Engineering Aspects of Terahertz Time-Domain Spectroscopy

by

Withawat Withayachumnankul

B Eng (Electronic Engineering, Honours)
King Mongkut's Institute of Technology Ladkrabang, Thailand, 2001
M Eng (Electronic Engineering)
King Mongkut's Institute of Technology Ladkrabang, Thailand, 2003

Thesis submitted for the degree of

Doctor of Philosophy

in

School of Electrical & Electronic Engineering
Faculty of Engineering, Computer & Mathematical Sciences
The University of Adelaide, Australia

December, 2009

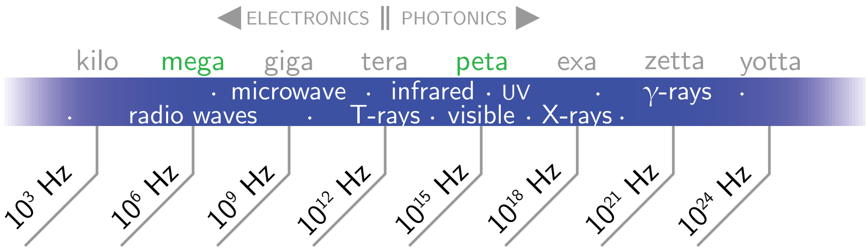
Chapter 1

Introduction

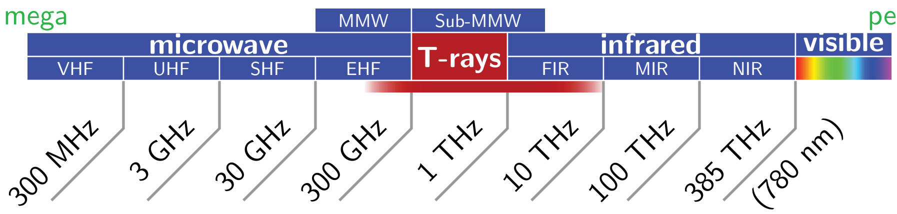
TERAHERTZ or T-ray radiation is a relatively unexplored frequency region of the electromagnetic spectrum, roughly situated on the border between electronics and photonics. The recent advent of terahertz time-domain spectroscopy (THz-TDS) endows researchers with a promising access to this frequency band. Engineering aspects of THz-TDS become important in improving its reliability and functionality. This introductory chapter offers a brief background of T-rays & THz-TDS, along with the thesis outline.

1.1 Definition of T-rays

Terahertz or T-ray radiation occupies the electromagnetic spectrum between 0.1 and 10 THz (Abbott and Zhang 2007), bridging the gap between the electronic and photonic worlds. Figure 1.1(a) depicts a broad view of the electromagnetic spectrum with the T-ray band in the centre. This frequency band overlaps with microwaves on one side and with the far-infrared (FIR) on the other side. A closeup in Figure 1.1(b) reveals the overlapping portions on the spectrum. On the longer-wavelength side, the conventional EHF microwave band ends at 300 GHz or 0.3 THz, whilst the T-ray band starts at 0.1 THz. At the other side, the FIR frequencies reach down to 1 THz, well below the upper boundary of T-rays at 10 THz. Overlapping of two frequency bands is not unusual, as the precedence already exists with, for example, the bands of X-rays and gamma rays (Abbott and Zhang 2007). Convention defines the terms millimetre and submillimetre waves spanning from 30 GHz to 300 GHz and from 300 GHz to 3 THz,



(a) Electromagnetic spectrum—broad view



(b) Electromagnetic spectrum—closeup to T-ray neighbours

Figure 1.1. Electromagnetic spectrum. T-rays are loosely defined between 0.1 to 10 THz, which is the gap between electronics and photonics. The lower and upper ends of the T-ray band overlap the EHF and FIR bands, respectively. A subset of T-rays are also referred to as submillimetre waves in passive detection systems in the 0.3 to 3 THz range. Note that panel (b) represents an expanded portion of panel (a) in the frequency range indicated in green coloured font.

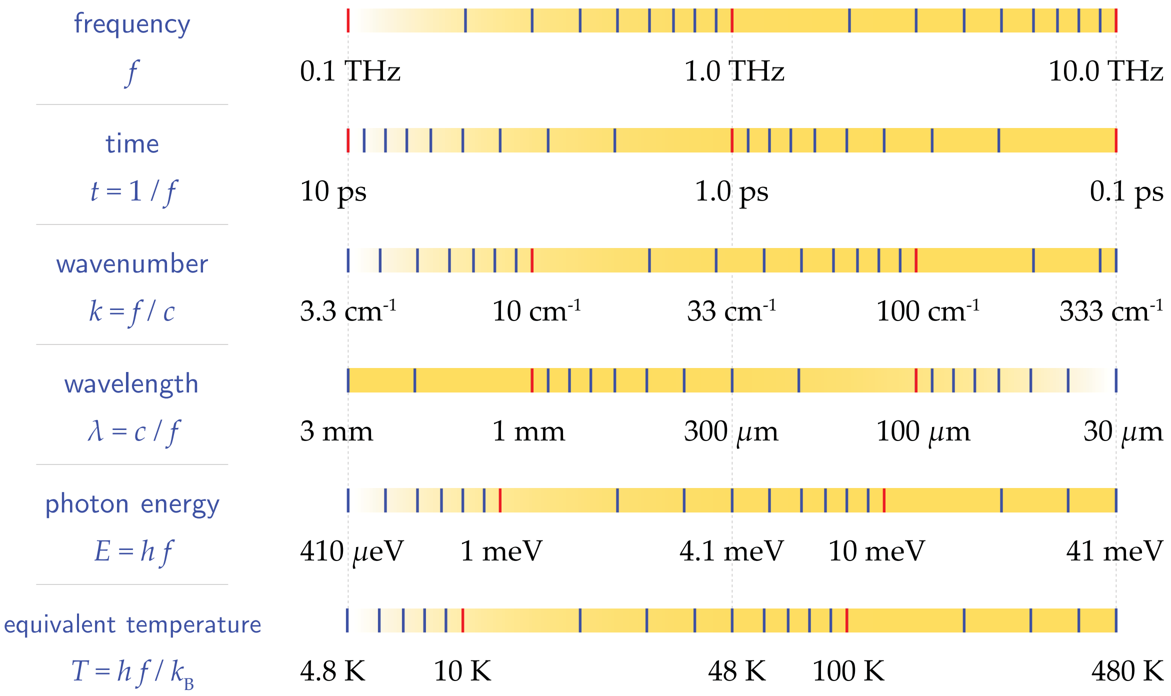


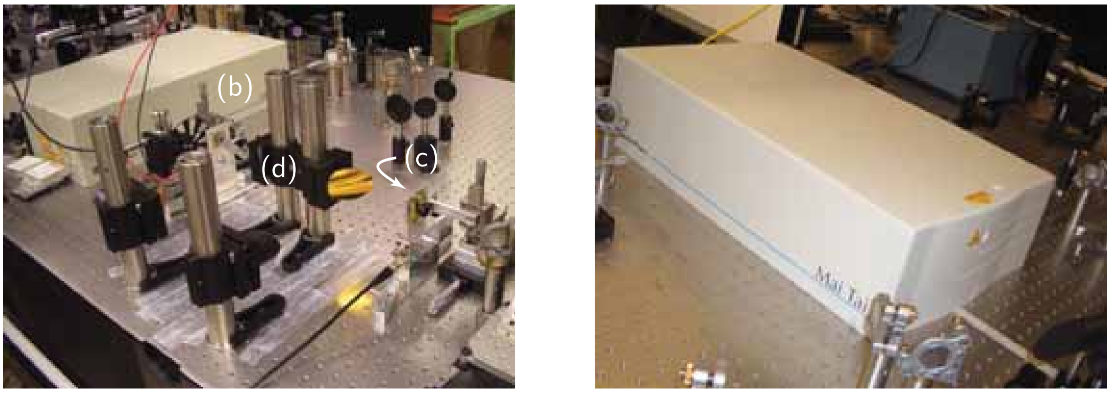
Figure 1.2. T-ray band in different units. The T-ray band is plotted in the units of frequency, time, wavenumber, wavelength, photon energy, and equivalent temperature. Here, h denotes Planck constant, equal to 4.135667×10^{-15} eV·s, and k_B denotes Boltzmann's constant, equal to 8.617343×10^{-5} eV/K. Note that wavenumber is conventionally used by spectroscopists, which is defined as $1/\lambda$.

respectively (Wiltse 1984). The term 'submillimetre wave' is usually reserved for passive detection systems, and is common in the field of astronomy. Figure 1.2 compared the T-ray band in different units. One terahertz is equivalent to 1 ps, 33 cm^{-1}, 300 μm, 4.1 meV, and 48 K.

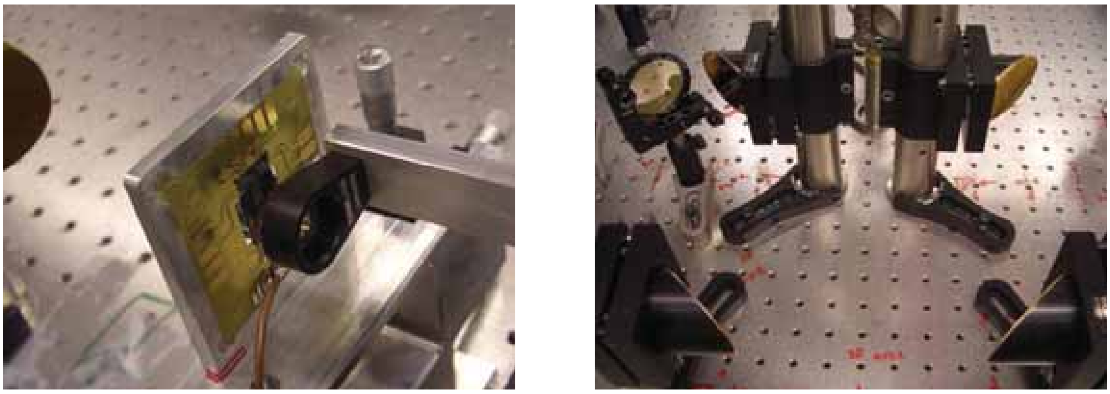
The T-ray spectrum was referred to as the 'terahertz gap' due to limited access to this frequency band with prior existing technologies. Approaching the band using electronic devices is hindered by the presence of parasitic capacitances, which function as lowpass filters. Incoherent radiation from thermal sources is swamped by thermal background radiation from surrounding objects at room temperature. Media with energy level transitions corresponding to T-ray frequencies are rare, deterring the realisation of stimulated emission. Furthermore, at room temperature, $k_B T/h = 5.6$ THz and thus energy levels become undesirably thermalised in the T-ray regime. Apart from these limitations, propagation of T-rays in the ambient environment suffers the problem of high atmospheric absorption (Barnes *et al.* 1935, van Exter *et al.* 1989b).

1.2 Generation and detection of T-rays using THz-TDS

A wide range of techniques for T-ray generation and detection have been proposed to date. One of the recently developed techniques is THz-TDS, which is a central theme of this thesis. Owing to its unprecedented capability in generating and detecting broadband coherent T-ray radiation, THz-TDS has received much attention from researchers. An example of a THz-TDS system is shown in Figure 1.3. A key component supporting the technique is an optical mode-locked laser, whose ultrashort pulses are utilised for both generation and detection of T-rays. At the emitter, femtosecond laser pulses or



(a) system overview (b) optical mode-locked laser



(c) photoconductive antenna (d) guiding optics

Figure 1.3. THz-TDS system. A THz-TDS system (a) is typically composed of a mode-locked laser (b), emitter/receiver, guiding optics (d), lock-in amplifier, and a controlling unit. The emitter and receiver for the system in the figure are photoconductive antennas (c). The T-ray beam path is usually purged with nitrogen to prevent interaction of T-rays with ambient water vapour. In this photo is one of the T-ray systems housed in the Adelaide T-ray laboratory.

pump pulses are down-converted to subpicosecond T-ray pulses via, for example, photocarrier transportation or optical rectification. A radiated T-ray pulse is composed of continuous frequency components, spanning from a few hundred gigahertz to a few terahertz or more. When the pulse interacts with a material, it is altered in amplitude and phase according to the dielectric response of the material in the T-ray regime. At the detection end of a THz-TDS system, incident T-ray pulses experience coherent sampling by a series of probe laser pulses, synchronous with pump pulses. Information registered by the system is a time-resolved T-ray pulse, which is ready for further processing. More discussion about the fundamentals of THz-TDS is provided in Chapter 2.

1.3 Applications for T-rays

The use of T-rays to investigate substances is a rapidly moving field. In general, nonpolar and nonmetallic materials, such as wood, ceramics, paper, plastics, or fabrics, are transparent or translucent to T-ray radiation. The transparency motivates the use of T-rays in security screening and quality control—contraband concealed beneath clothing (Ferguson *et al.* 2003, Zandonella 2003) or products contained in some types of packages (Morita *et al.* 2005, Jördens and Koch 2008) can be detected by T-ray sensors. Moreover, the wideband radiation of T-rays permits discrimination of a number of substances, including explosives and biological hazards (Choi *et al.* 2004), via their spectroscopic fingerprints. The strong absorption of T-ray energy by water molecules, which complicates measurements in open-air settings, has merits in biology, where T-rays are highly sensitive to the hydration level in biological tissue (Hadjiloucas *et al.* 1999, Han *et al.* 2000). In addition, applications of T-rays to medical and dental diagnoses are promising (Pickwell and Wallace 2006), chiefly because T-ray radiation is non-ionising and noninvasive. Further discussion on the safety issues of T-rays can be found in Section 3.5.

Regarding fundamental science, T-rays are of great importance to spectroscopy, as the frequencies underlie a wide range of significant physical phenomena. Essentially, molecules in the crystalline phase resonate at T-ray frequencies owing to intermolecular vibrational transitions (Fischer *et al.* 2002, Walther *et al.* 2003), whilst polar molecules in the gas phase resonate due to rotational transitions (Harde *et al.* 1997b, Bernath 2005)—the unique resonances of molecules cause spectral features, which are useful for material characterisation. Moderately doped semiconductors have plasma frequencies and damping rates within the T-ray frequency region, i.e., between 0.1 and 2.0 THz (van Exter and Grischkowsky 1990b, Jeon and Grischkowsky 1997). The plasma frequency and damping rate are proportional to the carrier density and mobility of semiconductors, respectively (Kudman 1963)—thus, T-ray spectroscopy is ideal for the study of carrier dynamics in semiconductors. Critical frequencies for Debye relaxation processes in liquids lie at T-ray frequencies (Kindt and Schmuttenmaer 1996), and the study of liquids with T-rays therefore enhances the understanding of the process. T-rays are also used to stimulate Rabi oscillations in two-level impurity states in semiconductors, which enables the manipulation of physical qubits (Cole *et al.* 2001, Ng and Abbott 2002, Brandi *et al.* 2003).

Since the emergence of the technology, a variety of T-ray applications have been reported. A more complete review of the applications based on THz-TDS is given in Chapter 3.

1.4 Prospects for T-rays

Researchers in T-ray groups worldwide have been working to improve T-ray hardware and investigate its potential applications. The hardware, particularly for spectroscopy and imaging applications, has been anticipated to provide higher radiation energy, finer spectral resolution, wider spectral range, and faster acquisition rate in a compact size. Moreover, the distinctive abilities of T-rays have been expected to become a solution to practical requirements not satisfied by other conventional technologies. Owing to these perspectives, significant funding has been attracted to the research and development through private sectors and government agencies.

In 2001, a European consortium initiated the ‘THz-bridge’ project that aimed for the investigation of the properties of and potential T-ray radiation damage to biological systems. The consortium convened and subsidised a number of universities and laboratories in Europe, which held existing T-ray systems. At the end of the programme in 2004, the project resulted in useful information made available to the public in the form of academic publications. The summary of the project can be found in Gallerano (2004). In 2004, The Ministry of Internal Affairs and Communications (MIC), Japan invited T-ray experts from Japan and overseas to conduct a study on the potential of T-rays. Two years later The National Institute of Information and Communications Technology (NICT), Japan, initiated the ‘Terahertz project’ with ambitious goals to develop

T-ray infrastructure, including hardware and material databases (Hosako *et al.* 2007). This project is anticipated to facilitate the development of future T-ray applications.

As of 2008, novel T-ray devices are already in use mainly for security purposes at military bases, airports, and financial centres. With strong financial support for T-ray research, it is envisaged that T-ray hardware will soon phase in as common equipment for industry, security, medical, and communication applications.

1.5 Thesis outline

As outlined in Figure 1.4, the thesis encompasses three major parts of the original contributions, including (**Part I**) signal enhancement and classification, (**Part II**) system evaluation and optimisation, and (**Part III**) T-ray optics. Each part contains two independent chapters. The fundamentals necessary for understanding the main idea

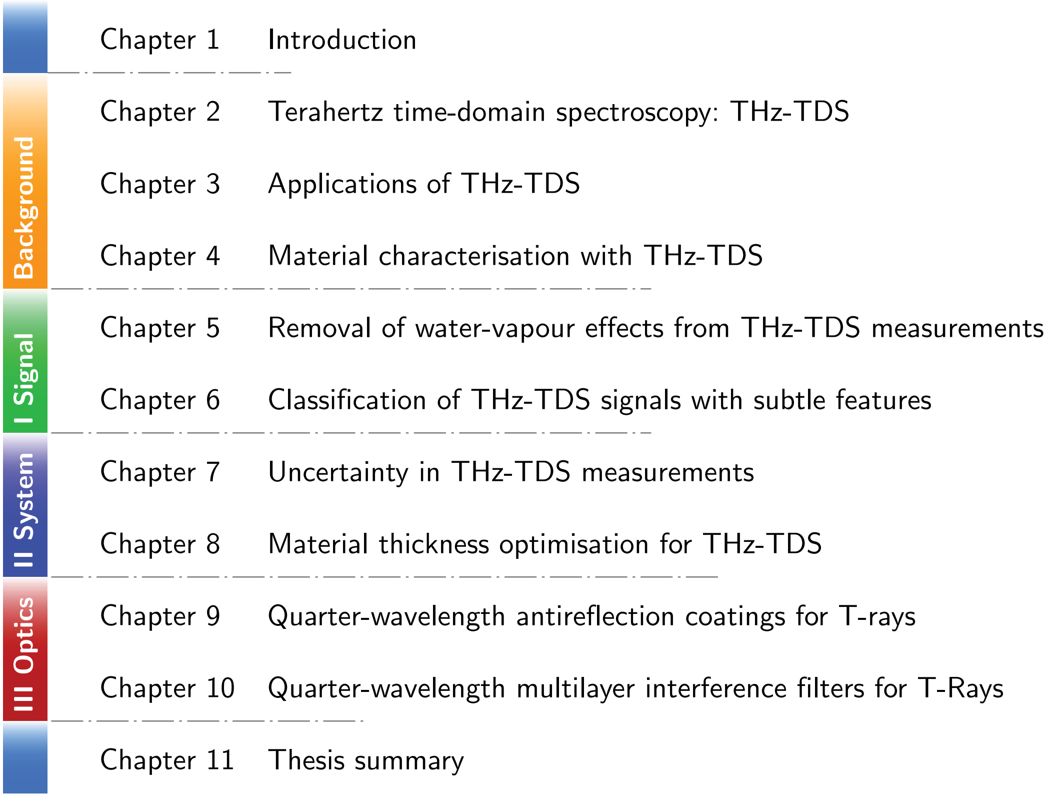


Figure 1.4. Thesis outline and original contributions. The thesis is composed of 11 chapters in total, divided into four major parts. The original contributions are distributed in three parts, from I to III. A supplementary part as a background of THz-TDS is offered. All chapters are virtually self-contained.

of each chapter is provided therein. Three additional background chapters revolving around THz-TDS are given to provide context. The detailed description for each part of the thesis is as follows:

Background provides the background of THz-TDS in three aspects as follows. Chapter 2 presents a historical review, principles underlying the operation, and variants of THz-TDS. Chapters 3 gives a survey on the emerging applications for THz-TDS, which include fundamental spectroscopic studies and advanced applications. Chapter 4 discusses theories for material characterisation via THz-TDS in various settings either in transmission or reflection modes, along with relevant issues on signal processing.

Part I—signal enhancement and classification is relevant to enhancement and classification of T-ray signals via digital signal processing. In Chapter 5, information underlying T-ray signals is enhanced through numerical removal of unwanted artefacts that are introduced by the response of water vapour during measurement. In Chapter 6, machine learning is recruited in classification of visually indistinguishable T-ray signals probing materials of the same general class.

Part II—system evaluation and optimisation focuses on THz-TDS systems with a particular interest in the measurement precision. In Chapter 7 an ISO standard for the evaluation of measurement uncertainty is adopted for assessing the uncertainty in THz-TDS measurements. The result is an analytical uncertainty model, which allows an improvement to the measurement precision through optimisation of a model parameter in the subsequent work presented in Chapter 8.

Part III—T-ray optics involves design, fabrication, and characterisation of THz-TDS hardware components, i.e., antireflection windows in Chapter 9 and multilayer interference filters in Chapter 10. The designs are based upon a conventional optical interference theory. Despite that, required materials and fabrication processes are completely different from those used in optics due to distinctive operating wavelengths that dictate the material response and structural dimensions.

1.6 Summary of original contributions

This thesis involves several original contributions in the field of terahertz technology, as declared in this section.

Typically, a THz-TDS measurement results in a signal plagued by absorption lines from ambient water vapour, unless carried out in a controlled environment. For the first time, a digital signal processing technique is implemented to remove those absorption lines (Withayachumnankul *et al.* 2008b). The technique allows effective field measurements using THz-TDS. In another contribution, machine learning is implemented to classify T-ray signals probing biological samples that have similar responses (Withayachumnankul *et al.* 2005). The automated classification broadens the usefulness of T-ray spectroscopy.

Another advancement contributing to T-ray measurement concerns characterisation and optimisation of measurement precision. In this contribution, the sources of error existing throughout the measurement process are analytically quantified for their impact on the output optical constants (Withayachumnankul *et al.* 2008c, Withayachumnankul *et al.* 2007b). The resulting uncertainty model offers a standard for evaluation of uncertainty in THz-TDS measurements. Furthermore, the model also enables analytical optimisation of the sample thickness, resulting in a higher measurement precision (Withayachumnankul *et al.* 2008a).

In addition, this thesis demonstrates the design and fabrication of two terahertz components, antireflection coatings (Withayachumnankul *et al.* 2007a) and multilayer interference filters (Withayachumnankul *et al.* 2008d), operated in the lower T-ray frequency regime, i.e., between 0.1 and 1 THz. Earlier, these components were designed and fabricated to serve astronomical observations in the neighbouring range of 1 to 10 THz. Although the fundamentals of optics are applicable in any frequency band, changing the operating range requires different materials and fabrication techniques.

These original contributions serve to advance subdisciplines of T-ray signal processing, T-ray measurement precision, and T-ray optics. Their broad impact may be envisaged in the applications of T-ray spectroscopy and communications.

Chapter 2

Terahertz Time-Domain Spectroscopy: THz-TDS

SINCE the first attempt to bridge the terahertz gap circa 1900, researchers struggled to come up with improved generation and detection techniques—their efforts resulted in limited advances due to technological difficulties. A new T-ray era is rapidly growing based on a strong foundation of emerging technologies and knowledge. THz-TDS is an example of modern T-ray hardware with fascinating and unique capabilities. This chapter reviews the ground-breaking developments of T-ray hardware, and focuses on the principles behind the success of THz-TDS.

2.1 Introduction

Since the previous century, attempts to fill the terahertz gap have emerged from both the far-infrared and microwave ends of the spectrum. The early activities in far-infrared research are evident by the publication of over 150 relevant articles during 1892-1922 (Palik 1977). Rubens and his colleagues, as major contributors at that period, utilised heat sources and reststrahlen plates to produce nearly monochromatic far-infrared radiation, whose frequency was selected via the narrow-band reflectance of the plates. It was this work of Rubens that led towards the discovery of the renowned Planck's law of black body radiation (Planck 1901). By 1911, Rubens and Baeyer managed to access the T-ray spectrum, the longer-wavelength side of the far-infrared, by using a mercury arc lamp as an incoherent source (Kimmitt 2003). During the same period, there existed a number of attempts to access the gap via radio-wave techniques. Hertzian spark gap generators (also known as Hertzian oscillators), essentially a metal dipole biased by a high-voltage step function, were anticipated for T-ray[1] generation. The spark gap generator prototype, constructed by Hertz around 1880s in order to prove Maxwell's equations, radiated microwave frequencies. Many years later, an *improved Hertzian oscillator* by Nichols and Tear (1923) reached the frequency of 0.167 THz. It was followed by a *paste radiator*, a refined spark gap generator design, by Glagolewa-Arkadiewa (1924), which pushed the radiated frequency towards 3.65 THz. This and other related experiments of Glagolewa-Arkadiewa could be regarded as the first success in bridging the gap between the electronic and photonic worlds, as indicated in Figure 2.1. Despite that, compared with thermal sources, Hertzian spark gap generators did not receive wide acceptance due to issues with instability (Wiltse 1984). Most experiments on T-ray molecular spectroscopy carried out afterwards relied on thermal sources. In 1925, T-ray molecular spectroscopy was first reported by Czerny for his observation of the rotational spectrum of HCl from 3-10 THz (Ginsburg 1977). Years after, the spectrum of atmospheric water vapour between 1.76-7.9 THz was reported by Barnes *et al.* (1935) and 4-16 THz by Randall *et al.* (1937).

Terahertz research activities were nearly abandoned after the eruption of World War II, since the efforts were put into the development of radar systems exploiting microwave frequencies. It is worth mentioning that microwave sources were invented during this period, as they were subsequently engineered to operate at higher frequencies later on. Examples of these sources include klystrons and cavity magnetrons (Robinson 1958).

---

[1] However, it was not until the mid-1990s that Bell Labs coined the term 'T-rays' itself.

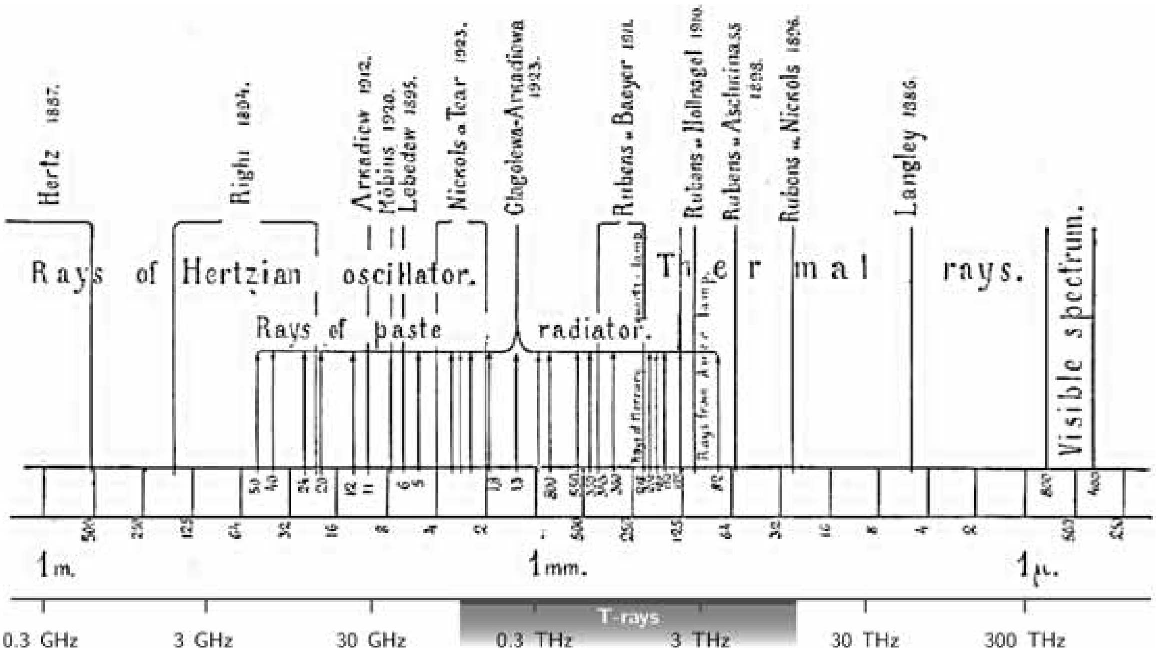


Figure 2.1. Experiments around the T-ray regime as of 1924. By 1923, the terahertz gap was populated from both sides of the spectrum. Radiation below 0.1 THz and above 1 THz was available from Hertzian spark gap generators and thermal sources, respectively. A paste radiator, invented by Glagolewa-Arkadiewa (1924), bridged the two separate regions. In brief, the paste radiator utilises small metal particles of brass and aluminium suspended in mineral oil as oscillators. The oil paste, once stimulated by a very high electrical current, radiates T-rays via these particles. After Glagolewa-Arkadiewa (1924)—the frequency axis is overlaid for ease of interpretation.

The first klystron, operated as an oscillator and amplifier, was invented by Varian and Varian (1939). Its operation is based on manipulation of an electron beam by an RF electric field. The first cavity magnetron was subsequently devised in 1939 by Boot and Randall. It generates microwave radiation by resonance induction from electrons circling inside a circular chamber (Boot and Randall 1976). The frequency emitted by magnetrons is not controllable. For this reason, years later they were retired from radar systems and instead found their destiny in microwave ovens. High-power klystrons are still in use for radars and also as a supporting unit for powerful T-ray sources.

Soon after the war, the T-ray research revived with the fundamentals being based on knowledge obtained from the earlier microwave research. A number of diode frequency multipliers were constructed in attempts to boost the frequencies from microwave sources towards the T-ray regime (Coleman 2000). Moreover, a backward wave oscillator (BWO), also known as carcinotron, was invented independently by Epsztein (1951) and Kompfner (1953). This oscillator produces coherent monochromatic microwave radiation, whose frequency is tunable over a wide spectral range. Its operating frequencies are now extended well into the T-ray band (Volkov *et al.* 1985, Gorshunov *et al.* 2005). This postwar period saw a significant development of broadband thermal detectors. In the earlier days, thermopiles were used as only far-infrared detectors, until the invention of bolometers by Langley (1881). By 1947, Golay cells, named after the inventor, were available with an improved sensitivity (Golay 1947), followed by cryogenic bolometers (Low 1961). The rapid advance of computer technology during this period permitted resolving Fourier components from complicated interferograms. Consequently, commercial Fourier transform infrared spectrometers (FTIR), exploiting a mercury lamp source and Golay detector, became available to the marketplace in the mid 1960s (Griffths 1983). Since then FTIR spectroscopy has been an important spectroscopic approach for the spectral band above a few terahertz.

The advent of the first working laser in 1960 attracted major research interest. A number of T-ray lasers were subsequently invented based on the principle of optical lasers. The first T-ray gas laser by Crocker *et al.* (1964) exploited DC-pumped water vapour, which produces stimulated emission at frequencies between 3.8-13 THz. Lower frequencies in the 0.4-4.3 THz range were reachable by other optically-pumped gas lasers (Chang and Bridges 1970, Chang *et al.* 1970). These gas lasers provide coherent and high power T-ray radiation ranging from few milliwatts to ten watts. However, they rely on vibrational or rotational transitions of gaseous molecules, and thus the radiated frequencies are discrete and fixed to molecular inherent transition levels. Note that at this stage some of the potential of terahertz radiation was realised (Senitzky and Oliner 1970). A powerful and tunable T-ray source, based on free electron lasers (FELs), was not ready until 1985. In general, an electron beam, accelerated at relativistic speeds in an FEL, is wiggled by a periodic magnetic field to produce coherent radiation with its frequency tunable via the electron beam energy or magnetic field strength. Elias *et al.* (1985) demonstrated an FEL operating at 0.3-0.77 THz and 10 kW peak power. Now FELs can span the entire T-ray region (Ramian 1992). Despite their excellent output power and unparalleled frequency tunability, FEL sources are large and expensive and hence not easily accessible.

A widespread resurrection of T-ray research activities followed the realisation of a photoconductive antenna (PCA) for T-ray radiation by Auston *et al.* (1984a). The antenna, composed of a pair of DC-biased metallic lines imprinted on a semi-insulating substrate, radiates picosecond T-ray pulses in response to excitation by subpicosecond optical pulses from a mode-locked laser. Interestingly, this invention is an advanced replication of a Hertzian spark gap generator, which dates back to the last century (Coleman 2000). The extension of this idea to metallic lines on a semi-insulating substrate was first carried out by Mourou *et al.* (1981b) in the gigahertz range, which then laid the foundation for Auston's work in the T-ray regime. Afterwards, many different types of efficient T-ray sources and detectors have been reported. The developments of these modern sources, including synchrotron sources (Carr *et al.* 2002), quantum cascade lasers (QCLs), frequency mixers, all-electronic systems (van der Weide *et al.* 1993, van der Weide *et al.* 2000), etc., have been propelled essentially by the advances in quantum mechanics, semiconductor science, and laser technology. This latest era of T-ray research and developments remains active till the present day.

This chapter reviews T-ray hardware emerging during the present era, with a focus on THz-TDS systems. Section 2.2 briefly discusses the fundamentals of some promising T-ray generation and detection techniques. The early developments of important THz-TDS sources/detectors, including PCAs and nonlinear crystals, are covered in Section 2.3. The basic principles underlying their operations are given in Section 2.4. Full transmission-mode THz-TDS systems utilising these types of sources and detectors are discussed in Section 2.5. Many other arrangements of THz-TDS realised to date are surveyed in Section 2.6.

2.2 Major T-ray generation and detection techniques

Some major T-ray techniques encompass quantum cascade lasers & quantum well photodetectors, frequency mixers, FTIR spectroscopy, and THz-TDS. The two latter techniques are different from the others in that they are complete spectroscopic systems, rather than being separate sources and detectors. This section provides a brief review of these techniques. More details of T-ray sources and detectors can be found in the excellent review papers of Coleman (2000) and Siegel (2002).

2.2.1 Quantum cascade lasers & quantum well photodetectors

Quantum cascade lasers (QCLs) are heterostructure semiconductor lasers, which can be engineered to lase at frequencies in the mid- to far-infrared regimes. A typical semiconductor laser exploits fundamental energy bandgaps in bulk semiconductors and interband transitions of electrons for lasing, but these bandgaps are too large for T-ray energies. A superlattice structure in QCLs proposed by Kazarinov and Suris (1971) splits a conduction band to subbands, which allow intersubband transitions at lower energy levels corresponding to T-ray photon energies. Furthermore, an electron experiencing an intersubband transition at a given layer of a superlattice can tunnel to another layer and experience a further transition. Electron tunnelling in a superlattice thus allows a single electron to produce multiple photons at different stages. The concept had not been realised for T-ray radiation until Köhler *et al.* (2002) used a GaAs/Al$_x$Ga$_x$As structure to lase at 4.4 THz with output powers more than 2 mW at the maximum operating temperature of 50 K. Successive developments offered both pulsed and CW T-ray QCLs at higher powers and more suitable operating temperatures. An extensive review of QCLs can be found in Williams (2007).

Quantum well photodetectors (QWPs), as a complement to QCLs, are semiconductor incoherent detectors that are able to respond to T-ray radiation. QWPs utilise a similar superlattice structure for splitting of the conduction band to subbands, where incident T-ray photons can cause the excitation of electrons and subsequently the flow of photocurrent. Owing to the compact size of a superlattice, multiple QWPs can be fabricated into an imaging array, with excellent 2D spatial resolution. The first QWPs that operated below the phonon energy of GaAs, or below 8 THz, were demonstrated (Liu *et al.* 2004a, Liu *et al.* 2005, Graf *et al.* 2004). Similar to QCLs, QWPs are operated at best at cryogenic temperatures, at which photocurrent resulting from background noise is suppressed.

2.2.2 Frequency mixers

Difference frequency generation (DFG) can be exploited to generate or detect CW T-ray radiation. Essentially, T-ray generation with this technique needs two CW optical 'pump' lasers with slightly different frequencies. The two optical beams from the pump lasers are mixed in a nonlinear crystal that has a high second-order susceptibility, e.g., quartz (Zernike and Berman 1965), $LiNbO_3$ (Auston *et al.* 1973, Ding and Khurgin 1998), GaP (Suto *et al.* 2005, Ding and Shi 2006, Tomita *et al.* 2006), under the phase-matching condition. The frequency of the output CW beam is equal to the difference between the laser frequencies. By selecting two appropriate laser frequencies, T-rays can be produced. The frequency tunability of the output depends on the function of the pump lasers. However, the conversion efficiency of this technique is limited by the Manley-Rowe relation to less than one percent (Brown *et al.* 2004). The process is reciprocal in nature and thus upconversion of T-rays to visible or infrared frequencies is also possible via DFG. By mixing CW T-ray radiation with an optical pump beam from a laser inside a nonlinear crystal, an optical idler wave, whose frequency equals the difference between the two input frequencies, is produced. This idler wave can be readily detected by an optical detector at room temperature. Some crystals used to demonstrate T-ray upconversion are, for example, GaP (Ding and Shi 2006), GaAs (Khan *et al.* 2007).

Alternative hardware capable of DFG is a photoconductive mixer or photomixer made of a semiconductor with lithographed electrodes. A photomixer illuminated by two CW pump lasers with different frequencies supplies free electron-hole pairs, a fraction of which oscillate at the difference frequency. Under the presence of DC bias, the free carriers radiate at their oscillating frequencies. The high frequency radiation is limited by the carrier lifetime of the semiconductor and the RC time constant of the electrode. Some semiconductors with short carrier lifetimes, widely used for photomixers, include LTG GaAs (McIntosh *et al.* 1995, Brown *et al.* 2004, Stone *et al.* 2005) and ErAs:GaAs (Bjarnason *et al.* 2004). Note that photomixers share a similar structure with PCAs that are used for pulsed T-ray generation. Further discussion of photomixers for CW T-ray radiation is available in book chapters from Matsuura and Ito (2005) and Duffy *et al.* (2004).

2.2.3 Fourier transform infrared spectroscopy

Fourier transform infrared spectroscopy (FTIR) typically comprises an incoherent thermal source such as a mercury arc lamp or SiC globar, and an incoherent thermal detector such as a helium-cooled bolometer (Griffths 1983, Griffths and de Haseth 1986), arranged in the form of a Michelson interferometer. The source/detector pair provides an excellent frequency response spanning from the far-infrared to visible light. The intensity measured at a detector results from interference between two beams from a single source taking two different paths; one path length is fixed and the other is varied. For a monochromatic source, a sinusoidal interferogram is measured by the detector as a function of the path length. On the other hand, a broadband source such as that used in FTIR spectroscopy results in a multiple-wavelength interferogram, which requires a Fourier analysis to resolve the spectral components. By placing a sample in one arm of the interferometer, an interferogram for the sample can be measured. The sample interferogram is then compared in the frequency domain to that of the free-path reference to determine the complex transmittance of the sample.

2.2.4 Terahertz time-domain spectroscopy: THz-TDS

THz-TDS is an efficient technique for generation and detection of broadband coherent T-ray radiation. As a vital part of THz-TDS, a femtosecond mode-locked laser is utilised for both generation and detection. Interaction of ultrashort laser pulses with a certain component results in burst of subpicosecond pulses spanning a lower frequency region of T-rays, i.e., from a few hundred gigahertz to a few terahertz. The emitted T-ray pulses are then guided through a sample, if any, to a detector, where the pulses induce a local change in the electric field. Via a beam splitter a fraction of the ultrashort laser pulse power is incident at the detector—together with a local electric field change, this leads to optical-gating of the T-ray pulses. This detection scheme resolves the temporal profile, or the amplitude and phase, of a coherent T-ray pulse with an excellent SNR. The complex dielectric constant of a sample can then be determined by deconvolving the reference from the sample measurement. More details of the generation and detection mechanisms in THz-TDS will be elaborated in Section 2.4 and a discussion of transmission-mode THz-TDS systems will be given in Section 2.5.

As mentioned earlier, the frequency span of THz-TDS is somewhere from a few hundred gigahertz to a few terahertz, whilst the span of FTIR spectroscopy is from a few terahertz onwards at room temperature. Thus, there exists an overlapping operating region of the two techniques. The overlapping region will be more significant, following the progress of hardware developments; recent trends indicate that THz-TDS can extend its sensitivity beyond ten terahertz, and a thermal source and detector of FTIR spectroscopy can reach down to a longer wavelength. Moreover, both techniques have comparable frequency resolutions of the order of a few gigahertz, dependent on the scanning range of a delay line (van Exter *et al.* 1989b, Grischkowsky *et al.* 1990). Though, it is worth mentioning the differences between THz-TDS and FTIR spectroscopy. First, THz-TDS, based on coherent generation and detection, delivers full amplitude and phase information of the signal, whilst FTIR spectroscopy loses the phase information due to its incoherent nature. Because FTIR spectroscopy is incoherent-based, thermal background noise sets the ultimate sensitivity of the system. In contrast, coherent time-gated detection of THz-TDS is insensitive to thermal background noise. Therefore, THz-TDS systems do not require cooling, and, moreover, the technique is very suitable for the measurement of a high-temperature sample (Cheville and Grischkowsky 1999, Han *et al.* 2001). In terms of flexibility, an FTIR spectroscopy setting is not convenient for specific measurements that require hydrated environments or optical-pump/terahertz-probe measurements, as does a THz-TDS setting (Markelz 2008). Other specific details of the comparison can be found in Han *et al.* (2001).

2.3 Early developments of THz-TDS sources & detectors

Advances in principles, materials, and techniques for ultrafast-response electrooptics, coupled with the advent of ultrashort pulse lasers, opened up a new spectroscopic modality, THz-TDS, for the T-ray regime in 1989. Behind the success of THz-TDS are decades of development of appropriate sources and detectors. As illustrated in Figure 2.2, two main electrooptic approaches, i.e., PCAs and nonlinear crystals, have been competing for the speed and functionality since 1970's. It started with T-ray generation using nonlinear crystals via optical rectification in 1970. Electrical transients generated from Auston switches via photocarrier transport reached picosecond rates and faster during 1975-1980. Later on in 1981, a system capable of coupling picosecond transients into free space was demonstrated. Subpicosecond electrooptic sampling via the Pockels effect was demonstrated in 1982, and soon after, effective T-ray generation based on Čherenkov-like radiation was realised in 1983. The first PCA, evolving from the Auston switch, made its debut in 1984, yet it was nonlinear crystals that performed the first T-ray contact spectroscopy in 1985. Tighter coupling of T-rays from a PCA transmitter to receiver by means of guiding optics in 1988 led to the first free-space THz-TDS

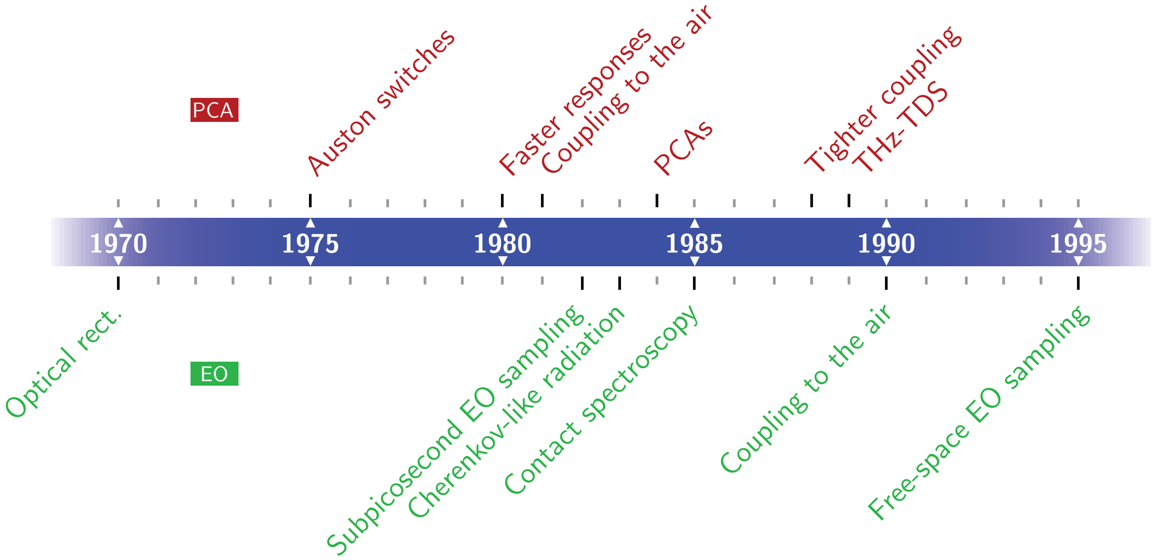


Figure 2.2. Timeline for THz-TDS development. The timeline lists some significant progresses in the field of THz-TDS research, starting from 1970's. The progresses relevant to PCAs are listed on the upper panel, whilst those for electrooptics (EO) are listed on the lower panel. The detailed discussion for PCAs and EO is given in Section 2.3.1 and 2.3.2, respectively.

system in 1989. The functionality of nonlinear crystals was fostered by the accomplishment of free-space coupling and free-space sampling in 1990 and 1995, respectively. Sections 2.3.1 and 2.3.2 cover these ground-breaking developments in details.

2.3.1 Photoconductive antennas

Auston switches—1975. The story started with the research on optoelectronic or photoconductive switching. A photoconductive switch is a fast-response switch, whose structure is essentially composed of two electrodes embedded on a semi-insulating substrate. In the absence of optical pulse excitation, the switch is off due to the high resistivity of the substrate. The illumination of an ultrashort optical pulse at the substrate between the electrode gap turns the switch on almost instantaneously, as the substrate conductivity suddenly changes from semi-insulator to a quasimetallic state via the photoconductive effect. The first switch of its kind, as shown in Figure 2.3, was devised by Auston (1975) at Bell Laboratories. This fast-response switch was made of high-resistivity silicon, and its operation was controlled via a mode-locked Nd:glass laser. The on-state rise time is within a few picoseconds, dependent on the optical

NOTE:
This figure is included on page 22
of the print copy of the thesis held in
the University of Adelaide Library.

Figure 2.3. First photoconductive switch. The operation of the switch relies on changes in the substrate conductivity via the application of ultrashort pulses from mode-locked lasers. A 0.53-μm pulse excites free carriers at the surface of the substrate, leading to the on state. A 1.06-μm pulse penetrates deeper, and hence excites free carriers to short the bias to the ground, causing the off state. Adapted from Auston (1975).

pulse duration. Because high-resistivity silicon has a relatively long carrier lifetime, changing to the off-state requires shorting the switch to ground by another laser pulse with a deeper penetration depth. The switch was implemented in conjunction with other components to control microwave (Johnson and Auston 1975) or optical signals (LeFur and Auston 1976). In an improved version, Lee (1977) changed the substrate material of the switch to Cr-doped semi-insulating GaAs, which has a faster carrier lifetime, i.e., less than 100 ps. This modified switch requires no second 'OFF' laser beam, as the fast carrier recombination rapidly increases the gap resistivity.

Faster response—1980. Afterwards, the switch was configured for a few other functions, but the main aim was to improve its response time—the rise time via a faster pump laser, and the fall time via a semiconductor with a shorter carrier lifetime. For example, the switch was operated as photodetector for detecting picosecond laser pulses. For this operation, Auston *et al.* (1980b) introduced a substrate made of amorphous silicon with a carrier lifetime of 30 ps, and Smith *et al.* (1981) further increased the speed by using ion-implanted SOS with a lifetime of only 8 ps. Basically, an operation of a biased switch causes a picosecond electrical transient to propagate down the circuit. It was noticed that this transient can be sampled coherently by another photoconductive switch connected in tandem on the same circuit. Auston *et al.* (1980a) demonstrated the sampling capability of the switch using evaporated amorphous silicon as a substrate with a carrier lifetime of 4 ps and a dye laser with a pulse duration of 3.5 ps.

Auston and Smith (1983) redesigned the electrodes of a generating switch to achieve 55-GHz transients, which were then sampled by another ion-implanted SOS switch with a lifetime shorter than 2.6 ps. Most of the research on fast-response switching during this initial period was pioneered or contributed by Auston at Bell Laboratories. Later, Auston was honoured with the switch bearing his name.

Coupling into the air—1981. The previous work demonstrated the ability of photoconductive switches to generate and sample guided picosecond electrical transients. Coupling of the resulting guided transients into free space is possible via an antenna. Example antenna structures for free-space coupling are those reported by Mourou *et al.* (1981a), Mourou *et al.* (1981b), and Heidemann *et al.* (1983). For Mourou *et al.* (1981a), a Cr-doped GaAs switch was connected to an X-band waveguide driving a dish-type antenna. Illuminating the switch by a 35-ps optical pulse from an Nd:YAG laser caused a free-space picosecond microwave burst with a FWHM of 50 ps. Mourou *et al.* (1981b) terminated a Cr-doped GaAs photoconductive switch at one end by a dipole antenna. As an optical pulse from a dye laser with a pulse width of 500 fs triggered the switch, microwave radiation with a pulse width of 2.8 ps was observed. In the design of Heidemann *et al.* (1983), the waveguide dispersion and coaxial-line loss affecting the guided electrical transients were ruled out altogether, as a Cr-doped GaAs switch was integrated into a slot-line antenna, which has a wideband flat response. A microwave pulse of 500 ps was radiated, following pumping of the switch by a 500-ps laser pulse. Note that the duration of an electrical pulse radiated by a coupling antenna essentially corresponds to the excitation laser pulse width.

Photoconductive antennas—1984. Nevertheless, among those earlier experiments on radiation coupling, the full capability of photoconductive switches for electromagnetic radiation had yet to be realised. Significant work stepping towards T-ray radiation with optoelectronic devices was performed by Auston *et al.* (1984a). In their experiment, as depicted in Figure 2.4, a pair of identical ion-implanted SOS switches, located on the opposite sides of an insulating slab, were designed to function as transmitting and receiving Hertzian dipole antennas. Excited by 100-fs optical pulses from a dye laser, the transmitting antenna radiated electrical pulses with a FWHM of 1.6 ps across the substrate. On the other side, the optical pulse beam, which was split from the excitation beam, gated the receiving antenna to sample the incoming electrical pulses.

NOTE:
This figure is included on page 24
of the print copy of the thesis held in
the University of Adelaide Library.

Figure 2.4. First photoconductive antennas. The two photoconductive antennas are located on the opposite sides of a 1.15-mm insulating slab. The transmitting antenna on the left is biased by DC, and the receiving antenna on the right is connected to a low-frequency amplifier. Excitation of the transmitting antenna by an ultrashort optical pulse results in radiation of an electrical pulse across the insulator to the receiving antenna, where the pulse is coherently sampled. Adapted from Auston *et al.* (1984a).

The configuration demonstrated for the first time generation and detection of freely propagating electrical pulses by photoconductive switches, which eliminate the need of dedicated waveguides, transmission lines, and antennas altogether. A photoconductive switch operated as an antenna is therefore referred to as a photoconductive antenna (PCA). Later on, Smith *et al.* (1988) studied the influence of the dipole length (50, 100, and 200 μm) of PCAs on the characteristics of radiated pulses. For each of the measurements, two identical ion-implanted SOS antennas, facing each other and separated by a 2-mm air gap, were used as a transmitter and receiver. Subpicosecond electrical transients were generated and detected via illumination from a dye laser with an optical pulse duration of 120 fs. It was found that a shorter dipole provides a shorter electrical pulse width and hence broader bandwidth. The broadest spectral band that they could achieve spanned from 100 GHz to 2 THz—well into the T-ray regime—for the 50-μm dipole antennas. In addition, it was shown that the amount of T-ray energy radiated into the substrate is much greater than that radiated into the air.

Tighter coupling—1988. Since the advent of the first photoconductive switch in 1975, most of the relevant research had been conducted by Auston and Smith at Bell Laboratories (named AT&T Bell Laboratories from 1984). Another prominent research group in this arena was led by van Exter, Fattinger, and Grischkowsky at IBM Watson Research Center, where optical techniques were applied to the manipulation of T-ray radiation for the first time (Fattinger and Grischkowsky 1988, Fattinger and Grischkowsky 1989, van Exter *et al.* 1989a). In Fattinger and Grischkowsky (1988), a gold-coated hemisphere sapphire mirror, attached to the backside of an ion-implanted SOS substrate, was used to collect 0.6-ps pulses from the transmitter and reflect them back onto the receiver, which was located on the same side as the transmitter. This experiment demonstrated an improved coupling efficiency via an optical mirror. In Fattinger and Grischkowsky (1989), two crystalline sapphire hemispherical lenses were attached to the backside of the separate transmitting and receiving PCAs with their dipoles located at the foci of the lenses. These lenses effectively collimated and focused a 0.6-ps T-ray beam that propagated across a distance of up to 1 m in free space with diffraction-limited divergence. No significant change in the pulse shape was observed, excepting the tail oscillations induced by ambient water vapour. Then van Exter *et al.* (1989a) redesigned the antenna structure using an ’H’ shape to boost the radiation power and bandwidth. In this new antenna design, photocurrent is able to flow only through the antenna arms whilst the coplanar transmission lines are more separated, causing higher radiation power and lower noise. Moreover, they used a pair of parabolic mirrors, in addition to the hemispherical lenses, to improve the transmission efficiency of T-rays in the free space. The first mirror helped to collimate the diverging beam from the source, whilst the second focused the collimated beam onto the detector. This new ’H’-type antenna with mirror/lens configuration has been widely adopted with virtually no change ever since.

Terahertz time-domain spectroscopy—1989. The free-space, tightly coupled T-ray system introduced by van Exter *et al.* (1989a) led to efficient T-ray spectroscopy. The first sample that was reported under measurement with this new system was water vapour, originally observed to cause the oscillations in the pulse tail. Via a Fourier analysis of the sample and reference pulses obtained from the system, van Exter *et al.* (1989b) were able to determine the strengths and positions of the nine strongest absorption lines of water vapour between 0.2 and 1.45 THz. In the same article, the term ’terahertz time-domain spectroscopy: THz-TDS’ was coined for this type of spectroscopic modality. Since then, THz-TDS has become very attractive for spectroscopy of materials in the T-ray frequency range. Nevertheless, the hardware development for

THz-TDS has not ceased, as researchers constantly aim for systems with wider bandwidth, faster scanning rate, higher transmitting power, and finer frequency resolution.

2.3.2 Nonlinear crystals

Optical rectification for picosecond pulses—1970. Shortly after the advent of working lasers in 1960, Bass *et al.* (1962) were the first to demonstrate the optical rectification process. In their experiment, KDP crystals were excited by 100-ns pulses produced by a ruby laser. The effect, in terms of induced potential, was observed via the electrodes attached to the crystal. Because of the second-order polarisation, the induced potential follows the intensity or envelope of a laser pulse. However, picosecond T-ray pulses generated via optical rectification had not been realised immediately afterwards due to the lack of fast-pulse lasers. Some years later, Yajima and Takeuchi (1970) of the University of Tokyo implemented the scheme using a mode-locked Nd:glass laser with a 10 ps pulse duration. Via optical rectification, the nonlinear crystals they investigated, including $LiNbO_3$, ZnTe, ZnSe, CdS, and quartz, under laser stimulation, produced radiated coherent broadband signals between 30 GHz and 0.6 THz in the absence of phase matching. Their experiment was further extended to meet phase-matching conditions, in Yajima and Takeuchi (1971). A $LiNbO_3$ crystal was oriented with respect to the laser beam direction to attain a phase match by means of birefringence. For the two specific orientations, the crystal produced narrow band signals peak at 0.47 and 0.58 THz. Nearly at the same time, Yang *et al.* (1971) at Berkeley performed similar experiments to generate T-rays. A mode-locked Nd:glass laser ejected 2-ps pulses into a $LiNbO_3$ crystal under the phase-matching conditions. The crystal rectified the laser pulses to produce coherent pulses, resonating somewhere between 0 and 0.42 THz, detectable by an FIR interferometer. The results confirmed their theoretical model published by Morris and Shen (1971). Though, the radiation efficiency for these early experiments was limited by the scarcity of fast laser sources.

Subpicosecond electrooptic sampling—1982. Valdmanis *et al.* (1982) at Rochester demonstrated subpicosecond EO sampling by using a nonlinear crystal or a *Pockels cell* made of $LiNbO_3$. Two synchronous optical beams from a colliding pulse mode-locked laser with a pulse duration of 120 fs were used for transient generation and EO detection. The first beam excited a Cr-doped semi-insulating GaAs photoconductive switch to produce an electric transient, which was then launched into the crystal via deposited electrodes. The transient amplitude was picked up by the second beam via the Pockels effect in the crystal. The phase matching condition was met by transmitting the second beam at an angle into the crystal, so that a component of the optical velocity matched the velocity of the electrical transient. Phase matching resulted in an improved temporal resolution, i.e., the system achieved a temporal resolution of < 4 ps, corresponding to a bandwidth in excess of 100 GHz. A voltage sensitivity of as low as 50 μV could be attained. In a similar experiment, Valdmanis *et al.* (1983) exploited $LiTaO_3$ for a Pockels cell. Note that $LiTaO_3$ is superior to $LiNbO_3$ in many aspects, but the important thing is that $LiTaO_3$ has an optical refractive index that is more suitable for velocity matching. In the experiment, the laser beam size was reduced to < 20 μm, compared with 100 μm in the earlier experiment. As a result, the subpicosecond temporal resolution of 850 fs could be achieved.

Čherenkov-like radiation—1983. Auston (1983) and Kleinman and Auston (1984) at Bell Laboratories proposed a model explaining the nature of an electrooptic transient produced by tightly focusing an optical beam into a nonlinear crystal. It was found that polarisation induced by a focused optical pulse can be approximated by a moving particle with a dipole moment, if the spatial and temporal distributions of the optical pulse are small enough. This localised polarisation radiates an electrical transient with T-ray frequencies. Since the phase velocity of T-rays is slower than the group velocity of the optical pulse, the appearance of T-ray radiation is characterised by a Čherenkov cone. Later on, Auston *et al.* (1984b) experimentally observed this phenomenon using a $LiTaO_3$ crystal. As illustrated in Figure 2.5, two parallel optical beams with a pulse duration of 100 fs from a mode-locked ring dye laser were focused into the crystal. The optical pump beam, focused to a waist of 3.8 μm, generated a Čherenkov cone of T-rays. The amplitude of this cone was detected via EO sampling by the copropagating probe pulse. The resolved T-ray signal had a peak amplitude of 10 V/cm and frequency range spanning from DC to 4 THz. It was concluded that the approach utilising a tightly focused ultrashort optical beam for Čherenkov-like radiation leads to high efficiency T-ray emission in a nonlinear crystal without the requirement of a phase-matching condition.

NOTE:
This figure is included on page 28
of the print copy of the thesis held in
the University of Adelaide Library.

Figure 2.5. Čherenkov-like radiation in a nonlinear crystal. Both pump and probe beams from an optical laser are focused into a nonlinear crystal. The pump beam produces a Čherenkov cone of T-rays, marked by the red shade. The apparent angle $\theta_c = 69^\circ$ of the cone is determined from the group velocity of the pump beam, $0.433c$, and the phase velocity of the radiated transient, $0.153c$, specifically for $LiTaO_3$. The probe beam samples the T-ray amplitude via the Pockels effect. Adapted from Auston *et al.* (1984b).

Spectroscopy of active crystals—1985. T-ray generation and detection with a nonlinear crystal had been developed to the point that the delivered energy and bandwidth were sufficient for phase-sensitive spectroscopy of materials in the T-ray regime. Realising this possibility, Cheung and Auston (1985) studied lattice vibrations in $LiTaO_3$ by observing its response via a T-ray waveform measured using a configuration similar to that shown in Figure 2.5. A detectable T-ray waveform contained damped oscillations confirming the existence of a phonon mode. Cheung and Auston (1986) determined complete absorption and dispersion spectra for $LiTaO_3$. A PCA receiver was attached to a crystal surface in parallel to the pump beam direction, as illustrated in Figure 2.6(a). By moving the pump beam relative to the PCA, a series of T-ray waveforms produced by the pump beam were registered as a function of the crystal thickness. From these waveforms, the absorption coefficient and refractive index of the crystal were calculated up to 1.6 THz. Instead of using PCA sampling, Auston and Nuss (1988) used EO sampling similar to that shown in Figure 2.5. The lateral distance between the two parallel beams was adjusted to observe the T-ray waveforms as a function of the crystal thickness. Also, EO sampling essentially responds faster than does PCA sampling, and thus extends an effective measurement bandwidth. From the extended spectra, a TO-phonon resonance at 6.23 THz and a two-phonon process at 2.7 THz were clearly identified.

Contact spectroscopy—1985. The approach to spectroscopy discussed in the previous paragraph is restricted to characterisation of active crystals where the emission process takes place. Auston and Cheung (1985) proposed the contact spectroscopy technique as a practice for the study of other materials. As illustrated in Figure 2.6(b), a target sample is in optical contact with a nonlinear crystal. A T-ray pulse produced inside the crystal is reflected at the crystal-sample interface back into the crystal, and detected by the optical probe pulse therein. According to Fresnel's law, this reflected pulse contains information about the sample's optical properties. In order to extract this information, deconvolution of the sample pulse with the reference pulse is necessary. With an identical setting, the reference measurement is performed with a material having a known T-ray response. In Auston and Cheung (1985), the technique was able to determine the electron density and mobility of a Sb-doped Ge sample, and the areal electron density and momentum relaxation time of GaAs/GaAlAs multi-quantum-well superlattice. Nuss *et al.* (1987) introduced a third optical beam to the

NOTE:
These figures are included on page 29 of the print copy of the thesis held in the University of Adelaide Library.

(a) spectroscopy of active crystal

(b) contact spectroscopy

Figure 2.6. EO spectroscopy in two configurations. (a) An active crystal, $LiTaO_3$ here, is a target of spectroscopy. A PCA receiver is attached to one side of the crystal and controlled via an optical gating beam. The receiver records a T-ray field generated by the optical pump beam positioned at a distance. The recorded waveform contains information about the absorption and dispersion of the crystal itself. Adapted from Cheung and Auston (1986). (b) The sample attached to a crystal is a target of spectroscopy. The T-ray field generated by the optical pump beam is partially reflected at the crystal-sample interface. This reflected pulse, containing information about the absorption and dispersion of the sample, is registered by the optical probe beam through the Pockel effect. Adapted from Auston and Cheung (1985).

setting in Figure 2.6(b). This beam, which was relatively powerful, hit the sample directly and caused a nonequilibrium state therein. Synchronisation between the third beam and the pump and probe beams allowed measurement of the dynamics of a sample following the optical injection. The sample under study was a GaAs/GaAlAs structure. Upon injection of a 2-eV optical pulse, hot carriers were created and their mobility could be observed as a function of the delay time after the injection. Although contact spectroscopy is applicable to a number of materials, free-space contactless spectroscopy is still advantageous.

Coupling to the air—1990. Typically, T-ray Čherenkov radiation, produced via optical rectification, makes a shallow angle of incidence to the crystal interface. Because of the high dielectric constant and hence the small critical angle of a nonlinear crystal crystal, the radiation suffers total internal reflection. This occurrence limits the application of the radiation to only contact spectroscopy and spectroscopy of the generating crystal itself, as discussed in the previous paragraphs. To overcome this limitation, Hu *et al.* (1990) derived a formula relating the angle of incidence of the pump beam to the refracted angle of the T-ray radiation. From this formula, an optical angle that permits the Čherenkov-like radiation to escape a nonlinear crystal can be determined. In their experiment with a $LiTaO_3$ crystal, the angle of incidence of the pump beam was set to 51°. With this angle, the propagation direction of the T-ray radiation was normal to the crystal surface, and therefore the radiation could be coupled out to the air. In addition, they suggested that a sapphire lens be attached to the crystal surface to confine the radiation to the diffraction limit. The radiated pulse coupled out of the crystal was detectable by a PCA receiver at many centimetres away in the free space.

Free-space electrooptic sampling—1995. As opposed to free-space radiation, free-space EO sampling via the Pockels effect had not been realised until many years after. The technique was reported independently by Wu and Zhang (1995) in the USA and Jepsen *et al.* (1996) in Germany nearly at the same time—one week after another. Wu and Zhang (1995) used a 500-μm-thick $LiTaO_3$ crystal equipped with a silicon lens for detection of freely propagating T-ray pulses emitted from a GaAs PCA transmitter located 10 cm away. Pulses from a mode-locked Ti:sapphire laser probed a refractive index change in the crystal caused by the incident T-ray pulses. The resolvable T-ray pulses had a rise time of 740 fs, followed by multiple reflections over a 100 ps time span. The reflections were developed inside the transmitting PCA and the detecting nonlinear crystal. Jepsen *et al.* (1996) carried out a similar experiment with a 1-mm-thick $LiTaO_3$ detecting crystal, in which the T-ray and optical probe beams copropagated. It was shown that the T-ray pulse obtained from this EO sampling system was consistent with the pulse obtained from photoconductive sampling. In addition, they elucidated the sampling mechanism in the situation that the phase matching condition between the copropagating T-ray and optical pulses is not met. In brief, the T-ray pulse propagating in the crystal results in a change in the optical refractive index, which introduces optical probe pulse phase retardation. Because the optical pulse travels in the crystal faster than the T-ray pulse, the measurable phase retardation of the optical pulse is related to the integral of the T-ray pulse. Both Wu and Zhang (1995) and Jepsen *et al.* (1996) suggested the use of EO sampling for characterisation of the 2D spatial distribution of a T-ray field. This suggestion was later adopted by Wu *et al.* (1996).

2.4 Principles for THz-TDS sources and detectors

Most, if not all, THz-TDS systems employ optical-electronic devices for generation and detection of T-rays. Potential devices for T-ray generation include PCAs, nonlinear crystals, bulk semiconductors, semiconductor quantum structures, etc. Also, PCAs and nonlinear crystals can be used at the detection side. These devices share some characteristics in general, but are different in terms of underlying mechanisms. Essentially, for the generation, the device radiates coherent T-ray pulses in reaction to the illumination of ultrashort pump laser pulses. For the detection, the device must have a subpicosecond linear response to incident T-ray pulses, and this response must be detectable via the application of probe laser pulses. This section presents a brief description of PCAs and nonlinear crystals for T-ray generation and detection.

2.4.1 Photoconductive antennas—photocarrier transport

Characteristics of the antennas

A photoconductive antenna (PCA) is central to T-ray generation or detection via the photocarrier transport process. The antenna can be configured for either the transmitter or receiver, depending on the connecting circuit, i.e., the DC supply or ammeter.

A common version of the antenna is composed of parallel microstrip metallic lines lithographically defined on a photoconductive substrate, as illustrated in Figure 2.7(a). The microstrip metallic lines actually function as coplanar transmission lines and a Hertzian dipole with a small photoconductive gap. In some applications, the same metallic lines also function as a waveguide, along which the generated T-rays can propagate. Typically, a hyperhemispherical lens, made of high-resistivity silicon that is non-absorptive and non-dispersive to T-rays, is placed in optical contact with the substrate as shown in Figure 2.7(b),(c). This lens, in addition to performing beam focusing, enhances coupling of the radiation out of or into the substrate by providing a better index matching.

The photoconductor, particularly at the dipole gap, is where the ultrafast process of photo-induced carrier production takes place. Hence, a semiconductor selected for a photoconductor must have a short carrier lifetime and high carrier mobility (Sakai and Tani 2005). In addition, it must have a high dark resistivity and high breakdown voltage. During the fabrication process, impurities are often added to the semiconductor, inside which they act as recombination centres for free carriers, and thus shorten the carrier lifetime. However, a shorter carrier lifetime, although increasing the speed of response, reduces the semiconductor sensitivity, because most of the carriers recombine before reaching the electrodes (Auston *et al.* 1980b). So far, many high-resistivity semiconductors have been trialled for ultrafast photoconductor. These include amorphous silicon, semi-insulating GaAs, InP, CdTe, ion-implanted SOS (Doany *et al.* 1987), etc. At present, LT-GaAs is perhaps the most popular semiconductor in use, as it prominently meets the requirements described above.

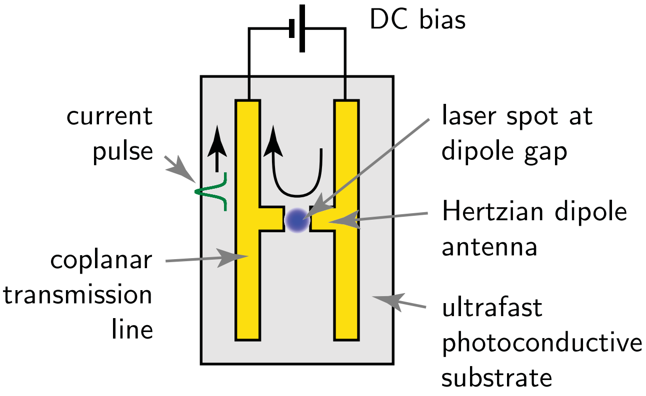
Generation process

During the generation process, the photoconductive region between the dipole is constantly pumped by a mode-locked laser with ultrashort optical pulses that have the level of photon energies higher than the semiconductor energy gap. Collision of an optical pulse at the photoconductor causes a rapid increment of free carriers. Under the influence of the biasing electrical field from the coplanar transmission lines, these photo-induced carriers are accelerated across the dipole gap before recombination, resulting in a transient photocurrent and subsequent radiation of a coherent broadband T-ray pulse from the dipole. Figure 2.7(b) shows the arrangement of a PCA for the transmission operation. In addition, an example T-ray signal and spectrum from a PCA is shown in Figure 2.8.

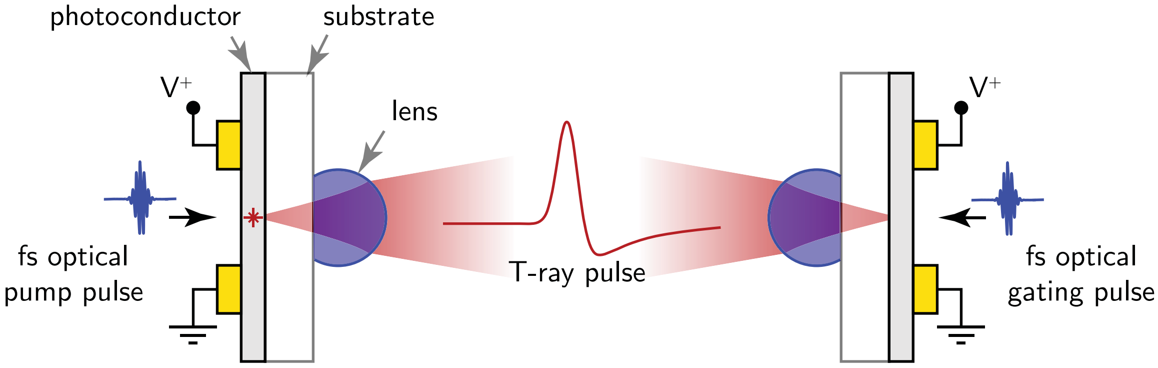
More specifically, the photocurrent density in a PCA, $J(t)$, is related to the convolution between the laser pump intensity, $P_{\mathrm{opt}}(t)$, and the impulse response of the PCA, $qn(t)v(t)$, (Duvillaret *et al.* 2001), or

$$J(t) \quad = \quad P_{\mathrm{opt}}(t) * [qn(t)v(t)] \; , \tag{2.1}$$

where $n(t)$ is the photocarrier density, as a function of the carrier lifetime; $v(t)$ is the photocarrier velocity, related to the biasing electric field; q is the electron charge; and

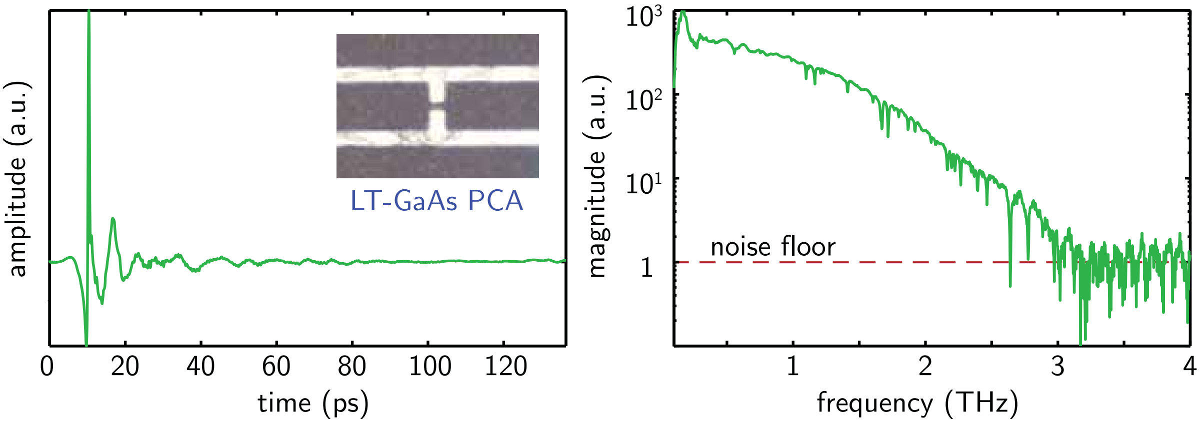


(a) top view



(b) side view, transmitting (c) side view, receiving

Figure 2.7. Photoconductive antenna. The antenna is composed of metallic coplanar transmission lines imprinted on a photoconductive substrate. A hyperhemispherical (truncated sphere) lens is attached to one side of the substrate to couple the radiated T-ray field out of or into the antenna. For the transmitter, the microstrip lines are connected to the biasing electric source. The electric source is substituted by a low-noise current preamplifier, lock-in amplifier, and analyser for the receiver. Modified from Sakai and Tani (2005).



(a) T-ray signal

(b) T-ray spectrum

Figure 2.8. Measured T-ray signal and spectrum from PCAs. The optical pump/probe pulses are supplied by a Ti:sapphire mode-locked laser with a pulse duration of 100 fs and central frequency of 800 nm. The inset shows a photograph of the H-type LT-GaAs PCA used in our lab for generation and detection. The signal has subpicosecond features, corresponding to the bandwith of 3 THz. The sharp dips in the spectrum are due to the absorption by ambient water vapour.

$*$ denotes the convolution operator. Thus, upon laser illumination, a rapid rise in the photocurrent can be observed, followed by a longer decay due to a finite carrier lifetime. This transient current radiates a coherent electric pulse, which, at the far field, is proportional to the first-order derivative of the photocurrent density (Auston *et al.* 1984a, Piao *et al.* 1999, Duvillaret *et al.* 2001),

$$E_{\mathrm{THz}}(t) \propto \frac{dJ(t)}{dt}\,, \tag{2.2}$$

and linearly polarised along the direction of the DC electric field. In this way, a *subpicosecond* transient spanning T-ray frequencies can be generated from a PCA with a *nanosecond* carrier lifetime via illumination of a *femtosecond* laser pulse (Sun *et al.* 1995), as high-frequency components at the rising edge of the photocurrent are enhanced by the far-field differentiation.

In terms of radiation energy, a significant amount of the energy is emitted from the dipole antenna during the rapid rise in the photocurrent, and not during the longer decay (van Exter and Grischkowsky 1990c). In addition, the radiation energy is coupled more into the substrate than into the air (Smith *et al.* 1988)—this explains why a hyperhemispherical lens is attached to the substrate opposite to the dipole antenna to collect the radiation. Due to the reflection losses at the substrate-lens and lens-air interfaces, the ultimate T-ray energy emitted from the transmitter is approximately 15% of the total radiation energy (van Exter and Grischkowsky 1990c).

Detection process with photoconductive sampling

The detection process using a PCA receiver is virtually the reverse of the generation process. As shown in Figure 2.7(c), the photoconductive gap of the antenna is constantly excited by a series of optical probe pulses from a mode-locked laser. In the absence of a T-ray pulse incident on the antenna, free carriers induced by the optical pulses recombine without causing a current flow between coplanar transmission lines. Once a T-ray pulse is focused onto the active region with its polarisation parallel to the dipole, the T-ray electric field sweeps free carriers across the gap to produce the photocurrent. Because free carriers exist for a short time span relative to the T-ray pulse duration, the measurable photocurrent is a sample of the T-ray amplitude at an instant; hence the process is termed photoconductive sampling. By adjusting the synchronisation between the optical probe pulse and T-ray pulse, the entire waveform of the T-ray pulse can be recorded. The equation for the photocurrent density in a PCA receiver is similar to Equation 2.1, but now the photocarrier velocity, $v(t)$, is a function of the incident T-ray field. Appendix B offers the complete analytical models for the PCA generation and detection processes.

Although the carrier lifetime is a factor that determines the antenna speed, the detection bandwidth of the antenna is not necessarily limited by the carrier lifetime (Katzenellenbogen and Grischkowsky 1991, Tani *et al.* 2002). Sun *et al.* (1995) showed that a GaAs photoconductor with a subnanosecond carrier lifetime is able to detect a subpicosecond pulse. Similarly, Kono *et al.* (2000) used a LT-GaAs photoconductor with a carrier lifetime of 1 ps and a mode-locked Ti:sapphire laser with a pulse width of 15 fs to resolve a T-ray signal with the fastest feature of 45 fs, corresponding to the cutoff frequency over 20 THz. The fast responses of PCAs despite much slower carrier lifetimes are attributed to ultrashort optical probe pulses, which cause the fast rising edge of the photocarrier density. Nevertheless, if the carrier lifetime is considerably long such that the carrier density approaches a step function originating from an incident optical pulse, the measured photocurrent is the integration of the T-ray amplitude arriving afterwards (Sun *et al.* 1995). In this case, the PCA receiver responds to the transient in an integration mode, which limits the detectable bandwidth because the temporal integration is equivalent to the multiplication of the spectrum by $1/\omega$.

2.4.2 Nonlinear crystals—nonlinear optical processes

Characteristics of nonlinear crystals

A key component for T-ray generation and detection via nonlinear optical processes is a nonlinear crystal. In practice, this crystal is commonly pumped/probed by ultrashort optical laser pulses to generate/detect T-ray pulses. A suitable nonlinear crystal should have a high transparency in both the optical and T-ray regimes to avoid energy dissipation (Auston and Nuss 1988). In addition, its electrooptic coefficient should be sufficiently large to achieve a high conversion efficiency. A number of materials satisfy these requirements, and these include dielectric crystals, e.g., $LiNbO_3$, $LiTaO_3$, semiconductors, e.g., ZnTe, GaAs, and organic materials, e.g., DAST (Sakai and Tani 2005). A factor that degrades the crystal performance is the optical phonon absorption resulting from lattice vibration, which lies in the T-ray frequency range, for example, ZnTe at 5.31 THz, GaAs at 8 THz, and $LiTaO_3$ at 6.23 THz. Further information about the far-infrared optical constants of some important nonlinear crystals, such as $LiNbO_3$, $LiTaO_3$, ZnTe and CdTe, is reported in Schall *et al.* (1999).

Generation process

Generation of T-ray pulses by using a nonlinear crystal relies on a nonlinear optical process, or more precisely, the inverse EO effect or optical rectification (Bass *et al.* 1962). In this process, an ultrashort optical pulse interacts with the crystal and induces the polarisation of dipoles therein. For a particular case, the polarisation has a time variation following the intensity envelope of the optical pulse—thus the name optical rectification. Subsequently, this low-frequency polarisation radiates a rectified electromagnetic transient, whose frequencies span from DC well into the T-ray regime.

The optical rectification requires second-order frequency mixing. Given that $E(t)$ represents an optical field from a mode-locked laser, the second-order dielectric polarisation inside a nonlinear crystal, excited by $E(t)$, is

$$P^{(2)}(t) \quad \propto \quad \chi^{(2)} E(t)^2 \, , \tag{2.3}$$

where $\chi^{(2)}$ is the second-order susceptibility of the crystal, related to the electrooptic coefficient. By assuming that the optical field comprises two frequency components,

$$E(t) \quad = \quad E_1 e^{-j\omega_1 t} + E_2 e^{-j\omega_2 t} + \text{c.c.} \, , \tag{2.4}$$

the second-order polarisation follows (Boyd 2003)

$$\begin{aligned} P^{(2)}(t) \quad \propto \quad & \chi^{(2)} \Big[E_1^2 e^{-2j\omega_1 t} + E_2^2 e^{-2j\omega_2 t} + 2E_1 E_2 e^{-j(\omega_1+\omega_2)t} \\ & + 2E_1 E_2^* e^{-j(\omega_1-\omega_2)t} + \text{c.c.} \Big] + 2\chi^{(2)} \left[E_1 E_1^* + E_2 E_2^* \right] \, , \end{aligned} \tag{2.5}$$

where c.c. represents the complex conjugate of all terms to its left. The frequency components of the polarisation encompass the second harmonics of E_1 and E_2, sum frequency, difference frequency, and DC. Only the difference-frequency component, $E(\omega_1 - \omega_2)$, is related to the optical rectification. Because an ultrashort optical pulse $E(t)$ contains a great number of frequency components, the optical rectification results in the polarisation oscillating at a wide band of difference frequencies spanning from DC to a few terahertz.

Fundamentally, the radiated T-ray waveform is related to the second-order derivative of the second-order polarisation, or

$$E_{\text{THz}}(t) \quad \propto \quad \frac{d^2 P^{(2)}(t)}{dt^2} \, . \tag{2.6}$$

By considering from Equation 2.3 that the second-order polarisation is proportional to the intensity of the generating laser pulse, the radiated T-ray waveform is thus related to the second-order derivative of the optical pulse envelope. However, it should be noted that a detectable pulse shape does not strictly follow optical rectification theory. Pulse reshaping is introduced by the reflection, absorption, and dispersion in both the generating and detecting crystals (Bakker *et al.* 1998).

In the case of a tightly focused optical beam, which is localised both in time and space, the polarisation is induced locally along the beam path in a nonlinear crystal. Because the velocity of the optical pulse inside the crystal is greater than that of the radiated T-rays, the overall T-ray electric field, as a sum of the local radiation, is described by the cone of a shockwave, analogous to the Čherenkov radiation of a relativistic charged particle travelling through an insulator. Though, in nonlinear optics, the radiating source is a dipole, which has a spatial extent, rather than a localised charged particle (Auston and Nuss 1988). The analysis shows that the pulse amplitude of the cone mainly depends on the optical pulse energy, the electrooptic coefficient of the crystal, and the optical refractive index (Auston *et al.* 1984b). Furthermore, the optical beam waist and optical pulse duration have a significant influence on the amplitude and frequency of the radiated T-rays (Auston *et al.* 1984b).

Note that nonlinear semiconductor crystals, such as GaAs, can produce T-rays via both optical rectification and photocarrier transport simultaneously (Zhang *et al.* 1992). As discussed in Section 2.4.1, if the energy bandgap of a semiconductor is narrower than the photon energy from the optical pump beam, free carriers are generated through optical excitation. A surface electric field in the semiconductor is able to drift these photo-induced carriers to produce a surge current leading to the T-ray radiation.

Detection process with EO sampling

The EO sampling technique is able to resolve the amplitude of a subpicosecond electrical transient via the linear EO effect or Pockels effect. The effect is related to a local change in the optical birefringence of a crystal following application of an electric field or here a T-ray pulse. The change in the birefringence modulates an optical probe pulse in terms of the phase retardation of the two orthogonal components. Analysis of the modulated optical probe pulse reveals the instantaneous field amplitude of the T-ray pulse. By adjusting the delay path of the optical probe beam, the entire T-ray waveform can be sampled. Note that a nonlinear crystal used for EO sampling is sometimes referred to as an EO crystal.

Figure 2.9 shows a common arrangement for a free-space EO sampling system. The polariser aligns the polarisation of the optical beam at an angle to the T-ray beam polarisation. The pellicle reflects the optical beam so that it copropagates with the T-ray beam into the nonlinear crystal. Due to the Pockels effect, the crystal introduces a phase retardation $\Delta\phi$ to the optical beam with the degree of retardation governed by the instantaneous T-ray amplitude E_{THz}, or

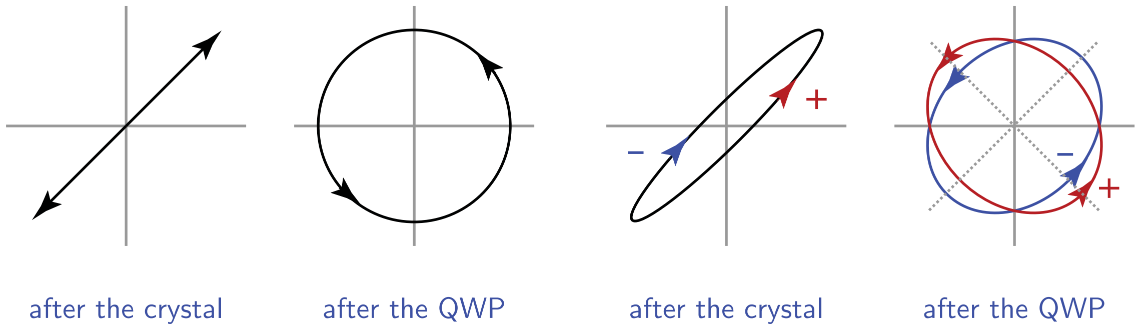
$$\Delta\phi \quad = \quad \frac{\omega d n_{\mathrm{opt}}^3 r_{41} E_{\mathrm{THz}}}{2c} (\cos\alpha \sin 2\varphi + 2\sin\alpha \cos 2\varphi)\ , \tag{2.7}$$

where ω is the optical frequency; d is the crystal thickness in the propagation direction; n_{opt} is the optical refractive index of the isotropic crystal; r_{41} is an electrooptic coefficient of the crystal; α and φ are the polarisation angles of T-ray and optical beams with respect to the $\langle 001 \rangle$ direction, respectively (Gorenflo 2006). In order to maximise the degree of retardation, both α and ϕ are adjusted to 90° (Gorenflo 2006). A thicker sample, although enhancing the degree of phase retardation, accentuates the phase mismatch between the optical and T-ray beams, resulting in a decrease detection bandwidth.

In the absence of the copropagating T-ray beam, the optical beam passes through the crystal without the phase retardation, provided that there is no static birefringence

NOTE:
This figure is included on page 39
of the print copy of the thesis held in
the University of Adelaide Library.

Figure 2.9. Free-space EO sampling system. The system is composed of a polariser, pellicle, nonlinear crystal, quarter-wave plate (QWP), Wollaston prism (WP), and pair of balanced photodiodes. The optical beam path is indicated by small red arrowheads, and the T-ray beam path is indicated by large green arrowheads. After Lu *et al.* (1997).



(a) polarisation in the absence of T-rays

(b) polarisation in the presence of T-rays

Figure 2.10. Polarisation state of the optical beam. (a) In the absence of a T-ray beam, the optical beam polarisation after the crystal remains linear. A quarter-wave plate (QWP) retards a component of the beam by 90°, causing the circular polarisation. (b) In the presence of the T-ray beam, the Pockels effect of the crystal slightly retards a component of the beam, resulting in elliptic polarisation. The rotation direction of the polarisation depends on the polarity of the T-ray amplitude (plus and minus signs). A QWP further retards a component of the beam by 90°. A Wollaston prism (WP) resolves the two orthogonal components with respect to the new axes (dotted lines).

in the crystal. Thus, the linear polarisation of the optical beam remains linear at the exit face of the crystal. The optical beam then enters the QWP with 45° polarisation angle with respect to the QWP's optical axis. As a result, the linearly polarised beam becomes circularly polarised after one component of the beam is retarded in the phase by 90° or a quarter of wavelength with respect to its orthogonal counterpart. The polarisation states before and after the optical beam enters the QWP are illustrated in Figure 2.10(a). The WP splits the two orthogonal components of the optical beam into two separated beams. The balanced photodiodes detect a difference in the intensities of the two orthogonally polarised beams. In this case the two beams have an equal intensity, causing net zero current flow from the photodiode ports.

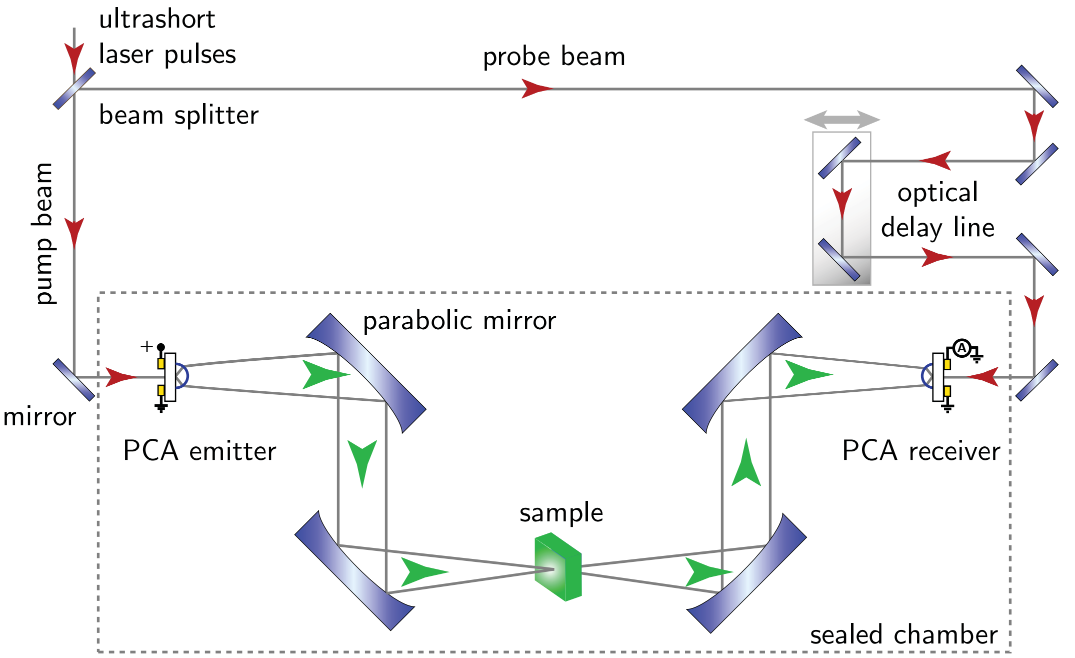
In the presence of the copropagating T-ray beam, the phase of the optical beam is retarded according to Equation 2.7. A small phase retardation changes the optical beam polarisation from linear to strong elliptical with the rotational direction dependent on the T-ray amplitude. The QWP, again, retards the phase of one optical component by 90° with respect to the other. The polarisation now becomes weakly elliptical. The change in the polarisation states is clearly depicted in Figure 2.10(b). The optical bias by the QWP sets the operation point of the system to the highest sensitivity and linear range. The WP separates the two optical components polarising in the major and minor axes of the ellipsoid and feeds them to the balanced photodiodes. The two optical components, which now have different intensities, result in a difference signal from the photodiodes. This difference signal is proportional to the degree of phase retardation and hence the T-ray amplitude. Note that the balanced photodiodes are able to reject correlated noise, and hence increase the sensitivity by twofold. The effect of uncorrelated noise can be reduced via temporal averaging at the expense of acquisition speed.

It is worth noting that the phase mismatch between the optical and T-ray pulses in the crystal is not considered here. In order to maximise the temporal resolution and hence the detection bandwidth, the condition of phase matching must be satisfied. Further discussion on EO sampling can be found in Gorenflo (2006).

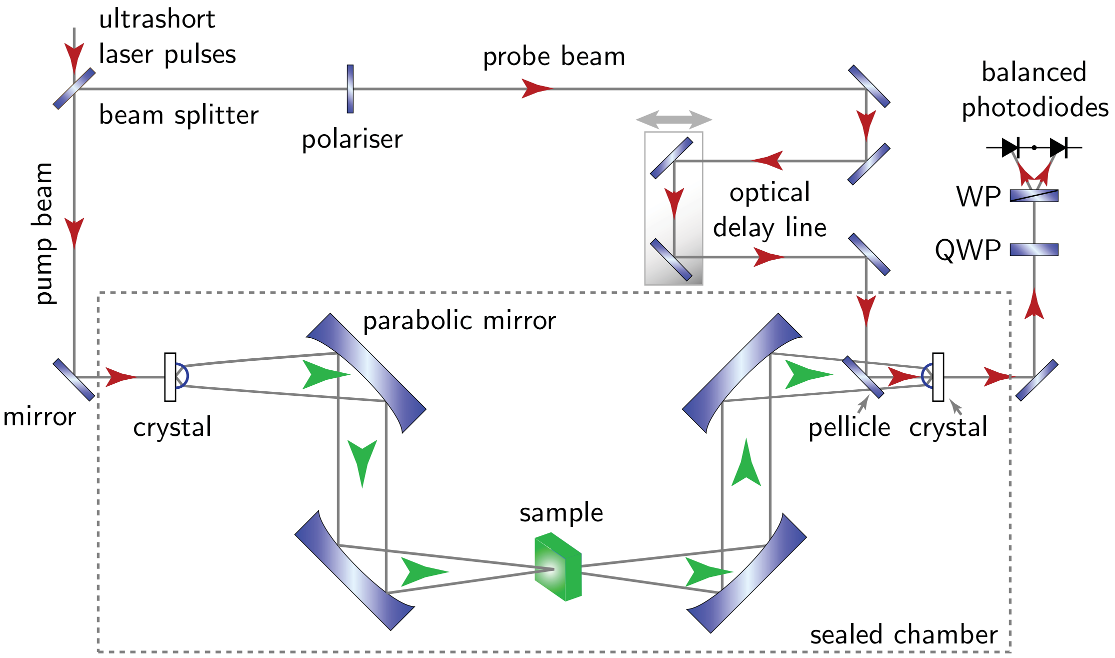
2.5 Transmission-mode THz-TDS

In Section 2.4 the important T-ray sources and detectors have been discussed individually in detail. This section gives the description of full THz-TDS systems in the transmission mode, which is the most prevalent setting due to its simplicity and generality. Figure 2.11 illustrates two transmission-mode systems, employing either PCAs or nonlinear crystals for their operation. There exist differences between these two systems particularly at the detection end, as discussed earlier in Section 2.4. But in general, a complete THz-TDS system comprises a femtosecond optical laser, T-ray emitter/receiver, optical delay line, and guiding optics. The entire T-ray beam path is usually purged with nonreactive atmosphere, e.g., nitrogen or vacuum, to suppress the effects of water vapour absorption. Historically, a desiccant, such as phosphorus pentoxide P_2O_5, was used to remove ambient water vapour from the path (Barnes *et al.* 1935).

In the operation of a THz-TDS system, the optical laser beam is divided by a beam splitter into probe and pump beams. At the T-ray generation side, the optical pump beam is focused by an objective lens onto either a PCA or nonlinear crystal. The respective photocarrier transport or optical rectification process therein is responsible for radiation of T-rays. The diverging T-ray radiation is then collimated and focused onto the sample by guiding optics, and shown in Figure 2.11 are a pair of off-axis parabolic mirrors. After passing through the sample, the T-ray beam is recollimated and focused



(a) THz-TDS with PCAs as emitter and receiver

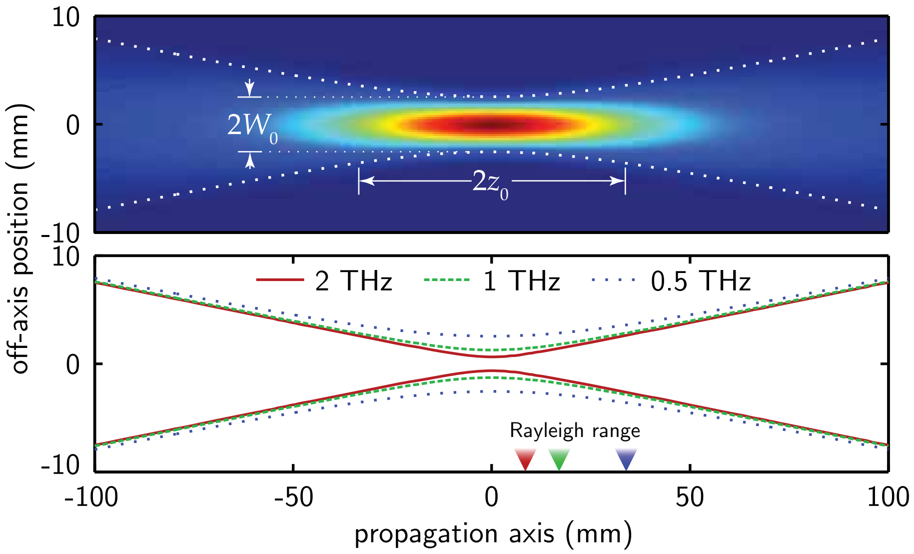


(b) THz-TDS with nonlinear crystals as emitter and receiver

Figure 2.11. THz-TDS systems in transmission mode. Either system generally consists of an ultrafast optical laser, T-ray emitter/receiver, optical delay line, and guiding optics. The optical beam paths are indicated by small red arrowheads, and the T-ray beam paths indicated by large green arrowheads. Typically, a sample under measurement is placed at the T-ray focus. The sample diameter should be greater than the beam waist to avoid edge diffraction.

onto the receiver by an identical set of guiding optics. At the detection side, the probe beam gates the T-ray receiver with a series of ultrashort optical pulses. Similarly to T-ray generation, the detection can be performed via either the photocarrier transport or the Pockels effect. Synchronisation between optical gating pulses and T-ray pulses allows coherent sampling of a T-ray waveform at a time instant. The output current from the receiver is amplified by a low-noise amplifier. A full temporal scan of a T-ray waveform is carried out via the discrete micromotion of a mechanical stage controlling the optical delay line. A complete T-ray waveform is thus represented by a plot of the detector current as a function of delay line position. This waveform, or a time-resolved T-ray pulse, is then readily convertible to its corresponding T-ray spectrum via a Fourier transform.

Spatial distribution of the T-ray beam. The spatio-temporal and spatio-frequency distributions of a T-ray beam are a subject of comprehensive studies, which are usually based on the diffraction integral (wave optics) and the Gaussian approximation (beam optics). For a large aperture transmitting PCA, assuming the Gaussian beam profile of a pump laser, the T-ray beam at far field reproduces a Gaussian shape with an initial beam width equal to the spot size of the pump beam (Gürtler *et al.* 2000). For a lens-coupled PCA, when the generated T-ray beam propagates across the lens-air interface, the effective aperture of the lens causes field diffraction, and the reflection at the interface results in the power transfer of only 21% (Jepsen and Keiding 1995). At a distance from the lens, the forward propagating T-ray beam has a Gaussian intensity distribution, and the beam divergence angle is proportional to the wavelength (Jepsen and Keiding 1995). Focusing of the beam leads to a frequency-dependent spatial distribution, where high frequency components are concentrated along the propagation axis and surrounded by low frequency components, as illustrated in Figure 2.12 (Gürtler *et al.* 2000, Hattori *et al.* 2002, Bitzer *et al.* 2007). In the time domain, beam focusing causes an observable ring shape of intensity at and near the focal plane, because high frequency components at the centre have temporal spans shorter than those of surrounding low frequency components (Rungsawang *et al.* 2003).



(a)
normalised
beam intensity
for 0.5-THz beam

(b)
1/e beam radius for
various frequencies

Figure 2.12. Simulated spatio-frequency distribution of focused T-ray Gaussian beam. A thin lens, or equivalently an off-axis parabolic mirror, with a focal length of $l_{\text{focus}} = 100$ mm, focuses an incident collimated beam with a diameter of $D = 15$ mm to the diffraction limit. The spot size at the focal plane, $2W_0$, equals $4cl_{\text{focus}}/\pi f D$, and the Rayleigh range z_0 equals $\pi f W_0^2/c$ (Saleh and Teich 1991). It is clear that the spot size is a function of the frequency.

2.6 Variants of THz-TDS

Although a transmission-mode THz-TDS system as discussed in Section 2.5 fulfils several general requirements, other variants of THz-TDS have been developed with certain purposes. These configuration variants may attain faster scanning, better sensitivity, higher-dimension measurement, etc. This section covers some THz-TDS variants proposed so far by presenting them in three major categories according to the measurement dimension: spectroscopy, imaging, and tomography.

2.6.1 Spectroscopy

Reflection-mode spectroscopy

For an optically dense sample, a transmission-mode THz-TDS system cannot deliver a promising measurement result, because the transmitted T-ray energy, particularly at high frequencies, is significantly absorbed or scattered by bulk material. A reflection-mode system, as that demonstrated in Figure 2.13, can circumvent the problem. The measured signal for this mode is influenced by the reflectivity at the sample surface, rather than the transmittivity through the sample. According to Fresnel's law, this reflectivity is a function of the complex refractive index of the sample, and thus it is possible to extract this index from the reflected signal. Two main operations of reflection-mode THz-TDS include single-reflection measurement (Jeon and Grischkowsky 1998a) and double-reflection measurement[2] (Thrane *et al.* 1995, Rønne *et al.* 1997, Jepsen *et al.* 2007). More details of these two operations and relevant parameter extraction processes are given in Chapter 4. Attenuated total reflection THz-TDS is another well-known mode of reflection spectroscopy. In this mode, the evanescent waves, caused by total internal reflection inside a prism, interact with the sample, and gives information about the complex optical constant and surface plasmon polaritons of the sample (Hirori *et al.* 2004, Nagai *et al.* 2006).

---

[2] Double-reflection measurement is also known as self referencing, because one measurement contains both the reference and sample signals.

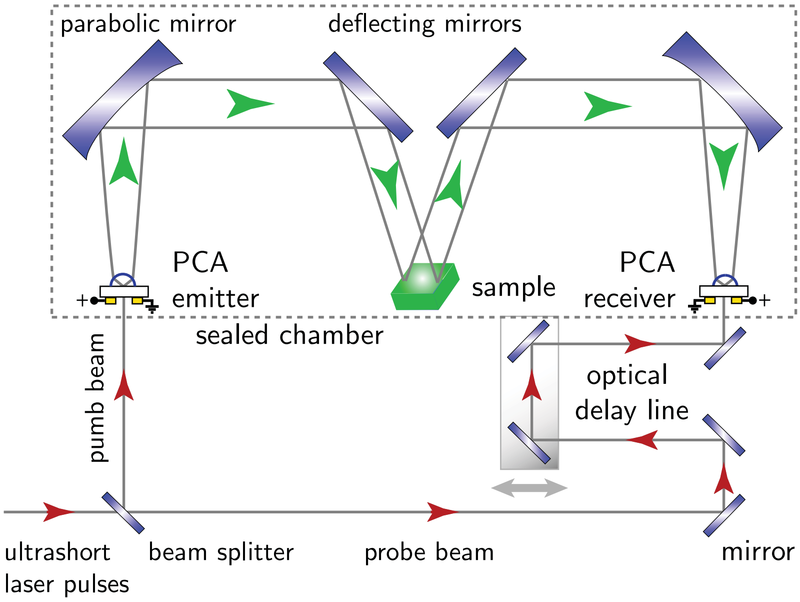


Figure 2.13. THz-TDS system in reflection mode with PCA transmitter/receiver. The optical beam paths are indicated by small red arrowheads, and the T-ray beam paths indicated by large green arrowheads. The T-ray beam is incident and reflected at an inclining angle to the sample surface, resulting in the polarisation and angle dependences of the measured reflected signal in addition to the reflectivity dependence.

Differential time-domain spectroscopy

Typically, noise in the system causes uncertainty in the measurement of reference and sample signals, and hence limits the highest sensitivity achievable from THz-TDS. This limitation greatly affects the characterisation of a thin sample. Differential time-domain spectroscopy or DTDS enhances the sensitivity of the system by mechanically swapping between the sample and reference targets in and out of the T-ray beam at a considerable rate. This technique yields the differential signal of the sample and reference with an exceptional sensitivity. The differential signal can be used to determine the optical constants of the sample. DTDS was initially proposed for thin-film characterisation, where a target film has a thickness of the order of a micrometre or less (Jiang *et al.* 2000, Lee *et al.* 2002, Mickan *et al.* 2002b, Mickan *et al.* 2002a). The high sensitivity of DTDS is also exploited for sensing biomolecular affinity (Mickan *et al.* 2002c, Menikh *et al.* 2004), changes in living cell monolayers (Liu *et al.* 2007a), polar liquids (Balakrishnan *et al.* 2009a), etc.

Waveguide spectroscopy

A parallel-plate metallic waveguide is known to be able to guide T-ray pulses without distortion (Mendis and Grischkowsky 2001a, Mendis and Grischkowsky 2001b). The structure consists of two metallic plates placed in parallel with a submillimetre gap in between, where T-rays are coupled in/out via lenses. This waveguide structure can be employed in T-ray spectroscopy in two configurations. In the first configuration, a submicron sample layer is deposited between the plates. The signal transmitted through the waveguide is thus a function of the sample dielectric constants. By this means, a nanometre-thick layer of water inside the waveguide was successfully characterised (Zhang and Grischkowsky 2004). In another configuration, a plate of an opaque sample substitutes one of the two metallic plates. The sample parameters are inferred from the transmitted signal based on the single-mode waveguide equation. A highly doped silicon wafer was characterised by this method (Mendis 2006). In either configuration, the accuracy of the determined optical constants is greatly influenced by the accuracy of the gap thickness. Metal-wire guiding of T-rays is another approach to waveguide spectroscopy. The wire is able to guide T-rays along its surface with low signal attenuation and dispersion (Wang and Mittleman 2004). A trace of polycrystalline lactose, dispersed on the surface of a platinum/iridium wire waveguide, was detected without interference from ambient water vapour (Walther *et al.* 2005).

Mechanical-free temporal scanning

A mechanical translation stage is a vital component of a conventional THz-TDS system as it enables sampling of different portions of a T-ray temporal profile. However, this moving part presents a limit to the scanning rates of THz-TDS systems. In order to overcome this speed limit, other sampling schemes have been proposed with the main idea of removing the translation stage. These schemes include chirped probe sampling, non-collinear sampling, and asynchronous optical sampling.

Chirped probe sampling can deliver a full temporal T-ray profile in a single optical shot via time-to-wavelength conversion (Jiang and Zhang 1998a, Sun *et al.* 1998, Ferguson *et al.* 2002a). A femtosecond probe pulse is chirped by a pair of gratings so that its frequency components spread over a time span of tens of picoseconds. This chirped probe pulse propagates collinearly with a T-ray pulse inside a non-linear crystal, resulting in a chirped pulse with its frequency spectrum modulated by the T-ray profile. The frequency spectrum of the modulated chirped pulse is then analysed by a dispersion grating and recorded onto an optical detector array. The unmodulated spectrum is subtracted from the modulated one to recover the T-ray profile.

Non-collinear sampling enables single-shot measurement of a T-ray signal via time-to-space conversion (Shan *et al.* 2000). To accomplish this, the optical probe beam must be expanded in its diameter. Non-collinear propagation of the T-ray beam and the expanded probe beam into a crystal detector results in the probe beam spatially modulated by a span of the T-ray amplitude. After the optical beam passes a QWP/WP, the spatial modulation can be detected by an optical camera.

Asynchronous optical sampling or AOS is an appealing approach aimed at eliminating a mechanical translation stage (Yasui *et al.* 2005, Bartels *et al.* 2006, Bartels *et al.* 2007). For this approach, two femtosecond mode-locked lasers with slightly different repetition rates are required. One laser emits a pump beam driving the emitter, and the other emits a probe beam gating the receiver. A slight mismatch of the repetition rates permits a series of probe pulses to scan through a T-ray profile generated by a pump pulse without moving parts. The system can achieve a very fine spectral resolution and rapid scanning rate.

2.6.2 Imaging

Raster-scan imaging

Raster-scan T-ray imaging is an implementation of THz-TDS designed to obtain spectroscopic data over an area of a sample. This mode of measurement is different from normal spectroscopy in that a sample is mounted on an x-y translation stage that is synchronised with the spectroscopic system. The mounted sample experiences repetitive measurement/translation to obtain spectroscopic scans over the entire region of interest. The first raster-scan T-ray system, introduced by Hu and Nuss (1995) and Mittleman *et al.* (1996) of Bell laboratories, attracted much public interest, owing to the see-through or *retection* capability without resorting to hazardous radiation. Ever since, raster-scan T-ray imaging has become the most prevailing T-ray imaging technique, being utilised in a wide range of applications, such as medical diagnostics (Fitzgerald *et al.* 2002, Knobloch *et al.* 2002) and semiconductor characterisation (Mittleman *et al.* 1997a, Herrmann *et al.* 2002a). Several images with different display modes of an object result, because a full time-resolved waveform is available at each scanned pixel (Herrmann *et al.* 2000). Due to the diffraction limit of a focused T-ray beam as discussed in Section 2.5, the spatial resolution of an image obtained from each frequency component is unique (Knobloch *et al.* 2002).

2D electrooptic imaging

Raster-scan imaging is relatively slow, since its operation involves mechanical scanning of an area of interest. As an alternative, an electrooptic crystal detector operated in conjunction with an optical camera enables local spectroscopic measurement over a large area at a time (Wu *et al.* 1996, Miyamaru *et al.* 2004, Usami *et al.* 2005, Rungsawang *et al.* 2005, Zhong *et al.* 2006). In brief, a T-ray beam is enlarged in order to illuminate an area of a sample. This sample T-ray beam then co-propagates with an optical probe beam, which also has its waist expanded accordingly, into a crystal detector. The optical probe beam, now modulated by the T-ray amplitude via the Pockels effect, is then transmitted through a QWP/WP and incident on a CCD or CMOS camera, where the spatial distribution of the optical intensity is registered. By scanning the delay line, the 2D temporal profiles of a T-ray field deduced from the optical intensity are attainable. The acquisition rate can be increased further to real-time by incorporating the system with a mechanical-free temporal sampling technique, such as chirped probe pulse (Jiang and Zhang 1998b, Jiang and Zhang 2000) or non-collinear propagation (Yasui *et al.* 2008). However, in most cases, one dimension of an image must be sacrificed for recording the temporal distribution.

Laser terahertz-emission microscopy

Laser terahertz-emission microscopy or LTEM is a T-ray imaging technique that is specific to investigation of certain samples, as the operation of LTEM requires T-ray generation from the sample being investigated. More specifically, a target sample irradiated by a pump pulse from a mode-locked laser must supply free carriers as a result of the photoconductive effect. A local electric field in the sample causes a photocurrent surge that leads to T-ray radiation, in a similar way to generation of T-rays from a PCA. The radiation, whose amplitude depends on the direction of the local electric field, is then collimated and focused onto a receiver. By scanning the laser pump beam across the sample, the local electric field direction can be mapped onto a 2D image. Kiwa *et al.* (2003) initially proposed LTEM for investigation of integrated circuits. It was demonstrated that the map of the electric field direction can help locate defects in an operating chip. A redesign of the system was able to resolve the spatial resolution down to a few microns, suitable for failure analysis of large-scale integrated circuits (Yamashita *et al.* 2005). LTEM is also employed in the studies of ferroelectric materials and superconductors (Murakami and Tonouchi 2008).

2.6.3 Tomography & other 3D imaging

Tomographic imaging or tomography is a general term referring to any technique that means to nondestructively image the internal of an object via a penetrative radiation. The principle is found in a wide range of imaging modalities, including X-ray computed tomography (X-ray CT), magnetic resonance imaging (MRI), ultrasonic imaging, etc. (Kak and Slaney 2001). The differences among these modalities are fundamentally inherited from the type (acoustic or electromagnetic) and frequency of the exploited radiation, which has unique abilities to penetrate and resolve different objects. T-ray tomography can image the internal structure of dielectrics with submillimetre resolution. The acquisition system for T-ray tomography is slightly different from that used for T-ray spectroscopy and imaging, and further data manipulation is needed. Various types of T-ray tomography are as follows:

Time-of-flight tomography

T-ray tomography was first demonstrated by Mittleman *et al.* (1997b) and Mittleman and Nuss (2000) using the time-of-flight technique in analogy to ultrasonic B-scan imaging (Fatemi and Kak 1980). It reconstructs the internal features of an object from time-resolved reflected pulses. At any dielectric interface underneath, a propagating T-ray pulse is reflected owing to the index mismatch. For a single pulse transmitted into an object, a series of pulses reflected from its internal layers are collected as a signal trace in the time domain. Further data processing of the trace reveals a refractive-index profile of the object in the line of propagation. Full 3D refractive-index data are thus obtained by raster-scanning an object in a single 2D plane, in a very similar way to raster-scan imaging. Due to its simplicity yet efficacy, this THz-TDS mode has been employed to inspect many targets, including space shuttle foam insulation (Zhong *et al.* 2005), pharmaceutical coatings (Fitzgerald *et al.* 2005), aluminium oxide layers (Geltner *et al.* 2002), teeth (Crawley *et al.* 2003b), etc.

Computed tomography

T-ray computed tomography (CT) shares a common principle with X-ray CT. Generally, for a single projection, a CT scanner illuminates the entire object with a penetrative wave and records the wave intensity that is transmitted through or reflected from the object by using a 2D array detector. To gain information sufficient for 3D reconstruction, similar projection measurements are carried out at different angles around the object. Subsequently, a numerical reconstruction technique, such as direct Fourier reconstruction (Stark *et al.* 1981), filtered backprojection (FBP) (Kak 1979) or the arithmetic reconstruction technique (ART) (Gordon *et al.* 1970), estimates the volumetric data of the object from the projection images. The source in present THz-TDS systems is not powerful enough to bathe the entire object of a moderate size with T-rays at one time. Raster scanning for each projection at a number of projections is thus necessary. The scanning time is therefore very prohibitive and imposes a limit to the utility of THz-TDS CT. T-ray CT in the transmission mode was reported in Ferguson *et al.* (2002b) and Ferguson *et al.* (2002c), where acquisition times were managed by using a chirped pulse technique. Furthermore, Yin *et al.* (2007) and Yin *et al.* (2009) were able to save acquisition time via adopting region-of-interest (ROI) T-ray CT. Further study taking into account the diffraction by the target was discussed in Wang *et al.* (2003) and Ferguson *et al.* (2006). It should be noted that T-ray CT is not efficient in terms of data utilisation—unlike time-of-flight tomography, CT discards the phase information available from the THz-TDS measurement.

Fresnel lens tomography

T-ray tomography with a Fresnel lens or binary lens was reported and patented by Wang and Zhang (2003) and Wang and Zhang (2006). A Fresnel lens (Wang *et al.* 2002), which has a frequency-dependent focal length, together with broadband T-ray radiation, enables resolving a series of 2D contrast images at different depths of the target from only one viewing angle. More precisely, the target is illuminated by a T-ray beam, which is subsequently focused by a Fresnel lens onto a 2D detector. Each image layer of the target is relevant to a frequency component of a received broadband pulse. This technique, although lacking quantitative information, can deliver full 3D reconstruction at a considerable speed, but the T-ray beam power limits the target size.

Additional information

Other T-ray tomographic and 3D imaging techniques reported so far are, for instance, holography (Wang and Zhang 2004, Crawley *et al.* 2006), wide-aperture reflection tomography (Pearce *et al.* 2005), and Kirchhoff migration (Dorney *et al.* 2001b, Dorney *et al.* 2002). The latter technique was implemented with tabletop seismic models, where T-rays are used as an alternative to high-frequency acoustic waves (Dorney *et al.* 2003). The idea behind the seismic implementation is that materials with a wide range of properties are available in the T-ray regime (Dorney *et al.* 2003). For further discussion, Zhang (2004) and Wang and Zhang (2004) provide extensive reviews on T-ray tomography, including B-scan tomography, computed tomography, diffraction tomography, binary lens tomography, and holography. Fundamentals of tomography and reconstruction algorithms can be found in the classical work of Kak and Slaney (2001).

2.7 Conclusion

In this chapter the hardware for T-ray generation and detection has been presented in several aspects. The introduction provides the historical viewpoint of T-ray research activities starting from the contributions of Hertz with his spark gap generator and

Rubens with thermal sources at the end of the nineteenth century. Since then a number of techniques have been developed in attempts to bridge the terahertz gap. Nevertheless, the development of T-ray systems is not as mature as for optical systems, and there is room for improved performance. This review is narrowed down to important T-ray systems being developed or in general use, which include QCL/QWP, frequency mixers, FTIR, and THz-TDS.

Emphasis is given to THz-TDS systems, which operate in a broad spectral band and yield full amplitude-phase information. THz-TDS has been advanced considerably, owing to the evolution of coherent electrooptic sources and detectors in the past 40 years. Two major types of the coherent sources and detectors for THz-TDS discussed in this chapter include PCAs and nonlinear crystals. Their operation relies on excitation by ultrashort optical pulses, but the underlying mechanisms of the two are completely different, i.e., free carriers play a major role in PCAs, whilst induced dipoles play a major role in nonlinear crystals. In addition to a typical arrangement of THz-TDS in the transmission mode, many variants of THz-TDS have been proposed and they are suitable for specific tasks. Future THz-TDS developments are likely to revolve around the power, sensitivity, bandwidth, spectral resolution, and compactness.

Having reviewed the principles of THz-TDS hardware, we now turn to several potential applications of THz-TDS in Chapter 3. The applications that harness the capability of THz-TDS hardware, range from material characterisation to medical diagnostics.

Chapter 3

Applications of THz-TDS

TERAHERTZ time-domain spectroscopy (THz-TDS) offers an efficient and unparalleled approach to observation of T-ray interaction with materials. Thus, THz-TDS has become attractive to researchers in a wide variety of disciplines, both in the aspects of fundamental material studies and practical applications. This chapter provides an extensive review on potential applications of THz-TDS, together with a discussion on safety issues.

3.1 Introduction

THz-TDS differs from traditional far-infrared spectroscopy in that the former is able to coherently generate and detect dynamic subpicosecond electric transients with high immunity to background radiation, and thus delivers the broadband response of materials in the T-ray frequency range. Coherent optical-gating detection inherent in THz-TDS systems directly accesses time-domain pulses, and hence reveals both the amplitude (absorption) and phase (dispersion) information at individual frequencies without resorting to the Kramers-Kronig relation. This information or a more material-specific parameter, the frequency-dependent complex refractive index[3], which is derivable from the amplitude and phase, is useful in characterising materials down to microscopic scales.

Many distinctive properties are associated with T-rays. Unlike waves at much higher energies, T-ray radiation is non-ionising (Berry 2003). This non-invasive property of T-rays is of great importance to applications that expose humans to radiation, e.g., medical diagnoses or security screening. A number of materials display unique patterns of absorption lines in response to T-rays, due to rotational and vibrational modes of transitions of molecules (Mittleman *et al.* 1996). With a moderate spectral resolution of THz-TDS, these absorption resonances can clearly be detected and identified (Shen *et al.* 2005, Huang *et al.* 2004, Fischer *et al.* 2005b). Non-polar, dry, and non-metallic materials are transparent to T-ray radiation. This see-through capability of T-rays finds applications in quality and security control (Herrmann *et al.* 2002b, Ferguson *et al.* 2003, Coward and Appleby 2003, Zandonella 2003). Weapons concealed beneath clothing or products contained in plastic packages can be seen by T-ray sensors. Due to their sensitivity to water molecules, T-rays can probe the state of hydration of a given sample.

The potential of THz-TDS enables unparalleled spectroscopic applications in various fields of science and technology. This chapter reviews several promising applications, divided into two major parts: fundamental spectroscopic studies in Section 3.2 and advanced applications in Section 3.3. Fundamental studies focus on the physical interactions of T-rays with materials, whilst advance applications exploit these interactions

---

[3] Or equivalently, the complex dielectric constant in nonmagnetic materials. Please refer to Appendix A for the relation between the complex index of refraction and the complex dielectric constant.

for practical implementations. In general, this review focuses on those T-ray applications utilising the capabilities of THz-TDS.

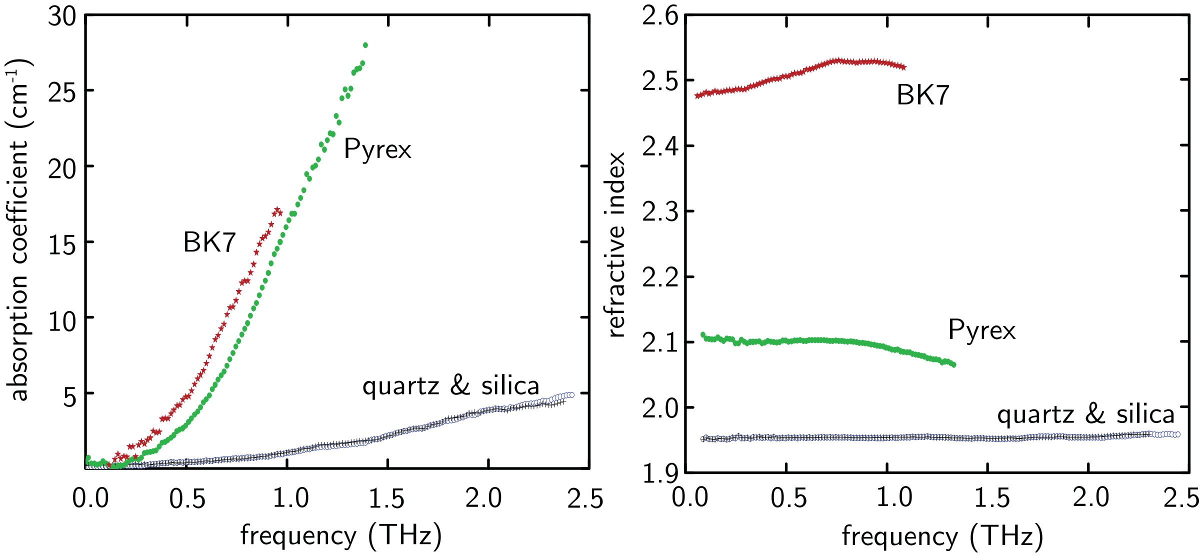
3.2 Fundamental spectroscopic studies

In the early days, ultrafast pulsed T-ray systems were not able to transmit and receive across free space due to technical difficulties. The spectroscopy applications of the systems were thus essentially limited to a few materials, which were relevant to T-ray generation and detection (Cheung and Auston 1986). Spectroscopy of ordinary materials by using free-space THz-TDS was first reported by van Exter *et al.* (1989b). A T-ray system with PCAs was employed to access the nine strongest absorption lines of water vapour in the range between 0.2 and 1.45 THz, with a frequency resolution of 5 GHz. Motivated by a requirement for a suitable T-ray lens material, Grischkowsky *et al.* (1990) reported the measurements of several dielectrics and semiconductors. The following years have seen a significant increase in THz-TDS in spectroscopic studies of a great number of materials, ranging from dielectrics, liquids, gases, semiconductors, to biomolecules.

The T-ray spectral region is where many significant microscopic motions take place within these materials. In general, bulk dielectric relaxations and intermolecular motions occur in this frequency range (Kindt and Schmuttenmaer 1996). Intermolecular hydrogen bonds in crystalline solids resonate with T-rays, causing distinctive absorption features (Fischer *et al.* 2005b). Critical frequencies for Debye relaxation processes in liquids fall into the T-ray regime (Kindt and Schmuttenmaer 1996). Pure rotational transitions can be observed when polar gases are stimulated by T-rays (Mittleman *et al.* 1998). Doped semiconductors have their plasma frequencies and damping rates defined between 0.1 and 2 THz (van Exter and Grischkowsky 1990b). Reports on the T-ray characteristics of these materials are surveyed in this section. An additional review can be found in Hangyo *et al.* (2005).

3.2.1 Regular dielectrics

Many common dielectrics have been measured with THz-TDS for their optical constants. These materials include a series of polymers (Jin *et al.* 2006), glasses (Kojima *et al.*

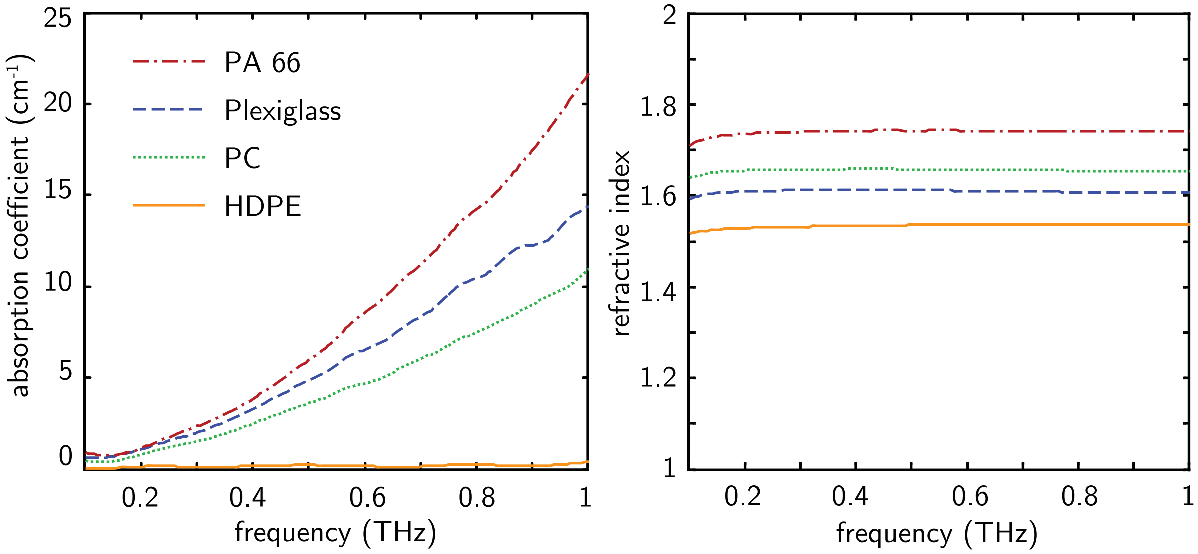


(a) absorption coefficients (b) refractive indices

Figure 3.1. Measured optical constants for some glasses. The optical constants are for polycrystalline quartz, amorphous silica, Pyrex and BK7 glasses. No difference between the optical constants of polycrystalline fused quartz and amorphous silica is observed. After Naftaly and Miles (2005).

2005, Naftaly and Miles 2005, Naftaly and Miles 2007b, Naftaly and Miles 2007a), ceramics (Bolivar *et al.* 2003, Berdel *et al.* 2005), ice crystal (Zhang *et al.* 2001), building materials (Piesiewicz *et al.* 2005, Piesiewicz *et al.* 2007), wood (Reid and Fedosejevs 2006), and many more.

For typical polymers, the refractive indices are constant around 1.4-1.8, and the absorption coefficients increase monotonically (Jin *et al.* 2006), and among those polymers, HDPE, PTFE, and PP show weak absorption, i.e., less than 3 cm^{-1} up to 2.5 THz. Degradation of ultra-high molecular weight polyethylene (UHMWPE) by γ irradiation can be observed through a change in its absorption at the terahertz frequencies (Yamamoto *et al.* 2004b). In the T-ray regime, the optical constants of polymers are affected by hygroscopicity, and thus it is suggested that polymer samples be baked in an oven first before a T-ray measurement (Balakrishnan *et al.* 2009b). Several glasses, including polycrystalline quartz, amorphous silica, Pyrex, BK7, and other glasses, were studied in terms of their optical constants (Naftaly and Miles 2005, Naftaly and Miles 2007b), and the constants for some of them are shown in Figure 3.1. The optical constants for common building materials, i.e., wood, brick, plastics, were measured with



(a) absorption coefficients

(b) refractive indices

Figure 3.2. Measured optical constants for some plastics. The optical constants are for polyamide 66 (PA 66), clear cast acrylic (Plexiglass), polycarbonate (PC), and high-density polyethylene (HDPE). Notice the exceptionally low absorption of HDPE. After Piesiewicz *et al.* (2007).

a purpose to model indoor propagation channels for wireless communication at carrier frequencies above 100 GHz (Piesiewicz *et al.* 2007). As shown in Figure 3.2, the absorption coefficients of some plastics increase with the frequency, owing to either scattering or real absorption. A thorough study of wood reveals that it exhibits birefringence ($\Delta n = 0.07 \pm 0.006$) and diattenuation[4] due to its fibre orientation (Reid and Fedosejevs 2006).

Low absorption and dispersion T-ray materials are required for constructing terahertz components, e.g., lenses, substrates, windows, filters. This requirement has led to an intensive search for materials with such properties. A first batch of dielectrics and semiconductors, including crystalline sapphire, crystalline quartz, fused silica, intrinsic silicon, gallium arsenide, and germanium, was characterised by THz-TDS for finding a low loss material (Grischkowsky *et al.* 1990). It was found that float-zone, high-resistivity silicon has a superior transparency ($\alpha < 0.05\ \mathrm{cm}^{-1}$ at 1 THz) and a constant refractive index beyond 2 THz. This exceptional transparency of float-zone silicon at T-ray frequencies was later confirmed by a well-designed THz-TDS experiment, which

---

[4] Diattenuation describes a difference in power absorption for different wave polarisations in material.

took into account a change in the system transfer function due to sample insertion into the beam path (Dai *et al.* 2004). Furthermore, in this later study a wide measurement bandwidth from 0.5 to 4.5 THz revealed a phonon absorption band at 3.6 THz (Dai *et al.* 2004). The temperature dependence of the refractive index of float-zone silicon is briefly reported in Rønne *et al.* (1997).

A more common material, which is known to have an adequately low absorption and dispersion, is high-density polyethylene (HDPE), which has $\alpha < 2.2\,\mathrm{cm}^{-1}$ and $n \approx 1.53$ at 1 THz. HDPE is currently exploited extensively in THz-TDS applications. However, the quality of HDPE samples relies heavily on the manufacturing process, which in fact varies widely. Another optional material with a low loss in transmission is alumina (Bolivar *et al.* 2003). For this particular ceramic, the refractive index is high, i.e., $n \approx 3.05$, making it very suitable for constructing photonic bandgap (PBG) crystals. Polystyrene foams have a low absorption and low refractive index, i.e., $\alpha < 0.6\ \mathrm{cm}^{-1}$ below 2 THz and $n \approx 1.016$-1.022, resulting from the low mass density of the material (Zhao *et al.* 2002). They are useful for blocking NIR/visible frequencies, while passing T-rays with low loss and low reflection (Zhao *et al.* 2002). Dielectrics that are transparent to UV/visible and T-ray radiation are cyclic olefin copolymer (COC) and picarin (Fischer 2005, Sengupta *et al.* 2006). The dual-band transparency of them aids alignment in pump/probe THz-TDS configurations. The absorption coefficients and refractive indices of these and other materials are detailed in Table 3.1.

3.2.2 Semiconductors

Among numerous kinds of materials, semiconductors have been studied extensively, owing to their unique characteristics in the T-ray frequency band and also the trend of electronics operating towards the sub-terahertz band. The carrier mobility and density in semiconductors are relevant to the plasma frequency and damping rate, which, for moderately doped semiconductors, fall into the T-ray frequency region, i.e., in between 0.1 to 2.0 THz (van Exter and Grischkowsky 1990b, van Exter and Grischkowsky 1990a, Jeon and Grischkowsky 1997). Below 2 THz the absorption of semiconductors is extremely sensitive to the number of free carriers but not to the crystal structure (van Exter and Grischkowsky 1990a). A deposited semiconductor layer with a typical thickness of $\approx$ 5-50 μm causes destructive interference at T-ray wavelengths—this permits access to both the thickness and doping properties of the

Table 3.1. Optical constants of some substances at 1 THz. The values are compiled from various sources as indicated. The refractive indices of these materials are nearly constant over the frequency ranges of measurement, and the absorption coefficients rise monotonically with no spectral feature. The symbols 'o' and 'e' refer to ordinary and extraordinary rays, respectively. Note the values for Bolivar *et al.* (2003) are converted from the complex permittivity and loss tangent. Ice I_h is frozen water in hexagonal crystal form.

Material	Refractive index	Abs. coefficient (cm^{-1})	Reference
Alumina	3.05	3.18	Bolivar *et al.* (2003)
COC	1.52	0.15	Sengupta *et al.* (2006)
Dry air	1.00027	–	Dai *et al.* (2004)
GaAs (undoped)	3.595	0.5	Grischkowsky *et al.* (1990)
HDPE	1.534	2.2	Jin *et al.* (2006)
	1.53	0.5	Piesiewicz *et al.* (2007)
Nitrogen	1.0006	–	Zhao *et al.* (2002)
PET	1.712	20	Jin *et al.* (2006)
PTFE	1.445	1.9	Jin *et al.* (2006)
Picarin	1.52	0.15	Sengupta *et al.* (2006)
Plexiglass	1.61	14	Piesiewicz *et al.* (2007)
PMMA	1.596	20	Jin *et al.* (2006)
Polycarbonate	1.651	10	Jin *et al.* (2006)
	1.66	11	Piesiewicz *et al.* (2007)
Polypropylene	1.510	1.9	Jin *et al.* (2006)
Polystyrene	1.59	2.0	Piesiewicz *et al.* (2007)
Polystyrene foam	1.016-1.022	<0.02	Zhao *et al.* (2002)
Quartz (crystalline)	2.108(o), 2.155(e)	0.1(o), 0.05(e)	Grischkowsky *et al.* (1990)
Quartz (polycrystalline)	1.95	1.0	Naftaly and Miles (2005)
Sapphire (crystalline)	3.07(o), 3.413(e)	1.0(o, e)	Grischkowsky *et al.* (1990)
Silica (amorphous)	1.95	1.0	Naftaly and Miles (2005)
	1.954	2.0	Grischkowsky *et al.* (1990)
Silicon (FZ, undoped)	3.418	0.03	Grischkowsky *et al.* (1990)
	3.415	0.215	Bolivar *et al.* (2003)
	3.4175	0.01	Dai *et al.* (2004)
Steatite	2.57	27	Bolivar *et al.* (2003)
Ice I_h (243 K)	1.793	8.5870	Zhang *et al.* (2001)

layer by using THz-TDS (Hashimshony *et al.* 2001). A conventional four-point probe semiconductor measurement gives a DC resistivity and Hall coefficient, from which the carrier mobility and concentration can be deduced. However, this contact method gives average values over a large area. Via THz-TDS, these quantities can be locally and non-invasively estimated (Herrmann *et al.* 2002a).

As mentioned previously, the first semiconductor characterisation with free-space THz-TDS was carried out by Grischkowsky *et al.* (1990) during a search for T-ray lens materials. Silicon, GaAs and Ge samples, which are intrinsic and crystalline, were studied in the 0.2 to 2.0 THz frequency range. Float-zone silicon shows a high transparency, in contrast to its higher absorption characteristic reported earlier. The transparency of float-zone silicon is attributed to its high purity and high resistivity. GaAs exhibits a monotonic rise in the absorption with frequency due to the resonance at higher frequencies. Germanium (Ge) , which has a considerable number of intrinsic carriers at room temperature, exhibits a carrier-related absorption, explainable by a Drude model. Also, van Exter and Grischkowsky (1990b) extended the investigation into n- and p-type silicon. Deduced from the THz-TDS results via a simple Drude model, in which the carrier damping is independent of the carrier energy, the measured electron and hole mobilities are higher than the literature values, whereas the measured carrier densities are lower than the manufacturer's specification. In addition, van Exter and Grischkowsky (1990a) found a considerable discrepancy between a simple Drude model and the experimental results of cold samples. Therefore, they suggested the application of a scattering model (Kinasewitz and Senitzky 1983), which takes into account the energy dependence of the damping rate. Some studies are summarised in Table 3.2.

Optionally, the study of semiconductor properties can be carried out by means of the terahertz Hall effect. Mittleman *et al.* (1997a) were the first to exploit the Hall effect for determining the carrier density and mobility in semiconductors. In the experiment, a sample was placed in a space between two permanent magnets, which have the flux direction parallel to the T-ray path. Once the T-ray beam impinges the sample, the electric field induces the current in the polarisation direction. Under the presence of the magnetic field, a Hall current polarises orthogonally according to the right hand rule, and radiates another T-ray field into the detectors. A change in the T-ray polarisation is used to determine the carrier density and mobility. Although it was claimed that the terahertz Hall effect aids the parameter calculation, the method appears to be

Table 3.2. Studies of semiconductors using THz-TDS. Drude (DR), Cole-Davidson (CD), and scattering (SC) represent conduction theories, which are used to analyse the measurable spectra. The notation 'i', 'n', and 'p' represents intrinsic, *n*-type, and *p*-type semiconductors, respectively.

Semiconductor	Type	Resistivity (Ω-cm)	Conduction model	Reference
Germanium	n	0.03	DR	Auston and Cheung (1985)
	i	42	DR	Grischkowsky *et al.* (1990)
Gallium arsenide	i	$>$10M	–	Grischkowsky *et al.* (1990)
	n/p	0.19/0.36	DR	Katzenellenbogen and Grischkowsky (1992)
	n	0.038	DR	Jeon and Grischkowsky (1998a)
	n/p	–	DR	Huggard *et al.* (2000)
Gallium nitride	n	–	DR	Zhang *et al.* (2003)
Indium antimonide	i	–	DR	Howells and Schlie (1996)
Indium arsenide	n	–	DR	Hirori *et al.* (2004)
Silicon	i	$>$10k (FZ)	–	Grischkowsky *et al.* (1990)
	n	1.1, 10, 285	DR	Hangyo *et al.* (2002)
	n/p	–	DR	Herrmann *et al.* (2002a)
	n	0.22	DR	Jeon and Grischkowsky (1998a)
	n	0.136	DR	Nashima *et al.* (2001)
	n/p	1	DR	van Exter and Grischkowsky (1990b)
	n/p	various	DR, CD	Jeon and Grischkowsky (1998b)
	n/p	0.1, 1, 10	DR, SC	van Exter and Grischkowsky (1990a)
	n/p	–	DR, CD, SC	Jeon and Grischkowsky (1997)

little used due to the complicated equipment setup and the low sensitivity to *p*-type semiconductors.

3.2.3 Liquids—relaxation processes

For polar liquids, two types of molecular interactions exist around T-ray frequencies. At frequencies lower than ≈6 THz, diffusive motions dominate the spectrum, causing a damped absorption, whilst at higher frequencies, resonant motions phase in (Kindt and Schmuttenmaer 1996). Despite that, most THz-TDS systems have frequency spans that limit the studies to the diffusive regime. In this low-frequency regime, the polarisation relaxation, following the orientation of permanent dipole moments in response to an applied electric field, is responsible for strong absorption (Pedersen and Keiding 1992). The dielectric function influenced by this mechanism can be described by a simple relaxation model, e.g. Debye, Cole-Cole, Cole-Davidson, which are characterised by relaxation time(s). This relaxation time is on a picosecond timescale, and therefore the inversion of a relaxation time—a critical frequency—falls within the T-ray regime (Kindt and Schmuttenmaer 1996, Keiding 1997). In this way, knowledge about the dynamics of a liquid can be inferred from a static THz-TDS measurement. Some polar liquids subject to such study include methanol, ethanol, 1-propanol, ammonia (Kindt and Schmuttenmaer 1996).

Water, a polar liquid of importance, has been studied greatly with THz-TDS. Liquid water has a remarkably high absorption, i.e., $\alpha \approx 200$ cm^{-1} at 1 THz, leading to discrepancies in the earlier measurements that used very low power T-ray sources. The first THz-TDS measurement for the optical constants of water as a function of temperature was reported in Thrane *et al.* (1995) in the frequency range between 50 GHz and 1 THz. Kindt and Schmuttenmaer (1996) proposed that two relaxation processes occur in liquid water in the frequency range below 1 THz, where a double Debye model nicely reproduces the measurement. However, they speculated that there are higher-frequency processes contributing to intermolecular relaxation. Later on, a similar study with a spectral coverage to 2 THz by Rønne *et al.* (1997) confirmed the parameters of a double Debye model for a room-temperature measurement reported in (Kindt and Schmuttenmaer 1996). In addition, it was found that both the absorption coefficient and the refractive index are proportional to the temperature. In Rønne *et al.* (1997), molecular dynamics simulations along with temperature dependence analyses were provided. Further understanding of the relaxation processes and their relevance to the liquid structure was acquired by investigating an isotope of water, D_2O, in comparison to H_2O at a range of temperature (Rønne *et al.* 1999, Rønne and Keiding 2002). Recently, from an analysis of the earlier measurements, including THz-TDS data provided by Rønne, Beneduci (2008) suggested that three relaxation processes occur in liquid water, rather than two, and that the resonance process does not contribute to the response lower than 2.5 THz.

A similar relaxation process can be found in non-polar liquids as a consequence of the orientation of transient dipole moments induced via collisions between molecules. In this case, a much weaker absorption is observed. In addition to a series of Debye relaxation models, some other models are used to explain the response, e.g., Langevin theory (Davies and Evans 1976). Note that THz-TDS studies for non-polar liquids have been performed with, for example, benzene, carbon tetrachloride, cyclohexane (Pedersen and Keiding 1992), carbon disulfide (Yu *et al.* 2003), and a group of substituted benzenes (Keiding 1997).

In addition to the THz-TDS studies of pure liquids, the studies of their mixtures have been carried out for various purposes. There is significant interest in the influence of mixing on the characteristics of liquids (Venables *et al.* 2000a). Flanders *et al.* (1996) studied the collisional interaction between polar and non-polar molecules, $CHCl_3$ and CCl_4, which basically induces transient dipole moments, leading to absorption. Kitahara *et al.* (2005) evaluated changes in the dielectric constants and relaxation times, following the molar variation of ethanol in water-ethanol mixtures. Gorenflo *et al.* (2006) reported that the complex refractive indices of synthetic oils are linearly dependent on the concentration of dissolved water for a small amount of water, and that hydrogen bonds are formed among oil/water molecules. Venables *et al.* (2000a) and Venables *et al.* (2000b) studied the far-infrared absorption spectra and molecular dynamics of acetone/methanol, acetonitrile/methanol, and acetone/acetonitrile mixtures. These mixtures exhibit the linear dependence of far-infrared spectra on the concentration of solvent/solute, and the linearity suggests small changes in the dynamics of these species upon dilution. The intra- and inter-molecular resonances were also observed at higher frequencies, i.e., in the mid-infrared regime.

3.2.4 Gases—rotational transitions

Polar gases resonate at discrete frequencies in response to stimulating radiation spanning from the microwave to infrared. The sharp resonances are caused by molecular rotational transitions, which are characteristic to the structure of molecules. Heavier gas molecules tend to have rotational resonances at longer wavelengths, whilst lighter gas molecules have those resonances at shorter wavelengths. The resonances of gaseous molecules can be broadened by a number of factors, among which the interesting factors at atmospheric pressure are self- and foreign-collision broadening. Rotational gas spectroscopy thus supports the study of molecular structures and dynamics, as well as enables the applications to gas sensing and recognition. Many decades before the advent of THz-TDS, rotational gas spectroscopy had long been performed in the microwave (Townes and Schawlow 1955), far-, mid-, and near-infrared regimes. However, THz-TDS offers additional features to rotational spectroscopy in the far-infrared regime. Time-resolved measurement provided by THz-TDS allows observation of the coherent transients or free induction decay of gaseous molecules following an impulse excitation. Furthermore, time-gating coherent detection can suppress thermal background radiation originating from high-temperature gas samples (Cheville and Grischkowsky 1995).

So far a number of gaseous molecules have been studied by THz-TDS for their transition frequencies and broadening effects. These molecules are, e.g., nitrous oxide: N_2O (Harde *et al.* 1991, Harde and Grischkowsky 1991), methyl chloride: CH_3Cl (Harde *et al.* 1994, Harde *et al.* 1995), methyl halides[5] (Harde *et al.* 1997b, Harde *et al.* 1997a), ammonia: NH_3 (Harde *et al.* 2001), methanol: CH_3OH (Yu *et al.* 2005). Most of the studied gas species exhibit periodical transient emissions, so-called *commensurate echoes*, after being excited by a T-ray pulse. The studies were typically accompanied by collisional-broadening theories that were used to explain the width of the spectral features.

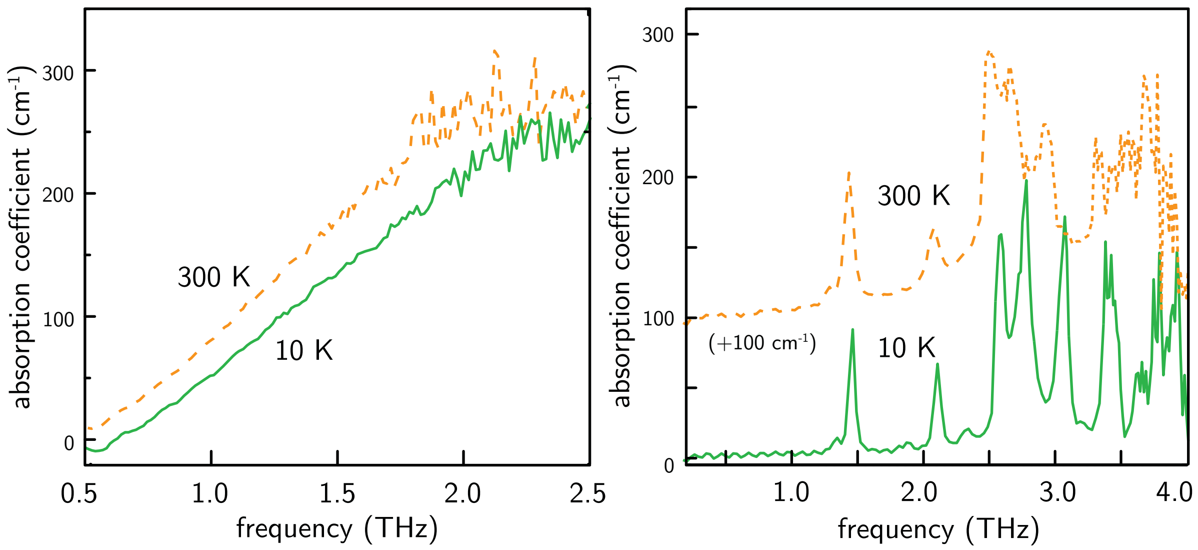
Water molecules have been studied greatly in the vapour phase with THz-TDS. During the early free-space THz-TDS measurements, water vapour was characterised under ambient temperature and pressure (van Exter *et al.* 1989b). Rotational resonances, along with line broadening, of high-temperature water vapour were observed (Cheville and Grischkowsky 1998, Cheville and Grischkowsky 1999). The response of water vapour was measured over a wide range of humidities (Yuan *et al.* 2003). For a theoretical treatment of water vapour, see Chapter 5.

3.2.5 Crystalline materials—vibrational transitions

T-rays can stimulate intra- and/or inter-molecular vibrational modes of transitions in solids, depending on the interaction forces of chemical bonds. The intramolecular vibrational modes are specific to molecules, whilst the intermolecular vibrational modes are sensitive to both the molecular configuration and overall conformation. Although a large number of intramolecular modes are predicted for a molecule in the

---

[5] The studied methyl halides include methyl fluoride: CH_3F, methyl chloride: CH_3Cl, and methyl bromide: CH_3Br.

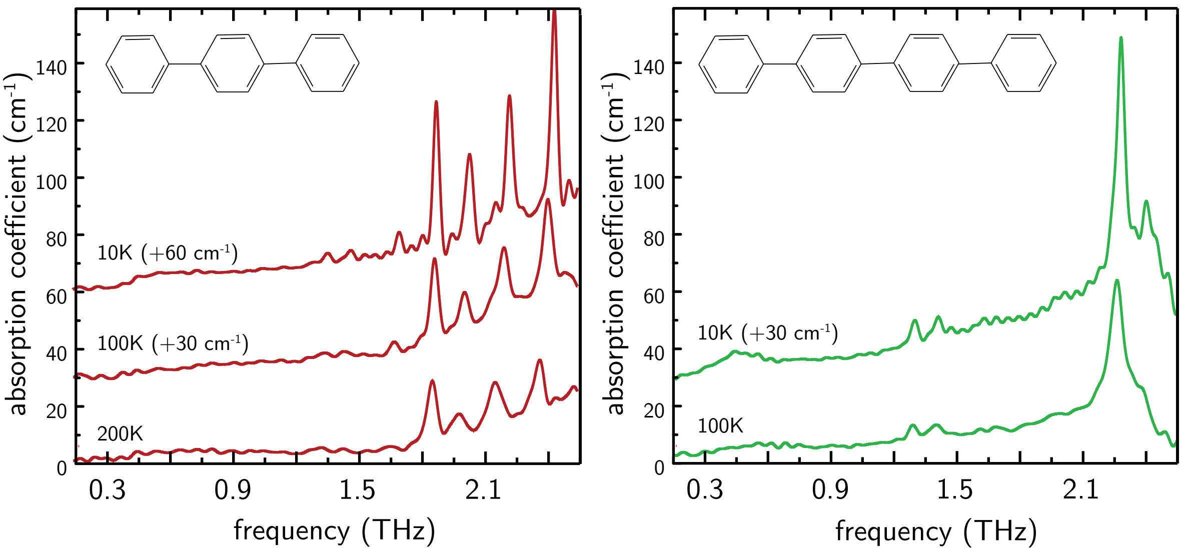


(a) amorphous glucose (b) polycrystalline glucose

Figure 3.3. Measured absorption spectra for amorphous and polycrystalline glucose. Distinct absorption resonances can be observed in the spectrum of polycrystalline glucose, but not in that of its amorphous counterpart. It can be deduced that these resonances are associated with the intermolecular motions. The resonances can be associated with the intramolecular interactions if and only if the resonances are present in the spectra for both the amorphous and crystalline phases. The vertical offset in (b) is intentionally inserted for clarity. After Walther *et al.* (2003).

solid phase, the modes are typically not observable in the T-ray spectrum, due to either their relatively strong interaction forces or broadening/damping effects. For intermolecular modes, their appearance in the T-ray frequency range partly depends on the crystallinity of a substance. Whereas in amorphous materials, these intermolecular modes are damped by structural irregularities, crystalline and polycrystalline structures can support the collective modes because of a long-range order of molecules (Fischer *et al.* 2005b). A comparison of spectra between amorphous and crystalline phases of the same substance are depicted in Figure 3.3. Often, the modes observed via THz-TDS are attributed to intermolecular hydrogen-bonding networks in the crystalline structures, as the energy of hydrogen bonds is comparable to the T-ray photon energy.

Crystalline or polycrystalline substances revealing vibrational resonances to THz-TDS, either in ambient or cryogenic temperatures, are, e.g., aspirin, benzoic acid and its derivatives (Walther *et al.* 2002), lactose, sucrose, cocaine, morphine, aspirin, tartaric acid (Fischer *et al.* 2005b), monosaccharides and disaccharides (Walther *et al.* 2003,



(a) *para*-terphenyl (b) *para*-quaterphenyl

Figure 3.4. Measured absorption spectra for polycrystalline phenylene oligomers. The resonances are assigned to either intra- or inter-molecular vibrations. The redshift of some resonances upon the temperature increment is a result of the coupling of the intramolecular modes to the lattice response. The vertical offsets are intentionally inserted for clarity. After Johnston *et al.* (2003).

Upadhya *et al.* 2004), acephate (Zhang *et al.* 2008c), phenylene oligomers (Johnston *et al.* 2003). Temperature-dependent spectra for phenylene oligomers are shown in Figure 3.4. Though, no resonance is observed in single crystals of ice Ih (hexagonal form) in the 0.25 to 1.0 THz range (Zhang *et al.* 2001), despite the ubiquitous hydrogen bonds constituting the crystal.

Theoretical calculations for vibrational modes based on a single molecule do not reflect experimental results in this frequency region, as the low-frequency modes involve the intermolecular force field and coupling between inter-/intra-molecular vibrations in the crystalline structure (Zhang *et al.* 2008c). Two alternatives are available for modelling these effects (Kleine-Ostmann *et al.* 2008): a calculation using periodic boundary conditions for a uniform and long-range system of molecules, and a cluster calculation by adding more molecules to a monomer. The latter also allows discrimination between intra- and inter-molecular motions.

As mentioned earlier, the spectral features at T-ray frequencies are usually dominated by intermolecular vibrational modes, which are sensitive the crystalline structure. Thus, THz-TDS is ideal for study and identification of polymorphism—the possibility of an identical substance to appear in different crystalline structures. A range of polymorphs have been studied with THz-TDS, for example, ranitidine hydrochloride (Taday *et al.* 2003b), carbamazepine and enalapril maleate (Strachan *et al.* 2004), α-glycine and γ-glycine (Shi and Wang 2005).

A molecule with the same chemical formula might have different isomers, i.e. different atomic arrangements therein. The presence of isomers leads to different crystalline structures, which can produce different features in the T-ray spectrum through vibrational modes. Moreover, THz-TDS studies of isomers have been carried out with, e.g., 2-, 3-, and 4-acetamidophenol (Taday 2004), 2-, 3-, and 4-hydroxybenzoic acid (Walther *et al.* 2002), all-*trans*, 9-*cis*, and 13-*cis* retinal (Walther *et al.* 2000)—note that the resonances for retinal isomers are from intramolecular torsional modes. Likewise, enantiomers of a chiral molecule and their racemic mixtures also produce largely different spectra because of their diverse crystalline structures. Furthermore, THz-TDS sensing between enantiomers and racemic compounds have been carried out with, e.g., polycrystalline L-, D-, and DL-alanine (Yamaguchi *et al.* 2005), polycrystalline D- and DL-cystine (Franz *et al.* 2006), polycrystalline L-, D-, and DL-tartaric acid (Nishikiori *et al.* 2008).

Polymers may have pronounced vibrational resonances in the T-ray frequency region, as their structures are composed of long-chained constituents. However, although long chains in polymers can support the modes, the chains are typically folded, twisted, and bound easily to their neighbours, rather than packed in an orderly manner. When disorder occurs, polymers become highly amorphous, resulting in the suppression of these modes. Particular polymers that have well-organised structures and exhibit clearly distinctive modes at T-ray frequencies are, for instance, UHMWPE with a lattice mode around 2.2 to 2.4 THz and PTFE with many lattice modes below 2 THz (Fischer 2005).

3.2.6 Biomolecules

Naturally, biomolecules, such as polysaccharides, proteins, or DNA, exist in the form of biopolymers, composed of a number of building blocks or monomers that are concatenated via particular covalent bonds. Table 3.3 shows a list of biopolymers and their respective monomers, which have been characterised by THz-TDS with a main interest in their structures. All listed monomers were investigated in the crystalline form, and thus their spectra are dominated by clear absorption peaks arising from intermolecular hydrogen bonds. The spectra for nucleobases and nucleosides, which are the monomers constituting DNAs, are given in Figure 3.5. Correlations between nucleobases and nucleosides, i.e., nucleobases bonded to either ribose or deoxyribose rings, can be observed in the spectral resonances (Fischer *et al.* 2002, Shen *et al.* 2004).

Large biomolecules tend to have featureless spectra in between 0 and 6 THz, due to the high density of vibrational modes and/or inhomogeneous resonance broadening (Zhang *et al.* 2004, Markelz 2008). Despite that, some biopolymers that have long-range molecular order are expected to support discrete resonance frequencies. A group of biopolymers, which show distinct vibrational resonances in T-rays, is for example, cellulose, chitin, and oligonucleotide at cryogenic temperatures (Fischer *et al.* 2005b), and polyglycine (Yamamoto *et al.* 2005b). An observable resonance for polyglycine is attributed to an interchain mode, whilst resonances for cellulose and chitin are from phonon-like modes along the polymer backbone.

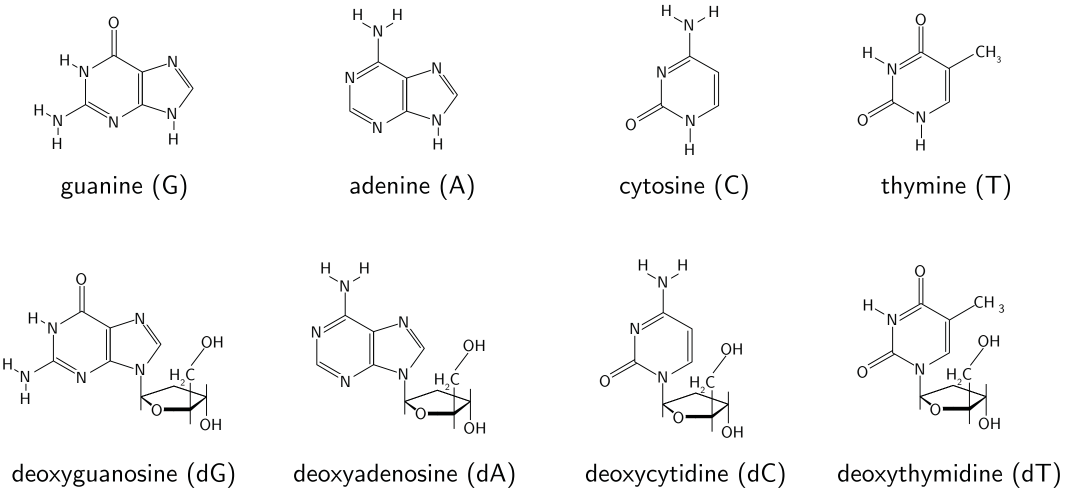
In addition to static spectroscopy, T-rays can perform functional spectroscopy of biomolecules by probing for hydration level, temperature, conformation, binding state, oxidation changes, etc (Markelz 2008). Brucherseifer *et al.* (2000) and Bolivar *et al.* (2002) showed that the optical constants of DNAs are strongly dependent on the DNA binding state, i.e., single stranded or double stranded. Woolard *et al.* (1997) proposed the possibility of detecting DNA mutagenesis by observing a change in specific vibrational modes. Markelz *et al.* (2007) found that the protein dynamic transition at 200 K is observable via the complex dielectric permittivity in the case of cytochrome *c* solutions. Markelz *et al.* (2000) studied a coupling between the T-ray absorption and hydration level of calf thymus DNA and bovine serum albumin, and so did Zhang *et al.* (2004) with myoglobin. Ebbinghaus *et al.* (2008) studied the relation between the T-ray response of proteins and the pH level.

3.3 Advanced applications

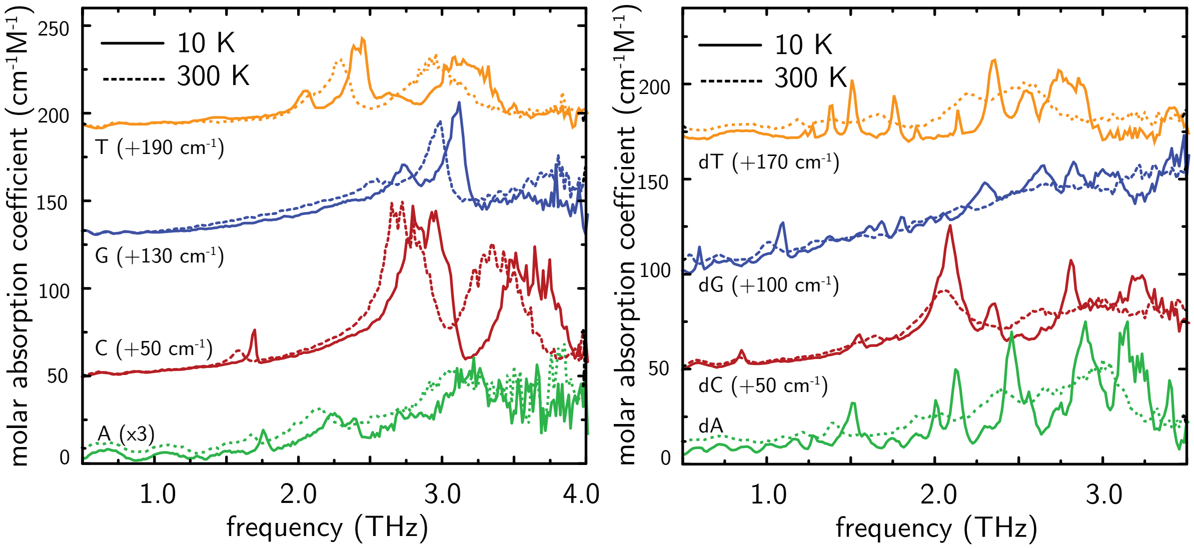
Advanced applications for THz-TDS take advantage of exceptional features from both the system and the radiation in combination. The striking features of the system include the noise immunity and coherent, broadband measurement capability. T-rays are of interest for many reasons including their low-photon energy, ability to penetrate

Table 3.3. Monomers and polymers of biomolecules studied by THz-TDS. All monomers are in the crystalline or polycrystalline form, and thus their spectra are dominated by intermolecular vibrations mediated by hydrogen bonds. Most polymers are typically in the amorphous phase, and the modes are damped. The asterisks indicate the absence of clear resonances. The horizontal lines group relevant monomers and polymers.

Monomer	Polymer
Monosaccharide	**Polysaccharide**
glucose (Walther *et al.* 2003, Upadhya *et al.* 2004)	cellulose, chitin (Fischer *et al.* 2005b)
fructose (Walther *et al.* 2003)	
mannose, galactose, fructose (Upadhya *et al.* 2004)	
Disaccharide	
sucrose (Walther *et al.* 2003)	
maltose, β-lactose (Upadhya *et al.* 2004)	
α-lactose (Fischer *et al.* 2005b)	
Nucleobase	
adenine, guanine, cytosine, thymine (Fischer *et al.* 2002)	
adenine (Shen *et al.* 2003)	
Nucleoside (deoxyribose ring)	**Deoxyribonucleic acid (DNA)**
deoxyadenosine, deoxyguanosine, deoxycytidine, deoxythymidine (Fischer *et al.* 2002)	calf thymus DNA* (Markelz *et al.* 2000) dA-dT-dA-dT-dA (Fischer *et al.* 2005b)
Nucleoside (ribose ring)	**Ribonucleic acid (RNA)**
adenosine, cytidine, thymidine, guanosine (Shen *et al.* 2004)	polyadenylic acid*, polycytidylic acid* (single strand) (Fischer *et al.* 2005c)
Amino acid	**Polypeptide**
glycine (Shi and Wang 2005)	polyglycine (Yamamoto *et al.* 2005b)
L-, D-, and DL-alanine (Yamaguchi *et al.* 2005)	poly-L-alanine* (Yamamoto *et al.* 2005b)
L-glutamic acid (Taday *et al.* 2003a)	
L-serine, L-cysteine (Korter *et al.* 2006)	
L,L-cystine, L-cysteine, L-methionine (Yamamoto *et al.* 2005a)	
L-Glu, L-Cys, L-His (Ueno *et al.* 2006, Rungsawang *et al.* 2006)	
	Protein
	type I collagen*, bovine serum albumin* (Markelz *et al.* 2000)
	myoglobin* (Zhang *et al.* 2004)



(a) molecular structures of nucleobases (first row) and nucleosides (second row)



(b) nucleobases A, C, G, and T

(c) nucleosides dA, dC, dG, and dT

Figure 3.5. Molecular structures and measured spectra for nucleobases and nucleosides. The distinct resonances for nucleobases are from intermolecular motions mediated by hydrogen bonds in the crystalline structure. The broad and intense resonances for nucleosides loosely replicate those of nucleobases, whilst the additional narrow resonances are from vibrational modes attributed to the deoxyribose rings. Vertical shifting is for clarity. After Fischer *et al.* (2002).

dry non-metallic materials, and their ability to probe molecular-level interactions—the latter is pinpointed in Section 3.2. A combination of these factors result in a multitude of advance applications unachievable by conventional technologies. This section provides a survey of mature and emerging applications for THz-TDS in several categories: gas recognition, quality control of pharmaceuticals, medical and dental diagnoses, biosensing, security screening, and miscellaneous areas of non-destructive testing.

3.3.1 Gas sensing and recognition

As mentioned in Section 3.2.4, the absorption spectra for polar gases possess unique patterns of sharp resonances spreading from the microwave to infrared, owing to the rotational modes of transitions. The fingerprint-like spectra allow the recognition of polar gases via either microwave, T-ray, or infrared spectroscopy. In the mid-infrared, both polar and nonpolar gases exhibit additional broad resonances as a result of vibrational transitions. The additional resonances are beneficial, unless the overlap of resonances causes the difficulty in recognition of compositions in a gas mixture (Mittleman *et al.* 1998). In the microwave region, the number of rotational modes becomes relatively low, resulting in a limited number of classification features. Moreover, the interaction between electromagnetic wave and polar molecules in the microwave region is three to six orders of magnitude weaker than that in the terahertz region (De Lucia 2003). Furthermore, microwave and infrared spectrometers are usually operated in a relatively narrow frequency band (Gopalsami *et al.* 1996, Leontakianakos 1992), and thus provide limited information. Switching to THz-TDS for gas recognition could settle many issues.

Moreover, THz-TDS was used to observe the spectral response of a propane-air flame at atmospheric pressure in the frequency range between 0.2 and 2.65 THz (Cheville and Grischkowsky 1995). From the spectrum of the flame it was possible to identify some combustion products, including H_2O, CH, and NH_3, which exhibit absorption resonances at this frequency range. However, the sources of some resonances were unidentified due to the scarcity of available published data. Further investigation of the absorption strength enabled estimation of the flame temperature and the concentrations of the combustion products. In another experiment, THz-TDS was employed in sensing of cigarette smoke in the frequency range between 0.3 and 1.2 THz

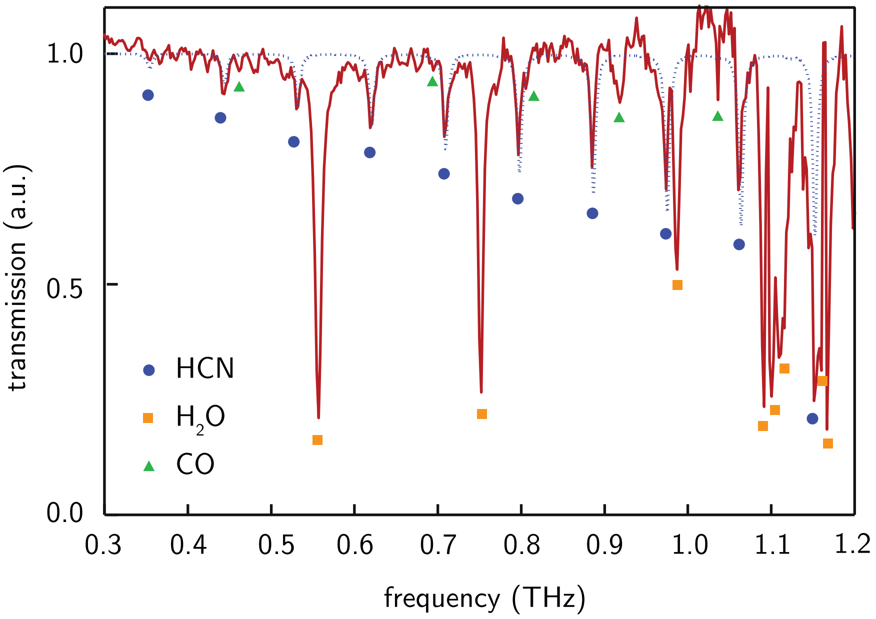


Figure 3.6. Measured spectrum for cigarette smoke at pressure of 950 hPa. The THz-TDS transmission spectrum in the red solid line reveals three major gas compositions as the products of the combustion process, including HCN, H_2O, and CO. The blue dotted line represents the theoretical spectrum of HCN. Dots, squares, and triangles mark the locations of the rotational resonances of HCN, H_2O, and CO, respectively. After Bigourd *et al.* (2007).

(Bigourd *et al.* 2007). The gas cell in use had a realistic pressure of 950 hPa and a propagation length of 580 mm. The measurement showed overlapping resonances caused by different gases in the smoke, which were unambiguously identified to be HCN, H_2O, and CO from the combustion process. Figure 3.6 depicts the spectrum for cigarette smoke, taken from this experiment. It was noted that the system failed to detect the presence of formaldehyde (H_2CO) and ammonia (NH_3). The sensitivity of a system can be further augmented by incorporating a White cell[6], inside which the wave propagation length is greatly extended through multiple reflections, for accommodating a gas sample. A THz-TDS system with a White cell configuration extending the propagation length to 5 m can sense a small fraction of gas at a pressure as low as 1 Pa (Harmon and Cheville 2004).

A procedure for automated gas recognition based on the time-domain analysis of THz-TDS waveforms has been proposed (Jacobsen *et al.* 1996, Mittleman *et al.* 1998). First,

---

[6] A White cell, named after J. U. White (White 1942), is a small gas chamber incorporating windows and mirrors. The windows enable spectroscopic measurement of a gas sample therein, whilst the mirrors facing each other extend the optical path length to several metres.

a codebook containing a series of vectors specific to a gas species is constructed from measurement data using a linear predictive coding (LPC) analysis. The LPC vector of the signal measuring an unknown gas is then compared against those vectors in the codebook to find its best match. A gas mixture is recognised and quantified for its compositions in this way, as the vector for a mixture is a geometrical summation of the vectors for individual gases. Recognition and quantification of a pure HCl and a NH_3/H_2O mixture have been demonstrated.

A technique for gas recognition via THz-TDS requires improvements in both hardware and software before it becomes competitive to the existing chemical analysis methods. A THz-TDS system needs a higher sensitivity and finer spectral resolution in order to sense a low concentration of polar gases. In addition, signal processing must be tolerable to measurement uncertainty to avoid false recognition. By addressing these issues, THz-TDS would be a promising option to *in situ* polar gas recognition.

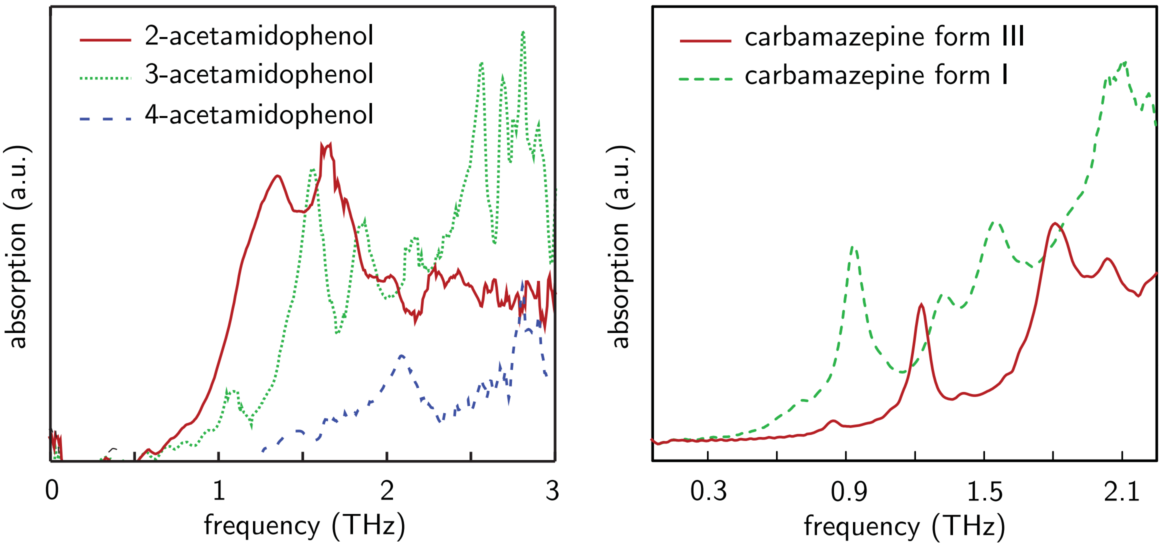
3.3.2 Pharmaceutical assessment

Pharmaceutical science gains limited benefit from existing optical assessment methodologies, either Raman or NIR spectroscopy. Both techniques, although proved to be sensitive to pharmaceutical active ingredients, have notable drawbacks, i.e., Raman spectroscopy induces chemical changes in the target through photochemical reactions, and NIR spectroscopy provides complicated spectra due to a combination of inter- and intra-molecular vibrational modes of transitions (Taday *et al.* 2003b, Taday 2004). As opposed to those conventional techniques, THz-TDS, with its relatively low output energy, is less likely to introduce any chemical or structural change in medicines (Taday 2004). Furthermore, T-ray radiation couples to intermolecular vibrations, resulting in a clear spectrum relevant to the structural crystallinity.

The sensitivity to the molecule and crystallinity allows the identification of active ingredients and their isomerism and polymorphism in tablets via the T-ray spectrum. The promise is further accentuated by the fact that amorphous substances do not compete with crystalline active ingredients in the spectrum. Moreover, THz-TDS has been proved to be successful in providing distinguishable spectra of many medicines, for example, three isomers of paracetamol (2-, 3-, and 4-acetamidophenol) (Taday 2004), two polymorphs of ranitidine hydrochloride (Taday *et al.* 2003b), polymorphs of carbamazepine and enalapril maleate (Strachan *et al.* 2004), polymorphs of theophylline (Upadhya *et al.* 2006). The example spectra are shown in Figure 3.7.

The T-ray spectrum, which is unique to the molecular/crystalline structure and dependent on concentration, aids in the exploration of tablet purity through the chemometric analysis. The technique successfully quantified paracetamol (4-acetamidophenol) and aspirin (acetylsalicylic acid), each mixed with excipients (lactose and cellulose) at different concentrations (Taday 2004). A binary mixture of an active ingredient in two different solid states was quantifiable in a similar fashion. Examples of binary mixtures quantified by using THz-TDS are two polymorphs of carbanazepine and enalapril maleate, crystalline and amophous forms of indomethacin, liquid-crystalline and crystalline forms of fenoprofen calcium (Strachan *et al.* 2005).

Not only providing information about the molecule, crystallinity, and purity of pharmaceutical ingredients, but THz-TDS measurements also provide information about the structure of pharmaceutical coatings. In contrast to molecular analyses, which rely on the characteristic spectra, the analysis of the coating structure is facilitated



(a) absorption spectra for acetamidophenol (b) absorption spectra for carbamazepine

Figure 3.7. Measured absorption spectra for some medicines. (a) The spectra for three isomers of paracetamol. After Taday (2004). (b) The spectra for two polymorphs of carbamazepine. After Strachan *et al.* (2004). The isomers and polymorphs have distinctive features in the T-ray spectrum due to the difference in crystallinity.

by the phase sensitivity of THz-TDS systems. Essentially, pulsed T-ray signals, capable of penetrating most of the coating materials, are reflected from internal interfaces of the coating layers. From reflected signals, it is possible to resolve the coating structure with a depth resolution of 30 μm and a lateral resolution 150 μm at 3 THz (Fitzgerald *et al.* 2005). A trial with with ibuprofen tablets showed that the depth profile of the coating resolved via THz-TDS is consistent with the microscope images of cutaway samples (Fitzgerald *et al.* 2005). This early study had a limitation on the scanning area due to the surface curvature of tablets. A robotic arm can alleviate the problem by re-positioning a mounted tablet for every THz-TDS scan, and the full 3D map of a tablet coating can be obtained (Zeitler *et al.* 2007).

It is evident that THz-TDS is greatly beneficial to pharmaceutical assessment, owing to its ability to analyse polymorphs, isomers, ingredients, and even coatings of medicines in a nondestructive way. The rapidness and nondestructive nature of the testing allow a seamless integration of THz-TDS systems into existing production lines of medicines.

3.3.3 Medical and dental diagnostics

T-rays and nearby frequency bands, including microwave, infrared, and visible radiation, are known to provide the contrast at skin-depth levels. This property form the basis of imaging and diagnostics of soft tissue. Infrared and visible imaging techniques operate at short wavelengths, and the image contrast of soft tissue offered by these techniques is relatively low due to Mie scattering (Han *et al.* 2000). The resolution of microwave imaging is limited by the longer wavelength. T-rays lie in the middle between the two extremes, and therefore offer a tradeoff between scattering and resolution (Abbott 2000). In addition, differences in T-ray spectra between different tissue types form the basis of a diagnostic capability (Arnone *et al.* 1999, Huang *et al.* 2009).

The interaction between T-ray radiation and biomolecular tissue was examined by Smye *et al.* (2001). It is known that the T-ray response of living soft tissue is dominated by water content contained therein, and the upper bound of the refractive index of the tissue is due to that of water (Fitzgerald *et al.* 2003). Via the sensitivity of T-rays to the hydration level in tissue, the dryness of human skin and the thickness of the *stratum corneum* can be determined *in vivo* (Pickwell *et al.* 2004b). In order to gain more knowledge about the interaction between the radiation and tissue, mathematical modelling of the electromagnetic propagation has been proposed. In Walker *et al.* (2003)

and Walker *et al.* (2004), the simulation results from two propagation models, i.e., the characteristic matrix and Monte Carlo, were compared to the experimental results on a skin-liked phantom. While the characteristic matrix is simpler to implement, the Monte Carlo model is more flexible for describing such a complex geometry. In Pickwell *et al.* (2004a), finite-difference time-domain (FDTD) was exploited for modelling propagation of T-ray pulses in tissue, and a double-Debye model was also incorporated to explain the behaviour of water content in tissue in response to the radiation. Later on, the developed FDTD model was used to elucidate the experimental results (Pickwell *et al.* 2004b).

Owing to the favourable levels of image contrast in the epidermal layer, it has been proposed that T-rays can assist skin cancer diagnoses. Basal cell carcinoma is one of the two most common types of nonmelanoma skin cancer. It often has a pearly, translucent appearance with a rolled border (Jensen *et al.* 2005). Identification and delineation of basal cell carcinomas by using THz-TDS have been studied extensively. Researchers are able to distinguish cancerous tissue, scar tissue, and inflammatory tissue from normal tissue using a reflection geometry (Woodward *et al.* 2002). The experiments were performed *in vitro* with excised tissue (Woodward *et al.* 2002, Woodward *et al.* 2003b, Woodward *et al.* 2003a), and later on *in vivo* (Pickwell *et al.* 2004b). On probing bulk tissue, the T-ray spectrum does not exhibit any distinct spectral features induced at the molecular level, as the features, if existing, would be obscured due to the inhomogeneity of bulk tissue. It is thus believed that the contrast mechanism that confers the classification ability is, in fact, ascribed to a difference in the water content in normal and cancerous cells (Markelz 2008). The histology confirmed that diseased tissue shows a rise in the water content, which contributes to increased absorption and refractive index compared with normal tissue (Pickwell *et al.* 2005, Wallace *et al.* 2006). Note that a similar relation between the tissue condition and dielectric constants is also held in breast tissue (Ashworth *et al.* 2009).

Another interesting application of THz-TDS can be found in dentistry. A preliminary *in vitro* study by Crawley *et al.* (2003a) indicated that T-ray imaging of teeth potentially results in much greater image contrast compared to X-ray imaging. More specifically, enamel, dentine, caries, and hypomineralisation respond differently to T-rays. Enamel and dentine are differentiable by their refractive indices. Caries cause relatively high

T-ray attenuation, probably via scattering due to their notable porosity. An application for THz-TDS to the early-stage detection of caries was thus anticipated. Crawley *et al.* (2003b) demonstrated that THz-TDS could be utilised in the determination of the enamel thickness, which was, at the time of study, inaccessible by any other non-destructive means. As mentioned earlier, enamel and dentine have different refractive indices. Thus, the reflection of a T-ray pulse at the air-enamel and enamel-dentine interfaces reveals information about the enamel thickness. The technique was envisaged for *in vivo* measurements to monitor enamel erosion. A further study by Pickwell *et al.* (2007) shows correlation between the mineral content profile obtainable from transmission microradiography and the refractive index profile from THz-TDS.

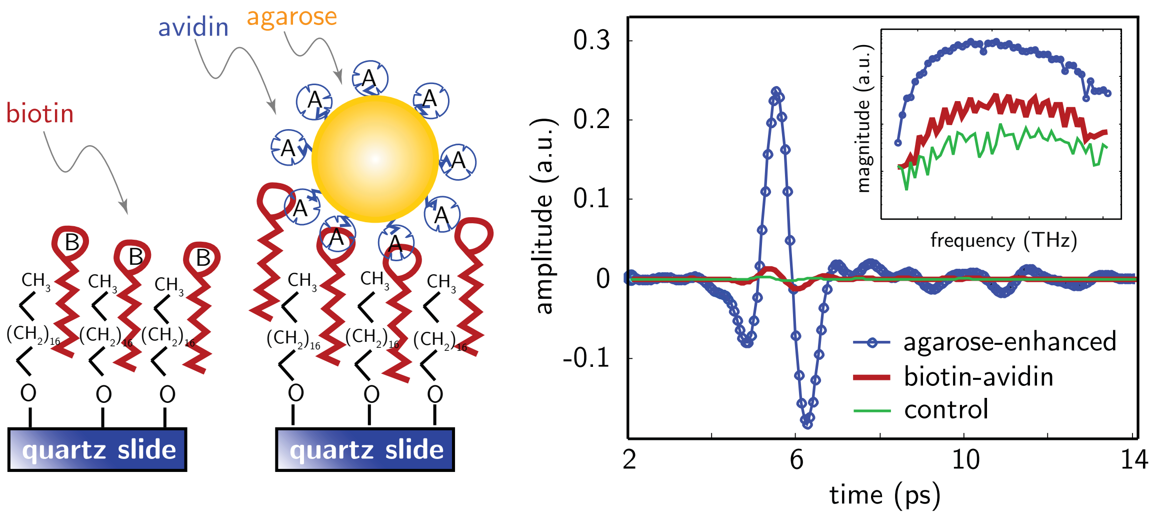
Extensive reviews of T-rays in biology and medicine can be found in Siegel (2004) and Pickwell and Wallace (2006).

3.3.4 Biosensing

Biosensing is typically performed through molecular binding, i.e., the specific affinity between a pair of biomolecules. Only target molecules or the ligand can be bound to sensor molecules or the analyte. In a test, a slide, containing the analyte, is exposed to the molecules of interest to find traces of the ligand. The exposed slide is then detected for the binding between the ligand and the analyte. Some methods for their detection are available commercially, and these include surface plasmon resonance (SPR) and fluorescence labelling.

It has been shown that THz-TDS is able to detect biomolecular affinity with an advantage of contactless sensing. Through the high sensitivity of double-modulated differential time-domain spectroscopy (DTDS), the binding of ultrathin avidin layers to biotin is detectable down to the level of 0.1 μg/cm^2 (Mickan *et al.* 2002c, Menikh *et al.* 2002, Menikh *et al.* 2004). The result from this experiment is shown in Figure 3.8. It is believed that the mechanism underlying the sensitivity of T-rays to the binding is attributed to a change in the structural vibrational modes and relaxation loss of the sample (Markelz 2008).

A sensitivity enhancement for biosensing is achievable by using a microstrip line resonator. The resonator is basically integrated into a transmission line terminated at both ends by a PCA transmitter and receiver. A minute quantity of biomolecules loaded



(a) biotin-avidin binding with enhancing agarose

(b) T-ray signals and spectra

Figure 3.8. Bioaffinity between biotin and avidin. (a) The sensor is composed of biotin, or the analyte, on a quartz slide. Exposure of the sensor to avidin or the ligand results in biotin-avidin binding. Agarose beads are used to enhance the difference between bound and unbound sensors. (b) The T-ray signals and spectra available from from DTDS measurements show different responses among the samples. It is clear that DTDS can detect the presence of the ligand. After Mickan *et al.* (2002c).

onto the resonator can be sensed by T-rays travelling the transmission line. The sensing ability is attributed to the resonance frequency of the resonator, which is inversely proportional to the refractive index of a loaded sample. This concept was first demonstrated in Nagel *et al.* (2002a) and Nagel *et al.* (2002b). In their experiment, a thin-film microstrip resonator as a high-Q bandpass filter with a centre frequency of 610 GHz was loaded with DNA samples. Difference between hybridised (double stranded, on-binding state) and denatured (single stranded, off-binding state) DNA was detectable by the technique, since hybridisation causes a higher refractive index of a DNA sample (Brucherseifer *et al.* 2000). The technique achieved molecular sensitivity in the order of femtomols, and could also detect single base mutations in DNA. Conceptual designs for alternative resonating structures were reported by Baras *et al.* (2003). Higher-sensitivity resonators, based on the same principle, were proposed (Stewing *et al.* 2004). Much research work is necessary in order to increase the sensitivity and throughput of T-ray biosensing, in order that it becomes competitive with SPR-based techniques.

Table 3.4. Performance among spectral ranges exploited in security screening. Only X-rays can penetrate metals, but the energetic radiation also leads to the ionisation of exposed molecules. Either T-rays or infrared provides the signature of materials, yet only T-rays have an ability to see-through most kinds of packages. After Federici *et al.* (2007).

Spectrum range	Transmission through metals	Transmission through dielectrics	Material signature	Non-ionising
Millimetre-wave	–	✓	–	✓
T-rays	–	✓	✓	✓
Mid- and near-infrared	–	–	✓	✓
X-rays	✓	✓	–	–

3.3.5 Security screening

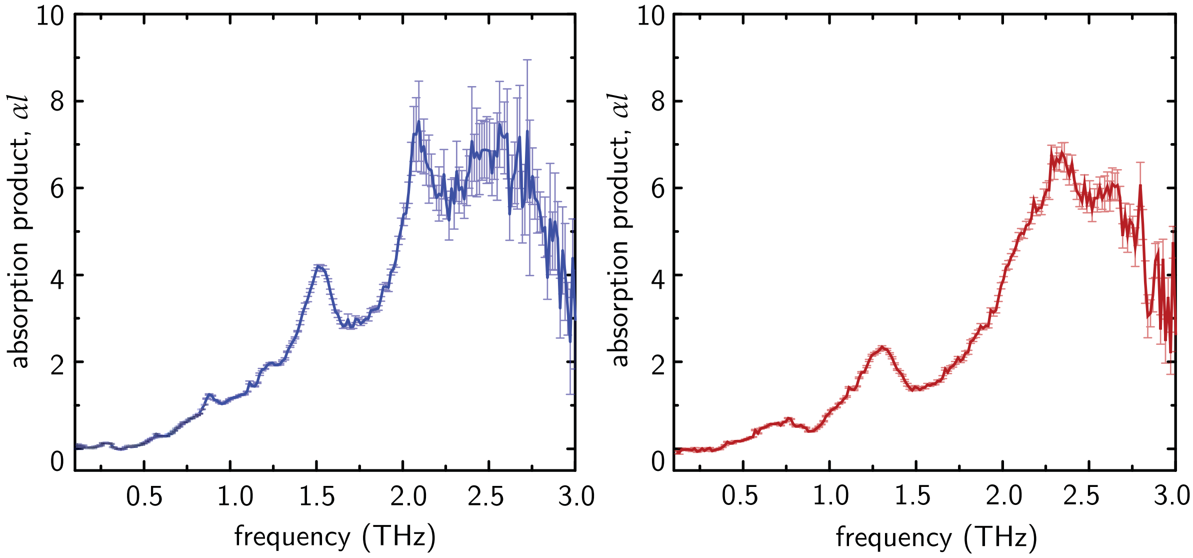
The failure of conventional security screening measures to counteract recent terrorist threats and drug smuggling has stimulated the search for complementary capable technologies (Singh and Singh 2003, Schubert and Rimski-Korsakov 2006). One potential solution is T-ray technology, which has been envisaged in a range of security applications (National Research Council (U.S.), Board of Chemical Sciences and Technology, National Academy of Sciences 2004). The outgrowth of T-rays in this arena is ascribed to three unprecedented advantages, as follows (Federici *et al.* 2005, Kemp *et al.* 2006): (i) several illicit substances exhibit characteristic resonances in the T-ray spectrum, which is useful for chemical recognition; (ii) T-rays can penetrate non-metallic, non-polar materials inclusive of clothes, plastics, and papers, and allows detection of concealed objects; and (iii) T-rays are generally considered harmless to living tissue for average power levels in the order of milliwatts, and therefore scanning of humans is viable. A comparison of T-rays with some frequency ranges in terms of screening performance is given in Table 3.4.

As opposed to household materials, numerous explosives have chemical fingerprints in the T-ray frequency band at room temperature (Kemp *et al.* 2006). Figure 3.9 demonstrates spectra for notorious explosives in comparison to spectra for common clothing. Some explosive materials possessing T-ray fingerprints are, for example, HMX (Allis *et al.* 2006), RDX (Huang *et al.* 2004, Shen *et al.* 2005, Liu *et al.* 2006), DNT (Chen *et al.* 2004), PETN (Allis and Korter 2006), γ-HNIW (Guo *et al.* 2006), TNPG (Zhou *et al.*

NOTE:
These figures are included on page 81 of the print copy of the thesis held in the University of Adelaide Library.

(a) absorption spectra for explosives (b) absorption spectra for clothes

Figure 3.9. Measured absorption spectra for some explosives and clothes. The absorption spectra for explosive materials in (a) exhibit characteristic resonances, whilst the spectra for clothes in (b) rise monotonically with no clear features. The fast oscillations in the spectra of clothing are due to Fabry-Pérot effects. The spectra in (a) are vertically offset, and both figures use different absorption scale. After Kemp *et al.* (2006).



(a) absorption spectra for cocaine (b) absorption spectra for morphine

Figure 3.10. Measured absorption spectra for cocaine and morphine. Cocaine and morphine under measurement are in the polycrystalline forms and sealed in PE bags. The spectral features for these substances are unique and not affected by the package. The error bars indicate one standard deviation. The vertical axis is the product of the absorption coefficient and the length, or αl. After Fischer *et al.* (2005b).

2008). Single crystalline HMX, RDX, and PETN are shown to have spectra that are dependent on the temperature and crystal orientation (Barber *et al.* 2005). Though, no characteristic spectral feature is found in the case of polymer-bonded explosives—PBX 9501 and PBX 9502 (Funk *et al.* 2004). In addition to explosive materials, many illicit drugs also produce unique spectra in the T-ray band, e.g., methamphetamine (Ning *et al.* 2005), cocaine, morphine (Fischer *et al.* 2005b). The spectra for cocaine and morphine are shown in Figure 3.10. The obvious fingerprint resonances result from phonon-like intermolecular modes in the crystalline structures (Federici *et al.* 2005, Liu *et al.* 2006). On top of pure material analyses, nondestructive *retection*[7] of hidden explosives or drugs via T-rays is possible. The case study includes methamphetamine, MDMA, and aspirin hidden in a mail envelop (Kawase *et al.* 2003, Watanabe *et al.* 2003), C-4 in an envelop (Yamamoto *et al.* 2004a), or RDX covered by paper, PE sheet, leather, or polyester cloth (Liu *et al.* 2006).

Although T-rays can easily penetrate most types of packages, in realistic situations the transparency of the concealed target and material beneath it needs to be considered. A chemical sample with a moderate thickness can appear opaque to T-rays. Likewise, T-rays generated by typical THz-TDS systems cannot penetrate human skin deeper than 1 mm (Woodward *et al.* 2002). The opaqueness of the target and materials nearby prohibits the utility of transmission-mode spectroscopy. In these situations, T-ray sensing in a reflection mode is favourable (Federici *et al.* 2005, Liu *et al.* 2006). Reflection-mode spectroscopy can produce reflection spectra of chemical substances that contain characteristic resonances similar to transmission spectra (Federici *et al.* 2005). Thus, as long as the package covering the target is penetrable, spectroscopy produces useful information about the target. A successful detection of RDX was demonstrated with specular reflection spectroscopy (Shen *et al.* 2005) and via a more realistic diffuse reflection spectroscopy (Liu *et al.* 2006).

Another problem which may be encountered in real-world applications is the lack of reference information. Typically, in a THz-TDS measurement, a reference signal accompanying a sample signal is required in order to remove the system dependency via signal processing. However, outside of the laboratory where the observer loses control over the environment, a reference scan is hardly attainable. Computational techniques based on various theoretical assumptions have been proposed to pinpoint the locations of resonances in the spectrum from only a sample signal. For a transmission-mode

---

[7] The word *retection* means detection of a concealed object (Abbott and Zhang 2007).

measurement, by differentiating the phase of a sample signal with respect to the frequency, an absorption-like spectrum, where the resonances are pronounced, is readily available (Zhang *et al.* 2008b). The procedure is similar for reflection measurements, but the second-order derivative of the phase is used instead (Zhong *et al.* 2008).

Threats that have no characteristic absorption at T-ray frequencies can still be detectable in specific circumstances. An example is a group of mostly nonpolar inflammable liquids that is indistinguishable from water by visual inspection. Ikeda *et al.* (2005) probed some inflammable liquids, including benzine, kerosene, gasoline, and gas oil, in PE and PET bottles using a transmission-mode THz-TDS. These inflammable liquids pass a fraction of the T-ray energy, whilst the same quantity of water totally blocks the beam path. Via differences in the absorption and dispersion, it is possible to roughly classify these liquids and distinguish them from water. Another example is a gun or knife that can present its shape through an otherwise opaque layer via T-ray imaging. A real-time T-ray imaging system operated in reflection mode successfully reveals the shape of a toy gun hidden beneath a fabric bag (Zhang *et al.* 2008a).

One of the most challenging aspects of security screening is standoff detection. A number of challenges in this area present themselves (Kemp *et al.* 2006), e.g., the absorptive atmosphere reducing the signal strength and the dynamic target and background hampering the measurement precision. In addressing such issues, CW systems are preferable to pulsed systems. With CW systems, it is possible to transmit only T-ray frequencies that fall into an atmospheric window, i.e., a spectral region with a low atmospheric absorption (Lee *et al.* 2006, Liu *et al.* 2007b). Particular frequencies are also selected on the basis of the material fingerprint and package transparency at that frequency window (Liu *et al.* 2007b). As a CW system requires no mechanical scanning delay line, the measurement time is remarkably fast. The rapid measurement can cope with the dynamics of the target and background. Despite these benefits, CW systems lack depth information and broadband sensitivity, normally available from pulsed systems.

3.3.6 Miscellaneous nondestructive testing

Nondestructive testing using THz-TDS can be carried out with any structure made of materials transparent to T-rays. The coherent and broadband nature of the system provides both the depth and spectroscopic information of the target. Some prominent THz-TDS applications to nondestructive testing include the inspection of the Space

Shuttle's insulation, the analysis of mural paintings, and the quality control of foods and wood products, etc.

It appears that THz-TDS technology came into public consciousness following the Columbia Space Shuttle disaster. After thoroughly investigating the accident, NASA and its contractors concluded that insulating foam shed from an external fuel tank struck the shuttle's left wing during the launch and caused a breach therein, allowing extreme heat to destroy the shuttle at its re-entry. This foam insulation problem needed to be resolved before any further shuttle launch. Accordingly, THz-TDS, along with X-ray backscatter radiography, were selected as nondestructive methods for inspection of sprayed-on foam insulation. Due to low T-ray absorption and scattering, voids and disbonds inside the foam can be revealed in a 2D T-ray map. The insulation foam was examined by T-ray reflection imaging both in pulsed (Zhong *et al.* 2005) and CW modes (Karpowicz *et al.* 2005), where pulsed imaging can add a third dimension to the map due to the phase sensitivity.

Another promising application of T-rays to nondestructive testing is the analysis of mural paintings. Although many frequency regimes have been exploited for painting analysis in the past, T-rays offer unprecedented performance. Whilst UV and IR radiation are able to resolve chemical components at the surface of paintings, they cannot penetrate lower layers; whereas X-ray and microwave radiation, with a penetration capability, lack depth resolution and fine spatial resolution, respectively (Jackson *et al.* 2008). As opposed to other modalities, T-rays can sense contrast in the chemicals of pigments, and can reveal hidden layers, i.e., sketches and antecedent paintings, in frescoes and murals (Jackson *et al.* 2008). Generally speaking, windowing the reflected T-ray time-domain signal selects a layer of a painting to be analysed, whilst the Fourier transform of a windowed signal reveals spectroscopic information about chemicals.

The food industry sees the potential advantages of THz-TDS technology. Jepsen *et al.* (2007) proposed a method for determining the concentration of solutes in a solution via THz-TDS. Provided with the complex dielectric constants of a solvent, the method can estimate the concentration of up to two solutes in a solution by using an empirical model, established from THz-TDS measurements of several controlled samples. They successfully determined the concentrations of ethanol/sucrose in water, and the concentration of alcohol content in a number of commercial liquors. Further, they reported that carbonation, which causes gas bubbles in a solution, does not affect the measurement results, given that the bubbles escape the measuring area rapidly.

The timber and paper industries can also benefit from THz-TDS. Reid and Fedosejevs (2006) showed that, in response to T-ray radiation, wood and paper exhibit a birefringent property because of the orientation of their microscale fibres. Hence, the fibre orientation can be measured as a function of the delay time of the T-ray pulse. The effective refractive index of wood or paper equals $n_\perp \cos\theta + n_\parallel \sin\theta$, where $n_\perp$ and $n_\parallel$ are the refractive indices in the perpendicular and parallel directions with respect to the fibre orientation, respectively, and θ is the angle that the T-ray polarisation makes with the vector perpendicular to the fibre orientation (Reid and Fedosejevs 2006). Koch *et al.* (1998) applied THz-TDS imaging to density mapping of wood. By scanning over a piece of wood, a transmission map with a sub-annual ring resolution can be attained. The correlation between the wood density and the T-ray absorption reveals the locations of earlywood, latewood, xylem, annual rings on the T-ray transmission map. The concept was later extended to the nondestructive evaluation of cork, of which cracks, voids, and defects are manifested on a T-ray map (Hor *et al.* 2008). Mie scattering in cork cells was also studied (Hor *et al.* 2008).

3.4 T-ray absorption of water

The T-ray absorption of water is a fundamental limit to many applications for THz-TDS. In order to understand more about this limitation, water has been characterised extensively through THz-TDS both in the liquid and gas phases. In the liquid phase the relaxation process of oriented dipole moments leads to a strong damped absorption in the T-ray regime (Thrane *et al.* 1995). A sample with a considerable hydration level hence appears opaque to T-ray radiation. The response of water vapour to T-rays is dictated by the rotational transitions of polar molecules (van Exter *et al.* 1989b). A T-ray spectrum measured in ambient atmosphere thus contains numerous water vapour resonances. It can be seen that water either in the liquid or gas phase can be troublesome in that it can hinder the observation of the spectroscopic features of a sample.

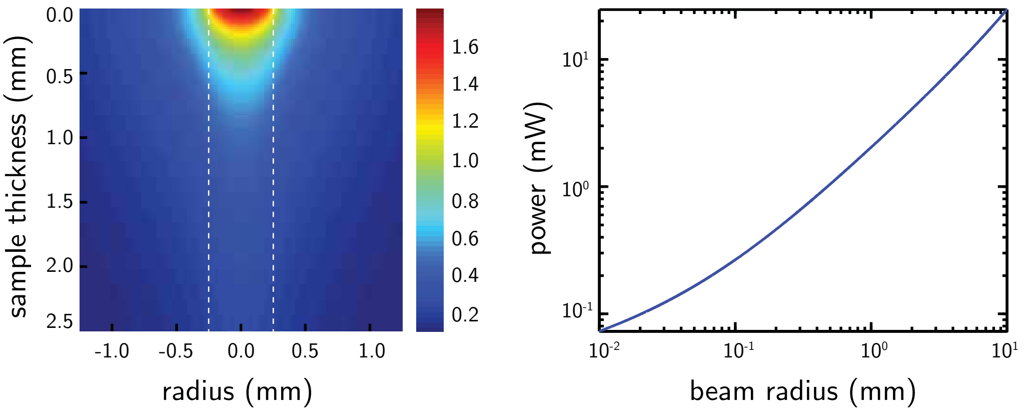
In the case that a specimen contains an excessive level of water, dehydration of the specimen can assist the measurement (Png *et al.* 2008). However, for biological measurements, dealing with water content is challenging and must be considered carefully (Png *et al.* 2008). A low level of water content can stop normal biological functions and destroy cellular structures, whereas a high level of water content can substantially attenuate a probing T-ray signal. Reflection-mode THz-TDS is a viable option to overcome the issue (Huang *et al.* 2009). Otherwise, the problem might be alleviated during the sample preparation stage. Several methods in preparation of biological samples have been suggested with different advantages and disadvantages (Markelz 2008). Some of them are, e.g., alcohol-dehydration, formalin fixing, lyophilising, polyethylene mixing, or hydrated film. A hydrated film is probably the best option in retaining the function and structure of biological samples. Even so, a film must be highly uniform to avoid the spatial-dependent response, which is always accentuated in a thin sample. In addition to those conventional methods, a specially-designed membrane, made of hydrophilic polyethersulfone or polypropylene, is appropriate for holding biological samples for THz-TDS measurements while preserving aqueous environment (Yoneyama *et al.* 2008).

Water vapour in ambient air can be reduced by purging the T-ray path with dry air or a nonpolar gas such as nitrogen. But this procedure is probably possible only in the laboratory environment. For an outdoor measurement or standoff detection, absorption of T-rays by water vapour, water droplets, and rain seems to be unavoidable (Liebe 1989). The situation becomes worse for long distance propagation, where most of the T-ray energy is absorbed. However, it is pointed out that frequencies around 0.8 THz are less absorbed by water vapour, and could be a potential atmospheric window (Yuan *et al.* 2003). For moderate water-vapour absorption appearing in T-ray spectra, Chapter 5 in this thesis addresses the problem by proposing a signal processing solution.

3.5 Safety concerns on T-rays and THz-TDS

T-ray photons have significantly lower energies (see photon energy in Figure 1.2) than do X-rays, and are unable to damage biological cells via ionisation. Furthermore, in spite of the nonlinear interaction between tissue and coherent T-ray radiation as verified by Fröhlich (1980) and Grundler and Kaiser (1992), T-rays are noninvasive to biological samples. This is evident from the low energy level involved in the absorption process, when T-rays are incident on the sample. For example, at a worse-case frequency f of 10 THz, the molar energy calculated from $E = N_A hf$, where N_A is Avogadro's number and h is Planck's constant, equals 4 kJ mol^{-1}. In comparison, the ATP process that is fundamental to biological systems takes up 31.8 kJ mol^{-1} (Guynn and Veech 1973). Thus the T-ray photon energy is below the level for activating any biologically significant process.

In terms of power density, from Planck's black body curve the worst-case T-ray background power level at room temperature is in the order of 10^{-4}W/THz/cm^2, whereas the average power from a typical laser-based THz-TDS system is in the order of 10^{-9} W/THz/cm^2 (Fischer *et al.* 2005b). It is obvious that typical THz-TDS systems have average power well below the background black body radiation level. An in-depth investigation of T-ray heating effects on a sample, based on Kirchhoff's heat equation, reveals some interesting facts as follows (Kristensen *et al.* 2010). For an incident CW beam at 1 THz with a spot size of 0.5 mm, the steady-state temperature increase in water is 1.78 K/mW. Hence, for a typical THz-TDS system, with source power in the μW region, this creates a temperature increase in the order of milli-Kelvins, which can be considered insignificant and harmless to biological cells. Furthermore, given the same total radiation power, the heating effect is more intense for a smaller beam width, but more penetrative for a larger beam width. Figure 3.11 shows the distribution of temperature increase in water exposed to T-rays and the power required to heat up water by 1 K. Note that water constitutes a significant portion of the soft tissue, and thus the



(a) distribution of temperature increase (b) power required to heat water by 1 K

Figure 3.11. Heating effects on water of T-rays. (a) The shading reveals the calculated steady-state temperature increase per milliwatt inside a water sample exposed to a T-ray beam indicated by the vertical dashed lines. (b) The power required to heat up water by 1 K as a function of the beam radius. In the calculation the initial temperature is 298.15 K (25°C), the frequency is 1 THz, and the beam radius is 0.25 mm. After Kristensen *et al.* (2010).

absorption coefficient of water at terahertz frequencies represents an upper bound of that for dry soft tissue, so the temperature increase estimated here is for the worst case.

Despite the positive aspects mentioned above, T-ray radiation has not been confirmed to be completely harmless to biological systems (Berry 2003). A mathematical analysis indicates that the T-ray field could aversely affect double-stranded DNA and thereby interfere with gene expression and DNA replication processes (Alexandrov *et al.* 2009). In addition, thermomechanical damage to tissue may be introduced by a pulse of less than 1 μs duration (International Commission on Non-Ionizing Radiation Protection 1996)—the subpicosecond T-ray pulses from THz-TDS fall into this case. A case in point is microwaves, which have a lower photon energy than that of T-rays, but possess known dangers. A direct exposure to microwaves can induce genetic changes (Banik *et al.* 2003), and hence a standard recommending the maximum permissible human exposure (MPE) to electromagnetic radiation from 3 kHz to 300 GHz is necessary (IEEE Standards Coordinating Committee 28 on Non-Ionizing Radiation Hazards 1999).

In the same fashion, a safety guideline for T-rays should be also established to prevent serious biological damage following an excessive exposure. However, owing to the early stage of T-ray technology, its effects on living tissue is not fully understood. In the meantime, the MPE to T-rays could be determined by drawing analogy from nearby frequency ranges (Berry *et al.* 2003).

When considering the safety issue of laser-based THz-TDS systems, the safety of high-power lasers must be taken into consideration. Even though these lasers are not utilised in the applications directly, they are indispensable in opto-electronic THz-TDS. The radiation from these lasers is considerably harmful because of their relatively high transmitted power and high photon energy. A pump/probe mode-locked laser typically emits a train of pulses with an average power up to 1 W in the NIR spectrum. According to International Electrotechnical Commission (2001), this type of laser falls into class 4, which is hazardous to eyes and skin.

3.6 Conclusion

In this chapter, a range of applications for THz-TDS have been surveyed. The survey is composed of two major parts: fundamental spectroscopic studies and advance applications. Fundamental studies involve the use of THz-TDS for observation of the intrinsic properties of materials in response to T-rays. These studies include spectroscopy of dielectrics and semiconductors, rotational spectroscopy of polar gases, vibrational spectroscopy of crystalline structures, and functional spectroscopy of biomolecules. On the other hand, advanced applications are relevant to those utilising the capabilities of THz-TDS for real-world tasks. Some promising applications of THz-TDS are anticipated for gas recognition, quality control of pharmaceuticals, medical and dental diagnostics, biosensing, security screening, and other forms of nondestructive testing.

Fundamental spectroscopic studies of material response via THz-TDS are maturing. It is considered that THz-TDS can plausibly fill the terahertz gap, and completes missing knowledge of material response in this frequency regime. Though, a wider system bandwidth and finer spectral resolution are always desired in order to resolve more spectral details. A fundamental limit to the study is the absorption of T-rays by water in samples, which often influences observable results. Future effort should be devoted to accurate theoretical modelling of molecular response, which is dominated by complex molecular structures in this low-frequency region.

As opposed to fundamental studies, almost all of the proposed advanced applications are currently in the development phase. These applications encounter a long acquisition time and low output power. In addition, the applications that require stand-off measurements suffer strong absorption by atmospheric humidity. When measuring arbitrary samples, the system may deliver spurious results due to scattering, reflection, and diffraction. In the near future, developments for improved THz-TDS systems will remain a central goal. But we can extrapolate from other mature technologies that signal processing techniques can be exploited to enhance and utilise the measurement data. Supported by open databases of material fingerprints in the T-ray frequency range (Graydon 2008), pattern recognition will foreseeably play a major role in bringing THz-TDS closer towards commercial applications.

Having discussed several potential applications of THz-TDS in this chapter, we now turn to Chapter 4 that reviews various techniques for extracting meaningful information from THz-TDS signals. These techniques are vital to both fundamental material studies and advanced applications that have been presented in this chapter.

Chapter 4

Material Characterisation with THz-TDS

TERAHERTZ time-domain spectroscopy is a technique capable of measuring optical constants of materials with broadband coherent T-ray radiation. A broadband T-ray signal transmitted through or reflected from a material sample is influenced by the absorption and dispersion of the sample. A change in the transmitted or reflected signal hence relates to the intrinsic properties of the sample, which can be extracted via the physical models of a propagating wave. This chapter provides a review of the material parameter extraction processes for several THz-TDS configurations, e.g., transmission or reflection, together with selected signal processing schemes commonly used with THz-TDS data.

4.1 Introduction

As discussed in Chapter 3, THz-TDS has been utilised in many applications, including medical diagnostics, industrial quality control, security screening, and so on. One of the most widely used applications is the characterisation of materials. The capabilities of THz-TDS allow access to the broadband response of materials in the T-ray regime. A number of solid, liquid, and gas samples have been characterised for their intrinsic properties by THz-TDS.

The broadband response of a material results in the modified amplitude and phase of a probe T-ray signal. This response can be described in terms of a complex refractive index, which is intrinsic to a material. However, other factors also contribute to the modification of the amplitude and phase of the signal—such as the shape and surface morphology of a sample under measurement and the propagation path. In addition, the measured signal is in fact a convolution between the incident T-ray field, the optical probe pulse, and the system response. Thus, in order to attain a complex refractive index from the measurement, it requires the geometrical analysis of wave propagation along with a reference measurement.

Although many THz-TDS measurement configurations with different geometries for wave propagation are available to date, some are more widespread than the others due to the simplicity or applicability to certain types of samples. This chapter reviews some common arrangements of THz-TDS. The mathematical models specific to these configurations are derived, relating the amplitude and phase changes in the measured signals to the optical constants of materials. Also some general signal processing methods typically required in the parameter extraction process are given.

The chapter is organised as follows. Section 4.3 give the introduction to the fundamental of wave propagation, including Fresnel equations and a model for wave propagation through a dielectric slab. Many configurations for transmission- and reflection-mode THz-TDS measurements are discussed in details in Section 4.4 and 4.5, respectively. Section 4.6 provides further information about signal processing methods usually accompanying the parameter extraction process.

4.2 Research objective and framework

Objective

To provide a review on material characterisation via transmission- and reflection-mode THz-TDS systems. The transmission-mode THz-TDS settings comprise (i) a single-sample measurement, (ii) a dual-sample measurement, (iii) a multiple-sample measurement, and (iv) a sample-in-cell measurement. The reflection-mode THz-TDS settings comprise (i) a single-reflection measurement and (ii) a double-reflection measurement. The mathematical relations, necessary for estimating the optical constants from measured T-ray signals, are provided in every case. Furthermore, the chapter includes signal processing topics, relevant to the parameter estimation, e.g., signal averaging, signal denoising, phase unwrapping, and reflection removal.

Framework

The following assumptions are made: A suitable material for characterisation can be either a dielectric or a semiconductor, in the homogeneous form, preformed into a slab with parallel and flat surfaces. The scattering of an incident T-ray beam is negligible. The transverse dimension of the sample is larger than the incident beam waist, so there is no edge diffraction. The sample under measurement is in free space with an index of refraction of unity. The goal of material characterisation, in this case, is extraction of optical constants, i.e., the refractive index and extinction coefficient. From these parameters, other relevant parameters, such as dielectric constants, can then be derived.

4.3 Propagation of electromagnetic wave in dielectrics

In this section, the background behind wave propagation based on ray optics is introduced. This includes wave reflection and refraction at a dielectric interface, wave propagation through a dielectric slab, and Fabry-Pérot effects. The background enables further analyses into the relation between the sample signal and the optical properties of a material for arbitrary measurement settings.

4.3.1 Propagation across an interface—Fresnel equations

A set of Fresnel equations defines the amplitude of the wave reflected and refracted at a dielectric interface. In Figure 4.1, when the plane wave propagates from medium a into medium b, and the polarisation of the electric field is perpendicular to the plane of incidence (σ-polarised, s-polarised, or TE wave), the Fresnel transmission coefficient is given by

$$\tau_\sigma = \frac{2\hat{n}_a \cos\theta_i}{\hat{n}_a \cos\theta_i + \hat{n}_b \cos\theta_t}, \tag{4.1a}$$

and the Fresnel reflection coefficient is

$$\rho_\sigma = \frac{\hat{n}_a \cos\theta_i - \hat{n}_b \cos\theta_t}{\hat{n}_a \cos\theta_i + \hat{n}_b \cos\theta_t}, \tag{4.1b}$$

where θ_i is the angle of incidence; θ_t is the angle of refraction; and $\hat{n}_a$ and $\hat{n}_b$ are the complex refractive indices of the media. The complex refractive index, $\hat{n}(\omega) = n(\omega) - j\kappa(\omega)$, comprises the index of refraction, $n(\omega)$, and the extinction coefficient, $\kappa(\omega)$, which, together, are referred to as the optical constants.

On the other hand, if the polarisation of the electric field is parallel to the plane of incidence (π-polarised, p-polarised, or TM wave), the transmission coefficient is given by

$$\tau_\pi = \frac{2\hat{n}_a \cos\theta_i}{\hat{n}_a \cos\theta_t + \hat{n}_b \cos\theta_i}, \tag{4.2a}$$

and the reflection coefficient is given by

$$\rho_\pi = \frac{\hat{n}_a \cos\theta_t - \hat{n}_b \cos\theta_i}{\hat{n}_a \cos\theta_t + \hat{n}_b \cos\theta_i}. \tag{4.2b}$$

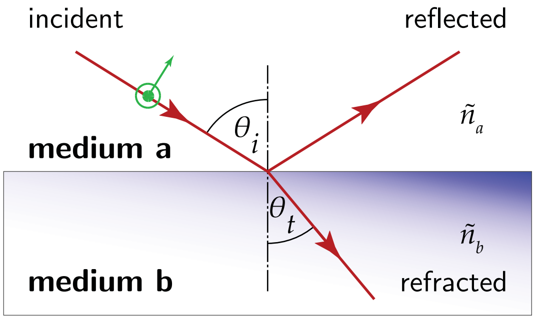


Figure 4.1. Reflection and refraction of an incident wave at interface. Regardless of the polarisation, the reflection angle is identical to the incident angle θ_i, and the refraction angle θ_t, according to Snell's law, equals $\arcsin(\sin\theta_i \cdot n_a/n_b)$. The amplitudes of the reflected and refracted waves depend on the polarisation, and are discussed in Section 4.3.

Clearly, if the angle of incidence is normal to the interface, $\tau_\sigma = \tau_\pi$ and $\rho_\sigma = \rho_\pi$.

If τ' and ρ' are the transmission and reflection coefficients when the wave is incident in medium b at the angle θ_t, two relations, independent of the wave polarisation, can be established as follows:

$$\rho' = -\rho\,, \tag{4.3a}$$
$$\tau\tau' = 1-\rho^2\,. \tag{4.3b}$$

Section 4.3.2 exploits these Fresnel equations for deriving the response of a dielectric slab to an incident wave.

4.3.2 Propagation through a dielectric slab

As shown by the Fresnel formulas in Section 4.3.1, while the complex plane wave propagates across the interface of two media with different refractive indices, a fraction of the wave energy is refracted into the second medium while the remaining is reflected back into the first medium. If the second medium is a slab, the traversing refracted wave would encounters the other interface, where reflection and refraction occur again. The reflection and refraction within the slab continue until the wave energy dissipates. This phenomenon ideally causes the splitting of the wave into an infinite number of intermediate paths, as illustrated in Figure 4.2.

The wave taking each path in Figure 4.2 has its expression determined by the multiplication between the incident wave $E_0(\omega)$ and a corresponding factor. At each interface, the wave characteristics are modified by the Fresnel transmission or reflection coefficient. Inside a bulk material, the exponential term, representing the complex response of bulk material, influences the propagating wave. Multiplication factors for all paths are given in Table 4.1.

In the case where the angle of incidence of the incoming wave is not strongly oblique, all the transmission (reflection) paths overlap. The total transmitted wave is therefore the summation of waves among the intermediate transmission paths, or following Table 4.1,

$$\begin{aligned} E_t(\omega) &= \tau\tau' \cdot \exp\left[-j\hat{n}_b(\omega)\frac{\omega l_\theta}{c}\right] \\ &\quad \cdot \left\{1 + \rho'^2 \exp\left[-2j\hat{n}_b(\omega)\frac{\omega l_\theta}{c}\right] + \rho'^4 \exp\left[-4j\hat{n}_b(\omega)\frac{\omega l_\theta}{c}\right] + \ldots\right\} \cdot E_0(\omega) \\ &= \tau\tau' \cdot \exp\left[-j\hat{n}_b(\omega)\frac{\omega l_\theta}{c}\right] \cdot \mathrm{FP}(\omega) \cdot E_0(\omega)\ , \end{aligned} \tag{4.4}$$

where $\mathrm{FP}(\omega)$ represents the Fabry-Pérot effect, as a consequence of reflections, and is given by,

$$\begin{aligned} \mathrm{FP}(\omega) &= \left\{1 + \rho'^2 \cdot \exp\left[-2j\hat{n}_b(\omega)\frac{\omega l_\theta}{c}\right] + \rho'^4 \exp\left[-4j\hat{n}_b(\omega)\frac{\omega l_\theta}{c}\right] + \ldots\right\} \\ &= \sum_{m=0}^{\infty} \left\{\rho'^2 \cdot \exp\left[-2j\hat{n}_b(\omega)\frac{\omega l_\theta}{c}\right]\right\}^m \\ &= \left\{1 - \rho'^2 \cdot \exp\left[-2j\hat{n}_b(\omega)\frac{\omega l_\theta}{c}\right]\right\}^{-1}\ . \end{aligned} \tag{4.5}$$

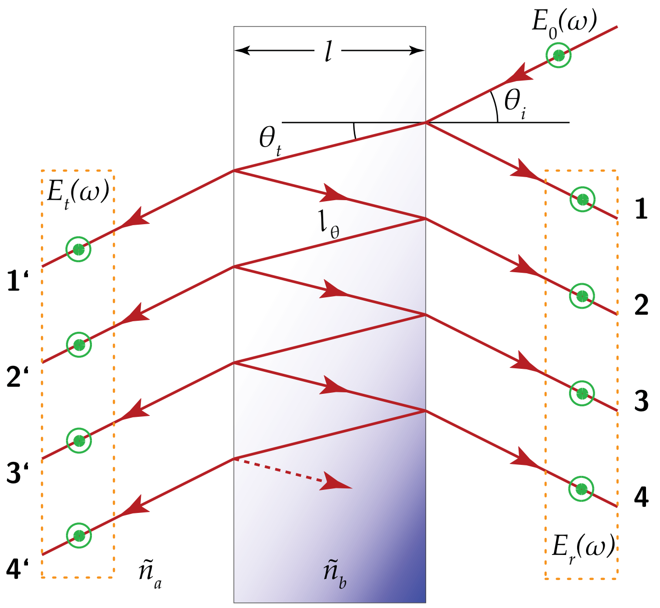


Figure 4.2. Wave propagation paths through a slab of a homogeneous dielectric. The refractive indices of surrounding and dielectric are denoted by $\hat{n}_a$ and $\hat{n}_b$, respectively. In addition, $E_0(\omega)$ is the incoming wave; θ_i is the angle of incidence; and θ_t is the refraction angle. The propagation distance l_θ is derived from the sample thickness l, and equals $l/\cos\theta_t$. The arrows show the propagation paths of the incident wave, which are grouped into the transmission, $1', 2', 3', \ldots$, and the reflection, $1, 2, 3, \ldots$.

Table 4.1. Multiplication factors for a propagating wave. Here, τ and ρ are the complex transmission and reflection coefficients for the wave incident in medium a. Likewise, τ' and ρ' are the transmission and reflection coefficients for the wave incident in medium b. These coefficients can be described in terms of the refractive indices, as discussed in Section 4.3.1. The path numbers and other variables are associated with Figure 4.2. After Klein (1970).

Reflection path	Multiplication factor	Transmission path	Multiplication factor
1	ρ	1'	$\tau\tau' \exp[-j\hat{n}_b \frac{\omega l_\theta}{c}]$
2	$\tau\tau'\rho' \exp[-2j\hat{n}_b \frac{\omega l_\theta}{c}]$	2'	$\tau\tau'\rho'^2 \exp[-3j\hat{n}_b \frac{\omega l_\theta}{c}]$
3	$\tau\tau'\rho'^3 \exp[-4j\hat{n}_b \frac{\omega l_\theta}{c}]$	3'	$\tau\tau'\rho'^4 \exp[-5j\hat{n}_b \frac{\omega l_\theta}{c}]$
4	$\tau\tau'\rho'^5 \exp[-6j\hat{n}_b \frac{\omega l_\theta}{c}]$	4'	$\tau\tau'\rho'^6 \exp[-7j\hat{n}_b \frac{\omega l_\theta}{c}]$
5	...	5'	...

In analogy to the total transmitted wave, the total reflected wave can be expressed as

$$\begin{aligned} E_r(\omega) &= \rho E_0(\omega) + \tau\tau'\rho' \cdot \exp\left[-2j\hat{n}_b(\omega)\frac{\omega l_\theta}{c}\right] \\ &\quad \cdot \left\{1 + \rho'^2 \cdot \exp\left[-2j\hat{n}_b(\omega)\frac{\omega l_\theta}{c}\right] + \rho'^4 \exp\left[-4j\hat{n}_b(\omega)\frac{\omega l_\theta}{c}\right] + \ldots\right\} \cdot E_0(\omega) \\ &= \rho E_0(\omega) + \tau\tau'\rho' \cdot \exp\left[-2j\hat{n}_b(\omega)\frac{\omega l_\theta}{c}\right] \cdot \mathrm{FP}(\omega) \cdot E_0(\omega)\ . \end{aligned} \tag{4.6}$$

These basic equations are vital in further analyses of wave propagation in different measurement geometries.

4.3.3 Fabry-Pérot effects

The Fabry-Pérot effects originate from multiple reflections and subsequent interference of an EM wave within a slab of a dielectric material. Figure 4.3 shows a simulated illustration of the Fabry-Pérot effect in the time domain, where the propagating wave is a picosecond pulse incident at a normal angle with respect to a dielectric slab. The transmitted pulse amplitude is plotted against the time and sample thickness. Clearly, the Fabry-Pérot effect in the time domain is simply a superposition of the main pulse along with the attenuated reflected pulses. In addition, the temporal separation of pulses is larger for a thicker sample, and the temporal separation of any two consecutive pulses, Δt, is approximately equal to the time that a pulse spends in propagating

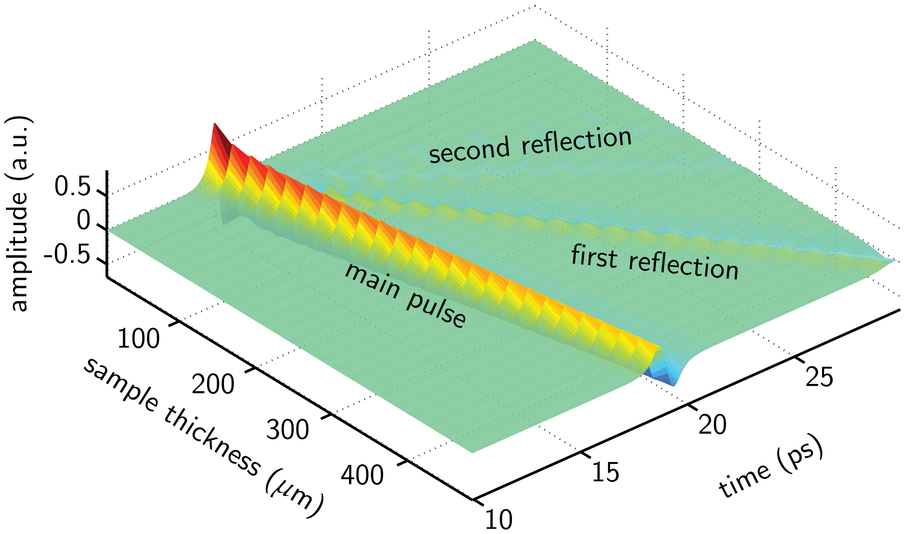


Figure 4.3. Simulated Fabry-Pérot effect in the time domain. The slab material has a complex refractive index of $\hat{n}_b = n_b - j\kappa_b = 3.42 - 0.1\omega j/(2\pi \times 10^{12})$. The signal is simulated on the basis of a PCA current density model for T-ray generation, discussed in Appendix B. Only two reflections after the main pulse are observable here.

back and forth the slab, or

$$\Delta t = \frac{n_g(\omega) l}{c}, \tag{4.7}$$

where the group refractive index $n_g(\omega) = n_b(\omega) + \omega(\partial n_b(\omega)/\partial\omega)$.

The plot in Figure 4.4 is the magnitude of the T-ray signal in the frequency domain. The apparent fringes are caused by the Fabry-Pérot reflections. As the sample thickness increases, the fringes are more compressed. The frequency separation between any adjacent maxima or minima of the fringes, or the so-called *free spectral range*, is given by

$$\Delta f = \frac{c}{2n_g(\omega) l}. \tag{4.8}$$

Clearly, from Equations 4.7 and 4.8, $\Delta f = 1/(2\Delta t)$, or the free spectral range is inversely proportional to the temporal separation between reflected pulses.

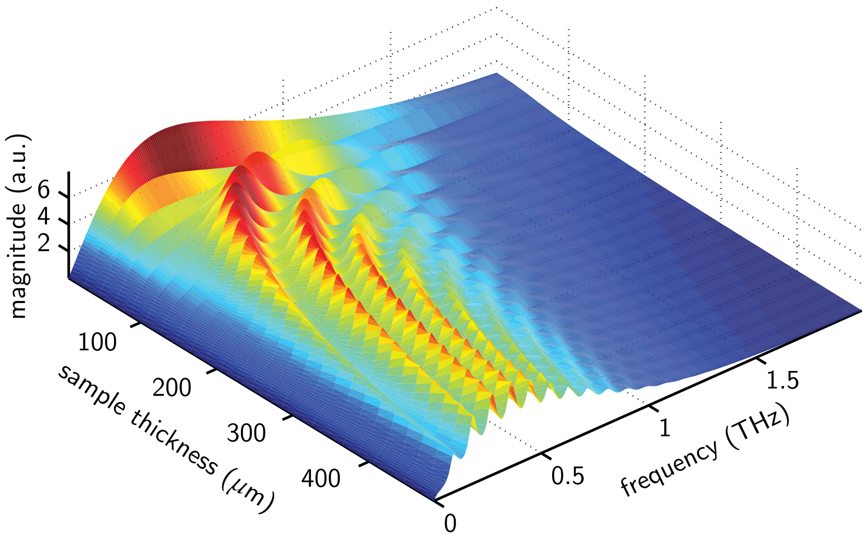


Figure 4.4. Simulated Fabry-Pérot effect in the frequency domain. The spectrum at each sample thickness is a Fourier pair of the signal shown in Figure 4.3. The fringes in magnitude are caused by Fabry-Pérot reflections.

4.4 Transmission-mode THz-TDS

Transmission-mode THz-TDS is the primary method for THz-TDS measurements, mainly due to its simplicity and versatility. This mode of measurement is widely used for T-ray characterisation of materials, including biosamples, semiconductors, dielectrics, and so on.

In this section, several transmission-mode configurations, widely employed in THz-TDS, are reviewed. These configurations include single-sample measurement, dual-sample measurement, multiple-sample measurement, and measurement of a sample in a cell. The methods for deriving the optical constants from the measurement data are given, along with some remarks and pointers to in-depth analysis in the literature.

4.4.1 Single-sample measurement

Single-sample THz-TDS measurement first appeared when Grischkowsky *et al.* (1990) sought dielectrics that are suitable for making T-ray lenses. Samples made of dielectrics preformed into slabs were measured in the transmission mode, and their optical constants were determined using curve fitting to the data. Later on, many iterative methods were proposed for extracting the parameters from the measurement. The method introduced by Duvillaret *et al.* (1996) models an error from the difference between the estimated and measured data at each frequency as a function of the optical constants. The error curve is then approximated by a paraboloid, the global minimum of which determines a correct solution. That solution can be obtained by an iterative numerical procedure. Duvillaret *et al.* (1999) and Dorney *et al.* (2001a) suggested a procedure to estimate the optical constants together with the sample thickness. The process simultaneously determines a set of the optical constants at various guessed thicknesses, and uses the criterion of the total variation of the optical constants to select the correct thickness. In this section, an approximated solution to the parameter extraction problem is derived (Jepsen and Fischer 2005).

Following Equation 4.4, a T-ray signal that passes through a dielectric sample with parallel surfaces at a normal angle of incidence, assuming no reflections, can be expressed in the frequency domain as

$$E_{\mathrm{sam}}(\omega) = \eta \frac{4\hat{n}_s(\omega) n_0}{[\hat{n}_s(\omega) + n_0]^2} \cdot \exp\left\{-j\hat{n}_s(\omega)\frac{\omega l}{c}\right\} \cdot E_0(\omega)\,, \tag{4.9}$$

where η is the transmission factor of free air surrounding the sample; $\hat{n}_s(\omega)$ is the complex refractive index of the sample; n_0 is the refractive index of free air; and l is the propagation length inside the sample, which equals the sample thickness for a normal angle of incidence. In addition to the sample spectrum,

$$E_{\mathrm{ref}}(\omega) = \eta \cdot \exp\left\{-jn_0\frac{\omega l}{c}\right\} \cdot E_0(\omega)\,, \tag{4.10}$$

is the complex frequency spectrum of a reference signal, i.e., a signal measured with the same settings but in the absence of the sample.

The material parameter extraction process requires these two spectra, $E_{\mathrm{sam}}(\omega)$ and $E_{\mathrm{ref}}(\omega)$, which are Fourier transformed from time-domain measurements. Normalising the sample spectrum by the reference, or $E_{\mathrm{sam}}(\omega)/E_{\mathrm{ref}}(\omega)$, yields the complex transfer function of the sample in the frequency domain:

$$H_0(\omega) = \frac{4\hat{n}_s(\omega) n_0}{[\hat{n}_s(\omega) + n_0]^2} \cdot \exp\left\{-\kappa_s(\omega)\frac{\omega l}{c}\right\} \cdot \exp\left\{-j[n_s(\omega) - n_0]\frac{\omega l}{c}\right\}\,. \tag{4.11}$$

Often, the complex refractive index, $\hat{n}(\omega)$, which is a result of the transmission at air-sample interfaces, is approximated to a real index, $n(\omega)$, giving

$$H(\omega) = \frac{4n_s(\omega) n_0}{[n_s(\omega) + n_0]^2} \cdot \exp\left\{-\kappa_s(\omega)\frac{\omega l}{c}\right\} \cdot \exp\left\{-j[n_s(\omega) - n_0]\frac{\omega l}{c}\right\}\,. \tag{4.12}$$

Taking the argument and logarithm of the simplified transfer function gives, respectively,

$$\angle H(\omega) = -[n_s(\omega) - n_0]\frac{\omega l}{c}\,, \tag{4.13a}$$

$$\ln|H(\omega)| = \ln\left[\frac{4n_s(\omega)n_0}{(n_s(\omega)+n_0)^2}\right] - \kappa_s(\omega)\frac{\omega l}{c}\,. \tag{4.13b}$$

The optical constants of the sample can be deduced from Equations 4.13a and 4.13b as

$$n_s(\omega) = n_0 - \frac{c}{\omega l}\angle H(\omega)\,, \tag{4.14a}$$

$$\kappa_s(\omega) = \frac{c}{\omega l}\left\{\ln\left[\frac{4n_s(\omega)n_0}{(n_s(\omega)+n_0)^2}\right] - \ln|H(\omega)|\right\}\,. \tag{4.14b}$$

Note that $H(\omega)$ is determined from the measurements of the sample and reference, and calculating $\angle H(\omega)$ requires a special phase unwrapping process, which is discussed in Section 4.6.3.

It is obvious that the optical constants extracted using Equation 4.14 are approximations, and also Fabry-Pérot effects are not taken into account in the calculation. A more precise method for determining the optical constants does not offer a closed-form solution. Accurate estimation requires an iterative approach as mentioned earlier.

4.4.2 Dual-sample measurement

It is possible that the sample with a different thickness be measured instead of the free path as a reference, as depicted in Figure 4.5. A measurement similar to this setting was first carried out by Auston and Nuss (1988). In their experiment, a T-ray Čherenkov cone was produced inside a $LiTaO_3$ crystal by using a focused optical pump beam via the optical rectification effect. The probe beam running in parallel to the pump beam resolved a full T-ray waveform through EO sampling. By changing the lateral distance between the probe and pump beams, two waveforms travelling through different thicknesses of the crystal were measured. A Fourier analysis and an analysis of wave propagation of the two waveforms yielded the complete absorption and dispersion spectra of $LiTaO_3$ between 0 and 4 THz. For a free-space THz-TDS system, the dual-sample measurement was employed in, e.g., Zhang *et al.* (2001) or Naftaly and Miles (2005), where, for the latter, each material was represented by five samples with different thicknesses; any two from five measurements were used to estimate the optical constants.

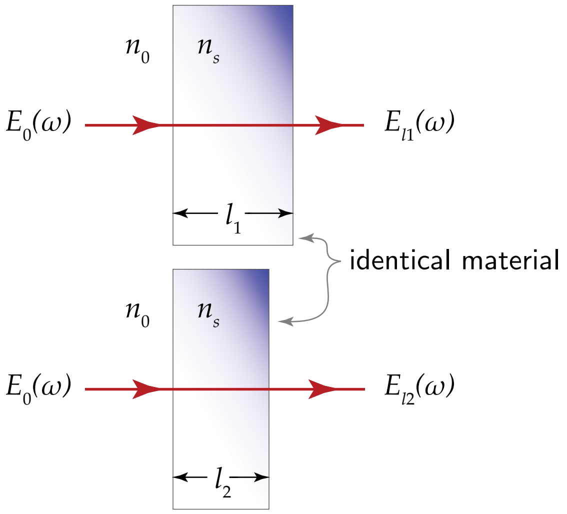


Figure 4.5. Geometry for dual-sample measurement. Two nominally identical samples, with different thicknesses l_1 and l_2, are measured separately for their complex responses. In this case the reference measured with free air is not necessary, because one of the two sample measurements represents the reference.

The derivation of the optical constants from the two sample measurements is as follows. According to Equation 4.4, the spectrum measured from the thick sample with a thickness l_1 is given by

$$E_{l_1}(\omega) = \eta \frac{4\hat{n}_s(\omega)n_0}{[\hat{n}_s(\omega)+n_0]^2} \cdot \exp\left\{-j\hat{n}_s(\omega)\frac{\omega l_1}{c}\right\} \cdot \mathrm{FP}_{l_1}(\omega) \cdot E_0(\omega) \; . \quad (4.15a)$$

Likewise, the spectrum for the thin sample with a thickness l_2, where $l_2 < l_1$, is

$$\begin{aligned} E_{l_2}(\omega) &= \eta \frac{4\hat{n}_s(\omega)n_0}{[\hat{n}_s(\omega)+n_0]^2} \cdot \exp\left\{-j\hat{n}_s(\omega)\frac{\omega l_2}{c}\right\} \\ &\quad \cdot \exp\left\{-jn_0\frac{\omega(l_1-l_2)}{c}\right\} \cdot \mathrm{FP}_{l_2}(\omega) \cdot E_0(\omega) \; , \end{aligned} \quad (4.15b)$$

where the exponential of n_0 is for the free space difference between the two thicknesses. By ignoring the Fabry-Pérot effect, the transfer function determined from $H(\omega) = E_{l_1}(\omega)/E_{l_2}(\omega)$ reads as

$$H(\omega) = \exp\left\{-j\hat{n}_s(\omega)\frac{\omega(l_1-l_2)}{c}\right\} \cdot \exp\left\{jn_0\frac{\omega(l_1-l_2)}{c}\right\} \; . \quad (4.16)$$

The phase and magnitude of the transfer function are therefore

$$\angle H(\omega) = -[n_s(\omega)-n_0]\frac{\omega(l_1-l_2)}{c} \; , \quad (4.17a)$$

$$\ln|H(\omega)| = -\kappa_s(\omega)\frac{\omega(l_1-l_2)}{c} \; . \quad (4.17b)$$

Hence, the optical constants can be determined from (Zhang *et al.* 2001)

$$n_s(\omega) = n_0 - \frac{c}{\omega(l_1 - l_2)} \angle H(\omega) \, , \tag{4.18a}$$

$$\kappa_s(\omega) = -\frac{c}{\omega(l_1 - l_2)} \ln |H(\omega)| \, . \tag{4.18b}$$

It is obvious that for this measurement setting, no approximation to the transmission coefficient is required, as the coefficient exists in both measurements and so is cancelled out. In addition, the setting permits simultaneous calculation of $n_s(\omega)$ and $\kappa_s(\omega)$, i.e. $\kappa_s(\omega)$ is not a function of $n_s(\omega)$ in Equation 4.18b. In such a case, the contribution of uncertainty in $n_s(\omega)$ towards $\kappa_s(\omega)$ vanishes. However, the thickness uncertainty and the imperfections of an additional sample could contribute to the uncertainty in the estimated optical constants.

4.4.3 Multiple-sample measurement

A series of nominally identical samples with multiple thicknesses can be used for characterisation of their optical properties. An early experiment using this technique was performed by Cheung and Auston (1986). In that work, an expanding T-ray Čherenkov cone generated within a $LiTaO_3$ crystal was coherently detected by a PCA receiver, attached to the crystal in parallel with the pump beam. A number of measurements were made, each with the unique distance between the pump beam and the antenna. As a result, a series of time-resolved T-ray signals were obtained as a function of the thickness of the $LiTaO_3$ crystal. From these results, the optical constants of $LiTaO_3$ were deduced. Exploiting a free-space THz-TDS system, Kindt and Schmuttenmaer (1996) employed a cell with a variable path length in the characterisation of polar liquids. A liquid sample filling a PE bag was placed in the cell composed of two transparent windows, one of which was movable with a micrometre resolution. THz-TDS measurements were made with different thicknesses of the liquid, yielding a set of T-ray signals as a function of the thickness.

In the estimation of the optical constants, a series of the signals are Fourier transformed into the magnitude and phase spectra. The magnitude or phase at each frequency component is plotted against the thickness. A linear regression of the plot has a slope related to an optical constant via an explicit function. A derivation for this function is described as follows. From Section 4.3, the measured spectrum for a single dielectric slab without Fabry-Pérot effects can be expressed as

$$\begin{aligned} E(\omega,l) &= \eta\frac{4\hat{n}_s(\omega)n_0}{[\hat{n}_s(\omega)+n_0]^2} \\ &\quad \cdot \exp\left\{-j\hat{n}_s(\omega)\frac{\omega l}{c}\right\}\cdot\exp\left\{-jn_0\frac{\omega(l_{\max}-l)}{c}\right\}\cdot E_0(\omega)\,. \end{aligned} \tag{4.19}$$

The second exponential compensates the delay in free space in the case of any sample that is not the thickest sample, and $l_{\max}$ is the maximum thickness among the samples. Notice the dependency of the spectrum on the thickness l. Taking the argument of $E(\omega,l)$ yields

$$\angle E(\omega,l) = \angle\left[\eta\frac{4\hat{n}_s(\omega)n_0}{[\hat{n}_s(\omega)+n_0]^2}\cdot E_0(\omega)\right]-n_s(\omega)\frac{\omega l}{c}-n_0\frac{\omega(l_{\max}-l)}{c}\,, \tag{4.20a}$$

and taking the logarithm of the absolute of $E(\omega,l)$ yields

$$\ln|E(\omega,l)| = \ln\left|\eta\frac{4\hat{n}_s(\omega)n_0}{[\hat{n}_s(\omega)+n_0]^2}\cdot E_0(\omega)\right|-\kappa_s(\omega)\frac{\omega l}{c}\,. \tag{4.20b}$$

Equation 4.20a or 4.20b is equivalent to plotting of the measured phase or magnitude of the signal at each frequency against the sample thickness. Assuming that the plots are linear, their slopes can be determined analytically by taking the first derivative of Equations 4.20a and 4.20b with respect to the thickness, or

$$\begin{aligned} M_{\text{arg}} &= \frac{\partial\angle E(\omega,l)}{\partial l} \\ &= -(n_s(\omega)-n_0)\frac{\omega}{c}\,, \end{aligned} \tag{4.21a}$$

$$\begin{aligned} M_{\text{abs}} &= \frac{\partial\ln|E(\omega,l)|}{\partial l} \\ &= -\kappa_s\frac{\omega}{c}\,, \end{aligned} \tag{4.21b}$$

where M_{arg} and M_{abs} are the slopes for the phase and magnitude plots, respectively. From Equation 4.21 the optical constants can be deduced as follows:

$$n_s(\omega) = n_0 - M_{\text{arg}}\frac{c}{\omega}\,, \tag{4.22a}$$

$$\kappa_s(\omega) = -M_{\text{abs}}\frac{c}{\omega}\,. \tag{4.22b}$$

It can be seen that, as in case of dual-sample measurement, the multiple-sample measurement removes from the estimation the transmission effects at dielectric interfaces, which are not relevant to a change in the sample thickness.

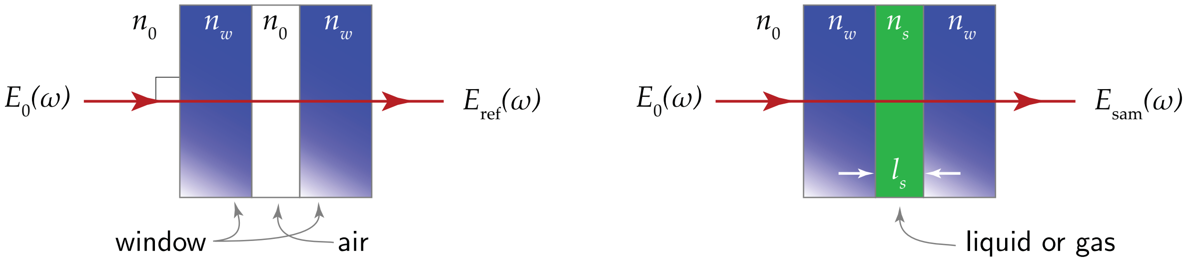
4.4.4 Measurement of a sample in a cell

Spectroscopy of a liquid or gas often requires a cell to confine the sample. In order to allow T-rays to probe the sample effectively, the cell is equipped with windows transparent to T-rays. The windows must be thick enough so that the reflections therein are well separated from the main pulse, and thus can be removed by post-processing. A number of THz-TDS measurements are carried out with a cell as a sample holder. Here the analysis of wave propagation through the cell is reviewed.

Figure 4.6 illustrates the propagation of T-rays through the cell during the measurements of the sample and reference data. By analysing the propagation geometry, we know that the sample spectrum is represented by

$$\begin{aligned} E_{\text{sam}}(\omega) \;=\;& \tau_{aw}\tau_{ws}\tau_{sw}\tau_{wa}\cdot \mathrm{FP}^2_{aws}(\omega)\cdot \mathrm{FP}_{wsw}(\omega) \\ & \cdot \exp\left[-j\hat{n}_s\frac{\omega l_s}{c}\right]\cdot \exp\left[-jn_w\frac{\omega l_w}{c}\right]\cdot E_0(\omega)\,, \end{aligned} \tag{4.23a}$$

where the subscripts a, s, and w are for air, sample, and window, respectively, and l_w is the total thickness of the two identical windows. The subscript of a transmission coefficient implies the propagation direction, e.g., τ_{aw} for the propagation from air to the window, τ_{ws} for the propagation from the window to the sample, etc. The subscript of a Fabry-Pérot term implies the location where the reflection takes place, e.g., $\mathrm{FP}_{aws}(\omega)$ for the reflection in a window, with air on one side and the sample on the other.



(a) reference measurement

(b) sample measurement

Figure 4.6. Geometry for sample-in-cell measurement. The cell is composed of two transparent windows, which maximise the transmission of T-rays. In a regular measurement, the propagation direction of T-rays is normal to the cell's surface. (a) T-rays propagate through an empty cell as the reference. (b) T-rays propagate through an identical cell, filled with the sample.

Again from Figure 4.6, it can be shown that the reference spectrum, measured when the sample is substituted by air, is

$$\begin{aligned} E_{\mathrm{ref}}(\omega) &= \tau_{aw}^2 \tau_{wa}^2 \cdot \mathrm{FP}_{awa}^2(\omega) \cdot \mathrm{FP}_{waw}(\omega) \\ &\quad \cdot \exp\left[-jn_0 \frac{\omega l_s}{c}\right] \cdot \exp\left[-jn_w \frac{\omega l_w}{c}\right] \cdot E_0(\omega)\,. \end{aligned} \tag{4.23b}$$

The transfer function of the sample can be resolved from Equations 4.23a and 4.23b using $H_0(\omega) = E_{\mathrm{sam}}(\omega)/E_{\mathrm{ref}}(\omega)$. By assuming that the reflections are removed from the sample and reference data, the transfer function is given by

$$H_0(\omega) = \frac{\tau_{ws}\tau_{sw}}{\tau_{wa}\tau_{aw}} \cdot \exp\left[-j(\hat{n}_s - n_0)\frac{\omega l_s}{c}\right]\,. \tag{4.24}$$

In a similar manner to the transfer function model of a single dielectric slab, this model requires an iterative calculation in order to extract the optical constants. However, for a general case, the transmission coefficients vanish by an approximation $\tau_{wa} \approx \tau_{ws}$ and $\tau_{aw} \approx \tau_{sw}$. Thus,

$$H(\omega) \approx \exp\left[-j(\hat{n}_s - n_0)\frac{\omega l_s}{c}\right]\,, \tag{4.25}$$

and then the optical constants can be found from

$$n_s(\omega) = n_0 - \frac{c}{\omega l_s}\angle H(\omega)\,, \tag{4.26a}$$

$$\kappa_s(\omega) = -\frac{c}{\omega l_s}\ln|H(\omega)|\,. \tag{4.26b}$$

The estimation of the optical properties with this measurement setting heavily relies on approximations. Using the iterative parameter estimation can improve the accuracy of the estimated optical constants.

4.5 Reflection-mode THz-TDS

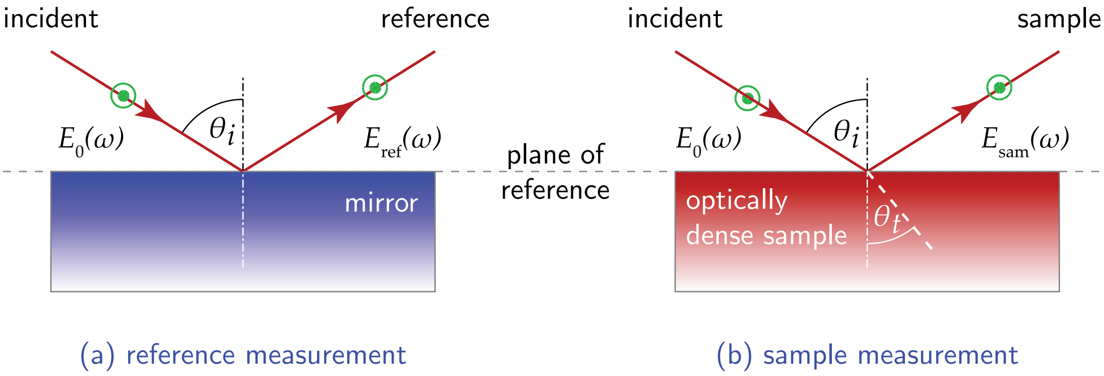
A transmission-mode measurement cannot be accomplished for optically dense samples due to a high energy loss within the samples. Reflection-mode THz-TDS can circumvent this limitation, as it does not require wave propagation through a bulk material. Jepsen and Fischer (2005) established a relation between the dynamic range of an absorption coefficient and the THz-TDS system characteristics. In transmission-mode THz-TDS the dynamic range is fundamentally limited by the SNR of the system. By contrast, in reflection-mode THz-TDS the dynamic range is only limited by the reproducibility of the signals. However, one must be aware that material properties extracted from a reflection-mode measurement do not necessarily represent the bulk properties of a material. This caveat was realised by Howells and Schlie (1996), commenting on the deficiency of reflection-mode systems in probing semiconductors. Due to Fermi level pinning at the semiconductor's surface, reflection data lead to an underestimate of carrier concentration.

We shall consider two THz-TDS reflection-mode configurations: (i) single-reflection and (ii) double-reflection. The single-reflection measurement requires two scans, one for the reference and the other for the sample. The double-reflection measurement requires only one scan, where the reference and sample data are combined into a series of reflections. Sections 4.5.1 and 4.5.2 describe both types of measurements in detail.

4.5.1 Single-reflection measurement

The single-reflection measurement was first demonstrated by Auston and Cheung (1985). A sample, such as Sb-doped Ge, was in optical contact with a nonlinear crystal, inside which T-rays were generated and detected coherently. A TE-mode Čerenkov T-ray wavefront, as a result of an optical pump pulse propagation, was incident at the crystal-sample interface. The wave reflected back into the crystal was detected by an optical probe pulse. This reflected wave was considered as the sample signal. The reference signal was measured with an identical setting, but with the sample being replaced with a material with a known T-ray response. The confinement of the T-ray field inside a crystal causes multiple internal reflections, which limit the longest temporal scan, and ultimately limit the frequency resolution of the system.

Later on, Howells and Schlie (1996) characterised a thick sample of undoped InSb by using TM-polarised T-rays in a free-space reflection geometry. The measured reflectance of the sample, relative to the reflectance of a silver-coated mirror, was fitted to a Drude model to find the sample's carrier mobility and concentration. A similar measurement setting was also utilised with a dual-layered sample (Hashimshony *et al.* 2001). A thin epitaxial layer of an extrinsic semiconductor grown on a semiconductor substrate was probed by a TM-polarised wave. The measured response, in comparison with that of the aluminium, was used to determine the doping concentration of the epitaxy layer, resistivity, and thickness, via fitting of a Drude model.



(a) reference measurement

(b) sample measurement

Figure 4.7. Geometry for single-reflection measurement. (a) For the reference measurement, a mirror is placed at a reference position, where it reflects nearly all incident T-ray energy into a receiver. (b) The sample measurement is similar to the reference, excepting that the mirror is replaced with a sample. The angle, θ_i, and polarisation, TE or TM, of the incident wave influence the measurable response of the sample.

As mentioned earlier, reflection-mode THz-TDS is very suitable for the measurement of optically dense samples, such as highly doped semiconductors or human tissue. Jeon and Grischkowsky (1998a) measured highly doped n-type GaAs and Si samples by using a reflection configuration with a TM wave. The reflection configuration allowed measuring of the absorption coefficient as high as 1800 cm^{-1}. Pickwell *et al.* (2004b) used this reflection configuration for *in vivo* measurement of human skin, which, altogether with the underlying soft tissue, is barely penetrable by T-rays. An intermediate layer of quartz was presented on top of either the sample (skin) or reference (air) to regulate the surface morphology.

Essentially, for the single-reflection setting, two THz-TDS scans are required, one for the reference and the other for the sample, as shown in Figure 4.7. The first scan is made with a metallic mirror, made of ,e.g., gold or aluminium, which has a nearly perfect reflectance to T-rays (Gatesman *et al.* 1995). The mirror positioned at the plane of reference reflects the T-ray field into a receiver, and the resulting signal is taken as the reference. With no change to the system arrangement, a sample replacing the mirror is scanned to obtain the sample signal.

For a metallic surface, the transmission coefficient, τ, approaches zero, and the reflection coefficient, ρ, approaches minus unity. Therefore, from Equation 4.6, the reflected signal for a metallic surface is expressed as

$$E_{\text{ref}}(\omega) \quad = \quad -E_0(\omega)\,. \tag{4.27}$$

For a non-metallic sample, by assuming no reflection from the back surface, the signal reflected from the front surface is given by

$$E_{\mathrm{sam}}(\omega) \quad = \quad \hat{\rho} E_0(\omega) \, , \tag{4.28}$$

where $\hat{\rho}$ is the reflection coefficient at the air-sample interface. The transfer function of a sample is defined as

$$H(\omega) = A \exp(j\phi) = \frac{E_{\mathrm{sam}}(\omega)}{E_{\mathrm{ref}}(\omega)} = -\hat{\rho} \, . \tag{4.29}$$

The transfer function is simply the negative of the reflection coefficient. Because the reflection coefficient $\hat{\rho}$ is a function of the complex refractive index of a sample, it is possible to extract this complex value from the measured transfer function. However, the reflection coefficient is also dependent on the polarisation of the incident wave and the angle of incidence. Therefore, the parameter extraction is established for three cases: (i) for normal angle of incidence, (ii) for TE waves, and (iii) for TM waves.

Normal angle of incidence

If the propagation direction of the incident wave is normal to the surface, from Equations 4.1b and 4.29, the transfer function of a sample is given by

$$\begin{aligned} H(\omega) &= -\hat{\rho} \\ &= \frac{\hat{n}_s(\omega) - 1}{\hat{n}_s(\omega) + 1} \\ &= \frac{n_s(\omega) - j\kappa_s(\omega) - 1}{n_s(\omega) - j\kappa_s(\omega) + 1} \, . \end{aligned} \tag{4.30}$$

Given that the measured transfer function is

$$H_{\mathrm{meas}}(\omega) \quad = \quad A \exp(j\phi) \, , \tag{4.31}$$

the optical constants can be expressed as (Jepsen and Fischer 2005)

$$n_s(\omega) \quad = \quad \frac{1 - A^2}{1 + A^2 - 2A\cos\phi} \, , \tag{4.32a}$$

$$\kappa_s(\omega) \quad = \quad \frac{-2A\sin\phi}{1 + A^2 - 2A\cos\phi} \, . \tag{4.32b}$$

TE wave

In the case that the incident wave is perpendicularly polarised, from Equations 4.1b and 4.29 the transfer function of a sample is given by

$$\begin{aligned} H(\omega) &= -\hat{\rho}_\sigma \\ &= -\frac{\cos\theta_i - \hat{n}_s\cos\hat{\theta}_t}{\cos\theta_i + \hat{n}_s\cos\hat{\theta}_t} , \end{aligned} \tag{4.33}$$

where the refraction angle now becomes complex (Born and Wolf 1999), and can be expressed as

$$\hat{\theta}_t = \arcsin\left(\frac{\sin\theta_i}{\hat{n}_s}\right) . \tag{4.34}$$

Finding a closed-form solution to the optical constants directly is intricate. Thus, the complex index of refraction in Equation 4.33 is substituted by the complex refraction angle in the form of

$$\hat{n}_s = \frac{\sin\theta_i}{\sin\hat{\theta}_t} . \tag{4.35}$$

Therefore, the transfer function is changed to (Born and Wolf 1999)

$$H(\omega) = \frac{\sin(\theta_i - \hat{\theta}_t)}{\sin(\theta_i + \hat{\theta}_t)} . \tag{4.36}$$

Solving Equation 4.36 for the complex refraction angle gives

$$\hat{\theta}_t = -\theta_i + \arctan\left[\frac{\sin 2\theta_i}{A\exp(j\phi) + \cos 2\theta_i}\right] . \tag{4.37}$$

The optical constants can be found by substituting Equation 4.37 into Equation 4.35.

TM wave

In the case that the incident wave is TM, from Equations 4.2b and 4.29 a sample transfer function is given by

$$\begin{aligned} H(\omega) &= -\hat{\rho}_\pi \\ &= \frac{\hat{n}_s\cos\theta_i - \cos\hat{\theta}_t}{\hat{n}_s\cos\theta_i + \cos\hat{\theta}_t} . \end{aligned} \tag{4.38}$$

Again, such an expression hinders deriving an explicit solution for the optical constants. The transfer function can be re-expressed in terms of the angles only, or

$$H(\omega) = \frac{\sin 2\theta_i - \sin 2\hat{\theta}_t}{\sin 2\theta_i + \sin 2\hat{\theta}_t} . \tag{4.39}$$

Solving Equation 4.39 for the complex refraction angle gives

$$\hat{\theta}_t = -\frac{1}{2}\arcsin\left[\frac{\sin 2\theta_i(A\exp(j\phi)-1)}{A\exp(j\phi)+1}\right] . \tag{4.40}$$

The optical constants can be found by substituting Equation 4.40 into Equation 4.35.

In the parameter extraction process, the sample thickness is not required, and thus the uncertainty in the thickness measurement does not contribute to the uncertainty in the optical constants. This should be an advantage of the single-reflection measurement. However, a very high precision for the mirror/sample positioning is crucial. Between two scans, the reflecting plane of a mirror and a sample must be as close as possible in the order of microns (Jeon and Grischkowsky 1998a). Otherwise, the estimated optical constants of a sample become inaccurate. A simple approach to precise positioning of a mirror/sample is suggested. At a particular reflecting plane any surface reflects the incident wave into a detector at the maximum level. Therefore, by moving a reflecting surface in the normal direction until the pulse amplitude is maximised, the plane of reference is found. By practising this procedure with both a metallic mirror and a sample, the reflecting surfaces of both materials are aligned.

Another means to circumvent difficulties in positioning of a mirror/sample is to discard the phase information available from the measurement, i.e., discard ϕ in Equation 4.31 (Liu *et al.* 2006). Instead, the phase is obtained through the Kramers-Kronig relation,

$$\phi(\omega) = \frac{2\omega}{\pi}\mathcal{P}\int_0^\infty \frac{\ln A(\hat{\omega})}{\hat{\omega}^2-\omega^2}d\hat{\omega} , \tag{4.41}$$

where $\mathcal{P}$ is Cauchy principal value. This calculated phase can be used in conjunction with the measured magnitude A to determine the optical constants, e.g., in Equation 4.32. However, this method is a very rough approximation and should be spared for only special circumstances.

4.5.2 Double-reflection measurement

A double-reflection measurement, or also known as self-referenced geometry, was presented by Thrane *et al.* (1995) and later expanded in Rønne *et al.* (1997) in experiments with liquid water, which has a very high absorption to T-rays. A cell containing liquid water was equipped with a flat window made of high-resistivity silicon, highly transparent to T-rays. As the T-ray beam is incident on the window at a normal angle of incidence, it is reflected twice—the first reflection is from the air-window interface and the second reflection is from the window-water interface. The first and second reflections are used as the reference and sample signals in the calculation of the optical constants of water. A diagram of this reflection setting is shown in Figure 4.8.

A window used in this type of measurement must have two special characteristics as follows. First of all, the window must be thick enough, so that a series of T-ray pulses reflected at the window interfaces are well separated in the time domain. In that case the reference and sample pulses can be extracted through temporal gating. Otherwise, overlapping among reflected pulses prevents finding of the reference and sample pulses. The second characteristic the window must have is transparency to T-rays. Its absorption must be negligible compared to the absorption of the sample. The window transparency facilitates an estimation of the optical constants. Apart from the window, the sample should satisfy the assumption of a semi-infinite slab. This requirement arises in order to avoid reflections from the back surface of the sample, which could interfere with the sample and reference signals. Either an optically dense or a thick sample is adequate, because an optically dense sample will absorb almost all traversing energy, while a thick sample will provide sufficient delay for pulses reflected from the back surface.

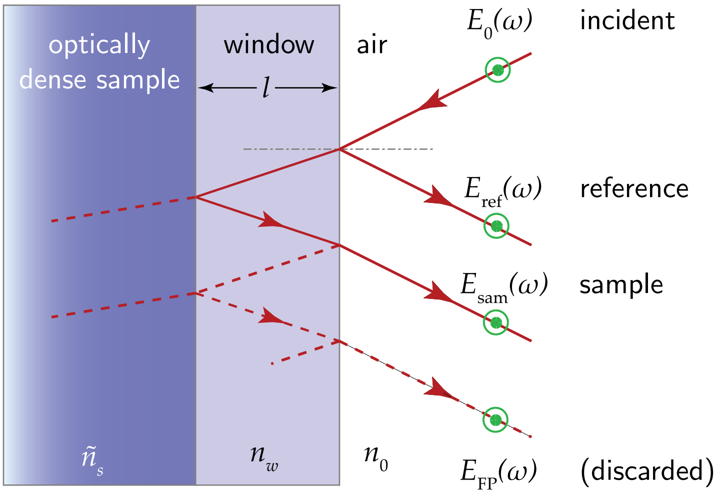


Figure 4.8. Geometry for double-reflection measurement. The first two reflections occur at the air-window and window-sample interfaces. These reflections are temporally separated and can be used as the reference and sample signals, respectively. The tilting of the incident beam is exaggerated.

For the case that the direction of the incident wave is normal to the window surface, a derivation for the optical constants of the sample is as follows. The reference signal is taken from the first reflection, which occurs at the air-window interface. Therefore, according to Equation 4.6, the reference spectrum is expressed as

$$E_{\text{ref}}(\omega) = \rho_{aw} E_0(\omega) , \tag{4.42}$$

where ρ_{aw} is the reflection coefficient for the air-window interface, and

$$\rho_{aw} = \frac{1 - n_w}{1 + n_w} . \tag{4.43}$$

Note that the absorption by the window is negligible, so only the refractive index of the window n_w appears in the reflection coefficient. The sample signal is taken from the second reflection, which occurs at the window-sample interface, and thus its spectrum is given by

$$E_{\text{sam}}(\omega) = \tau_{aw} \tau_{wa} \hat{\rho}_{ws} \exp\left[-2jn_w \frac{\omega l}{c}\right] E_0(\omega) , \tag{4.44}$$

where l is the window thickness, and τ_{aw} and τ_{wa} are the transmission coefficients for the air-window and window-air interfaces, respectively. Hence,

$$\tau_{aw} = \frac{2}{1 + n_w} , \tag{4.45a}$$

$$\tau_{wa} = \frac{2n_w}{1 + n_w} . \tag{4.45b}$$

The reflection coefficient $\hat{\rho}_{ws}$ is for the window-sample interface, and is given by

$$\hat{\rho}_{ws} = \frac{n_w - \hat{n}_s}{n_w + \hat{n}_s} , \tag{4.46}$$

where $\hat{n}_s = n_s - j\kappa_s$ is the complex refractive index of the sample. The transfer function is resolved by normalising the sample spectrum with the reference, or

$$\begin{aligned} H(\omega) &= \frac{E_{\text{sam}}(\omega)}{E_{\text{ref}}(\omega)} \\ &= \frac{\tau_{aw}\tau_{wa}}{\rho_{aw}} \exp\left[-2jn_w \frac{\omega l}{c}\right] \frac{n_w - \hat{n}_s}{n_w + \hat{n}_s} . \end{aligned} \tag{4.47}$$

Rearranging Equation 4.47 gives

$$\begin{aligned} \frac{n_w - \hat{n}_s}{n_w + \hat{n}_s} &= \frac{\rho_{aw}}{\tau_{aw}\tau_{wa}} \exp\left[2jn_w \frac{\omega l}{c}\right] H(\omega) \\ &= A \exp(i\phi) . \end{aligned} \tag{4.48}$$

The transfer function $H(\omega)$ can be found from the measurement, and all the parameters on the right-hand side of the above equation are presumably known. The optical constants can then be determined by solving Equation 4.48 (Thrane *et al.* 1995), giving

$$n_s(\omega) = \frac{n_w(1-A^2)}{1+A^2-2A\cos\phi} \ , \tag{4.49a}$$

$$k_s(\omega) = \frac{-2n_w A\sin\phi}{1+A^2-2A\cos\phi} \ . \tag{4.49b}$$

Thus, the optical constants of a sample can be derived from the measurement, provided that the optical properties of the window are well known. Note that the treatment of the optical constants presented here is for a normal-incidence geometry. For an arbitrary angle of incidence and angle of polarisation, an accurate treatment can be found in Jepsen *et al.* (2007).

4.6 Miscellaneous signal processing for THz-TDS

In this section, miscellaneous signal processing techniques, typically applied to THz-TDS data, are briefly discussed. These techniques include signal averaging, signal denoising, phase unwrapping, and removing of Fabry-Pérot effects.

4.6.1 Signal averaging in the time and frequency domains

Random noise in measurement can be reduced by averaging a series of signals obtained from multiple repeated measurements. The averaging is effective in the time domain, but not in the frequency domain (Fischer *et al.* 2005a). This contrast is depicted in Figure 4.9, where the noise floor reduces only in the case of time-domain averaging.

A mathematical explanation of the difference between averaging in the time and frequency domains is as follows. Given that $x_0(t)$ is a noise-free signal from a time-invariant system, and $n(t)$ is a random noise, which has a normal distribution around zero, the signal from the k^{th} measurement is expressed as

$$x_k(t) = x_0(t) + n_k(t) \ . \tag{4.50}$$

The expectation $\langle x\rangle$ is defined for the variable x_k. Since $\langle n_k(t)\rangle = 0$,

$$\langle x_k(t)\rangle = x_0(t) \ . \tag{4.51}$$

The expectation of a measured signal is noiseless. This implies that averaging the signals in the time domain decreases the noise.

Transforming the signal from Equation 4.50 into the frequency domain yields

$$X_k(\omega) = X_0(\omega) + N_k(\omega) , \tag{4.52}$$

where $X_k(\omega)$, $X_0(\omega)$, and $N_k(\omega)$ are the Fourier pairs of $x_k(t)$, $x_0(t)$, and $n_k(t)$, respectively. Analogous to the expectation of a signal in the time domain, the expectation of a spectrum in the frequency domain results in a noise-free spectrum, or

$$\langle X_k(\omega) \rangle = X_0(\omega) . \tag{4.53}$$

Again, it can be inferred that averaging of complex spectra in the frequency domain can reduce noise.

Now consider the magnitude of the spectrum in Equation 4.52, which is

$$|X_k(\omega)| = |X_0(\omega) + N_k(\omega)| . \tag{4.54}$$

The expectation of the magnitude is

$$\langle |X_k(\omega)| \rangle = \langle |X_0(\omega) + N_k(\omega)| \rangle . \tag{4.55}$$

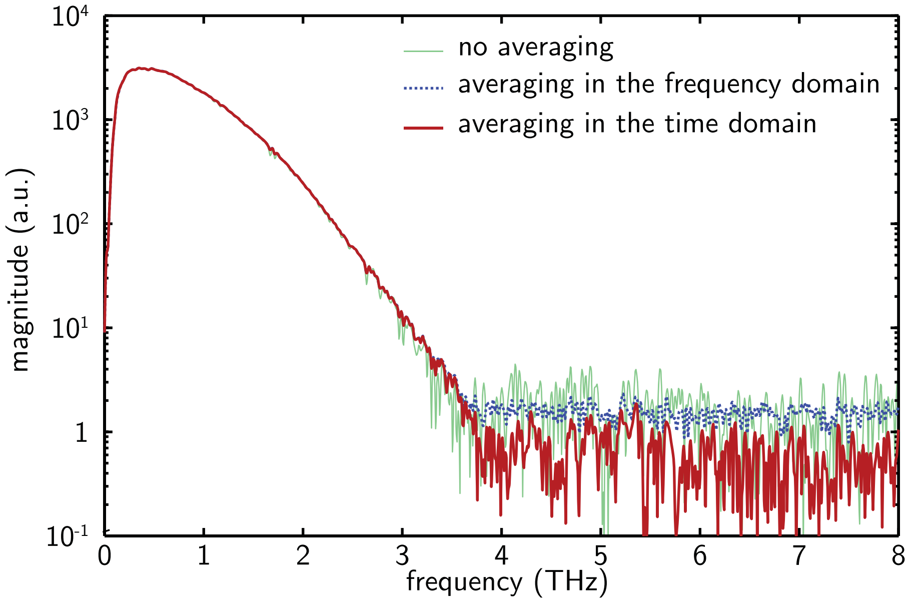


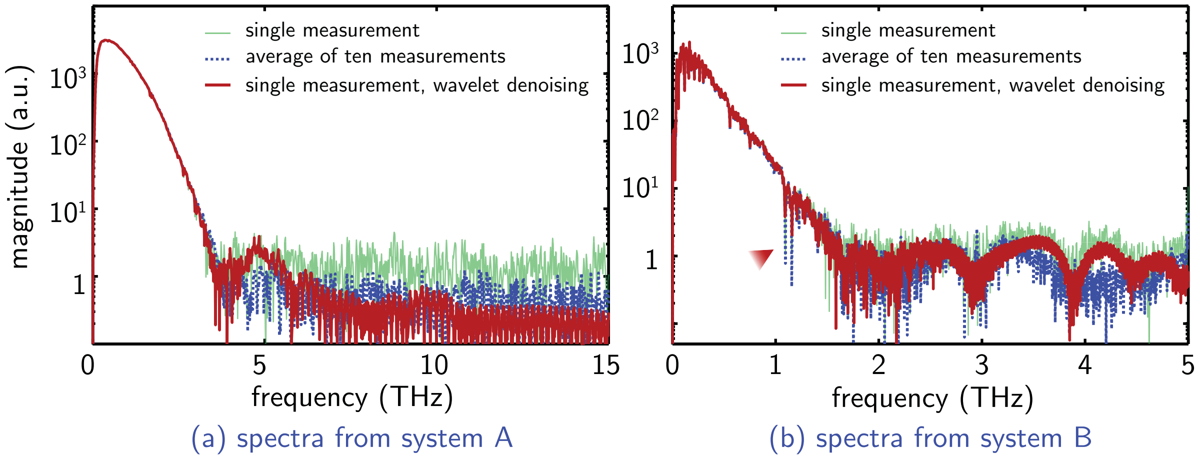
Figure 4.9. Spectra averaged by using different schemes. Time-domain averaging is carried out with ten measurements for the nitrogen atmosphere. Averaging in the frequency domain is performed with the magnitude of the spectra. A proper averaging technique can significantly lower the noise floor, and thus increase the dynamic range and bandwidth.

It is evident that noise persists in the magnitude although the averaging is carried out with an infinite number of measurements. A similar relation is held true when the averaging scheme is introduced at any further stage of the parameter extraction process. In conclusion, applying a nonlinear operator, e.g., an absolute, to the data before taking an average prohibits noise reduction.

4.6.2 Signal denoising

Noise in a T-ray signal could be reduced by a digital signal processing technique such as wavelet denoising. Ferguson and Abbott (2001b), and also Ferguson and Abbott (2001a), tested wavelet denoising by using experimental T-ray signals with simulated additive noise. Wavelet denoising was shown to improve the SNR of the signals, particularly when the signals are strongly absorbed, and the Coiflet order 4 wavelet was found to be optimal with up to 10 dB in noise reduction. Nevertheless, an optimal wavelet is not versatile, and the choice of wavelets is dependent on the T-ray pulse shape. Furthermore, wavelet filtering could introduce artefacts.

Here, wavelet denoising is trialled with real noisy T-ray data in the time domain, and the performance is observed via the improved dynamic range of the signals in the frequency domain. By following the suggestion of Ferguson and Abbott (2001a), Coiflet order 4 is utilised in decomposition of the signals before their detail coefficients are soft-thresholded. Figure 4.10 shows the resultant denoised signals in comparison with the signals obtained from averaging of several measurements. It is obvious that the performance of wavelet denoising on T-ray signals is satisfactory. The noise level of the denoised signals becomes comparable to that of the averages of ten measurements. Fine selection of wavelets could further improve the results. However, artefacts arise by the denoising. Indicated by the arrowhead in Figure 4.10(b), the water-vapour resonances near 1 THz are strongly reduced in magnitude after denoising. The reason for these artefacts is that in the time domain these sharp resonances manifest themselves as oscillations with a high frequency and low amplitude—soft-thresholding reckons any components with these characteristics as noise which needs to be removed. Another remark, which should be noted here, is that denoising of a signal obtained from averaging many measurements shows a relatively small improvement. Further investigation into these issues is required.



(a) spectra from system A

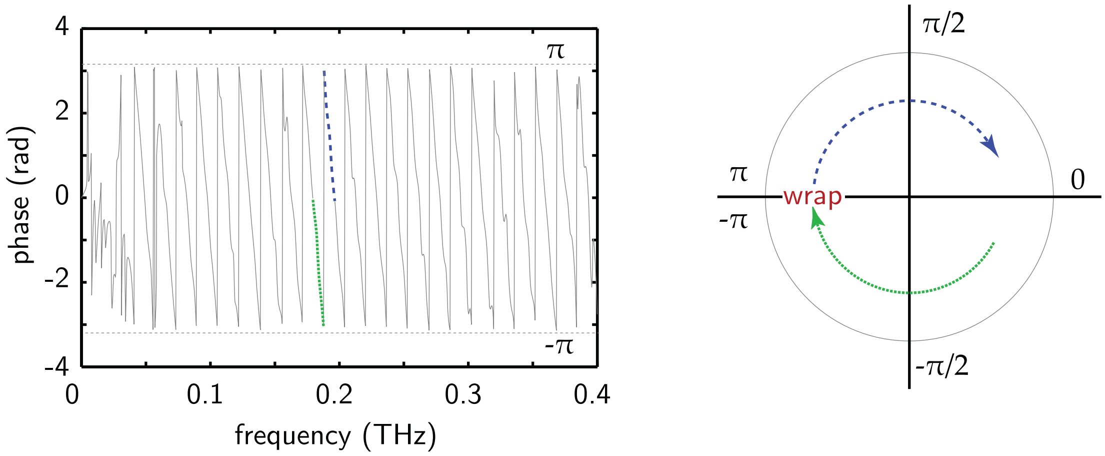
(b) spectra from system B

Figure 4.10. Wavelet-denoised spectra. The signal from a single measurement is denoised using soft-threshold wavelet denoising, and compared with the original signal and the averaged signal. The spectra in (a) and (b) are from two different THz-TDS systems. Furthermore, for (a) the T-ray path is purged with nitrogen, whereas for (b) the path contains ambient water vapour. The arrowhead in (b) indicates the missing resonances caused by wavelet denoising.

Another DSP approach to noise reduction is a spatially variant moving average filter (SVMAF) proposed by Pupeza *et al.* (2007). The SVMAF algorithm establishes the confidence interval of a transfer function via the measurement uncertainty. Intermediate optical constants extracted from the averaged transfer function are smoothed using an average filter over the frequency range. The new smoothed value at any frequency is accepted, if it constructs a new transfer function value that is confined within the confidence interval, and is rejected otherwise.

4.6.3 Phase unwrapping

The phase of a transfer function is a vital part in determining the optical constants. The phase spectrum obtained directly from the angle between the real and imaginary parts of a transfer function essentially wraps around $-\pi$ and π, i.e., whenever the absolute value of the phase is greater than π, it will jump to the opposite polarity by 2π. The jumping causes discontinuity artefacts in the phase. Figure 4.11 illustrates the phase wrapping of T-ray data. A standard means for phase unwrapping can resolve a wrapped phase by adding to or subtracting from the phase spectrum multiples of 2π,



(a) phase of transfer function (b) radian circle

Figure 4.11. Phase wrapping. (a) The phase of the transfer function, obtained from the measurement of a PMMA sample, is wrapped around $-\pi$ and π. (b) The radian circle illustrates the wrapping effect near 0.2 THz in (a).

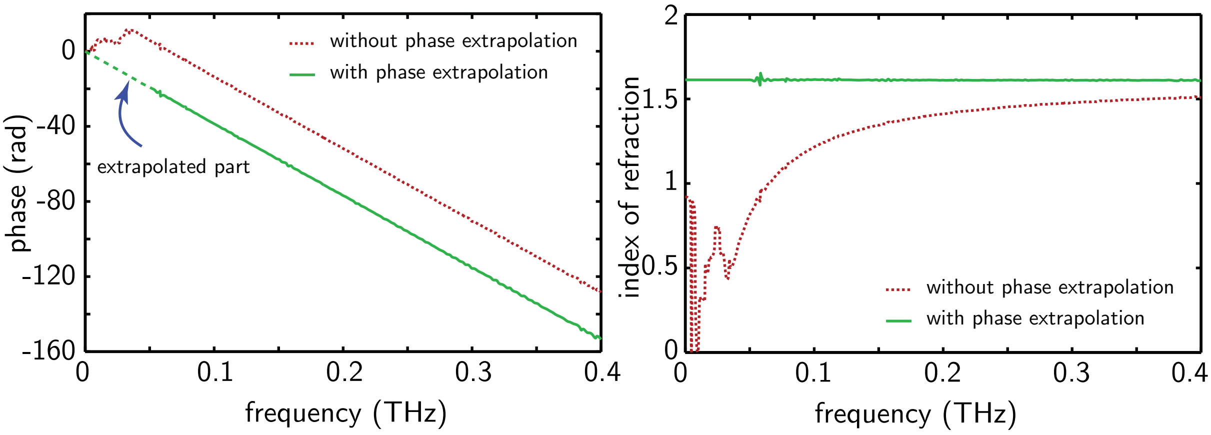
or mathematically

$$\phi(\omega) \quad = \quad \angle H(\omega) - 2\pi M(\omega)\,, \tag{4.56}$$

where M is an integer, counting the number of discontinuities at frequencies lower than ω.

A complication of phase unwrapping arises due to signal noise. For a T-ray spectrum, the reliable part at which the SNR is sufficiently high lies above 0.1 THz or so. At lower frequencies, noise plagues the magnitude and phase data. Unwrapping the phase starting from the noisy part at low frequencies causes error to propagate towards the phase at higher frequencies. An extra step must be taken to avoid false unwrapping.

The adapted phase unwrapping scheme discards the noisy phase at low frequencies, and carries out a normal unwrapping with the reliable phase part. A missing phase profile at low frequencies down to DC is then extrapolated from the unwrapped phase at higher frequencies. In most cases the assumption of a linear phase is sufficient (Duvillaret *et al.* 1996). The whole phase profile is then forced to start at zero. Figure 4.12 shows the comparison between the results obtained from normal unwrapping and extrapolating unwrapping. The refractive index of PMMA obtained from parameter extraction with no phase extrapolation exhibits large error (the correct index of PMMA is ≈1.6).



(a) unwrapped phase (b) index of refraction

Figure 4.12. Phase unwrapping. (a) The phase of the transfer function, obtained from the measurement of a PMMA sample, is unwrapped with and without phase extrapolation. The phase between 0.0 and 0.05 THz is linearly extrapolated from that between 0.05 and 0.1 THz. (b) The index of refraction is determined by using the phases in (a). The literature value of the refractive index of PMMA at T-ray frequencies is around 1.6 (Jin *et al.* 2006).

4.6.4 Removal of Fabry-Pérot effects

From Equation 4.4 in Section 4.3, it can be seen that a T-ray signal measuring a slab of material does not contain only a single primary pulse, but also additional delayed pulses resulting from the reflection at interfaces. The unwanted signals from these reflected pulses are known as Fabry-Pérot or etalon effects. In the parameter extraction process, the primary pulse is sufficient for determining the optical constants, and the reflected pulses are often unnecessary. In most cases, the reflected pulses must be removed from the signal, or otherwise they will corrupt the extracted optical constants. If the primary pulse is well separable from the reflected pulses, time windowing can remove the reflected pulses without difficulties. However, a problem arises if the primary pulse is not resolvable from the reflections. A complicated signal processing technique is required in this case. Here, one such technique is reviewed.

A Fabry-Pérot removal process proposed by Duvillaret *et al.* (1996) is for a single-sample setting, and can be implemented in conjunction with a basic material parameter extraction. Given the complete sample pulse and the reference, the process can cancel out most of the effects regardless of the thickness of a sample. A diagram of the process is shown in Figure 4.13. For the first iteration, the measured transfer function of the

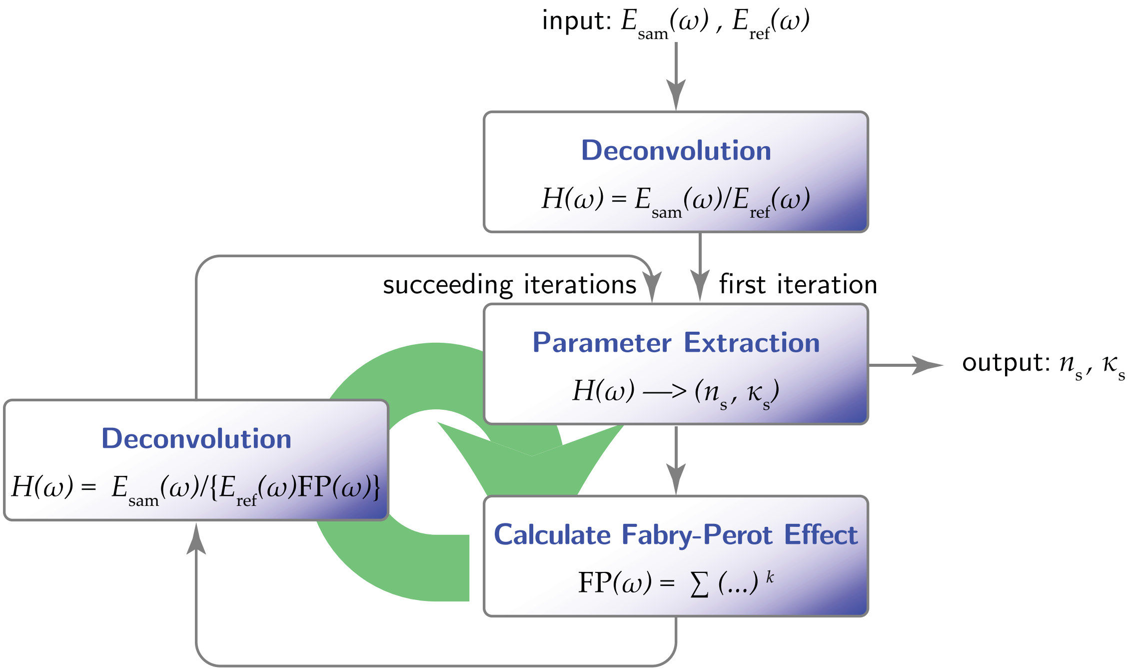


Figure 4.13. Repetitive Fabry-Pérot removal. The method calls for the reference and sample T-ray signals to be used to estimate the intermediate optical constants in the first iteration. Consequently, these constants are used to calculate the Fabry-Pérot term that is then deconvolved from the measured transfer function. The deconvolved transfer function is then fed into the parameter extraction process in the second iteration. The whole process is repeated until convergence is attained.

sample $H(\omega) = E_{\mathrm{sam}}(\omega)/E_{\mathrm{ref}}(\omega)$ is passed into the parameter extraction process directly. The intermediate extracted parameters $n_s(\omega)$ and $\kappa_s(\omega)$ from the first iteration are then used to determine the $\mathrm{FP}(\omega)$ term as described in Equation 4.5. This intermediate Fabry-Pérot term is deconvolved from the measured transfer function. The updated transfer function is sent into the parameter extraction process again. By running the Fabry-Pérot deconvolution repeatedly, improved parameters n_s and κ_s can be obtained.

Clearly, the method requires considerable computational time, as a result of the iterations in both parameter extraction and reflection removal processes. Note that the process is unsuccessful if some reflected pulses are rejected prior to the process, or noise plagues the signal. This process is applicable to the single-sample measurement but its principle can be adopted by other settings.

4.7 Conclusion

This chapter presents an introduction to material characterisation using THz-TDS. Emphasis is given to parameter extraction processes associated with some common measurement configurations, both in transmission and reflection modes. The transmission-mode THz-TDS configurations include single-sample, dual-sample, multiple-sample, and sample-in-cell measurements, and those belonging to the reflection mode include single-reflection and double-reflection measurements. For every setting, a relevant analytical model for parameter extraction is derived explicitly, on the basis of ray optics. In addition, this chapter discusses some related signal processing methods, which are regularly implemented along with the parameter extraction process. These signal processing methods encompass signal averaging, signal denoising, phase unwrapping, and reflection removal. Some examples are demonstrated for a better appreciation of the concepts.

In line with the signal processing topics presented in this chapter, the next chapter, starting Part I of the thesis, proposes a signal processing algorithm for removal of absorption lines in THz-TDS spectra introduced by ambient water vapour. This new method can be useful in some measurement configurations where dry-air purging is not feasible.

Chapter 5

Removal of Water-Vapour Effects from THz-TDS Measurements

ONE source of disturbance in a pulsed T-ray signal is attributed to ambient water vapour. Water molecules in the gas phase selectively absorb incident T-rays at discrete frequencies corresponding to their molecular rotational transitions. When T-rays propagate through the atmosphere, this results in prominent resonances spread over the T-ray spectrum. Furthermore, in the time domain, fluctuations after the main pulse are observed in the T-ray signal. These effects are generally undesired, since they may mask critical spectroscopic data. So, ambient water vapour is commonly removed from the T-ray path by using a closed chamber during the measurement. Yet, in some applications a closed chamber is not always easy to implement. This situation, therefore, motivates the need for an optional alternative method for reducing these unwanted artefacts. This chapter represents a study on a computational algorithm that is a step towards addressing the problem arising from water-vapour absorption over a moderate propagation distance.

5.1 Introduction

In the T-ray frequency region, many polar gases of general interest possess unique rotational transition energies, which give rise to spectral resonances (Mittleman *et al.* 1998). Because of these unique fingerprints, THz-TDS proves useful for gas classification and recognition. This property, on the other hand, affects the THz-TDS capability in an open air setting, in which water vapour is ubiquitous.

Water vapour, the third most abundant gas in nature (Tripoli 2003), is known to have numerous rotational resonances in the T-ray band (van Exter *et al.* 1989b). T-ray spectroscopy of a sample, in open air, therefore often results in a combination of the sample's spectral features and water vapour resonances in the frequency domain. In the time domain, this results in field fluctuations after the main T-ray pulse. Mostly, these effects are undesirable, since they can obscure spectroscopic results of interest.

These water-vapour effects are easily removable during the measurement by purging an enclosed T-ray path with dry air or a non-polar gas such as nitrogen, which does not have energy level transitions in the T-ray regime (Grischkowsky *et al.* 1990)—alternatively a vacuum is sometimes used. In some applications, it is not always possible to enclose the entire T-ray beam path, for example, in applications where stand-off detection is required (Federici *et al.* 2005, Kemp *et al.* 2006). The effects can be partially cancelled out by calculation, if two measurements, one for the sample and the other for the reference, are available. Still some traces of water absorption persist due to the variation of the humidity during the two measurements (Bigourd *et al.* 2007). Further, not all situations can provide both measurements. Thus, this motivates the need for a numerical method, to alleviate the affect of water vapour on the data (Kemp *et al.* 2006).

A knowledge of the resonance characteristics in the T-ray frequency range allows a numerical estimation of the complex response of water vapour. Theoretically, the estimated complex response can be deconvolved directly from a received T-ray signal. But in practice, fitting of a modelled complex response to a measured response is complicated by many factors, e.g., the limited dynamic range (Jepsen and Fischer 2005), limited frequency resolution (Xu *et al.* 2003), and measurement uncertainties[8]. Moreover, the exact model requires precise knowledge of geometric and atmospheric conditions during the measurement. Moreover, if direct deconvolution is carried out, there is no measure to validate the result.

---

[8] Further discussion on measurement uncertainties can be found in Chapter 7.

Fine-tuning the strengths of complex resonances based on a brute-force search is introduced in this work. Each accurately modelled resonance is tuned in magnitude within a predefined range and then deconvolved from the measured signal. A tuning criterion is met when the ratio of the fluctuation energy to the main pulse energy of an adjusted signal is minimised. Repeatedly tuning the strength line-by-line ultimately results in the mitigation of water-vapour-induced effects. Constrained by the condition of minimum fluctuation energy, the algorithm always provides optimum results.

The proposed algorithm based on signal processing has merits in that (i) it is sufficiently general that the spectral response of a sample is minimally affected in a number of cases of practical interest, and (ii) the exact measurement conditions, including humidity, temperature, pressure, and propagation length, are not crucial in order to remove the water-vapour response.

This chapter is organised as follows. Section 5.3 elaborates the effects of water vapour on T-ray signals and spectra. In Section 5.4, the model characteristics of each complex resonance is determined and validated with a measurement. Section 5.5 considers the entire THz-TDS measurement as a system with several complex responses, and some possibilities for direct numerical removal of the water-vapour response are analysed. In Section 5.6, the strength-tuning algorithm is introduced, and the fluctuation ratio is proposed as a measure of tuning. This algorithm is experimentally tested in Section 5.7.

5.2 Research objective and framework

Objective

To invent a numerical algorithm, which is capable of removing water-vapour-induced effects from pulsed T-ray signals, as an alternative to the conventional 'sealed chamber' methodology. The result is tested by showing that the shape of a pulse measured in an open air setting is comparable to that of a pulse measured in nitrogen gas setting. The algorithm must be generalised in such a way that it minimally affects the desired spectroscopic features of interest.

Framework

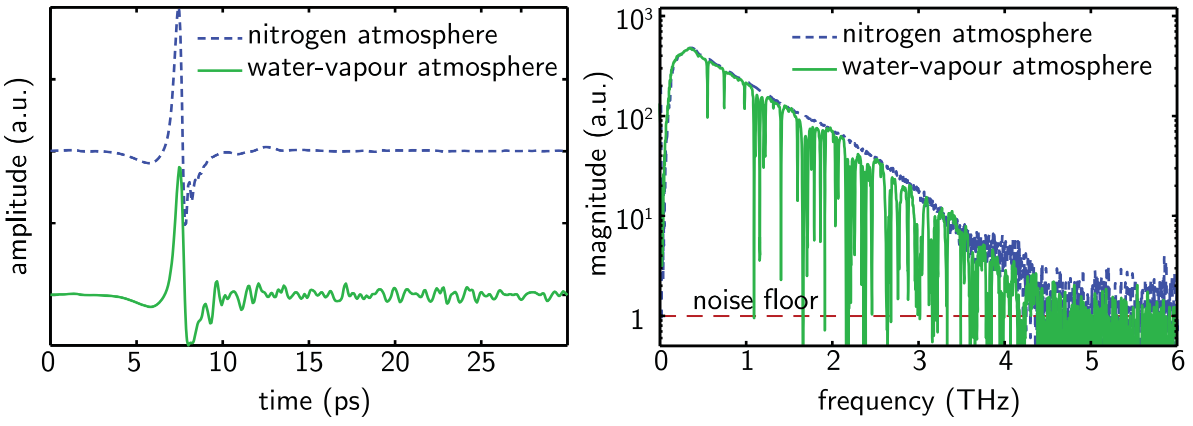
In this work it is presumed that an open-air measurement is performed in a normal atmosphere, i.e., in the range of room temperature and standard pressure. The propagation length is short enough that no continuum absorption (see Section 5.4.5) is involved in a measurement. The pure rotational resonances being investigated belong to the ground vibrational and electronic levels. A water molecule under measurement is H_2O, which is the most abundance in nature (Bievre *et al.* 1984).

5.3 Effects of water vapour on pulsed T-ray signal

Water vapour contains many frequency modes and the details of its interaction with an electromagnetic wave depend on the frequency. Water vapour exhibits pure rotational mode energy transitions spanning the millimetre wave to mid-infrared or approximately 3 GHz to 19.5 THz (Kemp *et al.* 1978, Bernath 2002). In the infrared region, both pure rotational line transitions and rotation-vibration bands are noticeable (Gasster *et al.* 1988). These transitions cause absorption and re-emission of wave energy in a number of narrow frequency bands, unique to water molecules.

In order to demonstrate the effects of water vapour transitions on the T-ray signal and spectrum, two T-ray propagation paths with different atmospheres are required, one containing water vapour and the other *in vacuo*. However, since non-polar gas molecules such as nitrogen do not demonstrate any T-ray rotational resonances, a nitrogen-filled path can serve as a replacement to an *in vacuo* path.

Figure 5.1 shows T-ray signals and spectra measured by THz-TDS under nitrogen and water-vapour atmospheres at ambient temperature and pressure. In the time domain, the signal recorded in the presence of water vapour undergoes long fluctuations after the main pulse. This is due to energy re-emission by rotational transitions of water molecules. In the frequency domain, the water vapour causes several sharp resonances at discrete frequencies, as a result of quantised rotational transition energies. In terms of Fourier theory, the pair of time-domain fluctuations and frequency-domain resonances is based on the principle that a sharp feature in one domain is related to a broad feature in the other domain.



(a) T-ray signals

(b) T-ray spectra

Figure 5.1. Effects of water vapour on T-ray pulse and spectrum. (a) T-ray signals recorded in nitrogen and water-vapour atmospheres. The sampling interval is 0.067 ps, and the total time duration is 136 ps. The vertical offset is for clarity. (b) Their corresponding spectra with the spectral resolution of 7.32 GHz. For this measurement the noise floor marks the cutoff frequency at approximately 4.0 THz.

In addition to the vapour-induced fluctuations in the time domain signal and the resonance effects in the frequency domain, the analysis shows that the T-ray energy loss by water absorption calculated between 0.0 and 4.0 THz is as high as 10% for the spectrum in Figure 5.1. The loss is likely due to the non-directional energy re-emission of water molecules. In the time domain, the ratio of the main pulse energy to the tail (vapour-induced fluctuation) energy is calculated for both of the time-domain signals (for more details on the calculation, see Section 5.6.3). The energy ratio for the nitrogen measurement is 429.98, whereas the ratio for the water vapour measurement is 18.82. This measure will be used later on to construct a criterion for fluctuation removal.

5.4 Model of water vapour absorption and dispersion

It is required that the water-vapour response be modelled appropriately before a further step in water-vapour removal is taken. Modelling of absorption and dispersion for a rotational resonance involves selecting a suitable lineshape and calculating the line position, line strength, and linewidth. Fortunately, the line position and line strength are, only to a small extent, dependent on atmospheric conditions, i.e., the line position is constant at low pressures (Townes and Schawlow 1955) and the peak absorption is not a function of the pressure (Townes and Schawlow 1955, Harde *et al.* 1997b). In fact, the line position and line strength have been well parameterised in many publications. Thus these two parameters are readily available, leaving only the linewidth and lineshape to be determined.

In the following subsections these parameters are discussed in detail. Section 5.4.1 gives a list of spectroscopic catalogues containing the line strength and line position of water molecules. The theories for linewidth and line broadening are briefly introduced in Section 5.4.2. The determined linewidth leads to a selection of the lineshape, shown in Section 5.4.3. Section 5.4.4 shows the complete absorption and dispersion models as a result of several rotational resonances. The effect of continuum absorption is discussed in Section 5.4.5.

5.4.1 Line strength and line position

Water vapour rotational resonances in the T-ray regime have been extensively measured and analysed via either conventional FTIR spectroscopy (Kauppinen *et al.* 1978, Messer *et al.* 1983), THz-TDS (van Exter *et al.* 1989b, Cheville and Grischkowsky 1999), or other techniques (Helminger *et al.* 1983, Matsushima *et al.* 1995, Chen *et al.* 2000)—the extensive study is largely due to the abundance of water vapour, which involves many physical processes in nature. Subsequently, the computer-accessible spectroscopic parameters of water vapour, including the line position, line strengths and linewidth, are available from many research groups[9], for example:

- Pickett *et al.* (1998), the Molecular Spectroscopy Team, at the Jet Propulsion Laboratory (JPL) report the spectral line intensity, strength, and other related parameters of 331 interstellar gas species in the submillimetre, millimetre, and microwave frequency ranges. Among these gases, H_2O reportedly contains 3086 spectral lines in the frequency range up to 29.8 THz. Later, Müller *et al.* (2005) founded the Cologne Database for Molecular Spectroscopy (CDMS), which added 280 more species to its own database, with an identical data format to that used at JPL.
- Rothman *et al.* (2005) with the HIgh-resolution TRANsmission molecular absorption database (HITRAN) provided a compilation of spectroscopic parameters of

---

[9] Some of them have two or more catalogue publications as a result of regular updating, and only the latest revision is cited here.

39 molecules including the water molecule and its isotopologues, spanning the microwave and ultraviolet bands. Apart from the line intensity and position, the list consists of the Einstein A-coefficients, statistical weight for upper and lower levels of the transitions, etc.

- Jacquinet-Husson *et al.* (1999) published the Gestion et Etude des Informations Spectroscopiques Atmosphériques (GEISA) database, which contains the spectroscopic parameters of 42 molecules (96 isotopes), including the water molecule, in the frequency range between 0 and 1076 THz. According to this database, H_2O contains 30,117 lines from 12 GHz to 680 THz.

Here in this work the catalogue published by the JPL group (Pickett *et al.* 1998) is adopted. As this catalogue has been regularly updated since before 1985 (Poynter and Pickett 1985), it is unlikely that the high-intensity lines of water vapour are absent from the list.

According to the JPL catalogue, the line strength is reported in terms of the integrated line intensity at 300 K, $I_a(300\text{ K})$, and is scalable with the temperature, T, via

$$I_a(T) \approx I_a(300\text{ K}) \left(\frac{300}{T}\right)^{n+1} \exp\left[-\left(\frac{1}{T} - \frac{1}{300}\right)\frac{E_i}{k_\mathrm{B}}\right], \tag{5.1}$$

where $n = 3/2$ for the water molecule, E_i is the lower state energy, also provided in the catalogue, and $k_\mathrm{B} = 8.617339 \times 10^{-5}$ eV/K is Boltzmann's constant. The approximation is valid for a small difference between the lower and upper state energy compared with $k_\mathrm{B}T$. Despite the temperature dependency, the line strength is independent of pressure (Townes and Schawlow 1955).

The line position is pressure- and temperature-dependent. Nevertheless, the shifting is an order of magnitude lower than the line broadening (Podobedov *et al.* 2004), and thus the position can be thought as constant (Townes and Schawlow 1955).

5.4.2 Linewidth

A rotational absorption resonance is possibly broadened by a number of factors (Townes and Schawlow 1955, Bernath 2005), for instance, self or foreign-gas collisions, molecule-wall collisions, the Doppler effect, and natural lifetime broadening. Among these factors, at a standard pressure the collisional broadening is predominant in determining the spectral linewidth (Gopalsami *et al.* 1996, Harde *et al.* 1997b).

Table 5.1. Lorentz, Vleck-Weisskopf, and Gross profiles. Here, ω_a denotes the a^{th} transition frequency, and $\Delta\omega_a$ denotes the HWHM of the profile at that frequency. The absorption and dispersion profiles are linked together via the Kramers-Kronig relation. Adapted from Kemp *et al.* (1978).

Lineshape	Absorption profile, $2\kappa/c$	Dispersion profile, $2(n-1)/c$
Lorentz	$\frac{\Delta\omega_a}{\pi}\left[\frac{1}{(\omega_a-\omega)^2+\Delta\omega_a^2}-\frac{1}{(\omega_a+\omega)^2+\Delta\omega_a^2}\right]$	$\frac{1}{\pi}\left[\frac{\omega_a-\omega}{(\omega_a-\omega)^2+\Delta\omega_a^2}-\frac{\omega_a+\omega}{(\omega_a+\omega)^2+\Delta\omega_a^2}\right]$
Vleck-Weisskopf	$\frac{\Delta\omega_a\omega}{\pi\omega_a}\left[\frac{1}{(\omega_a-\omega)^2+\Delta\omega_a^2}+\frac{1}{(\omega_a+\omega)^2+\Delta\omega_a^2}\right]$	$\frac{1}{\pi\omega_a}\left[\frac{\omega_a(\omega_a-\omega)+\Delta\omega_a^2}{(\omega_a-\omega)^2+\Delta\omega_a^2}-\frac{\omega_a(\omega_a+\omega)+\Delta\omega_a^2}{(\omega_a+\omega)^2+\Delta\omega_a^2}\right]$
Gross	$\frac{4\Delta\omega_a}{\pi}\frac{\omega\omega_a}{(\omega_a^2-\omega^2)^2+4\Delta\omega_a^2\omega^2}$	$\frac{2}{\pi}\frac{\omega_a(\omega_a^2-\omega^2)}{(\omega_a^2-\omega^2)^2+4\Delta\omega_a^2\omega^2}$

The collision broadened linewidth is dependent on temperature (Gasster *et al.* 1988, Cheville and Grischkowsky 1999, Podobedov *et al.* 2004) and pressure (Bernath 2005). This dependency yields the Benedict and Kaplan relationship (Benedict and Kaplan 1964, Kemp *et al.* 1978),

$$\Delta\omega \quad = \quad \Delta\omega_0\left(\frac{p}{p_0}\right)\left(\frac{T_0}{T}\right)^m , \tag{5.2}$$

where $\Delta\omega_0$ is the FWHM of the line at pressure p_0 and temperature T_0, $\Delta\omega$ is the FWHM at pressure p and temperature T, and m is the temperature index, varying between 0.5 and 1.0. An index of $m = 0.68$ can be used in the calculation (Rothman *et al.* 2005).

Determined from the experiment, the average FWHM, $\Delta\omega_0/2\pi$, of water vapour resonances at ambient temperature and standard pressure is 6 GHz. This value is in agreement with an estimate of 10 MHz per Torr, or 7.6 GHz at 1 atm (Bernath 2005, De Lucia 2003) within the limit of measurement uncertainty.

5.4.3 Lineshape

The rotational absorption lineshape can be modelled either by Lorentz (1906), Vleck-Weisskopf (Van Vleck and Weisskopf 1945), or Gross (Emery 1972, Kemp *et al.* 1978) profiles. The choice depends on the frequency range of interest and the nature of line broadening. The mathematical models of these profiles in case of absorption and dispersion are summarised in Table 5.1.

More complicated and general profiles are also available, e.g., the Voigt profile that is a convolution between the Lorentzian profile, as a result of collisional broadening, and the Gaussian profile, as a result of Doppler broadening[10] (Armstrong 1967, Bernath 2005). Voigt profile and Lorentzian profile give a similar fit in case that Doppler broadening is not significant (Gasster *et al.* 1988).

Notwithstanding this, no significant difference is observable among these profiles at or nearby a T-ray resonance, with the linewidth nearly three orders of magnitude lower than the resonance frequency. In this paper, a Lorentzian profile is adopted in modelling the water vapour response in the T-ray frequency range.

Figure 5.2 shows the measured absorption and dispersion profiles of water molecules, centred at 0.557 and 0.752 THz. In comparison to the measurement, the Lorentz model is precisely within the limit of uncertainty. A small frequency shift between the measurement and the model can be observed, but this shift is within the spectral resolution limit of measurement of 0.24 GHz.

5.4.4 Ensemble of rotational transition resonances

This subsection gives a complete model, in terms of optical constants, as a result of several rotational transition resonances in the frequency range of interest. The model is a function of the lineshape, linewidth, line strength, and line position, which are discussed in the previous subsections. Later on, the model is compared to experimental measurement in order to verify the accuracy.

The model extinction coefficient and index of refraction can be described as a summation between an offset and a set of Lorentzian profiles, or, respectively,

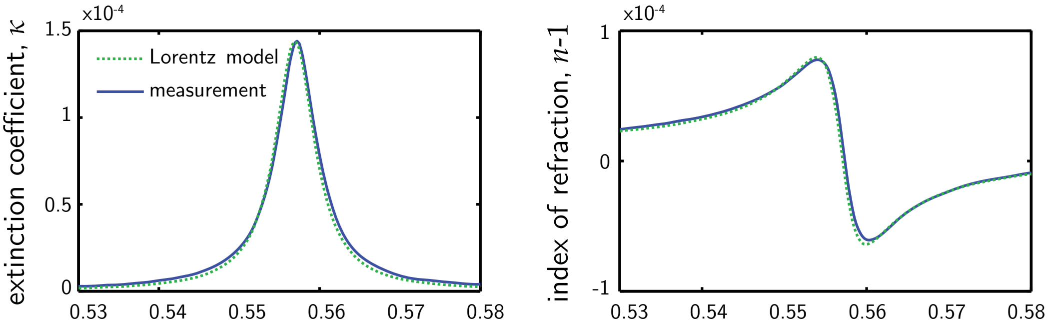
$$\kappa(\omega) = \kappa(0) + \sum_a m_a \kappa_a(\omega)\,, \tag{5.3a}$$

$$n(\omega) - 1 = n(0) + \sum_a m_a [n_a(\omega) - 1]\,, \tag{5.3b}$$

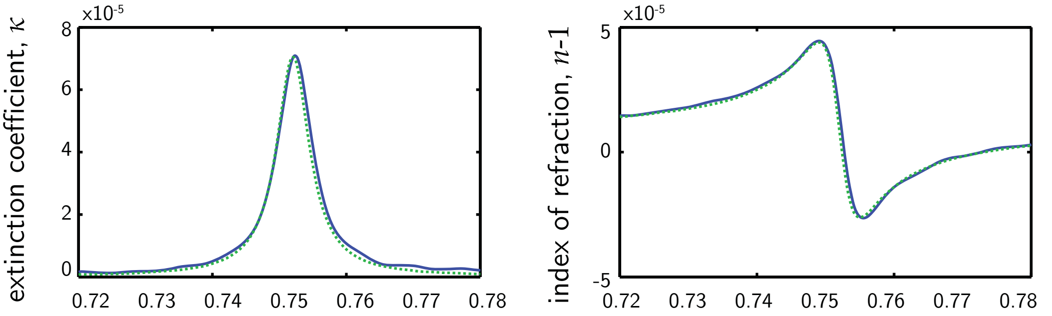
where κ_a and $n_a(\omega)$ are referred to the Lorentzian profiles in Table 5.1, the line strength m_a is proportional to the integrated line intensity, I_a, (Equation 5.1) divided by the line position, ω_a, or, $m_a \propto I_a/\omega_a$. The offset is partially a result of electronic and molecular vibrational resonances at high frequencies (Kemp *et al.* 1978), and also a contribution from the tails of other rotational resonances unaccounted for by the model.

---

[10] The Gaussian lineshape is inherited from the velocity distribution.



(a) absorption profile at 0.557 THz (b) dispersion profile at 0.557 THz



(c) absorption profile at 0.752 THz (d) dispersion profile at 0.752 THz

Figure 5.2. Model and measurement of H_2O absorption and dispersion profiles. The complex profiles are taken from two rotational resonances at 0.557 and 0.752 THz. The green dotted lines represent Lorentz models, whereas the blue solid lines represent measured profiles, which are extracted from the raw data shown in Figure 5.1. The model FWHM of both transitions is set to 6 GHz.

Typically, at the low frequency range from DC to a few hundred gigahertz, THz-TDS cannot produce sufficient energy to overcome the noise. As a result, the resolved optical constants in this frequency range are unreliable. As discussed in Section 4.6.3, a phase extrapolation technique is usually introduced to correct the unwrapping process, forcing the phase to start at zero. Thus, it is not necessary in the model to consider the index offset, $n(0)$, which is derived from the phase.

The absorption of water vapour, shown in Figure 5.3, is determined from the measured data (Figure 5.1) and the Lorentzian model in the frequency range between 0.0 and 4.0 THz. It can be clearly seen that below 1.6 THz the model closely resembles the measurement. This match is possible since the resonances in this low frequency range have their strengths beneath the maximum absorption coefficient measurable by the system. However, as the frequency goes beyond 1.6 THz, the model cannot track the measurement due to the dynamic range of the system. In this situation, the system

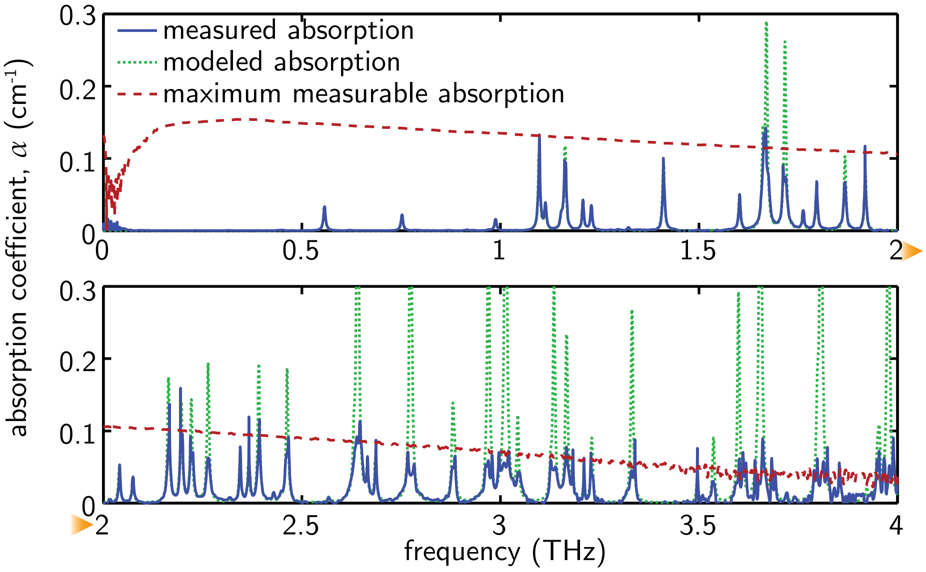


Figure 5.3. Model and measurement of H_2O absorption. The lower graph is a continuation of the upper graph, as indicated by the orange arrowheads. The measured and modelled curves match well in the range of 0.0 to 1.6 THz, but match poorly in the higher frequency range, where the absorption strengths become higher than the dynamic range of the system. The maximum measurable absorption coefficient of the system, represented by the red dashed line, is calculated using the method of Jepsen and Fischer (2005).

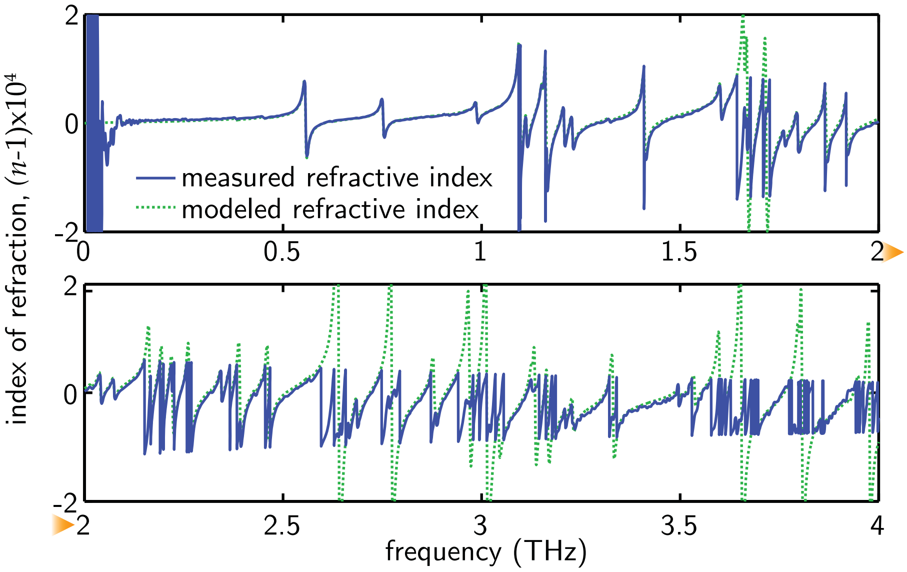


Figure 5.4. Model and measurement of H_2O dispersion. Similarly to the absorption curves, in Figure 5.3, the measurement and the model fit well together in the frequency range below 1.6 THz. In the higher frequency range, where those absorption lines are stronger than the maximum measurable absorption, the corresponding dispersion profiles here are also distorted. The spikes below 0.1 THz are due to low SNR. The measurement is offset in magnitude by 2.65×10^{-5} against the noise in the low frequency range.

can no longer measure the absorption coefficient correctly, and the clipped absorption peaks are obvious, in particular, in the range from 2.0 to 4.0 THz. Furthermore, the clipped lines become less distinct, when the T-ray power reaches the noise floor beyond 3.0 THz. The disagreement between the modelled and measured dispersion, in Figure 5.4, can be similarly explained.

5.4.5 Continuum absorption

The continuum absorption, unaccounted for so far, is defined as the excess measured absorption unable to be quantified by the resonance spectrum. The profile of the continuum absorption is smooth over a frequency range, rather than being sharp resonances. The mechanism underpinning the continuum absorption has not been fully understood (Podobedov *et al.* 2005). However, it is known that the continuum absorption is pressure and frequency dependent. A continuum absorption for H_2O collisions, extracted from an experiment in the T-ray frequency range, between 0.4 and 1.83 THz, is estimated to be 4.22×10^{-8} (dB/km)/(hPa GHz)^2 (Podobedov *et al.* 2005). In a normal atmosphere and with a relatively short path, the effect of the continuum absorption is negligible (Clough *et al.* 1989, Podobedov *et al.* 2005).

5.5 Removal of H_2O response by direct deconvolution

This section introduces a straightforward way to remove the water vapour response from T-ray signals. Direct deconvolution of the modelled water vapour response from the measured spectrum is theoretically possible. Nevertheless the capability of the method relies on the achievable SNR of the system and the precision of the physical conditions. Direct deconvolution is carried out here for purposes of comparison and highlighting drawbacks, thus motivating the need for our water-vapour removal method in Section 5.6.

In Section 5.5.1, the entire T-ray spectroscope is modelled as a system, in which the response of each part, including the water vapour, is substituted by a black box. A removal technique that exploits deconvolution of the sample signal with the water vapour model is discussed in Section 5.5.2. Section 5.5.3 explains the usual situation where the sample signal is deconvolved with the reference signal. The limitations of this technique are summarised in Section 5.5.4.

5.5.1 Water vapour as a black box

During a THz-TDS measurement, an emitted T-ray signal evolves based on several factors, e.g., the sample response, the water vapour response, the instrument response, and the noise. These can be modelled as a system, as illustrated in Figure 5.5.

Given that $x(t)$ denotes the input T-ray pulse, which is immediately deployed from a transmitter, $s(t)$ the impulse response of a sample, $w(t)$ the impulse response of water vapour, $u(t)$ the impulse response of the instrument, $n(t)$ the noise, and $r(t)$ the time windowing, the received pulse, $y(t)$, can be described as

$$y(t) = [x(t) * s(t) * w(t) * u(t)] \cdot r(t) + [n(t) * u(t)] \cdot r(t) \,, \tag{5.4}$$

where $*$ is the convolution operator. The noise is simply additive and can thus be treated as an extra term in Equation 5.4.

Via Fourier transform, the above equation in the frequency domain is

$$\begin{aligned} Y(\omega) &= [X(\omega) \cdot S(\omega) \cdot W(\omega) \cdot U(\omega)] * R(\omega) + [N(\omega) \cdot U(\omega)] * R(\omega) \\ &= [X(\omega) \cdot S(\omega) \cdot W(\omega) \cdot U(\omega)] * R(\omega) + \acute{N}(\omega) \,, \end{aligned} \tag{5.5}$$

where $Y(\omega)$, $X(\omega)$, $S(\omega)$, $W(\omega)$, $U(\omega)$, $R(\omega)$, and $N(\omega)$ represent complex frequency responses, as the Fourier pairs of $y(t)$, $x(t)$, $s(t)$, $w(t)$, $u(t)$, $r(t)$, and $n(t)$ respectively, and $\acute{N}(\omega) = [N(\omega) \cdot U(\omega)] * R(\omega)$.

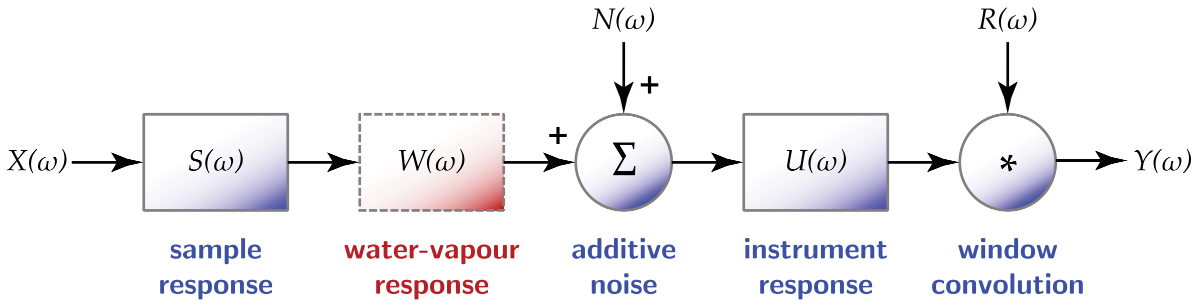


Figure 5.5. Model of spectroscopic system. The system response is the combination of the frequency responses of the sample, water vapour, and instrument, denoted by $S(\omega)$, $W(\omega)$, and $U(\omega)$, respectively. The symbols, $X(\omega)$, $Y(\omega)$, $N(\omega)$, and $R(\omega)$ represent input T-ray spectrum, measured spectrum, noise, and window (or limited recording duration) responses, respectively. The water-vapour response in the red dashed box is the target for removal.

The water-vapour frequency response can be expressed by

$$W(\omega) \quad = \quad \exp\left[-j\hat{n}_{\text{vap}}\frac{\rho(T)}{\rho_0(T)}\cdot\frac{\omega L}{c}\right] . \tag{5.6}$$

Here, L denotes the T-ray propagation length in free space, excluding the space occupied by sample(s). The water vapour's complex index of refraction, $\hat{n}_{\text{vap}} = n_{\text{vap}} - j\kappa_{\text{vap}}$, contains the refractive index n_{vap} and the extinction coefficient κ_{vap}, which are modelled in Equation 5.3. The water-vapour response is also determined by the humidity (Yuan *et al.* 2003), where $\rho(T)$ is the vapour density at temperature T and $\rho_0(T)$ is the saturation vapour density at the same temperature.

5.5.2 Deconvolution of model H$_2$O response from sample signal

In order to remove the effects of water vapour from a measurement, the water-vapour response, $W(\omega)$, in Equation 5.5 must be replaced by the vacuum response, $V(\omega)$, with the same propagation length,

$$V(\omega) \quad = \quad \exp\left[-j\frac{\omega L}{c}\right] . \tag{5.7}$$

In the case of a noise-free signal, the direct deconvolution can be performed,

$$\begin{aligned} Y_{\text{vacuum}}(\omega) &= Y(\omega)V(\omega)/W(\omega) \\ &= [X(\omega)\cdot S(\omega)\cdot V(\omega)\cdot U(\omega)] * R(\omega) . \end{aligned} \tag{5.8}$$

The deconvolution can be performed successfully with no hindrance from either the window response, $R(\omega)$, which causes the limit spectral resolution, or the system response, $U(\omega)$. However, it is only feasible with an ideal signal. If the deconvolution is carried out with a noisy signal, the outcome is degraded.

From Equation 5.5 the noisy response can be modelled in two parts,

$$Y(\omega) \approx \begin{cases} [X(\omega)\cdot S(\omega)\cdot W(\omega)\cdot U(\omega)] * R(\omega) & \text{, if} \quad |X\cdot S\cdot W| > |N| \\ \acute{N}(\omega) & \text{, otherwise.} \end{cases} \tag{5.9}$$

Assuming the noise level is unity and the model water-vapour response matches well with the measurement, performing the direct deconvolution would yield

$$Y(\omega) \approx \begin{cases} [X(\omega)\cdot S(\omega)\cdot V(\omega)\cdot U(\omega)] * R(\omega) & \text{, if} \quad |X\cdot S\cdot W| > 1 \\ V(\omega)/W(\omega) & \text{, otherwise.} \end{cases} \tag{5.10}$$

Thus, removal of the water vapour response fails at $|X \cdot S \cdot W| \leq |N|$ or $\mathrm{D}(\omega) \leq |S \cdot W|^{-1}$, where $\mathrm{D}(\omega) = |X/N|$ is the system's dynamic range. This usually occurs at the peaks of absorption resonances, as shown in Figure 5.3.

5.5.3 Deconvolution between sample and reference signals

In the case where the sample and reference signals are available, basically, deconvolution between these two signals would remove common features, including the water vapour response. However, consider these two equations for the sample and reference signals derived from Equation 5.5,

$$\begin{aligned} Y_{\text{sam}}(\omega) &= [X(\omega) \cdot S(\omega) \cdot W(\omega) \cdot U(\omega)] * R(\omega) \\ &\quad + [N_{\text{sam}}(\omega) \cdot U(\omega)] * R(\omega)\,, && (5.11\text{a}) \\ Y_{\text{ref}}(\omega) &= [X(\omega) \cdot W(\omega) \cdot U(\omega)] * R(\omega) \\ &\quad + [N_{\text{ref}}(\omega) \cdot U(\omega)] * R(\omega)\,. && (5.11\text{b}) \end{aligned}$$

Here, $\langle N_{\text{sam}} \rangle = \langle N_{\text{ref}} \rangle = N$. Again, each signal can be modelled by two parts,

$$Y_{\text{sam}}(\omega) \approx \begin{cases} [X(\omega) \cdot S(\omega) \cdot W(\omega) \cdot U(\omega)] * R(\omega) & \text{, if} \quad |X \cdot S \cdot W| > |N| \\ \acute{N}_{\text{sam}}(\omega) & \text{, otherwise,} \end{cases} \tag{5.12}$$

and

$$Y_{\text{ref}}(\omega) \approx \begin{cases} [X(\omega) \cdot W(\omega) \cdot U(\omega)] * R(\omega) & \text{, if} \quad |X \cdot W| > |N| \\ \acute{N}_{\text{ref}}(\omega) & \text{, otherwise.} \end{cases} \tag{5.13}$$

Dividing the sample spectrum by the reference yields (the windowing effect is ignored here)

$$\frac{Y_{\text{sam}}(\omega)}{Y_{\text{ref}}(\omega)} \approx \begin{cases} S(\omega) & \text{, if} \quad |X \cdot S \cdot W| > |N| \\ \acute{N}_{\text{sam}}(\omega) / \left[X(\omega) \cdot W(\omega) \cdot U(\omega)\right] & \text{, if} \quad |X \cdot W| > |N| > |X \cdot S \cdot W| \\ \acute{N}_{\text{sam}}(\omega) / \acute{N}_{\text{ref}}(\omega) & \text{, otherwise.} \end{cases} \tag{5.14}$$

It is clear from the above equation that the effects of the water vapour response still persist, even if they are common to both the sample and reference signals.

5.5.4 Limitation of the algorithm

From the two previous subsections, the deconvolution to remove the water vapour response from the sample signal is effective, if only the measured signal, $|X \cdot S \cdot W|$, is higher than the noise level, $|N|$, at all frequencies of interest. Only a small part of spectrum that violates this condition prohibits the whole deconvolution. This condition is not always attainable due to the sharp absorption resonances and nature of T-ray spectral magnitude, which is relatively low and tends to decline at higher frequencies.

Apart from the noise issue, the applicability of direct deconvolution in Section 5.5.2 critically depends on the knowledge of the atmospheric and geometric parameters during the measurement, i.e., the temperature, pressure, humidity, and propagation length, which are indispensable for modelling the water vapour response.

Although all water resonances are well above the noise floor and all required parameters are obtained exactly, significant discrepancies between the model and the measurement can still occur. In that case, there is no means to quantify the accuracy of the results produced by the direct deconvolution technique.

5.6 Removal of H_2O response by strength tuning

This section proposes an alternative to the direct deconvolution technique presented in the previous section. By introducing the strength-tuning algorithm, two difficulties, the noise and the requirement for atmospheric parameters are simultaneously mitigated.

Section 5.6.1 augments the spectroscopic system, in Figure 5.5, by decomposition of the water vapour response into numerous parts. The strength-tuning algorithm is introduced in Section 5.6.2, followed by the fluctuation ratio as a tuning criteria in Section 5.6.3. Discussion on the algorithm's generality can be found in Section 5.6.4.

5.6.1 Water vapour as black boxes

The water vapour response can be separated into several components, as shown in Figure 5.6, and the separation is formulated as

$$W(\omega) \quad = \quad V(\omega)W_0(\omega)W_1(\omega)\ldots W_a(\omega)\,. \tag{5.15}$$

Each component corresponds to a complex resonance, and has its frequency response derived from Equations 5.3 and 5.6, or

$$W_a(\omega) \quad = \quad \exp\left[-j\, m_a(n_a - j\kappa_a)\cdot\frac{\omega}{c}\right]\,. \tag{5.16}$$

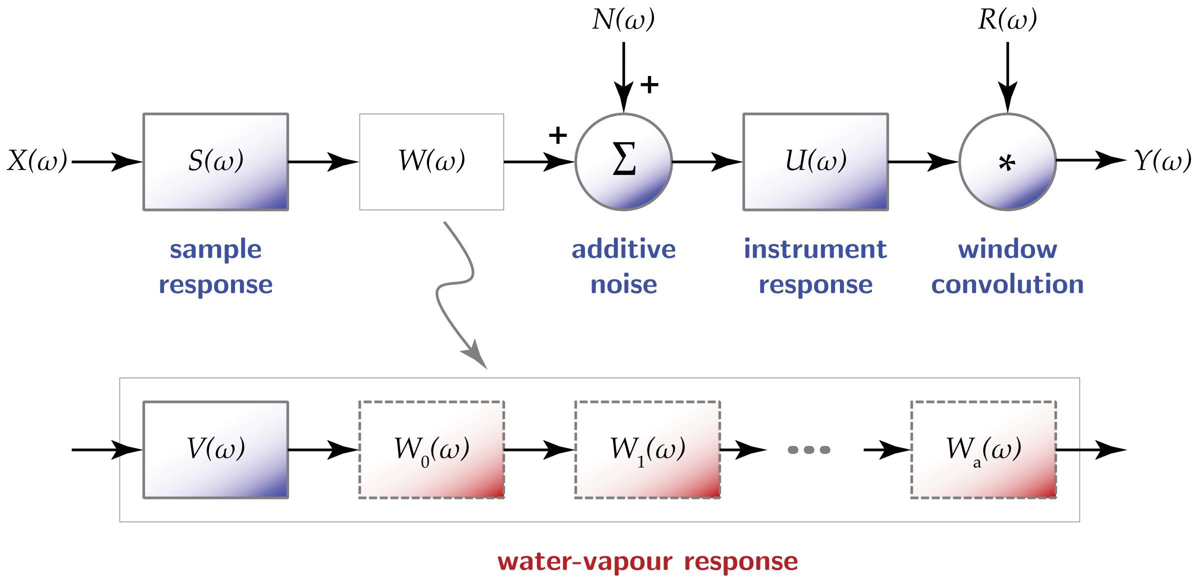


Figure 5.6. Model of spectroscopic system with decomposed H_2O response. The system is similar to that in Figure 5.5, excepting that the water vapour response is decomposed into many components according to Equation 5.15. Each component in a red dashed box corresponds to the complex response of a rotational resonance of water molecules, and is characterised by Equation 5.16. They are subject to removal by the strength-tuning algorithm, as discussed in Section 5.6.2. The $V(\omega)$ is the vacuum response, explained in Equation 5.7.

The line strength, m_a, now incorporates the humidity ratio, $\rho(T)/\rho_0(T)$, and the propagation length, L. It can be inferred from Equations 5.15 and 5.16 that all components or resonances can be removed from the measured spectrum separately and independently. Consequently, the deconvolution at the noisy parts of the spectrum can be avoided, and also it is possible to tune m_a of each resonance to the measurement without accurate information of the atmospheric and geometric parameters. In order to achieve this parameter relaxation, a criterion indicating the fit of tuning is essential—this is now addressed in the following subsection.

5.6.2 Strength-tuning algorithm

The strength-tuning algorithm is shown in Figure 5.7. The counters a and b for the line position and line strength are reset. The H_2O spectral line parameters, including intrinsic line positions and strengths, in the frequency range of interest, are then fetched from an existing database. A complex resonance profile, $n(\omega) - j\kappa(\omega)$, of a rotational transition at $\omega_{a=0}$ is modelled according to the Lorentzian profile in Table 5.1. The modelled profile multiplied by the initial strength factor $m_{a=0,b=0}$ is then temporarily deconvolved from the measured complex response, $Y(\omega)$. The deconvolved time-domain signal, $y(t)$, is estimated for its fluctuation ratio (see Section 5.6.3). The procedure is repeated to find the fluctuation ratio at other predefined line strengths $m_{a=0,b}$; $b \in \{1, 2, 3, \ldots\}$, which may increase by a discrete step in the vicinity of database's line strength.

Once the tuning range, $m_{a=0,b}$, is covered, the algorithm picks up an optimal strength within $m_{a=0,b}$, which gives the minimum fluctuation ratio, if any, and permanently removes that optimal complex resonance from the measured signal. It is possible that the change in the time-domain fluctuation is so subtle that the minimum fluctuation ratio cannot be found. In that case the removal of a current line is skipped. The tuning procedure starts again, but with the next transition resonance, $\omega_{a=1}$, and is repeated until all resonances are optimised and possibly removed. It should be noted that a strong resonance, which suffers from noise, i.e., $|X \cdot S \cdot W| < |N|$ at and around the peak of resonance, is avoided by the tuning process, because performing deconvolution of such an ill-defined resonance gives rise to a high fluctuation ratio. Since there might be dependencies among the resonances, with regard to time-domain fluctuations, the whole process starts again until the fluctuation ratio no longer decreases.

It is advisable not to perform zero padding prior to Fourier transform in the removal process, as the interpolated spectrum does not exactly reflect the reality—this is especially the case for points at resonances that have high variability. Otherwise, a discrepancy between the model and the spectrum would cause a large remnant in the spectrum after deconvolution. Also note that to speed up the algorithm, the resonances that have a strength lower than the amplitude uncertainty can be skipped.

5.6.3 Fluctuation ratio

A criterion related to the quality of a T-ray signal is necessary in selecting the optimal resonance line strengths. The 'quality' here is defined as having a high pulse energy with low fluctuations in the time domain, which implies the absence of water-vapour resonances in the frequency domain.

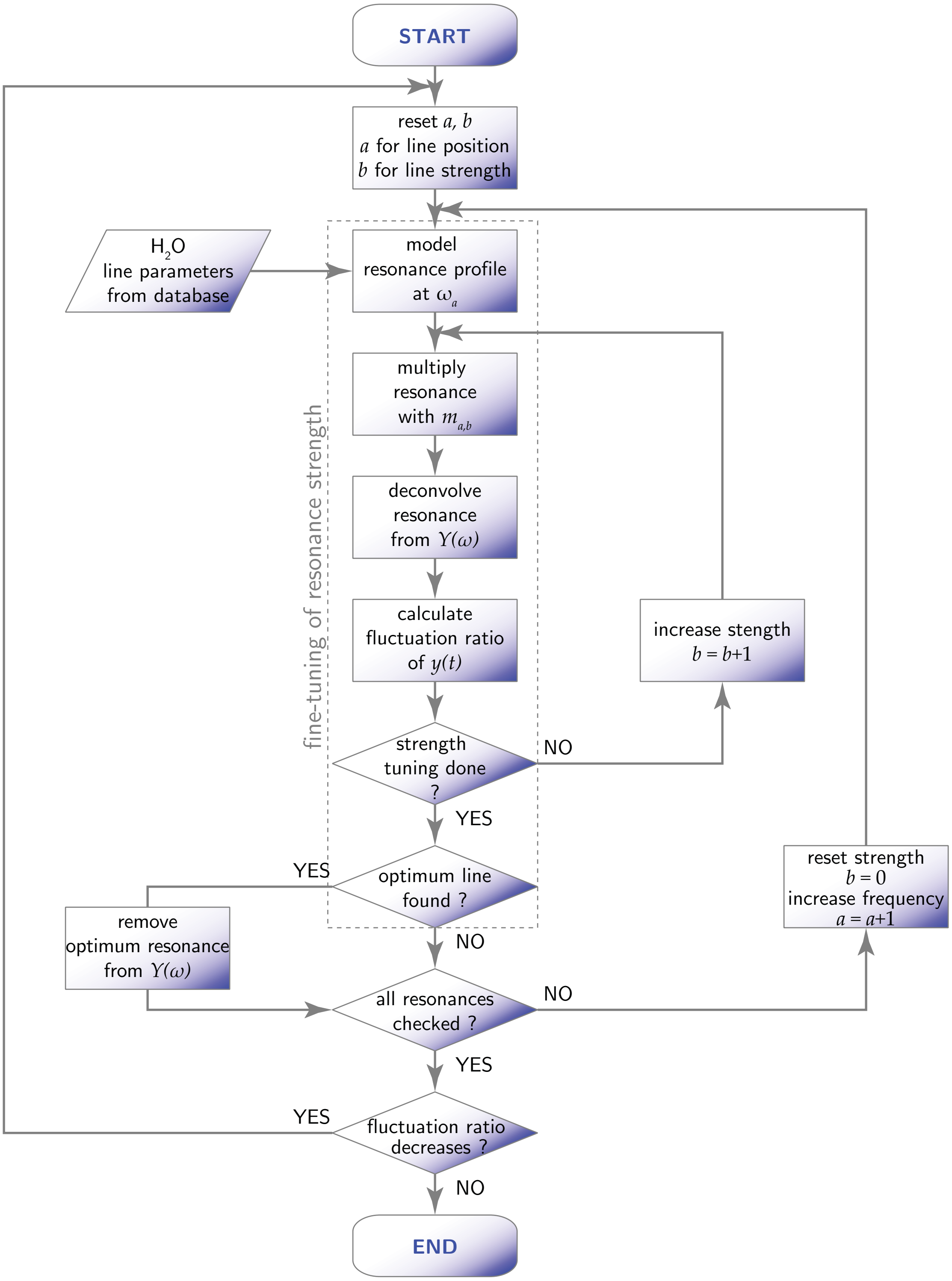
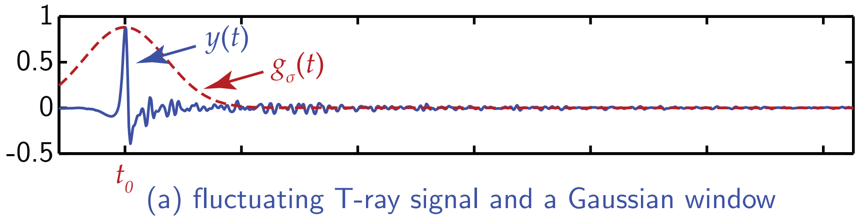
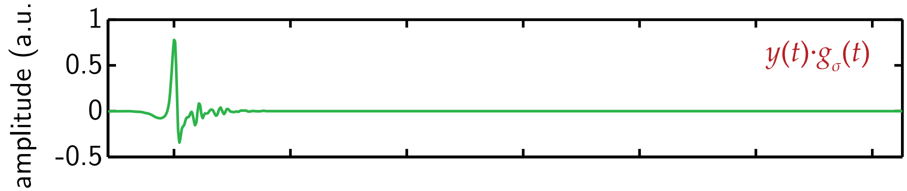


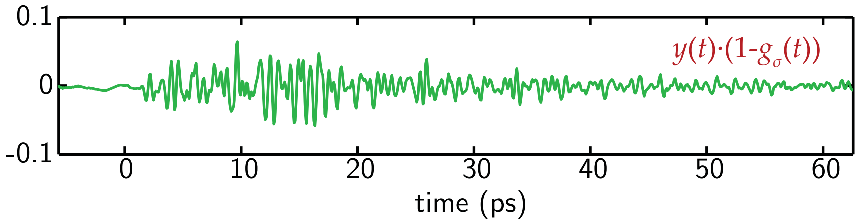
Figure 5.7. Strength-tuning algorithm. Each resonance in the T-ray frequency range is modeled and fine-tuned in its strength with the criterion of minimum fluctuation, and then the optimal resonance is removed from the measured complex response. The procedure is repeated for all resonances within the frequency range of interest, until the fluctuation ratio no longer decreases.



(a) fluctuating T-ray signal and a Gaussian window



(b) T-ray signal weighted by the Gaussian window



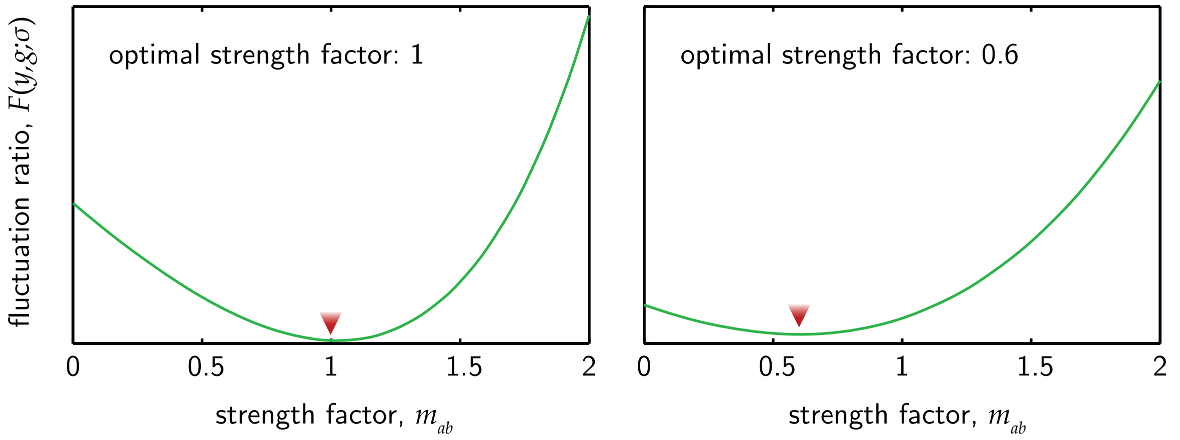
(c) T-ray signal weighted by the Gaussian window complement

Figure 5.8. Gaussian window applied to T-ray signal. A T-ray signal weighted by a Gaussian window, which has its peak position set exactly at the main pulse peak, yields the main pulse with suppressed fluctuations. A T-ray pulse weighted by the Gaussian window complement yields only fluctuations. The FWHM of the Gaussian window used here is 8.33 ps.

In order to quantify the quality one must be able to evaluate the total energies of the main pulse and of the fluctuations. One potential way is to window the time-domain signal with a Gaussian profile to obtain the amplitude at a desired portion. A Gaussian window is selected, because a T-ray main pulse is essentially derived from an optical pumping pulse, which takes on a Gaussian pulse profile (Duvillaret *et al.* 2001).

Figure 5.8(a) shows a Gaussian window, $g_\sigma(t)$, overlapping the T-ray signal, $y(t)$. A Gaussian window with appropriate width and position would eliminate the fluctuating tail of a signal, as shown in Figure 5.8(b). On the other hand, the complement of a Gaussian window could also be used to remove the main pulse, as in Figure 5.8(c). With the assistance of a Gaussian window, the fluctuation energy normalised by the main pulse energy—coined the *fluctuation ratio*—can be formulated as

$$F(y, g\,;\sigma) \;=\; \frac{\int_t [y(t)\cdot(1-g_\sigma(t))]^2 dt}{\int_t [y(t)\cdot g_\sigma(t)]^2 dt}\,. \tag{5.17}$$



(a) fluctuation ratio for a 0.75-THz resonance (b) fluctuation ratio for a 1.2-THz resonance

Figure 5.9. Variation of fluctuation ratio. Two example plots provide insight as to how the fluctuation ratio varies as the line strength is tuned away from its optimum. The optimal line strength results in the minimum fluctuation energy and maximum main pulse energy. The red arrowheads indicate the positions of optimal strength factors.

The integration is carried out over the time duration of a recorded T-ray signal. A Gaussian window is given by

$$g_\sigma(t) \quad = \quad \exp[-(t-t_0)^2/2\sigma^2]\,, \tag{5.18}$$

where t_0 is the peak position of the T-ray signal, $y(t)$, and σ multiplied by $2\sqrt{2\ln 2}$ is the FWHM of a Gaussian window.

Figure 5.9 demonstrates variation of the fluctuation ratio during the strength tuning procedure. The optimal strength is located at the local minimum, i.e., the point where the fluctuation energy is minimum and the main pulse energy is maximum. It is possible that the fluctuation ratio has no local minimum. In that case the current tuning step should be skipped.

5.6.4 Generality and limitation of the algorithm

The algorithm is plausibly general—in the sense that it can be applied to any T-ray signal with minimal disturbance to desired signal features. This generality is due to the following facts:

1. A T-ray spectrum is altered by the algorithm only at the frequencies at which water-vapour absorption lines are situated. The spectral resonances of other polar gases, occupying narrow frequency bands and usually not overlapping with water resonances, are not disturbed. In addition, the broad spectral features of a time-domain transient is unlikely to be significantly affected by narrow water line removal. The reflections, for example, which exhibit fringes over a broad frequency range, are not affected by narrow band resonance removal.
2. Uncorrelated fluctuations are ignored. Because the algorithm senses the fluctuation change, caused by strength tuning, rather than the fluctuation itself, any unrelated fluctuations cannot deceive the algorithm.
3. The causality of a signal is preserved. The absorption and dispersion profiles, used in the model of water vapour response, exactly comply with the Kramers-Kronig relation. Deconvolution of the measured spectrum by this causal response would yield the causal spectrum.

A limitation of the algorithm is when there is the overlapping of sharp resonances, i.e., a resonance of unrelated gas species exactly collocates with that of water vapour. In such a situation, both resonances are removed. Despite this limitation, the algorithm is nevertheless applicable to a large number of specific cases where overlap does not significantly occur. Scenarios of interest can always be benchmarked against sealed-chamber measurements in the laboratory.

5.7 Results

In this section, the proposed algorithm is tested with experimental T-ray data. The results from the algorithm are benchmarked directly with the experimental data recorded using a sealed chamber, which is an ideal case. This benchmarking is only to demonstrate the efficacy of the algorithm.

5.7.1 Free-path measurement

The first T-ray signal, subject to the water-vapour removal algorithm, is measured in free space without the presence of any material. The signal has a temporal resolution of

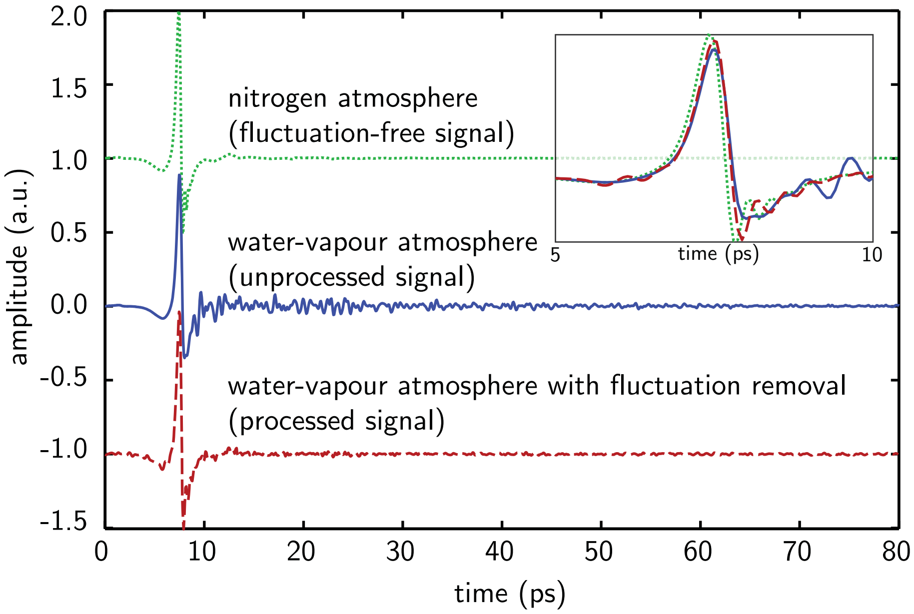
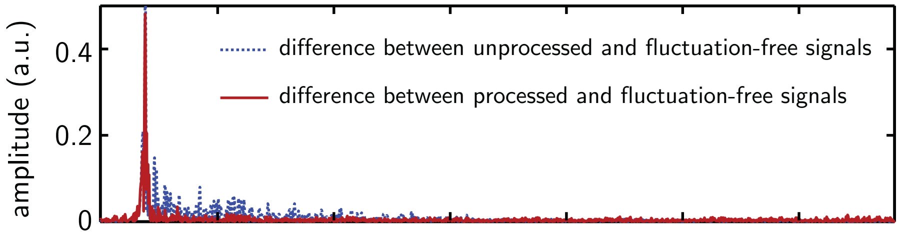


Figure 5.10. T-ray signals measured in a free path setting. The sampling interval is 66.7 fs with the total duration of 136.63 ps (only the first 80 ps is shown here). The inset shows a zoom-in of the signals between 5 and 10 ps. The fluctuation ratios of the fluctuation-free, unprocessed, and processed signals are 2.2, 9.2, and 3.3, respectively. The signals are vertically offset for clarity.

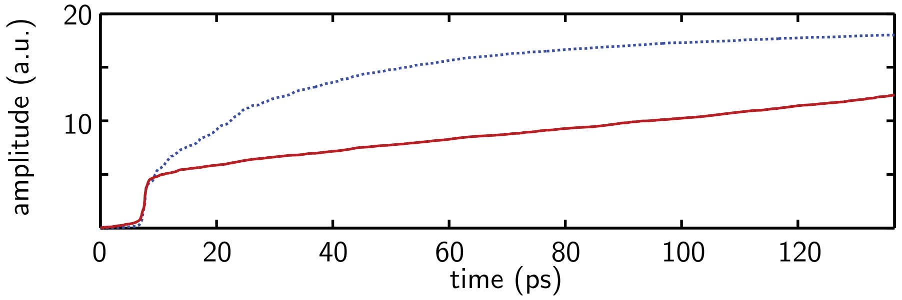
66.7 fs and a total duration of 136.63 ps, providing a spectrum with a spectral resolution of 7.3 GHz.

The strength-tuning algorithm is carried out with a set of complex resonances in the frequency range between 0 and 4 THz, where the T-ray magnitude is relatively high. The sequence, in which the resonances are interrogated and removed, follows the order of line strength, i.e., from high to low strengths. Only these resonances that have a strength higher than 0.01 of the maximum strength are inspected. The FWHM of the Gaussian window is 3 ps, and the iteration of the algorithm is set to five.

As shown by the processed signal in Figure 5.10, the algorithm can significantly reduce the time-domain fluctuations, which are previously located immediately after the main pulse, i.e., after 10 ps, of the unprocessed signal. The main pulses in the inset clearly illustrate the similarity between the fluctuation-free signal and the processed signal, despite a temporal shift that is not accounted for by the algorithm. The differences between the unprocessed/processed signal and the fluctuation-free signal are shown in Figure 5.11. The cumulative difference in Figure 5.11(b) clearly shows that the fluctuations after the main pulse are greatly reduced. In Figure 5.12, the spectrum of the



(a) signal difference



(b) cumulative difference

Figure 5.11. Signal differences for free-path measurement. (a) The absolute signal differences are calculated from the T-ray signals in Figure 5.10. The large amplitude around 8 ps in both cases is due to a temporal shift of the main pulse. (b) The cumulative differences of (a). The differences and their cumulative differences represent well the remaining fluctuation after the main pulse.

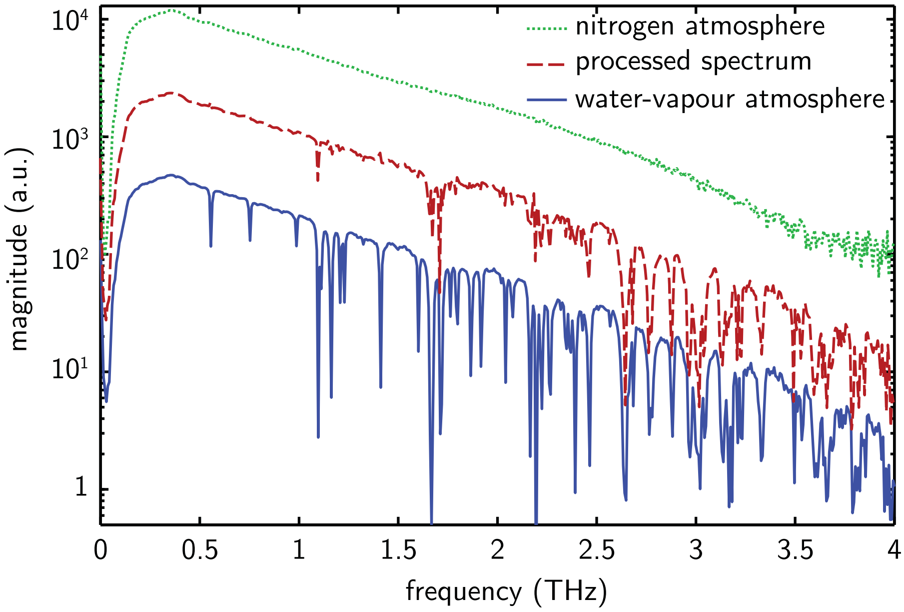


Figure 5.12. T-ray spectra measured in a free path setting. The spectral resolution is 7.3 GHz. Up to 2.5 THz most of the water resonances are removed, or at least reduced in strength, by the proposed algorithm. The magnitudes are offset for clarity.

processed signal demonstrates the success of the algorithm in removing, or reducing the strength of, the resonances in the frequency range from 0 to 2.5 THz. However, some resonances still persist, where there is spectral crowding or where the resonances are buried in noise

5.7.2 DL-phenylalanine measurement

The second experiment is performed with a signal measured from a sample of DL-phenylalanine at 220 K. The cryostat containing the sample is evacuated. The interacting water vapour is in the air in the remaining T-ray beam path. The total scan is 68.3 ps at the step size of 66.7 fs. The spectral resolution is 14.6 GHz.

This sample is selected because: (i) it has a typical distribution of vibrational modes, and (ii) the recorded signal particularly contains the reflections, caused by the cryostat's windows, made from a cyclo-olefin copolymer (Topas). Both cases are general to most spectroscopic samples, and hence this is to demonstrate that the method can be potentially applied to a large number of cases. However, a complete generality of the algorithm is the future challenge.

The parameters set for the algorithm are similar to the previous test, except for the interrogated frequency range that is in between 0 and 2 THz.

The processed signal is shown in Figure 5.13, along with the original. Again, it is obvious that the fluctuations located after the main pulse, beyond 15 ps, are remarkably reduced, in spite of an introduction of a small oscillation at the beginning. The reflection at 63 ps (point B) is not disturbed by the algorithm, and, interestingly, the reflection at 32 ps (point A), which was initially buried in fluctuations, is now recovered. Figure 5.14 provides a further comparison between the fluctuations of the unprocessed and processed signal. The cumulative difference in Figure 5.14(b) clearly illustrates the reduction of the fluctuation energy for the processed signal. In Figure 5.15, it can be clearly seen that for the processed spectrum, the algorithm can remove most of the H_2O resonances. However, a few resonances could not be completely removed. These persistent resonances are around 1.1 and 1.7 THz, exactly the same positions as the unremoved resonances in the previous case (see Figure 5.12). A closer look at 1.1 THz shows that two strong resonances overlap and merge together, but only one resonance is removed. This is because the removal of one resonance results in the deformation

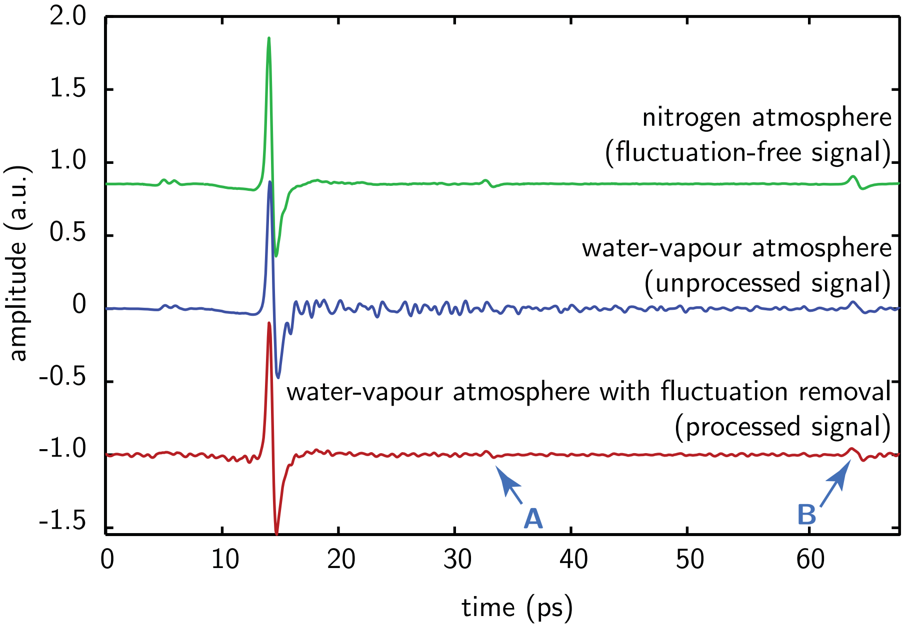
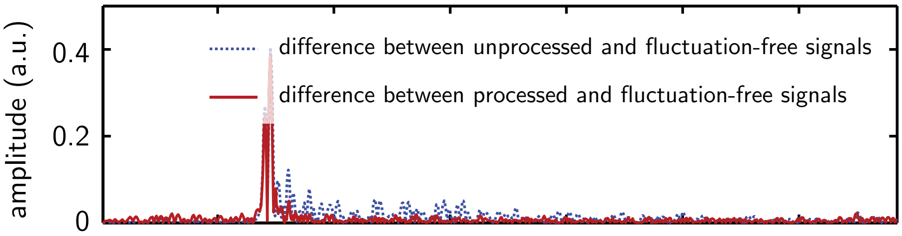
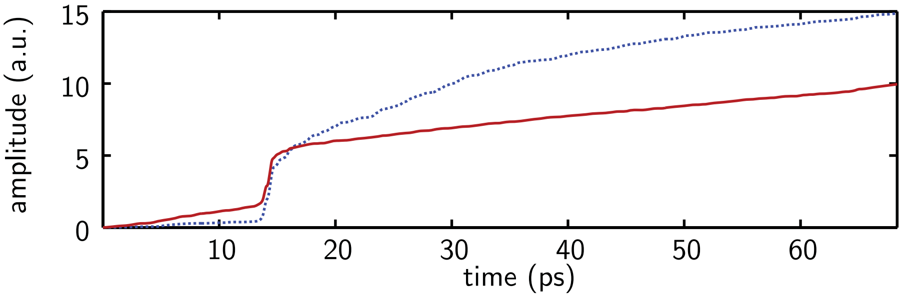


Figure 5.13. T-ray signals measured with DL-Phenylalanine sample in place. The sampling interval is 66.7 fs with the total duration of 68.38 ps. Interestingly, the reflections at 32 and 63 ps (points A and B), due to the cryostat windows, are clearly recovered from the fluctuations by our algorithm. The fluctuation ratios of the fluctuation-free, unprocessed, and processed signals are 2.6, 7.9, and 3.8, respectively. The signals are vertically offset for clarity.



(a) signal difference



(b) cumulative difference

Figure 5.14. Signal differences for DL-Phenylalanine measurement. (a) The absolute signal differences are calculated from the T-ray signals in Figure 5.13. The large amplitude around 15 ps in both cases is due to a temporal shift of the main pulse. (b) The cumulative differences of (a).

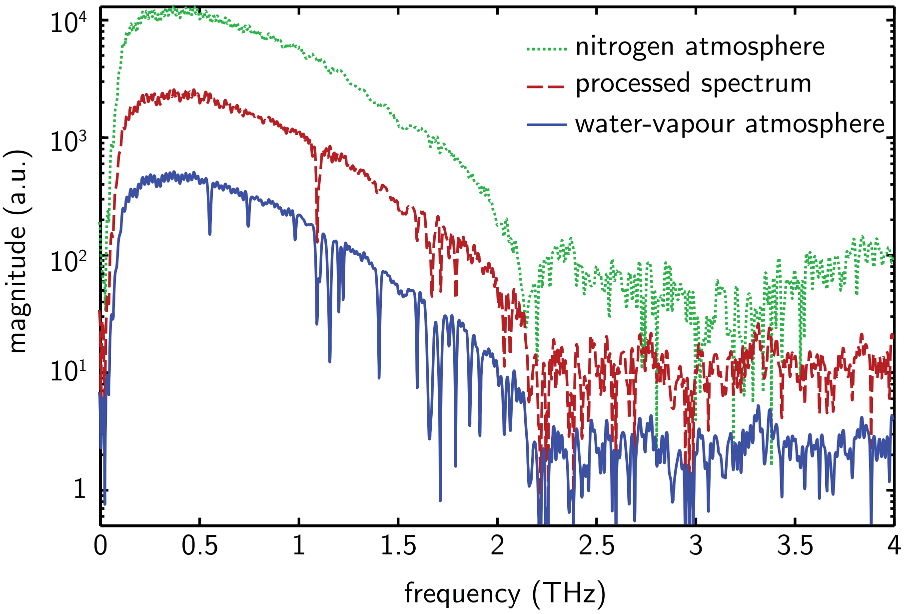


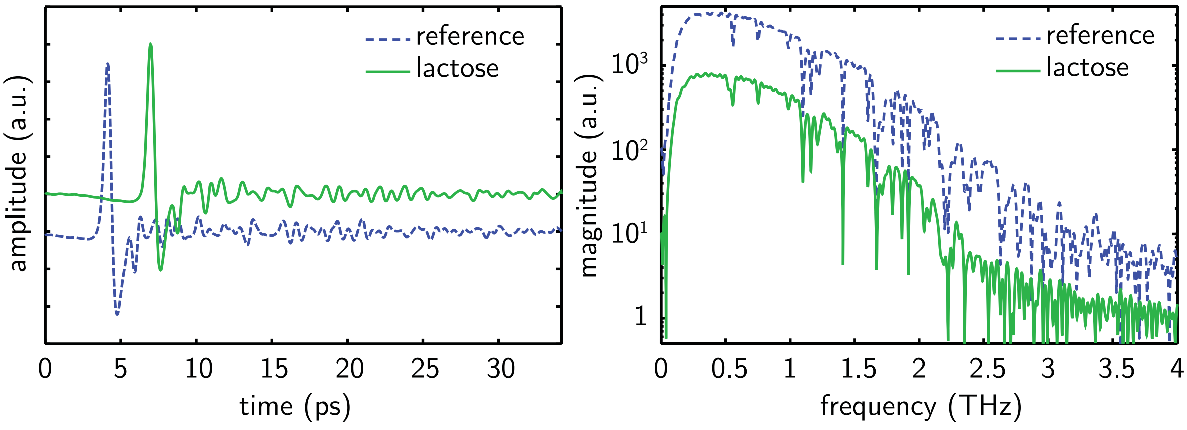
Figure 5.15. T-ray spectra measured with DL-Phenylalanine sample in place. The spectral resolution is 14.6 GHz. From 0 to 2 THz most of the resonances are removed, or at least reduced in the strength, by the proposed algorithm. The magnitudes are offset for clarity.

of the other resonance from a Lorentzian shape, and thus the remaining deformed resonance cannot be appropriately deconvolved by a Lorentzian model. In spite of that, the recovery of reflections suggests a promising application of the algorithm in T-ray range finding.

5.7.3 Lactose measurement

Another trial measurement, subject to water-effect removal and subsequent parameter extraction, is carried out with a 1.85-mm-thick lactose sample. The sample and reference signals have the time duration of 34.16 ps with the resolution of 16.68 fs. Without zero padding, the FFT provides the spectra with the finest resolution of 29.26 GHz—much wider than the estimated FWHM of a water resonance of 6 GHz. Figure 5.16 shows the signals and spectra of lactose and reference, both measured in water vapour atmosphere. The cutoff frequency is at 2 THz.

It is known that below 2 THz the lactose spectrum contains sharp and strong absorption features at 0.53 and 1.37 THz (Taday 2004, Fischer *et al.* 2005b, Ung *et al.* 2006), and the resonance at 0.53 has a FWHM of ≈23 GHz (Brown *et al.* 2007). These features are



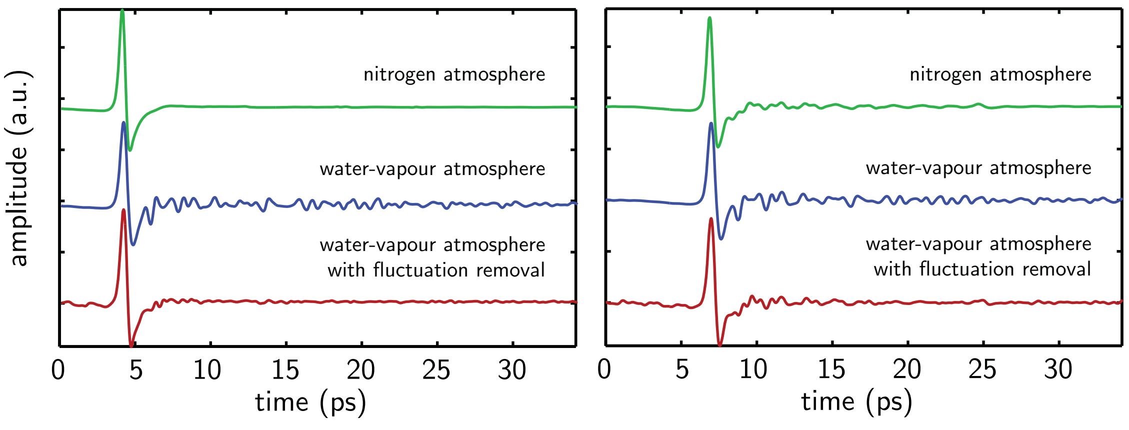
(a) T-ray signals (a) T-ray spectra

Figure 5.16. T-ray signals and spectra for reference and lactose. The sample measurement is carried out with a 1.85-mm lactose pellet in a water vapour atmosphere. Both the reference and sample signals have the time duration of 34.16 ps with the temporal resolution of 16.68 fs. The spectral resolution is 29.26 GHz. The signals and spectra are vertically offset for clarity.

reportedly associated with the intermolecular vibration modes of the hydrogen-bond network (Fischer *et al.* 2005b). The features overlap with two strong water-vapour resonances at 0.56 and 1.41 THz. Thus, the measurement poses three challenges against the success of the water removal algorithm—(i) the water resonances are poorly resolved by THz-TDS, (ii) and yet overlap with lactose's spectral features, (iii) which produce fluctuations in the time-domain signal as well.

For this trial, the water-vapour removal algorithm is applied to both the reference and sample spectra. The FWHM of the Gaussian window is 5 ps. The strength-tuning iteration is set to two. The interrogated spectral range is from 0 to 3 THz. Once the removal process is completed, the absorption coefficient of lactose is extracted.

The reference and lactose signals after applying the algorithm are shown in Figure 5.17. Here, it can be seen that most of the fluctuation energy induced by H_2O is suppressed by the algorithm. In Figure 5.18, the absorption coefficient of lactose extracted from the processed signals juxtaposes the coefficients extracted from signals measured in nitrogen and water-vapour atmospheres. The lactose absorption coefficient measured in water-vapour atmosphere contains only a few water vapour resonances, because the resonances are well above the noise level, and thus cancel themselves out in the direct deconvolution process. However, still some traces of water resonances can be observed at 0.04, 0.56, and 1.41 THz (marked by the arrows). The removal algorithm



(a) reference

(b) lactose

Figure 5.17. T-ray signals for reference and lactose in different conditions. The sampling interval is 16.68 fs with the total duration of 34.16 ps. For reference, the fluctuation ratios of the N_2-atmosphere, H_2O-atmosphere, and H_2O-removal measurements are 0.34, 5.48, and 0.69, respectively. For lactose, the fluctuation ratios of the N_2-atmosphere, H_2O-atmosphere, and H_2O-removal measurements are 1.16, 5.51, and 1.95, respectively. The signals are vertically offset for clarity.

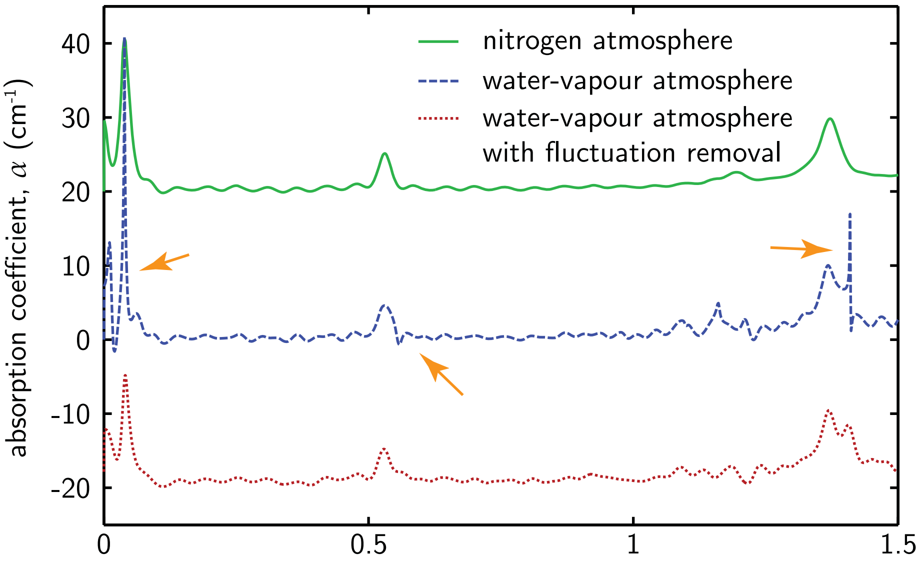


Figure 5.18. Absorption coefficients of lactose extracted from different signals. The orange arrows mark the traces of water lines at 0.04, 0.56, and 1.41 THz. The water-vapour removal algorithm can partially remove these traces with no affect on the spectral features of lactose. The curves are vertically offset for clarity.

Table 5.2. Performance evaluation of the strength-tuning algorithm. The fluctuation ratio and the mean squared error (MSE) are calculated using Equations 5.17 and 5.19, respectively. The fluctuation ratio and the spectral energy are normalised to those of the signal measured in nitrogen atmosphere. The notation $\mathrm{H_2O}$ and $\overline{\mathrm{H_2O}}$ indicates unprocessed and processed signals, respectively.

Signal	Fluctuation ratio		MSE (%)		Energy (a.u.)	
	$\mathrm{H_2O}$	$\overline{\mathrm{H_2O}}$	$\mathrm{H_2O}$	$\overline{\mathrm{H_2O}}$	$\mathrm{H_2O}$	$\overline{\mathrm{H_2O}}$
Free path	4.2	1.5	0.076	0.047	0.90	0.96
Phenylalanine	3.0	1.5	0.13	0.085	0.90	0.95
Lactose reference	16.1	2.0	0.31	0.19	0.90	0.98
Lactose sample	4.8	1.7	0.31	0.21	0.93	0.99

helps remove most of the water line remnants, yet preserves lactose's distinctive features at 0.53 and 1.37 THz. A small peak caused by water vapour remains beside the 1.37-THz lactose peak.

5.7.4 Performance evaluation of the strength-tuning algorithm

Table 5.2 shows the fluctuation ratio and mean squared error (MSE) of the signals and the total energy of the spectra, investigated in the previous subsections. The fluctuation ratio helps to measure a change in the fluctuation energy and main pulse energy, before and after the signals are processed. The MSE evaluates the fitness of the improved signals to a fluctuation-free signal (N_2-atmosphere measurement). The MSE is given by

$$\mathrm{MSE} = \frac{1}{m}\sum_{m}(\acute{y}_m - y_m)^2 \, , \tag{5.19}$$

where y is a fluctuation-free signal and $\acute{y}$ is a compared signal. The error is summed and averaged over m temporal points. The total spectrum energy shows the recovery of the energy from resonance absorptions. Note that the fluctuation ratio and the spectral energy are normalised to those of a fluctuation-free signal.

From the table the fluctuation ratio indicates that the fluctuation energy is greatly reduced in all cases, once the removal algorithm is implemented with the unprocessed signal. However, the MSE still reflects a considerable difference between the processed signals and the fluctuation-free signals, in spite of the great improvement visualised in the previous subsections. This would probably be due to the whole signal shift in the time domain, caused by a constant refractive index difference unaccounted for by the algorithm.

The total spectral energy in Table 5.2 reveals that the algorithm is successful in recovering a part of the signal's energy, which is absorbed by water vapour resonances during propagation. However, not all the energy is recoverable. With regard to the processed spectra shown in the previous subsections, the algorithm cannot completely remove the water vapour resonances in the region where the SNR is low and where the spectral lines are crowded. These two cases are analysed separately as follows.

At a low SNR region, distortions in the absorption and dispersion resonances by noise cause deviation of the model from the measurement. This situation is demonstrated in Figures 5.3 and 5.4. Applying the removal algorithm to a low SNR region would be effective, provided that the distortion is not severe. Otherwise, the large deviation inhibits the removal via a sub-optimal value of the fluctuation ratio. A possible means to alleviate the noise issue is to implement a signal enhancement methodology such as wavelet denoising (Ferguson and Abbott 2001a) prior to the removal process; this might better recover water resonances and subsequently improve the removal efficiency.

Crowded resonance features result in overlapping and merging between two or more resonances. If the spectral resolution is too low, these blended resonances are rendered incorrectly, resulting in a distorted spectral profile. The algorithm would see them as a single stronger and wider resonance, and as a result, the removal does not perform well. Improving the spectral resolution would allow better removal performance. For example, the free space measurement in Figure 5.12 shows a better result than the DL-Phenylalanine measurement in Figure 5.15 in the regions around 1.1 and 1.7 THz; the former and latter measurements have the spectral resolutions of 7.3 GHz and 14.6 GHz, respectively. In addition to the spectral resolution problem, blended resonances cannot be represented exactly by a simple sum between two Lorentzian lines (Pickett 1980).

5.8 Conclusion and potential extensions

A numerical algorithm for removal of water-vapour effects in T-ray measurements is investigated in this Chapter. Given only a sample signal with no reference, theoretically, the removal can possibly be carried out by simple deconvolution of the model complex water-vapour response from the signal. However, many factors limit the usability of the simple deconvolution. An exact model of water vapour resonance requires many physical parameters, including the pressure, temperature, humidity, and propagation length. In addition, if the noise level is sufficiently high such that some strong resonances are distorted, the whole deconvolution process cannot be performed.

The proposed algorithm fine-tunes the strength of each model spectral resonance. A criterion for strength tuning is met when the fluctuation ratio of the processed signal reaches the minimum. Once an optimal resonance is attained, it is removed from the signal. The algorithm then proceeds to removal of the next resonance. This tuning scheme relaxes the requirement for precise information of the delicate physical conditions during the measurement. An extension of the algorithm to cope with changes in the linewidth is possible by performing fine tuning of a resonance in the 2D space (linewidth and line strength spaces). Furthermore, the fluctuation ratio criterion inhibits any faulty deconvolution that might occur when the noise disturbs the quality of a measured signal.

In the experiments, the algorithm produces promising results. It can reduce a significant amount of vapour-induced fluctuations in the time domain signal and spectral resonances with small disturbances to other non-related features, such as reflections or sample-induced resonances. Moreover, we have demonstrated with experimental data that some features, which are initially masked by fluctuations, can be recovered by the algorithm. Optical constants extracted from these signals are demonstrated to hold their known values, within acceptable precision.

The proposed algorithm is general in the sense that any other polar gas response could be targeted, in principle, and hence removed from a measured spectrum if desired. It is not necessary that the parameters of a targeted gas be present in a spectroscopic catalogue, as long as its pure response in the frequency range of interest is fully estimable. This scheme might have benefits in some particular situations, for instance, where molecular contaminant(s) are unavoidable in a measurement.

Finally, it should be noted that the algorithm is not meant to serve as a full replacement to the 'sealed chamber' procedure executed during the measurement stage. Rather, it assists a measurement in cases where a sealed chamber is not feasible. Although the algorithm can apply to a large number of specific cases, it has its own limitations. Further improvement of the algorithm remains an attractive challenge.

The next chapter in Part I presents an application of machine learning to classification of terahertz signals probing different samples. The application of signal classification can be regarded as another signal processing scheme that is important to the development of terahertz technology.