path. All these contributions sum up to the uncertainty of the time transfer. On the other hand, if frequencies are to be compared as a consequence of the periodicity of the signal, the amount of delay need not be known but it is required that its value is constant during the comparison. The high performance of today's frequency standards, clocks and methods for intercomparison, moreover, makes it necessary to address thoroughly the constraints imposed by general relativity. We begin this chapter by giving a short description of time scales and their history (Section 12.1) before presenting a comprehensive treatment of general relativity (Section 12.2) necessary for comparison. We then discuss the methods and techniques currently used for dissemination and comparison of time and frequency. Finally, we present examples of the most demanding applications like pulsar timing and very long baseline interferometry.

12.1 Time Scales and the Unit of Time

12.1.1 Historical Sketch

For a long time the true solar day, i.e. the time between two successive meridian crossings of the sun, was the natural unit of time for mankind.[1] Because the Earth's orbit around the sun

is elliptical and the ellipse is tilted this duration is not uniform and can vary by as much as 50 seconds during the year. Hence, Universal Time (UT) now called UT0 as agreed by the International Astronomical Union (IAU) in 1928 was based on the mean solar day beginning at midnight on the Greenwich meridian. The day was divided into 86 400 seconds and, hence, the second was de facto coupled to the rotation of the Earth through this definition. Correcting UT0 for periodic contributions of the oscillations of the polar axis of the Earth resulted in the time scale UT1 which corresponds to the Earth's angular position but still reflects seasonal fluctuations of different origin. UT1 is used in celestial naviagation. To remove these fluctuations the more uniform time scale UT2 is derived from UT1 by applying corrections of up to 10^{-8}.

To obtain a time unit no longer depending on Earth's variable rotation, in 1956 the General Conference of Weights and Measures (CGPM) decided to adopt the definition of the ephemeris second which had already been defined in 1952 by the IAU. This ephemeris second later formed the base unit of time in the 1960 newly adopted International System of Units (SI). In short,[1] the ephemeris second was based on the period of revolution of the Earth around the sun which is more predictable than the rotation of Earth itself. Rather than using the siderial year,[2] the tropical year was chosen which is measured from one spring equinox to the next one.[3] The definition of the second "as 1/31 556 925.974 7 of the tropical year for December 31, 1899 at 12 hours ephemeris time" was replaced in 1967 by the definition: "The second is the duration of 9 192 631 770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the caesium 133 atom" (see page 203). This definition was based on the result of a measurement of the ephemeris second in terms of the period of the Cs transition that was conducted between 1955 and 1958 in a collaboration between the National Physical Laboratory (UK) [14] and the United States Naval Observatory.

12.1.2 Time Scales

A number of different time scales are in current use, some of them will be briefly described in the following.[3] Time scales as "ordered sets of scale markers with an associated numbering" [1] are divided into two categories, dynamic time scales and integrated time scales. Dynamic time scales are derived from a description of a dynamical physical system where the time t is used as a parameter to describe the evolution of the system. Examples of dynamic time scales are Universal Time (UT1) and Ephemeris Time (ET) which are derived from observing and modelling the rotation of the Earth around its polar axis and its revolution around the sun, respectively. The elapsed time can be determined from the observed position and the corresponding equation. Integrated time scales are based on a time interval, e.g., the second as derived from the frequency of the transition in ^{133}Cs. They are formed by a suitable starting point and by integrating the defined time units consecutively. The Temps Atomique

[1] A more detailed description of the history can be found, e.g., in [8].

[2] One siderial year is when the Earth after one orbit around the sun has the same position as measured with respect to the fixed reference system of distant stars.

[3] In the geocentric system used for astronomic time definition, the apparent path of the sun on the celestial globe is a great circle that intersects the projection of Earth's equator in the vernal and the autumnal equinox.

[3] For a more detailed description the reader is referred to [1, 8, 722].

International (International Atomic Time, TAI) and the Coordinated Universal Time (UTC) represent integrated time scales. Both time scales are established by the Time Section of the Bureau International des Poids et Mesures (International Bureau of Weights and measures, BIPM) in Paris by a complicated process. About 50 national timing institutes contribute to

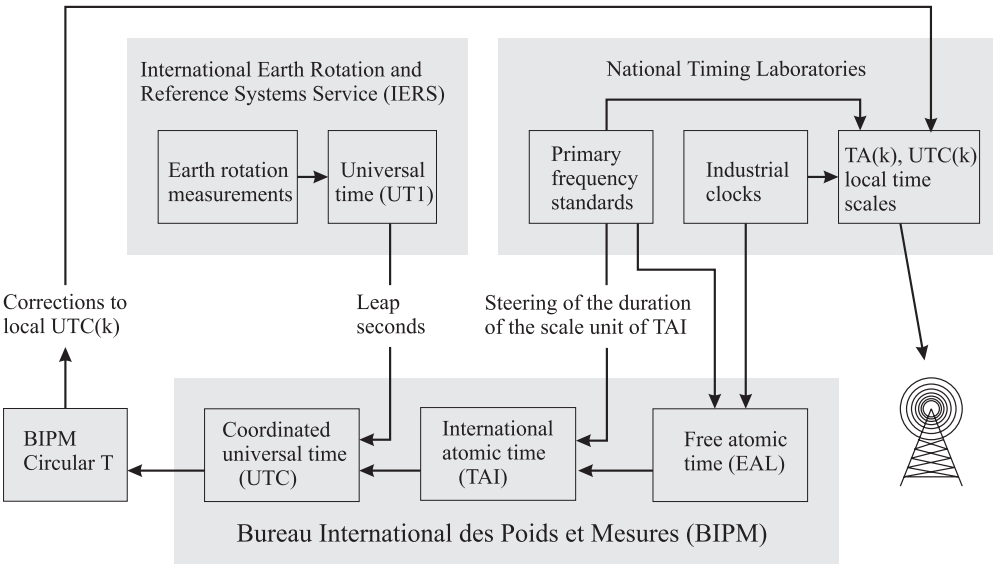


Figure 12.1: Simplified scheme of the formation of International Atomic Time (TAI) and Coordinated Universal Time (UTC) according to Jones [8].

this process (Fig. 12.1) with their clock ensembles that may consist of commercial clocks or primary clocks or both by forming their local atomic time scale TA(k) where k represents the acronym of the institute. Furthermore, the national timing institutes form a local representation UTC(k) that approximates the universal coordinated time scale UTC, to be discussed later. At the end of each month the participating institutes report the time differences between their individual clocks and their local UTC(k) to the BIPM. In addition to the clock data the results of regular time scale comparisons between the timing centres are communicated to the BIPM, too.

In a highly sophisticated scheme the BIPM uses these data to derive the free atomic time scale Echelle Atomique Libre (EAL, free atomic time scale) which represents a world mean. As a result of the large number of about 250 clocks contributing to EAL this is a very stable time scale. However, the scale unit of EAL does not necessarily coincide with that derived from the primary standards. Hence, TAI is derived from EAL by steering the duration of the EAL scale unit in accordance with the SI second as realised by primary clocks in some major timing laboratories. The applied corrections are kept below the short-term fluctuations of EAL. The origin of TAI is January, 1^{st} 1958 where it agreed with UT1.

Since our daily life and astronomical navigation are both governed by the rotation of the Earth, an atomic time scale called Universal Coordinated Time (UTC) was adopted in 1972. UTC is referenced to the Greenwich meridian. UTC is derived from TAI and kept in step with

the Earth rotation by insertion of leap seconds. Leap seconds [722] are inserted in such a way that UTC approximately follows UT1 and meets the condition $|\mathrm{UTC}(t) - \mathrm{UT1}(t)| < 0.9\,\mathrm{s}$. Hence, the scale unit of UTC is the same as that of TAI but UTC and TAI differ by an integer number n of seconds where n depends on the irregular angular velocity of the Earth

$$\mathrm{UTC}(t) = \mathrm{TAI}(t) - n. \tag{12.1}$$

Since UTC's scale unit is the second as it is defined in the SI, UTC is an atomic time scale. A leap second is inserted in UTC world-wide at the same epoch, preferably in the middle or at the end of a year. The time interval between two leap seconds depends on astronomical observations of the Earth's rotation. This information is provided by the International Earth Rotation and Reference Systems Service. As a result of the slowing down of Earth's rotation, UTC was more than half a minute behind TAI in the year 2000.

The atomic time scales TAI and UTC calculated by the BIPM and the deviations of the local time scales UTC(k) are distributed in a monthly bulletin called "Circular T" [723]. As a result of the time necessary to collect and process the data transferred to the BIPM, these time scales can be compared to other time scales "a posteriori". Hence, only the local approximations UTC(k) to UTC of the different timing centres are available in real time and they are used for numerous technical applications, different fields of navigation, telecommunication, space science and basic research.

It is requested that a particular UTC(k) does not deviate from UTC by more than 1 µs with a maximum deviation of 0.1 µs strived for in the future [1] which has been achieved already by about thirty UTC(k) laboratories in the year 2004. As a result the local UTC(k) have to be steered to follow UTC. For this purpose in the national timing centres a clock is selected that has proved to be particularly stable during the past months and its frequency is adjusted such that it is expected to realise a time scale that coincides as closely as possible to UTC.

There are other time scales, e.g., the system times used for the Global Positioning System (GPS) and the Russian Global Navigation Satellite System (GLONASS). Although GPS and GLONASS system time are maintained as independent time scales they are in fact steered with high update rates according to UTC(USNO) of the United States Naval Observatory and UTC(SU) of the Russian timing centre, respectively.

According to the definition of the SI second (see page 203), each clock realises the proper time in its local frame. When viewed from another local frame, the local time elsewhere is subject to the (different) gravitational potential there. General Relativity predicts (Section 12.2) that the time realised in the first frame appears to be shorter or longer depending on the sign of the difference of the respective gravitational potentials. As a result of this effect, the clocks of, e.g., the NIST at Boulder located at an altitude of about 1.6 km above sea level appear to run faster by about 2×10^{-13} when viewed from the site of PTB in Braunschweig at an altitude of 79.5 m. Hence, TAI was defined such that the gravitational red shift is taken into account as "TAI is a coordinate time scale defined in a geocentric reference frame (origin of the centre of the Earth) with the SI second as realised on the rotating geoid as the scale unit [1]."

12.2 Basics of General Relativity

The accuracy of frequency standards used in scientific and technical applications and clocks has long reached a stage where accurate time and frequency comparisons require the taking into account of relativistic effects. A clock in the vicinity of the Earth is subject to gravitational and rotational forces. Hence, the clock represents an accelerated system thereby requiring the description in terms of a curved space-time geometry of general relativity. The relationship between two infinitesimally close space-time events is given by the relativistic line element

$$ds^2 = g_{\alpha,\beta}\left(x^\mu\right) dx^\alpha dx^\beta \tag{12.2}$$

where $g_{\alpha,\beta}\left(x^\mu\right)$ is the coordinate-dependent metric tensor and $\left(x^\mu\right) \equiv \left(x^0 = ct, x^1, x^2,\ x^3\right)$ denote the four space-time coordinates with t termed the coordinate time and c the speed of light. In (12.2) Einstein's summation convention of repeated indices is used. In the solar system the space time-curvature due to the gravitational field is weak and the components of the metric tensor $g_{\alpha,\beta}\left(x^\mu\right)$ deviate from the Minkowski metric of special relativity ($g_{00} = -1, g_{ij} = \delta_{ij}$ with the Kronecker $\delta_{ij} = 1$ for $i = j$ and $\delta_{ij} = 0$ for $i \neq j$) by small corrections expressed as a power series in the gravitational potential [724]. In the vicinity of Earth the potential is weak and is approximated by the Newtonian gravitational potential U.[4] The metric in an Earth centred (non-rotating) inertial reference frame is

$$g_{00} = -\left(1 - \frac{2U}{c^2}\right), \qquad g_{0j} = 0, \qquad g_{ij} = \left(1 + \frac{2U}{c^2}\right)\delta_{ij} \tag{12.3}$$

where the non-diagonal elements of the metric tensor in a non-rotating geocentric coordinate system vanish. Hence, the line element can be approximated by

$$ds^2 = -\left(1 - \frac{2U}{c^2}\right)c^2 dt^2 + \left(1 + \frac{2U}{c^2}\right)\left[(dx^1)^2 + (dx^2)^2 + (dx^3)^2\right] \tag{12.4}$$

where the potential $U = U_E + U_T$ comprises the Newtonian gravitational potential U_E of the Earth and the tide-generating potential U_T of external bodies. A simple approximation to the gravitational potential of Earth [263]

$$U_E = \frac{GM_E}{r} + J_2 GM_E a_1^2 \frac{\left(1 - 3\sin^2\phi\right)}{2r^3}. \tag{12.5}$$

takes into account that Earth bulges around its equator [5] and, hence, the potential depends on the latitude characterised by the angle ϕ which is zero at the equator and is counted positive towards the north pole. The equatorial radius of the Earth is $a_1 = 6\,378\,136.5$ m; r is the distance from the origin of the geocentric coordinate frame and GM_E = 3.986 004 418

[4] We adopt here the notation of the "clock community" and the IAU where the potentials are taken with a positive sign. Note that no general convention exists for the signs of the spatial and temporal coordinates, here we follow the convention of references [263, 724].

[5] Note that (12.5) approximates the predominantly ellipsoidal shape of the Earth but does not take into account mass irregularities.

$\times 10^{14}\,\mathrm{m^3/s^2}$ is the product of the gravitational constant and Earth's mass. $J_2 = +\,1.082\,636 \times 10^{-3}$ is the quadrupole moment coefficient of the Earth. The potential of (12.5) leads to correct results for the gravitational red shift within a relative uncertainty of $\delta\nu/\nu < 10^{-14}$ of the frequency standards and clocks.

In a coordinate system co-rotating with the Earth, a coordinate transformation has to be performed, i.e. from the inertial system to a system rotating towards the east with a constant angular velocity ω as

$$\begin{aligned} x &= x' \cos(\omega t') - y' \sin(\omega t') \\ y &= x' \sin(\omega t') + y' \cos(\omega t') \\ z &= z' \\ t &= t' \end{aligned} \tag{12.6}$$

where $\omega = 7.292\,115 \times 10^{-5}\,\mathrm{rad/s}$ is the rotational angular velocity of the Earth. Here, we have restricted ourselves to the case where $\omega(x'^2 + y'^2) \ll c^2$. Insertion of (12.6) and its derivatives into the quadratic form of the interval $ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2$ of the inertial system leads to

$$\begin{aligned} ds^2 &= -\left[1 - \frac{\omega^2}{c^2}(x'^2 + y'^2)\right] c^2 dt'^2 - 2\omega y' dx' dt' + 2\omega x' dy' dt' + dx'^2 + dy'^2 + dz'^2 \\ &= g'_{\alpha,\beta}\left(x'^{\mu}\right) dx'^{\alpha} dx'^{\beta} \end{aligned} \tag{12.7}$$

where we have neglected the influence of the potential U for the time being. In the first term of (12.7) we find the additional term $\omega^2\rho^2$ which is related to the potential $U_{\mathrm{centr}} = \omega^2\rho^2/2$ due to the centrifugal force experienced on a massive body rotating at distance $\rho = \sqrt{x'^2 + y'^2}$ from the rotational axis with angular velocity ω. Hence, in a rotating coordinate system without gravitational potential one finds

$$g_{00} = -\left(1 - \frac{2U_{\mathrm{centr}}}{c^2}\right) \tag{12.8}$$

which is the same expression as has been used in the metric of (12.4) reflecting the equivalence of accelerational and gravitational potentials. From (12.7) one finds that there are non-diagonal terms of the metric tensor in the co-rotating geocentric coordinate system.

Using spherical coordinates r, ϕ and L for the distance from the geocentre, the angle of the latitude, and the angle of the longitude (which is counted positive towards the East), respectively, we use the transformations

$$\begin{aligned} x' &= r\cos\phi\cos L \\ y' &= r\cos\phi\sin L \\ z' &= r\sin\phi \\ t' &= t \end{aligned} \tag{12.9}$$

and end up with the metric [263]

$$ds^2 = -c^2dt^2 + \left[dr^2 + r^2d\phi^2 + r^2\cos^2\phi\left(\omega^2dt^2 + 2\omega dLdt + dL^2\right)\right]. \tag{12.10}$$

In contrast to (12.3) the metric in the co-rotating rotating frame with the gravitational potential can be written as

$$g_{00} = -\left(1 - \frac{2U}{c^2} - \frac{\left(\vec{\omega} \times \vec{r}\right)^2}{c^2}\right), \qquad g_{0j} = \frac{\left(\vec{\omega} \times \vec{r}\right)_j}{c}, \qquad g_{ij} = \left(1 + \frac{2U}{c^2}\right)\delta_{ij} \quad (12.11)$$

where the vector product term between the angular velocity $\vec{\omega}$ and the vector $\vec{r}$ pointing from the centre of the Earth to a point on the surface gives rise to the centrifugal potential and the Sagnac effect, to be discussed later, via the off-diagonal elements in the tensor (see (12.3) and (12.11)).

According to the definition of the SI second, the time displayed by a clock is the proper time τ, i.e. the time t measured in a coordinate system attached to the clock. Consider the transportation of a clock described in an external coordinate system between two infinitesimally close events from x^0, x^1, x^2, x^3 to $x^0 + dt, x^1 + dx^1, x^2 + dx^2, x^3 + dx^3$. The metric

$$d\tau = \frac{1}{c}\sqrt{-ds^2} \quad (12.12)$$

relates the increments of the proper time as measured by the clock with the increments of time dt as measured in the time t of the external coordinate system. The time t is called the coordinate time. The increment of the coordinate time dt_{clock} displayed by the clock as viewed from the coordinate system can be determined from

$$dt_{\mathrm{clock}} = d\tau_{\mathrm{clock}}\frac{dt}{d\tau} \quad (12.13)$$

by use of the metric (12.12) where all quantities have to be evaluated at the event x^0, x^1, x^2, x^3. Integration of (12.13) along the world line of the clock yields the coordinate time $t_{\mathrm{clock}}(t)$ of the clock. The relation $\frac{d\tau}{dt}$ can be found from (12.2) and (12.12) as

$$\frac{d\tau}{dt} = \sqrt{-g_{00}(x^0, x^1, x^2, x^3) - \frac{2}{c}g_{0i}(x^0, x^1, x^2, x^3)\frac{dx^i}{dt} - \frac{1}{c^2}g_{ij}(x^0, x^1, x^2, x^3)\frac{dx^i}{dt}\frac{dx^j}{dt}}. \quad (12.14)$$

In the vicinity of Earth the influence of the gravitational potential on the metric is small ($2U/c^2 \approx 1.4 \times 10^{-9} \ll 1$). Hence, it is suitable to consider only the deviation from flat space by defining a quantity $h(t)$ as

$$\frac{d\tau}{dt} \equiv 1 - h(t) \quad (12.15)$$

with $h(t)$ given as a power series in 1/c. The difference between the coordinate time and the proper time is given by

$$\Delta t \equiv t - \tau = \int_{t_0}^{t} h(t)dt. \quad (12.16)$$

Δt can be calculated by either using the metric in the geocentric system (12.4) or in the coordinate system rotating with the Earth (12.10). In the metric for a geocentric non-rotating coordinate system (12.4) the non-diagonal elements vanish and insertion into (12.14) leads to

$$h(t) = 1 - \sqrt{\left(1 - \frac{U}{2c^2}\right) - \frac{1}{c^2}\left(1 + \frac{U}{2c^2}\right)v^2}. \tag{12.17}$$

By expanding the root in (12.17) one ends up with

$$h(t) = \frac{U(t)}{c^2} + \frac{v(t)^2}{2c^2} + O(\frac{1}{c^4}). \tag{12.18}$$

The second term is known as the time-dilation shift of a clock moving with velocity $\vec{v}$ with respect to the centre of the coordinate system, i.e. the centre of Earth. The term $O(\frac{1}{c^4})$ of (12.18) typically contributes less than 10^{-18} and will be neglected in the following.

For a coordinate system co-rotating with Earth, Guinot [263] gives

$$h(t) = \frac{1}{c^2}\left[U_g + \Delta U(t) + \frac{V(t)^2}{2}\right] + \frac{2\omega}{c^2}\frac{dA_E}{dt} \tag{12.19}$$

which can be derived similarly when the metric (12.10) is used. V is the modulus of the coordinate velocity relative to the Earth. The last term results from the Sagnac effect [725]

$$\frac{1}{c^2}\int_{\mathcal{P}}^{\mathcal{Q}} (\vec{\omega} \times \vec{r}) \cdot d\vec{r} = \frac{1}{c^2}\int_{\mathcal{P}}^{\mathcal{Q}} \vec{\omega} \cdot (\vec{r} \times d\vec{r}) = 2\frac{1}{c^2}\int_{\mathcal{P}}^{\mathcal{Q}} \vec{\omega} \cdot d\vec{A}_E = \frac{2\omega A_E}{c^2}. \tag{12.20}$$

A_E is the area (Fig. 12.2) swept by the projection of the vector into the equatorial plane whose origin is at the centre of the Earth and which points to the clock that rests or moves slowly in the rotating frame (Fig. 12.2). $U_g = 6.263\,685\,75 \times 10^7\ \mathrm{m^2/s^2}$ is the constant

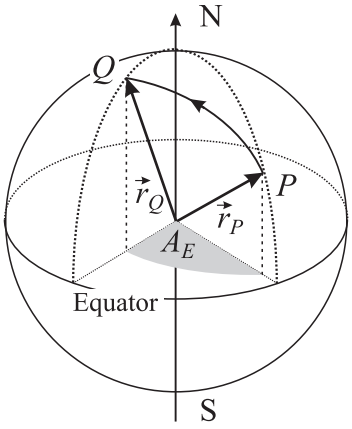


Figure 12.2: A clock moved from point P to Q on Earth experiences a time difference due to the Sagnac effect which is proportional to the area A_E.

potential in the geocentric rotating coordinate system on the rotating geoid which can be obtained from (12.5) augmented by the potential of the centrifugal force. The difference in the gravitational potential between a specific location and the geoid including the potential

due to the centrifugal force of the rotating geoid can be modelled if a relative uncertainty of 10^{-14} is sufficient as

$$\Delta U(\vec{r}) = \frac{GM_E}{r} + J_2 GM_E a_1^2 \frac{(1-3\sin^2\phi)}{2r^3} + (\omega^2 r^2 \cos^2\phi) - U_g. \tag{12.21}$$

An even better parametrisation as a function of the altitude b above the geoid and the latitude ϕ is given as follows [263]

$$\frac{\Delta U(b,\phi)}{c^2} = (-1.08821\times10^{-16} - 5.77\times10^{-19}\sin^2\phi)\frac{b}{\mathrm{m}} + 1.716\times10^{-23}\left(\frac{b}{\mathrm{m}}\right)^2 \tag{12.22}$$

This approximation is valid for an altitude $b < 15$ km above the geoid with the relative uncertainty of the potential being smaller than 10^{-15}.

The proper frequency of a frequency standard H is the frequency of the standard $\nu_H(\tau)$ where τ is the proper time at the location of the standard. This frequency may be referred to as the "nominal frequency" $\nu_{H,0}$. Often, the proper frequency $\nu_H(t)$ of the standard as function of a chosen coordinate time t is required. Following Guinot [263] we define the dimensionless quantity "proper normalised frequency" or "proper relative frequency" of a frequency standard H as

$$\Phi_H(t) \equiv \frac{\nu_H(t)}{\nu_{H,0}}. \tag{12.23}$$

Consider a clock resting on the surface of the geoid. This geoid rotates with respect to the geocentric reference system defined by Earth's centre of mass where the spatial axes do not show a rotation with respect to distant extragalactic objects. The coordinate time of the latter system is referred to as Temps Coordonnée Géocentrique (geocentric coordinate time, TCG). The proper normalised frequency of the clock in coordinate time TCG is

$$\Phi_H(TCG, x^\mu_{\mathrm{geoid}}) = \frac{d\tau}{dT_{TCG}} = 1 + \frac{U_g}{c^2} + O\left(\frac{1}{c^4}\right) = 1 + 6.969\,290\,3\times10^{-10}. \tag{12.24}$$

Hence, all clocks H operated according to the definition of the SI second on the surface of the geoid advance by this fraction with respect to TCG which adds up during one year to 22 ms. Thus a geocentric coordinate time called Terrestrial Time (TT) has been introduced whose scale unit is the SI second as realised on the rotating geoid. In both systems the normalised frequencies differ by the same amount. TAI realises TT.

12.3 Time and Frequency Comparisons

As simultaneity is not defined in general relativity, first one has to agree on how to synchronise clocks. Synchronised clocks show the same reading at the same instant of time. In contrast to Einstein's synchronism [6] today's convention [724] is to apply "coordinate synchronisation"

[6] Consider two separated clocks A and B. A time signal is sent from A at τ_A^{send} to B where it is received at τ_B^{rec}, reflected and received by A at τ_A^{rec}. A and B are (Einstein) synchronised if $\tau_B^{\mathrm{rec}} = 1/2(\tau_A^{\mathrm{send}} + \tau_A^{\mathrm{rec}})$ [266].

where two events described in a suitable coordinate system by their coordinates x_1^μ and x_2^μ are considered to be simultaneous if the values of the time coordinates are equal ($x_1^0 = x_2^0$).[7]

Clocks operating at different locations $\mathcal{P}$ and $\mathcal{Q}$ on Earth can be compared by different methods. The most common ones include transportation of a clock or exchanging electromagnetic signals between $\mathcal{P}$ and $\mathcal{Q}$. Both processes are readily described in a geocentric system where the origin of the coordinate system is at the centre of the Earth. The coordinate system can be chosen in two different ways, either in a locally inertial frame with fixed orientations in space (with respect to the most distant objects of the Universe) or as a co-rotating system fixed to the rotating Earth. The equations to be given below depend on this choice. Both cases are described in a recommendation of the International Telecommunication Union [726] and in [263, 725] together with useful examples.

12.3.1 Comparison by a Transportable Clock

When the time is transported from the point $\mathcal{P}$ to the point $\mathcal{Q}$ by means of a transportable clock, the coordinate time that is accumulated during this transport in the geocentric non-rotating reference frame is

$$\Delta t = \int_{\mathcal{P}}^{\mathcal{Q}} ds \left[1 + \frac{U(\vec{r}) - U_g}{c^2} + \frac{v^2}{2c^2}\right]. \tag{12.25}$$

$U(\vec{r})$ is the purely gravitational potential at the location of the clock and v is its velocity as viewed from the geocentric non-rotating reference frame. ds is the increment of proper time as measured in the rest frame of the moving clock.

In the rotating geocentric reference frame the time difference is

$$\Delta t = \int_{\mathcal{P}}^{\mathcal{Q}} ds \left[1 + \frac{\Delta U(\vec{r})}{c^2} + \frac{V^2}{2c^2}\right] + \frac{2\omega}{c^2} A_E \tag{12.26}$$

where V is the velocity of the clock with respect to the ground. $\vec{r}$ is the vector whose origin is at the centre of Earth and pointing to the position of the clock on its journey from point $\mathcal{P}$ to point $\mathcal{Q}$. The equatorial projection of the vector $\vec{r}$ sweeps an area A_E.

The three terms in (12.26) represent the influence of the gravitation, the time dilation and the Sagnac effect. The latter is actually nothing more than "time dilation shift in disguise" [8]. It results from the fact that the clocks on Earth co-rotate with the same angular velocity. Hence, their velocity is dependent on the latitude defining the distance from the polar axis. A_E is taken as positive when the path projected on the equatorial plane has an eastward component (the situation depicted in Fig. 12.2 leads to a negative A_E).

[7] In contrast to Einstein's synchronism, coordinate synchronism leads to transitivity, i.e. if clocks 1 and 2 are synchronised and clocks 2 and 3 are synchronised then clocks 1 and 3 are also synchronised.

12.3.2 Time Transfer by Electromagnetic Signals

To compare the readings of two or more distant clocks by electromagnetic signals with radio frequencies or optical frequencies, several techniques differing in expenditure and achievable accuracy are distinguished: one-way time transfer; common-view; and two-way time transfer. The time between the emission and the reception of an electromagnetic signal in the non-rotating geocentric reference frame is

$$\Delta t = \frac{1}{c} \int_{\mathcal{P}}^{\mathcal{Q}} d\sigma \left[1 + \frac{U(\vec{r}) - U_g}{c^2} + \frac{v^2}{2c^2} \right] \tag{12.27}$$

where $d\sigma$ is the increment of proper length along the transmission path between $\mathcal{P}$ and $\mathcal{Q}$ and all other quantities are defined as before.

In the rotating geocentric reference frame the time difference is

$$\Delta t = \frac{1}{c} \int_{\mathcal{P}}^{\mathcal{Q}} d\sigma \left[1 + \frac{\Delta U(\vec{r})}{c^2} \right] + \frac{2\omega}{c^2} A_E. \tag{12.28}$$

Here, $\Delta U(\vec{r})$ is the gravitational potential at the point $\vec{r}$ (diminished by the potential of the geoid, i.e. the Standard Earth) as monitored from a coordinate system rigidly attached to the Earth, and A_E is the area of the equatorial projection of the triangle with the corner points given by the centre of the Earth, the point P irradiating the signal and the point Q where the signal is recorded. The area A_E is taken positive if the signal has an eastbound component.

For the path from a position on Earth's surface to a satellite in a geostationary orbit and back, the second term including $\Delta U(\vec{r})/c^2$ leads to a correction of about a nanosecond corresponding to a distance ct of about 30 cm. The third term with $2\omega/c^2 = 1.6227 \times 10^{-6}$ ns/km^2 can be as large as a few hundreds of nanoseconds for large distances.

12.3.2.1 One-way Time Transfer

The most common way to disseminate time information uses the transmission of an electromagnetic signal with coded information. Examples include time signals that can be accessed from regular telephone and television service or the internet. Radio transmitters broadcasting short-wave transmissions with frequencies of several megahertz or long-wave transmission of a few tens of kilohertz (see 12.4) allow the user to access time signals within large areas. Clocks on board the satellites orbiting around the Earth, e.g., of the Global Positioning System (GPS) provide accurate time worldwide. Time information transmitted from terrestrial stations and from satellites is used to set clocks, computer time, or to discipline oscillators to be used as frequency standards and radio controlled clocks with particular impact on everyday life. The accuracy to which the clock of a user at the receiving location can be set, i.e. synchronised to the clock at the transmitting site, depends on the propagation time the signal takes to reach the user's clock. This time delay can be as large as several tenths of a second when the signal is transferred via the internet or a geo-synchronous satellite. Measuring the round-trip delay client-server-client and assuming equal path lengths in both directions, the

propagation time in the internet is taken into account for the largest contribution. For satellite links the delay time can be calculated and corrected for, using the speed of light and the known distance.

12.3.2.2 Common-view Time Transfer

The simultaneous observation of the same signal transmitted, e.g., by a satellite and received at different locations can be used to synchronise the clocks at the particular locations. Consider two stations A and B receiving the time signal t_S over the path S–A and over the path S–B with a delay time τ_{SA} and a delay time τ_{SB}, respectively. When stations A and B exchange the results of their measurements $\Delta t_A = (t_S - \tau_{SA}) - t_A$ and $\Delta t_B = (t_S - \tau_{SB}) - t_B$ they obtain

$$\Delta t_B - \Delta t_A = (t_A - t_B) - (\tau_{SA} - \tau_{SB}) \tag{12.29}$$

i.e. the time difference $t_A - t_B$ between their clocks and the difference of the path delays. This so-called "common-view method" avoids the necessity to know the exact time of the clock onboard the satellite since t_S cancels in the evaluation. This property was particularly useful before the year 2000 when the clocks signals in the GPS satellites (Section 12.5) were deteriorated on purpose by the so-called selective availability to degrade the positioning accuracy. The common-view method is frequently used to compare the clocks in different timing institutes.

12.3.2.3 Two-way Time Transfer

The most accurate means up to now to compare remote time scales is provided by a Two-Way Satellite Time and Frequency Transfer (TWSTFT). Consider two stations A and B each one having a clock, a transmitter and a receiver (Fig. 12.3). Each station transmits a signal to the satellite (uplink) which in turn transmits this signal to the other station (downlink). In order not to perturb the weak incoming signal by the strong outgoing signal in the same antenna of each station, different carrier frequencies are used for the uplink and downlink, e.g., of about 14 GHz for the uplink and about 12 GHz for the downlink at Ku-band frequencies. At a defined time t_A the clock of station A triggers a signal to be transmitted to station B via the satellite and at the same time starts the counter in station A. The same procedure is initiated at station B at t_B. The incoming signals from the satellite are used to stop the counters. Hence, the counters in station A and B measure the time differences

$$\Delta t_A = t_A - t_B + \delta_{B\to A} \tag{12.30}$$

$$\Delta t_B = t_B - t_A + \delta_{A\to B}. \tag{12.31}$$

If the transfer of the signals between the two stations was completely reciprocal the delay times $\delta_{B\to A}$ and $\delta_{A\to B}$ would be equal. The time difference ΔT between the two clocks in station A and B could be computed after exchanging the readings of both stations by subtracting (12.31) from (12.30) to yield $\Delta T = (\Delta t_A - \Delta t_B)/2$. There are, however, effects which lead to different delays in both directions. The time difference between the clocks in station A and

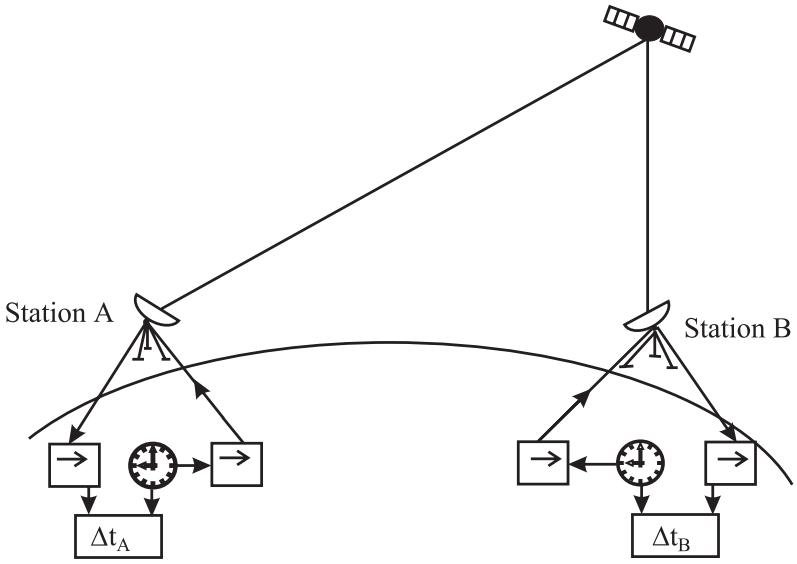


Figure 12.3: Two-way time transfer.

B is calculated as [727]

$$\begin{aligned}\Delta T &= \frac{\Delta t_A - \Delta t_B}{2} + \frac{\left(\tau_A^{\mathrm{up}} + \tau_B^{\mathrm{down}}\right) - \left(\tau_B^{\mathrm{up}} + \tau_A^{\mathrm{down}}\right)}{2} + \frac{\tau_{A\to B} - \tau_{B\to A}}{2} \\ &+ \frac{\left(\tau_A^T - \tau_A^R\right) - \left(\tau_B^T - \tau_B^R\right)}{2} + \Delta\tau_R. \end{aligned} \tag{12.32}$$

Apart from the measured time difference $(\Delta t_A - \Delta t_B)/2$ the second term on the right-hand side of (12.32) $\left[\left(\tau_A^{\mathrm{up}} + \tau_A^{\mathrm{down}}\right) - \left(\tau_B^{\mathrm{up}} + \tau_B^{\mathrm{down}}\right)\right]/2$ is the contribution of the different path delays of the uplink and downlink in both directions. In the case of quasi simultaneous transmission it can be neglected. The third term takes into account the different delays in the transponders of the satellite if different transponders are used for both directions. Different delays in the receiving and transmitting sections of the stations give rise to the fourth correction term of (12.32). The last term $\Delta\tau_R$ is a correction resulting from the Sagnac effect of the rotating Earth. Expressions to correct for these relativistic effects to picoseconds have been given by Petit and Wolf [728].

12.4 Radio Controlled Clocks

In several areas, e.g., in the United States, in Japan, or in Germany, long-wave transmitters are utilised to disseminate timing signals. As an example we refer to the long-wave transmitter called DCF77 operated in Mainflingen near Frankfurt, Germany, at a geographical location $50^\circ\,01'$ north and $09^\circ\,00'$ east by the German Telekom and supervised by the Physikalisch-Technische Bundesanstalt (PTB). Commercial Cs atomic clocks at the location of the plant

are used to steer the time signals of the transmitter. Time and frequency are transmitted by a carrier frequency of 77.5 kHz whose amplitude is modulated by second markers. At the beginning of each second (except for the 59$^{\text{th}}$ second) the amplitude of the carrier frequency is reduced to about 20 % for a duration of 0.1 s or 0.2 s, corresponding to a binary zero or one, respectively (Fig. 12.4). The trailing edge of the envelope represents the marker of the beginning of the second. To identify the beginning of a new minute the 59$^{\text{th}}$ second pulse is omitted.

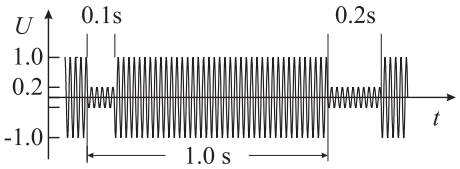


Figure 12.4: The amplitude reduction of the carrier of the 77.5 kHz signal of DCF77 to 20 % defines the beginning of a new second. The amplitude is reduced for 0.1 s or 0.2 s, corresponding to a binary zero or one, respectively.

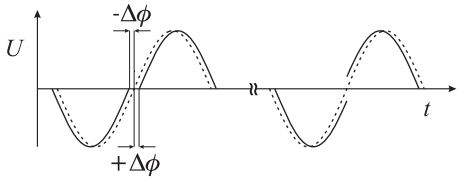


Figure 12.5: Phase modulation of DCF77 with a pseudo random phase shift keying.

The time information is encoded in a binary-coded decimal (BCD) system in a scheme given in Fig. 12.6 where the bits of information available at the 59 seconds of a minute have a specific meaning. The marker bits 21 to 27, e.g., are utilised to identify the current minute of the actual hour. Hence, the 47$^{\text{th}}$ minute of an hour is identified if the bits for the 40, 4, 2,

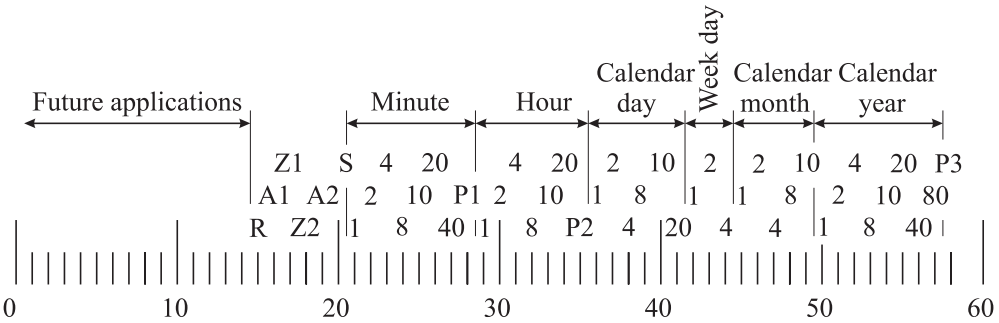


Figure 12.6: Coding scheme of DCF77 according to Becker and Hetzel. [729]

and 1 minutes are set, i.e. if the amplitudes of the 21$^{\text{st}}$, 22$^{\text{nd}}$, 23$^{\text{rd}}$, and 27$^{\text{th}}$ second markers are reduced for 0.2 s (one) and those for second numbers 24, 25, and 26 have a duration of 0.1 s (zero). Similarly, the information is encoded for the current hour, day, day of the week, month, or the last two digits of the number of the year. The coded time information is related to the legal time of Germany. Bits 1 to 14 are reserved for future use, e.g., to alert the public, and the bits 15 to 20 have a specific meaning. Bits Z1 and Z2 carry the information about the time zone (MEZ, i.e. standard European time with the bits Z1 = 0 and Z2 = 1 or MESZ, i.e. daylight saving time with the bits set to Z1 = 1 and Z2 = 0).

To achieve a more accurate time transfer and better use of the frequency spectrum available, in addition to the amplitude modulation of Fig. 12.4, the carrier is phase modulated with a pseudo random phase noise. The phase of the carrier is shifted by $\pm 13°$ (Fig. 12.5) according to a binary sequence, without changing the average phase of the carrier. The modulation frequency of $645.8\overline{3}$ Hz is a subharmonic (77 500/120) of the carrier frequency. Each pseudo noise cycle lasts 793 ms and conveys a bit where the inverted sequence corresponds to the state 1. The pseudo random phase shift keying of the carrier corresponds to a binary random sequence of 29 bits superimposed on the AM second markers. The binary information via the pseudo random noise corresponds to the AM information except for the minute identifier. In the receiver the pseudo random code can be reproduced as a signal and used for cross-correlation with the received pseudo random phase noise. As a result, the received time markers can be detected with higher precision. Despite the phase modulation, the reception of the AM time markers is not affected and the long-term properties of DCF77 as transmitter of a standard frequency is not degraded.

The electromagnetic long-wave signals of DCF77 can reach a receiver along different paths. One important part of the emitted electromagnetic radiation is guided along the surface (Fig. 12.7) and is called the ground wave. Alternatively, the electromagnetic wave can reach the same receiver via the so-called sky wave which is reflected by the ionosphere.[8] Due to the larger damping losses the ground wave becomes less important for distances larger than about 500 km and the maximum distance for the signals of DCF77 is reached when the sky wave leaves the ground tangentially (see Fig. 12.7).

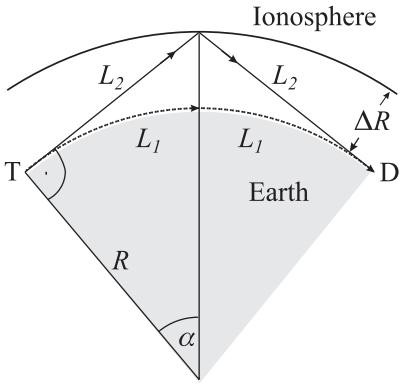


Figure 12.7: The signal from the transmitter T reaches the receiver D by a surface wave (dotted line) travelling the distance $2\,L_1$ and by a sky wave travelling the distance $2\,L_2$.

[8] In the ionosphere between about 70 km and 1000 km above ground the ultraviolet radiation from the sun is capable of ionising air molecules leading to a plasma of free mobile electrons of charge e and mass m_e, together with heavy almost-static ions. A perturbation of the plasma electrons results in free oscillations of the electrons around their equilibrium positions. The corresponding plasma (angular) frequency $\omega_p = \sqrt{\frac{n_0 e^2}{\epsilon_0 m_e}}$ depends on the square-root of the density n_e of the electrons; ϵ_0 is the permeability of the vacuum. Electromagnetic waves with angular frequencies $\omega \ll \omega_P$ force the electrons to oscillate with frequency ω leading to a reflection by the ionosphere.

The maximum distance is calculated from Fig. 12.7 as $L_1 = 2\alpha R \approx 2100$ km for a height of the ionosphere $\Delta R = 90$ km using $\cos\alpha = R/(R+\Delta R)$. With the transmitter located near Frankfurt, DCF77 can serve almost all of Western Europe with a signal of 100 µV/m or greater. The coverage area contracts during day-time and expands during night-time. It is furthermore affected by the available field strength since, e.g., state-of-the-art receivers for wrist watches require a minimum field strength between 15 and 20 µV/m for interference-free operation. The time delay between the sky wave and the surface wave is given as $\Delta L = 2L_2 - 2L_1 = 2R(\tan\alpha - \alpha)$ and leads to a time delay $\Delta t = \Delta L/c$ of about 70 µs for the maximum distance. The delay can easily go as high as 0.5 ms for small distances between the transmitter and the receiver. The DCF77 carrier frequency of 77.5 kHz is a standard frequency with a relative uncertainty averaged over one day of 1×10^{-12}. At the transmitting antenna, the phase time is kept in agreement with UTC(PTB) within the limits of approximately ± 25µs. Larger phase and frequency fluctuations at the location of the receiver are due to the superposition of sky wave and ground wave.

Similar long-wave systems (see Table 12.1) are operated, e.g., in the United States and in Japan. The US system (WWVB) located near Ft. Collins, Colorado operates at 60 kHz carrier frequency with a power of up to 50 kW and the signal is usable throughout most of the USA. The time code uses pulse-width modulation similarly to DCF77. The carrier power is reduced 10 dB at the beginning of each second with the leading edge of the negative pulse representing the marker. For a binary "0", "1", or a position marker full power is restored after 0.2, 0.5, or 0.8 seconds, respectively. A BCD format is used but the time code differs from the one shown in Fig. 12.6. Despite the increasing importance of GPS (to be described

Table 12.1: Characteristics of long-wave standard frequency and time signals. Others may be found in [730]. $\delta\nu/\nu$ represents the fractional uncertainty of the carrier frequency over one day (1σ).

Call sign	Location	Latitude	Longitude	Carrier frequency	$\delta\nu/\nu$
BPC	Pucheng China	34° 57′ N	109° 33′ E	68.6 kHz	
DCF77	Mainflingen Germany	50° 01′ N	09° 00′ E	77.5 kHz	$\pm 1 \times 10^{-12}$
HBG	Prangins Switzerland	46° 24′ N	06° 15′ E	75 kHz	$\pm 1 \times 10^{-12}$
JJY	Oktahadoyayama Japan	37° 22′ N	140° 51′ E	40 kHz	$\pm 1 \times 10^{-12}$
JJY	Haganeyama Japan	33° 28′ N	130° 11′ E	40 kHz	$\pm 1 \times 10^{-12}$
MSF	Rugby UK	52° 22′ N	01° 11′ W	60 kHz	$\pm 2 \times 10^{-12}$
WWVB	Fort Collins, CO USA	40° 40′ N	105° 03′ W	60 kHz	$\pm 1 \times 10^{-11}$

in Section 12.5), long-wave transmitters for dissemination of time will still have applications in the years to come. Advantages include cheaper prices, less power consumption and indoor use of the receivers. The applications are numerous, e.g., phase synchronisation of power plants, the control of traffic lights, air traffic guidance and the synchronisation of computer and telecommunication networks. Moreover, long-wave transmission time services are used for accurate billing of phone calls and stock trading and for radio controlled wrist watches.

12.5 Global Navigation Satellite Systems

Space-based navigational systems have outpaced most other earth-based navigation systems or are about to do so. The most widely known space-based navigational systems are the United States Navigation System with Timing and Ranging Global Positioning System (NAVSTAR GPS) , the Russian Global Navigation Satellite System (GLONASS) that evolved from dedicated military systems and the forthcoming purely civilian European GALILEO system.[9]

12.5.1 Concept of Satellite Navigation

A Global Navigation Satellite System (GNSS) can be divided into three segments often called the space segment, the operational control segment, and the user equipment segment. The space segment comprises a system-dependent number of satellites that deliver the ranging signals and other important data to the user. The operational control segment consists of monitor stations, ground antennas and a master control station. The monitor stations passively track all satellites in view, accumulating ranging data. This information is processed at the master control station to determine satellite orbits and is further transmitted to each satellite via the ground antennas to update each satellite's navigation message.

The satellites of the space segment are equipped with atomic clocks. Each satellite broadcasts a signal with the information on its position and status together with the time of its clock. The user determines his or her own position from the distances to several satellites of known positions using the time the signals take to travel to the user.

To determine its local position on Earth, a GNSS receiver simultaneously uses the signals with time stamps from different satellites and compares them with its local clock. If a signal is received by the user U at the coordinates X, Y, Z from a particular satellite "i" of known position x_i, y_i, z_i (see Fig. 12.8) the time delay between the transmission and reception of the signal is a measure of the distance between the satellite and the user.

If the clock in the user's receiver and the clock onboard the satellite were synchronised, the true range from the first satellite could be calculated from the propagation velocity c and the time delay Δt_1 as $R_1 = c \times \Delta t_1$. A similar evaluation with a second satellite immediately gives the position of the user in the plane containing the two satellites and the user, as one of the two intersection points of the two circles with the two ranges R_1 and R_2 (Fig. 12.8). For a three-dimensional location in space, a third satellite is necessary. However, in general, the quartz clock in the user's receiver will not be synchronised with the atomic clocks of the satellites to the required accuracy since a time difference $\delta t = 1\,\mu s$ would correspond

[9] The history that led to the current development of the GPS is described, e.g. in [731].

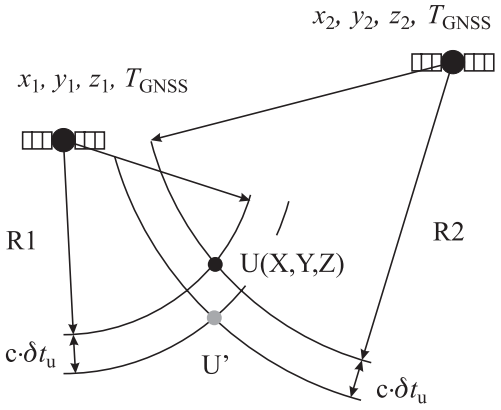


Figure 12.8: Concept of satellite navigation and time determination.

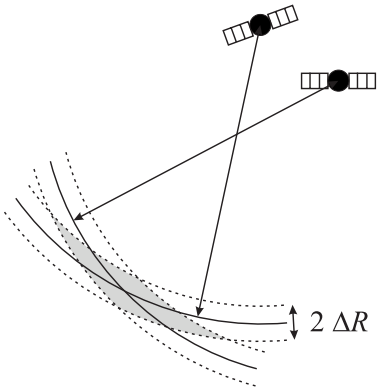


Figure 12.9: Geometric diffusion of precision.

to a systematic error of about 300 m. In the two-dimensional case depicted in Fig. 12.8 it is assumed that the time T_U of the user's clock advances the system time T_{GNSS} from the satellites by $\delta t_u = T_U - T_{\mathrm{GNSS}}$. Consequently the ranges are too big by $c \times \delta t_u$ and lead to the erroneous position U'. The ranges computed from the apparent time differences between the satellite clocks and the user clock, including the offsets resulting from the time difference δt_u, are referred to as "pseudo ranges" $P_i = R_i + c \times \delta t_u$.

If four different pseudo ranges P_i are used to set up four equations for four unknowns, the three spatial coordinates X, Y, Z and the offset of the user clock δt_u can be determined from the set of equations

$$\begin{aligned}
(x_1 - X)^2 + (y_1 - Y)^2 + (z_1 - Z)^2 &= (P_1 - c\delta t_u)^2, \\
(x_2 - X)^2 + (y_2 - Y)^2 + (z_2 - Z)^2 &= (P_2 - c\delta t_u)^2, \\
(x_3 - X)^2 + (y_3 - Y)^2 + (z_3 - Z)^2 &= (P_3 - c\delta t_u)^2, \\
(x_4 - X)^2 + (y_4 - Y)^2 + (z_4 - Z)^2 &= (P_4 - c\delta t_u)^2.
\end{aligned} \tag{12.33}$$

The system of nonlinear equations (12.33) can be solved for the unknowns either by linearisation, in a closed form or by Kalman filtering [731]. The linearised system of equations obtained by a Taylor expansion from (12.33) is iteratively used by employing estimates of the starting position and the clock offset. As a reference ellipsoid primarily the geocentric World Geodetic System of 1984 (WGS84) is employed.

In the following we describe the properties of a satellite based navigation system in more detail using the example of the GPS and keeping in mind that they similarly refer to GLONASS as well as GALILEO.

12.5.2 The Global Positioning System (GPS)

The clocks on board the GPS satellites are related to the GPS system time with a known time difference to UTC(USNO) and, hence, the satellite-based positioning system also disseminates an approximation to UTC. The GPS time scale is based on the readings from the

different atomic clocks in the satellites and in the ground stations in a complicated procedure of data processing. It is steered by the GPS control segment to coincide with UTC(USNO) of the US Naval Observatory to within 1 µs except for an integer number of seconds. At 0 hr January 6, 1980, both scales coincided but now they differ since in contrast to UTC(USNO) the GPS system time has not been adjusted to leap seconds.

The transmitted signal contains the time of the satellite, the difference between the time of the satellite clock and the system time of the GPS, the expected temporal variation of this difference (i.e. the drift of the clock aboard the satellite), the position of the satellite, the (Keplerian) data of the orbit of the respective satellite and the status of all other GPS satellites.

12.5.2.1 Satellite Constellation

The orbit of a satellite is determined by the equilibrium between Earth's gravitational force and the centrifugal force

$$G\frac{M_E M_S}{R^2} = M_S \omega^2 R \tag{12.34}$$

allowing an infinite number of re-entrant Keplerian orbits. $GM_E = 3.986\,004\,418 \times 10^{14}\,\mathrm{m^3/s^2}$ is the product of the gravitational constant G and the mass of the Earth. To achieve an optimised satellite constellation for the GPS, however, a number of constraints had to be taken into account. First, to allow for a continuous determination of the position and time at any spot on Earth, each spot has to be covered by the radiation cones from four satellites at a time. Second, the time to finish one orbit was chosen such that it corresponds to half a sidereal day equivalent to twelve hours minus two minutes to allow positioning of each satellite using distant stars. As a consequence, the satellites appear four minutes earlier each day at a particular location. For the chosen time of a satellite to orbit around Earth the semi-major axis of the respective ellipse with a focus in the centre of the Earth is calculated from Kepler's third law as 26 560 km. In order to keep the corrections due to the second-order Doppler shift and the gravitational red shift as constant as possible, the satellites are operated in almost circular orbits with eccentricities [10] no larger than $\epsilon = 0.02$.

Six parameters are required to describe the motion of the satellite at an epoch, i.e. a point in time. These six parameters could be chosen, for instance, as the three components of the spatial coordinates and the velocities, but since the satellite orbits in a Keplerian ellipse, it is more appropriate to represent the satellite's vector alternatively by the six so-called Keplerian parameters. The so-called terrestrial equatorial system defined by the equatorial plane of Earth is chosen as the reference system and the inertial reference with respect to distant stars is the direction to the vernal equinox Υ, i.e. the intersection point between the celestial equator and the ecliptic.[11]

The orientation of the satellite's plane in space with respect to the equatorial plane is given by two parameters: the inclination and the right ascension of the ascending node, i.e. that point where the satellite passes the equator on its way from south to north. The orbital ellipse of

[10] The eccentricity relates the semi-major axis a and the semi-minor axis b of the ellipse by $b = a\sqrt{1-\epsilon^2}$.

[11] The vernal equinox is that point in the sky which the sun reaches on its annual path at the beginning of spring.

the satellite is characterised by the length of the semi-major axis and the eccentricity ϵ of the ellipse.

The orientation of the ellipse in the orbital plane is described by an angle between the direction to the ascending node and the perigee. The sixth parameter is the time-dependent true anomaly, i.e. the angle between the direction to the perigee and the direction to the actual position of the satellite. The GPS satellite constellation consists of nominally 24 satellites with each four satellites orbiting in six fixed orbital planes inclined 55° from the equator.

12.5.2.2 Satellite Clocks and Signals

Aboard each satellite there are four clocks, either caesium clocks or rubidium clocks or both, used to transmit the satellite time. Since the clocks on board are in general less accurate compared to ground-based clocks, the information about the deviation between the time of the satellite clock and the system time is transmitted simultaneously.

From the onboard clocks four phase-coherent frequencies of 1.023 MHz, 10.23 MHz, L1 $= 1540 \times 1.023\,\text{MHz} = 1.575\,42\,\text{GHz}$ and L2 $= 1200 \times 1.023\,\text{MHz} = 1.227\,60\,\text{GHz}$, are generated. L1 and L2 are the carrier frequencies of two signals transmitted by each satellite in the so-called microwave L band. The carriers from each satellite are modulated by an individual spread code, i.e. a binary Pseudo Random Number (PRN) code (Fig. 12.10). Two code sequences are used, referred to as the Coarse Aquisition code (C/A code) and the Precision code (P code). The C/A code is a pseudo random number of $2^{10} - 1 = 1023$ bits and has a chipping rate [12] of 1.023 MHz and hence repeats itself after a millisecond. The P code repeats every 266.4 days. Each satellite is attributed an individual one-week segment of this code. Hence, the C/A code and the P code allow the user to identify unambiguously the satellite which transmits the signal with the aid of the codes stored in the GPS receivers.

Both, L1 and L2 are modulated with the P code and L1 is furthermore modulated with the C/A code. In order to transmit further information besides the PRN sequence, the PRN sequence of each code is inverted (state 1) or not inverted (state 0) with a bit rate of 50 Hz (Fig. 12.11). To modulate the high-frequency signal L1 with both codes it is split into two components shifted by $\pi/2$. One component is modulated by the C/A code, the other one by the P code and both components are superimposed again and transmitted. Hence, depending on the state of the P code and the C/A code the transmitted signal can exhibit four different phases (0/0, 0/1, 1,1, 1/0). The reader interested in more details is referred to reference [731].

12.5.2.3 Uncertainties Associated with GPS

The uncertainty with which users can determine their position, velocity or time from the GPS, depends primarily on a variety of effects that influence the determination of the pseudo range to a given satellite. The uncertainty of the pseudo range is referred to as the "User-Equivalent Range Error" (UERE). The uncertainty is furthermore increased by the combined effect of the geometry including the satellites and the user. As an example, illustrated in Fig. 12.9, the uncertainties in the pseudo ranges from the satellites can lead to a large uncertainty in the

[12] The term "chip" is used rather than "bit" to indicate that no information is conveyed in the sequence (except for the identification of the satellite).

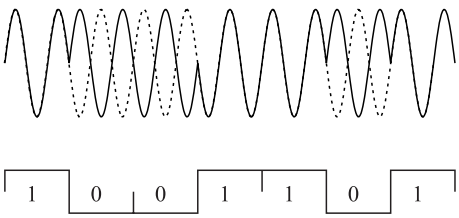


Figure 12.10: Phase modulation with a Pseudo Random Number Code.

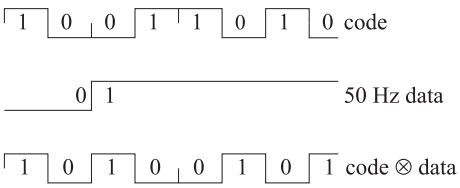


Figure 12.11: Mixing code and data in the GPS signal.

determination of the position along a particular direction if the satellites are seen by the user from a similar angle. This effect is termed "Geometrical Dilution of Precision" (GDOP) and it is taken into account by multiplying the UERE with a GDOP factor. This direction-dependent geometry factor is derived by solving the linearised pseudo range equations simultaneously for four satellites. From the analytical treatment it follows that the GDOP factor is inversely proportional to the volume of a polyhedron whose corners are formed by the positions of the four satellites and the user.

The uncertainty of the pseudo range depends on a variety of factors that lead to deviations of the pseudo ranges from the true ranges and the capability to correct them.

Ephemerides The exact knowledge of the position of each satellite is mandatory to determine the user's position and time. The satellite does not exactly follow a Keplerian orbit due to gravitational and other perturbations. These perturbations include drag from the outskirts of the atmosphere or forces resulting from the radiation pressure of the solar wind. The gravitational perturbations result, e.g., from the oblateness of the Earth and from the tidal effects of the sun and the moon (see also Fig. 13.4). The deviation from the sphericity of the Earth leads to a slow precession of the satellite orbit. As a result of these and other effects the orbit of a satellite is not stationary if not corrected for by the satellites's thrusters for manoeuvring and station keeping, i.e. keeping the satellites in their proper orbital positions. The positions of the satellites are determined by the monitor stations that use the measured pseudo range data together with their exactly known position and time, to determine the position of the satellite and the time of the satellite's clock. The master control station processes the data from the monitor stations to obtain accurate estimates of a satellite's ephemeris and time, together with predictions into the future. The ephemeris data, the formed almanac [13] and the clock data are transmitted by the satellite and are necessary for the user to determine his or her position and time.

Satellite Clock Uncertainty According to General Relativity the frequency of a clock depends on the gravitational potential (see (12.15) and (12.17)) which, according to the equivalence principle, is composed of the gravitational and centrifugal potential. The potential experienced by a clock orbiting around the centre of Earth at a distance R with velocity $v = \omega \times R$

[13] The almanac data comprise satellite constellation information including "health" of the satellites.

is

$$U = -\frac{GM_E}{R} - \frac{\omega^2 R^2}{2}. \tag{12.35}$$

From (12.35) and the data from reference [263] one calculates the potential $U_{\text{surface}} = -62.6\,(\text{km/s})^2$ for a clock located on the surface of the geoid. If the clock is onboard a satellite, the combination of (12.35) and (12.34) leads to

$$U_{\text{satellite}} = -\frac{GM_E}{R} - \frac{GM_E}{2R} = -\frac{3}{2}\frac{GM_E}{R}. \tag{12.36}$$

The difference in the potentials of a clock in a satellite and a clock on the surface of the Earth leads to a fractional frequency difference of

$$\frac{\Delta\nu}{\nu} = \frac{\Delta U}{c^2} = \frac{1}{c^2}\left(-\frac{3}{2}\frac{GM_E}{R} + 62.6 \times 10^6 \frac{\text{m}^2}{\text{s}^2}\right). \tag{12.37}$$

The corresponding time difference per day (Fig. 12.12) calculated by use of (12.37) is negative for low orbiting satellites and goes to zero at a height of about 3190 km above ground corresponding to half of Earth's radius. It becomes positive for satellites in higher orbits, e.g., for the GPS satellites or for geo-synchronous satellites. Hence, when viewed from the surface

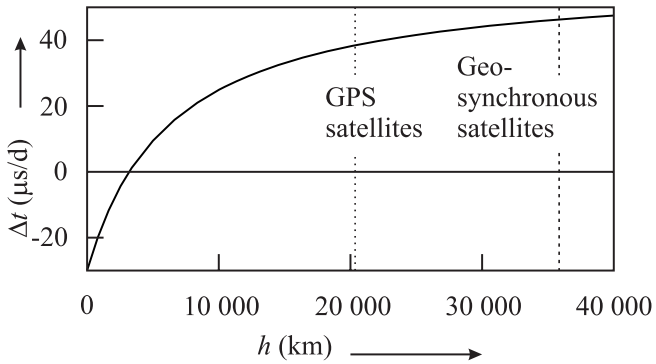


Figure 12.12: Time difference per day between a clock on board a satellite at height h above ground and a clock on the surface of Earth according to (12.37).

of the Earth the clocks onboard the GPS satellites at $R = 26\,600$ km advance by 38.5 µs/d. To compensate for this effect the atomic clocks on board are given a fixed fractional frequency offset of $-4.464\,733 \times 10^{-10}$ [725, 732] to deliver a frequency of 10.229 999 995 432 6 MHz rather than 10.23 MHz. This steering, however, does not take into account the slight eccentricity of the GPS satellite orbits. The satellites dip into lower gravitational potential and have a higher velocity at perigee. Both effects reduce the satellite's clock rate when viewed from Earth's surface. At apogee the satellite clock runs faster as a result of the lower velocity and the higher gravitational potential. This effect can lead to a maximum deviation of 70 ns [731].

Delay in the Atmosphere The electromagnetic waves transmitted by the satellites in the vacuum propagate differently when passing through Earth's atmosphere. The largest effects occur in the ionosphere.[8] The index of refraction n_p for the propagation of the phase of an electromagnetic signal of frequency ν can be represented to a good approximation [731] by

$$n_p = 1 + \frac{c_2}{\nu^2}. \tag{12.38}$$

The coefficient $c_2 = -40.3 \times n_e\,\mathrm{Hz}^2$ depends on the electron density n_e along the path from the satellite to the user. The electron density integrated along this path is called the Total Electron Count (TEC) representing the number of free electrons in a volume column of an area of $1\,\mathrm{m}^2$. The TEC varies between $10^{16}\,\mathrm{m}^{-2}$ and $10^{19}\,\mathrm{m}^{-2}$ depending on the location of the user, day time, satellite elevation, sunspot activity and others. Since the GPS signal is modulated it represents a frequency band of finite width. Hence, the group velocity is given by $n_g = n_p + \nu dn_p/d\nu$ as

$$n_g = 1 - \frac{c_2}{\nu^2}. \tag{12.39}$$

The (group velocity) delay time of the signal resulting from the influence of the ionosphere is given by

$$\Delta T = \frac{40.3 \times \mathrm{TEC}}{c\nu^2}. \tag{12.40}$$

The use of two different transmission frequencies L1 and L2 leads to a difference in the delay times of

$$\Delta\widetilde{T} \equiv \Delta T(\mathrm{L1}) - \Delta T(\mathrm{L2}) = \frac{40.3 \times \mathrm{TEC}}{c}\left(\frac{1}{\nu_1^2} - \frac{1}{\nu_2^2}\right) = \Delta T(\mathrm{L1})\frac{\nu_2^2 - \nu_1^2}{\nu_2^2} \tag{12.41}$$

and, hence, the delay ΔT_1 on the frequency L1 can be determined from the measured delay $\Delta\widetilde{T}$ by use of (12.41). The delay on L2 can be computed by multiplication with the ratio $\nu_1^2/\nu_2^2 = (77/60)^2$.

If only L1 is measured the influence of the ionosphere has to be corrected by referring to an empirical model. Parameters of the model are included in the broadcast message. The uncertainty after applying such a correction can be as large as 50 % of the effect itself.

The lower part of the atmosphere called the troposphere is nearly non-dispersive for frequencies up to 15 GHz. Thus, the delay of a signal within the troposphere cannot be measured by comparing the signals L1 and L2. In the troposphere the index of refraction depends on the temperature, pressure and humidity. The correction to the path length depending on these parameters and on the elevation angle of the satellite has to be taken into account by semi-empirical models. The corresponding correction of the path length corresponds to a few metres.

Accuracy of Time and Position Determination Between the years 1990 and 2000 the performance of the GPS system was degraded intentionally by the so-called Selective Availability (SA). The SA was accomplished by dithering the clock on board the satellites. Hence, the unperturbed data could be used only by military and other authorised personell with the knowledge of the regularities of the manipulations. Various effects contribute to the uncertainty

Table 12.2: Pseudo range uncertainty budget according to [731] grouped into the contributions from the space segment, the control segment and the user segment.

Source of uncertainty	Uncertainty
Satellite clock instability	3.0 m
Satellite perturbations	1.0 m
Other perturbations	0.5 m
Ephemeris prediction	4.2 m
Other	0.9 m
Ionospheric delay	2.3 m
Troposheric delay	2.0 m
Receiver noise	1.5 m
Multipath	1.2 m
Other	0.5 m
Total	6.6 m

Table 12.3: Uncertainty budget for the coarse acquisition SPS (C/A) code and the precision PPS (P) code using the uncertainty of the pseudo range (UERE) data as, e.g., taken from Table 12.2 and including the geometrical diffusion of precision (GDOP) (Fig. 12.9). The large difference between the SPS and the PPS mainly results if the Selective Availability is included.

	C/A	P
Position (3D)	95 m	17 m
Horizontal	56 m	10 m
Vertical	72 m	13 m
Time	100 ns	87 ns
Velocity	0.1 m/s	

of the determination of the pseudo range (see Table 12.2). The corresponding uncertainties in determining a position and time information (see Table 12.3) largely depend whether the selective availability is included or not.

To increase the accuracy of the determination of the position and the timing, sometimes a so-called "differential GPS" is applied. This method relies on an extension of the GPS system that uses land-based radio beacons to transmit differential position corrections from a fixed receiver at known position to mobile GPS receivers.

12.5.2.4 Time and Frequency Transfer with GPS

Table 12.4 shows the relative uncertainties that can be achieved for the time and frequency transfer with the different transfer methods via GPS. One-way GPS measurements rely on the data transmitted from the GPS satellite for calibration. In the single-channel common-view method each one of the two GPS receivers at the two different locations tracks the same GPS satellite at the same time. In the multi-channel common-view method each receiver collects data from all satellites in view. [14] As compared to the single-channel method more data are available leading to smaller statistical uncertainties.

So-called "geodetic" GPS receivers can also be used to perform time transfer. This is currently done on a world-wide basis in the framework of the International GPS Service for Geodynamics (IGS). These receivers process all GPS observables (Code PA, P1, P2, phase

[14] For this purpose the BIPM recommends a schedule for the reception of the signals from the 24 satellites in the different areas around the world. By averaging typically 20 to 30 comparisons per day, the time scales of two timing centres on the same continent can be compared with an uncertainty of a few nanoseconds which increases to about 10 ns – 20 ns for intercontinental comparisons.

Table 12.4: Uncertainties (2σ) achievable with GPS measurement techniques in a measurement time of 24 hours (according to [733]).

Technique	Fractional timing uncertainty	Fractional frequency uncertainty
One way	< 20 ns	$< 2 \times 10^{-13}$
Single-channel Common-view	≈ 10 ns	$\approx 1 \times 10^{-13}$
Multi-channel Common-view	< 5 ns	$< 5 \times 10^{-14}$
Carrier-phase Common-view	< 500 ps	$< 5 \times 10^{-15}$

L1, L2) including the phase of the carrier signal. Since the carrier phase is heavily distorted by the Doppler effect the Doppler-shifted frequency measurements have to be integrated to reconstruct the carrier phase. Observing the phase allows one to compare the frequencies of two distant clocks very accurately. If time scales are to be compared, the ambiguity of the phase has to be resolved, i.e. the accurate numbers of wavelengths between the satellite and the receivers have to be determined. This is achieved if a long series of uninterrupted measurements with the precision code is available. The measurements of the various geodetic receivers are evaluated together with the necessary satellite data in a network. In Europe this is done at the Centre for Orbit Determination in Europe (CODE) in Bern, Switzerland which is part of the IGS. In essence, the precision from the phase measurements corresponds to about 10 ps corresponding to a few millimetres.

Time transfer with uncertainties below a nanosecond requires that the location of the receiver is accurately determined in the station where the time signal is measured, as the travelling time in a 1 m coaxial cable already leads to a delay of 5 ns. The delay of each receiver D_A and D_B has to be determined and taken into account when calculating the time difference between two stations A and B from the measured time difference $\delta_{\mathrm{geod.rec.}}$

$$\delta_{\mathrm{geod.rec.}} = (T_A + D_A) - (T_B + D_B) = (T_A - T_B) + (D_A - D_B)\,. \tag{12.42}$$

The difference between the local delays $D_A - D_B$ in (12.42) can be determined in a so-called "common clock experiment" where two geodetic receivers are set up close to each other in the same laboratory and referenced to the same clock. In this case, $T_A - T_B = 0$ holds in (12.42). In such a so-called "zero-baseline experiment" the errors resulting from inadequate models of the troposphere or ionosphere cancel. For comparisons where the two station are separated again by a larger baseline the resulting differences contribute to the measured time difference.

The quality of time transfer in the year 2001 can be seen from a comparison between the Two-Way Satellite Time and Frequency Transfer method and a time transfer by GPS common-view (Fig. 12.13). As the TWSTFT leads to smaller uncertainties, the standard deviation of the difference data in Fig. 12.13 of 2.6 ns is a measure of the uncertainty achievable with the

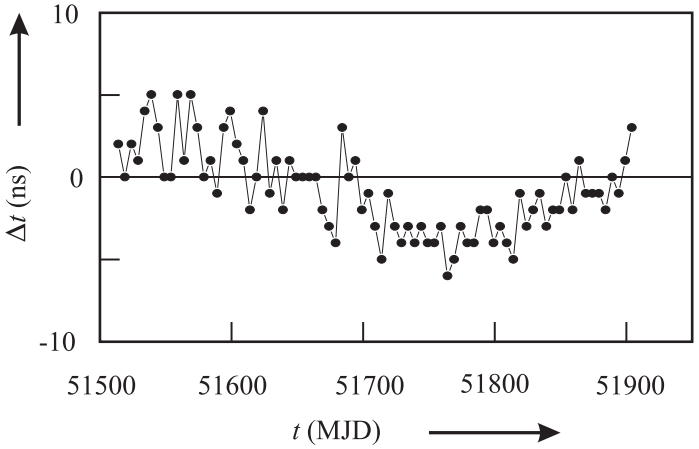


Figure 12.13: Measured differences between Two-Way Satellite Time and Frequency Transfer and GPS C/A-code common-view for a link between NPL and PTB measured for the Modified Julian Date MJD, i.e. a continuous count of the days. (MJD = 0 corresponds to 17 November 1858, 0 h.)

latter method. The long-term variations are probably also caused by GPS but this is difficult to verify.

In a transatlantic time and frequency comparison between the PTB and the USNO, [734] instabilities of 10^{-13} were reached after 300 s and the 10^{-14} level was reached after 30 000 s. The difference between the time transfer by geodetic receivers and by the two-way satellite time and frequency transfer, showed slight seasonal variation of a few nanoseconds amplitude attributed to multipath and thermal effects.

12.5.3 Time and Frequency Transfer by Optical Means

Few attempts have been made to use optical radiation for time and frequency transfer using either free-space or fibre-bound transmission. The LASSO experiment (Laser Synchronization from Stationary Orbit) [735] used satellites carrying a clock. Nd:YAG laser pulses were fired from stations at Grasse (France) and McDonald (Texas). 100 ps uncertainties of the time transfer were expected. However, later the uncertainties had to be raised to 1.5 ns. Time Transfer by Laser Link (T2L2) has been projected for the space stations Mir and the International Space Station (ISS) [736] but were not realised at the time of writing. This method needs a clear sky without clouds and, hence, cannot be utilised at every location and instant. It is, however, ideal for time and frequency transfer between satellites.

Optical fibre links have been used to transfer frequencies in local and regional fibre networks. Optical frequencies at 385 THz ($\lambda = 778$ nm) have been transferred by using a 3 km long single-mode 1.3 µm optical fibre to connect two laboratories in Paris [737]. The frequency shift introduced by the fibre was measured after sending back the light through the fibre that has been frequency shifted by means of an acousto-optic modulator (AOM). The

measured shift of 0.4 Hz was attributed to a time delay probably due to a temperature effect resulting, e.g., from a temperature variation of a few ten microkelvin per hour. Frequency jitter resulting from acoustic pressure can easily lead to phase variations of a several radians of phase noise and can easily lead to kilohertz broadening of the carrier. Ma *et al.* [738] have shown how this broadening can be reduced to the millihertz domain by using a double-pass heterodyne measurement where the measured phase excursion was divided by two in order to obtain a correction signal that was fed to a phase-compensating AOM. This approach is valid to the extent that the counterpropagating signals suffer from the same phase perturbations. This reciprocity is not necessarily given for long-haul transmissions.

Optical and radio frequency standards located at the US National Institute of Standards and Technology (NIST) and the JILA in Boulder, Colorado have been connected through a 3.45 km optical fibre link [739]. The comparison of the optical and microwave frequency of an iodine-stabilised Nd:YAG laser at 1064 nm and a H maser by means of optical femtosecond combs allowed the comparison of both frequencies before and after transmission in both laboratories.

An even longer distance optical link between the Laboratoire des Physique des Lasers in northern Paris and the BNM–SYRTE in the centre of Paris, separated by about 13 km, has been used to connect the frequency base of a femtosecond laser comb to the microwave standards of SYRTE [406]. The 1.55 µm laser beam, amplitude modulated by a 100 MHz signal, was transmitted through a modified commercial network where a few tens of network interconnection points of the standard single mode fibres were fused. The transmitted signal derived from a H maser was compared after a complete roundtrip of about 85 km with the original signal and the extra noise from the optical link was found to correspond to a relative Allan deviation $\sigma_y(\tau = 10\,000\,\mathrm{s}) < 10^{-15}$. More recent measurements have shown an order of magnitude lower noise level which is expected to be further reduced by using a higher modulation frequency of 1 GHz. These noise characteristics compare favourably with that of free-space microwave links.

12.6 Clocks and Astronomy

Accurate measurements of time and frequency and the synchronisation of clocks have a virtually unlimited number of applications for precise cosmology, interstellar medium physics, orbital evolution measurements and space exploration, to name only a few. In the following, some examples will be discussed.

12.6.1 Very Long Baseline Interferometry

Radio astronomy where telescopes gather and concentrate the radio waves from astronomical sources has contributed incredibly to today's knowledge about astronomical objects. The minimum resolvable angular separation θ of any imaging system is

$$\theta = \alpha \frac{\lambda}{b} \tag{12.43}$$

as a result of diffraction due to a limited aperture b. The constant α is of the order of unity and depends on the geometrical shape of the aperture and on the illumination across the aperture.

The desired highest angular resolution with radio waves of wavelength λ in the centimetre to metre range would require, according to (12.43), a size of a radio telescope which is impossible to construct with current technology.

The diffraction limit of (12.43) results from the interference of different partial waves originating from different spots on the telescope. Hence, the resolution can be increased if the signals from different single telescopes are combined with the proper phase relationship. The signals from two different receivers are then cross-correlated and the resulting fringe pattern can then be analysed to yield, e.g., an image of a distant astronomical object or a precise location of an astronomical radio source. In this way a so-called "aperture synthesis telescope" can be formed. For example, the Very Large Array (VLA) links 27 antennas of a radio telescope located near Socorro, New Mexico, to a maximum size of 36 km. The resolution of the VLA at the highest frequency of 43 GHz is 0.04 arcseconds.

For the so-called Very Long Baseline Interferometry (VLBI) the elements of the interferometer may be separated by thousands of kilometres or be spread over several continents (Fig. 12.14). The Very Large Baseline Array (VLBA), for instance, is an array of radio telescopes dedicated to VLBI and extending from the Hawaian to the Virgin islands. For such

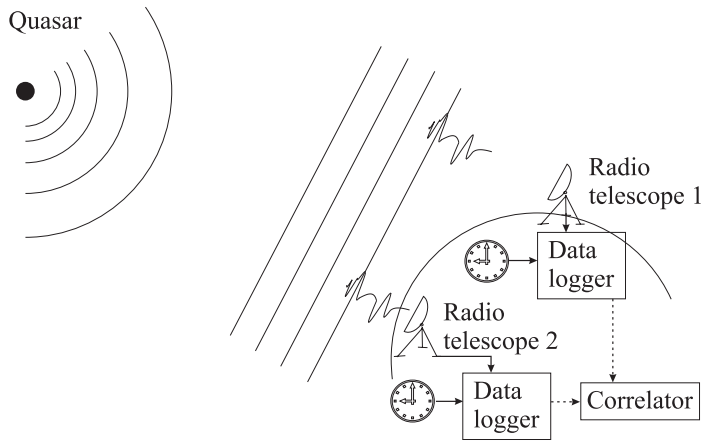


Figure 12.14: Principle of Very Long Baseline Interferometry (VLBI).

large separations it was no longer practical to physically combine the signals from the single telescopes in real time. Hence, the data are recorded first in digitised form with time stamps on magnetic tapes and the correlation is performed afterwards. The recording tapes are synchronised by the use of hydrogen masers in the respective stations. The correlator removes the Doppler shift and geometric delay due to the position and motion of the two stations and then cross-correlates them. Hence, VLBI can be thought of as a measurement of the time difference between the arrival of a radio signal at two telescopes from a very distant astronomical object e.g. a QUASAR.[15]

[15] The acronym QUASAR i.e. QUAsi StellAr Radio source was coined when during the sixties of the last century the positional accuracy of radio observations became high enough that visible objects could be attributed to known radio sources [740, 741]. The large red shift with $0.1 \lesssim v/c \lesssim 5$ indicated that these objects are several 10^9 light

The positions of these extragalactic radio sources (mostly QUASARs) as measured by VLBI to fractions of a milliarcsecond have been used to set up a celestial reference frame adopted by the International Astronomical Union. The great distances of these sources make their motions across the sky virtually undetectable and, hence, they form a true inertial reference frame. This system is used to relate the positions of stars in our galaxy and for measuring Earth's position and orientation very accurately. Theses data are used by geophysicists to derive models for the influence of the atmospheric angular momentum, ocean tides or the elastic response of the solid Earth. At the same time VLBI measurements allow the determination of the relative positions of the antennas for a one-day session as good as 1 mm in the horizontal and 3 mm in the vertical. From these data also valuable information about tectonic plate motion is derived.

The largest baseline on Earth is limited to its diameter of about 12 750 km. The baseline for VLBI can be further increased by implementing space-borne radio telescopes into the system. E.g., in the VLBI Space Observatory Programme (VSOP) the Japanese radio telescope called HALCA with its 8 metre radio antenna was launched in 1997 into an elliptical Earth orbit. In combination with surface-based antennas a baseline of over 30 000 km is achieved. At 5 GHz the VSOP mission allows imaging at resolutions lower than a milliarcsecond.

12.6.2 Pulsars and Frequency Standards

There was great excitement when in 1967 the first radio sources were observed that emitted periodic signals [742]. These radio sources were called pulsars and it was soon realised that they emit broad-band pulses of radiation with periods between about a millisecond to several seconds. At the end of 1998 more than 1000 pulsars had been detected. Since the time interval τ between the pulses is constant to about $\Delta\tau/\tau \approx 10^{-3}$ one is forced to assume that these pulses are emitted by a body of a rather rigid structure. One can think of a rapidly rotating body with a fixed radio source whose radiation cone sweeps across the Earth like a searchlight or the beam from a lighthouse. For a rotating body, one derives immediately an upper limit for its size by considering that the angular velocity at the surface cannot exceed the speed of light c. The radius R of a pulsar emitting one pulse each millisecond is thus less than 50 km. The pulsar PSR B1937+21 [743] [16] has a rotational period of 1.6 ms. It is not expected that much faster pulsars will be discovered since the balance between the centripetal and the gravitational force on the surface of the body sets an upper limit for the rotational

years away and hence, have a huge absolute luminosity. The temporal variation of the radiation of up to a factor of ten within a few days is compatible with a size not larger than a few light days since a source cannot fluctuate faster than a light beam takes to travel across the source. The true nature of QUASARs is still not known. A plausible explanation refers to a black hole in the centre of a galaxy with a 10^9-fold mass of the sun. When the black hole sucks in gas and stars from the vicinity, the accelerated ionised gas masses may generate huge magnetic fields, thereby emitting high energetic radiation.

[16] The position of any stellar object can be defined by two angles named as declination and right ascension. If we think of such a stellar object as being fixed on a celestial sphere whose centre is represented by Earth, the declination and right ascension correspond to the geographic longitude and latitude, respectively. The declination is the angle from the celestial equator to the object counted from zero to + 90 degrees north and -90 degrees to south. The right ascension is counted in hours from the spring point from west to east. Pulsars are identified by the prefix PSR. The pulsar named PSR B1937+21, for instance, can be found at the angular coordinates 19 hours 37 minutes declination and 21 degrees (north) right ascension.

frequency $\Omega = \sqrt{GM/R^3}$ depending on the radius R, the gravitational constant G, and the mass $M = 4\pi R^3\rho/3$. Using the highest density ρ that is known today, i.e. the typical density of a neutron star of $\rho \approx 10^{17}$ kg/m^3 one calculates a rotational period of 1.2 ms. Hence, it is assumed that these pulsars are rotating neutron stars.[17] If there was a magnetic field before the collapse of the star it is enhanced during the collapse. Consider a radius of $R_i \approx 7 \times 10^8$ m and $R_f \approx 5 \times 10^4$ m before and after the collapse, respectively. Conservation of the magnetic flux requires $B_i 4\pi R_i^2 = B_f 4\pi R_f^2$ and leads to an enhancement of the magnetic field B of about eight orders of magnitude with magnetic fields of about 10^8 T.[18] The origin of the pulsar radiation is readily explained in the "lighthouse model" (Fig. 12.15). As the

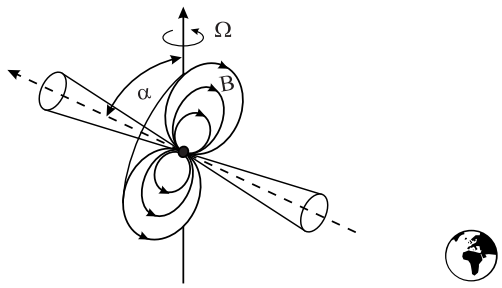


Figure 12.15: A rotating neutron star emits radiation into a rotating cone like a cosmic lighthouse.

neutron star is rotating with angular velocity Ω, charged particles are accelerated along the magnetic field lines in the magnetosphere. The accelerated particles emit electromagnetic radiation predominantly in the regions near the magnetic poles of the neutron star in a cone around the magnetic field axis of the neutron star. Since the magnetic field axis in general does not coincide with the rotation axis, the radiation beam can sweep across the observer once per rotation period like the light of a lighthouse. Hence, the period of the pulses is determined by the rotation period of the neutron star. The pulses are detected typically at radio frequencies between a few hundred megahertz and several gigahertz. Even if the power emitted by the pulsar is incredibly high only a tiny fraction can be detected on Earth. Typically, the spectral irradiance detected from the pulsars is at a low level between 10^{-29} W m^{-2} Hz^{-1} and 10^{-27} W m^{-2} Hz^{-1} at a reference frequency often chosen close to 400 MHz [745, 746]. Consequently, the individual radio pulses are often hidden in the noise. Since, however, the pulses are occurring periodically, standard techniques applying phase-sensitive detection can

[17] A neutron star may originate from a star which has used up its fusion fuel and has a mass $5M_\odot \leq M \leq 10M_\odot$ where $M_\odot$ is the stellar mass of our sun. In the steady state of a star there is a balance between the gravitational acceleration and the radiation pressure. When the star burns out the radiation pressure is reduced and the star collapses thereby heating up and blowing away its corona in a supernova explosion. The temperature of the remaining matter is high enough to ionise the atoms and to allow the protons (p^+) and electrons (e^-) to produce an inverse neutron decay ($p^+ + e^- \rightarrow n + \nu$). The neutrino ν is emitted and all the matter left consists of neutrons (n) leading to a so-called neutron star.

[18] In fact, even higher magnetic fields have been observed by a pulsar emitting bursts of low-energy gamma rays rotating with a period of about 7.4 s. A magnetic field of 8×10^{10} T has been deduced from the the rotation period and the slowdown of that rotation [744].

be used to reduce the noise and to recover the signals from the pulsars. The digitised signal from the telescope is coherently summed in slices of the expected pulse period. It is found experimentally that the average profile of the pulse shows a unique structure characteristic for each pulsar (Fig. 12.16) to some extent depending on the frequency of detection. About 3 %

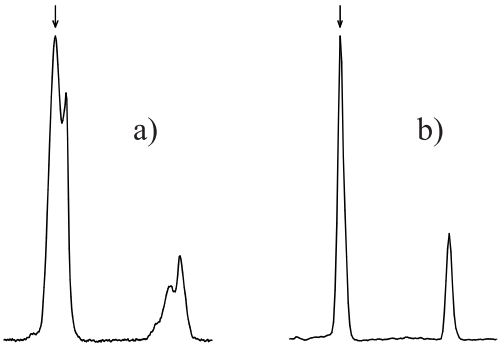


Figure 12.16: Average pulse profiles for PSR B1855+09 and PSR B1937+21 recorded at a frequency of 1.4 GHz and 2.4 GHz, respectively, [747]. Courtesy of V. Kaspi. These integrated pulse profiles show the "fingerprints" of the pulsars. More of them can be found in [746] and [748].

of the average pulse profiles show an interpulse at about half of the period after the main pulse (see e.g. that of PSR B1937+21 shown in Fig. 12.16 b). The main pulse and the interpulse can be explained if one assumes that both beam cones of the neutron star sweep across the observer on Earth, indicating that they result from the opposite magnetic poles of the neutron star. Double-pulse structures may also result from a hollow-cone shape of the emission beam [748].

The large number of known pulsars can be divided into two groups with distinct properties [745]. The larger group named "normal" or "slow pulsars" show pulse periods, i.e. rotation periods P of the order of one second (33 ms $< P <$ 5 s). The rotation period increases for most of the normal pulsars by typically $\dot{P} \approx 10^{-15}$ s/s. The second group called "millisecond pulsars" have periods of 1.5 ms $\lesssim P \lesssim$ 30 ms and spin down with $\dot{P} \approx 10^{-19}$ s/s. There are other differences between the slow and millisecond pulsars as their different ages of 10^5 years $\lesssim \tau < 10^9$ years [746] and $\tau \approx 10^9$ years or their surface magnetic field of $B \approx 10^8$ T and $B \approx 10^4$ T, respectively. Furthermore, about 80 % of the millisecond pulsars are observed to have orbital companions in contrast to the slow pulsars where this fraction is below 1 % [745]. Plausible physical models that have been developed to explain these observations and to derive these data describe a pulsar with its large magnetic field by a rotating magnetic dipole and an associated classical magnetic dipole moment M. The magnetic dipole moment rotates with the rotational frequency Ω and there is an angle α between the axis of the dipole and the rotation axis. According to classical electrodynamics the rotating magnetic dipole emits radiation and the total radiated power is given by

$$\frac{dE}{dt} = \frac{2(M \sin\alpha)^2 \Omega^4}{3c^2}. \tag{12.44}$$

The radiated power slows down the spinning neutron star and reduces the rotational energy

$$E_{\mathrm{rot}} = \tfrac{1}{2}\Theta\Omega^2, \tag{12.45}$$

where Θ is the moment of inertia of the neutron star. For a sphere of a radius $R \approx 15\,\mathrm{km}$ and a density of $\rho \approx 10^{17}\,\mathrm{kg/m^3}$ the moment of inertia is $\Theta = 2/5MR^5 = 8/15\pi\rho R^5 \approx 1.3 \cdot 10^{38}\,\mathrm{kg\,m^2}$. The loss of rotational energy can be calculated from the observed rotational frequency $\Omega = 2\pi/P$ and its derivative $\dot{\Omega} = -2\pi\dot{T}/P^2$ as

$$\frac{dE_{\mathrm{rot}}}{dt} = \Theta\Omega\dot{\Omega} = -4\pi^2\Theta\frac{\dot{P}}{P^2}. \tag{12.46}$$

For 29 slow pulsars Camilo and Nice [746] conclude $10^{23}\,\mathrm{W} \lesssim \dot{E}_{\mathrm{rot}} \lesssim 10^{26}$ W. The latter value corresponds roughly to the power irradiated by our sun resulting from the nuclear fusion processes. Equating the loss of rotational energy (12.46) and the total energy radiated by the magnetic dipole (12.44) leads to

$$\dot{\Omega} = \frac{2(M\sin\alpha)^2}{3\Theta c^3}\Omega^3. \tag{12.47}$$

As an example, from the magnetic dipole moment M (12.47) the surface magnetic induction B can be estimated as $B \propto \sqrt{P\dot{P}}$.

12.6.2.1 Pulsar Timing

For accurate determinations of parameters that should reflect the properties of the pulsar, the arrival times of the pulsar signals have to be corrected for a number of effects affecting the signals that are detected by an antenna on the moving Earth. As a first step, the data are referenced to an inertial observer. To a good approximation the centre of gravity of the solar system (barycentric frame) can be regarded as such an inertial system to get rid of the sinusoidal yearly variation due to Earth's revolution around the sun and the sinusoidal monthly variation due to the revolution around the Earth-Moon barycentre. The arrival time at the solar barycentre t_b is obtained from the arrival time t measured at the telescope by

$$t_b = t + \frac{\vec{r}\cdot\hat{n}}{c} + \frac{(\vec{r}\cdot\hat{n})^2 - |\,\vec{r}\,|^2}{2cd} - \frac{D}{f^2} + \Delta_{E\odot} + \Delta_{S\odot} + \Delta_{A\odot} \tag{12.48}$$

where t is the observed topocentric time of arrival, $\vec{r}$ is a vector from the barycentre to the telescope, and $\hat{n}$ is a unit vector from the barycentre to the pulsar. c is the speed of light, d is the distance to the pulsar, D is the dispersion constant of the interstellar medium resulting from the ionised interstellar plasma, and f is the radio frequency. $\Delta_{E\odot}$ is the Einstein delay resulting from the red shift and the time dilation, $\Delta_{S\odot}$ is the so-called Shapiro delay resulting from the curvature of space-time in the vicinity of the sun and $\Delta_{A\odot}$ is the aberration resulting from Earth's rotation [749]. The correction is performed by using a planetary ephemeris that has been computed, e.g., in the so-called JPL-DE200 or JPL-DE450 codes [750].

After performing the transformations (12.48) the pulsar's rotational parameters can be determined from the rotational phase $\phi(t)$ given by a Taylor expansion

$$\Phi(t) = \phi(t_0) + \Omega(t-t_0) + \tfrac{1}{2}\dot{\Omega}(t-t_0)^2 + \tfrac{1}{6}\ddot{\Omega}(t-t_0)^3 + \cdots \tag{12.49}$$

with Ω the angular velocity of the pulsar. From these data, valuable information on the properties of pulsars themselves can be obtained, e.g., on the structure of the neutron star, the equations of state, or the pulsar evolution. Besides the investigation of the pulsars themselves, they can be used as tools to test fundamental theories and for other applications [751]. The fundamental physics tests include [752] relativistic precession, Einstein and Shapiro delay, gravitational waves, variation in G, Chandrasekhar mass, strong equivalence principle, Lorentz invariance and conservation laws. Hulse and Taylor have received the Nobel prize for their investigation of the 59 ms binary pulsar 1913+16 comprising a neutron star and a companion [5, 753]. These investigations allowed the most accurate tests of General Relativity. The periastron rotation, more than four orders of magnitude larger than in the case of the mercur perihelion, has been well confirmed. The observed reduction $\dot{P}/P \approx -3 \times 10^{-12}$ could be explained by the emission of gravitational waves. In the mean time, a number of more precise fundamental tests with increasing accuracy have been performed [752]. Measurement of the orbital period derivative allows the measurement of the temporal change of G, Newton's gravitational constant. Currently, for instance, $\dot{G}/G = (-22 \pm 775) \times 10^{-12}$ a^{-1} has been observed for the 4.57 ms binary pulsar PSR J1713+0747.

12.6.2.2 Pulsars as Frequency Standards

Pulsars have been referred to as "nature's most stable clocks" [5]. In fact, the pulsars 1937+21 and 1855+09 have been observed to have stabilities of $\dot{P} = 1.05 \times 10^{-19}$ s/s and $\dot{P} = 1.78 \times 10^{-20}$ s/s, respectively [754]. The stability is affected also by the measurement noise, which is often considered as white phase noise with a measurement uncertainty of the order of a microsecond.

To determine the stability, however, the deterministic but *a priori* unknown drift present in the timing data has to be removed in order not to be confused with long-term random fluctuations. In order to ignore fixed frequency drifts, Matsakis *et al.* [755] and Vernotte [756] have suggested and applied third-order differences leading to a variance referred to as "pulsar variance" or σ_z^2. The fractional instabilities of pulsars come down to a $\sigma_z \approx 10^{-15}$ regime for measurement times of a few years [5, 755].

There are, however, several effects that can change the rotational frequency of a pulsar. First, if the axis of the magnetic field is not collinear to the axis of rotation as depicted in Fig. 12.15 electromagnetic waves will carry away energy. Similarly, any deviation of the rotational symmetry of the mass distribution will lead to the emission of gravitational waves. In binary systems, a decrease in the orbital frequency has been observed that is consistent with the emission of gravitational waves according to the theory of general relativity. Both effects will reduce the frequency of the signal from the pulsars. As a result, older pulsars should become slower. On the other hand, pulsars in binary systems could increase their rotational frequency by an effect where matter is sucked by the pulsar from the companion star. In this case the sucked plasma is collected in an accretion disk in the orbital plane of the binary system. When this plasma eventually is trapped by the neutron star its angular momentum is transferred to the latter thereby increasing its rotational frequency. Furthermore, the steady decrease in the frequencies of some pulsars has been observed to be interrupted by a sudden increase. The interpretation of these so-called glitches is based on the assumption that neutron stars consist of a fluid interior and a solid brittle crust. The glitches are explained either by

cracking of the crust or by fluctuations of the angular momentum resulting from the quantised vortex properties in the neutron superfluid of the interior. As a result, pulsars with their built-in precise clocks are extremely helpful tools for astronomy but presently it seems safe to conclude that they will not be used for time keeping.

13 Technical and Scientific Applications

The high accuracy of time and frequency metrology based on the use of atomic clocks and frequency standards has challenged and inspired researchers and engineers to transfer this accuracy to the measurements of other physical quantities. Together with the widespread availability of easy-to-use and cost-effective reference frequencies, on virtually each level of accuracy, frequency-based measurement techniques are applied to various fields of technology. Prominent examples in the low-accuracy regime are the various sensors based on quartz where the eigenfrequency of a quartz resonator (Section 4.1) changes in a defined way under an external influence. Sensitive and accurate thermometers, pressure gauges and accelerometers of microbalances are constructed using this principle. The microbalance where the eigenfrequency changes with the mass deposited on the quartz crystal can be used, e.g., to sense quantitatively the adsorption of organic molecules in vapour or liquids. In this chapter, however, examples are presented for the measurements of dimensional, electric and magnetic quantities with highest accuracy together with applications in technology and fundamental physics.

13.1 Length and Length-related Quantities

13.1.1 Historical Review and Definition of the Length Unit

During the time of the French revolution, the need was felt to set up a general system of units [1] that could be derived from properties of the Earth. The unit of length, the metre, was chosen as the tenth million part of the quadrant of an Earth meridian. From geodetic length measurements the metre was determined and transferred to a length bar called the "Mètre des Archives". It was only in 1889 when the 1^{st} Conférence Générale des Poids et Mesures (CGPM; General Conference on Weights and Measures) defined the length unit "metre" as the length of the "International Prototype of the Metre" made of a platinum iridium alloy whose length could be related to that of the Mètre des Archives. Besides the definition of a unit, like the metre, one is interested in a procedure to "realise" this unit, i.e. to perform practical measurements based on this definition. The relative uncertainty to realise the metre derived from the International Prototype was about 10^{-7} (see Fig. 13.1) which was limited, e.g., by the quality of the edges of the grooves used as the length markers in the International Prototype.

Already by the beginning of the 20^{th} century Michelson and others had created "secondary standards of length" by determining the wavelengths of various emission lines, e.g.,

[1] A more detailed review can be found in [757].

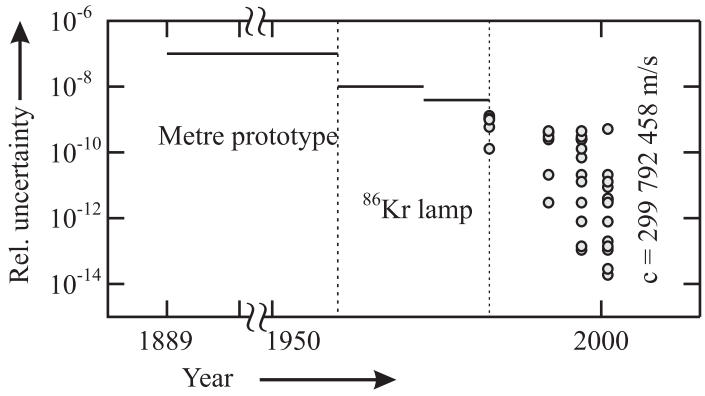


Figure 13.1: Progress in the accuracy of the realisation of the length unit in the International System of Units (SI) where, according to the definitions of the CGPM, the metre was defined by the International Prototype (1889 – 1960; [758]), by the Krypton wavelength (1960 – 1983; [759]) and by the speed of light using realisations by laser frequency standards (from 1983; [95, 370, 760–762]).

of Cadmium and Mercury, by comparison with the International Prototype thereby following an earlier suggestion of Maxwell.[2] In the fifties of the past century, Engelhard in the Physikalisch-Technische Bundesanstalt developed a special Krypton lamp whose orange wavelength at 605.78 nm was superior to any artefact standard with respect to stability and reproducibility. The lamp used the transition $5d_5 \rightarrow 2p_{10}$ (designation according to Paschen; see footnote 1 on page 257) of the ^{86}Kr isotope excited in a gas discharge. To reduce the Doppler width, the discharge was immersed in liquid nitrogen. In 1960 about six decades after Michelson's first attempts, the CGPM defined the metre via the wavelength of the Krypton lamp.

In the same year, however, the invention of the laser initiated the development of frequency-stabilised light sources that were superior to the Krypton lamp with respect to the output power, the coherence length or the reproducibility of their frequencies. It was clearly foreseen that through the years to come new and more stable lasers would continuously be developed that, in principle, could be used to realise the metre with increasing accuracy. To make full use of the best of these lasers for accurate length measurements, however, would require a frequent re-definition of the SI unit metre. To solve this problem, the 17$^{\text{th}}$ CGPM adopted in 1983 a new definition of the metre as

> the metre is the length of the path travelled by light in a vacuum during a time interval of 1/299 792 458 of a second.

This definition utilises the impressive accuracy that can be achieved by measurements of time with atomic frequency standards together with the value of a fundamental constant, the speed

[2] "If, then, we wish to obtain standards of length, time, and mass which shall be absolutely permanent, we must seek them not in the dimensions, or the motion, or the mass of our planet, but in the wavelength, the period of vibration, and the absolute mass of the imperishable and unalterable and perfectly similar molecules" quoted from Petley [763], page 15.

of light c. With the 1983 definition of the metre this value of c = 299 792 458 m/s is now fixed. Its value had been determined very accurately in the laboratory by a measurement of a distance and a frequency (see e.g [681]). The invariance of the speed of light is not only a postulate of Einstein's theory of relativity based on Michelson and Morley's experiment but has been shown experimentally to be valid with very low uncertainty [764].

To realise the metre according to the 1983 definition, the Comité International des Poids et Mesures (CIPM) adopted a recommendation that the metre be realised by one of the following methods:

a) by means of the length l of the path travelled in vacuum by a plane electromagnetic wave in a time t; this length is obtained from the measured time t, using the relation

$$l = c \times t \tag{13.1}$$

and the value of the speed of light in vacuum c = 299 792 458 m/s;

b) by means of the wavelength in vacuum λ of a plane electromagnetic wave of frequency f; this wavelength is obtained from the measured frequency f, using the relation $\lambda = c/f$ and the value of the speed of light in vacuum c = 299 792 458 m/s;

c) by means of one of the radiations from a list (see Table 13.1), whose stated wavelength in vacuum, or whose stated frequency, can be used with the uncertainty shown, provided that the given specifications and accepted good practice are followed and that in all cases any necessary corrections be applied to take account of actual conditions such as diffraction, gravitation, or imperfection in the vacuum [761].

The different methods a) to c) will be discussed in the following in more detail.

13.1.2 Length Measurement by the Time-of-flight Method

The time-of-flight method a) of Section 13.1.1 is particularly suited to the measurement of large distances. Examples of this approach include satellite navigation (Section 12.5) or distance measurements on astronomic scales where the measure of distance is the light year.

13.1.2.1 Lunar Ranging

As an example, consider the distance between Earth and Moon which has been measured regularly over more than thirty years by directing a pulsed laser beam towards the moon. The pulses are reflected back by reflectors placed on the surface of the moon during the American space missions Apollo 11, 14, and 15 and the Soviet mission Luna 21. The measured travelling time allows one to measure this distance with an uncertainty of a few centimetres [765, 766]. These data have been used as precise tests for gravitational theories. Gravitational forces can lead to a precession of gyrating tops. A particular example is the precession of a gyro freely falling in a gravitational field. The system, Earth and Moon, exhibits an angular momentum and, hence, can be considered as a gyro orbiting around the Sun. De Sitter has recognised the corresponding precession as early as 1916 as a consequence of General Relativity. The calculated de Sitter angle of precession of the lunar orbit is about 2″ per century. The theoretical value has been confirmed with an uncertainty of about 1 % from lunar ranging data and a comparison with a model of the orbits of Earth and Moon [767, 768].

13.1.2.2 Deep Space Network

The art of ranging based on precise timing and advanced clocks has culminated to paramount accuracy in deep space navigation. As an example, consider the Cassini project. In late 1997 the Cassini spacecraft lifted off heading towards a seven year's journey to Saturn with four gravity assisted flybys of Venus (twice), Earth and Jupiter. The orbiter Cassini is expected to deliver in 2004 ESA's probe Huygens in the Saturn system for a descent to Saturn's moon Titan which is known to have a substantial atmosphere. The winds will cause perturbations to Huygens's local horizontal velocity and, by Doppler tracking the Huygens entry probe, the zonal wind profile on Titan is expected to be measured [769]. Similar measurements of the deep zonal winds at Jupiter have been obtained from the changes in the speed of the Galileo probe during descent with wind speeds as high as 200 m/s [770]. For these purposes the Cassini spacecraft carries an ultrastable SC cut quartz crystal oscillator with 4.79 MHz with an Allan deviation of $\sigma(\tau = 1\,\mathrm{s}) = 2 \times 10^{-13}$. On board the Huygens probe there is a rubidium clock with $\sigma(\tau = 1\,\mathrm{s}) = 6 \times 10^{-11}$ [771].

The spacecraft telemetry is acquired by NASA's Deep Space Network with three different tracking stations about 120 degrees apart at Goldstone (California), Canberra (Australia) and Madrid (Spain). To track the spacecraft, a radio frequency signal is sent to the spacecraft which transmits a phase coherent signal back to Earth thereby allowing the ground stations to determine the Doppler shift and, hence, the velocity of the space vehicle. The ranging is performed by sending a pseudo-random code on the uplink signal which is transmitted back on the downlink. The correlation of the received code with the replica of the (uplink) ranging signal allows the ground stations to measure the round trip time in order to determine the range.

To appreciate the need for the very precise telemetry required for accurate manoeuvres consider the flyby at Venus. During these flybys kinetic energy is transferred to the spacecraft which needs to arrive at the destination with minimum lift-off fuel. The mission requires that the spacecraft passes venus at an altitude of (300 ± 25) km. Another challenge occurs when the spacecraft is to fly by Titan with a required accuracy of 10 km at a distance of about 1.5×10^9 kilometres from Earth.

13.1.3 Interferometric Distance Measurements

The time-of-flight method a) of Section 13.1.1 is not very accurate on the scale of dimensions relevant to everyday life. Consider a distance measurement of 1 m with a desired fractional uncertainty of 1×10^{-7} which corresponds to the realisation uncertainty that has been achieved already with the International Metre Prototype (Fig. 13.1). This moderate uncertainty would require a measurement of the time-of-flight of 3 ns with an uncertainty of 0.3 fs.

Hence, interferometric methods are applied for distance measurements on the laboratory scale where the distance to be measured is compared to the number of wavelengths of the particular radiation used in the interferometer. A laser whose frequency is stabilised to an atomic, molecular or ionic transition represents a light source whose frequency, and, hence its vacuum wavelength, is largely independent of ambient conditions. Provided that this wavelength is known with sufficient accuracy the small wavelength of visible radiation of $\lambda \approx 0.5\,\mu\mathrm{m}$ provides an accurate ruler. Recommendations b) and c) of Section 13.1.1 give methods showing how to arrive at the vacuum wavelength with low uncertainty for a particular source of radiation.

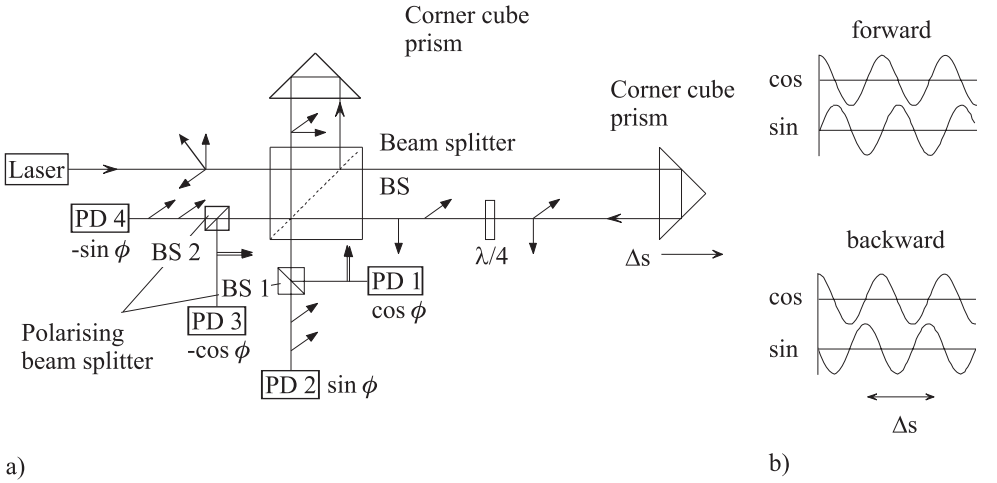


Figure 13.2: Michelson-type interferometer for displacement measurements Δs. a) Schematic set-up with the direction of polarisation indicated, PD: photodetectors. b) Photocurrent of the detectors PD 1 and PD 2.

For interferometric distance measurements often a two-beam interferometer according to Michelson (Fig. 13.2) is used. In the classical laser interferometer design, corner cube prisms are used instead of plane mirrors because this configuration is less sensitive to tilts during movement and it minimises the light back-reflected to the laser which can change the frequency of the laser source. Due to the corner cube prism, the measurement beam and reference beam are always parallel to the incoming beam. In the beam splitter the wave is split into two partial waves of field amplitude E_1 and E_2 that are recombined at the beam splitter after travelling along the different paths r_1 and r_2 yielding

$$\begin{aligned} E_1 &= E_{01}\,\cos(kr_1 - \omega t + \phi_1) \qquad \text{and} \\ E_2 &= E_{02}\,\cos(kr_2 - \omega t + \phi_2). \end{aligned} \tag{13.2}$$

The waves being reflected and transmitted by the beam splitter suffer from independent phase shifts which are taken into account by the additional phases ϕ_1 and ϕ_2. The time dependent power impinging on a detector is

$$\begin{aligned} I(t) &\propto (E_1 + E_2)^2 = {E_1}^2 + {E_2}^2 + 2E_1E_2 \qquad (13.3)\\ &= {E_1}^2 + {E_2}^2 + 2\,E_{01}E_{02}\,[\cos(kr_1 + \phi_1)\cos\omega t + \sin(kr_1 + \phi_1)\sin\omega t] \\ &\quad \times [\cos(kr_2 + \phi_2)\cos\omega t + \sin(kr_2 + \phi_2)\sin\omega t] \\ &= E_1^2 + E_2^2 + 2E_{01}E_{02}\,\big[\cos(kr_1 + \phi_1)\,\cos(kr_2 + \phi_2)\,\cos^2\omega t \\ &\quad + \sin(kr_1 + \phi_1)\,\cos(kr_2 + \phi_2)\,\sin\omega t\,\cos\,\omega t \\ &\quad + \cos(kr_1 + \phi_1)\,\sin(kr_2 + \phi_2)\,\sin\omega t\,\cos\,\omega t \\ &\quad + \sin(kr_1 + \phi_1)\,\sin(kr_2 + \phi_2)\,\sin^2\omega t\big] \end{aligned}$$

where we have made use of the identity

$$\cos(\alpha - \beta) = \cos\alpha\cos\beta + \sin\alpha\sin\beta \tag{13.4}$$

to separate the space-dependent $kr + \phi$ terms and the time-dependent ωt terms in the phases. Since photodetectors are not fast enough to respond to the high carrier frequency of light, the fast oscillating $\sin \omega t \cos \omega t$ terms integrate to zero when time averaged and one ends up with

$$\begin{aligned} I &= \langle I(t) \rangle \\ &\propto \tfrac{1}{2} {E_{01}}^2 + \tfrac{1}{2} {E_{02}}^2 \\ &\quad + 2(E_{01}E_{02})\tfrac{1}{2} \left[\cos(kr_1 + \phi_1)\cos(kr_2 + \phi_2) + \sin(kr_1 + \phi_1)\sin(kr_2 + \phi_2)\right] \\ &= \frac{{E_{01}}^2}{2} + \frac{{E_{02}}^2}{2} + E_{01}\ E_{02}\ \cos(kr_1 + \phi_1 - kr_2 - \phi_2) \\ &\propto I_1 + I_2 + 2\sqrt{I_1 I_2}\cos(kr_1 - kr_2 + \phi_1 - \phi_2) \end{aligned} \tag{13.5}$$

where we have used again (13.4). I_1 and I_2 are the time averaged irradiances of the partial waves.

If the retro-reflector is moved between two positions with a separation Δs and the reference mirror is kept at a fixed position the power varies periodically with the path difference. The visibility or contrast $V \equiv (I_2 - I_1)/(I_1 + I_2)$ is optimal ($V = 1$) if $I_1 = I_2$ holds. Counting the number $N(\Delta s)$ of the power maxima or minima allows one to determine the separation Δs by

$$2\Delta s = N(\Delta s)\ \lambda. \tag{13.6}$$

The periodicity of the signal allows an unambiguous determination of a displacement of the measurement reflector only within a period of $\lambda/4$. To measure larger displacements it is necessary to trace the signal continuously and to count the number of periods due to the zero crossings of the cosine term. With a single periodic signal alone, the direction of the reflector movement cannot be determined. Hence, often a second interference signal is generated with constant 90° phase shift. Fig. 13.2 shows a possible optical set-up of such a "homodyne interferometer". In a principal beam splitter BS a measurement beam and a reference beam are generated from the laser beam whose linear polarisation vector is decomposed into two orthogonal components of equal amplitudes. In the measurement beam a $\lambda/4$ wave plate with the principle axes rotated 45° relative to the polarisation of the beam produces a circular polarisation state, where the perpendicular polarisation states have a phase shift of 90°. In the reference beam these two perpendicular polarisation states are in phase. After superposition of measurement beam and reference beam in the principal beam splitter a secondary polarising beam splitter BS 1 is used to generate two 90° phase shifted interference signals detected by PD 1 and PD 2. The direction of movement of the corner cube in the measurement beam can now be determined at the zero crossing of the first interference signal by using the sign of the second signal (Fig. 13.2 b). There are two other photodetectors (PD 3 and PD 4) and a second secondary beam splitter (BS 2) behind the second output of the principal beam splitter generating two more interference signals. These additional 180° and 270° phase shifted signals are used to minimise the offsets of the interference signals. Simply counting the zero crossings of both interference signals gives a resolution of $\lambda/8$. For higher precision the resolution can be improved by an interpolation where the phase ϕ of the interference signal is determined from the irradiances I_0 of the 0° signal and I_{90} of the 90° signal with $\phi = \arctan(I_0/I_{90})$.

This method to determine the displacement of the retro-reflector by the optical path length suffers from the dependence of the wavelength $\lambda(n) = n\ \lambda_{\mathrm{vac}}$ on the index of refraction n

of the air. The index of refraction of air $n_{\mathrm{air}} \approx 1.00027$ depends largely on the temperature, pressure, humidity and on the content of other gases, in particular of CO_2. In practical interferometers, often the temperature, pressure and the humidity are measured and the deviation of the wavelength from the vacuum wavelength is determined by use of an empirically determined formula. The corrections determined originally by Edlén [772] have been refined later (see, e.g., [773] and references therein). Under ideal circumstances these corrections are good enough to allow one to achieve a relative uncertainty $\Delta n/n \approx 10^{-8}$. For large distances or in a harsh environment like an industrial plant, however, the achievable uncertainty may be considerably higher due to the less accurately known composition of the air.

In applications demanding higher accuracy the light path which changes during the movement is kept in vacuum. Interferometric measurements have been performed with fractional uncertainties as low as 2×10^{-11} [774] where a displacement of $\Delta s \approx 4\,\mathrm{m}$ was used. The necessary interpolation ($< 10^{-4}$) between the zeros of the interference fringes was performed by locking a tuneable laser to an interference minimum and measuring the beat frequency between this laser and a frequency standard. With such an impressive accuracy for distance and time measurements at hand also measurements of velocities and accelerations can be performed with unprecedented accuracy as shown by the example of gravimetry to be discussed in the following.

13.1.3.1 Gravimetry

Highly accurate absolute values of Earth's gravitational acceleration g are needed in a wide variety of scientific and technical applications. These include the determination of crustal deformation of the Earth, the measurement of changing sea levels, variations of ice masses in Greenland or Antarctica or the measurements of fundamental constants and units, e.g., for the Watt balance [775]. Gravimeters for precise measurements of g also find applications in geophysical explorations to locate crude oil or other natural resources.

The gravitational acceleration g experienced by a body of mass m at the position $\vec{r}_0$ as a result of Earth's gravitational force is

$$g(\vec{r}_0) = \frac{\vec{F}}{m} = \mathrm{G} \int \frac{dM_{\mathrm{Earth}}}{(\vec{r} - \vec{r}_0)^2} = \mathrm{G} \int \frac{\rho(\vec{r}_0) dV}{(\vec{r} - \vec{r}_0)^2}. \tag{13.7}$$

Here, $\mathrm{G} = 6.67 \times 10^{-11}$ m^3/(s^2 kg) is the Newtonian gravitational constant and the integration has to be performed over the locally varying density ρ of the Earth. Local variations of g witnessed on the surface of the Earth are thus hints to variations of the density that might be caused by bubbles of natural gas, deposits of ores or oil. As a result of the $1/r^2$ dependence of (13.7) nearby density variations contribute more to the local gravitational acceleration than do more distant ones. For precise measurements of the local gravitational acceleration, gravimeters are in current use (see Fig. 13.3).

A gravimeter of this type basically represents a Michelson interferometer with a vertical arm whose retro-reflector is a corner cube. g is determined by dropping the retro-reflector inside an evacuated vertical tube several times in a minute and monitoring at the same time the interference fringes as a function of time. The measured temporal sequence of the interference

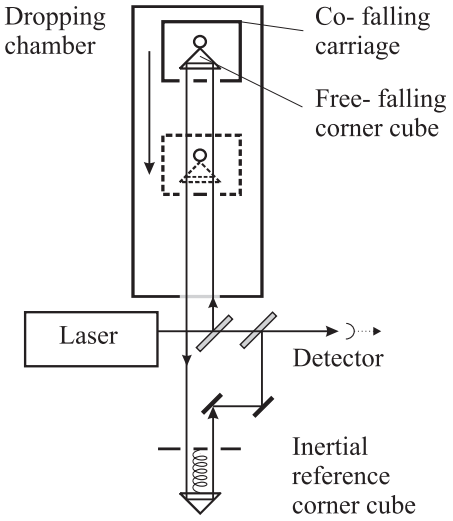


Figure 13.3: Schematic diagram of a gravimeter for the measurement of the local gravitational acceleration.

maxima allows one to determine the height $h(t)$ making use of the free-fall condition

$$h(t - t_0) = \tfrac{1}{2} g \, (t - t_0)^2 \tag{13.8}$$

to determine g. The two beams in the interferometer are displaced in order not to be retro-reflected into the laser. The corner cube falls drag-free inside a co-falling carriage in order to eliminate the influence of the friction of the residual gas in the apparatus on the falling corner cube which would reduce the value of the measured g.

The interferometer of the most advanced types use iodine stabilised He-Ne lasers whereas also two-mode stabilised He-Ne lasers have been used. The measured path difference in the Michelson interferometer critically depends on the inertial reference of the second corner cube. In the system described in [776] a two-stage spring-isolated platform carries the reference corner cube with the second stage employing a "super spring" system. The concept of the super spring [777] simulates a very long spring with a very small spring constant despite using a relatively short spring. This short spring can be considered as the end part of the long spring where the properties of the "missing part" are simulated by an active electronic stabilisation scheme. The relative uncertainty of determining g that can be obtained with such a device can be as low as 10^{-9} g [776, 778].

Novel types of gravimeters are based on the technique of a Ramsey-Bordé atom interferometer [779–781] (Section 6.6.2.3). These devices employ laser-cooled Cs atoms in an atomic fountain where three pairs of Raman pulses are used to split, redirect and recombine the atomic wave packets by a $\pi/2$, π, $\pi/2$ pulse, respectively. Under the influence of gravity the third pulse that is used to recombine the partial waves must have a phase shift (see also Section 9.4.4.2)

$$\Delta\Phi = g k_{\mathrm{eff}} T^2. \tag{13.9}$$

$k_{\text{eff}} = 2\pi/(\lambda_1 + \lambda_2)$ hold for counterpropagating Raman beams. A comparison with a gravimeter employing a falling corner cube gave agreement within $(7 \pm 7) \times 10^{-9}$ [781]. At this level of accuracy the measured acceleration has to be corrected for the temporal variation resulting from the tides caused by the moon and the sun (see Fig. 13.4). Atom interferometers have also been used to measure gravity gradients [782].

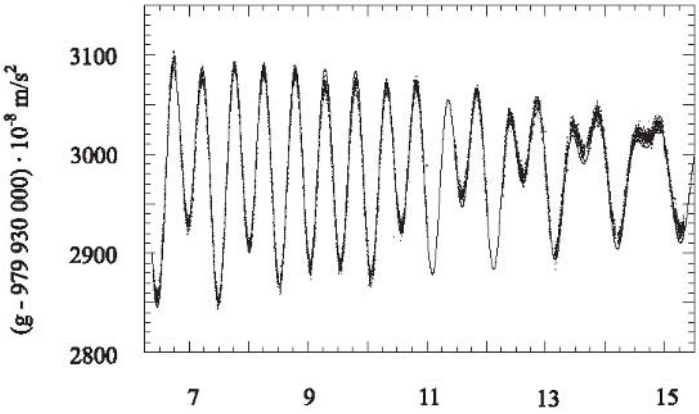


Figure 13.4: The variation of Earth's gravitational acceleration g measured with an atom interferometer (dots) [779–781] and a theoretical model (line) of the gravity tides at the measurement site at Stanford University, California. Courtesy of A. Peters.

13.1.4 Mise en Pratique of the Definition of the Metre

The full accuracy possible with interferometric methods can be achieved only if the wavelength of the radiation used in the interferometer is known with the same accuracy. Optical frequency standards with well known vacuum wavelengths at various spectral regimes have been developed (see Section 9). Method b) of Section 13.1.1 allows one to determine the wavelength of any source of monochromatic electromagnetic radiation in conformity to the definition of the metre with an accuracy that ultimately depends on the accuracy of the frequency measurement. The measurement of optical frequencies can be performed now by direct comparison to the frequency of the primary standard of time and frequency (see Section 11). Despite the continuous development of optical frequency measurements culminating in reduced complexity like the femtosecond comb, these methods are still the domain of a small number of well equipped laboratories. Hence, the CIPM recommended a number of validated and selected radiations of such stabilised lasers as references for the realisation of the metre and for precision spectroscopy (see Table 13.1 [370]). This list (method c) of Section 13.1.1 of the realisation of the metre, is referred to as the "Mise en Pratique of the Definition of the Metre". It contains a number of approved wavelength standards where the frequency measurement has been performed and the accuracy has been evaluated. After its first edition in 1983 [761] this list was updated in 1992 [762], 1997 [95], 2001 [370] and

Table 13.1: Radiations recommended by the CIPM for the realisation of the metre [370].

Quantum absorber	Transition	Wavelength (nm)	Rel. stand. uncertainty
$^{115}In^+$	$5s^2\ ^1S_0 - 5s5p\ ^3P_0$	236.540 853 549 75	3.6×10^{-13}
^1H	1S – 2S	243.134 624 626 04	2.0×10^{-13}
$^{199}Hg^+$	$5d^{10}6s\ ^2S_{1/2}\ (F = 0)$ – $5d^96s^2\ ^2D_{5/2}\ (F = 2)\ \Delta m_F = 0$	281.568 867 591 969	1.9×10^{-14}
$^{171}Yb^+$	$6s\ ^2S_{1/2}\ (F = 0)$ – $5d\ ^2D_{3/2}\ (F = 2)$	435.517 610 739 69	2.9×10^{-14}
$^{171}Yb^+$	$^2S_{1/2}\ (F = 0, m_F = 0)$ – $^2F_{7/2}\ (F = 3, m_F = 0)$	466.878 090 061	4.0×10^{-12}
$^{127}I_2$	R(56) 32–0, a_{10}	532.245 036 104	8.9×10^{-12}
$^{127}I_2$	R(127) 11–5, a_{16} or (f)	632.991 212 58	2.1×10^{-11}
^{40}Ca	$^1S_0 - {}^3P_1;\ \Delta m_J = 0$	657.459 439 291 67	1.1×10^{-13}
$^{88}Sr^+$	$5\ ^2S_{1/2} - 4\ ^2D_{5/2}$	674.025 590 863 1	7.9×10^{-13}
^{85}Rb	$5S_{1/2}\ (F_g = 3) - 5D_{5/2}\ (F_e = 5)$	778.105 421 23	1.3×10^{-11}
$^{13}C_2H_2$	P(16) $(\nu_1 + \nu_3)$	1 542.383 712	5.2×10^{-10}
CH_4	$F_2^{(2)}$ comp., P(7) ν_3, (7–6) trans.	3 392.231 397 327	3.0×10^{-12}
OsO_4	coinciding with $^{12}C^{16}O_2$ laser line R(10) $(00^01) - (10^00)$	10 318.436 884 460	1.4×10^{-13}

contains some of the lasers stabilised to atomic (see Table 5.2), ionic (see Table 10.2)) or molecular transitions (see Table 5.4 and Table 9.1). The wavelengths of these standards range from the near ultraviolet (243 nm) to the infrared (10.3 µm).

13.1.4.1 Frequency Standards for Optical Telecommunication

Optical reference frequencies with less demanding accuracy such as the acetylene transition at 1.54 µm given in Table 13.1 find important applications also in optical telecommunications where Wavelength Division Multiplexing (WDM) in optical fibres provides many wavelength channels. In the important wavelength region of the erbium-doped amplifier from about 1.540 µm to 1.56 µm the International Telecommunication Union (ITU) recommends the use of a frequency grid of 50 GHz or 100 GHz channel spacing with the possibility of narrower spacings implemented in the near future.

Reference cells filled with suitable absorbers are available as transfer standards, e.g., from the national standards institutes. Light from broadband sources, like light emitting diodes, passing these cells exhibits the absorption dips characteristic for the molecular spectrum.

These lines can be used to calibrate optical spectrum analysers or wavelength meters and to characterise tuneable lasers and channel wavelengths with an uncertainty of a few tenths of a picometre. The $\nu_1 + \nu_3$ ro-vibrational transitions in acetylene $^{12}C_2H_2$ (Fig. 13.5) and $^{13}C_2H_2$ have each more than 50 strong absorption lines in the region between about 1510 nm and 1550 nm (Fig. 5.7, Fig. 13.5). The spectrum of the $2\nu_3$ overtone transitions in hydrogen

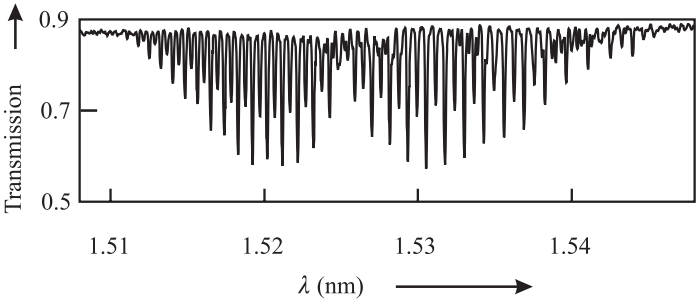


Figure 13.5: Ro-vibrational absorption lines of acetylene used as frequency references in the telecommunication bands. Courtesy of F. Bertinetto.

cyanide $H^{13}C^{14}N$ span the region between approximately 1525 nm and 1565 nm [473]. In the wavelength division multiplexing L-band between about 1565 nm and 1625 nm the spectrum of hydrogen iodide (HI) and carbon monoxide ($^{12}C^{16}O$) cover the high frequency range, whereas $^{13}C^{16}O$ has about 35 lines between 1595 nm and 1628 nm. Semiconductor distributed feedback (DFB) lasers, extended cavity diode lasers and DFB fibre lasers have been used in this region. The latter, stabilised to CO near 1.58 µm, had a frequency stability of a few megahertz over some minutes [783]. CO in general has weaker lines as compared to HI [473]. For higher accuracy the subharmonic of the Rb two-photon transition (Section 9.4.3) or the He-Ne laser operating at 1523 nm has been used.

Table 13.2: Optical frequency standards in the telecommunication bands. Other candidates can be found, e.g. in [47].

Band	Range (nm)	Absorber	Reference
O-band	1260 to 1360		[784]
E-band	1360 to 1460		
S-band	1460 to 1530	$^{12}C_2H_2$	
C-band	1530 to 1565	HCN	[473]
L-band	1565 to 1625	HI, $^{12}C^{16}O$	[473]
U-band	1625 to 1675		

13.2 Voltage Standards

The measurement of a voltage can be traced back to a frequency measurement via the Josephson effect. In 1962 Brian D. Josephson described effects [785] which can occur in a "Josephson junction" where two superconducting layers are separated by an isolating layer with a thickness of a few nanometres. The superconducting state on either side of the barrier is characterised by "Cooper pairs" comprised of two electrons with opposite spins and $\vec{k}$ vectors and is described by a single macroscopic wave function with a macroscopic phase for all Cooper pairs. If the barrier is thin enough Cooper pairs can tunnel through it and the two wave functions on either side of the barrier are weakly coupled. This coupling of the two quantum mechanical states leads to a current through the barrier that sinusoidally depends on the phase difference ϕ of the two states [786] when the Josephson junction is connected to a current source. Furthermore, the evolution of the phase difference is related to a voltage U applied between the superconductors which leads to an alternating current of frequency

$$f = \frac{1}{2\pi}\frac{d\phi}{dt} = \frac{2e}{h}U \equiv K_J U. \tag{13.10}$$

The Josephson constant K_J is the reciprocal value of the elementary flux quantum in superconductors

$$\Phi_0 = h/2e \tag{13.11}$$

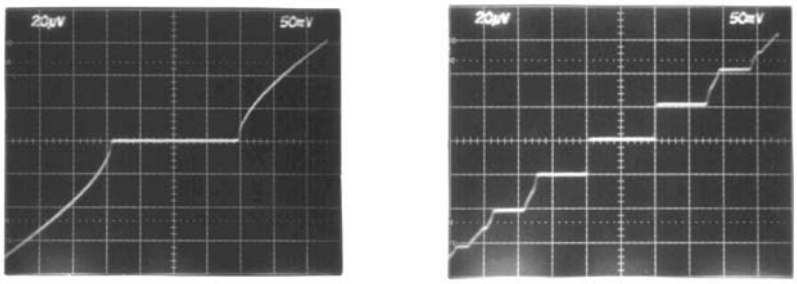
and can be calculated from the Planck constant h and the elementary charge e. Hence, the Josephson effect can be thought of as a voltage controlled oscillator that links a voltage to a frequency via fundamental constants. If this oscillator is locked to an external frequency f_e the non-linear dc characteristic of the Josephson junction can lead to higher harmonics of the oscillation frequency and, hence, to steps of constant voltage at the voltages

$$U_n = n\frac{h}{2e}f \tag{13.12}$$

where n = 1,2,$\cdots$.

In real Josephson devices, the ideal Josephson junction is shunted by a capacitance, the high frequency source and an ohmic resistance, which leads to a damping of the oscillator. The voltage steps (Fig. 13.6) are used as reproducible reference frequencies. The low voltage of the steps in Fig. 13.6 b) has been increased to 1 V and 10 V by connecting up 20 000 Josephson junctions in series [787, 788] which have shown over several years that a reproducibility of a 10 V standard of as good as 5×10^{-11} can be reached [788]. Series junctions as well as measurement systems are commercially available. An interlaboratory comparison between sixteen national, industrial and military laboratories with four travelling Zener diode standards have shown relative differences at 10 V for most of the laboratories of less than 2×10^{-8} [789].

It has to be emphasised, however, that this is not the accuracy that can be attributed to a voltage measurement in the international system of units (SI) since the values of the fundamental constants are not known with sufficient accuracy. Hence, in 1990 a value of $K_{J-90} = 483\,597.9$ GHz/V was recommended by the Consultative Committee of Electricity (CCE) for maintaining the unit volt. This agreed value allows one to reproduce the voltage much better than to know its value in SI units where according to [790, 791], $h/2e =$



a) b)

Figure 13.6: a) dc voltage-to-current characteristic of a highly damped Josephson junction from Nb-PdAu-Nb without microwave. b) Voltage-to-current characteristic with an applied microwave frequency of 10 GHz. Courtesy of J. Niemeyer.

$483\,597\,879(41) \times 10^9$ Hz/V with a relative standard uncertainty of 8.5×10^{-8}. Cryogenic voltage standards have been developed that rely on standard frequencies transmitted either by DCF77 (see Section 12.4) or GPS (see Section 12.5).

13.3 Measurement of Currents

Extending Maxwell's suggestion (see footnote 2 of this chapter) to electric units, there is considerable interest in reproducing the ampere as the SI unit of electric current also from quantum standards. At first sight the route seems to be straightforward since the periodic transportation of a defined number N of elementary charges e with a well defined frequency f, immediately leads to a current

$$I = Nef. \tag{13.13}$$

The difficulty rests with the smallness of the current associated with the elementary charge and with the requirement to determine the number unambiguously in a real device. Hence several possibilities are explored to relate the measurement of a current to a frequency using (13.13).

13.3.1 Electrons in a Storage Ring

It has been proposed to use the electrons that circulate with a well defined frequency in an electron storage ring [792] as a standard of current. Due to their high kinetic energy of the order of several hundred megaelectronvolts to several gigaelectronvolts in a storage ring, the relativistic electrons travel with a velocity very close to the speed of light c. In the magnets used to bend the trajectories of the electrons to a closed orbit the electrons are radially accelerated thereby emitting synchrotron radiation into a small cone along the trajectory. In a storage ring the loss of kinetic energy due to the emission of synchrotron radiation is compensated for by an accelerating electromagnetic field in a microwave cavity inserted into the

ring. The radio frequency f_{rf} of the microwave field is synchronised with the frequency of revolution f_e of the electrons. In general, $f_{\mathrm{rf}} = nf_e$ holds and the frequency of revolution is an exact sub-harmonic $1/n$ of the radio frequency and, thus, can be measured precisely. For a few thousands orbiting electrons, the actual number N can be determined directly from the irradiated power of the synchrotron radiation measured by a photodetector (Fig. 13.7).

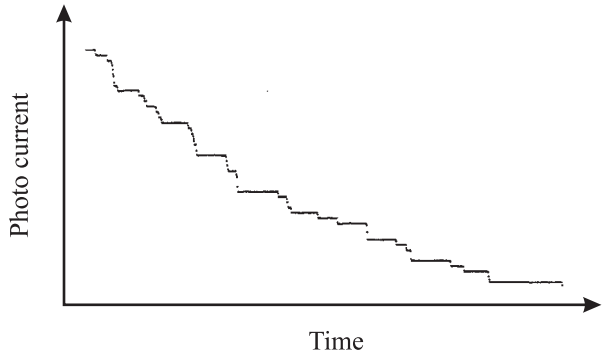


Figure 13.7: The steps in the measured synchrotron radiative power can be used to identify the exact number of electrons circulating in the storage ring [792].

When a single electron is removed from a number (e.g. one hundred) of stored electrons, the measured photo current of a detector that monitors the synchrotron radiation power is reduced by 1 %, provided the detector is sufficiently linear and the noise level is not too high. The photo current of Fig. 13.7 [792] shows the steps corresponding to the last forty electrons that were successively removed from the electron beam down to a single electron. The current associated with a single electron was about $I_e = ec/L \approx 0.77 \times 10^{-12}$ A for a perimeter of the storage ring of $L = 62.4$ m. The accuracy attainable with this method depends on the accuracy that can be obtained when the electron current in the storage ring is compared to an external current to be calibrated. Such a comparison can be done, e.g., by means of a cryogenic current comparator where the difference of the magnetic field of the two currents is measured with a SQUID magnetometer (see Section 13.4.1). A current resolution as low as 6 fA/$\sqrt{\mathrm{Hz}}$ to 65 fA/$\sqrt{\mathrm{Hz}}$ depending on the frequency range has been measured [793]. An even smaller resolution of 0.1 fA/$\sqrt{\mathrm{Hz}}$ is expected in the white noise region for an optimised device [794]. Such a device could allow to compare a current with the current of 10 nA carried by 1300 circulating electrons in a dedicated storage ring of 6 m perimeter in a measuring time of 1 s to 10^{-8}.

13.3.2 Single Electron Devices

Single electron devices are based on the long-range Coulomb interaction between singly charged carriers that contribute to the electronic transport in circuits based on small conducting areas. The effects of charging small neutral bodies with a few additional quanta of charge has a long history dating back to the well known experiment of Millikan at the beginning of the twentieth century. Modern single electron devices use isolated islands charged with a few

elementary charges at the most. The charge flow from a particular island is accomplished by sequential tunnelling of electrons through the barriers between the islands. The probability that an electron tunnels from one island to the next one depends on its kinetic energy determined by the temperature T of the device and on the energy the electron gains in this process. The transport of an electron to an island is equivalent to charging a capacitor. For a small capacitance C of the island the Coulomb energy $E_c = eU$ can be much higher compared to the thermal energy

$$E_c = \frac{e^2}{2C} \gg kT. \tag{13.14}$$

Under this condition the tunnelling is blocked (Coulomb blockade) as long as the voltage U is not compensated by an externally applied voltage U_{ext}. As a consequence, the current in single-electron devices results from the tunnelling of single electrons and can be controlled effectively by applying external voltages. They allow to develop a current source, where the movement of single charges can be gated by an accurate radio frequency. The basic circuit referred to as a single electron tunnelling pump may consist of two islands and a gate electrode for each island (Fig. 13.8). Since the transportation of single electrons depends on the instan-

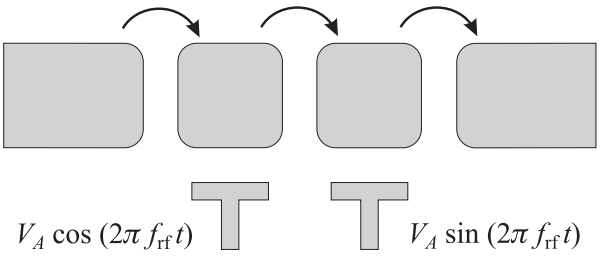


Figure 13.8: Single electron tunnelling pump based on three tunnelling sections.

taneous values of the two gate voltages, a periodic temporally shifted modulation of both gate voltages leads to a current $I = ef_{\mathrm{rf}}$ where a single electron is transported per period of the applied radio frequency. The radio frequency of a few megahertz leads to a low current of a few picoamperes. Single electron tunnelling devices have been utilised also to set up a capacitance standard based on counting electrons with a relative uncertainty of 3×10^{-7} [795, 796]. In this work a single electron tunnelling element comprising seven tunnelling contacts was employed in order to reduce the so-called co-tunnelling. This effect results from higher-order quantum mechanical tunnelling processes through the complete chain that lead to erroneous counting of electrons.

For a metrological dc current source the current in the picoampere regime is too low to be compared with the required accuracy to higher currents using state-of-the art SQUID-based current comparators. It has been shown, however, that the required sensitivity can be achieved in the nanoampere range [794]. Another method that may be suited to reaching this level makes use of surface acoustic waves in a semiconducting material where single electrons are transported in travelling potential minima [797, 798]. As a consequence of the piezoelectric effect in GaAs, a propagating modulation of the electrostatic potential can be achieved in a

two-dimensional electron gas close to the surface. The higher frequencies of several gigahertz and the associated higher currents may lead to an alternative standard of electric current [799] based on fundamental constants and frequency standards only.

13.4 Measurements of Magnetic Fields

Several methods are used to measure magnetic fields very accurately via frequency measurements.

13.4.1 SQUID Magnetometer

A Superconducting QUantum Interferometric Device (SQUID) comprises a superconducting ring where the flux is quantised according to

$$\Phi = n\Phi_0 = n\frac{h}{2e} \tag{13.15}$$

and may be realised in two different configurations [800]. In the dc SQUID two Josephson junctions divide the ring into two halves (Fig. 13.9). The dc SQUID can be operated with

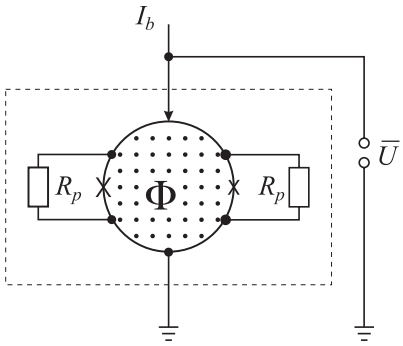


Figure 13.9: The voltage $\bar{U}$ measured using a SQUID driven by a bias current I_b is a measure of the magnetic flux Φ enclosed by the two halves of a superconducting ring separated by two Josephson junctions $(\times)$ shunted with parallel resistance R_p. The dashed box indicates the low-temperature environment necessary to sustain the superconducting state.

a dc bias current. In the rf-SQUID the superconducting ring contains a single Josephson junction which is read out by an inductively coupled resonant circuit. Quantisation of the magnetic flux in the superconducting loop and the Josephson effect lead to a flux-to-voltage characteristic with flux quantum periodicity [800]. In order to obtain a large dynamic range, a null-detector scheme with flux quanta counting can be used. dc SQUID magnetometers can achieve a flux noise depending on the bandwidth of $10^{-6}\Phi_0/\sqrt{\text{Hz}}$ [800, 801]. The SQUID is by far the most sensitive sensor of magnetic flux. Magnetic field resolutions in the region of $50\,\text{fT}/\sqrt{\text{Hz}}$ are possible with an area of the pickup coil of only a few square millimetres.

SQUIDs require cooling to low temperatures, at least to the temperature where high temperature superconductors such as YBaCuO can be employed that are cooled by liquid nitrogen or by cryocoolers. SQUIDs can measure low magnetic fields (about 20 fT) and low magnetic field gradients (<1 pT/cm).

13.4.2 Alkali Magnetometers

Compact laser magnetometers can be built which use the magnetic field-dependent splitting and shift of energy levels in the alkaline metals caesium (Fig. 7.1), rubidium (Fig. 8.8) or potassium. Several schemes have been used based on zero-field level crossing resonances [802] or non-linear magneto-optical (Faraday) rotation [803]. Sensitive magnetometers can also be built based on magnetic field sensitive dark resonances. The dark resonances can occur in a so-called Λ system in an alkaline atom such as Rb or Cs when two coherent laser fields couple the two closely spaced ground states to a common third level (see Fig. 5.12 a). The schematic shown in Fig. 8.11 allows one to set up a compact magnetometer operated at room temperature and requiring low power with a sensitivity of a few picotesla in one second integration time [804].

13.4.3 Nuclear Magnetic Resonance

Nuclear Magnetic Resonance (NMR) methods use the nuclear spin as a probe to sensitively detect structure and dynamics of the electronic cloud and that of other nuclei in the neighbourhood of the probe spin. From its invention [805, 806] the technique has contributed invaluably to molecular spectroscopy, medical diagnostics and other fields. Its precision results from the fact that the probe responds with frequency shifts that can be read out with the associated accuracy. The interaction energy

$$E_{\mathrm{mag}} = -\vec{\mu}_I \vec{B}_0 = g_I \mu_N B_0 m_I \tag{13.16}$$

between an external magnetic field B_0 and the magnetic moment of the nucleus

$$\vec{\mu}_I = \frac{g_I \mu_N}{\hbar} \vec{I} \tag{13.17}$$

depends on the angular momentum (spin) $\vec{I}$ of the nucleus and its orientation described by the magnetic quantum number $m_i = I, I-1, \cdots -I$. In contrast to the Landé g_J factor of the electron shell, the nuclear g_I factor cannot be calculated from other quantum numbers but rather has to be determined experimentally for each nucleus with $I \neq 0$. It furthermore can be negative or positive. The interaction energy (13.16) is much smaller than the interaction energy (5.8) of the magnetic moment of the electron shell as a result of the nuclear magneton $\mu_N = 5.051 \times 10^{-27}$ A m^2 being m_e/m_p times smaller than Bohr's magneton μ_B.

Consequently, the frequency $\nu = \Delta E_{\mathrm{mag}}/h$ required to excite transitions between the Zeeman split magnetic sub-states is much smaller than in the case of the electronic ground state. For hydrogen nuclei ($g_I = 5.585\,6912$; $\Delta m_I = 1$) the NMR frequency is about 42.576 MHz/T which has to be contrasted with the ground state Zeeman splitting of hydrogen in the gigahertz regime (Fig. 5.22).

The sample is brought into a magnetic field with $B_0 \lesssim 20\,\mathrm{T}$. High fields are desirable to achieve large oscillation frequencies and, hence, good resolution. A high-frequency coil produces a field pulse that induces transitions between the Zeeman split energy levels of the nuclei contained in the sample. The absorbed signal is due to the difference between the absorption and the stimulated emission between the two energetically separated states. Due to the low fractional population difference between the two states

$$\frac{N_1 - N_2}{N_1 + N_2} = \frac{1 - \exp\left(-g_I \mu_N B_0 / k_B T\right)}{1 + \exp\left(-g_I \mu_N B_0 / k_B T\right)} \approx \frac{g_I \mu_N B_0}{2 k_B T} \tag{13.18}$$

only a very small fraction of the total number of spins of the sample contributes to the polarisation of the sample. Owing to the long wavelength which in general is larger than the extension of the sample under investigation, a short high-frequency pulse excites all dipoles in phase which results in a macroscopic magnetic polarisation of the sample. This magnetisation decays by two different mechanisms. The component along B_0 (longitudinal component) depends on the population difference in the energetically different states and decays, e.g., by thermal fluctuations with a time constant that can be as long as $10^{-4} \lesssim T_1 \lesssim 10\,\mathrm{s}$ in liquid samples and $10^{-2} \lesssim T_1 \lesssim 1000\,\mathrm{s}$ in solid samples. The longitudinal relaxation, hence, can be used to derive information about the binding forces between the atoms in the lattice of a crystalline sample. The Larmor precession of the individual magnetic moments around the magnetic field B_0 corresponds to sinusoidal fields in the transverse directions. Local fluctuations of the magnetic field that can also be induced by the sample itself lead to different precession frequencies of the individual magnetic dipoles. Hence, after some time referred to as the transverse relaxation time T_2 the transverse components are randomly phased. The transverse relaxation can be used to determine the mean diffusion of the atoms.

13.4.3.1 Nuclear Magnetic Resonance Magnetometer

Methods based on the nuclear magnetic resonance (NMR) are employed for high precision measurements of static magnetic fields [807]. In NMR magnetometry a sample of a few cubic centimetres containing nuclei with spin different from zero (commonly hydrogen or deuterium) is placed in the static magnetic field B_0 to be measured. The frequency necessary to excite resonance transitions is directly proportional to the magnetic field B_0 (see (13.16)) and B_0 is then determined by the resonance frequency of the absorption line.

In the field range between $0.01\,\mathrm{T} \lesssim B_0 \lesssim 20\,\mathrm{T}$ commercial NMR magnetometers show fractional inaccuracies of the order of 10^{-6} as a result of the narrow absorption line. Furthermore, the temperature dependence of the resonance frequency is extremely weak being below about 10^{-6} from -20 ° C to +70 ° C. High field probes often use deuterated samples (^{2}H; $I = 1; g_I = 0.857\,437\,6$). On the other hand, these magnetometers are not very well suited to measure magnetic fields with relative inhomogeneities $\Delta B_0/(B_0 \Delta x)$ much larger than $10^{-4}\,\mathrm{cm}^{-1}$.

13.5 Links to Other Units in the International System of Units

Motivated by the success of the current definition of the metre which allows one to relate length measurements to a measurement of time or frequency via a fixed value of the speed of light in vacuum, there are a number of proposals to follow the same route for other basic units in the International System of Units (SI). The kilogram is currently still defined as the mass of the International Prototype which is kept at the BIPM in Paris. The prototype, as the last material artefact for the realisation of a basic unit, has prompted several proposals for a frequency based definition. Some of them are based on the equations $E = h\nu$ and $E = mc^2$ therefore relating the mass of a particle $m = h/c^2 \times \nu$ to a frequency via Planck's constant h and the speed of light c. Since c is already fixed in the present definition of the SI, attributing a fixed value to Planck's constant would also allow one to measure the mass of a microscopic particle via a frequency measurement.

Wignall has proposed an absolute atomic definition of the mass where the mass of a particle is defined by its de Broglie (angular) frequency $mc^2/\hbar$ [808]. The mass could then be determined by measuring the reduced de Broglie wavelength $\lambda/(2\pi) = \hbar/(m\gamma v)$ of a beam of monoenergetic particles with known speed v where $\gamma = (1 - v^2/c^2)^{-1/2}$.

One route to a new definition of the kilogram as a mass standard based on fundamental constants is via the Avogadro constant N_A. The work involves an accurate measurement of the mass and volume of a crystalline silicon sphere together with the lattice constant and other material properties of silicon. Some of the measurements, e.g., that of the lattice constants and the volume can be related to a wavelength and hence to a frequency measurement. The achieved relative uncertainty of higher than 10^{-7} [809, 810] has to be reduced by at least an order of magnitude to become competitive with the uncertainty that is routinely achieved with the current definition.

Taylor and Mohr [811] have suggested to employ a moving-coil watt balance [775] that has been used recently to determine the Planck constant [812] and proposed a definition as "The kilogram is the mass of a body at rest whose equivalent energy equals the energy of a collection of photons whose frequencies sum to 135639274×10^{42} Hz". In this way the unit of mass could be related directly to a frequency without making any statement about the realisation to be used.

13.6 Measurement of Fundamental Constants

The need to measure the fundamental physical constants with increasing accuracy is driven by different reasons. First, these constants can be used to realise units that no longer depend on the environment, on local conditions or on material artefacts. The use of an agreed value of the Josephson constant (see 13.2) gives an example of this kind allowing one to use the available high accuracy in applied metrology for the benefit of industry and trade. Second, these constants often show up in different branches of natural sciences with their specific theories. The accurate determination of the relevant physical constants in the different subfields leads to a test of the consistency of these theories and to their limitations. An example of this kind is given by discussing the determination of the fine structure constant (see Section 13.6.2).

13.6.1 Rydberg Constant

The Rydberg constant (5.6) defines the energy levels in an atom and is furthermore related to other fundamental constants $m_e, e, \hbar$ and c. The Rydberg constant has been measured preferably in hydrogen since it is the most simple atom whose energy levels can be computed with the highest accuracy and which has suitable transitions accessible to highest resolution laser spectroscopy. In the hydrogen atom the finite mass of the nucleus leads to the largest deviation of the reduced mass (see (5.7)) from the rest mass m_e of the electron in all atoms. However, the ratio m_e/m_p of the masses of the electron and the proton can be measured with high accuracy in ion traps (Section 10.4.1) where a recent measurement [650] gives $m_p/m_e = 1836,152\,664\,6(58)$. The simplicity of (5.4) allows one to determine the Rydberg constant R_∞ from a measurement of the transition frequency

$$\nu_{m,n} = \frac{E_m - E_n}{h} = c\, Z^2\, R_\infty \left(\frac{1}{n^2} - \frac{1}{m^2}\right) \tag{13.19}$$

between two levels with the principal quantum numbers n and m, to a relative uncertainty of about $\Delta R_\infty/R_\infty \approx 10^{-5}$. To obtain a more accurate value various theoretical corrections have to be applied. There are relativistic effects, interaction between the electron spin and the nuclear spin, contributions from quantum electrodynamics and corrections resulting from the finite size of the nucleus. The quantum electrodynamic corrections lead to the Lamb shift between the S level and the P level thereby lifting the degeneracy of the orbital angular momentum in the case of identical principal quantum number and identical total angular momentum. The extended charge distribution of the nucleus leads to an additional shift of the S level since the S electron spends some time at the position of the nucleus.

Measurements of the Rydberg constant have been performed over the years with increasing accuracy in the groups of T. W. Hänsch (Garching) [93,813] and F. Biraben (Paris) [103,814]. Hänsch's group measured the two-photon transitions 1S – 2S and 2S – 4S, the Paris group 2S – 8D, 2S – 12D. Earlier measurements have determined the transition frequencies in the hydrogen atom by comparison with the difference frequency of two optical frequency standards, i.e. the methane stabilised He–Ne laser and the iodine stabilised He–Ne laser. Recently, a direct comparison with the frequency of the Cs atomic clock has been obtained [93]. These measurements have been used with a high weight to obtain a value of R_∞ = 10 973 731.568 525 (73) m^{-1} for the 2002 CODATA evaluation of fundamental constants [791]. The relative uncertainty of 6.6×10^{-12} represents one of the most accurate measurements of a fundamental constant. Similar measurements have determined the Lamb shift of the 1S ground state in hydrogen and deuterium ([102, 394]). In order to perform a meaningful comparison between measurement and theory, it is necessary to include quantum electrodynamic corrections and the nuclear charge distribution. Since the experimentally achieved accuracy is currently much higher compared to the corrections, these measurements represent a means of testing the validity of the quantum electrodynamics or the nuclear charge distribution, depending on which contribution has the highest uncertainty [815].

13.6.2 Determinations of the Fine Structure Constant

The fine structure constant α represents one of the most fundamental constants of nature since it scales the electromagnetic interaction. Its value [3] can be determined in a variety of independent experiments conducted in different branches of physics via the von Klitzing effect (quantum Hall effect), the ac Josephson effect, the $g-2$ value of the electron, the de Broglie wavelength of the neutron [816] or atom interferometry [817]. The relative uncertainties achieved so far range from about 4.2×10^{-9} for the $g-2$ experiment (Section 10.4.2) including a QED estimation [653] to about an order of magnitude higher. All methods essentially rely on frequency measurements as will be pointed out for the two latter experiments. The fine structure constant can be related to the determination of the Rydberg constant R_∞ described in Section 13.6.1 and the measurement of h/m_e as follows

$$\alpha^2 = \frac{2R_\infty}{c}\frac{h}{m_e}. \tag{13.20}$$

The determination of the quantity h/m_e can be related to the determination of any other ratio $h/m = h/m_e \times m_e/m$ since, in general, mass ratios between microscopic particles again can be determined via frequency measurements in ion traps (Section 10.4). In the neutron experiment [816] $h/M_n = \lambda_n v$ has been determined using the de Broglie wavelength $\lambda_n \approx 0.25\,\text{nm}$ of a neutron wave packet, Bragg reflected from a silicon crystal, and the neutron's velocity v_n. The interplanar spacing a of the atomic planes of the silicon crystal needed to derive $\lambda_n = 2a\sin\theta$ (Bragg's law) has been determined interferometrically by reference to a laser wavelength standard (Section 9). The measurement of the velocity v_n has also been referred to an interferometric measurement of the flight path and the time of flight. The latter was determined by a periodic modulation of the polarisation of the neutron beam before and a detection of this modulation after travelling along the path.

An atom interferometric measurement of the fine structure constant with Caesium atoms is performed in the group of S. Chu (Stanford University) [817–819] where h/m_{Cs} is measured which is related to the fine structure constant by

$$\alpha^2 = \frac{2R_\infty}{c}\frac{h}{m_{\text{Cs}}}\frac{m_p}{m_e}\frac{m_{\text{Cs}}}{m_p} = 2R_\infty\frac{c\Delta\nu_{\text{rec}}}{\nu_{\text{Cs}}}\frac{h}{m_{\text{Cs}}}\frac{m_p}{m_e}\frac{m_{\text{Cs}}}{m_p}. \tag{13.21}$$

Here,

$$\Delta\nu_{\text{rec}} = 2\frac{hk^2}{2m4\pi^2} \tag{13.22}$$

is the frequency separation between the two recoil components that shows up in saturation spectroscopy with narrow spectral lines or in Ramsey-Bordé atom interferometry (Section 6.6.2.3). (13.21) shows that only frequency measurements are needed to derive the fine structure constant and hence allows for a high accuracy.

[3] From the CODATA evaluation [791] its value is $1/\alpha = 137.035\,999\,11 \pm 3.3 \times 10^{-9}$.

13.6.3 Atomic Clocks and the Constancy of Fundamental Constants

The question of whether fundamental constants such as the fine structure constant α are really constant or vary with time, was raised as early as 1937 by Dirac [820] in his large number hypothesis [4]. Dirac's large number hypothesis was based on the observation that most of the dimensionless constants such as the fine structure constant $\alpha \approx 1/137$, are close to unity, but other dimensionless ratios are really large and come up with a value around 10^{40}. Examples include the Coulomb electrostatic force divided by the gravitational force between electron and proton, the length scale of the universe divided by the classical radius of the electron, or the age of the universe divided by the time it takes for light to travel the distance equal to the classical electron radius. If this coincidence was not by accident but all values were to be proportional, they would increase with time since the radius of the universe varies with time. Depending on which of the "constants" can be regarded as constant, different variations of the others are deduced. As an order-of-magnitude for any of the constants β the variation expected from the known value of the Hubble constant is $\dot{\beta}/\beta \approx 10^{-11}$ /year. Today Dirac's large-number hypothesis can be ruled out by experimental data (see [821] and Table 13.3). There are, however, other theories that require the fundamental constants to change with time.

Such a temporal variation of the non-gravitational fundamental constants is excluded by the equivalence principle of General Relativity. Theories attempting to unify gravitation and other interactions, however, may violate this principle. In the concept of string theory as well as in the Kaluza-Klein theories that use extra spatial dimensions [822] new fields namely the scalar "dilaton field" or "moduli fields" are proposed as partners of Einstein's tensor field $g_{\mu\nu}$. These fields couple to matter and might lead to time-varying fundamental constants [823,824] and to a violation of the universality of free fall according to which all bodies fall with the same acceleration in an external gravitational field. These ideas have attracted new interest from the report of frequency shifted absorption spectra from distant quasars [825] that were interpreted as evidence for the cosmological evolution of the fine structure constant α. There are, however, stringent experimental bounds that limit these violations. From the Lunar Laser Ranging experiment [767] one finds that Earth and Moon fall towards the Sun with the same acceleration to better than 10^{-12}. Nuclear data together with the Oklo phenomenon, astrophysical data and clock comparisons, introduce even more stringent limits, to be discussed in the following. The Oklo phenomenon is attributed to the existence of a natural fission reactor moderated by water in Gabon (Western Africa). The evidence that this fission reactor operated about two billion years ago for about a million years has been derived from ores of the Oklo mine that contained much less ^{149}Sm, ^{151}Eu, ^{155}Gd and ^{157}Gd than the usual natural abundancy. As an example, the isotope ratio of the ^{149}Sm/^{147}Sm from this site was measured to be about 0.02 compared to the ratio of natural ores from other sites of about 0.9. This phenomenon has been investigated by the French Commissariat à l'Energie Atomique and it is believed that the missing isotopes like the ^{149}Sm isotope, which is a good neutron absorber, were burned up by the neutron flux from the uranium fission. The relevant transition ^{149}Sm +n $\rightarrow$^{150}Sm + γ has a cross-section which, due to a resonance, is about two orders of magnitude larger than the corresponding transition ^{147}Sm +n $\rightarrow$^{148}Sm + γ. From

[4] An extensive review of the meaning and earlier measurements of the fundamental physical constants has been given by Petley [763]

the ^{149}Sm/^{147}Sm ratio one can deduce a maximal variation of the position of the resonance from the time when the reactor was operational until today. The smallness of the variation of the energy of the resonance places an upper limit to a possible variation of α, as has been pointed out by Shlyakter [826]. The data have been re-evaluated by Damour and Dyson [827] who have obtained stringent limits for $\dot{\alpha}/\alpha < 5 \times 10^{-17}$ /year.

Another completely different source for the determination of any possible time dependence of the fundamental constants is given by the absorption lines from QUASAR spectra. The large distance to these quasi-stellar objects of up to 10^{10} light years, means that the measured absorption spectra contained the information about the value of a fine structure constant 10^{10} years ago. A possible shift of these lines with respect to the corresponding absorption spectra obtained at the present time in the laboratory, however, may be buried in the huge red shift of the radiation from the QUASARs. Comparing the spectra of heavy and light atoms from the same quasi-stellar objects, or the gross structure with the fine structure of the same element, allows one to surpass this difficulty. Evaluations of the spectra of Fe^+ and Mg^+ with respect to the relativistic correction were performed [828–830] and yielded relative variations in $\alpha^2 < 10^{-14}$.

The rapid progress in the development of clocks allows one to investigate possible variations of fundamental constants by comparing the frequencies of clocks based on different physical principles or transitions. Turneaure and Stein performed a twelve-day comparison between the frequencies of a superconductive cavity-stabilised oscillator near 8.6 GHz and a Cs beam clock and observed a relative drift rate of $(-0.4 \pm 3.4) \times 10^{-14}$ /day. This comparison between a macroscopic clock and a microscopic one monitors the Bohr radius determining the dimensions of the cavity with respect to the hyperfine structure splitting of Cs. An upper limit for the variation of a combination of fundamental constants including α^3 was $< 1.5 \times 10^{-12}$ [60, 61, 831]. Godone *et al.* [832] compared, for about a year, the 601 MHz frequency of the fine structure transition $^3P_0 \rightarrow {}^3P_1, \Delta m_j = 0$ of ^{24}Mg with a commercial Cs clock referenced to PTB's primary Cs standard by satellite time comparisons. These authors derived a limit for the time stability of the fine structure constant of $\dot{\alpha}/\alpha < 2.7 \times 10^{-13}$ /year. Prestage *et al.* [833] compared the frequency (40.5 GHz) of the hyperfine structure transition of Hg^+ stored in an ion trap to the hyperfine frequency of a hydrogen maser and obtained a limit of $\dot{\alpha}/\alpha < 3.7 \times 10^{-14}$/a that might be slightly corrected due to the comments of Karshenboim [831]. The sensitivity of such clock rate comparisons to a variable fine structure constant, results from the fact that the relativistic contributions of the hyperfine splitting are a function of α times the nuclear charge Z which increases for heavier atoms or ions.

The rapid progress with fountain clocks and optical frequency standards makes it very likely that the current limits can be reduced dramatically in the years to come. The frequencies of a Cs and a Rb fountain clock have been compared over about five years and yielded $\dot{\alpha}/\alpha = (0.4 \pm 16) \times 10^{-16}$/year [835].

In contrast to the microwave frequency standards, there is no analytical formula for the dependence of the frequency of optical transitions on α. For some of the most promising optical frequency standards, however, the relativistic corrections have been calculated [830, 838]. From these calculations one expects the largest contribution in the ^{199}Hg ion, a smaller contribution in the ^{171}Yb ion and the smallest effect in ^{40}Ca and in ^{1}H. These belong to the optical frequency standards with the currently lowest uncertainties with the best prospects of further significant reduction and, hence, may be the candidates to achieve lower limits on a

Table 13.3: Selected experimental limits for the temporal variation of fundamental constants $\beta \in G, \alpha$ per year a.

Method	β	$\dot{\beta}/\beta$ (a^{-1})	Reference
Lunar orbit	G	$(1 \pm 1) \times 10^{-12}$	[766]
Oklo natural fission reactor	α	$< 5 \times 10^{-17}$	[827]
Quasar spectra	α	$(-2.2 \pm 5.1) \times 10^{-16}$	[828]
Cosmic background	α	$< 7 \cdot 10^{-13}$	[834]
Mg vs Cs clock	α	$< 2.7 \times 10^{-13}$	[832]
H-maser vs Hg^+	α	$< 3.7 \times 10^{-14}$	[833]
Cs vs Rb clock	α	$(-0.04 \pm 1.6) \times 10^{-15}$	[835]
$^{199}Hg^+$ (1.064 THz) vs Cs clock	α	$< 1.2 \times 10^{-15}$	[639]
$^{171}Yb^+$ (1.064 THz) vs Cs clock	α	$< 2 \times 10^{-15}$	[836]
^1H (2.466 THz) vs Cs clock	α	$< 2.9 \times 10^{-15}$	[837]

possible time dependence by comparing their frequencies over some time. A measurement of the optical $^{199}Hg^+$ standard (Section 10.3.2.4) with respect to the hyperfine transition of the Cs atomic clock over a two year duration by Bize *et al.* [639] yielded an upper bound for $|\dot{\alpha}/\alpha| < 1.2 \times 10^{-15}$/year. Combining this result with measurements of the Yb^+ standard [836] (Section 10.3.2.2) or the ^1H standard [837] (Section 9.4.5) similar upper bounds could be derived, however, with an analysis that was using less *a priori* assumptions.

14 To the Limits and Beyond

From the temporal evolution of the quality of frequency standards and clocks demonstrated in Fig. 1.2 one would expect that the tremendous improvement in the accuracy and stability of such devices will not lose its pace in the near future. It is therefore interesting to discuss the constraints that eventually limit the performance of frequency standards. Excellent stability is always prerequisite to achieving the utmost accuracy of a frequency standard since the detection of small systematic frequency shifts is only possible to the extent they can be separated from the frequency fluctuations of the standard. In this chapter we will first recall the limitations to the stability that arise from the quantum nature of the radiation field and the absorbers and discuss ideas that may allow some of the limitations to be overcome. We will end this book with speculations on novel technical developments that may be envisaged to lead to standards with improved accuracy.

14.1 Approaching the Quantum Limits

In this section we assume that technical noise associated with the oscillator in the frequency standard has been reduced to such a level that the stability is limited by fundamental quantum-mechanical fluctuations. There are different limiting cases that depend on the experimental realisation of the frequency standard.

Consider a device where a weak radiation field interrogates a large number of quantum absorbers in the scheme of Fig. 1.3 and the scattered light is used for the detection. Here the fluctuations of the detected radiation may limit the achievable signal-to-noise ratio and, hence, the stability of the frequency standard. For uncorrelated fluctuations of the radiation field the fundamental limit is then often given by the shot noise of the photons (Section 14.1.2).

In other frequency standards the radiation field may interact with a small number of quantum absorbers. Here, the fluctuations of the detected radiation field comprising a large number of photons may be negligible with respect to the fluctuations associated with the interaction process between the quantum absorbers and the photons. In such a case the stability is affected by the quantum projection noise [89], (Section 14.1.3.1) which is already limiting the stability of single-ion frequency standards or the best Cs fountains.

Suitable preparation of either the radiation field or the absorbers with correlated fluctuations, however, may allow one to overcome these limits and approach the ultimate Heisenberg limit based on the quantum mechanical uncertainty relations.

14.1.1 Uncertainty Relations

According to quantum mechanics two non-commuting operators $\hat{A}$ and $\hat{B}$ that have no common eigenstates can be described by

$$\left[\hat{A},\hat{B}\right] \equiv \hat{A}\hat{B} - \hat{B}\hat{A} = i\hat{C} \tag{14.1}$$

which leads to the uncertainty relation

$$\sqrt{\langle\hat{A}^2\rangle\langle\hat{B}^2\rangle} \geq \frac{1}{2}\left|\langle\hat{C}\rangle\right|. \tag{14.2}$$

As an example of (14.1) consider

$$\begin{aligned} [\hat{q}_i,\hat{p}_{i'}] &= i\hbar\delta_{ii'}\mathbf{1} \\ [\hat{p}_i,\hat{p}_{i'}] &= [\hat{q}_i,\hat{q}_{i'}] = 0 \end{aligned} \tag{14.3}$$

with $\hat{C} = \hbar\mathbf{1}$ for the commutator relation between the operator $\hat{q}_i$ and $\hat{p}_{i'}$ of the spatial coordinate and the momentum of a free particle, respectively. The uncertainty relation (14.2) holds for arbitrary non-commuting operators and, hence, also for the operators of the mean deviation, i.e. the uncertainties of the observables $\hat{A}$ and $\hat{B}$

$$\widehat{\Delta A} = \hat{A} - \langle\hat{A}\rangle \qquad \text{and} \qquad \widehat{\Delta B} = \hat{B} - \langle\hat{B}\rangle \tag{14.4}$$

leading to the Heisenberg uncertainty relation

$$\sqrt{\langle\widehat{\Delta A}^2\rangle\langle\widehat{\Delta B}^2\rangle} \geq \frac{1}{2}\left|\langle\hat{C}\rangle\right|. \tag{14.5}$$

Equation (14.5) sets lower limits for the minimal fluctuations that can be achieved when the expectation values of two conjugated quantities are measured that are represented by two non-commuting operators. The limit derived from (14.5) is commonly referred to as the Heisenberg limit. Examples include the uncertainty of the simultaneous measurement of a spatial coordinate and momentum of a particle

$$\Delta x \Delta p_x \geq \frac{\hbar}{2} \tag{14.6}$$

or the uncertainty of the simultaneous measurement of the number of photons n and the phase ϕ of an electromagnetic field (see [39] [1])

$$\Delta n \Delta\phi \geq \frac{1}{2}. \tag{14.7}$$

[1] Note, that there is no Hermitian operator corresponding to the classical phase variable and, hence, the derivation in [39] is semi-classical.

14.1.2 Quantum Fluctuations of the Electromagnetic Field

14.1.2.1 Quantisation of the Field

The classical electromagnetic field based on Maxwell's equations is quantised in textbooks on quantum optics [133, 135] by starting with the electric field in a linear cavity [2] of length L. With the cavity axis chosen as the z axis the electric field which is linearly polarised along the x direction can be expanded in the normal modes of the cavity as

$$E_x(z,t) = \sum_j A_j q_j(t) \sin(k_j z) = \sum_j \sqrt{\frac{2\omega_j^2 m_j}{\epsilon_0 V}} q_j(t) \sin(k_j z). \tag{14.8}$$

Here, the amplitude $A_j q_j(t)$ of the mode is split into a factor A_j with the unit V/m^2 and a "mechanical" amplitude $q_j(t)$ whose unit is m. V is the mode volume of the resonator and $k_j = \omega_j c = j\pi/L$ with $j = 1, 2, \cdots$. From (14.8) and Maxwell's equation (4.23) the magnetic field is obtained as

$$H_y(z,t) = \sum_j \sqrt{\frac{2\omega_j^2 m_j}{\epsilon_0 V}} \frac{\epsilon_0}{k_j} \dot{q}_j(t) \cos(k_j z). \tag{14.9}$$

Insertion of (14.8) and (14.9) into the classical Hamiltonian of the electromagnetic field

$$\mathcal{H} = \frac{1}{2} \int \left[\epsilon_0 E_x^2 + \mu_0 H_y^2\right] dV \tag{14.10}$$

leads to

$$\mathcal{H} = \frac{1}{2} \sum_j \left[m_j \omega_j^2 q_j^2 + \frac{\dot{q}_j^2}{m_j} \right] \tag{14.11}$$

which is equivalent to a Hamiltonian comprising a sum of harmonic oscillators characterised by an amplitude q_j and a velocity $\dot{q}_j$ (momentum p_j/m_j). The electromagnetic field can be quantised by interpreting the amplitude q and momentum p of the j^{th} oscillator in (14.11) as operators which obey the commutator relations (14.3).

The Schrödinger equation with the Hamiltonian of a single harmonic oscillator representing the j^{th} mode in (14.11) can be solved [839] by the correspondence $p = -i\hbar d/dq$ leading to eigenfunctions in terms of the Hermite polynomials and to the eigenenergies

$$W_n = \hbar\omega_j \left(n + \tfrac{1}{2}\right). \tag{14.12}$$

Alternatively, following the treatment of Scully and Zubairy [135] one defines annihilation and creation operators $\hat{a}$ and $\hat{a}^\dagger$, respectively, by

$$\begin{aligned} \hat{a}e^{-i\omega t} &= \frac{1}{\sqrt{2m\hbar\omega}} (m\omega\hat{q} + i\hat{p}) \\ \hat{a}^\dagger e^{i\omega t} &= \frac{1}{\sqrt{2m\hbar\omega}} (m\omega\hat{q} - i\hat{p}) \end{aligned} \tag{14.13}$$

[2] For the purpose of this chapter the differences in the results of the quantisation of the field in unbounded free space do not matter.

with the commutator relations following from (14.3)

$$\begin{aligned}\left[\hat{a}_i, \hat{a}_{i'}^\dagger\right] &= i\hbar\delta_{ii'} \\ [\hat{a}_i, \hat{a}_{i'}] &= \left[\hat{a}_i^\dagger, \hat{a}_{i'}^\dagger\right] = 0. \end{aligned} \tag{14.14}$$

Insertion of (14.13) into (14.10) the Hamiltonian for a single mode reads

$$H = \hbar\omega_j \left(\hat{a}\hat{a}^\dagger + \tfrac{1}{2}\right). \tag{14.15}$$

With the help of (14.13) the electric and magnetic fields (14.8) and (14.9) now read

$$E_x(z,t) = \sum_j \sqrt{\frac{\hbar\omega_j}{\epsilon_0 V}} \left(\hat{a}e^{-i\omega t} + \hat{a}^\dagger e^{i\omega t}\right) \sin(k_j z) \tag{14.16}$$

$$H_y(z,t) = -i\epsilon_0 c \sum_j \sqrt{\frac{\hbar\omega_j}{\epsilon_0 V}} \left(\hat{a}e^{-i\omega t} - \hat{a}^\dagger e^{i\omega t}\right) \cos(k_j z). \tag{14.17}$$

By introducing the Hermitian operators

$$\begin{aligned}\hat{X}_1 &= \frac{1}{2}\left(\hat{a} + \hat{a}^\dagger\right) \qquad \text{and} \\ \hat{X}_2 &= \frac{1}{2i}\left(\hat{a} - \hat{a}^\dagger\right) \end{aligned} \tag{14.18}$$

and the commutation relation

$$\left[\hat{X}_1, \hat{X}_2\right] = \frac{i}{2} \tag{14.19}$$

that follows from (14.18) and (14.14) the electric field for the j^{th} mode now reads

$$E_x(z,t) = 2\sqrt{\frac{\hbar\omega_j}{\epsilon_0 V}} \left(\hat{X}_1 \cos\omega t + \hat{X}_2 \sin\omega t\right) \sin(k_j z). \tag{14.20}$$

The two non-commuting operators $\hat{X}_1$ and $\hat{X}_2$ in (14.20) correspond to the amplitudes of the classical quadratures E_1 and E_2 (see (2.7)) of the field of the classical monochromatic electromagnetic field which are out of phase by $\pi/2$. Hence, the quadrature components of the electromagnetic field cannot both be determined with zero uncertainty at the same time. The uncertainty principle sets a lower limit for the product of the uncertainties in E_1 and E_2. From (14.19) follows the uncertainty relation (see Section 14.1.1)

$$\Delta\hat{X}_1 \Delta\hat{X}_2 \leq \frac{1}{4} \tag{14.21}$$

for the quadrature operators of the electric field. As a classical approximation the electric field is sometimes [39] written as

$$\begin{aligned} E(t) &= E_1(t) + iE_2(t) \qquad\qquad \text{with} \\ E_1(t) &= \langle E_1(t)\rangle + \Delta E_1(t) \qquad \text{and} \\ E_2(t) &= \langle E_2(t)\rangle + \Delta E_2(t) \end{aligned} \tag{14.22}$$

where the quantities $E_1(t)$ and $E_2(t)$ are given by their expectation values $\langle E_1(t)\rangle$ and $\langle E_2(t)\rangle$ which can be obtained as the mean of an (infinite) number of measurements and $\Delta E_1(t)$ and $\Delta E_2(t)$ are the uncertainties. Hence,

$$\Delta E_1 \Delta E_2 \geq \frac{\hbar\omega}{2\epsilon_0 V}. \tag{14.23}$$

14.1.2.2 States of the Light Field

Number States The eigenstates $|n\rangle$ corresponding to the energies W_n of (14.12) represent states with well defined numbers of photons in the j^{th} mode and are referred to as number states or Fock states. The annihilation and creation operators reduce or raise the number of photons, respectively,

$$\hat{a}\,|n\rangle = \sqrt{n}\,|n-1\rangle \tag{14.24}$$

$$\hat{a}^\dagger\,|n\rangle = \sqrt{n+1}\,|n+1\rangle. \tag{14.25}$$

The lowest number state $|n=0\rangle$ is called the vacuum state. The number states form a complete, orthogonal, and normalised set of states.

Coherent States In general it is not easily feasible to prepare states with an exact number of photons n but rather to have a definite average number $\langle n\rangle$ of photons. A particularly useful class of states that have these properties are the coherent states [840] which are eigenstates of the annihilation operator $\hat{a}$ [135, 841]

$$\hat{a}\,|\alpha\rangle = \alpha\,|\alpha\rangle \tag{14.26}$$

where in general α is a complex number. Even though two different coherent states are not orthogonal, the coherent states form a basis which is over complete and, hence, can be used as a basis to expand any state. In contrast to the Fock state where each measurement of the photon number gives the same result, the probability of measuring a number n of photons in a coherent states is given by the Poissonian probability distribution [840, 841]

$$p(n) = \frac{\langle n\rangle^n e^{-n}}{n!}. \tag{14.27}$$

The variance of the Poissonian distribution $(\Delta n)^2$ is equal to the mean value $\langle n\rangle$ and, hence,

$$\Delta n = \sqrt{\langle n\rangle}. \tag{14.28}$$

The fluctuations of the photons according to (14.27) and (14.28), that also show up in the fluctuations of the current from a photodetector used to measure the radiation field, are often termed as shot noise.

An important property of the coherent state is that the uncertainties of the quadrature components (14.21) are minimal and equal. This limit

$$\langle \Delta\hat{X}_1{}^2\rangle = \langle \Delta\hat{X}_2{}^2\rangle = \tfrac{1}{4} \tag{14.29}$$

is referred to as the "standard quantum limit". For E_1 and E_2 the standard quantum limit reads

$$\Delta E_1 = \Delta E_2 = \sqrt{\left(\frac{\hbar\omega}{2\epsilon_0 V}\right)}. \tag{14.30}$$

Hence, the coherent state of the electric field is represented by a phasor whose uncertainty contour is a circle (see Fig. 14.1 a).

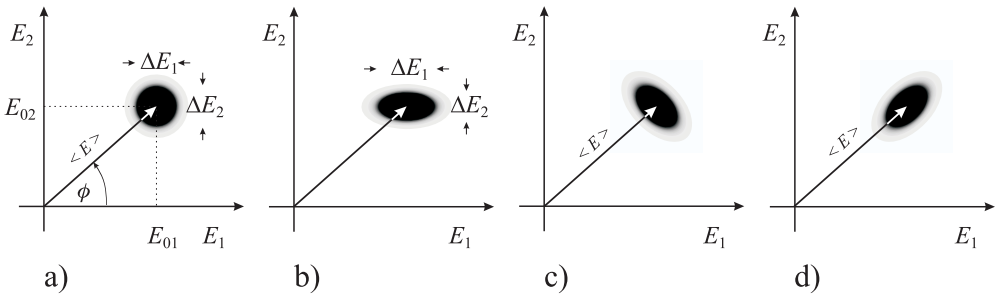


Figure 14.1: a) The phasor of the electromagnetic field in a coherent state represented by the expectation value $\langle E \rangle$ with the uncertainty ΔE_1 and ΔE_2 divided equally between the quadrature components E_1 and E_2, respectively. b) Squeezed field with $\Delta E_1 > \Delta E_2$. c) Squeezed state with amplitude fluctuation less than a coherent state. d) Squeezed state with phase fluctuation less than a coherent state.

14.1.2.3 Squeezed States

A coherent state of the electric field is said to be "squeezed" if the fluctuations of one quadrature component are reduced below the standard quantum limit at the expense of the conjugate one (see Fig. 14.1 b) while the uncertainty relation (14.23) is still preserved. Such cases are depicted in Fig. 14.1 c) and d). In squeezed light the fluctuations of the respective quantities, e.g., phase and field amplitude (photon number) or the quadrature components are correlated. There are several processes that can be used to correlate the quadrature components and to generate squeezed light. As an example, consider a medium exhibiting the optical Kerr effect where the index of refraction depends on the irradiance of the light wave (see (11.28)). Since the phase of the light wave is affected by the index of refraction there is a coupling between the number of photons and the phase and, hence, a correlation between the fluctuations of either quantity at the output of the medium.

14.1.2.4 Possible Applications of Squeezed Light to Frequency Standards

Squeezed light might find applications in optical frequency standards in cases where the quantum fluctuations of the light interrogating the microscopic or macroscopic absorbers, limit the achievable stability of the standard. Which of the quadrature components has to be used depends on the method of interrogation. For standards employing an absorption line as frequency reference for the optical oscillator, amplitude squeezed light could be used. Several methods

are at hand [841] and have been used to produce amplitude squeezed light, e.g., by a second harmonic process [842], in an optical parametric oscillator [843] or in diode lasers [844]. Quadrature phase squeezed light can be generated, e.g., in the four-wave mixing process where the amplitude noise is increased and the phase noise is reduced.

Using amplitude squeezed light produced by an optical parametric oscillator, Polzik *et al.* [843] were able to reduce the fluctuations of the current from a photodetector by 5 dB below the vacuum-state level. The squeezed light was used to measure the Doppler-free saturation signal of atomic Cs at 852 nm with an enhanced sensitivity, thereby reducing the noise by more than 3 dB.

The short-term stability of optical frequency standards is governed in general by the short-term stability of a pre-stabilised laser (Section 9) locked, e.g., to a Fabry-Pérot interferometer where the wavelength of the laser is stabilised with respect to the distance between the end mirrors. The precision with which an interferometer can measure the distance between two or more mirrors is also Heisenberg limited. From the Heisenberg uncertainty relation $\Delta z \Delta p_z \approx \Delta z(m\Delta z/\tau) \geq \hbar/2$ a minimum uncertainty of the position measurement of an end mirror of mass m in a duration τ can be derived as

$$\Delta z_{\mathrm{HL}} = \sqrt{\frac{\hbar\tau}{2m}}. \tag{14.31}$$

For a mass of 0.5 kg, a duration $\tau = 1$ ms the limit according to (14.31) is $\Delta z_{\mathrm{HL}} \approx 3 \times 10^{-19}$ m. In general, the shot noise of the detected photocurrent behind one output of the interferometer leads to fluctuations that impose much higher limitations to the length measurement of an interferometer. Assuming that the photons and photoelectrons obey Poisson statistics, the fractional fluctuations scale inversely with the radiation power and, hence, can be reduced by applying higher power to the interferometer. However, the increasing power is associated with increasing pressure fluctuations of the radiation impinging onto the interferometer mirrors [845]. There is an optimum power where the fluctuations due to the radiation pressure and to the photon counting are equal. The minimum of the combined fluctuations approaches the standard quantum limit of the interferometer. Often, the available laser power is far too low to reach the standard quantum limit and the shot noise with its $1/\sqrt{n}$ level of fractional phase fluctuations for a coherent laser source dominates the measurement precision. To approach the Heisenberg limit of $1/n$ rad in the interferometer at low laser power it has been suggested to use non-classical states of light [845], e.g., by illuminating the unused input port of the interferometer with squeezed light. Following this suggestion an increase in the signal-to-noise ratio of 3 dB with respect to the shot-noise limit has been demonstrated in a Mach-Zehnder interferometer [841, 846]. It has to be emphasised, however, that the reduction of the fluctuations in the photon current has to be paid for by increased fluctuations in the radiation pressure. There have been several other suggestions as to how phase measurements in the interferometer [847] can be improved using entangled or squeezed states [848–850] or by driving an interferometer with two Fock states containing equal numbers of photons [851].

14.1.3 Population Fluctuations of the Quantum Absorbers

14.1.3.1 Quantum Projection Noise

We now turn to a different regime where the fluctuations of the signal obtained by interrogating an ensemble of quantum absorbers by a radiation field are no longer limited by the quantum fluctuations of the field itself.

Consider a suitable two-level system with narrow transition and resonance frequency ν_0 that is interrogated by an electromagnetic field with frequency ν close to the resonance frequency. For two-pulse Ramsey excitation the probability that a quantum absorber is found in state $|2\rangle$ is (see (6.44))

$$p_2 = \frac{1 + \cos(\omega - \omega_0)T}{2}. \tag{14.32}$$

Here, T is the free evolution time and we have assumed that the duration τ of the Ramsey pulses is small compared to T. By scanning the frequency of the oscillator the frequency dependence of the power absorbed by an ensemble of N two-level systems results in a curve like the one shown in Fig. 14.2. Near the maxima of this curve the probability p_2 of finding

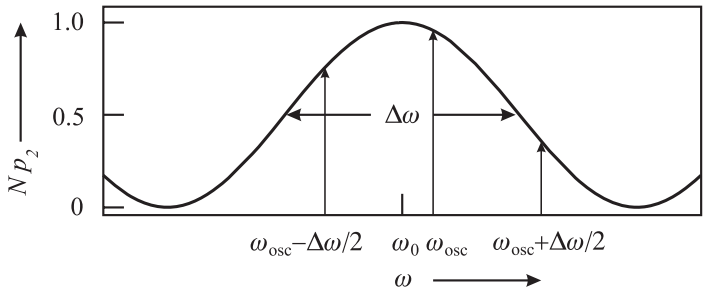


Figure 14.2: Stabilisation of the angular frequency ω of an oscillator to the resonance ω_0 of a suitable absorber is often performed by modulating the angular frequency of the oscillator by $\pm\Delta\omega/2$. For a symmetrical resonance curve, $\omega_{\rm osc}$ is an estimate of ω_0 if the mean signals obtained at the two shifted frequencies are equal.

the atoms in the state $|2\rangle$ is close to unity whereas the minima of the resonance curve reflect the cases when this probability goes to zero. The internal state is in general a superposition of the two eigenstates

$$|\psi\rangle = c_1\,|1\rangle + c_2\,|2\rangle \tag{14.33}$$

where according to Section 5.3.1 $|c_1|^2 + |c_2|^2 = 1$ and $p_1 = |c_1|^2$ and $p_2 = |c_2|^2$ are the probabilities of finding the two-level system in state $|1\rangle$ and in state $|2\rangle$, respectively. When a measurement is performed it is defined whether the two-level system has absorbed a photon or not. Consider the case where the frequency ν is close to the inflection points of the curve in Fig. 14.2. This situation arises, e.g., if the frequency of the interrogating oscillator is square-wave modulated with the modulation width chosen close to the halfwidth of the resonance

curve (Fig. 14.2) in order to lock the oscillator to the centre frequency ω_0. In this and in all other cases except when c_1 or c_2 are zero quantum mechanics requires that the occurrence of the absorption process cannot be predicted without uncertainty. The measurement process projects the two-level atom either in state $|1\rangle$ or into $|2\rangle$ where the photon has been absorbed or not absorbed, respectively. To determine the variance σ^2 of the measurement of the state $|2\rangle$ we follow Itano *et al.* [89] and define a projection operator $\hat{P}_2 \equiv |2\rangle\langle 2|$. The probability of finding the atom in $|2\rangle$ is given by the expectation value $\langle\psi|\hat{P}_2|\psi\rangle = |c_2|^2 = p_2$. The variance is then calculated as

$$\begin{aligned}\sigma^2 = \left(\Delta\hat{P}_2\right)^2 &= \langle\left(\hat{P}_2 - \langle\hat{P}_2\rangle\right)^2\rangle \\ &= \langle\hat{P}_2^2 - 2\langle\hat{P}_2\rangle\hat{P}_2 + \langle\hat{P}_2\rangle^2\rangle \\ &= \langle\hat{P}_2^2\rangle - \langle\hat{P}_2\rangle^2.\end{aligned} \tag{14.34}$$

Using $\hat{P}_2^2 = (|2\rangle\langle 2|)(|2\rangle\langle 2|) = |2\rangle\langle 2| = \hat{P}_2$ it follows that

$$\begin{aligned}\sigma^2 &= \langle\hat{P}_2\rangle - \langle\hat{P}_2\rangle^2 \\ &= \langle\hat{P}_2\rangle\left(1 - \langle\hat{P}_2\rangle\right) \\ &= p_2(1-p_2).\end{aligned} \tag{14.35}$$

As follows from (14.35) the quantum projection noise, described by σ, i.e. the uncertainty of the measurement to find the atom in state $|2\rangle$ is zero when either $p_2 = 1$ or $p_2 = 0$ and has its maximum value $\sigma(p_2 = 1/2) = 1/2$ at the inflection points of Fig. 14.2. For an ensemble of N particles the variance of the number of atoms in the excited state is given as

$$\langle\Delta N\rangle^2 = Np_2(1-p_2). \tag{14.36}$$

The situation is equivalent to the case when photons from a laser beam impinge on a beam splitter characterised by power reflectivity p_2 and transmissivity p_1. The number of photons detected behind either exit port shows the same indeterminism as does the number of two-level atoms in the ground state or excited state. Assuming that the number of quantum absorbers projected onto the excited state $|2\rangle$ in a frequency standard can be detected with probability unity, the uncertainty in the estimated value of ω_0 can be derived using (14.36) as

$$|\delta\omega_0| = \frac{\sqrt{Np_2(1-p_2)}}{\left|\dfrac{d(Np_2)}{d\omega}\right|}. \tag{14.37}$$

From (14.32), in Fig. 14.2 at the inflection points determined by $(\omega - \omega_0)t = \pi/2$ one finds $p_2 = 1/2$ and, hence,

$$|\delta\omega_0| = \frac{1}{\sqrt{N}\tau} \tag{14.38}$$

holds. Again, we encounter a shot-noise limit $\propto 1/\sqrt{N}$. The fundamental limit of the quantum projection noise has been reached using ion traps [89] either with a single $^{199}Hg^+$ ion or with $^9Be^+$ ions. The latter experiment performed with a few to up to 385 ions clearly showed the enhanced noise near the inflection points of Fig. 14.2 according to (14.35). Santarelli *et al.* [64] have observed the quantum projection noise limit also in the case of a caesium fountain clock.

14.1.3.2 Absorbers with Quantum Mechanical Correlations

As in the case of the shot noise level $1/\sqrt{n}$ of the photons, the $1/\sqrt{N}$ level of the fluctuations resulting in the interaction of a radiation field with N two-level atoms, i.e. the quantum projection noise need not be a rigid limit. The use of particularly prepared quantum states with well chosen quantum mechanical correlations could in principle allow one to overcome this limit.

Entangled States Quantum mechanical systems often have to be described by a wave function that cannot be separated into product states, i.e. that cannot be written as a product of the wave systems of the sub-systems. The sub-states are said to be "entangled". Hence, it is possible to prepare two or more particles in an entangled state where a particular quantum mechanical quantity has a well defined value but where, however, the state of the same variable is not well defined for each single particle.

A well known example for an entangled two-particle state represents the so-called Einstein-Podolsky-Rosen state $|\Psi_{\text{EPR}}\rangle$ whose name refers to the Gedanken experiment of Einstein, Podolsky and Rosen [852]. In the version of Bohm a spin-1 particle decays into a pair of spin-1/2 particles described by the singlet state

$$|\Psi_{\text{EPR}}\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow_1,\downarrow_2\rangle - |\downarrow_1,\uparrow_2\rangle\right) \tag{14.39}$$

$$= \frac{1}{\sqrt{2}}\left(|+\vec{r}_1,-\vec{r}_2\rangle - |-\vec{r}_1,+\vec{r}_2\rangle\right). \tag{14.40}$$

Here, $|\uparrow\rangle$ und $|\downarrow\rangle$ are eigenstates of the spins for each particle 1 and 2 along the z axis defining the magnetic quantisation field. $|\vec{r}_1\rangle$ and $|\vec{r}_2\rangle$ are the states in any arbitrary direction $\vec{r}$. Before a measurement is performed neither particle 1 nor particle 2 are in a defined state. However, when the spin of particle 1 has been measured and found to point in a defined direction $\vec{r}$, the direction of the spin of particle 2 is fixed and points along $-\vec{r}$. Consequently, there are correlations between particle 1 and 2 that are independent of the chosen basis. Not only that these correlations are counter-intuitive when asking classical questions such as "How does the second particle know that one has measured the spin of the first one along a particular direction and why can it respond to that measurement with a velocity higher than the speed of light?". Moreover, the quantum mechanical statistical predictions violate the predictions of local theories (see e.g. [853]) expressed by the Bell inequalities [854]). These contradictions are even more visible in an entangled state formed by the triplet of spin-1/2 particles investigated by Greenberger, Horne and Zeilinger (GHZ) [855]

$$|\Psi_{\text{GHZ}}\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow_1,\uparrow_2,\uparrow_3\rangle + ||\downarrow_1,\downarrow_2,\downarrow_3\rangle\right) \tag{14.41}$$

where a single ideal experiment yields completely different results for quantum mechanics and local theories [856]. Mermin [857] investigated GHZ states with N Spin-1/2 particles and pointed out that these states are a superposition of two states differing in all N degrees of freedom. From the fact that only the mean values of the N particle operators show interference effects he concluded that non-locality of quantum mechanics here occurs as a direct consequence of the interference effects between macroscopically distinct states.

Spin-squeezed Atomic States A particular example of entangled particle states are the "spin squeezed states" that can be engineered in such a way that the correlations lead to fluctuations below the shot noise limit (standard quantum limit).

A two-level atom interacting with radiation can be described mathematically in the same way as a spin-1/2 system in a magnetic field (Section 5.3.1), [139]. Consequently, the latter framework is often used to conveniently describe the collective observables of a large number $N = 2S$ of identical two-level atoms. As an example, consider the difference in the number of atoms populating the two internal states of the individual atoms given by the J_z component of the total spin $J = S = N/2$. Applying the commutator relation (14.1) for the angular momentum operators of a spin-1/2 system

$$\left[\widehat{J}_i, \widehat{J}_j\right] = i\hbar\widehat{J}_k \tag{14.42}$$

and cyclic permutations, Heisenberg's uncertainty relation for the Cartesian components follows from (14.5) as

$$\widehat{\Delta J}_x \widehat{\Delta J}_y \geq \frac{\hbar}{2}\left|\langle \widehat{J}_z \rangle\right|. \tag{14.43}$$

Hence, each one of the N individual spin-1/2 systems contributes to the macroscopic spin. If all individual spin systems are in the "spin-up" state, the state of the macroscopic system is an eigenstate $|J_z = S\rangle$ and the total spin vector S of length S(S+1) spans a cone (Fig. 14.3 a).

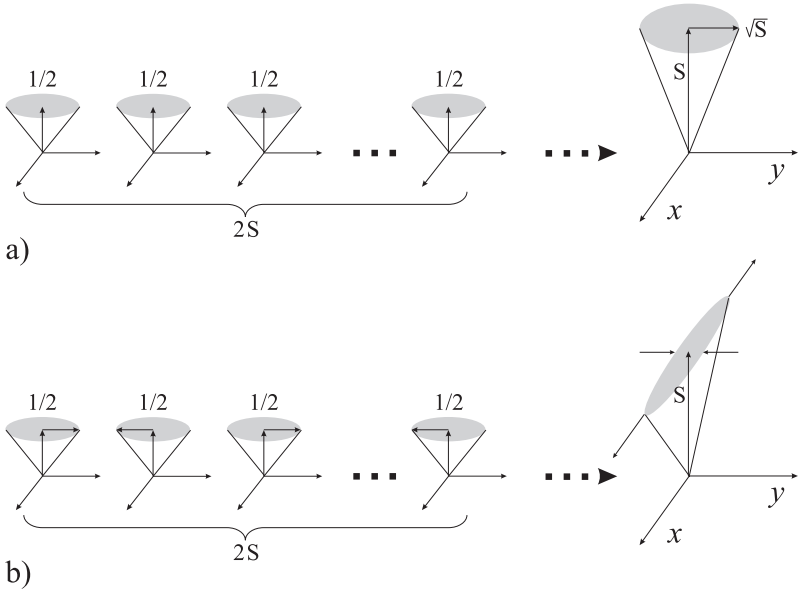


Figure 14.3: a) $2S$ uncorrelated spin-1/2 states lead to a coherent state. b) Correlations between the y components lead to a spin squeezed state with reduced fluctuations along the y direction and increased fluctuations along the x direction (after [858]).

In the coherent spin state the directions of the particular spin vectors add up incoherently in such a way that the fluctuations ΔJ_x and ΔJ_y are equal and satisfy the minimum value of $\hbar/2\,|J_z|$ (see (14.43)).

In contrast, consider a case where the individual spins are correlated in the way depicted in Fig. 14.3 b) which leads to a squeezing of the fluctuations of the total spin in the y direction with enhanced fluctuations in the x direction.

Preparation of Spin-squeezed Atomic States Methods have been developed to prepare spin-squeezed atomic states similarly to the one in Fig. 14.3 b). We consider in the following primarily entangled states of very few trapped ions, e.g., in a linear rf trap since the ions can be prepared easily in well prepared motional states and the dissipative interaction with the environment is small. A collection of N spin-1/2 particles can be prepared in an arbitrary entangled state

$$\begin{aligned} |\psi\rangle &= a_0|\downarrow\rangle_1|\downarrow\rangle_2\cdots|\downarrow\rangle_N \\ &+ a_1|\downarrow\rangle_1|\downarrow\rangle_2\cdots|\uparrow\rangle_N \\ &+ +\cdots \\ &+ a_k|\downarrow\rangle_1|\downarrow\rangle_2\cdots|\downarrow\rangle_k\cdots|\uparrow\rangle_{N-1}|\uparrow\rangle_N \\ &+ \cdots \\ &+ a_{2^N-1}|\uparrow\rangle_1|\uparrow\rangle_2\cdots|\uparrow\rangle_N \end{aligned} \tag{14.44}$$

where the a_k are the amplitudes with k atoms in the up-state. Such a state can be constructed, e.g., by use of the Cirac-Zoller scheme [859] that has been used to demonstrate [860] an efficient entanglement using a pair of ions trapped in an ion trap.

The maximally entangled state

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|\downarrow\rangle_1|\downarrow\rangle_2\cdots|\downarrow\rangle_N + e^{i\Phi}|\uparrow\rangle_1|\uparrow\rangle_2\cdots|\uparrow\rangle_N\right) \tag{14.45}$$

as the generalisation of the GHZ state (see (14.41)) has the consequence that a measurement on any atom immediately determines the value of all other atoms. Mølmer and Sørensen have devised a scheme [861] to produce large-scale maximally entangled states of (14.45) with a single laser pulse.

Consider an atomic or ionic system like the one shown in Fig. 14.4 a) where two spin-1/2 particles are kept in a harmonic potential. Such a system is realised by two identical ions in the harmonic well of an ion trap where the angular oscillation frequencies are ω_m. We assume that both ions are initially entangled in the $|\downarrow\downarrow\rangle$ internal state since they are in a collective motional state n where the centre-of-mass motion of the two ions has an energy of $n\hbar\omega_m$. Two Raman pulses of (optical) angular frequencies $(\omega_0-\omega_m)$ and $(\omega_0+\omega_m)$ are used to bring the ions from the state $|\downarrow\downarrow\rangle$ to the state $|\uparrow\uparrow\rangle$ via two interfering paths. Due to the detuning δ, neither one of the frequencies is resonant with a *single-particle* transition but the sum of both frequencies is resonant with the *two-particle* transition. The transition amplitudes for the two paths are $(\eta\Omega_R\sqrt{n+1})^2/\delta$ and $-(\eta\Omega_R\sqrt{n})^2/\delta$ where η is the Lamb-Dicke parameter and Ω_R is the single-ion Rabi frequency on resonance. The n dependence of the two paths is different and can be calculated using the properties of the creation and annihilation operators

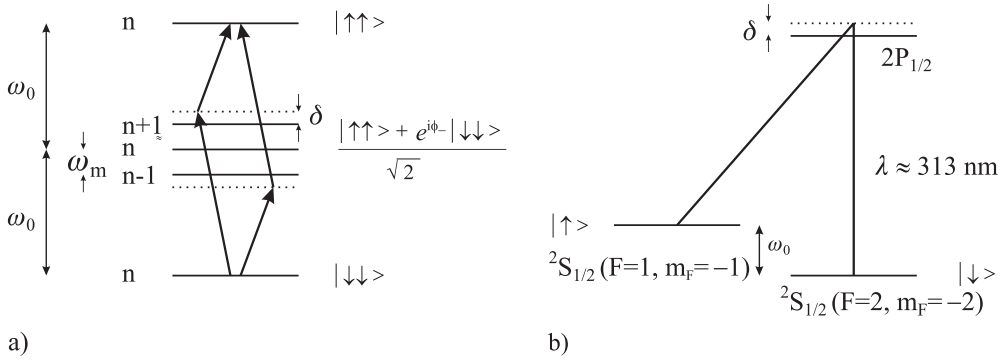


Figure 14.4: a) Entanglement scheme for two particles according to Mølmer and Sørensen [861]. ϕ_- is the difference in the two laser phases at the positions of the two ions. b) Adaptation of the scheme of Fig. 14.4 a) to $^9Be^+$ ions by Sackett *et al.* [862].

given in (14.24) and (14.25) for a harmonic oscillator. The – sign results from the blue and red detunings along the two different paths. When the amplitudes are added the interesting feature occurs that the total transition amplitude is $\eta^2\Omega_R^2/\delta$ which does not depend on the quantum number n of the intermediate motional state for ions in the Lamb-Dicke regime ($\eta^2(n+1) \ll 1$). By driving the transition with the appropriate $\pi/2$ pulse, the entangled state

$$\psi_2 = \frac{|\uparrow\uparrow\rangle + e^{i\phi_+}|\downarrow\downarrow\rangle}{\sqrt{2}} \tag{14.46}$$

can be generated where ϕ_+ is the sum of the two laser phases at the positions of the two ions. Analogously in the state of N entangled ions ϕ_+ contains the sum of all N laser phases and the probability of finding an ion in the excited state oscillates with

$$p_2^{(N)} = \frac{1 + \cos\left[N(\omega - \omega_0)\tau\right]}{2} \tag{14.47}$$

for excitation with two Ramsey pulses which differs from (14.32) by the factor of N in the argument. Consequently, this affects the uncertainty in estimating the frequency ω_0 which now is

$$|\delta\omega_0| = \frac{1}{N\tau}. \tag{14.48}$$

In contrast to (14.32) the measurement with the entangled state now decreases with $1/N$ rather than with $1/\sqrt{N}$ which for large numbers N could improve the stability of a frequency standard with entangled absorbers considerably. The scheme works for any number of even ions and can be extended to odd numbers [863]. It has been applied in a slightly modified system (Fig. 14.4 b, [862]) to two and four $^9Be^+$ ions with two ground states.

Applications of Entangled States to Frequency Standards Larger systems with N entangled atomic or ionic states have been suggested to lead to an improved quantum limited signal-to-noise ratio in frequency standards [864–866]. The basic idea from is taken from [865] and

shown in Fig. 14.5 for the evolution of the pseudo-spin vector in the two-pulse Ramsey excitation described in more detail in Section 6.6.1. We start with a state where all atoms are in a

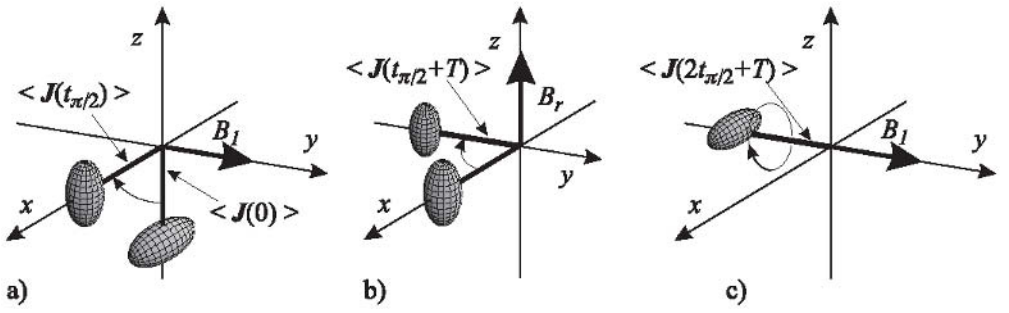


Figure 14.5: Ramsey spectroscopy with spin squeezed states.

squeezed ground state and where the correlation between the quantum mechanical absorbers is such that the uncertainties $\Delta J_x, \Delta J_y, \Delta J_z$ of the combined state are correlated and form an uncertainty ellipsoid which is squeezed for $\Delta J_y(0)$. Applying the first $\pi/2$ interaction pulse the pseudo spin is rotated in the $x-z$ plane around the y axis which is also the direction of the B_1 field driving the clock transition, e.g., between the two ground states of Cs. After this short interaction the expectation value $\langle J(t_{\pi/2})\rangle$ is oriented along the x axis (Fig. 14.5 a). In the time between the first and second Ramsey pulses the magnetic moment and, hence, the pseudo spin precesses around the C field (B_r) which is assumed to be parallel to the z axis. For a detuning $\omega - \omega_0 = -T\pi/2$ corresponding to a red detuning of the exciting radiation field from the atomic resonance of half a linewidth, $\langle J(t_{\pi/2}+T)\rangle$ now points along the $-y$ axis (Fig. 14.5 b). The second Ramsey pulse rotates the pseudo spin by $\pi/2$ around the B_1 field. Interestingly, the fluctuations $\Delta J_z = \Delta J_y(0)$ are now squeezed (Fig. 14.5 c) and the quantum projection noise is smaller than for a coherent state.

It has been pointed out [867], however, that the advantage presented in the maximally entangled states [866] is lost if the influence of the decoherence is taken into account. In contrast to other, more robust entangled states [867, 868] the maximally entangled states are particularly fragile. Huelga *et al.* [867] came to the conclusion that maximally entangled states of N particles lead to the same minimal uncertainty as is obtained with standard Ramsey spectroscopy, but at a $\sqrt{N}$ shorter time. The minimum measurement time is therefore given by the decoherence time of the spin phases resulting from collisions, magnetic, electric stray fields, or fluctuations of the radiation source.

Meyer *et al.* [869] have demonstrated how entangled states can be used to improve the precision when determining the frequency in a Ramsey excitation scheme below the standard quantum limit (shot noise limit) in a coherent spin state with two ions (Fig. 14.6). In the case of a few ions the small gain margin between the standard quantum limit and the Heisenberg limit $\propto 1/\sqrt{N}$ will not be as large as with neutral atom ensembles.

For neutral atom ensembles, several different routes to entanglement are explored. Two-atom correlations also have been observed from a thermal source of ultra-cold atoms [870] but far from quantum degeneracy the small overlap of the wave functions makes these correlations not very suitable to be used directly. It has been proposed also to couple neutral atoms

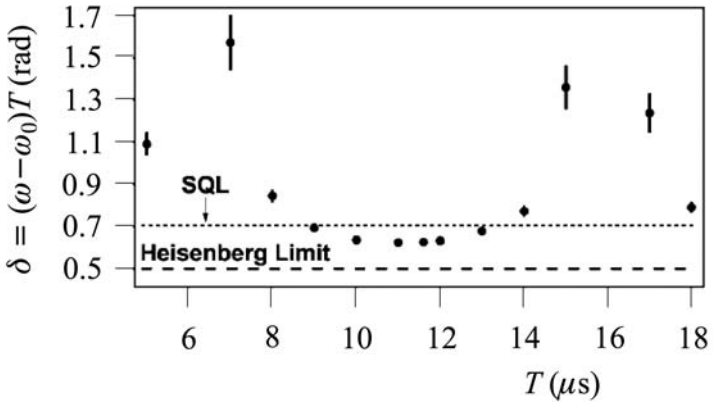


Figure 14.6: Improvement in the precision to determine the transition frequency in a Ramsey interrogation scheme [869]. SQL: Standard quantum limit. Courtesy of D. Wineland.

in an optical lattice by phonon-like excitations [871] but it remains to be shown that quantum entanglement of neutral atoms will lead to a reduced signal-to-noise ratio in frequency standards. Squeezed atomic samples have been produced by irradiating atoms with squeezed light [872], or with an off-resonant laser beam [873]. The latter method led to a spin noise reduction of 70 % below the standard quantum limit expected for a coherent spin state. Atom-number squeezed states have been prepared in an optical lattice populated by atoms from a Bose–Einstein condensate [874].

The application of non-classical states of light or entangled atomic absorbers may allow one to overcome the standard quantum limit for coherent states for a given light power or number of absorbers but not for the Heisenberg limit. Nevertheless, the potential is huge as can be seen from an example of neutral atom frequency standards. The instability of a Cs fountain clock employing 10^5 atoms can be decreased by a factor of ≈ 300 from the $1/\sqrt{N}$ dependent standard quantum limit to the $1/N$ dependent Heisenberg limit in the same measurement time. Alternatively, in the latter case the same level of instability can be reached by a 10^5 fold reduced measurement time.

It has been shown theoretically [875] that quantum entanglement and squeezing can also be used to overcome the classical limits in clock synchronisation or ranging.

14.2 Novel Concepts

14.2.1 Ion Optical Clocks Using an Auxiliary Readout Ion

There are excellent candidate ions not yet used for optical frequency standards since their transitions for cooling the ion and detecting the excitation by the electron shelving technique are in the deep ultraviolet. A method has been devised [644] that allows one to overcome these limitations by using two ions in the same trap where besides the clock ion there is a second

one, called the logic [3] ion, which is used for cooling the clock ion and detecting transitions in the clock ion. The method makes use of the entanglement of the internal states of the clock and logic ion with their external degrees of freedom. Consider the case where both, the clock ion and the logic ion, are described by a two-level system with the two states $|\uparrow\rangle$ and $|\downarrow\rangle$. The motion of the two ions in the trap is described by the motional quantum state $|n\rangle_M$ with $n = 0, 1, 2, \cdots$. The clock ion is sympathetically cooled by the logic ion which in turn is cooled by conventional laser cooling to the initial state prepared such that both atoms being are their ground states and in the quantised motional ground state $|0\rangle_M$ (see Fig. 14.7 a)

$$|\psi_0\rangle = |\downarrow\rangle_L|\downarrow\rangle_C|0\rangle_M. \tag{14.49}$$

Application of a coherent pulse of radiation tuned to the clock transition puts the two elec-

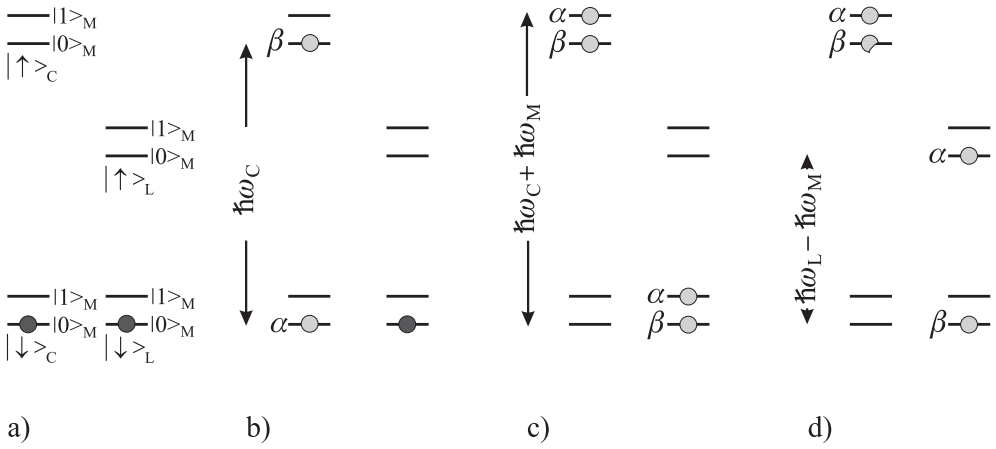


Figure 14.7: Scheme for excitation of a clock ion (index: C) and read-out by a logic ion (index: L) in different electronic states ($\uparrow$ and $\downarrow$) and motional states (index: M) according to [644]. a) Clock ion and logic ion are in their electronic and motional ground states. b) Excitation of the clock ion. c) A π pulse detuned to the blue motional sideband maps the excitation amplitudes α and β of the clock ion onto the motional states of the clock ion and, as a result of the entanglement, the logic ion. d) A π pulse detuned to the red motional sideband maps α and β on the electronic states of the logic ion.

tronic states of the clock ion into a coherent superposition with the amplitude coefficients α and β leading to (Fig. 14.7 b)

$$\begin{aligned} |\psi_0\rangle \rightarrow |\psi_1\rangle &= |\downarrow\rangle_L\left[\alpha|\downarrow\rangle_C + \beta|\uparrow\rangle_C\right]|0\rangle_M \\ &= |\downarrow\rangle_L\left[\alpha|\downarrow\rangle_C|0\rangle_M + \beta|\uparrow\rangle_C|0\rangle_M\right]. \end{aligned} \tag{14.50}$$

If now a π pulse which tuned to the blue motional sideband is applied to the clock ion only the state is $|\downarrow\rangle_C$ is affected whereas the state $|\uparrow\rangle_C$ is not since there is no state $|\downarrow\rangle_C|-1\rangle_M$. Hence,

$$\begin{aligned} |\psi_1\rangle \rightarrow |\psi_2\rangle &= |\uparrow\rangle_L\left[\alpha|\uparrow\rangle_C|1\rangle_M + \beta|\uparrow\rangle_C|0\rangle_M\right] \\ &= |\downarrow\rangle_L|\uparrow\rangle_C\left[\alpha|1\rangle_M + \beta|0\rangle_M\right]. \end{aligned} \tag{14.51}$$

[3] This name reflects the fact that the applied scheme is very similar to the ones used in quantum information processing with ions.

By comparing (14.50) and (14.51) one notices that application of this blue-detuned π pulse has mapped the clock states onto the motional states (Fig. 14.7 c). Due to to the very nature of the entanglement of the internal and external states of both ions this mapping affects both the clock and the logic ion. In the next step (see Fig. 14.7 d) the motional state is mapped onto the state of the logic ion by applying a π pulse which is tuned to the red motional sideband as

$$|\psi_2\rangle \rightarrow |\psi_{\text{final}}\rangle = [\alpha|\uparrow\rangle_L + \beta|\downarrow\rangle_L]\,|\uparrow\rangle_C|0\rangle_M. \tag{14.52}$$

Now, the probability β^2 to find the logic ion in the ground state can be read out, e.g., by a measurement of the number of quantum jumps using the electron-shelving technique. At the same time, this measurement gives the probability that the clock ion was excited.

An experiment using this technique is performed at the NIST, Boulder (USA) using the clock transition $^1\mathrm{S}_0 \rightarrow ^3\mathrm{P}_0$ of the $^{27}\mathrm{Al}^+$ ion ($\lambda = 267.44\,\text{nm}$) with an expected lifetime of $\tau(^3\mathrm{P}_0 = 284\,\text{s})$ [644]. The cooling of the $^{27}\mathrm{Al}^+$ ion and the detection of the clock transition is performed by using Be^+ as the logic ion.

14.2.2 Neutral-atom Lattice Clocks

A promising candidate for an optical frequency standard has been proposed by Hidetoshi Katori [163, 493] that combines the advantages of single ion standards and an ensemble with a large number of neutral atoms, namely long interaction times and high short-term stability. In this approach ultra-cold Sr atoms are trapped in the potential wells of an optical lattice (Section 6.4.2) and the clock transition is excited in the trapped ensemble. Even though the radiation producing the optical lattice will shift the atomic levels connected by the clock transition, a so-called "magic" wavelength can be found, where the light shifts for both states cancel each other out. A similar concept was also used in the ^{9}Be Penning trap microwave standard where the huge Zeeman shift of the two states involved in the microwave clock transition, cancels at a well defined ("magic") magnetic field (Section 10.3.1.1). The basic idea relies on the fact that it is not necessary to avoid all perturbations, but rather to control them in a defined way. Katori proposed to use the $5\mathrm{s}^2\,^1\mathrm{S}_0(F = 9/2)$–$5\mathrm{s}5\mathrm{p}\,^3\mathrm{P}_0(F = 9/2)$ transition of the ^{87}Sr isotope. Due to the nuclear spin-orbit interaction, the strictly forbidden $J = 0 \rightarrow J = 0$ transition acquires a dipole-allowed transition probability corresponding to an excited-state lifetime of about 160 s by hyperfine mixing of the $^3\mathrm{P}_0(F = 9/2)$ state with the $^1\mathrm{P}_1$ and $^3\mathrm{P}_1$ states. The clock transition of ^{87}Sr atoms has been observed trapped in a one-dimensional optical lattice [876] with a magic wavelength of $(813.5 \pm 0.9)\,\text{nm}$. Its frequency has been determined in a ballistic flight to be 429 228 004 235(20) kHz [877]. Trapping in an optical lattice allowed Ido and Katori [878] to confine ^{88}Sr atoms in the Lamb-Dicke regime where the first-order Doppler effect is suppressed and no recoil shift occurs (Fig. 14.8). Calculations of ac multipole polarisabilities and dipole hyperpolarisabilities for the clock transition performed, indicate that the contribution of the higher-order light shifts can be reduced to less than 1 mHz, allowing for a projected fractional inaccuracy of better than 10^{-17} [876]. Besides strontium there are other suitable candidates such as ^{171}Yb [879, 880] or ^{43}Ca. The possibility to store a large number of atoms in an optical lattice at the magic wavelength and in the motional ground states may furthermore allow one to make use of the full potential of the other promising methods discussed in this chapter.

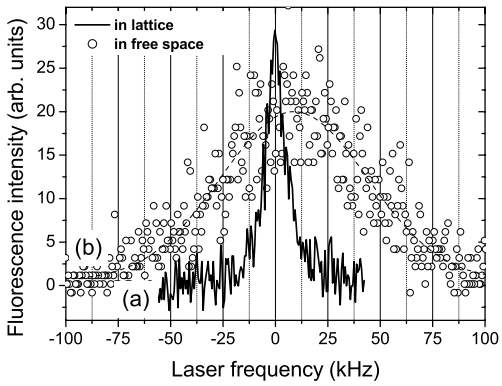


Figure 14.8: Measured intercombination transition of ^{88}Sr atoms, a) using atoms trapped in an optical lattice operated at a "magic" wavelength, b) using atoms in ballistic flight [878]. Courtesy of H. Katori.

14.2.3 On the Use of Nuclear Transitions

Up to this point in frequency standards, only clock transitions have been used that connect states of the electronic shell of the atoms. In the nucleus, however, there are extremely long-lived states that may lead to narrow linewidths that have been favourably used in Mössbauer spectroscopy. In comparison to electronic states, nuclear energy states might be more immune to a number of external perturbations, e.g., collisions or blackbody radiation. In general, the energy separations of the nuclear states are much larger as compared to the energy states in the electronic shell and, hence, suitable Mössbauer transitions have to be excited by the radiation of X-ray sources whose coherence properties are far from those of the oscillators used in frequency standards. Even though the coherence of radiation generated by X-ray lasers, higher harmonic generation from pulsed laser sources, or from future free-electron lasers is steadily increasing, truly phase coherent sources currently exist only up to the optical regime.

There might, however, exist other possibilities of using nuclear transitions for optical frequency standards. From γ-ray spectroscopy, a long-lived isomeric state of the ^{229}Th nucleus has been deduced [633, 881, 882]. This state has an estimated lifetime of a few hours and has an energy of (3.5 ± 1.0) eV above the ground state. Peik and Tamm [883] have proposed a method of detecting the laser excitation of the nucleus in a double resonance method by probing hyperfine transitions in the electronic shell. These authors show that the frequency of the nuclear transition is independent of external magnetic and electric fields to first and second order, respectively, making the ^{229}Th absorber a novel candidate for a high-accuracy optical clock.

14.3 Ultimate Limitations Due to the Environment

The surface of Earth is not the most ideal location for the accurate and stable clocks one can envisage for the future. A variety of perturbations both of man-made and natural origin may

limit the performance and, hence, the achievable accuracy and stability will depend on the extent to which one can control these perturbations. Examples include seismic perturbations that limit the optimal stability of macroscopic references, temperature radiation leading to the black-body shift (Section 7.1.3.4), or changes in the gravitational potential.

According to the definition each frequency standard and clock realises its proper time and can be used for numerous applications, e.g., those described in Section 12 and Section 13. When operated in a gravitational potential, however, the influence of this potential has to be taken into account if the frequency of the standard has to be compared with other clocks on a different potential. As was shown (see (12.24)) this influence can be as large as 7×10^{-10} in the vicinity of the geoid. The uncertainty of the potential on the geoid is of the order of $1\mathrm{m}^2/\mathrm{s}^2$ [263] which contributes to an uncertainty of about 1×10^{-17} in s/s(TCG). If the height above the geoid can be determined by the use of geodetic GPS receivers to about 1 m this leads to an uncertainty of about 10^{-16} in s/s(TCG). Differential GPS together with levelling networks can now give levelling uncertainties of a few centimetres. For this regime of accuracy Earth can no longer be regarded as a stable platform. Tidal effects from the gravitational potentials of the Moon and Sun can change local elevation by several tens of centimetres. Even a continental drift of about 1 cm/year would lead to a relative frequency shift of 1×10^{-18} via the first-order Doppler effect. Thus synchronisation of clocks with respect to coordinate time TCG on the surface of Earth seems to be limited to a few parts in 10^{-17} [724].

With this order of uncertainty a typical Cs fountain of 1 m height can no longer be regarded as a local system since according to (12.22) the proper normalised frequency of a much smaller clock would vary by 1.1×10^{-16} over this vertical distance.

The deep gravitational potential well of Earth makes its surface not the most ideal place for clocks and one may speculate on future "master clocks" in a space region with a more flat gravitational potential. For frequency standards and clocks on board terrestrial satellites Wolf [724] has shown that synchronisation with respect to TCG can be achieved within a few parts in 10^{18} limited by the accuracy of the orbits. The required accuracy of position and velocity are 1 cm and 1×10^{-5} m/s for a satellite at 1000 km altitude and about 0.4 m and 3×10^{-5} m/s for a geostationary satellite, respectively.

If the rapid evolution of the quality of frequency standards and clocks demonstrated in Fig. 1.2 will go on also in the future one might have to be prepared to put the best clocks in a microgravity environment. First experiments on the International Space Station (ISS) are already under way (Section 7.4) but such a multi-purpose space station like the ISS which, e.g. is loosing height with a rate of about 1 cm/s due to the air drag and needs periodic re-boosts [884] is still not the ideal microgravity environment. One may envisage that eventually such atomic "master clocks in space" will be placed on dedicated satellites.

Bibliography

[1] R. L. Sydnor and D. W. Allan, editors. *Handbook Selection and Use of Precise Frequency and Time Systems*. Radiocommunication Bureau of the International Telecommunication Union, ITU, Place des Nations, CH-1211 Geneva 20, Switzerland, 1997.

[2] J. R. Vig. Quartz crystal oscillators and resonators. http://www.ieee-uffc.org/fc, October 1999. SLCET-TR-88-1 (Rev. 8.3.9).

[3] Guide to the expression of uncertainty in measurement. ISO/TAG 4. Published by ISO, 1993 (corrected and reprinted, 1995) in the name of the BIPM, IEC, IFCC, ISO, IUPAC, IUPAP and OIML, 1995. ISBN number: 92-67-10188-9, 1995.

[4] J. W. Wells. Coral growth and geochronometry. *Nature*, 197:948–950, 1963.

[5] J. H. Taylor, Jr. Millisecond pulsars: Nature's most stable clocks. *Proc. IEEE*, 79:1054–1062, 1991.

[6] Dava Sobel. *Longitude*. Walker Books, New York, 1995.

[7] A. Scheibe and U. Adelsberger. Schwankungen der astronomischen Tageslänge und der astronomischen Zeitbestimmung nach den Quarzuhren der Physikalisch–Technischen Reichsanstalt. *Physikal. Zeitschrift*, 37:185–203, 1936.

[8] Tony Jones. *Splitting the Second: The Story of Atomic Time*. Institute of Physics, Bristol and Philadelphia, 2000.

[9] Norman F. Ramsey. History of atomic clocks. *J. Res. NBS*, 88:301–320, 1983.

[10] N. F. Ramsey. Experiments with separated oscillatory fields and hydrogen masers. *Rev. Mod. Phys.*, 62:541–552, 1990.

[11] Jacques Vanier and Claude Audoin. *The Quantum Physics of Atomic Frequency Standards*. Adam Hilger, Bristol and Philadelphia, 1989.

[12] Norman F. Ramsey. Fifty years of atomic frequency standards. In P. Gill, editor, *Frequency Standards and Metrology, Proceedings of the Sixth Symposium*, pages 8–17, Singapore, 2002. World Scientific.

[13] L. Essen and J. V. L. Parry. The caesium frequency standard. In *NPL News*, volume 65. National Physics Laboratory, Teddington, UK, September 1955.

[14] L. Essen and J. V. L Parry. The Caesium resonator as a standard of frequency and time. *Phil. Trans. Roy. Soc.*, A 250:45–69, 1957.

[15] Paul Forman. Atomichron: The atomic clock from concept to commercial product. *Proc. IEEE*, 73:1181–1204, 1985.

[16] Andreas Bauch, K. Dorenwendt, B. Fischer, Thomas Heindorff, E. K. Müller, and Roland Schröder. CS2: The PTB's new primary clock. *IEEE Trans. Instrum. Meas.*, IM–36:613–616, 1987.

[17] A. Clairon, C. Salomon, S. Guellati, and W. D. Phillips. Ramsey resonance in a Zacharias fountain. *Europhys. Lett.*, 16:165–170, 1991.

[18] Pierre Lemonde, Philippe Laurent, Giorgio Santarelli, Michel Abgrall, Yvan Sortais, Sébastian Bize, Christophe Nicolas, Shougang Zhang, André Clairon, Noël Dimarcq, Pierre Petit, Anthony G. Mann, Andre N. Luiten, Sheng Chang, and Christophe Salomon. Cold-atom clocks on earth and in space. In Andre N. Luiten, editor, *Frequency Measurement and Control*, volume 79, pages 131–152. Springer, Berlin, Heidelberg, New York, 2001.

[19] S. Weyers, U. Hübner, R. Schröder, Chr. Tamm, and A. Bauch. Uncertainty evaluation of the atomic caesium fountain CSF1 of the PTB. *Metrologia*, 38:343–352, 2001.

[20] S. R. Jefferts, J. Shirley, T. E. Parker, T. P. Heavner, D. M. Meekhof, C. Nelson, F. Levi, G. Costanzo, A. De Marchi, R. Drullinger, L. Hollberg, W. D. Lee, and F. L. Walls. Accuracy evaluation of NIST-F1. *Metrologia*, 39:321–336, 2002.

[21] S. A. Diddams, Th. Udem, J. C. Bergquist, E. A. Curtis, R. E. Drullinger, L. Hollberg, W. M. Itano, W. D. Lee, C. W. Oates, K. R. Vogel, and D. J. Wineland. An optical clock based on a single trapped $^{199}Hg^{+}$ ion. *Science*, 293:825–828, 2001.

[22] Paul Horowitz and Winfield Hill. *The Art of Electronics*. Cambridge University Press, Cambridge, New York, Melbourne, second edition, 1989.

[23] P. Kersten. Ein transportables optisches Calcium-Frequenznormal. PTB-Bericht PTB-Opt–59, Physikalisch-Technische Bundesanstalt, Braunschweig, 1998.

[24] John L. Hall, Matthew S. Taubman, and Jun Ye. Laser stabilization. In Michael Bass, Jay M. Enoch, Eric W. Van Stryland, and William L. Wolfe, editors, *Handbook of Optics*, pages 27.1 – 25.24. McGraw-Hill, New York, 2001.

[25] J. Rutman. Characterization of phase and frequency instabilities in precision frequency sources: fifteen years of progress. *Proc. IEEE*, 66:1048–1075, 1978.

[26] D. W. Allan. Statistics of atomic frequency standards. *Proc. IEEE*, 54:221–230, 1966.

[27] J. A. Barnes, A. R. Chi, L. S. Cutler, D. J. Healey, D. B. Leeson, T. E. McGunigal, J. A. Mullan, W. L. Smith, R. L. Sydnor, R. F. C. Vessot, and G. M. R. Winkler. Characterization of frequency stability. *IEEE Trans. Instrum. Meas.*, IM-20:105–120, 1971.

[28] John A. Kusters, Leonard S. Cutler, and Edward D. Powers. Long-term experience with cesium beam frequency standards. In *Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum and The IEEE International Frequency Control Symposium*, pages 159–163, 1999.

[29] Andreas Bauch. Caesium atomic clocks: Function, performance and applications. *Meas. Sci. Technol.*, 14:1159–1173, 2003.

[30] Anthony G. Mann. Ultrastable cryogenic microwave oscillators. In A. N. Luiten, editor, *Frequency Measurement and Control*, volume 79 of *Topics in Applied Physics*, pages 37–66. Springer, Berlin, Heidelberg, New York, 2001.

[31] B. C. Young, F. C. Cruz, W. M. Itano, and J. C. Bergquist. Visible lasers with subhertz linewidths. *Phys. Rev. Lett.*, 82:3799–3802, 1999.

[32] C. W. Oates, E. A. Curtis, and L. Hollberg. Improved short-term stability of optical frequency standards: approaching 1 Hz in 1 s with the Ca standard at 657 nm. *Opt. Lett.*, 25:1603–1605, 2000.

[33] Angela Duparré, Josep Ferre-Borrull, Stefan Gliech, Gunther Notni, Jörg Steinert, and Jean M. Bennett. Surface characterization techniques for determining the root-mean-square roughness and power spectral densities of optical components. *Appl. Opt.*, 41:154–171, 2002.

[34] David W. Allan and James Barnes. A modified "Allan variance" with increased oscillator characterization ability. In *Proceedings of the 35th Ann. Freq. Control Symposium*, pages 470–475, Ft. Monmouth, NJ 07703, May 1981. Electronic Industries Association.

[35] John L. Hall and Miao Zhu. An introduction to phase-stable optical sources. In *Laser Manipulation of Atoms and Ions*, volume Course CXVIII of *Proceedings Internat. School of Physics "Enrico Fermi"*, pages 671–702. North Holland–Elsevier, Amsterdam, 1992.

[36] D. S. Elliott, Rajarshi Roy, and S. J. Smith. Extracavity laser band-shape and bandwidth modification. *Phys. Rev. A*, 26:12–26, 1982.

[37] H. R. Telle. Stabilization and modulation schemes of laser diodes for applied spectroscopy. *Spectrochimica Acta Rev.*, 15:301–327, 1993.

[38] A. Godone and F. Levi. About the radiofrequency spectrum of a phase noise modulated carrier. In *Proceedings of the 1998 European Frequency and Time Forum*, pages 392–396, 1998.

[39] Amnon Yariv. *Optical Electronics in Modern Communications*. Oxford University Press, New York, Oxford, fifth edition, 1997.

[40] Christian Koch. Vierwellen-Mischung in Laserdioden. PTB-Bericht Opt-43, Physikalisch-Technische Bundesanstalt, Braunschweig, 1994.

[41] M. J. O'Mahony and I. D. Henning. Semiconductor laser linewidth broadening due to $1/f$ carrier noise. *Electron. Lett.*, 19:1000–1001, 1983.

[42] H. Telle:. Lecture notes; unpublished, 2003.

[43] Fred L. Walls. Phase noise issues in femtosecond lasers. In John L. Hall and Jun Ye, editors, *Proceedings of SPIE: Laser Frequency Stabilization, Standards, Measurement, and Applications*, volume 4269, pages 170–177, P.O. Box 10, Bellingham, Washington 98227–0010 USA, 2001. SPIE.

[44] Klaus H. Sann. The measurement of near-carrier noise in microwave amplifiers. *IEEE Trans Microw. Theory Tech.*, MTT-16:761–766, 1968.

[45] Eugene N. Ivanov, M. E. Tobar, and R. A. Woode. Microwave interferometry: Application to precision measurements and noise reduction techniques. *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, 45:1526–1536, 1998.

[46] L. E. Richter, H. I. Mandelberg, M. S. Kruger, and P. A. McGrath. Linewidth determination from self-heterodyne measurements with subcoherence delay times. *IEEE J. Quantum Electron.*, QE-22:2070–2074, 1986.

[47] Tetsuhiko Ikegami, Shoichi Sudo, and Yoshihisa Sakai. *Frequency Stabilization of Semiconductor Laser Diodes*. Artech House, Boston, London, 1995.

[48] G. Kramer. Noise in passive frequency standards. In *CPEM 74; Conference on Precision Electromagnetic Measurements, 1-5 July London*, pages 157–159. IEE Conference Publication 113, 1974.

[49] Claude Audoin, Vincent Candelier, and Noel Dimarcq. A limit to the frequency stability of passive frequency standards due to an intermodulation effect. *IEEE Trans. Instrum. Meas.*, 40:121–125, 1991.

[50] G. J. Dick, J. Prestage, C. Greenhall, and L. Maleki. Local oscillator induced degradation of medium-term stability in passive atomic frequency standards. In *Proceedings of the 22nd Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, Vienna VA, USA*, pages 487–509, 1990.

[51] Giorgio Santarelli, Claude Audoin, Ala'a Makdissi, Philippe Laurent, G. John Dick, and André Clairon. Frequency stability degradation of an oscillator slaved to a periodically interrogated atomic resonator. *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, 45:887–894, 1998.

[52] Alain Joyet, Gaetano Mileti, Gregor Dudle, and Pierre Thomann. Theoretical study of the Dick effect in a continuously operated Ramsey resonator. *IEEE Trans. Instrum. Meas.*, 50:150–156, 2001.

[53] Charles A. Greenhall and G. John Dick. Local oscillator limited frequency stability for passive atomic frequency standards using square wave frequency modulation. *IEEE Trans. Ultrason. Ferroelec. Freq. Contr.*, 47:1593–1600, 2000.

[54] E. Philippot, Y. V. Pisarevsky, B. Capelle, and J. Détaint. Present state of the development of the piezoelectric materials. In *Proceedings of the 15th European Frequency and Time Forum*, pages 33–37, Rue Jaquet-Droz 1, Case postale 20, CH-2007 Neuchâtel, Switzerland, 2001. FSRM Swiss Foundation for Research in Microtechnology.

[55] Raymond A. Heising, editor. *Quartz Crystals for Electrical Circuits*. Van Nostrand, New York, 1947.

[56] R. J. Besson. A new "electrodeless" resonator design. In *Proceedings of the 31st Annual Symposium on Frequency Control*, pages 147–152, Fort Monmouth, New Jersey, 1977. U.S. Army Electronics Command.

[57] R. J. Besson, M. Mourey, S. Galliou, F. Marionnet, F. Gonzalez, P. Guillemot, R. Tjoelker, W. Diener, and A. Kirk. 10 MHz hyperstable quartz oscillators performances. In *Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum and The IEEE International Frequency Control Symposium*, pages 326–330, 26 Chemin de l'Epitaphe, 25030 BESANCON CEDEX - FRANCE, 1999. EFTF co/Société Francaise des Microtechniques et de Chronométrie (SFMC).

[58] John David Jackson. *Classical Electrodynamics*. John Wiley & Sons, New York, third edition, 1998.

[59] M. Abramowitz and I. A. Stegun, editors. *Handbook of Mathematical Functions*. Dover Publications, New York, 1968.

[60] J. P. Turneaure and S. R. Stein. An experimental limit on the time variation of the fine structure constant. In J. H. Sanders and A. H. Wapstra, editors, *Atomic Masses and Fundamental Constants*, volume 5, pages 636–642. Plenum Press, New York, London, 1976.

[61] J. P. Turneaure, C. M. Will, B. F. Farrell, E. M. Mattison, and R. F. C. Vessot. Test of the principle of equivalence by a null gravitational red-shift experiment. *Phys. Rev. D*, 27:1705–1714, 1983.

[62] S. Buchmann, J. P. Turneaure, J. A. Lippa, M. Dong, K. M. Cumbermack, and S. Wang. A superconducting microwave oscillator clock for use on the space station. In *Proceedings of the 52th Annual IEEE International Frequency Control Symposium, Pasadena, USA*, pages 534–539, 1998.

[63] John G. Hartnett and Michael E. Tobar. Frequency-temperature compensation techniques for high-Q microwave resonators. In A. N. Luiten, editor, *Frequency Measurement and Control*, volume 79 of *Topics in Applied Physics*, pages 67–91. Springer, Berlin, Heidelberg, New York, 2001.

[64] G. Santarelli, Ph. Laurent, P. Lemonde, A. Clairon, A. G. Mann, S. Chang, A. N. Luiten, and Ch. Salomon. Quantum projection noise in an atomic fountain: A high stability cesium frequency standard. *Phys. Rev. Lett.*, 82:4619–4622, 1999.

[65] Rabi T. Wang and G. John Dick. Cryocooled sapphire oscillator with ultrahigh stability. *IEEE Trans. Instrum. Meas.*, 48:528–531, 1999.

[66] Dana Z. Anderson, Josef C. Frisch, and Carl S. Masser. Mirror reflectometer based on optical cavity decay time. *Appl. Opt.*, 23:1238–1245, 1984.

[67] G. Rempe, R. J. Thompson, H. J. Kimble, and R. Lalezari. Measurement of ultralow losses in an optical interferometer. *Opt. Lett.*, 17:363–365, 1992.

[68] H. Kogelnik and T. Li. Laser beams and resonators. *Appl. Opt.*, 5:1550–1567, 1966.

[69] Anthony E. Siegman. *Lasers*. University Science Books, Mill Valley, California, 1986.

[70] J. Durnin. Exact solutions for nondiffracting beams. I. The scalar theory. *J. Opt. Soc. Am. A*, 4:651–654, 1987.

[71] I. Kimel and L. R. Elias. Relations between Hermite and Laguerre Gaussian modes. *IEEE J. Quantum Electron.*, 29:2562–2567, 1993.

[72] V. B. Braginsky, M. L. Gorodetsky, and V. S. Ilchenko. Quality-factor and nonlinear properties of optical whispering-gallery modes. *Phys. Lett A*, 137:393–397, 1989.

[73] M. L. Gorodetsky, A. A. Savchenkov, and V. S. Ilchenko. Ultimate Q of optical microsphere resonators. *Opt. Lett.*, 21:453–455, 1996.

[74] V. V. Vassiliev, V. L. Velichansky, V. S. Ilchenko, M. L. Gorodetsky, L. Hollberg, and A. V. Yarovitsky. Narrow-line-width diode laser with a high-Q microsphere resonator. *Opt. Commun.*, 158:305–312, 1998.

[75] M. L. Gorodetsky and V. S. Ilchenko. High-Q optical whispering-gallery microresonators: precession approach for spherical mode analysis and emission patterns with prism couplers. *Opt. Commun.*, 113:133–143, 1994.

[76] Mitchell H. Fields, Jürgen Popp, and Richard K. Chang. Nonlinear optics in microspheres. In Emil Wolf, editor, *Progress in Optics*, volume 41, pages 1–95. Elsevier, Amsterdam, 2000.

[77] V. S. Ilchenko, P. S. Volikov, V. L. Velichansky, F. Treussart, V. Lefèvre-Seguin, J.-M. Raimond, and S. Haroche. Strain-tunable high-Q optical microsphere resonator. *Opt. Commun.*, 145:86–90, 1998.

[78] Wolf von Klitzing, Romain Long, Vladimir S. Ilchenko, John Hare, and Valérie Lefèvre-Seguin. Frequency tuning of the whispering-gallery modes of silica microspheres for cavity quantum electrodynamics and spectroscopy. *Opt. Lett.*, 26:166–168, 2001.

[79] M. L. Gorodetsky and V. S. Ilchenko. Optical microsphere resonators: optimal coupling to high-Q whispering-gallery modes. *J. Opt. Soc. Am. B*, 16:147–154, 1999.

[80] Ming Cai and Kerry Vahala. Highly efficient optical power transfer to whispering-gallery modes by use of a symmetrical dual-coupling configuration. *Opt. Lett.*, 25:260–262, 2000.

[81] F. Bayer-Helms, H. Darnedde, and G. Exner. Längenstabilität bei Raumtemperatur von Proben der Glaskeramik "Zerodur' '. *Metrologia*, 21:49–57, 1985.

[82] J. Helmcke, J. J. Snyder, A. Morinaga, F. Mensing, and M. Gläser. New ultra-high resolution dye laser spectrometer utilizing a non-tunable reference resonator. *Appl. Phys. B*, 43:85–91, 1987.

[83] F. Riehle. Use of optical frequency standards for measurements of dimensional stability. *Meas. Sci. Technol.*, 9:1042–1048, 1998.

[84] L. Marmet, A. A. Madej, K. J. Siemsen, J. E. Bernard, and Bradley G. Whitford. Precision frequency measurement of the $^2S_{1/2}-^2D_{5/2}$ transition of Sr^+ with a 674-nm diode laser locked to an ultrastable cavity. *IEEE Trans. Instrum. Meas.*, 46:169–173, 1997.

[85] S. Seel, R. Storz, G. Ruoso, J. Mlynek, and S. Schiller. Cryogenic optical resonators: A new tool for laser frequency stabilization at the 1 Hz level. *Phys. Rev. Lett.*, 78:4741–4744, 1997.

[86] R. Storz, C. Braxmaier, K. Jäck, O. Pradl, and S. Schiller. Ultrahigh long-term dimensional stability of a sapphire cryogenic optical resonator. *Opt. Lett.*, 23:1031–1033, 1998.

[87] James D. Bjorken and Sidney Drell. *Relativistic Quantum Mechanics*. Mc Graw-Hill, New York, 1965.

[88] B. Cagnac. Progress on the Rydberg constant: The hydrogen atom as a frequency standard. *IEEE Trans. Instrum. Meas.*, 42:206–212, 1993.

[89] W. M. Itano, J. C. Bergquist, J. J. Bollinger, J. M. Gilligan, D. J. Heinzen, F. L. Moore, M. G. Raizen, and D. J. Wineland. Quantum projection noise: Population fluctuations in two-level systems. *Phys. Rev. A*, 47:3554–3570, 1993.

[90] J. Vanier and R. Larouche. A comparison of the wall shift of TFE and FEP teflon coatings in the hydrogen maser. *Metrologia*, 14:31–37, 1978.

[91] S. Bize, Y. Sortais, M. S. Santos, C. Mandache, A. Clairon, and C. Salomon. High-accuracy measurement of the ^{87}Rb ground-state hyperfine splitting in an atomic fountain. *Europhys. Lett.*, 45:558–564, 1999.

[92] Bureau International des Poids et Mesures, editor. *Comptes Rendus des séances de la 13^e CGPM*, Pavillon de Breteuil, F-92310 Sévres, France, 1967/1968. BIPM.

[93] M. Niering, R. Holzwarth, J. Reichert, P. Pokasov, Th. Udem, M. Weitz, T. W. Hänsch, P. Lemonde, G. Santarelli, M. Abgrall, P. Laurent, C. Salomon, and A. Clairon. Measurement of the hydrogen 1S – 2S transition frequency by phase coherent comparison with a microwave cesium fountain clock. *Phys. Rev. Lett.*, 84:5496–5499, 2000.

[94] Charlotte E. Moore. *Atomic Energy Levels*. Number 35/V.I in Nat. Stand. Ref. Data, Nat. Bur. Stand. (US). National Bureau of Standards, U.S. Government Printing Office, Washington D.C. 20402, 1971.

[95] T. J. Quinn. Practical realization of the definition of the metre (1997). *Metrologia*, 36:211–244, 1999.

[96] G. Ferrari, P. Cancio, R. Drullinger, G. Giusfredi, N. Poli, M. Prevedelli, C. Toninelli, and G. M. Tino. Precision frequency measurement of visible intercombination lines of strontium. *Phys. Rev. Lett.*, 91:243002–1–4, 2003.

[97] P. L. Larkins and P. Hannaford. Precision measurement of the energy of the $4d^9\ 5s^2\ ^2d_{5/2}$ metastable level in Ag I. *Z. Phys. D*, 32:167–172, 1994.

[98] Charlotte E. Moore. *Atomic Energy Levels*. Number 35/V.III in Nat. Stand. Ref. Data, Nat. Bur. Stand. (US). National Bureau of Standards, U.S. Government Printing Office, Washington D.C. 20402, 1971.

[99] J. L. Hall, M. Zhu, and P. Buch. Prospects for using laser-prepared atomic fountains for optical frequency standards applications. *J. Opt. Soc. Am. B*, 6:2194–2205, 1989.

[100] M. Walhout, U. Sterr, A. Witte, and S. L. Rolston. Lifetime of the metastable 6s'$[1/2]_0$ clock state in xenon. *Opt. Lett.*, 20:1192–1194, 1995.

[101] S. A. Webster, P. Taylor, M. Roberts, G. P. Barwood, P. Blythe, and P. Gill. A frequency standard using the $^2S_{1/2} - -^2F_{7/2}$ octupole transition in $^{171}Yb^+$. In Patrick Gill, editor, *Proceedings of the Sixth Symposium on Frequency Standards and Metrology*, pages 115–122, New Jersey, London, Singapore, Hong Kong, 2002. World Scientific.

[102] M. Weitz, A. Huber, F. Schmidt-Kaler, D. Leibfried, and T. W. Hänsch. Precision measurement of the hydrogen and deuterium 1S ground state Lamb shift. *Phys. Rev. Lett.*, 72:328–331, 1994.

[103] B. de Beauvoir, F. Nez, L. Julien, B. Cagnac, F. Biraben, D. Touahri, L. Hilico, O. Acef, A. Clairon, and J. J. Zondy. Absolute frequency measurement of the $2S - 8S/D$ transitions in hydrogen and deuterium: New determination of the Rydberg constant. *Phys. Rev. Lett.*, 78:440–443, 1997.

[104] G. Uhlenberg, J. Dirscherl, and H. Walther. Magneto-optical trapping of silver atoms. *Phys. Rev. A*, 62:063404–1–4, 2000.

[105] S. Guérandel, T. Badr, M. D. Plimmer, P. Juncar, and M. E. Himbert. Frequency measurement, isotope shift and hyperfine structure of the $4d^95s^2\ ^2D_{5/2} \rightarrow 4d^{10}6p\ ^2P_{3/2}$ transition in atomic silver. *Eur. Phys. J. D*, 10:33–38, 2000.

[106] C. H. Townes and A. L. Schawlow. *Microwave Spectroscopy*. Dover Publications, New York, 1975.

[107] Hermann Haken and Hans Christoph Wolf. *Molekülphysik und Quantenchemie*. Springer, Berlin, Heidelberg, New York, third edition, 1998.

[108] S. Gerstenkorn and P. Luc. Atlas du spectre d'absorption de la molécule d'iode; 14 000 cm^{-1} – 15 600 cm^{-1} (1978); 15 600 cm^{-1} – 17 600 cm^{-1} (1977); 17 500 cm^{-1} – 20 000 cm^{-1} (1977). Technical report, Laboratoire Aimé–Cotton CNRS II, Centre National de la Recherche Scientifique, 15, quai Anatole-France, 75700 Paris, 1977-1978.

[109] S. Gerstenkorn, J. Verges, and J. Chevillard. Atlas du spectre d'absorption de la molécule d'iode; 11 000 cm^{-1} – 14 000 cm^{-1} (1982). Technical report, Laboratoire Aimé–Cotton CNRS II, Centre National de la Recherche Scientifique, 15, quai Anatole-France, 75700 Paris, 1977-1978.

[110] H. Kato. Doppler-free high resolution spectral atlas of iodine molecule. Technical report, Japan Society for the Promotion of Science, 2000.

[111] B. Bodermann, H. Knöckel, and E. Tiemann. Widely usable interpolation formulae for hyperfine splittings in the $^{127}I_2$ spectrum. *Eur. Phys. J. D*, 19:31–44, 2002.

[112] H. Knöckel, B. Bodermann, and E. Tiemann. High precision description of the rovibronic structure of the I_2 B–X spectrum. *Eur. Phys. J. D*, 28:199–209, 2004.

[113] J. L. Dunham. The energy levels of a rotating vibrator. *Phys. Rev.*, 41:721–731, 1932.

[114] S. Gerstenkorn and P. Luc. Description of the absorption spectrum of iodine recorded by means of Fourier transform spectroscopy: the (B-X) system. *J. Physique*, 46:867–881, 1986.

[115] C. R. Vidal. Accurate determination of potential energy curves. *Comments At. Mol. Phys.*, 17:173–197, 1986.

[116] M. Broyer, J. Vigué, and J. C. Lehmann. Effective hyperfine Hamiltonian in homonuclear diatomic molecules. Application to the B state of molecular Iodine. *J. de Physique*, 39:591–609, 1978.

[117] M. Gläser. Hyperfine components of iodine for optical frequency standards. PTB–Bericht Opt-25, Physikalisch–Technische Bundesanstalt, Braunschweig, 1987.

[118] Mark L. Eickhoff and J. L. Hall. Optical frequency standard at 532 nm. *IEEE Trans. Instrum. Meas.*, 44:155–158, 1995.

[119] K. Hedfeld and R. Mecke. Das Rotationsschwingungsspektrum des Acetylens. I. *Z. Phys.*, 64:151–161, 1930.

[120] W. H. J. Childs and R. Mecke. Das Rotationsschwingungsspektrum des Acetylens. II. *Z. Phys.*, 64:162–172, 1930.

[121] R. Mecke. Das Rotationsschwingungsspektrum des Acetylens. III. *Z. Phys.*, 64:173–185, 1930.

[122] E. K. Plyler, E. D. Tidwell, and T. A. Wiggins. Rotation-vibration constants of Acetylene. *J. Opt. Soc. Am.*, 53:589–593, 1963.

[123] A. Baldacci, S. Ghersetti, and K. Narahari Rao. Interpretation of the Acetylene spectrum at 1.5 μm. *J. Mol. Spectrosc.*, 68:183–194, 1977.

[124] K. Nakagawa, M. de Labachelerie, Y. Awaji, and M. Kourogi. Accurate optical frequency atlas of the 1.5-μm bands of acetylene. *J. Opt. Soc. Am. B*, 13:2708–2714, 1996.

[125] J. Cariou and P. Luc. Atlas du spectre d'absorption de la molecule de Tellure; partie1: 17 500 – 20300 cm^{-1}, temperature: 680 °C. Technical report, Laboratoire Aimé–Cotton CNRS II, Centre National de la Recherche Scientifique, 15, quai Anatole-France, 75700 Paris, 1980.

[126] J. Cariou and P. Luc. Atlas du spectre d'absorption de la molecule de Tellure; partie3: 20 900 -23 700 cm^{-1}, temperature: 600 °c. Technical report, Laboratoire Aimé–Cotton CNRS II, Centre National de la Recherche Scientifique, 15, quai Anatole-France, 75700 Paris, 1980.

[127] G. P. Barwood, W. R. C. Rowley, P. Gill, J. L. Flowers, and B. W. Petley. Interferometric measurements of $^{130}Te_2$ reference frequencies for 1S–2S transitions in hydrogenlike atoms. *Phys. Rev. A*, 43:4783–4790, 1991.

[128] Ph. Courteille, L. S. Ma, W. Neuhauser, and R. Blatt. Frequency measurement of $^{130}Te_2$ resonances near 467 nm. *Appl. Phys. B*, 59:187–193, 1994.

[129] Y. Awaji, K. Nakagawa, M. de Labachelerie, M. Ohtsu, and H. Sasada. Optical frequency measurement of the $H^{12}C^{14}N$ Lamb-dip-stabilized 1.5-μm diode laser. *Opt. Lett.*, 20:2024–2026, 1995.

[130] B. Bodermann, M. Klug, H. Knöckel, E. Tiemann, T. Trebst, and H. R. Telle. Frequency measurement of I_2 lines in the NIR using Ca and CH_4 optical frequency standards. *Appl. Phys. B*, 67:95–99, 1998.

[131] M. Roberts, P. Taylor, G. P. Barwood, P. Gill, H. A. Klein, and W. R. C. Rowley. Observation of an electric octupole transition in a single ion. *Phys. Rev. Lett.*, 78:1876–1879, 1997.

[132] Harold J. Metcalf and Peter van der Straten. *Laser Cooling and Trapping*. Springer, New York, Berlin, Heidelberg, 1999.

[133] Pierre Meystre and Murray Sargent III. *Elements of Quantum Optics*. Springer, Berlin, Heidelberg, New York, second edition, 1991.

[134] Harald Schnatz and Friedhelm Mensing. Iodine-stablized, frequency-doubled Nd:YAG lasers at $\lambda = 532$ nm; design and performance. In John L. Hall and Jun Ye, editors, *Proceedings of SPIE: Laser Frequency Stabilization, Standards, Measurement, and Applications*, volume 4269, pages 239–247, P.O. Box 10, Bellingham, Washington 98227–0010 USA, 2001. SPIE.

[135] Marlan O. Scully and M. Suhail Zubairy. *Quantum Optics*. Cambridge University Press, Cambridge, New York, Melbourne, Madrid, 1997.

[136] A. R. Edmonds. *Angular momentum in quantum mechanics*. Princeton University Press, Princeton, New Jersey, 1957.

[137] F. Bloch and A. Siegert. Magnetic resonance for nonrotating fields. *Phys. Rev.*, 57:522–527, 1940.

[138] L. Allen and J. H. Eberly. *Optical Resonance and Two-Level Atoms*. Dover Publications Inc., New York, 1987.

[139] Richard P. Feynman, Frank L. Vernon, Jr., and Robert W. Hellwarth. Geometrical representation of the Schrödinger equation for solving maser problems. *J. Appl. Phys.*, 28:49–52, 1957.

[140] I. I. Rabi, N. F. Ramsey, and J. Schwinger. Use of rotating coordinates in magnetic resonance problems. *Rev. Mod. Phys.*, 26:167–171, 1954.

[141] Alfred Kastler. Quelques suggestions concernant la production optique et la détection optique d'une inégalité de population des niveaux de quantification spatiale des atomes. Application a l'expérience de Stern et Gerlach et a la résonance magnétique. *J. Phys. Radium*, 11:255–265, 1950.

[142] E. Arimondo. Coherent population trapping in laser spectroscopy. In E. Wolf, editor, *Progress in Optics*, volume XXXV, pages 257–354. Elsevier, Amsterdam, 1996.

[143] G. Janik, W. Nagourney, and H. Dehmelt. Doppler-free optical spectroscopy on the Ba^+ mono-ion oscillator. *J. Opt. Soc. Am. B*, 2:1251–1257, 1985.

[144] P. L. Kelley, P. J. Harshman, O. Blum, and T. K. Gustafson. Radiative renormalization analysis of optical double resonance. *J. Opt. Soc. Am. B*, 11:2298–2302, 1994.

[145] Yves Stahlgies. Lichtverschiebung und Fano-Resonanzen in einem einzelnen Ba^+-Ion. Master's thesis, Universität Hamburg, 1993.

[146] R. C. Hilborn. Einstein coefficients, cross sections, f values, dipole moments, and all that. *Am. J. Phys.*, 50:982–986, 1982. Erratum in: Am. J. Phys., **51** (1983), 4710.

[147] R. E. Walkup, A. Spielfiedel, and D. E. Pritchard. Observation of non-Lorentzian spectral line shapes in Na-noble-gas systems. *Phys. Rev. Lett.*, 45:986–989, 1980.

[148] Rudolf Grimm, Matthias Weidemüller, and Yurii B. Ovchinnikov. Optical dipole traps for neutral atoms. *Adv. At. Mol. Opt. Phys.*, 42:95–170, 2000.

[149] A. L. Schawlow and C. H. Townes. Infrared and optical masers. *Phys. Rev.*, 112:1940–1949, 1958.

[150] John W. Farley and William H. Wing. Accurate calculation of dynamic Stark shifts and depopulation rates of Rydberg energy levels induced by blackbody radiation. Hydrogen, helium, and alkali-metal atoms. *Phys. Rev. A*, 23:2397–2424, 1981.

[151] S. L. Rolston and W. D. Phillips. Laser-cooled neutral atom frequency standards. *Proceedings IEEE*, 79:943–951, 1991.

[152] R. H. Dicke. The effect of collisions upon the Doppler width of spectral lines. *Phys. Rev.*, 89:472–473, 1953.

[153] R. H. Romer and R. H. Dicke. New technique for high-resolution microwave spectroscopy. *Phys. Rev.*, 99:532–536, 1955.

[154] Stephan Briaudeau, Solomon Saltiel, Gerard Nienhuis, Daniel Bloch, and Martial Ducloy. Coherent Doppler narrowing in a thin vapor cell: Observation of the Dicke regime in the optical domain. *Phys. Rev. A*, 57:R3169–R3172, 1998.

[155] N. F. Ramsey. *Molecular Beams*. Clarendon Press, Oxford, 1956.

[156] P. Jacquinot. Atomic beam spectroscopy. In K. Shimoda, editor, *High-Resolution Laser Spectroscopy*, volume 13 of *Topics in Applied Physics*, pages 52–93. Springer, Berlin, Heidelberg, New York, 1976.

[157] A. Huber, B. Gross, M. Weitz, and T. W. Hänsch. High-resolution spectroscopy of the $1S - 2S$ transition in atomic hydrogen. *Phys. Rev. A*, 59:1844–1851, 1999.

[158] D. J. Wineland and Wayne M. Itano. Laser cooling of atoms. *Phys. Rev. A*, 20:1521–1540, 1979.

[159] T. W. Hänsch and A. L. Schawlow. Cooling of gases by laser radiation. *Opt. Commun.*, 13:68–69, 1975.

[160] D. Wineland and H. Dehmelt. Proposed $10^{14}\ \Delta\nu/\nu$ laser fluorescence spectroscopy on Tl^+ mono-ion oscillator III. *Bull. Am. Phys. Soc.*, 20:637, 1975.

[161] S. Chu, L. Hollberg, J. E. Bjorkholm, A. Cable, and A. Ashkin. Three-dimensional viscous confinement and cooling of atoms by resonance radiation pressure. *Phys. Rev. Lett.*, 55:48–51, 1985.

[162] P. D. Lett, W. D. Phillips, S. L. Rolston, C. E. Tanner, R. N. Watts, and C. I. Westbrook. Optical molasses. *J. Opt. Soc. Am. B*, 6:2084–2107, 1989.

[163] Hidetoshi Katori, Tetsuya Ido, Yoshitomo Isoya, and Makoto Kuwata-Gonokami. Magneto-optical trapping and cooling of strontium atoms down to the photon recoil temperature. *Phys. Rev. Lett.*, 82:1116–1119, 1999.

[164] T. Binnewies, G. Wilpers, U. Sterr, F. Riehle, J. Helmcke, T. E. Mehlstäubler, E. M. Rasel, and W. Ertmer. Doppler cooling and trapping on forbidden transitions. *Phys. Rev. Lett.*, 87:123002–1–4, 2001.

[165] E. A. Curtis, C. W. Oates, and L. Hollberg. Quenched narrow-line laser cooling of ^{40}Ca to near the photon recoil limit. *Phys. Rev. A*, 64:031403(R)–1–4, 2001.

[166] A. Aspect, E. Arimondo, R. Kaiser, N. Vansteenkiste, and C. Cohen-Tannoudji. Laser cooling below the one-photon recoil energy by velocity-selective coherent population trapping. *Phys. Rev. Lett.*, 61:826–829, 1988.

[167] T. Esslinger, F. Sander, M. Weidemüller, A. Hemmerich, and T. W. Hänsch. Subrecoil laser cooling with adiabatic transfer. *Phys. Rev. Lett.*, 76:2432–2435, 1996.

[168] M. Kasevich and S. Chu. Laser cooling below a photon recoil with three-level atoms. *Phys. Rev. Lett.*, 69:1741–1744, 1992.

[169] Jonathan D. Weinstein, Robert de Carvalho, Thierry Guillet, Bretislav Friedrich, and John M. Doyle. Magnetic trapping of calcium monohydride molecules at millikelvin temperatures. *Nature*, 395:148–150, 1998.

[170] Hendrick L. Bethlem, Giel Berden, Floris M. H. Crompvoets, Rienk T. Jongma, André J. A. van Roij, and Gerard Meijer. Electrostatic trapping of ammonia molecules. *Nature*, 406:491–494, 2000.

[171] Hendrick L. Bethlem, Giel Berden, André J. A. van Roij, Floris M. H. Crompvoets, and Gerard Meijer. Trapping neutral molecules in a traveling potential well. *Phys. Rev. Lett.*, 84:5744–5747, 2000.

[172] A. Fioretti, D. Comparat, A. Crubellier, O. Dulieu, F. Masnou-Seeuws, and P. Pillet. Formation of cold Cs_2 molecules through photoassociation. *Phys. Rev. Lett.*, 80:4402–4405, 1998.

[173] T. Takekoshi, B. M. Patterson, and R. J. Knize. Observation of optically trapped cold Cesium molecules. *Phys. Rev. Lett.*, 81:5105–5108, 1998.

[174] A. N. Nikolov, J. R. Ensher, E. E. Eyler, H. Wang, W. C. Stwalley, and P. L. Gould. Efficient production of ground-state potassium molecules at sub-mK temperatures by two-step photoassociation. *Phys. Rev. Lett.*, 84:246–249, 2000.

[175] Roahn Wynar, R. S. Freeland, D. J. Han, C. Ryu, and D. J. Heinzen. Molecules in a Bose–Einstein condensate. *Science*, 287:1016–1019, 2000.

[176] William H. Wing. On neutral particle trapping in quasistatic electromagnetic fields. *Prog. Quant. Electr.*, 8:181–199, 1984.

[177] W. Ketterle and D. E. Pritchard. Trapping and focusing ground state atoms with static fields. *Appl. Phys. B*, 54:403–406, 1992.

[178] A. Ashkin and J. P. Gordon. Stability of radiation-pressure particle traps: an optical Earnshaw theorem. *Opt. Lett.*, 8:511–513, 1983.

[179] Alan L. Migdall, John V. Prodan, William D. Phillips, Thomas H. Bergeman, and Harold J. Metcalf. First observation of magnetically trapped neutral atoms. *Phys. Rev. Lett.*, 54:2596–2599, 1985.

[180] William H. Wing. Electrostatic trapping of neutral atomic particles. *Phys. Rev. Lett.*, 45:631–634, 1980.

[181] T. Bergeman, Gidon Erez, and Harold J. Metcalf. Magnetostatic trapping fields for neutral atoms. *Phys. Rev. A*, 35:1535–1546, 1987.

[182] Ettore Majorana. Atomi orientati in campo magnetico variabile. *Il Nuovo Cimento*, 9:43–50, 1932.

[183] Wolfgang Petrich, Michael H. Anderson, Jason R. Ensher, and Eric A. Cornell. Stable, tightly confining magnetic trap for evaporative cooling of neutral atoms. *Phys. Rev. Lett.*, 74:3352–3355, 1995.

[184] J. D. Miller, R. A. Cline, and D. J. Heinzen. Far-off-resonance optical trapping of atoms. *Phys. Rev. A*, 47:R4567–R4570, 1993.

[185] Nir Davidson, Heun Jin Lee, Charles S. Adams, Mark Kasevich, and Steven Chu. Long atomic coherence times in an optical dipole trap. *Phys. Rev. Lett.*, 74:1311–1314, 1995.

[186] H. J. Lee, C. S. Adams, M. Kasevich, and S. Chu. Raman cooling of atoms in an optical dipole trap. *Phys. Rev. Lett.*, 76:2658–2661, 1996.

[187] E. L. Raab, M. Prentiss, A. Cable, S. Chu, and D. E. Pritchard. Trapping of neutral sodium atoms with radiation pressure. *Phys. Rev. Lett.*, 59:2631–2634, 1987.

[188] A. M. Steane, M. Chowdhury, and C. J. Foot. Radiation force in the magneto-optical trap. *J. Opt. Soc. Am. B*, 9:2142–2158, 1992.

[189] C. Monroe, W. Swann, H. Robinson, and C. Wieman. Very cold trapped atoms in a vapor cell. *Phys. Rev. Lett.*, 65:1571–1574, 1990.

[190] G. Zinner. Ein optisches Frequenznormal auf der Basis lasergekühlter Calciumatome. PTB-Bericht PTB-Opt–58, Physikalisch-Technische Bundesanstalt, Braunschweig, 1998.

[191] C. W. Oates, F. Bondu, R. W. Fox, and L. Hollberg. A diode-laser optical frequency standard based on laser-cooled Ca atoms: Sub-kilohertz spectroscopy by optical shelving detection. *Eur. Phys. J. D*, 7:449–460, 1999.

[192] B. P. Anderson and M. A. Kasevich. Enhanced loading of a magneto-optic trap from an atomic beam. *Phys. Rev. A*, 50:R3581–3584, 1994.

[193] W. D. Phillips and H. Metcalf. Laser deceleration of an atomic beam. *Phys. Rev. Lett.*, 48:596–599, 1982.

[194] William D. Phillips, Jon V. Prodan, and Harold J. Metcalf. Laser cooling and electromagnetic trapping of neutral atoms. *J. Opt. Soc. Am. B*, 2:1751–1767, 1985.

[195] T. Kurosu, M. Morinaga, and F. Shimizu. Observation of the Ca $4s^1S_0$-$4p^3P_1$ transition in continuous free-falling cold atomic flow from an atom trap. *Jpn. J. Appl. Phys.*, 31:L273–L275, 1992.

[196] Th. Kisters, K. Zeiske, F. Riehle, and J. Helmcke. High-resolution spectroscopy with laser-cooled and trapped calcium atoms. *Appl. Phys. B*, 59:89–98, 1994.

[197] F. Riehle, H. Schnatz, B. Lipphardt, G. Zinner, P. Kersten, and J. Helmcke. Optical frequency standard based on laser-cooled Ca atoms. In J. C. Bergquist, editor, *Proceedings of the Fifth Symposium on Frequency Standards and Metrology*, pages 277–282, Singapore, New Jersey, London, Hong Kong, 1996. World Scientific.

[198] R. L. Cavasso-Filho, D. A. Manoel, D. R. Ortega, A. Scalabrin, D. Pereira, and F. C. Cruz. Optical frequency standards based on cold calcium atoms. In P. Gill, editor, *Frequency Standards and Metrology, Proceedings of the Sixth Symposium*, pages 546–548, Singapore, 2002. World Scientific.

[199] A. Witte, Th. Kisters, F. Riehle, and J. Helmcke. Laser cooling and deflection of a calcium atomic beam. *J. Opt. Soc. Am. B*, 9:1030–1037, 1992.

[200] A. Hemmerich and T. W. Hänsch. Two-dimensional atomic crystal bound by light. *Phys. Rev. Lett.*, 70:410–413, 1993.

[201] G. Grynberg, B. Lounis, P. Verkerk, J.-Y. Courtois, and C. Salomon. Quantized motion of cold cesium atoms in two- and three-dimensional optical potentials. *Phys. Rev. Lett.*, 70:2249–2252, 1993.

[202] Marshall T. DePue, Colin McCormick, S. Lukman Winoto, Steven Oliver, and David S. Weiss. Unity occupation of sites in a 3D optical lattice. *Phys. Rev. Lett.*, 82:2262–2265, 1999.

[203] Hidetoshi Katori. Spectroscopy of strontium atoms in the Lamb-Dicke confinement. In P. Gill, editor, *Proceedings of the Sixth Symposium on Frequency Standards and Metrology*, pages 323–330, Singapore, 2002. World Scientific.

[204] W. Ketterle, D.S. Durfee, and D.M. Stamper-Kurn. Making, probing and understanding Bose-Einstein condensates. In *Bose-Einstein condensation in atomic gases*, volume Course CXL of *Proceedings Internat. School of Physics "Enrico Fermi"*, pages 67–176, Amsterdam, Oxford, Tokyo, Washington DC, 1999. IOS Press.

[205] Dieter Meschede. *Optics, Light and Lasers*. Wiley-VCH, Weinheim–New York, 2004.

[206] Paul D. Lett, Richards N. Watts, Christoph I. Westbrook, William D. Phillips, Phillip L. Gould, and Harold J. Metcalf. Observation of atoms laser cooled below the Doppler limit. *Phys. Rev. Lett.*, 61:169–172, 1988.

[207] A. S. Arnold and P. J. Manson. Atomic density and temperature distributions in magneto-optical traps. *J. Opt. Soc. Am. B*, 17:497–506, 2000.

[208] A. Kastberg, W. D. Phillips, S. L. Rolston, R. J. C. Spreeuw, and P. S. Jessen. Adiabatic cooling of Cesium to 700 nK in an optical lattice. *Phys. Rev. Lett.*, 74:1542–1545, 1995.

[209] A. M. Steane and C. J. Foot. Laser cooling below the Doppler limit in a magneto-optical trap. *Europhys. Lett.*, 14:231–236, 1991.

[210] P. Kohns, P. Buch, W. Süptitz, C. Csambal, and W. Ertmer. On-line measurement of sub-Doppler temperatures in a Rb magneto-optical trap by trap centre oscillations. *Europhys. Lett.*, 22:517–522, 1993.

[211] Mark Kasevich, David S. Weiss, Erling Riis, Kathryn Moler, Steven Kasapi, and Steven Chu. Atomic velocity selection using stimulated Raman transitions. *Phys. Rev. Lett.*, 66:2297–2300, 1991.

[212] V. S. Letokhov. Saturation spectroscopy. In K. Shimoda, editor, *High-Resolution Laser Spectroscopy*, volume 13 of *Topics in Applied Physics*, pages 95–171. Springer, Berlin, Heidelberg, New York, 1976.

[213] W. R. Bennett, Jr. Hole burning effects in a He-Ne optical maser. *Phys. Rev.*, 126:580–593, 1962.

[214] Christian Bordé. Progress in understanding sub-Doppler lineshapes. In J. L. Hall and J. L. Carlsten, editors, *Laser Spectroscopy III*, volume 21 of *Springer Series in Optical Sciences*, pages 121–134, Berlin, 1977. Springer.

[215] Ch. J. Bordé, Ch. Salomon, S. Avrillier, A. Van Lerberghe, Ch. Bréant, D. Bassi, and G. Scoles. Optical Ramsey fringes with traveling waves. *Phys. Rev. A*, 30:1836–1848, 1984.

[216] J. Ishikawa, F. Riehle, J. Helmcke, and Ch. J. Bordé. Strong-field effects in coherent saturation spectroscopy of atomic beams. *Phys. Rev. A*, 49:4794–4825, 1994.

[217] J. L. Hall, Ch. J. Bordé, and K. Uehara. Direct optical resolution of the recoil effect using saturated absorption spectroscopy. *Phys. Rev. Lett.*, 37:1339–1342, 1976.

[218] F. Riehle, J. Ishikawa, and J. Helmcke. Suppression of a recoil component in nonlinear Doppler-free spectroscopy. *Phys. Rev. Lett.*, 61:2092–2095, 1988.

[219] F. Riehle, A. Witte, Th. Kisters, and J. Helmcke. Interferometry with Ca atoms. *Appl. Phys. B*, 54:333–340, 1992.

[220] T. Kurosu and A. Morinaga. Suppression of the high-frequency recoil component in optical Ramsey-fringe spectroscopy. *Phys. Rev. A*, 45:4799–4802, 1992.

[221] S. N. Bagayev, V. P. Chebotayev, A. K. Dmitriyev, A. E. Om, Y. V. Nekrasov, and B. N. Skvortsov. Second-order Doppler-free spectroscopy. *Appl. Phys.*, B 52:63–66, 1991.

[222] C. Chardonnet, F. Guernet, G. Charton, and Ch. J. Bordé. Ultrahigh-resolution saturation spectroscopy using slow molecules in an external cell. *Appl. Phys. B*, 59:333–343, 1994.

[223] L. S. Vasilenko, V. P. Chebotaev, and A. V. Shishaev. Line shape of two-photon absorption in a standing-wave field in a gas. *JETP Lett.*, 12:113–116, 1970.

[224] N. Bloembergen and M. D. Levenson. Doppler-free two-photon absorption spectroscopy. In K. Shimoda, editor, *High-Resolution Laser Spectroscopy*, volume 13 of *Topics in Applied Physics*, pages 315–369. Springer, Berlin, Heidelberg, New York, 1976.

[225] Malcolm Geoffrey Boshier. *Precise Laser Spectroscopy of the Hydrogen 1S – 2S Transition*. PhD thesis, University of Oxford, 1988.

[226] F. Biraben, M. Bassini, and B. Cagnac. Line-shapes in Doppler-free two-photon spectroscopy. The effect of finite transit time. *J. Physique*, 40:445–455, 1979.

[227] N. F. Ramsey. A molecular beam resonance method with separated oscillating fields. *Phys. Rev.*, 78:695–699, 1950.

[228] Norman F. Ramsey. Molecular beam resonances in oscillatory fields of nonuniform amplitudes and phases. *Phys. Rev.*, 109:822–825, 1958.

[229] G. Kramer. Linear optical "Ramsey" resonance by means of a spacially modulated molecular beam. *J. Opt. Soc. Am.*, 68:1634–1635, 1978.

[230] Ye. V. Baklanov, B. Ya. Dubetsky, and V. P. Chebotayev. Non-linear Ramsey resonance in the optical region. *Appl. Phys.*, 9:171–173, 1976.

[231] G. Kramer, C. O. Weiss, and B. Lipphardt. Coherent frequency measurements of the hfs-resolved methane line. In A. De Marchi, editor, *Frequency Standards and Metrology*, pages 181–186, Springer, Berlin, Heidelberg, New York, 1989.

[232] Ye. V. Baklanov, V. P. Chebotayev, and B. Ya. Dubetsky. The resonance of two-photon absorption in separated optical fields. *Appl. Phys.*, 11:201–202, 1976.

[233] M. M. Salour and C. Cohen-Tannoudji. Observation of Ramsey's interference fringes in the profile of Doppler-free two-photon resonances. *Phys. Rev. Lett.*, 38:757–760, 1977.

[234] S.-A. Lee, J. Helmcke, and J. L. Hall. High-resolution two-photon spectroscopy of Rb Rydberg levels. In H. Walther and K. W. Rothe, editors, *Laser Spectroscopy IV*, volume 21 of *Springer Series in Optical Sciences*, pages 130–141, Berlin, 1979. Springer.

[235] A. Huber, B. Gross, M. Weitz, and T. W. Hänsch. Two-photon optical Ramsey spectroscopy of the $1S-2S$ transition in atomic hydrogen. *Phys. Rev. A*, 58:R2631–R2634, 1998.

[236] J. C. Bergquist, S. A. Lee, and J. L. Hall. Saturated Absorption with Spatially Separated Laser Fields: Observation of Optical "Ramsey" Fringes. *Phys. Rev. Lett.*, 38:159–162, 1977.

[237] R. L. Barger, J. C. Bergquist, T. C. English, and D. J. Glaze. Resolution of photon-recoil structure of the 6573-Å calcium line in an atomic beam with optical Ramsey fringes. *Appl. Phys. Lett.*, 34:850–852, 1979.

[238] R. L. Barger. Influence of second-order Doppler effect on optical Ramsey fringe profiles. *Opt. Lett.*, 6:145–147, 1981.

[239] M. Baba and K. Shimoda. Observation of Ramsey resonance absorption in three separated laser fields produced by a corner reflector. *Appl. Phys.*, 24:11–12, 1981.

[240] Ch. J. Bordé, S. Avrillier, A. van Lerberghe, Ch. Salomon, Ch. Breant, D. Bassi, and G. Scoles. Observation of optical Ramsey fringes in the 10 μm spectral region using a supersonic beam of SF_6. *Appl. Phys. B*, 28:82–83, 1982.

[241] J. Helmcke, D. Zevgolis, and B. Ü. Yen. Observation of high contrast, ultra narrow optical Ramsey fringes in saturated absorption utilizing four interaction zones of travelling waves. *Appl. Phys. B*, 28:83–84, 1982.

[242] Ch. J. Bordé. Atomic interferometry with internal state labelling. *Phys. Lett. A*, 140:10–12, 1989.

[243] Ch. J. Bordé. Atomic interferometry and laser spectroscopy. In M. Ducloy, E. Giacobino, and G. Camy, editors, *Laser Spectroscopy*, pages 239–245, Singapore, 1992. World Scientific.

[244] U. Sterr, K. Sengstock, J. H. Müller, D. Bettermann, and W. Ertmer. The magnesium Ramsey interferometer: Applications and prospects. *Appl. Phys. B*, 54:341–346, 1992.

[245] U. Sterr, K. Sengstock, W. Ertmer, F. Riehle, and J. Helmcke. Atom interferometry based on separated light fields. In P. Berman, editor, *Atom Interferometry*, pages 293–362, San Diego, 1997. Academic Press.

[246] F. Riehle, Th. Kisters, A. Witte, J. Helmcke, and Ch. J. Bordé. Optical Ramsey spectroscopy in a rotating frame: Sagnac effect in a matter-wave interferometer. *Phys. Rev. Lett.*, 67:177–180, 1991.

[247] J. Jacobsen, G. Björk, and Y. Yamamoto. Quantum limit for the atom-light interferometer. *Appl. Phys. B*, 60:187–191, 1995.

[248] H. Hinderthür, F. Ruschewitz, H.-J. Lohe, S. Lechte, K. Sengstock, and W. Ertmer. Time-domain high-finesse atom interferometry. *Phys. Rev. A*, 59:2216–2219, 1999.

[249] A. Bauch, B. Fischer, T. Heindorff, and R. Schröder. Performance of the PTB reconstructed primary clock CS1 and an estimate of its current uncertainty. *Metrologia*, 35:829–845, 1998.

[250] A. Bauch, B. Fischer, T. Heindorff, P. Hetzel, R. Schröder, and P. Wolf. Comparisons of the PTB primary clocks with TAI in 1999. *Metrologia*, 37:683–692, 2000.

[251] A. De Marchi, J. Shirley, D. J. Glaze, and R. Drullinger. A new cavity configuration for cesium beam primary frequency standards. *IEEE Trans. Instrum. Meas.*, 37:185–190, 1988.

[252] Robert E. Drullinger, David J. Glaze, J. L. Lowe, and Jon H. Shirley. The NIST optically pumped Cesium frequency standard. *IEEE Trans. Instrum. Meas.*, 40:162–164, 1991.

[253] W. D. Lee, R. E. Drullinger, J. H. Shirley, C. Nelson, D. A. Jennings, L. O. Mullen, F. L. Walls, T. E. Parker, A. Hasegawa, K. Fukuda, N. Kotake, M. Kajita, and T. Morikawa. Accuracy evaluations and frequency comparisons of NIST-7 and CRL-01. In *Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum and The IEEE International Frequency Control Symposium*, pages 62–65, 1999.

[254] A. De Marchi, G. D. Rovera, and A. Premoli. Pulling by neighbouring transitions and its effects on the performance of caesium-beam frequency standards. *Metrologia*, 20:37–47, 1984.

[255] Ho Seong Lee, Taeg Yong Kwon, Hoon-Soo Kang, Young-Ho Park, Cha-Hwan Oh, Sang Eon Park, Hyuck Cho, and V. G. Minogin. Comparison of the Rabi and Ramsey pulling in an optically pumped caesium-beam standard. *Metrologia*, 40:224–231, 2003.

[256] L. S. Cutler, C. A. Flory, R. P. Giffard, and A. De Marchi. Frequency pulling by hyperfine σ transitions in cesium beam atomic frequency standards. *J. Appl. Phys.*, 69:2780–2792, 1991.

[257] Andreas Bauch and Roland Schröder. Frequency shifts in a cesium atomic clock due to Majorana transitions. *Ann. Physik*, 2:421–449, 1993.

[258] Wayne M. Itano, L. L. Lewis, and D. J. Wineland. Shift of $^2S_{1/2}$ hyperfine splittings due to blackbody radiation. *Phys. Rev. A*, 25:1233–1235, 1982.

[259] V. G. Pal'chikov, Yu. S. Domnin, and A. V. Novoselov. Black-body radiation effects and light shifts in atomic frequency standards. *J. Opt. B: Quantum Semiclass. Opt.*, 5:S131–S135, 2003.

[260] A. Bauch and R. Schröder. Experimental verification of the shift of the Cesium hyperfine transition frequency due to blackbody radiation. *Phys. Rev. Lett.*, 78:622–625, 1997.

[261] E. Simon, P. Laurent, and A. Clairon. Measurement of the Stark shift of the Cs hyperfine splitting in an atomic fountain. *Phys. Rev. A*, 57:436–439, 1998.

[262] Nikolaos K. Pavlis and Marc A. Weiss. The relativistic redshift with 3×10^{-17} uncertainty at NIST, Boulder, Colorado, USA. *Metrologia*, 40:66–73, 2003.

[263] B. Guinot. Application of general relativity to metrology. *Metrologia*, 34:261–290, 1997.

[264] Jon H. Shirley. Velocity distributions calculated from the Fourier transforms of Ramsey lineshapes. *IEEE Trans. Instrum. Meas.*, 46:117–121, 1997.

[265] Ala'a Makdissi and Emeric de Clercq. A signal approach analysis of the Ramsey pattern in Cesium beam frequency standards. *IEEE Trans. Instrum. Meas.*, 46:112–116, 1997.

[266] Claude Audoin and Bernard Guinot. *The Measurement of Time: Time, Frequency and the Atomic Clock*. Cambridge University Press, Cambridge, New York, 2001.

[267] Jon H. Shirley, W. D. Lee, G. D. Rovera, and R. E. Drullinger. Rabi pedestal shifts as a diagnostic tool in primary frequency standards. *IEEE Trans. Instrum. Meas.*, 44:136–139, 1995.

[268] A. Bauch, T. Heindorff, R. Schröder, and B. Fischer. The PTB primary clock CS3: type B evaluation of its standard uncertainty. *Metrologia*, 33:249–259, 1996.

[269] Jean-Louis Picqué. Hyperfine optical pumping of a cesium atomic beam, and applications. *Metrologia*, 13:115–119, 1977.

[270] G. Avila, V. Giordano, V. Candelier, E. de Clercq, G. Theobald, and P. Cerez. State selection in a cesium beam by laser-diode optical pumping. *Phys. Rev. A*, 36:3719–3728, 1987.

[271] A. Makdissi and E. de Clercq. Evaluation of the accuracy of the optically pumped caesium beam primary frequency standard of the BNM-LPTF. *Metrologia*, 38:409–425, 2001.

[272] S.-I. Ohshima, Y. Nakadan, T. Ikegami, Y. Koga, R. Drullinger, and L. Hollberg. Characteristics of an optically pumped Cs frequency standard at the NRLM. *IEEE Trans. Instrum. Meas.*, 38:533–536, 1989.

[273] Ken Hagimoto, S. Ohshima, Y. Nakadan, and Yasuki Koga. Accuracy evaluation of the optically pumped Cs frequency standard at NRLM. *IEEE Trans. Instrum. Meas.*, 48:496–499, 1999.

[274] G. D. Rovera, E. de Clercq, and A. Clairon. An analysis of major frequency shifts in the LPTF optically pumped primary frequency standard. *IEEE Trans. Ultrason. Ferroelec. Freq. Contr.*, 41:245–249, 1994.

[275] W. D. Lee, J. H. Shirley, J. P. Lowe, and R. E. Drullinger. The accuracy evaluation of NIST-7. *IEEE Trans. Instrum. Meas.*, IM 44:120–123, 1995.

[276] E. de Clercq and A. Makdissi. Current status of the LPTF optically pumped Cs beam standard. In J. C. Bergquist, editor, *Proceedings of the Fifth Symposium on Frequency Standards and Metrology*, pages 409–410, Singapore, New Jersey, London, Hong Kong, 1996. World Scientific.

[277] R. E. Drullinger, S. L. Rolston, and W. M. Itano. Primary atomic-frequency standards: New developments. In W. Ross Stone, editor, *Review of Radio Science 1993–1996*, pages 11–41, Oxford, New York, 1996. Oxford University Press.

[278] M. A. Kasevich, E. Riis, S. Chu, and R. G DeVoe. rf spectroscopy in an atomic fountain. *Phys. Rev. Lett.*, 63:612–615, 1989.

[279] A. Clairon, S. Ghezali, G. Santarelli, Ph. Laurent, S. N. Lea, M. Bahoura, E. Simon, S. Weyers, and K. Szymaniec. Preliminary accuracy evaluation of a cesium fountain frequency standard. In J. C. Bergquist, editor, *Proceedings of the 5th Symposium on Frequency Standards and Metrology*, pages 49–59, Singapore, 1996. World Scientific.

[280] D. M. Meekhof, S. R. Jefferts, and T. E. Parker. Accuracy evaluation of a cesium fountain primary frequency standard at NIST. *IEEE Trans. Instrum. Meas.*, 50:507–509, 2001.

[281] Stefan Weyers, Andreas Bauch, Udo Hübner, Roland Schröder, and Christian Tamm. First performance results of PTB's atomic caesium fountain and a study of contributions to its frequency instability. *IEEE Trans. Ultrason. Ferroelec. Freq. Contr.*, 47:432–437, 2000.

[282] E. Burt, T. Swanson, and C. Ekstrom. Cesium fountain development at USNO. In *Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum and The IEEE International Frequency Control Symposium*, pages 20–23, 1999.

[283] P. B. Whibberley, D. Henderson, and S. N. Lea. Development of a caesium fountain primary frequency standard at the NPL. In *Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum and The IEEE International Frequency Control Symposium*, pages 24–26, 1999.

[284] M. S. Huang, A. Yao, J. L. Peng, C. C. Chen, Shau Wei Hsu, Jen Ming Hsiao, C. S. Kou, and Chia Shu Liao. Compact cesium atomic fountain clock. In *Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum and The IEEE International Frequency Control Symposium*, pages 27–29, 1999.

[285] W. Liji, W. Changhua, H. Bingying, L. Mingshou, Q. Jin, and J. Wangxi. Design & preliminary results of NIM cesium fountain primary frequency standard. In *Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum and The IEEE International Frequency Control Symposium*, pages 30–33, 1999.

[286] Y. Sortais, S. Bize, C. Nicolas, A. Clairon, C. Salomon, and C. Williams. Cold collision frequency shifts in a ^{87}Rb atomic fountain. *Phys. Rev. Lett.*, 85:3117–3120, 2000.

[287] Ch. Fertig, R. Legere, W. Süptitz, and K. Gibble. Laser-cooled Rb fountain clocks. In *Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum and The IEEE International Frequency Control Symposium*, pages 39–42, 1999.

[288] A. Joyet, G. Mileti, P. Thomann, and G. Dudle. Continuous fountain Cs standard: Stability and accuracy issues. In P. Gill, editor, *Frequency Standards and Metrology, Proceedings of the Sixth Symposium*, pages 273–280, Singapore, 2002. World Scientific.

[289] Filippo Levi, Luca Lorini, Davide Calonico, and Aldo Godone. Systematic shift uncertainty evaluation of IEN CSF1 primary frequency standard. *IEEE Trans. Instrum. Meas.*, 52:267–271, 2003.

[290] D. Boiron, A. Michaud, P. Lemonde, Y. Castin, C. Salomon, S. Weyers, K. Szymaniec, L. Cognet, and A. Clairon. Laser cooling of cesium atoms in grey optical molasses down to 1.1 μK. *Phys. Rev. A*, 53:R3734–R3737, 1996.

[291] K. Gibble and S. Chu. Laser-cooled Cs frequency standard and a measurement of the frequency shift due to ultracold collisions. *Phys. Rev. Lett.*, 70:1771–1774, 1993.

[292] S. Ghezali, Ph. Laurent, S. Lea, and A. Clairon. An experimental study of the spin-exchange frequency shift in a laser-cooled cesium fountain frequency standard. *Europhys. Lett.*, 36:25–30, 1996.

[293] Paul J. Leo, Paul S. Julienne, Fred H. Mies, and Carl J. Williams. Collisional frequency shifts in ^{133}Cs fountain clocks. *Phys. Rev. Lett.*, 86:3743–3746, 2001.

[294] F. Pereira Dos Santos, H. Marion, S. Bize, A. Clairon, and C. Salomon. Controlling the cold collision shift in high precision atomic interferometry. *Phys. Rev. Lett.*, 89:233004–1–4, 2002.

[295] Chad Fertig and Kurt Gibble. Measurement and cancellation of the cold collison frequency shift in an ^{87}Rb fountain clock. *Phys. Rev. Lett.*, 85:1622–1625, 2000.

[296] Shin-Ichi Ohshima, Takayuki Kurosu, Takeshi Ikegami, and Yasuhiro Nakadan. Multipulse operation of cesium atomic fountain. In J. C. Bergquist, editor, *Proceedings of the Fifth Symposium on Frequency Standards and Metrology*, pages 60–65, Singapore, New Jersey, London, Hong Kong, 1996. World Scientific.

[297] Kurt Gibble. Collisional effects in cold alkalis. In J. C. Bergquist, editor, *Proceedings of the Fifth Symposium on Frequency Standards and Metrology*, pages 66–73, Singapore, New Jersey, London, Hong Kong, 1996. World Scientific.

[298] R. Legere and K. Gibble. Quantum scattering in a juggling atomic fountain. *Phys. Rev. Lett.*, 81:5780–5783, 1998.

[299] P. Berthoud, E. Fretel, and P. Thomann. Study of a bright, slow, and cold cesium source for a continuous beam frequency standard. In *Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum and The IEEE International Frequency Control Symposium*, pages 88–91, 1999.

[300] Patrick Berthoud, Emmanuel Fretel, Allan Joyet, Gregor Dudle, and Pierre Thomann. Toward a primary frequency standard based on a continuous fountain of laser-cooled cesium atoms. *IEEE Trans. Instrum. Meas.*, 48:516–519, 1999.

[301] G. Dudle, A. Joyet, E. Fretel, P. Berthoud, and P. Thomann. An alternative cold cesium frequency standard: The continuous fountain. In *Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum and The IEEE International Frequency Control Symposium*, pages 77–80, 1999.

[302] P. Lemonde, P. Laurent, E. Simon, G. Santarelli, A. Clairon, C. Salomon, N. Dimarcq, and P. Petit. Test of a space cold atom clock prototype in the absence of gravity. *IEEE Trans. Instrum. Meas.*, 48:512–515, 1999.

[303] Ph. Laurent, P. Lemonde, M. Abgrall, G. Santarelli, F. Pereira Dos Santos, A. Clairon, P. Petit, and M. Aubourg. Interrogation of cold atoms in a primary frequency standard. In *Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum and The IEEE International Frequency Control Symposium*, pages 152–155, 1999.

[304] Ph. Laurent, A. Clairon, P. Lemonde, G. Santarelli, C. Salomon, C. Sirmain, F. Picard, Ch. Delaroche, O. Grosjean, M. Saccoccio, M. Chaubet, L. Guillier, and J. Abadie. The space clock PHARAO: Functioning and expected performances. In *Proceedings of the 2003 IEEE International Frequency Control Symposium and PDA Exhibition Jointly with the 17th European Frequency and Time Forum*, pages 179–184, 2003.

[305] S. R. Jefferts, T. P. Heavner, L. W. Hollberg, J. Kitching, D. M. Meekhof, T. E. Parker, W. Phillips, S. Rolston, H. G. Robinson, J. H. Shirley, D. B. Sullivan, F. L. Walls, N. Ashby, W. M. Klipstein, L. Maleki, D. Seidel, R. Thompson, S. Wu, L. Young, R. F. C. Vessot, and A. De Marchi. PARCS: A primary atomic reference clock in space. In *Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum and The IEEE International Frequency Control Symposium*, pages 141–144, 1999.

[306] Chad Fertig, Kurt Gibble, Bill Klipstein, James Kohel, Lute Maleki, David Seidel, and Rob Thompson. Laser-cooled microgravity clocks. In *Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum and The IEEE International Frequency Control Symposium*, pages 145–147, 1999.

[307] G. J. Dick, W. M. Klipstein, T. P. Heavner, and S. R. Jefferts. Design concept for the microwave interrogation structure in PARCS. In *Proceedings of the 2003 International Frequency Control Symposium and PDA Exhibition Jointly with the European Frequency and Time Forum*, pages 1032–1036, 2003.

[308] N. G. Basov and A. M. Prokhorov. Application of molecular beams to radiospectroscopic investigations of rotational molecular spectra. *Sov. Phys. JETP*, 27:431–438, 1954. (In Russian).

[309] J. P. Gordon, H. J. Zeiger, and C. H. Townes. Molecular microwave oscillator and new hyperfine structure in the microwave spectrum of NH_3. *Phys. Rev.*, 95:282–284, 1954.

[310] H. M. Goldenberg, D. Kleppner, and N. F. Ramsey. Atomic hydrogen maser. *Phys. Rev. Lett.*, 5:361–365, 1960.

[311] Daniel Kleppner, H. Mark Goldenberg, and Norman F. Ramsey. Theory of the hydrogen maser. *Phys. Rev.*, 126:603–615, 1962.

[312] D. Kleppner, H. C. Berg, S. B. Crampton, N. F. Ramsey, R. F. C. Vessot, H. E. Peters, and J. Vanier. Hydrogen-maser principles and techniques. *Phys. Rev. A*, 138:972–983, 1965.

[313] J. Vanier. Atomic frequency standards: Basic physics and impact on metrology. In *Recent Advantages in Metrology and Fundamental Constants*, volume Course CXLVI of *Proceedings Internat. School of Physics "Enrico Fermi"*, pages 397–452, Amsterdam, Oxford, Tokyo, Washington DC, 2001. IOS Press Ohmsha.

[314] P. L. Bender. Effect of hydrogen-hydrogen exchange collisions. *Phys. Rev.*, 132:2154–2158, 1963.

[315] Howard C. Berg. Spin exchange and surface relaxation in the atomic hydrogen maser. *Phys. Rev.*, 137:A1621–A1635, 1965.

[316] H. Friedburg and W. Paul. Optische Abbildung mit neutralen Atomen. *Naturwissenschaften*, 38:159–160, 1951.

[317] Helmut Friedburg. Optische Abbildung mit neutralen Atomen. *Z. Phys.*, 130:493–512, 1951.

[318] Aaron Lemonick, Francis M. Pipkin, and Donald R. Hamilton. Focusing atomic beam apparatus. *Rev. Sci. Instrum.*, 26:1112–1119, 1955.

[319] R. L. Christensen and D. R. Hamilton. Permanent magnet for atomic beam focusing. *Rev. Sci. Instrum.*, 30:356–358, 1959.

[320] W. G. Kaenders, F. Lison, I. Müller, A. Richter, R. Wynands, and D. Meschede. Refractive components for magnetic optics. *Phys. Rev. A*, 54:5067–5075, 1996.

[321] F. G. Major. *The Quantum Beat*. Springer, New York, Berlin, Heidelberg, 1998.

[322] H. Bryan Owings, Paul A. Koppang, Collin C. MacMillan, and Harry E. Peters. Experimental frequency and phase stability of the hydrogen maser standard output as affected by cavity auto-tuning. In *Proceedings of the 46th Annual IEEE International Frequency Control Symposium, 27–29 May 1992, Hershey, Pa, USA*, pages 92–103, 1992.

[323] A. Boyko, G. Yolkin, N. Gestkova, G. Kurnikov, and V. Paramzin. Hydrogen maser with improved short-term frequency stability. In *Proceedings of the 15th European Frequency and Time Forum*, pages 406–408, Rue Jaquet-Droz 1, Case postale 20, CH-2007 Neuchâtel, Switzerland, 2001. FSRM Swiss Foundation for Research in Microtechnology.

[324] Helmut Hellwig, Robert F. C. Vessot, Martin W. Levine, Paul W. Zitzewitz, David W. Allan, and David J. Glaze. Measurement of the unperturbed hydrogen hyperfine transition frequency. *IEEE Trans. Instrum. Meas.*, IM-19:200–209, 1970.

[325] Robert F. C. Vessot, Edward M. Mattison, George U. Nystrom, Laurence M. Coyle, David Boyd, and Thomas E. Hoffman. High precision time transfer to test an H-maser on Mir. In J. C. Bergquist, editor, *Proceedings of the Fifth Symposium on Frequency Standards and Metrology*, pages 39–45, Singapore, New Jersey, London, Hong Kong, 1996. World Scientific.

[326] R. S. Van Dyck, Jr., P. B. Schwinberg, and H. G. Dehmelt. New high-precision comparison of electron and positron g factors. *Phys. Rev. Lett.*, 59:26–29, 1987.

[327] C. Audoin. Fast cavity auto-tuning systems for hydrogen masers. *Revue Phys. Appl.*, 16:125–130, 1981.

[328] J. M. V. A. Koelman, S. B. Crampton, H. T. C. Stoof, O. J. Luiten, and B. J. Verhaar. Spin-exchange frequency shifts in cryogenic and room-temperature hydrogen masers. *Phys. Rev.*, 38:3535–3547, 1988.

[329] Ronald L. Walsworth, Isaac F. Silvera, Edward M. Mattison, and Robert C. Vessot. Measurement of a hyperfine-induced spin-exchange frequency shift in atomic hydrogen. *Phys. Rev. A*, 46:2495–2512, 1992.

[330] Michael E. Hayden, Martin D. Hürlimann, and Walter N. Hardy. Atomic hydrogen spin-exchange collisions in a cryogenic maser. *IEEE Trans. Instrum. Meas.*, 42:314–319, 1993.

[331] S. J. J. M. F. Kokkelmans and B. J. Verhaar. Discrepancies in experiments with cold hydrogen atoms. *Phys. Rev. A*, 56:4038–4044, 1997.

[332] S. B. Crampton. Spin-exchange shifts in the hydrogen maser. *Phys. Rev.*, 158:57–61, 1967.

[333] L. S. Cutler and C. L. Searle. Some aspects of the theory and measurement of frequency fluctuations in frequency standards. *Proc. IEEE*, 54:136–154, 1966.

[334] Thomas E. Parker. Hydrogen maser ensemble performance and characterization of frequency standards. In *Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum and The IEEE International Frequency Control Symposium*, pages 173–176, 1999.

[335] N. A. Demidov, A. V. Pastukhov, and A. A. Uljanov. Progress in the development of IEM KVARZ passive hydrogen masers. In *Proceedings of the 31th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, December 7 – 9, 1999 Laguna Cliffs Marriott, Dana Point, California*, volume 31, (Compact Disk only), pages 579–587, http://tycho.usno.navy.mil/ptti/ptti99/PTTI_1999_579.PDF, 1999.

[336] L. Mattioni, M. Belloni, P. Berthoud, I. Pavlenko, H. Schweda, Q. Wang, P. Rochat, F. Droz, P. Mosset, and H. Ruedin. The development of a passive hydrogen maser clock for the Galileo navigation system. In *Proceedings of the 34th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, December 3–5, 2002, PTTI 2002, The Hyatt Regency, Reston Town Center, Reston, Virginia*, volume 34, pages 579–587, http://tycho.usno.navy.mil/ptti/ptti2002/paper14.pdf, 2002.

[337] G. Busca and Q. Wang. Time prediction accuracy for a space clock. *Metrologia*, 40:S265–S269, 2003.

[338] R. F. C. Vessot, E. M. Mattison, R. L. Walsworth, and I. F. Silvera. The cold hydrogen maser. In A. De Marchi, editor, *Frequency standards and Metrology*, pages 88–93. Springer, Berlin, Heidelberg, 1989.

[339] S. B. Crampton. Introduction to cold hydrogen masers. In A. De Marchi, editor, *Frequency standards and metrology*, pages 86–87. Springer, Berlin, Heidelberg, New York, 1989.

[340] Harald F. Hess, Greg P. Kochanski, John M. Doyle, Thomas P. Greytak, and Daniel Kleppner. Spin-polarized hydrogen maser. *Phys. Rev. A*, 34:1602–1604, 1986.

[341] Ronald L. Walsworth, Jr., Isaac F. Silvera, H. P. Godfried, C. C. Agosta, Robert C. Vessot, and Edward M. Mattison. Hydrogen maser at temperatures below 1 K. *Phys. Rev. A*, 34:2550–2553, 1986.

[342] M. D. Hürlimann, W. N. Hardy, A. J. Berlinsky, and R. W. Cline. Recirculating cryogenic hydrogen maser. *Phys. Rev. A*, 34:1605–1608, 1986.

[343] M. E. Hayden and W. N. Hardy. Spin exchange and recombination in a gas of atomic hydrogen at 1.2 K. *Phys. Rev. Lett.*, 76:2041–2044, 1996.

[344] R. F. C. Vessot, M. W. Levine, E. M. Mattison, E. L. Blomberg, T. E. Hoffman, G. U. Nystrom, B. F. Farrel, R. Dechel, P. B. Eby, C. R. Baugher, J. W. Watts, D. L. Teuber, and F. D. Wills. Test of relativistic gravitation with a space-borne hydrogen maser. *Phys. Rev. Lett.*, 45:2081–2084, 1980.

[345] A. Bauch and S. Weyers. New experimental limit on the validity of local position invariance. *Phys. Rev. D*, 65:081101–1–4, 2002.

[346] D. F. Phillips, M. A. Humphrey, E. M. Mattison, R. E. Stoner, R. F. C. Vessot, and R. L. Walsworth. Limit on Lorentz and CPT violation of the proton using a hydrogen maser. *Phys. Rev. A*, 63:111101–1–4, 2001.

[347] D. Colladay and V. A. Kostelecký. Lorentz-violating extension of the standard model. *Phys. Rev. D*, 58:116002–1–23, 1998.

[348] H. Weaver, D. R. W. Williams, N. H. Dieter, and W. T. Lum. Observations of a strong unidentified microwave line and of emission from the OH molecule. *Nature*, 208:29–31, 1965.

[349] James M. Moran. Cosmic masers: A powerful tool for astrophysics. In J. Hamelin, editor, *Modern Radio Science 1996*, pages 245–262. Oxford University Press, 1996.

[350] Normand Cyr, Michel Têtu, and Marc Breton. All-optical microwave frequency standard: A proposal. *IEEE Trans. Instrum. Meas.*, 42:640–649, 1993.

[351] N. Vukičević, A. S. Zibrov, L. Hollberg, F. L. Walls, J. Kitching, and H. G. Robinson. Compact diode-laser based Rubidium frequency reference. In *Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum and The IEEE International Frequency Control Symposium*, pages 133–136, 1999.

[352] E. B. Alexander, M. V. Balabas, D. Budker, D. English, D. F. Kimball, C.-H. Li, and V. V. Yashchuk. Light-induced desorption of alkali-metal atoms from paraffin coating. *Phys. Rev. A*, 66:024903–1–12, 2002.

[353] M. Stephens, R. Rhodes, and C. Wieman. Study of wall coatings for vapor-cell laser traps. *J. Appl. Phys.*, 76:3479–3488, 1994.

[354] Csaba Szekely and Robert Drullinger. Improved rubidium frequency standards using diode lasers with AM and FM noise control. *Proc. of the SPIE*, 1837:299–305, 1992.

[355] Y. Koyama, H. Matsuura, K. Atsumi, Y. Nakajima, and K. Chiba. An ultra-miniature rubidium frequency standard with two-cell scheme. In *Proceedings of the 49th Annual IEEE International Frequency Control Symposium, 31 May – 2 June 1995, San Francisco, USA*, pages 33–38, 1995.

[356] Claire Couplet, Pascal Rochat, Gaetano Mileti, Hartmut Schweda, Pierre Thomann, and Giovanni Busca. Miniaturized rubidium clocks for space and insdustrial applications. In *Proceedings of the 49th Annual IEEE International Frequency Control Symposium, 31 May – 2 June 1995, San Francisco, USA*, pages 53–59, 1995.

[357] T. McClelland, I. Pascaru, I. Shtaerman, C. Szekely, J. Zacharski, and N. D. Baskar. Subminiature rubidium frequency standard: Manufacturability and performance results from production units. In *Proceedings of the 49th Annual IEEE International Frequency Control Symposium, 31 May – 2 June 1995, San Francisco, USA*, pages 39–52, 1995.

[358] A. Jeanmaire, P. Rochat, and F. Emma. Rubidium atomic clock for Galileo. In *Proceedings of the 31th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, December 7 – 9, 1999, Laguna Cliffs Marriott Dana Point, California*, NASA Conference Publication, pages 627–636, U.S. Naval Observatory, 3450 Massachusetts Ave., N.W. Washington, D.C. 20392-5420, USA, 1999.

[359] Pascal Rochat and Bernard Leuenberger. A new synchronized miniature rubidium oscillator with an auto-adaptive disciplining filter. In *Proceedings of the 33th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, November 27-29, 2001, Hyatt Regency Hotel Long Beach, California*, pages 627–636, U.S. Naval Observatory, Time Service, 3450 Massachusetts Ave., N.W. Washington, DC 20392-5420, USA, 2001.

[360] J. G. Coffer and J.C. Camparo. Long-term stability of a rubidium atomic clock in geosynchronous orbit. In *Proceedings of the 33th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, December 7–9, 1999, Laguna Cliffs Marriott Dana Point, California*, pages 65–74, U.S. Naval Observatory, Time Service, 3450 Massachusetts Ave., N.W. Washington, DC 20392-5420, USA, 1999.

[361] Y. Saburi, Y. Koga, S. Kinugawa, T. Imamura, H. Suga, and Y. Ohuchi. Short-term stability of laser-pumped rubidium gas cell frequency standard. *Electron. Lett.*, 30:633–635, 1994.

[362] G. Mileti, J. Deng, F. L. Walls, D. A. Jennings, and R. E. Drullinger. Laser-pumped Rubidium frequency standards: new analysis and progress. *IEEE J. Quantum Electron.*, 34:233–237, 1998.

[363] J. Kitching, S. Knappe, N. Vukičević, L. Hollberg, R. Wynands, and W. Weidmann. A microwave frequency reference based on VCSEL-driven dark line resonances in Cs vapor. *IEEE Trans. Instrum. Meas.*, 49:1313–1317, 2000.

[364] M. Zhu and L. S. Cutler. Theoretical and experimental study of light shift in a CPT-based Rb vapor cell frequency standard. In *Proceedings of the 31th Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting, November 28–30, 2000, Washington DC, USA*, volume 2220 of *NASA Conference Publication*, pages 311–324, U.S. Naval Observatory, Time Service, 3450 Massachusetts Ave., N.W. Washington, DC 20392-5420, USA, 2001.

[365] J. Kitching, S. Knappe, and L. Hollberg. Miniature vapor-cell atomic frequency references. *Appl. Phys. Lett.*, 81:553–555, 2002.

[366] R. Lutwak, D. Emmons, T. English, W. Riley, A. Duwel, M. Varghese, D. K. Serland, and G. M. Peake. The chip-scale atomic clock – recent development progress. In *Proceedings of the 34th Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting, December 2-4, 2003, The Hilton Resort on Mission Bay, San Diego, California*, pages 539–550, U.S. Naval Observatory, Time Service, 3450 Massachusetts Ave., N.W. Washington, DC 20392-5420, USA, 2003.

[367] Li-Anne Liew, Svenja Knappe, John Moreland, Hugh Robinson, Leo Hollberg, and John Kitching. Microfabricated alkali atom vapor cells. *Appl. Phys. Lett.*, 84:2694–2696, 2004.

[368] Li Hua. Verbesserung der Kohärenzeigenschaften der Emission von Halbleiterlasern durch Rückkopplung von einem Resonator hoher Finesse. PTB–Bericht PTB–Opt–33, Physikalisch-Technische Bundesanstalt, Braunschweig, Juni 1990.

[369] H. Gerhardt, H. Welling, and A. Güttner. Measurements of the laser linewidth due to quantum phase and quantum amplitude noise above and below threshold. I. *Z. Physik*, 253:113–126, 1972.

[370] T. J. Quinn. Practical realization of the definition of the metre, including recommended radiations of other optical frequency standards (2001). *Metrologia*, 40:103–133, 2003.

[371] R. Balhorn, H. Kunzmann, and F. Lebowsky. Frequency stabilization of internal-mirror Helium-Neon lasers. *Appl. Opt.*, 11:742–744, 1972.

[372] U. Brand, F. Mensing, and J. Helmcke. Polarization properties and frequency stabilization of an internal mirror He-Ne laser emitting at 543.5 nm wavelength. *Appl. Phys.*, B48:343–350, 1989.

[373] T. M. Niebauer, J. E. Faller, H. M. Godwin, J. L. Hall, and R. L. Barger. Frequency stability measurements on polarization-stabilized He-Ne lasers. *Appl. Opt.*, 27:1285–1289, 1988.

[374] T. Baer, F. V. Kowalski, and J. L. Hall. Frequency stabilization of a 0.633 μm He-Ne longitudinal Zeeman laser. *Appl. Opt.*, 19:3173–3177, 1980.

[375] R. A. McFarlane, W. R. Bennett, Jr., and W. E. Lamb, Jr. Single mode tuning dip in the power output of an He-Ne optical maser. *Appl. Phys. Lett.*, 2:189–190, 1963.

[376] Willis E. Lamb, Jr. Theory of an optical maser. *Phys. Rev.*, 134:A1429–A1450, 1964.

[377] G. R. Hanes and C. E. Dahlstrom. Iodine hyperfine structure observed in saturated absorption at 633 nm. *Appl. Phys. Lett.*, 14:362–364, 1969.

[378] H. Darnedde, W. R. C. Rowley, F. Bertinetto, Y. Millerioux, H. Haitjema, S. Wetzels, H. Pirée, E. Prieto, M. Mar Pérez, B. Vaucher, A. Chartier, and J.-M. Chartier. International comparisons of He-Ne lasers stabilized with $^{127}I_2$ at $\lambda \approx 633$ nm (July 1993 to September 1995). *Metrologia*, 36:199–206, 1999.

[379] A. Lassila, K. Riski, J. Hu, T. Ahola, S. Naicheng, L. Chengyang, P. Balling, J. Blabla, L. Abramova, Yu. G. Zakharenko, V. L. Fedorin, A. Chartier, and J.-M. Chartier. International comparison of He-Ne lasers stabilized with $^{127}I_2$ at $\lambda \approx 633$ nm. *Metrologia*, 37:701–707, 2000.

[380] A. A. Madej, J. E. Bernard, L. Robertsson, L.-S. Ma, M. Zucco, and R. S. Windeler. Long-term absolute frequency measurements of 633 nm iodine-stabilized laser standards at NRC and demonstration of high reproducibility of such devices in international frequency measurements. *Metrologia*, 41:152–160, 2004.

[381] Matthew S. Taubman and John L. Hall. Cancellation of laser dither modulation from optical frequency standards. *Opt. Lett.*, 25:311–313, 2000.

[382] D. A. Tyurikov, M. A. Gubin, A. S. Shelkovnikov, and E. V. Koval'chuk. Accuracy of the computer-controlled laser frequency standards based on resolved hyperfine structure of a methane line. *IEEE Trans. Instrum. Meas.*, 44:166–169, 1995.

[383] M. A. Gubin, D. A. Tyurikov, A. S. Shelkovnikov, E. V. Koval'chuk, G. Kramer, and B. Lipphardt. Transportable He-Ne/CH_4 optical frequency standard and absolute measurements of its frequency. *IEEE J. Quantum Electron.*, 31:2177–2182, 1995.

[384] S. N. Bagayev, A. K. Dmitriyev, and P. V. Pokasov. Transportable He-Ne/CH_4 frequency standard for precision measurement. *Laser Physics*, 7:989–992, 1997.

[385] O. Acef, A. Clairon, G. D. Rovera, F. Ducos, L. Hilico, G. Kramer, B. Lipphardt, A. Shelkovnikov, E. Kovalchuk, E. Petrukhin, D. Tyurikov, M. Petrovskiy, M. Gubin, R. Felder, P. Gill, and S. Lea. Absolute frequency measurements with a set of transportable Methane optical frequency standards. In *Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum and The IEEE International Frequency Control Symposium*, pages 742–745, 1999.

[386] P. S. Ering, D. A. Tyurikov, G. Kramer, and B. Lipphardt. Measurement of the absolute frequency of the methane E-line at 88 THz. *Opt. Commun.*, 151:229–234, 1998.

[387] Claus Braxmaier. *Fundamentale Tests der Physik mit ultrastabilen optischen Oszillatoren*. PhD thesis, Universität Konstanz, Konstanz, 2001.

[388] K. M. Evenson, G. W. Day, J. S. Wells, and L. O. Mullen. Extension of absolute frequency measurements to the cw He-Ne laser at 88 THz (3.39 μ). *Appl. Phys. Lett.*, 20:133–134, 1972.

[389] D. J. E. Knight, G. J. Edwards, P. R. Pearce, and N. R. Cross. Frequency of the methane-stabilized He-Ne laser at 88 THz measured to ± 3 parts in 10^{11}. *Nature*, 285:388–390, 1980.

[390] Yu. S. Domnin, N. B. Koshelyaevskii, V. M. Tatarenkov, and P. S. Shumyatskii. Measurement of the frequency of a He-Ne/CH_4 laser. *JETP Lett.*, 34:167–170, 1981.

[391] B. G. Whitford and G. R. Hanes. Frequency of a methane-stabilized helium-neon laser. *IEEE Trans. Instrum. Meas.*, 37:179–184, 1988.

[392] C. O. Weiss, G. Kramer, B. Lipphardt, and E. Garcia. Frequency measurement of a CH_4 hyperfine line at 88 THz / "optical clock". *IEEE J. Quantum Electron.*, 24:1970–1972, 1988.

[393] M. Gubin, E. Kovalchuk, E. Petrukhin, A. Shelkovnikov, D. Tyurikov, R. Gamidov, C. Erdogan, E. Sahin, R. Felder, P. Gill, S. N. Lea, G. Kramer, and B. Lipphardt. Absolute frequency measurements with a set of transportable He-Ne/CH_4 optical frequency standards and prospects for future design and applications. In P. Gill, editor, *Frequency Standards and Metrology, Proceedings of the Sixth Symposium*, pages 453–460, Singapore, 2002. World Scientific.

[394] A. Huber, Th. Udem, B. Gross, J. Reichert, M. Kourogi, K. Pachucki, M. Weitz, and T. W. Hänsch. Hydrogen-deuterium $1S - 2S$ isotope shift and the structure of the deuteron. *Phys. Rev. Lett.*, 80:468–471, 1998.

[395] J. von Zanthier, J. Abel, Th. Becker, M. Fries, E. Peik, H. Walther, R. Holzwarth, J. Reichert, Th. Udem, T. W. Hänsch, A. Yu. Nevsky, M. N. Skvortsov, and S. N. Bagayev. Absolute frequency measurement of the $^{115}In^{+}$ $5s^2\ ^1S_0$–$5s5p\ ^3P_0$ transition. *Opt. Commun.*, 166:57–63, 1999.

[396] E. N. Bazarov, G. A. Gerasimov, K. I. Guryev, V. L. Derbov, M. A. Kovner, Yu. I. Posudin, S. K. Potapov, and V. A. Chenin. Vibration-rotational super-high resolution spectrum of OsO_4 and its theoretical interpretation. *J. Quant. Spectrosc. Radiat. Transfer*, 17:7–12, 1977.

[397] Yu. S. Domnin, N. B. Koshelyaevskii, V. M. Tatarenkov, P. S. Shumyatskii, O. N. Kompanets, A. R. Kukudzhanov, V. S. Letokhov, and E. L. Mikhailov. CO_2: $^{192}OsO_4$ laser: Absolute frequency of optical oscillations and new possibilities. *JETP Lett.*, 30:249–252, 1979.

[398] André Clairon, Alain Van Lerberghe, Christophe Salomon, Michel Ouhayoun, and Christian J. Bordé. Towards a new absolute frequency reference grid in the 28 THz range. *Opt. Commun.*, 35:368–372, 1980.

[399] A. Clairon, O. Acef, C. Chardonnet, and C. J. Bordé. State-of-the-art for high accuracy frequency standards in the 28 THz range using saturated absorption resonances of OsO_4 and CO_2. In A. De Marchi, editor, *Frequency Standards and Metrology*, pages 212–221, Berlin, Heidelberg, New York, 1989. Springer-Verlag.

[400] Yu. S. Domnin, N. B. Koshelyaevskii, A. N. Malimon, V. M. Tatarenkov, and P. S. Shumyatskii. Infrared frequency standard based on osmium tetraoxide. *Sov. J. Quantum Electron.*, 17:801–803, 1987.

[401] K. Stoll. Perspektiven für ein OsO_4 - Frequenznormal. PTB-Bericht PTB-Opt–49, Physikalisch-Technische Bundesanstalt, Braunschweig, 1995.

[402] O. Acef. Metrological properties of CO_2/OsO_4 optical frequency standard. *Opt. Commun.*, 134:479–486, 1997.

[403] O. Acef, F. Michaud, and Giovanni Daniele Rovera. Accurate determination of OsO_4 absolute frequency grid at 28/29 THz. *IEEE Trans. Instrum. Meas.*, 48:567–570, 1999.

[404] O. Acef. CO_2/OsO_4 lasers as frequency standards in the 29 THz range. *IEEE Trans. Instrum. Meas.*, IM–46:162–165, 1997.

[405] Giovanni Daniele Rovera and O. Acef. Absolute frequency measurement of mid-infrared secondary frequency standard at BNM-LPTF. *IEEE Trans. Instrum. Meas.*, 48:571–573, 1999.

[406] A. Amy-Klein, A. Goncharov, C. Daussy, C. Grain, O. Lopez, G. Santarelli, and C. Chardonnet. Absolute frequency measurement in the 28-THz spectral region with a femtosecond laser comb and a long-distance optical link to a primary standard. *Appl. Phys. B*, 78:25–30, 2004.

[407] Ch. Chardonnet and Ch. J. Bordé. Hyperfine interactions in the ν_3 band of osmium tetroxide: Accurate determination of the spin-rotation constant by crossover resonance spectroscopy. *J. Mol. Spectr.*, 167:71–98, 1994.

[408] T. W. Hänsch and B. Couillaud. Laser frequency stabilization by polarization spectroscopy of a reflecting reference cavity. *Opt. Commun.*, 35:441–444, 1980.

[409] Eugene Hecht and Alfred Zajac. *Optics*. Addison-Wesley, Reading MA, Amsterdam, London, 6 edition, 1980.

[410] R. W. P. Drever, J. L. Hall, F. V. Kowalski, J. Hough, G. M. Ford, A. J. Munley, and H. Ward. Laser phase and frequency stabilization using an optical resonator. *Appl. Phys. B*, 31:97–105, 1983.

[411] R. V. Pound. Electronic frequency stabilization of microwave oscillators. *Rev. Sci. Instrum.*, 17:490–505, 1946.

[412] J. Helmcke, S. A. Lee, and J. L. Hall. Dye laser spectrometer for ultrahigh spectral resolution: Design and performance. *Appl. Opt.*, 21:1686–1694, 1982.

[413] N. C. Wong and J. L. Hall. Servo control of amplitude modulation in frequency-modulation spectroscopy: demonstration of shot-noise-limited detection. *J. Opt. Soc. Am. B*, 2:1527–1533, 1985.

[414] Edward A. Whittaker, Chi Man Shum, Haim Grebel, and Haim Lotem. Reduction of residual amplitude modulation in frequency-modulation spectroscopy by using harmonic frequency modulation. *J. Opt. Soc. Am. B*, 5:1253–1256, 1988.

[415] G. C. Bjorklund. Frequency-modulation spectroscopy: a new method for measuring weak absorptions and dispersions. *Opt. Lett.*, 5:15–17, 1980.

[416] J. L. Hall, L. Hollberg, T. Baer, and H. G. Robinson. Optical heterodyne saturation spectroscopy. *Appl. Phys. Lett.*, 39:680–682, 1981.

[417] Ady Arie and Robert L. Byer. Laser heterodyne spectroscopy of $^{127}I_2$ hyperfine structure near 532 nm. *J. Opt. Soc. Am. B*, 10:1990–1997, 1993.

[418] Jun Ye, Long-Sheng Ma, and John L. Hall. Ultrastable optical frequency reference at 1.064 μm using a C_2HD molecular overtone transition. *IEEE Trans. Instrum. Meas.*, 46:178–182, 1997.

[419] P. Cordiale, G. Galzerano, and H. Schnatz. International comparison of two iodine-stabilized frequency-doubled Nd:YAG lasers at λ = 532 nm. *Metrologia*, 37:177–182, 2000.

[420] A. Yu. Nevsky, R. Holzwarth, J. Reichert, Th. Udem, T. W. Hänsch, J. von Zanthier, H. Walther, H. Schnatz, F. Riehle, P. V. Pokasov, M. N. Skvortsov, and S. N. Bagayev. Frequency comparison and absolute frequency measurement of I_2-stabilized lasers at 532 nm. *Opt. Commun.*, 192:263–272, 2001.

[421] G. C. Bjorklund, M. D. Levenson, W. Lenth, and C. Ortiz. Frequency modulation (fm) spectroscopy: Theory of lineshapes and signal-to-noise analysis. *Appl. Phys. B*, 32:145–152, 1983.

[422] Jon H. Shirley. Modulation transfer processes in optical heterodyne saturation spectroscopy. *Opt. Lett.*, 7:537–539, 1982.

[423] J. J. Snyder, R. K. Raj, D. Bloch, and M. Ducloy. High-sensitivity nonlinear spectroscopy using a frequency-offset pump. *Opt. Lett.*, 5:163–165, 1980.

[424] R. K. Raj, D. Bloch, J. J. Snyder, G. Camy, and M. Ducloy. High-frequency optically heterodyned saturation spectroscopy via resonant degenerate four-wave mixing. *Phys. Rev. Lett.*, 44:1251–1254, 1980.

[425] A. Schenzle, R. G. DeVoe, and R. G. Brewer. Phase-modulation laser spectroscopy. *Phys. Rev. A*, 25:2606–2621, 1982.

[426] G. Camy, Ch. J. Bordé, and M. Ducloy. Heterodyne saturation spectroscopy through frequency modulation of the saturating beam. *Opt. Commun.*, 41:325–330, 1982.

[427] M. Ducloy and D. Bloch. Theory of degenerate four-wave mixing in resonant Doppler-broadened media. II. Doppler-free heterodyne spectroscopy via collinear four-wave mixing in two- and three-level systems. *J. Physique*, 43:57–65, 1982.

[428] Ma Long-Sheng and J. L. Hall. Optical heterodyne spectroscopy enhanced by an external optical cavity: Toward improved working standards. *IEEE J. Quantum Electron.*, 26:2006–2012, 1990.

[429] Esa Jaatinen. Theoretical determination of maximum signal levels obtainable with modulation transfer spectroscopy. *Opt. Commun.*, 120:91–97, 1995.

[430] R. W. Fox, L. Hollberg, and A. S. Zibrov. Semiconductor diode lasers. In F. B. Dunning and Randall G. Hulet, editors, *Atomic, Molecular, and Optical Physics: Electromagnetic Radiation*, volume 29C, pages 77–102. Academic Press, San Diego, 1997.

[431] Ulrich Strößner, Jan-Peter Meyn, Richard Wallenstein, Pavel Urenski, Ady Arie, Gil Rosenman, Jürgen Mlynek, Stephan Schiller, and Achim Peters. Single-frequency continuous-wave optical parametric oscillator system with an ultrawide tuning range of 550 to 2830 nm. *J. Opt. Soc. Am. B*, 19:1419–1424, 2002.

[432] Andrew Dienes and Diego R. Yankelevich. Continuous wave dye lasers. In F. B. Dunning and Randall G. Hulet, editors, *Atomic, Molecular, and Optical Physics: Electromagnetic Radiation*, volume 29C, pages 45–75. Academic Press, San Diego, 1997.

[433] Wolfgang Demtröder. *Laser Spectroscopy: Basic Concepts and Instrumentation*. Springer, Berlin, Heidelberg, New York, 2003.

[434] J. C. Bergquist and L. Burkins. Efficient single mode operation of a cw ring dye laser with a Mach–Zehnder interferometer. *Opt. Commun.*, 50:379–385, 1984.

[435] Ch. H. Henry. Theory of the linewidth of semiconductor lasers. *IEEE J. Quantum Electron.*, QE–18:259–264, 1982.

[436] Motoichi Ohtsu. *Highly Coherent Semiconductor Lasers*. Artech House, Boston, London, 1992.

[437] A. Celikov, F. Riehle, V. L. Velichansky, and J. Helmcke. Diode laser spectroscopy in a Ca atomic beam. *Opt. Commun.*, 107:54–60, 1994.

[438] V. Vassiliev, V. Velichansky, P. Kersten, T. Trebst, and F. Riehle. Subkilohertz enhanced-power diode-laser spectrometer in the visible. *Opt. Lett.*, 23:1229–1231, 1998.

[439] B. Bodermann, H. R. Telle, and R. P. Kovacich. Amplitude-modulation-free optoelectronic frequency control of laser diodes. *Opt. Lett.*, 25:899–901, 2000.

[440] V. L. Velichanskii, A. S. Zibrov, V. S. Kargopol'tsev, V. I. Molochev, V. V. Nikitin, V. A. Sautenkov, G. G. Kharisov, and D. A. Tyurikov. Minimum line width of an injection laser. *Sov. Tech. Phys. Lett.*, 4:438–439, 1978.

[441] M. W. Fleming and A. Mooradian. Spectral characteristics of external-cavity controlled semiconductor lasers. *IEEE J. Quantum Electron.*, QE–17:44–59, 1981.

[442] K. Petermann, editor. *Laser Diode Modulation and Noise*. Kluwer Academic Publishers, The Hague, Netherlands, 1988.

[443] J. Mørk, B. Tromborg, J. Mark, and V. Velichansky. Instabilities in a laser diode with strong optical feedback. *Proc. of the SPIE*, 1837:90–104, 1992.

[444] Bjarne Tromborg, Henning Olesen, Xing Pan, and Shigeru Saito. Transmission line description of optical feedback and injection locking for Fabry-Perot and DFB lasers. *IEEE J. Quantum Electron.*, QE-23:1875–1889, 1987.

[445] Rudolf F. Kazarinov and Charles H. Henry. The relation of line narrowing and chirp reduction resulting from the coupling of a semiconductor laser to a passive resonator. *IEEE J. Quantum Electron.*, QE-23:1401–1409, 1987.

[446] Ph. Laurent, A. Clairon, and Ch. Bréant. Frequency noise analysis of optically self-locked diode lasers. *IEEE J. Quantum Electron.*, 25:1131–1142, 1989.

[447] Dag Roar Hjelme, Alan Rolf Mickelson, and Raymond G. Beausoleil. Semiconductor laser stabilization by external optical feedback. *IEEE J. Quantum Electron.*, QE-27:352–372, 1991.

[448] K. Liu and M. G. Littman. Novel geometry for single-mode scanning of tunable lasers. *Opt. Lett.*, 6:117–118, 1981.

[449] T. Day, F. Luecke, and M. Brownell. Continuously tunable diode lasers. *Lasers and Optronics*, June 1993:15–17, 1993.

[450] F. Favre, D. Le Guen, J. C. Simon, and B. Landousies. External-cavity semiconductor laser with 15 nm continuous tuning range. *Electron. Lett.*, 22:795–796, 1986.

[451] W. R. Trutna, Jr. and L. F. Stokes. Continuously tuned external cavity semiconductor laser. *J. Lightwave Technol.*, 11:1279–1286, 1993.

[452] V. Vassiliev, V. Velichansky, P. Kersten, and F. Riehle. Injection locking of a red extended-cavity diode laser. *Electron. Lett.*, 33:1222–1223, 1997.

[453] B. Dahmani, L. Hollberg, and R. Drullinger. Frequency stabilization of semiconductor lasers by resonant optical feedback. *Opt. Lett.*, 12:876–878, 1987.

[454] A. Hemmerich, D. H. McIntyre, D. Schropp, Jr., D. Meschede, and T. W. Hänsch. Optically stabilized narrow linewidth semiconductor laser for high resolution spectroscopy. *Opt. Commun.*, 75:118–122, 1990.

[455] S. Kremser, B. Bodermann, H. Knöckel, and E. Tiemann. Frequency stabilization of diode lasers to hyperfine transitions of the iodine molecule. *Opt. Commun.*, 110:708–716, 1994.

[456] M. Ebrahimzadeh and M. H. Dunn. Optical parametric oscillators. In Michael Bass, Jay M. Enoch, Eric W. Van Stryland, and William L. Wolfe, editors, *Handbook of Optics*, pages 22.1 – 22.72. McGraw-Hill, New York, 2001.

[457] E. V. Kovalchuk, D. Dekorsy, A. L. Lvovsky, C. Braxmaier, J. Mlynek, A. Peters, and S. Schiller. High-resolution Doppler-free molecular spectroscopy using a continuous-wave optical parametric oscillator. *Opt. Lett.*, 26:1430–1432, 2001.

[458] G. Hollemann, E. Peik, A. Rusch, and H. Walther. Injection locking of a diode-pumped Nd:YAG laser at 946 nm. *Opt. Lett.*, 20:1871–1873, 1995.

[459] Thomas J. Kane and Robert L. Byer. Monolithic, unidirectional single-mode Nd:YAG ring laser. *Opt. Lett.*, 10:65–67, 1985.

[460] W. R. Trutna, Jr., D. K. Donald, and Moshe Nazarathy. Unidirectional diode-laser-pumped Nd:YAG ring laser with a small magnetic field. *Opt. Lett.*, 12:248–250, 1987.

[461] Ady Arie and Eran Inbar. Laser spectroscopy of molecular cesium near 1064 nm enhanced by a Fabry–Perot cavity. *Opt. Lett.*, 20:88–90, 1995.

[462] E. Inbar, V. Mahal, and A. Arie. Frequency stabilization of Nd:YAG lasers to $^{133}Cs_2$ sub-Doppler lines near 1064 nm. *J. Opt. Soc. Am. B*, 13:1598–1604, 1996.

[463] Shie-Chang Jeng, De-Yuan Chung, Chi-Chang Liaw, Fu-Hsiang Yang, Jow-Tsong Shy, Tyson Lin, and Sen-Yen Shaw. Absolute frequencies of the $^{133}Cs_2$ transitions near 1064 nm. *Opt. Commun.*, 155:263–269, 1998.

[464] Peter Fritschel and Rainer Weiss. Frequency match of the Nd:YAG laser at 1.064 μm with a line in CO_2. *Appl. Opt.*, 31:1910–1912, 1992.

[465] Jun Ye, Lennart Robertsson, Susanne Picard, Long-Sheng Ma, and John L. Hall. Absolute frequency atlas of molecular I_2 lines at 532 nm. *IEEE Trans. Instrum. Meas.*, 48:544–549, 1999.

[466] Feng-Lei Hong, Yun Zhang, Jun Ishikawa, Youichi Bitou, Atsushi Onae, Jun Yoda, Hirokazu Matsumoto, and Ken'ichi Nakagawa. Frequency reproducibility of I_2-stabilized Nd:YAG lasers. In John L. Hall and Jun Ye, editors, *Proceedings of SPIE: Laser Frequency Stabilization, Standards, Measurement, and Applications*, volume 4269, pages 248–254, P.O. Box 10, Bellingham, Washington 98227 - 0010 USA, 2001. SPIE.

[467] Jun Ye, Long Shen Ma, and John L. Hall. Molecular iodine clock. *Phys. Rev. Lett.*, 87:270801–1–4, 2001.

[468] M. de Labachelerie, K. Nakagawa, and M. Ohtsu. Ultranarrow $^{13}C_2H_2$ saturated-absorption lines at 1.5 μm. *Opt. Lett.*, 19:840–842, 1994.

[469] Jun Ye, Long-Sheng Ma, and John L. Hall. Sub-Doppler optical frequency reference at 1.064 μm by means of ultrasensitive cavity-enhanced frequency modulation spectroscopy of a C_2HD overtone transition. *Opt. Lett.*, 21:1000–1002, 1996.

[470] Takayuki Kurosu and Uwe Sterr. Frequency-stabilization of a 1.54 micrometer DFB-laser diode to Doppler-free absorption lines of acetylene. In John L. Hall and Jun Ye, editors, *Proceedings of SPIE: Laser Frequency Stabilization, Standards, Measurement, and Applications*, volume 4269, pages 143–154, P.O. Box 10, Bellingham, Washington 98227–0010 USA, 2001. SPIE.

[471] Chikako Ishibashi, Kotaro Suzumura, and Hiroyuki Sasada. Sub-Doppler resolution molecular spectroscopy in the 1.66-μm region. In John L. Hall and Jun Ye, editors, *Proceedings of SPIE: Laser Frequency Stabilization, Standards, Measurement, and Applications*, volume 4269, pages 32–40, P.O. Box 10, Bellingham, Washington 98227-0010 USA, 2001. SPIE.

[472] Livio Gianfrani, Richard W. Fox, and Leo Hollberg. Cavity-enhanced absorption spectroscopy of molecular oxygen. *J. Opt. Soc. Am. B*, 16:2247–2254, 1999.

[473] Sarah S. Gilbert, William C. Swann, and Tasshi Dennis. Wavelength standards for optical communications. In John L. Hall and Jun Ye, editors, *Proceedings of SPIE: Laser Frequency Stabilization, Standards, Measurement, and Applications*, volume 4269, pages 184–191, P.O. Box 10, Bellingham, Washington 98227 - 0010 USA, 2001. SPIE.

[474] Y. Millerioux, D. Touahri, L. Hilico, A. Clairon, R. Felder, F. Biraben, and B. de Beauvoir. Towards an accurate frequency standard at λ = 778 nm using a laser diode stabilized on a hyperfine component of the Doppler-free two-photon transitions in rubidium. *Opt. Commun.*, 108:91–96, 1994.

[475] D. Touahri, O. Acef, A. Clairon, J.-J. Zondy, R. Felder, L. Hilico, B. de Beauvoir, F. Biraben, and F. Nez. Frequency measurement of the $5S_{1/2}$ (F = 3) – $5D_{5/2}$ (F = 5) two-photon transition in rubidium. *Opt. Commun.*, 133:471–478, 1997.

[476] J. L. Hall, J. Ye, L.-S. Ma, S. Swartz, P. Jungner, and Steve Waltman. Optical frequency standards – some improvements, some measurements, and some dreams. In J. C. Bergquist, editor, *Proceedings of the fifth symposium on Frequency Standards and Metrology*, pages 267–276, Singapore, 1996. World Scientific.

[477] G. Hagel, C. Nesi, L. Jozefowski, C. Schwob, F. Nez, and F. Biraben. Accurate measurement of the frequency of the 6S–8S two-photon transitions in cesium. *Opt. Commun.*, 160:1–4, 1999.

[478] N. Beverini and F. Strumia. High precision measurements of the Zeeman effect in the Calcium metastable states. In *Interaction of Radiation with Matter, A Volume in honour of A. Gozzini*, Quaderni della Scuola Normale Superiore de Pisa, pages 361–373, Pisa, 1987.

[479] K. Zeiske, G. Zinner, F. Riehle, and J. Helmcke. Atom interferometry in a static electric field: Measurement of the Aharonov–Casher phase. *Appl. Phys. B*, 60:205–209, 1995.

[480] G. M. Tino, M. Barsanti, M. de Angelis, L. Gianfrani, and M. Inguscio. Spectroscopy on the 689 nm intercombination line of Strontium using an extended-cavity InGaP/InGaAlP diode laser. *Appl. Phys. B*, 55:397–400, 1992.

[481] A. Celikov, P. Kersten, F. Riehle, G. Zinner, L. D'Evelyn, A. Zibrov, V. L. Velichansky, and J. Helmcke. External cavity diode laser high resolution spectroscopy of the Ca and Sr intercombination lines for the development of a transportable frequency/length standard. In *Proceedings of the 49th Annual IEEE International Frequency Control Symposium, 31 May – 2 June 1995, San Francisco, USA*, pages 153–160, 1995.

[482] A. M. Akulshin, A. A. Celikov, and V. L. Velichansky. Nonlinear Doppler-free spectroscopy of the $6\,^1S_0$ – $6\,^3P_1$ intercombination transition in barium. *Opt. Commun.*, 93:54–58, 1992.

[483] A. Morinaga, F. Riehle, J. Ishikawa, and J. Helmcke. A Ca optical frequency standard: Frequency stabilization by means of nonlinear Ramsey resonances. *Appl. Phys. B*, 48:165–171, 1989.

[484] P. Kersten, F. Mensing, U. Sterr, and F. Riehle. A transportable optical calcium frequency standard. *Appl. Phys. B*, 68:27–38, 1999.

[485] N. Ito, J. Ishikawa, and A. Morinaga. Frequency locking a dye laser to the central optical Ramsey fringe in a Ca atomic beam and wavelength measurement. *J. Opt. Soc. Am. B*, 8:1388–1390, 1991.

[486] N. Ito, J. Ishikawa, and A. Morinaga. Evaluation of the optical phase shift in a Ca Ramsey fringe stabilized optical frequency standard by means of laser-beam reversal. *Opt. Commun.*, 109:414–421, 1994.

[487] A. S. Zibrov, R. W. Fox, R. Ellingsen, C. S. Weimer, V. L. Velichansky, G. M. Tino, and L. Hollberg. High-resolution diode-laser spectroscopy of calcium. *Appl. Phys. B*, 59:327–331, 1994.

[488] N. Beverini, E. Maccioni, D. Pereira, F. Strumia, and G. Vissani. Production of low-velocity Mg and Ca atomic beams by laser light pressure. In G. C. Righini, editor, *Quantum Electronics and Plasma Physics 5th Italian Conference*, pages 205–211, Bologna, Italy, 1988. Italian Physical Society.

[489] K. Sengstock, U. Sterr, G. Hennig, D. Bettermann, J. H. Müller, and W. Ertmer. Optical Ramsey interferences on laser cooled and trapped atoms, detected by electron shelving. *Opt. Commun.*, 103:73–78, 1993.

[490] F. Ruschewitz, J. L. Peng, H. Hinderthür, N. Schaffrath, K. Sengstock, and W. Ertmer. Sub-kilohertz optical spectroscopy with a time domain atom interferometer. *Phys. Rev. Lett.*, 80:3173–3176, 1998.

[491] T. Kurosu and F. Shimizu. Laser cooling and trapping of alkaline earth atoms. *Jpn. J. Appl. Phys.*, 31:908–912, 1992.

[492] T. P. Dinneen, K. R. Vogel, E. Arimondo, J. L. Hall, and A. Gallagher. Cold collisions of Sr^*-Sr in a magneto-optical trap. *Phys. Rev. A*, 59:1216–1222, 1999.

[493] Hidetoshi Katori, Tetsuya Ido, Yoshitomo Isoya, and Makoto Kuwata-Gonokami. Laser cooling of strontium atoms toward quantum degeneracy. In E. Arimondo, P. DeNatale, and M. Inguscio, editors, *Atomic Physics*, volume XVII, pages 382–396, Woodbury, New York, 2001. American Institute of Physics.

[494] T. Binnewies, U. Sterr, J. Helmcke, and F. Riehle. Cooling by Maxwell's demon: Preparation of single-velocity atoms for matter-wave interferometry. *Phys. Rev. A*, 62:011601(R)–1–4, 2000.

[495] F. Riehle, H. Schnatz, B. Lipphardt, G. Zinner, T. Trebst, and J. Helmcke. The optical calcium frequency standard. *IEEE Trans. Instrum. Meas.*, 48:613–617, 1999.

[496] T. Kurosu, G. Zinner, T. Trebst, and F. Riehle. Method for quantum-limited detection of narrow-linewidth transitions in cold atomic ensembles. *Phys. Rev. A*, 58:R4275–R4278, 1998.

[497] Guido Wilpers, Tomas Binnewies, Carsten Degenhardt, Uwe Sterr, Jürgen Helmcke, and Fritz Riehle. Optical clock with ultracold neutral atoms. *Phys. Rev. Lett.*, 89:230801–1–4, 2002.

[498] W. Nagourney, J. Sandberg, and H. Dehmelt. Shelved optical electron amplifier: Observation of quantum jumps. *Phys. Rev. Lett.*, 56:2797–2799, 1986.

[499] Leo Hollberg, Chris W. Oates, E. Anne Curtis, Eugene N. Ivanov, Scott A. Diddams, Thomas Udem, Hugh G. Robinson, James C. Bergquist, Robert J. Rafac, Wayne M. Itano, Robert E. Drullinger, and D. J. Wineland. Optical frequency standards and measurements. *IEEE J. Quantum Electron.*, 37:1502–1513, 2001.

[500] F. Riehle, H. Schnatz, B. Lipphardt, G. Zinner, T. Trebst, T. Binnewies, G. Wilpers, and J. Helmcke. The optical Ca frequency standard. In *Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum and The IEEE International Frequency Control Symposium*, pages 700–705, 26 Chemin de l'Epitaphe, 25030 BESANCON CEDEX - FRANCE, 1999. EFTF co/Société Francaise des Microtechniques et de Chronométrie (SFMC).

[501] Th. Udem, S. A. Diddams, K. R. Vogel, C. W. Oates, E. A. Curtis, W. D. Lee, W. M. Itano, R. E. Drullinger, J. C. Bergquist, and L. Hollberg. Absolute frequency measurement of the Hg^+ and Ca optical clock transitions with a femtosecond laser. *Phys. Rev. Lett.*, 86:4996–4999, 2001.

[502] G. Wilpers. Ein Optisches Frequenznormal mit kalten und ultrakalten Atomen. PTB-Bericht PTB-Opt–66 (ISBN 3-89701-892-6), Physikalisch-Technische Bundesanstalt, Braunschweig, 2002. Dissertation, University of Hannover.

[503] G. Wilpers, C. Degenhardt, T. Binnewies, A. Chernyshov, F. Riehle, J. Helmcke, and U. Sterr. Improvement of the fractional uncertainty of a neutral atom calcium optical frequency standard to $2 \cdot 10^{-14}$. *Appl. Phys. B*, 76:149–156, 2003.

[504] H. Schnatz, B. Lipphardt, J. Helmcke, F. Riehle, and G. Zinner. First phase-coherent frequency measurement of visible radiation. *Phys. Rev. Lett.*, 76:18–21, 1996.

[505] J. Stenger, T. Binnewies, G. Wilpers, F. Riehle, H. R. Telle, J. K. Ranka, R. S. Windeler, and A. J. Stentz. Phase-coherent frequency measurement of the Ca intercombination line at 657 nm with a Kerr-lens mode-locked laser. *Phys. Rev. A*, 63:021802(R), 2001.

[506] T. Trebst, T. Binnewies, J. Helmcke, and F. Riehle. Suppression of spurious phase shifts in an optical frequency standard. *IEEE Trans. Instrum. Meas.*, 50:535–538, 2001.

[507] T. Binnewies. Neuartige Kühlverfahren zur Erzeugung ultrakalter Ca-Atome. PTB-Bericht PTB-Opt-65, Physikalisch–Technische Bundesanstalt, Braunschweig, 2001.

[508] E. Anne Curtis, Christopher W. Oates, and Leo Hollberg. Quenched narrow-line second- and third-stage laser cooling of ^{40}Ca. *J. Opt. Soc. Am. B*, 20:977–984, 2003.

[509] B. Gross, A. Huber, M. Niering, M. Weitz, and T. W. Hänsch. Optical Ramsey spectroscopy of atomic hydrogen. *Europhys. Lett.*, 44:186–191, 1998.

[510] F. Schmidt-Kaler, D. Leibfried, S. Seel, C. Zimmermann, W. König, M. Weitz, and T. W. Hänsch. High-resolution spectroscopy of the $1S - 2S$ transition of atomic hydrogen and deuterium. *Phys. Rev. A*, 51:2789–2800, 1995.

[511] R. G. Beausoleil and T. W. Hänsch. Ultrahigh-resolution two-photon optical Ramsey spectroscopy of an atomic fountain. *Phys. Rev. A*, 33:1661–1670, 1986.

[512] I. D. Setija, H. G. C. Werij, O. J. Luiten, M. W. Reynolds, T. W. Hijmans, and J. T. M. Walraven. Optical cooling of atomic hydrogen in a magnetic trap. *Phys. Rev. Lett.*, 70:2257–2260, 1993.

[513] K. S. E. Eikema, J. Walz, and T. W. Hänsch. Continuous wave coherent Lyman-α radiation. *Phys. Rev. Lett.*, 83:3828–3831, 1999.

[514] W. Ertmer, R. Blatt, and J. L. Hall. Some candidate atoms and ions for frequency standards research using laser radiative cooling techniques. In W. D. Phillips, editor, *Laser Cooled and Trapped Atoms*, pages 154–161. U.S. National Bureau of Standards special publication Vol. 653, Reading, Massachusetts, 1983.

[515] T. Badr, S. Guérandel, Y. Louyer, S. Challemel Du Rozier, M. D. Plimmer, P. Juncar, and M. E. Himbert. Towards a silver atom optical clock. In P. Gill, editor, *Frequency Standards and Metrology, Proceedings of the Sixth Symposium*, pages 549–551, Singapore, 2002. World Scientific.

[516] J. Dirscherl and H. Walther. Towards a silver frequency standard. In *Digest of the 14th International Conference on Atomic Physics (ICAP 94)*, page Poster 1H3, Boulder, 1994.

[517] H. G. Dehmelt. Mono-ion oscillator as potential ultimate laser frequency standard. *IEEE Trans. Instrum. Meas.*, IM-31:83–87, 1982.

[518] R. Blatt, P. Gill, and R. C. Thompson. Current perspectives on the physics of trapped ions. *J. Mod. Opt.*, 39:193–220, 1992.

[519] R. C. Thompson. Spectroscopy of trapped ions. In D. Bates and B. Bederson, editors, *Advances in Atomic, Molecular, and Optical Physics*, volume 31, pages 63–136, Boston, 1993. Academic Press.

[520] P. T. H. Fisk. Trapped-ion and trapped-atom microwave frequency standards. *Rep. Prog. Phys.*, 60:761–817, 1997.

[521] Alan A. Madej and John E. Bernard. Single-ion optical frequency standards and measurement of their absolute optical frequency. In Andre N. Luiten, editor, *Frequency Measurement and Control*, volume 79 of *Topics in Applied Physics*, pages 153–194. Springer, Berlin, Heidelberg, New York, 2001.

[522] W. Paul and M. Raether. Das elektrische Massenfilter. *Z. Phys.*, 140:262–273, 1955.

[523] F. M. Penning. Die Glimmentladung bei niedrigem Druck zwischen koaxialen Zylindern in einem axialen Magnetfeld. *Physica III*, 9:873–894, 1936.

[524] H. G. Dehmelt. Radiofrequency spectroscopy of stored ions I: Storage. In D. R. Bates and I. Estermann, editors, *Advances in Atomic and Molecular Physics*, volume 3, pages 53–72. Academic Press, New York, London, 1967.

[525] W. Paul. Electromagnetic traps for charged and neutral particles. *Rev. Mod. Phys.*, 62:531–540, 1990.

[526] T. Tamir. Characteristic exponents of Mathieu functions. *Math. Comp.*, XVI:100–106, 1962.

[527] D. A. Church. Storage-ring ion trap derived from the linear quadrupole radio-frequency mass filter. *J. Appl. Phys.*, 40:3127–3134, 1969.

[528] I. Waki, S. Kassner, G. Birkl, and H. Walther. Observation of ordered structures of laser-cooled ions in a quadrupole storage ring. *Phys. Rev. Lett.*, 68:2007–2010, 1992.

[529] D. J. Berkeland, J. D. Miller, J. C. Bergquist, W. M. Itano, and D. J. Wineland. Laser-cooled mercury ion frequency standard. *Phys. Rev. Lett.*, 80:2089–2092, 1998.

[530] M. G. Raizen, J. M. Gilligan, J. C. Bergquist, W. M. Itano, and D. J. Wineland. Ionic crystals in a linear Paul trap. *Phys. Rev. A*, 45:6493–6501, 1992.

[531] J. D. Prestage, G. J. Dick, and L. Maleki. New ion trap for frequency standard applications. *J. Appl. Phys.*, 66:1013–1017, 1989.

[532] P. T. H. Fisk, M. J. Sellars, M. A. Lawn, and C. Coles. Performance of a prototype microwave frequency standard based on laser-detected, trapped $^{171}Yb^{+}$ ions. *Appl. Phys. B*, 60:519–527, 1995.

[533] E. Fischer. Die dreidimensionale Stabilisierung von Ladungsträgern in einem Vierpolfeld. *Z. Physik*, 156:1–26, 1959.

[534] R. S. Van Dyck, Jr., P. B. Schwinberg, and H. G. Dehmelt. Electron magnetic moment from geonium spectra: Early experiments and background concepts. *Phys. Rev. D*, 34:722–736, 1986.

[535] L. S. Brown and G. Gabrielse. Precision spectroscopy of a charged particle in an imperfect Penning trap. *Phys. Rev. A*, 25:2423–2425, 1982.

[536] J. N. Tan, J. J. Bollinger, and D. J. Wineland. Minimizing the time-dilation shift in Penning trap atomic clocks. *IEEE Trans. Instrum. Meas.*, IM 44:144–147, 1995.

[537] J. J. Bollinger, J. D: Prestage, W. M. Itano, and D. J. Wineland. Laser-cooled-atomic frequency standard. *Phys. Rev. Lett.*, 54:1000–1003, 1985.

[538] F. Plumelle, M. Desaintfuscien, M. Jardino, and P. Petit. Laser cooling of magnesium ions: Preliminary experimental results. *Appl. Phys. B*, 41:183–186, 1986.

[539] R. C. Thompson, G. P. Barwood, and P. Gill. Progress towards an optical frequency standard based on ion traps. *Appl. Phys. B*, 46:87–93, 1988.

[540] H. Walther. Phase transitions of stored laser-cooled ions. In D. Bates and B. Bederson, editors, *Advances in Atomic, Molecular, and Optical Physics*, volume 31, pages 137–182, Boston, 1993. Academic Press.

[541] F. Diedrich, E. Peik, J. M. Chen, W. Quint, and H. Walther. Observation of a phase transition of stored laser-cooled ions. *Phys. Rev. Lett.*, 59:2931–2934, 1987.

[542] D. J. Wineland, J. C. Bergquist, W. M. Itano, J. J. Bollinger, and C. H. Manney. Atomic-ion Coulomb clusters in an ion trap. *Phys. Rev. Lett.*, 59:2935–2938, 1987.

[543] J. N. Tan, J. J. Bollinger, B. Jelenkovic, and D. J. Wineland. Long-range order in laser-cooled, atomic-ion Wigner crystals observed by Bragg scattering. *Phys. Rev. Lett.*, 75:4198–4201, 1995.

[544] M. Drewsen, C. Brodersen, L. Hornekaer, J. S. Hangst, and J. P. Schiffer. Large ion crystals in a linear Paul trap. *Phys. Rev. Lett.*, 81:2878–2881, 1998.

[545] R. Alheit, C. Hennig, R. Morgenstern, F. Vedel, and G. Werth. Observation of instabilities in a Paul trap with higher-order anharmonicities. *Appl. Phys. B*, 61:277–283, 1995.

[546] Th. Gudjons, F. Kurth, P. Seibert, and G. Werth. Ca^+ in an Paul trap. In *Proceedings of the workshop frequency standards based on laser-manipulated atoms and ions*, volume Opt. 51, pages 59–66, Braunschweig, 1996.

[547] R. Alheit, S. Kleineidam, F. Vedel, M. Vedel, and G. Werth. Higher-order non-linear resonances in a Paul trap. *Int. J. Mass Spectrom. Ion Processes*, 154:155–169, 1996.

[548] H. Schnatz, G. Bollen, P. Dabkiewicz, P. Egelhof, F. Kern, H. Kalinowsky, L. Schweikhard, H. Stolzenberg, and H.-J. Kluge. In-flight capture of ions into a Penning trap. *Nucl. Instrum. Meth.*, A 251:17–20, 1986.

[549] R. B. Moore and G. Rouleau. In-flight capture of an ion beam in a Paul trap. *J. Mod. Opt.*, 39:361–371, 1992.

[550] G. Gabrielse, X. Fei, K. Helmerson, S. L. Rolston, R. Tjoelker, T. A. Trainor, H. Kalinowsky, J. Haas, and W. Kells. First capture of antiprotons in a Penning trap: A kiloelectronvolt source. *Phys. Rev. Lett.*, 57:2504–2507, 1986.

[551] J. D. Miller, D. J. Berkeland, J. C. Bergquist F. C. Cruz, W. M. Itano, and D. J. Wineland. A cryogenic linear ion trap for $^{199}Hg^+$ frequency standards. In *Proceedings of the 1996 IEEE International Frequency Control Symposium*, volume IEEE catalog number 96CH35935, 36CB35935, pages 1086–1088, IEEE Service Center, Piscataway, NJ, 1996.

[552] L. S. Cutler, R. P. Giffard, and M. D. McGuire. Thermalization of ^{199}Hg ion macromotion by a light background gas in an rf quadrupole trap. *Appl. Phys. B*, 36:137–142, 1985.

[553] M. H. Holzscheiter. Cooling of particles stored in electromagnetic traps. *Physica Scripta*, T22:73–78, 1988.

[554] W. M. Itano, J. C. Bergquist, J. J. Bollinger, and D. J. Wineland. Cooling methods in ion traps. *Phys. Scripta*, T59:106–120, 1995.

[555] H. G. Dehmelt. Radiofrequency spectroscopy of stored ions II: Spectroscopy. In D. R. Bates and I. Estermann, editors, *Advances in Atomic and Molecular Physics*, volume 5, pages 109–154. Academic Press, New York, London, 1982.

[556] N. Beverini, V. Lagomarsino, G. Manuzio, F. Scuri, G. Testera, and G. Torelli. Experimental verification of stochastic cooling in a Penning trap. *Physica Scripta*, T22:238–239, 1988.

[557] R. F. Wuerker, H. Shelton, and R. V. Langmuir. Electrodynamic containment of charged particles. *J. Appl. Phys.*, 30:342–349, 1959.

[558] F. G. Major and H. G. Dehmelt. Exchange-collision technique for the rf spectroscopy of stored ions. *Phys. Rev.*, 170:91–107, 1968.

[559] A. Bauch, D. Schnier, and Chr. Tamm. Microwave spectroscopy of ^{171} Yb^{+} stored in a Paul trap. In J. C. Bergquist, editor, *Proceedings of the Fifth Symposium on Frequency Standards and Metrology*, pages 387–388, Singapore, New Jersey, London, Hong Kong, 1996. World Scientific.

[560] W. Neuhauser, M. Hohenstatt, P. Toschek, and H. Dehmelt. Optical-sideband cooling of visible atom cloud confined in parabolic well. *Phys. Rev. Lett.*, 41:233–236, 1978.

[561] D. J. Wineland, R. E. Drullinger, and F. L. Walls. Radiation-pressure cooling of bound resonant absorbers. *Phys. Rev. Lett.*, 40:1639–1642, 1978.

[562] L. R. Brewer, J. D. Prestage, J. J. Bollinger, W. M. Itano, D. J. Larson, and D. J. Wineland. Static properties of a non-neutral ^{9}Be^{+}-ion plasma. *Phys. Rev. A*, 38:859–873, 1988.

[563] Q. A. Turchette, D. Kielpinski, B. E. King, D. Leibfried, D. M. Meekhof, C. J. Myatt, M. A. Rowe, C. A. Sackett, C. S. Wood, W. M. Itano, C. Monroe, and D. J. Wineland. Heating of trapped ions from the quantum ground state. *Phys. Rev. A*, 61:063418–1–8, 2000.

[564] F. Diedrich, J. C. Bergquist, W. M. Itano, and D. J. Wineland. Laser cooling to the zero-point energy of motion. *Phys. Rev. Lett.*, 62:403–406, 1989.

[565] R. E. Drullinger, D. J. Wineland, and J. C. Bergquist. High-resolution optical spectra of laser cooled ions. *Appl. Phys.*, 22:365–368, 1980.

[566] D. J. Larson, J. C. Bergquist, J. J. Bollinger, W. M. Itano, and D. J. Wineland. Sympathetic cooling of trapped ions: A laser-cooled two-species nonneutral ion plasma. *Phys. Rev. Lett.*, 57:70–73, 1986.

[567] H. G. Dehmelt and F. L. Walls. "Bolometric" technique for the rf spectroscopy of stored ions. *Phys. Rev. Lett.*, 21:127–131, 1968.

[568] R. M. Weisskoff, G. P. Lafyatis, K. R. Boyce, E. A. Cornell, R. W. Flanagan, Jr., and D. E. Pritchard. rf SQUID detector for single-ion trapping experiments. *J. Appl. Phys.*, 63:4599–4604, 1988.

[569] S. R. Jefferts, T. Heavner, P. Hayes, and G. H. Dunn. Superconducting resonator and a cryogenic GaAs field-effect transistor amplifier as a single-ion detection system. *Rev. Sci. Instrum.*, 64:737–740, 1993.

[570] R. Iffländer and G. Werth. Optical detection of ions confined in a rf quadrupole trap. *Metrologia*, 13:167–170, 1977.

[571] D. J. Wineland, J. C. Bergquist, W. M. Itano, and R. E. Drullinger. Double-resonance and optical-pumping experiments on electromagnetically confined, laser-cooled ions. *Opt. Lett.*, 5:245–247, 1980.

[572] E. Peik. Laserspektroskopie an gespeicherten Indium-Ionen. Dissertation MPQ 181, Max-Planck-Institut für Quantenoptik, 1993.

[573] D. J. Bate, K. Dholakia, R. C. Thompson, and D. C. Wilson. Ion oscillation frequencies in a combined trap. *J. Mod. Opt.*, 39:305–316, 1992.

[574] E. C. Beaty. Simple electrodes for quadrupole ion traps. *J. Appl. Phys.*, 61:2118–2122, 1987.

[575] W. Neuhauser, M. Hohenstatt, P. E. Toschek, and H. Dehmelt. Localized visible Ba^+ mono-ion oscillator. *Phys. Rev. A*, 22:1137–1140, 1980.

[576] J. C. Bergquist, D. J. Wineland, W. M. Itano, H. Hemmati, H.-U. Daniel, and G. Leuchs. Energy and radiative lifetime of the $5d^9 6s^2\ ^2\mathrm{D}_{5/2}$ state in Hg II by Doppler-free two-photon laser spectroscopy. *Phys. Rev. Lett.*, 55:1567–1570, 1985.

[577] Chr. Tamm and D. Engelke. Optical frequency standard investigations on trapped, laser-cooled ^{171}Yb ions. In J. C. Bergquist, editor, *Proceedings of the Fifth Symposium on Frequency Standards and Metrology*, pages 283–288, Singapore, New Jersey, London, Hong Kong, 1996. World Scientific.

[578] H. Straubel. Zum Öltröpfchenversuch von Millikan. *Naturwiss.*, 42:506–507, 1955.

[579] N. Yu, W. Nagourney, and H. Dehmelt. Demonstration of new Paul-Straubel trap for trapping single ions. *J. Appl. Phys.*, 69:3779–3781, 1991.

[580] R. G. Brewer, R. G. DeVoe, and R. Kallenbach. Planar ion microtraps. *Phys. Rev. A*, 46:6781–6784, 1992.

[581] C. A. Schrama, E. Peik, W. W. Smith, and H. Walther. Novel miniature ion traps. *Opt. Commun.*, 101:32–36, 1993.

[582] J. Walz, I. Siemers, M. Schubert, W. Neuhauser, R. Blatt, and E. Teloy. Ion storage in the rf octupole trap. *Phys. Rev. A*, 50:4122–4132, 1994.

[583] J. D. Prestage, R. L. Tjoelker, and L. Maleki. Hg^+ frequency standards. In Daniel H. E. Dubin and Dieter Schneider, editors, *Trapped charged particles and fundamental physics*, volume 457 of *AIP Conference Proceedings*, pages 357–363, American Institute of Physics, Woodbury, New York, 1999.

[584] John D. Prestage, Robert L. Tjoelker, and Lute Maleki. Higher pole linear traps for atomic clock applications. In *Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum and The IEEE International Frequency Control Symposium*, pages 121–124, 1999.

[585] G. Werth. Hyperfine structure and g-factor measurements in ion traps. *Physica Scripta*, T59:206–210, 1995.

[586] J. J. Bollinger, S. L. Gilbert, W. M. Itano, and D. J. Wineland. Frequency standards utilizing Penning traps. In A. De Marchi, editor, *Frequency Standards and Metrology*, pages 319–325, Berlin, Heidelberg, New York, 1989. Springer.

[587] F. Arbes, M. Benzing, T. Gudjons, F. Kurth, and G. Werth. Precise determination of the ground state hyperfine structure splitting of ^{43}Ca II. *Z. Phys. D*, 29:27–30, 1994.

[588] R. Blatt and G. Werth. Precision determination of the ground-state hyperfine splitting in $^{137}Ba^{+}$ using the ion-storage technique. *Phys. Rev. A*, 25:1476–1482, 1982.

[589] H. Knab, K.-D. Niebling, and G. Werth. Ion trap as a frequency standard. Measurement of Ba^{+} HFS frequency fluctuations. *IEEE Trans. Instrum. Meas.*, IM–34:242–245, 1985.

[590] U. Tanaka, H. Imajo, K. Hayasaka, R. Ohmukai, M. Watanabe, and S. Urabe. Laser microwave double-resonance experiment on trapped $^{113}Cd^{+}$ ions. *IEEE Trans. Instrum. Meas.*, IM 46:137–140, 1997.

[591] R. B. Warrington, P. T. H. Fisk, M. J. Wouters, and M. A. Lawn. A microwave frequency standard based on laser -cooled ^{171} Yb^{+} ions. In Patrick Gill, editor, *Proceedings of the Sixth Symposium on Frequency Standards and Metrology*, pages 297–304, New Jersey, London, Singapore, Hong Kong, 2002. World Scientific.

[592] C. Tamm, D. Schnier, and A. Bauch. Radio-frequency laser double-resonance spectroscopy of trapped ^{171}Yb ions and determination of line shifts of the ground-state hyperfine resonance. *Appl. Phys. B*, 60:19–29, 1995.

[593] P. T. H. Fisk, M. J. Sellars, M. A. Lawn, and C. Coles. Accurate measurement of the 12.6 GHz "clock" transition in trapped $^{171}Yb^{+}$ ions. *IEEE Trans. Ultrason. Ferroelec. Freq. Contr.*, 44:344–354, 1997.

[594] J. J. Bollinger, D. J. Heinzen, W. M. Itano, S. L. Gilbert, and D. J. Wineland. A 303 MHz frequency standard based on trapped Be^{+} ions. *IEEE Trans. Instrum. Meas.*, 40:126–128, 1991.

[595] R. Blatt, H. Schnatz, and G. Werth. Ultrahigh-resolution microwave spectroscopy on trapped $^{171}Yb^{+}$ ions. *Phys. Rev. Lett.*, 48:1601–1603, 1982.

[596] F. G. Major and G. Werth. High-resolution magnetic hyperfine resonance in harmonically bound ground-state ^{199}Hg ions. *Phys. Rev. Lett.*, 30:1155–1158, 1973.

[597] R. Casdorff, V. Enders, R. Blatt, W. Neuhauser, and P. E. Toschek. A 12-GHz standard clock on trapped Ytterbium ions. *Ann. Phys.*, 7:41–55, 1991.

[598] K. Sugiyama and J. Yoda. Study of Yb^{+} trapped in a rf trap with light buffer gas by irradiation with resonant light. *IEEE Trans. Instrum. Meas.*, 42:467–473, 1993.

[599] K. Sugiyama and J. Yoda. Characteristics of buffer-gas-cooled and laser-cooled Yb^{+} in rf traps. In J. C. Bergquist, editor, *Proceedings of the Fifth Symposium on Frequency Standards and Metrology*, pages 432–433, Singapore, New Jersey, London, Hong Kong, 1996. World Scientific.

[600] K. Sugiyama and J. Yoda. Production of YbH^{+} by chemical reaction of Yb^{+} in excited states with H_2 gas. *Phys. Rev. A*, 55:R10–R13, 1997.

[601] D. J. Seidel and L. Maleki. Efficient quenching of population trapping in excited Yb^{+}. *Phys. Rev. A*, 51:2699–2702, 1995.

[602] R. Blatt, R. Casdorff, V. Enders, W. Neuhauser, and P. E. Toschek. New frequency standards based on Yb^{+}. In A. De Marchi, editor, *Frequency Standards and Metrology*, pages 306–311, Berlin, Heidelberg, New York, 1989. Springer-Verlag.

[603] A. Bauch, D. Schnier, and Chr. Tamm. Collisional population trapping and optical deexcitation of ytterbium ions in a radiofrequency trap. *J. Mod. Opt.*, 39:389–401, 1992.

[604] P. Gill, H. A. Klein, A. P. Levick, M. Roberts, W. R. C. Rowley, and P. Taylor. Measurement of the $^2S_{1/2}$–$^2D_{5/2}$ 411-nm interval in laser-cooled trapped $^{172}Yb^+$ ions. *Phys. Rev. A*, 52:R909–R912, 1995.

[605] P. T. H. Fisk, M. A. Lawn, and C. Coles. Laser cooling of $^{171}Yb^+$ ions in a linear Paul trap. *Appl. Phys. B*, 57:287–291, 1993.

[606] V. Enders, Ph. Courteille, R. Huesmann, L. S. Ma, W. Neuhauser, R. Blatt, and P. E. Toschek. Microwave-optical double resonance on a single laser-cooled $^{171}Yb^+$ ion. *Europhys. Lett.*, 24:325–331, 1993.

[607] M. Jardino, M. Desaintfuscien, R. Barillet, J. Viennet, P. Petit, and C. Audoin. Frequency stability of a mercury ion frequency standard. *Appl. Phys.*, 24:107–112, 1981.

[608] C. Meis, M. Jardino, B. Gely, and M. Desaintfuscien. Relativistic Doppler effect in $^{199}Hg^+$ stored ions atomic frequency standard. *Appl. Phys. B*, 48:67–72, 1989.

[609] L. S. Cutler, R. P. Giffard, and M. D. McGuire. A trapped mercury 199 ion frequency standard. In *Proceedings of the 13th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, December 1-3, 1981, Washington DC, USA*, volume 2220 of *NASA Conference Publication*, pages 563–578, U.S. Naval Observatory, Time Service, 3450 Massachusetts Ave., N.W. Washington, DC 20392-5420, USA, 1981.

[610] L. S. Cutler, R. P. Giffard, P. J. Wheeler, and G. M. R. Winkler. Initial operational experience with a mercury ion storage frequency standard. In *Proceedings of the 41st Annual Frequency Control Symposium May 27-29, 1987, Philadelphia*, pages 12–17, National Technical Information Service, Springfield, VA 22161, USA, 1987.

[611] D. N. Matsakis, A. J. Kubik, J. A. DeYoung, R. P. Giffard, and L. S. Cutler. Eight years of experience with mercury stored ion devices. In *Proceedings of the 49th Annual IEEE International Frequency Control Symposium, 31 May – 2 June 1995, San Francisco, USA*, pages 86–108, 1995.

[612] R. L. Tjoelker, J. D. Prestage, and L. Maleki. Record frequency stability with mercury in a linear ion trap. In J. C. Bergquist, editor, *Proceedings of the Fifth Symposium on Frequency Standards and Metrology*, volume 31, pages 33–38, Singapore, New Jersey, London, Hong Kong, 1996. World Scientific.

[613] M. G. Raizen, J. M. Gilligan, J. C. Bergquist, W. M. Itano, and D. J. Wineland. Linear trap for high-accuracy spectroscopy of stored ions. *J. Mod. Opt.*, 39:233–242, 1992.

[614] Alan A. Madej, Klaus J. Siemsen, John D. Sankey, Richard F. Clark, and Jacques Vanier. High-resolution spectroscopy and frequency measurement of the midinfrared $5d^2D_{3/2}$–$5d^2D_{5/2}$ transition of a single laser-cooled barium ion. *IEEE Trans. Instrum. Meas.*, IM–42:234–241, 1993.

[615] A. A. Madej, K. J. Siemsen, B. G. Whitford, J. E. Bernard, and L. Marmet. Precision absolute frequency measurements with single atoms of Ba^+ and Sr^+. In J. C. Bergquist, editor, *Proceedings of the Fifth Symposium on Frequency Standards and Metrology*, pages 165–170, World Scientific, Singapore, New Jersey, London, Hong Kong, 1996.

[616] W. Nagourney, N. Yu, and H. Dehmelt. High resolution Ba^+ monoion spectroscopy with frequency stabilized color-center laser. *Opt. Commun.*, 79:176–180, 1990.

[617] H. S. Margolis, G. Huang, G. P. Barwood, S. N. Lea, H. A. Klein, W. R. C. Rowley, P. Gill, and R. S. Windeler. Absolute frequency measurement of the 674-nm $^{88}Sr^{+}$ clock transition using a femtosecond optical frequency comb. *Phys. Rev. A*, 67:032501–1–5, 2003.

[618] A. A. Madej, J. E. Bernard, P. Dubé, L. Marmet, and R. S. Windeler. Absolute frequency measurement of the $^{88}Sr^{+}$, 5s $^{2}S_{1/2}$ – 4d $^{2}D_{5/2}$ reference transition at 445 THz and evaluation of systematic shifts. *Phys. Rev. A*, 2004. Accepted for publication.

[619] S. Urabe, K. Hayasaka, M. Watanabe, H. Imajo, and R. Ohmukai. Laser cooling of Ca^{+} ions and observation of collision effects. *Jpn. J. Appl. Phys.*, 33:1590–1594, 1994.

[620] M. Knoop, M. Vedel, and F. Vedel. Lifetime, collisional-quenching, and j-mixing measurements of the metastable ^{3}D levels of Ca^{+}. *Phys. Rev. A*, 52:3763–3769, 1995.

[621] G. Ritter and U. Eichmann. Lifetime of the Ca^{+} $3^{2}D_{5/2}$ level from quantum jump statistics of a single laser-cooled ion. *J. Phys. B*, 30:L141–L146, 1997.

[622] Chr. Tamm, D. Engelke, and V. Bühner. Spectroscopy of the electric-quadrupole transition $^{2}S_{1/2}$ (F = 0) - $^{2}D_{3/2}$ (F = 2) in trapped $^{171}Yb^{+}$. *Phys. Rev. A*, 61:053405–1–9, 2000.

[623] Chr. Tamm, T. Schneider, and E. Peik. Comparison of two single-ion optical frequency standards at the sub-hertz level. In P. Hannaford, A. Sidorov, H. Bachor, and K. Baldwin, editors, *Laser Spectroscopy, Proceedings of the XVI International Conference*, pages 40–48, New Jersey, 2004. World Scientific. eprint physics/0402120.

[624] P. Taylor, M. Roberts, S. V. Gateva-Kostova, R. B. M. Clarke, G. P. Barwood, W. R. C. Rowley, and P. Gill. Investigation of the $^{2}S_{1/2}$–$^{2}D_{5/2}$ clock transition in a single ytterbium ion. *Phys. Rev. A*, 56:2699–2704, 1997.

[625] M. Roberts, P. Taylor, S. V. Gateva-Kostova, R. B. M. Clarke, W. R. C. Rowley, and P. Gill. Measurement of the $^{2}S_{1/2} - ^{2} D_{5/2}$ clock transition in a single $^{171}Yb^{+}$ ion. *Phys. Rev. A*, 60:2867–2872, 1999.

[626] Jörn Stenger, Christian Tamm, Nils Haverkamp, Stefan Weyers, and Harald R. Telle. Absolute frequency measurement of the 435.5-nm $^{171}Yb^{+}$-clock transition with a Kerr-lens mode-locked femtosecond laser. *Opt. Lett.*, 26:1589–1591, 2001.

[627] E. Peik, G. Hollemann, and H. Walther. Laser cooling and quantum jumps of a single indium ion. *Phys. Rev. A*, 49:402–408, 1994.

[628] W. Nagourney, J. Torgerson, and H. Dehmelt. Optical frequency standard based upon single laser-cooled Indium ion. In Daniel H. E. Dubin and Dieter Schneider, editors, *Trapped charged particles and fundamental physics*, volume 457 of *AIP Conference Proceedings*, pages 343–347, Woodbury, New York, 1999. American Institute of Physics.

[629] J. von Zanthier, Th. Becker, M. Eichenseer, A. Yu. Nevsky, Ch. Schwedes, E. Peik, H. Walther, R. Holzwarth, J. Reichert, Th. Udem, T. W. Hänsch, P. V. Pokasov, M. N. Skvortsov, and S. N. Bagayev. Absolute frequency measurement of the In^{+} clock transition with a mode-locked laser. *Opt. Lett.*, 25:1729–1731, 2000.

[630] J. Helmcke, A. Morinaga, J. Ishikawa, and F. Riehle. Optical frequency standards. *IEEE Trans. Instrum. Meas.*, IM 38:524–532, 1989.

[631] G. P. Barwood, G. Huang, H. A. Klein, P. Gill, and R. B. M. Clarke. Subkilohertz comparison of the single-ion optical clock $^2S_{1/2}$ – $^2D_{5/2}$ transition in two $^{88}Sr^+$ traps. *Phys. Rev. A*, 59:R3178–R3181, 1999.

[632] G. P. Barwood, P. Gill, H. A. Klein, and W. R. C. Rowley. Clearly resolved secular sidebands on the $^2S_{1/2}$ – $^2D_{5/2}$ 674-nm clock transition in a single trapped Sr^+ ion. *IEEE Trans. Instrum. Meas.*, 46:133–136, 1997.

[633] G. P. Barwood, K. Gao, P. Gill, G. Huang, and H. A. Klein. Observation of the hyperfine structure of the $^2s_{1/2}-^2d_{5/2}$ transition in $^{87}Sr^+$. *Phys. Rev. A*, 67:013402–1–5, 2003.

[634] Wayne M. Itano. External-field shifts of the $^{199}Hg^+$ optical frequency standard. *J. Res. NIST*, 105:829–837, 2000.

[635] H. Dehmelt. Proposed $10^{14}\delta\nu < \nu$ laser fluorescence spectroscopy on Tl^+ mono-ion oscillator. *Bull. Am. Phys. Soc.*, 18:1521, 1973.

[636] A. Yu. Nevsky, M. Eichenseer, J. von Zanthier, and H. Walther. Narrow linewidth laser system for precise spectroscopy of the indium clock transition. In Patrick Gill, editor, *Proceedings of the Sixth Symposium on Frequency Standards and Metrology*, pages 409–416, New Jersey, London, Singapore, Hong Kong, 2002. World Scientific.

[637] Th. Becker, J. v. Zanthier, A. Yu. Nevsky, Ch. Schwedes, M. N. Skvortsov, H. Walther, and E. Peik. High-resolution spectroscopy of a single In^+ ion: Progress towards an optical frequency standard. *Phys. Rev. A*, 63:051802–1–4, 2001.

[638] J. C. Bergquist, U. Tanaka, R. E. Drullinger, W. M. Itano, D. J. Wineland, S. A. Diddams, L. Hollberg, E. A. Curtis, C. W. Oates, and Th. Udem. A mercury-ion optical clock. In P. Gill, editor, *Frequency Standards and Metrology, Proceedings of the Sixth Symposium*, pages 99–105, Singapore, 2002. World Scientific.

[639] S. Bize, S. A. Diddams, U. Tanaka, C. E. Tanner, W. H. Oskay, R. E. Drullinger, T. E. Parker, T. P. Heavner, S. R. Jefferts, L. Hollberg, W. M. Itano, and J. C. Bergquist. Testing the stability of fundamental constants with the $^{199}hg^+$ single-ion optical clock. *Phys. Rev. Lett.*, 90:150802–1–4, 2003.

[640] R. J. Rafac, B. C. Young, F. C. Cruz, J. A. Beall, J. C. Bergquist, W. M. Itano, and D. J. Wineland. $^{199}Hg^+$ optical frequency standard: Progress report. In *Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum and The IEEE International Frequency Control Symposium*, pages 676–681, 1999.

[641] R. J. Rafac, B. C. Young, J. A. Beall, W. M. Itano, D. J. Wineland, and J. C. Bergquist. Sub-dekahertz ultraviolet spectroscopy of $^{199}Hg^+$. *Phys. Rev. Lett.*, 85:2462–2465, 2000.

[642] M. Block, O. Rehm, P. Seibert, and G. Werth. 3d $^2D_{5/2}$ lifetime in laser cooled Ca^+: Influence of cooling laser power. *Eur. Phys. J. D*, 7:461–465, 1999.

[643] D. J. Wineland, C. Monroe, W. M. Itano, D. Leibfried, B. E. King, and D. M. Meekhof. Experimental issues in coherent quantum-state manipulation of trapped atomic ions. *J. Res. Nat. Inst. Stand. Technol.*, 103:259–328, 1998.

[644] D. J. Wineland, J. C. Bergquist, J. J. Bollinger, R. E. Drullinger, and W. M. Itano. Quantum computers and atomic clocks. In Patrick Gill, editor, *Proceedings of the Sixth Symposium on Frequency Standards and Metrology*, pages 361–368, New Jersey, London, Singapore, Hong Kong, 2002. World Scientific.

[645] G. Audi and A. H. Wapstra. The 1993 atomic mass evaluation. *Nucl. Phys.*, A 565:1–397, 1993.

[646] F. DiFilippo, V. Natarajan, K. R. Boyce, and D. E. Pritchard. Accurate atomic masses for fundamental metrology. *Phys. Rev. Lett.*, 73:1481–1484, 1994.

[647] G. Bollen and The ISOLTRAP Collaboration. Mass determination of radioactive isotopes with the ISOLTRAP spectrometer at ISOLDE, CERN. *Physica Scripta*, T 59:165–175, 1995.

[648] R. S. Van Dyck Jr., F. L. Moore, D. L. Farnham, and P. B. Schwinberg. Mass ratio spectroscopy and the proton's atomic mass. In A. De Marchi, editor, *Frequency Standards and Metrology*, pages 349–355, Berlin, Heidelberg, New York, 1989. Springer-Verlag.

[649] P. B. Schwinberg, R. S. Van Dyck, Jr., and H. G. Dehmelt. New comparison of the positron and electron g factors. *Phys. Rev. Lett.*, 47:1679–1682, 1981.

[650] R. S. Van Dyck Jr., D. L. Farnham, and P. B. Schwinberg. Precision mass measurements in the UW-PTMS and the electron's "atomic mass". *Physica Scripta*, T59:134–143, 1995.

[651] Tomas Beier, Hartmut Häffner, Nikolaus Hermanspahn, Savely Karshenboim, H.-Jürgen Kluge, Wolfgang Quint, Stefan Stahl, Joseé Verdú, and Günther Werth. New determination of the electron's mass. *Phys. Rev. Lett.*, 88:011603–1–4, 2002.

[652] H. Häffner, T. Beier, N. Hermanspahn, H.-J. Kluge, W. Quint, S. Stahl, J. Verdú, and G. Werth. High-accuracy measurement of the magnetic moment anomaly of the electron bound in hydrogenlike carbon. *Phys. Rev. Lett.*, 85:5308–5311, 2000.

[653] Toichiro Kinoshita. Fine-structure constant obtained from an improved calculation of the electron $g-2$. *IEEE Trans. Instrum. Meas.*, IM–46:108–111, 1997.

[654] Toichiro Kinoshita. Improvement of the fine-structure constant obtained from the electron $g-2$. *IEEE Trans. Instrum. Meas.*, IM–50:568–571, 2001.

[655] A.-M. Jeffery, R. E. Elmquist, L. H. Lee, J. Q. Shields, and R. F. Dziuba. NIST comparison of the quantized Hall resistance and the realization of the SI OHM through the calculable capacitor. *IEEE Trans. Instrum. Meas.*, 46:264–268, 1997.

[656] G. Gabrielse, D. Phillips, W. Quint, H. Kalinowsky, G. Rouleau, and W. Jhe. Special relativity and the single antiproton: Fortyfold improved comparison of $\bar{p}$ and p charge-to-mass ratios. *Phys. Rev. Lett.*, 74:3544–3547, 1995.

[657] G. Gabrielse, A. Khabbaz, D. S. Hall, C. Heimann, H. Kalinowsky, and W. Jhe. Precision mass spectroscopy of the antiproton and proton using simultaneously trapped particles. *Phys. Rev. Lett.*, 82:3198–3201, 1999.

[658] H. Dehmelt, R. Mittleman, R. S. Van Dyck, Jr., and P. Schwinberg. Past electron-positron $g-2$ experiments yielded sharpest bound on CPT violation for point particles. *Phys. Rev. Lett.*, 83:4694–4696, 1999.

[659] J. D. Prestage, J. J. Bollinger, W. M. Itano, and D. J. Wineland. Limits for spatial anisotropy by use of nuclear-spin-polarized $^9Be^+$ ions. *Phys. Rev. Lett.*, 54:2387–2390, 1985.

[660] St. Weinberg. Precision tests of quantum mechanics. *Phys. Rev. Lett.*, 62:485–488, 1989.

[661] J. J. Bollinger, D. J. Heinzen, W. M. Itano, S. L. Gilbert, and D. J. Wineland. Test of the linearity of quantum mechanics by rf spectroscopy of the $^9Be^+$ ground state. *Phys. Rev. Lett.*, 63:1031–1034, 1989.

[662] D. J. Wineland, J. J. Bollinger, D. J. Heinzen, W. M. Itano, and M. G. Raizen. Search for anomalous spin-dependent forces using stored-ion spectroscopy. *Phys. Rev. Lett.*, 67:1735–1738, 1991.

[663] L. O. Hocker, A. Javan, D. Ramachandra Rao, L. Frenkel, and T. Sullivan. Absolute frequency measurement and spectroscopy of gas laser transitions in the far infrared. *Appl. Phys. Lett.*, 10:147–149, 1967.

[664] L. M. Matarrese and K. M. Evenson. Improved coupling to infrared whisker diodes by use of antenna theory. *Appl. Phys. Lett.*, 17:8–10, 1970.

[665] O. Acef, L. Hilico, M. Bahoura, F. Nez, and P. De Natale. Comparison between MIM and Schottky diodes as harmonic mixers for visible lasers and microwave sources. *Opt. Commun.*, 109:428–434, 1994.

[666] H. H. Klingenberg and C. O. Weiss. Rectification and harmonic generation with metal-insulator-metal diodes in the mid-infrared. *Appl. Phys. Lett.*, 43:361–363, 1983.

[667] Carl O. Weiss, G. Kramer, B. Lipphardt, and H. Schnatz. Optical frequency measurement by conventional frequency multiplication. In A. N. Luiten, editor, *Frequency Measurement and Control*, volume 79 of *Topics in Applied Physics*, pages 215–247. Springer, Berlin, Heidelberg, New York, 2001.

[668] Martin M. Fejer, G. A. Magel, Dieter H. Jundt, and Robert L. Byer. Quasi-phase-matched second harmonic generation: Tuning and tolerances. *IEEE J. Quantum Electron.*, 28:2631–2653, 1992.

[669] M. Houé and P. D. Townsend. An introduction to methods of periodic poling for second-harmonic generation. *J. Phys. D: Appl. Phys.*, 28:1747–1763, 1995.

[670] J.-P. Meyn and M. M. Fejer. Tunable ultraviolet radiation by second-harmonic generation in periodically poled lithium tantalate. *Opt. Lett.*, 22:1214–1216, 1997.

[671] Ch. Koch and H. R. Telle. Bridging THz-frequency gaps in the near IR by coherent four-wave mixing in GaAlAs laser diodes. *Opt. Commun.*, 91:371–376, 1992.

[672] R. Kallenbach, B. Scheumann, C. Zimmermann, D. Meschede, and T. W. Hänsch. Electro-optic sideband generation at 72 GHz. *Appl. Phys. Lett.*, 54:1622–1624, 1989.

[673] Motonobu Kourogi, Ken'ichi Nakagawa, and Motoichi Ohtsu. Wide-span optical frequency comb generator for accurate optical frequency difference measurement. *IEEE J. Quantum Electron.*, 29:2693–2701, 1993.

[674] M. Kourogi, T. Enami, and M. Ohtsu. A monolithic optical frequency comb generator. *IEEE Phot. Techn. Lett.*, 6:214–217, 1994.

[675] Harald R. Telle and Uwe Sterr. Generation and metrological application of optical frequency combs. In A. N. Luiten, editor, *Frequency Measurement and Control: Advanced Techniques and Future Trends*, pages 295–313. Springer; Berlin, Heidelberg, New York, 2001.

[676] M. Kourogi, B. Widiyatomoko, Y. Takeuchi, and M. Ohtsu. Limit of optical-frequency comb generation due to material dispersion. *IEEE J. Quantum Electron.*, 31:2120–2126, 1995.

[677] L. R. Brothers and N. C. Wong. Dispersion compensation for terahertz optical frequency comb generation. *Opt. Lett.*, 22:1015–1017, 1997.

[678] Motonobu Kourogi, Kazuhiro Imai, Bambang Widiyatmoko, and Motoichi Ohtsu. Generation of expanded optical frequency combs. In A. N. Luiten, editor, *Frequency Measurement and Control: Advanced Techniques and Future Trends*, pages 315–335. Springer; Berlin, Heidelberg, New York, 2001.

[679] Th. Udem, J. Reichert, T. W. Hänsch, and M. Kourogi. Accuracy of optical frequency comb generators and optical frequency interval divider chains. *Opt. Lett.*, 23:1387–1389, 1998.

[680] J. Ye, L.-S. Ma, T. Daly, and J. L. Hall. Highly selective terahertz optical frequency comb generator. *Opt. Lett.*, 22:301–303, 1997.

[681] Z. Bay, G. G. Luther, and J. A. White. Measurement of an optical frequency and the speed of light. *Phys. Rev. Lett.*, 29:189–192, 1972.

[682] D. A. Jennings, C. R. Pollock, F. R. Petersen, R. E. Drullinger, K. M. Evenson, J. S. Wells, J. L. Hall, and H. P. Layer. Direct frequency measurement of the I_2-stabilized He-Ne 473-THz (633-nm) laser. *Opt. Lett.*, 8:136–138, 1983.

[683] O. Acef, J. J. Zondy, M. Abed, D. G. Rovera, A. H. Gérard, A. Clairon, Ph. Laurent, Y. Millerioux, and P. Juncar. A CO_2 to visible optical frequency synthesis chain: accurate measurement of the 473 THz HeNe/I_2 laser. *Opt. Commun.*, 97:29–34, 1993.

[684] J. E. Bernard, A. A. Madej, L. Marmet, B. G. Whitford, K. J. Siemsen, and S. Cundy. Cs-based frequency measurement of a single, trapped ion transition in the visible region of the spectrum. *Phys. Rev. Lett.*, 82:3228–3231, 1999.

[685] G. Kramer, B. Lipphardt, and C. O. Weiss. Coherent frequency synthesis in the infrared. In *Proc. 1992 IEEE Frequency Control Symposium*, pages 39–43, Hershey, Pennsylvania, USA, 1992. IEEE catalog no. 92CH3083-3.

[686] Harald R. Telle, Burghard Lipphardt, and Jörn Stenger. Kerr-lens mode-locked lasers as transfer oscillators for optical frequency measurements. *Appl. Phys. B*, 74:1–6, 2002.

[687] Peter A. Jungner, Steve Swartz, Mark Eickhoff, Jun Ye, J. L. Hall, and S. Waltman. Absolute frequency of the molecular iodine transition R(56)32-0 near 532 nm. *IEEE Trans. Instrum. Meas.*, 44:151–154, 1995.

[688] H. R. Telle, D. Meschede, and T. W. Hänsch. Realization of a new concept for visible frequency division: phase locking of harmonic and sum frequencies. *Opt. Lett.*, 15:532–534, 1990.

[689] K. Nakagawa, M. Kourogi, and M. Ohtsu. Proposal of a frequency-synthesis chain between the microwave and optical frequencies of the Ca intercombination line at 657 nm using diode lasers. *Appl. Phys. B*, 57:425–430, 1993.

[690] Th. Udem, A. Huber, B. Gross, J. Reichert, M. Prevedelli, M. Weitz, and T. W. Hänsch. Phase-coherent measurement of the hydrogen 1S–2S transition frequency with an optical frequency interval divider chain. *Phys. Rev. Lett.*, 79:2646–2649, 1997.

[691] N. C. Wong. Optical-to-microwave frequency chain utilizing a two-laser-based optical parametric oscillator network. *Appl. Phys. B*, 61:143–149, 1995.

[692] T. Ikegami, S. Slyusarev, S. Ohshima, and E. Sakuma. A cw optical parametric oscillator for optical frequency measurement. In J. C. Bergquist, editor, *Proceedings of the Fifth Symposium on Frequency Standards and Metrology*, pages 333–338, Singapore, New Jersey, London, Hong Kong, 1996. World Scientific.

[693] T. Ikegami, S. Slyusarev, S. Ohshima, and E. Sakuma. Accuracy of an optical parametric oscillator as an optical frequency divider. *Opt. Commun.*, 127:69–72, 1996.

[694] D. H. Sutter, G. Steinmeyer, L. Gallmann, N. Matuschek, F. Morier-Genoud, U. Keller, V. Scheuer, G. Angelow, and T. Tschudi. Semiconductor saturable-absorber mirror-assisted Kerr-lens mode-locked Ti:Sapphire laser producing pulses in the two-cycle regime. *Opt. Lett.*, 24:631–633, 1999.

[695] U. Siegner and U. Keller. Nonlinear optical processes for ultrashort pulse generation. In Michael Bass, Jay M. Enoch, Eric W. Van Stryland, and William L. Wolfe, editors, *Handbook of Optics*, pages 25.1 – 25.31. McGraw-Hill, New York, 2001.

[696] R. Fluck, I. D. Jung, G. Zhang, F. X. Kärtner, and U. Keller. Broadband saturable absorber for 10-fs pulse generation. *Opt. Lett.*, 21:743–745, 1996.

[697] I. D. Jung, F. X. Kärtner, N. Matuschek, D. H. Sutter, F. Morier-Genoud, G. Zhang, U. Keller, V. Scheuer, M. Tilsch, and T. Tschudi. Self-starting 6.5-fs pulses from a Ti:sapphire laser. *Opt. Lett.*, 22:1009–1011, 1997.

[698] Ferenc Krausz, Martin E. Fermann, Thomas Brabec, Peter F. Curley, Martin Hofer, Manfred H. Ober, Christian Spielmann, Ernst Winter, and A. J. Schmidt. Femtosecond solid-state lasers. *IEEE J. Quantum Electron.*, 28:2097–2122, 1992.

[699] V. Magni, G. Cerullo, S. De Silvestri, and A. Monguzzi. Astigmatism in Gaussian-beam self-focussing and in resonators for Kerr-lens mode-locking. *J. Opt. Soc. Am. B*, 12:476–485, 1995.

[700] R. L. Fork, O. E. Martinez, and J. P. Gordon. Negative dispersion using pairs of prisms. *Opt. Lett.*, 9:150–152, 1984.

[701] Robert Szipöcs, Kárpát Ferencz, Christian Spielmann, and Ferenc Krausz. Chirped multilayer coatings for broadband dispersion control in femtosecond lasers. *Opt. Lett.*, 19:201–203, 1994.

[702] F. X. Kärtner, N. Matuschek, T. Schibli, U. Keller, H. A. Haus, C. Heine, R. Morf, V. Scheuer, M. Tilsch, and T. Tschudi. Design and fabrication of double-chirped mirrors. *Opt. Lett.*, 22:831–833, 1997.

[703] J. C. Knight, T. A. Birks, P. St. J. Russell, and D. M. Atkin. All-silica single-mode optical fiber with photonic crystal cladding. *Opt. Lett.*, 21:1547–1549, 1996.

[704] A. B. Fedotov, A. M. Zheltikov, L. A. Mel'nikov, A. P. Tarasevitch, and D. von der Linde. Spectral broadening of femtosecond laser pulses in fibers with a photonic-crystal cladding. *JETP Lett.*, 71:281–284, 2000.

[705] Jinendra K. Ranka, Robert S. Windeler, and Andrew J. Stentz. Visible continuum generation in air-silica microstructure optical fibers with anomalous dispersion at 800 nm. *Opt. Lett.*, 25:25–27, 2000.

[706] A. Husakou, V. P. Kalosha, and J. Hermann. Nonlinear phenomena with ultrabroadband optical radiation in photonic crystal fibers and hollow waveguides. In K. Porsezian and V. C. Kuriakose, editors, *Optical Solitons. Theoretical and Experimental Challenges*, Lecture Notes in Physics, pages 299–325. Springer, 2003.

[707] K. L. Corwin, N. R. Newbury, J. M. Dudley, S. Coen, S. A. Diddams, K. Weber, and R. S. Windeler. Fundamental noise limitations to supercontinuum generation in microstructure fiber. *Phys. Rev. Lett.*, 90:113904–1–4, 2003.

[708] David J. Jones, Scott A. Diddams, Matthew S. Taubman, Steven T. Cundiff, Long-Sheng Ma, and John L. Hall. Frequency comb generation using femtosecond pulses and cross-phase modulation in optical fiber at arbitrary center frequencies. *Opt. Lett.*, 25:308–310, 2000.

[709] H. R. Telle, G. Steinmeyer, A. E. Dunlop, J. Stenger, D. H. Sutter, and U. Keller. Carrier-envelope offset phase control: A novel concept for absolute optical frequency measurement and ultrashort pulse generation. *Appl. Phys. B*, 69:327–332, 1999.

[710] J. Reichert, R. Holzwarth, Th. Udem, and T. W. Hänsch. Measuring the frequency of light with mode-locked lasers. *Opt. Commun.*, 172:59–68, 1999.

[711] Jun Ye, John L. Hall, and Scott A. Diddams. Precision phase control of an ultrawide-bandwidth femtosecond laser: a network of ultrastable frequency marks across the visible spectrum. *Opt. Lett.*, 25:1675–1677, 2000.

[712] Jörn Stenger and Harald R. Telle. Intensity-induced mode shift in femtosecond lasers via the nonlinear index of refraction. *Opt. Lett.*, 26:1553–1555, 2000.

[713] Jun Ye, Tai Hyun Yoon, John L. Hall, Alan A. Madej, John E. Bernard, Klaus J. Siemsen, Louis Marmet, Jean-Marie Chartier, and Annick Chartier. Accuracy comparison of absolute optical frequency measurement between harmonic-generation synthesis and a frequency-division femtosecond comb. *Phys. Rev. Lett.*, 85:3797–3800, 2000.

[714] Scott A. Diddams, L. Hollberg, Long-Sheng Ma, and Lennart Robertsson. Femtosecond-laser-based optical clockwork with instability $\leq 6.3 \times 10^{-16}$ in 1 s. *Opt. Lett.*, 27:58–60, 2002.

[715] Jörn Stenger, Harald Schnatz, Christian Tamm, and Harald R. Telle. Ultra-precise measurement of optical frequency ratios. *Phys. Rev. Lett.*, 88:073601–1–4, 2002.

[716] Hidemi Tsuchida. Timing-jitter reduction of a mode-locked Cr:LiSAF laser by simultaneous control of cavity length and pump power. *Opt. Lett.*, 25:1475–1477, 2000.

[717] Florian Tauser, Alfred Leitenstorfer, and Wolfgang Zinth. Amplified femtosecond pulses from an Er:fiber system: Nonlinear pulse shortening and self-referencing detection of the carrier-envelope phase evolution. *Optics Express*, 11:594–600, 2003.

[718] Th. Udem, J. Reichert, R. Holzwarth, and T. W. Hänsch. Accurate measurement of large optical frequency differences with a mode-locked laser. *Opt. Lett.*, 24:881–883, 1999.

[719] Scott A. Diddams, David J. Jones, Long-Sheng Ma, Steven T. Cundiff, and John L. Hall. Optical frequency measurement across a 104-THz gap with a femtosecond laser frequency comb. *Opt. Lett.*, 25:186–188, 2000.

[720] S. A. Diddams, D. J. Jones, J. Ye, S. T. Cundiff, and J. L. Hall. Direct link between microwave and optical frequencies with a 300 THz femtosecond laser comb. *Phys. Rev. Lett.*, 84:5102–5105, 2000.

[721] J. Reichert, M. Niering, R. Holzwarth, M. Weitz, Th. Udem, and T. W. Hänsch. Phase coherent vacuum-ultraviolet to radio frequency comparison with a mode-locked laser. *Phys. Rev. Lett.*, 84:3232–3235, 2000.

[722] R. A. Nelson, D. D. McCarthy, S. Malys, J. Levine, B. Guinot, H. F. Fliegel, R. L. Beard, and T. R. Bartholomew. The leap second: its history and possible future. *Metrologia*, 38:509–529, 2001.

[723] Bureau International de Poids et Mesures. Circular T can be found in http://www.bipm.fr/en/scientific/tai/.

[724] Peter Wolf. Relativity and the metrology of time. Monographie 97/1, Bureau International des Poids et Mesures, Pavillon de Breteuil, F-92312 Sevres Cedex, 1997.

[725] Robert A. Nelson. *Relativistic Effects in Satellite Time and Frequency Transfer and Dissemination*. International Telecommunication Union, Geneva, 2004. to be published.

[726] International Telecommunication Union, ITU, Place des Nations, CH-1211 Geneva 20, Switzerland. *ITU-R Recommendations: Time Signals and Frequency Standards Emissions*, 1997.

[727] D. Kirchner. Two-way satellite time and frequency transfer (TWSTFT): Principle, implementation, and current performance. In W. Ross Stone, editor, *Review of Radio Science 1996–1999*, pages 27–44, Oxford, New York, 1999. Oxford University Press.

[728] G. Petit and P. Wolf. Relativistic theory for picosecond time transfer in the vicinity of the earth. *Astron. Astrophys.*, 286:971–977, 1994.

[729] G. Becker and P. Hetzel. Kodierte Zeitinformation über den Zeitmarken- und Normalfrequenzsender DCF77. *PTB-Mitteilungen*, 83:163–164, 1973.

[730] International Telecommunication Union, ITU, Place des Nations, CH-1211 Geneva 20, Switzerland. *ITU-R Recommendation TF-768-3: Standard Frequencies and Time Signals*, 1997.

[731] Elliot D. Kaplan, editor. *Understanding GPS: Principles and Applications*. Artech House, Boston, London, 1996.

[732] Neil Ashby and Marc Weiss. Global positioning system receivers and relativity. Technical Report NIST Technical Note 1385, National Institute of Standards and Technology, USA, 1999.

[733] Michael L. Lombardi, Lisa M. Nelson, Andrew N. Novick, and Victor S. Zhang. Time and frequency measurements using the global positioning system (GPS). *Cal. Lab. Int. J. Metrology*, pages 26–33, July–September 2001.

[734] G. Dudle, F. Overney, Th. Schildknecht, T. Springer, and L. Prost. Transatlantic time and frequency transfer by GPS carrier phase. In *Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum and The IEEE International Frequency Control Symposium*, pages 243–246, 1999.

[735] P. Fridelance and C. Veillet. Operation and data analysis in the LASSO experiment. *Metrologia*, 32:27–33, 1995.

[736] E. Samain and P. Fridelance. Time transfer by laser link (T2L2) experiment on Mir. *Metrologia*, 35:151–159, 1998.

[737] B. de Beauvoir, F. Nez, I. Hilico, L. Julien, F. Biraben, B. Cagnac, J. J. Zondy, D. Touahri, O. Acef, and A. Clairon. Transmission of an optical frequency through a 3 km long optical fiber. *Eur. Phys. J. D*, 1:227–229, 1998.

[738] Long-Sheng Ma, Peter Jungner, Jun Ye, and J. L. Hall. Accurate cancellation (to milli-Hertz levels) of optical phase noise due to vibration or insertion phase in fiber transmitted light. In Yaakov Shevy, editor, *Proceedings of SPIE: Laser Frequency Stabilization and Noise Reduction*, volume 2378, pages 165–175, P.O. Box 10, Bellingham, Washington 98227–0010 USA, 1995. SPIE.

[739] Jun Ye, Jin-Long Peng, R. Jason Jones, Kevin W. Holman, John L. Hall, David J. Jones, Scott A. Diddams, John Kitching, Sebastien Bize, James C. Bergquist, Leo W. Hollberg, Lennart Robertsson, and Long-Sheng Ma. Delivery of high-stability optical and microwave frequency standards over an optical fiber network. *J. Opt. Soc. Am. B*, 20:1459–1467, 2003.

[740] C. Hazard, M. B. Mackey, and A. J. Shimmins. Investigation of the radio source 3C 273 by the method of lunar occultations. *Nature*, 4872:1037–1039, 1963.

[741] M. Schmidt. 3C 273: A star-like object with large red-shift. *Nature*, 197:1040, 1963.

[742] A. Hewish, S. J. Bell, J. D. H. Pilkington, P. F. Scott, and R. A. Collins. Observation of a rapidly pulsating radio source. *Nature*, 217:709–713, 1968.

[743] D. C. Backer, Shrinivas R. Kulkarni, Carl Heiles, M. M. Davis, and W. M. Goss. A millisecond pulsar. *Nature*, 300:615–618, 1982.

[744] C. Kouveliotou, S. Dieters, T. Strohmayer, J. van Paradijs, G. J. Fishman, C. A. Meegan, K. Hurley, J. Kommers, I. Smith, D. Frail, and T. Murakami. An x-ray pulsar with a superstrong magnetic field in the soft γ-ray repeater SGR1806–20. *Nature*, 393:235–237, 1998.

[745] D. R. Lorimer. Binary and millisecond pulsars. http://www.livingreviews.org/ Articles/Volume1/ 1998-10lorimer, 1998.

[746] F. Camilo and D. J. Nice. Timing parameters of 29 pulsars. *Astrophys. J.*, 445:756–761, 1995.

[747] V. M. Kaspi. High-precision timing of millisecond pulsars and precision astrometry. In E. Høg and P. K. Seidelmann, editors, *Proceedings of the IAU Symposium 166: Astronomical and Astrophysical Objectives of Sub-Milliarcsecond Optical Astrometry*, page 163, The Hague, Netherlands, 1995. Kluwer.

[748] Michael Kramer. Determination of the geometry of the PSR B1913+16 system by geodetic precession. *Astrophys. J.*, 509:856–860, 1998.

[749] J. H. Taylor and J. M. Weisberg. Further experimental tests of relativistic gravity using the binary pulsar PSR 1913+16. *Astrophys. J.*, 345:434–450, 1989.

[750] Arnold Rots. JPL DE200 and DE405 in FITS, Barycenter Code. ftp://heasarc.gsfc.nasa.gov/xte/calibdata/clock/bary, 2001.

[751] J. F. Bell. Radio pulsar timing. *Adv. Space Res.*, 21:137–147, 1998.

[752] J. Bell. Tests of relativistic gravity using millisecond pulsars. In *Pulsar Timing, General Relativity, and the Internal Structure of Neutron Stars*, pages 31–38, 1996. Amsterdam.

[753] R. A. Hulse and J. H. Taylor. Discovery of a pulsar in a binary system. *Astrophys. J.*, 195:L51–L53, 1975.

[754] G´erard Petit. Limits to the stability of pulsar time. In *Proceedings of the 27th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, November 29 - December 1, 1995, San Diego, California*, volume 3334 of *NASA Conference Publication*, pages 387–396, Goddard Space Flight Center, Greenbelt, Maryland 20771, 1995.

[755] Demetrios N. Matsakis, J. H. Taylor, and T. Marshall Eubanks. A statistic for describing pulsar and clock stabilities. *Astron. Astrophys.*, 326:924–928, 1997.

[756] Francois Vernotte. Estimation of the power spectral density of phase: Comparison of three methods. In *Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum and The IEEE International Frequency Control Symposium*, pages 1109–1112, 1999.

[757] Jürgen Helmcke and Fritz Riehle. Physics behind the definition of the meter. In *Recent Advances in Metrology and Fundamental Constants*, volume Course CXLVI of *Proceedings Internat. School of Physics "Enrico Fermi"*, pages 453–493, Amsterdam, Oxford, Tokyo, Washington DC, 2001. IOS Press Ohmsha.

[758] L'Ecole Polytechnique, du Bureau des Longitudes, editor. *Comptes Rendus des séances de la 1er CGPM 1889*, Quai des Grands-Augustins, 55, France, 1890. Gauthier-Villars et Fils.

[759] Bureau International des Poids et Mesures, editor. *Comptes Rendus des séances de la 11^{e} CGPM*, Quai des Grands-Augustins, 55, France, 1960. Gauthier-Villars & C^{ie}.

[760] Bureau International des Poids et Mesures, editor. *Comptes Rendus des séances de la 17^{e} CGPM*, Pavillon de Breteuil, F-92310 Sévres, France, 1983. BIPM.

[761] Editor's note. Documents concerning the new definition of the metre. *Metrologia*, 19:163–177, 1984.

[762] T. J. Quinn. Mise en pratique of the definition of the Metre (1992). *Metrologia*, 30:523–541, 1993/94.

[763] Brian William Petley. *The fundamental physical constants and the frontier of measurement*. Adam Hilger, Bristol, 1985.

[764] B. E. Schaefer. Severe limits on variations of the speed of light with frequency. *Phys. Rev. Lett.*, 82:4964–4966, 1999.

[765] DFG. M. Schneider, editor. *Satellitengeodäsie: Ergebnisse aus dem gleichnamigen Sonderforschungsbereich der TU München*. VCH-Verlag Weinheim, 1990.

[766] Kenneth Nordtvedt. Lunar laser ranging – a comprehensive probe of the post-Newtonian long range interaction. In F. W. Hehl C. Lämmerzahl, C. W. F. Everitt, editor, *Gyros, Clocks, Interferometers...: Testing Relativistic Gravity in Space*, pages 317–329, Springer, Berlin, Heidelberg, New York, 2001.

[767] J. G. Williams, X. X. Newhall, and J. O. Dickey. Relativity parameters determined from lunar laser ranging. *Phys. Rev. D*, 53:6730–6739, 1996.

[768] Jürgen Müller and Kenneth Nordtvedt. Lunar laser ranging and the equivalence principle signal. *Phys. Rev. D*, 58:062001–1–13, 1998.

[769] D. H. Atkinson, J. B. Pollack, and A. Seiff. Measurement of a zonal wind profile on Titan by Doppler tracking of the Cassini entry probe. *Radio Science*, 25:865–881, 1990.

[770] David H. Atkinson, James B. Pollack, and Alvin Seiff. Galileo Doppler measurements of the deep zonal winds at Jupiter. *Science*, 272:842–843, 1996.

[771] Sami Asmar. Trends in performance and characteristics of ultra-stable oscillators for deep space radio science experiments. In Lute Maleki, editor, *Proceedings of the Workshop on the Scientific Applications of Clocks in Space, November 7-8, 1996*, volume JPL Publication 97-15, pages 195–199, 1997.

[772] B. Edlén. The refractive index of air. *Metrologia*, 2:71–80, 1966.

[773] G. Bönsch and E. Potulski. Measurement of the refractive index of air and comparison with modified Edlén's formulae. *Metrologia*, 35:133–139, 1998.

[774] G. Bönsch, A. Nicolaus, and U. Brand. Wavelength measurement of a 544 nm FM-I_2-stabilised He-Ne laser. *Optik*, 107:127–131, 1998.

[775] Richard L. Steiner, David B. Newell, and Edwin R. Williams. A result from the NIST watt balance and an analysis of uncertainties. *IEEE Trans. Instrum. Meas.*, 48:205–208, 1999.

[776] T. M. Niebauer, G. S. Sasagawa, J. E. Faller, R. Hilt, and F. Klopping. A new generation of absolute gravimeters. *Metrologia*, 32:159–180, 1995.

[777] Peter G. Nelson. An active vibration isolation system for inertial reference and precision measurement. *Rev. Sci. Instrum.*, 62:2069–2075, 1991.

[778] L. Robertsson, O. Francis, T. M. vanDam, J. Faller, D. Ruess, J.-M. Delinte, L. Vitushkin, J. Liard, C. Gagnon, Guo You Guang, Huang Da Lun, Fang Yong Yuan, Xu Jin Yi, G. Jeffries, H. Hopewell, R. Edge, I. Robinson, B. Kibble, J. Makinen, J. Hinderer, M. Amalvict, B. Luck, H. Wilmes, F. Rehren, K. Schmidt, M. Schnull, G. Cerutti, A. Germak, Z. Zabek, A. Pachuta, G. Arnautov, E. Kalish, Y. Stus, D. Stizza, J. Friederich, J.-M. Chartier, and I. Marson. Results from the fifth international comparison of absolute gravimeters, ICAG'97. *Metrologia*, 38:71–78, 2001.

[779] Achim Peters. *High Precision Gravity Measurements using Atom Interferometry*. PhD thesis, Stanford University, Stanford, CA, 1998.

[780] A. Peters, K. Y. Chung, and S. Chu. Measurement of gravitational acceleration by dropping atoms. *Nature*, 400:849–852, 1999.

[781] A. Peters, K. Y. Chung, and S. Chu. High-precision gravity measurements using atom interferometry. *Metrologia*, 38:25–61, 2001.

[782] M. J. Snadden, J. M. McGuirk, P. Bouyer, K. G. Haritos, and M. A. Kasevich. Measurement of the Earth's gravity gradient with an atom interferometer-based gravity gradiometer. *Phys. Rev. Lett.*, 81:971–974, 1998.

[783] Harald Simonsen, Jes Henningsen, and Susanne Søgaard. DFB fiber lasers as optical wavelength standards in the 1.5 μm region. *IEEE Trans. Instrum. Meas.*, 50:482–485, 2001.

[784] T. Dennis, E. A. Curtis, C. W. Oates, L. Hollberg, and S. L. Gilbert. Wavelength references for 1300-nm wavelength-division multiplexing. *J. Lightw. Technol.*, 20:804–810, 2002.

[785] B. D. Josephson. Possible new effects in superconductive tunneling. *Phys. Lett.*, 1:251–253, 1962.

[786] R. P. Feynman, editor. *The Feynman Lectures on Physics*. Adison Wesley, Reading MA, 1965.

[787] Jürgen Niemeyer. Counting of single flux and single charge quanta for metrology. In J. Hamelin, editor, *Modern Radio Science 1996*, pages 85–109. Oxford University Press, Oxford, 1996.

[788] J. Niemeyer. Das Josephsonspannungsnormal – Entwicklung zum Quantenvoltmeter. *PTB-Mitt.*, 110:169–177, 2000.

[789] David Deaver, William B. Miller, Leonardo Pardo, Klaus Jaeger, Dennis Plowman, and Clark A. Hamilton. Interlaboratory comparison of Josephson voltage standards. *IEEE Trans. Instrum. Meas.*, 50:199–202, 2001.

[790] Peter J. Mohr and Barry N. Taylor. Codata recommended values of the fundamental physical constants: 1998. *Rev. Mod. Phys.*, 72:351–495, 2000.

[791] P. J. Mohr and B. N. Taylor. The 2002 CODATA recommended values of the fundamental physical constants. Web Version 4.0, available at physics.nist.gov/constants (National Institute of Standards and Technology, Gaithersburg, MD 20899, 9 December 2003), 2004. to be published in 2004.

[792] F. Riehle, S. Bernstorff, R. Fröhling, and F. P. Wolf. Determination of electron currents below 1 nA in the storage ring BESSY by measurement of the synchrotron radiation of single electrons. *Nucl. Instr. Meth. Phys. Res.*, A268:262–269, 1988.

[793] Andreas Peters, Wolfgang Vodel, Helmar Koch, Ralf Neubert, Hannes Reeg, and Claus Herrmann Schroeder. A crogenic current comparator for the absolute measurement of nA beams. In Robert O. Hettel, Stephen R. Smith, and Jennifer D. Masek, editors, *AIP Conference Proceedings of the Beam Instrumentation Workshop, May 1998, Stanford, CA, USA*, volume 451, pages 163–180, 1998.

[794] J. Sesé, G. Rietveld, A. Camôn, C. Rillo, L. Vargas, G. Christian S. Brons, M. G. H. Hiddink, J. Flokstra, H. Rogalla, W. Jaszczuk, and H. Altenburg. Design and realization of an optimal current sensitive CCC. *IEEE Trans. Instrum. Meas.*, 48:370–374, 1999.

[795] Mark W. Keller, Ali L. Eichenberger, John M. Martinis, and Neil M. Zimmerman. A capacitance standard based on counting electrons. *Science*, 285:1706–1709, 1999.

[796] Mark W. Keller. Standards of current and capacitance based on single-electron tunneling devices. In *Recent Advances in Metrology and Fundamental Constants*, volume Course CXLVI of *Proceedings of the Internat. School of Physics "Enrico Fermi"*, pages 291–316, Amsterdam, Oxford, Tokyo, Washington DC, 2001. IOS Press Ohmsha.

[797] J. M. Shilton, V. I. Talyanskii, M. Pepper, D. A. Ritchie, J. E. F. Frost, C. J. B. Ford, C. G. Smith, and G. A. C. Jones. High-frequency single-electron transport in a quasi-one-dimensional GaAs channel induced by surface acoustic waves. *J. Phys.: Condens. Matter*, 8:L531–L539, 1996.

[798] J. Cunningham, V. I. Talyanskii, J. M. Shilton, M. Pepper, M. Y. Simmons, and D. A. Ritchie. Single-electron acoustic charge transport by two counterpropagating surface acoustic wave beams. *Phys. Rev. B*, 60:4850–4855, 1999.

[799] J. Cunningham, V. I. Talyanskii, J. M. Shilton, M. Pepper, A. Kristensen, and P. E. Lindelof. Quantized acoustoelectric current – an alternative route towards a standard of electric current. *J. Low Temp. Phys.*, 118:555–569, 2000.

[800] H. Weinstock, editor. *SQUID Sensors: Fundamentals, Fabrication and Applications*. Kluwer Academic Publishers, Dordrecht, Boston, London, 1996.

[801] Volkmar Kose and Friedmund Melchert. *Quantenmaße in der elektrischen Meßtechnik*. VCH, Weinheim, New York, Basel, Cambridge, 1991.

[802] J. Dupont-Roc, S. Haroche, and C. Cohen-Tannoudji. Detection of very weak magnetic fields (10^{-9} Gauss) by ^{87}Rb zero-field level crossing resonances. *Phys. Lett.*, 28A:638–639, 1969.

[803] D. Budker, D. F. Kimball, S. M. Rochester, V. V. Yashchuk, and M. Zolotorev. Sensitive magnetometry based on nonlinear magneto-optical rotation. *Phys. Rev. A*, 62:043403–1–7, 2000.

[804] R. Wynands and A. Nagel. Precision spectroscopy with coherent dark states. *Appl. Phys. B*, 68:1–25, 1999.

[805] F. Bloch. Nuclear induction. *Phys. Rev.*, 70:460–474, 1946.

[806] F. Bloch, W. W. Hansen, and M. Packard. The nuclear induction experiment. *Phys. Rev.*, 70:474–485, 1946.

[807] R. Prigl, U. Haeberlen, K. Jungmann, G. zu Putlitz, and P. von Walter. A high precision magnetometer based on pulsed NMR. *Nucl. Instr. and Meth.*, A 374:118–126, 1996.

[808] J. W. G. Wignall. Proposal for an absolute, atomic definition of mass. *Phys. Rev. Lett.*, 68:5–8, 1992.

[809] Paul De Bièvre, Staf Valkiers, Rüdiger Kessel, Philip D. P. Taylor, Peter Becker, H. Bettin, Anna Peuto, Savino Pettorruso, K. Fujii, A. Waseda, M. Tanaka, R. D. Deslattes, H. S. Peiser, and M. J. Kenny. A reassessment of the molar volume of silicon and of the Avogadro constant. *IEEE Trans. Instrum. Meas.*, 50:593–597, 2001.

[810] P. Becker. The molar volume of single-crystal silicon. *Metrologia*, 38:85–86, 2001.

[811] B. N. Taylor and P. J. Mohr. On the redefinition of the kilogram. *Metrologia*, 36:63–64, 1999.

[812] E. R. Williams, R. I. Steiner, D. B. Newell, and P. T. Olson. Accurate measurement of the Planck constant. *Phys. Rev. Lett.*, 81:2404–2407, 1998.

[813] T. Andreae, W. König, R. Wynands, D. Leibfried, F. Schmidt-Kaler, C. Zimmermann, D. Meschede, and T. W. Hänsch. Absolute frequency measurement of the hydrogen 1S – 2S transition and a new value of the Rydberg constant. *Phys. Rev. Lett.*, 69:1923–1926, 1992.

[814] C. Schwob, L. Jozefowski, B. de Beauvoir, L. Hilico, F. Nez, L. Julien, F. Biraben, O. Acef, and A. Clairon. Optical frequency measurement of the $2S - 12D$ transitions in hydrogen and deuterium: Rydberg constant and Lamb Shift determinations. *Phys. Rev. Lett.*, 82:4960–4963, 1999.

[815] K. Pachucki, D. Leibfried, M. Weitz, A. Huber, W. König, and T. W. Hänsch. Theory of the energy levels and precise two-photon spectroscopy of atomic hydrogen and deuterium. *J. Phys. B: At. Mol. Opt. Phys.*, 29:177–195, 1996.

[816] Eckhard Krüger, Wolfgang Nistler, and Winfried Weirauch. Determination of the fine-structure constant by measuring the quotient of the Planck constant and the neutron mass. *IEEE Trans. Instrum. Meas.*, 46:101–103, 1997.

[817] Andreas Wicht, Joel M. Hensley, Edina Sarajlic, and Steven Chu. A preliminary measurement of h/M_{Cs} with atom interferometry. In P. Gill, editor, *Frequency Standards and Metrology, Proceedings of the Sixth Symposium*, pages 193–212, Singapore, 2002. World Scientific.

[818] D. S. Weiss, B. C. Young, and S. Chu. Precision measurement of the photon recoil of an atom using atomic interferometry. *Phys. Rev. Lett.*, 70:2706–2709, 1993.

[819] D. S. Weiss, B. C. Young, and S. Chu. Precision measurement of $\hbar/m_{Cs}$ based on photon recoil using laser-cooled atoms and atomic interferometry. *Appl. Phys. B*, 59:217–256, 1994.

[820] P. A. M. Dirac. The cosmological constants. *Nature*, 264:323, 1937.

[821] P. Sisterna and H. Vucetich. Time variation of fundamental constants: Bounds from geophysical and astronomical data. *Phys. Rev. D*, 41:1034–1046, 1990.

[822] Oskar Klein. Quantentheorie und fünfdimensionale Relativitätstheorie. *Z. Phys.*, 37:895–906, 1926.

[823] William J. Marciano. Time variation of the fundamental "constants" and Kaluza–Klein theories. *Phys. Rev. Lett.*, 52:489–491, 1984.

[824] T. Damour. Equivalence principle and clocks. In J. D. Barrow, editor, *Proceedings of the 34$^{\rm th}$ Rencontres de Moriond, "Gravitational waves and Experimental Gravity"*, pages 1–6, gr-qc/9711084, 1999.

[825] J. K. Webb, M. T. Murphy, V. V. Flambaum, V. A. Dzuba, J. D. Barrow, C. W. Churchill, J. X. Prochaska, and A. M. Wolfe. Further evidence for cosmological evolution of the fine structure constant. *Phys. Rev. Lett.*, 87:091301–1–4, 2001.

[826] A. I. Shlyakther. Direct test of the constancy of fundamental nuclear constants. *Nature*, 264:340, 1976.

[827] Thibault Damour and Freeman Dyson. The Oklo bound on the time variation of the fine-structure constant revisited. *Nucl. Phys. B*, 480:37–54, 1996.

[828] J. K. Webb, V. V. Flambaum, Ch. W. Churchill, M. J. Drinkwater, and J. D. Barrow. Search for time variation of the fine structure constant. *Phys. Rev. Lett.*, 82:884–887, 1999.

[829] V. A. Dzuba, V. V. Flambaum, and J. K. Webb. Space-time variation of physical constants and relativistic corrections in atoms. *Phys. Rev. Lett.*, 82:888–891, 1999.

[830] V. A. Dzuba, V. V. Flambaum, and J. K. Webb. Calculations of the relativistic effects in many-electron atoms and space-time variation of fundamental constants. *Phys. Rev. A*, 59:230–237, 1999.

[831] Savely G. Karshenboim. Some possibilities for laboratory searches for variations of fundamental constants. *Canad. J. Phys.*, 78:639–678, 2000.

[832] A. Godone, C. Novero, P. Tavella, and K. Rahimullah. New experimental limits to the time variations of $g_p(m_e/m_p)$ and α. *Phys. Rev. Lett.*, 71:2364–2366, 1993.

[833] John D. Prestage, Robert L. Tjoelker, and Lute Maleki. Atomic clocks and variations of the fine structure constant. *Phys. Rev. Lett.*, 74:3511–3514, 1995.

[834] Steen Hannestad. Possible constraints on the time variation of the fine structure constant from cosmic microwave background data. *Phys. Rev. D*, 60:023515–1–5, 1999.

[835] H. Marion, F. Pereira Dos Santos, M. Abgrall, S. Zhang, Y. Sortais, S. Bize, I. Maksimovic, D. Calonico, J. Grünert, C. Mandache, P. Lemonde, G. Santarelli, Ph. Laurent, A. Clairon, and C. Salomon. Search for variations of fundamental constants using atomic fountain clocks. *Phys. Rev. Lett.*, 90:150801–1–4, 2003.

[836] E. Peik, B. Lipphardt, H. Schnatz, T. Schneider, Chr. Tamm, and S. G. Karshenboim. New limit on the present temporal variation of the fine structure constant. arXiv:physics/0402132, 2004. to be published in 2004.

[837] M. Fischer, N. Kolachevsky, M. Zimmermann, R. Holzwarth, Th. Udem, T.W. Haensch, M. Abgrall, J. Gruenert, I. Maksimovic, S. Bize, H. Marion, F.Pereira Dos Santos, P. Lemonde, G. Santarelli, P. Laurent, A. Clairon, C. Salomon, M. Haas, U.D. Jentschura, and C.H. Keitel. New limits to the drift of fundamental constants from laboratory measurements. arXiv:physics/0312086, 2004. to be published in 2004.

[838] V. A. Dzuba and V. V. Flambaum. Atomic optical clocks and search for variation of the fine-structure constant. *Phys. Rev. A*, 61:034502–1–3, 2000.

[839] Leonard I. Schiff. *Quantum Mechanics*. Mc Graw-Hill, New York, 1968.

[840] Roy J. Glauber. Coherent and incoherent states of the radiation field. *Phys. Rev.*, 131:2766–2788, 1963.

[841] Hans-A. Bachor. *A Guide to Experiments in Quantum Optics*. Wiley-VCH, Weinheim–New York, 1998.

[842] R. Paschotta, M. Collett, P. Kürz, K. Fiedler, H. A. Bachor, and J. Mlynek. Bright squeezed light from a singly resonant frequency doubler. *Phys. Rev. Lett.*, 72:3807–3810, 1994.

[843] E. S. Polzik, J. Carri, and H. J. Kimble. Spectroscopy with squeezed light. *Phys. Rev. Lett.*, 68:3020–3023, 1992.

[844] J. Kitching, A. Yariv, and Y. Shevy. Room temperature generation of amplitude squeezed light from a semiconductor laser with weak optical feedback. *Phys. Rev. Lett.*, 74:3372–3375, 1995.

[845] Carlton M. Caves. Quantum-mechanical noise in an interferometer. *Phys. Rev. D*, 23:1693–1708, 1981.

[846] Min Xiao, Ling-An Wu, and H. J. Kimble. Precision measurement beyond the shot-noise limit. *Phys. Rev. Lett.*, 59:278–281, 1987.

[847] Andrew J. Stevenson, Malcolm B. Gray, Hans-A. Bachor, and David E. McClelland. Quantum-noise-limited interferometric phase measurements. *Appl. Opt.*, 32:3481–3493, 1993.

[848] Bernard Yurke, Samuel L. McCall, and John R. Klauder. SU(2) and SU(1,1) interferometers. *Phys. Rev. A*, 33:4033–4054, 1986.

[849] B. Yurke. Input states for enhancement of fermion interferometer sensitivity. *Phys. Rev. Lett.*, 56:1515–1517, 1986.

[850] B. C. Sanders and G. J. Milburn. Optimal quantum measurements for phase estimation. *Phys. Rev. Lett.*, 75:2944–2947, 1995.

[851] M. J. Holland and K. Burnett. Interferometric detection of optical phase shifts at the Heisenberg limit. *Phys. Rev. Lett.*, 71:1355–1358, 1993.

[852] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be complete? *Phys. Rev.*, 48:777–780, 1935.

[853] Alain Aspect, Philippe Grangier, and Gérard Roger. Experimental realization of Einstein–Podolsky–Rosen–Bohm Gedankenexperiment: A new violation of Bell's inequalities. *Phys. Rev. Lett.*, 49:91–94, 1982.

[854] J. S. Bell. On the Einstein Podolsky Rosen paradox. *Physics*, 1:195–200, 1964.

[855] Daniel M. Greenberger, Michael A. Horne, Abner Shimony, and Anton Zeilinger. Bell's theorem without inequalities. *Am. J. Phys.*, 58:1131–1143, 1990.

[856] Jian-Wei Pan, Dik Bouwmeester, Matthew Daniell, Harald Weinfurter, and Anton Zeilinger. Experimental test of quantum nonlocality in three-photon Greenberger-Horne-Zeilinger entanglement. *Nature*, 403:515–519, 2000.

[857] N. David Mermin. Extreme quantum entanglement in a superposition of macroscopically distinct states. *Phys. Rev. Lett.*, 65:1838–1840, 1990.

[858] Masahiro Kitagawa and Masahito Ueda. Squeezed spin states. *Phys. Rev. A*, 47:5138–5143, 1993.

[859] J. J. Cirac and P. Zoller. Quantum computations with cold trapped ions. *Phys. Rev. Lett.*, 74:4091–4094, 1995.

[860] Q. A. Turchette, C. S. Wood, B. E. King, C. J. Myatt, D. Leibfried, W. M. Itano, C.Monroe, and D. J. Wineland. Deterministic entanglement of two ions. *Phys. Rev. Lett.*, 81:1525–1528, 1998.

[861] Klaus Mølmer and Anders Sørensen. Multiparticle entanglement of hot trapped ions. *Phys. Rev. Lett.*, 82:1835–1838, 1999.

[862] C. A. Sackett, D. Kielpinski, B. E. King, C. Langer, V. Meyer, C. J. Myatt, M. Rowe, Q. A. Turchette, W. M. Itano, D. J. Wineland, and C. Monroe. Experimental entanglement of four particles. *Nature*, 404:256–259, 2000.

[863] C. Monroe, C. A. Sackett, D. Kielpinski, B. E. King, C. Langer, V. Meyer, C. J. Myatt, M. Rowe, Q. A. Turchette, W. M. Itano, and D. J. Wineland. Scalable entanglement of trapped ions. In E. Arimondo, P. deNatale, and M. Inguscio, editors, *AIP Conference Proceedings*, volume 551, pages 173–186, American Institute of Physics, Melville, New York, 2001.

[864] D. J. Wineland, J. J. Bollinger, W. M. Itano, F. L. Moore, and D. J. Heinzen. Spin squeezing and reduced quantum noise in spectroscopy. *Phys. Rev. A*, 46:R6797–R6800, 1992.

[865] D. J. Wineland, J. J. Bollinger, W. M. Itano, and D. J. Heinzen. Squeezed atomic states and projection noise in spectroscopy. *Phys. Rev. A*, 50:67–88, 1994.

[866] J. J. Bollinger, Wayne M. Itano, D. J. Wineland, and D. J. Heinzen. Optimal frequency measurements with maximally correlated states. *Phys. Rev. A*, 54:R4649–R4652, 1996.

[867] S. F. Huelga, C. Macchiavello, T. Pellizzari, A. K. Ekert, M. B. Plenio, and J. I. Cirac. Improvement of frequency standards with quantum entanglement. *Phys. Rev. Lett.*, 79:3865–3868, 1997.

[868] W. Dür. Multipartite entanglement that is robust against disposal of particles. *Phys. Rev. A*, 63:020303–1–4, 2001.

[869] V. Meyer, M. A. Rowe, D. Kielpinski, C. A. Sackett, W. M. Itano, C. Monroe, and D. J. Wineland. Experimental demonstration of entanglement-enhanced rotation angle estimation using trapped ions. *Phys. Rev. Lett.*, 86:5870–5873, 2001.

[870] Masami Yasuda and Fujio Shimizu. Observation of two-atom correlation of an ultracold neon atomic beam. *Phys. Rev. Lett.*, 77:3090–3093, 1996.

[871] A. Hemmerich. Quantum entanglement in dilute optical lattices. *Phys. Rev. A*, 60:943–946, 1999.

[872] A. Kuzmich, Klaus Mølmer, and E. S. Polzik. Spin squeezing in an ensemble of atoms illuminated with squeezed light. *Phys. Rev. Lett.*, 79:4782–4785, 1997.

[873] A. Kuzmich, L. Mandel, and N. P. Bigelow. Generation of spin squeezing via continuous quantum nondemolition measurement. *Phys. Rev. Lett.*, 85:1594–1597, 2000.

[874] C. Orzel, A. K. Tuchman, M. L. Fenselau, M. Yasuda, and M. A. Kasevich. Squeezed states in a Bose–Einstein condensate. *Science*, 291:2386–2389, 2001.

[875] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. Quantum-enhanced positioning and clock synchronization. *Nature*, 412:417–419, 2001.

[876] Hidetoshi Katori, Masao Takamoto, V. G. Pal'chikov, and V. D. Ovsiannikov. Ultrastable optical clock with neutral atoms in an engineered light shift trap. *Phys. Rev. Lett.*, 91:173005–1–4, 2003.

[877] I. Courtillot, A.Quessada, R.P. Kovacich, A. Brusch, D. Kolker, J.-J. Zondy, G. D. Rovera, and P. Lemonde. Clock transition for a future optical frequency standard with trapped atoms. *Phys. Rev. A*, 68:030501–1–4, 2003.

[878] Tetsuya Ido and Hidetoshi Katori. Recoil-free spectroscopy of neutral Sr atoms in the Lamb-Dicke regime. *Phys. Rev. Lett.*, 91:053001–1–4, 2003.

[879] Chang Yong Park and Tai Hyun Yoon. Efficient magneto-optical trapping of Yb atoms with a violet laser diode. *Phys. Rev. A*, 68:055401–1–4, 2003.

[880] S. G. Porsev, A. Derevianko, and E. N. Fortson. Possibility of an ultra-precise optical clock using the $6\,{}^1S_0 \rightarrow 6\,{}^3P_0^o$ transition in ^{171,173}Yb atoms held in an optical lattice. *Phys. Rev. A*, 69:021403(R)–1–4, 2004.

[881] R. G. Helmer and C. W. Reich. An excited state of ^{229}Th at 3.5 eV. *Phys. Rev. C*, 49:1845–1858, 1994.

[882] E. V. Tkalya. Properties of the optical transition in the ^{229}Th nucleus. *Physics – Uspekhi*, 46:315–324, 2003.

[883] E. Peik and Chr. Tamm. Nuclear laser spectroscopy of the 3.5 eV transition in Th-229. *Europhys. Lett.*, 61:161–186, 2003.

[884] P. Uhrich, P. Guillemot, P. Aubry, F. Gonzalez, and C. Salomon. ACES microwave link requirements. In *Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum and The IEEE International Frequency Control Symposium*, pages 213–216, 1999.

Index