Heuristic approach to take up the challenge of terahertz time-domain spectroscopy for biology

Melanie Lavancier

To cite this version:

Melanie Lavancier. Heuristic approach to take up the challenge of terahertz time-domain spectroscopy for biology. Electronics. Université de Lille, 2021. English. NNT : 2021LILUN008 . tel-03539445

HAL Id: tel-03539445
https://theses.hal.science/tel-03539445
Submitted on 21 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ÉCOLE DOCTORALE SCIENCES POUR L'INGÉNIEUR (SPI-ENGSYS)

UNIVERSITÉ DE LILLE

DOCTORAL THESIS

Heuristic approach to take up the challenge of terahertz time-domain spectroscopy for biology

A thesis submitted in fulfilment of the requirements for the degree
of Doctor of Philosophy of the University of Lille, written by

Mélanie LAVANCIER

Presented on October 15, 2021

Jury:

M. Frederic GARET	*Université de Savoie, IMEP-LAHC*	Reviewer
M. Guilhem GALLOT	*Ecole Polytechnique, LOB*	Reviewer
Mme. Dominique COQUILLAT	*Université de Montpellier, L2C*	President of the jury
Mme. Sophie ELIET BAROIS	*IEMN*	Examiner
M. Jean-François LAMPIN	*IEMN*	Thesis supervisor
M. Romain PERETTI	*IEMN*	Thesis advisor

Institut d'Electronique, Microélectronique et Nanotechnologie (IEMN)

École doctorale Sciences Pour l'Ingénieur (SPI-ENGSYS)

Université de Lille

Thèse de doctorat

Approche heuristique pour relever les défis de la spectroscopie terahertz dans le domaine temporel pour la biologie

---

Thèse préparée et soutenue publiquement, pour obtenir le grade de Docteur de l'université de Lille en sciences pour l'ingénieur, par

Mélanie LAVANCIER

Presentée le 15 Octobre 2021

Jury:

M. Frederic GARET	*Université de Savoie, IMEP-LAHC*	Rapporteur
M. Guilhem GALLOT	*Ecole Polytechnique, LOB*	Rapporteur
Mme. Dominique COQUILLAT	*Université de Montpellier, L2C*	Présidente du jury
Mme. Sophie ELIET BAROIS	*IEMN*	Examinatrice
M. Jean-François LAMPIN	*IEMN*	Directeur de thèse
M. Romain PERETTI	*IEMN*	Encadrant de thèse

Institut d'Electronique, Microélectronique et Nanotechnologie (IEMN)

Abstract / Résumé

There is growing interest for the terahertz (THz) frequency range due to several specificities. Located at the highest frequencies reached in electronics, there are applications for faster data bit rate in telecommunications. Moreover, due to the low energy of the photons, there are also applications in security, non-destructive monitoring for the industry and art conservation. Since this radiation is non-ionizing, another field of interest is biology.

These last two decades have seen the rise of published articles in the literature, made possible by the advances in terahertz time-domain spectroscopy setups. Many studies have been conducted on sugars, solvents, and even DNA and proteins. However, there is still no consensus in the Terahertz (THz) community on the experimental or data analysis protocols to implement.

This thesis displays three advances towards this direction. First, it presents a reliable and quantitative time-domain fitting process with a comparison metric that makes it possible to trust the results obtained and compare them to the theory. This enables us to move forward and analyze more and more complex samples, to come closer to *in vivo* conditions for biological samples. Indeed, I was able to provide more insight on the analysis of liquid water, which has been a great topic of interest since the beginning of THz time-domain spectroscopy, and on the influence of ions and proteins on the structure of its network. Furthermore, we made it possible for smaller samples, either in size or in quantity, to be analyzed as well thanks to our butterfly device. The protocols as well as the fitting software can still be improved, but they provide trustworthy results and give a criterion for comparison of data. We foresee that this will lead the way to a better understanding and more intricate experiments with THz time-domain spectroscopy for biology.

Le domaine TeraHertz (THz) est de plus en plus attractif grâce à ses propriétés particulières. Il se situe à la limite haute des fréquences pour l'électronique et entre naturellement en jeu pour de nombreuses applications en télécommunications. De plus, la faible énergie du photon THz le rend non ionisant, ce qui crée des applications dans des domaines variés tels que la sécurité, le contrôle non-destructif dans l'industrie, ou la conservation d'art. Un autre important domaine d'application est la biologie.

Le nombre d'articles publiés dans la littérature ces vingt dernières années n'a fait qu' augmenter, ce qui a été rendu possible par les avancées sur les dispositifs de spectroscopie THz dans le domaine temporel. De nombreuses études ont été réalisées sur des sucres, des solvants, ainsi que sur de l'ADN et des protéines. Cependant, il n'y a toujours pas de consensus dans la communauté THz concernant les protocoles expérimentaux et d'analyse des données à implémenter.

Cette thèse expose trois avancées dans cette direction. Tout d'abord, elle présente un protocole de "fit" (ou d'ajustement des données) dans le domaine temporel fiable et quantitatif, ainsi qu'un critère de comparaison, ce qui permet une plus grande confiance dans les résultats obtenus et de les comparer à la théorie. Nous pouvons donc avancer et analyser des échantillons toujours plus complexes, de manière à être de plus en plus proche des conditions *in vivo* pour les échantillons biologiques. En effet, j'ai pu fournir de nouvelles perspectives à l'analyse de l'eau liquide, qui est un sujet phare depuis les débuts de la spectroscopie TeraHertz à domaine temporel, et sur l'influence des ions et des protéines sur la

structure de son réseau. En outre, nous avons rendue possible l'analyse de petits échantillons, que ce soit en taille ou en quantité, grâce à notre dispositif appelé "papillon". Les protocoles ainsi que le logiciel de fit peuvent encore être améliorés, mais ils ont permis de fournir des résultats fiables et donnent un critère pour la comparaison d'information. Nous espérons que ce travail ouvrira la voie à une meilleure compréhension et à de plus complexes expériences sur les échantillons biologiques grâce à la spectroscopie TeraHertz dans le domaine temporel.

Acknowledgments

Je souhaite avant tout remercier Romain Peretti pour ton encadrement et ton soutien durant ces trois années de thèse. Merci de m'avoir aiguillée et aidée, ce qui m'a permis de progresser tout en prenant mon indépendance petit à petit. Merci aussi pour ton enthousiasme et ta gentillesse qui m'ont permis de rester motivée même quand les résultats n'étaient pas encore là. Je garde tes conseils en tête précieusement, je suis certaine qu'ils m'aideront dans cette nouvelle étape. Je remercie aussi Jean-François Lampin pour les connaissances que tu m'as apportées sur les points théoriques, pour les discussions, et pour ta disponibilité quand je venais toquer à ton bureau.

Un grand merci aussi aux membres du jury pour le temps que vous avez accordé à rapporter et examiner mon travail. Merci à Frédéric Garet et Guilhem Gallot d'avoir accepté d'être rapporteurs et d'évaluer mes travaux de thèse. Merci également à la présidente du jury Dominique Coquillat, ainsi qu'à Sophie Eliet qui ont accepté d'examiner ma thèse.

Je te remercie aussi Sophie pour tes conseils de manips, les discussions sur les aspects de caractérisation, mais aussi les discussions musique. Merci aussi à Flavie Braud d'avoir passé tant de temps à optimiser le design des papillons et à trouver la meilleure manière de les découper. Merci à David Guérin pour tes conseils de chimie et de choix des échantillons ioniques. Merci à Coralie Bompard pour tes conseils sur la cristallisation du lysozyme ainsi que pour ta disponibilité et les matériaux que tu nous as fournis.

Je remercie aussi le personnel de l'IEMN, Nora et Andy pour leur aide et leur efficacité sur les points administratifs, ainsi que Valérie pour l'organisation des déplacements (malheureusement réduits ces temps-ci).

Merci à Nabil, Juliette, Clémence et Eliot, que j'ai eu le plaisir de co-encadrer pendant leur stage et dont la motivation et l'envie d'apprendre m'ont inspirée à mon tour.

Merci à l'équipe Photonique THz pour vos questions et points de vue lors des présentations, pour vos nombreux partages de connaissances et pour les "boissons et snakes" après les réunions.

Merci à la team escalade ! Kevin, même si tu nous énervais tous à pouvoir te soulever sans problème et que tu progressais plus vite que nous en bloc, on a passé de supers séances ensemble. Jean-Marc, pour toutes les séances faites à progresser ensemble, pour ta motivation à aller dans les bars avec moi, ton humour et ta force mentale. Théo, même si je me demande toujours si tu as un réel instinct de survie en escalade, merci de m'avoir suivie dès que j'avais envie d'y aller, merci aussi pour les soirées bars et guitar hero, et surtout de m'avoir supportée dans le bureau... Kathia, même si c'était pas facile de te motiver on a fini par te voir à l'escalade ! Et merci pour les séances shopping entre filles et les pauses café pour discuter. Cybelle, merci pour ta bonne humeur, tes blagues et toutes les sorties (et maintenant les voyages à Grenoble), c'était un plaisir de venir t'embêter pour te faire sortir de ton bureau. Je n'oublie pas non plus Ariel, qu'on a rendue addict à l'escalade. Et évidemment Arthur, merci pour ta gentillesse et les repas à l'IEMN entre deux confinements. En bref, je suis extrêmement chanceuse de vous avoir rencontrés, vous avez rendu mes trois années à l'IEMN plus belles et ça me fait bien bizarre maintenant que la plupart d'entre vous êtes partis.

Évidemment, merci aux petits nouveaux d'égayer mes repas du midi et pour les afterworks (principalement au beersquare...): Valentin, Walter (qui me supporte aussi au quotidien dans le bureau..), Adrien, Elsa, Adèle, Alexandre, Alexandra, Eliot, Roman, Victor, Christiane et Jeyan. Et merci à Burger Party pour toutes ces belles soirées et ces vacances !

Enfin, merci à mes parents pour leur amour et leur soutien. Je n'en serais pas arrivée là sans vous, vous m'avez poussée à m'améliorer, vous avez cru en moi et m'avez aidée à atteindre mes objectifs. J'espère que vous savez à quel point je suis reconnaissante de vous avoir comme parents. Merci à Adrien, le meilleur petit frère, d'être là et de me supporter depuis tant d'années. Pour finir, merci à Romain, mon amour, qui m'a supportée, soutenue, aidée, et qui a rendu ces mois de rédaction beaucoup plus supportables. Je suis extrêmement chanceuse de t'avoir dans ma vie.

Contents

Abstract / Résumé **iii**

Acknowledgments **vi**

General introduction **1**

0 Introduction **3**
- 0.1 Definition and applications of the terahertz domain . . . 3
- 0.2 A brief history of terahertz time-domain spectroscopy . . . 4
- 0.3 What is terahertz time-domain spectroscopy? . . . 5
  - 0.3.1 Principle . . . 6
  - 0.3.2 THz generation and detection . . . 7
    - Generation of broadband THz radiation using photoconductive antennas . . . 8
    - Detection of broadband THz radiation using photoconductive switches 9
  - 0.3.3 The delay line . . . 10
  - 0.3.4 Applications of terahertz time-domain spectroscopy . . . 10
- 0.4 Biology and THz . . . 13
  - 0.4.1 Molecular crystals . . . 13
  - 0.4.2 Water . . . 14
  - 0.4.3 Analyzing complex biological samples . . . 15

1 Recovering information in THz **time-domain spectroscopy (THz-TDS)** **17**
- 1.1 Information in a THz-TDS experiment . . . 17
  - 1.1.1 The transfer function . . . 17
  - 1.1.2 Extraction of optical parameters . . . 21
  - 1.1.3 Permittivity models . . . 21
    - The Drude model . . . 22
    - The Lorentz model . . . 23
    - The Debye model . . . 24
- 1.2 Retrieving information in the frequency domain . . . 26
- 1.3 Retrieving information in the time-domain . . . 28
  - 1.3.1 Optimization algorithms . . . 29
    - Swarm particle algorithms . . . 29
    - Gradient descent optimization algorithms . . . 30
    - Dual annealing algorithms . . . 30
  - 1.3.2 Validation of the method on a simulated sample . . . 30
- 1.4 Improving the extraction of information by taking experimental noise into account . . . 32

1.4.1 Implementing a new error function for the fitting process 32
1.4.2 Experimental noise is deterministic 33
Dark noise (absence of IR excitation on both antennas) 34
THz-dark noise 36
Delay noise 37
1.4.3 There is a small noise depending on the signal left 37
1.5 An overview of fit@TDS 39
1.6 An example: Analysis of an α-lactose monohydrate pellet 41
1.6.1 Experimental protocol 41
1.6.2 Retrieval of the noise matrices 42
1.6.3 The fitting process 43
1.6.4 The final fit 46
1.7 Conclusion 47

2 Studying liquids with THz-TDS **49**
2.1 Liquid water, the most basic liquid sample? 49
2.1.1 The unique structure of liquid water 49
2.1.2 Studying liquid water in the terahertz range 50
2.1.3 Experimental protocol 55
2.1.4 Experimental results 56
2.1.5 Fitting process 57
Comparing models at room temperature 58
Applying this model to other temperatures 62
2.2 Heavy water 62
2.2.1 Experimental results 63
2.3 Disruption of the H-bond network of liquid water with ions 65
2.4 Study of a protein: the lysozyme 67
2.4.1 Experimental protocol 68
Crystallization 68
THz-TDS experiment 69
2.4.2 Results 70
2.5 Conclusion 73

3 The challenge of analysing small samples with THz-TDS **75**
3.1 General introduction and motivations 75
3.1.1 Confining light for the analysis of biosamples 75
3.1.2 Objectives of the confinement device 77
3.2 The first realization of the "butterfly" device 77
3.2.1 Design of the device 77
3.2.2 Fabrication process of the butterfly 78
3.2.3 Performances 80
3.2.4 Test on α-lactose monohydrate 81
3.3 New version of the butterfly to achieve smaller volumes 83
3.3.1 The new fabrication process 83
Laser cutting 83
Cleaning with hydrogen fluoride 84
Sputtering process 84

3.3.2 Blocking the parasitic pulse 85
3.3.3 Test on α-lactose monohydrate and glutamine 86
3.3.4 Understanding the coupling issue 87
Setups of the microscopy and far-field experiments 88
Results 89
3.4 Conclusion 91

4 Conclusion and perspectives **93**

Bibliography **97**

List of Figures

1 The terahertz domain in the electromagnetic spectrum . . . 3
2 Unit conversion in the THz frequency range [1] . . . 3
3 The water vapour spectrum: hand-traced by Rubens in 1921 using a mercury vapor lamp and a grating spectrometer and then the first one retrieved by THz-TDS in 1989. . . . 5
4 Schematic of the principle of THz-TDS . . . 6
5 An example of a THz pulse measured by a TDS setup (on the left) and its spectrum (on the right). The measurement is taken in a nitrogen purged environment without any sample between the emission and detection antennas. 7
6 Principle of THz radiation from optical rectification and photoconductive antennas . . . 7
7 Schematic of the emission of a THz pulse with a photoconductive antenna, in blue is the semiconductor substrate, in yellow are the two metal electrodes, in pink the incoming femtosecond pulse, in dark blue the exiting THz pulse. $\vec{E}$ is the radiated electric far-field, $\frac{\mathrm{d}\vec{J}}{\mathrm{d}t}$ is the time-derivative of the current density. 8
8 Schematic of the detection of a THz pulse with a photoconductive antenna, in blue is the semiconductor substrate, in yellow are the two electrodes, in dark blue the incoming THz pulse. The ammeter measures the photocurrent. $\vec{E}$ is the radiated electric field to detect, $\vec{J}$ is the current density. . . . 9
9 Absorption spectra of explosives compounds RDX (Cyclotrimethylenetrinitramine) and HMX (Cyclotetramethylene tetranitramine) covered with different materials (plastic, cotton and leather). The black curve is the transmission spectra of pure samples, the red one is the transmission spectra of covered samples, the blue one is the transmission spectra of the covers, and the green one is the sum of pure samples and covers. The THz spectrum of ambient air is used as reference. [36] . . . 11
10 On the left is the THz image of a fresh leaf. On the right is the same leaf 48 hours later. The THz radiation is highly attenuated by the water inside the leaf, which explains that the dryer the leaf is the less absorption there is. This is illustrated by the colorbar on the right. [47] . . . 13

1.1 Propagation of the electric field through material 2 surrounded by two different materials (1 and 3). e_{Fi} is the forward propagating field, e_{Bi} is the backward propagating field, T_{ij} is the transmission matrix between materials i and j, R_{ij} is the reflection matrix at the interface between materials i and j, and d_2 is the thickness of the second layer. . . . 18

1.2 Propagation of the electric field through a 3-layer sample. e_{Fi} is the forward propagating field, e_{Bi} is the backward propagating field, T_{ij} is the transmission matrix between materials i and j, R_{ij} is the reflection matrix at the interface between materials i and j, and d_i is the thickness of the i^{th} layer. 20
1.3 Simplified drawing of the Drude model and an example of a permittivity curve calculated with it. 22
1.4 Simplified drawing of the Lorentz model. The electrons are represented in blue, the nucleus is the pink circle containing protons and neutrons. 24
1.5 Example of the permittivity of a Lorentz model with a plasma frequency ω_p of 100 THz, a resonance frequency ω_0 of 100 THz and a width of 10 THz. On the left is the real part of the permittivity, the imaginary part is on the right. . 24
1.6 (A) The water molecule, its dipolar moment is represented in blue. (B) Simplified drawing of the Debye model. The blue arrows represent the dipoles reorienting inside the sample. 25
1.7 Example of the relative permittivity calculated with a Debye model with $\Delta\epsilon$ = 1000 and a time constant of 0.01 ps. 26
1.8 The fitting process in the frequency domain 28
1.9 (A) Time-traces of the experiment. In black is the reference time-trace, in pink the time trace of the simulated sample, in blue dots the resulting fitted time-trace and in grey the difference between the pink and blue curves (times 100) and shifted for clarity. (B) The corresponding spectra to the time-traces on the left. 31
1.10 Permittivity retrieved with the frequency domain formula from section 1.1.2 (in pink) and retrieved from the fit (in blue). (A) is the real part and (B) is the imaginary part. 32
1.11 THz-TDS setup and its sources of experimental noise 34
1.12 The most contributing noises to the THz-TDS experimental noise. 35
1.13 Five measurements of the THz-dark noise, the shape is very repeatable. . . . 36
1.14 THz-dark noises 'A'and 'B', before and after the filtering process. On the left is the frequency domain and on the right the time domain. 37
1.15 Experimental noise recorded on our Menlosystems THz-TDS setup after having removed the contribution from other sources of noise and the term proportional to the signal derivative. A zoom between −5 ps and 5 ps is also provided. 38
1.16 Initialization tab of the software 39
1.17 Model parameters tab of the software 40
1.18 Optimization tab of the software 40
1.19 A pellet - disc composed of pressed powder - on the left, and its holder on the right. The holder was designed with Solidworks and built with aluminum in the laboratory workshop. 41
1.20 Time-trace and spectrum of the reference (*i.e.* the THz pulse goes through the holder without pellet inside, in black) and the sample (in pink). The time-trace measurement was actually made on a larger time range (between -80 and 120 ps but it is zoomed in for visibility. 42
1.21 Remaining noise after the pre-processing of the reference time traces (on the left) and the sample time traces (on the right) 43

1.22 (A) Noise convolution matrix (NCM) of the lactose experiment. I traced the logarithm for better contrast. The big cross corresponds to the contribution of the reference noise in the principal pulse and the Fabry-Perot echoes, which correspond to the oscillations after the main pulse on the sample time trace. (B) Plot of two of the columns of the noise matrix. In pink is column 1350, which is located in the middle of the cross, and in blue is column 5000, which is located far from the cross. 44
1.23 Results of the fit for the first two steps of the process. 45
1.24 Results of the fitting process, illustrating the parameters retrieved in table 1.2 47

2.1 Water molecule and its dipolar moment. In pink is the oxygen atom and in yellow the two hydrogen atoms. 49
2.2 Fit results for each model, on the frequency range studied, compared to a set of data from the literature [121] 53
2.3 Absorption coefficients of all the permittivity models proposed in the literature, over a wide frequency range (0.2 GHz-2000 THz), compared to data from [121]. 54
2.4 Experimental setup for the measurement of liquid water with THz-TDS. . . . 55
2.5 Experimental process for the study of liquid water with THz-TDS. 56
2.6 Time traces of the measurements of liquid water, of the reference, and the empty microfluidic cell. On the right is provided a zoom of the main pulses of the liquid water measurements. 57
2.7 Spectra of the liquid water measurements, of the reference and empty microfluidic cell retrieved via a Fourier transform. On the right, the spectra are normalized by the spectrum of the reference. 58
2.8 (A) Time traces of the measurement of liquid water at 20 °C, with a zoom on the main pulse of the liquid water trace. (B) Spectra retrieved from the time-traces. 58
2.9 Absorption coefficients retrieved after taking into account inertial effects in the model of the dielectric permittivity of liquid water. 60
2.10 Result of the fitting process for a Rocard-Powles-Debye model. (A) Time-traces of the experimental measurements and fitted time-trace. (B) Retrieved spectra of the measurements and the fitted time-trace. 61
2.11 Result of the fitting process for a Titov model. (A) Time-traces of the experimental measurements and fitted time-trace. (B) Retrieved spectra of the measurements and the fitted time-trace. 62
2.12 Time traces measured for the study of heavy water. In black the reference, in dark blue the empty microfluidic cell, and in other colors heavy water at different temperatures. On the right, a zoom on the main pulse of the heavy water measurements is provided. 63
2.13 (A) Spectra retrieved from the measured time traces for the study of heavy water. (B) Spectra normalized by the spectrum of the reference. 64
2.14 Comparison of the normalized spectra of D_2O and H_2O at all temperatures. . 64
2.15 Hofmeister series classification of ions. 65
2.16 Normalized spectra of the five ionic solutions for each concentration. Liquid water is also added in yellow for comparison. 66

2.17 Normalized spectra of the highest concentration for each studied ionic solution compared with liquid water. 67
2.18 (A) Take a coverslip and lay it flat. (B) Deposit 2 µL of lyzozyme solution. (C) Deposit 2 µL of stock solution. 69
2.19 Schematic of a Linbro plate well after all the steps of the protocol have been done. 69
2.20 Sample placed inside the THz-TDS setup to study the crystallized lysozyme in solution. On the left is the front view of the mount, on the right is the side view of the mount with the layered quartz wafers. 70
2.21 Photograph of the setup for the lysozyme analysis. The emitted pulse is collimated by a parabolic mirror and then focused by a TPX lens on the sample (the mount is offset on the photograph to take the reference time trace). The symmetric setup focuses the pulse onto the detector after the sample. 70
2.22 (A) Photograph of the lysozyme crystals, their diameter is approximately 50 µm. (B) Photograph of other lysozyme crystals, where the solutions had previously been stored in a refrigerator, resulting in bigger crystals of around 300 µm diameter. 71
2.23 Time traces measured for the small crystals (A) and for the big crystals (B). For each experiment were measured a reference time trace, the time trace of the solution without crystals and the solution alone. 71
2.24 (A) Spectra retrieved from the time traces measured. (B) Spectra of the crystalline solutions and crystals normalized by the spectrum of the reference. (C) normalized spectra of the crystals divided by the normalized spectra of the crystalline solutions. 72
2.25 Mean of the absorption coefficients of all the measurements, with the standard deviation in each case (with and without the presence of lysozyme crystals). 73

3.1 Expansion of the mode (in blue) in the vertical direction of a parallel-plate metallic waveguide, producing a strong astigmatism. 76
3.2 (A) Solidworks drawing of the light confinement device. L is the length of the waveguide, w the gap between the sides, t its thickness, and D the opening at the entrance of the tapered antenna. The inset provides a map of the electric field inside the waveguide. (B) The light confinement device fabricated and mounted on a micrometer displacement stage. 78
3.3 Characterization of the laser cutting process. **(a)** Assembly of the two wings, showing the TSLW slot. The minimum slot width of 30 µm is determined by two integrated spacers cut at each extremity of the wings, as shown in the inset. **(b)** SEM view of the laser cut sidewalls; **(c)** Magnified view of the surface roughness of region 1; **(d)** Magnified view of the surface roughness of region 2. 79
3.4 Schematic of the THz-TDS setup used for the characterization of the butterfly device. 80
3.5 (A) Power losses for various values of w (width), the negative values are due to diffraction. (B) Dispersion curves for various values of w (Figure from [150]). 81
3.6 Results of TDS experiments on lactose with the butterfly. **(a)** Time-domain results. **(b)** Associated calculated losses and effective refractive index. 82

3.7 Beginning and end result of the laser cutting process 83
3.8 (A) View of the stacked butterfly wings with an optical microscope after laser cutting. (B) View of the stacked butterfly wings with the optical microscope after the first ultrasonic bath. 84
3.9 (A) View of the stacked butterfly wings with the optical microscope after the HF bath. (B) View of the stacked butterfly wings with the optical microscope after the second ultrasonic bath. 84
3.10 (A) The wings of the butterfly device on their holder, attached to the the sputtering device holder. (B) View of the butterfly wings with the optical microscope after the sputtering process. 85
3.11 (A) New version of the butterfly in its plastic holder. (B) THz absorber blocker above and below the butterfly. 85
3.12 (A) Time traces measured when the butterfly device is opened to 50 µm versus when it is closed, after adding the absorbers. (B) Spectra retrieved when the butterfly device is opened to 50 µm versus when it is closed, after adding the absorbers. 86
3.13 Absorbance of three samples of α-lactose monohydrate. 87
3.14 Schematic of the far-field experiment. 88
3.15 Schematic of the microscopy experiment. 88
3.16 Image recorded by the THz camera of a variable diaphragm with the microscopy setup. 89
3.17 Image recorded by the THz camera of the beam inside the butterfly waveguide with the microscopy setup. On the right is a magnified image of the spot. 90
3.18 Far-field radiation pattern of the butterfly device, the x axis is shown in figure 3.11a. 90
3.19 Reconstructed far-field radiation pattern of the butterfly on the left. On the right is the result of the CST simulation for comparison. The green curve corresponds to a frequency of 5 THz, the red curve to 2.6 THz, and the black curve to 0.2 THz. 91

List of Tables

1.1 The optimized parameters for the possible models for the first two absorption peaks of lactose. 45
1.2 The optimized model parameters of lactose. Other retrieved information is: $\epsilon_\infty = \mathbf{2.739}$; thickness of the pellet is **599.5** $\boldsymbol{\mu}$**m** (coherent with the measured one) ; Average delay is $-17.5\,\mathrm{fs}$; Coefficients of the noise proportional to the signal and its derivatives are a = 3.393e-05 and b = 6.944e-04 ; Final Akaike criterion is **2.028e+06**. 46

2.1 Parameters for the modeling of the dielectric permittivity of water. 52
2.2 Parameters of all the tried models, found after fitting the time traces from the liquid water experiment. 59

General introduction

In the 16^{th} century, a controversy between D'Alembert, Euler and Bernoulli put the spotlight on the study of objects resonant vibrations. This led to many consequences, at very different scales, up until these days. Indeed, this enabled researchers to link an object's size to its vibration frequency; from the Earth's mantle, to violin cords, or to molecular bonds scanned with infrared spectroscopy. As for biological samples, their resonant frequencies lie in the terahertz range (from 0.1 to 10 THz). Moreover, terahertz waves cover photon energies that are orders of magnitude smaller (0.4 to 40 meV) than the visual spectrum. The low energy carried by such waves indicates that they are not ionizing and are overall considered to be weakly hazardous for fragile samples such as proteins, cells or viruses. Hence, there is a strong motivation for developing THz spectroscopy.

This technique is highly promising as it takes advantage of recent progress in THz systems (sources, detectors,...) and nanotechnological advances. It already has some applications in security, such as detection of explosives or contraband products, or in the food industry, with the detection of foreign bodies in food products, or even in art conservation, with the analysis of paint layers. However, there is some struggle when analysing biological samples.

Biological objects, such as proteins, cells, or viruses, in addition to having a complex molecular structure, change their conformation according to the environment they are in. Therefore, in order to observe their structural dynamics under conditions as close as possible to living conditions, they have to be kept in buffer solutions during the investigation. These solutions are mostly composed of liquid water, which is highly absorbing in the terahertz range, making it already difficult to have any usable signal. Therefore, there is a need to adapt the experimental setup first. Moreover, it is challenging in itself to analyze liquid water alone. There has been countless studies on liquid water, but, even if some progress has been made, its dynamical behavior is still only partially understood, and there is no consensus on how to model the way the THz pulse is modified when going through it. As a result, we cannot dissociate the contribution of liquid water from the contribution of the biological samples inside the solution. Next, biological objects are often very small and available in small volumes, which means that the signal to detect should be very small and hidden by the high absorption of the buffer solution. Therefore, there is also work to be done on data analysis, in order to find a correct way to model liquid water's contribution, if we want to analyze biological samples. Finally, analyzing small volumes, of the order of the nL, in THz-TDS (Terahertz Time-Domain Spectroscopy) is needing a way to enhance light matter interactions between the THz pulse and the sample.

In this work, we will review the challenges undertaking such studies bring, whether it be from a data analysis point of view or the modifications of experimental protocols, or even new devices to help increase interactions with biological samples. The outline of this thesis is organized as follows:

- Chapter 0 is devoted to introducing the THz domain, its history, and the variety of applications it offers. In this chapter, we will review the principle of THz-TDS and the biological samples that were studied with it in the past.

- In chapter 1, we will discuss the concept of information in THz-TDS and how to retrieve optical constants from the recorded data. We will show that we can fit these data in the time-domain with permittivity models and present our fitting software fit@TDS. We will also discuss the importance of taking experimental noise into account when analysing THz-TDS data and illustrate it on a α-lactose monohydrate pellet sample.

- Chapter 2 will be focused on the analysis of liquids with THz-TDS. We will look at one of most discussed samples in the community, liquid water. We will review the permittivity models proposed in the literature, compare them when fitting our own data and look at other models to expand to higher frequencies. We will also check how temperature affects the models and therefore the microscopic processes in liquid water. Next, we will analyze a very similar sample, heavy water. We will also show that THz-TDS is sensitive to the disruption of the structure of water when components such as ions are added to it. Finally we will analyse a biological sample in liquid water: hen egg white lysozyme.
- Chapter 3 will present the challenge of analysing small biological samples and how to overcome this. We will present our light confinement device, from its first design to the new improved version, and show its usefulness for the analysis of biosamples with a volume of the order of the nanolitre. We will also present a quasi-optics setup for the characterization of the far-field radiation patterns of THz antennas or metamaterials, used on our device.
- Chapter 4 will present an overall conclusion as well as guidelines and perspectives for analysing even more complex samples with THz-TDS.

Chapter 0

Introduction

0.1 Definition and applications of the terahertz domain

The THz frequency range does not have a standard definition yet, but it is commonly defined as the region of the electromagnetic spectrum between 100 GHz and 10 THz (or between 3.33cm^{-1} and 333cm^{-1}). It is also called *far infrared* and is located between the microwave domain and the infrared: it is at the crossroads of electronics and optics, as illustrated in figure 1.

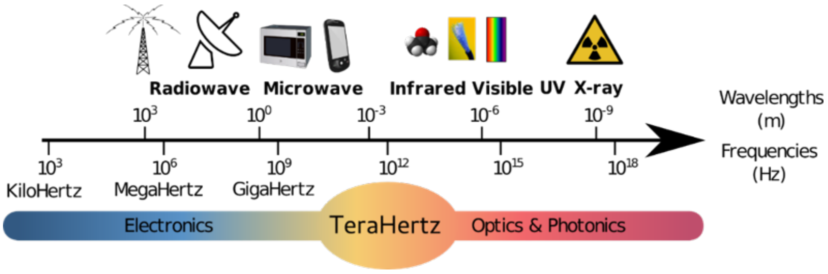


FIGURE 1: The terahertz domain in the electromagnetic spectrum

The terahertz domain being set between 100 GHz and 10 THz, it means that its wavelength borders are 30 µm and 3 mm, or 0.4 meV and 40 meV in energy (see figure 2).

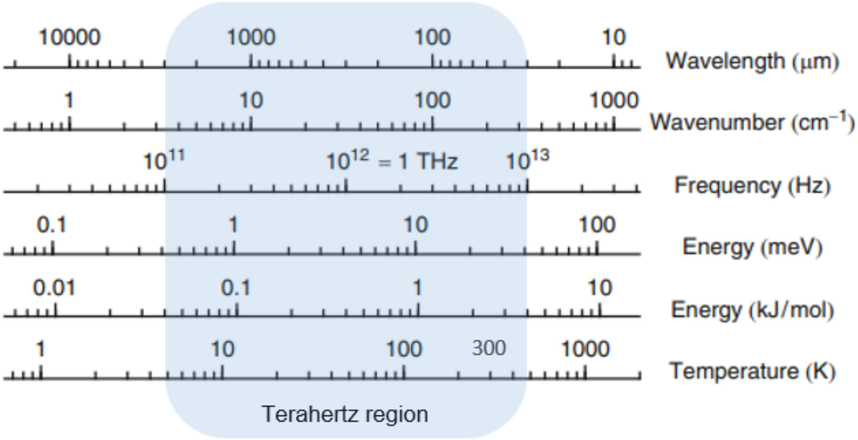


FIGURE 2: Unit conversion in the THz frequency range [1]

This part of the electromagnetic spectrum remains the least explored region mainly due to the technical difficulties involved in making efficient and compact THz sources and detectors. The lack of suitable technologies led to the THz band being referred as the "terahertz gap", which remained unexplored for decades. However, this technological gap has been rapidly diminishing for the last thirty years with the advances both in optics and electronics.

Nowadays, the THz region is an active research domain. Indeed, it is crowded by spectral features associated with fundamental physical processes such as rotational transitions of molecules, large-amplitude vibrational motions of organic compounds, lattice vibrations in solids, intraband transitions in engineered semiconductors, and energy gaps in superconductors. Therefore, it has applications in multiple domains:

Astronomy It is possible to identify the compounds of cosmic clouds, retrace their history and discover their residual lifetime [2]. It is also possible to image objects through dust disks. In fact, several satellites have been sent to space with onboard equipment that can detect terahertz and sub-millimetre signatures of celestial objects. Moreover, there are also THz radiotelescopes on earth, such as the Atacama Large Millimeter/submillimeter Array (ALMA) [3].

Industry It is used for non-destructive quality control. With THz, we can measure the drying time of car paint as well as its thickness for instance [4, 5].

Defense It is now possible to detect weapons through clothes or to detect the signatures of contraband products or explosives [6, 7].

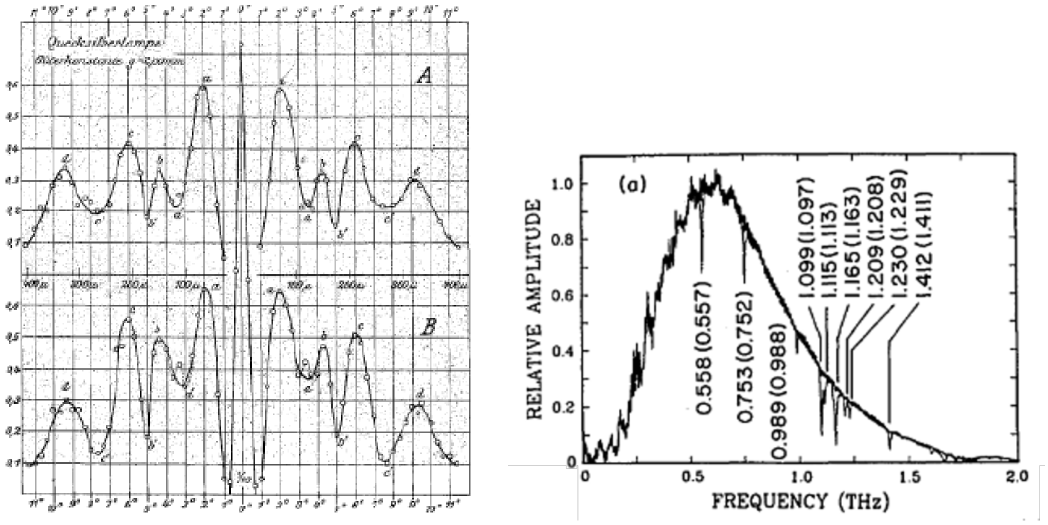
Telecommunications Since more than a century a general trend in wireless telecommunications is an increase in carrier frequency that allows higher data rates. Since ten years, systems moved from cm to mm wavelength and might reach the submm range which have free spectral ranges. Also, increasing the frequency of the telecommunications signals increases data speed. Hence, a few systems have demonstrated high speed in free space over tens of meters, which proves the feasibility [8, 9].

Biology Since the energy of the THz waves are of the same order of magnitude than the energies of molecular rotations and vibrations, we can identify materials or molecules[10–12]. One of the tools to perform such identification is terahertz time-domain spectroscopy.

0.2 A brief history of terahertz time-domain spectroscopy

While the term "terahertz" is quite recent, research in this frequency range dates back to more than a century. One of the pioneers of terahertz research is Heinrich Rubens, who spent 30 years studying the far infrared. Among more than 150 publications, he studied far infrared vibrations in crystals [13], gave a description of black bodies [14] and made the first measurement of a water vapour absorption spectrum [15] (see in figure 3a).

Progress in telecommunications was made in the 1940s with new techniques and new detectors that are the basis to current terahertz detectors using semiconductors. Then, the 1960s saw numerous breakthroughs such as the first CO_2 laser [17] or the first measurement of cosmic black-body radiation [18]. In 1968, the first TDS (time-domain spectroscopy), in the microwave range, was developed [19].



(A) Water vapour absorption spectrum measured and hand-traced by Rubens [15].

(B) Water vapour absorption spectrum retrieved with THz-TDS [16].

FIGURE 3: The water vapour spectrum: hand-traced by Rubens in 1921 using a mercury vapor lamp and a grating spectrometer and then the first one retrieved by THz-TDS in 1989.

For terahertz time-domain spectroscopy, research really started in the 1980s with the advent of picosecond and femtosecond lasers. Research groups led by Dan Grischkowsky (IBM Watson Research Center), David Auston (Columbia University), Gerard Mourou (Rochester University), and Martin Nuss (Bell Labs) were trying to generate and detect ultrashort electrical transients as they propagated down a transmission line. It was realized that these transmission lines were also radiating short bursts of electromagnetic radiation: this was the beginning of photoconductive antennas. In the late 1980s, reports were published wherein these radiated pulses were propagated through free space from a generator to a detector [20–22]. Thanks to these discoveries as well as mode-locked lasers, THz spectroscopy was born [23]. The first measurement of water vapour in the terahertz range, thanks to this technique, was made possible in 1989 [16] (see figure 3b).

Since then, advances in system designs and materials have promoted much growth in THz generation and detection and we can now find efficient commercial setups of THz-TDS that can scan physical and chemical properties of samples over frequencies ranging from 0.1 to 6 THz. Some research setups with very short pulses can even cover a frequency range up to 30 THz [24].

0.3 What is terahertz time-domain spectroscopy?

THz-TDS is a coherent technique with a high signal-to-noise ratio and good stability, so it can determine the complex electromagnetic response of a material accurately. It is also broadband, as it is one of the few techniques that cover several decades. It investigates physical and chemical properties of materials in the THz range. Moreover, since this frequency domain covers both rotational transitions from the microwave regime and vibrational modes from the infrared, multiple phenomena may be observed and abundant data can be extracted and interpreted. Polar gases will prominently exhibit rotational transitions; liquids typically have broad absorption due to strong interactions between molecules and hydrogen bonding (for polar liquids containing hydrogen) and lower-energy vibrations; and solids may either absorb broadly (if they are amorphous with polar bonds, such as glasses) or can have sharper-absorbing peaks (in the case of crystals) due to low frequency molecular vibrations or lattice vibrations (phonons).

0.3.1 Principle

THz-TDS is a spectroscopic technique in which the properties of matter are probed with short pulses of terahertz radiation. By measuring in the time-domain, this technique provides information on the sample's effect on both the amplitude and the phase of the terahertz radiation contrary to other conventional techniques such as Fourier-transform spectroscopy with the standard approach (sample in frond of the detector).

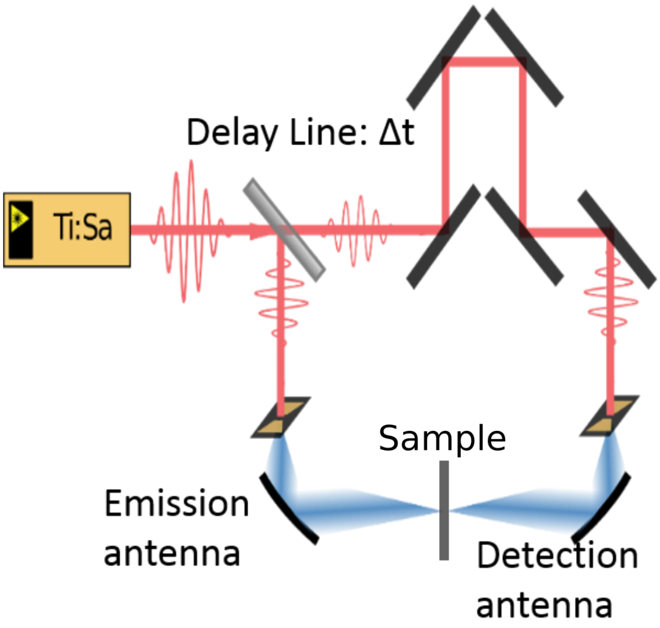


FIGURE 4: Schematic of the principle of THz-TDS

As illustrated in figure 4, a femtosecond laser (here a Ti:Sapphire) emits an ultrafast pulse train that is split in two by a beamsplitter and directed to photoconductive (PC) antennas. When the ultrashort pulse excites the emission antenna, it creates charges carriers that are accelerated by a DC electric field to create the radiated terahertz pulse. The delay line allows to adjust the path length of the probe pulse before reaching the detection antenna. The probe pulse strobes the detector, which is sensitive to the electric field when excited by the optical pulse. Hence, by varying the path length, the signal is reconstructed as a function of time with the same principle as an equivalent-time sampling oscilloscope. Finally, a single terahertz pulse will contain frequency components covering a large part of the terahertz range. However, the spectral resolution is limited as it is linked to the time-window of the measurement (limited by the travel range of the delay line) and ultimately by the repetition rate

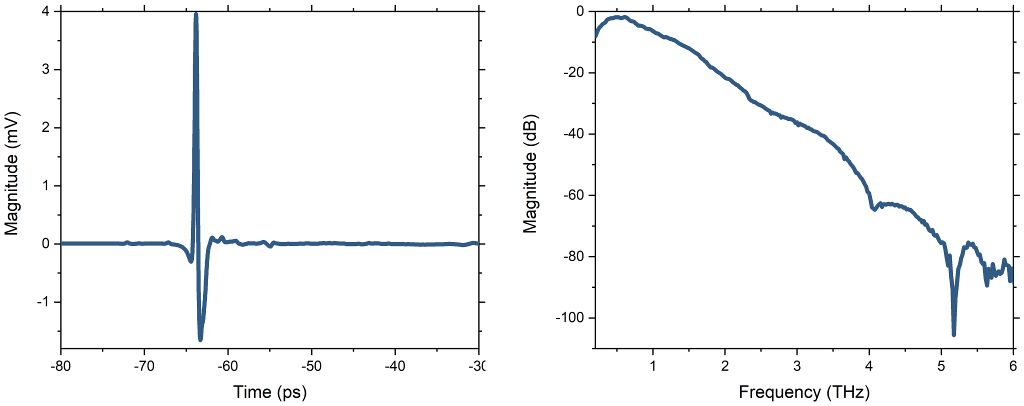


FIGURE 5: An example of a THz pulse measured by a TDS setup (on the left) and its spectrum (on the right). The measurement is taken in a nitrogen purged environment without any sample between the emission and detection antennas.

of the pulsed laser. An example of a time-domain pulse as well as its Fourier transformed spectrum can be observed in figure 5.

0.3.2 THz **generation and detection**

There are numerous techniques for the generation and detection of broadband THz radiation. A few examples are PC antennas, second-order non-linear electro-optic crystals, plasma oscillations and electronic non-linear transmission lines. In THz-TDS, the most commonly used are photoconductive and electro-optical rectification techniques schematized in figure 6. Since our setup contains PC antennas, I will focus on the generation and detection of THz radiation using this technique.

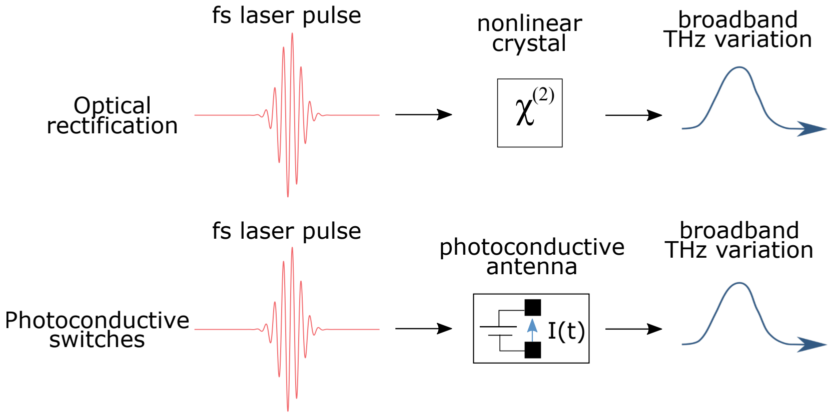


FIGURE 6: Principle of THz radiation from optical rectification and photoconductive antennas

Generation of broadband THz radiation using photoconductive antennas

A PC antenna, also called photoconductive switch, presented in figure 7, is formed by two metal electrodes separated by a gap of a few micrometers, both laid on a semiconductor substrate. A DC bias is applied between the electrodes. Focused on the gap, the laser beam triggers the creation of electron-hole pairs when the energy of the incident photons exceeds the bandgap energy of the semiconductor. The static bias accelerates the free carriers and, simultaneously, the charge density declines by trapping carriers in defect sites on the time scale of carrier lifetimes. The acceleration and decay of free carriers give rise to a photocurrent that is the source of the subpicosecond pulses of electromagnetic radiation.

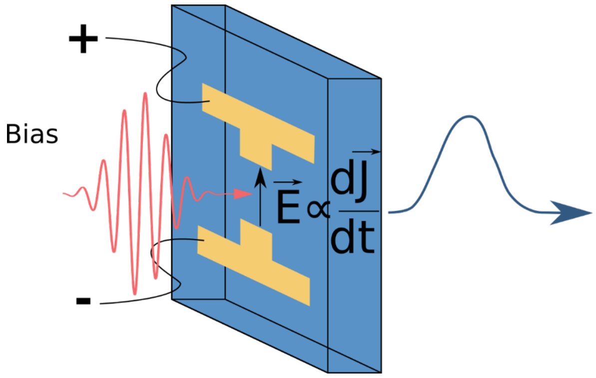


FIGURE 7: Schematic of the emission of a THz pulse with a photoconductive antenna, in blue is the semiconductor substrate, in yellow are the two metal electrodes, in pink the incoming femtosecond pulse, in dark blue the exiting THz pulse. E is the radiated electric far-field, $\frac{d\vec{J}}{dt}$ is the time-derivative of the current density.

The radiation source of a PC emitter can be modelled as a Hertzian dipole antenna whose size is smaller than the wavelength of the emitted radiation (depending eventually on the size of the metal electrodes). The radiated electric far-field is proportional to the time derivative of the current, I_p, (or equivalently to the current density J_p) as [25]:

$$E_{THz}(t) \propto \frac{\delta J_p(t)}{\delta t} \tag{1}$$

The photocurrent is generated by the transition of electrons from the valence band to the conduction band under laser illumination. Hence:

$$J_p(t) = -e \cdot n(t) \cdot v(t) \tag{2}$$

where e is the electron charge, $n(t)$ is the density of photocarriers, and $v(t)$ is their velocity. The same relation holds for the holes, but since their effective mass is much larger than that of electrons their contribution to the THz current and radiation is much smaller and can be neglected.

Several factors can affect the characteristics of the generated transient field, as well as its spectral range, such as the antenna structure and shape, the characteristics of its substrate, the bias voltage, and the duration of the pulsed laser. The bandgap energy of the substrate must match the laser pulse characteristics to trigger the generation of photocarriers. Moreover, The switch-on time is a function of the laser pulse duration, and the switch-off time is mainly determined by the photoexcited carrier lifetime in the semiconductor substrate of the antenna; therefore, in addition to a short laser pulse duration, a short carrier lifetime is a vital property for fast generation and particularly for detection. Since Ti:Sapphire femtosecond lasers with wavelengths around 800 nm are often used, the PC substrates are usually made of low-temperature grown Gallium Arsenide (LT-GaAs), of which its short carrier lifetime of around 1 ps or below is very suitable, and that can also operate at 1550 nm [26]. Other common semiconductor materials used are irradiated silicon and indium gallium arsenide (InGaAs).

Detection of broadband THz **radiation using photoconductive switches**

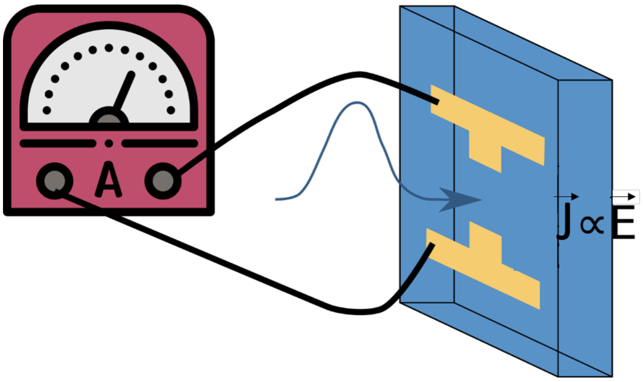


FIGURE 8: Schematic of the detection of a THz pulse with a photoconductive antenna, in blue is the semiconductor substrate, in yellow are the two electrodes, in dark blue the incoming THz pulse. The ammeter measures the photocurrent. $\vec{E}$ is the radiated electric field to detect, $\vec{J}$ is the current density.

The detection of THz waves by PC antennas relies on the same principle as their generation. The structure of the detector is similar to the PC antenna used as an emitter, as is shown in figure 8. A femtosecond laser beam is focused on the gap between the electrodes of the detector to generate photocarriers. In this case, the electrodes of the receiver are not biased. The generated photocarriers are accelerated by the THz electric field to be detected, meaning that the photocurrent is proportional to the THz electric field. The THz pulse shape is mapped out in the time domain by measuring the photocurrent while varying the time delay between the THz pulse and the optical probe. A current-voltage converter with high gain is necessary to convert the weak current signals into measurable voltages. In order to enhance the signal-to-noise ratio, the signal is processed by a lock-in amplifier synchronized with an optical intensity modulator such as an optical chopper. The current can be also directly measured without any lock-in amplifier, in this case the signal/noise ratio is enhanced thanks to the averaging of a lot of sweeps obtained with a fast delay line. The detected photocurrent at a time delay of t can be explained based upon Ohm's law:

$$J_{det}(t) = E_{THz}(t) * \sigma_{det}(t) = E_{THz}(t) * [e \cdot \mu_e \cdot n_r(t) + S_{\sigma_{det}}(t)] \tag{3}$$

where σ_{det} represents the conductivity of the detector, E_{THz} is the received THz signal, μ_e is the detector's semiconductor electron mobility, and $n_r(t)$ is the generated photocarrier density. Finally, $S_{\sigma_{det}}(t)$ corresponds to the shot noise of the detector.

For photoconductive materials with ultra-short carrier lifetime, we can consider $n_r(t) = \delta(t)$ (the Dirac function) which gives:

$$J_{det}(f) \propto \tilde{E}_{THz}(f) \tag{4}$$

With f the frequency. On the other hand, if the carrier lifetime is long, $n_r(t)$ behaves like a step function which gives:

$$J_{det}(f) \propto \frac{\tilde{E}_{THz}(f)}{f} \tag{5}$$

This demonstrates that characteristics of the photo-carrier density in the detector antenna and its decay behaviour affect the bandwidth of the detected signal. In between these two cases, distortion effect of the detector on the incident THz field is considered by convolving the detector response with the THz field (in the time-domain). Thus, the detected photocurrent is linked to the spectral shape of the laser pulses, $I_{laser}(f)$, and to the frequency response of excited photocarriers, $B(f)$:

$$J_{det}(f) \propto I_{laser}(f) \cdot B(f) \cdot \tilde{E}_{THz}(f) \tag{6}$$

More generally, the response time of the detector determines the amount of detected signal at high frequencies, which is governed by the carrier lifetime and the RC time constant related to the device capacitance. But usually the capacitance of the typical structures showed in figure 8 is quite low and the carrier lifetime is the main limitation.

0.3.3 The delay line

The position of the optical delay line, usually a retroreflective mirror, is controlled by a computer allowing detection at different time steps. It allows to record the temporal variation of the detected photocurrent, as its path length Δx corresponds to a time interval between successive measurements:

$$\Delta t = \frac{2\Delta x}{c} + \tau(t) \tag{7}$$

where c is the celerity of light in vacuum and $\tau(t)$ is the sampling error due to mechanical positioning errors [27]. Thus, using the Fourier Transform, the variation of both the amplitude and phase of the electric field can be measured over a broadband spectral range.

0.3.4 Applications of terahertz time-domain spectroscopy

Throughout the last 30 years, THz-TDS has been employed to investigate gases, liquids and solid-state samples. There are many applications to THz-TDS, more being researched as the technique gains more recognition. Here are listed some of these applications.

Solid-state materials Solids are probably the most extensively studied materials in THz-TDS [28]. Semiconductors, conducting polymers [29], insulators [30] and superconductors [31] have been studied. Electronic processes [32, 33] as well as lattice modes (phonons) [34], have been investigated. THz spectroscopy is well adapted for these studies because scattering events of carriers in semiconductors tend to be on time scales of tens to hundreds of femtoseconds and phonon modes in solids are often found in the range of several THz to tens of THz. Thus, they are all located in the range covered by THz-TDS.

Security THz radiation has the ability to penetrate many materials including clothes and shoes. Because many threatening weapons are metal-based, a THz signal would be reflected when encountering such objects. Furthermore, even a ceramic knife can be detected. Moreover, The low-power and non-ionizing nature of the THz beam ensures no health hazards. Another use for THz-TDS in the field of security has been the detection of explosives [35, 36] and contraband items [37]. Indeed, numerous materials, components of these items, display absorption peaks in the THz range allowing qualitative identification. Moreover, these materials can be detected through covering layers such as cotton, leather or plastic as shown in figure 9. However, there are still limitations to the technique, the most important one being the absorption of the THz signal by the water vapour in the atmosphere.

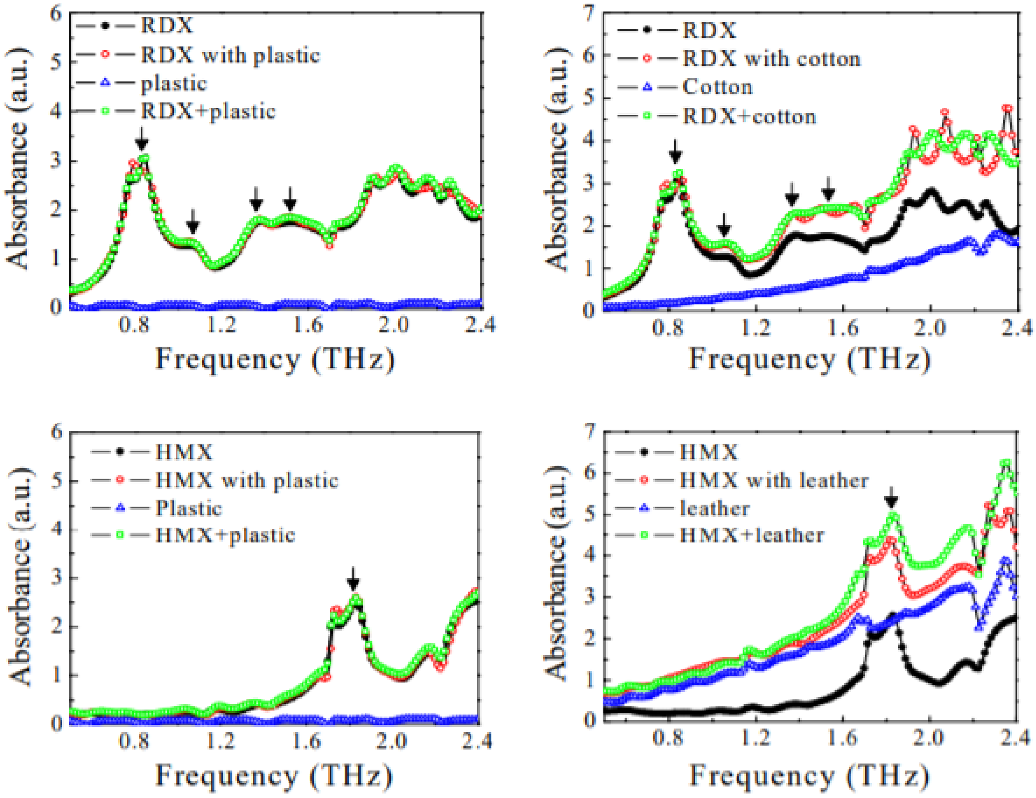


FIGURE 9: Absorption spectra of explosives compounds RDX (Cyclotrimethylenetrinitramine) and HMX (Cyclotetramethylene tetranitramine) covered with different materials (plastic, cotton and leather). The black curve is the transmission spectra of pure samples, the red one is the transmission spectra of covered samples, the blue one is the transmission spectra of the covers, and the green one is the sum of pure samples and covers. The THz spectrum of ambient air is used as reference. [36]

Food industry The detection of foreign bodies in food products plays an important role in the food industry. Currently, the techniques used are metal detectors and X-ray systems. Yet, the former can only detect metallic contaminations while the use of X-ray scanners is limited due to the fear that consumers might not buy products tested with ionizing radiation and because big screens are needed to protect the staff. Therefore, THz systems may offer an alternative for the non-ionizing nature of THz waves. However, due to the water content of several edible goods, the technique cannot be applied to all. For instance, THz-TDS works well for the investigation of chocolate, as it contains little water and mostly weakly absorbing molecules [38]. Another promising application is dry powders, or dry ingredients such as pasta, which are quite transparent to THz waves. It has been demonstrated that it is possible to detect foreign bodies in different powdered eatables such as sugar or wheat flour [39].

Paper industry In this field, THz-TDS systems are used to monitor the thickness and moisture content of the paper during the production process [40]. [41] demonstrated a quantitative method for determining the moisture content in paper based on extinction and phase measurements in the lower THz range while [42] looked for a simultaneous determination of thickness and moisture content of paper sheets. While this application is quite new, its low cost and results are promising for the paper industry.

Art conservation Infrared and X-ray techniques have been used for non destructive detection of images hidden below the top layers of paint in valuable paitings and murals. However, these techniques are limited by their resolution and penetration depth. Using THz-TDS, it has been demonstrated that different paints have different absorption depths and can therefore be distinguished [43]. Next, [44] generated a database on the THz properties of different pigments, binders and their mixtures using THz-FTIR (Fourier-Transform InfraRed spectroscopy) and extended their investigations on a medieval manuscript using THz-TDS. There has also been studies in reflection mode showing that we could recover a graphite drawing through paint and plaster layers (since graphite possesses a high reflectivity in the THz range) [45]. Finally, results of reflection images of hidden paint layers in a painting were compared with X-ray radiography results, which showed that only THz imaging provided information on the thickness of the hidden paint layers [46].

Imaging Because far-infrared radiation passes through many plastics when visible and infrared radiation cannot, it is possible, for example, to image the layout inside integrated chip packages [47]. There have also been some studies of semiconductor wafers: it is possible to map out the doping density semiconductor wafers with submillimeter resolution [48]. Images are often obtained by raster-scanning the sample through the focused THz beam. It is also possible to collect a two-dimensional image using a CCD camera as a detector (using the electro-optic effect instead of the photoconduction), which allows very high data acquisition rates [49]. A more recent imaging technique is THz tomography, wherein the echo or return of a pulse reflected off an object is captured and analyzed or the transmitted beam is analyzed as the sample is rotated [5, 50]. Finally, this techniques appears to be of utmost interest for biology, and THz imaging of plant and animal material debuted in 1995 [47] (see example in figure 10). It has been possible to image biological samples, especially of specimens that have a high contrast between aqueous and non-aqueous regions [51].

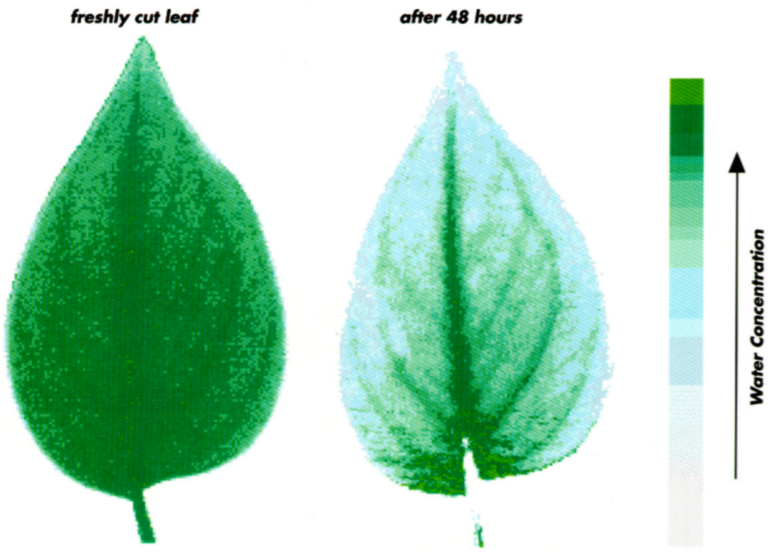


FIGURE 10: On the left is the THz image of a fresh leaf. On the right is the same leaf 48 hours later. The THz radiation is highly attenuated by the water inside the leaf, which explains that the dryer the leaf is the less absorption there is. This is illustrated by the colorbar on the right. [47]

In the next section, we will focus on biological samples as it is the application domain of this thesis.

0.4 Biology and THz

THz-TDS is a low-power technique that is favorable in many applications where samples are delicate. It is non-ionizing but strongly interacts with molecules, creating potential for use in biological samples. Moreover, to understand the dynamics and structure of biological samples, it is interesting to analyze collective displacements of their atoms at the nanometer and micrometer scales. Indeed, the intermolecular interactions energies are comparable to the thermal energy $kT \sim 6\text{THz}$, which is accessible with THz-TDS. That is why this technique has been widely used to analyze several types of biological samples since the 1990s.

0.4.1 Molecular crystals

In the THz region, crystals mainly exhibit phonon vibrations due to their lattice structure. This structure has both intramolecular vibrations, occurring within a single molecule, and intermolecular vibrations, arising from interactions between adjacent molecules. These intermolecular vibrations are probed in the THz region while individual molecular vibrations are seen in the infrared and pure rotations are seen at microwave frequencies (in large molecules, low frequency intramolecular vibrations can reach the THz range). The easiest way to analyze crystalline structured samples that can be produced in large quantities is to turn them into pellets, namely discs composed of pressed powder.

Structurally simple samples have first been studied, such as saccharides. Glucose has been a sample of choice in THz-TDS analysis because of the unique absorption features its spectrum exhibits. It has been compared to sucrose as the glucose molecule essentially comprises half of the entire sucrose molecule. Interestingly, it was found that their spectra were quite different with little overlap of absorption bands, highlighting the difference in their microscopic structures [11]. Another "golden sample" is lactose, which has proven to be more complex than generally thought. Using periodic DFT simulations of lactose molecular crystals, [52] investigated the role of water in the lactose THz spectrum, presenting both direct and indirect contributions. They showed that, firstly, THz modes are spread over more than one molecule and require periodic computation rather than a gas-phase one. Secondly, hydration water does not only play a perturbative role but also participates in the facilitation of the THz vibrations [53].

The pharmaceutical industry can also rely on THz-TDS analysis, as several drugs such as aspirin, palatinose, and riboflavin were scanned [54]. Component analysis has been used to identify concentrations of aspirin, palatinose, and riboflavin in pellets, which could be useful for non-destructive quality investigations. Another use of THz-TDS for this industry is its ability to differentiate a crystalline form from another. Indeed, different isomers lead to different crystalline structures with different spectral fingerprints in the THz range. Hence, an organic compound can come in different polymorphic forms that typically differ in their physico-chemical properties and ultimately in their pharmaceutical performance. For example, different polymorphs may exhibit varying rates of dissolution in the body, thereby affecting the bioavailability of that molecule [55, 56]. Therefore it is imperative to identify which crystalline forms are most effective and stable when establishing new drugs to treat illnesses.

0.4.2 Water

Since the beginning of THz research, it has been clear that water was the main obstacle due to its high absorption in this frequency range. In gas phase, it limits the propagation of THz radiation in free space as well as in samples that contain water in condensed phase, which concerns most biological samples [57]. Indeed, the great majority of biological tissues contain water, whether it be in small quantity (for fat or teeth for example) or the major part of their weight (for soft tissues, such as lungs, skin or muscles) [58]. However, it can also be seen as a great informative feature when comparing different levels of hydration inside samples and it also makes a very interesting subject on its own.

Water molecules can be found in two states, free or bound, depending on if they are a structural unit of an enzyme molecule or a solvent for example. In biological tissues, water is bound to other molecules, which modifies its structure and thus its THz vibrational response. Therefore, knowing liquid water properties in its free or bound state, in the terahertz range, is fundamental for the study of any biological sample [59, 60].

Many researchers have studied the dielectric response of water since P. Debye introduced the rotational diffusion model of polar liquids in 1929. If progress has been made on modelizing the dielectric response of liquid water at THz frequencies, there is still no consensus on the correct model to apply and the molecular-level mechanisms controlling its behavior remain essentially unknown. This will be the subject of chapter 2.

0.4.3 Analyzing complex biological samples

After analyzing simple biological samples, researchers have wanted to investigate more complex ones, such as DNA. Calculations of vibrational modes in the millimeter-wave or THz ranges have been published for a long time but in 1997, researchers observed vibrational modes in herring DNA for the first time [61]. This created hope for cancer research as it was thought that defects or damage in DNA strands could be identified because the local-mode absorbance frequencies would shift. From this, they extrapolated that the degree of tissue damage could be determined from these DNA modes. Furthermore, it has been discovered that the methylation of DNA, which is an alteration of DNA during the development of cancer, has a vibrational frequency at 1.65 THz. As a result, several research groups have been working on the ability to demethylize DNA in the hope of curing cancer [62].

Other biological samples of importance are proteins. They were first studied in 2000 [10] and numerous studies have followed since. In order to be closer to *in vivo* condition, where the protein is in an aqueous environment composed of water, salts, small organic molecules and other macromolecules (some of which may interact with it), the protein can be studied in a simpler defined substitute called a buffer solution. In this type of solution, researchers were able to probe the conformational changes in these proteins. They found that the major contribution to absorption of THz radiation is the hydration level in the protein and that the hydration level can indirectly detect conformational change [63]. Below 10 THz, the analysis of the structure of proteins is possible in theory if the collective modes can be distinguished from the hydrogen bond and other modes [64]. However, as was said in the previous section, liquid water is a strong absorber of low energy photons masking the absorbance due to the protein. A way to overcome this problem would be to remove water by lyophilisation [65], however proteins have been demonstrated to lose their structural integrity while in the dry form. Another way is to freeze the protein in the solution as ice absorbs far less than liquid water in the THz range but, still, its absorbance prevents much of the spectrum to be used.

All sorts of other biological samples were also investigated. Indeed, slices of skin, adipose tissue, striated muscle, artery, vein, nerve, and portions of blood were scanned from 0.5 to 1.5 THz with THz in transmission mode [66]. The hope of this study was to gain knowledge in order to be able to interpret scans of complete skin samples.

This succint review of the biological applications for THz-TDS shows its potential. However, if many studies have been done, the information retrieved is mainly only phenomenological. In this work, we will focus on the retrieval of information in THz-TDS studies in order to exploit the potential of this technique in gaining knowledge on all sorts of biological samples. After a thorough explication of our data analysis method, we will apply it to the study of biological samples. First, by analyzing liquid water that is the foundation of every biological sample and then by working with more complex samples in order to be able to study virus samples in the future.

Chapter 1

Recovering information in THz-TDS

Commercial THz-TDS systems are becoming more and more accessible, which attracts many researchers and industrialists because of the broad range of applications they offer. However, there are no standardized measurement protocols that ensure traceable results, and no standardized data analysis either. In this chapter, I will focus on the concept of information in THz-TDS, explaining what type of information we can retrieve from experiments, the limitations of current data analysis processes and how to overcome them thanks to our home-made fitting software.

1.1 Information in a THz-TDS experiment

The detection antenna in a THz-TDS setup records a time-reconstructed THz pulse, as explained in section 0.3.1 on the principle of the technique. As a result, at the end of an experiment, the data we obtain is a recording of the pulse, a time-domain trace, that we are going to analyse to get further information on the material properties.

1.1.1 The transfer function

The ability to measure directly the electric field of the THz pulse gives access to both the phase and the amplitude of the waveform. Therefore, it provides information on the absorption coefficient and the refractive index of the sample versus the frequency. This is achieved by comparing two signals: a reference and a sample time-traces, which propagate through the setup without and with sample in place respectively. Then, by performing a Fourier transform of the two time-traces and taking the ratio between them, we obtain what is called the transfer function of the experiment that can be written as:

$$\tilde{T}(\omega) = \frac{\tilde{E}_s(\omega)}{\tilde{E}_{ref}(\omega)} \tag{1.1}$$

Where $\tilde{E}_s(\omega)$ and $\tilde{E}_{ref}(\omega)$ are the Fourier transforms of time-domain signals $E_s(t)$ (sample time trace) and $E_{ref}(t)$ (reference time trace) respectively.

In order to characterize the propagation of the waveforms inside the setup and model the transfer function, we can use the transfer matrices inside a material and at the interface between two materials. In figure 1.1 is illustrated the propagation of the electric field in a material, at normal incidence, where e_{Fi} corresponds to the field propagating to the right while e_{Bi} corresponds to the field propagating to the left, where $i \in \mathbb{N}$ is the number of the layer (or material). T_{ij} and R_{ij} correspond to the transmission and reflection matrices respectively, with $\{i, j\} \in \mathbb{N}$.

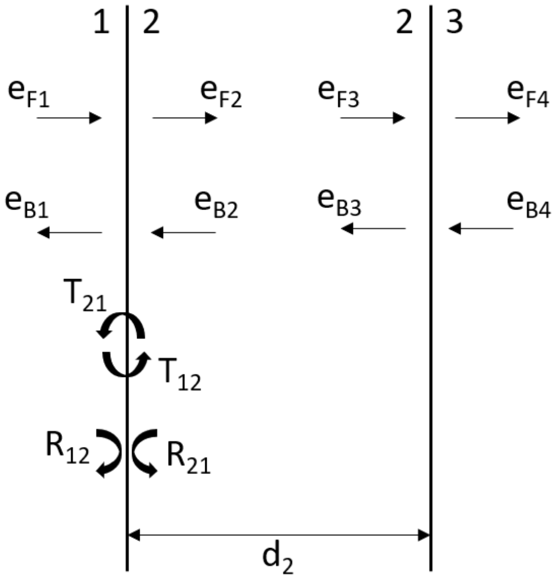


FIGURE 1.1: Propagation of the electric field through material 2 surrounded by two different materials (1 and 3). e_{Fi} is the forward propagating field, e_{Bi} is the backward propagating field, T_{ij} is the transmission matrix between materials i and j, R_{ij} is the reflection matrix at the interface between materials i and j, and d_2 is the thickness of the second layer.

The relation between the fields at the interface between materials 1 and 2 is:

$$\begin{cases} e_{F2} = T_{12}e_{F1} + R_{21}e_{B2} \\ e_{B1} = R_{12}e_{F1} + T_{21}e_{B2} \end{cases} \tag{1.2}$$

Now, we can write the fields from material 2 considering material 1:

$$\begin{cases} e_{F2} = T_{12}e_{F1} - R_{21}{T_{21}}^{-1}R_{12}e_{F1} + R_{21}{T_{21}}^{-1}e_{B1} \\ e_{B2} = {T_{21}}^{-1}e_{B1} - {T_{21}}^{-1}R_{12}e_{F1} \end{cases} \tag{1.3}$$

As a result, the transfer matrix between 1 and 2 is:

$$T_{1\to 2} = \begin{pmatrix} T_{12} - R_{21}{T_{21}}^{-1}R_{12} & R_{21}{T_{21}}^{-1} \\ -{T_{21}}^{-1}R_{12} & {T_{21}}^{-1} \end{pmatrix} \tag{1.4}$$

Where

$$\begin{pmatrix} e_{F2} \\ e_{B2} \end{pmatrix} = T_{1\to 2} \begin{pmatrix} e_{F1} \\ e_{B1} \end{pmatrix} \tag{1.5}$$

Similarly, we can write the transfer matrix of the propagating wave inside a material:

$$\begin{pmatrix} e_{F3} \\ e_{B3} \end{pmatrix} = \begin{pmatrix} \exp(-j\tilde{n}_2\omega d_2/c) & 0 \\ 0 & \exp(j\tilde{n}_2\omega d_2/c) \end{pmatrix} \begin{pmatrix} e_{F2} \\ e_{B2} \end{pmatrix} = P_2 \begin{pmatrix} e_{F2} \\ e_{B2} \end{pmatrix} \tag{1.6}$$

Where $\tilde{n}_2$ is the complex refractive index of material 2 where the real part corresponds to a delay and the imaginary part to an absorption in the material, d_2 is its thickness, and ω is the angular frequency.

In a THz-TDS experiment, material 2 is the sample and materials 1 and 3 are the layers surrounding the sample, which will be considered as air here. Thus, there are 3 transfer matrices to consider: the two transmission matrices at the interfaces between materials and the propagation matrix inside 2. The experiment can then be modelled as:

$$\begin{pmatrix} e_{F4} \\ e_{B4} \end{pmatrix} = T_{2\rightarrow \text{air}} P_2 T_{\text{air}\rightarrow 2} \begin{pmatrix} e_{F1} \\ e_{B1} \end{pmatrix} \tag{1.7}$$

The THz beam arriving on the sample is e_{F1} and we suppose that no pulse arrives from the right which means that $e_{B4} = 0$. Since the objective is to obtain the transfer function, it is more convenient to write equation 1.7 differently:

$$\begin{pmatrix} e_{F1} \\ e_{B1} \end{pmatrix} = T^{-1}_{\text{air}\rightarrow 2} P_2^{-1} T^{-1}_{2\rightarrow \text{air}} \begin{pmatrix} e_{F4} \\ e_{B4} \end{pmatrix} \tag{1.8}$$

Since the beam arrives at normal incidence, the transmission and reflection matrices are singletons which allows us to simplify the transfer matrices:

$$T^{-1}_{\text{air}\rightarrow 2} = \begin{pmatrix} {T_{\text{air}2}}^{-1} & -{T_{\text{air}2}}^{-1} R_{2\text{air}} \\ R_{\text{air}2} {T_{\text{air}2}}^{-1} & T_{2\text{air}} - R_{\text{air}2} {T_{\text{air}2}}^{-1} R_{2\text{air}} \end{pmatrix} = \frac{1}{t_{\text{air}2}} \begin{pmatrix} 1 & -r_{2\text{air}} \\ r_{\text{air}2} & 1 \end{pmatrix} \tag{1.9}$$

Resolving equation 1.8 allows us to find the transfer function of the sample:

$$\tilde{T}(\omega) = \frac{\tilde{E}_{F4}}{\tilde{E}_{F1}} = t_{\text{air}2} t_{2\text{air}} \frac{\exp(-j\tilde{n}_2 \omega d_2/c)}{1 + r_{\text{air}2} r_{2\text{air}} \exp(-2j\omega d_2 \tilde{n}_2/c)} \tag{1.10}$$

For a multilayer sample, the method stays the same. In figure 1.2 is shown the example of a 3-layer material, which is a common case in THz-TDS experiments as it depicts, for instance, a liquid sample held inside a microfluidic cell or a cuvette.

In this case, the relation between the entering and exiting field is:

$$\begin{pmatrix} e_{F1} \\ e_{B1} \end{pmatrix} = {T_{\text{air}\rightarrow 1}}^{-1} {P_1}^{-1} {T_{1\rightarrow 2}}^{-1} {P_2}^{-1} {T_{2\rightarrow 3}}^{-1} {P_3}^{-1} {T_{3\rightarrow \text{air}}}^{-1} \begin{pmatrix} e_{F8} \\ e_{B8} \end{pmatrix} \tag{1.11}$$

Which gives the following result:

$$\begin{pmatrix} e_{F1} \\ e_{B1} \end{pmatrix} = \frac{1}{t_{3\text{air}} t_{23} t_{12} t_{\text{air}1}} \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} e_{F8} \\ e_{B8} \end{pmatrix} \tag{1.12}$$

With the A, B, C, D coefficients:

$$\begin{aligned} A = {} & \exp\left(j\omega(n_3 d_3 + n_2 d_2 + n_1 d_1)/c\right) + r_{3\text{air}} r_{23} \exp\left(j\omega(-n_3 d_3 + n_2 d_2 + n_1 d_1)/c\right) \\ & + r_{12} r_{23} \exp\left(j\omega(n_3 d_3 - n_2 d_2 + n_1 d_1)/c\right) + r_{12} r_{3\text{air}} \exp\left(j\omega(-n_3 d_3 - n_2 d_2 + n_1 d_1)/c\right) \\ & + r_{12} r_{\text{air}1} \exp\left(j\omega(n_3 d_3 + n_2 d_2 - n_1 d_1)/c\right) + r_{\text{air}1} r_{12} r_{3\text{air}} r_{23} \exp\left(j\omega(-n_3 d_3 + n_2 d_2 - n_1 d_1)/c\right) \\ & + r_{\text{air}1} r_{23} \exp\left(j\omega(n_3 d_3 - n_2 d_2 - n_1 d_1)/c\right) + r_{\text{air}1} r_{3\text{air}} \exp\left(j\omega(-n_3 d_3 - n_2 d_2 - n_1 d_1)/c\right) \end{aligned}$$

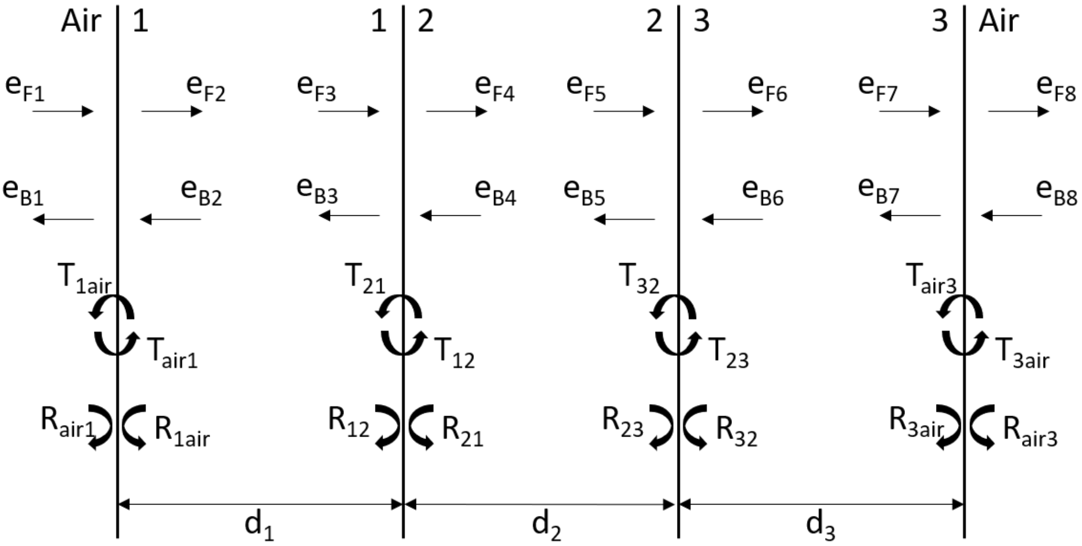


FIGURE 1.2: Propagation of the electric field through a 3-layer sample. e_{Fi} is the forward propagating field, e_{Bi} is the backward propagating field, T_{ij} is the transmission matrix between materials i and j, R_{ij} is the reflection matrix at the interface between materials i and j, and d_i is the thickness of the i^{th} layer.

$$\begin{aligned}
B =& - r_{\text{air1}} \exp\left(j\omega(n_3d_3 + n_2d_2 + n_1d_1)/c\right) - r_{\text{air1}} r_{\text{3air}} r_{23} \exp\left(j\omega(-n_3d_3 + n_2d_2 + n_1d_1)/c\right) \\
& - r_{\text{air1}} r_{12} r_{23} \exp\left(j\omega(n_3d_3 - n_2d_2 + n_1d_1)/c\right) - r_{\text{air1}} r_{12} r_{\text{3air}} \exp\left(j\omega(-n_3d_3 - n_2d_2 + n_1d_1)/c\right) \\
& - r_{12} \exp\left(j\omega(n_3d_3 + n_2d_2 - n_1d_1)/c\right) - r_{12} r_{\text{3air}} r_{23} \exp\left(j\omega(-n_3d_3 + n_2d_2 - n_1d_1)/c\right) \\
& - r_{23} \exp\left(j\omega(n_3d_3 - n_2d_2 - n_1d_1)/c\right) - r_{\text{3air}} \exp\left(j\omega(-n_3d_3 - n_2d_2 - n_1d_1)/c\right)
\end{aligned}$$

$$\begin{aligned}
C =& - r_{\text{3air}} \exp\left(j\omega(n_3d_3 + n_2d_2 + n_1d_1)/c\right) - r_{23} \exp\left(j\omega(-n_3d_3 + n_2d_2 + n_1d_1)/c\right) \\
& - r_{12} r_{23} r_{\text{3air}} \exp\left(j\omega(n_3d_3 - n_2d_2 + n_1d_1)/c\right) - r_{12} \exp\left(j\omega(-n_3d_3 - n_2d_2 + n_1d_1)/c\right) \\
& - r_{\text{air1}} r_{12} r_{\text{3air}} \exp\left(j\omega(n_3d_3 + n_2d_2 - n_1d_1)/c\right) - r_{\text{air1}} r_{12} r_{23} \exp\left(j\omega(-n_3d_3 + n_2d_2 - n_1d_1)/c\right) \\
& - r_{\text{air1}} r_{23} r_{\text{3air}} \exp\left(j\omega(n_3d_3 - n_2d_2 - n_1d_1)/c\right) - r_{\text{air1}} \exp\left(j\omega(-n_3d_3 - n_2d_2 - n_1d_1)/c\right)
\end{aligned}$$

$$\begin{aligned}
D =& r_{\text{air1}} r_{\text{3air}} \exp\left(j\omega(n_3d_3 + n_2d_2 + n_1d_1)/c\right) + r_{\text{air1}} r_{23} \exp\left(j\omega(-n_3d_3 + n_2d_2 + n_1d_1)/c\right) \\
& + r_{\text{air1}} r_{12} r_{23} r_{\text{3air}} \exp\left(j\omega(n_3d_3 - n_2d_2 + n_1d_1)/c\right) + r_{\text{air1}} r_{12} \exp\left(j\omega(-n_3d_3 - n_2d_2 + n_1d_1)/c\right) \\
& + r_{12} r_{\text{3air}} \exp\left(j\omega(n_3d_3 + n_2d_2 - n_1d_1)/c\right) + r_{12} r_{23} \exp\left(j\omega(-n_3d_3 + n_2d_2 - n_1d_1)/c\right) \\
& + r_{23} r_{\text{3air}} \exp\left(j\omega(n_3d_3 - n_2d_2 - n_1d_1)/c\right) + \exp\left(j\omega(-n_3d_3 - n_2d_2 - n_1d_1)/c\right)
\end{aligned}$$

Finally, the transfer function of a 3-layer material is:

$$\tilde{T}(\omega) = \frac{e_{F8}}{e_{F1}} = \frac{t_{\text{3air}} t_{23} t_{12} t_{\text{air1}}}{A} \tag{1.13}$$

The transfer matrix method can be applied to any n-layer material.

1.1.2 Extraction of optical parameters

Equation 1.1 sets the "forward problem": knowing $\tilde{E}_{ref}(\omega)$ and $\tilde{n}(\omega)$ one can obtain $\tilde{E}_s(\omega)$. Since the experiment gives $\tilde{E}_{ref}(\omega)$ and $\tilde{E}_s(\omega)$, the actual interest is the "inverse problem", that is, with knowledge of $\tilde{E}_{ref}(\omega)$ and $\tilde{E}_s(\omega)$, one can determine an estimate of the frequency-dependent optical properties of a sample (complex index of refraction, permittivity, or permeability). The complex refractive index is $\tilde{n}(\omega) = n(\omega) + j\kappa(\omega)$ (or $\tilde{n}(\omega) = n(\omega) - j\kappa(\omega)$ according to the conventions, we will consider the former here), where $n(\omega)$ is what we call the refractive index and $\kappa(\omega)$ is the extinction coefficient of the sample.

As can be seen in the previous section, three parameters impact the transfer function: the real and imaginary parts of the complex refractive index and the thickness d of the sample. If we write the transfer function as: $\tilde{T}(\omega) = \left|\tilde{T}(\omega)\right| \exp(-j\phi(\omega))$, knowing both the experimental expression of the transfer function, $\tilde{T}(\omega)_{\text{exp}}$, from equation 1.1 and its analytical expression, $\tilde{T}(\omega)_{\text{model}}$, from equations 1.10 and 1.13 according to the sample scanned, we can deduce the real and imaginary parts of the complex refractive index with:

$$n(\omega) = n_{\text{air}} + \left|\frac{\phi(\omega)c}{d\omega}\right| \tag{1.14}$$

$$\kappa(\omega) = \frac{c}{d\omega} \ln\left(\frac{\left|\tilde{T}(\omega)\right|}{\tilde{T}(\omega)}\right) \tag{1.15}$$

Moreover, we can recover the absorption coefficient $\alpha(\omega)$ of the sample with:

$$\alpha(\omega) = \frac{2\kappa(\omega)\omega}{c} \tag{1.16}$$

Finally, the complex dielectric permittivity $\tilde{\epsilon}(\omega) = \epsilon'(\omega) + j\epsilon''(\omega)$ of the sample, where $\epsilon'(\omega)$ is its real part and $\epsilon''(\omega)$ its imaginary part, can also be retrieved with:

$$\epsilon'(\omega) = n(\omega)^2 - \kappa(\omega)^2 \tag{1.17}$$

$$\epsilon''(\omega) = 2n(\omega)\kappa(\omega) \tag{1.18}$$

In order to gain more knowledge on the studied sample, one can fit these retrieved curves to a permittivity model.

1.1.3 Permittivity models

When a particle is placed in a constant field E, the polarizability, α, is defined as:

$$p = \epsilon_0 \alpha E \tag{1.19}$$

Where p is the polarization of the particle. When E is not stationary, the polarization is then:

$$p(t) = \epsilon_0 \int \alpha(t - t')E(t')dt' \iff \tilde{p}(\omega) = \epsilon_0 \tilde{\alpha}(\omega)\tilde{E}(\omega) \tag{1.20}$$

The interaction between a material and the electromagnetic field is described by the dielectric permittivity $\tilde{\epsilon}$ and the linear dielectric susceptibility $\tilde{\chi}$. The susceptibility is a dimensionless proportionality constant that indicates the degree of polarization of a dielectric material in response to an applied electric field. By making the approximation that the dielectric susceptibility is linear, we have:

$$P(t) = \epsilon_0 \int \chi(t-t')E(t')dt' \iff \tilde{P}(\omega) = \epsilon_0 \tilde{\chi}(\omega)\tilde{E}(\omega) \tag{1.21}$$

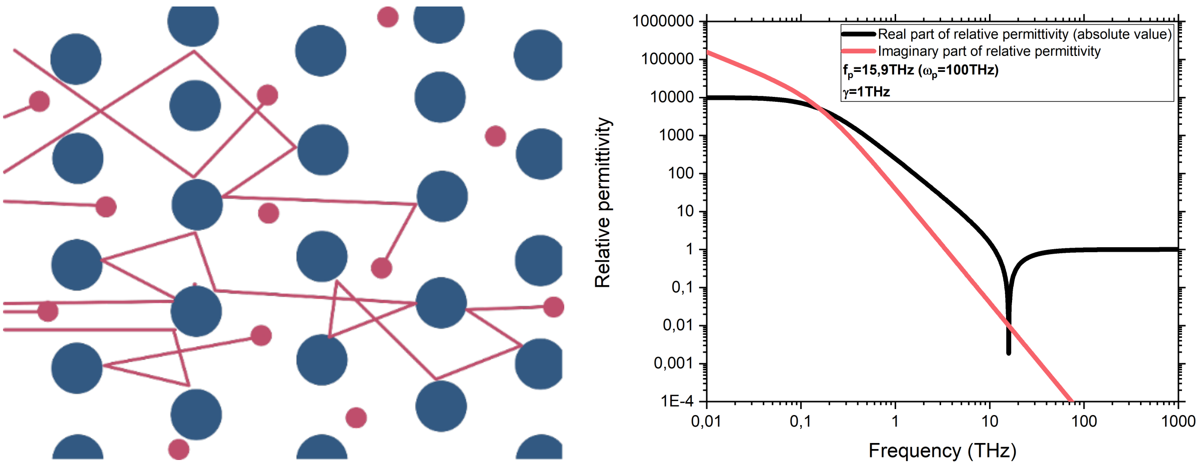
Where P is the polarization density, that can also be defined as:

$$\tilde{P}(\omega) = -ne\tilde{r}(\omega) \tag{1.22}$$

With n the number of dipole moments (per volume), e the elementary charge and r the displacement from the negative to the positive charge. From equation 1.21, we see that $\chi(t)$ is the impulse response of the material to the electromagnetic wave, and because of causality it is equal to 0 for $(t-t') < 0$; which is translated in the frequency domain by Kramers-Kronig relations.

The first step is to understand how these macroscopic values depict microscopic processes, that is how the charges in the material and the forces acting on them are linked to the refractive index, the permittivity or the susceptibility.

The Drude model



(A) Simplified drawing of the Drude model. In blue are the fixed positive charges and in pink are the electrons propagating inside the material and colliding with the charges.

(B) Example of the permittivity calculated with a Drude model with a plasma frequency of 100 THz and a width γ of 1 THz, where $\gamma = \frac{1}{\tau}$.

FIGURE 1.3: Simplified drawing of the Drude model and an example of a permittivity curve calculated with it.

This model was developed by Paul Drude to represent the behavior of conduction electrons in a metal. The metal is considered as a network of fixed positive charges in which the negatively charged electrons travel, as illustrated in figure 1.3a. There are no interactions between electrons and the collisions are inelastic. If we suppose that the electric field is polarized linearly (and we neglect the magnetic field), the equation of motion is:

$$m_e \frac{d\left\langle v(t)\right\rangle}{dt} + \frac{m_e}{\tau}\left\langle v(t)\right\rangle = -eE(t) \tag{1.23}$$

With m_e the electron effective mass, E the electric field, $\langle v(t)\rangle$ the mean electron velocity, and τ the average time between two collisions. This can be written in the frequency domain as:

$$m_e j\omega\left\langle v(\omega)\right\rangle + \frac{m_e}{\tau}\left\langle v(\omega)\right\rangle = -e\tilde{E}(\omega) \tag{1.24}$$

We can then deduce the dielectric susceptibility of the material in the frequency domain:

$$\tilde{\chi}_{Drude}(\omega) = \frac{\tilde{P}(\omega)}{\epsilon_0\tilde{E}(\omega)} = \frac{\dfrac{-nev(\omega)}{j\omega}}{\epsilon_0\tilde{E}(\omega)} \tag{1.25}$$

When replacing $\tilde{E}(\omega)$ with equation 1.24 we find:

$$\tilde{\chi}_{Drude}(\omega) = \frac{\dfrac{ne^2}{m_e\epsilon_0}}{j\dfrac{\omega}{\tau} - \omega^2} = \frac{w_p^2}{j\dfrac{\omega}{\tau} - \omega^2} \tag{1.26}$$

Where the plasma frequency ω_p is defined as usual: $\omega_p = \sqrt{\dfrac{ne^2}{m_e\epsilon_0}}$.

This shows that the radiation has an effect on the charges via the Drude-Lorentz forces, which drive the electron motion and then lead to the dielectric susceptibility. This model is applied to materials containing free charges, such as plasma, metals and doped semiconductors. Its dispersion is shown in figure 1.3b. At low frequency (below the plasma frequency), the materials act as a reflector (very high permittivity) while its behaviour at high frequencies is the one of a dielectric.

The Lorentz model

The Lorentz harmonic oscillator model is another very famous model. Here, the electrons are linked to the nucleus with a spring force and the oscillator movement is damped. This is represented on figure 1.4. The equation of motion, projected on the polarization axis (which we suppose is the same as the spring axis) is:

$$m_e \frac{d\left\langle v(t)\right\rangle}{dt} + m_e\Gamma\left\langle v(t)\right\rangle + m_e\omega_0^2\left\langle r(t)\right\rangle = -eE(t) \tag{1.27}$$

With m_e the effective mass of the electrons, Γ the damping coefficient, ω_0 the resonance pulsation frequency corresponding to the square-root of the ratio between the spring force and the effective mass of the electron. In the frequency domain, we have:

$$m_e(j\omega)^2\left\langle r(\omega)\right\rangle + m_e\Gamma j\omega\left\langle r(\omega)\right\rangle + m_e\omega_0^2\left\langle r(\omega)\right\rangle = -e\tilde{E}(\omega) \tag{1.28}$$

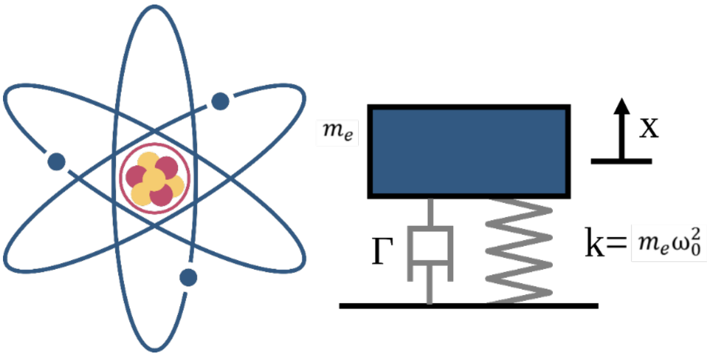


FIGURE 1.4: Simplified drawing of the Lorentz model. The electrons are represented in blue, the nucleus is the pink circle containing protons and neutrons.

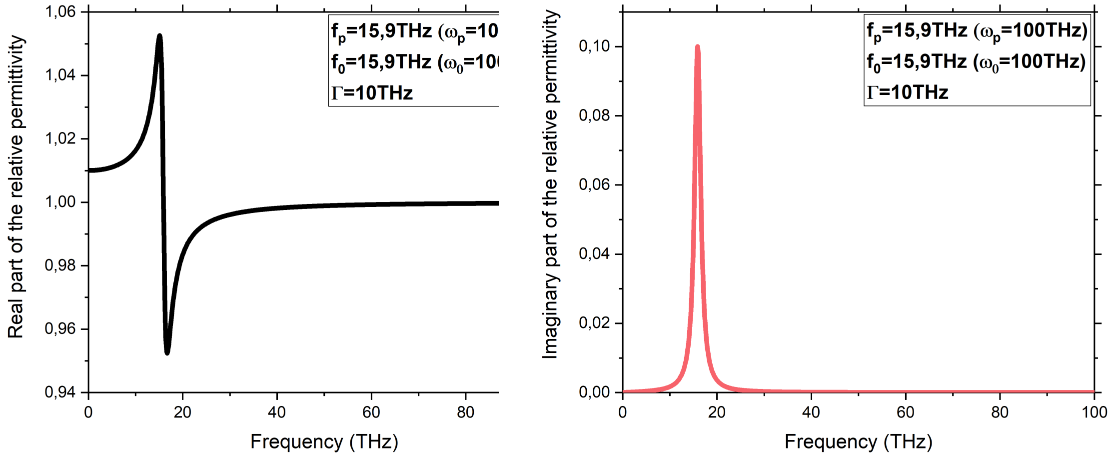


FIGURE 1.5: Example of the permittivity of a Lorentz model with a plasma frequency ω_p of 100 THz, a resonance frequency ω_0 of 100 THz and a width of 10 THz. On the left is the real part of the permittivity, the imaginary part is on the right.

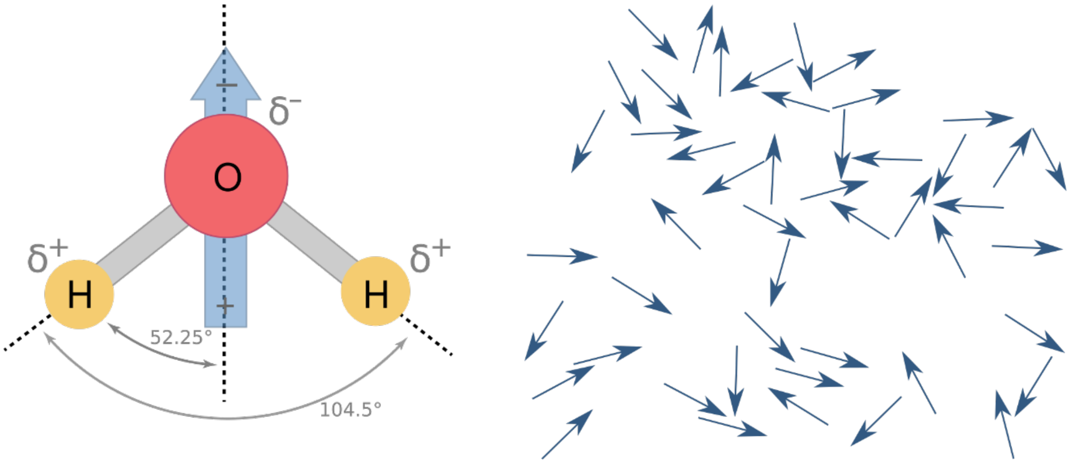
Therefore, using the same method as with the Drude model, we can deduce the susceptibility:

$$\tilde{\chi}_{Lorentz}(\omega) = \frac{\frac{ne^2}{m_e\epsilon_0}}{\omega_0^2 - \omega^2 + j\omega\Gamma} = \frac{\omega_p^2}{\omega_0^2 - \omega^2 + j\omega\Gamma} \tag{1.29}$$

Its dispersion is traced on figure 1.5. A Lorentz model corresponds to a resonant mode, its width is the inverse of the damping coefficient in the damped oscillator model.

The Debye model

Another case to cover is the one of polar molecules in gas or liquid phase. In this case, the molecule has a dipolar moment $\vec{\mu}$ and the electromagnetic response is triggered by the



(A)

(B)

FIGURE 1.6: (A) The water molecule, its dipolar moment is represented in blue. (B) Simplified drawing of the Debye model. The blue arrows represent the dipoles reorienting inside the sample.

dipole orientation. The usual way to represent this is the Debye model. It gives the dielectric constant of a liquid composed of polar molecules diluted in a liquid composed of non-polar molecules. An important hypothesis to keep in mind is that there are no dipole-dipole interactions inside the liquid in this model.

The equation of motion, in this case, is a Euler Langevin equation [67], which describes the rotational Brownian motion in the space of a sphere, when the inertia of the sphere is neglected:

$$-\Gamma\overrightarrow{\omega(t)} = \overrightarrow{\lambda(t)} + \overrightarrow{\mu(t)} \wedge \overrightarrow{E(t)} \tag{1.30}$$

With $\Gamma\overrightarrow{\omega(t)}$ the damping torque due to Brownian movement, $\overrightarrow{\lambda(t)}$ the white noise driving torque due to Brownian movement (so that $<\overrightarrow{\lambda(t)}>= 0$ and $<\overrightarrow{\lambda_i(t)}.\overrightarrow{\lambda_j(t')}>= 2kT\Gamma\delta_{i,j}\delta(t-t')$, where the indices $i, j = 1, 2, 3$ in Kronecker's delta $\delta_{i,j}$ correspond to the Cartesian coordinate axes X, Y, Z), and $\overrightarrow{\mu(t)}$ the dipole. Since this dipole is considered rigid:

$$\frac{d\overrightarrow{\mu(t)}}{dt} = \overrightarrow{\mu(t)} \wedge \overrightarrow{\omega(t)} \tag{1.31}$$

Equation 1.30 becomes, after multiplying by $\overrightarrow{\mu(t)}$:

$$-\Gamma\dot{\mu}(t) = \overrightarrow{\lambda(t)} \wedge \overrightarrow{\mu(t)} + \mu^2\overrightarrow{E(t)} - \left(\overrightarrow{\mu(t)}.\overrightarrow{E(t)}\right) \wedge \overrightarrow{\mu(t)} \tag{1.32}$$

After a few steps [67], this results in a Fokker-Planck equation from which we find the response to a step function (the field $E(t)$ is null at $t = 0$):

$$P(t > 0) =< \vec{\mu}(t > 0) >= \frac{n\mu^2 E_0}{3kT}\exp(-\frac{t}{\tau_D}) \tag{1.33}$$

With E_0 the amplitude of E before $t = 0$, T the temperature, k the Boltzmann constant, and τ_D the Debye relaxation time constant where $\tau_D = \dfrac{\Gamma}{2kT}$.

In any integrable linear system, the response to a step function is the integral of the impulse response, which is the susceptibility. Therefore, we can calculate the susceptibility by deriving the response to a step function. So, in the case of a polar liquid:

$$\tilde{\chi}(\omega) = \int_{-\infty}^{\infty} \left(\int_{0}^{\infty} \frac{n\mu^2}{3kt} \frac{1}{\tau_D} \exp\left(\frac{-t}{\tau_D}\right) \exp(j\omega t) dt \right) \exp(j\omega t) d\omega = \frac{n\mu^2}{3kT} \frac{1}{1 + j\omega\tau_D} \tag{1.34}$$

$$= \frac{\Delta\epsilon}{1 + j\omega\tau_D} \tag{1.35}$$

Where $\Delta\epsilon = \dfrac{n\mu^2}{3kT}$.

This formula does not take into account the inertial term, by taking it into account, $3kT$ [67] becomes $2kT$. An example of dispersion of the Debye model is traced on figure 1.7.

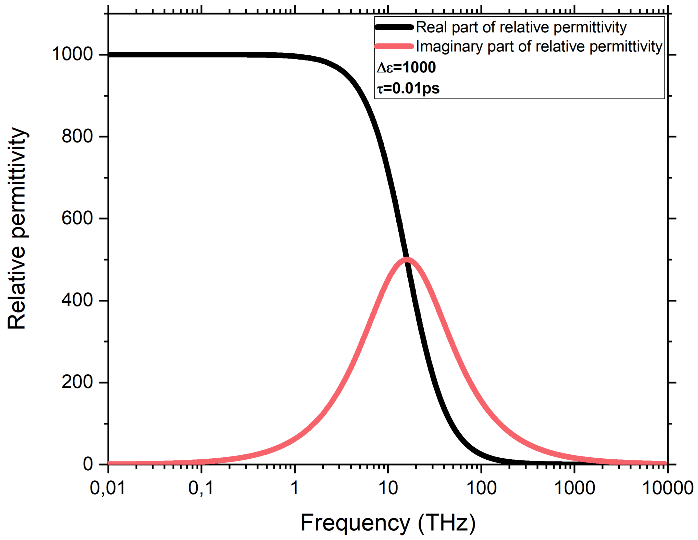


FIGURE 1.7: Example of the relative permittivity calculated with a Debye model with $\Delta\epsilon = 1000$ and a time constant of 0.01 ps.

All of these examples illustrate how the motion of electric charges inside a material builds the refractive index and, reciprocally, how the excitation of a material with a beam will trigger a motion of charges.

1.2 Retrieving information in the frequency domain

We have seen that it is possible to deduce the complex refractive index from the experimental data. However, the inverse problem stated in section 1.1.2 can only be solved analytically by ignoring the Fabry-Perot term (*i.e.* by ignoring reflections inside the sample) by using a temporal filter for instance [68] and only for a sample without absorption. These assumptions imply that there is no phase term in the transmission coefficients. Hence, the method consists in the extraction of the unwrapped phase from the THz-TDS data in the frequency domain in order to obtain the refractive index with equation 1.14. Another implication of these assumptions is that there are no losses during the pulse propagation. As a consequence, this method is limited to optically-thick ($nd > 1.5\,\text{mm}$), non-absorbing samples for a typical THz-TDS setup.

Nevertheless, it is possible to determine the real part of the refractive index from the unwrapped phase, and then add the losses contribution by adding an imaginary term to the refractive index. This creates the need to compensate the phase term in the transmission due to this imaginary part, thus doing these two steps several times until a satisfactory result is reached. However, this method does not guarantee the convergence or any reliability of the obtained results. Since this technique does not fulfill the Kramers-Kroenig relation (therefore it is intrinsically non-causal), it can be improved by solving the inverse problem with a numerical approach. To do so, one defines an error function:

$$Err(\tilde{n})(\omega) = (|\tilde{T}_{exp}(\omega)| - |\tilde{T}_{model}(\omega)|)^2 + \psi(\phi_{exp}(\omega) - \arg(\tilde{T}(\omega)_{model}))^2 \tag{1.36}$$

This error function adds the errors from both the modulus and the phase between the modeled transmission coefficient and the experimental one. ψ is a weighting coefficient enabling the addition of the phase and modulus errors, its value is usually set to 1 amplitude unit/rad.

Once the error function has been defined, a minimization algorithm must be implemented. For example, one can use the simplex method (a gradient free method) [69] or a quasi-Newton algorithm [70, 71]. This has to be applied to every single frequency and gives as a result the real and imaginary parts of the complex refractive index $\tilde{n}(\omega)$ and $\tilde{\kappa}(\omega)$. This method is fast and works regardless of whether the Fabry-Perot effect is taken into account or not. However, the result does not respect the Kramers-Kroenig relations, which creates a significant issue during the unwrapping step that strongly depends on the dynamic range.

It has been shown that it is possible to partially solve this problem by including a correction of the unwrapped measured phase using partial Kramers-Kroenig relations [72]. Moreover, in the error function 1.36, both the modulus and the phase errors have the same weight. Any weighting other than 1 amplitude unit/rad could improve or diminish the efficiency and accuracy of the algorithm.

Both the iterative and the optimization techniques show good results, but they are limited: there is still a need for a precise measurement of the sample thickness or the implementation of an additional optimization step to retrieve it [73]. In addition, a low dynamic range of the measured data or a strong absorption in the sample lead to difficulties in obtaining the refractive index following the Kramers-Kroenig relations. Indeed, the phase is lost in the frequency range where the signal value is below the noise level [72], which implies an additional step while performing the phase unwrapping. This step includes more assumptions of the number and the shape of the absorption peaks. Furthermore, the arbitrary weighting between the phase and the amplitude, as well as an arbitrary limitation of the bandwidth to avoid the range of low dynamic range, limits the robustness and thus the expansion of the optimization method.

Nevertheless, with the complex refractive index retrieved by these techniques, it is possible to extract further information about the material itself. For this purpose, the refractive

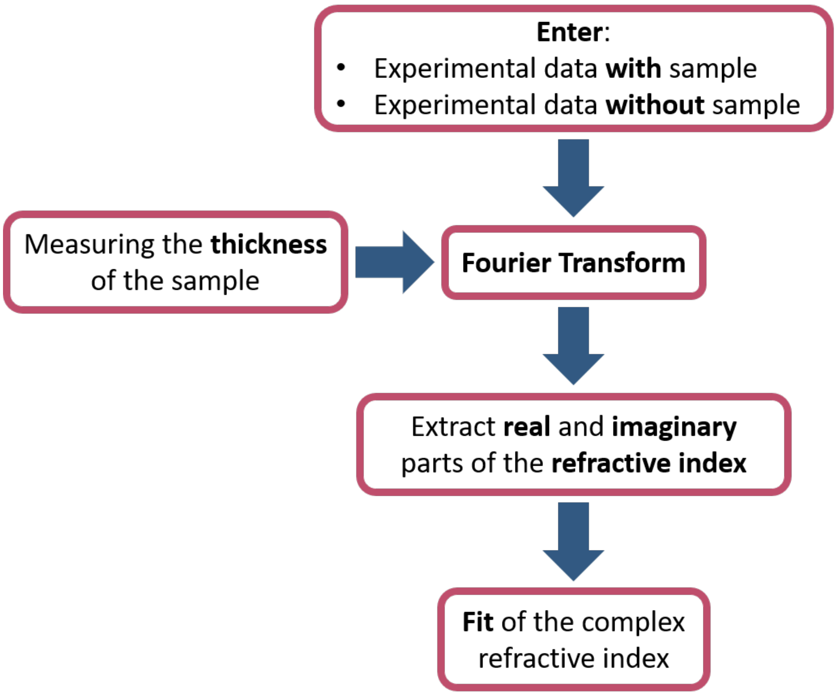


FIGURE 1.8: The fitting process in the frequency domain

index is fitted using permittivity models such as described in section 1.1.3 [74–80], or another possibility is to fit directly the dielectric permittivity [81–85]. Indeed, the parameters of these permittivity models give insight on the motion and the state of the charges inside a sample, or on the signature of a compound. To retrieve them with a fitting process in the frequency domain, optimization routines are used [86, 87], such as genetic algorithms [88], or nonlinear least square fitting [76, 78, 84, 89] for example based on the Levenberg-Marquardt algorithm [81, 90, 91]. Some groups are also using home-made softwares [92, 93]. The principle of the fitting process is shown in figure 1.8, it has shown promising results, but has several drawbacks as previously described, this is why a few researchers are looking into fitting the data in the time-domain instead.

1.3 Retrieving information in the time-domain

THz-TDS outputs a signal as a function of time, which means that a Fourier transform has to be applied in order to fit the data in the frequency domain. In addition to losing the phase for strong absorption, this produces artifacts that can degrade the quality of the fit and yield misleading results. As a result, we decided to create our own fitting software, coded with the Python language, in which the fits are done in the time-domain [94].

In this case, we need two types of information to begin the optimization routine:

1. The experimental data, that is two time-traces: one without sample, $E_{ref}(t)$, and one with sample, $E_{sample}(t)$.

2. A permittivity model depending on a set of parameters $\{p_i\}$ depicting how the sample transforms the reference pulse into the modeled one ($E_{model}\{p_i\}(t)$).

Then, the error function to minimize is:

$$Err\{p_i\} = \sum_{t=0}^{t=t_{max}} (E_{model}\{p_i\}(t) - E_{sample}(t))^2 dt \tag{1.37}$$

The result given by this function changes according to the values the model parameters $\{p_i\}$ take. The goal of the optimization algorithm is to find the parameters that minimize it. This function is proportional to the electromagnetic energy, which helps the user interpret the results and understand any discrepancies in either the experiment or the model. For instance, it facilitates the understanding of any divergence or convergence of the fit algorithm to some local minimum, or validates the choice of a model during the optimization. Moreover, Parseval's theorem states that the norm of a function is the same as the norm of its Fourier transform, meaning that:

$$Err\{p_i\} = \sum_{\omega=\omega_{min}}^{\omega_{max}} \left|\tilde{E}_{model}\{p_i\}(\omega) - \tilde{E}_{sample}(\omega)\right|^2 d\omega \tag{1.38}$$

This is extremely convenient as it allows the calculation of the error function in both the time and frequency domain without performing a Fourier transform at each iteration. Specifically, since all the refractive index models are defined in the frequency domain, this formula allows one to perform the time-domain optimization while computing the error function in the frequency domain.

1.3.1 Optimization algorithms

To perform the optimization, we implemented several routines in order to have a versatile tool adapted to any parameter retrieval. As a result we decided to use a library offering several optimization algorithms (the Python-based optimization package called PyOpt [95]), which is designed to formulate and solve nonlinear constrained optimization problems. The main advantage of this package is that it includes 20 different optimization algorithms, allowing the users of Fit@TDS (the name of our time-fitting software) to change the algorithm according to their specific needs. In addition, PyOpt allows parallelization, which is extremely useful for decreasing the computation duration. Here is a short description of the different algorithms we use in the software.

Swarm particle algorithms

A swarm particle algorithm solves the problem by having a population of candidate solutions, the particles, and moving these particles around in the search-space according to a formula over the particle's position and velocity. Each particle's motion is influenced by its local best known position, but is also guided towards the best known positions in the search-space, which are updated as better positions are found by other particles. Therefore, the swarm moves slowly towards the best solutions. This algorithm ensures to find a global minimum of the error function.

The swarm routine we implemented is called "ALPSO" (Augmented Lagrangian Particle Swarm Optimizer), which is an improvement on the basic "swarm particle optimization" approach used in constrained engineering design and the optimization problem [96]. It is a parallel augmented Lagrange multiplier particle swarm optimizer developed in Python [95]. Its advantage is that it is a versatile method that makes no or few assumptions about the problem being optimized, and can search very large spaces of candidate solutions. However, it is not optimized in terms of computation time, and is the slowest out of all the proposed algorithms.

Gradient descent optimization algorithms

The principle of a gradient descent is to take repeated steps in the opposite direction of the gradient of the function at the current point, because this is the direction of steepest descent. The derivative (or slope) at the starting point will be steeper, but as new parameters are generated, the steepness should gradually reduce until it reaches the lowest point on the curve, which is the point of convergence.

We use two different gradient descent algorithms. The first one is called "L-BFGS-B" (Limited-memory Broyden–Fletcher–Goldfarb–Shannon Boxed algorithm), from the scipy library, and the second one is called "SLSQP" (Sequential Least-Squares Programming). They both use quasi-Newton methods [95]. Contrary to the "ALPSO" algorithm, they have a very low computation time, but they only provide a local minimum, which means that the first guess of the parameters must be close to their optimized value.

Dual annealing algorithms

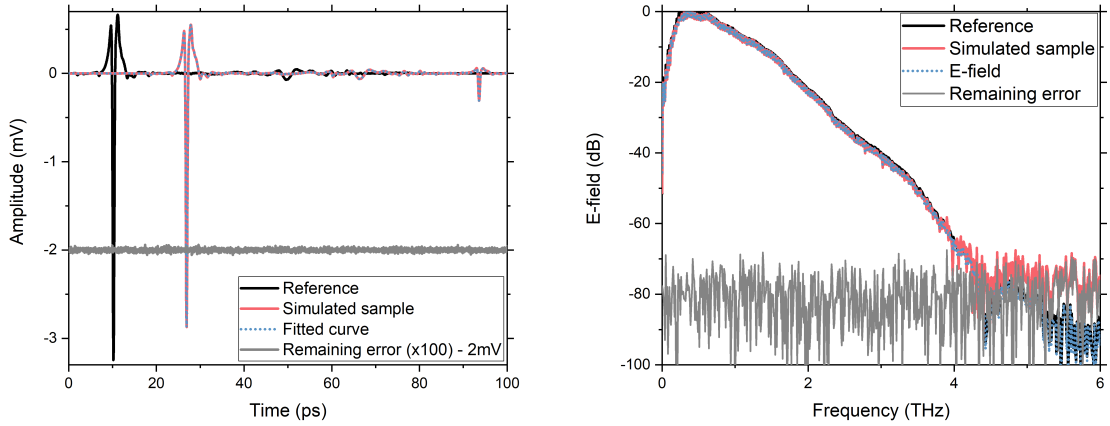
There is also a dual annealing algorithm provided in fit@TDS, which is based on the simulated annealing optimization algorithm. In this a simulated annealing routine, a candidate solution is modified in a random way and the modified solutions are accepted to replace the current candidate solution probabilistically. This means that it is possible for worse solutions to replace the current candidate solution. The probability of this type of replacement is high at the beginning of the search and decreases with each iteration, controlled by the "temperature" hyperparameter.

Among several modifications to the simulated annealing approach, in dual annealing, a local search algorithm can be applied to the solution found by the simulated annealing search [97]. This way, the algorithm can locate the search-space where the optimal solution is and then locally search the optima in the area. This diminishes the computation time and gives an alternative to a swarm particle algorithm when there is no previous knowledge on the parameters to retrieve.

1.3.2 Validation of the method on a simulated sample

To showcase the efficiency of our software, let us use it on a simulated sample. To do so, I recorded a reference spectrum (hence there is no sample inside the setup) with the TeraSmart THz-TDS setup by Menlo Systems GmbH (with 1000 accumulations). Then, I simulated the time-trace we would obtain if a 5 mm thick sample was placed in this setup by convoluting the transfer function of equation 1.10 with the reference time-trace. The complex permittivity of the simulated sample was set to follow a Lorentz model with the following parameters: $\epsilon_\infty = 4$; $\Delta\epsilon = 0.001$; $\omega_0 = 0.5\text{THz}$; $\gamma = 0.1\text{THz}$, where:

$$\epsilon(\omega) = \epsilon_\infty + \frac{\Delta\epsilon\omega_0^2}{\omega_0^2 - \omega^2 + j\omega\gamma} \tag{1.39}$$



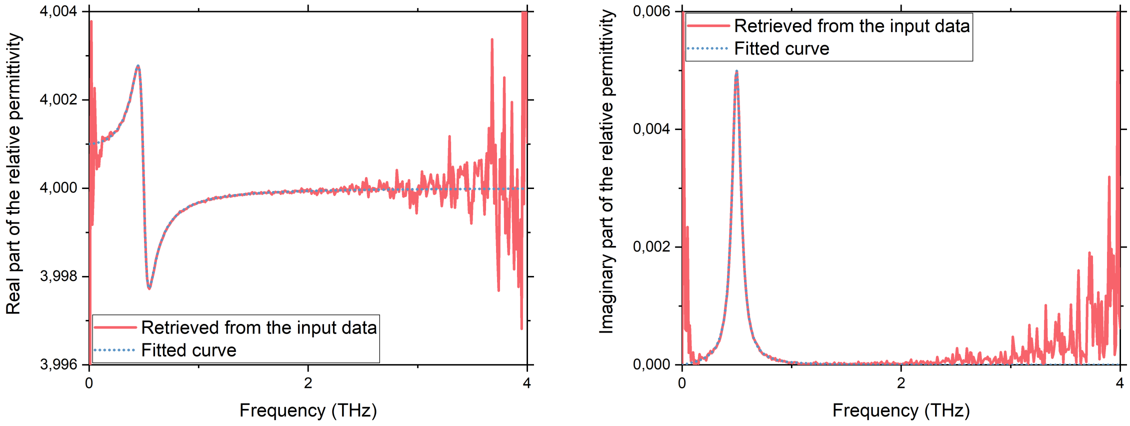
(A)

(B)

FIGURE 1.9: (A) Time-traces of the experiment. In black is the reference time-trace, in pink the time trace of the simulated sample, in blue dots the resulting fitted time-trace and in grey the difference between the pink and blue curves (times 100) and shifted for clarity. (B) The corresponding spectra to the time-traces on the left.

Finally, to be closer to real experimental conditions, I added a Gaussian white noise to the sample time-trace, with a magnitude equivalent to the level of the high frequency noise observed in figure 1.9 (5.10^{-4} of the amplitude unit used, and approximately -80 dB compared to maximum power spectral density). The data at the end of the reference time-trace, on a segment length corresponding to the time delay between the two THz pulses (with and without sample), was set to zero to avoid any folding effect due to the periodicity of the fast Fourier transform (FFT) ("Zero padding").

These two time-traces were used as inputs into fit@TDS, their time-traces and spectra are shown in figure 1.9. The retrieved permittivity is shown in figure 1.10. Then, after entering guesses and bounds for all the parameters (±1% for the thickness, ±50% for the other parameters), I chose the dual annealing algorithm with 1000 iterations. The optimization took around 10 s on a common personal computer (Intel(R) Core$^{\text{TM}}$ i5-8250U CPU @ 1.60 GHz) and retrieved the correct parameters, with a final error equal to 0.086%. In fact, the computation time strongly depends on the algorithm chosen, the size of the search-space, and thus on the bounds we specify. In the difference between the fit and the simulated data, we can see an extremely small discrepancy between the targeted sample time-trace and the fitted one. This discrepancy is due to the noise added to the simulated sample time-trace, meaning that the algorithm converged and that the L^2 residual error is simply the L^2 of the added noise. The results for the Lorentz model parameters are very close to the targeted ones, their retrieved value is: $\epsilon_\infty = 4.00000506$; $\Delta\epsilon = 0.000998788733$; $\omega_0 = 0.499980203$THz; $\gamma = 0.0998637352$THz and thickness = 4.99999319mm. Moreover, the advantage of fitting in the time-domain is highlighted by figure 1.10 in which we see the artifacts that appeared due to the frequency domain retrieval on the pink curve. The blue curve is obtained with the time-domain fitting process.



(A) (B)

FIGURE 1.10: Permittivity retrieved with the frequency domain formula from section 1.1.2 (in pink) and retrieved from the fit (in blue). (A) is the real part and (B) is the imaginary part.

1.4 Improving the extraction of information by taking experimental noise into account

Our time-domain fitting software works well for simulated samples or samples where a Gaussian random noise has been added. However, as for other fitting methods, the permittivity model used for the fit is an ideal noiseless one. This is not a problem if we make the approximation that the experimental noise is a Gaussian white noise but, when correlated and non-Gaussian noise is also present, the algorithm considers its distribution as part of the data to fit and can mistake it for real physical features. To sum up, it offers poor discrimination between good models and bad ones, as studied by [98]. This is problematic when analyzing all sorts of samples. For example, the model chosen in the case of gases will change the shape of the absorption lines in the spectrum, leading to different interpretations as the broadening can be caused by collisions or by the Doppler effect. In liquids, and more precisely water, the question is even more important as there is no consensus on the correct model in the THz range. Several models are discussed, hence different physical interpretations, so a good criterion is necessary to pick the right one.

1.4.1 Implementing a new error function for the fitting process

So, how do we take experimental noise into account during the fitting process? Our first idea was to enter a new information, the experimental noise envelope, corresponding to the difference between two measurements of the sample inside the setup. This information, $E_{noise}(t)$, is used to define a new error function:

$$Err\{p_i\} = \sum_{t=0}^{t=t_{max}} \frac{(E_{model}\{p_i\}(t) - E_{sample}(t))^2}{E_{noise}(t)} dt \tag{1.40}$$

However, if it indicates where the uncertainties are located, it doesn't take into account the input noise and correlations: when the input noise is convoluted by the transfer function, its contribution is found in all of the Fabry-Perot echoes in the sample time-trace. That is why we looked into the maximum likelihood estimator method explained in [98]. For this method, there is an additional step of creating a noise matrix that will be implemented later into the error function calculation. This noise matrix contains information about the noise both in the reference and sample time-traces but also information about the correlations. The creation of such a noise matrix is explained later in section 1.6.2.

Now, it is possible to use a least-square algorithm to minimize an error function taking into account this noise convolution matrix (NCM) as all of the information about the noise is present. This new error function is [98]:

$$Q(\boldsymbol{p_i}) = [\boldsymbol{r}(\boldsymbol{p_i})]^{\intercal}\boldsymbol{r}(\boldsymbol{p_i}) \tag{1.41}$$

where

$$\boldsymbol{r}(\boldsymbol{p_i}) = [\boldsymbol{M}_{noise}(\boldsymbol{p_i})]^{-\frac{1}{2}}[(\boldsymbol{E}_{model}(\boldsymbol{p_i}) - \boldsymbol{E}_{sample})] \tag{1.42}$$

and $\boldsymbol{M}_{noise}(\boldsymbol{p_i})$ is the noise convolution matrix, $\boldsymbol{E}_{model}(\boldsymbol{p_i})$ is the vector containing the fitted time-trace data points, $\boldsymbol{E}_{sample}$ is the vector containing the experimental sample time-trace data points.

With this error function, the software is only fitting the data. A last step to compare two models between one another is to use the Akaike criterion [98]:

$$Q_{compare}(\boldsymbol{p_i}) = Q_{optimum}(\boldsymbol{p_i}) + 2N_{p_i} \tag{1.43}$$

where N_{p_i} is the number of free parameters in the model and $Q_{optimum}(\boldsymbol{p_i})$ is the result of equation 1.41 for the optimum parameters. The smaller the result, the better the model. Moreover, this way, the optimal model is the one that has the best fit while not being over-parametrized.

This equation is the last one from [98], and has certain limitations. Indeed, their method relies on a Monte-Carlo study that cannot be applied on real experiments as we would have to record a large number of spectra consecutively (their example shows 50 waveforms), which is impossible for stability reasons (temperature variations, etc...). Moreover, their transfer function is equal to 1, meaning that they don't have a sample inside their simulated experiment, they only compare two reference time-traces. Finally, they suppose that experimental noise is non-deterministic, which is not true in a real THz-TDS setup, as we will see in the next paragraph.

1.4.2 Experimental noise is deterministic

Experimental noise arises from different parts of the setup. It is the combination of different types of noises, such as dark noise or delay noise. The causes are numerous, such as fluctuations of the femtosecond laser intensity, vibrations in the delay line, noise in the bias of the photoconductive antennas, reflections within components, noise in the transimpedance amplifier, or influence of the temperature on the optical fibres as illustrated on figure 1.11. The sample can also be seen as a source of noise, or more generally non-reproducibility, as it may differ from the model due to its inhomogeneity, its scattering properties, or simply

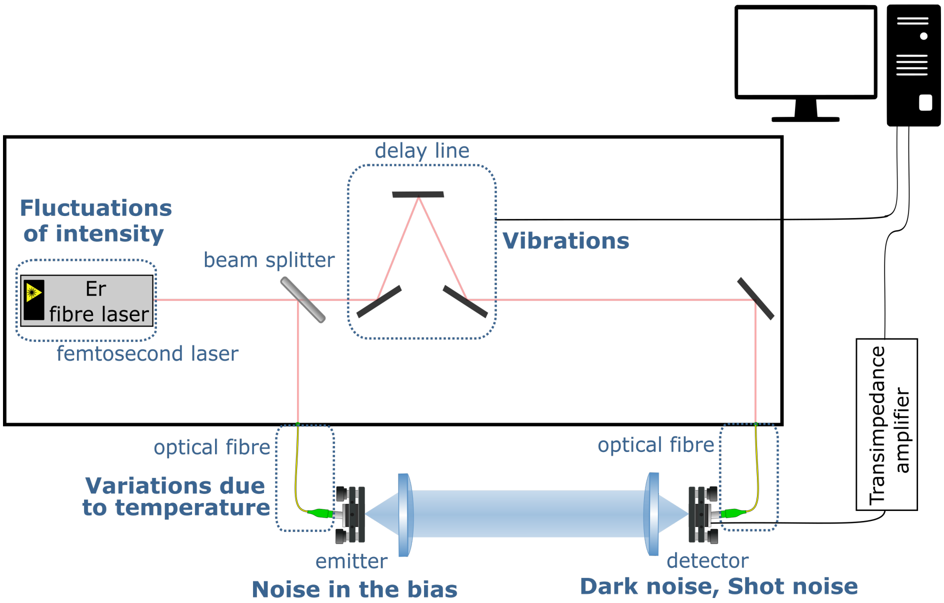


FIGURE 1.11: THz-TDS setup and its sources of experimental noise

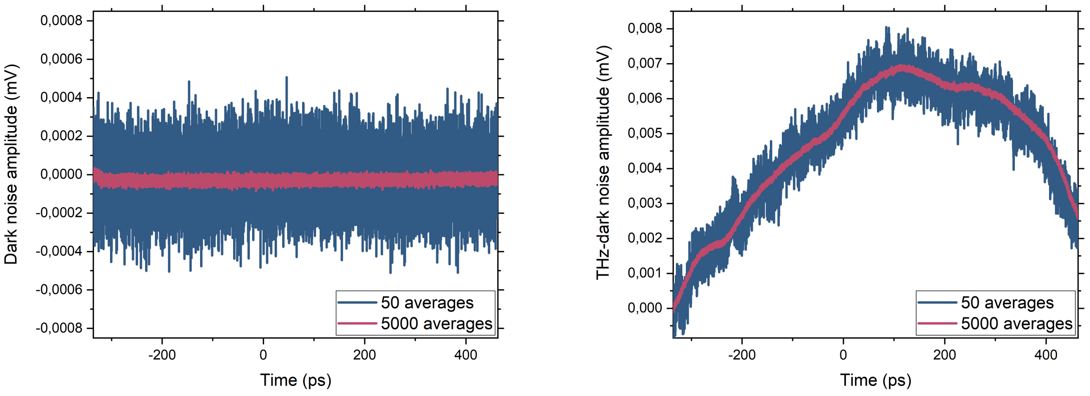
because of the temperature variations of its environment during the experiments. Uncorrelated Gaussian white noises can be minimized by increasing the number of averaging when data is collected but averaging during a long time can lead to other artifacts due to the variations in the environment so that a compromise has to be made. Moreover, other noises are correlated to the signal or determinated by the signal and thus independent from the averaging and cannot be minimized by accumulation during the experiment but should instead be modelled.

There are a lot of different contributing noises and, if some are uncorrelated and can indeed be depicted as Gaussian noises, others are to be analyzed before approximating that all experimental noises are Gaussian noises. The approximation is valid for noises such as thermal noise, Johnson noise (or thermal noise), Shot noise, amplification noise or laser noise because, no matter their shape, the error bars of THz-TDS experiments are greater than their contribution amplitude for an adapted averaging. However, other noises contribute more to the experimental noise and these are the ones we are going to analyze here. It is important to keep in mind that the noise profiles are also extremely dependent on the setup.

We analyse the four noises shown in figure 1.12 separately in order to find a way to remove them from the experimental signal so that they don't impair the fitting process. If they can't be removed, a procedure will be implemented in the fitting software to take their contribution into account.

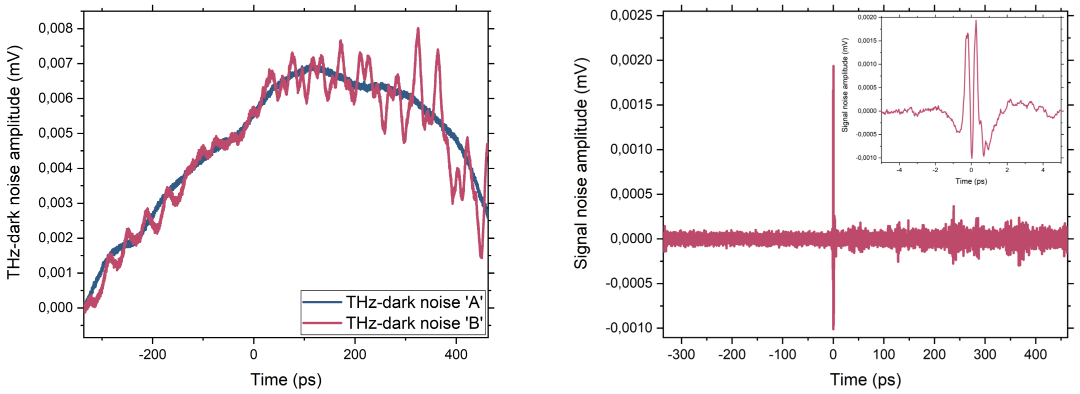
Dark noise (absence of IR excitation on both antennas)

We call "dark noise" the random signal that is generated by the detector. Even when no THz (or IR) pulse is shone onto the detector, some charge carriers exist in the semiconductor and hole-electron pairs can be created, generating a small current leading to this dark noise. We recorded it by unplugging the two optical fibres that link the femtosecond laser to the PC antennas so that they are no longer excited. The result is shown on figure 1.12a. This



(A) Dark noise (no optical excitation on both photoconductive antennas) recorded with two different averages, its contribution decreases when the averaging increases, independently of the time window.

(B) THz-dark noise (no optical excitation on the emitter) recorded with two different averages, its contribution decreases when the averaging increases. However, it is a deterministic signal that depends on the chosen time window.



(C) The two types of THz-dark noises, in blue without excitation on the emitter, in pink with optical excitation on the emitter and an absorber between the two antennas.

(D) Remaining noise after removing the contributions of the THz-dark noise and fixing the delay. A zoom between -5 ps and 5 ps is provided.

FIGURE 1.12: The most contributing noises to the THz-TDS experimental noise.

noise is not deterministic and its amplitude decreases as we increase the averaging number (or the integration time). Indeed, its noise spectral density goes from $5.10^{-6}\mu\text{V}/\sqrt{\text{Hz}}$ for 50 averages to $1.10^{-6}\mu\text{V}/\sqrt{\text{Hz}}$ for 5000 averages. As expected it does not depend on the time window scanned. The approximation that it is a Gaussian white noise can be made. In conclusion, it impacts the fitting process, as it will be present in the Fabry-Perot echoes, but does not invalidate the fit because of its amplitude and the fact that it is non deterministic. It will be neglected in the rest of the study.

THz-dark noise

We call “THz-dark noise” the signal recorded by the THz-TDS in the absence of emission of a THz pulse. Hence, the detector is excited by the femtosecond laser but there is no THz pulse received as the optical fibre going to the emitter is unplugged. As we can see on figure 1.12b, the excitation has an impact on the noise recorded, its amplitude is now one order of magnitude greater than the previous recording (Dark noise 5000 averages). Moreover, it seems that there are two components to this noise: an uncorrelated variable one and a deterministic uncorrelated one (that we can call artefact).

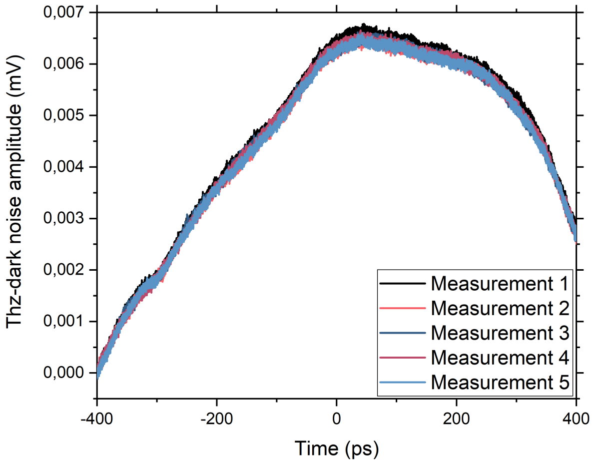


FIGURE 1.13: Five measurements of the THz-dark noise, the shape is very repeatable.

Figure 1.13 shows 5 measurements of this THz-dark noise with an averaging number of 5000 (*i.e.* an integration time of around 20 min). It is clear that its shape is very different from a Gaussian noise and is repeatable, as well as its amplitude. This artefact varies with time, and its amplitude contribution is most important at low frequencies, below 200 GHz, that is close to the frequency limit of the system. Hence, it is of utmost importance to remove this artefact, as it is a signal that does not go through the sample, otherwise it is easy to mistake it for real physical features when it is added to the pulse signal. This artefact is not present in standard THz-TDS setups like the ones based on Titanium-sapphire lasers and slow mechanical delay lines. It seems to be typical of the Menlo TeraSmart design.

In a usual experiment, both of the PC antennas are excited by the femtosecond laser. Thus, we recorded the THz-dark noise signal again by blocking the slot occupied by the sample during an experiment with a THz absorber (microwave absorber panel from Microwave Absorbers Inc., reference MA-F025K). We will call this THz-dark noise signal ’B’ while calling the first one ’A’. The result is shown in figure 1.12c. Ideally, this signal should have been the same as the previous one but it is clearly not the case. It implies that some parasitic signal arrives to the detector even when we block the path of the main signal. Therefore, it is most probably purely optical. The contribution of this parasitic signal, as repeatable as the previous one, is also most important at frequencies below 200 GHz, meaning that we should filter the data in order to get rid of both THz-dark noises. That is why we implemented a high-pass filter into fit@TDS in order to get rid of the unwanted data. This filter is a smooth step function, with the following formula:

$$y[n] = 0.5 + 0.5\tanh\left(\frac{(f[n] - f_{cut})\alpha}{f_{cut}}\right) \tag{1.44}$$

Where f is the frequency, f_{cut} the cut frequency (200 GHz in our case), and α the sharpness of the filter, set at 10 here. This can be adjusted to each setup according to the company's claims considering the frequency range. With a cut frequency of 200 GHz, we obtain the result shown in figure 1.14. The deterministic component of the noise (the artefact) is no longer there and the remaining THz-dark noise signals, without and with excitation on the emitter, have, after this filtering, a contribution of the same order of magnitude as the dark noise. In conclusion, this perturbation is no longer deterministic and can be neglected after the filtering.

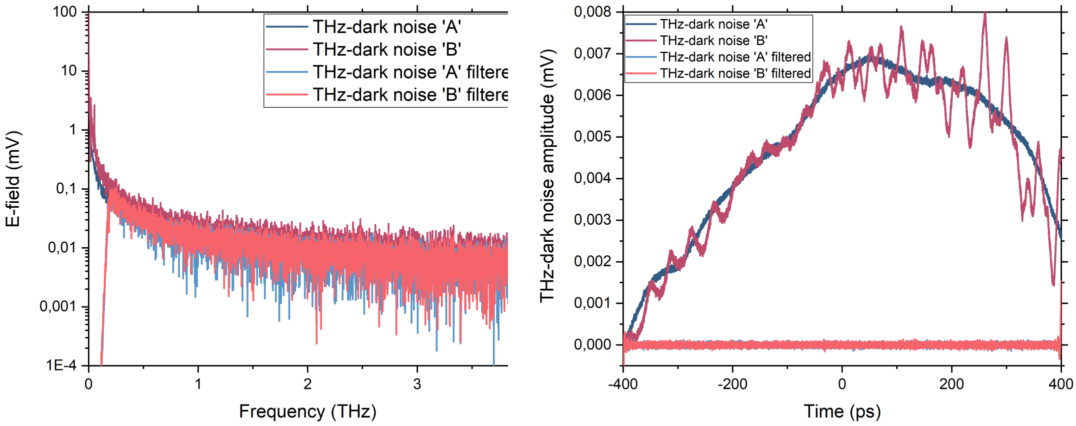


FIGURE 1.14: THz-dark noises 'A'and 'B', before and after the filtering process. On the left is the frequency domain and on the right the time domain.

Delay noise

When recording the same time trace repeatedly, it appears that there is a small delay between each pulse of the order of a few femtoseconds, smaller than the laser pulsewidth. In fact, the variations in the delay line mechanical positioning can cause a mechanical jitter and a small time drift during an experiment. Then, this delay corresponds to the average of the time drift during the time the spectrum was recorded. This small delay can lead to imprecision when retrieving the refractive index and/or the thickness of a sample. In order to lessen its contribution to the experimental noise, we implemented a feature in fit@TDS that allows the user to readjust two pulses so that there is no more delay between them.

1.4.3 There is a small noise depending on the signal left

The other sources of experimental noise cannot be recorded separately. The only possibility is to find out what is left in the noise when all of the previous ones have been fixed. To do that, we take 5 reference measurements, that is with both antennas excited by the femtosecond laser and no absorber inside the setup. Then, we use the high pass filter with a 200 GHz frequency cut and readjust the average delay between the pulses. Finally, we subtract each of the 5 signals to the mean. Figure 1.12d shows that there is still a small noise left, which is not deterministic but correlated to the signal.

This noise corresponds to a sum of signals coming from the other sources of noise in the setup, such as delay drift (we only adjusted the mean time drift), shot noise and laser power fluctuations. This noise is correlated to the signal. It means that noise at one particular time contains information about the noise at neighbouring times. We found that this leftover noise is proportional to the signal and its second derivative. The first derivative corresponding to the average delay drift, it has already been fixed. Thus, before any fitting process can be done on the data, one has to fit this leftover noise, $y(t)$, with the following model:

$$y(t) = a \times x(t) + b \times \ddot{x}(t) \tag{1.45}$$

Where $x(t)$ is the time-domain signal and $\ddot{x}(t)$ its second derivative. a and b are the proportionality coefficients.

Once this final step has been done, we can see on figure 1.15, that most of the experimental noise has been eliminated. There is only a small contribution left. We found that this signal corresponds to the flickered noise, which is non deterministic. Therefore, by taking the envelope of this noise, we will now apply the method proposed by [98] and retrieve the corresponding convolution noise matrix and use the new error function during the fitting process.

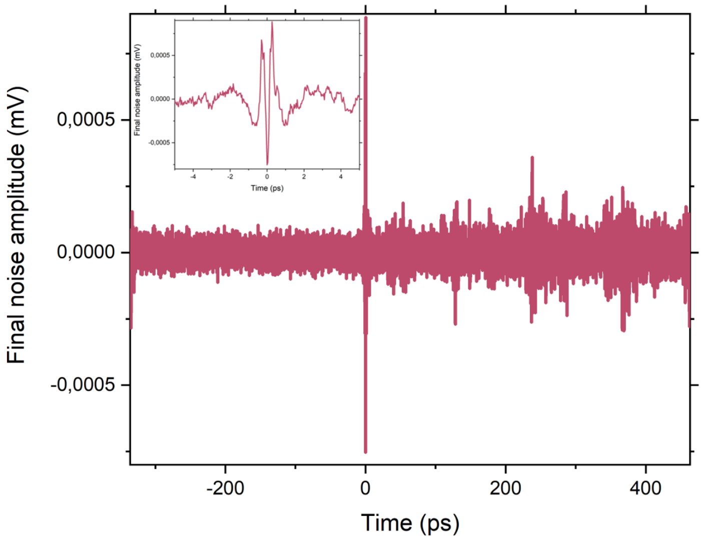
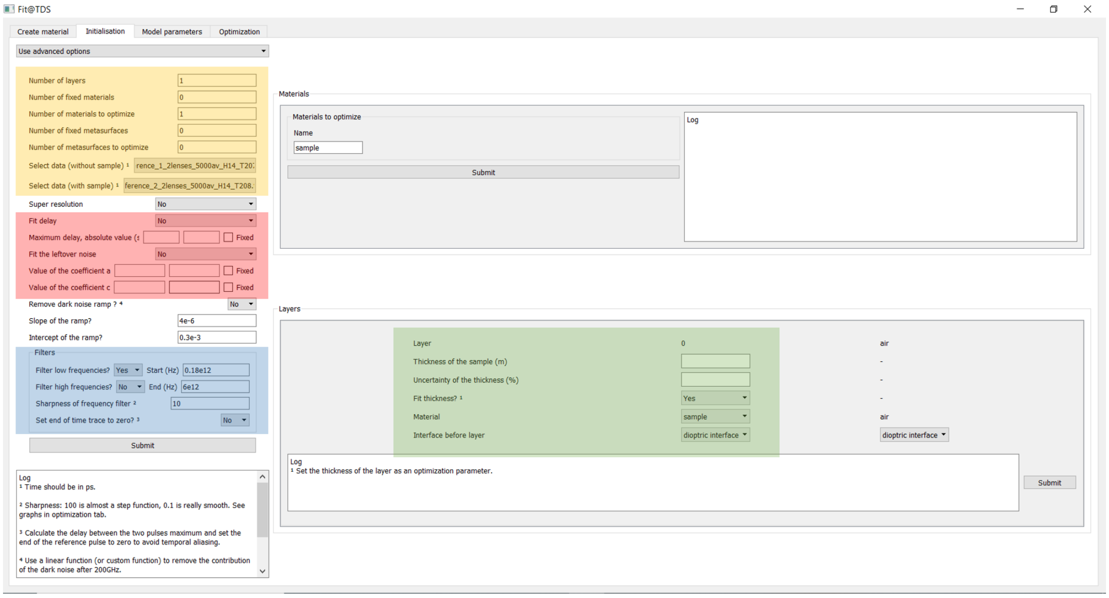


FIGURE 1.15: Experimental noise recorded on our Menlosystems THz-TDS setup after having removed the contribution from other sources of noise and the term proportional to the signal derivative. A zoom between −5 ps and 5 ps is also provided.

1.5 An overview of fit@TDS

Now that the theory behind the software has been explained and all the different features mentioned, I will give an overview of the interface of the software. It has greatly improved since [94], and this overview corresponds to its current state, three years later. The interface is now composed of four tabs. The first one is called "create material" and allows the user to create a file containing the model parameters of a material (he chooses). This file is useful if we want to use a fixed material in the study.



Use advanced options

Number of layers: 1
Number of fixed materials: 0
Number of materials to optimize: 1
Number of fixed metasurfaces: 0
Number of metasurfaces to optimize: 0
Select data (without sample) ¹: rence_1_2lenses_5000av_H14_T20
Select data (with sample) ¹: ference_2_2lenses_5000av_H14_T208.
Super resolution: No
Fit delay: No
Maximum delay, absolute value (s ☐ Fixed
Fit the leftover noise: No
Value of the coefficient a ☐ Fixed
Value of the coefficient c ☐ Fixed
Remove dark noise ramp ? ⁴: No
Slope of the ramp?: 4e-6
Intercept of the ramp?: 0.3e-3
Filter low frequencies?: Yes Start (Hz): 0.18e12
Filter high frequencies?: No End (Hz): 6e12
Sharpness of frequency filter ²: 10
Set end of time trace to zero? ³: No

Log
¹ Time should be in ps.

² Sharpness: 100 is almost a step function, 0.1 is really smooth. See graphs in optimization tab.

³ Calculate the delay between the two pulses maximum and set the end of the reference pulse to zero to avoid temporal aliasing.

⁴ Use a linear function (or custom function) to remove the contribution of the dark noise after 200GHz.

Name: sample

Log

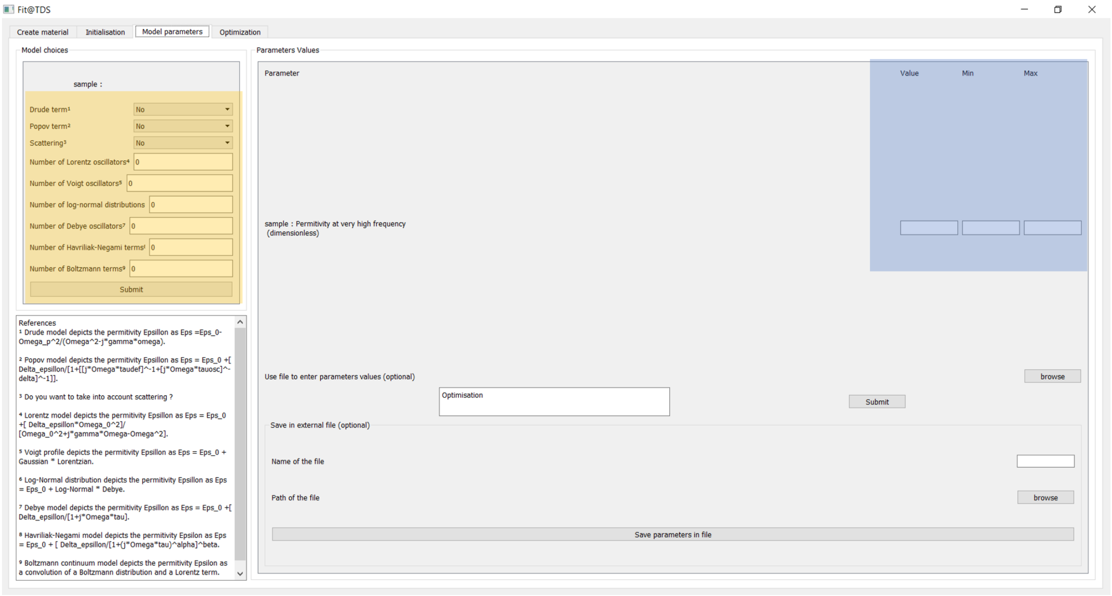
Layer	0	air
Thickness of the sample (m)		-
Uncertainty of the thickness (%)		-
Fit thickness? ¹	Yes	-
Material	sample	air
Interface before layer	dioptric interface	dioptric interface

Log
¹ Set the thickness of the layer as an optimization parameter.

FIGURE 1.16: Initialization tab of the software

The second tab is called "Initialisation" and is shown on figure 1.16. The yellow part consists in entering information on the sample, such as the number of layers it contains, the number of materials we want to optimize and the number of fixed materials (meaning that we already know the parameters of this layer). The two bottom lines correspond to the path of the files containing the time trace data of the reference and of the sample. The red part gives a choice to fit the delay and the noise proportional to the signal and its second derivative. If we choose to do so, we have to enter a guess for the parameters as well as boundaries for their search. In the blue part, the user chooses to apply filters on the signal or not. There are a high pass filter, a low pass filter, and a feature to set to zero the final points of the time trace of the reference, on a length corresponding to the delay between the reference and the sample pulses. This last feature is useful to avoid folding due to the periodicity of the fast Fourier transform. Finally, in green, the user enters a guess for the thickness of the sample and the uncertainty on this value.

The next tab is called "Model parameters" and is shown on figure 1.17. The yellow part corresponds to the permittivity model choice, where are displayed all the models available, such as Drude, Lorentz, or Debye. Any other model can be added to the software easily. On the right, in the blue part, the user has to enter a guess and boundaries for each parameter of the chosen model. There is also the possibility to directly enter a file containing this information, which is useful when the model contains a large number of parameters.



sample :

Drude term[1] No

Popov term[2] No

Scattering[3] No

Number of Lorentz oscillators[4] 0

Number of Voigt oscillators[5] 0

Number of log-normal distributions 0

Number of Debye oscillators[7] 0

Number of Havriliak-Negami terms[8] 0

Number of Boltzmann terms[9] 0

References

[1] Drude model depicts the permitivity Epsillon as Eps =Eps_0-Omega_p^2/(Omega^2-j*gamma*omega).

[2] Popov model depicts the permitivity Epsillon as Eps = Eps_0 +[Delta_epsillon/[1+[[j*Omega*taudef]^-1+[j*Omega*tauosc]^-delta]^-1]].

[3] Do you want to take into account scattering ?

[4] Lorentz model depicts the permitivity Epsillon as Eps = Eps_0 +[Delta_epsillon*Omega_0^2]/[Omega_0^2+j*gamma*Omega-Omega^2].

[5] Voigt profile depicts the permitivity Epsillon as Eps = Eps_0 + Gaussian * Lorentzian.

[6] Log-Normal distribution depicts the permitivity Epsillon as Eps = Eps_0 + Log-Normal * Debye.

[7] Debye model depicts the permitivity Epsillon as Eps = Eps_0 +[Delta_epsillon/[1+j*Omega*tau].

[8] Havriliak-Negami model depicts the permitivity Epsilon as Eps = Eps_0 + [Delta_epsillon/[1+(j*Omega*tau)^alpha]^beta.

[9] Boltzmann continuum model depicts the permitivity Epsilon as a convolution of a Boltzmann distribution and a Lorentz term.

sample : Permitivity at very high frequency (dimensionless)

Use file to enter parameters values (optional)

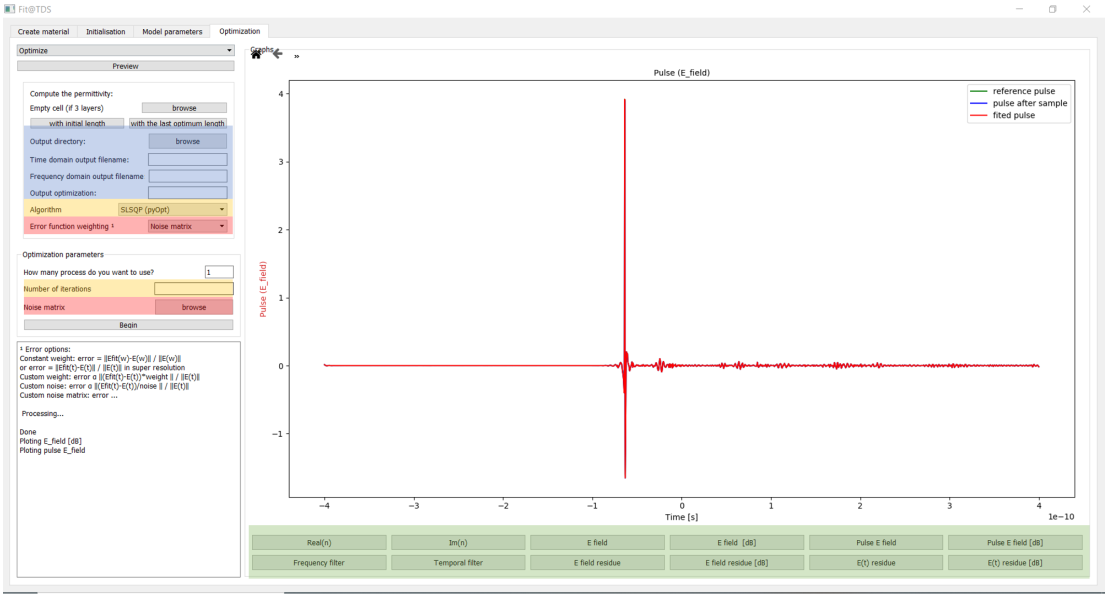
Optimisation

Save in external file (optional)

Name of the file

Path of the file

FIGURE 1.17: Model parameters tab of the software



Compute the permittivity:

Empty cell (if 3 layers)

Output directory:

Time domain output filename:

Frequency domain output filename

Output optimization:

Algorithm SLSQP (pyOpt)

Error function weighting [1] Noise matrix

Optimization parameters

How many process do you want to use? 1

Number of iterations

Noise matrix

```
¹ Error options:
Constant weight: error = ||Efit(w)-E(w)|| / ||E(w)||
or error = ||Efit(t)-E(t)|| / ||E(t)|| in super resolution
Custom weight: error α ||(Efit(t)-E(t))*weight || / ||E(t)||
Custom noise: error α ||(Efit(t)-E(t))/noise || / ||E(t)||
Custom noise matrix: error ...

 Processing...

Done
Ploting E_field [dB]
Ploting pulse E_field
```

FIGURE 1.18: Optimization tab of the software

The final tab is called "Optimization" and is shown on figure 1.18. In blue, the user chooses the path and the name for the output files (the time domain data of the fitted pulse, its frequency domain data and the values of the optimal parameters). The yellow part corresponds to the algorithm choice and the number of iterations associated to it (as well as the size of the search-space for the swarm particle algorithm). The red part corresponds to the weighting of the fit, where we can enter the noise convolution matrix of the experiment. Finally, the red part highlights all the possible curves the software is able to display, such as the refractive index of the sample, its time domain waveform, its spectrum, or the differences between the experimental and fitted curves (residuals).

1.6 An example: Analysis of an α-lactose monohydrate pellet

The monohydrate-crystallized form of lactose, called α-lactose monohydrate, has become the "gold standard" in THz-TDS thanks to its easily recognizable absorption peaks at 0.53, 1.19 and 1.37 THz. Moreover, due to its availability and low cost, lactose samples are often used to test THz spectroscopy equipment [99], serve as a "first-try" sample in minor-volume detection techniques [94, 100], and are used as a mixture compound in approving methods for content quantification [101]. Furthermore, the study of lactose is important for pharmacology [102], medicine [103] and the food industry [104].

We chose to illustrate our fitting software with this sample because, first, it is a "textbook case". However, even if there are numerous studies in the literature, lactose is a complex sample due to its polymorphism making it a molecular crystal hard to simulate. The theoretical explanation of the lactose THz spectrum is rather rare in literature and is mostly based on gas phase DFT calculations considering a single lactose molecule, and not crystals, meaning the collective vibrations are not taken into account [105]. As a result, proper comparison of numerical simulation results and experiments has yet to be done. Finally, providing a good fit of lactose means that we can use the same method for other complex samples.

1.6.1 Experimental protocol

FIGURE 1.19: A pellet - disc composed of pressed powder - on the left, and its holder on the right. The holder was designed with Solidworks and built with aluminum in the laboratory workshop.

The lactose powder (α-lactose monohydrate) was purchased from Sigma-Aldrich Co. Ltd. (($\geq$ 99.9% total lactose basis including less than 4% β-lactose)) and was pressed into a pellet (there was no dilution) with a 13 mm diameter and 600 µm thickness, such as shown on figure 1.19. The measurement was made with a commercial Terasmart setup by Menlosystems, shown in figure 1.11, placed inside a nitrogen-purged box. To the two collimating lenses with a 50 mm focal length we added two focusing lenses with the same focal length since the diameter of the pellet is smaller than the diameter of the collimated THz beam. The sample was placed in the middle of the setup, held inside a home-made metallic holder (see figure 1.19). The reference was taken with the same holder without any pellet inside. In order to build the noise convolution matrices, we took two consecutive measurements in both cases. A time trace of the reference and one of the sample and their corresponding spectra are shown in figure 1.20. The oscillations we see on the lactose time trace are due to the echos inside the pellet as well as the absorption, visible under the form of a peak on the spectrum.

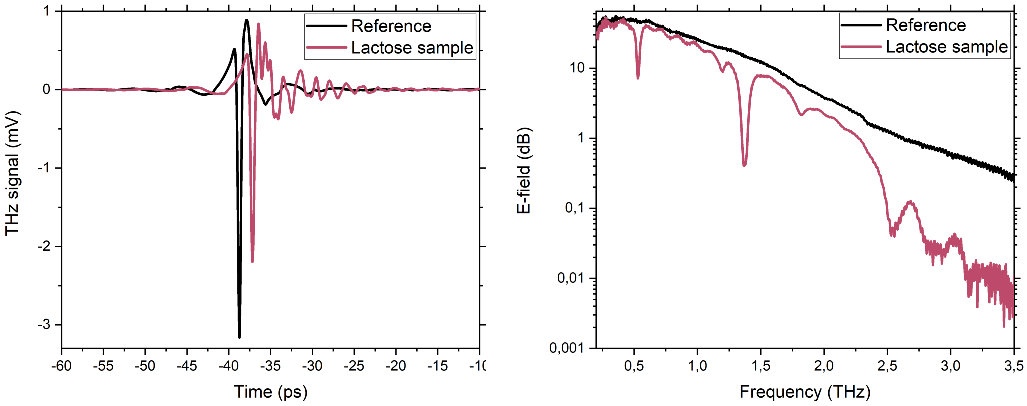


FIGURE 1.20: Time-trace and spectrum of the reference (*i.e.* the THz pulse goes through the holder without pellet inside, in black) and the sample (in pink). The time-trace measurement was actually made on a larger time range (between -80 and 120 ps but it is zoomed in for visibility.

1.6.2 Retrieval of the noise matrices

The first step after having acquired the data is to do the pre-processing explained in section 1.4.2. Since we have two consecutive measurements for the reference and two consecutive measurements of the sample, we can retrieve the remaining noise for each set by subtracting the two and applying the high-pass filter, correcting the delay and fixing the noise proportional to the signal and its second derivative. We are left with the two signals in blue shown in figure 1.21.

The next step consists in retrieving the envelopes of the signals in figure 1.21, which are displayed in pink. The remaining signal inside these envelopes is a Gaussian noise, thus all the information is contained in the envelope. Then, we take the square of this envelope. Finally, I created a Matlab program with the following steps:

1. With the two time-traces (from the reference and the lactose sample), we start by calculating the experimental transfer function of the experiment: $\tilde{\mathrm{TF}}(\omega) = \dfrac{\tilde{E}_{lactose}(\omega)}{\tilde{E}_{ref}(\omega)}$
2. We begin a Monte Carlo study, meaning that we repeat the following steps a large number of times:
    - We multiply each noise envelope by a random vector, to obtain $\sigma_{noise,ref}(t)$ and $\sigma_{noise,sample}(t)$.

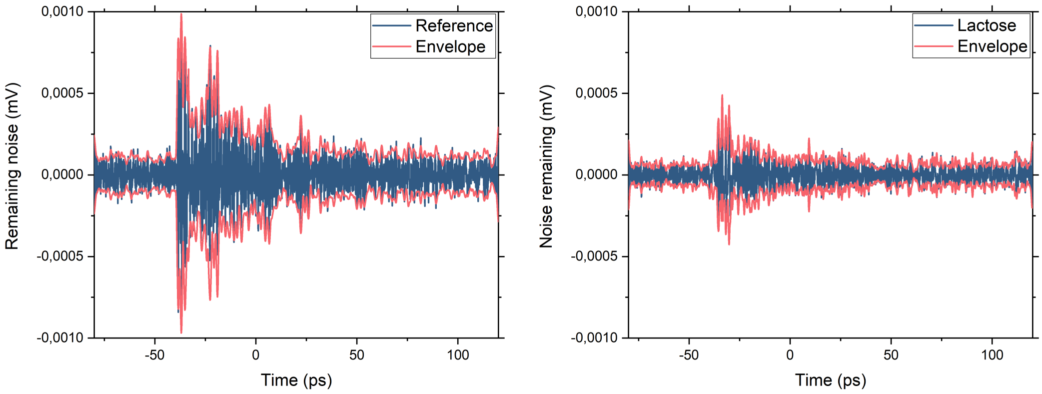


FIGURE 1.21: Remaining noise after the pre-processing of the reference time traces (on the left) and the sample time traces (on the right)

- For the reference, we apply the transfer function to this result in the frequency domain, and then go back to time domain: $\sigma_{noise,TFref}(t) = \mathrm{FT}^{-1}\left[\mathrm{FT}\left[\sigma_{noise,ref}(t)\right].\tilde{\mathrm{TF}}(\omega)\right]$ (where FT is short for Fourier Transform). This indicates the contribution of the reference noise in the Fabry Perot echoes as well as in the principal pulse.
- We multiply each result by its transpose in order to obtain a matrix: $U_{ref} = \sigma_{noise,TF(ref)}(t).\sigma^{\mathrm{T}}_{noise,TF(ref)}(t)$ and $U_{sample} = \sigma_{noise,sample}(t).\sigma^{\mathrm{T}}_{noise,sample}(t)$.

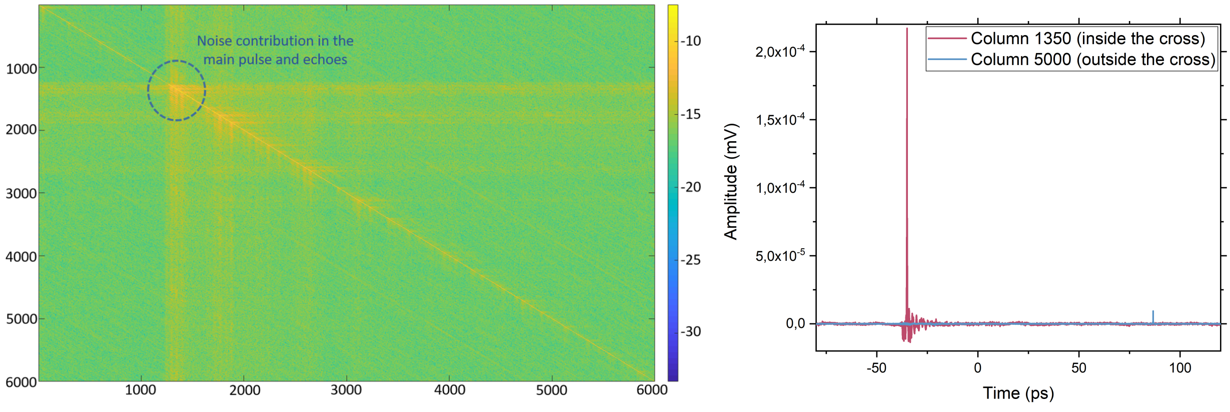
3. Once the Monte Carlo study over, we add all the results to obtain the final matrices and finally add the reference and sample noise convolution matrices: $U_{tot} = U_{ref} + U_{sample}$.
4. Following equation 1.42, we invert and take the square-root of this matrix to obtain the final NCM: $U_{final} = [U_{tot}]^{-\frac{1}{2}}$.

The result is shown in figure 1.22. To sum up, the diagonal terms correspond to the sum of the two noise envelopes and the crosses (off-diagonal non-zero terms) correspond to the contribution in the echoes of the noise on the reference. Indeed, the pink curve extracted from the noise matrix possesses oscillations, which means that the noise has a contribution in those. However, the blue curve is flat where the pulse is located. Now, we can use it as an input in fit@TDS to perform the fit of the lactose sample.

1.6.3 The fitting process

In the literature, the lactose absorption peaks are usually modelled by Lorentz oscillators. However, since there are different dynamics behind each peak, there is no reason why they should all have the same shape. Therefore, our goal is to use our new metric to compare Lorentz and Voigt models for each peak, which will give us more information on the microscopic structure of the sample. Indeed, an ideal crystal would have Lorentz absorption peaks, but defects inside the structure broaden the peaks, creating the need for a Voigt model. Their formula are given in equations 1.46 and 1.47.

$$\chi_{Lorentz}(\omega) = \frac{\Delta\epsilon{\omega_0}^2}{{\omega_0}^2 - \omega^2 + j\omega\gamma} \tag{1.46}$$



(A)

(B)

FIGURE 1.22: (A) Noise convolution matrix (NCM) of the lactose experiment. I traced the logarithm for better contrast. The big cross corresponds to the contribution of the reference noise in the principal pulse and the Fabry-Perot echoes, which correspond to the oscillations after the main pulse on the sample time trace. (B) Plot of two of the columns of the noise matrix. In pink is column 1350, which is located in the middle of the cross, and in blue is column 5000, which is located far from the cross.

$$\chi_{Voigt}(\omega) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}\frac{\exp(\frac{-\omega^2}{2\sigma^2})\Delta\epsilon {\omega_0}^2}{{\omega_0}^2-\omega^2+j\omega\gamma}d\omega = \chi_{Lorentz}(\omega) * \frac{1}{\sqrt{2\pi}\sigma}\exp(\frac{-\omega^2}{2\sigma^2}) \qquad (1.47)$$

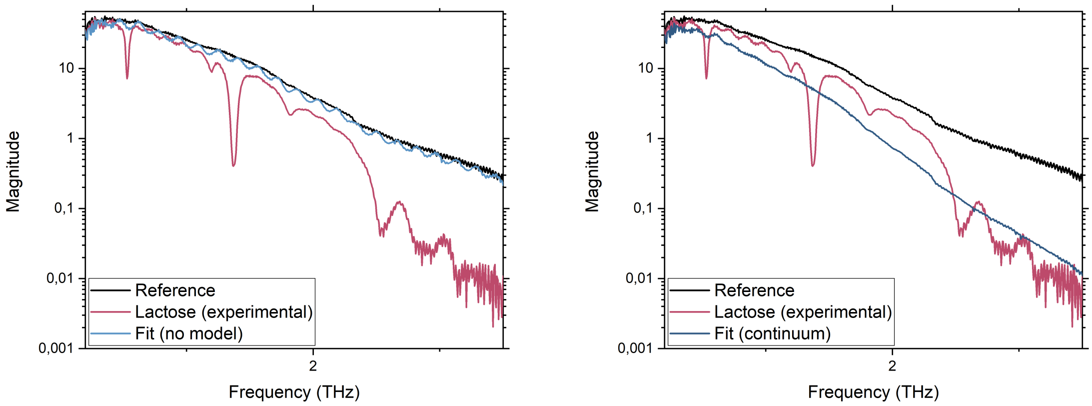
where $\Delta\epsilon$ is the strength of the oscillator, ω_0 is the centre frequency of the Lorentz oscillator, γ is the width of the Lorentz oscillator and σ is the width of the Gaussian broadening.

We fitted our data peak by peak, from the ones that absorb the most to the ones that absorb less. As explained previously, we used the Akaike criterion to determine which model to choose at every step.

The first step of the fitting process is to run a fit without any model in order to have a base criterion. If the fit was perfect, the Akaike criterion would be equal to the number of points (in our case **6.008e+03**). When no model is entered (only a constant value for the permittivity and a value for the thickness of the pellet), the fit is far from perfect and the value of the criterion is **6.096e+08**. The corresponding fitted spectrum is shown on figure 1.23a. Now, our goal is to reduce this criterion and, more importantly, to find the model that reduces it the most. We found that there was a need to fit the broad high frequency losses first. Thus, we decided to fit it phenomenologically with a continuum model (see equation 1.48), which reduced the Akaike criterion to **2.522e+08**, the result is shown in figure 1.23b. The continuum model compensates for all absorption losses since no peaks have been added yet, but its parameters will adjust as the fitting process progresses.

$$\chi(\nu) = \chi_{Lorentz}(\nu) * \left[1-\left(1-\frac{\nu h}{kT}-\frac{(\nu h)^2}{2(kT)^2}\right).\exp(-\frac{\nu h}{kT})\right]H(\nu-\nu_0,1) \qquad (1.48)$$

Where ν is the frequency, h is Planck constant, k is the Boltzmann constant, T is the temperature, and H is the Heaviside function, meaning that $H(\nu-\nu_0,1)=1$ when $\nu \geqslant \nu 0$.



(A) Fitted spectrum when no dispersion model is entered (the only parameter is $\epsilon_\infty = 3.890$) in blue. The reference spectrum is shown in black and the lactose spectrum is shown in pink.

(B) Fitted spectrum when a continuum model is entered (the optimal parameters are $\epsilon_\infty = 3.130$, $\Delta\epsilon = 0.912$, $\omega_0 = 0.951$THz, $\gamma = 5.202$THz, T = 7.4e-5K) in blue. The reference spectrum is shown in black and the lactose spectrum is shown in pink.

FIGURE 1.23: Results of the fit for the first two steps of the process.

Now, we can focus on the two main absorption peaks of lactose. In table 1.1 we showcase the results for the peaks that absorb the most, which are located at 0.53 and 1.37 THz. In all cases, the criterion has been reduced significantly, which proves the need to fit these two absorption peaks. The two bottom configurations are the ones that possess the lowest Akaike results, which means that they are the best suited models. However, we ended up choosing the Lorentz model for the second peak because of the fact that the width of the Gaussian in the Voigt model is really small (0.1 GHz) thus negligible considering the frequency resolution (5 GHz).

Model	ϵ_∞	$\Delta\epsilon$	ω_0 (THz)	γ (GHz)	σ (GHz)	Akaike criterion
1st peak: Lorentz	2.614	0.050	0.530	25.20		6.303e+06
2nd peak: Lorentz		0.029	1.370	47.52		
1st peak: Lorentz	2.616	0.050	0.530	25.20		6.304e+06
2nd peak: Voigt		0.029	1.370	47.50	9.246	
1st peak: Voigt	2.611	0.049	0.530	23.32	3.820	6.211e+06
2nd peak: Lorentz		0.029	1.370	47.36		
1st peak: Voigt	2.611	0.049	0.530	23.32	3.820	6.211e+06
2nd peak: Voigt		0.029	1.370	47.36	0.100	

TABLE 1.1: The optimized parameters for the possible models for the first two absorption peaks of lactose.

The parameters of the continuum are now: $\Delta\epsilon = 0.588$, $\omega_0 = 3.389$ THz, $\gamma = 2.207$ THz, $T = 1.0$ K (this temperature has no physical meaning, it is a purely phenomenological model). The next steps are to add absorption peaks one by one until the criterion does not decrease any more.

1.6.4 The final fit

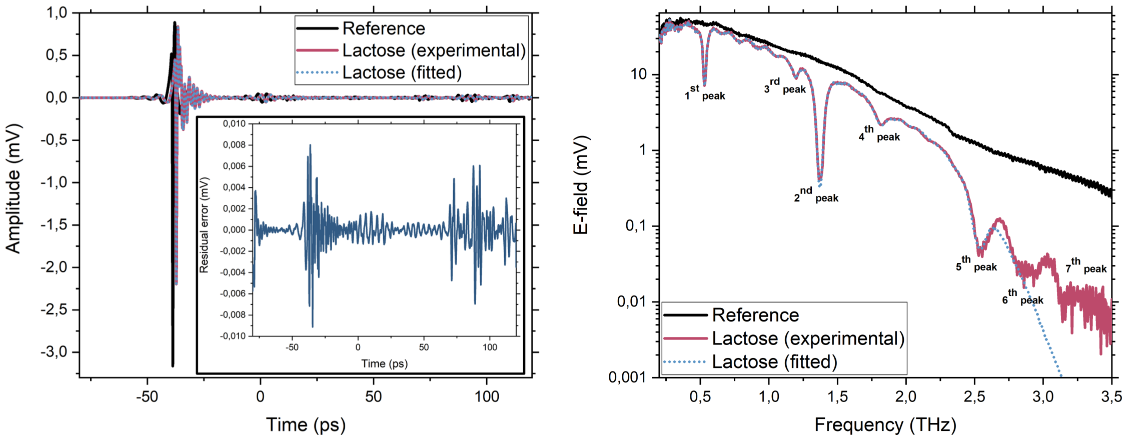
Feature	Model	$\Delta\epsilon$	ω_0 (THz)	γ (GHz)	σ (GHz)	T (K)
1st peak	Voigt	0.046	0.531	23.17	4.280	
3rd peak	Voigt	3.822e-03	1.195	43.88	0.837	
2nd peak	Doublet (Voigt)	3.839e-03	1.264	316.0	3.670	
		0.029	1.370	46.56	10.47	
4th peak	Doublet (Voigt)	9.764e-06	1.767	1.138e-03	695.5	
	Voigt	5.379e-03	1.818	104.97	6.362	
5th peak	Doublet (Voigt)	5.962e-03	2.527	10.68	49.28	
		0.067	2.594	51.05e-07	1.417e+03	
6th peak	Voigt	0.085	2.901	0.274	889.2	
7th peak	Voigt	0.119	3.409	44.04	475.7	
High frequency absorption	Continuum	0.184	0.258	5.460		0.1

TABLE 1.2: The optimized model parameters of lactose. Other retrieved information is: ϵ_∞ = **2.739** ; thickness of the pellet is **599.5** μ**m** (coherent with the measured one) ; Average delay is −17.5 fs ; Coefficients of the noise proportional to the signal and its derivatives are a = 3.393e-05 and b = 6.944e-04 ; Final Akaike criterion is **2.028e+06**.

To continue taking the experimental noise into account, it is important to keep in mind that the delay noise and the noise proportional to the signal and its second derivative still have to be taken into account. We only fixed them for the two consecutive measurements of the reference, and then of the sample. What we correct is the average value of the drift between two measurements, which means that if we look at two different measurements there is no reason why this value should be the same. As a result, we have to incorporate it in our fit as well. Furthermore, it is important to follow the Akaike criterion, as if we add these parameters too soon in the fitting process, they tend to compensate other losses.

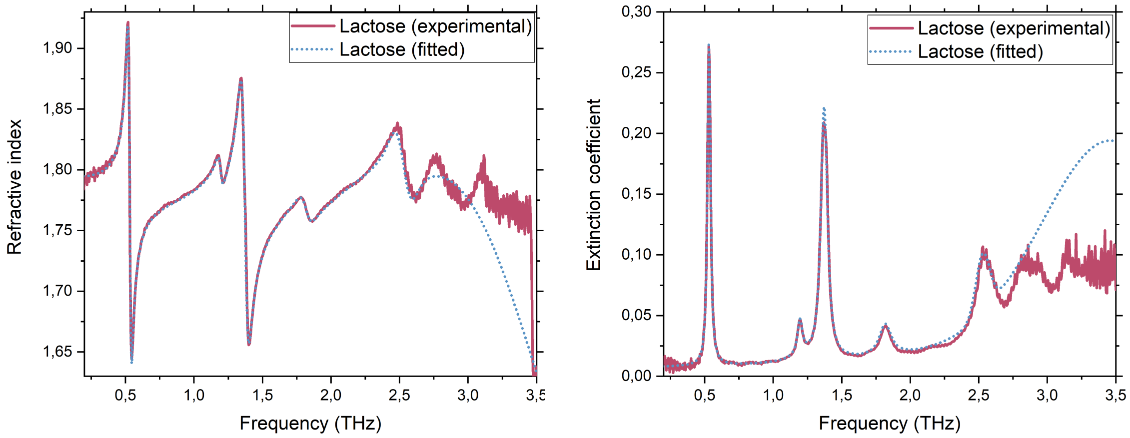
After having fitted the two main absorption peaks as well as the high frequency absorption, according to Akaike criterion, the noise terms can be added to the fitting process. Finally, due to theoretical considerations [52], we decided to replace the Lorentz model for the 1.37 THz absorption peak by a doublet (which was observed in DFT simulations), as well as the 1.8 THz peak and fitted them. The final model parameters are given in table 1.2. The final Akaike criterion achieved is **2.028e+06**, which is strongly lower than the base criterion. However, we have not yet reached a "perfect" fit, and it could certainly be improved by a better modelling of the noise and a more accurate model for the high frequency losses.

The fitted spectrum, as well as the real and imaginary parts of the refractive index are shown in figure 1.24. Above 2.6 THz, the fitted curve strays from the experimental data, which is expected since they become more noisy. Moreover, the absorption peaks are wider, which can mean that the Lorentz and Voigt model do not describe the dynamics correctly



(A) Reference (in black), experimental lactose sample (in pink), and fitted lactose sample (in light blue dots) time traces. In the inset is the residual error of the fit.

(B) Reference (in black), experimental lactose sample (in pink), and fitted lactose sample (in blue dots) spectra. The number of the absorption peaks correspond to the parameters given in table 1.2.



(C) Refractive index retrieved from experimental data and from the fitted data.

(D) Extinction coefficient retrieved from experimental data and from the fitted data.

FIGURE 1.24: Results of the fitting process, illustrating the parameters retrieved in table 1.2

any more and another model may be more appropriate. With more input from simulations we would be able to adjust this part of the fit.

1.7 Conclusion

In this chapter, I have explained the concept of information in THz-TDS and how to recover it. The study shows that time-domain fitting, in addition to being closer to the experiment, gives also more reliable results. Moreover, taking experimental noise and artifacts into account during the fitting process is of utmost importance. Hence, in order to analyze data rigorously in THz-TDS, it is important to have an extensive knowledge of the experimental noise and artifacts arising from the setup. By taking into account this experimental noise in the pre-processing of the data as well as during the fitting process, we ensure that the retrieved permittivity models describe accurately the analyzed sample. The application of this method on the α-lactose monohydrate pellet gave insight on the molecular dynamics behind the absorption peaks and will be even more insightful when compared to simulation results. This shows that THz-TDS is a great tool to analyze biological samples in the form of pressed powder and we should now focus on samples in conditions closer to *in vivo* conditions for most biological samples: liquids.

Chapter 2

Studying liquids with THz-TDS

Now that we are able to properly recover the information in a THz-TDS experiment, we can focus on applying it to more complex biological samples. In order to study these biological samples in conditions that are the closest to *in vivo* conditions, they have to be preserved in an aqueous solution composed, mostly, of water and sugars. Indeed, freezing or lyophilizing the samples can compromise their structural integrity and thus reveal features that aren't present in their natural state, or hide others. Therefore, this chapter is focused on the study of these solutions the biological samples are in, beginning with liquid water and then investigating the changes occurring in its structure when other components are added to it.

2.1 Liquid water, the most basic liquid sample?

2.1.1 The unique structure of liquid water

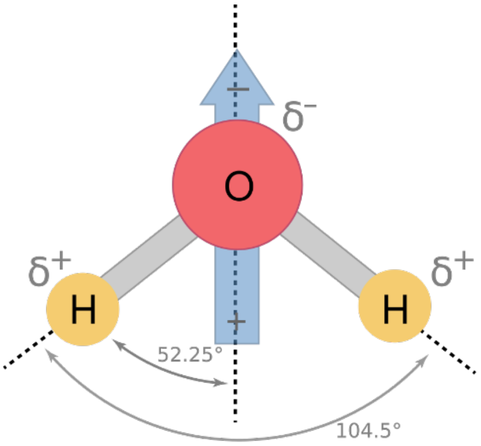


FIGURE 2.1: Water molecule and its dipolar moment. In pink is the oxygen atom and in yellow the two hydrogen atoms.

Water (H_2O) consists of two light hydrogen atoms attached to a single heavier oxygen atom (see figure 2.1). To form the molecule, the single electrons donated by the two hydrogen atoms and the eight electrons donated by the oxygen atom pair up into five orbitals, one pair closely associated with the oxygen atom, two pairs forming each of the two identical covalent bonds between the oxygen and hydrogen atoms, and two pairs associated with the oxygen atom spread out away from these covalent bonds. Each molecule is electrically neutral but polar, with the center of positive and negative charges located in different places.

Liquid water, despite a seemingly simple formula, has a very unique structure. Hydrogen bonding between neighbouring water molecules creates a complex network, and the high density of molecules present due to their small size is responsible for the liquid nature of water at ambient temperatures. In the liquid state, in spite of 80% of the electrons in H_2O being concerned with bonding, the three atoms do not stay together. The hydrogen atoms are constantly exchanging between water molecules, due to protonation and deprotonation processes (see equation 2.1) [106].

$$2H_2O(aq) \rightleftharpoons H_3O^+(aq) + OH^-(aq) \tag{2.1}$$

In liquid water, all water molecules have at least one hydrogen bond to the neighboring water molecules, which means that there are no free water molecules under ambient conditions. Water consists primarily of a mixture of clusters of water molecules with different degrees of hydrogen bonding existing in rapid and complex equilibria [107]. It contains by far the densest hydrogen bonding of any solvent with almost as many hydrogen bonds as there are covalent bonds. These hydrogen bonds can rapidly rearrange in response to changing conditions and environments. The hydrogen bonding patterns are apparently random in water, thus for any water molecule chosen at random, there is equal probability that the four hydrogen bonds are located at any of the four sites around the oxygen.

Finally, another notable physical property of water is the opposite influences of hot and cold water. In particular, several properties of water change at about 4 °C [106] and also at about 50 °C [108]. As cold liquid water is heated, individual molecules shrink, hence bulk water shrinks and becomes less easy to compress, and its refractive index increases. In contrast, as hot liquid water is heated, it expands, it becomes easier to compress, and its refractive index decreases.

2.1.2 Studying liquid water in the terahertz range

As indicated in the introduction (see 0.4.2), water is probably the most studied liquid. Different methods have been used to study the properties of water but its dynamical behavior is still only partially understood. To retrieve information on molecular processes in water, and in particular intermolecular interactions, the study has to be conducted in the terahertz range, which covers the hydrogen-bond (H-bond) response [109]. A variety of setups have been used in order to gain more knowledge, each one covering a different part of the terahertz range:

0.006 − 1.12 THz: Vector network analyzer based dielectric spectrometers [91, 110]

0.05 − 4.0 THz: THz-TDS in transmission mode [83, 93, 111–113]

0.06 − 1.5 THz: fs THz transmission spectrometer [84, 114]

0.1 − 2.0 THz: THz-TDS in reflection mode [115]

0.2 − 3.5 THz: THz time-domain attenuated total reflection spectroscopy [91, 111, 116, 117]

1.5 − 6.6 THz: Dispersive Fourier transform spectrometer [85]

3.0 – 12.0 THz: Far-infrared Fourier transform attenuated total reflection [91]

However, the experimental uncertainty of the reported data is often not estimated explicitly and direct comparisons between different data sets is difficult because few authors monitor the temperature during the experiment. Still, as a broadband technique with high resolution, THz-TDS is a perfect candidate for the investigation. However, despite the many studies of liquid water, there is still no consensus on the correct permittivity model to use to fit the data. Several models have been proposed, starting with the well-known Debye model (see equation 2.2) that works for frequencies below a few tens of gigahertz [118]:

$$\tilde{\epsilon}(\omega) = \epsilon_\infty + \frac{\Delta\epsilon}{1 + j\omega\tau} \tag{2.2}$$

Where $\Delta\epsilon = \frac{n\mu^2}{3kT}$ and $\tau = \frac{\Gamma}{2kT}$.

However, the Debye model works in the hypothesis that there are no interactions between dipoles and that there is no inertia. Therefore, it does not depict accurately the collective vibrations that THz-TDS scans and other models have been proposed. Most of the time, these other models add an extension to the Debye model when they extend to higher frequencies. Some add another Debye term [82–84, 93, 112–115] and others even add one [91, 111, 117] or two Lorentz terms [116, 119]:

$$\tilde{\epsilon}(\omega) = \epsilon_\infty + \sum_{i=1}^{2} \frac{\Delta\epsilon_{Di}}{1 + j\omega\tau_{Di}} + \sum_{i=1}^{2} \frac{\Delta\epsilon_{Li}\omega_{0,i}^2}{\omega_{0,i}^2 - \omega^2 + j\omega\gamma_i} \tag{2.3}$$

In some articles, we can even find a third Debye term [90, 110, 120]. Other versions have also been proposed, such as replacing the Debye terms by Havriliak-Negami derived models:

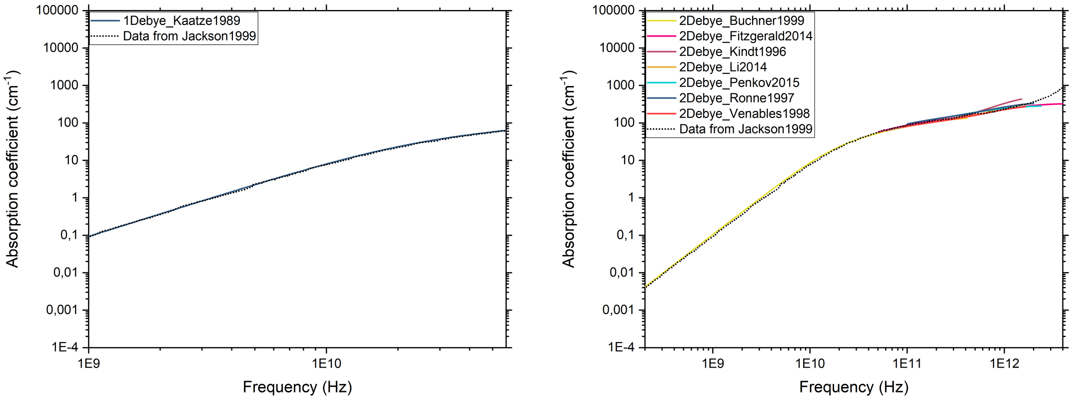
$$\tilde{\epsilon}(\omega) = \epsilon_\infty + \frac{\Delta\epsilon}{\left[1 + (j\omega\tau)^\alpha\right]^\beta} \tag{2.4}$$

Table 2.1 contains the model parameters for all the published papers aforementioned. It is clear that, from an experiment to another, the parameters values change, in part due to the temperature changes and because of the problems in the fitting process explained in the previous chapter. Still, in figure, **??**, we can see that all of these model fit quite well the frequency domain they scan. However, researchers point out that there is a struggle to fit the high frequency part of the dielectric permittivity of liquid water. Moreover, if the proposed models are accurate for a small part of the frequency domain, the fact that the parameters values change when the scanned frequencies shift indicates that the models do not depict the whole range correctly. Indeed, when we look at the overall picture, such as shown on figure 2.3, none of these models seem to correspond to the way the experimental absorption coefficient decreases from 20 THz.

As we made progress on the fitting protocol, we will apply our knowledge to the study of water in the terahertz range with THz-TDS. We will begin by comparing the proposed models thanks to their Akaike criterion and we will look at other models that could be better adapted to the big picture fit. We will also scan different temperatures as the dielectric spectrum of liquid water changes when the temperature is modified, thus the correct model should take this temperature dependence into account.

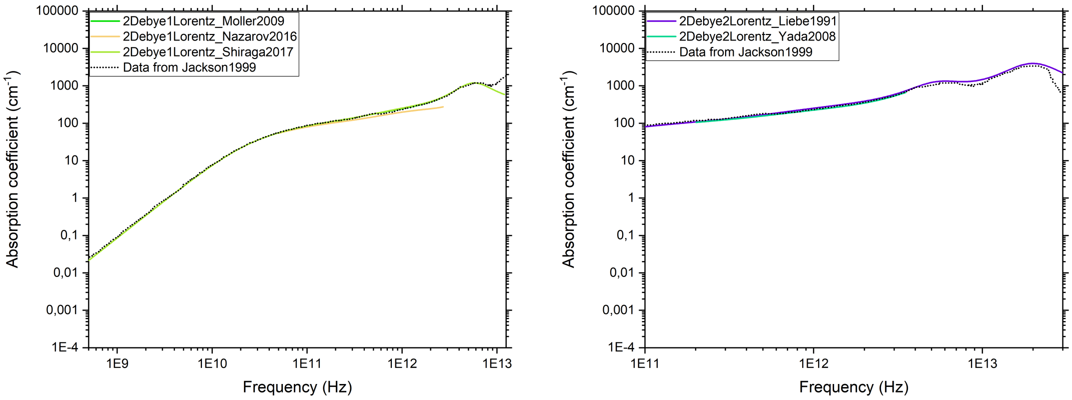
Frequency (THz)	Temperature (°C)	ϵ_∞	$\Delta\epsilon_{D1}$	τ_1 (ps)	$\Delta\epsilon_{D2}$	τ_2 (ps)	$\Delta\epsilon_{D3}$	τ_3 (ps)	$\Delta\epsilon_{L1}$	$\omega_{0,1}$ (THz)	γ_1 (THz)	$\Delta\epsilon_{L2}$	$\omega_{0,2}$ (THz)	γ_2 (THz)	Ref.
0.001 - 0.057	25.0	5.2	78.36	8.27											[118]
0.0002 - 0.410	20.0	4.42	73.78	9.6	2.11	1.2									[82]
0.05 - 4	22.0	3.2	73.6	8.0			1.6	0.18							[93]
0.06 - 1.5	N.A.	3.48	78.36	8.24			4.93	0.18							[84]
0.2 - 2.0	20.0	3.35	72.72	8.17			2.29	0.169							[83]
0.18 - 2.4	25.0	2.5	78.4	8.25			1.4	0.31							[113]
0.1-2.0	30.0	3.4	76.6	7.0			2.0	0.2							[115]
0.09 - 1.65	18.0-20.0	3.392	72.93	8.794			1.978	0.212							[114]
0.2 - 2.5	25.0	2.68	72.3	8.34			2.12	0.36	1.13	5.01	7.06				[117]
0.05 - 2.7	22.0	2.5	75	9.5			1.47	0.23	1.14	5.3	5.35				[111]
0.0005 - 12	27.0	2.38	72.04	7.93			2.01	0.273	1.25	5.28	5.35				[91]
0.1 - 30	20.0	3.34	74.70	9.39			2.03	0.236	0.96	5.11	4.46	0.85	18.2	15.4	[119]
0.2 - 3.5	20.0	2.0	74.9	9.47			1.67	0.248	1.12	5.3	5.35	0.50	14.7	8.08	[116]
0.0059 - 1.12	25.0	3.2	72.33	8.37	1.14	1.05	1.71	0.178							[110]
0 - 1.0	20.0	1.7	75.15	9.53	1.37	2.3	2.0	0.125							[90]
1.5 - 6.6	25.0	2.2	71.49	8.31 $\alpha = 1.0$ $\beta = 1.0$	2.8	1.0 $\alpha = 1.0$ $\beta = 0.77$	1.6	0.1 $\alpha = 0.9$ $\beta = 0.8$	0.92	5.26	0.025				[85]

TABLE 2.1: Parameters for the modeling of the dielectric permittivity of water.



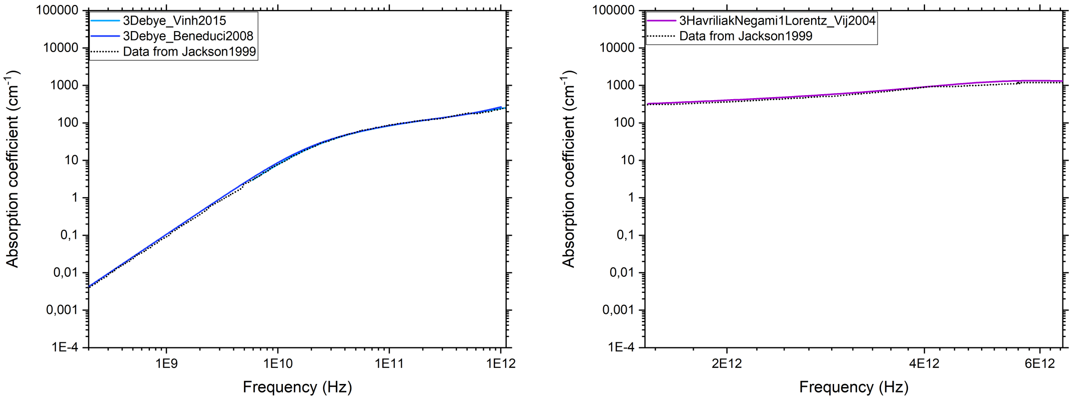
(A) 1 Debye model fit result on a 0.001 - 0.057 THz range

(B) 2 Debye model fit result on a 0.0002 - 4.0 THz range



(C) 2 Debye and 1 Lorentz model fit result on a 0.0005 - 12.0 THz range

(D) 2 Debye and 2 Lorentz model fit result on a 0.1 - 30.0 THz range



(E) 3 Debye model fit result on a 0.0002 - 1.12 THz range

(F) 3 Havriliak-Negami and 1 Lorentz model fit result on a 1.5 - 6.6 THz range

FIGURE 2.2: Fit results for each model, on the frequency range studied, compared to a set of data from the literature [121]

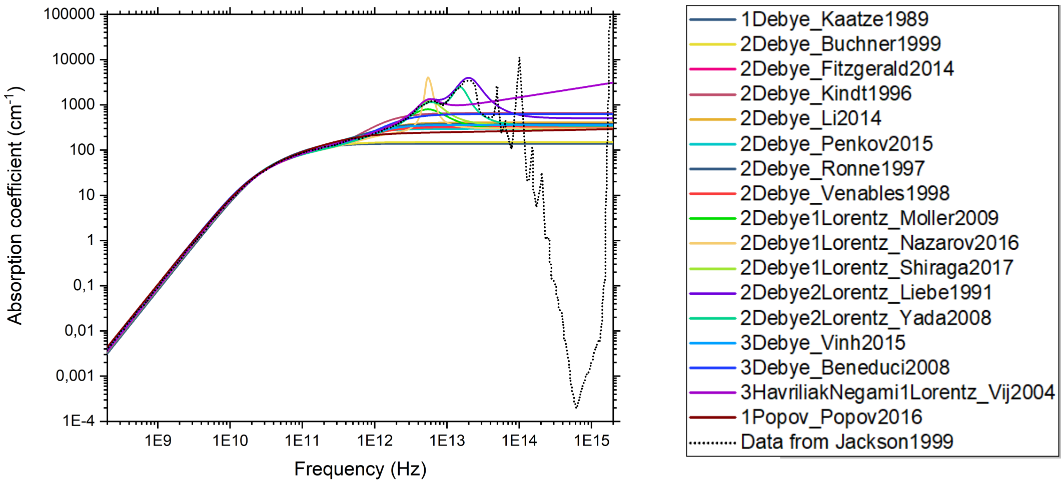
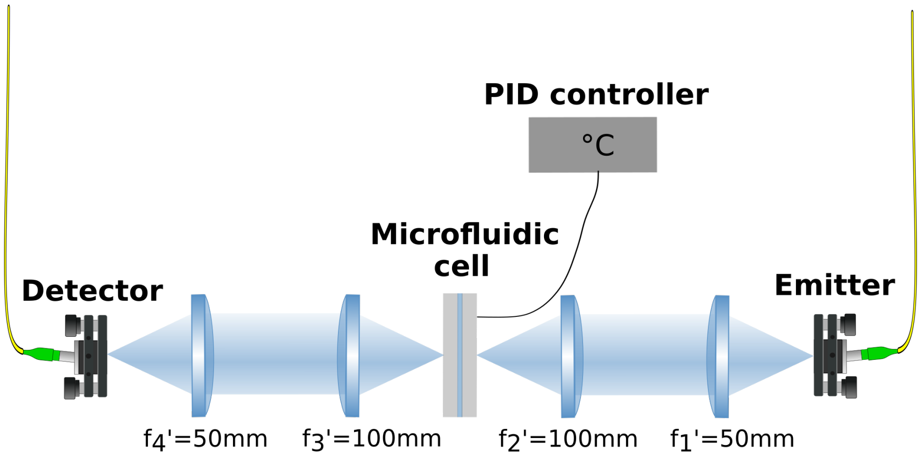
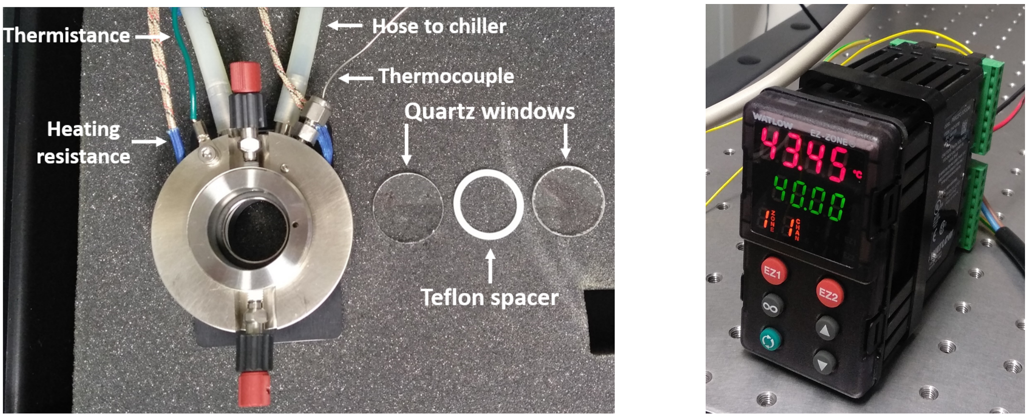


FIGURE 2.3: Absorption coefficients of all the permittivity models proposed in the literature, over a wide frequency range (0.2 GHz-2000 THz), compared to data from [121].

2.1.3 Experimental protocol



(A) Schematic of the experiment, the THz beam travels from the emitter to the detector, going through the TPX lenses and the microfluidic cell. The microfluidicell is composed of two z-cut quartz windows surrounding the liquid water and separated by Teflon spacers. It is linked to a PID controller monitoring the temperature.



(B) Picture of the microfluidic cell unmounted. The quartz windows and spacer are shown beside the cell.

(C) Picture of the PID controller. In red is the current measured temperature and in green is the temperature setpoint to reach.

FIGURE 2.4: Experimental setup for the measurement of liquid water with THz-TDS.

The THz-TDS setup is the same as for the the study of lactose: a Menlo System Terasmart in a Plexiglas box, purged with nitrogen. The setup between the emitter and the detector is shown on figure 2.4a. The terahertz beam is collimated by a TPX lens with a 50 mm focal length and focused on the sample with a TPX lens with a 100 mm focal length. We need the focusing length as the diameter of the cell is slightly smaller than the diameter of the pulse. We made sure that the depth of field was longer than the thickness of the cell. Then, the beam goes through a symmetric system to the detector. The microfluidic cell, shown on figure 2.4b, is composed of two z-cut quartz windows, with a thickness of 2 mm each, separated by a Teflon spacer with a thickness of 100 µm. It is heated by heating resistances on the sides and chilled with liquid deionized (DI) water going through a chiller. The temperature response is ensured by a thermocouple via a PID controller that monitors it thanks to a thermistance. The PID controller is shown on figure 2.4c, where the temperature setpoint is in green and the current temperature in red. Finally, the microfluidic cell is filled with liquid water thanks to syringes connected at the top and bottom on the cell. It is important to avoid having air inside the syringes in order to avoid air bubbles in the cell. The measurements are all done inside the Plexiglas box, purged with nitrogen (less than 2% humidity) and the temperature is increased by 10 °C steps. In order to have a stable temperature during the experiment (less than 0.1 °C variations), we waited 30 min after the setpoint had been entered to measure the time trace. The experimental protocol during the measurements is explained in figure 2.5. As explained in the previous chapter, we take two spectra each time in order to use the noise matrices during data analysis.

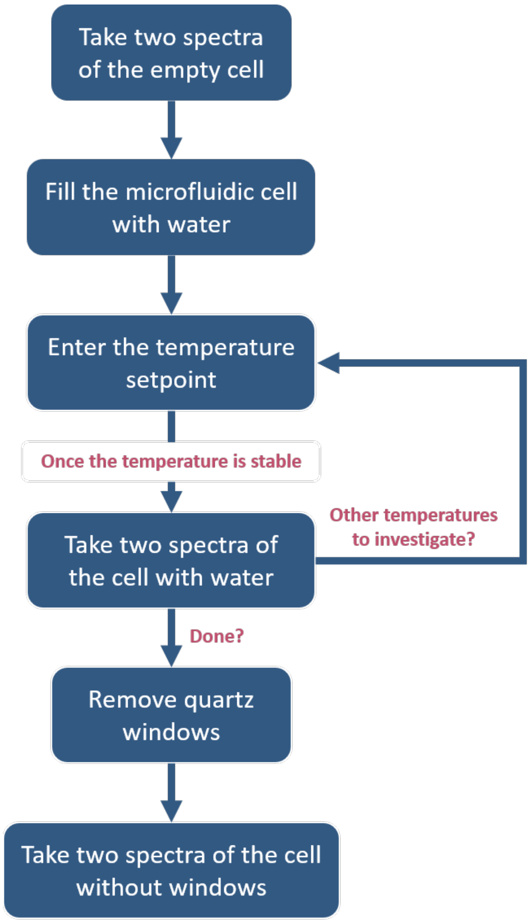


FIGURE 2.5: Experimental process for the study of liquid water with THz-TDS.

2.1.4 Experimental results

Figure 2.6 showcases the time traces obtained following our experimental protocol. The reference time trace, in black, corresponds to a measurement of the THz pulse with an empty mount (*i.e.* the windows of the cell have been removed). Then, the empty cell time trace corresponds to a measurement of the empty microfluidic cell (*i.e.* with air between the quartz windows) and finally the other time traces correspond to the measurements of the microfluidic cell filled with water at various temperatures. For each measurement where the beam goes through the cell, Fabry Perot echoes appear because of the reflections inside the quartz windows and inside the gap between them. The visible echoes separated by 30 ps correspond to the reflections inside the windows while the reflections inside the gap are hidden by the main pulse, as the gap is really small thus the delay between them is really small as well (around 1.4 ps for air in the gap, and 2.7 ps if there is water). On the right of figure 2.6, I zoomed in on the main pulse in order to see the differences between the liquid water time traces. As the temperature increases, the amplitude of the pulse decreases, which indicates stronger absorption by the liquid water. Also, the pulse shifts to the right, meaning that the refractive index is higher for higher temperatures.

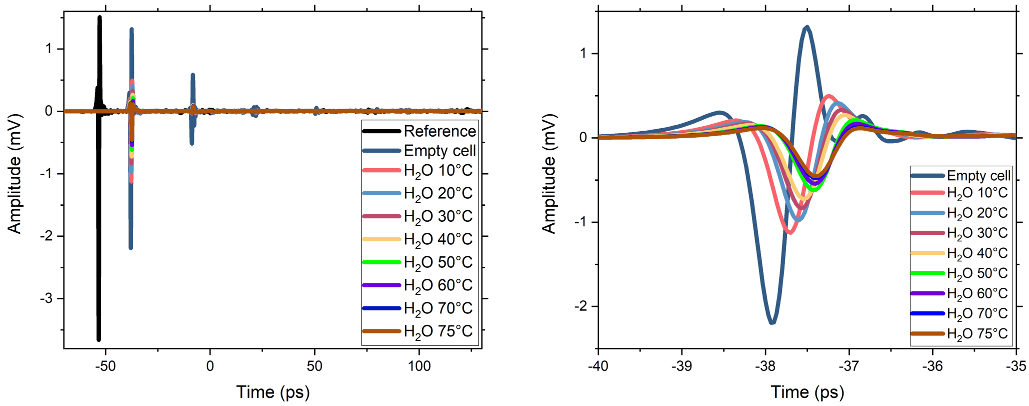


FIGURE 2.6: Time traces of the measurements of liquid water, of the reference, and the empty microfluidic cell. On the right is provided a zoom of the main pulses of the liquid water measurements.

After a Fourier transform on all these time traces, we obtain the spectra on the left of figure 2.7. On the right, I divided each spectrum by the spectrum of the reference in order to normalize them. Once again, the amplitude decreases as the temperature increases, which indicates stronger absorption at high temperatures. Also, the amplitude difference between two consecutive measurements decreases as the temperature rises, which may indicate that the H-bond network structure is less subject to change at temperatures above 50 °C. The deep at around 2.9 THz for 10 °C that shifts with the temperature may be due to Fabry Perot echoes that are not taken into account. Therefore, it creates aliasing and leads to the appearance of oscillations in the spectrum.

A fit of these data is now necessary to understand the underlying microscopic processes better.

2.1.5 Fitting process

Before comparing the measurements at different temperatures, we have to decide on a permittivity model to describe the dielectric response of liquid water. To do that, the data at 20 °C is fitted with all the models proposed in the literature and their retrieved Akaike criteria are compared. Then, we will look at other models proposed by theoreticians that should

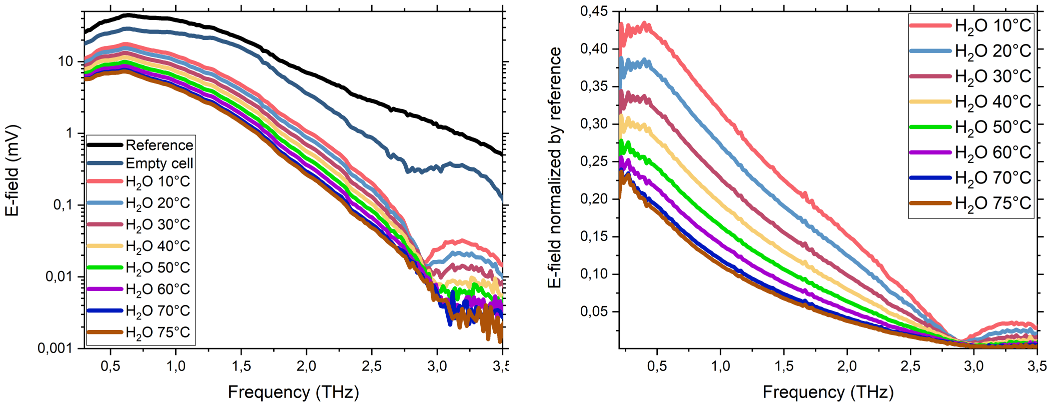
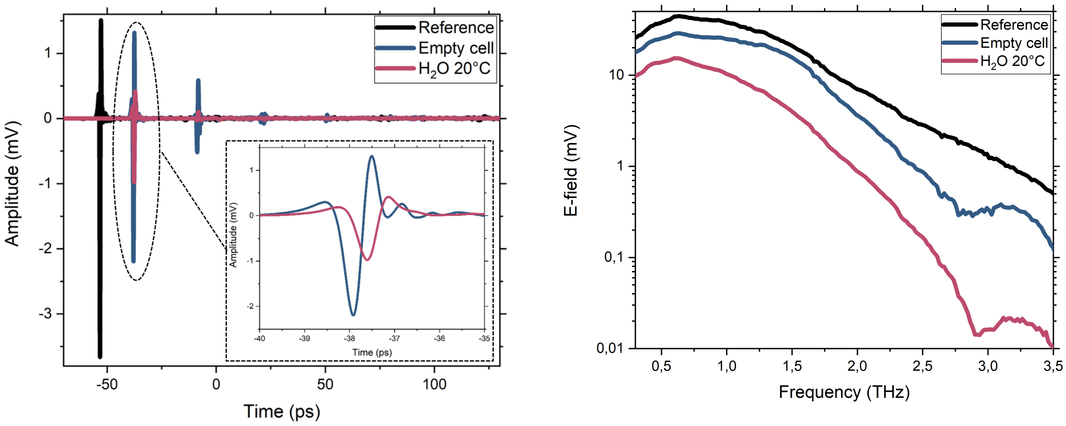


FIGURE 2.7: Spectra of the liquid water measurements, of the reference and empty microfluidic cell retrieved via a Fourier transform. On the right, the spectra are normalized by the spectrum of the reference.

take into account more of the collective processes in water. Finally, we will fit all data in order to look at the variations of the parameters according to the temperature.

Comparing models at room temperature

The time trace and spectrum of liquid water at 20 °C are displayed on figure 2.8. The parameters and Akaike criteria found when we fit this data with the permittivity models mentioned in section 2.1.2 are gathered in table 2.2.



(A)

(B)

FIGURE 2.8: (A) Time traces of the measurement of liquid water at 20 °C, with a zoom on the main pulse of the liquid water trace. (B) Spectra retrieved from the time-traces.

On the frequency range scanned (from 0.2 THz to 3.5 THz), the best model is the 2 Debye and 1 Lorentz model according to the Akaike criterion. Indeed, it is first reduced when we add an additional Debye term to the original Debye model. Adding a third Debye term has

Model	ϵ_∞	$\Delta\epsilon_{D1}$	τ_1 (ps)	$\Delta\epsilon_{D2}$	τ_2 (ps)	$\Delta\epsilon_{D3}$	τ_3 (ps)	$\Delta\epsilon_{L1}$	$\omega_{0,1}$ (THz)	γ_1 (THz)	$\Delta\epsilon_{L2}$	$\omega_{0,2}$ (THz)	γ_2 (THz)	Akaike criterion
1 Debye	3.666	50.21	3.26											1.668e+07
2 Debye	2.792	113.6	11.07			0.97	0.058							6.505e+06
3 Debye	2.792	85.39	11.04	27.99	11.09	0.97	0.058							6.505e+06
2 Debye 1 Lorentz	3.123	106.74	10.31			0.00	0.73	0.61	3.03	4.48				6.429e+06
2 HN	2.792	113.34	11.04			0.97	0.058							6.505e+06
			α = 1.0				α = 1.0							6.505e+06
			β = 1.0				β = 1.0							6.505e+06
3 HN	2.792	113.34	11.04	0.00	18.64	0.97	0.058							6.505e+06
			α = 1.0		α = 1.0		α = 1.0							6.505e+06
			β = 1.0		β = 1.0		β = 1.0							6.505e+06

TABLE 2.2: Parameters of all the tried models, found after fitting the time traces from the liquid water experiment.

no effect, since the fit finds a term that has a time constant very close to the low frequency Debye term, which indicates that there is no other relaxation process occurring in this frequency range. Changing from a Debye model to Havriliak-Negami derived models has no effect either, as the fit results in α and β parameters equal to 1. However, when we add a Lorentz term, the fit founds one with a central frequency close to the one found in the literature. But, adding this Lorentz term removes one Debye term, resulting in a final model composed of 1 Debye and 1 Lorentz term. This is not very satisfactory considering that most articles mention a second Debye term for the model but, as we said earlier, the parameters and the models change according to the frequency range scanned. Therefore, we decided to focus on the big picture.

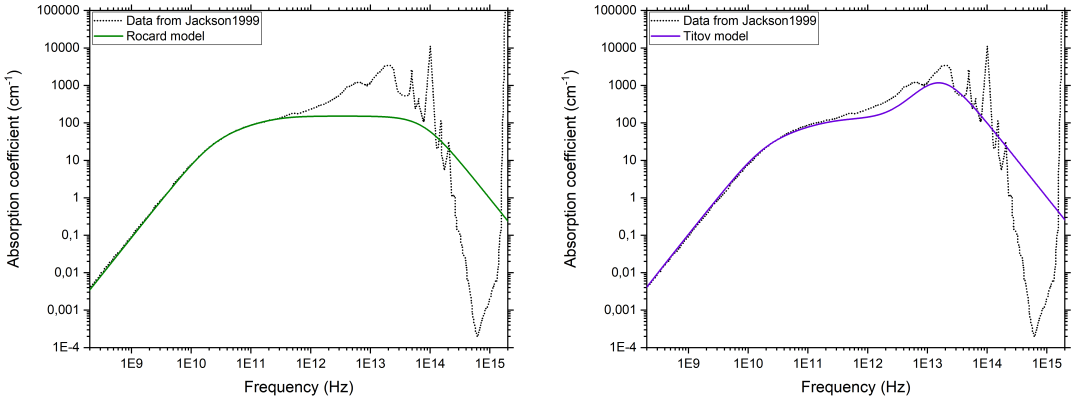
The problem of most of the models found in the literature is that they are phenomenological, and they are based on the Debye model, which does not take inertia or dipole-dipole interactions into account. We found another model, very close to the Debye model, but taking into account inertial effects, which is the Rocard-Powles-Debye model [122, 123]:

$$\tilde{\epsilon}(\omega) = \epsilon_\infty + \sum_{k=1}^{N} \frac{a_k}{(1 - j\omega\tau_k)(1 - j\omega\tau_{fk})} \tag{2.5}$$

Where N denotes the total number of Debye relaxation times τ_k, each with an associated friction time τ_{fk}. ϵ_∞ is still the high frequency limit of the relative dielectric permittivity and $\sum_{k=1}^{N} a_k = \epsilon_s - \epsilon_\infty$, where ϵ_s is the static permittivity.

For this model, as for the Debye model, they considered a single molecule and examined how the orientation of the axis of its dipole changed on the average, when the external field had been switched off. Then, they studied the rotational Brownian motion of a sphere. Hence, they added inertial effects but no dipole-dipole interactions. This results in a model similar to the Debye model, but with an additional time constant due to the frictional force. An example absorption coefficient obtained from this model is shown on figure 2.9a. The parameters are chosen by hand to resemble the data curve, but no fitting process has been conducted yet. This is encouraging as the curve decreases from a certain point, which is closer than the Debye model to the overall behavior of liquid water permittivity.

We investigated another model taking inertial effects into account, proposed by [124].



(A) Absorption coefficient retrieved from modeling the dielectric permittivity with a Rocard-Powles-Debye model.

(B) Absorption coefficient retrieved from modeling the dielectric permittivity with a Titov model.

FIGURE 2.9: Absorption coefficients retrieved after taking into account inertial effects in the model of the dielectric permittivity of liquid water.

This model, that we will call "Titov model" in the rest of the study, adds inertial effects to Budó's generalization of the Debye rotational diffusion model of dielectric relaxation of polar molecules to an assembly with internal interacting polar groups [125]. Indeed, Budó's model already extended the Debye model to molecules consisting of two similar polar groups, which cannot rotate freely relatively to one another owing to their mutual potential energy. Thus, they introduced some interaction between molecules. The Titov model adds inertial effects to this extension, creating a more complete model for polar liquids. The susceptibility is expressed with:

$$\epsilon(\omega) = \epsilon_\infty + \frac{\chi}{1 + j\eta\omega\tau_D}\left[\frac{2\beta_2^2}{2\beta_2^2 + j\eta\omega\tau_D} + \frac{j\eta\omega\tau_D}{\sigma_V - 2(\eta\omega/\gamma_1)^2 + 2j\beta_1\eta\omega/\gamma_1}\right] \tag{2.6}$$

With η a characteristic time constant, τ_D the Debye relaxation time of the second (heavy) polar group (from the two that interact), β_1 and β_2 normalized friction coefficients (they call β_2^{-2} the inertial parameter), σ_V the interaction parameter, and $\gamma_1 = \frac{\eta}{\eta_1}$ a ratio of characteristic times.

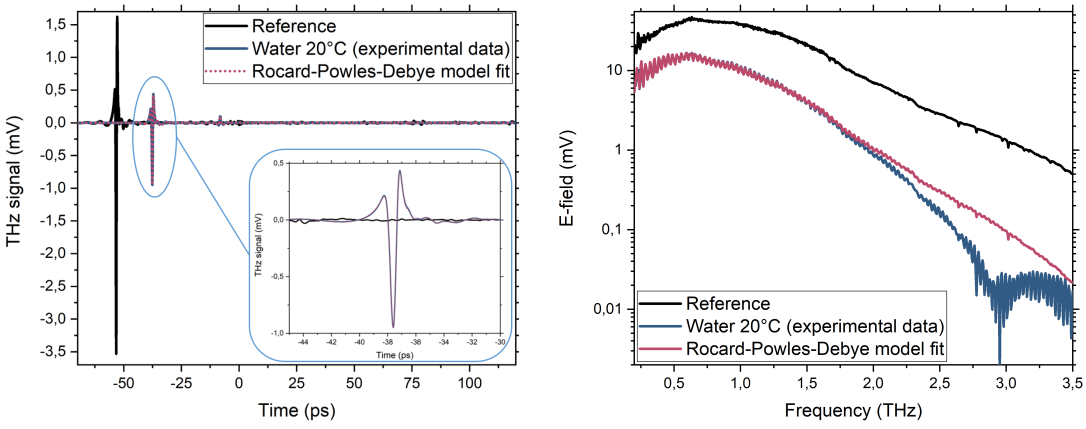
The absorption coefficient retrieved from modeling the dielectric permittivity with the Titov model is displayed on figure 2.9b. As for Rocard-Powles-Debye model, the parameters were chosen by hand and no fit has been conducted yet. Still, we observe that this model is even closer to the absorption coefficient retrieved from experimental data, which shows the importance of taking all effects, both inertial and interactions, in the THz range.

Now, we fitted our data with both of these models. First, the Rocard-Powles-Debye model fit results are shown in figure 2.10. The optimized parameters found for the Rocard-Powles-Debye model are:

- $\epsilon_\infty = 3.097$
- $a_1 = 23.63$

- τ_1 = 0.75ps
- τ_{f1} = 1.24ps
- a_2 = 3.65
- τ_2 = 0.24ps
- τ_{f2} = 0.24fs

The final Akaike criterion for this model is 5.228e+06, which is lower than for the Debye and Lorentz model. Therefore, taking inertial effects into account not only is important for the big picture fit, but also improves the fit at terahertz frequencies.



(A)

(B)

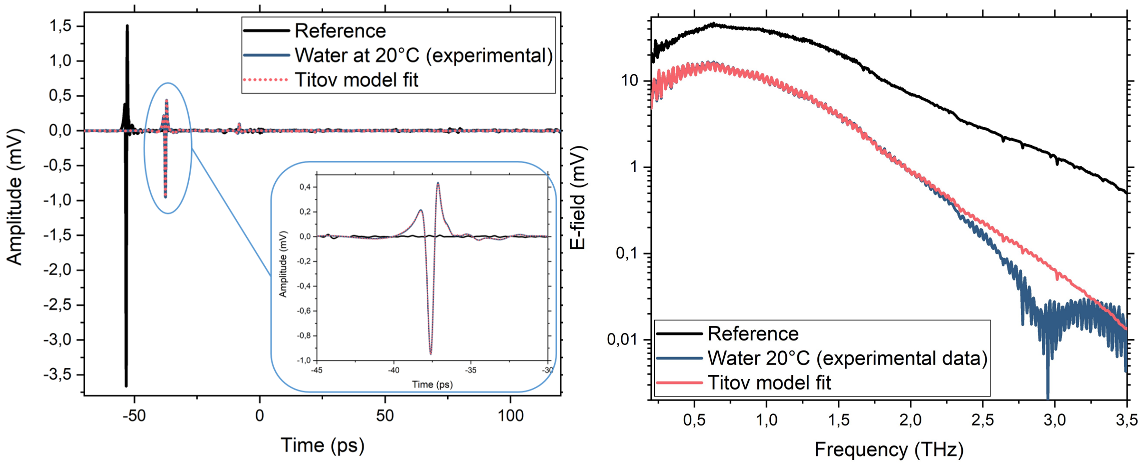
FIGURE 2.10: Result of the fitting process for a Rocard-Powles-Debye model. (A) Time-traces of the experimental measurements and fitted time-trace. (B) Retrieved spectra of the measurements and the fitted time-trace.

Then, we fitted the data with the Titov model, which results are shown in figure 2.11. The optimized parameters found for the Titov model are:

- ϵ_∞ = 2.796
- χ = 1.806
- η = 3.425e-08 s
- τ_D = 2.275e-06
- β_1 = 9.242e-18
- β_2 = 5.525
- σ_V = 8.892e-03
- γ_1 = 3.887e+05

First, we note that in all the models we tried, the value of ϵ_∞ is higher than 1.76, which is its value in the visible range. Therefore, if we only look at this coefficient, all the models seem valid. However, the problem resides in how to decrease to reach this value, hence the problem we have with the other models.

Here, the final Akaike criterion for this model is 3.010e+06, which is lower than any of the previous criteria. As for the Rocard-Powles-Debye model, taking inertial effects also has an impact on the fit in our frequency range, as well as interactions between molecules. Still, this model only takes into account interactions between two molecules, and we would need to solve a n-body problem in order to characterize all interactions. This problem is still discussed among theoreticians and will need more time and work to be solved. For this study, we will use the Titov model as the best model for liquid water.



(A)

(B)

FIGURE 2.11: Result of the fitting process for a Titov model. (A) Time-traces of the experimental measurements and fitted time-trace. (B) Retrieved spectra of the measurements and the fitted time-trace.

Applying this model to other temperatures

The fitting of liquid water data at other temperatures is ongoing. Since we encountered the Titov model recently, we prefer to reserve the fit results for the defence, where we will have taken a step back to understand the tendencies we observe. Still, we can say that the Akaike criterion decreases when the temperature increases. In other words, the less signal there is, the smaller the Akaike criterion (hence the better the fit). This is due to the fact that the model is not perfect, which we know as explained above.

2.2 Heavy water

Heavy water, deuterated water, or deuterium oxide (D_2O), is composed of the same chemical elements as normal liquid water, but its hydrogen atoms are heavy isotopes. This isotope is called deuterium, its nucleus contains a neutron in addition to the proton also present in any hydrogen atom. The presence of the heavier hydrogen isotope gives water different nuclear properties, and the increase of mass gives it slightly different physical and chemical properties when compared to normal water. For example, the melting point of deuterated water is 3.82 K higher than that of normal water. It has also been shown that the structural differences between liquid H_2O and D_2O display an appreciable variation over the temperature range of the liquid state and increase significantly near the melting point [126, 127]. Moreover, their viscosity is different (0.89 mPas for H_2O and 1.25 mPas for D_2O at 20 °C). However, their dipole moment is very similar (1.85 D for H_2O and 1.87 D for D_2O).

As a result, we expect that the structure of the hydrogen bond network is modified when compared to "light" water. Indeed, quantum dynamical simulations found that, when comparing the D_2O radial distribution functions with analogous results for liquid H_2O, the impact of nuclear quantum effects on the structural arrangement of the hydrogen-bond network is noticeable at all temperatures. Moreover, they found that, as in liquid H_2O, the hydrogen-bond dynamics in heavy water is effectively a collective process [128], which means that the permittivity models should be similar but the parameters should differ.

2.2.1 Experimental results

The experimental protocol is the same as in section 2.1.3. The D_2O sample was purchased at Sigma Aldrich (99.9 atom % D). The time traces measured during the experiment are shown in figure 2.12. As for H_2O, the time trace of the reference (measurement where the sample is just the mount of the microfluidic cell, without windows) is in black, the empty cell (the microfluidic cell is not filled with heavy water) is in medium blue and the measurement of deuterated water at temperatures between 10 °C and 75 °C are displayed in other colors. The zoom provided on the right shows the same evolution as for "light" water, meaning that as the temperature increases, the amplitude decreases and the pulse shifts to the right of the graph. Thus, we have the same conclusions, that is that the absorption increases and the refractive index as well when the temperature increases.

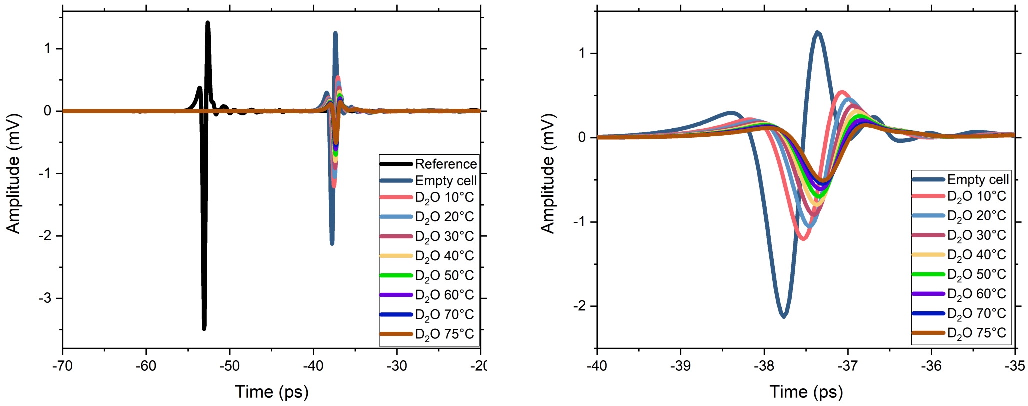
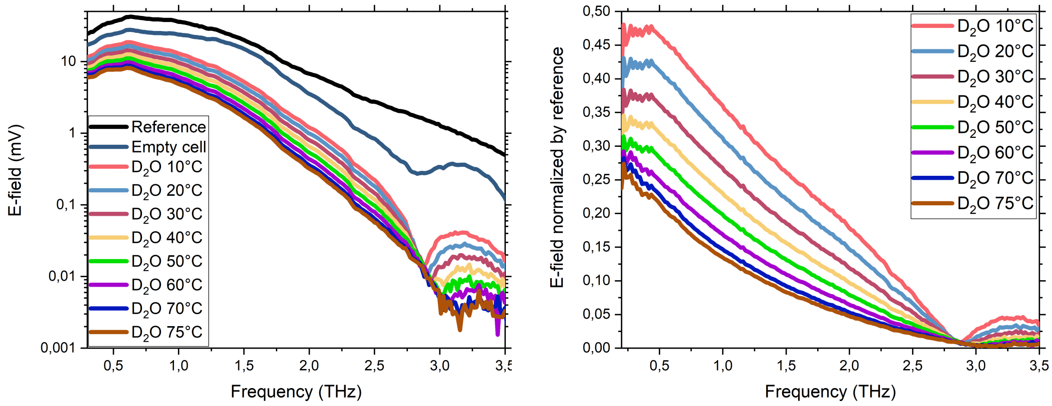


FIGURE 2.12: Time traces measured for the study of heavy water. In black the reference, in dark blue the empty microfluidic cell, and in other colors heavy water at different temperatures. On the right, a zoom on the main pulse of the heavy water measurements is provided.

The spectra obtained after a Fourier transform are displayed on the left of figure 2.13. On the right, the spectra of D_2O at different temperatures are normalized by the spectrum of the reference. Once again, we can observe the increase in absorption as the temperature increases.



(A) (B)

FIGURE 2.13: (A) Spectra retrieved from the measured time traces for the study of heavy water. (B) Spectra normalized by the spectrum of the reference.

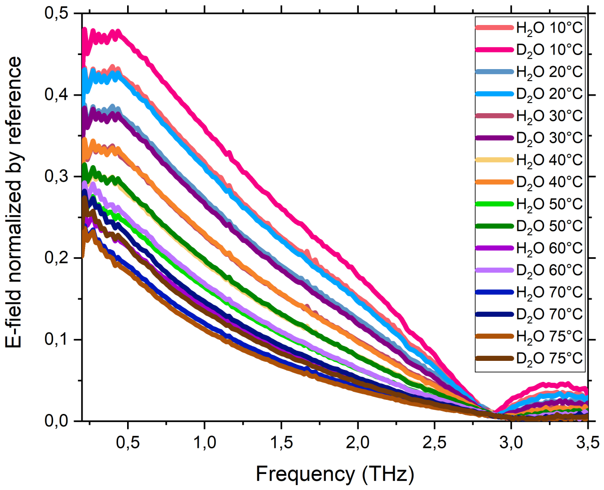


FIGURE 2.14: Comparison of the normalized spectra of D_2O and H_2O at all temperatures.

Figure 2.14 displays the normalized spectra of all temperatures of heavy and light water. As mentioned before, both heavy and light water absorb more signal as the temperature increases. However, at the same temperature, heavy water has less absorption than light water. This indicates that the refractive index of heavy water is smaller than the one of light water. As the structure of their hydrogen bond network is similar but slightly differs, we attribute this change in absorption to the modification of the network. Indeed, with the knowledge gained in the following section, we are able to say that the H-bond network of heavy water is more structured (more bonds are made) than the one of light water. The heaviness of the hydrogen atoms might prevent them from making and breaking bonds as quickly as light water does. To gain more understanding, the data have to be fitted with the Titov model as well.

2.3 Disruption of the H-bond network of liquid water with ions

From the differences seen between H_2O and D_2O, we can conclude that THz-TDS is very sensitive to the structural changes of the H-bond network. This is particularly interesting because, when biological objects are added to liquid water, they should be included in this network and thus modify its structure. In order to verify this hypothesis, we added ions to a liquid water solution. We decided to use ions classified as "chaotropes" and "kosmotropes". Chaotropic and kosmotropic ions have the property of breaking or making hydrogen bonds respectively. Thus, they have the property of decreasing or increasing the structuring of liquid water H-bond network. Figure 2.15 gives a classification of different ions, called the Hofmeister series, from the most chaotropic on the left to the most kosmotropic on the right. In addition to seeing how much THz-TDS is sensitive to the structural changes that occur, we can then use this study later to compare the effect of biological objects that are added to water to these results. Indeed, we will be able to understand how bond-making and breaking affect the THz pulse and how the ions are included in the network.

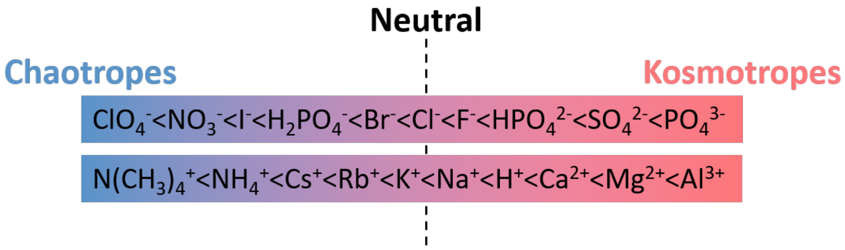
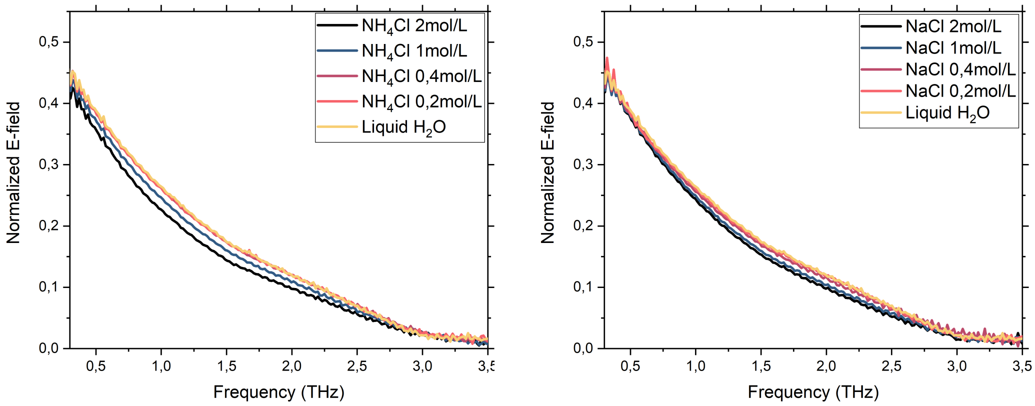


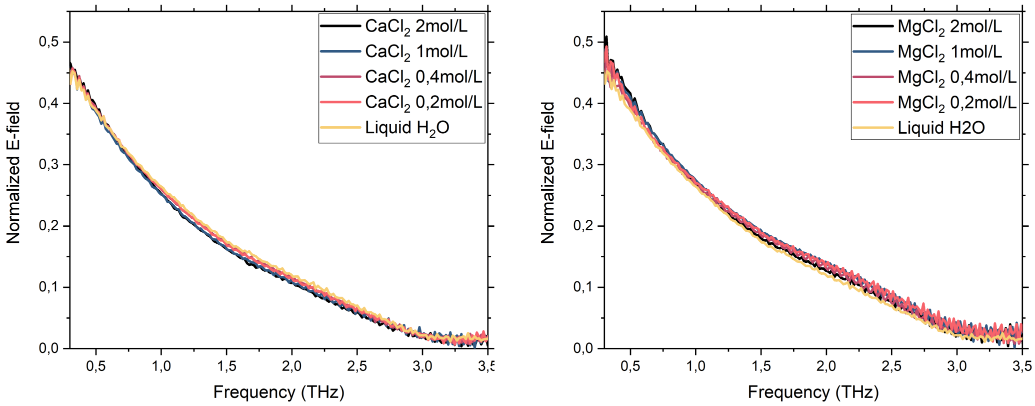
FIGURE 2.15: Hofmeister series classification of ions.

For this study, we chose a neutral anion, Cl^-, and several cations that should impact differently the network: NH_4^+, Na^+, Ca^{2+}, Mg^{2+} and Al^{3+}. For each ionic solution, we recorded the time traces of 4 concentrations: 2 mol/L, 1 mol/L, 0.4 mol/L, and 0.2 mol/L with THz-TDS. The experimental protocol is the same as for the analysis of liquid H_2O and D_2O with the exception that we only recorded the time trace at 20 °C. In figure 2.16 are displayed the retrieved spectra, normalized by the spectrum of the reference (that is the time trace of the microfluidic cell without windows). The normalized spectrum of liquid water is also added to all graphs for comparison.

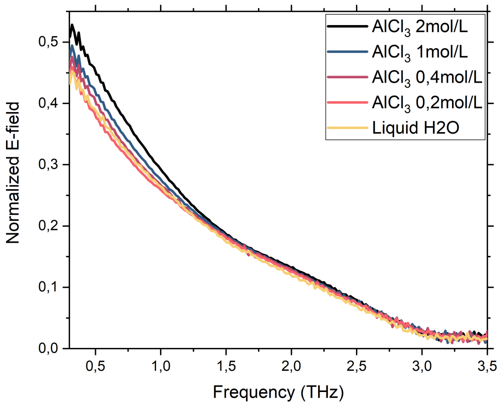
As expected, the transmission of the THz pulse is impacted by the presence of ions. Furthermore, the higher the concentration of the ionic solution, the greater the impact is. According to these results, the most neutral solutions are $CaCl_2$ and MgCl2, which is slightly off if we compare to the classification. As a result, the chaotropic or kosmotropic nature of the ions is not the only factor impacting the absorption of the sample. Still, the solutions of NH_4Cl and NaCl increase the absorption of the THz pulse while the solution of $AlCl_3$ decreases absorption. These differences are due to the changes of structure of the H-bond network. Indeed, chaotropic ions have a tendency to reduce the amount of order of the



(A) NH_4Cl normalized spectra (B) NaCl normalized spectra



(C) $CaCl_2$ normalized spectra (D) $MgCl_2$ normalized spectra



(E) $AlCl_3$ normalized spectra

FIGURE 2.16: Normalized spectra of the five ionic solutions for each concentration. Liquid water is also added in yellow for comparison.

structure, which causes more absorption, while kosmotropic ions increase the amount of order of the network, which results in less absorption. For more clarity, figure 2.17 displays the normalized spectra of the most concentrated solutions compared to liquid water. We can observe that the increase in absorption follows the same tendency as the Hofmeister series.

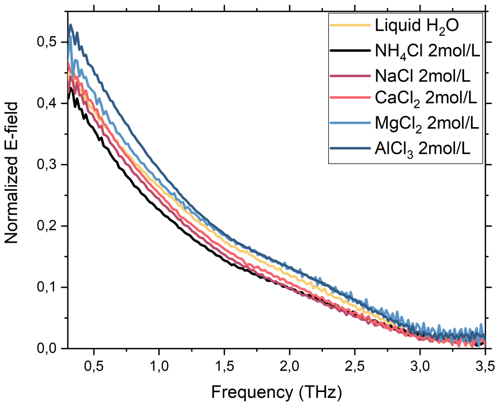


FIGURE 2.17: Normalized spectra of the highest concentration for each studied ionic solution compared with liquid water.

To sum up, THz-TDS enables us to characterize how molecules impact the H-bond network of water. For the ions that are the most chaotropic or kosmotropic, the changes in absorption clearly correspond to their nature, however for the ions that have less impact on the network, the data measures with THz-TDS is less clear. Therefore, there has to be other factors impacting the results, such as the weight of the ions added to the solution, or the interactions between the ions. Fitting the time traces from these data will give us even more information. According to how the parameters of the permittivity model of liquid water change when these ions are present, we will be able to deduce some of the microscopic dynamics and understand these experimental results better. However, these experiments were done before collecting the data about water at different temperatures, as we did not have to right microfluidic cell yet. We only have data about the temperature inside the purged plexiglas box. Since the influence of the temperature on the absorption is undeniable, we will need to redo these experiments with a rigorous monitoring of the temperature with our new microfluidic cell to be able to compare the results with the ones of liquid water.

2.4 Study of a protein: the lysozyme

We were able to record the disruption of the H-bond network by ions with THz-TDS. However, we analyzed specific ions that have a strong ability to modify the bonds. Biological objects should still disrupt the H-bond network but less intensely. Therefore, we begin by analyzing a more simple protein, the lysozyme.

The lysozyme is an antimicrobial enzyme produced by animals that forms part of the innate immune system. It acts as a biological catalyst, meaning that it increases the velocity of a chemical reaction without being altered. Almost all metabolic processes in the cell need enzymes in order to occur at rates fast enough to sustain life. Lysozyme is abundant in human secretions but can also be found in large amounts in egg whites, which makes it easily accessible and low cost.

In the THz range, the best way to study such a biosample is to use its crystalline form. Indeed, firstly, the concentration of sample is higher that way and all the proteins are oriented in the same direction, which avoids polarization issues. Moreover, the samples have the same symmetrical arrangement from one experiment to another, which ensures repeatability. Finally, this study has to be conducted in an aqueous environment because the crystals grow in this type of environment (hence removing them from it would risk damaging them), and because it is a natural environment for proteins to be in. As a result, it makes sense to stay close to *in vivo* conditions.

2.4.1 Experimental protocol

The experimental protocol contains the crystallization part, which is deemed easy for the lysozyme but still tricky for us as we are not biologists, and the study with THz-TDS.

Crystallization

The crystallization method we used is the vapour phase method in hanging drops. The steps to follow are:

1. Preparing the stock solutions: one 50 mmol/L sodium acetate solution, and solutions of NaCl from 0.6 mol/L to 1.6 mol/L.
2. Preparing the lysozyme solutions: two solutions of 50 mg/mL and 25 mg/mL.
3. Filling the wells of a Linbro plate with 1 mL of stock solution. Following the schematic below, we fill the reservoirs of column 1 with the 0.6 mol/L NaCl solution, reservoirs from column 2 with the 0.8 mol/L NaCl solution and so on until we reach 1.6 mol/L.

	1	2	3	4	5	6
A	○	○	○	○	○	○
B	○	○	○	○	○	○
C	○	○	○	○	○	○
D	○	○	○	○	○	○

4. On a coverslip (which is actually a quartz wafer of 24 mm diameter and a thickness of 2 mm), mixing 2 µL of the lysozyme solution (50 mg/mL for line A and 25 mg/mL for line B) with the NaCl stock solution from the associated well, such as drawn on figure 2.18.
5. Flipping the quartz coverslip and putting it on the associated well.
6. Repeating the last two steps for all wells.

In the end, one well of the Linbro plate contains stock solution and is covered by a quartz coverslip where the drops of lysozyme and stock solution are hanging, the impermeability being ensured by the immersion oil (see figure 2.19). Then, the crystals are obtained within 24 hours.

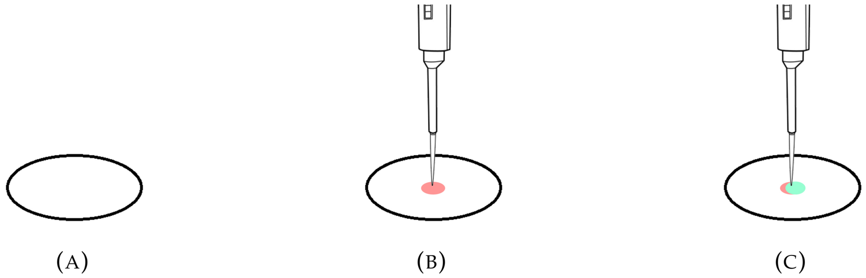


FIGURE 2.18: (A) Take a coverslip and lay it flat. (B) Deposit 2 µL of lyzozyme solution. (C) Deposit 2 µL of stock solution.

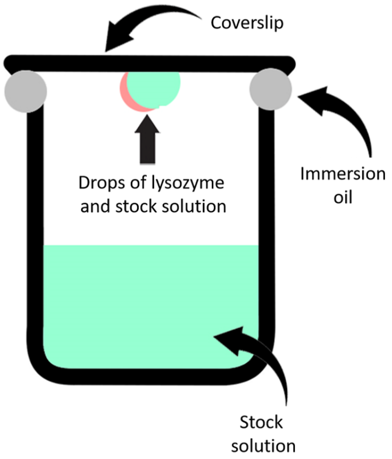


FIGURE 2.19: Schematic of a Linbro plate well after all the steps of the protocol have been done.

THz-TDS experiment

The experimental protocol for the THz-TDS experiment slightly differs from the previous ones. Once the lysozyme has crystallized on the quartz wafer, we remove and flip the wafer from the well. We also remove most of the immersion oil from the sides of the wafer. Then, we place a Teflon spacer on top of the wafer (the spacer thickness varies according to the size of the crystals grown), and finally another quartz wafer on top of it. This creates a setup close to the microfluidic cell used previously. In order to hold this layered sample, we created a mount, shown on figure 2.20.

Finally, the THz-TDS measurements are similar to the previous ones. We use the same THz-TDS setup from Menlo System, placed inside a nitrogen purged plexiglas box. However, the collimating lenses have been replaced with two parabolic mirrors in order to avoid the TPX lenses absorption. The two focusing lenses have a focal length of 100 mm. The humidity and temperature inside the box are monitored. A reference is taken (without any sample on the path of the beam), then the measurements of the lysozyme inside its "cell" are taken, and finally we take a measurement of the same cell, but without crystals inside (only the stock solution). A photograph of the setup is displayed in figure 2.21.

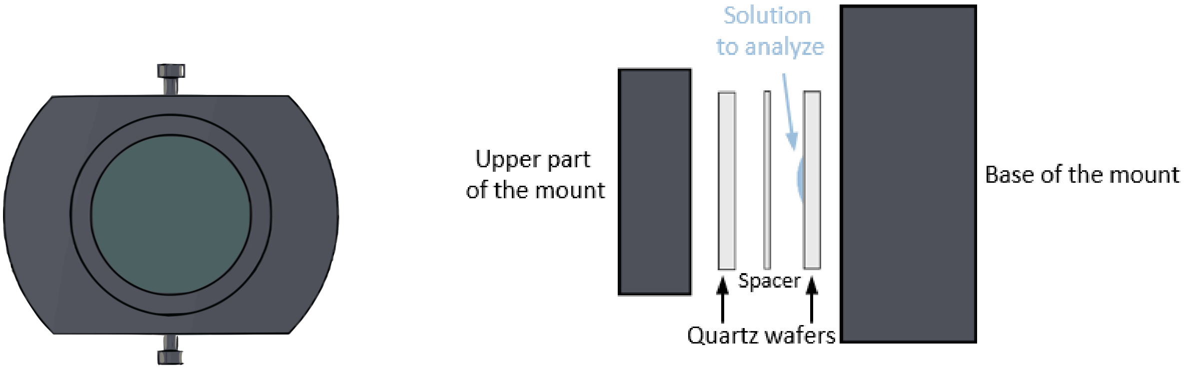


FIGURE 2.20: Sample placed inside the THz-TDS setup to study the crystallized lysozyme in solution. On the left is the front view of the mount, on the right is the side view of the mount with the layered quartz wafers.

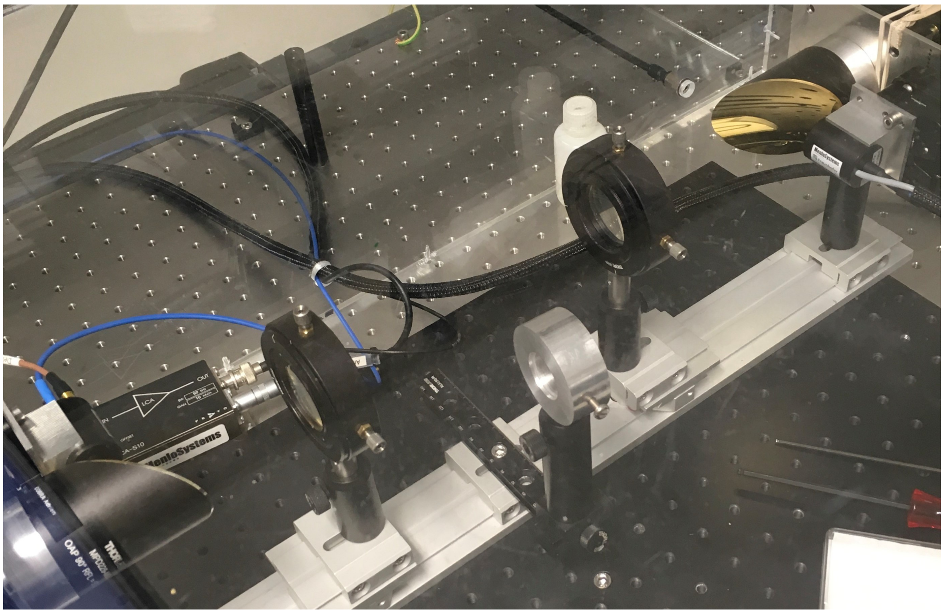
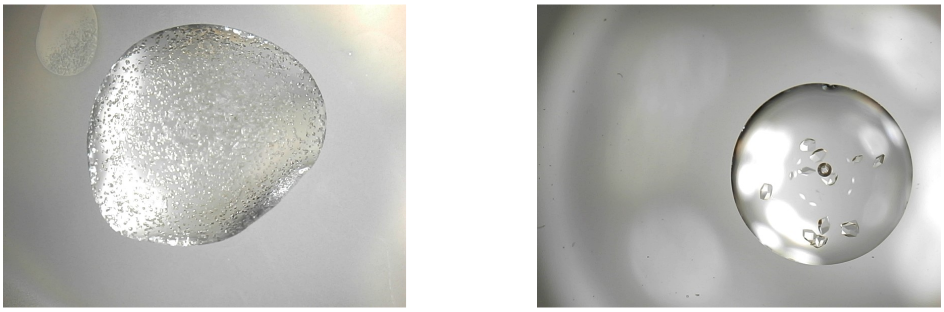


FIGURE 2.21: Photograph of the setup for the lysozyme analysis. The emitted pulse is collimated by a parabolic mirror and then focused by a TPX lens on the sample (the mount is offset on the photograph to take the reference time trace). The symmetric setup focuses the pulse onto the detector after the sample.

2.4.2 Results

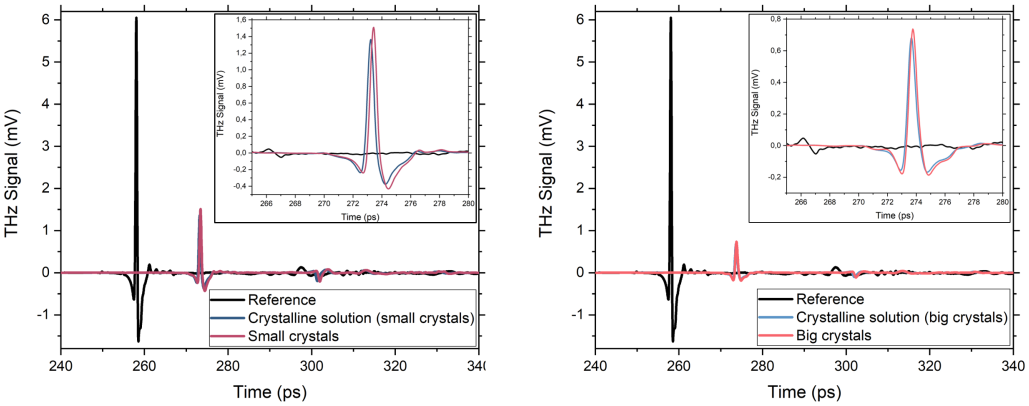
Figure 2.22 shows the crystals obtained after following the protocol explained above. We were able to grow two sizes of crystals that will be called “small” (for the 50 µm ones) and “big” (for the 300 µm ones) in the rest of this study. The small crystals are obtained by following the protocol explained above, and the big crystals are obtained adding a first step to this protocol, that is storing the stock solutions in the refrigerator beforehand. Since the temperature is cooler, the crystallization process takes more time and the crystals are therefore bigger. As the order of magnitude of the wavelength of the THz pulse is 300 µm, we decided to investigate both sizes.



(A) (B)

FIGURE 2.22: (A) Photograph of the lysozyme crystals, their diameter is approximately 50 µm. (B) Photograph of other lysozyme crystals, where the solutions had previously been stored in a refrigerator, resulting in bigger crystals of around 300 µm diameter.

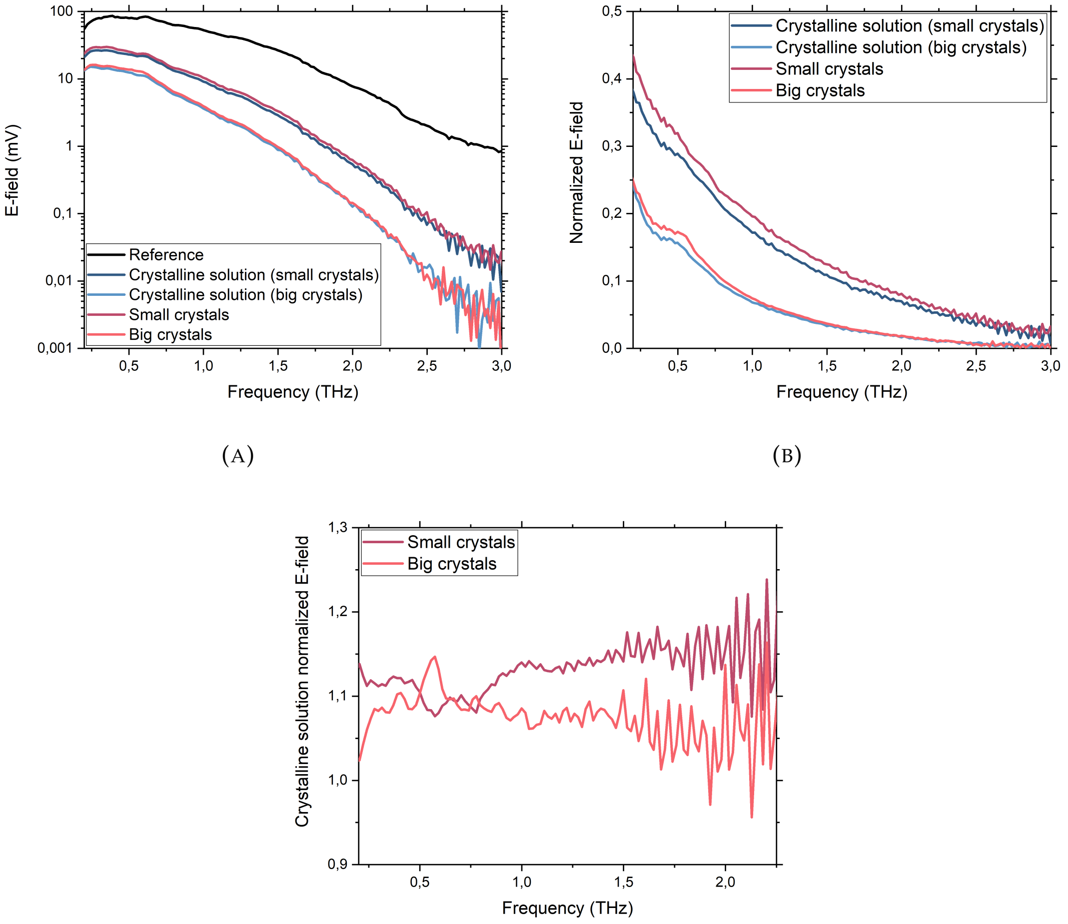
Figure 2.23 shows the time traces measured for the reference, the crystalline solution alone, and the lysozyme crystals in their solution. There is a slight delay between the time trace of the solution alone and the solution with crystals and a difference in amplitude. This can be due to a difference in the refractive index, which would indicate a signature of the lysozyme in the THz range, or simply a difference in the thickness of the space between the quartz wafers.



(A) (B)

FIGURE 2.23: Time traces measured for the small crystals (A) and for the big crystals (B). For each experiment were measured a reference time trace, the time trace of the solution without crystals and the solution alone.

To investigate this further, we retrieved the spectra of the measured time traces, showcased on figure 2.24. The first graph shows the raw spectra, where we can see that there is less absorption when the crystals are present. When we normalize the spectra of the crystalline solution alone and of the crystalline solution with lysozyme with the spectrum of the reference, this difference in absorption is even more visible. In order to investigate if small features were visible on the spectra, we traced the ratio between the crystalline solution with lysozyme and the associated crystalline solution alone. Even if some peaks and dips are present, they are not repeatable.



(A)

(B)

(C)

FIGURE 2.24: (A) Spectra retrieved from the time traces measured. (B) Spectra of the crystalline solutions and crystals normalized by the spectrum of the reference. (C) normalized spectra of the crystals divided by the normalized spectra of the crystalline solutions.

Finally, in order to investigate whether the difference in absorption (and the delay in the time domain) was due to a different refractive index or a different thickness, we used the fit@TDS software for a quick fit (without models, that is only a parameter for a constant refractive index and a parameter for the thickness, but with pre-processing of the data). We found that there was indeed a difference in thickness between the two measurements, which is due to the fact that the mount holding the layered sample does not allow to control the force applied to close it and thus the pressure put on the wafers is never the same. Hence, some are well compressed and the thickness is maintained by the spacer, while others are looser. This is not a problem when we use the microfluidic cell since the upper part of the mount screws onto the other part, always applying the same pressure.

Still, when we trace the absorption coefficients (see figure 2.25), there is a slight decrease in absorption at low frequencies that is repeatable, which is a great motivation to keep improving the experimental setup and the data analysis process.

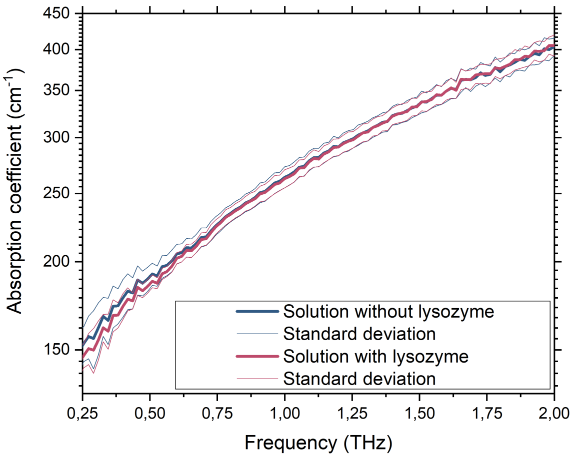


FIGURE 2.25: Mean of the absorption coefficients of all the measurements, with the standard deviation in each case (with and without the presence of lysozyme crystals).

In conclusion, we were able to grow lysozyme crystals and design a setup in order to analyze them with THz-TDS. We measured the time traces of the crystals in solution but we were unable to differentiate them from the time traces of the crystalline solution. Still, we think that if there is an effect of the presence of lysozyme crystals on the THz pulse, this effect must be small. Therefore, we need to fit these time traces in order to find if there are some differences in the parameters retrieved. The results of the fit will be shown during the presentation.

2.5 Conclusion

In this chapter, we focused on the analysis of liquid samples, as most biological samples are preserved in a buffer solution, composed mostly of liquid water. Hence, in order to be able to obtain any signal from a biological object, we have to characterize the buffer solution first.

First, we focused on liquid water, one of the most studied, but also one of the more complex, samples. The dynamical behaviour of the hydrogen-bond network, responsible for the liquid nature of water at ambient temperatures, is still only partially understood. In the literature, many permittivity models have been proposed in order to fit THz-TDS data in their scanned frequency range but, when we look at the permittivity of liquid water over a wide frequency range, from 0.1 GHz to 2000 THz, none of these models describe accurately its behaviour. Since these models do not take into account interactions between molecules and inertial effects, we investigated models that do take these effects into account and found that the most accurate one was one we called the "Titov model", which considers inertial effects as well as interactions between two polar groups. However, taking into account all interactions is a lot more complex and more work, as well as a collaboration with theoreticians will be needed. Still, THz-TDS is a great tool to analyse liquid water and is sensitive to the temperature changes. Moreover, the analysis of heavy water showed that, when the H-bond network is modified, the measured data changes as well.

When biological samples are added to liquid water, they will inherently alter the structure of the H-bond network. If we know now that THz-TDS is sensitive to the changes in the network, we decided to investigate bond-making and bond-breaking ions as a way to better interpret THz-TDS data. We found that when bonds are broken, the absorption increases, and when more bonds are mare, the absorption decreases.

Finally, we analysed a biological sample in liquid water, hen egg white lysozyme. After crystallizing the lysozyme and recording its THz-TDS time-trace, we found that its contribution was too small to be observed on the experimental curves. However, a fit of the data will uncover any impact of the lysozyme on the THz pulse. From an experimental point of view, scanning a sample with less water would absorb less and thus be easier to analyse. To this end, we can use the device presented on the next chapter.

Doing these experiments on liquid samples taught us several important things to keep in mind for the future. Firstly, it is absolutely necessary to be stable in temperature, and to monitor it during an experiment. The studies of H_2O and D_2O showed that any temperature variation changes the absorption values. Therefore, if we want to compare any data and ensure the repeatability of the measurements, we need to have a precise measurement of the temperature of the liquids. The second lesson is the importance of chemical stability. For any sample, but particularly for biological objects, it is important to have a good baseline in order to compare data since their effect is weak. Therefore, we have to analyse the exact buffer solution the samples are in, as the ionicity of the solution has an influence on the absorption of the solution.

Chapter 3

The challenge of analysing small samples with THz-TDS

In these last chapters, we have showed the use of THz-TDS for analysing some biological samples, either in a pressed powder form, or in liquid solutions. However, THz-TDS suffers from the size of many biological objects, such as bacteria, cells or viruses, which are a lot smaller than the size of the THz wavelength (and subsequent focal spot). Indeed, the THz wavelength is typically between 0.1 and 1 mm, which means that the diffraction limit is at best 50 µm, which is one to two orders of magnitude too big. To overcome this problem, we could analyse a large quantity but the production of large volumes of biosamples, such as viruses, is not possible. As a result, pellets cannot be made and we can either produce less than a mg of powder or having a small concentration of samples in a buffer solution, of which we have seen the challenges in the previous chapter. Therefore, interactions with these photons are highly reduced, making it hard to analyse samples with a volume smaller than a microlitre. We need to increase light-matter interactions between them and THz pulses. This chapter is divided in two main parts: a first part, explaining the first realization of a light confinement design (section 3.2), that took place before and at the very beginning of my thesis work, and a second part explaining my improvements of the design (section 3.3).

3.1 General introduction and motivations

3.1.1 Confining light for the analysis of biosamples

According to Purcell's theory, confining light to a small volume increases light-matter interactions. In materials with a high refractive index, light confinement is attainable using a photonic crystal structure for instance. However, the refractive index of biosamples is often relatively low, which makes light confinement a challenge for several spectral ranges. In the visible domain and the near infrared, the use of dielectric structures for light confinement has shown interesting results but the reached volumes are still too large (of the order of magnitude of the cube of the wavelength) [129–131]. In order to achieve smaller volumes, metallic structures based on plasmonic effects have been used but they generate additional losses [132–134].

In the THz community, most of the light confinement methods exploit the fact that, in the THz domain, metals have a better response than in other frequency ranges. Two approaches emerge:

- The use of a near-field microscope adapted to THz-TDS [135]. This part I will not focus on but there is ongoing work in our THz biophotonics team on the subject [136].
- The use of a sub-wavelength metallic waveguide.

For the second approach, one can, for instance, use a metallic wire [137] around which the THz wave is going to propagate. In this case, the sample is placed around the wire but the light-matter interaction is not fully enhanced and using actual metallic edges instead will lead to a better confinement. Therefore, in order to confine light in one direction, a natural approach is to use a parallel-plate metallic waveguide [138]. In such a waveguide, two kinds of modes exist: transverse electric (TE) and transverse magnetic (TM) (which are in fact transverse electromagnetic mode (TEM) if the dielectric is homogeneous) depending on the direction of the electric field relative to the metallic walls. The TEM mode has no frequency cut-off [139], moderate losses, and, consequently, weak dispersion (only linked to the metallic material's dispersion), which is ideal for THz-TDS experiments [16]. However, it is necessary to use a large bandwidth coupling scheme to properly excite the mode of interest.

A first coupling method consists in using silicon cylindrical lenses [138], which introduces a coupling that is very dependent on the waveguide's geometry, and adds reflections and Fresnel losses because of the high refractive index of silicon. Despite this, the technique has been used inside a cryostat and showed encouraging results [140]. To avoid these limitations, it is possible to use tapered coupling antennas [141], which improve the coupling and facilitate the use of the device. However, the main limitation of the parallel-plate metallic waveguide is that we suppose that its height is quasi-infinite, meaning that there is no vertical confinement (see figure 3.1). Thus, during the propagation inside the waveguide, the mode is going to expand in the vertical direction (where there is no confinement), reducing the light-matter interactions and producing a strong astigmatism when the beam is coupled to free space (illustrated on figure 3.1).

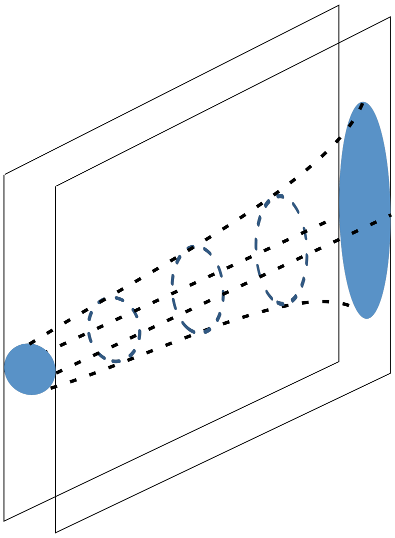


FIGURE 3.1: Expansion of the mode (in blue) in the vertical direction of a parallel-plate metallic waveguide, producing a strong astigmatism.

3.1.2 Objectives of the confinement device

In order to have fewer losses and fewer dispersion, we decided to design a device to 2D confine THz pulses into a 1D waveguide. To do so, we combined the curved tapered antenna approach by [141] and the thick slot line waveguide (TSLW) (a map of the field of the mode is shown in figure 3.2). The TSLW has an intermediate shape between the parallel plate waveguide [142, 143] and the slot line waveguide [144]. If no dielectric substrate is close to the slot (homogeneous dielectric), both are known to support a TEM mode, and thus have little dispersion [145, 146], even up to optical wavelengths [147].

In order to achieve our goal, the device has to complete four requirements:

1. The guiding part, where occurs the confinement and light matter interaction enhancement should be long enough (several mm, that is typically several tens of wavelength).
2. The guiding part has to remain accessible in order to fill it with the samples.
3. The THz signal should be injected and extracted from the waveguide with standard THz lenses or mirrors.
4. There should not be significant dispersion introduced by the device nor Fabry Perot echoes resulting from reflections.

As a result, the cross section of the waveguide should be open, with a gap below the wavelength scale, that is permitted by the TEM mode. Moreover, we need the same non-dispersive antennas for injection and extraction of the pulse (a symmetric device).

3.2 The first realization of the "butterfly" device

3.2.1 Design of the device

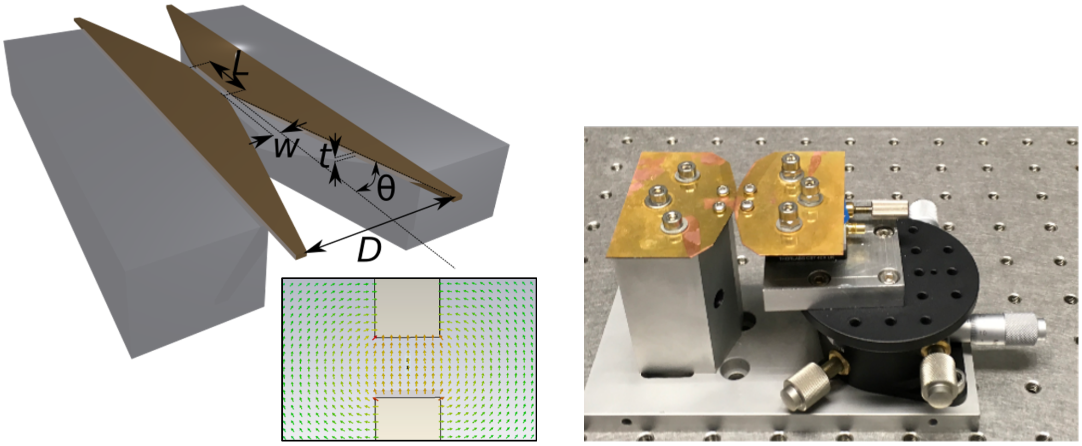
The chosen design is shown in figure 3.2, the device is composed of three parts: an injection exponentially tapered antenna (ETA) (that can be seen as a Vivaldi antenna [148]), a TSLW, and an extraction ETA. We gave it the name “butterfly” because of its two wings.

The physical dimensions of the device are displayed on the schematic of the butterfly in figure 3.2. The thickness, t, of the metallic plates of the TSLW was chosen to be 500 µm, in order to be thick enough to insert a sample but thin enough to confine the THz field with moderate losses. The width w is variable, as one of the wings of the butterfly is mobile thanks to a translation stage with a micrometric screw. The length L of the waveguide is chosen to be 10 mm, which is comfortable for THz solid-state samples but allows the evaluation of losses in the device.

Thanks to the exponentially tapered antenna [149], we avoid reflections at the guide/antenna interface as the curve is continuous, with a continuous derivative. Its shape follows the parametric equation:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} s \\ \beta(\exp(\frac{s}{\alpha}) - 1) \end{bmatrix} \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} = \begin{bmatrix} s\cos(\theta) - \sin(\theta)\beta(\exp(\frac{s}{\alpha}) - 1) \\ s\sin(\theta) + \cos(\theta)\beta(\exp(\frac{s}{\alpha}) - 1) \end{bmatrix} \tag{3.1}$$

Where $\theta = -\arctan(\frac{\beta}{\alpha})$, the opening angle of the antenna away from the waveguide, is set to keep the curve derivative continuous at the interface with the waveguide. The parameter α



(A) (B)

FIGURE 3.2: (A) Solidworks drawing of the light confinement device. L is the length of the waveguide, w the gap between the sides, t its thickness, and D the opening at the entrance of the tapered antenna. The inset provides a map of the electric field inside the waveguide. (B) The light confinement device fabricated and mounted on a micrometer displacement stage.

corresponds to how fast the exponential is going to end up following its asymptotic linear curve, and the parameter β is defined through the opening angle θ. As a result, the larger α is, the smoother and longer the exponential part is, which will introduce some uncertainty on the length of the waveguide (as the end of the waveguide is smoothed with the beginning of the antenna). We fixed α to 10 µm and β to 30 µm, resulting in an opening angle θ of 18.4 degrees (displayed in figure 3.2).

For the validation of the design, electromagnetic (EM) time-domain simulations were performed with the CST software. These results are presented in [150]. It was found that, in the waveguide, the losses are inversely proportional to the width w and originate naturally from ohmic losses in the metal. The results also shown negligible dispersion as expected for a TEM mode. For the antenna, it was found that there was a reflection at the end of the antenna, meaning that there was a low frequency cut-off as in most of large-bandwidth antennas. However, for an opening length D of 5 cm, the cut-off frequency was around 12 GHz, which is perfectly usable for THz-TDS. It is important to note that, due to computational limitations, we were not able to simulate the whole device in all of the frequency range (0.1-4 THz). Indeed, the device has a size, in wavelength units, of several hundreds of wavelength. The number of mesh points needed for this simulation was not supported by CST on our computer.

3.2.2 Fabrication process of the butterfly

The main step of the fabrication of the device is cutting of the wings from a 500-μm-thick copper plate (purity 99.9%). We need a technique that is able to cut a complex curved trajectory (from the tapered exponential profile of the antennas), and that is able to cut almost vertically the TSLW edges, with low roughness (small compared to the working wavelength) to avoid ohmic losses and propagation scattering effects. The classical machining of such a thin copper sheet is rather difficult and therefore, laser ablation was chosen.

The laser setup used to cut wings in copper plates is based on a Tangerine laser source from Amplitude-Systèmes that generates 350 fs pulses at an infrared wavelength of 1030 nm. The laser beam trajectory is controlled using a galvanometric scanning head followed by a final telecentric focusing lens of a focal distance of 100 mm. The focused spot diameter is estimated to be around 10 µm at the considered wavelength. An optical attenuator composed of a half-wave plate and a polarizer allow fine-tuning of the beam power independently from the laser source parameters. After crossing the entire optical path, the maximum available average power amounts to 12 W at a repetition rate of 200 kHz. The copper cutting was performed at maximum average power, the scanning speed was adjusted to 20 mm/s, and 150 laser beam passes were used. The optimization of the parameters and the cut were performed thanks to the expertise of Flavie Braud (IEMN engineer). I would like to note that the parameters are under continuous improvements and were given for reproducibility reasons.

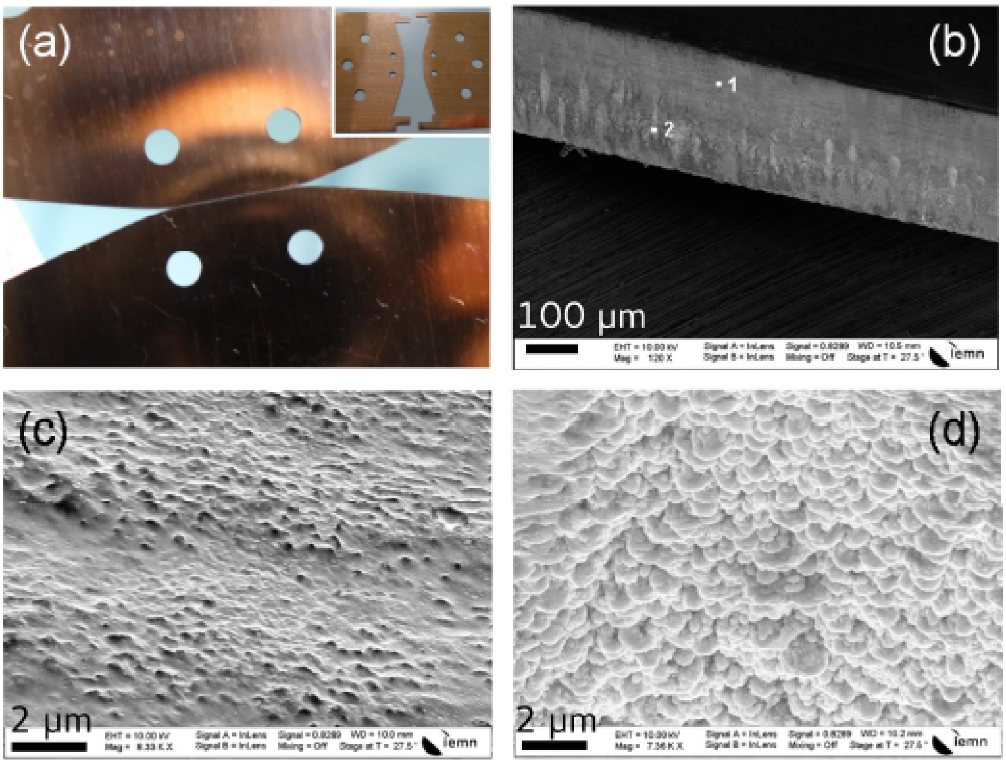


FIGURE 3.3: Characterization of the laser cutting process. **(a)** Assembly of the two wings, showing the TSLW slot. The minimum slot width of 30 µm is determined by two integrated spacers cut at each extremity of the wings, as shown in the inset. **(b)** SEM view of the laser cut sidewalls; **(c)** Magnified view of the surface roughness of region 1; **(d)** Magnified view of the surface roughness of region 2.

In figure 3.3**(a)**, we can see the waveguide part of the device, where we were able to obtain a 30 µm gap between the wings. From the scanning electron-microscope (SEM) images **(b)**, we make two observations. Firstly, no defects induced by overheating or fusion are observed, in contrast with the use of continuous wave and pulsed lasers of the nanosecond range. Secondly, the roughness is different in the upper part labelled ’1’, corresponding to the laser beam entrance, and the lower part labelled ’2’. Figures **(c)** and **(d)** correspond to zoomed in views of the upper and lower parts respectively. The difference in roughness is due to more thermal losses in the lower part. Indeed, even though the femtosecond regime avoids normally excessive heating of the matter (athermal ablation), the heat accumulation resulting from a high repetition rate at the bottom of the kerf line can partially activate the thermal ablation regime and is consistent with the formation of a crater-like rougher surface [151–153]. Nonetheless, the overall roughness remains on the micron scale (around 3 µm RMS), which should not impact too much the THz propagation in theory.

3.2.3 Performances

As for the simulations, the characterization of this device is given in details in [150]. Our Menlo Systems THz-TDS inside the purged plexiglas box was used, with the setup displayed in figure 3.4. It is composed of four TPX lenses with a focal length of 50 mm, of which two are collimating lenses for the photoconductive antennas and the last two inject and collect the THz pulse in and from the butterfly. The butterfly was aligned so that the focal plane of the lenses was placed at the entrance and exit of the TSLW part. According to the EM simulations, the butterfly should ideally not need focusing lenses in order to couple the pulse inside the waveguide but after some testing we found that they were needed. This problem will be addressed in section 3.3.4.

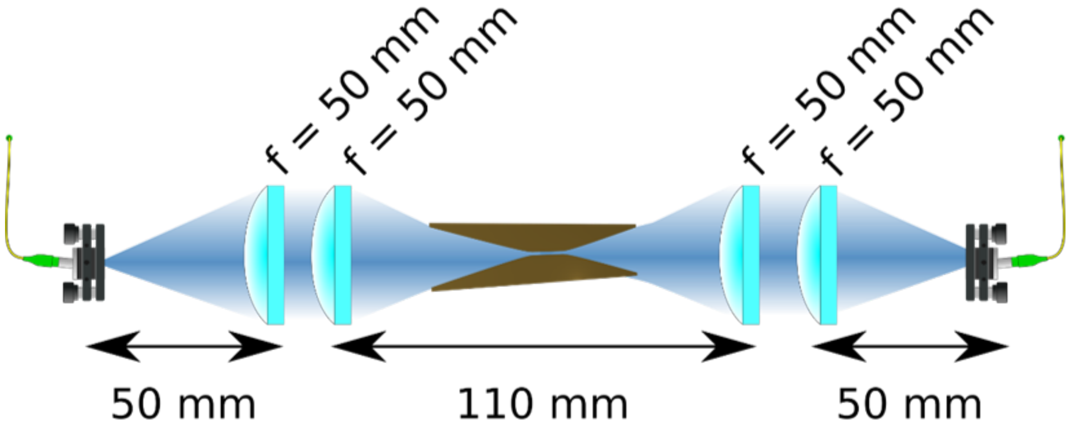
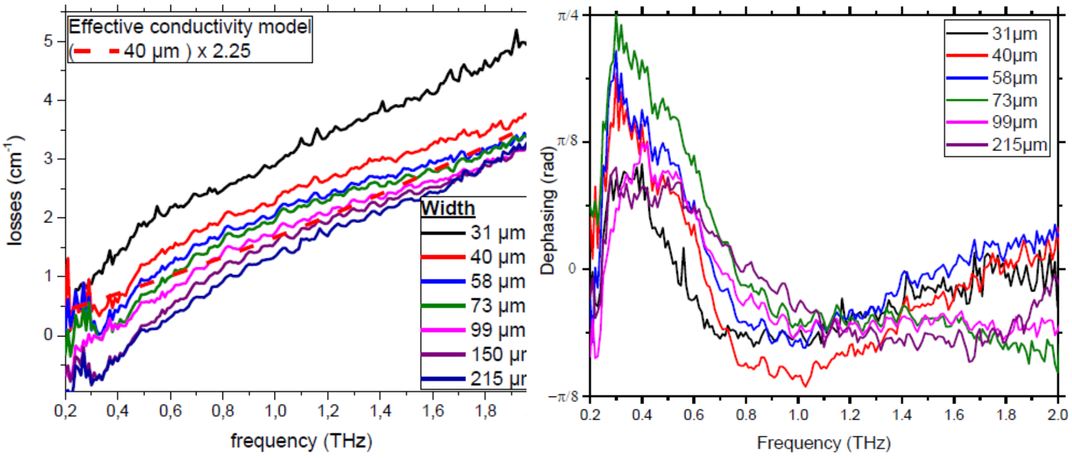


FIGURE 3.4: Schematic of the THz-TDS setup used for the characterization of the butterfly device.

With the help of the lenses, we found that the coupling efficiency was close to unity. Indeed, most of the energy ($> 90\%$ for a gap of w =31 µm and $> 98\%$ for w =215 µm over the whole frequency range) goes through the waveguide, especially for the higher frequencies. However, a parasitic pulse that does not go through the guide was observed. When the gap of the butterfly is closed, there is a small pulse that can be detected, which is hidden by the reference pulse going through the device otherwise. Another way to observe this parasitic pulse is to put a sample inside the butterfly. The beam that goes through the butterfly (hence that goes through the sample) is delayed while the parasitic pulse is not. In figure 3.6, we can see the two pulses, one is located at around 2 ps (aligned with the reference), and the other one is located at around 12 ps. The parasitic pulse is a part of the THz beam that goes above and below the waveguide. It is likely to be surface waves, created at the metal/air interface, due to the fact that the THz-TDS antennas are not perfectly polarized (there is still 20% of the vertical polarization).

Regarding power losses, they account for a few inverse centimeters, increasing with the frequency and decreasing with the gap length w, as is shown in figure 3.5. Experimentally, they are four times higher than in simulations, which is due to the roughness of the sidewalls in the gap. Finally, it was found that the dispersion was very low in the whole spectrum (below $\pi/4$).



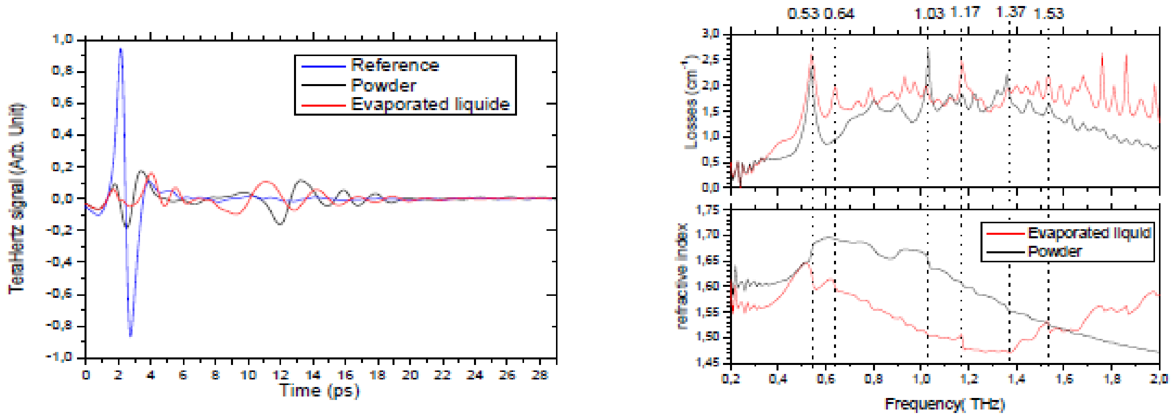
(A) (B)

FIGURE 3.5: (A) Power losses for various values of w (width), the negative values are due to diffraction. (B) Dispersion curves for various values of w (Figure from [150]).

3.2.4 Test on α-lactose monohydrate

Once again, we tested the device with the “gold” standard, and simple biosample, α-lactose monohydrate. This was done in two steps. First, on a powder sample, and second on an evaporated lactose solution. For the powder sample, the gap was opened at $w = 150\,\mu\text{m}$, which corresponds to a waveguide volume of 0.75 µL that was filled by inserting the powder from the top. The sample stayed in place thanks to agglomeration and edge surface adhesion effects. For the liquid sample, we made a saturated solution of lactose in deionized (DI) water (200 g/L). The gap was filled with the solution from the top using a 10 µL micropipette and let dry. The residue left after drying was estimated to weight 200 µg. The measured time traces, as well as the retrieved losses and refractive index are showcased in figure 3.6.

Comparing with the results found in chapter 1, the main peak at 530 GHz is still present. Therefore, with the butterfly, we are able to retrieve not only the refractive index with the delay, but also the absorption lines from the sample. In addition, the parasitic pulse is still here (visible at 2 ps for the powder and evaporated liquid time traces in figure 3.6). Thus, to avoid any fake interference effects, the entire signal before 6 ps was removed, before performing the Fourier transform, only keeping the actual data from the sample. From the spectrum, we extracted the effective refractive index and the losses. The effective refractive



(a)

(b)

FIGURE 3.6: Results of TDS experiments on lactose with the butterfly. **(a)** Time-domain results. **(b)** Associated calculated losses and effective refractive index.

index and the losses found are slightly lower than the ones found in the first chapter. This is due to the fact that the gap is not completely filled with lactose. Moreover, the pressure exerted by the butterfly walls is certainly weaker than pressure in the press used to fabricate the pellets. Consequently, more voids are present here. Also, some peaks are different from the ones found earlier (at 1.53 THz and 1.03 THz). These are Mie resonances coming from the scattering and random nature of the medium such as Mie or effects that are more complex. In fact, there are many other tiny peaks in the spectra, that are not reproducible, especially in the evaporated liquid one, which we interpreted as such.

To sum up, the butterfly device allowed us to perform THz-TDS on a small powder biosample (less than a µL) and on an evaporated liquid biosample (less than 10 µL of liquid) and we were able to obtain the main spectral peaks recorded in the literature. Still, several improvements or modifications are needed:

- The important roughness has to be reduced because it prevents from closing the gap even more to analyse smaller volumes. It is also responsible for the limited bandwidth (2.5 THz) accessible with the device.
- The cost of a copper plate is high and will be a problem if several devices have to be produced at once.
- The parasitic pulse has to be blocked.
- The coupling issue makes the use of focusing lenses mandatory and questions the efficiency of the tapered antenna.

All of these issues have to be resolved in order to analyse smaller volumes of samples.

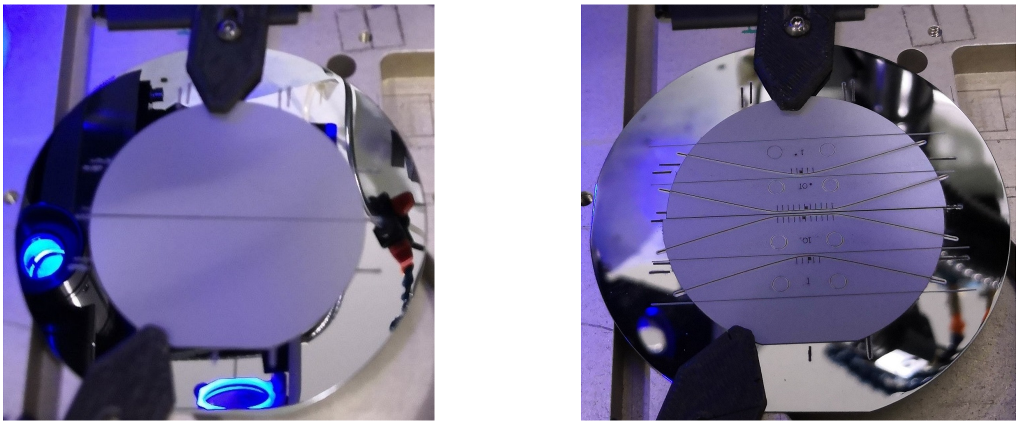
3.3 New version of the butterfly to achieve smaller volumes

The previous sections gave a summary of what was done by the team on the butterfly device before I arrived (and at the very beginning of my thesis work). The next sections will focus on my work to improve the performances of the device. Due to the limitations listed above, we decided to change the fabrication process. The goal was to have a broader bandwidth, get rid of the parasitic pulse, lower the cost, and reduce the roughness to be able to do experiments on smaller volumes of samples.

3.3.1 The new fabrication process

For this new version, we switched from a 500-µm-thick copper plate to a 330-µm-thick silicon wafer. We kept the same butterfly profile, with the same values for the length of the guide and the opening angle.

Laser cutting

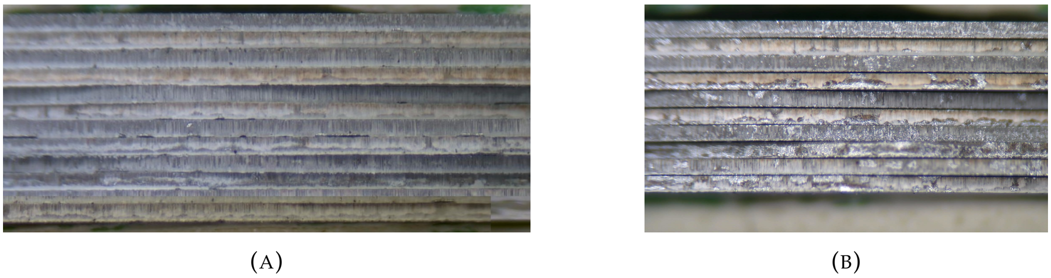


(A) Experimental setup of the laser cutting process (B) Silicon wafer cut after the process

FIGURE 3.7: Beginning and end result of the laser cutting process

The butterfly wings are now obtained by laser cutting silicon wafers. The laser and its parameters are the same as for the first version. The cut is performed by the same engineer as for the first version. The silicon wafer is placed onto a silicon holder with an internal diameter slightly smaller than the one of the wafer. Two clamps hold the wafer from both sides, as shown in figure 3.7. This figure also shows the finished cutting process. After they have been detached from the wafer, we observed them with an optical microscope and noted the presence of a white deposit, as observed in figure 3.8a. In order to improve the quality of the waveguide, the wings have to be cleaned.

To clean the butterfly wings, they were immersed in acetone in an ultrasonic bath for 30 min. The temperature of the ultrasonic bath was set to 45 °C, and the power was set to 100 %. In figure 3.8b, we can observe that the guides have lost some of the white deposit but there are still defects to remove.

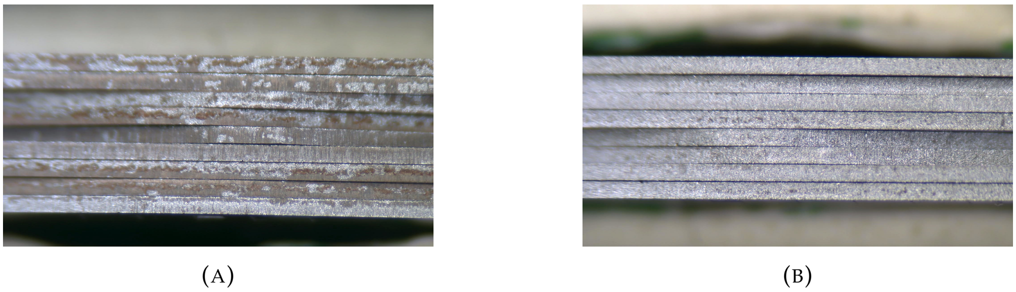

(A)

(B)

FIGURE 3.8: (A) View of the stacked butterfly wings with an optical microscope after laser cutting. (B) View of the stacked butterfly wings with the optical microscope after the first ultrasonic bath.

Cleaning with hydrogen fluoride

We suspected that the white deposit was formed during the cutting process by a silicon compound (such as SiO_2). Hence, we decided to use HF (Hydrogen Fluoride) to remove it. The butterfly wings were immersed in a solution composed of 10 mL of 50 % HF and 40 mL of deionized (DI) water for 5 min. After that, they were rinsed thoroughly with DI water and dried.


(A)

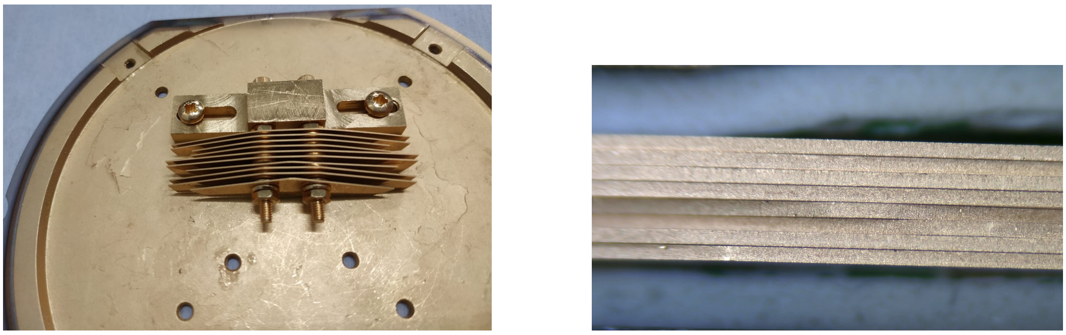
(B)

FIGURE 3.9: (A) View of the stacked butterfly wings with the optical microscope after the HF bath. (B) View of the stacked butterfly wings with the optical microscope after the second ultrasonic bath.

Figure 3.9a shows that the deposit was partially removed, and the quality of the waveguide has greatly improved. To get rid of the final imperfections, we used another ultrasonic bath with the same procedure as the first one. Figure 3.9b highlights the fact that the deposit was removed thanks to the second bath, the guide looks more homogeneous, which is confirmed by the profilometer. The average arithmetic roughness is estimated to be around 1.5 µm RMS, which is better than with the copper plates design.

Sputtering process

Finally, we deposited gold on the sides of butterfly wings. To do so, we placed the wings upwards on the holder shown in figure 3.10a. We used plastic spacers between the wings to prevent any damage. Finally, with a DP650 sputtering machine by Alliance concept, we deposited a 20 nm layer of titanium for adhesion and a 1000 nm layer of gold. The procedure, done one time, took approximately 40 min. The final waveguides are shown in figure 3.10b.



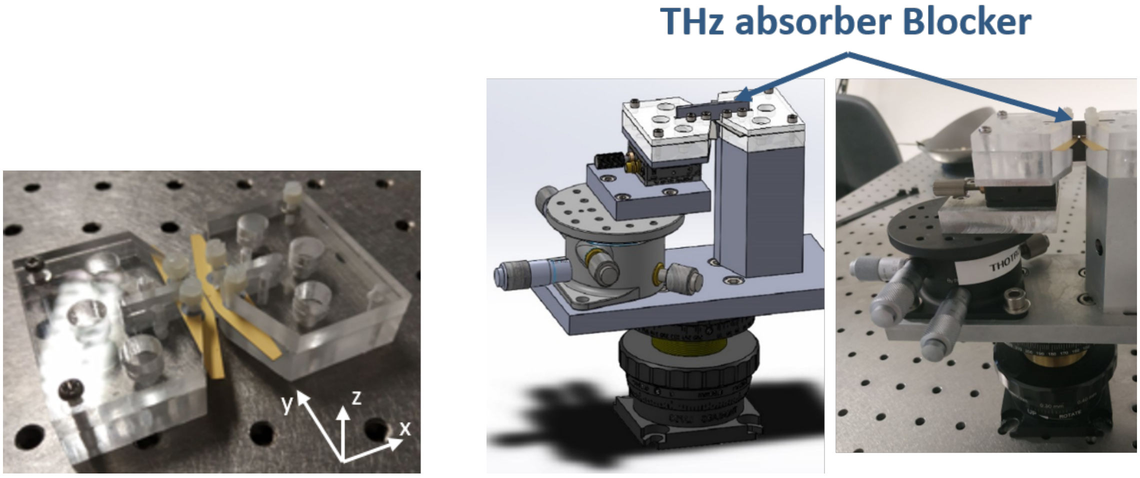
(A) (B)

FIGURE 3.10: (A) The wings of the butterfly device on their holder, attached to the the sputtering device holder. (B) View of the butterfly wings with the optical microscope after the sputtering process.

Now that the roughness has been reduced thanks to this fabrication process, the first point of the limitations list has been checked. The butterfly device has to be characterised to check on the other limitations.

3.3.2 Blocking the parasitic pulse

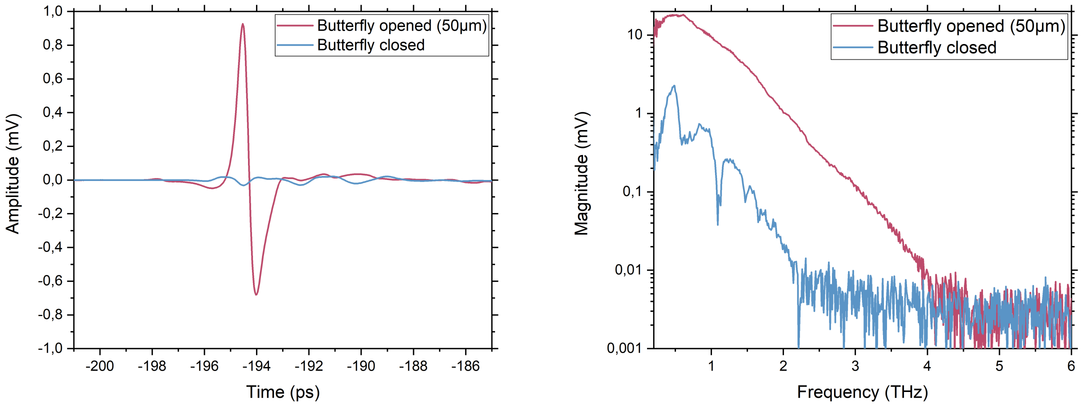
Since the wings are smaller now and more fragile, we used a protective plastic cover that is fixed to the same mount as for the first version (see in figure 3.11a). The THz-TDS setup for the characterisation is the same as for the first version. We found that the bandwidth was broadened, with now 4 THz instead of 2.5 THz before. However, the parasitic pulse is still present.



(A) (B)

FIGURE 3.11: (A) New version of the butterfly in its plastic holder. (B) THz absorber blocker above and below the butterfly.

To get rid of it, we decided to use an absorber. However, putting an absorber on the top of the butterfly was not enough, as we found that the parasitic pulse travelled really close



(A)

(B)

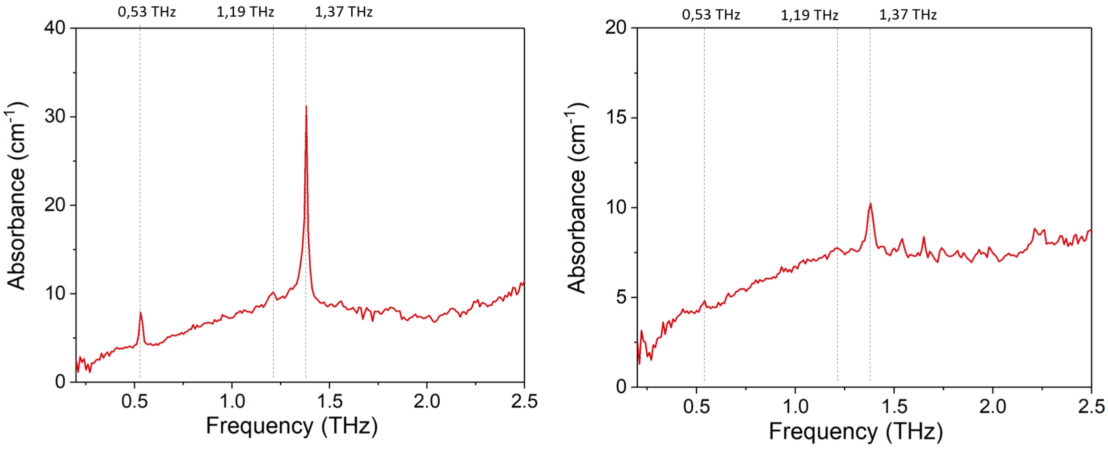
FIGURE 3.12: (A) Time traces measured when the butterfly device is opened to 50 µm versus when it is closed, after adding the absorbers. (B) Spectra retrieved when the butterfly device is opened to 50 µm versus when it is closed, after adding the absorbers.

to the waveguide portion, both above and below. Therefore, we adjusted the mount of the wings in order to be able to place an absorber both above and below the waveguide, such as displayed in figure 3.11b. This way, we only detect the pulse that has gone through the butterfly. Figure 3.12 displays the time traces obtained when the absorber has been added. Compared to the previous version, the amplitude of the parasitic pulse is highly reduced and will interfere less with the data. Moreover, the spectra showcase the broad bandwidth available even with a small width of 50 µm for the gap.

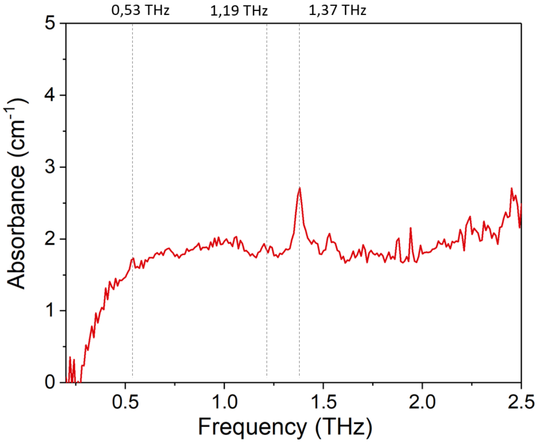
3.3.3 Test on α-lactose monohydrate and glutamine

Finally, to compare the results with the first version, we measured an α-lactose monohydrate sample. Three experiments were conducted, the first one on a 15 nL powder sample (a length of 1 mm was filled in the waveguide), the second one on a 5 nL powder sample (a length of 0.5 mm was filled in the waveguide), and a third one on a 3 nL liquid evaporated sample. In all cases, the gap was opened to 50 µm (adjusted with a calibrated 50 µm-thick Kapton sheet). We were able to reach smaller volumes of sample than with the previous version because the device is thinner, and because we are able to reach a smaller width of the waveguide. Finally, the presence of graduation on the wing along the waveguide allows us to better measure the quantity of sample inside it.The retrieved absorbance curves for the three samples are displayed in figure 3.13.

The absorption peaks of lactose are sharper with this new version, with more intensity (and less powder) and there are less small peaks due to scattering effects. For the 15 nL sample, we can distinguish the first three peaks at 530 GHz, 1.19 THz and 1.37 THz. Even if the 1.19 THz peak disappears in the noise for smaller volumes, the two other peaks are still visible (particularly the 1.37 THz peak), which demonstrates the sensitivity and broad bandwidth of the device.



(A) 15 nL α-lactose monohydrate powder sample. (B) 5 nL α-lactose monohydrate powder sample.



(C) 3 nL α-lactose monohydrate evaporated liquid sample.

FIGURE 3.13: Absorbance of three samples of α-lactose monohydrate.

To conclude, we managed to improve the spectral range to 4 THz, to remove the parasitic pulse and to reduce fabrication time and cost. Thanks to the improved roughness, we can now achieve a gap opening of 10 µm. Moreover, we were able to measure successfully the THz spectrum of a few-nL volume biosample. However, the new fabrication process did not help to improve the coupling of the device and we still need focusing lenses for the THz pulse to enter the gap.

3.3.4 Understanding the coupling issue

In order to understand how the device confines and couples light, we designed a quasi-optics experiment for the characterization of the near-field and far-field radiation patterns of the butterfly device based on a THz laser source and a THz camera. Quasi-optics antenna are specific because their size is a lot bigger than that of the wavelength range. The near-field and far-field definitions in this case are for a quasi-optics experiment, meaning that we want to image the THz field at the focal point ("near-field") and the THz field outside the butterfly ("far-field").

Setups of the microscopy and far-field experiments

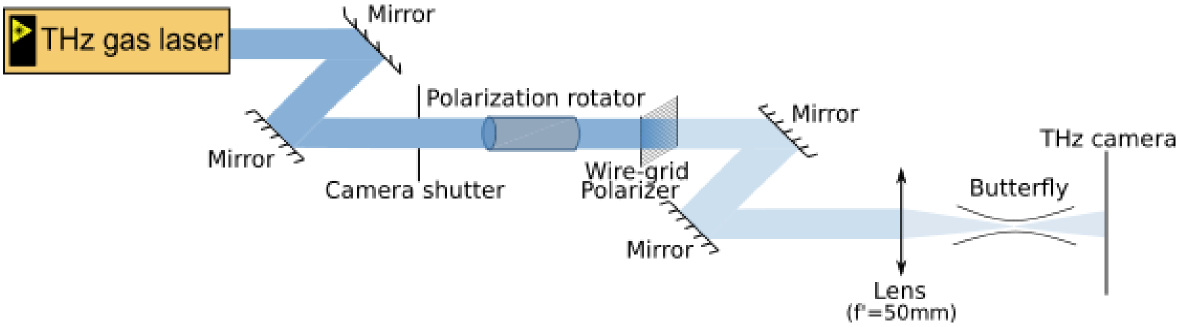


FIGURE 3.14: Schematic of the far-field experiment.

Figure 3.14 represents the schematic of the far-field experiment. We use a CO_2 pumped methanol gas laser emitting at 2.5 THz (which is around the middle of the bandwidth achieved in THz-TDS with the butterfly). The beam goes through a polarization rotator in order to change its polarization from vertical to horizontal, and then through a wire-grid polarizer used as an attenuator in order to avoid damage on the THz camera. Then, a TPX lens is used to focus the beam at the entrance of the waveguide and the THz camera is placed far enough from the butterfly to image the far field pattern but close enough to be able to image it considering the sensor size. The THz camera we use is the model TZCAM from i2S, it is sensitive from 0.1 THz to 5 THz, its pixel size is 50 µm × 50 µm. Its resolution is 320 × 240 pixels (the sensor size is 16 × 12 mm).

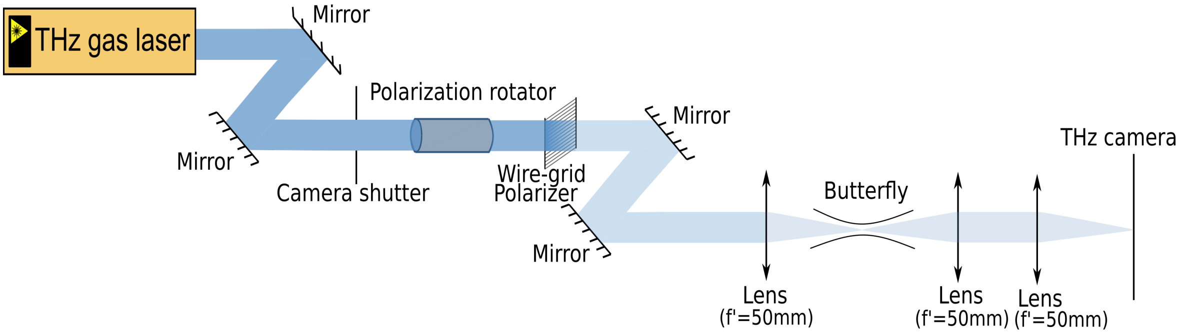
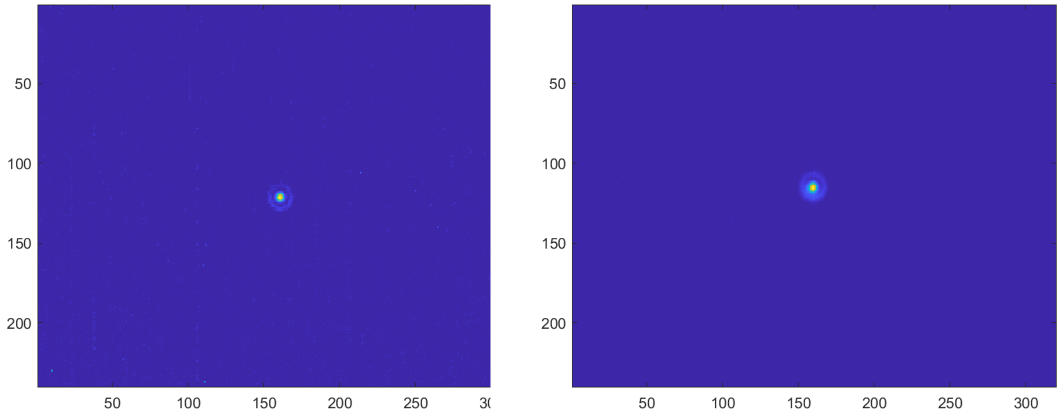


FIGURE 3.15: Schematic of the microscopy experiment.

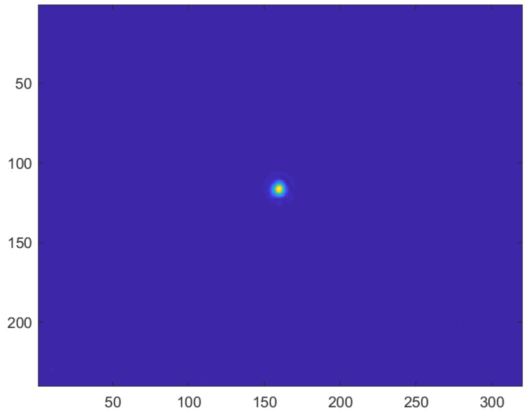
In addition to the far-field setup, we decided to add two lenses after the butterfly device, such as shown in figure 3.15, to design a microscopy setup. This will enable us to image the beam inside the butterfly and thus gain more knowledge on the way the light is confined inside it. Finally, in all experiments, the absorber blockers were used.

Results

Firstly, in order to verify the setup, we put a variable diaphragm in place of the butterfly device. With the setup from figure 3.15, we imaged the diaphragm for three openings: closed (diameter of 0.8 mm), half-opened (diameter of around 5 mm), and fully opened (diameter of 15.0 mm). The result is displayed in figure 3.16. As expected, the size of the spot on the camera increases as we open the diaphragm, and the intensity is higher at the center of the spot.



(A) diameter of 0.8 mm (B) diameter of around 5 mm



(C) diameter of 15.0 mm

FIGURE 3.16: Image recorded by the THz camera of a variable diaphragm with the microscopy setup.

Next, we removed the diaphragm, put the butterfly in place and adjusted the lenses in order for their focal planes to be at the entrance and exit of the waveguide. The gap between the wings of the butterfly was opened to 50 µm thanks to a calibrated Kapton sheet. We recorded the image shown in figure 3.17. There is a very small spot on the camera, showing a good confinement of the THz pulse. When we zoom in on this spot, we can distinguish the two sides of the waveguide, as it is oval-shaped, and straighter on the sides (the waveguide has a cross section of 50 × 330 μm^2).

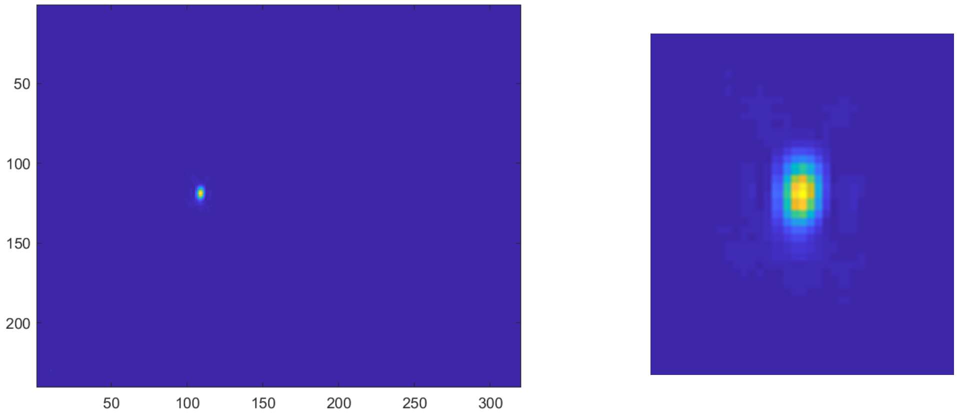
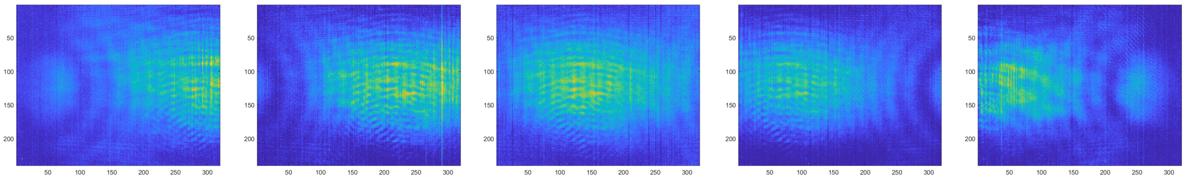


FIGURE 3.17: Image recorded by the THz camera of the beam inside the butterfly waveguide with the microscopy setup. On the right is a magnified image of the spot.

Since the microscopy setup worked well, we removed the two lenses before the THz camera in order to record the far-field. We had to move the camera really close to the exit of the tapered antenna in order to have a usable image, as the size of the exiting beam increases rapidly. Still, we were able to image the far field radiation pattern by taking several pictures while moving the camera from left to right. The result is displayed on figure 3.18.



(A) $x = -7$mm (B) $x = -3$mm (C) $x = 0$mm (D) $x = +3$mm (E) $x = +7$mm

FIGURE 3.18: Far-field radiation pattern of the butterfly device, the x axis is shown in figure 3.11a.

Figure 3.19 shows a reconstruction of the far-field radiation pattern from these 5 images (hence the slight differences in intensity). We see a principal beam which is oval with a larger divergence in the horizontal plane than in the vertical plane as can be expected from the diffraction (the waveguide is smaller in the horizontal direction than in the vertical one). When we compare this image to the one obtained by CST simulations (still, we have to remember that the simulation were not done on the whole device), we see some similarities. Indeed, they both have a main lobe in the middle and two lobes on the sides. However, the butterfly is less directive than quantified. Our hypothesis was that the THz pulse was guided by the tapered antenna. However, the diffraction angle (considering the output aperture) is smaller than the opening angle of the butterfly, which leads to the fact that the antenna cannot play its role. Hence, the THz pulse diverges. This explains why we had to

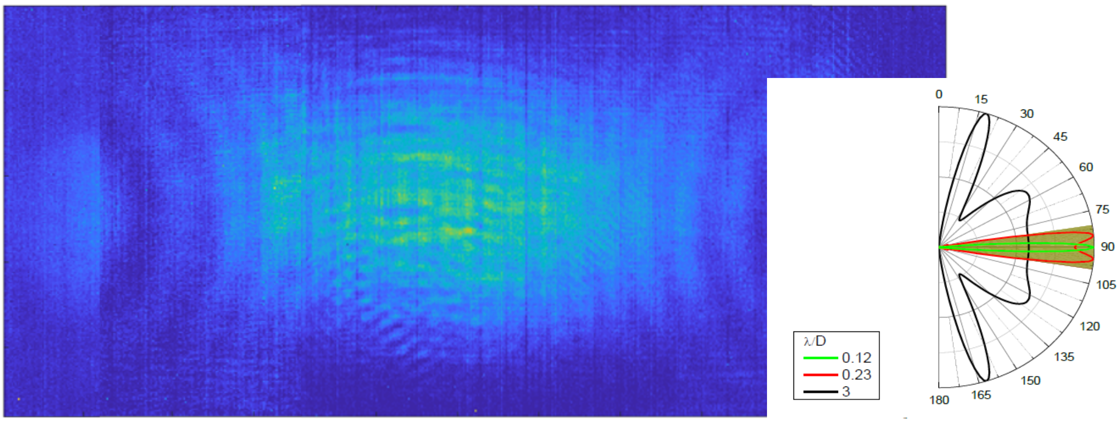


FIGURE 3.19: Reconstructed far-field radiation pattern of the butterfly on the left. On the right is the result of the CST simulation for comparison. The green curve corresponds to a frequency of 5 THz, the red curve to 2.6 THz, and the black curve to 0.2 THz.

be so close to the antenna to see the far field pattern. It also explains why the coupling to the far-field was different than expected and why we had to use focusing lenses. In fact, when one compares the opening angle of the lens (its half opening angle equal to around 20°) to the one of the butterfly (equal to 18.4°), they almost match. Moreover, we found experimentally that the most efficient lens was the one with the closest opening angle.

In order to resolve this problem, we have to modify the design of the tapered part of the device. The next design will consist in three parts: the same waveguide section in the middle, the same tapered antenna with the parametric equation at the end, and an additional tapered part between the two that follows a square-root curve ($f(x) = \sqrt{\lambda x}$). After that, we have two choices. Either we modify the design of the butterfly to have the same "escaping" angle (angle where the pulse is no longer guided) at all frequencies, but it means a different diffraction angle at each frequency, or we keep the same diffraction angle but with a different escaping angle. In the latter case, we have to reduce the opening angle of the butterfly to less than the diffraction angle in order to control the exiting of the THz pulse.

3.4 Conclusion

This chapter presented a THz light confinement device, called the butterfly. The first version showed great potential, with good confinement, large bandwidth, and sensitivity. However, it had several limitations: important roughness only allowing a minimum gap size of 30 µm maximum, a limited bandwidth (of around 2.5 THz), coupling issues, the presence of a parasitic pulse and an important cost and long fabrication time. Therefore, we changed the fabrication process. The performances of the second version were improved, with a broader bandwidth (around 4 THz) and improved roughness that allowed to close the gap to 10 µm. We also removed the parasitic pulse thanks to a THz absorber blocker. The test on α-lactose monohydrate demonstrated its usability on biosamples, and especially on small volumes of the order of the nL, which is the smaller volume reached compared to the state of the art [154]. Still, one limitation remained: the coupling issue. Indeed, the far-field experiments showed good light confinement, which explains the good experimental results, but lower directivity than expected. As a result, the butterfly design will need to be modified once again. The next steps, of which we hope to show the results during the thesis presentation, will be to test several devices with smaller opening angles. This will enable us to verify the hypothesis that the diffraction angle of the THz beam is smaller than the angle of the butterfly wings. A new design is also under work. It consists in adding a transitional part between the waveguide and the tapered curved antenna (both for the injection and extraction of the pulse). It will be presented during the defence of the thesis work. Another lead for the improvement of the device is to use elliptic 3D-printed waveplates [155] to inject and extract the beam of the device.

Chapter 4

Conclusion and perspectives

Investigating biological samples with THz-TDS brings up a lot of challenges. There has been a lot of enthusiasm for this technique at first, and a lot of biological samples were measured. However, researchers realized that they lacked knowledge on the basics and that they needed better tools for data analysis. In this thesis, we focused on what we think are the three main challenges to overcome: analysing biosamples as close to *in vivo* conditions as possible, taking into account their small size and the small volumes they are available in, and finally how to retrieve information from the THz-TDS data.

THz-TDS is a relatively new technique, as it has been around for about thirty years. There has been a few studies on biological samples, especially under a crystallized form, but there is not much knowledge on what kind of feature to expect when analysing them under conditions close to *in vivo* conditions. Therefore, we have to make sure that our data analysis process is sufficiently robust so that the features or parameters retrieved are entirely due to the presence of biological samples. To this end, we showed that time-domain fitting, in addition to being closer to experiments than frequency-domain fitting, gives also more reliable results. Moreover, taking experimental noise into account during the fitting process is of utmost importance so that we do not mistake noise artifacts for real data. We showed the accuracy of our fit@TDS software on a α-lactose monohydrate pellet sample and were able to retrieve the same main peaks as in the literature, and precisely fitting the refractive index up until 2.6 THz. The bandwidth limitation can be pushed further with a thinner pellet or a mixed one with polyethylene.

The method, from the experimental and the data analysis points of view, works well for studying biological samples under the powder form, pressed in pellets. This also means that they have to be available in large quantities. However, it was found that the structure of biological objects, such as proteins, cells or virus, can be altered by the conditions they are in. Hence we have to analyse them closest to *in vivo* conditions, which means investigating them in an aqueous solution. This is a great challenge because of the high absorption of liquid water in the terahertz range. Still, it could be done if we were able to characterize liquid water itself. However, if the experimental protocol to put in place is not straightforward (microfluidic cell to be able to analyze small amounts of liquid water, monitoring of the temperature) but widespread in the community, it is the data analysis part that is tricky. Indeed, there is no consensus on the permittivity model to use when fitting liquid water data. The models proposed for the terahertz range add several Debye terms, but they cannot be extended to higher frequencies. As a result, we found that there was a need for a new model, taking into account inertial effects in water and interactions between molecules, which the Debye model does not. The Titov model takes into account some of these effects (inertial and interactions between two polar groups) and we showed that, in addition to being extendable to higher frequencies, it also provided a better fit at terahertz frequencies. With this information, we showed that a modification of the H-bond network structure results in an absorption difference on the THz-TDS data and therefore a modification of the parameters of the model. We also characterized this absorption difference, meaning that an increase in absorption was linked to a breakage of bonds while a decrease in absorption was linked to a making of bonds. Finally, we applied our knowledge to the study of a protein in aqueous solution, the lysozyme. This data still needs to be fitted but is highly promising. However, it shown that the effect of the biological object was weak, it is not observable from the experimental curves, and we need further analysis to be able to determine whether THz-TDS is able to probe it or not.

The last challenge concerns the size of biological objects. Indeed, many of them are a lot smaller than the size of the THz wavelength, which is typically between 0.1 and 1 mm. To overcome this problem, we would need to analyze a large volume of biosamples but, most of the time, they are not available in large quantities as they are hard to produce. As a result, interactions between biosamples and the THz waves are reduced, making it hard to analyze samples with a volume smaller than a microlitre. As a result, we decided to confine the THz waves in order to increase interactions with the sample. To achieve this goal, we designed a device called the butterfly, composed of a thick slot line waveguide in the middle, and of two curved tapered antennas to inject and extract the THz pulse. The first version, made with copper plates, showed promising results but was limited in bandwidth and has too much roughness in the waveguide part, preventing the analysis of very small volumes of samples. Therefore, a new version was designed, made of silicon wafers sputtered with titanium and gold layers, which showed better performances and allowed us to analyze very small volumes. However, it is still difficult to couple efficiently the device to the THz-TDS, hindering it to reach its full capability. Still, this type of device allows us to overcome the challenge of analyzing small biosamples, and small volumes of any type of sample.

To conclude, this work provided an accurate data analysis method and tool, with improvements on taking experimental noise into account, improving parameter extraction, and the ability to compare permittivity models. This work also presented an overview of the research done on liquid water and showed the importance of monitoring rigorously the temperature of the sample during the experiment, as well as making sure that the permittivity models used for the fit can be extended to higher frequencies. Also, we demonstrated how the modifications of the bonds inside a H-bond network influenced the THz-TDS data and showed that the concentration of an ionic solution had an impact on the absorption. Moreover, if we want to be able to analyse biosamples in liquid solutions, we have to take into account all of the above and realize an experiment with a very low uncertainty to be able to observe, or extract, information on the biosample itself. Finally, we explained that the confinement of the THz pulse enhanced interactions with the sample, which is made possible by our butterfly device. We were able to analyse volumes of samples of the order of the nanolitre, which is a first in the community.

As for the future, THz-TDS is multidisciplinary. As we have seen in this thesis, there needs to be collaborations to understand all aspects: biology, biochemistry, experimental physics, theoretical physics, and signal processing. The challenges of THz-TDS are great, and overcoming them will require more work on the foundations, such as the study of liquid water and simpler biosamples, like the lysozyme studied here. Studies on complex proteins, the methylation of DNA or cancer research are enticing but they seem out of reach for now considering the results we obtained. The three chapters of this work will need to be carried on, but they all lay the ground for future experiments. There is already work done on the following of chapter 1 that will help reducing experimental noise. Indeed, part of the noise present in the data is due to the delay between each measurement, as we already pointed out. Therefore, we decided to accumulate the THz-TDS data ourselves instead of doing that with the software provided. Using a python routine, we now save a large number of time traces separately, readjust the delay, and then proceed to taking the mean. This step added to the high pass filter still necessary for values under 200 GHz produces data that are a lot easier to analyse afterwards. Moreover, there is ongoing talks with theoreticians in order to improve the permittivity model of liquid water and the theory behind it to understand the molecular dynamics better. The ionic solutions will also need to be remeasured so that the temperature monitoring matches the one of the liquid water experiments. Finally, by fixing the problem of coupling with the butterfly device and some technological improvements (such as working on the thickness), we will be able to reach even smaller volumes of sample, one objective being 0.1 nL, the volume of a human cell.

This thesis manuscript is the result of three years of work, done under covid-19 conditions. Inevitably, experimental work has been impacted by lockdowns, therefore the focus has been shifted to data analysis and leaves some questions unsolved, notably concerning the analysis of very small volumes of samples. Still, we think that this work will be a great foundation for future analysis of biological samples and shows the potential of the technique.

Bibliography

[1] Erik Bründermann, Heinz-Wilhelm Hübers, and Maurice F. Kimmitt. “Optical Principles at Terahertz Frequencies”. In: *Springer Series in Optical Sciences*. Springer Berlin Heidelberg, 2011, pp. 23–49. DOI: `10.1007/978-3-642-02592-1_2`.

[2] Craig Kulesa. “Terahertz Spectroscopy for Astronomy: From Comets to Cosmology”. In: *IEEE Transactions on Terahertz Science and Technology* 1.1 (2011), pp. 232–240. DOI: `10.1109/tthz.2011.2159648`.

[3] Alwyn Wootten. “Atacama Large Millimeter Array (ALMA)”. In: *Large Ground-based Telescopes*. Ed. by Jacobus M. Oschmann and Larry M. Stepp. SPIE, 2003. DOI: `10.1117/12.456705`.

[4] Takeshi Yasui et al. “Terahertz paintmeter for noncontact monitoring of thickness and drying progress in paint film”. In: *Applied Optics* 44.32 (2005), p. 6849. DOI: `10.1364/ao.44.006849`.

[5] J. B. Perraud et al. “Terahertz imaging and tomography as efficient instruments for testing polymer additive manufacturing objects”. In: *Applied Optics* 55.13 (2016), p. 3462. DOI: `10.1364/ao.55.003462`.

[6] J. E. Bjarnason et al. “Millimeter-wave, terahertz, and mid-infrared transmission through common clothing”. In: *Applied Physics Letters* 85.4 (2004), pp. 519–521. DOI: `10.1063/1.1771814`.

[7] Roger Appleby and H. Bruce Wallace. “Standoff Detection of Weapons and Contraband in the 100 GHz to 1 THz Region”. In: *IEEE Transactions on Antennas and Propagation* 55.11 (2007), pp. 2944–2956. DOI: `10.1109/tap.2007.908543`.

[8] Tadao Nagatsuma, Guillaume Ducournau, and Cyril C. Renaud. “Advances in terahertz communications accelerated by photonics”. In: *Nature Photonics* 10.6 (2016), pp. 371–379. DOI: `10.1038/nphoton.2016.65`.

[9] Guillaume Ducournau. “Terahertz wireless communications”. In: *Photoniques* (Apr. 2017), pp. 32–37. DOI: `10.1051/photon/2017s232`.

[10] A.G Markelz, A Roitberg, and E.J Heilweil. “Pulsed terahertz spectroscopy of DNA, bovine serum albumin and collagen between 0.1 and 2.0 THz”. In: *Chemical Physics Letters* 320.1-2 (2000), pp. 42–48. DOI: `10.1016/s0009-2614(00)00227-x`.

[11] A.G. Davies P.C. Upadhya Y.C. Shen and E.H Linfield. “Terahertz Time-Domain Spectroscopy of Glucoseand Uric Acid”. In: *Journal of Biological Physics* 29 (2003), pp. 117–121.

[12] Philip F. Taday. “Applications of terahertz spectroscopy to pharmaceutical sciences”. In: *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 362.1815 (2003). Ed. by A. Giles Davies, Edmund H. Linfield, and Michael Pepper, pp. 351–364. DOI: `10.1098/rsta.2003.1321`.

[13] Heinrich Rubens and E. F. Nichols. "Heat Rays Of Great Wave Length". In: *Physical Review (Series I)* 4.4 (1897), pp. 314–323. DOI: `10.1103/physrevseriesi.4.314`.

[14] Max Planck. *Vorlesungen über die Theorie der Wärmestrahlung*. 1913.

[15] Heinrich Rubens. *Gittermessungen im langwelligen Spektrum*. Sitzungsberichteder Preussischen Akademie derWissenschaften, 1921.

[16] Martin van Exter, Ch. Fattinger, and D. Grischkowsky. "Terahertz time-domain spectroscopy of water vapor". In: *Optics Letters* 14.20 (1989), p. 1128. DOI: `10.1364/ol.14.001128`.

[17] C. K. N. Patel. "Continuous-Wave Laser Action on Vibrational-Rotational Transitions of CO2". In: *Physical Review* 136.5A (1964), pp. 1187–1193. DOI: `10.1103/physrev.136.a1187`.

[18] R. H. Dicke et al. "Cosmic Black-Body Radiation." In: *The Astrophysical Journal* 142 (1965), p. 414. DOI: `10.1086/148306`.

[19] A. M. Nicolson. "Broad-Band Microwave Transmission Characteristics from a Single Measurement of the Transient Response". In: *IEEE Transactions on Instrumentation and Measurement* 17.4 (1968), pp. 395–402. DOI: `10.1109/tim.1968.4313741`.

[20] P. R. Smith, D. H. Auston, and M. C. Nuss. "Subpicosecond photoconducting dipole antennas". In: *IEEE Journal of Quantum Electronics* 24.2 (1988), pp. 255–260. DOI: `10.1109/3.121`.

[21] Ch. Fattinger and D. Grischkowsky. "Terahertz beams". In: *Applied Physics Letters* 54.6 (1989), pp. 490–492. DOI: `10.1063/1.100958`.

[22] J. Valdmanis, G. Mourou, and C. Gabel. "Subpicosecond electrical sampling". In: *IEEE Journal of Quantum Electronics* 19.4 (1983), pp. 664–667. DOI: `10.1109/jqe.1983.1071915`.

[23] D. H. Auston and K. P. Cheung. "Coherent time-domain far-infrared spectroscopy". In: *Journal of the Optical Society of America B* 2.4 (1985), p. 606. DOI: `10.1364/josab.2.000606`.

[24] T. Seifert et al. "Efficient metallic spintronic emitters of ultrabroadband terahertz radiation". In: *Nature Photonics* 10.7 (2016), pp. 483–488. DOI: `10.1038/nphoton.2016.91`.

[25] Yun-Shik Lee. *Principles of Terahertz Science and Technology*. Springer-Verlag GmbH, 2009. ISBN: 9780387095400. URL: `https://www.ebook.de/de/product/12469854/yun_shik_lee_principles_of_terahertz_science_and_technology.html`.

[26] Maximilien Billet. "Photodétecteurs rapides à la longueur d'onde de 1550 nm pour la génération et la détection d'ondes sub-THz et THz". PhD thesis. University of Lille, 2018.

[27] Arno Rehn et al. "Periodic sampling errors in terahertz time-domain measurements". In: *Optics Express* 25.6 (2017), p. 6712. DOI: `10.1364/oe.25.006712`.

[28] Vincent P. Wallace Jean-Louis Coutaz Frederic Garet. *Terahertz Spectroscopy : an Introductory Textbook*. Singapore: Pan Stanford Publishing, 2018. ISBN: 9789814774567.

[29] Tae-In Jeon et al. "Electrical characterization of conducting polypyrrole by THz time-domain spectroscopy". In: *Applied Physics Letters* 77.16 (2000), pp. 2452–2454. DOI: 10.1063/1.1319188.

[30] Ming Li; J. Fortin; J.Y. Kim; G. Fox; F. Chu; T. Davenport; Toh-Ming Lu; Xi-Cheng Zhang. "Dielectric constant measurement of thin films using goniometric terahertz time-domain spectroscopy". In: *IEEE Journal of Selected Topics in Quantum Electronics* 7.4 (2001), pp. 624–629.

[31] R. D. Averitt et al. "Ultrafast Conductivity Dynamics in Colossal Magnetoresistance Manganites". In: *Physical Review Letters* 87.1 (2001). DOI: 10.1103/physrevlett.87.017401.

[32] Tae-In Jeon and D. Grischkowsky. "Observation of a Cole–Davidson type complex conductivity in the limit of very low carrier densities in doped silicon". In: *Applied Physics Letters* 72.18 (1998), pp. 2259–2261. DOI: 10.1063/1.121271.

[33] P. Uhd Jepsen et al. "Ultrafast carrier trapping in microcrystalline silicon observed in optical pump–terahertz probe measurements". In: *Applied Physics Letters* 79.9 (2001), pp. 1291–1293. DOI: 10.1063/1.1394953.

[34] G. C. Cho et al. "Optical phonon dynamics of GaAs studied with time-resolved terahertz spectroscopy". In: *Optics Letters* 25.21 (2000), p. 1609. DOI: 10.1364/ol.25.001609.

[35] John F. Federici et al. "THz imaging and sensing for security applications—explosives, weapons and drugs". In: *Semiconductor Science and Technology* 20.7 (2005), pp. 266–280. DOI: 10.1088/0268-1242/20/7/018.

[36] H. Zhao G. J. Bastiaans J. Chen Y. Chen and X. C. Zhang. "Absorption coefficients of selected explosives and related compounds in the range of 0.1–2.8 THz". In: *Optics Express* 15.19 (2007), pp. 12060–12067.

[37] Z. Chen K. Kang Y. Wang Z. Zhao and B. Feng. "Terahertz absorbance spectrum fitting method for quantitative detection of concealed contraband". In: *Journal of Applied Physics* 102 (2007), p. 113108.

[38] Christian Jördens. "Detection of foreign bodies in chocolate with pulsed terahertz spectroscopy". In: *Optical Engineering* 47.3 (2008), p. 037003. DOI: 10.1117/1.2896597.

[39] M. Herrmann et al. "Terahertz imaging of objects in powders". In: *IEE Proceedings - Optoelectronics* 149.3 (2002), pp. 116–120. DOI: 10.1049/ip-opt:20020185.

[40] Cyril Guers. "Optimisation des matériaux cellulosiques pour des applications en radiofréquences et térahertz". PhD thesis. Université Grenoble Alpes (ComUE), 2019.

[41] D. Banerjee et al. "Diagnosing water content in paper by terahertz radiation". In: *Optics Express* 16.12 (2008), p. 9060. DOI: 10.1364/oe.16.009060.

[42] Payam Mousavi et al. "Simultaneous composition and thickness measurement of paper using terahertz time-domain spectroscopy". In: *Applied Optics* 48.33 (2009), p. 6541. DOI: 10.1364/ao.48.006541.

[43] U. Klotzbach E. Beyer S. Winnerl M. Helm F. Rutz C. Jordens M. Koch W. Kohler M. Panzner and H. Leitner. "Non-Destructive Investigation of Paintings with THz-Radiation". In: *Proceedings of the European Conference of Non-Destructive Testing*. 2006.

[44] Shin'ichiro Hayashi Kaori Fukunaga Yuichi Ogawa and Iwao Hosako. "Application of terahertz spectroscopy for character recognition in a medieval manuscript". In: *IEICE Electronics Express* 5.7 (2008), pp. 223–228.

[45] J.B. Jackson et al. "Terahertz imaging for non-destructive evaluation of mural paintings". In: *Optics Communications* 281.4 (2008), pp. 527–532. DOI: `10.1016/j.optcom.2007.10.049`.

[46] Aurèle J. L. Adam et al. "TeraHertz imaging of hidden paint layers on canvas". In: *Optics Express* 17.5 (2009), p. 3407. DOI: `10.1364/oe.17.003407`.

[47] B. B. Hu and M. C. Nuss. "Imaging with terahertz waves". In: *Optics Letters* 20.16 (1995), p. 1716. DOI: `10.1364/ol.20.001716`.

[48] D. M. Mittleman et al. "Noncontact semiconductor wafer characterization with the terahertz Hall effect". In: *Applied Physics Letters* 71.1 (1997), pp. 16–18. DOI: `10.1063/1.119456`.

[49] Zhiping Jiang and Xi-Cheng Zhang. "Terahertz imaging via electrooptic effect". In: *IEEE Transactions on Microwave Theory and Techniques* 47.12 (1999), pp. 2644–2650. DOI: `10.1109/22.809019`.

[50] D.M. Mittleman et al. "Recent advances in terahertz imaging". In: *Applied Physics B: Lasers and Optics* 68.6 (1999), pp. 1085–1094. DOI: `10.1007/s003400050750`.

[51] Franz Schuster et al. "A broadband THz imager in a low-cost CMOS technology". In: *2011 IEEE International Solid-State Circuits Conference*. IEEE, 2011. DOI: `10.1109/isscc.2011.5746211`.

[52] Yue Bai Theo Hannotte Melanie Lavancier Djamila Hourlier Goedele Roos Sergey Mitryukovskiy Danny E. P. Vanpoucke and Romain Peretti. "On the influence of water on THz vibrational spectral features of molecular crystals". In: *PCCP (submitted)* (2021).

[53] Marie-Claire Bellissent-Funel et al. "Water Determines the Structure and Dynamics of Proteins". In: *Chemical Reviews* 116.13 (2016), pp. 7673–7697. DOI: `10.1021/acs.chemrev.5b00664`.

[54] Y. Watanabe et al. "Component analysis of chemical mixtures using terahertz spectroscopic imaging". In: *Optics Communications* 234.1-6 (2004), pp. 125–129. DOI: `10.1016/j.optcom.2004.02.032`.

[55] P. F. Taday et al. "Using terahertz pulse spectroscopy to study the crystalline structure of a drug: A case study of the polymorphs of ranitidine hydrochloride". In: *Journal of Pharmaceutical Sciences* 92.4 (2003), pp. 831–838. DOI: `10.1002/jps.10358`.

[56] Clare J. Strachan et al. "Using Terahertz Pulsed Spectroscopy to Quantify Pharmaceutical Polymorphism and Crystallinity". In: *Journal of Pharmaceutical Sciences* 94.4 (2005), pp. 837–846. DOI: `10.1002/jps.20281`.

[57] G. Albiser, A. Lamiri, and S. Premilat. "The A–B transition: temperature and base composition effects on hydration of DNA". In: *International Journal of Biological Macromolecules* 28.3 (2001), pp. 199–203. DOI: `10.1016/s0141-8130(00)00160-4`.

[58] H.H. Mitchell et al. "The Chemical Composition Of The Adult Human Body And Its Bearing On The Biochemistry Of Growth". In: *Journal of Biological Chemistry* 158.3 (1945), pp. 625–637. DOI: `10.1016/s0021-9258(19)51339-4`.

[59] U. Heugen et al. “Solute-induced retardation of water dynamics probed directly by terahertz spectroscopy”. In: *Proceedings of the National Academy of Sciences* 103.33 (2006), pp. 12301–12306. DOI: 10.1073/pnas.0604897103.

[60] Gerhard Schwaab, Federico Sebastiani, and Martina Havenith. “Ion Hydration and Ion Pairing as Probed by THz Spectroscopy”. In: *Angewandte Chemie International Edition* 58.10 (2019), pp. 3000–3013. DOI: 10.1002/anie.201805261.

[61] D. L. Woolard et al. “Millimeter Wave-induced Vibrational Modes in DNA as a Possible Alternative to Animal Tests to Probe for Carcinogenic Mutations”. In: *Journal of Applied Toxicology* 17.4 (1997), pp. 243–246. DOI: 10.1002/(sici)1099-1263(199707)17:4<243::aid-jat436>3.0.co;2-6.

[62] Hwayeong Cheon et al. “Effective demethylation of melanoma cells using terahertz radiation”. In: *Biomedical Optics Express* 10.10 (2019), p. 4931. DOI: 10.1364/boe.10.004931.

[63] E. Castro-Camus and M.B. Johnston. “Conformational changes of photoactive yellow protein monitored by terahertz spectroscopy”. In: *Chemical Physics Letters* 455.4-6 (2008), pp. 289–292. DOI: 10.1016/j.cplett.2008.02.084.

[64] Robert J. Falconer and Andrea G. Markelz. “Terahertz Spectroscopic Analysis of Peptides and Proteins”. In: *Journal of Infrared, Millimeter, and Terahertz Waves* 33.10 (2012), pp. 973–988. DOI: 10.1007/s10762-012-9915-9.

[65] J. Jeff Schwegman, John F. Carpenter, and Steven L. Nail. “Infrared microscopy for in situ measurement of protein secondary structure during freezing and freeze-drying”. In: *Journal of Pharmaceutical Sciences* 96.1 (2007), pp. 179–195. DOI: 10.1002/jps.20630.

[66] N.n. Zinov’ev S.homer-vanniasinkam R.e.miles J.m. Chamberlain A.j. Fitzgerald E.berry and M.a.smith. “Catalogue of Human Tissue Optical Properties at Terahertz Frequencies”. In: *Journal of Biological Physics* 129 (2003), pp. 123–128.

[67] William T. Coffey. *The Langevin Equation*. World Scientific Publishing Company, July 2012. 854 pp. ISBN: 9814355666. URL: https://www.ebook.de/de/product/15424804/william_t_coffey_the_langevin_equation.html.

[68] Peter Uhd Jepsen. “Phase Retrieval in Terahertz Time-Domain Measurements: a “how to” Tutorial”. In: *Journal of Infrared, Millimeter, and Terahertz Waves* 40.4 (2019), pp. 395–411. DOI: 10.1007/s10762-019-00578-0.

[69] Jeffrey C. Lagarias et al. “Convergence Properties of the Nelder–Mead Simplex Method in Low Dimensions”. In: *SIAM Journal on Optimization* 9.1 (1998), pp. 112–147. DOI: 10.1137/s1052623496303470.

[70] William C. Davidon. “Variable Metric Method for Minimization”. In: *SIAM Journal on Optimization* 1.1 (1991), pp. 1–17. DOI: 10.1137/0801001.

[71] Joseph-Frédéric Bonnans et al. *Numerical Optimization*. Springer-Verlag GmbH, Oct. 2006. 494 pp. ISBN: 9783540354475. URL: https://www.ebook.de/de/product/19292732/joseph_frederic_bonnans_jean_charles_gilbert_claude_lemarechal_claudia_a_sagastizabal_numerical_optimization.html.

[72] Maxime Bernier et al. “Accurate Characterization of Resonant Samples in the Terahertz Regime Through a Technique Combining Time-Domain Spectroscopy and Kramers–Kronig Analysis”. In: *IEEE Transactions on Terahertz Science and Technology* 6.3 (2016), pp. 442–450. DOI: `10.1109/tthz.2016.2535244`.

[73] Ioachim Pupeza, Rafal Wilk, and Martin Koch. “Highly accurate optical material parameter determination with THz time-domain spectroscopy”. In: *Optics Express* 15.7 (2007), p. 4335. DOI: `10.1364/oe.15.004335`.

[74] Osman S Ahmed et al. “Efficient Optimization Approach for Accurate Parameter Extraction With Terahertz Time-Domain Spectroscopy”. In: *Journal of Lightwave Technology* 28.11 (2010), pp. 1685–1692. DOI: `10.1109/jlt.2010.2047936`.

[75] Zhi Li et al. “Linewidth Extraction From the THz Absorption Spectra Using a Modified Lorentz Model”. In: *Journal of Infrared, Millimeter, and Terahertz Waves* 34.10 (2013), pp. 617–626. DOI: `10.1007/s10762-013-0007-2`.

[76] Roberto Cota et al. “Evidence for Reduced Hydrogen-Bond Cooperativity in Ionic Solvation Shells from Isotope-Dependent Dielectric Relaxation”. In: *Physical Review Letters* 120.21 (2018), p. 216001. DOI: `10.1103/physrevlett.120.216001`.

[77] Yuko UENO et al. “Quantitative Analysis of Amino Acids in Dietary Supplements Using Terahertz Time-domain Spectroscopy”. In: *Analytical Sciences* 27.4 (2011), p. 351. DOI: `10.2116/analsci.27.351`.

[78] Masaaki Tsubouchi Masato Kondoh Yasuhiro Ohshima. “Ion effects on the structure of water studied by terahertz time-domain spectroscopy”. In: *Chemical Physics Letters* 591 (2014), pp. 317–322.

[79] Matthias Krüger et al. “Uncertainty and Ambiguity in Terahertz Parameter Extraction and Data Analysis”. In: *Journal of Infrared, Millimeter, and Terahertz Waves* 32.5 (2010), pp. 699–715. DOI: `10.1007/s10762-010-9669-1`.

[80] Oleksandr Sushko et al. “Revised metrology for enhanced accuracy in complex optical constant determination by THz-time-domain spectrometry”. In: *Journal of the Optical Society of America A* 30.5 (2013), p. 979. DOI: `10.1364/josaa.30.000979`.

[81] Keiichiro Shiraga et al. “Broadband dielectric spectroscopy of glucose aqueous solution: Analysis of the hydration state and the hydrogen bond network”. In: *The Journal of Chemical Physics* 142.23 (2015), p. 234504. DOI: `10.1063/1.4922482`.

[82] R. Buchner, J. Barthel, and J. Stauber. “The dielectric relaxation of water between 0°C and 35°C”. In: *Chemical Physics Letters* 306.1-2 (1999), pp. 57–63. DOI: `10.1016/s0009-2614(99)00455-8`.

[83] Ruoyu Li et al. “Mesoscopic Structuring and Dynamics of Alcohol/Water Solutions Probed by Terahertz Time-Domain Spectroscopy and Pulsed Field Gradient Nuclear Magnetic Resonance”. In: *The Journal of Physical Chemistry B* 118.34 (2014), pp. 10156–10166. DOI: `10.1021/jp502799x`.

[84] J. T. Kindt and C. A. Schmuttenmaer. “Far-Infrared Dielectric Properties of Polar Liquids Probed by Femtosecond Terahertz Pulse Spectroscopy”. In: *The Journal of Physical Chemistry* 100.24 (1996), pp. 10373–10379. DOI: `10.1021/jp960141g`.

[85] J.K. Vij, D.R.J. Simpson, and O.E. Panarina. "Far infrared spectroscopy of water at different temperatures: GHz to THz dielectric spectroscopy of water". In: *Journal of Molecular Liquids* 112.3 (2004), pp. 125–135. DOI: `10.1016/j.molliq.2003.12.014`.

[86] Timothy D. Dorney, Richard G. Baraniuk, and Daniel M. Mittleman. "Material parameter estimation with terahertz time-domain spectroscopy". In: *Journal of the Optical Society of America A* 18.7 (2001), p. 1562. DOI: `10.1364/josaa.18.001562`.

[87] Osman S. Ahmed et al. "Efficient material parameters estimation with terahertz time-domain spectroscopy". In: *Terahertz Technology and Applications IV*. Ed. by Laurence P. Sadwick and Créidhe M. M. O'Sullivan. SPIE, 2011. DOI: `10.1117/12.873336`.

[88] Wanhua Zhu et al. "The calculation of dielectric dispersive models in THz range with GA". In: *Microwave and Optical Technology Letters* 49.10 (2007), pp. 2540–2545. DOI: `10.1002/mop.22776`.

[89] Seckin Sahin, Niru K. Nahar, and Kubilay Sertel. "Dielectric Properties of Low-Loss Polymers for mmW and THz Applications". In: *Journal of Infrared, Millimeter, and Terahertz Waves* 40.5 (2019), pp. 557–573. DOI: `10.1007/s10762-019-00584-2`.

[90] Amerigo Beneduci. "Which is the effective time scale of the fast Debye relaxation process in water?" In: *Journal of Molecular Liquids* 138.1-3 (2008), pp. 55–60. DOI: `10.1016/j.molliq.2007.07.003`.

[91] Keiichiro Shiraga et al. "Characterization of the hydrogen-bond network of water around sucrose and trehalose: Microwave and terahertz spectroscopic study". In: *The Journal of Chemical Physics* 146.10 (2017), p. 105102. DOI: `10.1063/1.4978232`.

[92] Daniel C. Elton. "The origin of the Debye relaxation in liquid water and fitting the high frequency excess response". In: *Physical Chemistry Chemical Physics* 19.28 (2017), pp. 18739–18749. DOI: `10.1039/c7cp02884a`.

[93] Anthony J. Fitzgerald, Emma Pickwell-MacPherson, and Vincent P. Wallace. "Use of Finite Difference Time Domain Simulations and Debye Theory for Modelling the Terahertz Reflection Response of Normal and Tumour Breast Tissue". In: *PLoS ONE* 9.7 (2014). Ed. by Irene Georgakoudi, e99291. DOI: `10.1371/journal.pone.0099291`.

[94] Romain Peretti et al. "THz-TDS Time-Trace Analysis for the Extraction of Material and Metamaterial Parameters". In: *IEEE Transactions on Terahertz Science and Technology* 9.2 (2019), pp. 136–149. DOI: `10.1109/tthz.2018.2889227`.

[95] Ruben E. Perez, Peter W. Jansen, and Joaquim R. R. A. Martins. "pyOpt: a Python-based object-oriented framework for nonlinear constrained optimization". In: *Structural and Multidisciplinary Optimization* 45.1 (2011), pp. 101–118. DOI: `10.1007/s00158-011-0666-3`.

[96] P.W. Jansen and R.E. Perez. "Constrained structural design optimization via a parallel augmented Lagrangian particle swarm optimization approach". In: *Computers & Structures* 89.13-14 (2011), pp. 1352–1366. DOI: `10.1016/j.compstruc.2011.03.011`.

[97] Y Xiang et al. "Generalized simulated annealing algorithm and its application to the Thomson model". In: *Physics Letters A* 233.3 (1997), pp. 216–220. DOI: `10.1016/s0375-9601(97)00474-x`.

[98] Laleh Mohtashemi et al. “Maximum-likelihood parameter estimation in terahertz time-domain spectroscopy”. In: *Optics Express* 29.4 (2021), p. 4912. DOI: `10.1364/oe.417724`.

[99] Kai-Erik Peiponen, Axel Zeitler, and Makoto Kuwata-Gonokami. *Terahertz Spectroscopy and Imaging*. Springer Berlin Heidelberg, 2013. DOI: `10.1007/978-3-642-29564-5`.

[100] Sergey Mitryukovskiy et al. “Shining the Light to Terahertz Spectroscopy of nL-Volume Biological Samples”. In: *Conference on Lasers and Electro-Optics*. OSA, 2019. DOI: `10.1364/cleo_at.2019.atu3k.6`.

[101] Lingbo Qiao et al. “Identification and quantitative analysis of chemical compounds based on multiscale linear fitting of terahertz spectra”. In: *Optical Engineering* 53.7 (2014), p. 074102. DOI: `10.1117/1.oe.53.7.074102`.

[102] Graham Clydesdale et al. “Modeling the Crystal Morphology of α-lactose Monohydrate”. In: *Journal of Pharmaceutical Sciences* 86.1 (1997), pp. 135–141. DOI: `10.1021/js950496w`.

[103] Andrew Szilagyi. “Adult Lactose Digestion Status and Effects on Disease”. In: *Canadian Journal of Gastroenterology and Hepatology* 29.3 (2015), pp. 149–156. DOI: `10.1155/2015/904686`.

[104] John A. Lucey, Don Otter, and David S. Horne. “A 100-Year Review: Progress on the chemistry of milk and its components”. In: *Journal of Dairy Science* 100.12 (2017), pp. 9916–9932. DOI: `10.3168/jds.2017-13250`.

[105] Peter Banks, Luke burgess, and Michael Ruggiero. “The Necessity of Periodic Boundary Conditions for the Accurate Calculation of Crystalline Terahertz Spectra”. In: *ChemRxiv. Cambridge: Cambridge Open Engage* (2020). DOI: `10.26434/chemrxiv.13473063.v1`.

[106] Martin F. Chaplin. *Structure and Properties of Water in its Various States*. Ed. by P. Maurice. Encyclopedia of Water, 2019, pp. 1–19. DOI: `10.1002/9781119300762.wsts0002`.

[107] Ildikó Pethes and László Pusztai. “Reverse Monte Carlo investigations concerning recent isotopic substitution neutron diffraction data on liquid water”. In: *Journal of Molecular Liquids* 212 (2015), pp. 111–116. DOI: `10.1016/j.molliq.2015.08.050`.

[108] L.M. Maestro et al. “On the existence of two states in liquid water: impact on biological and nanoscopic systems”. In: *International Journal of Nanotechnology* 13.8/9 (2016), p. 667. DOI: `10.1504/ijnt.2016.079670`.

[109] A. H. Narten, M. D. Danford, and H. A. Levy. “X-ray diffraction study of liquid water in the temperature range 4–200°C”. In: *Discuss. Faraday Soc.* 43.0 (1967), pp. 97–107. DOI: `10.1039/df9674300097`.

[110] N. Q. Vinh et al. “High-precision gigahertz-to-terahertz spectroscopy of aqueous salt solutions as a probe of the femtosecond-to-picosecond dynamics of liquid water”. In: *The Journal of Chemical Physics* 142.16 (2015), p. 164502. DOI: `10.1063/1.4918708`.

[111] M M Nazarov, O P Cherkasova, and A P Shkurinov. “Study of the dielectric function of aqueous solutions of glucose and albumin by THz time-domain spectroscopy”. In: *Quantum Electronics* 46.6 (2016), pp. 488–495. DOI: `10.1070/qel16107`.

[112] A. De Ninno et al. “Dielectric permittivity of aqueous solutions of electrolytes probed by THz time-domain and FTIR spectroscopy”. In: *Physics Letters A* 384.34 (2020), p. 126865. DOI: `10.1016/j.physleta.2020.126865`.

[113] Nikita Penkov et al. “Terahertz Spectroscopy Applied for Investigation of Water Structure”. In: *The Journal of Physical Chemistry B* 119.39 (2015), pp. 12664–12670. DOI: `10.1021/acs.jpcb.5b06622`.

[114] D. S. Venables and C. A. Schmuttenmaer. “Far-infrared spectra and associated dynamics in acetonitrile–water mixtures measured with femtosecond THz pulse spectroscopy”. In: *The Journal of Chemical Physics* 108.12 (1998), pp. 4935–4944. DOI: `10.1063/1.475902`.

[115] Cecilie Rønne et al. “Investigation of the temperature dependence of dielectric relaxation in liquid water by THz reflection spectroscopy and molecular dynamics simulation”. In: *The Journal of Chemical Physics* 107.14 (1997), pp. 5319–5331. DOI: `10.1063/1.474242`.

[116] Hiroyuki Yada, Masaya Nagai, and Koichiro Tanaka. “Origin of the fast relaxation component of water and heavy water revealed by terahertz time-domain attenuated total reflection spectroscopy”. In: *Chemical Physics Letters* 464.4-6 (2008), pp. 166–170. DOI: `10.1016/j.cplett.2008.09.015`.

[117] Uffe Møller et al. “Terahertz reflection spectroscopy of Debye relaxation in polar liquids [Invited]”. In: *Journal of the Optical Society of America B* 26.9 (2009), A113. DOI: `10.1364/josab.26.00a113`.

[118] Udo Kaatze. “Complex permittivity of water as a function of frequency and temperature”. In: *Journal of Chemical & Engineering Data* 34.4 (1989), pp. 371–374. DOI: `10.1021/je00058a001`.

[119] Hans J. Liebe, George A. Hufford, and Takeshi Manabe. “A model for the complex permittivity of water at frequencies below 1 THz”. In: *International Journal of Infrared and Millimeter Waves* 12.7 (1991), pp. 659–675. DOI: `10.1007/bf01008897`.

[120] W. J. Ellison. “Permittivity of Pure Water, at Standard Atmospheric Pressure, over the Frequency Range 0–25THz and the Temperature Range 0–100°C”. In: *Journal of Physical and Chemical Reference Data* 36.1 (2007), pp. 1–18. DOI: `10.1063/1.2360986`.

[121] John David Jackson. *Classical Electrodynamics*. Ed. by American Association of Physics Teachers. Vol. 67. 841-842 9. WILEY, Aug. 1999. 832 pp. ISBN: 047130932X. URL: `https://www.ebook.de/de/product/3240907/john_david_jackson_classical_electrodynamics.html`.

[122] K.E. Oughstun. “Dynamical evolution of the Brillouin precursor in Rocard-Powles-Debye model dielectrics”. In: *IEEE Transactions on Antennas and Propagation* 53.5 (2005), pp. 1582–1590. DOI: `10.1109/tap.2005.846452`.

[123] James McConnell J. T. Lewis and B. K. P. Scaife. “Relaxation Effects in Rotational Brownian Motion”. In: *Proceedings of the Royal Irish Academy. Section A: Mathematical and Physical Sciences* 76 (1976), pp. 43–69.

[124] S. V. Titov et al. “Anomalous diffusion of molecules with rotating polar groups: The joint role played by inertia and dipole coupling in microwave and far-infrared absorption”. In: *Physical Review E* 102.5 (2020), p. 052130. DOI: `10.1103/physreve.102.052130`.

[125] A. Budó. “Dielectric Relaxation of Molecules Containing Rotating Polar Groups”. In: *The Journal of Chemical Physics* 17.8 (1949), pp. 686–691. DOI: `10.1063/1.1747370`.

[126] R. T. Hart et al. “Temperature Dependence of Isotopic Quantum Effects in Water”. In: *Physical Review Letters* 94.4 (2005), p. 047801. DOI: `10.1103/physrevlett.94.047801`.

[127] A. K. Soper and C. J. Benmore. “Quantum Differences between Heavy and Light Water”. In: *Physical Review Letters* 101.6 (2008), p. 065502. DOI: `10.1103/physrevlett.101.065502`.

[128] Francesco Paesani. “Hydrogen bond dynamics in heavy water studied with quantum dynamical simulations”. In: *Physical Chemistry Chemical Physics* 13.44 (2011), p. 19865. DOI: `10.1039/c1cp21863h`.

[129] Nicolas Descharmes et al. “Observation of Backaction and Self-Induced Trapping in a Planar Hollow Photonic Crystal Cavity”. In: *Physical Review Letters* 110.12 (2013), p. 123601. DOI: `10.1103/physrevlett.110.123601`.

[130] C. Koos et al. “All-optical high-speed signal processing with silicon–organic hybrid slot waveguides”. In: *Nature Photonics* 3.4 (2009), pp. 216–219. DOI: `10.1038/nphoton.2009.25`.

[131] Clément Sieutat et al. “Strong confinement of light in low index materials: the Photon Cage”. In: *Optics Express* 21.17 (2013), p. 20015. DOI: `10.1364/oe.21.020015`.

[132] Stefan A. Maier and Harry A. Atwater. “Plasmonics: Localization and guiding of electromagnetic energy in metal/dielectric structures”. In: *Journal of Applied Physics* 98.1 (2005), p. 011101. DOI: `10.1063/1.1951057`.

[133] R. Quidant and C. Girard. “Surface-plasmon-based optical manipulation”. In: *Laser & Photonics Reviews* 2.1-2 (2008), pp. 47–57. DOI: `10.1002/lpor.200710038`.

[134] Ye Pu et al. “Nonlinear Optical Properties of Core-Shell Nanocavities for Enhanced Second-Harmonic Generation”. In: *Physical Review Letters* 104.20 (2010), p. 207402. DOI: `10.1103/physrevlett.104.207402`.

[135] Stefan Mastel et al. “Terahertz Nanofocusing with Cantilevered Terahertz-Resonant Antenna Tips”. In: *Nano Letters* 17.11 (2017), pp. 6526–6533. DOI: `10.1021/acs.nanolett.7b01924`.

[136] T. Hannotte et al. “s-SNOM imaging of a THz photonic mode”. In: *2020 45th International Conference on Infrared, Millimeter, and Terahertz Waves (IRMMW-THz)*. IEEE, 2020. DOI: `10.1109/irmmw-thz46771.2020.9370928`.

[137] Markus Walther, Mark R. Freeman, and Frank A. Hegmann. “Metal-wire terahertz time-domain spectroscopy”. In: *Applied Physics Letters* 87.26 (2005), p. 261107. DOI: `10.1063/1.2158025`.

[138] R. Mendis and D. Grischkowsky. “Undistorted guided-wave propagation of subpicosecond terahertz pulses”. In: *Optics Letters* 26.11 (2001), p. 846. DOI: `10.1364/ol.26.000846`.

[139] N. Marcuvitz. *Waveguide Handbook*. Institute of Electrical Engineers, Sept. 2000. 448 pp. ISBN: 0863410588. URL: `https://www.ebook.de/de/product/21469238/n_marcuvitz_waveguide_handbook.html`.

[140] N. Laman et al. "High-Resolution Waveguide THz Spectroscopy of Biological Molecules". In: *Biophysical Journal* 94.3 (2008), pp. 1010–1020. DOI: `10.1529/biophysj.107.113647`.

[141] Alisha J. Shutler and D. Grischkowsky. "Gap independent coupling into parallel plate terahertz waveguides using cylindrical horn antennas". In: *Journal of Applied Physics* 112.7 (2012), p. 073102. DOI: `10.1063/1.4754846`.

[142] Charles Vassallo. *Théorie des guides d'ondes électromagnétiques*. Vol. 1. Eyrolles Paris, 1985.

[143] Simon Ramo, John R Whinnery, and Theodore Van Duzer. *Fields and waves in communication electronics*. John Wiley & Sons, 1994.

[144] S.B. Cohn. "Slot Line on a Dielectric Substrate". In: *IEEE Transactions on Microwave Theory and Techniques* 17.10 (1969), pp. 768–778. DOI: `10.1109/tmtt.1969.1127058`.

[145] T. Kitazawa et al. "Slot Line with Thick Metal Coating (Short Papers)". In: *IEEE Transactions on Microwave Theory and Techniques* 21.9 (1973), pp. 580–582. DOI: `10.1109/tmtt.1973.1128075`.

[146] T. Kitazawa, Y. Hayashi, and M. Suzuki. "Analysis of the Dispersion Characteristic of Slot Line with Thick Metal Coating". In: *IEEE Transactions on Microwave Theory and Techniques* 28.4 (1980), pp. 387–392. DOI: `10.1109/tmtt.1980.1130081`.

[147] Georgios Veronis and Shanhui Fan. "Guided subwavelength plasmonic mode supported by a slot in a thin metal film". In: *Optics Letters* 30.24 (2005), p. 3359. DOI: `10.1364/ol.30.003359`.

[148] P.J. Gibson. "The Vivaldi Aerial". In: *1979 9th European Microwave Conference*. IEEE, 1979. DOI: `10.1109/euma.1979.332681`.

[149] H. Oraizi and S. Jam. "Optimum design of tapered slot antenna profile". In: *IEEE Transactions on Antennas and Propagation* 51.8 (2003), pp. 1987–1995. DOI: `10.1109/tap.2003.811090`.

[150] Romain Peretti et al. "Broadband Terahertz Light–Matter Interaction Enhancement for Precise Spectroscopy of Thin Films and Micro-Samples". In: *Photonics* 5.2 (2018), p. 11. DOI: `10.3390/photonics5020011`.

[151] C Momma et al. "Precise laser ablation with ultrashort pulses". In: *Applied Surface Science* 109-110 (1997), pp. 15–19. DOI: `10.1016/s0169-4332(96)00613-7`.

[152] S. Nolte et al. "Ablation of metals by ultrashort laser pulses". In: *Journal of the Optical Society of America B* 14.10 (1997), p. 2716. DOI: `10.1364/josab.14.002716`.

[153] Rudolf Weber et al. "Processing constraints resulting from heat accumulation during pulsed and repetitive laser materials processing". In: *Optics Express* 25.4 (2017), p. 3966. DOI: `10.1364/oe.25.003966`.

[154] Sergey Mitryukovskiy et al. "Device for Broadband THz Spectroscopy of 1-nL-Volume Samples". In: *2019 44th International Conference on Infrared, Millimeter, and Terahertz Waves (IRMMW-THz)*. IEEE, 2019. DOI: `10.1109/irmmw-thz.2019.8874026`.

[155] David Rohrbach, Bong Joo Kang, and Thomas Feurer. “3D-printed THz wave- and phaseplates”. In: *Optics Express* 29.17 (2021), p. 27160. DOI: `10.1364/oe.433881`.