UNIVERSITY OF EXETER

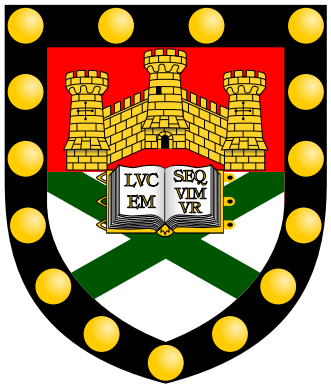
DOCTORAL THESIS

Enhancing Nonlinear Optics in Thin Materials

Submitted by Justus Bohn to the University of Exeter as a thesis for the degree of Doctor of Philosophy in Physics, October 2021.

Supervisors:
Prof. Euan HENDRY
Prof. William L BARNES

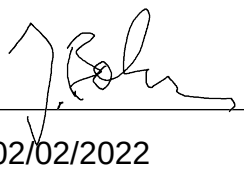
Examiners:
Dr. Sebastian A SCHULZ
Prof. Robert J HICKEN



Declaration of Authorship

I, Justus BOHN, declare that this thesis titled, "Enhancing Nonlinear Optics in Thin Materials" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: 

Date: 02/02/2022

"If I have seen further it is by standing on the shoulders of Giants."

Isaac Newton

UNIVERSITY OF EXETER

Abstract

School of Physics
College of Engineering, Mathematics and Physical Sciences

Doctor of Philosophy

Enhancing Nonlinear Optics in Thin Materials

by Justus BOHN

This thesis examines various nonlinear optical processes to either control the frequency or switch the amplitude of light all-optically. This is achieved in sub-wavelength thin samples designed to overcome issues such as phase-matching. In order to achieve a significant nonlinear optical response, we utilise plasmonic resonances and nano-antennas.

First, we enhance the high-harmonic generation in graphene. A heterostructure of gold nano-ribbons on graphene with a thin insulator in-between enhances the incoming electric field and enormously enhances the high-harmonic generation. A thousandfold enhancement of the fifth-harmonic generations is measured.

Next, we study the all-optical switching of an epsilon-near-zero plasmon in indium tin oxide (ITO). The resonance provides near-perfect absorption and occurs for layers of just 60 nm thickness. Utilising the Kretschmann-geometry in a pump-probe scheme enables phase-matching to the resonance and studying the nonlinear changes of the reflection. Significant absolute changes of 45 % are measured, with an initial reflection of ~1 % thanks to the near-perfect absorption resonance. A novel two-beam coupling contribution is identified and will be essential to take into account in other studies.

The addition of cross-shaped nano-antennas allows for improvement on various critical issues. Optical switching is now possible for lower intensities, normal-incidence and better control of polarisation. The symmetric cross-shape enables a nonlinear dichroic response by which only the probe polarisation parallel to the pumped cross-bar is undergoing a significant nonlinear shift. We analyse the complete polarisation ellipses and identify two wavelength regimes in which the amplitude or phase can be modulated independently.

Finally, we study the angle and frequency-dependent phase-modulation in ITO films of various thicknesses. We present good agreement with the temporal refraction for low angle of incidence. However, we find an additional dependence for higher angles. Close to the epsilon-near-zero case, we even report an opposing frequency shifting contribution. We show that the additional phase contribution origins from the temporal changes to the spatial refractive index boundary. Hence, we refer to the process as spatiotemporal refraction. This contribution can tailor the frequency shift and allow for better designability of simultaneous ultrafast changes in amplitude and phase.

In summary, both graphene and ITO have proven themselves as useful nonlinear active media. The enhanced response makes them promising materials for controlling frequency and amplitude in applications such as optical communication. Beyond that, the work on ITO even unveils novel effects such as spatiotemporal refraction.

Acknowledgements

I acknowledge financial support from the Engineering and Physical Sciences Research Council (EPSRC) of the United Kingdom, via the EPSRC Centre for Doctoral Training in Metamaterials (Grant No. EP/L015331/1).

Personal

On my path to earning the PhD title, I have had the great pleasure of sharing my life with many helpful, passionate and supporting people. Without them, I would have only achieved a mere fraction of my goals. My gratitude goes to far more people than I can list in this section, but I hope I will mention most of the PhD-related colleagues and friends.

Firstly, and most relevant to the PhD project, I want to thank my supervisors, Euan and Bill, for their continued support throughout the last few years. Bill, I want to thank you for your support in understanding and questioning physical concepts, especially regarding plasmonics. Even though the overlap in research field was not always given, you always made time when I needed it. Euan, we have shared an exceptional few years. From questioning our initial research field to building up new research directions and projects with a success that seemed almost unreachable from my first-year student perspective. Your continued trust in my capabilities and freedom to explore issues with the setup where others might have pushed for progress has enabled me to grow research projects that I am proud of. Your outstanding, hands-on support with, e.g. laser issues and intuitive understanding of physical problems have been extremely helpful qualities that I am sure not every PhD student could enjoy.

As the only other PhD student working on nonlinear optics, I had the great pleasure to spend my first two to three years researching together with Craig Tollerton. Sharing the supervisor meetings with you was good fun and made the transition into my PhD a lot easier than being completely alone. However, I also want to thank Sam Hornett for supporting me in understanding electronic issues and the like. Even more, I want to thank Sam Hutchings for his continued remote support in dealing with all the python controlled stages and buggy scripts that I inherited. I also want to thank Ros for existing now and then in the lab and making research in the basement less lonely. I further enjoyed being part of Euans group, which includes even more awesome people like Lauren and Michal.

During my PhD, I characterised the nonlinear optical response in the complex optical setup I built in Exeter. To do so efficiently, I had to rely on collaborators for providing state of the art samples or sophisticated theoretical modelling. In chronological order, I want to first thank someone I only shared time within the very beginning of my research endeavour in Exeter, Freddie Page. My initial project aimed to investigate plasmonic gain in graphene as a potential explanation for signals measured in experiments that go far beyond theoretical expectations. Within my first months in England, Freddie invited me over for a few days to the Imperial College in London to help me model the predicted impact of the gain on our experimental measurements. Despite "only" resulting in a negative result, it was a crucial and significant discovery to make early on and shift focus for my scientific hunt for explanations.

Next, I want to thank all the people from ICFO that were involved in the graphene research projects. Again, the initial help by Darrick and David L. in understanding the second-order nonlinear simulations of graphene. Joel Cox has been incredibly

helpful in further extending towards modelling third-order effects such as four-wave mixing. David A. I. provided the sample, advice and linear characterisation required to do my first nonlinear optics research on nano-structured graphene samples. These samples continued to stay relevant and became part of my first thesis chapter on high-harmonic generation enhancement for graphene. This collaboration was only possible thanks to the help of our colleagues in Vienna which agreed to do the nonlinear measurements in their setup upon realising that our laser has too little repetition rate to be a sensible choice. Most honour goes to Irati and Lee for continuing the measurements for various sample parameters, polarisation and wavelength long after I left. The still ongoing theoretical modelling spans complex problems tackled from multiple perspectives thanks to Phillip, Eduardo.

The switch to successful switch to epsilon-near-nonlinear optics was crucially enabled by the quick and high-quality supply of ITO samples fabricated by Willie. These simple but well fabricated thin layer samples of varying thickness became my central piece of research and are the foundation to both chapters in this thesis that I am by far the proudest of. After having initial difficulties with ENZ samples supplied by industry, these samples were the basis for leading research with my novel setup. Additionally, I want to thank Simon Horsley for his kind, quick and very helpful input on modelling the time-resolved nonlinear optical response of our thin layers.

Finally, I have the great pleasure to collaborate with Bob Boyd's group thanks to a research exchange grant we won together with the help of Jeremy. I thank Zahir for his proactive approach to collaborating on antenna enhanced ENZ nonlinear optics. My thanks also go to Mohammad and Yaswant for fabricating the samples and characterising their linear properties.

Beyond the direct involvement in my research project, I want to thank the CDT PhD students of my cohort and beyond for making this endeavour a joined adventure. Feeling and emphasising the pain of different yet similar research projects certainly makes this journey more endurable and socially valuable.

Special thanks also go to the more organisational staff that made this journey possible such as Anja, Alastair and many more. I am grateful for their work leading to exciting opportunities for passionate students within the CDT.

I also want to thank all the people who supported me outside of the realm of science with their continuous friendship. I want to thank especially my go-to social circle of Conor, Marziyeh, Sean and Maria. The weekends, dinners and social trips we enjoyed together have been some of the greatest highlights of my life in Exeter.

Finally, I want to thank the backbone of my successes in life: my family's continued and indestructible support. My parents motivated and enabled me to the best of their abilities to achieve whatever goals I may have. It will always be one of my greatest pleasures to visit my parents' house, but also meeting the entire "Hecht-Clan". Despite the growing challenge in doing so, I am sure we will find ways to prevail and nurture this social network that I was most privileged to enjoy throughout my life.

To Judith. After sharing our life in Jena, Sydney and Exeter, I can't wait for our next chapter in Switzerland. While we have all options to design the life we want, I can always be sure it will be the best possible version, knowing you are part of it.

Publications

In preparation

J. Bohn, Y. Vaddi , R. Boyd, W. Barnes, E. Hendry, M. Z. Alam "**Time-varying dichroic response of a symmetric plasmonic nanoparticle array**" (minor additional measurements required) [see chapter 6]

I. A. Calafell, L. A. Rozema, A. Trenti, **J. Bohn**, D. A. Iranzo, J. D. Cox, F. J. Garcia de Abajo, F. H. L. Koppens, E. Hendry, P. Walther "**High-harmonic generation enhancement with graphene heterostructures**" (minor additional measurements required) [see chapter 4]

Published

J. Bohn, T. S. Luk, S. Horsley, E. Hendry "**Spatiotemporal refraction of light in an epsilon-near-zero ITO layer: frequency shifting effects arising from interfaces**," OPTICA, **8**, 1532-1537 (2021) [see chapter 7]

J. Bohn, T. S. Luk, C. Tollerton, S. W. Hutchings, I. Brener, S. Horsley, W. Barnes, E. Hendry "**All-optical switching of an epsilon-near-zero plasmon resonance in indium tin oxide**," NATURE COMMUNICATIONS, **12**, 1017 (2021) [see chapter 5]

C. J. Tollerton, **J. Bohn**, T.J. Constant, S. A. R. Horsley, W. L. Barnes, E. Hendry, D. E. Chang, Z. Li, "**Origins of All-Optical Generation of Plasmons in Graphene**," SCIENTIFIC REPORTS, **9**, 3267 (2019) [see Appendix A]

C. Zou, A. Komar, S. Fasold, **J. Bohn**, A. A. Muravsky, A. A. Murauski, T. Pertsch, D. N. Neshev, I. Staude, "**Electrically Tunable Transparent Displays for Visible Light Based on Dielectric Metasurfaces**," ACS PHOTONICS **6** (6), 1533-1540 (2019) [M. Sc. project]

J. Bohn, T. Bucher, K. Chong, A. Komar, D. Neshev, Y. Kivshar, T. Pertsch, I. Staude, "**Active tuning of spontaneous emission by Mie-resonant dielectric metasurfaces**," NANO LETTERS , **18**, 6, 3461–3465 (2018) [M. Sc. project]

A. Vaskin, **J. Bohn**, K. Chong, M. Zilk, T. Bucher, D. Choi, D. Neshev, T. Pertsch, I. Staude, "**Directional and Spectral Shaping of Light Emission with Mie-Resonant Silicon Nanoantenna Arrays**," ACS PHOTONICS (2018) [M. Sc. project]

A. Komar, Z. Fang, **J. Bohn**, J. Sautter, M. Decker, A. Miroshnichenko, T. Pertsch, I. Brener, Y. Kivshar, I. Staude, D. Neshev, "**Electrically tunable all-dielectric optical metasurfaces based on liquid crystals**," APPLIED PHYSICS LETTERS, **110**, 071109 (2017) [M. Sc. project]

Contents

List of Figures

List of Tables

List of Abbreviations

2D	**T**wo-**d**imensional
ALD	**A**tomic **l**ayer **d**eposition
AZO	**A**luminium doped **z**inc **o**xide
BBO	**B**arium **bo**rate
CVD	**C**hemical **v**apor **d**epostion
EBL	**E**lectron **b**eam **l**ithography
ENZ	**E**psilon-**n**ear **z**ero
EUV	**E**xtreme **u**ltra **v**iolett
FDTD	**F**inite **d**ifference **t**ime **d**omain (method)
FEM	**F**inite **e**lement **m**ethod
FH(G)	**F**ifth **h**armonic (**g**eneration)
FWHM	**F**ull-**w**idth-**h**alf-**m**aximum
GPP	**G**raphene **p**lasmon **p**olariton
HWP	**H**alf **w**ave **p**late
HHG	**H**igh **h**armonic **g**eneration
IR	**I**nfra**r**ed
ITO	**I**ndium **t**in **o**xide
LSP	**L**ocalized **s**urface **p**lasmon polariton
MIR	**M**id **i**nfra**r**ed
NIR	**N**ear **i**nfra**r**ed
NLO	**N**on **l**inear **o**ptics
OPA	**O**ptical **p**arametric **a**mplifier
OPO	**O**ptical **p**arametric **o**scillator
PA	**P**erfect **a**bsorption
pol	**pol**arisation
SEM	**S**canning **e**lectron **m**icroscope
SH(G)	**S**econd **h**armonic (**g**eneration)
SNSP	**S**uperconducting **N**anowire **S**ingle-**P**hoton (detector)
SP(P)	**S**urface **p**lasmon **p**olariton
SPM	**S**elf **p**hase **m**odulation
TBC	**T**wo-**b**eam **c**oupling
TCO	**T**ransparent **c**onductive **o**xide
TH(G)	**T**hird **h**armonic (**g**eneration)
TMDC	**T**ransistion **m**etal **dic**halcognides
TMM	**T**ransfer **m**atrix **m**ethod
vdW	**v**an **d**er **W**aals
VIS	**Vis**ible (wavelength range)
UV	**U**ltra **v**iolett
XPM	**C**ross **p**hase **m**odulation

Chapter 1

Introduction

Abstract

This thesis details experimental investigations on enhancing the nonlinear optical (NLO) response of thin materials. The fundamental strategy is to utilise the plasmonic resonances such as surface plasmon polaritons (SPP) or epsilon-near-zero (ENZ) resonances to enhance the nonlinear response of the active layer.

1.1 History and Motivation

The studies of optics range back 1000s of years, possibly even longer, with examples such as the development of lenses by the ancient Egyptians and Mesopotamians. The laser, however, is an invention of the 20th century and was first built in 1960 by Theodore H. Maiman[1], based on theoretical work by Charles Hard Townes and Arthur Leonard Schawlow[2]. With its invention, multifold optical research areas were born, one of them being nonlinear optics (NLO). Peter Franken and coworkers were the first to observe second-harmonic generation in a quartz crystal[3]. Since then, NLO has become a mature field of science and engineering. The scope of this discipline includes all phenomena in which optical parameters of materials are changed by irradiation with light. In optical labs and industry, it has become a standard tool to create new wavelengths.

With the advances of nano-fabrication, an ever-growing demand for the miniaturising of optical setups has developed. Micrometre sized waveguides have enabled to guide, control and manipulate light on photonic chips. With the discovery of one atom thick layer of graphene in 2004 by Andre Geim and Kostya Novoselov[4], a new opportunity has risen: Nonlinear optics in 2D materials. While the one atom thick layer offers the advantage of ultimate miniaturisation, it also presents the significant challenge of producing nonlinear effects within this one layer. Especially considering graphene only absorbs about 2.3 % of the incident light, so it barely interacts with electromagnetic waves. Despite this, its zero band-gap nature, unusually large chemical and electrical tunability, and effects such as Pauli blocking provide unique functionalities for photonic nanodevices. Euan Hendry and collaborators presented first that graphene offers a strong coherent nonlinear optical response by investigating four-wave mixing of graphene flakes [5]. Properties such as ultrafast photoexcitation dynamics, high chemical and mechanical stability, and large thermal and optical threshold damage of graphene make it an ideal test-bed for studying nonlinear optics phenomena in 2D physical systems. A multifold of transition metal dichalcogenides (TMDCs) have also been identified and studied as promising candidates such as MoS_2, WSe_2 and hexagonal boron nitride (h-BN)[6]. However, to establish 2D materials as viable nonlinear materials, further optimisation is required.

So-called surface polaritons may present one solution. Surface polaritons are modes of light that are confined to the surface of a material. In the case of a graphene plasmon polariton (GPP), the light is bound to the surface by constantly interacting with quantised carrier oscillations, plasmons. Due to this interaction, the electromagnetic mode has a much smaller effective wavelength than vacuum propagating photons with the same energy would have. Electric field confinement down to 1/100 of the wavelength and less have been documented for graphene surface plasmons [7]. The electric fields also decay exponentially away from the surface, leading to strong fields directly on the surface.

While monolayer materials present the ultimate confinement goal, a thickness smaller than the wavelength of light is often more than enough to combat issues such as phase-matching. Only very recently, the field of epsilon-near-zero (ENZ) nonlinear optics arose, utilising a vanishing real part of the permittivity for huge nonlinear effects. For the example of indium-tin-oxide (ITO), the nonlinear change in refractive index was measured to be of unity order ($\Delta n \sim 1$), which not only is large by itself but also larger than the low initial refractive index ($n_0 \sim 0$) [8]. This can at least be partially explained by the electric field enhancement of the ENZ region near the plasma frequency. With ENZ resonances being particularly interesting for ultra-thin cases of less than 100 nm, they present a promising candidate for sub-wavelength

nonlinear optics and will form the main part of this thesis. The huge nonlinear refractive index combined with sub-ps cooling dynamics make ENZ materials an ideal platform for ultrafast all-optical switching.

1.2 Outline

This thesis will form an experimental investigation into combining plasmonic or ENZ materials and nonlinear optics in thin layered materials.

In chapter 4, studies of enhanced high harmonic generation in graphene-based devices are presented. Combining graphene with nanostructures allows for field enhancement of the pump to increase the high harmonic generation. The investigation includes various structuring parameters such as antenna width and distance to the graphene layer. We find more than three orders of magnitude fifth-harmonic generation enhancement.

In chapter 5, all-optical switching of an epsilon-near-zero plasmon resonance in indium tin oxide will be studied. The near-perfect absorption of the ENZ plasmon is combined with the red-shifting of the bulk plasma frequency in the ITO layer to produce almost two orders of magnitude sub-ps reflection switching.

In chapter 6, a cross-shaped antenna is added on top of an ITO layer to create a transient polarization-dependent response. The sample was designed only to induce a nonlinear response along the bar that has been pumped. We find a more than seven times stronger nonlinear response of a probe beam parallel polarised to the pump compared to the perpendicular case.

In chapter 7, we investigate "spatiotemporal refraction" inside thin ITO layers. Just below the ENZ wavelength, we find strong blue shifting for thin samples, only explainable via temporal changes to the spatial boundaries of the ITO film.

Finally, we summarize our findings in chapter 8 and discuss possible future directions.

Chapter 2

Fundamentals

Abstract

Here, fundamental theories and previous research activities are shown. First, the path from Maxwell equations to nonlinear optical (NLO) concepts are presented. Then, the materials used in this thesis, ITO and Graphene, are introduced. Different types of plasmon polaritons and their potential to enhance NLO are outlined. Finally, reports on state-of-the-art research will signify the importance of the projects within this thesis.

2.1 Electromagnetics

This section will provide the theoretical basis for describing and understanding the optical properties studied in this thesis. We will start defining a dielectric medium, adding free carriers, and finally, considering the case of high intensities, crucial for nonlinear optical processes.

2.1.1 Maxwell's Equations in a Dielectric Medium

To understand electromagnetic processes, the Maxwell equations form the basis of almost every problem and describe the relation between electric (E) and magnetic (H) fields:

$$\vec{\nabla} \cdot \vec{D} = 0 \tag{2.1}$$

$$\vec{\nabla} \cdot \vec{B} = 0 \tag{2.2}$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \tag{2.3}$$

$$\vec{\nabla} \times \vec{H} = \frac{\partial \vec{D}}{\partial t}. \tag{2.4}$$

This representation assumes the absence of free charges or currents. The electric (D) and magnetic (B) flux densities describe the material response, taking into account the materials polarisation (P) and magnetisation densities (M). For this thesis we will work with non-magnetic materials, resulting in $\vec{H} = \frac{1}{\mu_0}\vec{B}$. However, the displacement field $\vec{D}$ is different from the electric field $\vec{E}$. The very basis of an optical material response resides in the polarisation field $\vec{P}$, which appears mathematically as:

$$\vec{D} = \epsilon_0 \vec{E} + \vec{P}. \tag{2.5}$$

For standard dielectric optics such as lenses, the polarisation field $\vec{P}$ follows as (linear, homogeneous, isotropic)

$$\vec{P} = \epsilon_0 \chi \vec{E}, \tag{2.6}$$

where the scalar χ is called the *electric susceptibility*. For this example we see, $\vec{D} \propto \vec{E}$ with the proportionality factor given as the *electric permittivity* ε of the material:

$$\varepsilon = \epsilon_0(1 + \chi). \tag{2.7}$$

2.1.2 Conductive Medium

Materials with free electric charges pose an additional contribution to the permittivity that can be described by the so-called *Drude model*. It describes the conduction electrons as independent particles in an ideal gas that freely move between scattering events inside the medium. To determine the permittivity of this electron gas $\varepsilon_{\text{Drude}}$, we investigate the equation of electron motion upon applying an external field $\vec{E}$:

$$m\ddot{\vec{x}} + m\gamma\dot{\vec{x}} = -e\vec{E}. \tag{2.8}$$

Here, m corresponds to the optical mass and γ to the scattering rate of the electrons. For the case of a monochromatic electric field $\vec{E}(t) = \vec{E}_0 \exp^{-i\omega t}$ it follows:

$$\vec{x} = \frac{e}{m(\omega^2 + i\gamma\omega)} e\vec{E}(t). \tag{2.9}$$

With the macroscopic polarisation given as $\vec{P} = -ne\vec{x}$ and n being the electron density we get

$$\vec{P} = -\frac{ne^2}{m(\omega^2 + i\gamma\omega)} e\vec{E}(t). \tag{2.10}$$

This finally leads to the complex permittivity of the free electron gas as

$$\varepsilon_{\text{Drude}}(\omega) = 1 - \frac{\omega_p^2}{\omega^2 + \mathrm{i}\omega\gamma}, \tag{2.11}$$

where $\omega_p^2 = \frac{ne^2}{\varepsilon_0 m}$ is the *plasma frequency* of the free electron gas.

2.1.3 Nonlinear Medium

For high field intensities $\langle \vec{E} \times \vec{H} \rangle_T \propto |\vec{E}|^2$ the frequency-dependent polarisation field of a material is often approximated by introducing higher-order electric field dependencies:

$$\vec{P}(\vec{r}, \omega) = \varepsilon_0 \left(\overbrace{\chi^{(1)} \cdot \vec{E}}^{\text{linear}} + \overbrace{\underbrace{\chi^{(2)} : \vec{E}\vec{E}}_{2^{\text{nd}}\text{ order}} + \underbrace{\chi^{(3)} \vdots \vec{E}\vec{E}\vec{E}}_{3^{\text{rd}}\text{ order}} + \cdots}^{\text{nonlinear}} \right). \tag{2.12}$$

For the general case of an anisotropic medium, $\chi^{(1)}$ forms a matrix. The nonlinear susceptibility $\chi^{(n)}$ corresponds in general to a tensor of rank $n+1$. The wavelength dependence of the susceptibilities can be explicitly written in the form of $\chi^{(n)}(\omega_{\text{out}} | \omega_1, \omega_2, \dots)$. The incoming frequencies ω_i can be positive or negative and their sum results in the outgoing frequency $\omega_{\text{out}} = \sum_{i=1}^{n} \omega_i$. Depending on the material, these $\chi^{(n)}$ values can change dependent on every single frequency choice. However, often the complexity of these tensors drastically reduces due to various symmetries and loss-less characteristics.

One of the most important symmetry aspects is that centrosymmetric systems can only have $\chi^{(odd)}$ components that are non-zero. This can quickly be understood by considering the spacial inversion $\vec{r} \to -\vec{r}$, which implies $\vec{P} \to -\vec{P}$ and $\vec{E} \to -\vec{E}$. If spatial inversion is symmetric, meaning the relation between $\vec{P}$ and $\vec{E}$ is unchanged after inversion, this implies the even terms must be 0: $\chi^{(even)} = 0$. For insights into further symmetries of nonlinear optics, we refer to e.g. [9].

To see the impact of this polarisation field, we take the rotation of Equation 2.3 and replace $\vec{\nabla} \times \vec{B}$ with Equation 2.4, leading to:

$$\vec{\nabla} \times \vec{\nabla} \times \vec{E} + \frac{1}{c^2} \frac{\partial^2 \vec{E}}{\partial t^2} = -\mu_0 \frac{\partial^2 \vec{P}^{(1)}}{\partial t^2} - \mu_0 \frac{\partial^2 \vec{P}^{(\text{NL})}}{\partial t^2}. \tag{2.13}$$

Transitioning to the frequency domain to include our frequency dependent nonlinear polarisation leads to:

$$\vec{\nabla} \times \vec{\nabla} \times \vec{E}(\vec{r},\omega) - \frac{\omega^2}{c^2}\varepsilon_r(\vec{r},\omega)\vec{E}(\vec{r},\omega) = \omega^2 \mu_0 \vec{P}^{(\mathrm{NL})}(\vec{r},\omega), \tag{2.14}$$

with ε_r being the relative permittivity of the material.

Most importantly, one can see now that the nonlinear polarisation field $P^{(\mathrm{NL})}$ can act as a driving field of this wave equation. Different nonlinear susceptibilities can be utilised in various processes to mix frequencies or control light propagation depending on its intensity, as described in the next section.

2.1.4 Nonlinear Processes

Here, the most relevant examples of nonlinear processes for our research activities are presented. To establish relevance, we will use the laser system used for the research projects as a reference. Specifically, we use a "Legend Coherent" Ti-sapphire laser system with an 800 nm centre wavelength. This feeds into "TOPAS" optical parametric amplifiers to generate new wavelengths ranging from 200 nm to 10 μm through 2$^{\mathrm{nd}}$ order nonlinear processes.

2nd Order Processes ($\chi^{(2)}$)

In the 2$^{\mathrm{nd}}$ order case the wave mixing process consists of two incoming frequencies ω_1,ω_2 and one outgoing frequencies ω_3 as depicted in Figure 2.1.

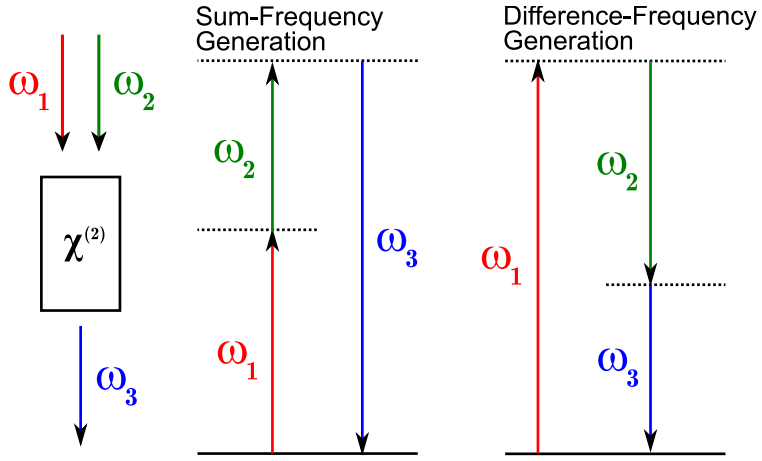


FIGURE 2.1: **Schematic of the 2$^{\mathrm{nd}}$ order processes.** The incoming frequencies ω_1, ω_2 can generate ω_3 to be the sum- or difference-frequency, depending on the materials frequency dependent $\chi^{(2)}$.

The *difference-frequency* process forms the basis of optical parametric amplifiers. For our case, the pump wavelength is $\lambda_{pump} = \lambda_1 = 800$ nm and signal wavelength is $\lambda_{signal} = \lambda_2 > 800$ nm leading to the generation of the idler wavelength $\lambda_{idler} = \lambda_3$ and, more importantly in the case of an optical parametric *amplifier*, to the stimulated emission of a signal photon. This process is being used in our TOPAS setup to generate near-infrared (NIR) wavelength ranging ~ 1...3 μm. To further generate wavelength up to 10 μm efficiently, one can use an additional nonlinear crystal, in our case $AgGaS_2$, to generate the difference frequency of the signal and idler.

To reach higher frequencies, the *sum-frequency generation* is utilised. In the case of our TOPAS setup, two BBO crystals are used to reach wavelengths as low as 200 nm by double second harmonic generation, meaning $\lambda_1 = \lambda_2$. The non-degenerate sum-frequency processes can then be used to tune between 200 and 800 nm.

3rd Order Processes ($\chi^{(3)}$)

In principle, all materials could posses a non-frequency degenerate four-wave-mixing nonlinearity. Considering the vast permutation capabilities of three incoming frequencies ω_1, ω_2 and ω_3 we will only shortly introduce the most common case: $\omega_1 = \omega_2 = \omega_3$ as depicted in Figure 2.2.

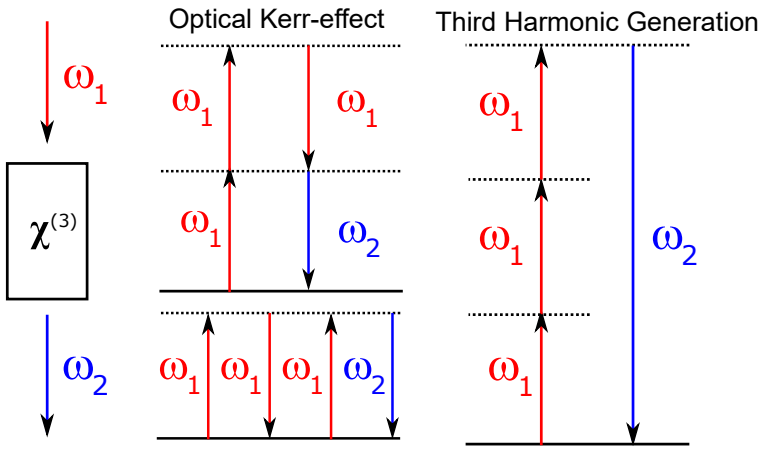


FIGURE 2.2: **Schematic of the 3rd order processes.** Based on a single incoming frequency ω_1 there are two well known effects: The optical Kerr effect, which leads to the same outgoing frequency $\omega_2 = \omega_1$ and the third harmonic generation case, where $\omega_2 = 3\omega_1$.

The optical *Kerr-effect* is a special case where the incoming frequency and outgoing frequency are equivalent. This offers the interesting effect of generating a frequency-dependent refractive index, often presented in the form of $n(I) = n_0 + n_2 I$. This effect can, for example, be utilised to focus beams with high intensity, which is commonly used in femtosecond lasers to mode-lock by placing an aperture in the beam path to favour a pulsed, high-intensity mode in the cavity. We will discuss further details on the intensity-dependent refractive index in subsection 2.1.5.

Finally, *third harmonic generation* is similar to second-harmonic generation since it is essentially just the following order. The real challenge lies within *phase matching*. To generate new frequencies in a bulk medium, the incoming beams and newly generated wavelength beams must not walk off from each other, or the newly generated frequencies would eventually destructively interfere. To reach this condition, birefringent phase matching is typically used in 2nd order processes. With increasing differences in frequency for 3rd or higher-order processes, it becomes harder to resolve this issue. Hence, high harmonic generation applications in solid-state materials are even more problematic. We will be investigating the option to avoid bulk propagation altogether and use 2D materials instead.

High Harmonic Generation (HHG)

High harmonic generation generally describes the process of producing photons with the energy $\hbar\omega_n = n\hbar\omega_1$, where $\hbar\omega_1$ resembles the energy of the incoming photon. While second and third-order generation has been discussed before, this effect usually refers to $n \gg 1$. This field of research has become crucial mainly due to the following two aspects.

First, and most obviously, high harmonics have been utilised to drive the generation of high-frequency coherent radiation sources. This enabled the generation of extreme ultraviolet (EUV) radiation and beyond.

Second, a broad range of frequencies can be covered due to the potentially high number of harmonic orders. All these orders have a fixed-phase relation and have enabled the generation of attosecond pulses.

2.1.5 Intensity-Dependent Refractive Index

When light with high intensity propagates through a medium, a change in the refractive index effects occurs via a so-called optical Kerr-effect:

$$\Delta n = n_2 I, \tag{2.15}$$

where n_2 is the nonlinear (refractive) index. This nonlinear index originates from changes to the fields as seen in the nonlinear effective permittivity [10]

$$\varepsilon = \varepsilon^{(1)} + 3\chi^{(3)}|E|^2, \tag{2.16}$$

assuming 3rd order contributions only. It relates to the refractive index as exactly

$$n = \sqrt{\varepsilon} = \sqrt{\varepsilon^{(1)} + 3\chi^{(3)}|E|^2}. \tag{2.17}$$

However, within the standard definition one would Taylor expand the root term to the first order:

$$n = \sqrt{n_0^2 + 2n_0\, n_2\, I} \approx n_0 + n_2\, I. \tag{2.18}$$

Considering the optical field intensity of

$$I = 2\mathrm{Re}(n_0)\varepsilon\, c\, |E|^2, \tag{2.19}$$

one can now extract the standard definition of the nonlinear index as

$$n_2 = \frac{3\chi^{(3)}}{4n_0\,\mathrm{Re}(n_0)\varepsilon_0\, c}. \tag{2.20}$$

Values of n_2 for typical glasses substrates such as fused silica and $CaFl_2$ are of the order $10^{-16}\,\mathrm{cm}^2\,\mathrm{GW}^{-1}$ [11], and satisfy the Taylor expansion requirement of $|2n_2\, I/n_0| \ll 1$[9]. Recent findings in nonlinear optics have shown that this approximation does not hold for materials such as indium tin oxide[10] as seen in Figure 2.3. Hence, we will model our experiments in chapter 5 and chapter 7 via changes to the permittivity and utilise the root dependence $n = \sqrt{\varepsilon}$ to directly extract nonlinear effects rather than approximating via Taylor expansion.

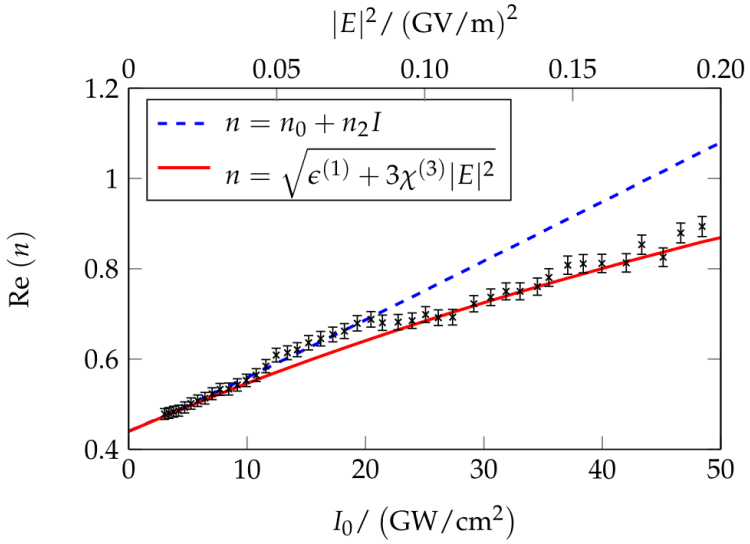


FIGURE 2.3: **Intensity-dependent index of the refraction of ITO.** For the case of $\lambda = 1240\,\text{nm}$ and $\varepsilon \approx 0$, Equation 2.18 performs poorly at describing the refractive index at most intensities (dashed blue line). Using Equation 2.17, a much-improved agreement with the measurement without additional fit parameters is obtained (red line). (Reproduced from [10])

Frequency shifting

In a bulk medium, the propagation phase of light is given as

$$\Phi = n\,k_0\,z - \omega\,t. \tag{2.21}$$

The previously discussed temporal changes to the refractive index Δn across a nonlinear medium of the length L lead to temporal phase modulations equal to

$$\Delta\Phi = \Delta n\,k_0\,L. \tag{2.22}$$

Considering the frequency of light corresponds to the temporal phase change of the optical beam given as [12]

$$\omega = -\frac{\partial\Phi}{\partial t}, \tag{2.23}$$

the nonlinear refractive index leads to a frequency shift given as

$$\Delta\omega = -\frac{\partial n}{\partial t}\frac{L}{c}\omega. \tag{2.24}$$

The temporal dependence of the refractive index changes, given by the pulse and material parameters, determines the frequency shift. This effect has been long known as self-phase modulation (SPM), e.g. highly intense laser beams travelling through a glass fibre, or cross-phase modulation (XPM), e.g. for an intense beam (pump) modulating the refractive index of a material passed by a weaker beam (probe). This could, in principle, also be achieved by non-optical pumping. However, conventional heating or similar methods do not achieve the ultrafast time requirements to induce a

noticeable effect.

Recently, ENZ nonlinear materials were reported to provide extreme frequency shifting effects, e.g. ~6 % for intensities of ~500 GW cm^{-2} and 620 nm of ITO[13]. The significant ultrafast changes to the refractive index were interpreted as a temporal refractive index boundary $\Delta n(t)$. This time translation symmetry breaking has to be compensated by a change of the optical property of light. While spatial refraction leads to a change of the oscillation of light occupied in space (wavelength), *temporal refraction* leads to a change of the oscillation of light in time (frequency).

So far, the research community has discussed frequency shifting effects mainly in the context of these propagation phase changes inside a bulk medium, as seen in Equation 2.22. In chapter 7, we show that phase changes occurring at spatial boundaries can also lead to frequency shifting effects upon temporal modulation of the involved optical parameters. We will refer to this effect as *spatiotemporal refraction*.

2.1.6 Nonlinear Optical Mechanisms

The nonlinear optical properties can appear in different ways. Table 2.1 gives an overview of nonlinear refractive indices and the corresponding time scales for various mechanisms of origin taken from [9] and [14].

The electronic polarisation component is given by electron-cloud distortions that are fast (10^{-15}s), but also relatively low in effect (10^{-16} cm^2 W^{-1}).

The molecular orientation denotes the additional electric polarisation contribution from an optical field-induced reorientation of anisotropic molecules in a liquid. The response time of this process is dependent on the rotational viscosity of molecules in the liquid and is approximately 10^{-12} s.

There is an induced acoustic motion inside the medium for the electrostriction case, and the process is no longer localised. The induced refractive index change behaviour depends on the acoustic velocity and the interaction size between the light and induced acoustic waves.

In Opto-thermal processes, the absorbed optical energy will partially transfer to the thermal energy through radiation-less transitions, resulting in a temperature change. For most optical media, the refractive index is sensitive to temperature change. The corresponding changes in refractive index can be large (10^{-6} cm^2 W^{-1}), but are slow (10^{-3} s).

TABLE 2.1: Typical values of the nonlinear refractive index for different mechanisms. (Reproduced from [9])

Mechanism	n_2 (cm^2 W^{-1})	Response time (s)
Electronic polarisation	10^{-16}	10^{-15}
Molecular orientation	10^{-14}	10^{-12}
Electrostriction	10^{-14}	10^{-9}
Opto-thermal effect	10^{-6}	10^{-3}

This thesis will concern mainly two cases. In graphene, the measured phase modulation effects arise from a more complex phenomenon referred to as saturable photoexcited-carrier refraction [15]. Large and ultrafast optical nonlinearities are measured, surpassing that of typical glasses by eight orders of magnitude with fs-time dynamics [5].

Second, TCOs, which exhibits electron population changes causing a nonlinear optical response. The intraband processes in TCOs can provide strong nonlinear effects, as seen for the case of ITO (10^{-10} cm^2 W^{-1}) with a sub-ps time-scale [8].

2.2 Materials

This section introduces the materials used in the research project, transparent conductive oxides (TCOs such as ITO) and graphene.

2.2.1 Transparent Conductive Oxides

In this thesis, thin layers of transparent conductive oxides (TCO) will be used to study nonlinear optical effects. They can be designed to have their epsilon-near-zero (ENZ) regime in the telecommunication wavelength range. This corresponds to the wavelength range where the real part of the permittivity (and refractive index) is close to zero. Considering the nonlinear refractive index dependence of $n_2 \propto 1/n^2$ from Equation 2.20, this promises a highly nonlinear response and has been shown by M. Alam et al. in "Large optical nonlinearity of indium tin oxide in its epsilon-near-zero region" [8].

Oxide semiconductors such as zinc oxide, cadmium oxide and indium oxide can be highly doped to conduct electricity [16]. Since these semiconductors have a large bandgap, they are transparent in the visible range. Hence, these materials are known as transparent conducting oxides. Such TCOs are famous for display panels where they form the electrical contacts to the pixel circuitry. One of the most popular TCOs is indium tin oxide (ITO) [17].

The free carrier concentration can be easily manipulated between 10^{19} cm^{-3} and 10^{21} cm^{-3} through doping or electronic gating, which offers excellent controllability and electrical tunability of the optical properties in the telecom wavelength range. The optical properties of TCOs can be tuned from behaving similar to a dielectric or metallic material. During the transition, the real permittivity crosses zero, and the absolute permittivity reaches a minimum. TCOs then become epsilon-near-zero (ENZ) materials. The electric field will concentrate in the ENZ TCO due to the electric displacement continuity at the interface leading to $E_{\mathrm{ENZ},\perp} = E_{\mathrm{air},\perp}/\varepsilon_{\mathrm{ENZ}}$, which increases the light-matter interaction in the TCOs. Such properties make TCOs promising materials for building active electro-optical devices [18].

Indium Tin Oxide (ITO)

Here, we present the optical properties of ITO as one example of many similar TCOs. Depending on doping levels and layer thickness, different optical parameters can be reached. Due to the large doping concentration, the optical properties of ITO at near-infrared wavelengths resemble those of a free electron gas. Considering also that the bandgaps of the TCOs are big compared to the wavelength range of interest, their response can be appropriately estimated by the Drude-response give by [8]

$$\varepsilon_{\mathrm{Drude}}(\omega) = \varepsilon_\infty - \frac{\omega_\mathrm{p}^2}{\omega^2 + \mathrm{i}\omega\gamma}, \tag{2.25}$$

with ε_∞ as the high-frequency permitivitty of the tin-doped indium oxide, ω_p as the bulk plasmon frequency and γ as the damping rate.

To give an example, for the nonlinear studies of M. Alam et al. [8] that will be discussed in subsection 2.4.2, a 310 nm ITO film was used leading to $\varepsilon_\infty = 3.8055$, $\omega_\mathrm{p}/2\pi = 473$ THz and a damping rate of $\gamma = 0.0468\omega_\mathrm{p}$. The epsilon-zero wavelength results in $\lambda_0 = 1240$ nm as seen in Figure 2.4.

There are many exciting applications for ENZ materials, such as nonlinear optics, flexible photonics, quantum information processing and heat management, to name

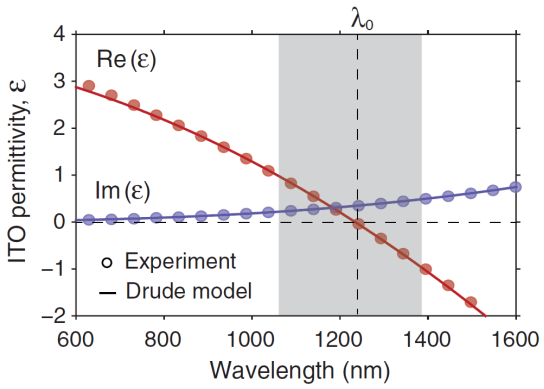


FIGURE 2.4: **Drude response of an ITO film.** Linear relative permittivity of a 310 nm ITO film measured via spectroscopic ellipsometry (symbols) and fitted using the Drude model (lines). The shaded region highlights the ENZ region around λ_0 = 1240 nm, where $\Re(\varepsilon) = 0$. (Reproduced from [8])

the most promising [19]. This thesis will focus on the prospects for nonlinear optics, which will be further discussed in the nonlinear part presented in subsection 2.4.2.

2.2.2 Graphene

Graphene, a 2D honeycomb lattice of carbon atoms, has been a theoretical concept for various decades [20] and has been isolated as a single layer in 2004 [21]. The interest in studying graphene is driven by its linear, massless band structure and the many unusual electrical, thermal, mechanical, and optical properties [22, 23] . The linear electronic dispersion around the Dirac points also leads to strongly nonlinear optical behaviour since resonant interband optical transitions occur at all photon energies [5].

The honeycomb lattice of graphene leads to so-called Dirac-points in the corners (K and K' points) of its Brillouin zone (as shown in Figure 2.5). The Dirac cones are probably graphenes most intriguing feature considering the linear, gap-less Energy dispersion that comes with it:

$$E(\vec{k}) \approx \pm \hbar v_F |\vec{k}|, \tag{2.26}$$

with $\vec{k}$ as the momentum measured relatively to the Dirac points and $v_F \approx 10^6$ m/s as the constant Fermi velocity.

The linear dispersion charge carriers in graphene are called massless Dirac fermions, where Equation 2.26 describes a 2D analogue of the Dirac equation, with the Fermi velocity playing the role of the speed of light. This allows the investigation of relativistic quantum phenomena in a benchtop experiment [23].

The carbon atoms of graphene have 6 electrons, 2 in the inner shell and 4 in the outer shell. The four outer shell electrons in an individual carbon atom are available for chemical bonding. However, in graphene, each atom is sp^2 hybridised and connected to 3 other carbon atoms on the two-dimensional plane, leaving 1 electron freely available in the third dimension for electronic conduction. These highly mobile electrons are called pi (π) electrons and are located above and below the graphene sheet. These π orbitals overlap and help to enhance the carbon to carbon bonds in

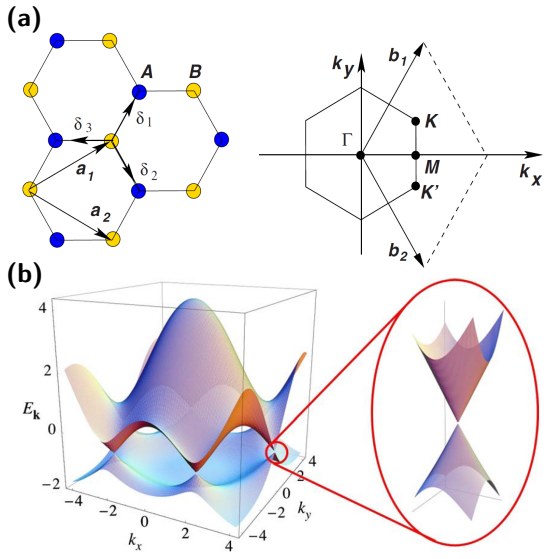


FIGURE 2.5: **Graphene electrical properties. a**, The honeycomb lattice structure of graphene (left) and the corresponding Brillouin zone (right). The Dirac cones, as seen in the electrical dispersion presented in **b**, are located at the K and K′ points. (Reproduced from [22])

graphene. Fundamentally, the electronic properties of graphene are dictated by the bonding and anti-bonding (the valance and conduction bands) of these pi orbitals. [24]

While graphene offers outstanding capabilities for applications such as a thin transparent conductor, we want to concentrate on the optical properties and the potential for nonlinear optical applications. The optical properties of graphene are described by an effective susceptibility, calculated based on graphene's surface conductivity $\sigma(\omega)$ [25]:

$$\chi^{(\mathrm{eff})}(\vec{r},\omega) = \mathrm{i}\frac{\sigma(\omega)}{\omega\varepsilon_0 d}, \tag{2.27}$$

with d as the thickness of the effective layer.

The surface conductivity can be derived by the Kubo model as [25]

$$\sigma(\omega) = \frac{e^2(\omega+\mathrm{i}\gamma_c)}{\mathrm{i}\pi\hbar^2}\left[\int_{-\infty}^{\infty}\frac{|\epsilon|}{(\omega+\mathrm{i}\gamma_c)^2}\frac{\mathrm{d}f_0(\epsilon)}{\mathrm{d}\epsilon}\mathrm{d}\epsilon - \int_0^{\infty}\frac{f_0(-\epsilon)-f_0(\epsilon)}{(\omega+\mathrm{i}\gamma_c)^2-4(\epsilon/\hbar)^2}\mathrm{d}\epsilon\right], \tag{2.28}$$

where $-e$ is the electron charge, $\hbar$ is the reduced Planck constant, γ_c is a phenomenological carrier scattering rate that is assumed to be independent of the energy ϵ, $f_0(\epsilon) = \{\exp[(\epsilon-\mu_c)/k_BT]+1\}^{-1}$ is the Fermi function, μ_c is the chemical potential (controlled with the help of a gate voltage), k_B is Boltzmann's constant and T is the ambient temperature. The first term corresponds to intraband electron-photon scattering processes, whereas the second term corresponds to the direct interband electron transitions.

For the often-made assumption that temperatures are relatively low ($\mu_c \gg k_B T$), the *intraband conductivity* term results in [25]

$$\sigma^{\text{intra}}(\omega) = \frac{\mathrm{i}2e^2|\mu_c|}{\pi\hbar^2(\omega + \mathrm{i}\gamma_c)}, \tag{2.29}$$

and the *interband conductivity* term can be written as [25]

$$\sigma^{\text{inter}}(\omega) = \frac{e^2}{4\hbar}\left[\Theta(\hbar\omega - 2|\mu_c|) + \frac{\mathrm{i}}{\pi}\ln\left|\frac{\hbar(\omega + i\gamma_c) - 2|\mu_c|}{\hbar(\omega + i\gamma_c) + 2|\mu_c|}\right|\right]. \tag{2.30}$$

Considering the high-frequency limit, $\hbar\omega \gg (k_B T, \mu_c)$, the total conductivity takes the universal value of $\sigma_0 = e^2/4\hbar$ leading to the well known ~ 2.3% absorption over the optical wavelength range.

2.3 Plasmons

In this section, an introduction to diverse plasmon resonances is given. It focuses on understanding the origin and properties of surface plasmon polaritons (SPP) and further describes the newly developed field of 2D van der Waals polaritons. The primary resources used for the SPP basics are the review "Surface plasmon subwavelength optics" by W. Barnes et al. [26] and the book "Plasmons: Fundamentals and Applications" by S. Maier [27]. To learn about polaritons in 2D layered van der Waals materials, the reviews from D. Basov et al. [28] and T. Low et al. [29] can be recommended. The paper of S. Campione et al. [30] gives a good introduction into epsilon-near-zero modes in ultrathin films.

2.3.1 Surface Plasmon Polaritons

SPPs are waves that propagate along the interface of a conductor and a dielectric. They still form an electromagnetic mode, but they are now bound to the surface as depicted in Figure 2.6a. These electromagnetic surface waves arise via the coupling of the electromagnetic fields to oscillations of the conductor's electron plasma. This interaction leads to strong confinement of the electric field on the surface, as seen in Figure 2.6b by the exponential decay of field strength away from the interface. Furthermore, the effective wavelength of the SPP-mode decreases the more pronounced the coupling is and the less "light like" the mode behaves. This can be seen in Figure 2.6c, for increasing ω the SPP wavevector k_{SPP} increases stronger compared to the light line.

To calculate k_{SPP} we investigate the flat interface between a dielectric, non-absorbing half-space ($z > 0$) with positive real dielectric constant ε_2 and an adjacent conducting half-space ($z < 0$) described via a dielectric function $\varepsilon_1(\omega)$. The requirement of metallic character implies that $\Re(\varepsilon_1) < 0$. For conductors, this condition is fulfilled at frequencies below the bulk plasma frequency ω_p, as seen in the Drude response given by Equation 2.25. We want to look for propagating wave solutions confined to the interface, i.e., evanescent decay in the perpendicular z-direction. Choosing the TM-mode approach, which provides a finite E_z-component reveals the propagation constant plotted in Figure 2.6c [31]:

$$k_{SPP} = k_0\sqrt{\frac{\varepsilon_1(\omega)\varepsilon_2}{\varepsilon_1(\omega) + \varepsilon_2}}. \tag{2.31}$$

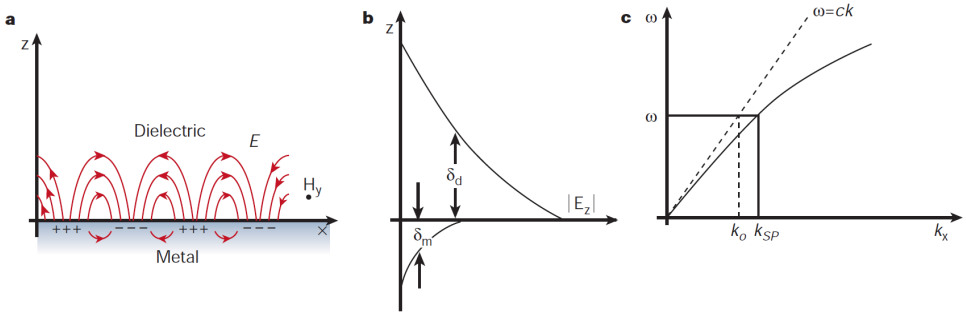


FIGURE 2.6: **Surface plasmon polaritons on a conductor-dielectric interface: a**, Schematic representation of the energy exchange between the electromagnetic wave and the surface charge oscillations while propagating along with the interface. **b**, The electric fields of the surface mode decay exponentially away from the medium. **c**, The dispersion relation of the surface plasmon in comparison to freely propagating light. (Reproduced from [26])

Structuring of the surface allows for further tuning of the plasmonic properties. Metal nano-resonators have become a standard tool to generate plasmonic antennas with structure-dependent resonances. Most established for working in the visible wavelength range are metals such as Gold or Silver. They have their ENZ characteristics at the short-wavelength end of the visible wavelength spectrum. However, going to the NIR range, the imaginary part of metals get too high, and an alternative material is desired. For the telecommunications range, the material of choice is less clear. Transparent conductors seem to have become the material of choice since they can be tuned by doping to shift their ENZ area over the whole range of the telecom wavelength [32].

Confining light to a surface and decreasing the materials structure size offers essential opportunities for the ever-growing demand for miniaturisation in the telecommunications industry. Plasmonics could offer a link between photonic and electrical communication [33]. Furthermore, enhancing local fields using plasmonic tips or surfaces has become a standard tool in biosensing applications [34]. Even high harmonic generation in gases have been reported to be enhanced by the use of gold nano-antennas [35].

2.3.2 Thin Layer Resonances

A more common scenario than the half-space case discussed earlier in the case of a thin layer. To understand the possible modes, we first consider the case of a plasmonic layer surrounded by air (see Figure 2.7). For very thick films ($d \to \infty$), the mode dispersion approximates a half-space surface plasmon dispersion (red solid line). As the film's thickness is decreased past the skin depth, the surface plasmons at each interface begin to interact and form two branches, the long-range and short-range surface plasmon polaritons, whose dispersion varies with the thickness (dashed orange curves for 150 nm films). For ultrathin films, such as the $d = 2$ nm film corresponding to the green curve, the long-range mode becomes nearly flat and asymptotically approaches the dispersion of the plasma frequency within a range of wave numbers. This mode occurs at precisely the frequency at which the permittivity is approximately zero and is referred to as the ENZ mode [36]. The field profiles in Figure 2.7b show the emergence of two surface plasmons into one flat ENZ mode as the film becomes thinner. This ultra-thin layer ENZ mode provides substantial

field enhancement, which grows inversely proportional to the ITO thickness [30]. The bulk plasmon branch also exhibits a flat dispersion above the plasma frequency and is denoted as either the Berreman or the Brewster mode. It can lead to directional dependent perfect absorption (PA) at frequencies corresponding to $\Im(\omega) = 0$ but requires a gold substrate.

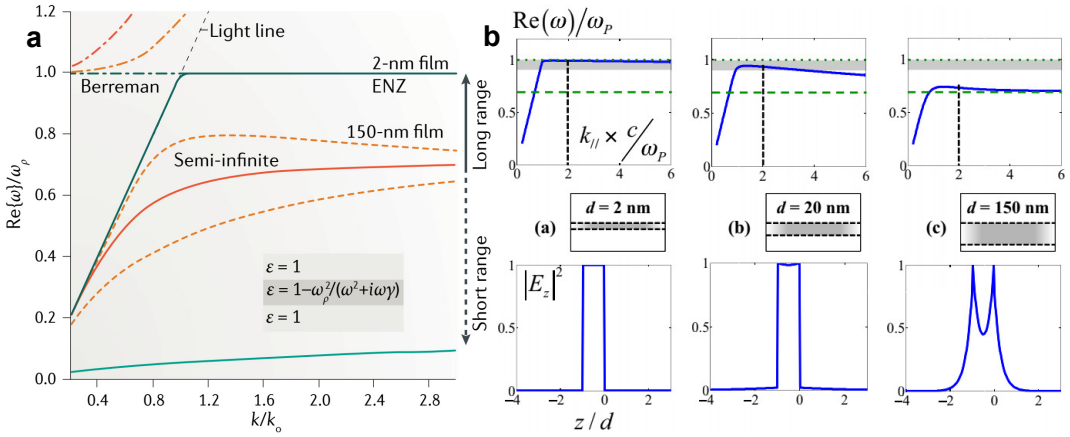


FIGURE 2.7: **Thin layer resonances: a**, The plasmon dispersion branches of thin plasmonic layers of various thicknesses. (Reproduced from [36]) **b**, Dispersion (top) and corresponding profile of the intensity of the normal component of the electric field vs the spatial variable z normalised by the three given thicknesses (bottom). The shaded grey area describes the ENZ mode validity range. (Reproduced from [30])

2.3.3 Directional Perfect Absorption

Ting S. Luk et al. showed that "[d]irectional perfect absorption [can be achieved] using deep subwavelength low-permittivity films" [37]. They demonstrated perfect absorption for deep subwavelength films of indium tin oxide (ITO) on metal substrates at frequencies just above the ENZ frequency, where both the real and the imaginary parts of the ITO permittivity are less than 1. This can be understood within a simple theoretical framework based on the three-medium layer Fresnel reflection:

$$R = \frac{r_{12} + r_{23}e^{2i\Phi}}{1 + r_{12}r_{23}e^{2i\Phi}}, \tag{2.32}$$

where r_{ij} is the reflection coefficient at the i-j-interface, $\Phi = k_{z2}d$ the single pass phase shift and d the thickness of the layer (medium 2). For the assumption of a perfect electric conductor as the third medium ($r_{23} = 1$), we get $r_{12} = e^{2i\Phi}$ as the required condition for zero reflection. This ultimately leads to the condition of:

$$\frac{2\pi d_{\mathrm{PA}}}{\lambda_{\mathrm{PA}}} = \left(\frac{\varepsilon_2'^2 + \varepsilon_2''^2}{\varepsilon_1^{3/2}\varepsilon_2''}\right)\frac{1}{\tan\theta_{\mathrm{PA}}\sin\theta_{\mathrm{PA}}} \tag{2.33}$$

Furthermore, the perfect absorption (PA) regime corresponds to a vanishing imaginary part in the excited mode. At these points, the field in medium 1 is neither growing nor decaying. Instead, it is the field of an incoming plane wave. Thus, at these points, the field structure of this mode, which is an exact solution to Maxwell's

equations, is indistinguishable from the field structure of an incoming and absorbed plane wave. Hence, a plane wave can phase match the mode at these points (and in their close vicinity).

Epsilon-Near-Zero Plasmon

Introducing a glass substrate instead of a reflective metal can lead to similar features in the case of total internal reflection, as demonstrated by S. Campione et al. [38]. The so-called *ENZ plasmon* feature appears just past the critical angle, as can be seen in Figure 2.8. The abrupt absorption change to ~ 100 % right above the critical angle is due to the constructive interference of the incident and reflected waves around the excitation conditions of the mode. The near-perfect absorption and field enhancement make this mode a promising pathway to enhance nonlinear optical processes and open up directions for ultrafast, tunable thermal emission.

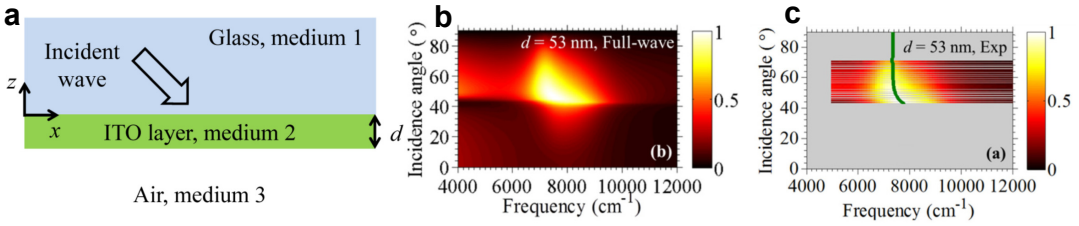


FIGURE 2.8: **Experimental verification of the epsilon-near-zero plasmon.** **a**, Schematic of an incident wave from the glass medium to excite the high wavevector ENZ plasmon (Kretschmann geometry). **b**, Simulation of a 53 nm ITO film with the ENZ plasmon as the high absorption feature just past the light line. **c**, Experiment that confirms the strong absorption feature. (Reproduced from [38])

2.3.4 Graphene Plasmon Polaritons

Polaritons are hybrids of light-matter oscillations that can originate in different physical phenomena. In the case of 2D materials, one could have conduction electrons in graphene and topological insulators (surface plasmon polaritons), infrared-active phonons in boron nitride (phonon polaritons), excitons in dichalcogenide materials (exciton-polaritons), superfluidity in FeSe- and Cu-based superconductors with high critical temperature T_c (Cooper-pair polaritons), and magnetic resonances (magnon polaritons). The family of van der Waals (vdW) materials supports all of these polaritons, as shown in Figure 2.9. The matter oscillation component (e.g. electron plasma oscillation) results in negative permittivity of the polaritonic material (previously given by the conductor half space), giving rise to optical-field confinement at the interface with a positive-permittivity environment. The vdW polaritons exhibit strong confinement, as defined by the ratio of incident light wavelength λ_0 to polariton wavelength λ_p. [28]

Graphene plasmonics has led to demonstrating proof-of-principle device concepts such as mid-infrared optoelectronics [39], bio-sensing [40] and fingerprinting [41]. It is anticipated that these applications should continue to develop over the following years and could also enable free-space beam shaping and steering with graphene metasurfaces [42]. Going forward, these high-quality GPPs can also provide an excellent platform for realising tunable 2D mid-infrared nano-photonics circuits with novel functionalities not previously attainable.

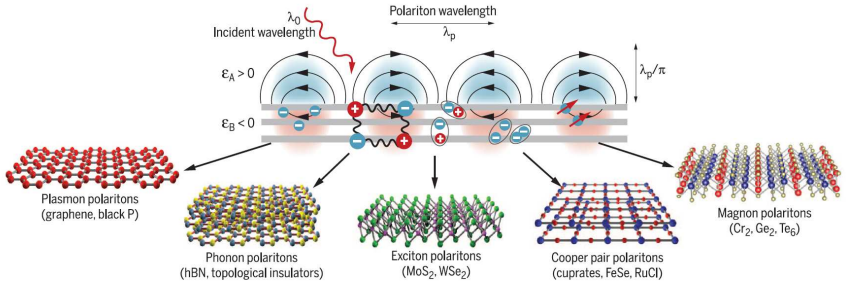


FIGURE 2.9: **Polaritons in van der Waals materials.** Various examples of polaritons in van der Waals materials. (Reproduced from [28])

Graphene plasmons themself provide excellent confinement, but recent research pushed the vertical field confinement even further to the limits. In the work of D. Alcaraz et al. on "Probing the ultimate plasmon confinement limits with a van der Waals heterostructure" [43], they present a graphene-insulator-metal hybrid device that enables vertical plasmon confinement down to the ultimate limit of the length scale of one atom ($\sim \lambda/26,000$). They use far-field light to couple to these strongly confined plasmons and find a vertical mode length down to 0.3 nm, whereas higher-order Fabry-Pérot resonances reveal that the propagating character of the plasmons is preserved. Furthermore, the metal rods facilitate efficient coupling between far-field light and the strongly confined plasmons, where the width of the rods defines the resonant conditions for the plasmon modes.

2.3.5 Excitation of Polaritons

The confinement factors of these 2D layer polaritonic modes reach values up to $\lambda_0/\lambda_p \sim 1000$. While offering excellent field confinement, this also creates an even more substantial wavevector mismatch with freely propagating radiation. Various approaches to overcome this issue have been presented (see Figure 2.10). Polaritons can, for example, be excited at periodic structures [44], using nonlinear wave mixing [45] or antenna-like nanotips [46]. These options maintain phase coherence with respect to the external illumination, in contrast to the inelastic processes of using electron beams [47], localized emitters [48] or electron tunneling [49]. The coherence is lost through intermediate processes such as phonon scattering. Localised polaritons confined to nanoislands can also be resonantly excited by incident light interaction [28].

While these examples provide a suite of good choices for 2D materials, we turn now to more conventional methods, well known from metallic thin film studies over the last decades. Exciting, e.g. a surface plasmon polariton on an air/gold interface is less challenging as the wavevector only slightly surpasses the light line as can be seen in Figure 2.6) for relatively low frequencies. The presented techniques are taken primarily from the book "Plasmonics: Fundamentals and Applications" by Stefan A Maier [27] and a lecture on plasmonics by Shalaev [50].

Prism coupling

To slightly increase the incoming light wavevector for a constant frequency, one can utilise an optically dense medium with $n > 1$. For the same angle of incidence, the in-plane wavevector component is now increased by $k_{\parallel} = n\,k_{\parallel,\mathrm{air}}$. One common way

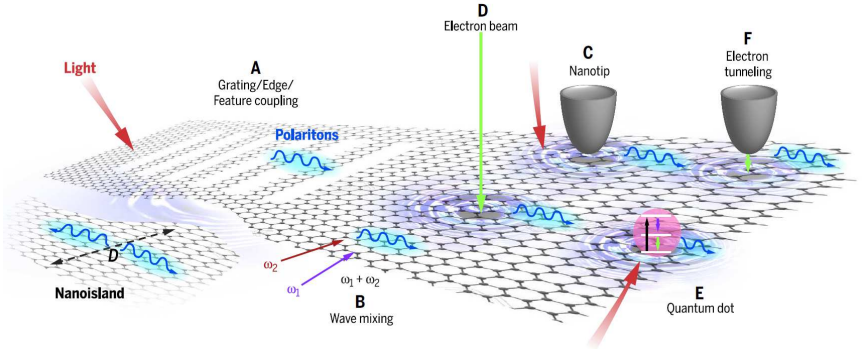


FIGURE 2.10: **Excitation of polaritons in van der Waals materials.** Techniques to excite plasmons as displayed: (A) Periodic structuring, (B) nonlinear wave mixing, (C) antenna-like nanotips, (D) electron beams, (E) quantum dots and localised emitters, and (F) electron tunnelling. (Reproduced from [28])

of implementing this trick is utilising a glass prism, as seen in Figure 2.11. The light beam incident from the prism side now surpasses the maximum air light-line impulse and can match the surface plasmon. All frequency wavevector combinations within the new prism light-line range are now able to be phase-matched. For that, the metal film has to be thin enough to allow for the evanescent fields inside the thin layer to be non-zero at the air/metal interface, on which the SP will be excited.

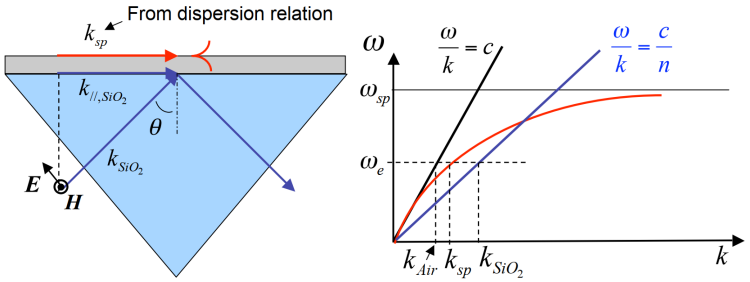


FIGURE 2.11: **Excitation of surface plasmon polaritons via prism coupling.** The refractive index of the prism increases the parallel wavevector component $k_{\parallel}$ and can match that of the surface plasmon k_{sp}. The surface plasmon dispersion is now partly inside the prism light cone (blue) and can be excited. (Reproduced from [50])

Note that the surface plasmon at the prism/metal interface can not be excited as the dispersion relation is shifted and slightly surpasses the prism light line, similar to the in-air case. Further, the excited SPPs using the prism phase-matching method inherently leaky waves, i.e. they lose energy due to the inherent absorption inside the metal and leakage of radiation into the prism: the excited propagation constants lie within the prism light cone (see Figure 2.11). The minimum intensity of the reflected beam is due to destructive interference between this leakage radiation and the reflected part of the excitation beam. For an optimum metal film thickness, the destructive interference can be perfect, providing a zero in the reflected beam intensity, so that leakage radiation cannot be detected. [27]

Grating coupling

Another method to overcome the phase difference is to structure the surface. Here, we investigate a grating pattern with the periodicity of P. According to Bloch, the periodic dielectric constant couples waves for which the k-vectors differ by a reciprocal lattice vector $|\vec{G}| = 2\pi/P$. Hence, phase matching takes place if the wavevectors fulfil the condition of

$$k_{sp} = k_{\parallel} \pm mG, \tag{2.34}$$

with m being an integer. The corresponding coupling mechanism is sketched in Figure 2.12.

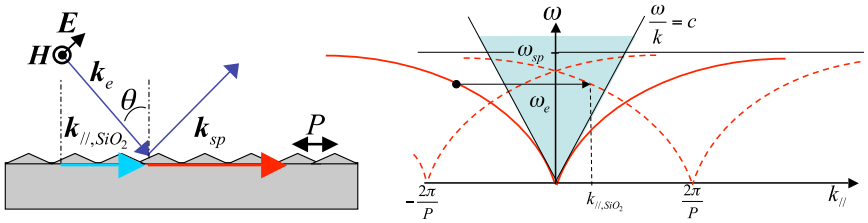


FIGURE 2.12: **Excitation of surface plasmon polaritons via grating coupling.** The grating of periodicity P provides a strong Fourier component at $G = 2\pi/P$. Freely propagating light can now excite the surface plasmon via multiple scattering events on the grating if the shifted surface plasmon branch overlaps with the light cone. (Reproduced from [50])

The reverse process can also occur: SPPs propagating along a surface modulated with a grating can couple to light and thus radiate. The gratings need not be milled directly into the metal surface but can also consist of dielectric material.

For one-dimensional gratings, significant changes to the SPP dispersion relation occur if the gratings are sufficiently deep so that the modulation can no longer be treated as a small perturbation of the flat interface. Appreciable band gaps appear already for a groove depth on the order of 20 nm for metallic gratings. For even larger depths, localised modes inside the grooves lead to distortions of the first higher-order band folded back at the Brillouin zone boundary, enabling coupling even for short pitches $P < \lambda/2$ upon normal incidence due to a lowering in the frequency of the modified SPP dispersion curve.

Antenna coupling

More generally, SPPs can also be excited on films in areas with random surface roughness or manufactured localised scatterers. Momentum components Δk_x are provided via scattering so that the phase-matching condition

$$k_{sp} = k_{\parallel} \pm \Delta k_x \tag{2.35}$$

can be fulfilled. For local scatterers, the spatial Fourier components can reach values up to

$$\Delta k_x \sim 2\pi/d, \tag{2.36}$$

where d corresponds to the scattering body dimension along the electric field direction. In Figure 2.13, the fluorescence images of a singular bar and a dot are shown. The dot, in particular, shows a dipole like radiation pattern. It can excite a surface plasmon as indicated with an arrow in Figure 2.13b.

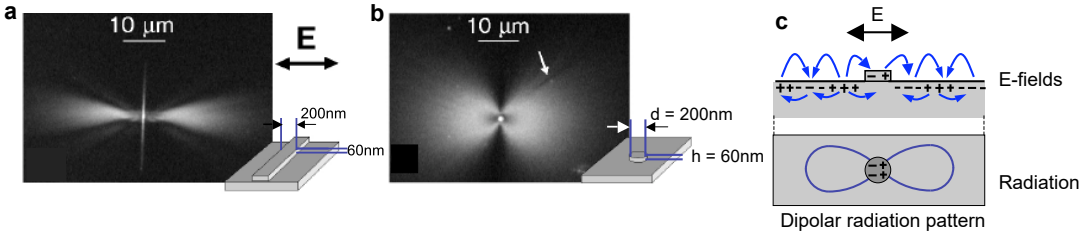


FIGURE 2.13: **Excitation of surface plasmon polaritons via antenna coupling.** Fluorescence images of the field intensity of SPs excited via a silver wire (width 200 nm, height 60 nm, length 20 mm) (**a**) and a silver particle (diameter 200 nm, height 60 nm) (**b**). **c**, The schematic dipolar radiation pattern of the silver particle in side and top view. (Reproduced from [51, 50])

2.3.6 Localised Plasmons in Nano-Antennas

So far, discussions have revolved around the excitation of propagating, dispersive plasmon modes that require phase matching to be excited. Here, we discuss the non-propagating, localised surface plasmons on nanostructures. The interaction of a sphere with radius R with an electromagnetic field can be analysed using the simple quasi-static approximation provided that $R \ll \lambda$, i.e. the particle is much smaller than the wavelength of light. In this case, the phase of the harmonically oscillating electromagnetic field is practically constant over the particle volume so that one can calculate the spatial field distribution by assuming the simplified problem of a particle in an electrostatic field. The harmonic time dependence can then be added to the solution once the field distributions are known. The absorption cross-section for a sphere is given as [52]

$$C_{\text{abs}} = k\text{Im}(\alpha) = 4\pi k R^3 \text{Im}\left(\frac{\varepsilon - \varepsilon_m}{\varepsilon + 2\varepsilon_m}\right), \tag{2.37}$$

with surrounding medium permittivity ε and the metal permittivity ε_m.

L. Novotny [53] showed that these nano-antennas basically act like scaled down versions of their radio-frequency counterparts as sketched in Figure 2.14. However, an effective wavelength scaling must be applied. For example, for a half-wave antenna made of a gold rod (length $L = 110\,\text{nm}$, radius $R = 5\,\text{nm}$) one calculates $\lambda_{\text{eff}} = \lambda/5.3$ ($L = \lambda/10.6$). More generally, the effective wavelength may be estimated as

$$\lambda_{\text{eff}} = a + b\frac{\lambda}{\lambda_{\text{p}}}, \tag{2.38}$$

where λ_{p} is the plasma wavelength and a, b are coefficients with dimensions of length that depend on antenna geometry and static dielectric properties. The assumptions are that the antenna is made of linear segments with radius $R \ll \lambda$ and that the metal can be described by a free electron gas according to the Drude model.

Nano-antennas have become a standard tool to enhance (nonlinear) optical processes. They provide a unique capability to enhance electric fields and direct optical energy to defined local hot spots.

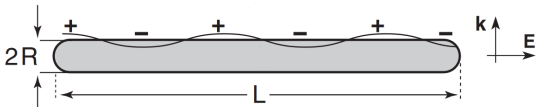


FIGURE 2.14: **Schematic of a nano-antenna.** A single antenna element is represented by a metal rod. Incident light polarises the rod ends and gives rise to a standing surface charge wave. (Reproduced from [53])

2.4 Nonlinear Optics in Plasmonic Thin Layers

Nonlinear plasmonics has become a fertile ground for developing and testing new ideas on light-matter interaction under extreme field conditions. These ideas have found many applications in surface science, active photonic nanodevices, near-field optical microscopy, and nonlinear integrated photonics [54]. For catching up with the field of ENZ nonlinear optics, I recommend the reviews of Reshef et al. [55] and Kinsey et al. [36]. Here, we explore the latest developments of nonlinear optics in graphene nanostructures and thin layer ENZ materials. We focus on research directly related to the projects presented in this thesis.

2.4.1 Graphene Nonlinear Optics

Graphene has a hexagonal lattice structure, which makes it centrosymmetric. Hence, no second-order nonlinearity such as SHG is expected under normal incidence. THG and other odd-order interactions are allowed and particularly strong in graphene, as will become apparent in this section.

The potential of graphene for nonlinear optics was first shown by E. Hendry et al. in the "Coherent Nonlinear Optical Response of Graphene" [5]. They performed the first measurements of the coherent nonlinear optical response of single- and few-layer graphene using four-wave mixing. Graphene exhibits an extreme nonlinear optical response in the near-infrared spectral region. The sizeable optical nonlinearity originates from the interband electron transitions and is eight orders of magnitude larger than the nonlinearities observed for dielectric materials without such transitions.

For our research projects, we are interested in two perspectives of combining GPPs and nonlinear optics, similar to the case of metal before. First, the tool of nonlinear optics is utilised to excite GPPs, as seen in the wave-mixing example (B) in Figure 2.9b. Second, the direct excitation of GPPs in patterned graphene samples is utilised to confine incoming electric fields to the nonlinear medium and, by that, further increasing the nonlinear signal.

Nonlinear Optical excitation of Graphene Surface Plasmons

Exciting the ultra confined graphene plasmon polaritons requires phase matching. An interesting approach was chosen by T. Constant el al. by investigating "All-Optical Generation of Surface Plasmons in Graphene" [45]. They take advantage of the intrinsic second-order nonlinear optical response of graphene for angled incidence. Free-space, visible light pulses are used to generate surface plasmons in a planar

graphene sheet using difference frequency wave mixing to match the wavevector and energy of the surface wave. By controlling the phase-matching conditions, they showed that one could excite surface plasmons with a defined wavevector and direction across an extensive frequency range, as sketched in Figure 2.15.

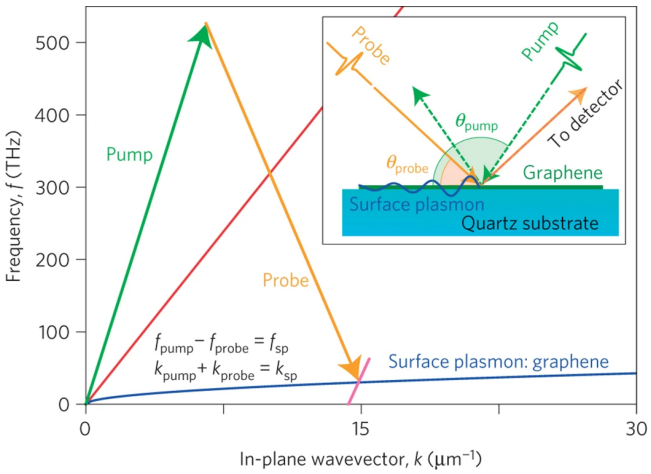


FIGURE 2.15: **All-optical generation of surface plasmons in graphene.** The difference frequency generation of the pump (green arrow) and probe (orange arrow) allows access to wavevectors outside of the light line (red line). This permits phase matching to the surface plasmon modes in graphene (blue line). (Reproduced from [45])

This concept has been used by B. Yao et al. [56] to generate "Broadband gate-tunable terahertz plasmons in graphene heterostructure". They demonstrated terahertz plasmon generation and control in chip-scale integrated graphene. The coherent difference frequency excitation is gate-tunable for both graphene layers via electrical doping and manipulation of the chemical potential. Together with the symmetric–antisymmetric frequency crossing between the two layers a frequency range of 4.7 to 9.4 THz can be tuned across. The incoming wavelength was chosen to be in the range of 1530 to 1610 nm, making this device an excellent example for future telecommunications applications.

However, both of these studies presented $\chi^{(2)}$ values that are roughly three orders of magnitude higher than perturbation theory currently suggests. In comparison, other theoretical studies presented in B. Yao et al. [56], but originally proposed by X. Yao et al. [57], coincide roughly with the experimental data, their model shows wrong qualitative behaviour with an experimentally not verifiable divergence for difference frequencies approaching zero. This issue was highlighted and discussed in "Origins of All-Optical Generation of Plasmons in Graphene" by C. Tollerton, **J. Bohn**, et al. [58], which can be found in Appendix A. This work has also been extensively featured in chapter 5 of C. Tollertons doctoral thesis [59].

Enhanced high-harmonic generation

Graphene and its versatile properties have been utilised to generate new frequency components, including the modulation and enhancement of said process. The "Gate-tunable third-order nonlinear optical response of massless Dirac fermions in graphene" was studied by T. Jiang et al. [60]. They demonstrated that the third-order nonlinearity of graphene can be varied by orders of magnitude with the help of gate-controlled doping or shift of the chemical potential. In general, the optical

nonlinearity of linear-band materials with the chemical potential close to the Dirac or Weyl point tends to diverge when the input frequency combination approaches zero. The resulting giant nonlinearity of such materials, particularly graphene, can be of great use in future optoelectronic devices.

H. Hafez et al. reported an "Extremely efficient terahertz high-harmonic generation in graphene by hot Dirac fermions" [61]. They report the generation of terahertz harmonics up to the seventh order in single-layer graphene at room temperature and under ambient conditions, driven by terahertz fields of only tens of kilovolts per centimetre, and with field conversion efficiencies over 10^{-3}, 10^{-4} and 10^{-5} for the third, fifth and seventh terahertz harmonics, respectively. These conversion efficiencies are remarkably high, given that the electromagnetic interaction occurs in a single atomic layer. The effective nonlinear optical coefficients of graphene for the third, fifth and seventh harmonics exceed the respective nonlinear coefficients of typical solids by 7-18 orders of magnitude. This provides a direct pathway to highly efficient terahertz frequency synthesis using the present generation of graphene electronics, which operate at much lower fundamental frequencies of only a few hundreds of gigahertz. However, for the NIR/VIS wavelength range, orders of magnitude less efficient high harmonic generation has been presented with efficiencies of 10^{-10}, 10^{-13} for third and fifth harmonics for 5-monolayer graphene [62].

"Plasmon-Enhanced Nonlinear Wave Mixing in Nanostructured Graphene" has been theoretically studied by J. Cox et al. [63]. They investigated patterned graphene to realise nonlinear wave mixing on the nanoscale with extraordinarily high efficiencies, surpassing metal nanoparticles of similar lateral sizes. The large magnitudes of the predicted nonlinear polarizabilities are attributed to plasmonic enhancement in these nanostructures. These plasmons can be tuned by changing the number of doping charge carriers, adding another advantage for the conventional plasmonic metal response. These effects can lead to tunable plasmonic response and plasmon-induced enhanced nonlinearities within the visible and near-infrared spectral ranges. These results configure a platform for developing nanoscale nonlinear optical devices based upon graphene nanostructures with lateral dimensions of only a few nanometers.

As described in subsection 2.3.4, D. Alcaraz et al. [43] even further decreased the lateral confinement down to the dimensions of an atom, suggesting an even more substantial effect on the nonlinear efficiency. While not yet being able to couple to graphene plasmons, I. Alonso et al. showed giant enhancement of third-harmonic generation in these graphene–metal heterostructures, as seen in Figure 2.16 [64]. In chapter 4, we jointly investigated enhanced fifth harmonic generation in similar devices.

2.4.2 Epsilon-Near-Zero Nonlinear Optics

Recently, an unprecedented nonlinear refractive index change of unity order has been observed in ITO [8]. To get an intuition of why epsilon-near-zero (ENZ) materials have a strong optical Kerr effect, one can recall the standard relation between n_2 and $\chi^{(3)}$ as [9]:

$$n_2 = \frac{3\chi^{(3)}}{4\varepsilon_0 c n_0^2}, \tag{2.39}$$

with ε_0 as the permittivity of free space. Considering $n = \sqrt{(\varepsilon)}$, it becomes clear that n_2 gets large for a vanishing denominator given via $\mathrm{Re}(\varepsilon) \approx 0$. Alam et al. [8] report values of $n_2 \approx 10^{-10}$ cm^2/W and $\Delta n \approx 0.72$, which is even more impressive

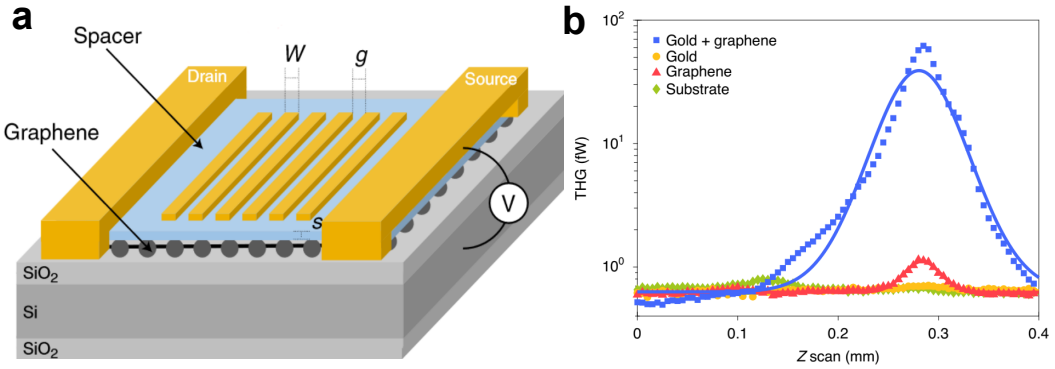


FIGURE 2.16: **Giant enhancement of third-harmonic generation in graphene–metal heterostructures . a**, graphene is encapsulated by a few-nm-thick Al_2O_3 or a monolayer h-BN film, which sets the space s between the graphene and gold nanoribbons. The gold nanoribbon arrays are characterised by the ribbon width W and the inter-ribbon gap g. **b**, Third harmonic generation measurements on four different regions of the sample depending on the sample displacement relative to the optical focus along the optical axis (symbols). Gaussian fits to the data (curves) provide visual guides of the THG signal. (Reproduced from [64])

considering the initial refractive index of $n_0 \approx 0.4$ is smaller then the nonlinear change itself (see Figure 2.17). This corresponds to a permittivity change from $\varepsilon = 0 + i\,0.352$ to $\varepsilon = 1.22 + i\,0.61$. Hence, ITO exhibits a reversible transition from metallic to a lossy dielectric state with a subpicosecond time response at wavelengths slightly longer than the bulk plasmon wavelength.

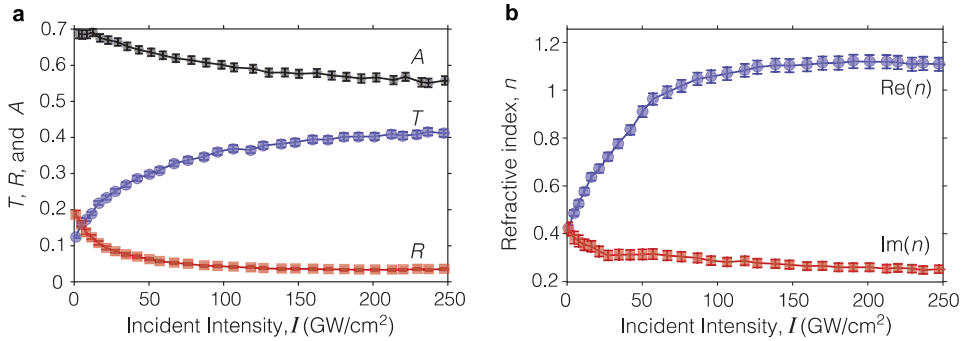


FIGURE 2.17: **Nonlinear optical properties of an ITO thin film. a**, Intensity-dependent transmittance (T), reflectance (R), and absorptance (A) of an 310 nm ITO-glass structure (see Figure 2.4) for $\theta = 30°$. **b**, Complex effective refractive index of ITO extracted from the measured values in (A) using a transfer-matrix method. (Reproduced from [8])

These results challenge the notion that the nonlinear optical response is only a perturbation to the linear response and has also been found in other ENZ materials such as AZO [65, 66] or CdO[67]. Materials with a sizeable nonlinear response are expected to enable exotic nonlinear dynamics and allow all-optical control of metasurface and active plasmonic devices. This allows for a completely new paradigm in nonlinear optics and opens new avenues for developing optical nanostructures with large nonlinearity for applications in nano-photonics, plasmonics, and nonlinear

nano-optics.

Physical origin of nonlinear refractive index changes

The significant refractive index modulation originates from ultrafast changes to the electron distribution of the TCO. Upon pumping the material, one of two typical reflection changes can be expected depending on the pump energy, as depicted in Figure 2.18. These can best be understood by considering changes to the bulk plasmon frequency given as

$$\omega_p^2 = \frac{e^2 n}{\varepsilon_0 m_{\text{eff}}}, \tag{2.40}$$

with e as the elementary charge of an electron, n as the density of electrons, ε_0 as the permittivity of free space and m_{eff} as the effective mass.

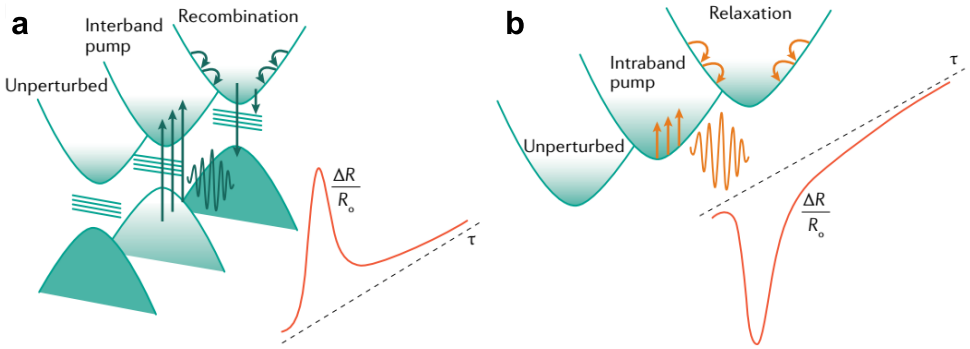


FIGURE 2.18: **Nonlinear dynamics in transparent conductive oxides.** **a**, For a pump pulse with an energy larger than the bandgap, additional electrons are promoted to the conduction band. The increase of the plasmon frequency leads to a more metallic behaviour, typically resulting in an increased $\Delta R/R_0$. **b**, For a pump pulse with energy smaller than the bandgap, the electrons inside the conduction band are heated up. The redistribution of carriers in the non-parabolic band leads to an increased effective mass. Hence, the reflection is expected to decrease due to the reduction of the plasma frequency. (Reproduced from [36])

First, for optical pulse energies larger than the bandgap, the dominant contribution to ultrafast transients comes from interband transitions. Electrons that reside in the valence band are promoted into the conduction band and increase the total free-carrier density (n), blueshifting the plasma frequency. Second, and most relevant for the studies in this thesis, when an ultrafast optical pulse with energy smaller than the bandgap excites a metal oxide, the conduction-band electrons undergo intraband transitions via free-carrier absorption. Before the pulse arrives, the conduction-band electrons are in equilibrium and described by a room temperature Fermi distribution, in which the Fermi energy resides inside the conduction band. Immediately following excitation, the electrons are energised, and their distribution is highly non-thermal; however, within a few femtoseconds, the excited plasma relaxes via electron-electron scattering to a smeared-electron distribution described by a quasi-Fermi distribution and an elevated electron temperature T_e. The result is an increase in the average effective mass of the electron sea due to an overall shift to higher energy, higher-mass

(m_{eff}) states within a non-parabolic conduction band, thereby producing a redshift of the Drude plasma frequency. [36]

F. Zhang et al. studied the origin of the fast temporal dynamics in ITO [68]. The relaxation dynamics are more than an order of magnitude faster compared to more common plasmonic material like gold, as seen in Figure 2.19. The recovery process is found to be determined mainly by the electron-phonon scattering process. They propose that the two-order magnitude higher electron-phonon scattering rate in ITO, as well as in AZO, results in a sharper free electron temperature drop and hence the faster optical response compared with Au. The sub-ps switching capabilities make TCOs an auspicious material choice for ultrafast all-optical switching applications.

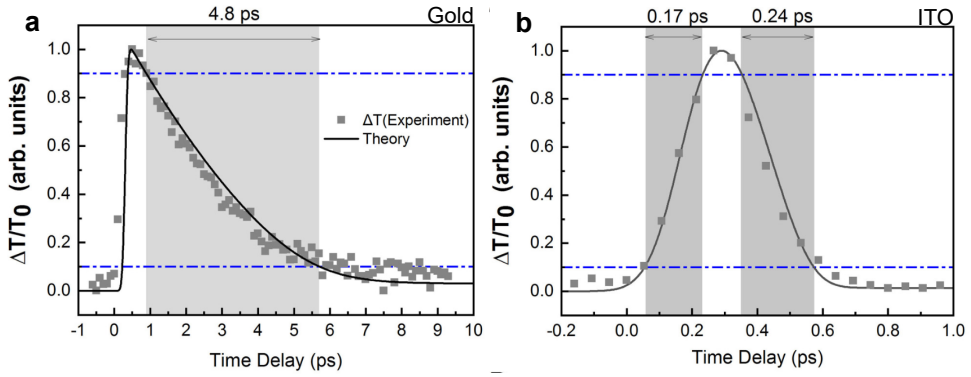


FIGURE 2.19: **Relaxation time: Gold vs ITO.** Time dependence of the normalised optical response (denoted by the square dot) and the convolution of the time evolution of the electron temperature and the temporal envelope of the probe pulse for (**a**) gold and (**b**) ITO, respectively. (Reproduced from [68])

All-optical switching

The shifting of the plasma frequency and corresponding resonances opened an existing pathway for all-optical switching. Already Alam et al. [8] showed significant absolute reflection and transmission changes for simply utilizing a 310 nm ITO film as seen in Figure 2.17.

To maximise this effect, Yang et al. [67] used CdO, which offers higher carrier mobility, in combination with a gold substrate. They measured massive absolute reflection changes of 86.3 % (see Figure 2.20). However, this was measured for 2.08 µm and required an additional gold film, acting as a cavity.

In chapter 5, we utilise a 60 nm ITO film in a prism geometry to excite the ENZ plasmon, leading to near-total absorption at wavelength compatible with the telecom sector. Significant absolute reflection changes of 45 % have been measured, corresponding to switching across almost two orders of magnitude.

Antenna coupling

To further enhance the nonlinear response, nanostructures have been added on top of the ITO layer. Z. Alam et al. [69] investigated the "Large optical nonlinearity of nanoantennas coupled to an epsilon-near-zero material". They measured a broadband (~ 400 nm bandwidth) and intensity-dependent refractive index n_2 as large as $-3.73 \pm$

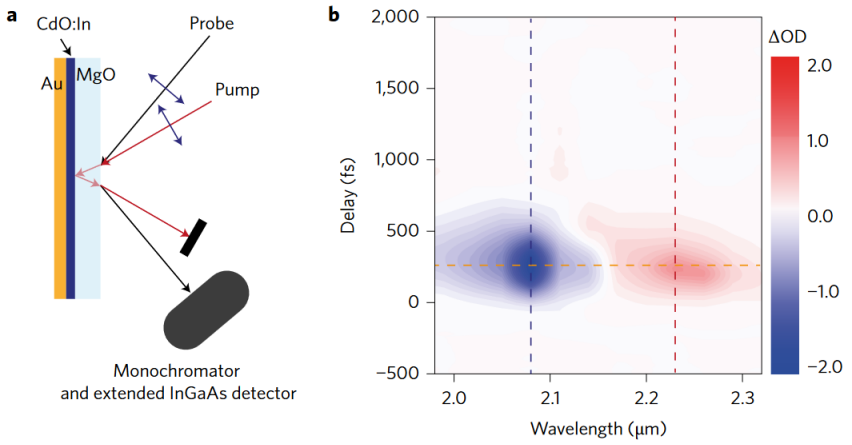


FIGURE 2.20: **Optical switching using a cadmium oxide-based perfect absorber. a**, Schematic of the pump-probe measurement setup. **b**, The ΔOD map as a function of wavelength and delay time. (Reproduced from [67])

$0.56\,\mathrm{cm^2\,GW^{-1}}$. This leads to a maximum optically induced refractive index change of ±2.5 over a spectral range of ~ 200 nm.

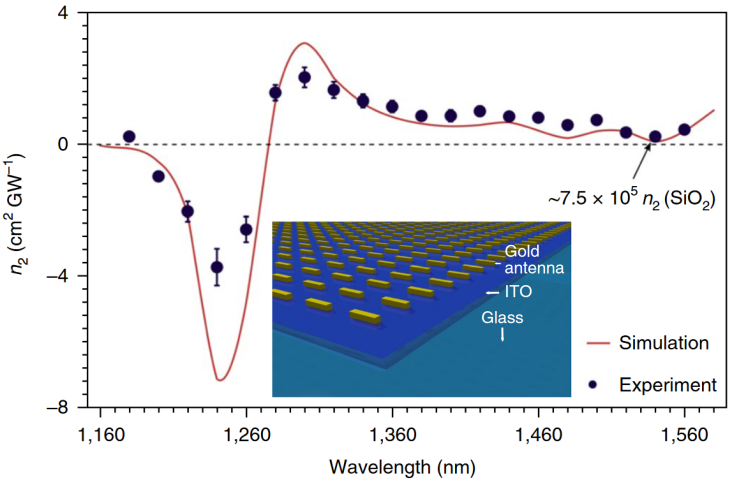


FIGURE 2.21: **Large optical nonlinearity of nano-antennas coupled to an epsilon-near-zero material.**, Effective nonlinear refractive index n_2, measured by displacing the sample along the optical axis relative to a tight optical focus. The metasurface structure comprises a gold dipole antenna array of 27 nm thickness on a 23-nm-thick indium tin oxide (ITO) layer. The intensity-dependent refractive index of the metasurface is almost six orders of magnitude larger than that of SiO_2, even for a highly red-detuned wavelength (λ = 1540 nm). (Reproduced from [69])

Bruno et al. [70] studied "Negative Refraction in Time-Varying Strongly Coupled Plasmonic-Antenna–Epsilon-Near-Zero Systems". They utilise the plasmonic resonance of the metal resonators to strongly interact with the optical ENZ modes of the AZO film. The optical pumping induced a nonlinear polarisation oscillating at double the frequency, responsible for an efficient generation of a phase conjugate and

a negative refracted beam with a conversion efficiency of more than four orders of magnitude greater than the bare ENZ film.

However, these designs only work for one polarisation. In chapter 6, we manipulate both polarisations independently using a cross-shaped nano-antenna and measure a more than seven times stronger response for the probe being parallel to the pump polarisation compared to the perpendicular case.

Frequency control

The nonlinear response of ENZ TCO materials can additionally be utilised for frequency manipulation. T. S. Luk et al. studied "Enhanced third-harmonic generation from the epsilon-near-zero modes of ultrathin films." The conversion efficiency of more than 10^{-6} is achieved by exploiting the field enhancement properties of the epsilon-near-zero mode with an enhancement factor of 200 [71]. Utilising the high mobility CdO on gold, similar to the previously discussed polarisation switch, Y. Yang et al. [72] measured significant high harmonic generation up to the 9th order. Additionally, they find that ENZ-assisted harmonics exhibit a pronounced spectral redshift and linewidth broadening, resulting from the photoinduced electron heating and the consequent time-dependent ENZ wavelength of the material.

Beyond harmonic generation, ENZ materials have only recently become prominent for efficient frequency shifting via temporal refraction. The refraction at a temporal boundary was found to cause a frequency shift of up to 6 % in a 620 nm ITO sample for a broad frequency range around the ENZ case [13]. Similar effects have been reported for AZO [73, 74]. Nano-antennas have been employed to reduce the intensity requirements by more than two orders of magnitude [75].

However, in this temporal refraction case, the frequency shifts are always directly related to the temporal changes to the effective refractive index of the active medium. In chapter 7, we study the effect of spatiotemporal refraction. In addition to the bulk's temporal response, the boundary conditions change as the pulse propagates through the thin layer. These spatiotemporal effects can create a counteracting frequency shift and tailor the desired frequency shift without adjusting the nonlinear refractive index or thickness.

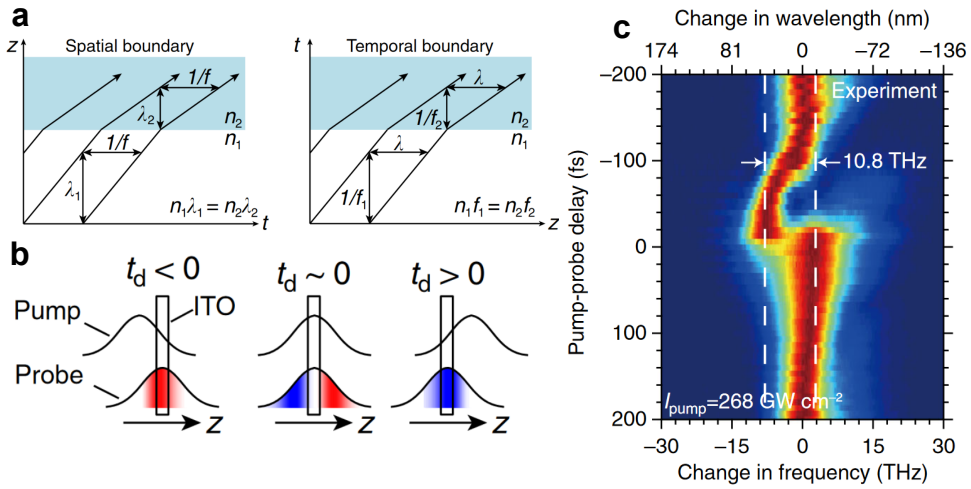


FIGURE 2.22: **Frequency translation through time refraction in an epsilon-near-zero material. a**, A comparison of light refraction at a spatial (left) or a temporal (right) refractive index boundary. Note that the horizontal (left) and vertical axis (right) represent a temporal and not a second spatial coordinate as commonly used in similar representations of spatial refraction.**b**, In experiments, the frequency of the probe redshifts (blueshifts) if the pump beam lags (leads) the probe, considering the rise of the refractive index upon pumping. At near-zero delay, both redshift and blueshift can occur. **c**, Experimental probe spectrum as a function of the pump-probe delay time for 1235 nm. The spectral magnitude for each pump-probe delay is normalised individually. (Reproduced from [13])

Chapter 3

Methodology

Abstract

The methods required to carry out and model the experiments are presented. For this thesis, we utilised two different approaches to characterise the nonlinear response of the active material. A z-scan setup was used to measure high harmonic generation exclusively in the intense focus position. For ITO, the pump-probe setup allowed for a time-resolved study of ultrafast changes to the transmission and reflection, including shifts in their respective spectra. Regarding modelling, we will start with the response of a singular interface and end with the multi-layer transfer matrix method (TMM) applied to measurement examples.

3.1 High Harmonic Generation in Graphene: z-Scan Setup

In this section, we first describe the basic idea of the z-scan setup for high-harmonic generation. The experimental setup described here is part of the laboratories at the University of Vienna and used to do the final measurements presented in chapter 4.

3.1.1 The Concept of a z-Scan Measurement

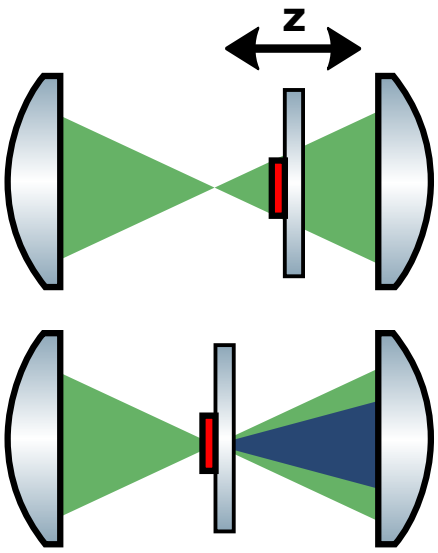


FIGURE 3.1: **z-Scan schematic.** Nonlinear optical effects require large intensities. The sample is moved along the propagation direction (z) through a tight focus. Only when the sample is centred in the intense focus high harmonic signals are generated.

Nonlinear optics describes the optical properties of materials for high intensities. For the example of high harmonic generation, the efficiency of the n-th order is expected to scale with an intensity dependence of $I^{n/2}$ (see subsection 2.1.3). Hence, isolation of HHG signals may be achieved by comparing a high-intensity measurement to a low-intensity reference. In the case of a z-scan setup, this is done by moving the sample of interest through a tightly focussed beam as depicted in Figure 3.1. The strongest HHG signal will be achieved when the active material is precisely in focus. The Rayleigh length is proportional to the spatial resolution of the setup, which can be estimated via gaussian optics. For graphene, it will prove helpful to compare the response of the front and back of the sample to isolate the nonlinear contribution of the monolayer material.

3.1.2 The Measurement Setup

To measure the nonlinear signals of our samples, we used a modified z-scan setup, where the nonlinear signal (either third-harmonic generation or fifth-harmonic generation) is measured while the sample is moved along the z-axis through the focus of the laser beam (see Figure 3.2).

Our pump beam is a linearly-polarized pulsed laser with a ~ 260 fs pulse width, a central wavelength of 3900 nm (0.225 eV), and an a 76.2 MHz repetition rate. This beam is generated by an Optical Parametric Oscillator (OPO), fed by a mode-locked femtosecond Ti:sapphire laser. We use a Half-Wave Plate (HWP) to rotate the polarisation of the incoming beam to that set by the polariser (pol). A lens with a 5.95 mm

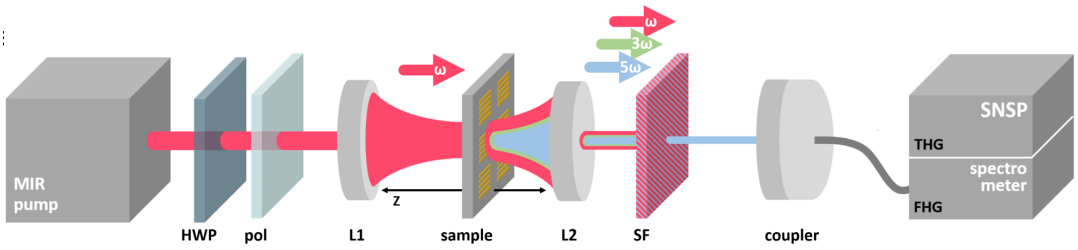


FIGURE 3.2: **z-Scan setup for high-harmonic generation.** Experimental setup of the modified z-scan measurement. The difference-frequency generator (DFG) provides linearly-polarised femtosecond laser pulses at an 76.2 MHz repetition rate. The half-wave plate (HWP), together with the polariser (pol), rotate the polarisation of the light that is then focused down with a short-focal lens (L1). A second collimation lens (L2) collimates both the initial light and the TH signal excited in the sample that is moved in the z-direction through the beam's focal point. Finally, a spectral filter (SF) isolates the nonlinear signal sent to a superconducting nanowire single-photon detector (SNSP) and a spectrometer when measuring THG and FHG, respectively.

focal length focuses the pump down to a waist of ≈ 13.5 µm. When the sample is moved parallel to the pump beam (along the z axis), the nonlinear emission occurs most efficiently where the fluence is maximum, i.e. at the focal point. Afterwards, a lens with an 11 mm focal length collimates the beam, which is then sent through a set of spectral filters that separate the nonlinear signals from the pump beam. The resulting nonlinear signal is coupled into a multimode fibre, which can be sent to two detectors. The third-harmonic signal, at ≈ 1300 nm, is measured with a large-area Superconducting Nanowire Single-Photon (SNSP) detector from PhotonSpot with a ≈ 60 % detection efficiency at this wavelength. We can combine this detector with a Gemini interferometer from NIREOS to measure the spectrum of the third-harmonic. The fifth-harmonic signal, which is centred around 780 nm, is sent to a single-photon sensitive silicon Andor spectrometer with a resolution of 1 nm, and detection efficiency of ≈ 10 %.

3.1.3 Extracting the Third-Order and Fifth-Order Susceptibility

Experimentally $\chi^{(3)}$ and $\chi^{(5)}$ are estimated following the procedure reported in [60]. The measured average power is proportional to the modulus square of the electric field through the following relation:

$$P(\omega) = \frac{1}{8}\left(\frac{\pi}{\ln 2}\right)^{3/2} f\tau W^2 n\epsilon_0 c\frac{|E(\omega)|^2}{2} \tag{3.1}$$

where f is the repetition rate of the laser, τ is the temporal pulse width, W is the Gaussian beam diameter, n is the refractive index, and ϵ_0 and c are the permittivity and speed of light in vacuum. In the experiment here reported f = 76.2 MHz, τ = 260 fs (FWHM), W = 27 µm and the refractive index is considered to be constant, n = 2.4.

The generated THG and FHG electric fields in a graphene monolayer of thickness d_{gr} = 0.33 nm are related to the electric field at the fundamental pump frequency $E(\omega_0)$ through the following relations [9, 76]:

$$E(3\omega_0) = i\frac{\omega_0}{8nc}\chi^{(3)} d_{gr} E(\omega_0)^3 \tag{3.2}$$

$$E(5\omega_0) = i\frac{\omega_0}{8nc}\chi^{(5)} d_{gr} E(\omega_0)^5 \tag{3.3}$$

From Eq. 3.2-3.3 it is thus possible to calculate the effective $\chi^{(3)}$ and $\chi^{(5)}$. Note that all the quantities in Eq. 3.1-3.2-3.3 are considered in standard unit (SI) system, while the $\chi^{(3)}$ and $\chi^{(5)}$ values we report later, e.g. in Fig. 4.6c-d, are in electrostatic unit of charge (esu). The relative conversion factors are given by:

$$\chi^{(3)}(\text{SI}) = \frac{4\pi}{(3 * 10^4)^2}\ \chi^{(3)}(\text{esu}) \tag{3.4}$$

$$\chi^{(5)}(\text{SI}) = \frac{4\pi}{(3 * 10^4)^4}\ \chi^{(5)}(\text{esu}) \tag{3.5}$$

3.2 Time-Resolved Nonlinear Optics in ITO: Pump-Probe Setup

This section will first introduce the basic concepts of time-resolved measurement using a pump and a probe beam. Then we show and explain the experimental implementation in our lab.

3.2.1 The Concept of a Pump-Probe Measurement

To do time-resolved measurements of nonlinear processes, it is typically required to use two beams. One highly intense beam (pump) interacts with the sample. In the case of ITO, electrons would absorb the energy and be redistributed inside the conduction band. The resulting changes to the optical properties can then be tested with a much weaker beam (probe) to prevent optical induced changes by the probe itself. The two beams are delayed relative to each other to resolve the material changes temporally, allowing for a time resolution limited by the pulse length (~ 100 fs in this study). The process is sketched in Figure 3.3 and shows the example of an ultrafast increase of transmission upon pumping.

3.2.2 The Measurement Setup

For the pump-probe measurements (see Figure 3.5), we used an amplified Ti:sapphire laser (Legend Elite, Coherent), with a central wavelength of 800 nm and a repetition rate of 1050 Hz, feeding two identical OPAs (TOPAS, Light Conversion). The final pulse duration was measured via autocorrelation of our pulses at 1200 nm using an APE pulseCheck autocorrelator (see Figure 3.4). The measured pulse length is 107 ± 5 fs, with no significant change in pulse length over the range in wavelengths employed in this work. The measurement does not indicate any chirp like behaviour.

The signal output of one OPA was used as the pump, and the signal output of the other OPA was used as the probe, allowing us independent control of pump and probe frequencies. The combination of a variable $\lambda/2$ wave plate followed by a polariser is used to maximise the initial power of the light sources for an arbitrary incoming frequency. The time delay between the two beams is implemented with a delay line in the pump path, consisting of a linear translation stage with a retro-reflecting pair of mirrors.

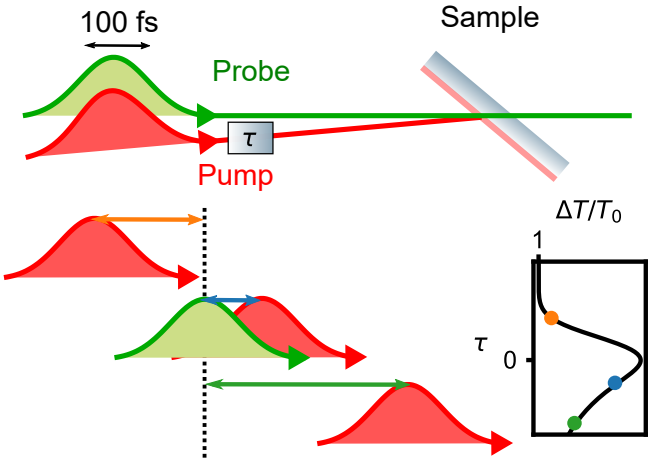


FIGURE 3.3: **Pump-probe schematic.** A pump pulse excites the sample. The much weaker probe pulse arrives with a time delay τ relative to the pump. This way, the ultrafast pump induced changes to the optical transmission of the probe can be resolved down to ~ 100 fs (pulse length).

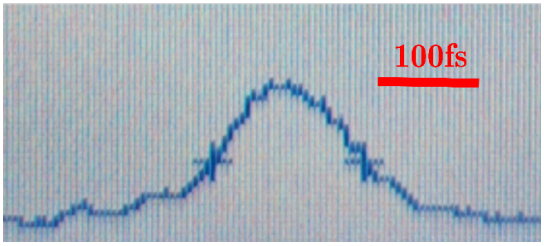


FIGURE 3.4: **Autocorrelation for pulse length and quality check.** The autocorrelation was measured for a wavelength of 1200 nm. The pulse has been measured multiple times, leading to a pulse time estimate of 107 ± 5 fs.

We measure the power of a glass reflection (pump) and ND filter reflection (probe) to track the stability of both optical parametric amplifiers. We used the pump power reference combined with a computer-controlled variable density filter wheel to set a frequency-independent pump intensity on the sample. To make sure the intensity of the probe is significantly smaller than the pump, we used several additional OD filters to decrease the probe power and tested that the nonlinear signal was independent of adding/removing filters. Hence, problems such as a saturated nonlinear electronic response in the detector are avoided. The pump was focused using either a 30 or 40 cm BK7 lens, leading to a 480 µm or 800 µm Gaussian beam diameter (1/e) at the sample. The probe is focussed with a 25 cm $CaFl_2$ lens to a 250 µm diameter.

The probe beam was chopped and measured with a fibre-connected detector. We used a PbS Amplified Detector (PDA30G-EC) combined with an electronic lock-in setup for precise measurements of ultrafast changes to the transmission or reflection. Ideally, one would chop the pump, but chopping the probe is helpful to study the absolute changes if you have a well known and accessible to measure reference, such as total internal reflection. Also, by adding a rotatable quarter wave-plate in front of the analyser, one can extract all four Stokes parameters and analyse the complete polarisation ellipses [77]. We employ the method of Schaefer et al. [78] to extract the Stokes parameters via an 8 point rotation of the quarter-wave plate.

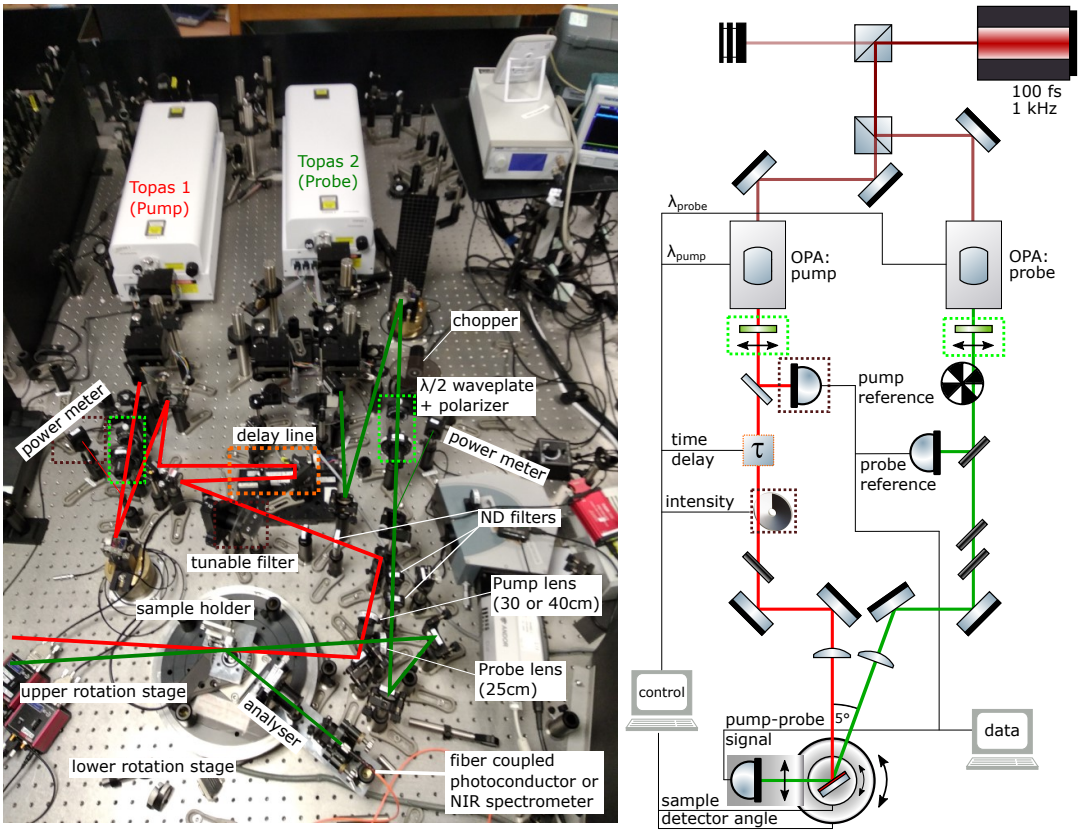


FIGURE 3.5: **Pump-probe setup for ultrafast ITO studies.** The setup as found in the lab can be seen on the left. Experimental devices and the beam paths are marked. On the right, one can see a schematic diagram. The labelling focuses on the electrical control and data collection of the setup.

For spectral measurements, we used an Andor Shamrock 163 spectrograph with a DU490A-1.7 camera. The spectra can be integrated over wavelength to get a good estimate of the transmission and reflection.

The angle of incidence of the pump is 5° smaller than that of the probe. However, for the Kretschmann geometry used to excite the ENZ plasmon, as seen in chapter 5 , we slightly adjusted the setup. The ITO thin layer was positioned on top of a 180 µm coverslip and attached to the front of a right-angle prism (EKSMA, UV-FS 5x5 mm) by applying an index matching fluid (Olympus IMMOIL-F30CC). We used the total internal reflection of a blank coverslip as a reference to obtain absolute reflection measurement. For this case, the pump angle of incidence is ~ 3.4° smaller than the probe.

3.2.3 Spectroscopic Ellipsometry: Thin-Film Properties (Linear)

This section concerns the linear characterisation of the ENZ thin films using a spectroscopic ellipsometer. The corresponding measurements were taken by our group member Thomas Phillips. His PhD thesis regarding measurements with the corresponding ellipsometer is taken as the primary reference for this section [79].

Ellipsometry measures the complex reflectance ratio ρ of p- and s-polarized light in terms of the parameters ψ and Δ:

$$\rho = \frac{r_p}{r_s} = \tan(\psi)e^{i\Delta}, \tag{3.6}$$

where r_p and r_s are the Fresnel reflection coefficients fo p- and s- polarized light. $\tan(\psi)$ is the amplitude of ρ (amplitude ratio of r_p and r_s) and Δ is the phase of ρ (phase difference of r_p and r_s). Measuring the ratio eliminates a lot of noise, e.g. from the light source, thereby making it a very sensitive and powerful tool for optical characterizations.

Ellipsometry is most commonly used to determine the optical (n, k and ε) and physical (thickness, roughness) properties of thin layered films. This is achieved by taking a series of spectra at different angles of incidence and fitting the spectra to an appropriate model (e.g. Fresnel coefficients, effective medium theories or Mueller-Jones matrices).

The typical setup consists of a light source with either spectral selection initially, as sketched in the form of a monochromator in Figure 3.6, or with a spectrally resolved detector, as is the case for the modern Woollam M-2000XI used in our studies. The light is prepared using a polariser and a compensator to modify the polarisation as needed before reflecting off the sample. Finally, the beam is analysed before being detected.

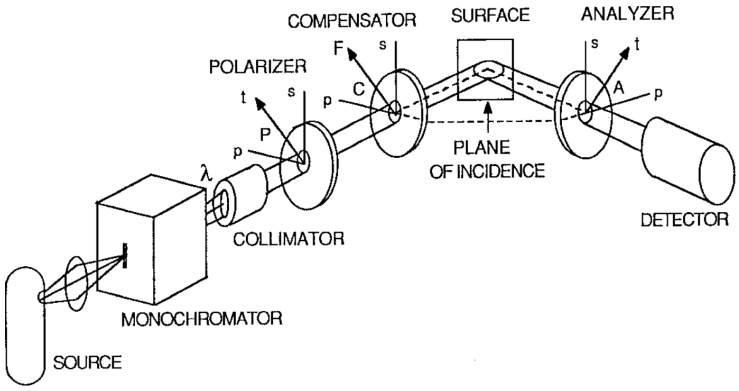


FIGURE 3.6: **Ellipsometer set-up schematic.** A generic ellipsometer design that includes a spectroscopic measurement capability. (Ref.: [80])

The spectroscopic ellipsometry of the ITO thin films was carried out using a J. A. Woollam Co. M-2000XI with which we can measure the ellipsometric parameters ψ and Δ in the wavelength range 210-1690 nm, with a wavelength step of 1.5 nm for 210-1000 nm and 3.5 nm for 1000-1690 nm. The unfocussed spot was approximately 3–5 mm in diameter depending on the angle of incidence. When studying smaller regions, it is possible to use focussing optics with numerical aperture 0.1, reducing the spot diameter to approximately 50 µm.

3.3 Multi-Layer Modelling

To understand the electromagnetic response of a thin layer composite, we investigate the changes to electric fields travelling through the sample, starting with a singular

refractive index interface and finishing with an arbitrarily large multilayer system. Born and Wolfs "Principle of Optics" book [81] is recommended for further insight beyond the introduction given here.

3.3.1 Interface (Fresnel Coefficients)

Propagation of electromagnetic waves across a spatial boundary presents one of the most well known problems in optics. The basis is given by the boundary conditions of the electric and magnetic fields:

$$\begin{array}{ll} \text{Field } \perp \text{ Interface} & \text{Field } \parallel \text{ Interface} \\ D_2 - D_1 = 0, & E_2 - E_1 = 0, \\ B_2 - B_1 = 0, & H_2 - H_1 = 0. \end{array} \tag{3.7}$$

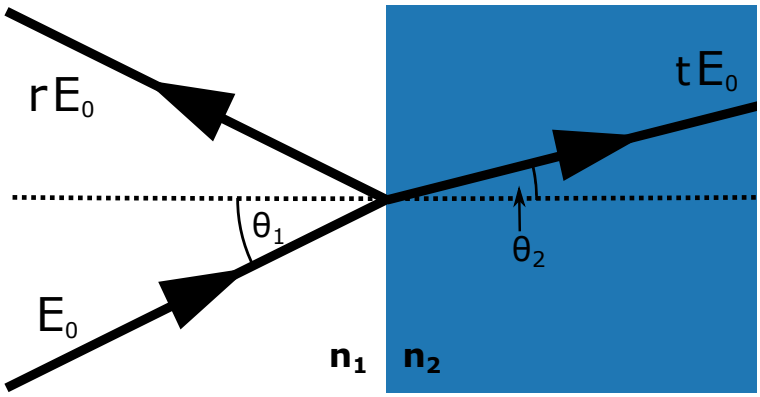


FIGURE 3.7: **Refraction of light at a spatial refractive index boundary.** At the interface the beam splits into a reflected and a transmitted beam with amplitudes modulated by the factor r and t respectively. The change in angle follows Snell's law $n_1 \sin\theta_1 = n_2 \sin\theta_2$.

Solving these for the two polarization cases of the transverse magnetic field (TM) and the transverse electric field (TE) for the case of a refractive index boundary (as seen in Figure 3.7) results in the following reflection and transmission coefficients, also known as Fresnel coefficients:

$$\begin{array}{ll} \text{TM} - \text{polarisation} & \text{TE} - \text{polarisation} \\ t_{12} = \dfrac{2n_1 \cos\theta_1}{n_2 \cos\theta_1 + n_1 \cos\theta_2}, & t_{12} = \dfrac{2n_1 \cos\theta_1}{n_1 \cos\theta_1 + n_2 \cos\theta_2}, \\ r_{12} = \dfrac{n_2 \cos\theta_1 - n_1 \cos\theta_2}{n_2 \cos\theta_1 + n_1 \cos\theta_2}, & r_{12} = \dfrac{n_1 \cos\theta_1 - n_2 \cos\theta_2}{n_1 \cos\theta_1 + n_2 \cos\theta_2}. \end{array} \tag{3.8}$$

Here, the refractive index (n) and angle of light relative to the surface normal (θ) are indexed corresponding to the case of the incident medium (1) and the transmitted medium (2) at the given boundary. These relate to the reflection and transmission as follows:

$$R = |r|^2, \quad T_{1,\mathrm{f}} = \frac{n_f \cos\theta_\mathrm{f}}{n_1 \cos\theta_1} |t_{1,\mathrm{f}}|^2. \tag{3.9}$$

The latter assumes a dielectric initial and final medium ($f = 2$ for the interface case). Additionally, phase changes to the reflected and transmitted beams can be studied by extracting the complex phase of the Fresnel coefficients. Hence, the reflection and transmission coefficients provide a powerful tool to study the response of optical interface systems.

3.3.2 Single-layer (Airy Formula)

To extend from the singular interface to a thin layer, we add a second interface, as seen in Figure 3.8.

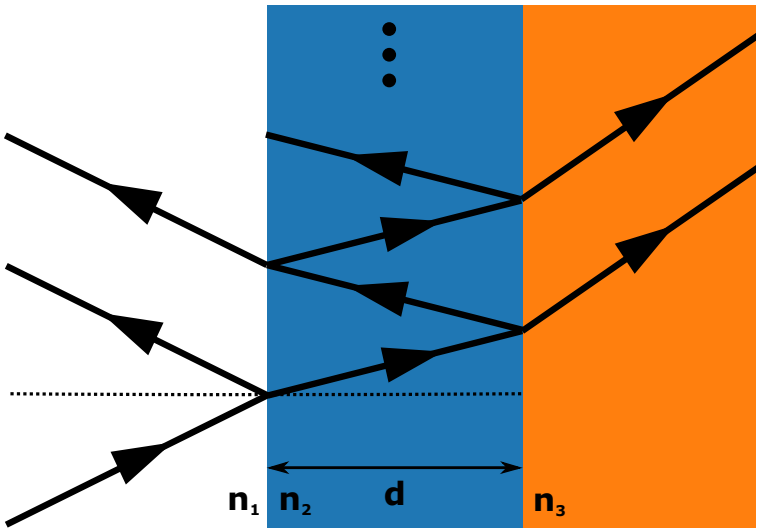


FIGURE 3.8: **Refraction of light at a refractive index layer.** An additional refractive event is added due to the second interface. Multiple internal reflections can occur until the light eventually follows the reflected or transmitted path.

To extract the reflection and transmission coefficients of this thin layer case one has to trace and sum up all possible contributions. For the transmission case the beam is always transmitted at the initial interface (t_{12}) and then propagates through the thin layer medium ($e^{ik_2 d}$). Finally, the beam is transmitted through the back interface (t_{23}), however, it can be internally reflected and propagate through the layer several times, adding up to various contributions given as:

$$t = t_{12}t_{23}e^{ik_2 d}\left(1 + r_{23}e^{ik_2 d}r_{21}e^{ik_2 d} + \ldots + (r_{21}r_{23}e^{2ik_2 d})^n\right). \tag{3.10}$$

Identifying this infinite sum as the geometric series and utilising the fact that $r_{12} = -r_{21}$ we get the total transmission coefficient of the thin layer, also known as the Airy formula:

$$t = \frac{t_{12}t_{23}e^{ik_2 d}}{1 + r_{12}r_{23}e^{2ik_2 d}}. \tag{3.11}$$

The total transmission coefficient (t) of the layer can be calculated using the sample thickness (d), the in-plane wavevector component inside the layer ($k_2 = n_2 k_0 \cos\theta_2$) and the Fresnel coefficients of the front (r_{12}, t_{12}) and back interface (r_{23}, t_{23}).

The reflection coefficient can be extracted in a similar fashion. Using the relation $t_{12} = t_{21}^*$ we get:

$$r = \frac{r_{12} + r_{23}e^{2ik_2d}}{1 + r_{12}r_{23}e^{2ik_2d}}. \tag{3.12}$$

The expressions for t and r can be easily implemented in scripts and vectorised to quickly analyse large parameter spaces of amplitude and phase changes inside the material of interest.

3.3.3 Multi-Layer (Transfer Matrix Method)

To extend this idea to an arbitrary number of interfaces, as indicated in Figure 3.9, a more elaborate approach will be utilised.

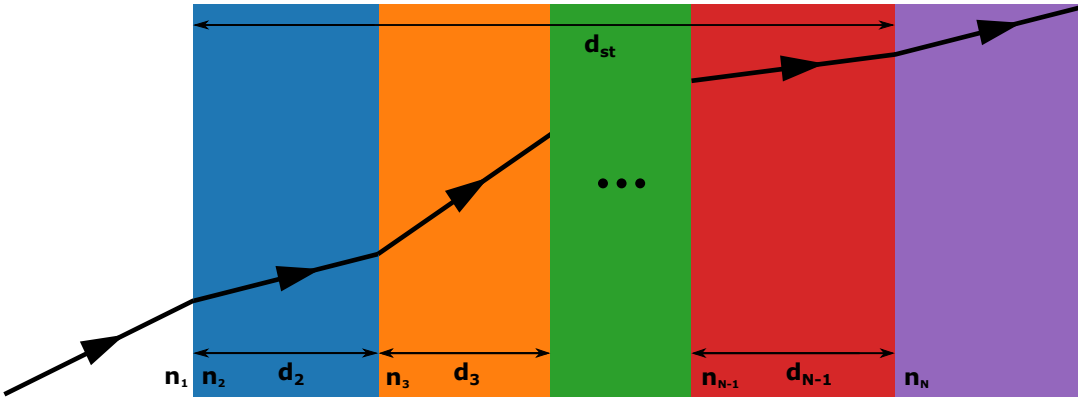


FIGURE 3.9: **Refraction of light at a multi-layer.** A stack of multiple refractive index layers lead to several internal reflection events along the plotted primary transmission path. The transfer matrix method can be employed to calculate the total reflection and transmission coefficient.

The transfer matrix method (TMM) describes the propagation of fields in a stratified medium. This section follows the concise introduction given in [82]. For a much more in-depth guide, see, e.g. Born and Wolfs "Principle of Optics" [81]. The electric field is considered to be harmonic in time ($e^{-i\omega t}$) and in space. The superposition of a forward- and backwards-moving electric field is defined as:

$$E_x(z) = E_f e^{ikz} + E_b e^{-ikz}, \tag{3.13}$$

with $k = nk_0 \cos\theta$ and $E_{f,b}$ as the constant initial forward and backward electric field amplitude. The core idea of the TMM is that the parallel field components E_x and H_y at the front and back of a slab are relatable via a matrix:

$$\begin{pmatrix} E_x(d) \\ H_y(d) \end{pmatrix} = M(k_d, d) \begin{pmatrix} E_x(0) \\ H_y(0) \end{pmatrix}, \tag{3.14}$$

where d is the slab thickness and $M(k_d, d)$ is referred to as the transfer matrix method. Utilising the interdependence of the fields, it follows that:

$$M(k_d, d) = \begin{pmatrix} \cos(k_d d) & \frac{1}{k}\sin(k_d d) \\ -k\sin(k_d d) & \cos(k_d d) \end{pmatrix}, \tag{3.15}$$

with k_d being the wavevector component of the slab. Extending to a stack of N slabs simply follows by matrix multiplication

$$\begin{aligned}\begin{pmatrix} E_x(d_{\mathrm{st}}) \\ H_y(d_{\mathrm{st}}) \end{pmatrix} &= M_{\mathrm{st}} \begin{pmatrix} E_x(0) \\ H_y(0) \end{pmatrix} \\ &= M_N \cdot M_{N-1} \ldots M_2 \cdot M_1 \begin{pmatrix} E_x(0) \\ H_y(0) \end{pmatrix},\end{aligned} \tag{3.16}$$

with d_{st} being the thickness of the entire stack. Once the final matrix (M_{st}) is calculated, the reflection and transmission coefficients (r and t) can be extracted by implementing the Ansatz:

$$\begin{aligned} E_{\mathrm{L}}(z) &= E_0 \left(e^{\mathrm{i}kz} + r e^{-\mathrm{i}kz} \right) \quad z < 0 \\ E_{\mathrm{R}}(z) &= E_0 \left(t e^{\mathrm{i}k(z-d_{\mathrm{st}})} \right) \quad z > d_{\mathrm{st}} \end{aligned}. \tag{3.17}$$

Finally, r and t can be extracted by solving the equation system given as

$$\begin{pmatrix} t \\ \mathrm{i}kt \end{pmatrix} = M_{\mathrm{st}} \begin{pmatrix} 1+r \\ \mathrm{i}k(1-r) \end{pmatrix}. \tag{3.18}$$

For numerical simulations in Python, the TMM-package by Steven J. Byrnes [83] provides an exemplary implementation of this method and will be used as the basis for multi-layer calculations in this thesis.

3.3.4 ENZ-layers With a Time-Varying Bulk Plasma Frequency

In this section, we examine time-dependent changes to the optical properties by applying the TMM model. As described in section 2.4, the excitation of ITO with a NIR pump leads to a red-shift of the bulk plasma frequency. For the temporal dynamics of ITO plotted in Figure 2.19 one can see that the shape of the signal closely matches that of the pump pulse with an additional sub-ps decay.

Hence, we model the bulk plasma frequency ω_{p} with the temporal shape Ansatz of the Gaussian pump convolved with an exponential decay as seen in Figure 3.10.

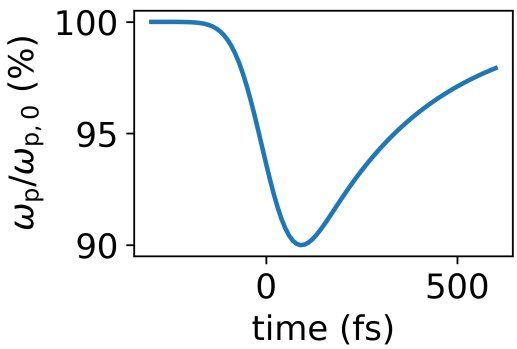


FIGURE 3.10: **Temporal shape Ansatz for the bulk plasma frequency.** The temporal shape Ansatz for the bulk plasma frequency ω_{p}. It takes the form of the gaussian pump (107 fs) convolved with an exponential decay (300 fs). The maximum change in ω_{p} is used to fit the nonlinear signals. For this example, a typical reduction of 10% is chosen.

This ω_p-change results in variations of the permittivity and refractive index. To demonstrate, we use the Drude parameters of $\varepsilon_\infty = 4$ as the high-frequency permittivity, $\omega_p = 3 \times 10^{15}$ rad/s as the bulk plasma frequency and $\gamma = 1.5 \times 10^{14}$ rad/s as the scattering rate. Considering the frequency dependence, we only show the example of $f \approx f_{ENZ}$ in Figure 3.11. The large refractive index changes even exceed the initial refractive index and will form the basis of large nonlinear optical effects, as discussed.

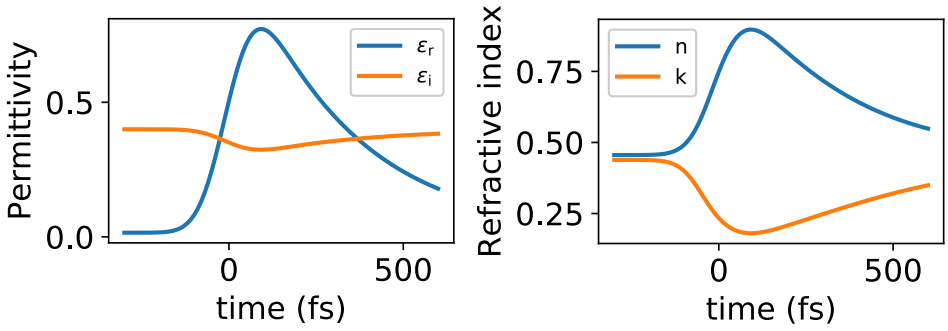


FIGURE 3.11: **Expected temporal changes to the permittivity and refractive index.** The frequency was chosen to be 237 THz to closely match the ENZ case. The resulting refractive index changes are larger than the initial refractive index itself as has been experimentally verified in [8].

The temporal refractive index is now applied to a 600 nm layer via the TMM model. The calculated changes in transmission and phase are shown in Figure 3.12 for the case of normal incidence. Note that the transmission phase has been defined as $\Phi(t \to -\infty) = 0$. Additionally, any phase jumps potentially arising due to the limits of the trigonometric function definitions are removed.

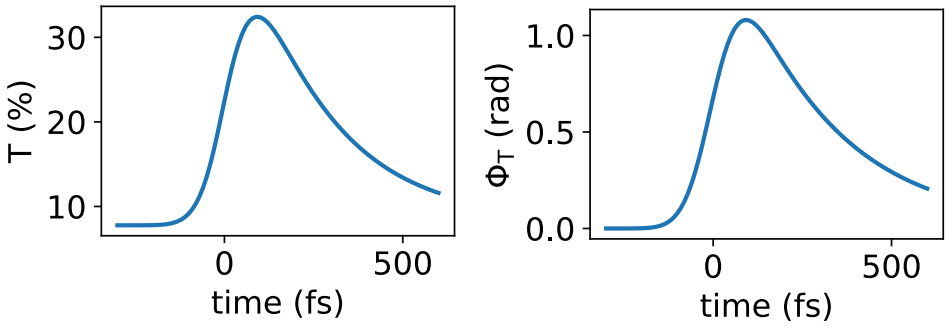


FIGURE 3.12: **Expected temporal changes to the transmission and phase.** Already for normal incidence, the absolute transmission change through the 600 nm ENZ is large (~ 25 %). Additionally, a phase increase by 1 rad is predicted, corresponding to the temporal changes of the bulk propagation phase through the layer.

These techniques form the basis of modelling the reflection changes in chapter 5 and frequency changes in chapter 7. The latter directly follows as the temporal derivative of the phase as seen in Equation 2.23.

Chapter 4

Enhanced High Harmonic Generation in Graphene

Abstract

Light weakly interacts with its environment. This feature makes it extraordinarily suitable for many optical applications, but, at the same time, this peculiarity makes light challenging to handle and manipulate. Nonlinear photonic platforms combine the strong field confinement provided by nanostructures, together with strong optical nonlinear materials, to spectrally and spatially control light. Among the most successful and accurate architectures, metallic plasmons have been shown to be a promising candidate. However, the short lifetime of metallic plasmons hinders any scalability. Here, we try to merge the strong field confinement of metallic nanostructures together with the long-lived, electrically tunable and extremely nonlinear graphene plasmons. Despite not being able to utilise the graphene surface plamons dut to experimental limitation we report third-order and fifth-order nonlinear signals three orders of magnitude larger than that of planar graphene. Our findings show that graphene–insulator–metal is a promising heterostructure for optical frequency conversion as a nano-optoelectronic component.

Author Contribution Statement

For this chapter, I designed and built a z-Scan setup. However, due to a strong background in the nonlinear signals for the high peak intensity laser system in Exeter, we took the final measurements presented here collaboratively at the University of Vienna. Additionally, I studied the thermal background in the fifth harmonic spectra. The work is in preparation to be submitted as I. Alonso et al. “High-harmonic generation enhancement with graphene heterostructures”.

4.1 Introduction

High harmonic generation (HHG) has been intensely investigated as a route for the generation of coherent attosecond radiation in the extreme ultraviolet and x-ray spectral regions. However, thus far, the most successful demonstrated systems for HHG have been atomic gasses [84, 85, 86, 87], which require high vacuum, making them impractical for many applications. As a result, there has been a concerted drive to develop solid-state systems, and HHG has been reported in various crystalline materials [88, 89, 90, 91, 92, 93]. However, the mechanism of HHG in solid-state appears to be fundamentally different from that in atomic gases and is highly sensitive, e.g. to crystal and polarisation orientations [94, 95].

Graphene is interesting in that it can demonstrate both intraband and interband conductivity in the infrared [22, 96, 97, 98]. More importantly, as a 2D material, it doesn't suffer the same phase matching restrictions, which generally diminishes HHG in thick crystals. The potential of graphene as a frequency conversion device was already presented for the third-harmonic generation (THG) case [99, 100] with additional features such as electrical tunability [101, 60]. High harmonic generation in graphene has been measured for VIS/IR frequencies [102, 62] and THz radiation [61, 103, 104, 105]. Plasmonic resonances have been suggested for enhancing HHG in graphene [106, 107, 108]. Graphene-heterostructures can even reach atom scale confinement levels of plasmons [7, 43] and have been utilised to enhance THG [64].

In this chapter, we show that the fifth-harmonic generation (FHG) in graphene can be enhanced by more than three orders of magnitude. We accomplish this by using a heterostructure made of graphene and gold nanoribbons separated by a nanometric insulating layer. The nanoribbons serve to enhance the incident pump field in the graphene layer, exhibiting no THG or FHG signal without the graphene layer. This allows us to fully utilize graphene's unique nonlinear optical properties, achieving HHG with very modest pump fluencies from a single atomic layer. We study the dependence of the THG and FHG efficiency on insulator material and ribbon widths, finding that the optimal ribbon width for THG generation is different from the optimal width for FHG. This is surprising, since the enhancement of the incoming electric field was meant to be optimized as the dominant contribution. Enhancement factors surpassing three orders of magnitude are measured for THG and FHG.

4.2 Experimental Setup

The samples were fabricated on top of lightly p-doped (1 - 10 Ωcm) Si with a 285 nm thermally grown layer of SiO_2. The Si and SiO_2 are mid-IR transparent. CVD graphene was wet transferred. For the samples with Al_2O_3, a thin layer of Al_2O_3 (5 nm) was grown on top of the graphene by atomic layer deposition (ALD), and for the samples with h-BN, the CVD h-BN monolayer was wet transferred. The masks were defined by electron beam lithography. Metallic rod arrays with a length of 9 µm covering approximately areas of 100 µm by 100 µm were made by EBL and lift-off (see Figure 4.1d). We investigate regions with and without graphene for reference purposes. Figure 4.1b shows a typical Scanning Electron Microscopy (SEM) image of the resulting high-quality metal edges of a one sample. The samples studied here are similar to [43]. An in-depth description of the manufacturing process and further details can be found in the corresponding Supplementary information of [43].

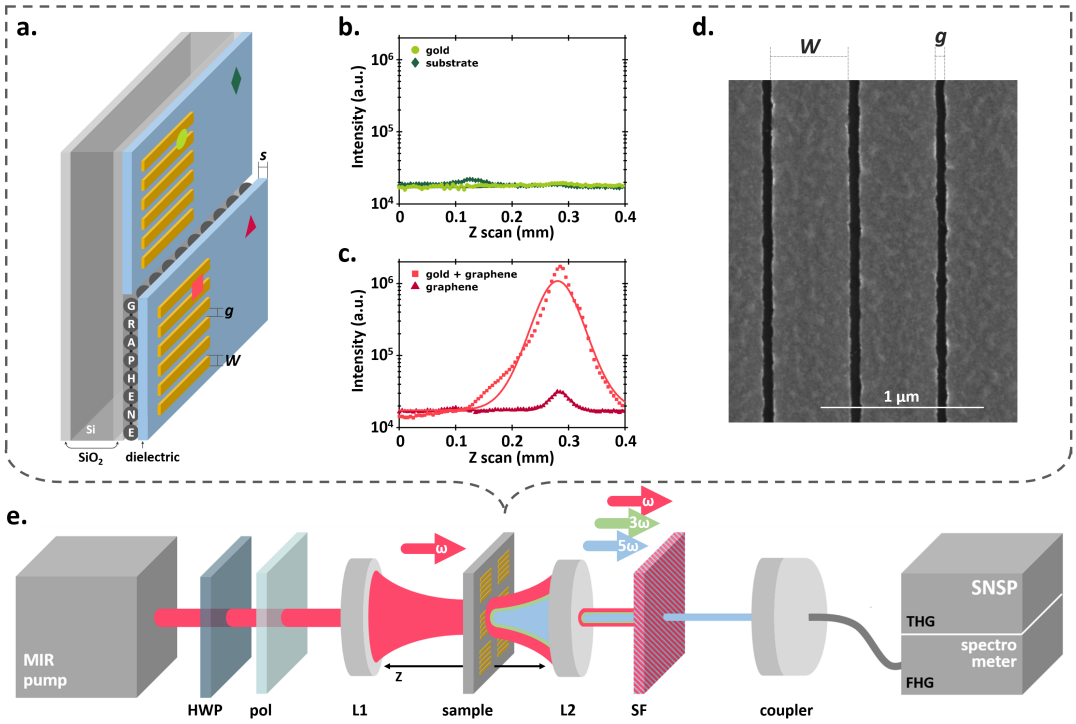


FIGURE 4.1: **Studied graphene heterostructure sample and setup.** **a,** Graphene is encapsulated by a nm-thick dielectric material (h-BN or Al_2O_3), that sets the space s between the graphene and the gold nanoribbons are grown on top. These arrays of nanoribbons are characterized by their width W and the gap g in between them. The incident light is shined perpendicular to the sample and the THG is collected in transmission. The sample that allows us to study the individual contribution of each component of the sample shows **b,** no nonlinear signal when there are only substrate or gold nanoribbons and **c,** shows a significant signal with bare graphene and an enormously enhanced signal when gold nanoribbons are added on top of the graphene. **d,** Scanning electron microscopy (SEM) image of an array. **e,** Experimental setup of the modified z-scan measurement, introduced in subsection 3.1.1.

To measure our the nonlinear signal of our sample, we use a modified z-scan setup. The nonlinear signal (either THG or FHG) is measured while the sample is moved along the z axis through the focus of the laser beam (see Figure 4.1e). Our pump beam is a linearly-polarized pulsed laser with a ~ 260 fs pulse width, a central wavelength of 3900 nm (0.225 eV), and a 76.2 MHz repetition rate. This beam is generated by an Optical Parametric Oscillator (OPO), fed by a mode-locked femtosecond Ti:saphhire laser. We use a Half-Wave Plate (HWP) to rotate the polarization of the incoming beam to that set by the polarizer (pol). A lens with a 5.95 mm focal length focuses the pump down to a waist of ~ 13.5 µm. When the sample is moved parallel to the pump beam (along the z axis), the nonlinear emission occurs most efficiently where the fluence is maximum, i.e. at the focal point. Afterwards, a lens with a 11 mm focal length collimates the beam, which is then sent through a set of spectral filters that separate the nonlinear signals from the pump beam. The resulting nonlinear signal is coupled into a multimode fibre, which can be sent to one of two detectors. The THG signal, at ~ 1300 nm, is measured with a large-area Superconducting Nanowire Single-Photon (SNSP) detector from PhotonSpot with a ~ 60 % detection efficiency at this wavelength. We can combine this detector with a Gemini interferometer

from NIREOS to measure the spectrum of the THG signal. The FHG signal, which is centred around 780 nm, is sent to a single-photon sensitive silicon Andor spectrometer with a resolution of 1 nm, and detection efficiency of ~ 10 %.

To verify the origin of our nonlinear signals, we perform z-scans on various regions of our samples. As shown in Figure 4.1b, neither the substrate (green diamonds) nor the gold nanoribbons without the graphene layer (green circles) display any measurable nonlinear response. However, both graphene (red triangles) and graphene with nanoribbons (red squares) show significant nonlinear signals at the focal point of the pump beam (see Figure 4.1c). These control measurements show that the graphene layer contributes to any nonlinearity measured in our setup. Moreover, they already show the huge nonlinear enhancement provided by the gold nanoribbons.

4.3 High-Harmonic Spectrum

We further verify the nature of our nonlinear signals by measuring their spectra. As displayed in Figure 4.2, we find a THG signal ~ 1300 nm and a FHG signal ~ 780 nm. Both signals fall at the expected wavelengths for 3900 nm pump beam. Furthermore, the full width at half maximum (FWHM) of the two nonlinear processes scale as expected, taking into account the λ^2 factor. The FWHM of the THG signal is about 50 nm, while the FWHM of the FHG signal is about 20 nm. In addition to these two signals, we observe broadband white light generation, which we attribute thermal photoluminescence from the graphene layer. This has been previously observed in [109], and will be discussed further in the next section.

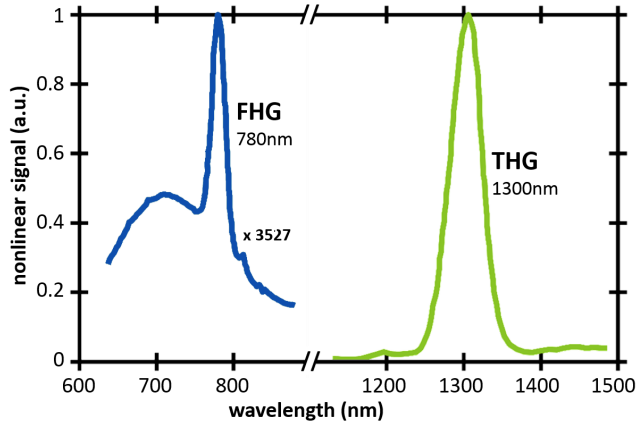


FIGURE 4.2: **Characterization of the nonlinear spectrum.** Spectrum of the THG signal measured with a Fourier transform spectrometer (15 nm resolution), together with the spectrum of the FHG signal mesured with a single-photon spectrometer (1 nm resolution).

Thermal Radiation Background

In this section, the physical origin of the background signal seen in the FHG spectrum of Figure 4.2 will be identified as ultrafast photoluminescence from graphene. The significant broadband light emission from graphene under excitation by femtosecond laser pulses has been studied by C. Lui et al. [109] with great detail and is employed here to model our measurements. The ultrafast heating of the electron gas is modelled with a two-temperature model, consisting of the electronic temperature T_{el} and the

temperature of strongly coupled optical phonons T_{op}:

$$\begin{aligned} \frac{dT_{el}(t)}{dt} &= \frac{I(t) - \Gamma\left(T_{el}, T_{op}\right)}{c_e\left(T_{el}\right)} \\ \frac{dT_{op}(t)}{dt} &= \frac{\Gamma\left(T_{el}, T_{op}\right)}{c_{op}\left(T_{op}\right)} - \frac{T_{op}(t) - T_0}{\tau_{op}}, \end{aligned} \tag{4.1}$$

with the phonon decay time $\tau_{op} = 1$ ps and the specific heat capacity of the phonons $c_{op}\left(T_{op}\right)$ and electrons $c_{el}\left(T_{el}\right)$.

In our case, we define the electron-phonon energy exchange rate Γ similar to C. Lui et al. [109] as

$$\begin{aligned} \Gamma\left(T_{el}, T_{op}\right) = \alpha \cdot (n_{op}(T_{op}) + 1) \int E \cdot (E - E_{op}) \cdot f_{el}(E, T_{el}) \cdot [1 - f_{el}((E - E_{op}), T_{el})] dE \\ - \alpha \cdot (n_{op}(T_{op})) \int E \cdot (E - E_{op}) \cdot f_{el}(E - E_{op}, T_{el}) \cdot [1 - f_{el}((E), T_{el})] dE, \end{aligned} \tag{4.2}$$

where $E_{op} = 0.18$ eV is the phonon energy, $n\left(T_{op}\right) = \left[\exp\left(E_{op}/kT_{op}\right) - 1\right]^{-1}$ represents the phonon population at temperature T_{op} and $f\left(E, T_{el}\right) = \left[\exp\left(E/kT_{el}\right) + 1\right]^{-1}$ is the Fermi-Dirac distribution for electrons at T_{el}. The phonon coupling constant α corresponds to:

$$\alpha = 3\frac{9\left(\frac{\partial t}{\partial b}\right)^2}{\pi\rho E_{op}\hbar^3 v_F^4}, \tag{4.3}$$

with $v_F = 10^6\,\mathrm{ms}^{-1}$ $\frac{\partial t}{\partial b} = 45$ eV/nm and $\rho = 7.6 \times 10^{-7}$ kg/m^2 as the density of graphene [110].

The calculated electron temperature in graphene for 260 fs pulses with an incoming peak intensity of 100 MW/cm^2 can be seen in the inset of Figure 4.3. To get the corresponding spectral radiant fluency, we use Plancks's law:

$$\mathcal{F}\left(\lambda, T_{el}\right) = \varepsilon(\lambda)\frac{2hc^2}{\lambda^5}\int\left[\exp\left(\frac{hc/\lambda}{kT_{el}(t)}\right) - 1\right]^{-1} dt. \tag{4.4}$$

The emissivity $\varepsilon(\lambda)$ of graphene was calculated via the absorption of light coming from air, taking into account the wavelength-dependent modulation via the cavity formed by the 285 nm layer of glass between the graphene and Si. Assuming that the illuminated graphene area acts roughly like a thermal Lambert emitter, we estimate our lens to collect up to the angle of

$$\theta = \arctan\left(\frac{D_{eff}/2}{f}\right) \tag{4.5}$$

, where the focus length is $f = 11$ mm and the effective diameter is scaled to match the experimental signal via $D_{eff} = 30\% \times 5.5$ mm. This is the only scaling parameter that has been tweaked to match the amplitude of the thermal estimate (not the wavelength dependence) with the background we see in the experiment, as shown in Figure 4.3. It compensates for alignment issues, as the apparatus is not designed to measure white light generation but HHG. It also compensates filter reflection losses, coupling

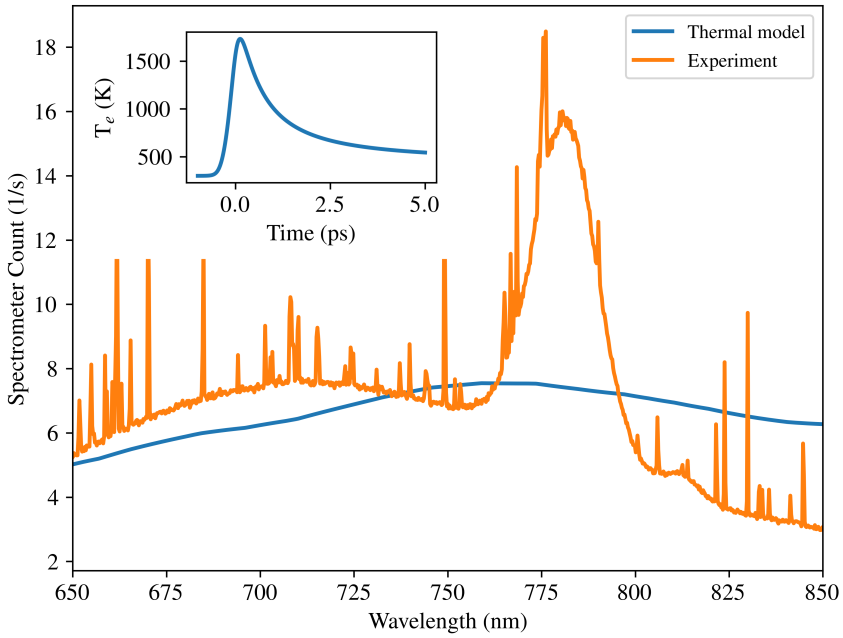


FIGURE 4.3: **Ultrafast thermal emission from planar graphene.** Spectral emission of monolayer graphene on Si-wafer, pumped by 260 fs pulses with a peak intensity of 100 MW/cm^2. The experimental data (orange) shows the FH signal around 780 nm, and a broadband background. This background is modelled as ultrafast thermal emission by hot carriers in graphene (blue). The corresponding time-dependent electron temperature can be seen in the inset.

into the spectrometer and the like. It does not compensate wavelength-dependent behaviour, which we assume to be most influenced by the rough estimate of our absorbed mid-infrared fluency (~ 1.5%), but also by the collection lens having a chromatic alignment and efficiency dependence. Considering all of these aspects, our estimate matches sufficiently well to identify the ultrafast thermal emission of hot electrons as the dominant source of the background seen in the FH measurements.

4.4 Polarization Dependence

We also study the polarization dependence of the measured nonlinear signals. Owing to the geometry of the gold nanoribbons, the field absorption, and thus metallic plasmon excitation, is strongly dependent on the polarization of the input light. As shown in Figure 4.4, the enhancement is maximum when the light polarization is perpendicular to the direction of the nanoribbons. However, when the polarization is shifted from the perpendicular case, the nonlinear signal decreases as $\cos^6\theta$ and $\cos^{10}\theta$, for THG and FHG signals, respectively, reaching a minimum when the polarization is parallel to the direction of the nanoribbons. At this point the signal strength is even lower than that of planar graphene without gold nanoribbons because of the screening of the nanoribbons.

The Substrate Transmission

While the polarization dependence induced by the ribbons clearly matches our expectation, there are some subtle details to be found for the planar graphene case. The

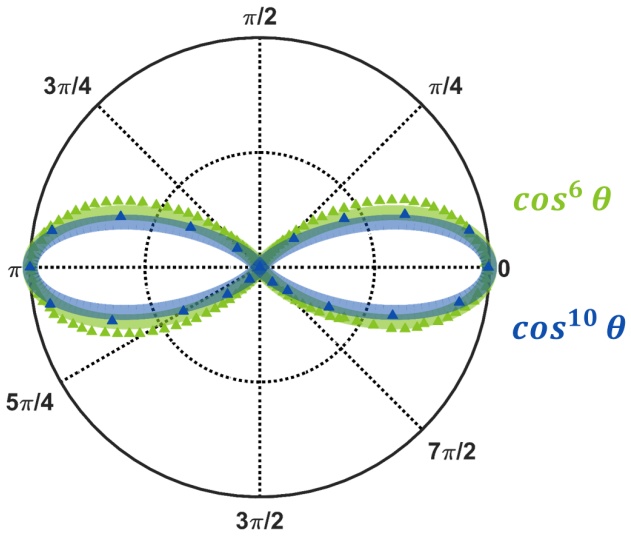


FIGURE 4.4: **Characterization of the nonlinear polarization.** Study of the THG and FHG signal as a function of the input polarization. As expected the gold nanoribbons impose a $\cos^6\theta$ and $\cos^{10}\theta$ dependence for THG and FHG signals, respectively, maximizing the signal when the incident light is perpendicular to the direction of the nanoribbons.

upper panel of Figure 4.5 shows the measured pump power normalized transmission through the sample as a function of the orientation of the linear polarization. With reference to Figure 4.1a, far-field incident light with carrier wavelength of 3900 nm is shined perpendicular to the sample on the side of the substrate, composed by a stack of two layers of silica and one layer of silicon. Silica layers have a thickness of 285 nm, while the silicon layer has a thickness of 500 µm (details of sample fabrication are reported in [43]). Here, the investigated sample area does not present the array of gold nanoribbons on top of the encapsulated graphene (see the violet triangle in Figure 4.1a). Red points in Figure 4.5 show a periodic pump transmission through the sample as a function of the incoming linear polarization angle, tuned with a calibrated half-wave plate. The measured relative difference in transmission is about 5%. In the lower panel, the green points show the polarization-dependent normalized third-harmonic power and the normalized cube of the transmitted power (blue points). The quantitative matching between the two explains the polarization dependent third-harmonic generated signal, which depends cubically on the input power.

The physical origin of the polarization effect can be related to the slight silicon birefringence reported in [111], to which a different absorption as a function of the polarization is likely to be associated. Another possible explanation can be due to the Fresnel polarization dependent reflection in the case of non-normal incidence [112]. Indeed, in the case of an interface between air and silica, an angle of ~ 7 ° compare to the normal of the interface, is enough to induce a relative reflectance difference of about 5 % between *s* and *p* polarization and likely to have a noticeable effect due to imperfect alignment.

4.5 Dielectric Spacer Dependence

To better understand the nature of the enhanced high harmonic generation, we first study the material and thickness of the dielectric spacer between the graphene and

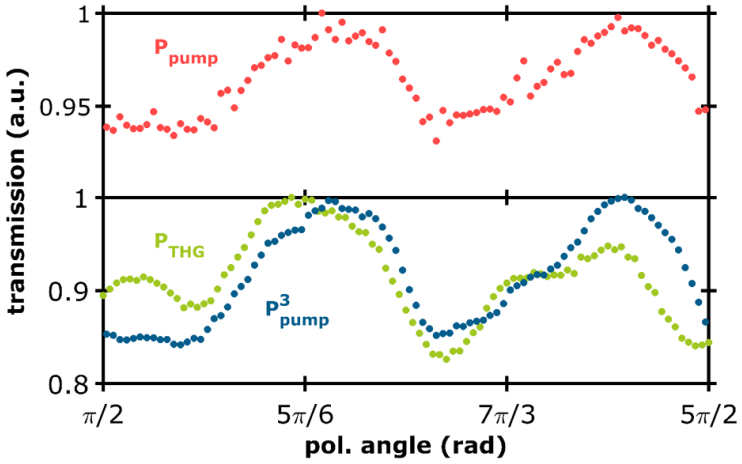


FIGURE 4.5: **Polarization dependent transmission through the substrate.** Red points show the normalized polarization-dependent transmission through the substrate. Green points are the normalized third-harmonic generated power as a function of the input polarization. Blue points are the cube of the transmitted pump power (red points). All the measurements are performed with carrier wavelength of 3900 nm.

gold nanoribbons. We later proceed to characterize the dependence of the THG and FHG signals on the nanoribbon width. To quantify the effect of the spacer material and thickness, in Figure 4.6a we show the enhancement of the TH signals for a 5 nm Al_2O_3 spacer (dark green squares) and a monolayer h-BN spacer (bright green triangles) with respect to the signal from bare graphene (red squares). We measure the nonlinear signal as a function of the input power and fit a linear curve, where the slope is fixed to three and the interception point is the only fitting parameter. The quotient of the interception points then gives the enhancement. In this case, we find a factor of 26 enhancement with the 5 nm Al_2O_3 spacer and a factor 1600 enhancement with the h-BN spacer. The effective thickness of the monolayer h-BN is ~0.3 nm. In Figure 4.6b, we show the same power scaling for the FH signal, where the slope of the fitting curve is set to 5, and find an enhancement factor of 300 for the 5 nm Al_2O_3 spacer (dark blue squares) and a factor 4400 enhancement with the h-BN spacer (bright blue squares). While the scaling of the FHG relative to the THG makes sense for the Al_2O_3 spacer, the h-BN spacer leads to significantly lower enhancement then expected: $1600^{5/3} \sim 220,000$. The origin of this discrepancy will be studied and discussed in more detail at the end of this chapter.

Note: Problems with the high peak power system in Exeter

The initial idea to carry out the nonlinear measurements in Exeter was limited by saturation effects. The plus of our laser system was higher and more stable fluencies, which limited the search for high harmonics beyond the fifth-order. However, the laser system in Exeter is a high pulse energy system with a repetition frequency of just 1 kHz compared to the 76.2 MHz system in Vienna. Hence, the peak intensity for the same fluency is increased by a factor of ~80,000. Considering the third harmonic signal saturates already for the maximum fluency of the Vienna system, as seen, e.g. in Figure 4.6a or Figure 4.7a, the graphene high harmonic signal turned out to be low compared to the surface nonlinearity. A strategic change towards sizeable focal length and thereby reduced peak intensities allowed for discriminating the nonlinear

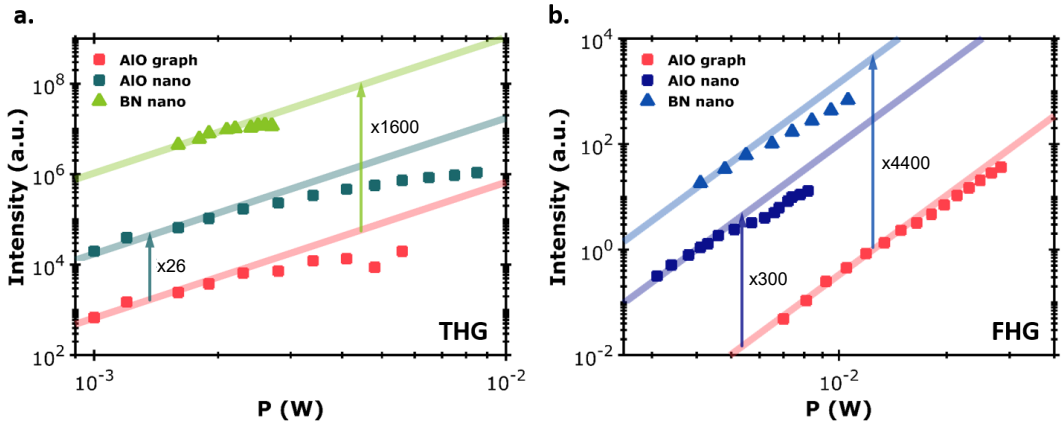


FIGURE 4.6: **Spacer dependence of nonlinear enhancement. a,** and **b,** show the TH/FH signal enhancement with gold nanoribbons over bare graphene (red squares) for a 5 nm Al_2O_3 spacer (dark green/blue squares) and a monolayer h-BN spacer (bright green/blue triangles). In the device with Al_2O_3 spacer, the TH and FH signals are enhanced by a factor 26 and 300, respectively. In the case of the h-BN spacer, the TH and FH signals are enhanced by a factor 1600 and 4400.

origins by x-y-mapping rather than z-scanning for the THG case; however, it proved difficult for higher harmonics. At that point, we decided to move back to the high repetition system in Vienna. We further optimized the Vienna setup with, e.g. the VIS sensitive spectrometer, brought from Exeter to Vienna during my research visit, to measure fifth and potentially seventh harmonics.

4.6 Nanoribbon Width Dependence

To quantify the effect of the nanoribbon width on the TH and FH signal enhancement we compile a series of measurements performed with the device with the h-BN spacer, where we study this dependence (see Figure 4.7). In panels a and b, we show the third- and fifth harmonic signals as a function of the input power for nanoribbon widths from 20 to 500 nm (yellow to blue) and bare graphene (red squares). The scatter points are experimental data points, and the lines are fits, where only the interception points are free parameters, and the slopes are set to three and five, respectively. In both cases, some saturation effect can be observed. Non-destructive saturation originates from blocking both intra and interband transitions due to an increased temperature, also referred to as bleaching. Note that, due to the lower damage threshold of the gold nanoribbons, the power scaling of the 20, 50 and 500 nm nanoribbon widths are missing in panel b. Using these interception points, we calculate a maximum enhancement of 1600 for the THG signal when gold nanoribbons are placed on top of the bare graphene and 4400 for the FHG signal. In panels Figure 4.7c and d, we calculate the effective third-order and fifth-order susceptibilities of the graphene heterostructures and find the maxima to be $\chi^{(3)} = 5.6 \cdot 10^{-6}$ and $\chi^{(5)} = 1.8 \cdot 10^{-14}$ esu. The derivation of the effective nonlinearities is reported in subsection 3.1.1. Note that the resonant nanoribbon width is different for TH and FH.

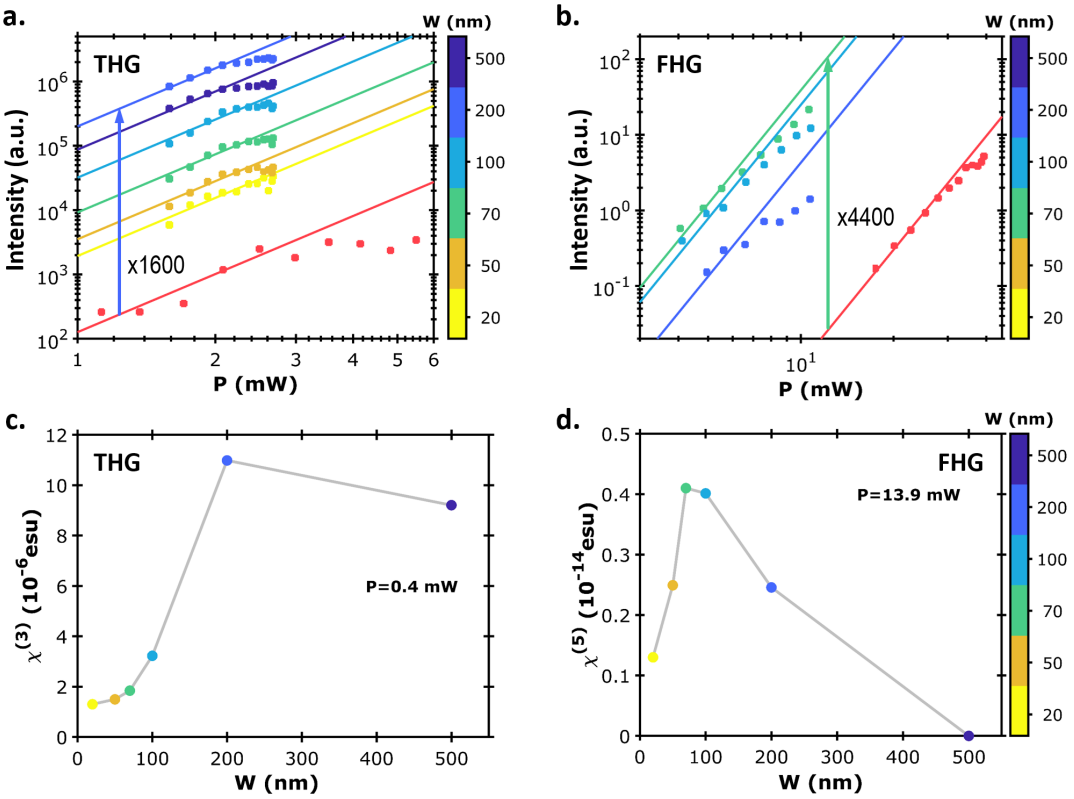


FIGURE 4.7: **Nanoribbon width dependence of nonlinear enhancement. a,** and **b,** are the power dependence of the THG and FHG signals for bare graphene (red) and graphene with gold nanoribbons of different widths (yellow through blue). The interception points found with the fits, whose slopes are set to 3 and 5, respectively, enable calculating the enhancement of the signals due to the field confinement. In the case of THG, the maximum enhancement it 1580 for W = 200 nm and for FHG it is 4400 for W = 70 nm. **c,** and **d,** summarize $\chi^{(3)}$ and $\chi^{(5)}$ for different nanoribbon widths and given powers P = 0.4 mW and P = 10.6 mW, respectively.

4.7 Simulations

As currently being verified by simulations, the most efficient nanoribbon width for THG is around 200 to 500 nm, whereas for FHG, it is around 70 to 100 nm. The theoretical group of Prof Javier García de Abajo from ICFO, Barcelona, combined with students from the University of Vienna, most notably Philipp Jenke, have tried modelling the observed behaviour in various ways. While the THG generation can be simulated reasonably well, as already shown in Irati's publication on THG [64], the FHG process is much harder to model.

In a first attempt, FDTD simulations were carried out using Lumerical. A fundamental problem here is the absence and inability of implementing a non-local conductivity model for graphene. Furthermore, $\chi^{(5)}$ simulations via the Green function approach[113] did not converge.

Currently, FEM simulations are carried out using COMSOL. The Green function approach has not been precise enough due to singularity issues. Next, the derivation of an analytical form for $\sigma^{(5)}$ (and therewith $\chi^{(5)}$) was done, following the derivation in [114]. This model also enabled the inclusion of cascaded processes, as plotted in Figure 4.8. Further tests included the impact of round nanoribbon edges, different

thicknesses of the adhesion layer and the like. So far, none of the predictions point to a peak around ~100 nm, as suggested by measurements.

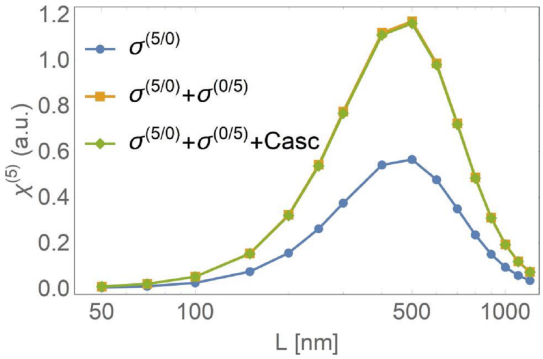


FIGURE 4.8: **Preliminary wavelength-dependent predictions of $\chi^{(5)}$.** The $\chi^{(5)}$ corresponding to the analytically derived $\sigma^{(5)}$ suggests a peak around 500 nm. Additionally, we show that the cascaded processes lead to no noticeable, similar to many other approaches tested such as round nanoribbon edges, different thicknesses of the adhesion layer and the like.

Given the apparent absence of non-local effects, our best working explanation for the differing dependence of THG vs FHG is quenching of the fifth-order emission by the metallic plasmon. Due to the higher frequency of the fifth-order emission, it will sit near the plasmon resonances of the gold in the visible spectral region. Quenching of emission would also explain the lower than expected enhancement of FHG: if considering only incident field enhancement, and one can achieve a factor of 1600 enhancement of THG, then one might expect an enhancement factor for FHG of $1600^{5/3} \sim 220,000$, i.e. significantly higher than the 4400 observed in experiment.

Initial modelling of the nanostructures carried out by our colleagues in Vienna suggests that the FHG emission band should lie just outside the plasmon resonance region, hence predicting a minor quenching effect. However, one would require a relatively small shift in the expected plasmon resonance wavelengths predicted by modelling to explain the experimental data, something that often occurs due to a slightly different physical geometry or Drude metal parameters achieved during fabrication. At the time of submission of this thesis, plans were being drawn up to measure the optical spectra of the samples to try and confirm this hypothesis.

4.8 Conclusion

Merging the highly accurate control and manipulation of the electromagnetic fields in nanoplasmonics, together with the long-lived and electrically tunable plasmons in graphene, offers a new promising approach to perform a wide range of nonlinear optical processes, with the potential to achieve single-photon-level nonlinearities. Our observation of up to ~1600 fold enhancement of the THG signal and ~4400 fold enhancement of the FHG signal demonstrates the suitability of this novel architecture for frequency conversion nonlinear optical applications. Reaching higher harmonics was hindered by experimental difficulties of going to longer wavelength, required for exploiting graphene plasmon polariton and their extreme field confinement. Future research directions might investigate even more substantial enhancements via graphene plasmon generation. Such graphene-based heterostructures would thus be

of great importance for many nonlinear processes at the single-photon level, such as quantum technology.

Chapter 5

All-Optical Switching of an Epsilon-Near-Zero Plasmon Resonance in Indium Tin Oxide

Abstract

Nonlinear optical devices and their implementation into modern nanophotonic architectures are constrained by their usually moderate nonlinear response. Recently, epsilon-near-zero (ENZ) materials have been found to have a strong optical nonlinearity, which can be enhanced through the use of cavities or nano-structuring. Here, we study the pump dependent properties of the plasmon resonance in the ENZ region in a thin layer of indium tin oxide (ITO). Exciting this mode using the Kretschmann-Raether configuration, we study reflection switching properties of a 60 nm layer close to the resonant plasmon frequency. We demonstrate a thermal switching mechanism, which results in a shift in the plasmon resonance frequency of 20 THz for a TM pump intensity of 70 GW/cm^2. We highlight an additional two-beam coupling contribution not previously isolated in ENZ nonlinear optics studies for degenerate pump and probe frequencies, which leads to an overall pump-induced change in reflection from 1 % to 45 %.

Author Contribution Statement

For this chapter, I designed the sample and experiment. I carried out the nonlinear measurements and numerical analysis. The work has been published as J. Bohn et al. "**All-optical switching of an epsilon-near-zero plasmon resonance in indium tin oxide**," NATURE COMMUNICATIONS, **12**, 1017 (2021)[115].

5.1 Introduction

Nonlinear optics is utilised for a wide range of photonic applications such as quantum all-optical data processing [116, 117], information technology [118, 119] and telecommunication applications. With the rise of new computational demands such as artificial intelligence, all-optical signal processing is often seen as a breakthrough technology for the next generation of computation and communication devices [120]. However, such applications are limited by the interaction of light signals, with extremely weak optical nonlinearity exhibited by most materials. This leads to high power consumption and large physical size of optical circuitry, making integration into existing nanophotonic platforms challenging [121, 122]. Moreover, most optical switching materials and geometries are not compatible with existing complementary metal–oxide–semiconductor (CMOS) fabrication technologies, which is preferential for implementation into existing platforms [123].

Epsilon-near-zero (ENZ) materials have recently attracted much attention, not only for their intriguing linear properties [19], but also because they exhibit large optical nonlinearities [10, 36]. Moreover, a subset of ENZ materials, transparent conductive oxides, exhibit resonance frequencies in the near-infrared, thereby allowing for integrated telecom applications [124, 125, 126, 127, 128, 129]. Indium tin oxide (ITO) has been shown to undergo a refractive index change of order unity upon optical pumping of a thin film [8]. This effect is thought to arise from electron heating, which leads to a change in effective mass due to the non-parabolic electron dispersion [130, 131]. Similar optical nonlinearities have been measured for doped zinc oxides [65, 66] and CdO [67]. These materials are also tunable, with variability in their doping level, giving control over the ENZ resonance wavelength spanning the infrared range [132, 133]. To further increase the optical switching properties of transparent conducting oxides different strategies have been utilized, including additional structuring [134, 69, 128] or the design of cavity modes [135, 136, 67, 137].

It is also well known that near the ENZ frequency of a thin transparent conducting oxide, one can also excite a plasmon resonance, enhancing the incoming field and near-perfect absorption[38]. We will refer to this resonance as the ENZ plasmon. One can excite plasmons using a high index incident prism in the Kretschmann-Raether configuration, circumventing the need for nano-structuring or the additional support of a cavity. This approach has been employed to study plasmon based nonlinear optical dynamics in gold films[138, 139, 140]. However, while ENZ plasmon excitation has been employed to enhance third harmonic radiation[71], there has been no study of all-optical switching in transparent conducting oxides in this desirable geometry.

Here, we investigate optically induced shifting of the ENZ plasmon frequency via pump-probe experiments in the Kretschmann-Raether configuration. This geometry provides a potential switching platform from near-total absorption to total internal reflection upon tuning the plasmon resonance into and out of the spectral range. We identify two contributions to the nonlinear signal: A dominant thermal switching process results in a shift in the plasmon resonance frequency of 20 THz for a TM pump intensity of 70 GW cm^{-2} when pumping resonantly, resulting in a change in reflection of the probe from 1 % to 30 %. Exclusively for the TM pump polarisation, we isolate an additional two-beam coupling (TBC) contribution. These two mechanisms combine to enable reflection switching of a 60 nm layer by more than an order of magnitude, with a measured change in reflection of the probe from 1 % to 45 % for a pump intensity of 70 GW cm^{-2}.

5.2 Optical Properties of the ITO Sample

For our study ITO was sputtered onto a cover glass at room temperature using 90/10 In_2O_3/SnO_2 Kurt Lesker target and sputtering tool. The base pressure before the deposition was in low 10^{-6} torr, but raised to 3 mTor of Ar only during deposition with a RF power of 145 W. In order to achieve high carrier density, both deposition and annealing were performed in the lowest possible residual oxygen environment. The samples are post-annealed in forming gas for 3 min at temperatures between 425-525°C in a rapid thermal annealer. Considering the thickness is much thicker than ~10 nm the layer is assumed to be a fully connected layer and we did not carry out structural studies. However, the varying results of the polycrystalline structure will have an effect on the scattering rate[141]. To obtain the optical properties, relevant for the studies in this thesis, the near-infrared permittivity of the ITO film was measured with an ellipsometer and described using the Drude model:

$$\varepsilon_{\mathrm{ITO}}(\omega) = \varepsilon_\infty - \frac{\omega_\mathrm{p}^2}{\omega^2 + i\omega\gamma} \tag{5.1}$$

with $\varepsilon_\infty = 3.43$ as the high-frequency permittivity, $\omega_\mathrm{p} = 2.86 \times 10^{15}$ rad/s as the bulk plasmon frequency and $\gamma = 2.24 \times 10^{14}$ rad/s as the damping rate (see Figure 5.1a). These parameters were determined experimentally by ellipsometry and subsequent fitting.

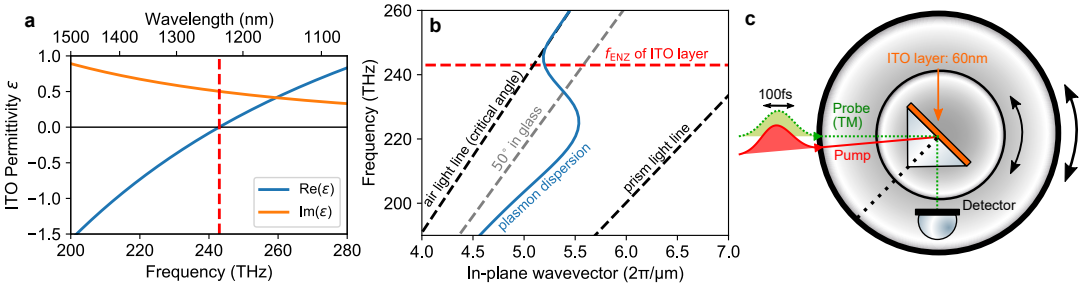


FIGURE 5.1: **Basic material properties and schematic setup. a**, Optical permittivity of the ITO film used in this study, measured using ellipsometry, with an epsilon-near-zero frequency of 243 THz (red dashed). **b**, Plasmon dispersion branch of the 60 nm ITO thin film closest to the air light line (blue), calculated using Equation 5.2. **c**, Schematic pump-probe setup with the ITO sample index matched to a prism in order to probe the ENZ plasmon beyond the critical angle (here, $\theta = \theta_{\mathrm{pr}} = 45°$). In this geometry the relative beam angles are fixed such that the pump angle is always given as $\theta_{\mathrm{pm}} = \theta_{\mathrm{pr}} - 3.4°$, enabling a quasi co-linear geometry while reducing pump collection in the detector.

The resulting dispersion of the ENZ plasmon modes can be found using the dispersion relation taken from ref. [142]:

$$F(k,\omega) = \tanh(\alpha_2 d) + \frac{\epsilon_2\alpha_2\left(\epsilon_1\alpha_3 + \epsilon_3\alpha_1\right)}{\alpha_2^2\epsilon_1\epsilon_3 + \alpha_1\alpha_3\epsilon_2^2} = 0, \tag{5.2}$$

with $\alpha_j^2 = k^2 - k_0^2\varepsilon_j$. The permittivities are given by $n_1 = 1$ (air), $\varepsilon_2 = \varepsilon_{\mathrm{ITO}}(\omega)$ and $n_3 = 1.43$ (substrate). We solve for a complex wavevector $k = k_r + ik_i$ and real frequency ω. The solution closest to the air light line is plotted in Figure 5.1b.

Wavevector matching is required to excite a plasmon. In our design, we follow the Kretschmann-Raether configuration as seen in Figure 5.1c. The prism enables the pump and probe beams to be incident upon the ITO layer beyond the critical angle, resulting in a wavevector beyond the air light-line and enabling plasmon excitation in the near-perfect absorption regime. Of interest here is the area around epsilon being zero, which coincides with the backbend segment of the plasmon dispersion, shown in Figure 5.1c for a 60 nm thick film. We refer to this plasmon segment in the remainder of this chapter as the ENZ plasmon region. We are specifically interested in the non-radiative ENZ plasmon, which lies beyond the air light line – in this region, one expects to observe a near-perfect absorption and enhanced fields[38].

5.3 Nonlinear Model: Intensity Dependent Reflection

We introduce two models that we use to interpret our experimental pump-probe data presented and discussed in section 5.4. The first "static" model treats our sample as a homogeneous layer, solving for excited state permittivity parameters. The second "dynamic" model predicts the complete dynamic behaviour of samples assuming a time-dependent plasmon resonance frequency and allows us to predict the temporal dynamics observed in our pump-probe measurements.

Angle and frequency-dependent reflection (Static model):

We use a transfer matrix approach to calculate the reflection and absorption from our sample, assuming a given permittivity. The initial absorption is calculated for the TE polarized pump beam $A_{\mathrm{pm}}\left(\theta_{\mathrm{pm}}, f_{\mathrm{pm}}\right)$, which can be seen in Figure 5.2. The incoming medium is assumed to be substrate/prism (index n = 1.43), while the medium on the far side is air (n=1). The calculated $A_{\mathrm{pm}}\left(\theta_{\mathrm{pm}}, f_{\mathrm{pm}}\right)$ is used to determine the angle and

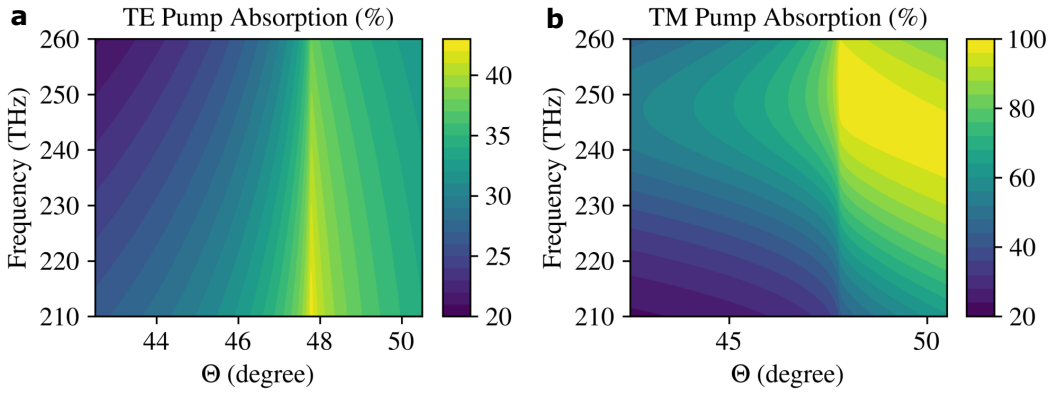


FIGURE 5.2: **Pump absorption. a**, The calculated TE pump absorption based on ellipsometric data. Due to $\theta = \theta_{\mathrm{pr}} = \theta_{\mathrm{pm}} + 3.4°$ we find the largest absorption close to 48°, the critical angle of the pump beam. **b**, The calculated TM pump absorption, which provides a strong absorption through the resonant ENZ plasmon.

frequency dependent absorbed pump intensity

$$I_{\mathrm{abs}} = \cos\left(\theta_{\mathrm{pm}}\right) A_{\mathrm{pm}}\left(\theta_{\mathrm{pm}}, f_{\mathrm{pm}}\right) I_{\mathrm{external}}. \tag{5.3}$$

The intensity-dependent ITO is modelled as an effective medium layer. We assume a linear intensity dependence of the Drude permittivity parameters, such that the

redshifted bulk plasma frequency can be approximated as:

$$\omega_p(I) = (1 + \omega_{p,2} I_{abs})\, \omega_{p,0}. \tag{5.4}$$

We use this to calculate the pump induced reflection of the probe $R_{pumped}(\theta_{pr}, f_{pr}, I)$. By least square fitting of the proportionality factor, $\omega_{p,2}$, to the experimental data in Figure 5.6c, we find $\omega_{p,2} = -0.38\ \%\,cm^2\,GW^{-1}$. The decrease in ω_p with increasing pump intensity can be understood as an increase of effective mass upon exciting the electrons in the non-parabolic conduction band of ITO, as discussed in various references [8, 67, 143, 131].

Several references in the literature have also reported simultaneous intensity-dependent changes to the scattering rate γ [8, 67]. However, using our full dynamic model (introduced below), the most convincing reproduction of our measured pump-probe dynamics is found when γ is independent or weakly dependent ($\gamma_2 <= 0.5\,\%\,cm^2\,GW^{-1}$) on intensity. We see this by comparing the experimental measurements (Figure 5.6b and Figure 5.3a) with the simulation cases of $\gamma_2 = 0$ (Figure 5.3b), $\gamma_2 = 5\,\%\,cm^2\,GW^{-1}$ (Figure 5.3c), $\gamma_2 = 0.5\,\%\,cm^2\,GW^{-1}$ (Figure 5.3d) and $\gamma_2 = -0.5\,\%\,cm^2\,GW^{-1}$ (Figure 5.3e). Moreover, fitting our angle and frequency-dependent results assuming both $\omega_{p,2}$ and a similarly defined γ_2 leads to an unphysical effect: a predicted small decrease in γ with increasing intensity, which arises due to a narrowing of the resonance feature observed in the R_{pumped} data in Figure 5.7 for increasing intensity. This is an artefact of our data analysis due to the oscillatory features for case II (where the ENZ resonance shifts spectrally through the probe), which become more prominent for increasing intensity. In any case, our results suggest that the nonlinear switching we observe in our samples is primarily due to intensity-dependent changes to ω_p, while intensity-dependent changes to γ are not identifiable in our experimental data.

Finally, we can contrast the magnitude of the observed optical nonlinearity with the most comparable data previously reported in the literature. Our fit result $\omega_{p,2} = -0.38\,\%\,cm^2\,GW^{-1}$ is surprisingly similar to that report by Alam et al. [8], corresponding to $\omega_{p,2} = -0.33\,\%\,cm^2\,GW^{-1}$, found from the absorbed intensity dependence plotted in Figure 5.4a. However, this is unexpected, as the nonlinear response should be dictated by absorbed energy density, and the sample thickness is different by a factor of ~ 5 (310 nm, while here 60 nm). In Figure 5.4b we plot the calculated absorption per length using transfer matrix modelling[83] as an indicator of absorbed energy density, which eventually causes the electron heating [131]. For our Kretschmann-Raether geometry with a film thickness of 60 nm, the absorbed energy density is reasonably constant across the layer (see blue line in Figure 5.4b). However, for direct transmission from air with a 310 nm thick film, the geometry used in reference [8], a significant exponential decay of energy density is expected across the film, with the decay length of 168 nm combined with the increased thickness lead to significant variation across the sample. This means that the Drude parameters, in this case, are not expected to be homogeneous, but will be strongly dependent on spatial position in the sample. Nevertheless, as a guide one would expect $\omega_{p,2}$ to be lower, by around a factor of 3, for the geometry used in reference [8] compared to the Kretschman-Raether geometry used here. We do not fully understand this discrepancy, but it may arise due to complications in the homogeneous analysis used in reference [8], or due to variations between ITO samples.

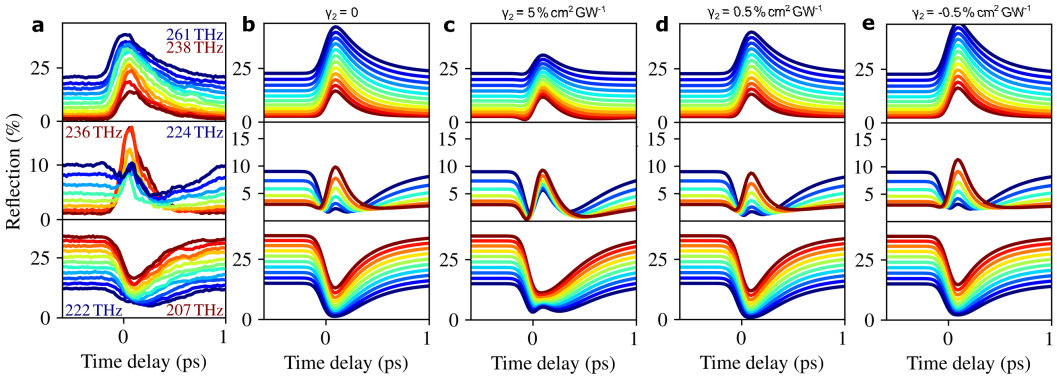


FIGURE 5.3: **Impact of γ_2 on model.** **a**, Experimental time delay scans presented in Figure 5.6b (pr: TM, 48.3°, $f_{pr} = 261 \ldots 207$ THz; pm: TE, 44.9°, $f_{pm} = f_{pr}$). **b**, Dynamic model results for $\omega_{p,2} = -0.38\,\%\,\mathrm{cm}^2\,\mathrm{GW}^{-1}$, estimated by fitting Figure 5.6c with an intensity dependent ω_p only. **c**, Adding a $\gamma_2 = 5\,\%\,\mathrm{cm}^2\,\mathrm{GW}^{-1}$, increases the estimated pumped resonance width. This leads to a reduced absorption depth of the switched plasmon (reflection doesn't go to zero), which is in line with experiments, but also leads to a decreased nonlinear reflection for the high frequencies, which is in this size incompatible with our findings. **d**, A more moderate effect given by up to $\gamma_2 \sim 0.5\,\%\,\mathrm{cm}^2\,\mathrm{GW}^{-1}$ could still be in-line with our measurements. **e**, $\gamma_2 = -0.5\,\%\,\mathrm{cm}^2\,\mathrm{GW}^{-1}$, has the opposite effect by even further increasing the nonlinear reflection for the high frequency case. Moreover, a scattering rate which decreases with increasing electron temperature seems unphysical.

Refractive index

To enable comparison to other studies using the refractive index and n_2 as the nonlinear optical parameters we provide the nonlinear refractive index depending on the absorbed intensity:

$$n(I_{abs}) = n_0 + n_2 I_{abs} + O(I_{abs}^2) \tag{5.5}$$

with

$$n_0 = \sqrt{\varepsilon_\infty - \frac{\omega_{p,0}^2}{\omega^2 + i\omega\gamma_0}} \tag{5.6}$$

$$n_2 = \frac{\omega_{p,0}^2}{2n_0} \frac{2\omega^2\omega_{p,2} - i\omega\gamma_0(\gamma_2 - 2\omega_{p,2})}{(\omega^2 - i\omega\gamma_0)^2} \tag{5.7}$$

To give an example, n_2 takes the value of $(0.01 - i * 0.016)\,\mathrm{cm}^2\,\mathrm{GW}^{-1}$ for our ITO sample case and 1250 nm wavelength. This equation may result in an approximate conversion for other ITO sample designs, but $\omega_{p,2} = 0.38\,\%\,\mathrm{cm}^2\,\mathrm{GW}^{-1}$ is expected to be different for materials with different band-structures such as AZO. Furthermore, to compare between different layer thicknesses and/or pulse lengths, one should also consider that the absorbed energy density is the expected scaling parameter of the thermal nonlinearity, not the absorbed intensity.

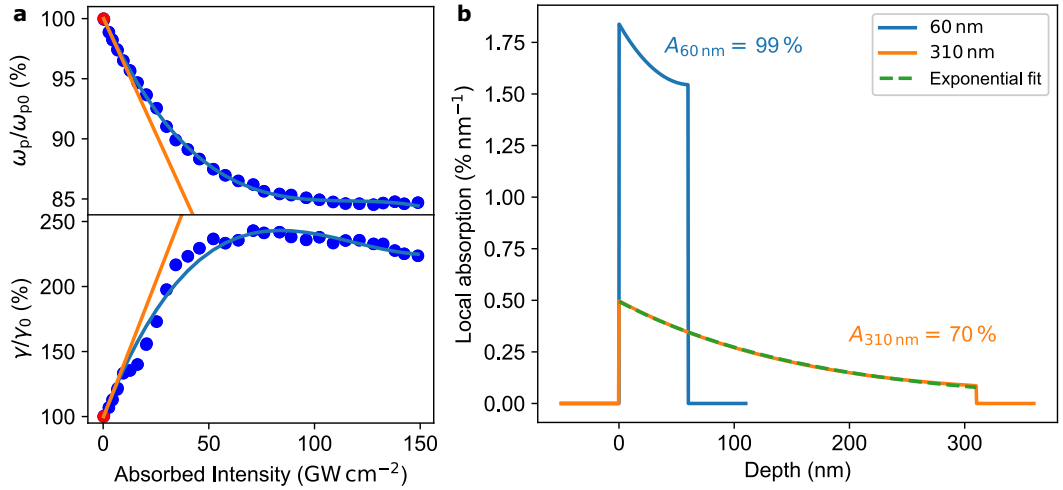


FIGURE 5.4: **Comparison of different sample thickness.** We compare to the previous studies of Alam et al. [8] which features a 310 nm ITO layer with similar optical properties to ours (TM polarized). **a**, We convert the given intensity-dependent complex refractive index given in [8] to ω_p and γ by assuming a Drude model. Linear contributions for the low intensity fit region give $\omega_{p,2} = -0.33\,\%\,\mathrm{cm}^2\,\mathrm{GW}^{-1}$ and $\gamma_2 = 4.6\,\%\,\mathrm{cm}^2\,\mathrm{GW}^{-1}$. **b**, The 310 nm film (in air, $\theta = 30°$, $\lambda = 1240$ nm) produces an absorbed energy density which varies significantly with distance and is described by an exponential decay length of $\tau = 168$ nm. However, the 60 nm thin film (prism, $\theta = 45°$, $\lambda = 1200$ nm) shows a relatively constant absorption throughout the layer as the decay length is noticeably longer than the thickness ($\tau \gg 60$ nm).

Time varying medium (Dynamic model)

In this model we concentrate on the full temporal dynamics of a probe pulse experiencing a time-dependent modulation of the medium's dielectric function, induced by the pump pulse. Without any time modulation of the permittivity, Maxwell's equations for an incident TM polarised wave are

$$\nabla \times \mathbf{E} = -\mu_0 \frac{\partial H}{\partial t}\hat{\mathbf{z}} \tag{5.8}$$

and

$$\nabla H \times \hat{\mathbf{z}} = \frac{\partial \mathbf{D}}{\partial t} = \epsilon_0 \frac{\partial}{\partial t}\left[\mathbf{E}(t) + \int_0^\infty \chi\left(\frac{t'}{\Delta}\right)\mathbf{E}\left(t - t'\right)dt'\right] \tag{5.9}$$

where $\chi\left(\frac{t'}{\Delta}\right)$is the time domain susceptibility, $H = H_z$, and Δ is the characteristic timescale of the material response. In our case the susceptibility function χ itself varies in time due to the heating induced changes of the plasmon frequency ω_p (see Figure 5.5). The latter are modelled as a convolution of the pump pulse and an exponential decay with $\tau = 300$ fs, normed such that ω_p given in Equation 5.4 corresponds to the extremum. We thus modify (5.9) so that the susceptibility has an additional dependence on $t - t'$,

$$\nabla H \times \hat{\mathbf{z}} = \epsilon_0 \frac{\partial}{\partial t}\left[\mathbf{E}(t) + \int_0^\infty \chi\left(\frac{t - t'}{\Delta_1}, \frac{t'}{\Delta_2}\right)\mathbf{E}\left(t - t'\right)dt'\right] \tag{5.10}$$

We now have two timescales Δ_1 (the modulation of the permittivity, such as electron

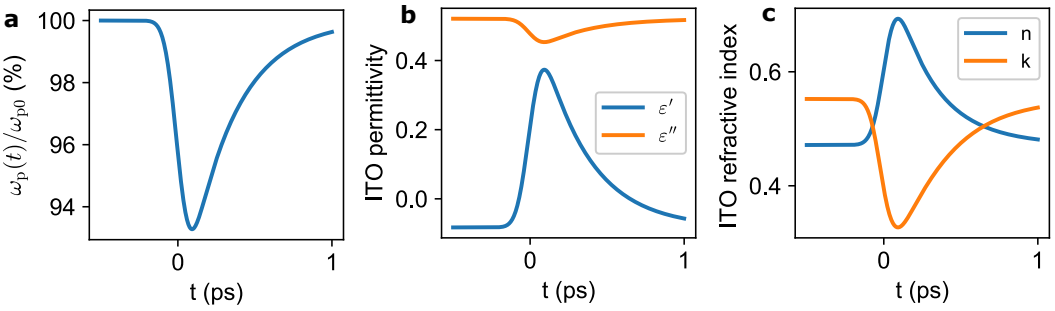


FIGURE 5.5: **Time dependent bulk plasmon frequency.** The convolution of the pump pulse and an exponential decay with $\tau = 300\,\text{fs}$ is normed such that $\omega_p(I)$, given in Equation 5.4, corresponds to the extremum. **a**, The $\omega_p(t, I)$ is used to calculate the time dependent material parameters such as χ during interaction with the pump beam. For the plotted case we assume $\theta = 48.3°$, $I_{\text{external}} = 70\,\text{GW cm}^{-2}$ and a TE polarized pump (as seen in Figure 5.6b). For the case of $f_{pr} = 240\,\text{THz}$ we plot the corresponding permittivity (**b**) and refractive index (**c**).

heating and cooling effects), and Δ_2 (the timescale over which the material responds). These timescales are assumed to be such that the modulation is much slower than the material response $\Delta_1 \gg \Delta_2$ and the material response tends to zero at large delay times $\chi(\tau_1, \tau_2 \to \infty) = 0$. Taking the curl of (5.10) and applying (5.8) we find the equation for the out of plane magnetic field in a homogeneous region of space

$$\nabla^2 H - \frac{1}{c^2}\frac{\partial^2 H}{\partial t^2} - \frac{1}{c^2}\frac{\partial}{\partial t}\int_0^\infty \chi\left(\frac{t-t'}{\Delta_1}, \frac{t'}{\Delta_2}\right)\frac{\partial H(t-t')}{\partial t}dt' = 0 \tag{5.11}$$

In general, this is an integro-differential equation that is difficult to solve. To make progress we use multiple scales perturbation theory[144], and define two time variables $\tau_1 = t/\Delta_1$ and $\tau_2 = t/\Delta_2$. The out of plane magnetic field is approximated as a function of these two time variables

$$H(t) = H\left(\frac{t}{\Delta_1}, \frac{t}{\Delta_2}\right) = H(\tau_1, \tau_2) \tag{5.12}$$

and the time derivatives can be written as

$$\frac{\partial}{\partial t} = \frac{1}{\Delta_1}\frac{\partial}{\partial \tau_1} + \frac{1}{\Delta_2}\frac{\partial}{\partial \tau_2} \tag{5.13}$$

Given that $\chi(\tau_1 - \tau_1', \tau_2')$ decays with increasing τ_2' over which time τ_1' has not increased appreciably from zero we can expand the integral kernel around $\tau_1' = 0$ with the leading order approximation being $\chi(\tau_1 - \tau_1', \tau_2') \sim \chi(\tau_1, \tau_2')$. Further terms are proportional to increasing powers of Δ_2/Δ_1. Equation (5.11) is thus to leading order in the small quantity Δ_2/Δ_1

$$\nabla^2 H(\tau_1, \tau_2) - \frac{1}{\Delta_2^2 c^2}\frac{\partial^2 H(\tau_1, \tau_2)}{\partial \tau_2^2} - \frac{1}{\Delta_2^2 c^2}\int_0^\infty \chi(\tau_1, \tau_2')\frac{\partial^2 H(\tau_1, \tau_2 - \tau_2')}{\partial \tau_2^2}d\tau_2' = 0 \tag{5.14}$$

Equation (5.14) is different from (5.11) in an important way. Having dropped the terms involving the timescale $\tau_1 = t/\Delta_1$ from the integral and the derivatives, the

modulation of the dielectric function neither appears in the derivatives nor the integration. The time variable τ_1 is now merely a label, and we may solve (5.14) as we would in the time-independent situation, with the dispersive permittivity $\varepsilon(\tau_1, \omega)$. Note that to improve upon this approximation, one can add back in the terms proportional to the ratio of timescales Δ_2/Δ_1, expanding perturbatively.

We now apply this result to the experiment. We have a total internal reflection geometry with fixed permittivities, $\varepsilon_1 = 2.1$ in the region $x < 0$, and $\varepsilon_3 = 1$ in the region $x > d$. In the remaining region $0 < x < d$ we have a layer of material with a permittivity that is modulated over time $\varepsilon_2(\tau_1, \omega)$ with the same separation of timescales assumed above. A pulse is incident at a fixed angle θ from the region $x < 0$. The reflection coefficient (which, as with the permittivity is now a function of modulation time τ_1, and incidence frequency ω) is given by

$$r(\tau_1, \omega) = \frac{\cos(k_2 d)\left(\frac{k_2}{k_3\epsilon_2} - \frac{k_2\epsilon_1}{k_1\epsilon_2}\right) - i\sin(k_2 d)\left(1 - \frac{k_2^2\epsilon_1}{k_1 k_3 \epsilon_2^2}\right)}{\cos(k_2 d)\left(\frac{k_2\epsilon_1}{k_1\epsilon_2} + \frac{k_2}{k_3\epsilon_2}\right) - i\sin(k_2 d)\left(1 + \frac{k_2^2\epsilon_1}{k_1 k_3 \epsilon_2^2}\right)}, \tag{5.15}$$

where $k_0 = \omega/c$, $k_1 = k_0\sqrt{\varepsilon_1}\cos(\theta)$, $k_2 = k_0\sqrt{\varepsilon_2(\tau_1, \omega) - \varepsilon_1\sin^2(\theta)}$, and $k_3 = k_0\sqrt{1 - \varepsilon_1\sin^2(\theta)}$. If the spectrum of the incident pulse is $a(\omega)$ then the total pulse in the first region $x < 0$ is the sum of the incident spectrum plus the incident spectrum weighted by the modulated reflection coefficient $r(\tau_1, \omega)$:

$$\begin{aligned} H(\tau_1, \tau_2) &= H_{\text{inc}}(\tau_1, \tau_2) + H_r(\tau_1, \tau_2) \\ &= \int_{-\infty}^{\infty} a(\omega)\left[e^{ik_0(\cos(\theta)x - c\Delta_2\tau_2)} + r(\tau_1, \omega)\, e^{-ik_0(\cos(\theta)x + c\Delta_2\tau_2)}\right]\frac{d\omega}{2\pi} \\ &= \int_{-\infty}^{\infty} a(\omega)\left[e^{ik_0(\cos(\theta)x - ct)} + r\left(\frac{t}{\Delta_1}, \omega\right)e^{-ik_0(\cos(\theta)x + ct)}\right]\frac{d\omega}{2\pi}. \end{aligned} \tag{5.16}$$

Here, the $+ik_0$ correspond to the forward propagating terms, while $-ik_0$ correspond to the backward propagating terms. The spectrum of the reflected pulse is found by simply taking the Fourier transform of the second term in the square brackets of (5.16), at $x = 0$:

$$H_r(\omega) = \int_{-\infty}^{\infty} H_r(t)e^{i\omega t}dt = \int_{-\infty}^{\infty}\frac{d\omega'}{2\pi}a(\omega')\, r(\omega - \omega', \omega'), \tag{5.17}$$

where

$$r(\omega_1, \omega_2) = \int_{-\infty}^{\infty} r\left(\frac{t}{\Delta_1}, \omega_2\right)e^{i\omega_1 t}dt. \tag{5.18}$$

Equation (5.17) indicates that the spectrum of the reflected pulse is the convolution of the incident spectrum with the two frequency reflection coefficient $r(\omega_1, \omega_2)$.

Finally, to compare to the experiment, we calculate the relative reflected power given in the main text. To do this we then take the ratio of the integrated incident and reflected Poynting vectors. For instance, the time-integrated incident power is

$$P_{\text{inc}} = \int_{-\infty}^{\infty}\hat{x}\cdot S(t)dt = \int_{-\infty}^{\infty}E_{y,\text{inc}}(t)H_{\text{inc}}(t)dt = \frac{\cos(\theta)}{c\epsilon_0\sqrt{\epsilon_1}}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}|a(\omega)|^2, \tag{5.19}$$

and therefore the relative reflected power is

$$R = \frac{\int_{-\infty}^{\infty} \frac{d\omega}{2\pi} |H_r(\omega)|^2}{\int_{-\infty}^{\infty} \frac{d\omega}{2\pi} |a(\omega)|^2}, \tag{5.20}$$

where $H_r(\omega)$ is calculated using Equation 5.17, and the time-frequency reflection coefficient in Eq. (5.15). Equation (5.20) is used to calculate the reflected power for the different pump pulse arrival times in Figure 5.6b and was already applied for the discussion surrounding the γ_2-dependence in Figure 5.3.

5.4 Nonlinearity Near the ENZ Plasmon Resonance

Recent measurements have shown that pumping below the band-gap of a transparent conducting oxide leads to carrier heating which subsequently results in an increased effective mass and decreased plasma frequency[134]. Such an intensity-dependent plasma frequency should materialise as a shifting resonance frequency of an ENZ plasmon. We study this effect in the ENZ plasmon dispersion of a 60 nm ITO thin film using a pump-probe scheme. To begin, we pump with TE polarisation and probe with TM polarisation – this removes coherent interference of pump and probe pulses at the cost of less efficient absorption. Figure 5.6a shows three typical pump-probe measurements where we pump and probe different regions of the plasmon resonance. For a probe frequency of 240 THz (red, case I) the ENZ resonance redshifts away from the probe, the absorption decreases and we see an increase in reflectivity. For 214 THz (green, case III) the ENZ resonance shifts spectrally towards the probe, the absorption increases and we see a decrease in reflectivity. Both cases display expected temporal dynamics: a fast, ~100 fs (pulse limited) initial change, followed by a slower , ~1 ps thermal relaxation, similar to previous studies [145, 65, 8].

For 227 THz (orange, case II) the ENZ resonance shifts spectrally through the probe, and we observe some rather unusual dynamics in the pump-probe signal that, to the best of our knowledge, have not previously been observed and discussed. For this case, one initially observes a decrease in reflection, followed by an increase in reflection, as heating causes the resonance to shift through the probe frequency. Subsequent cooling then returns the resonance to its starting frequency. These combined effects cause multiple oscillations in the pump-probe dynamics. To better describe these effects, we introduced a dynamic model, which captures effects arising due to rapid time-dependent changes to the reflection coefficient (details in section 5.3). With this model, we can reproduce the features seen in our pump-probe signals: the three broad types of behaviour are shown for various frequencies in Figure 5.6b and modelled using the time-varying reflection coefficient implied directly from experimental observations. For the first two cases, it is straightforward to define maxima/minima in reflection (R_{pumped}) relative to the initial reflection (R_0), as labelled in Figure 5.6a. Defining a maximal response in the oscillatory case is problematic - for simplicity, we define $R_{\text{pumped}} = R_0 + \Delta R_{\text{max}} + \Delta R_{\text{min}}$ for all measured time delay scans. While not being able to resolve the transient features seen in Figure 5.6b, this reduces the complex dynamical information contained in one complete time delay scan into two quantities of interest: the reflection before the arrival of the pump (R_0) and a quantity describing the maximally changing reflection on photoexcitation (R_{pumped}). This allows us to condense information from many measurements into one colour plot.

By varying both the incident angle and the degenerate pump/probe frequency, one can observe the plasmon dispersion shift (Figure 5.6c). We illuminate the sample

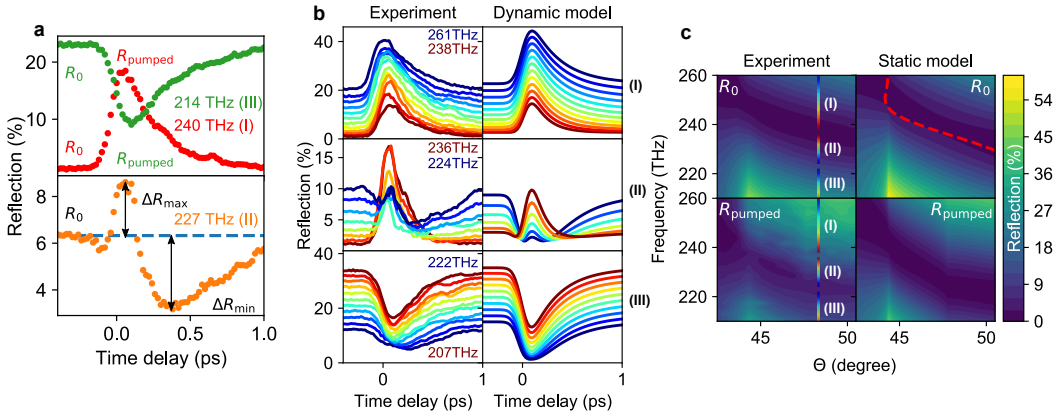


FIGURE 5.6: **Time dependence of the nonlinear reflection for degenerate pump and probe frequencies. a**, Three typical pump-probe measurements, where: I) the probe frequency (240 THz, red) is larger than the plasmon resonance frequency, II) the resonance shifts spectrally through the probe during pumping (227 THz, orange), and III) the probe frequency is smaller than the pumped resonance (214 THz, green). The probe is TM polarized to study the plasmon response, while the 70 GW/cm^2 pump is TE polarized to avoid coherent contributions. (pr: TM, 48.3°, f_{pr} = 240, 227, 214 THz; pm: TE, 44.9°, $f_{pm} = f_{pr}$). **b**, Various pump-probe measurements divided into the previously discussed three case types. The dynamic model shows the reflection coefficient for a time varying effective medium. (pr: TM, 48.3°, f_{pr} = 261 . . . 207 THz; pm: TE, 44.9°, $f_{pm} = f_{pr}$). **c**, Scans of the probe reflection over incoming angle and degenerate frequency, showing the initial R_0 (top) and the pumped case R_{pumped} (bottom). The vertical jet coloured lines at 48.3° indicate the measurements presented in b. The transfer matrix model on the right shows the expected initial reflection at the top with the ENZ plasmon dispersion from Figure 5.1c (red dashed) with R_{pumped} based on the carrier heating nonlinearity at the bottom. (pr: TM, 42.5 . . . 50.5°, f_{pr} = 261 . . . 210 THz; pm: TE, 39.1 . . . 47.1°, $f_{pm} = f_{pr}$).

with a TE pump of 70 GW/cm^2 intensity. Note that we avoided high intensities only because of the index-matching fluid, which could be avoided by depositing ITO directly on a prism. We measure a red-shift of the ENZ plasmon resonance, as seen for the pumped reflection (R_{pumped}) relative to the initial reflection (R_0). We interpret these changes using a simplified "static" model, which calculates the reflection using a transfer matrix model assuming an effective medium with an intensity-dependent permittivity for the ITO layer (details in section 5.3, static model). We assume a linear intensity-dependent shift of the plasma frequency, $\omega_p(I) = (1 + \omega_{p,2} I)\,\omega_{p,0}$, where I is the calculated absorbed intensity and $\omega_{p,2}$ is the nonlinear fit parameter. This approach gives good agreement for the red-shifting behaviour of the resonance, while also confirming some more subtle features, such as the critical angle feature for the probe pulse just below 45°, as well as a second critical angle feature near 48°, which arises due to the 3.4° difference in angle between pump and probe. We find best agreement with the data for an intensity-dependent red-shifting of the plasma frequency, described by $\omega_{p,2} = -0.38\,\%/\frac{GW}{cm^2}$. This is the general behaviour expected for heating of the electron plasma in ITO, an effect that arises due to the non-parabolicity of the conduction band in this material [8, 67, 143]. However, as discussed in section 5.3, comparison between experiment and modelling suggests

that the electron heating is weaker in our ITO compared to that previously reported in [8], as our extracted value for $\omega_{p,2}$ is similar despite an increased local intensity arising due to the Kretschman geometry. We do not fully understand this discrepancy, but it may arise due to complications in the homogeneous layer assumption of the z-scan analysis used in ref [8] (highlighted here surrounding Figure 5.4), or due to variations between ITO samples. ITO is well known for its structure and resulting optical properties being sensitive to slight changes in manufacturing[141]. As further discussed in section 5.3, we can not identify changes to the scattering rate as readily as the changes to the plasma frequency, as the scattering rate affects mainly the width of the plasmon resonance, which is impacted by an artefact of our analysis using our static model (discussed below). However, it is expected that heating should have minimal effect on scattering rate in transparent conducting oxides due to the dominance of impurity scattering [131]. As shown explicitly in Figure 5.3, the comparison between our dynamic model and experimental data suggests that pump induced changes to scattering rate are negligible.

5.5 Intensity Dependent Resonance Shift

In Figure 5.7, we vary pump intensity while fixing the pump frequency to 240 THz and incident angle to just beyond the critical angle of the pump (48.3°). The pump polarisation is TE, i.e. non-resonant with the ENZ plasmon. On increasing pump intensity, we observe a clear redshift of the ENZ plasmon, with approximately linear intensity dependence. We also observe a more subtle effect: an apparent slight narrowing of the resonance. This is an artefact of our analysis, arising from the oscillatory features for case II, which become more prominent for increasing intensity. Since we are unable to remove completely these effects in our static analysis, they give rise to a slight distortion of the resonance lineshape for high intensities.

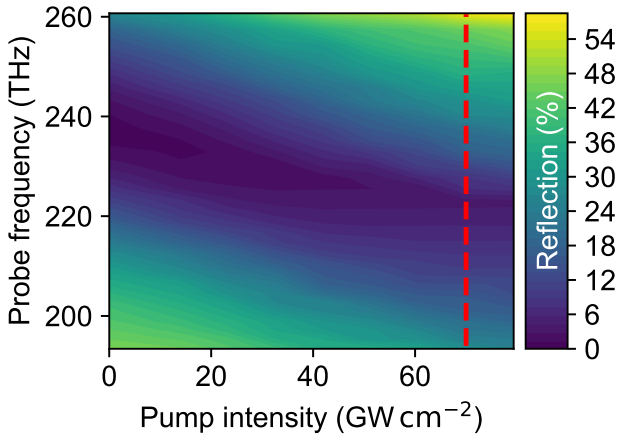


FIGURE 5.7: **Intensity dependence of the ENZ-resonance position.** Intensity dependent R_{pumped} for a constant incident probe angle of Θ = 48.3° and a pump with a constant frequency of 240 THz (TE polarized, i.e. non-resonant with the ENZ plasmon). The ENZ plasmon red shifts with increasing pump intensity. The red dashed line indicates the 70 GW/cm^2 used for the measurements in Figure 5.6. (pr: TM, 48.3°, f_{pr} = 260 . . . 190 THz; pm: TE, 44.9°, f_{pm} = 240 THz)

5.6 Coherent vs Incoherent Response

Finally, we aim to maximise switching: by exciting with a TM pump, a more efficient energy deposition via the ENZ plasmon resonance is expected. We investigate the non-degenerate frequency dependence of the nonlinear material response by fixing the external pump intensity to 70 GW/cm^2 and incident pump angle to 48.3°, and compare TE and TM excitation. Firstly, the pump absorption for TE excitation is only weakly frequency-dependent. For this reason, we observe a pump frequency-independent shift of the ENZ resonance of 14 THz in Figure 5.8a. For TM excitation, seen in Figure 5.8b, we see two additional interesting aspects. For on-diagonal (degenerate) measurements, we can identify a noticeably larger reflection due to two-beam coupling (TBC). We note that the TBC contribution may be concealed, and not readily separated, within other degenerate pump-probe measurements in the literature. In Figure 5.8c we compare the coherent and incoherent contributions to our signal. For the non-degenerate case, we observe thermal switching behaviour resulting in a large change in reflection from $R_0 \sim 1\,\%$ to $R_{\text{pumped}} \sim 30\,\%$. This corresponds to a shift of the plasmon resonance frequency of 20 THz, which is more than four times the spectral width of a 100 fs pulse. In the degenerate case, we see a further increase of the differential reflection to $\Delta R \sim 45\,\%$ due to the interference between the pump and probe beams (details in subsection 5.7.1). Only by systematically varying both pump and probe frequencies independently have we been able to identify this coherent contribution to switching.

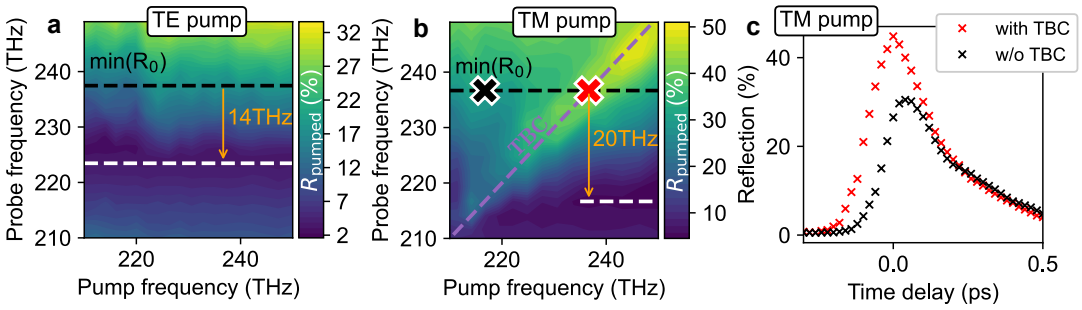


FIGURE 5.8: **Non-degenerate frequency dependence and coherent contribution.** We examine the pump-probe frequency dependence of R_{pumped} for 70 GW/cm^2. **a**, Using a TE pump, the frequency shift of the resonance is 14 THz, and independent of pump frequency. The dashed lines indicate the initial (black) and pumped (white) ENZ plasmon resonance frequency. (pr: TM, 48.3°, $f_{\text{pr}} = 250 \ldots 210$ THz; pm: TE, 44.9°, $f_{\text{pm}} = 250 \ldots 210$ THz). **b**, A TM pump highlights the two-beam coupling (TBC) for equivalent polarization and frequency. The higher pump absorption through the plasmon resonance leads to a larger, 20 THz shift of the resonance frequency. (pr: TM, 48.3°, $f_{\text{pr}} = 250 \ldots 210$ THz; pm: TM, 44.9°, $f_{\text{pm}} = 250 \ldots 210$ THz). Markers indicate the time delay scans compared in **c**, where the probe frequency is 237 THz, while the pump frequency is either 237 THz (red, with TBC) or 217 THz (black, without TBC). (pr: TM, 48.3°, $f_{\text{pr}} = 237$ THz; pm: TM, 44.9°, $f_{\text{pm}} = 237, 217$ THz)

5.7 Additional Checks

5.7.1 Two-Beam Coupling Contribution

When two coherent, co-polarised, near-degenerate beams impinge upon a plane, they generate an interference pattern. This gives rise to a spatially dependent change to the index of the material, which can cause diffraction generated signals in the experiment. The description we give below refers specifically to our pump-probe geometry, though it is important to realise that similar coherent signals (i.e. resulting from interference and diffraction) can result in many types of nonlinear optical measurement, even those employing a single focused beam [146].

For the geometry used in Figure 5.8c, we model beams in glass (prism) with wavelength 1250 nm wavelength, $\theta_{pr} = 44.9°$ and $\theta_{pm} = 48.3°$. These beams will result in the interference pattern inside the ITO, as shown in Figure 5.9. For the beams used in our experiment ($I_{pm} = 70\,GW\,cm^{-2}$ and $I_{pr} = 0.13\,GW\,cm^{-2}$) we expect the interference pattern shown in Figure 5.9b, with a spatially dependent oscillation in intensity of $\pm 4\,GW\,cm^{-2}$. Due to the intensity-dependent index of refraction in ITO, one also expects a spatial dependence to the local index of refraction. Assuming the linear intensity dependence present in Equation 5.4, one can expect the refractive index of the ITO layer to roughly resemble that shown in Figure 5.9c.

A spatial profile in the index of refraction will act as a diffraction grating, scattering pump light into the direction of the probe beam, and subsequently into our detector. While the spatial modulation in the index is relatively small, leading to a relatively weak scattering effect, we only require a small intensity of the much stronger pump beam scattered in the direction of the probe to give a large switching signal. To obtain an estimate of this contribution, we use COMSOL to calculate the diffraction pattern expected from the spatially varying index shown in Figure 5.9c, assuming a 60 nm ITO layer with a uniform refractive index in the z-direction. This predicts a 1st order diffraction of ~ 0.1 % of the pump beam that will be scattered in the direction of the probe beam. Hence, the signals we measure in our detector are not only based on the zero-order reflection of the probe, but also a contribution of the first-order diffraction of the pump. Considering the spatial overlap required for the pump to be scattered, we can directly estimate the contribution of the scattered pump to the "differential reflection" signal: this corresponds to

$$\Delta R_{CC} \sim \frac{0.1\,\% * 70\,GW\,cm^{-2}}{0.13\,GW\,cm^{-2}} \sim 54\,\% \tag{5.21}$$

of the probe intensity. This is even larger than the difference in peak TM and TE experimental signals, which is 15 % (see Figure 5.6c). However, due to their different temporal dynamics, the thermal and coherent contributions to the signal are expected to add sub-linearly. Moreover, we note that the refractive index relation used to calculate the scattering of the pump assumes full thermalisation, whereas the ITO will heat up during the evolution of the pump pulse. Thus the 54 % predicted above is an overestimate, and we believe it to be consistent with the measured value of 15 %.

For a non-degenerate pump and probe, one expects a non-stationary interference pattern, i.e. one that changes quickly with time. For our geometry, one can easily show that the effects of the grating will be washed out within the ~ 100 fs of our pulses when the pump and probe differ by only a few nm. Again, this is in agreement with our experiments presented in Figure 5.6b, which show coherent signals only for near-degenerate measurements.

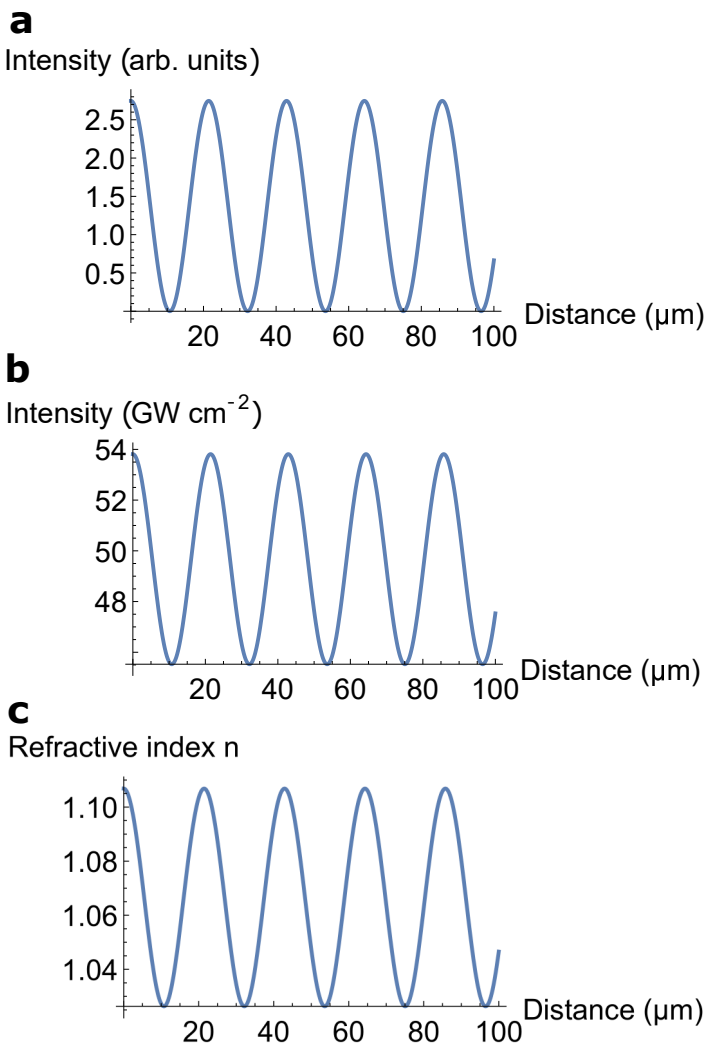


FIGURE 5.9: **Diffraction through interference induced refractive index grating. a**, Interference pattern of two equally intense beams (44.9°, 1250 nm) and (48.3°, 1250 nm) in glass. **b**, Interference pattern of 70 GW/cm^2 pump (44.9°, 1250 nm) and 0.13 GW cm^{-2} probe (48.3°, 1250 nm) at the glass/ITO interface. **c**, Corresponding refractive index distribution resulting from B, based on Equation 5.4.

5.7.2 Probe Polarisation Dependence of the Nonlinear Effect

In Figure 5.10, we compare the dynamical model of Figure 5.6b with the corresponding TE probe case. The nonlinear optical effects on a TE polarised probe are small, as the shifting resonance feature is only accessible for TM polarised light (see e.g. in Figure 5.2). However, these small nonlinear changes to the TE reflection/absorption are beneficial as changes to the TE pump absorption are negligible, enabling easier modelling of time-dependent effects.

5.7.3 TE-TE Measurement

Due to the weak nonlinear response of the TE probe, we expect only small thermal contributions to nonlinear measurements as discussed in subsection 5.7.2. However, if the pump is also TE polarised one expects a notable two-beam coupling contribution as the refractive index modulation inside the ITO layer remains similar to the case

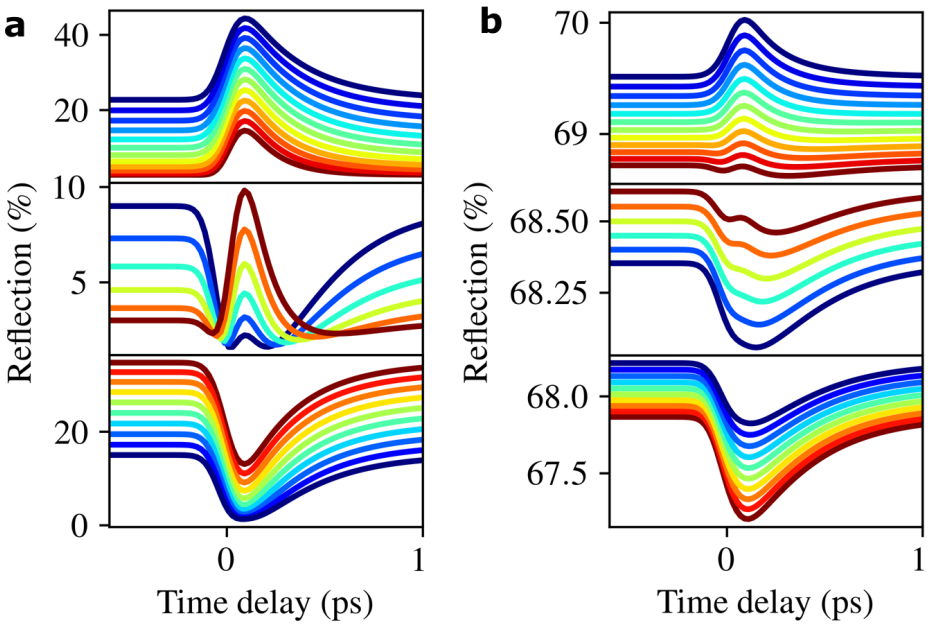


FIGURE 5.10: **Polarization dependent nonlinear reflection. a**, The TM probe case as seen in Figure 5.6b or Figure 5.3 (pr: TM, 48.3°, $f_{pr} = 261 \ldots 207$ THz; pm: TE, $f_{pm} = f_{pr}$). **b**, The same case except for the probe being TE polarized, leading to only small reflection changes of $\Delta R < 1\,\%$. (pr: TE, 48.3°, $f_{pr} = 261 \ldots 207$ THz; pm: TE, $f_{pm} = f_{pr}$)

discussed in subsection 5.7.1. This is indeed what we observe in the TE-TE experiment as seen in Figure 5.11. The resulting TBC can be seen as a 10 % relative reflection increase when near degeneracy of pump and probe, and is significantly larger than the thermal effect in the same measurement.

5.7.4 Coverslip Measurement

We have performed measurements at 45°, with both beams TM polarised and both wavelengths set to 1200 nm (see Figure 5.12). We first insert the ITO sample, align and measure the time delay scan (1). After that, we exchange the sample for a coverslip and only adjust the tilt of the sample holder and the angle of the sample holder to maximise reflection and compensate changes due to slight changes in index matching fluid binding to the prism. Then we measure time delay scan (2). Finally, we change back to ITO, again, only adjusting tilt and angle to retrieve the maximum reflection (for 1500 nm, due to higher reflection). With the then measured scan (3) we show that the nonlinear response was perfectly reproduced. Importantly, the response of the coverslip alone (2) is considerably smaller.

We have checked the absence of signal for case (2) for other angles, and signals are always considerably smaller than when the ITO is present, hence we can unequivocally attribute the bulk of the signal to the ITO film.

5.8 Conclusions

In conclusion, we present an investigation into the all-optical switching of ENZ plasmons. We identify two contributions to our nonlinear signal: a thermally driven switching process results in a shift in the plasmon resonance frequency of 20 THz for a

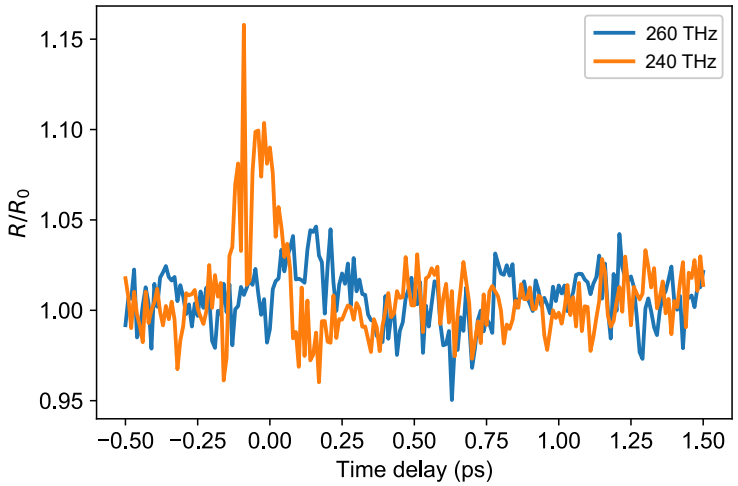


FIGURE 5.11: **TE-TE measurement.** The absence of the thermal response for a TE probe leads to an isolated clear TBC feature when pumping with a TE polarized beam. (pr: TE, 45°, f_{pr} = 240 THz; pm: TE, f_{pm} = 240, 260 THz)

pump intensity of 70 GW/cm^2, while the additional two-beam coupling is observed for degenerate pump and probe frequencies. In total we observe switching of more than one order of magnitude, from $R_0 \sim 1$ % to $R_{pumped} \sim 45$ % , resulting entirely from resonant conditions allowed by the geometry. For comparison, switching of ITO in air with 70 GW/cm^2 has been shown to result in a change in transmission from 12 % to 34 % [8], for samples with five times the material thickness than those studied here. Considering the 16.5 dB extinction ratio achieved in this proof of principle study, switching from near-perfect absorption to high reflection is useful from a signal processing point of view , and could pave the way towards optical plasmon switching at telecom frequencies. Currently, achievable switching rates are limited by further experimental requirements such as the need for high-pulse-energy femtosecond laser amplifiers. While the switching performance of the layer and the requirement for high intensities could be further improved by more elaborate layer designs and/or further material improvements such as in carrier mobility, achievable switching rates will be ultimately restricted by limitations in the repetition rate of pulsed optical sources. Finally, the compatibility with CMOS fabrication technique makes thin TCO layers and their ENZ plasmon feature a compelling new route for nonlinear integrated photonics applications without the need for nanostructure or building additional cavities, while better matching the spatial modes used in photonic circuit systems.

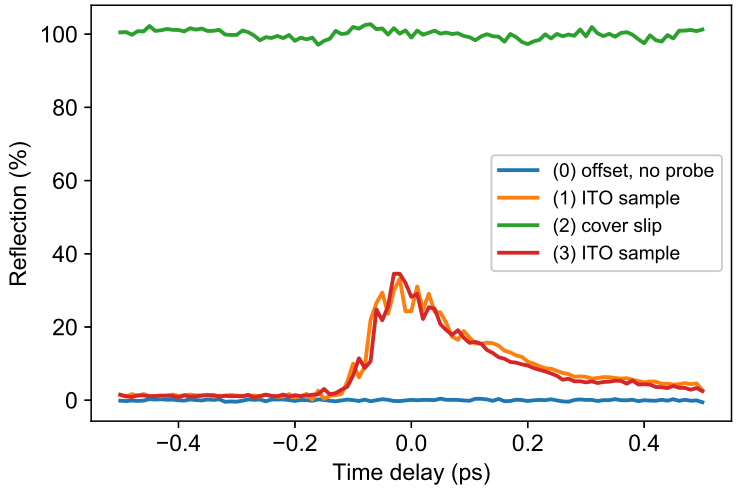


FIGURE 5.12: **Coverslip measurement.** To show the negligible nonlinear response of the prism, index matching fluid and prism we present a delay scan without ITO layer (2). To ensure the alignment is not being influenced we take ITO measurements before and after with the same alignment procedure (1,3). We also show the electronic offset without a probe, corresponding to $R = 0$ %. (pr: TM, 45°, $f_{\mathrm{pr}} = 250$ THz; pm: TM, $f_{\mathrm{pm}} = 250$ THz)

Chapter 6

Transient Dichroic Response of a Symmetric Plasmonic Nanoparticle Array

Abstract

Ultrafast polarisation control of the light is crucial for various optical applications. However, designs are often limited to specific incoming polarisations, angled incidence, or high intensities. We employ cross-shaped gold nano-antennas on top of a thin indium tin oxide layer to improve these issues. Large amplitude changes of $\Delta T/T_0 = -50\,\%$ are measured for an intensity of $5\,\mathrm{GW\,cm^{-2}}$. The transient dichroic design enables a more than seven times stronger response for the probe being parallel to the pump relative to a perpendicular orientation. Additionally, a phase-shifting regime of up to 8.7° was measured. Combined phase and amplitude modulation present a promising elliptical polarisation switching in the telecommunications wavelength range.

Author Contribution Statement

For this chapter, I performed the nonlinear optical measurements and analysis. The corresponding paper is prepared for submission as "Time-varying dichroic response of a symmetric plasmonic nanoparticle array".

6.1 Introduction

Linear optical responses of plasmonic nanoparticles depend on three properties: refractive index of the nanoparticles, shape and size of the nanoparticles, and the refractive index of the surrounding material. Thus a plasmonic nanoparticle that exhibits four-fold rotational symmetry in the transverse plane and is embedded in a homogeneous medium exhibits polarisation-independent optical response. An induced nonlinear optical response that may sufficiently change the refractive index of the nanoparticle or the surrounding medium in an inhomogeneous manner can be used to weaken the relationship between geometric and optical symmetries. However, the maximum possible magnitude of refractive index change of a typical noble metal or a dielectric is only perturbative to its linear refractive index [147]. Consequently, a strong, transient breaking of optical symmetry through only a weak, nonlinear modification of the refractive index of metal is impossible to achieve.

Recently, ENZ layers have become a prominent choice as a nonlinear optic environment as already introduced and investigated in the context of an ENZ plasmon in chapter 5. Planar ENZ films already present great promise in terms of switching depth ($\Delta I/I_0$, with I as the measured signal intensities) and nano-scale size. However, several issues remain to be addressed, including high-intensity requirements, oblique incidence, and polarisation dependence. The potential of utilising nano-antennas with thin ITO layers for reduced-intensity requirements and normal incidence has already been demonstrated[69]. Strongly coupled nano-antennas with ENZ modes have been used for effects such as negative refraction [70]. High contrast elliptical polarisation manipulation in ITO coupled to a plasmonic nano-antenna array have been presented with an asymmetric array and a 310 nm thick film [148].

Here, we show that by incorporating an epsilon-near-zero (ENZ) material made of a 23-nm-thick indium tin oxide (ITO) film as a substrate, we can strongly and reversibly break the optical symmetry of a Greek-cross-shaped plasmonic particle when the structure is excited by a pulsed laser. Specifically, we demonstrate large, ultrafast dichroic responses from a symmetric plasmonic nanoparticle array.

We present the concept of our polarisation switching device in Figure 6.1. The resonance wavelength of a plasmonic dipole antennas array with constant length and width depends on the refractive index of the surrounding medium. Unlike a simple dipole antenna, a Greek-cross-shaped plasmonic antenna has two principal axes. Such an antenna embedded in an isotropic host medium exhibits polarisation-invariant resonances in the linear regime. Under a sufficiently strong optical excitation, the nearfield of the antenna induces a nonlinear optical response to the surrounding medium only in the vicinity of the antenna arm that is parallel to the polarisation vector of the incident light beam. Consequently, a structural birefringence is induced in the surrounding medium. In response to the induced birefringence, the resonance wavelengths of the antenna become spectrally non-degenerate. That amplifies the induced structural birefringence and introduces frequency- and polarisation-dependent absorption. In our implementation, the induced response is governed by the nonlinear response of the surrounding dielectric medium, i.e. the thin ENZ substrate layer. Thus the overall nonequilibrium dynamics are rooted in the ultrafast (sub-ps) and large nonlinear optical response of the ITO layer.

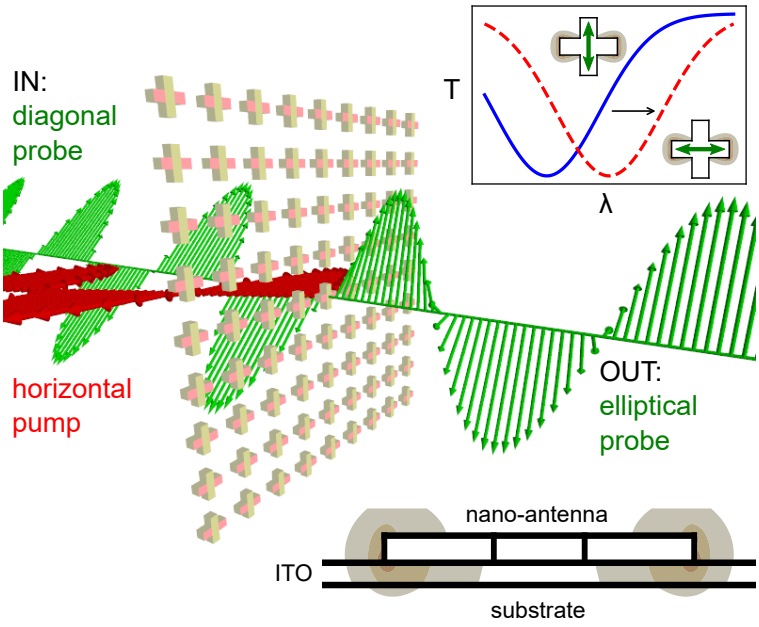


FIGURE 6.1: **Schematic of the optical excitation.** The schematic shows the pump polarisation parallel to the horizontal bar of the cross-shaped nano-antenna. The plasmonic dipole excites the ITO locally around the edges of that axis. Hence, only the parallel polarised probe (red, dashed) sees a red-shifted resonance, while the perpendicularly polarised probe (blue, solid) remains unchanged. The pump-induced modulation of the horizontal probe amplitude and phase leads to the probe transitioning from a diagonal to elliptical polarisation.

6.2 Experimental Setup

For the experimental realisation, the nano-antennas sample was produced as follows: The polarisation-insensitive plasmonic antenna array is fabricated using e-beam lithography on top of the ITO film. We choose the dimensions of the plasmonic antenna array (inset of Figure 6.2a) such that the resonance wavelength, in the absence of the ITO layer, is ~1150 nm. Figure 6.2b shows the experimentally obtained linear responses of the metasurface. The antenna array geometry is chosen to be cross-shaped such that there is no polarisation dependence in the linear response. We note that the measurements shown in Figure 6.2b are insensitive to the input light linear polarisation. In the presence of the ITO layer, the localised plasmonic mode of the antenna array interacts with the ENZ mode of the ITO layer. This nearfield mediated interaction leads to a strong coupling-induced resonance splitting of ~400 nm. This splitting is larger than the line widths of both the plasmonic resonance and the ENZ mode. The strong coupling-induced resonance splitting results in two distinct dips in the linear transmittance response of the metasurface. We observe a dominant resonance at ~1150 nm and a weaker resonance at ~1475 nm. The main dip is clearly visible in the measured linear transmission of the metasurface.

The ITO sample itself was purchased from Präzisions Glas & Optik GmbH. It has a thickness of 23 nm with Drude parameters of $\varepsilon_\infty = 3.9$ as the high-frequency permittivity, $\omega_p = 2.67 \times 10^{15}$ rad/s as the bulk plasmon frequency and $\gamma = 0.0873\omega_p$ as the damping rate (see Figure 5.1a). 1415 nm is the corresponding ENZ wavelength.

The nonlinear characterisation was performed in a transmission geometry as seen in Figure 6.2c. For the pump-probe measurements, we used an amplified Ti:sapphire

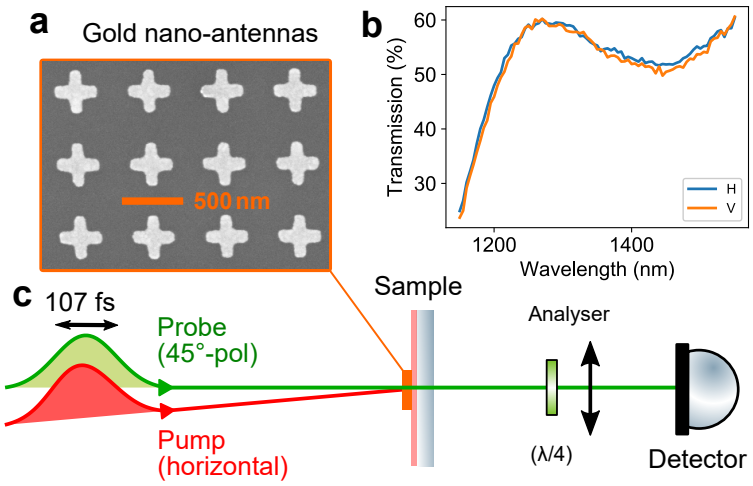


FIGURE 6.2: **Sample and setup details. a,** SEM image of a cross-shape antenna array. In this study the cross has a bar-length of l = 361 nm , a bar-width of w = 108 nm and a height of h = 32 nm. The crosses are periodically aligned with a lattice constant of p = 609 nm. **b**, The polarisation independent linear transmittance of the array (H: Horizontal, V: Vertical polarisation). **c**, To investigate the polarisation response of the antennas on a 23 nm ITO thin film, we pump at a 5° angle of incidence with the polarisation aligned along the horizontal axis. The probe comes in at a normal incidence with a 45° polarisation. The analyser allows for studying both the horizontal and vertical polarisation components. Finally, the quarter-wave plate is used to extract the Stokes parameters and plot the polarisation ellipses.

laser (Legend Elite, Coherent), with a central wavelength of 800 nm, pulse duration of 107 fs and repetition rate of 1 kHz, feeding two identical OPAs (TOPAS, Light Conversion). The signal output of one OPA was used as the pump, and the signal output of the other OPA was used as the probe, allowing us independent control of pump and probe frequencies. The pump was focused using a 30 cm BK7 lens, the probe with a 25 cm $CaFl_2$ lens. The pump beam diameter (1/e) was measured to be 800 µm, while the probe was 250 µm. We used a PbS Amplified Detector (PDA30G-EC) for detecting the probe.

The pump polarisation is aligned along the horizontal bar of the cross-shaped nano-antennas. Plasmon excitation in the nano-antennas leads to a dipole-like field profile and the resulting localised intensity pattern (see Figure 6.3). Electrons inside the ITO layer close to the edges of that bar are excited and facilitate a nonlinear refractive index change via a redshifting of the materials plasma frequency [8]. The normal incidence probe has a diagonal polarisation to investigate the pumped and unpumped nano-antenna-axis transmittance changes by choosing the analyser orientation. A rotatable quarter wave-plate in front of the analyser allows us to extract all four Stokes parameters and analyse the complete polarisation ellipses later on [77].

6.3 Transient Dichroic Response

First, we analyse the transient dichroic response by studying the nonlinear transmittance changes. We compare changes to the parallel and perpendicular probe polarisation relative to the horizontal pump. In Figure 6.4a we present transmittance changes dependent on the probe delay for several probe wavelengths. The pump

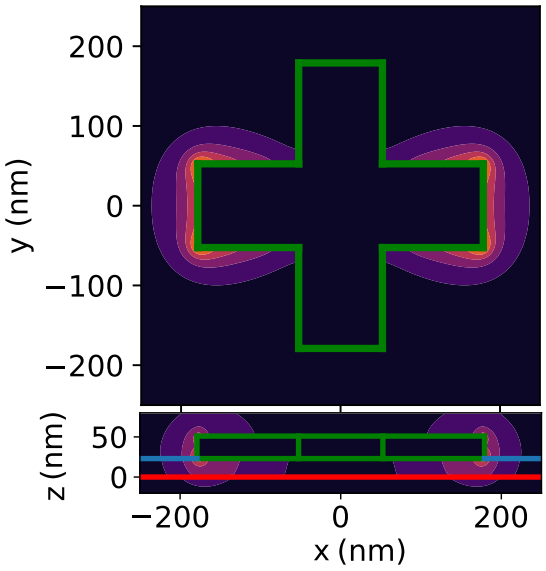


FIGURE 6.3: **Electric fields of the cross-shaped nano-antenna.** The y-x-crossection shows the nano-antenna shape highlighted in green and the electric fields surrounding the edges of the horizontal bar for a horizontally polarised pump. The z-x-crosssection highlights the penetration of the fields into the nonlinear ITO layer (between blue and red line). It further reveals that the E-fields below the central part, shared between both cross axis, are negligible.

is set to a wavelength of 1220 nm and an intensity of 5 GW cm^{-2} to avoid damaging the sample. Upon pumping, the ITO layer increases its electron temperature where the electric field strength is large (see Fig. S1). The bulk plasma frequency redshifts upon heating, leading to a shift of the coupled resonance seen in Figure 6.2b. For the parallel pump and probe polarisation, we measure large transmittance decreases by up to ~18 % for probe wavelength around 1150-1240 nm. This coincides with the plasmonic resonance at ~1150 nm red shifting upon pumping. For larger wavelengths, the sign flips, leading to small increases in transmission. This behaviour corresponds to the resonance associated with the epsilon-near-zero thin layer around 1450 nm (see Figure 6.2b), which also red-shifts, increasing the transmittance just below 1450 nm. These two scenarios can happen subsequently, as seen in the case of ~1270 nm. The sub-ps switching dynamics occur similarly for the perpendicularly polarised probe, as seen in the lower plot. However, the maximum reduction of transmission does not exceed -2.5 % compared to the -18 % of the parallel case.

To test the pump dependence, we study the frequency-resolved response of the sample. The transmittance changes plotted Figure 6.4b are measured for a pump intensity of 5 GW cm^{-2}. The pump shows a quasi wavelength-independent behaviour, which can be explained by weak wavelength dependence of the absorption. The probe transmission, however, shows a strong wavelength dependence, as already seen in Figure 6.4a. The maximum ΔT is marked just below 1200 nm, corresponding to the maximum slope position of the ~1150 nm resonances longer wavelength tail.

To test the sample's dichroic response, we compare the nonlinear transmission changes depending on the intensity for both probe polarisation cases. In Figure 6.4c, we plot the intensity-dependent $\Delta T/T_0$ for the case of 1195 nm (for further detail we also plot the spectrally resolved intensity dependencies in Figure 6.5). The intensities range up to 5 GW cm^{-2} to avoid sample damage and enable a good linear fit. The $\Delta T/T_0$ for the parallel component reaches -50 %. Comparing the two slopes (a), the

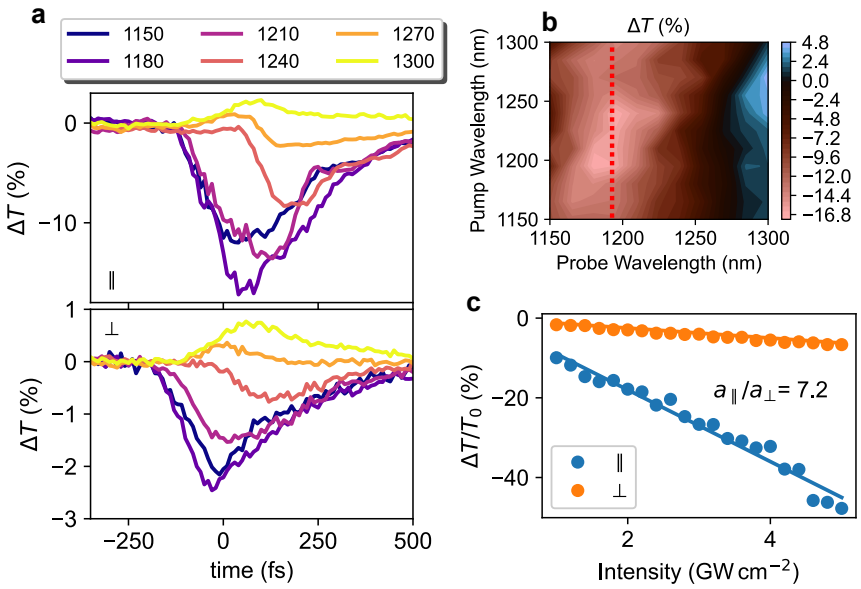


FIGURE 6.4: **Polarization dependent resonance shift. a,** The measured differential transmission (ΔT) is plotted depending on the time delay between pump and probe. The pump is fixed to a wavelength of 1220 nm and intensity of 5 GW cm^{-2}. The probe is analysed parallel ($\parallel$, top) and perpendicular ($\perp$, bottom) to the horizontal pump. **b**, The frequency-resolved transmittance changes highlight the process to be quasi-independent of the pump wavelength. The maximum differential transmission occurs for a probe wavelength marked at 1195 nm. **c**, For the case of 1195 nm we plot the intensity-dependent differential transmission and find a large maximum $\Delta T/T_0$, while suppressing the perpendicular nonlinear transmittance changes by a factor of 7.2 as indicated by the slope-ratio of the linear fits.

polarisation-dependent performance becomes very clear: The parallel contribution is modulated 7.2 times more efficiently than the perpendicular is. The local ITO excitation caused by the pump orientation translates to great differences in the nonlinear response depending on the probe polarisation. The corresponding transient dichroic ratio is calculated via $(\Delta T_\parallel - \Delta T_\perp)/|\Delta T_\parallel + \Delta T_\perp| = 76\,\%$, which is more than three times bigger than that of recent similar studies [147]. The absolute dichroic transmission difference of $\Delta T_\parallel - \Delta T_\perp \sim 16\,\%$ compares to be over an order of magnitude larger.

6.4 Transient Elliptical Polarisation

In addition to amplitude modulation, we also observe significant phase modulation. We investigate the complete polarisation by measuring the Stokes parameters (S_n). For that, we turn back to our analyser design shown in Figure 6.2c. The polariser is now fixed horizontally, and we add a quarter-wave plate. We employ the method of Schaefer et al. [78] to extract the Stokes parameters, $S_0 = I_H + I_V$, $S_1 = I_H - I_V$, $S_2 = I_{45°} - I_{135°}$ and $S_3 = I_{↻} - I_{↺}$, via an eight-point rotation of the quarter-wave plate. We reference S_1 and S_3 to the total average power $S_{ref} = \sqrt{S_1^2 + S_2^2 + S_3^2}$ before the pump arrival. The resulting S_1 plot seen in the upper case of Figure 6.6a highlights once more the previously discussed amplitude switching regime for $\lambda \sim 1200$ nm. The decrease in S_1 corresponds to a decrease in the horizontal transmission. The switch from diagonal towards vertical polarisation can best be seen by plotting the ellipses

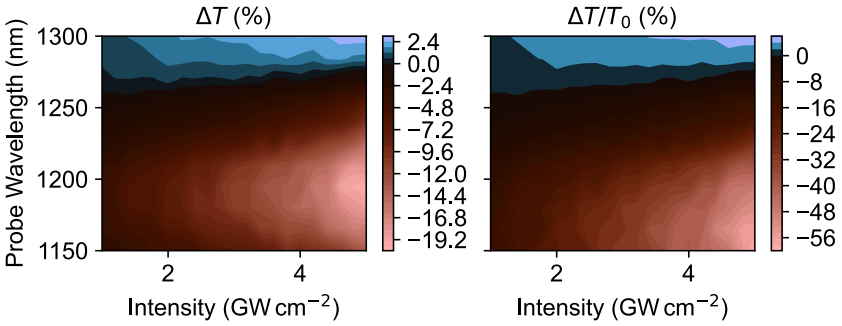


FIGURE 6.5: **Intensity dependent spectral changes of ΔT and $\Delta T/T$.** The intensity-dependent ΔT (left) and $\Delta T/T_0$ (right) for the case of a 1195 nm horizontal pump. The probe analyser is also horizontally aligned, while the probe wavelength sweeps from 1150 nm to 1300 nm. The maximum $\Delta T/T_0$ is closer to the resonance than ΔT, due to T_0 being small.

for the initial versus the pumped case plotted in the upper case of Figure 6.6b. We utilise the "py-pol" python package to plot the polarized light.

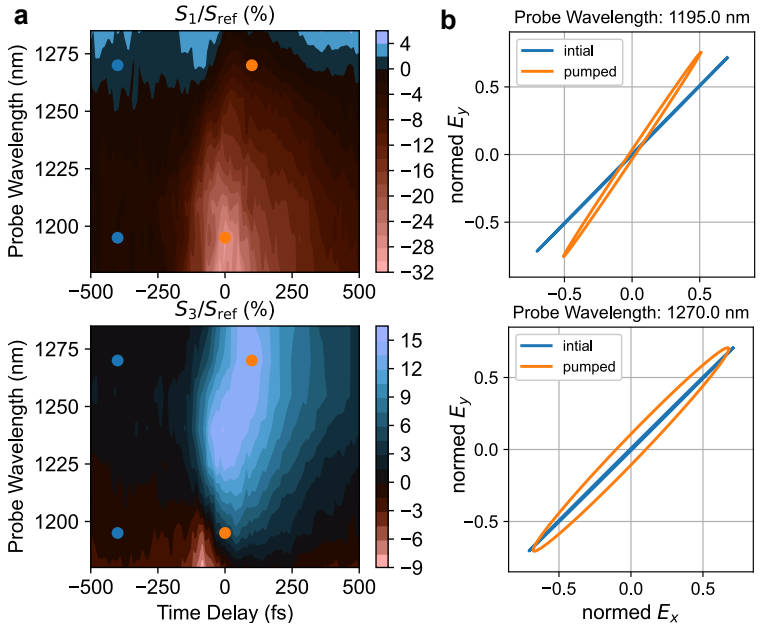


FIGURE 6.6: **Amplitude and phase control.** **a,** The time and wavelength-dependent Stokes parameters $S_1 = I_H - I_V$ and $S_3 = I_{\circlearrowright} - I_{\circlearrowleft}$ are plotted. The reference power S_{ref} corresponds to the total transmitted power S_0 in the absence of the pump. The pump is fixed to a wavelength of 1220 nm and intensity of 5 GW cm^{-2}. The markers correspond to the polarisation ellipses plotted in **b**. The initial case is representative of no pump excitation and hence a diagonal polarisation. The pumped case represents the maximum nonlinear effect, corresponding to a horizontal amplitude reduction (upper) or a horizontal phase shift (lower).

Next we investigate the time-dependent changes of the circular polarization component $S_3 = I_{\circlearrowright} - I_{\circlearrowleft}$ (Figure 6.6a, bottom). We find a change of up to 15 % around

1270 nm, corresponding to a horizontal phase change of 8.7°. Interestingly, pump-induced horizontal amplitude changes are small (S_1 constant) in this region, and instead, only the circular component is changing (S_3 increasing). This is emphasised in the bottom plot of Figure 6.6b by comparing the initial and pumped cases: the diagonal polarisation is turning elliptical without significant amplitude modulation. Finally, one can choose a probe wavelength between these two extremes to tune across a wide range of elliptical polarisation options.

6.5 Qualitative Modelling

To gain a qualitative expectation for the transient elliptical polarisation changes we investigate a simple model with the commercial software COMSOL (FEM). We define an effectively pumped ITO region around the pumped cross bar region as indicated in Figure 6.7. The region is not optimised nor calculated, but simply ensures two criteria are meet: The edges of one cross bar are covered, while the other bar is not in direct contact with the heated region. The ITO layer is modelled with exemplary optical parameters of $\varepsilon_\infty = 3.8055$, $\omega_{p,0} = 2\pi \cdot 473 \cdot 10^{12}$ rad/s and $\gamma_0 = \omega_{p,0} \cdot 0.0468$ [8], not perfectly matching the experimental case.

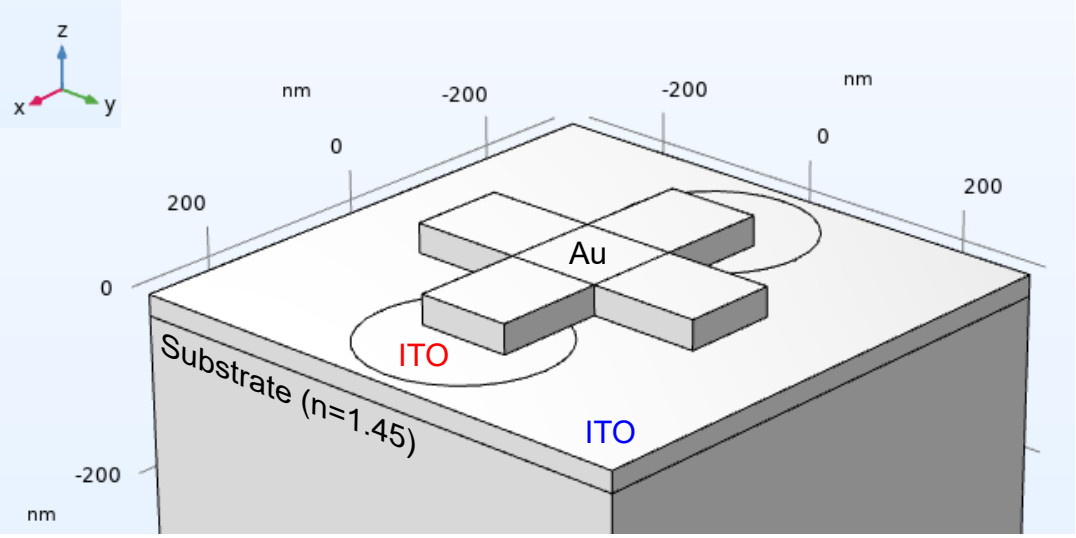


FIGURE 6.7: **Cross-shaped nano-antenna array geometry for effective bulk plasma frequency reduction simulation.** In this study the cross has a bar-length of $l = 361$ nm , a bar-width of $w = 108$ nm and a height of $h = 32$ nm. The crosses are periodically aligned with a lattice constant of $p = 609$ nm. The ITO layer is 23 nm thick and is placed on top of a coverslip ($n = 1.45$). We create indicative cylindrical "heated" ENZ patch around the pumped bar edges. The reduction of the plasma frequency in these volumes gives an impression of the expected nonlinear optical behaviour.

The effectively heated area is modelled with a 15 % reduction in bulk plasma frequency ($\omega_{p,0}$). Probing with a polarisation aligned along that bar one can see a significant redshifting of the resonant feature around 1150 nm, as seen for the red line Figure 6.8 compared to the blue line representing the initial homogeneous layer case. For a specific wavelength of interest, we add the transmitted polarization of a *diagonally polarised probe* input. These elliptical polarization plots highlight the horizontal phase shift induced for the off-resonant cases, leading to a shift towards

circular polarization. At the same time amplitude switching dominates roughly at the position of maximum transmission slope of the initial resonance position. This precisely matches our expectation and the qualitative behaviour of the two switching cases presented in the previous section.

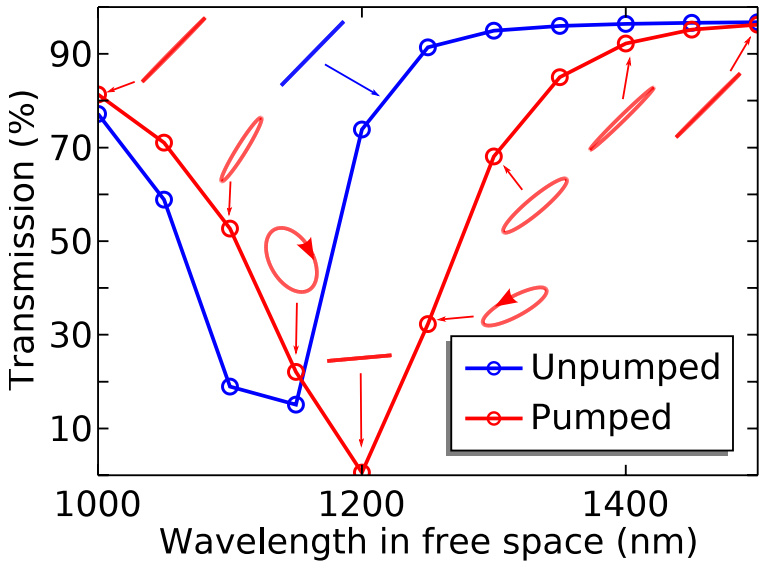


FIGURE 6.8: **Elliptical polarisation shifting in symmetric nano-antenna arrays.** The lines correspond to the probe polarization aligned along the pumped bar. The unpumped case (blue line) and "pumped" case (red line) are plotted. The insets correspond to the transmitted polarisation of a diagonally polarised input for the marked wavelength. The unpumped case remains diagonal independent of wavelength.

The simulations highlight the most promising circular polarization switching regime at the crossing point of the initial and the pumped transmission resonance. Minor modulation to the absolute transmission will be expected. However, large modulation of the phase will lead to the strong shift towards circular polarization indicated for ~1150 nm in Figure 6.8.

Unfortunately, our quarter wave-plate is only suitable for wavelength beyond 1200 nm. Also, the shifts measured in experiment are not as extreme: the crossing point occurs for only slight shift in wavelength and transmission relative to the original transmission minimum. This indicates a less pronounced nonlinear phase shift is to be expected in our measurements.

At the time of submission of this thesis, plans were being drawn up to measure the precise geometrical and (linear) optical parameters of the sample. Subsequent more elaborate simulations are expected to grant a better fit.

6.6 Conclusions

In summary, we present a compelling strategy of using cross-shaped nano-antennas on top of a highly nonlinear epsilon-near-zero layer to gain ultrafast control of the polarisation. We have shown that pumping along one nano-antenna axis allows for modulation of the amplitude or phase of the probe along that specific axis. This proof of principle study already finds a 7.2 times stronger nonlinear amplitude change for the parallel probe polarisation. This corresponds to a dichroic performance that surpasses similar studies by up to an order of magnitude. Additional changes to the

circular components of the probe have been identified and highlight the ability to control the phase as well. This is achieved for normal incidence and can be done for horizontal and vertical alignment equally and quasi-independently due to the sample symmetry. Polarisation control within the given sub-ps time scales is crucial for future ultrafast optical switching devices. Combined with the given wavelength range, we pave the way towards complete, ultrafast, and all-optical polarisation control for telecommunications applications and beyond.

Chapter 7

Spatiotemporal Refraction of Light in an Epsilon-Near-Zero ITO Layer

Abstract

When light travels through a medium in which the refractive index is rapidly changing with time, the light will undergo a shift in its frequency. Significant frequency shifting effects have recently been reported for transparent conductive oxides. These observations have been interpreted as emerging from temporal changes to the propagation phase in a bulk medium resulting from temporal variations in the refractive index, an effect referred to as temporal refraction. Here, we show that the frequency shift in an epsilon-near-zero (ENZ) layer made of indium tin oxide (ITO) originates not only from this bulk response but includes a significant effect resulting from temporal changes to the spatial boundary conditions. This boundary effect can lead to a dominant, opposing shift to the bulk effect for certain angles. Hence, this process gives rise to a frequency shift that can be tailored through the angle, decoupling the amplitude and phase modulation.

Author Contribution Statement

For this chapter, I designed the sample and experiment. I carried out the nonlinear measurements and numerical analysis. The work has been published as J. Bohn et al. "**Spatiotemporal refraction of light in an epsilon-near-zero ITO layer: frequency shifting effects arising from interfaces**," OPTICA, **8**, 1532-1537 (2021) .

7.1 Introduction

All-optical signal processing requires controlling various parameters of light waves such as amplitude, phase and frequency. Recent work has painted a promising picture for all-optical switching, showing sub-ps and large amplitude modulation for various platforms based on amorphous silicon [149], gallium phosphide [150], plasmonic waveguides [151], and epsilon near zero (ENZ) layers[65, 66, 67], including indium tin oxide (ITO) [8, 115]. Even cavity-based optical transistors have been demonstrated [152, 153]. However, while the amplitude of light can be straightforwardly controlled using these different modulator materials, controlling the frequency of light is more challenging. Control over the frequency is required for applications such as laser tuning, temporal [154] and spectral [155] pulse compression and optical switching of channels in telecoms [156]. Meanwhile, the ability to control both amplitude and frequency of a wave independently is thought to be important for generating the spatiotemporal modulation required for non-reciprocal devices [157] and time crystals [158]. To modulate the frequency of an optical signal, one can change its phase on ultrafast time scales, effectively creating a temporal refractive index boundary [159, 160, 161]. However, in many circumstances, changes to frequency are necessarily and directly coupled to changes in amplitude, making independent modulation difficult.

Recently, thin films of epsilon-near-zero (ENZ) materials have offered a promising route to frequency modulation [73, 13, 74, 75]. In such materials, significant (> unity) changes to the refractive index can be induced on sub-100 fs timescales due to ultrafast heating of the electron gas[8, 66], an effect which can result in frequency shifts by up to a few per cent of the carrier frequency[13]. However, while observed frequency shifts are large when temporal refraction occurs in these systems, they are also directly linked to refractive index changes, which lead to simultaneous amplitude modulation, preventing independent optimisation.

Here, we investigate spatiotemporal refraction for tailored frequency shifting in thin indium tin oxide layers. We show that the frequency shift arises not only from a bulk response but includes a significant contribution from temporal changes to the spatial boundary conditions. The frequency shift arising from boundary effects can oppose the bulk effect and can even be the dominant contribution for sub-μm layer thicknesses. We further show that for high incident angles, it gives rise to a dominant, opposing shift, i.e. shifting to higher rather than the usual lower frequencies, while maintaining an increase in differential transmission. This competition between surface and bulk responses could have applications where tuning the amplitude and direction of frequency shift is useful, or decoupling of amplitude and phase modulation is required.

7.2 ITO Sample

In this study, we utilise ITO samples of different thicknesses. For the 107 nm sample, the ITO was sputtered onto a coverslip at room temperature using 90/10 In_2O_3/SnO_2 Kurt Lesker target and sputtering tool. The base pressure before the deposition was in low 10^{-6} torr but raised to 3 mT of Ar only during deposition with an RF power of 145 W. In order to achieve high carrier density, both deposition and annealing were performed in the lowest possible residual oxygen environment. The samples are post-annealed in forming gas for 3 min at temperatures between 425-525°C in a rapid thermal annealer. The 407 nm sample was obtained from UQG Ltd.

The optical properties and thicknesses of the indium tin oxide (ITO) samples are characterised using an ellipsometer. The ellipsometry data was fitted using Tauc-Lorentz and Drude models. The extracted optical parameters are listed in Table 7.1, where the epsilon-near-zero (ENZ) frequency is calculated as $f_{\mathrm{ENZ}} = \frac{1}{2\pi}\sqrt{\omega_{\mathrm{p}}^2/\varepsilon_\infty - \gamma^2}$. The corresponding wavelength-dependent permittivity and refractive index are plotted in Figure 7.1.

TABLE 7.1: Drude parameters for the different ITO film thicknesses.

t_{ITO} (nm)	ε_∞	$\omega_{\mathrm{p}}(10^{15}$ rad/s)	$\omega_{\mathrm{p}}/\gamma$	f_{ENZ} (THz)
407	3.45	2.5	11.5	212
115	3.81	3.03	16.9	246

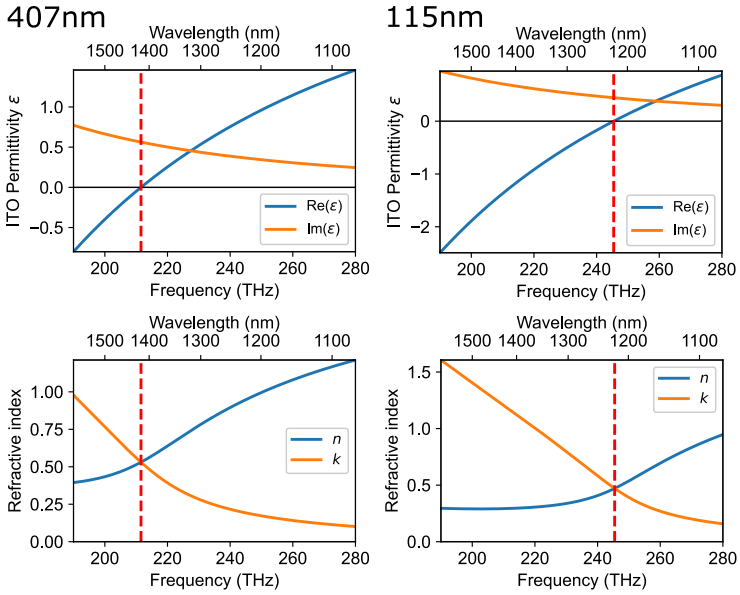


FIGURE 7.1: **Optical properties of indium thin films.** Optical permittivity (top) and refractive index (bottom) are plotted for the two different samples. The ENZ frequency is marked at 212 THz for the 407 nm sample (left), 246 THz for 115 nm (right).

7.3 ITO Layer Model

For temporal refraction, a bulk medium of homogeneous refractive index n undergoes a temporal change in the index [162]. The temporal shape of the phase change is determined by the nonlinear refractive index change $\Delta n = n_2 I$, which depends on the nonlinear material characteristics n_2 and the pump pulse intensity I. Time translation symmetry breaking leads to a change in frequency Δf, determined by $nf = (n + \Delta n)(f + \Delta f)$, predicting that the fractional frequency shift $\Delta f/f$ should depend only on the material indices through $\Delta n/(n + \Delta n)$.

However, nonlinear media can be considerably thinner than the spatial extent of a laser pulse so that only part of a pulse is present in the medium at any one moment in time. Under such conditions, the time translation symmetry argument outlined above breaks down. This is particularly true for films of ITO, which are highly absorptive close to the ENZ frequency, so transmission is very low for film

thicknesses > 1 µm. When films of several hundred nm thicknesses are excited by femtosecond pulses, the frequency shifts measured in experiment are considerably smaller than those predicted by time translation symmetry breaking [13]. Under these circumstances, one should instead turn to the more generally applicable phase modulation picture, familiar from effects such as self-phase modulation. Here, we calculate time-dependent changes to the transmitted phase of the laser pulse. Consider the change in phase for normal incidence arising from propagation through the bulk of the film. For a time-dependent change to index $\Delta n(t)$ one predicts a change in phase given by $\Delta\Phi(t) = \Delta n(t) k_0 d$ (see Figure 7.2), where k_0 is the wavenumber of the incident radiation in free space. We see from this that the thickness of the film d is expected to limit the phase change, and therefore the observable frequency shift, which is determined by

$$\Delta\omega = -\mathrm{d}\Phi/\mathrm{d}t. \tag{7.1}$$

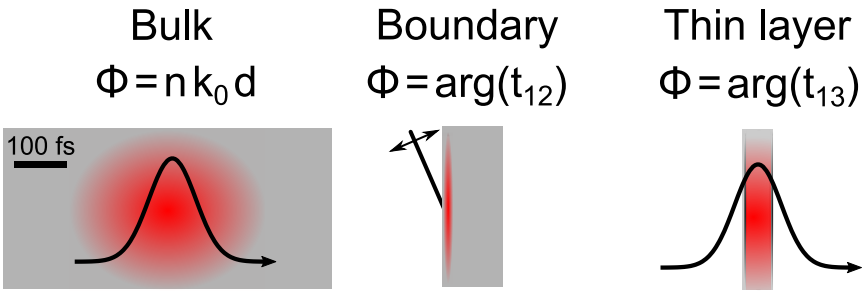


FIGURE 7.2: **Temporal changes in an optical medium.** The schematic of the bulk propagation case highlights that the nonlinear refractive index facilitates a temporally changing phase for intense laser pulses. The Fresnel coefficients can be used to extract phase changes at a spatial boundary. Temporal changes to the interface induced phase-jumps also shift the frequency. Finally, the thin layer case studied here, which combines these phenomena.

However, in thin films, interfaces also play an essential role in determining the transmission phase. Effects arising from a temporal change of the Fresnel coefficients describing interfaces have been suggested to contribute to frequency shifting [74] and are expected to be independent of the film thickness. When the complex ITO refractive index, or equivalently permittivity ε, varies with time, as in Figure 7.3a, the Fresnel coefficients for the interface of the ITO will also vary with time. This will result in time-dependent spatial refraction at the interface, referred to here as spatiotemporal refraction. The temporal change to this boundary induced phase can be associated with a frequency shift, as per Equation 7.1 above, alongside the temporal refraction induced by the bulk.

We can predict temporal changes to the ITO thin layer transmission coefficient (t_{13}) using the Airy formula, given as

$$t_{13} = \overbrace{\frac{1}{1 + r_{12} r_{23} \mathrm{e}^{2\mathrm{i}k_2 d}}}^{\text{internal reflections}} \underbrace{t_{12} t_{23}}_{\text{interface transmission}} \overbrace{\mathrm{e}^{\mathrm{i}k_2 d}}^{\text{bulk propagation}} . \tag{7.2}$$

with the Fresnel coefficients of the front (r_{12}, t_{12}) and back interface (r_{23}, t_{23}). Here, we treat the time-dependent changes to permittivity as being homogeneous throughout

the ITO layer. We will later revisit this assumption and show that spatial inhomogeneity due to the exponential decay of the pump throughout the sample leads to relatively minor quantitative differences. For non-normal incidence at angle θ, we label the angle of the wavevector inside the ITO layer as θ_2. Then, the bulk phase change (Φ_b) is determined by the multiplication of the ITO thickness (d) and the normal wavevector component inside the ITO layer ($k_2 = nk_0 \cos\theta_2$). Further phase contributions appear from the Fresnel transmission coefficients (Φ_{12}, Φ_{23}) and the internal reflections term $\Phi_{ir} = \arg(1/(1 + r_{12}r_{23}e^{2ik_2d}))$.

We consider a beam incident from air on an ITO layer followed by a coverslip. Exciting such a system near the ENZ frequency heats the electron gas, introducing strong optical nonlinearity thought to arise predominately from changes to the plasma frequency [8, 66]. To gain qualitative insights into the expected behaviour, we define temporal changes to the bulk plasma frequency as a convolution of our Gaussian pump pulse, with a pulse length of 107 fs, and an exponential decay of 300 fs, determined from optical pump-probe measurements using the analysis presented in [115]. For a sample with an epsilon-near-zero frequency at 211.5 THz, undergoing an estimated 10 % redshift of the plasma frequency ω_p, a probe frequency f_{pr} = 200 THz will experience the temporal changes to permittivity and refractive index plotted in Figure 7.3a. For such a redshifting ω_p, the real part of the refractive index n initially increases on pumping, while the imaginary part initially decreases. For the plotted example, the real part of the permittivity crosses zero shortly after pumping: this corresponds to the ENZ condition.

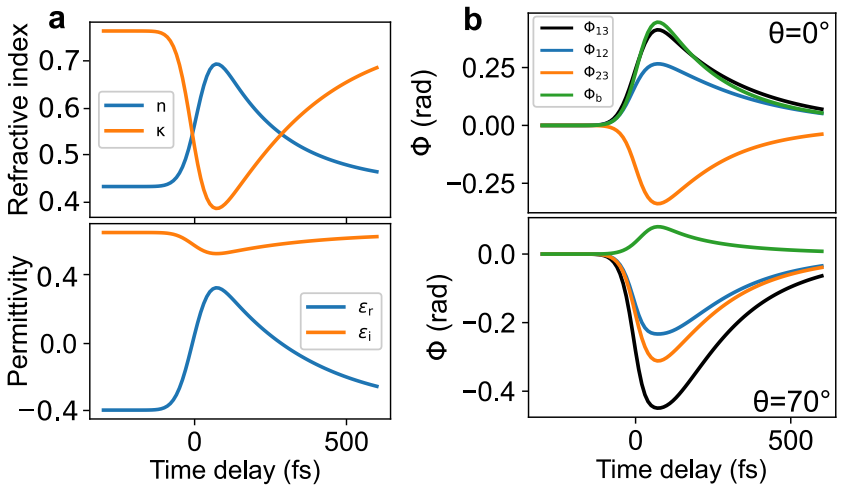


FIGURE 7.3: **Spatiotemporal refraction simulation.** We study the impact of a typical ω_p reduction by 10 % for the 407 nm ITO sample parameters at 200 THz. **a**, The upper panel shows the real part of the refractive index n initially increases until the electron gas reaches the minimum ω_p, while the imaginary part κ decreases. They cross roughly at the maximum gradient, corresponding to the ENZ point as seen below for the real part of the permittivity being zero (ε_r). The thermal decay is modelled with a time constant of 300 fs. The lower panel presents the same changes in terms of permittivity. **b**, For the case of normal incidence, the simulated thin layer transmission phase (Φ_{13}) agrees very well with the bulk phase ($\Phi_b = k_2 d$). However, for an incoming angle of 70° a total thin layer phase shift of opposite sign is expected.

Using Equation 7.2, and for a sample with $d = 407\,\text{nm}$, we can extract the transmission coefficient of our thin layer based on the time-dependent refractive index and plot the corresponding phase in Figure 7.3b. For normal incidence (upper panel), we see that the phase change due to the interfaces, Φ_{12} and Φ_{23}, are opposite in sign and approximately cancel. The resulting total time-dependent phase for this normal case, Φ_{13}, is then similar to that predicted from Φ_b, the change in phase expected due to propagation through the bulk.

However, increasing the angle of incidence leads to a very different phase response. For larger angles, the temporal gradient of the transmission phase can even change sign, as seen for the example of 70° incidence presented in the lower panel of Figure 7.3b. This behaviour arises due to the interface contributions, which no longer cancel each other. Moreover, these interface contributions also lead to a change in phase that is opposite in sign to the normal incidence case, i.e. a decrease in phase for the permittivity transitioning from negative to positive real values. This effect is fundamentally different from the bulk response, demonstrating the importance of the interfaces in determining changes in phase and frequency. Interestingly, such a contrasting behaviour of the phase results in an apparent experimental signature: a blueshift of the frequency instead of the redshift expected from a bulk.

7.3.1 In-Depth Details

Considering the multiplication of complex numbers for Equation 7.2, one can extract phase estimates of the transmitted light via the addition of phases as follows:

$$\Delta\Phi_{13} = \overbrace{\Delta\Phi_{ir} + \Delta\Phi_{12} + \Delta\Phi_{23}}^{\text{interface effects}} + \overbrace{\Delta\Phi_b}^{\text{bulk propagation}} \quad . \tag{7.3}$$

The interface contributions appear from the Fresnel transmission coefficients (Φ_{12}, Φ_{23}) and the internal reflections term $\Phi_{ir} = \arg(1/(1 + r_{12}r_{23}e^{2ik_2d}))$. Moreover, the expected frequency shifts are:

$$\Delta f_{13} = \overbrace{\Delta f_{ir} + \Delta f_{12} + \Delta f_{23}}^{\text{spatiotemporal refraction}} + \overbrace{\Delta f_b}^{\text{temporal refraction}} \quad . \tag{7.4}$$

Equation 7.4 shows that temporal refraction can be considered a special case of (cross or self) phase modulation, where phase phase modulation effects from interfaces can be neglected. This chapter investigates the case where the temporal refraction is no longer large compared to the frequency shift contributions induced by the spatial boundary. To simulate the temporal changes to the transmission amplitude and phase, we introduce a time-varying bulk plasmon frequency of up to 10 % reduction. The temporal shape is given by a convolution of a Gaussian with a full-width-half-maximum pulse length of 107 fs and a 300 fs exponential decay (see Figure 7.4a) and corresponds to the example plotted in Figure 7.3 of the manuscript. Additionally, we plot the time-dependent permittivity, highlighting the ENZ crossing. The temporal changes of the phase plotted in Figure 7.3b result in frequency shifts that are shown in Figure 7.4b.

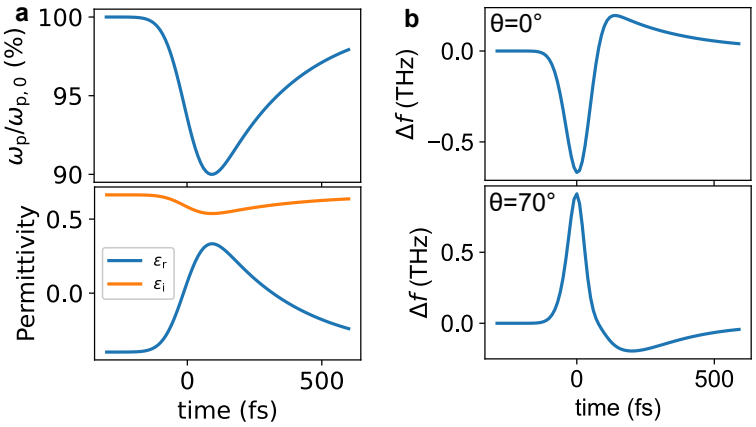


FIGURE 7.4: **Spatiotemporal refraction: frequency shifts.** We study the impact of a typical ω_p reduction by 10 %, as plotted **a**. The resulting permittivity is plotted for the case of the 407 nm ITO sample parameters at 200 THz. The ENZ case is passed at roughly the maximum slope. **b**, The expected frequency shifts for the transmitted probe are calculated using Δf_{13} and plotted for the normal and high angle incidence case.

Frequency dependence

To study the frequency shifting behaviour over large frequency and angle ranges, we exclusively investigate the maximum frequency shift (see Figure 7.5). For both the 407 nm and 115 nm sample, the blueshifted feature appears only for a large angle of incidence, as established earlier. Additionally, this blueshift occurs only for frequencies just below the (initial) ENZ case. The region of maximum blueshift sits right between the initial and pumped ENZ frequency case. This can be understood by considering the Fresnel coefficient of the front interface, which dominantly affects this feature plotted in Figure 7.4b. For the ENZ case Taylor expansion slightly below 90° the coefficient is given as ($\theta \lessapprox \pi/2$, $\theta_2 \ll \theta \to n_{\mathrm{ITO}} \ll 1$, TM-polarization):

$$t_{12}(\tau) \propto \sqrt{\frac{\varepsilon_{\mathrm{ITO}}(\tau)}{\varepsilon_{\mathrm{ITO}}(\tau)-1}} \overset{|\varepsilon_{\mathrm{ITO}}|\ll 1}{\approx} \mathrm{i}\sqrt{\varepsilon_{\mathrm{ITO}}(\tau)}, \tag{7.5}$$

with τ being the time delay between pump and probe pulse. One can see, t_{12} reduces in phase upon crossing the ENZ frequency (starting from negative ε_r values as seen in Figure 7.4a). For the TE case in the same limit, one simply calculates $t_{12,\mathrm{TE}} \propto 1/\sqrt{\varepsilon_{\mathrm{ITO}}(\tau)-1}$. Hence, the blueshift feature around the ENZ wavelength is specific to the TM polarization.

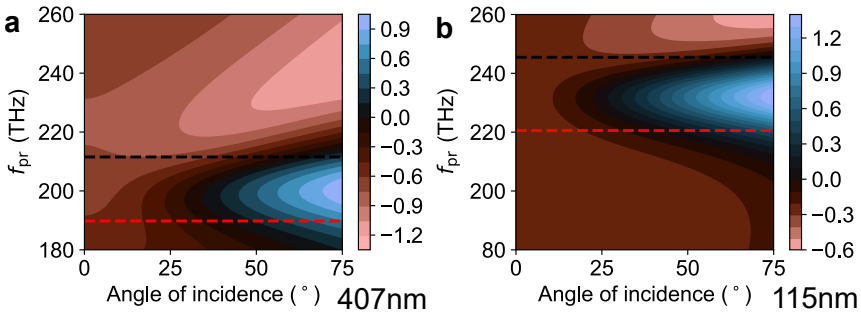


FIGURE 7.5: **Angle and frequency-dependent spatiotemporal refraction behaviour.** We plot the frequency shifts as seen in Figure 7.4b for the maximum case of $\tau = 0$ fs and study the probe frequency and angle dependence for the case of the 407 nm (**a**) and the 115 nm (**b**) parameters. The black dashed line indicates the initial f_{ENZ}, while the red dashed line indicates the pumped case.

7.4 Experimental Measurements

For the nonlinear optical characterisation via pump-probe measurements, we used an amplified Ti:sapphire laser (Legend Elite, Coherent), with a central wavelength of 800 nm, pulse duration of 107 fs and repetition rate of 1 kHz, feeding two identical OPAs (TOPAS, Light Conversion). The signal output of one OPA was used as the pump, and the signal output of the other OPA was used as the probe, allowing us independent control of pump and probe frequencies. The pump was focused using a 30 cm BK7 lens, the probe with a 25 cm $CaFl_2$ lens. The pump beam diameter (1/e) was measured to be 480 µm in air, while the probe was 250 µm. To make sure the incident intensity of the probe is significantly smaller than the pump, we used several additional OD filters to decrease the probe power and tested that the nonlinear reflection was independent of adding/removing filters. The angle of incidence of the pump is 5° smaller than that of the probe. For the spectral analysis, we used an Andor Shamrock 163 spectrograph with a DU490A-1.7 camera.

The experimental schematic is depicted in Figure 7.6a, defined by the parameters I_0 (the peak intensity of the pump pulse in the incident air), f_{pm} (the central frequency of the pump pulse) and f_{pr} (the central frequency of the probe pulse). Unless otherwise stated, we use $I_0 = 400\,\mathrm{GW\,cm^{-2}}$, resulting in frequency redshifts of ~1 THz. This intensity corresponds to the maximum intensity for the given geometry and focus size of our setup. We note that our observed frequency shifts reported below are lower than those reported in [13]. This is due to a higher Drude scattering rate and a reduced thickness for our ITO samples.

A TE pump polarisation is chosen as it provides only a slight angle and frequency dependence in absorption. We use a probe polarised at 45° to check the frequency shifting behaviour of either polarisation by rotating an analyser in front of a spectrometer. Spectra are recorded for different pump delay times. The spectra form the data basis, which we translate into changes in frequency and transmission of the central frequency of the pulse, both presented relative to the initial probe spectrum without pumping. The experimental data plotted in Figure 7.6b show typical behaviour for a probe frequency slightly below the initial ENZ frequency: a strong initial increase of the transmission as a function of time for all three angles (5°, 50°, 70°), up to 300 % of the initial transmission. Simultaneously, very different frequency shifting effects are measured: For the low angle case (5°), the typical redshifting behaviour is observed,

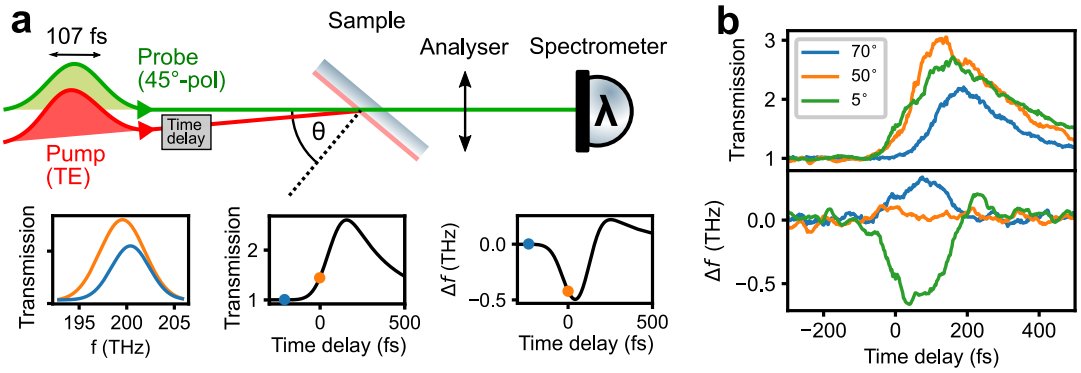


FIGURE 7.6: **Experimental setup and measurement. a**, The setup consists of a TE polarised pump and a 45° probe that enables quick spectral measurement of either polarisation by choice of the analyser. The samples of interest are ITO films of 407 nm or 115 nm thickness on top of a coverslip. The lower plots show schematically how the measured spectra (left) at different time delays (blue, orange) can be used to extract changes in both transmission (middle) and central frequency (left). **b**, Measurements taken for three different angles. They correspond to an incident pump intensity of $I_0 = 400\,\mathrm{GW\,cm^{-2}}$,$f_{\mathrm{pm}} = 250\,\mathrm{THz}$, $f_{\mathrm{pr}} = 200\,\mathrm{THz}$ and TM polarisation. The transmission increases up to 200 to 300 % for all angles, while the frequency shift is either negative (5°), negligible (50°) or positive (70°).

while the high angle case (70°) presents a blueshift - as described above, this is a signature dominated by the change in phase at the interface. This interface effect can act as an opposing shift and may be utilised to tailor the frequency shift or even suppressing it entirely, as seen for 50°.

7.4.1 Thickness Dependence

Intuitively, one would expect the effect of interfaces to be more important for thinner samples. To investigate this, we repeat the initial predictions shown in Figure 7.3 for a varying thickness of ITO. In Figure 7.7a, we plot the maximal frequency shift for the redshift feature ($\theta = 0°$) and blueshift feature ($\theta = 70°$) as a function of ITO thickness, predicted using the quasi-time harmonic model introduced below. The redshift feature is predicted to scale linearly with the thickness of the sample, which is in line with expectations that the signal is predominantly determined by the bulk phase. However, the high angle blueshift feature is dominated by surface effects and is relatively thickness independent for $d > 200\,\mathrm{nm}$. For $d < 200\,\mathrm{nm}$, internal reflections counteract the blueshift, leading to a gradual decrease of the shift as $d \to 0$ (see Figure 7.8). For very thick films with $d > 1000\,\mathrm{nm}$, the blueshift again decreases through the counteracting bulk contribution, which is comparatively small for large angles, since $\Phi_{\mathrm{b}} \propto \cos(\theta_2)$.

To demonstrate this effect experimentally, we compare the measured frequency shifts for the $d = 407\,\mathrm{nm}$ sample (Figure 7.7b) to a considerably thinner one with $d = 115\,\mathrm{nm}$ (Figure 7.7c). Figure 7.7c clearly shows a strongly reduced redshift feature, agreeing with the expectation of a reduced bulk effect. Additionally, with the decline of the redshift, the blueshift feature has become more prominent. In the experiment, we observe that the frequency shift in this region is even larger than that measured for the 407 nm thick film, an effect explained primarily by a lower scattering rate for this particular sample.

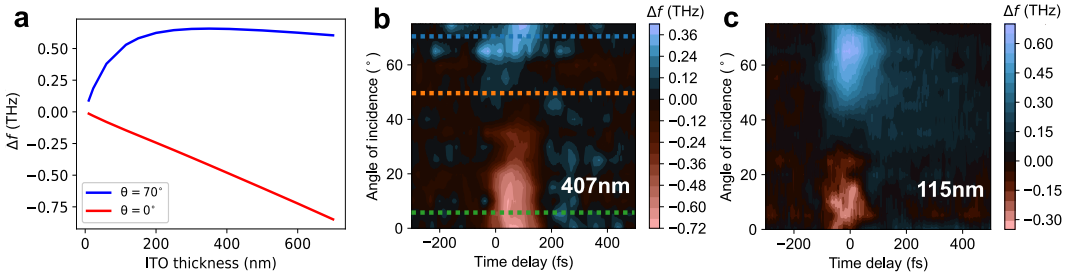


FIGURE 7.7: **Thickness dependence. a,** To further investigate the thickness dependence we model the ITO for the two angle extremes depending on the layer size. The quickly saturating blueshift (high angle) and a linearly increasing redshift (low angle) are plotted (ω_p reduction by 10 %, 417 nm sample parameters, f_{pr} = 200 THz). **b**, The experimentally measured frequency shifts of the 407 nm layer, similar to the measurements in Figure 7.6b with angles marked by dashed lines of the corresponding colours (I_0 = 400 GW cm^{-2},f_{pm} = 250 THz, f_{pr} = 200 THz, TM polarisation). **c** , The experimentally measured frequency shift of a 115 nm layer (I_0 = 400 GW cm^{-2}, f_{pm} = 214 THz, f_{pr} = 240 THz, TM polarisation).

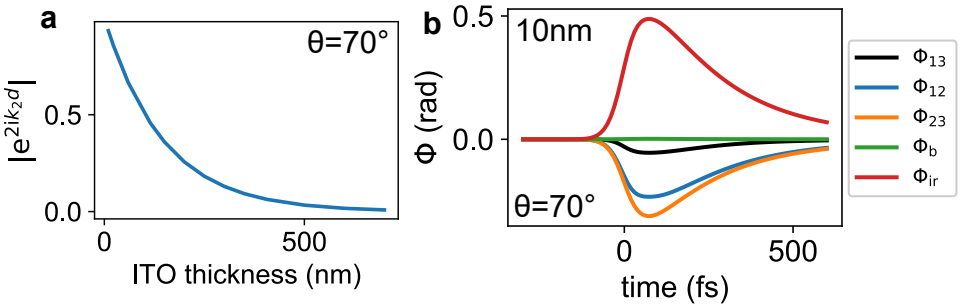


FIGURE 7.8: **Thickness dependent phase components. a,** We plot the absolute value of the internal reflection transmission coefficient term, showing that it only becomes comparable to 1, and therefore relevant, for $d \lesssim 200$ nm. **b**, For the very thin case of 10 nm ($e^{2ik_2d} \sim 1$) the internal reflection based phase shifts becomes large and compensates transmission-based effects.

Blueshifted spectrum

We present the example of the strongest measured blueshift feature, given by the 115 nm film as seen in the spectrum of Figure 7.9 and corresponding to the maximum case plotted in Figure 7.7c.

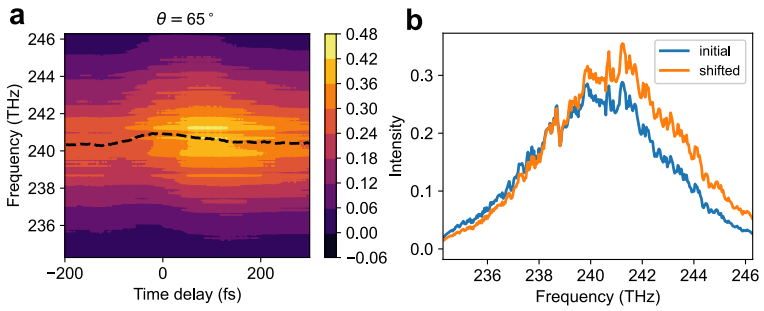


FIGURE 7.9: **Blueshifted spectrum. a**, For the 115 nm sample, the time-dependent spectrum was measured. The incident pump intensity is 400 GW cm^{-}2. The dashed line indicates the traced central frequency position. **b**, The pump free initial spectrum (blue) is plotted along with the maximum shifted case (orange).

7.5 Comparison Between Experiment and Models

One can calculate the expected frequency shift by relating measured transmission changes to expected temporal changes in phase, which can then be used to predict expected frequency shifts. The simplest possible way to achieve this is using a quasi-time harmonic model. We first calculate expected temporal changes to the permittivity of the material, relate this to expected temporal changes to the transmission phase, and directly predict from these expected changes to the frequencies contained in a probe pulse using Equation 7.1. This is possible via the following steps: firstly, we use a standard pump-probe measurement to determine the temporal dynamics of transmission amplitude vs time and fit it with the convolution of a Gaussian and exponential decay. We then calculate t_{13} for a given incident frequency f and angle θ via Equation 7.2 for different bulk plasma frequencies. The results from this calculation are used as a spline function to convert the measured change in the amplitude of transmission into corresponding changes of the bulk plasma frequency. We then convert from plasma frequency vs time to transmission phase vs time, again using Equation 7.2. Finally, we calculate the temporal derivative of the extracted phase and use this to estimate the expected frequency shift as per Equation 7.1. Note that this model assumes all frequencies in the pulse respond identically. Hence we refer to it as quasi-time harmonic - we come back to this approximation later.

The frequency shifts predicted by this quasi-time harmonic model are shown in Figure 7.10a for the case of the 407 nm sample. The qualitative behaviour predicted in Figure 7.10a, defined by the relative amplitudes of red- and blue- frequency shifts and the angle at which the frequency shift changes sign, is in good agreement with our direct measurement in Figure 7.7b. This is striking, especially considering there are no fitting parameters used in our approach. For this simple quasi-time harmonic model, we expect the most significant frequency shifts to occur at times corresponding to the maximal change in the optical properties of the material. As shown in figure Figure 7.10c (blue line), we see the maximum in predicted frequency shift corresponds to the maximum temporal gradient in the index (blue arrow in Figure 7.10c). This can also be seen in Figure 7.6b, where the most significant frequency shifts are measured for the largest slope of the transmission change. Similarly, we expect zero frequency change for the times corresponding to zero gradients in the index, as is seen for the time corresponding to the extremum in plasma frequency (red arrow). These two behaviours explain the narrow temporal widths of the frequency shift features seen in both model (Figure 7.10a) and experiment (Figure 7.7b).

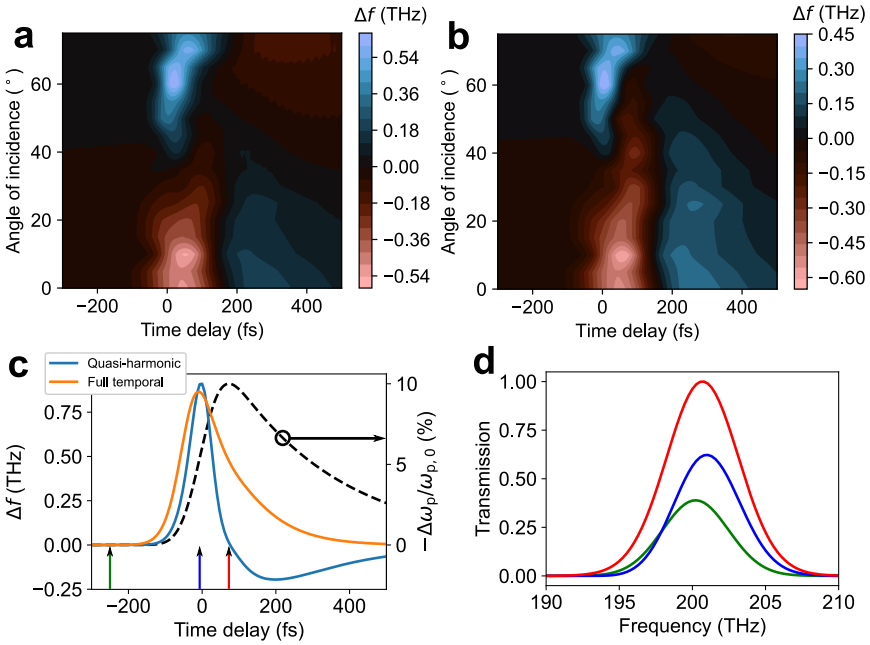


FIGURE 7.10: **Modelling of experimental data.** Predicted frequency shifts under the quasi-time harmonic approximation assuming homogeneous (**a**) and in-homogeneous (**b**) layer approximations, calculated assuming parameters: (I_0 = 400 GW cm^{-2},f_{pm} = 250 THz, f_{pr} = 200 THz, TM polarisation). **c**, A comparison between the quasi-time harmonic (blue solid) and the full temporal (orange solid) model. The temporal response of the optical material assumes a reduction in plasma frequency of 10 % (dashed black). The full temporal model considers the spectral width of 5.3 THz full width half maximum. The time-dependent absorption changes cause an effective redshift of the frequency distribution, removing the blueshift feature seen for the quasi-harmonic case. The times marked by arrows correspond to: before interaction (green), the maximum gradient in the index (blue) and zero gradients in the index (red). **d**, The frequency distributions corresponding to the three marked times in **c**, as predicted by the full temporal model. Modelled frequency shifts are lower than those reported in [13] due to a higher Drude scattering rate and a reduced thickness in our samples.

In the next section we will look a bit closer into the thickness dependent refractive index model before turning to the time-dependent model and an in-depth discussion of Figure 7.6c and d.

Thickness dependent refractive index

We note that there are some approximations used in our quasi-time harmonic modelling approach. Firstly, the finite penetration depths of the laser pulses results in an inhomogeneous refractive index of the ITO throughout the thickness of the film. We can take this effect into account by treating the ITO as a multi-layer structure, with the index change in each layer determined by the predicted pump intensity in that layer.

To test the model, we put the implicit assumption under test that the transmission changes and phase changes have the same origin, a redshift in ω_p. The transmission changes are linked to the redshifting of ω_p. These can be used to calculate all-optical

parameter changes of interest. In Figure 7.10a and Figure 7.11a, we show the case of using a uniform ITO layer excitation and find that the blueshift seems to be a bit overestimated. For a first improvement, we assume a spatial dependence of the ω_p in ITO. This results in a graded refractive index (GRIN) profile. We use the transfer matrix method to calculate the depth-dependent absorption of the pump inside the ITO layer as seen in Figure 7.11b and scale the distribution. The front will be most affected by pumping, and t_{12} driven effects may saturate sooner, while the overall blueshift may be smaller as the angle independent t_{23} contribution (see Figure 7.8a) becomes smaller. This can be seen in Figure 7.10a or Figure 7.11c of our thickness-dependent model application, which behaves qualitatively very similar to the uniform layer plotted in a, but has a significantly smaller blueshift. This also corresponds better to what is measured in experiment. While this depth-dependent plasma frequency should still be seen as a rough assumption, it appears to be noticeably more accurate than simply taking a uniform layer.

We also present the extracted ω_p values averaged over the layer thickness in Figure 7.11d. The modulation of up to ~10 % are in good agreement with similar studies [115]. Also, the increased angle leads to a reduced incident intensity due to a stretched out elliptical beam-sample cross-section area. Hence, a minor bulk plasma modulation is seen for higher angles.

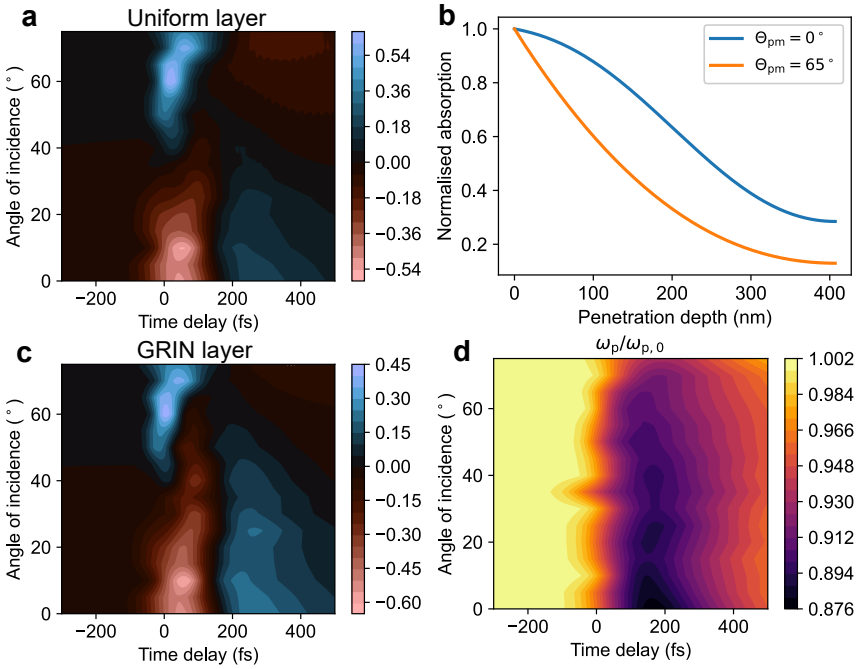


FIGURE 7.11: **Uniform vs GRIN layer.** **a,** The plotted frequency shift estimates are based on the time-dependent transmission phase, calculated via the uniform layer assumption (I_0 = 400 GW cm^{-2},f_{pm} = 250 THz, f_{pr} = 200 THz, TM analyser). **b**, The normalised depth-dependent absorption of the TE polarised pump for the low and high angle case plotted in Figure 7.6b. **c**, The GRIN based frequency shift estimates, utilising the depth-dependent plasma frequency. **d**, The plasma frequencies plotted correspond to the transmission changes measured in the experiment Figure 7.7b.

The resulting prediction (see Figure 7.10b), while qualitatively very similar, predicts a smaller blueshift and larger redshift than the single-layer model presented in Figure 7.10a, in better agreement with the measurement in Figure 7.7b.

Calculation of transmitted pulse spectrum

For changes in refractive index that occur on the timescale of the oscillating field, one should also convolve the probe field with the time-dependent response of the medium. This approach describes the full temporal dynamics of the interaction, including the broadband nature of the probe pulse. However, it is challenging to solve Maxwell's equations in a time-dependent dispersive medium, even a homogeneous one. To obtain an approximate form of the transmitted pulse in our system, we reinterpret our incident pulse $H_i(t)$ as a sum of delta function pulses each of amplitude $H_i(\tau)$, arriving at time τ, i.e. (similar to the dynamic model in section 5.3)

$$H_i(t) = \int_{-\infty}^{\infty} d\tau \delta(t-\tau) H_i(\tau). \tag{7.6}$$

the spectral decomposition of which is

$$\tilde{H}_i(\omega) = \int_{-\infty}^{\infty} d\tau e^{i\omega\tau} H_i(\tau) \tag{7.7}$$

While the delta pulse of amplitude $H_i(\tau)$ propagates through the structure, the frequency-dependent transmission coefficient is taken to have the approximately constant value $t(\tau,\omega)$. This assumes that the permittivity remains approximately constant during the ring-down time of the structure. Using (7.7), the spectrum of the transmitted pulse can then be written as

$$\tilde{H}_t(\omega) = \int_{-\infty}^{\infty} d\tau\, t(\tau,\omega)\, e^{i\omega\tau} H_i(\tau). \tag{7.8}$$

Using the inverse Fourier transform (7.7) we can then rewrite the transmitted spectrum (7.7) in terms of the incident one, rather than the incident amplitude $H(\tau)$

$$\tilde{H}_t(\omega) = \int_{-\infty}^{\infty} \frac{d\omega'}{2\pi}\, t(\omega-\omega',\omega)\, \tilde{H}_i(\omega'), \tag{7.9}$$

where

$$t(\omega-\omega',\omega) = \int_{-\infty}^{\infty} t(\tau,\omega) e^{i(\omega-\omega')\tau} \tag{7.10}$$

We numerically implemented (7.9) to calculate the transmitted spectra of our pulses, using a Fast Fourier transform to calculate $t(\omega-\omega',\omega)$, then treating the convolution (7.8) discretely as a matrix multiplication, for a fixed frequency resolution $\Delta\omega$.

As shown in Figure 7.10c, this full temporal model (orange line) shows only a slight difference in terms of the maximum predicted frequency shift. However, the response is slightly broadened in time than that predicted by the quasi-time harmonic response, curtailing the negative frequency shift seen previously for times beyond 200 fs, again more similar to the observation in experiment in Figure 7.7b. This can be understood to arise from two effects: Firstly, the temporal width of the probe pulse spreads the frequency shifting event in time. Second, the Gaussian frequency distribution of the probe pulse is reshaped not only by spatiotemporal refraction but also by the temporally shifted spectral absorption peak of the ITO layer. This last point is often ignored but very important for absorbing materials, where it can be difficult to separate changes in the absorption of the material from effects arising from the temporal gradient in the index. Further, the reshaping explains why we only measure a blueshift in experiment, but not a red-shift. Additionally, in Figure 7.10d

we see that the spectrum corresponding to the maximal frequency shift (blue line) does not correspond to the maximum change in transmission intensity (red line). In such circumstances where the frequency shifts are small, the two main signatures discussed above (i.e. the observation of the maximum shift corresponding to the maximum gradient in the index and the change in sign of frequency shift with angle) are clear indications that can help distinguish spatiotemporal refraction.

Finally, we point out that a complete model that simultaneously takes into account both the convolution between the temporal dynamics and the field, as well as the complete spatial dependence of the refractive index, is a complex computational task and lies beyond the scope of this project.

7.6 Additional Checks

Here, we present further experimental data to support the previously discussed scenarios.

7.6.1 Wavelength Dependence

The ENZ wavelength of the 407 nm sample is 1417 nm. We show in Figure 7.12 that the blueshift does not appear above the ENZ case. Additionally, we show that it is reserved for TM polarisation as predicted in the discussion surrounding Equation 7.5.

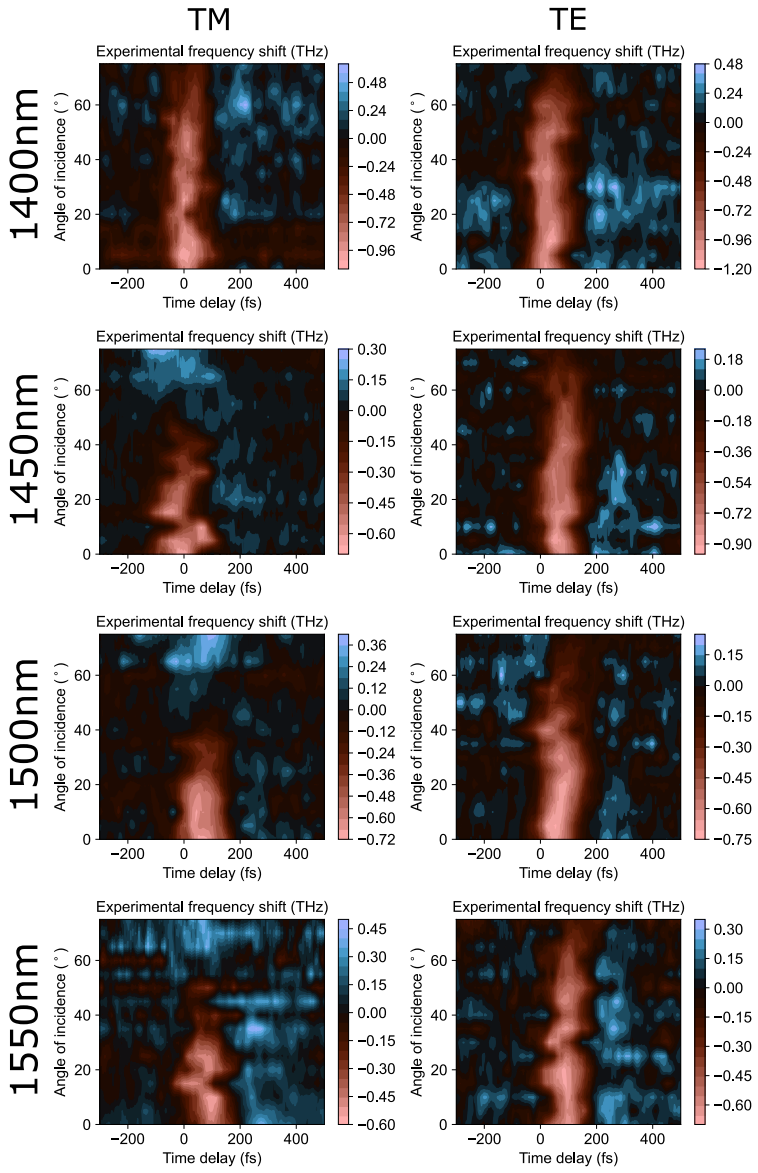


FIGURE 7.12: **Wavelength and polarisation dependence.** The measured frequency shifts show the wavelength dependence. Additionally, the TE case is measured by rotating the analyser. (I_0 = 400 GW cm^{-2}, f_{pm} = 250 THz).

7.6.2 Intensity Dependence

The intensity-dependent measurements are done for 100 to 400 GW cm^{-}2 for both samples and plotted in Figure 7.13. The redshift feature appears to saturate slightly faster for the 115 nm sample than for the 407 nm. The blueshift feature appears to start saturating for both samples roughly at 200 GW cm^{-}2 incident intensity.

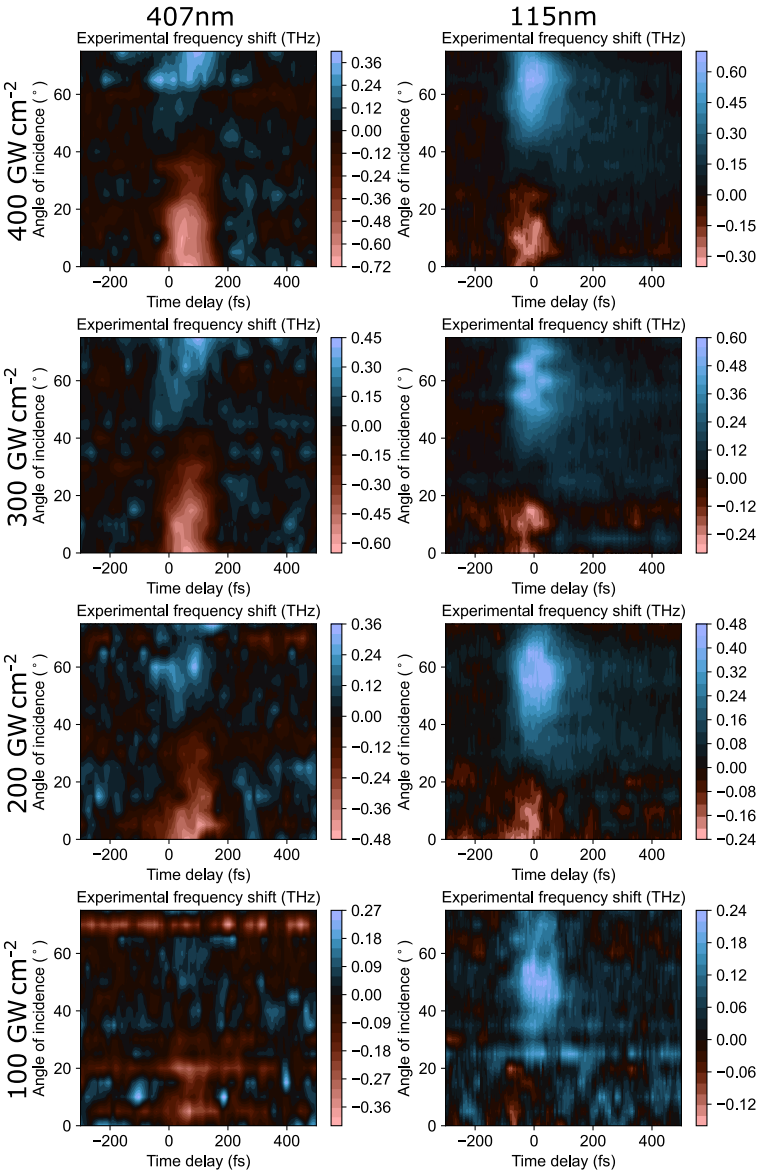


FIGURE 7.13: **Intensity dependence.** The measured frequency shifts show the incident intensity dependence of the TM contribution. This was done for both samples: 407 nm (f_{pm} = 250 THz, f_{pr} = 200 THz) and 115 nm (f_{pm} = 214 THz, f_{pr} = 240 THz).

7.6.3 Polarisation Dependence

We repeat the measurements of Figure 7.7b,c for the TE polarisation and confirm the absence of the blueshift feature as seen in Figure 7.14.

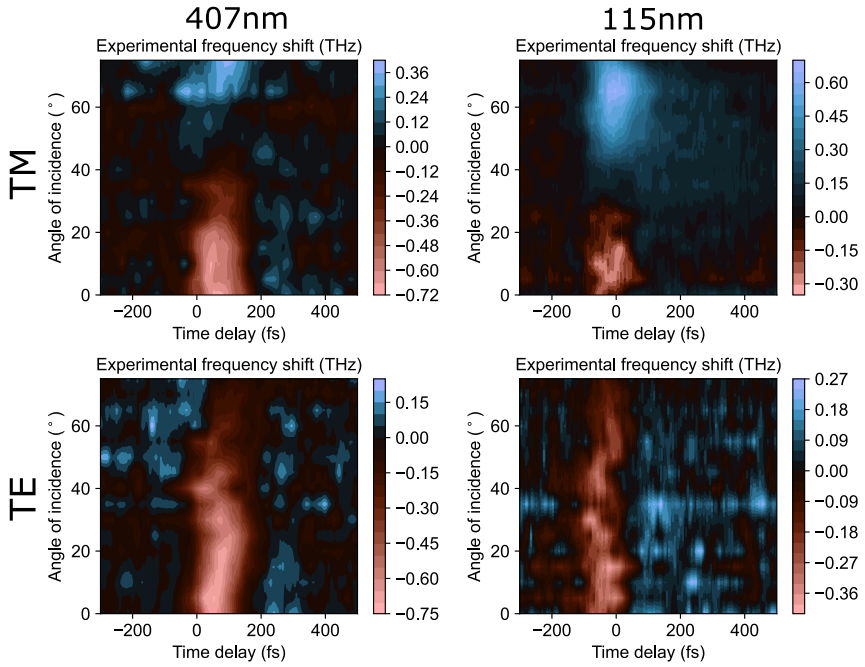


FIGURE 7.14: **Polarization dependence for both thicknesses.** The experimentally measured frequency shifts of the 407 nm layer, equivalent to the measurements in Figure 7.7b in the TM case (I_0 = 400 GW cm^{-2},f_{pm} = 250 THz, f_{pr} = 200 THz, TM polarisation). The experimentally measured frequency shift of a 115 nm layer, equivalent to the measurements in Figure 7.7c in the TM case (I_0 = 400 GW cm^{-2}, f_{pm} = 214 THz, f_{pr} = 240 THz, TM polarisation).

7.7 Conclusions

In conclusion, we show that spatiotemporal refraction provides a frequency shift that is relatively strong compared to the temporal refraction for sub-µm samples. This contribution remains roughly constant down to film thicknesses of ~100 nm. Compared to similar ITO studies [13] with 15 THz shifts, the 1 THz shift may appear small, but is based on surface effects and enables the use of much thinner samples. Most importantly, we demonstrate that the frequency shift can now be controlled not only by the temporal refractive index changes but also by the angle. This enables tailoring the nonlinear frequency shift in thin layers independent of other nonlinear modulations such as the transmission. In future, more versatile layer choices or combining multiple layers could provide a plethora of ultrafast amplitude and frequency switching devices by tuning interface effects to obtain the desired results. These effects are expected similarly for various other TCOs. For example, slightly weaker for AZO [73, 74] or at different wavelength and possibly stronger for CdO [67, 72].

Chapter 8

Summary

This PhD thesis involved various strategies for enhancing different nonlinear optical processes in thin materials. For that, we utilised two material platforms, graphene and ITO.

In chapter 4, we investigated the potential of generating new frequencies with graphene as a one-atom-thick active medium. We utilised the electric field confinement effects of gold nano-ribbons acting as an antenna close to a graphene film for enhancing the incoming electric field. The generated fifth harmonic signal is more than three orders of magnitude stronger than bare graphene's response.

We also studied the dimensional dependencies of our graphene-insulator-metal heterostructure. The thickness of the insulator spacer was varied by either utilising a 5 nm Al_2O_3 layer or a monolayer h-BN. We found that the thinner, monolayer h-BN sample provides a THG and FHG response that is both more than an order of magnitude stronger than their 5 nm Al_2O_3 counterpart.

The antenna geometry dependence was investigated by varying the ribbon width. The optimum enhancement reached for THG corresponds to an enhancement of 1600 relative to bare graphene for the antenna width of 200 nm. The FHG reached enhancement factors of 4400. However, these occurred for 100 nm antenna width, highlighting that the optimum antenna design changes for different high-harmonic orders.

While graphene-based devices make for incredibly thin frequency mixing devices, we studied epsilon-near-zero materials next. We investigated the ultrafast nonlinear modulation of light parameters such as absorption, reflection, transmission and the corresponding phase and frequency shifts.

In chapter 5, we utilised the near-perfect absorption of an epsilon-near-zero plasmon in a 60 nm ITO film to induce almost two orders of magnitude reflection changes in the telecom wavelength range. This switching performance stands in stark contrast to the original layered ITO studies. The latter may also show notable absolute changes, but the resonant character of the ENZ plasmon presented here enabled a much deeper initial reflection of just ~1 %.

Additionally, a thorough experimental study of the frequency dependence for both polarisations has highlighted a novel contribution to the nonlinear signal. An increased reflection is measured for the coherent case of the pump and probe sharing the same polarisation and frequency. The interference of both beams leads to a static intensity pattern along the thin film. The resulting spatially dependent refractive index pattern resulted in efficient scattering of the pump beam into the probe reflected beam direction, leading to the experimentally measured increase in signal. This effect is known as two-beam coupling and has been identified for epsilon-near-zero materials simultaneously and independently by J. Paul et al. [163].

A typical issue of (surface) plasmon based devices is the reliance on TM-polarised light. In chapter 6, we investigated the possibility of achieving a more holistic control

of the polarisation. A dichroic response is designed via a cross-shaped antenna on top of a thin ITO film. Aligning the pump polarisation along one nano-antenna bar led to nonlinear changes predominantly for a probe polarisation aligned parallel to the same bar. The electric field confinement effects to the antenna edge induced a more than seven times stronger nonlinear response for a parallel probe compared to the perpendicular case.

We added a Stokes analyser in the setup to measure the full, elliptical polarisation of light. It enabled the extraction of all four Stokes parameters depending on the pump-probe delay. While confirming the amplitude switching, we also identify a second regime of phase switching. This occurs for a slightly longer wavelength, further away from the resonance. In the previously discussed amplitude switching case, the original diagonal polarisation becomes more vertical due to the suppression of the horizontal transmission. In this phase switching regime, the pump does not affect the probe amplitude but modulates the horizontal phase exclusively, leading to a switch towards elliptical polarisation.

While significant changes in amplitude are very desirable, additional optical parameters need to be considered. The previously mentioned ultrafast changes in phase, for example, also lead to shifts in instantaneous frequency. In chapter 7, we studied the ultrafast phase and frequency shifting dynamics of light in ITO layers of varying thickness and angle of incidence.

For the case of a 407 nm ITO layer, we found a redshift of ~1 THz for low angle of incidence. This feature aligns with the standard description of frequency shifting through ultrafast fast changes to the propagation phase inside a nonlinear medium. However, for a high angle of incidence, a blueshift is measured upon pumping through the ENZ case, counteracting the propagation phase.

This surface-based blueshift also referred to as spatiotemporal refraction, is dominant for sub-1 µm layers and quasi thickness independent down to ~100 nm. In thinner cases, multiple internal reflections occur, and the frequency shift vanishes.

Discussion

For nonlinear optics in graphene, the ultimate goal would be combining the graphene plasmon resonance and electric field confinements to a single atom thin vertical resolution as suggested in [43]. This limit has been challenging to utilise for high-harmonic generation, as presented in chapter 4, due to experimental difficulties. These GPP resonances typically lie in the MIR around 6 µm or longer. Both light sources and detectors are less efficient and more expensive for this region of interest. Experimentally, we could not trace high harmonic signals for fundamental wavelength beyond 5 µm, yet. Most significantly due to reduced laser fluencies for higher wavelength.

Another exciting aspect of graphene is the difference in origin of the nonlinear signals. For short-wavelength, the graphene response follows the photo-excited carrier refraction [15]. For long-wavelength, high-harmonics are generated by thermal currents [61]. The local dependence of the electron temperature and the resulting dynamics have shown huge high-harmonic generation efficiencies beyond classical expectations. We believe this thermal nonlinearity plays a significant role in nonlinear processes even beyond HHG, e.g. in the all-optical excitation of graphene surface plasmons [45, 58] (see also Appendix A). Even though the incoming light beams are in the visible wavelength range, their interference pattern moves slowly relative to the thermal decay of the electron distribution and enables difference frequency and phase matching with the graphene surface plasmons.

For ITO and TCOs in general, we have presented an extensive toolkit for ultrafast control of optical properties and advancing the status quo in ultrafast amplitude, phase and frequency switching devices.

Via switching an ENZ plasmon resonance as studied in chapter 5 we decreased the thin-layer size and increased the amplitude switching response. To reach similar performances, typical studies in the ENZ nonlinear optics community either use relatively thick few hundred nm thin films [65, 8] or sub-30 nm films that rely on nano-antennas to couple with the ENZ mode.

In chapter 6, we used such cross-shaped nano-antenna on top of a 23 nm ITO film to generate a transient dichroic response. This enables two well-separated polarisation channels for optical switching. Even more, we measure both amplitude and phase switching, enabling the conversion of diagonally polarised light towards an elliptical or vertical polarisation. This design could pave the way towards ultrafast complex polarisation modulation.

Finally, we identified a new frequency shifting effect that origins from ultrafast phase changes at a spatial refractive index boundary. This spatiotemporal refraction provides a thickness independent phase modulation and frequency shifting contribution. The well-known bulk propagation contribution, also known as temporal refraction, is directly proportional to the refractive index. This new interface component offers a tailoring option via the angle without changing the nonlinear refractive index. This is crucial for independent optimisation of the nonlinear amplitude and phase changes. Additionally, the thickness independence enables the use of thinner layers, thereby avoiding large absorption close to the ENZ region.

While all these effects paint a promising picture for ENZ nonlinear optical devices, the progress in the field is still hindered by some obstacles that need to be addressed. Most notably, the ENZ case should not only provide $\mathrm{Re}(\varepsilon) = 0$, but $|\varepsilon| = 0$. Hence, the losses of ITO and similar materials severely hinder nonlinear enhancement, as already seen, e.g. in chapter 7 the increased frequency shift for a reduced scattering rate. More technical problems also include easy implementation with established CMOS devices and precise reproducibility in sample manufacturing. With the constant advances in fabrication capabilities, future ENZ materials will present an exciting platform to demonstrate novel nonlinear effects and show promise for implementation into industrial applications such as lasers or novel ultrafast optical switching devices.

Outlook

To directly follow up on my work we will investigate more carefully the linear optical and structural characteristics of both nanostructured samples. With that, we will be able to make more qualified statements regarding the suggested quenching process of fifth harmonic generation in our graphene device studied in chapter 4. For the plasmonic nano-antennas on ENZ thin films (see chapter 6), want to use those characteristics to model our samples better. Not only to describe the current measurements but also to optimize future devices.

In general, more complex modelling methods than the TMM method often used here will need to be utilized in the future. Methods such as finite element modelling (FEM), as seen for very rough estimates in subsection 5.7.1 or section 6.5, should be used to model the propagation of pulses in space- and time-varying media. Such methods would allow for an easy extension to structured or antenna enhanced models, not able to be simulated with the layer-based TMM method.

Beyond that, various ideas come to mind following the research direction presented here. From an experimental point of view, my setup's further advancement

to fully automatically measure all polarisation combinations desired would be a thrilling addition. The first step towards this goal was done by implementing an automated Stokes analyser to analyse the whole polarisation ellipses, rather than just a conventional linear analyser.

While we already achieved optical switching by almost two orders of magnitude using an ENZ plasmon resonance as studied in chapter 5, we believe this performance could be even better by optimising the material choice and thickness. A higher reflection for the "on"-state and a lower reflection for the "off"-state would be feasible similar to CdO studies [67] and more.

The polarisation switching with the cross-shaped antenna could be optimised with more complex geometries. Adding a cavity enhancement via replacing the dielectric substrate with a mirror could be an obvious extension. Clearly, structuring the ENZ material itself would hold great promise for diverse new research routes to follow. However, fabrication has proven to be difficult.

Similar difficulties persist for multilayer designs. We focused on studying basic phenomena with one active nonlinear medium layer. Combining our two material categories may present a promising route to follow once feasible from a manufacturing and characterisation point of view. With the strong nonlinear response of long-wavelength ENZ materials such as CdO, a combination with graphene surface plasmons or similar 2D effects could provide promising pathways for even more pronounced effects. These may even be tunable via the electrical doping of graphene.

Appendix A

Paper: Origins of All-Optical Generation of Plasmons in Graphene

Author Contribution statement

In the context of this paper I measured four wave mixing and pump-probe signals of graphene. Calculations and discussions around the measurements adjusted the focus of the paper, which was later published in Scientific Reports [58].

SCIENTIFIC REPORTS

OPEN

Origins of All-Optical Generation of Plasmons in Graphene

C. J. Tollerton[1], J. Bohn[1], T. J. Constant[1], S. A. R. Horsley[1], D. E. Chang[2,3], E. Hendry[1] & D. Z. Li[2]

Received: 5 September 2018
Accepted: 4 February 2019
Published online: 01 March 2019

Graphene, despite its centrosymmetric structure, is predicted to have a substantial second order nonlinearity, arising from non-local effects. However, there is disagreement between several published theories and experimental data. Here we derive an expression for the second order conductivity of graphene in the non-local regime using perturbation theory, concentrating on the difference frequency mixing process, and compare our results with those already published. We find a second-order conductivity ($\sigma^{(2)} \approx 10^{-17}$ AmV^{-2}) that is approximately three orders of magnitude less than that estimated from recent experimental results. This indicates that nonlinear optical coupling to plasmons in graphene cannot be described perturbatively through the electronic nonlinearity, as previously thought. We also show that this discrepancy cannot be attributed to the bulk optical nonlinearity of the substrate. As a possible alternative, we present a simple theoretical model of how a non-linearity can arise from photothermal effects, which generates a field at least two orders of magnitude larger than that found from perturbation theory.

Graphene, with its linear dispersion and a linear density of states for electrons[1], exhibits remarkable optical properties such as universal, linear optical conductivity[2]. Moreover, while a single layer of graphene is relatively transparent (due to its mono-layer thickness), the nonlinear optical conductivity has been shown to be surprisingly large[3,4]. This is particularly true for the second order nonlinearity, which is unexpected within the dipole approximation for a centrosymmetric material[5], but can be substantial in graphene due to non-locality[6–11].

The 2D nature of graphene also gives rise to plasmons with wavelengths that are substantially smaller than free-space electromagnetic radiation of the same frequency by approximately two orders of magnitude[12], generating large non-local effects. Of particular interest here are the nonlinearities in the infrared spectral region, which may be enhanced due to the presence of plasmons[12,13]. Recently, B. Yao[7] and Constant[14] have independently reported experimental measurements of a frequency mixing process, with a difference frequency generation (DFG) in the mid-infrared, that implied enhancement due to the presence of plasmons. Such an all-optical coupling scheme for plasmon generation in graphene holds great promise, for example, in the design of plasmon sensors or new THz sources.

Given the large nonlinearities observed in graphene and its potential for optical devices, it would be highly beneficial to develop a quantitative, microscopic understanding of its origin. There already exist a number of calculations in literature of the second-order conductivity in graphene for DFG conditions, in a regime where the incident fields are assumed to only weakly perturb the equilibrium Fermi carrier distribution. However, the conclusions are not all consistent with one another, and differing models or assumptions have been used to point to consistency with experimental results. Here, our primary objectives are to show definitively that the correct perturbative model of graphene nonlinearities does not describe well existing experiments, and to propose an alternative non-perturbative mechanism based upon photothermal effects, whose predicted strength is closer to experimental values. First, we obtain a second-order conductivity of graphene for DFG that is different from the first theoretical calculation specifically for DFG[6], but provides confirmation to a set of other published theoretical results[8,9,11]. While the final result in itself is not new, we derive it in a different fashion based upon the Peierls substitution[15], which avoids issues that arise in the calculation of the linear conductivity using a vector potential with the Dirac Hamiltonian[11]. We then present a set of original results, beginning with an application of this theoretical result to the experimental conditions of Constant[14] and a quantitative comparison between theory and experiment highlighting a large discrepancy. We explicitly show that the experimental effects observed by

[1]Department of Physics and Astronomy, University of Exeter, Exeter, UK. [2]ICFO-Institut de Ciencies Fotoniques, The Barcelona Institute of Science and Technology, 08860, Castelldefels, Barcelona, Spain. [3]ICREA-Institució Catalana de Recerca i Estudis Avançats, 08015, Barcleona, Spain. C. J. Tollerton and D. Z. Li contributed equally. Correspondence and requests for materials should be addressed to C.J.T. (email: cjt220@exeter.ac.uk)

Constant[14] and B. Yao[7] cannot be attributed to the non-linearity of the substrate. Finally, we discuss other possible contributions to the wave mixing signals observed. In particular, we derive a model showing how such a signal could arise from photothermal effects, and estimate a difference frequency field that is two orders of magnitude larger than that from perturbation theory.

Results and Discussion

Perturbation theory. Formally, the interaction of an electron with a vector potential $\vec{A}$ can be incorporated into a Hamiltonian via the substitution $\hat{\vec{p}} \to \hat{\vec{p}} + e\vec{A}$, where $\hat{\vec{p}}$ is the canonical momentum and e is the elemental charge. Absent the vector potential, an electron in a periodic crystal potential $V(\vec{r})$, with Hamiltonian $\hat{H} = \hat{\vec{p}}^2/2m + V(\vec{r})$, where m is the bare mass of an electron, can be formally diagonalized to produce a band structure. In the case of graphene, the Hamiltonian is typically taken to be of a tight-binding form. The Peierls substitution[15] formally enables one to incorporate the vector potential into such a postulated model, avoiding the need to actually solve for the eigenstates of the Hamiltonian with the replacement $\hat{\vec{p}} \to \hat{\vec{p}} + e\vec{A}$. The tight-binding Hamiltonian of graphene with $\vec{A}$ thus reads:

$$\hat{H} = \mathcal{T}\sum_{i=1,\ldots,N}\left(\hat{a}^{\dagger}_{\vec{R}_i}\hat{a}_{\vec{R}_i} + \hat{b}^{\dagger}_{\vec{R}_i+\vec{\tau}_1}\hat{b}_{\vec{R}_i+\vec{\tau}_1}\right) - \mathcal{T}'\sum_{\substack{i=1,\ldots,N\\ l=1,2,3}}\left(e^{i\frac{e}{\hbar}\vec{A}\cdot\vec{\tau}_l}\,\hat{a}^{\dagger}_{\vec{R}_i}\hat{b}_{\vec{R}_i+\vec{\tau}_l} + e^{-i\frac{e}{\hbar}\vec{A}\cdot\vec{\tau}_l}\,\hat{b}^{\dagger}_{\vec{R}_i+\vec{\tau}_l}\hat{a}_{\vec{R}_i}\right). \tag{1}$$

Here $\mathcal{T}$ and $\mathcal{T}'$ are the diagonal and nearest-neighbor off-diagonal matrix elements of the Hamiltonian respectively in the basis of atomic orbitals in absence of $\vec{A}$; $\left(\hat{a}_{\vec{R}_i}, \hat{a}^{\dagger}_{\vec{R}_i}\right)$ and $\left(\hat{b}_{\vec{R}_i+\vec{\tau}_l}, \hat{b}^{\dagger}_{\vec{R}_i+\vec{\tau}_l}\right)$ are annihilation and creation operators for the two sublattices in graphene, with $\vec{R}_i$, $i = 1, \ldots, N$ denoting the sublattice sites, and τ_l, $l = 1, 2, 3$ denoting the vectors from a lattice site to its three nearest neighbors. The current density operator can then be obtained by $\hat{\vec{j}}(\vec{R}_i) = \partial\hat{H}/\partial\vec{A}$[16].

In weak electromagnetic (EM) fields, both $\hat{H}$ and $\hat{\vec{j}}$ can be expanded in terms of $\vec{A}$. $\hat{H}$ can be broken into a non-interacting part $\hat{H}_0$ and an interacting part $\hat{H}_I$, with $\hat{H}_I = \hat{H}_I^{(1)} + \hat{H}_I^{(2)} + \cdots$ and the superscripts indicating the order of $\vec{A}$:

$$\hat{H}_0 = \mathcal{T}\sum_{i=1,\ldots,N}\left(\hat{a}^{\dagger}_{\vec{R}_i}\hat{a}_{\vec{R}_i} + \hat{b}^{\dagger}_{\vec{R}_i+\vec{\tau}_1}\hat{b}_{\vec{R}_i+\vec{\tau}_1}\right) - \mathcal{T}'\sum_{\substack{i=1,\ldots,N\\ l=1,2,3}}\left(\hat{a}^{\dagger}_{\vec{R}_i}\hat{b}_{\vec{R}_i+\vec{\tau}_l} + \hat{b}^{\dagger}_{\vec{R}_i+\vec{\tau}_1}\hat{a}_{\vec{R}_i}\right), \tag{2}$$

$$\hat{H}_I^{(1)} = -i\frac{e\mathcal{T}'}{\hbar}\sum_{\substack{i=1,\ldots,N\\ l=1,2,3}}\vec{A}(\vec{R}_i, t)\cdot\vec{\tau}_l\left(\hat{a}^{\dagger}_{\vec{R}_i}\hat{b}_{\vec{R}_i+\vec{\tau}_l} - \hat{b}^{\dagger}_{\vec{R}_i+\vec{\tau}_l}\hat{a}_{\vec{R}_i}\right), \tag{3}$$

$$\hat{H}_I^{(2)} = \frac{e^2\mathcal{T}'}{2\hbar^2}\sum_{\substack{i=1,\ldots,N\\ l=1,2,3}}\left[\vec{A}(\vec{R}_i, t)\cdot\vec{\tau}_l\right]^2\left(\hat{a}^{\dagger}_{\vec{R}_i}\hat{b}_{\vec{R}_i+\vec{\tau}_l} + \hat{b}^{\dagger}_{\vec{R}_i+\vec{\tau}_l}\hat{a}_{\vec{R}_i}\right). \tag{4}$$

and similarly $\hat{\vec{j}} = \hat{\vec{j}}_0 + \hat{\vec{j}}^{(1)} + \hat{\vec{j}}^{(2)} + \cdots$:

$$\hat{\vec{j}}_0(\vec{R}_i) = i\frac{e\mathcal{T}'}{\hbar}\sum_{l=1,2,3}\vec{\tau}_l\left(\hat{a}^{\dagger}_{\vec{R}_i}\hat{b}_{\vec{R}_i+\vec{\tau}_l} - \hat{b}^{\dagger}_{\vec{R}_i+\vec{\tau}_l}\hat{a}_{\vec{R}_i}\right), \tag{5}$$

$$\hat{\vec{j}}^{(1)}(\vec{R}_i) = -\frac{e^2\mathcal{T}'}{\hbar^2}\sum_{l=1,2,3}\left[\vec{A}(\vec{R}_i, t)\cdot\vec{\tau}_l\right]\vec{\tau}_l\left(\hat{a}^{\dagger}_{\vec{R}_i}\hat{b}_{\vec{R}_i+\vec{\tau}_l} + \hat{b}^{\dagger}_{\vec{R}_i+\vec{\tau}_l}\hat{a}_{\vec{R}_i}\right), \tag{6}$$

$$\hat{\vec{j}}^{(2)}(\vec{R}_i) = -i\frac{e^3\mathcal{T}'}{2\hbar^3}\sum_{l=1,2,3}\left[\vec{A}(\vec{R}_i, t)\cdot\vec{\tau}_l\right]^2\vec{\tau}_l\left(\hat{a}^{\dagger}_{\vec{R}_i}\hat{b}_{\vec{R}_i+\vec{\tau}_l} - \hat{b}^{\dagger}_{\vec{R}_i+\vec{\tau}_l}\hat{a}_{\vec{R}_i}\right). \tag{7}$$

Crucially, the Peierls substitution yields terms in Eqs (4), (6) and (7), which cannot be obtained by replacing $\hat{\vec{p}}$ with $\hat{\vec{p}} + e\vec{A}$ in the Dirac Hamiltonian (as done in previous works[8,9,11]). In fact, it is the term in Eq. (6) that cancels out a term in Eq. (5) that could otherwise cause a divergence in the linear conductivity[11] (see details in the Methods).

At the Dirac points, following standard procedures, one can derive equivalent spinor forms of the above operators. The expectation value of the current density in the presence of the fields can be calculated as $\langle\hat{\vec{j}}\rangle = \mathrm{Tr}(\hat{\rho}\,\hat{\vec{j}})$, where $\hat{\rho}$ is the (self-consistent) single-particle density matrix. The matrix elements can be calculated by using the time evolution equation $\mathrm{d}\hat{\rho}/dt = (i/\hbar)[\hat{\rho}, \hat{H}_0 + \hat{H}_I]$ and solving the density matrix $\hat{\rho} = \sum_n \hat{\rho}^{(n)}$ perturbatively in powers of $\hat{H}_I$. We leave the detailed derivations of the spinor formalism and the currents to the Methods, and only quote the result for the nonlinear current here. We consider the response to EM fields described by a potential

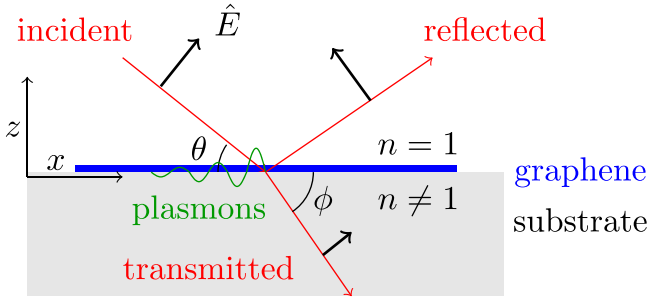


Figure 1. Illustration of electromagnetic fields ($\overrightarrow{E}$) (applicable to pump, probe, and DFG) propagating in the x-z plane. All the fields are p-polarized and the directions of propagation and polarizations are indicated by the red and black arrows respectively. The angles of incidence and transmission are defined in the figure as θ and ϕ.

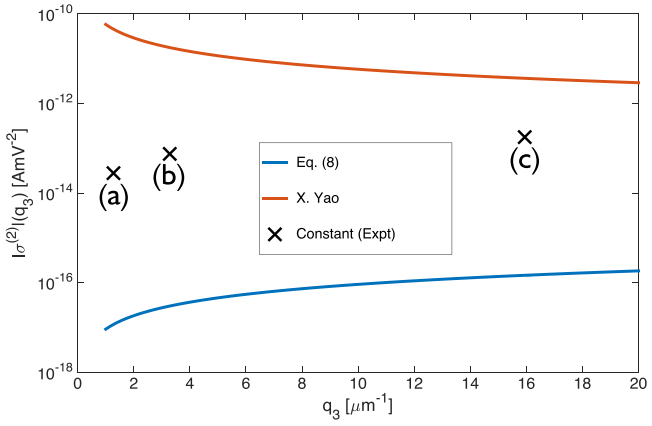


Figure 2. Comparison of the nonlinear conductivity $\sigma^{(2)}(\omega_3, q_3)$ derived here from Eq. (8) with other results derived using perturbation theory: Eq. (5) from X. Yao[6]. All theoretical curves are plotted for a Fermi energy of 500 meV and difference frequency of 15.3 THz. [Experimental estimates from Constant[14], for the different experimental geometries (a), (b), (c) (Table 1) are indicated by black crosses].

Geometry	θ_{pump}	θ_{probe}	λ_{pump} (nm)	$\frac{\omega_3}{2\pi}$ (THz)	$\frac{\Delta R}{R\Phi}$ (mJ^{-1} cm^2)	$\sigma^{(2)}$ (fAmV^{-2})
(a)	45°	55°	607	7.0	−0.0097	24
(b)	50°	70°	597	15.3	−0.025	75
(c)	125°	15°	587	23.8	0.062	180

Table 1. Differential reflectivity, normalized to pump fluence Φ, and experimentally determined $\sigma^{(2)}$ extracted using the model of[14] for three geometries (a), (b) and (c). For all geometries $\lambda_{probe} = 617.53$ nm.

$\overrightarrow{A}(\overrightarrow{r}, t) = (1/2)\sum_{m=1,2}[A_m \hat{x}\, e^{i(q_m x - \omega_m t)} + \text{c. c.}]$, where the electric field components parallel to the graphene layer are related to the potentials by $E_{mx} = -\partial A_m/\partial t$; q_m and ω_m are the wavevectors and angular frequencies, and $\hat{x}$ is a unit vector along the x-direction. Following relevant experiment[14] the fields for $m = 1, 2$ are called "pump" and "probe" respectively, and we illustrate their configurations in Fig. (1). Here we are interested in the case of DFG and look for the nonlinear current at difference frequency $\omega_3 = \omega_2 - \omega_1$ and wavevector $q_3 = q_2 - q_1$, which can be formally written as $j_x^{(2)} = \sigma^{(2)} E_{1x} E_{2x}$, where $\sigma^{(2)}$ defines the second-order conductivity. Under the relevant experimental conditions[7,14] $\omega_1 \approx \omega_2 \gg \omega_3$,

$$\sigma^{(2)}(\omega_3, q_3) \approx -\frac{2e^3 v_F^2}{\pi\hbar^2}\left(\frac{q_3}{\omega_3}\right)\frac{\omega_F^4}{(\omega_3^2 - 4\omega_F^2)\omega_2^4}, \tag{8}$$

where in ω_F and v_F are the Fermi angular frequency and velocity.

Comparison of perturbation theory with experiment. X. Yao[6] were the first to derive the nonlinear conductivity in graphene relevant to DFG. In Fig. 2, we plot this derived second order nonlinear response (Eq. (5) of X. Yao[6] converted to $\sigma^{(2)}$) along with experimental results from Constant[14], recent perturbative calculations[11] and our own Eq. (8). It is important to note first that the results of Wang[11] and Cheng[9] strongly agree with our

own suggesting this is the correct prediction from pertubation theory (in addition Tokman[8] and Rostami[17] achieve the same result except a factor of 2 which may be due to definitions we could not clearly identify). However there is a rather large discrepancy between Eq. (8) and the model derived by X. Yao[6], which were derived for identical conditions using perturbation theory. Moreover the conductivity derived by X. Yao[6] has a non-physical divergence for $q_3 \to 0$. While it is not clear from where this unphysical behavior arises, in a centrosymmetric material such as graphene this behavior is paradoxical. Meanwhile, the conductivity from Eq. (8) tends to zero as $q_3 \to 0$, as it must in graphene. We note that, depending on the value of q_3 in Fig. 2, the magnitude of $\sigma^{(2)}$ predicted by Eq. (8) is at least 4 orders lower than that found by X. Yao[6].

The first experimental signatures attributed to DFG of plasmons were found by Constant[14]. In this experiment, by illuminating the graphene with two tunable, femtosecond laser pulses ("pump" and "probe") with well-defined angles of incidence but different frequencies, Constant[14] was able to phase-match to the plasmon. The geometry of the experiment is the same as that chosen for our theoretical calculation defined in Fig. 1. The graphene supports tightly guided plasmons with a dispersion relation $\omega_{pl}(k)$. The differential reflectance of the probe pulse ΔR was seen to change significantly whenever the difference frequency and wavevector were aligned to the plasmon dispersion relation, $\omega_{pl}(\overrightarrow{k_1} - \overrightarrow{k_2}) = \omega_1 - \omega_2$, suggesting efficient plasmon excitation via DFG. In practice, a range of difference frequencies and wavevectors were scanned by continuously varying the pump wavelength, and by choosing different discrete incident beam angles.

Constant investigated three experimental geometries (noted in Table 1) with different angles of incidence, θ. We examine one of the resonant conditions for each of the three experimental geometries, as defined in Table 1. Assuming the differential reflection signals arise from DFG, one can use the model introduced in the Supplementary Information of Constant[14] (briefly reviewed in the Methods section) to estimate a value for $\sigma^{(2)}$ for each measurement. The results of this analysis, i.e. values of $\sigma^{(2)}$ which describe the experimental signals, are also shown in Table 1.

Figure 2 compares the experimental values of $\sigma^{(2)}$ from Constant[14] and theoretical predictions from Eq. (8), Wang[11] and X. Yao[6]. Firstly, the q_3 dependence of Yao[6] clearly differs greatly from that of both the experiment and the near-linear predictions of other theoretical derivations. The experimental magnitudes of $\sigma^{(2)}$ are also significantly lower than the prediction of X. Yao[6], and several orders higher than those from recent perturbative works[8,9,11]. As found by[17], it is only possible to find agreement between the experiments[14] and perturbative second order calculations if one invokes an unphysically low decay rate for the plasmon (resulting in extraordinarily narrow resonances). More recently, a similar experiment has been carried out by B. Yao[7] in a waveguiding geometry, and the theory from X. Yao[6] was used to model the experimental signals. While the geometries of B. Yao[7] and Constant[14] are significantly different, similar signals were observed in each experiment. Therefore, ignoring the unphysical results in X. Yao[6], the large discrepancy between both experiments and the theoretical consensus points to a second order response that is not purely perturbative, as originally interpreted. In the remainder of this paper, we therefore discuss other possible contributions which might account for the discrepancy between perturbation theory and the experiments of B. Yao[7] and Constant[14].

Substrate Response. In this section we consider contribution of the second order nonlinearity of the quartz substrate used in experiment[14]. The analysis significantly simplifies if the nonlinear polarization is generated far from a phase-matching condition of the bulk, and depletion can be ignored, as should be the situation for Constant[14]. In this case, the pump and probe fields generate a polarization in quartz are given by

$$P_3(\overrightarrow{r}, t) = \varepsilon_0 \chi^{(2)} e^{i\left[(\overrightarrow{k}_{T1} - \overrightarrow{k}_{T2})\cdot\overrightarrow{r} - (\omega_1 - \omega_2)t\right]} t_1 t_2^* E_{I1} E_{I2}^*, \tag{9}$$

where $\chi^{(2)}$ is the second order susceptibility of the substrate, $\overrightarrow{k}_{Ti}$, t_i, and E_{Ii} denote the wavevector on the transmitted (substrate) side, transmission coefficient, and the incident field amplitude of the pump ($i = 1$) and the probe ($i = 2$) fields respectively. The transmission coefficients t_i are given in the Methods. The subscript $i = 3$ indicates quantities corresponding to the difference frequency signal at $\omega_3 = \omega_1 - \omega_2$. As charge density waves in graphene are driven by an electric field, we must relate the nonlinear polarization to the field generated in the quartz, which satisfies the wave equation

$$-\nabla^2 E_{3s} + \frac{\varepsilon_3}{c^2}\frac{\partial^2 E_{3s}}{\partial t^2} = -\mu_0 \frac{\partial^2 P_3}{\partial t^2}. \tag{10}$$

Here, the subscript "s" denotes that this is an effective source field that will later drive a response in the graphene (distinct from the resulting plasmon field). Also, $\varepsilon_3 = n(\omega_3)^2$ indicates the permittivity of quartz evaluated at the difference frequency, with the model of the frequency dependent $n(\omega)$ given in the Appendix. Due to the plane-wave nature of P_3, E_{3s} takes on the same spatial and frequency dependence. In our regime of interest, the spatial derivative of the field, $|\nabla^2 E_{3s}| = |\overrightarrow{k}_{T1} - \overrightarrow{k}_{T2}|^2 E_{3s}$, is significantly larger than the time derivative. This is because the pump and probe fields are chosen to phase-match with surface plasmons in graphene (thus the associated wavevectors are much larger than free-space fields at the difference frequency). Thus the field amplitude created by the nonlinear polarization is well approximated by

$$E_{3s} \approx \frac{(\omega_1 - \omega_2)^2}{c^2 |\overrightarrow{k}_{T1} - \overrightarrow{k}_{T2}|^2} \chi^{(2)} t_1 t_2^* E_{I1} E_{I2}^*. \tag{11}$$

In particular, it should be noted that a large wavevector mismatch results in strong suppression of the field. To simplify the discussion, we will assume the scenario which produces the highest field, i.e. in which E_{3s} is

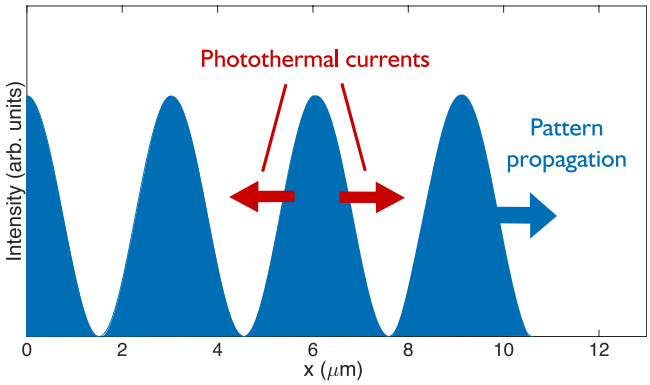


Figure 3. Intensity pattern generated from the interference of two beams in geometry (b) ($\theta_{pump} = 70°$, $\theta_{probe} = 50°$ $\lambda_{pump} = 587$ nm, $\lambda_{probe} = 617.53$ nm). The temperature gradient in the sample follows this pattern and generates photothermal currents when thermalising. The pattern propagates, phase-matched to the difference frequency field, with wavevector $q_3 = q_1 - q_2$.

completely polarized along $\hat{x}$ (parallel to the graphene sheet) so that it maximally drives a charge density wave in graphene. As we see below, even in this best case scenario, the generated field is rather small.

Since the nonlinear response is considered here to be completely within the substrate, which provides an effective source field E_{3s}, the remaining part of the calculation is completely linear in its nature. Using the conventions in Fig. 1, we take "reflected" and "transmitted" field components of unknown amplitude, which correspond to the plasmon field on the vacuum and substrate sides. The wavevector along $\hat{x}$ for these fields is equal to $q_3 = q_{T1} - q_{T2}$, where q_{T1} and q_{T2} are the in-plane components of $\vec{k}_{T1}$ and $\vec{k}_{T2}$, while the perpendicular components must satisfy the respective dispersion relations for each side of the interface, *e.g.*, $k_{T3z}^2 = \varepsilon_3(\omega_3/c)^2 - q_3^2$. Similar to the procedures to solve the pump (probe) field laid out in the methods, the two unknown field amplitudes can be readily solved by taking E_{3s} to be the incident field on the substrate side, and enforcing electromagnetic boundary conditions at the vacuum-graphene-quartz interface, which yields the following parallel-field component on the substrate side, evaluated at the graphene layer (z = 0),

$$E_{pl} = -E_{3s}\frac{(c\varepsilon_0 + \sigma^{(1)}(\omega_3)\sin\theta_3)\sin\phi_3}{c\varepsilon_0 \sin\phi_3 + \sin\theta_3(c\varepsilon_0 n_3 + \sigma^{(1)}(\omega_3)\sin\phi_3)}. \tag{12}$$

Here, $\sigma^{(1)}(\omega_3)$ is the linear conductivity of graphene evaluated at frequency ω_3.

Specifically we can numerically evaluate E_{pl} for geometry (b) in[14]. Taking a value of $\chi^{(2)} = 0.3$ pmV^{-1} for quartz[18], we find that $E_{pl} \approx 15$ Vm^{-1}. The modeling in[14] predicts a considerably larger value for the inferred plasmon field in experiment of $\approx 8 \times 10^4$ Vm^{-1}. We therefore do not believe that the substrate nonlinearity contributes significantly to the signals observed in[14]. However, we note that our substrate model does not consider any surface enhanced nonlinearity. Its theoretical modeling would require experimental measurements of surface nonlinear coefficients relevant to our system, which we were unable to find in literature.

Photothermal Effect. Here, we present an alternative mechanism by which a plasmon field at the difference frequency and wavevector can be generated. Fundamentally, the effect discussed below arises from the linear Seebeck effect, so that the total current is described by

$$j_x(\omega_3, q_3) = \sigma^{(1)}(\omega_3, q_3)E_{3x} + \sigma^{(2)}E_{1x}E_{2x} + \sigma^{(1)}S\frac{dT}{dx}. \tag{13}$$

The first term on the right hand side describes the normal linear relationship between the current and field, while the second describes the conventional second order electronic nonlinearity. The third term, and most important here, arises due to photothermal effects, and accounts for the Seebeck current emerging due to a temperature gradient $\left(\frac{dT}{dx}\right)$ in a material described by Seebeck coefficient S. As we discuss below, this term can give rise difference frequency currents even in the absence of a nonlinear conductivity (i.e. even when $\sigma^{(2)} = 0$).

It is known that excitation of graphene carriers by intense femtosecond pulses ($\Phi \approx 0.1$ mJcm^{-2} with pulse width ~100 fs in[14]) is not perturbative in nature. The electron temperature is raised by several thousand kelvin under such excitation[19] and is not in equilibrium with the phonon temperature. Furthermore, under geometries similar to those used in B. Yao[7] and Constant[14], heating due to optical illumination is not homogeneous. When two or more light sources of similar frequency are incident on an interface at oblique angles, the result will be a near stationary interference pattern such as that shown in Fig. 3. When the frequencies are slightly different, the pattern will propagate in the plane with a velocity equal to ω_3/q_3. The spatially dependent interference pattern will give rise to a temperature modulation that propagates along the graphene sheet, which, via the Seebeck effect, can also generate a current and thereby drive a plasmon excitation when phase-matching conditions are satisfied.

Below, we present a theoretical model of this effect, and show that it can generate a difference frequency field that is orders of magnitude larger than that calculated from the perturbation theory.

The intensity pattern imprinted on the graphene sheet can be expressed as

$$I(x,\,t) = \frac{c\varepsilon_0}{2}\left|E_{1x}e^{i(q_1x-\omega_1t)} + E_{2x}e^{i(q_2x-\omega_2t)}\right|^2 \approx \frac{c\varepsilon_0 {E_{1x}}^2}{2}\left[1 + 2\left|\frac{E_{2x}}{E_{1x}}\right|\cos(q_3x - \omega_3t)\right] + \mathcal{O}(E_{2x}^2). \tag{14}$$

Here E_{1x} and E_{2x} are the in-plane components of the pump and probe fields respectively, and we have assumed the probe field is much weaker than the pump field. The ratio of the in-plane components can be calculated as $|E_{2x}/E_{1x}| = (t_2^{(L)}\sin\phi_2)/(t_1^{(L)}\sin\phi_1)\sqrt{I_2/I_1}$, with I_1 and I_2 being the incident intensities of the pump and probe beams respectively, and the linear transmission coefficient $t_i^{(L)}(i = 1,\,2)$ is given by Eq. (38) considering only the linear optical conductivity of graphene Eq. (40). Using the parameters in the experiment of Constant[14], we obtain $|E_{2x}/E_{1x}| \approx 0.1$.

The intensity pattern acts as a heat source for the temperature distribution which, in linear response theory, satisfies a diffusion equation:

$$\frac{\partial T}{\partial t} - \alpha\frac{\partial^2 T}{\partial x^2} = \beta I(x,\,t) - y(T - T_0), \tag{15}$$

where α is the diffusivity, β is the heating rate due to the intensity pattern, and y is the relaxation rate back to the equilibrium temperature T_0. Due to the linearity of the equation, it can be readily solved in the Fourier domain, in which when taking into account Eq. (14) the solution takes the form

$$T(x,\,t) = T_0 + T_{dc} + T_{ac}\cos(q_3x - \omega_3t - \Psi). \tag{16}$$

Here, T_{dc} is the (large) position- and time-independent temperature increase arising from the incident lasers, while T_{ac} represents a position- and time-varying temperature oscillation that must necessarily be generated in the presence of moving intensity interference pattern. Ψ denotes a phase offset between the intensity and temperature modulations, whose specific form is not relevant here. Substituting $T(x, t)$ into Eq. (15), in the regime of interest $y \ll \omega_3$, one obtains $|T_{ac}|/T_{dc} \approx 0.2y/\omega_3$. As expected, the temperature modulation T_{ac} is reduced significantly as the oscillation frequency ω_3 increases with respect to the damping rate. It is known that intense, femtosecond pulses similar to those in[14] lead to T_{dc} of approximately 2000 K[19]. The relaxation rate y is due to electron-phonon scattering, and we take a value of $y \approx 1/(100\,\text{fs})$[19]. At $\omega_3 = 2\pi \times 10\,\text{THz}$ these parameters give a temperature modulation $T_{ac} = 60\,\text{K}$.

The Seebeck effect enables the generation of a source current in the presence of a temperature gradient, and this can be described by $j_s(\omega_3,\,q_3) = \sigma^{(1)}(\omega_3)q_3S|T_{ac}|$, where S is called the Seebeck coefficient. In principle, the Seebeck coefficient could be frequency and wavevector dependent. However, this dependence has not been measured carefully in literature, nor is it straightforward to calculate from first principles. There have been several measurements of the Seebeck effect in graphene, both in DC experiments ($S \approx 5 \times 10^{-5}\,\text{V/K}$[20], $S \approx 8 \times 10^{-5}\,\text{V/K}$[21]) and under illumination from 100 fs pulses ($S \approx 10^{-4}\,\text{V/K}$[22]). Whilst it is hard to predict how the Seebeck effect behaves on 10 fs timescales relevant here (corresponding to peak to peak propagation time of the intensity pattern in Fig. 3), it is likely that photothermal effects will be higher on ballistic timescales, as with other materials[23]. Here we use a conservative value of $S \approx 10^{-4}\,\text{V/K}$ reported in[22].

Now using the standard EM boundary conditions at the graphene layer (see Fig. 1), with the aid of the charge continuity equation, one can find the relation between the electric fields and the surface current density at the difference frequency ω_3. Note now the Seebeck effect contribution needs to be added to the surface current density:

$$j_x(\omega_3,\,q_3) = \sigma^{(1)}(\omega_3)E_{3x} + j_s(\omega_3,\,q_3). \tag{17}$$

Then solving the equations of the boundary conditions (see Methods for more details), we obtain for the electric field at the difference frequency:

$$E_{3x} = -\frac{t_3^{(L)}}{2c\varepsilon_0}\sigma^{(1)}(\omega_3)q_3S|T_{ac}|\sin\phi_3, \tag{18}$$

where $t_3^{(L)}$ is the linear transmission coefficient at frequency and wavevector ω_3, q_3. For geometry (b) of Constant[14] we find a magnitude of E_{3x}, when on plasmon resonances, of $E_{pl} \approx 2.3 \times 10^3\,\text{V/m}$.

Just as the pump and probe fields can generate a plasmon field through the Seebeck effect, a back-action effect (involving Seebeck mixing of the plasmon and pump fields) results in a change of the probe differential reflectance. In principle, this could be rigorously calculated in a manner similar to above, but this would require knowledge of the Seebeck coefficient at *optical* frequencies, which has never been measured or calculated. However, we can nonetheless obtain an approximate value for the differential reflection, by exploiting conservation of energy. In particular, in steady state, the number of plasmons dissipated per unit time must equal the rate of photons removed from (added to) the pump (probe) beam. There are two contributions to the energy dissipation at the difference frequency: both the graphene layer and the substrate will exhibit absorption. For graphene, the power loss per unit area from absorption can be found by $P_g = (1/2)\,\text{Re}\,\sigma^{(1)}(\omega_3)|E_{3x}|^2$[24]. In the experiment, the substrate

itself (quartz) can provide a non-negligible loss, through coupling with phonons. The corresponding power loss per unit area can be calculated as $P_s = (\varepsilon_0/2)\int_{-\infty}^{0} dz\ \omega_3\ \mathrm{Im}\,[n^2(\omega_3)]\ |E_3|^2$, where our model for the frequency-dependent refractive index $n(\omega)$ is provided in the Methods. Then, the number of photons, at the difference frequency, absorbed per unit time and area is $\Gamma_d = (P_g + P_s)/\hbar\omega_3$. At the individual photon level of a DFG process, an incoming pump photon breaks down to an outgoing probe photon and a plasmon, and therefore the number of plasmons created is also equal to the number of newly generated probe photons that enter either the reflected or transmitted beam. The number of photons per unit time and area in the incident probe beam is simply $\Gamma_{in} = I_2\ \sin\theta_2/\hbar\omega_2$. Thus the order of magnitude of the differential reflectance of the probe beam can be estimated as $\Delta R/Rn \sim \Gamma_d/\Gamma_{in}$. For configuration (b) in Table 1, we estimate the peak differential reflectance after normalized by the fluence (0.1 mJ cm^{-2}) to be ~7×10^{-7} mJ^{-1} cm^2. Although we emphasize that the Seebeck effect and the electronic nonlinearity $\sigma^{(2)}$ are completely independent effects, nonetheless to facilitate a better comparison, one can ask what hypothetical value of nonlinear conductivity $\sigma_S^{(2)}$ would be required, in order to produce the same current as predicted from the Seebeck effect, i.e. $j_s = \sigma_S^{(2)} E_{1x} E_{2x}^*$. We extract a value of $\sigma_S^{(2)} \approx 3.1 \times 10^{-15}$ AmV^{-2}.

While this model predicts a value slightly smaller than experiment ($\sigma^{(2)} \approx 7.5 \times 10^{-14}$AmV^{-2}[14]), the Seebeck coefficient could be larger on ballistic timescales relevant here (≈10 fs peak to peak propagation time) as is expected for other materials[23]. Nevertheless such an effect is fundamental to the experiments and is significantly larger than the predictions of perturbation theory.

We note that photothermal effects will be prominent in the waveguiding geometry of B. Yao[7] (in such a geometry, even though the absolute field intensities are lower, the considerably larger propagation length can compensate). Interestingly, the power dependence of such a photothermal signal would not necessarily follow that of conventional difference frequency generation ($j = \sigma^{(2)} E_1 E_2^*$), and could explain those observed by Constant[25]. Investigating the intensity dependences of these nonlinear signals could provide great insight into the origins of these effects. We also note that the photo-Dember effect can similarly induce local intensity dependent currents and is surprisingly large in graphene on ultrafast timescales[26]. However, since the photo-Dember effect depends on mobility asymmetry between electrons and holes, it will be sample and substrate specific, making it difficult to estimate.

Conclusions

We have derived a second order conductivity of planar graphene ($\sigma^{(2)} \approx 10^{-17}$ AmV^{-2}) with non-local perturbation theory, addressing the long wavelength divergence in[6,7] and divergent linear current in[11]. However, while our result is in agreement with recent calculations[8,9,11,17], it is insufficient to explain observations from experiment[7,14]. We also show that this discrepancy cannot be attributed to the bulk nonlinearity of the substrate.

We also discuss the possibility of photothermal effects in experiments, wherein a spatial intensity pattern resulting from interference of incident beams leads to local inhomogeneous heating of the sample and show that these effects will give rise to frequency mixing currents. We derive a rigorous model for DFG arising from photothermal effects (with the only uncertainties arising from knowledge of material properties such as the Seebeck coefficient), and conservatively estimate a DFG current which is at least two orders of magnitude larger than that found from perturbation theory, significantly closer to experimental estimates from[14]. Microscopic modeling of such local photothermal effects (and other non-equilibrium processes) presents a considerable challenge, and it would be interesting to develop theoretical techniques to do so. We believe that such efforts would shed further light on discrepancies between recent experiments[7,14] and theory[6–11] for all-optical plasmon generation processes in graphene, and enable the strengths of these nonlinear processes to be optimized for future nonlinear optical applications.

Methods

Spinor formalism. We can perform a Fourier expansion on the operators ($\hat{a}_{\vec{R}_i}$, $\hat{a}^\dagger_{\vec{R}_i}$) and ($\hat{b}_{\vec{R}_i+\vec{\eta}_l}$, $\hat{b}^\dagger_{\vec{R}_i+\vec{\eta}_l}$) in the tight-binding Hamiltonian in terms of the operators in the reciprocal lattice space:

$$\hat{a}_{\vec{R}_i} = \frac{1}{\sqrt{N}} \sum_{\vec{k}\in\Omega_B} \hat{a}_{\vec{k}}\, e^{i\vec{k}\cdot\vec{R}_i}, \quad \text{and} \quad \hat{b}_{\vec{R}_i+\vec{\eta}_l} = \frac{1}{\sqrt{N}} \sum_{\vec{k}\in\Omega_B} \hat{b}_{\vec{k}}\, e^{i\vec{k}\cdot(\vec{R}_i+\vec{\eta}_l)}, \tag{19}$$

in which Ω_B denotes the first Brillouin zone, $\vec{k}$ is the electron momentum and N is the number of sites in one sublattice. One can then substitute these expansions into the expressions of the Hamiltonian Eqs (2–4) and current density operators Eqs (5–7). As usual, near the two Dirac points $\vec{K}$ and $-\vec{K}$[27], the operators can be expanded in orders of $\vec{k}$ referenced from $\vec{K}$ or $-\vec{K}$: $\vec{k} \mp \vec{K} \to \vec{k}$. One can then derive equivalent spinor forms for Eqs (2–7) in the first quantization picture. If only the terms to lowest order of $\vec{k}$ are kept, then

$$\hat{H}_0 \to v_F \hat{\vec{\sigma}} \cdot \hat{\vec{p}} \quad \text{at } \vec{K}, \quad \text{and} \quad -v_F \hat{\vec{\sigma}}^* \cdot \hat{\vec{p}} \quad \text{at } -\vec{K}, \tag{20}$$

in which the Fermi velocity $v_F = \sqrt{3}\,a\mathcal{T}'/2\hbar$, with a being the lattice constant of the underlying sublattices, and $\hat{\vec{\sigma}} \equiv \hat{\sigma}_x\hat{x} + \hat{\sigma}_y\hat{y}$ with $\hat{\sigma}_{x,y}$ representing the Pauli spin matrices. Meanwhile, for $\hat{H}_I$:

$$\hat{H}_I^{(1)} \to \pm\, e v_F A(\vec{r}, t)\hat{\sigma}_x, \tag{21}$$

with "+" at $\vec{K}$ and "−" at $-\vec{K}$, and we have for simplicity assumed $\vec{A}$ is along the x-axis. The second order component is meanwhile given by

$$\hat{H}_I^{(2)} \to -\frac{e^2a^2\mathcal{T}}{8\hbar^2}\frac{A_1A_2^*}{4}e^{\mathrm{i}(qx-\omega t)}\hat{\sigma}_x, \tag{22}$$

at both $\vec{K}$ and $-\vec{K}$, where we have taken $\vec{A}$ to be in the form $\vec{A}(\vec{r}, t) = (1/2)\sum_{m=1,2}[A_m\hat{x}\,e^{\mathrm{i}(q_m x-\omega_m t)} + \text{c.c.}]$, where q_m is the in-plane component of the momentum and only kept the terms that give rise to a perturbation at $\omega = \omega_1 - \omega_2$ and $q = q_1 - q_2$ (DFG). For $\hat{H}_0$ we find that the single-particle eigenenergies $\mathcal{E}_s = sv_F\hbar k$ and eigenstates are

$$\psi_{\vec{k}s} = \frac{1}{\sqrt{2\xi}}e^{\mathrm{i}\vec{k}\cdot\vec{r}}\chi_s, \quad \text{with} \quad \chi_s = \begin{pmatrix}1\\ se^{i\theta_{\vec{k}}}\end{pmatrix} \text{ at } \vec{K}, \quad \text{and} \quad \chi_s = \begin{pmatrix}-se^{i\theta_{\vec{k}}}\\ 1\end{pmatrix} \text{ at } -\vec{K}. \tag{23}$$

Here $s = \pm 1$ is the band index, ξ is the area of the graphene sheet, and $\theta_{\vec{k}}$ is the polar angle of $\vec{k}$. Similarly, the spinor forms of the current densities can be found as

$$\hat{\vec{j}}_0(\vec{r}) \to -ev_F\hat{\vec{\sigma}} \quad \text{at } \vec{K}, \quad \text{and} \quad ev_F\hat{\vec{\sigma}}^* \quad \text{at } -\vec{K}, \tag{24}$$

$$\begin{aligned}\hat{j}_x^{(1)} &\to \frac{e^2a^2\mathcal{T}'}{4\hbar^2}A(\vec{r}, t)\hat{\sigma}_x \quad \text{at } \pm\vec{K}, \quad \text{and} \quad \hat{j}_x^{(2)}\\ &\to \pm\frac{e^3a^2v_F}{32\hbar}A_1A_2^*e^{\mathrm{i}(qx-\omega t)}\hat{\sigma}_x \quad \text{“}+\text{”at } \vec{K} \text{ and “}-\text{”at } -\vec{K}.\end{aligned} \tag{25}$$

It should be noted that in addition to the "typical" contribution to the current density, Eq. (24), which is associated with the Bloch momentum, there is also the "diamagnetic" term of $\hat{j}_x^{(1)}$ in Eq. (25). We will see in the following that this term, and in fact a higher correction to this term of order $\vec{k}$, cancels the term that could otherwise cause a divergence in the linear conductivity.

Density matrix method. The time evolution of the density matrix under the interaction with EM fields described by $\hat{H}_I$ is given by $\mathrm{d}\hat{\rho}/dt = (i/\hbar)[\hat{\rho}, \hat{H}_0 + \hat{H}_I]$. From perturbation theory, the matrix elements $\rho_{nm} \equiv \langle n|\hat{\rho}|m\rangle$ of the i th order perturbation of $\hat{\rho}$ satisfy

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_{nm}^{(i)} = -\mathrm{i}\,\omega_{nm}\,\rho_{nm}^{(i)} - \frac{\mathrm{i}}{\hbar}[\hat{H}_I, \hat{\rho}^{(i-1)}]_{nm} - \gamma\,\rho_{nm}^{(i)}, \quad (i \geq 1) \tag{26}$$

in which n, m are dummy indices denoting both the band index s and wavevector $\vec{k}$, $\omega_{nm} = (\mathcal{E}_n - \mathcal{E}_m)/\hbar$, and γ is a phenomenological dissipation term introduced universally for all matrix elements. We take the density matrix in absence of fields to be the Fermi distribution f_n with Fermi energy $\hbar\omega_F$ at zero temperature, $\rho_{nm}^{(0)} = \Theta(\hbar\omega_F - \mathcal{E}_n)\delta_{mn}$, where Θ is the Heaviside step function. A solution to Eq. (26) can be written as

$$\rho_{nm}^{(i)}(t) = -\frac{\mathrm{i}}{\hbar}\int_{-\infty}^{t}\mathrm{d}t'[\hat{H}_I(t'), \hat{\rho}^{(i-1)}(t')]_{nm}e^{-(\mathrm{i}\omega_{nm}+\gamma)(t-t')}. \tag{27}$$

Linear conductivity. In previous works calculating graphene conductivities using the vector potential[8,9,11], the authors typically replace $\hat{\vec{p}}$ by $\hat{\vec{p}} + e\vec{A}$ in the Dirac Hamiltonian. However, the linear current thus calculated has a term that diverges when the integration limit of the electronic momenta is taken to be infinity. This issue was fixed in[11] by adding an artificial quadratic term to the Dirac Hamiltonian. In this section we show that this problematic term is actually canceled by a term in Eq. (7) in the Results section, thus no artificial term needs to be introduced to regularize the calculation.

To begin with we consider the current response to an in-plane electric field described by a vector potential $\vec{A}(\vec{r}, t) = (1/2)A\hat{x}e^{\mathrm{i}(\vec{q}\cdot\vec{r}-\omega t)} + \text{c. c.}$, where $\vec{q} = q\hat{x}$. The current generated at $\vec{q}$ and ω is calculated through the expectation value of $\hat{j}_x(\vec{q}) = 2/\xi\,e^{-iqx}\,\hat{j}_x(\vec{r})$, applied to the density matrix

$$\langle\hat{j}_x(\vec{q})\rangle = \mathrm{Tr}(\hat{\rho}\,\hat{j}_x(\vec{q})) \approx \sum_{n,m}\rho_{nm}^{(1)}j_x^{(0)}(\vec{q})_{mn} + \rho_{nm}^{(0)}j_x^{(1)}(\vec{q})_{mn}, \tag{28}$$

where the superscripts (0), (1),… denote the order of $\vec{A}$ included in the terms. According to Eq. (27), the first-order density matrix is given by

$$\rho_{nm}^{(1)}(t) = \frac{eAv_F}{2\hbar}\frac{f_m - f_n}{\omega - \omega_{nm} + i\gamma}\langle n|\hat{\sigma}_x e^{iqx}|m\rangle e^{-i\omega t}. \tag{29}$$

The matrix elements of the current density operators can be obtained by using Eqs (23–25). Now we can substitute these results into Eq. (28). We then replace the summation on states by an integral over Bloch momenta $\vec{k}$,

introducing an upper bound on the range of integration $k < k_c$ (which approximately captures the edge of the Brillouin zone). The first term of Eq. (28) becomes

$$\sum_{n,m} \rho^{(1)}_{nm} j^{(0)}_x(\vec{q})_{mn} = \frac{e^2 A v_F}{8\pi\hbar} e^{-i\omega t}\left\{2(k_c - k_F) + \frac{\omega}{2v_F}\left[\ln\frac{\omega + 2\omega_F}{|\omega - 2\omega_F|} + i\pi\,\Theta(\omega - 2\omega_F)\right]\right\}, \tag{30}$$

where k_F is the Fermi wavevector, and the second term gives zero. If k_c is extended to infinity like in a free-electron gas, Eq. (30) will yield a divergent linear current, as discussed by Wang[11]. We show next how this strong dependence on k_c is cancelled by the lowest order non-zero contribution to the second term in Eq. (28).

We expand the linear current density operator Eq. (6) to first order of $\vec{k}$, and obtain an additional term to $\hat{j}^{(1)}_x$ in Eq. (25) which we label as $\hat{j}^{(1)'}_x$:

$$\hat{j}^{(1)'}_x = \frac{e^2 a^2}{4\hbar^2} v_F A(\vec{r}, t)\left(\pm\hat{\sigma}_x\hat{p}_x + \frac{1}{3}\hat{\sigma}_y\hat{p}_y\right), \tag{31}$$

where "+" or "−" sign corresponds to the Dirac point $\vec{K}$ or $-\vec{K}$. Now the second term in Eq. (28) has an additional term which gives a finite contribution:

$$\sum_{n,m} \rho^{(0)}_{nm} j^{(1)'}_x(\vec{q})_{mn} = -\frac{e^2 A v_F}{8\pi\hbar} e^{-i\omega t}\frac{a^2(k_c^3 - k_F^3)}{9}. \tag{32}$$

We note $k_F \ll k_c$ and $k_F \gg a^2 k_F^3$, thus Eq. (32) cancels with the k_c term in Eq. (30) at $k_c = 3\sqrt{2}/a \approx |\vec{K}|$, the edge of the Brillouin zone[27]. An exact calculation beyond the Dirac cone approximation would also result in the same qualitative cancellation and a small correction depending on the details of the entire band structure.

Eliminating A by $\vec{E} = -\partial\vec{A}/\partial t$ in the expression of $\langle\hat{j}_x(\vec{q})\rangle$, using the definition $\vec{j} = \sigma^{(1)}\vec{E}$ where $\sigma^{(n)}$ denotes the n th order conductivity, and multiplying the valley and spin degeneracy factor of 4, we finally reach

$$\sigma^{(1)}(\omega) = \frac{e^2}{\hbar}\left[\frac{1}{4}\,\Theta\,(\omega - 2\omega_F) + \frac{i}{\pi}\left(\frac{\omega_F}{\omega} + \frac{1}{4}\ln\frac{|\omega - 2\omega_F|}{\omega + 2\omega_F}\right)\right], \tag{33}$$

which is in agreement with the result derived by various other theoretical approaches, *e.g.* using a scalar potential. We have seen both $\hat{\vec{j}}_0$ and $\hat{\vec{j}}^{(1)}$ (Eqs (5) and (6) in Results) play an important role in obtaining the correct linear conductivity; they are actually analogous to the paramagnetic and diamagnetic parts of current respectively in the case of free electrons coupled to a vector potential. However the replacement $\hat{\vec{p}} \to \hat{\vec{p}} + e\vec{A}$ in the Dirac Hamiltonian would only yield the "paramagnetic" part and therefore give incorrect result for the linear current. Thus the Dirac Hamiltonian is insufficient when using a vector potential and one has to start with the original tight-binding Hamiltonian. We note that the issue with using a vector potential with the Dirac Hamiltonian has been known before, as was pointed out in the works studying optical sum rules[27,28].

Nonlinear conductivity. We consider the nonlinear current in response to the EM fields described by $\vec{A}(\vec{r}, t) = (1/2)\sum_{m=1,2}[A_m\hat{x}\, e^{i(q_m x - \omega_m t)} + \text{c.c.}]$, at difference frequency $\omega_3 = \omega_1 - \omega_2$ and wavevector $q_3 = q_1 - q_2$. Similar to Eq. (28) the nonlinear current can be calculated using the density matrix as

$$\langle\hat{j}_x(\vec{q_3})\rangle = \mathrm{Tr}(\hat{\rho}\hat{j}_x(\vec{q_3})) \approx \sum_{n,m} \rho^{(0)}_{nm} j^{(2)}_x(\vec{q_3})_{mn} + \rho^{(1)}_{nm} j^{(1)}_x(\vec{q_3})_{mn} + \rho^{(2)}_{nm} j^{(0)}_x(\vec{q_3})_{mn}, \tag{34}$$

Using Eqs (23) and (25), we find the first term in Eq. (34) vanishes due to the angular integral. For the second term in the summation, one notes using Eqs (23), (25) and (29), $j^{(1)}_x(\vec{q_3})_{mn}$ flips sign at $\vec{K}$ and $-\vec{K}$ whilst $\rho^{(1)}_{nm}$ stays the same, thus the two contributions at $\vec{K}$ and $-\vec{K}$ cancel. There are two distinctive contributions to the matrix element $\rho^{(2)}_{nm}$ in the third term: there is one contribution coming from the nonlinear vector potential interaction Eq. (22), which when acting upon the equilibrium density matrix $\hat{\rho}^{(0)}$ (see Eq. (27)) produces density matrix $\hat{\rho}^{(2)}$ oscillating at the difference frequency:

$$\rho^{(2)}_{nm}(t) = -\frac{e^2 A^2 t}{32\hbar^2} A_1 A_2^* \frac{f_m - f_n}{\omega - \omega_{nm} + i\gamma}\langle n|\hat{\sigma}_x e^{iqx}|m\rangle e^{-i\omega_3 t}. \tag{35}$$

This term flips sign at $\vec{K}$ and $-\vec{K}$, while $j^{(0)}_x(\vec{q_3})_{mn}$ does not. Thus the two contributions at $\vec{K}$ and $-\vec{K}$ cancel. There is a further contribution coming from the frequency 1(2) component in the linear interaction Eq. (21)) acting upon $\hat{\rho}^{(0)}$ through Eq. (27) to generate a first order perturbation $\hat{\rho}^{(1)}$, and then the frequency 2(1) component in $\hat{H}^{(1)}_I$ acting upon $\hat{\rho}^{(1)}$, and generating a nonlinear perturbation $\hat{\rho}^{(2)}$ at the difference frequency. Using Eqs (21), (27), and (29) one gets

$$\rho_{nm}^{(2)}(t) = \frac{e^2 v_F^2 A_1 A_2^*}{4\hbar^2} e^{-i\omega_3 t} \frac{1}{\omega_3 - \omega_{nm} + i\gamma} \times \sum_l \left(\frac{f_m - f_l}{-\omega_2 - \omega_{lm} + i\gamma} - \frac{f_l - f_n}{\omega_1 - \omega_{nl} + i\gamma} \right) \times \langle n|\hat{\sigma}_x e^{iq_1 x}|l\rangle\langle l|\hat{\sigma}_x e^{-iq_2 x}|m\rangle + (q_1 \leftrightarrow -q_2, \omega_1 \leftrightarrow -\omega_2). \tag{36}$$

One can then carry out the summation in Eq. (34). All band combinations need to be considered (s_n, s_m, $s_l = \pm 1$). The summation can be transformed to an integral, which in general needs to be evaluated numerically. However we can also expand the kernel in terms of q_3, and extract the leading order contributions. Under the experimental conditions of both[7] and[14]: (i.e. $\omega_3 < \omega_F \ll \omega_1 \approx \omega_2$), we obtain

$$\langle \hat{j}_x^{(2)}(\vec{q})\rangle = -\frac{e^3 v_F^2 A_1 A_2^*}{2\pi\hbar^2} \left(\frac{q_3}{\omega_3}\right) \frac{\omega_F^4}{(\omega_3^2 - 4\omega_F^2)\omega_2^2}; \tag{37}$$

the current contribution at $-\vec{K}$ is the same. Thus, changing the potentials to electric fields, using the definition $\langle \hat{j}_x^{(2)}(\vec{q_3})\rangle = \sigma^{(2)} E_1 E_2^*$, and introducing the spin and valley degeneracy factor of 4, we finally reach Eq. (8) in the Results.

Transmission and reflection coefficients. To make this contribution self-contained, in this subsection we briefly review the theoretical model developed in the Supplementary Information of Constant[14] to describe the input and nonlinear electric fields in the experiment. This model assumes a linear frequency dependence of the second order nonlinear graphene conductivity. However, we have tested that this assumption does not significantly modify the result, since the differential reflection signal in the model is predominantly determined by absorption of the difference frequency field, and hence by the magnitude of the conductivity at the difference frequency. Throughout this paper we follow Constant[14] and take a plasmon linewidth of order ~10 THz.

The convention to define the field polarizations and beam angles is illustrated in Fig. 1. In general, the transmission coefficients can be found by imposing boundary conditions at the graphene interface (continuity of the normal electric displacement and tangential electric field). For the transmission coefficient of field i ($i = 1, 2$),

$$t_i = \frac{2\sin\theta_i}{n_i \sin\theta_i + \sin\phi_i} - \left(\frac{\rho_{is}}{\varepsilon_0 E_{Ii}}\right) \frac{\sin\theta_i}{n_i \cos\phi_i (n_i \sin\theta_i + \sin\phi_i)}, \tag{38}$$

where E_{Ii} are the incident field amplitudes, $n_i = n(\omega_i)$ denotes the index of refraction of the substrate at the field frequency, and $\rho_{is} = \rho(\omega_i, q_i)$ is the graphene surface charge density at frequency ω_i and in-plane wavevector $q_i = (\omega_i/c)\cos\theta_i$. The reflection coefficients r_i can be obtained via the relation $r_i + (\sin\phi_i/\sin\theta_i)t_i = 1$. As in Constant[14], we take the substrate model given by Luxmoore[29]:

$$n^2(\omega) = \varepsilon_\infty + \sum_{j=1}^{3} \frac{f_j \omega_{TO,j}^2}{\omega_{TO,j}^2 - \omega^2 - i\omega\gamma_{TO,j}}. \tag{39}$$

The high-frequency dielectric constant $\varepsilon_\infty = 2.4$, and $\omega_{TO} = 2\pi \times (13.44, 23.75, 33.84)$ THz, $\gamma_{TO} = 2\pi \times (0.80, 1.27, 1.27)$ THz, and $f = (0.7514, 0.1503, 0.6011)$ are the frequencies, damping rates, and oscillator weights of the three transverse optical phonon modes respectively. In practice, Eq. (39) is only relevant at low difference frequency of $\omega_3 = \omega_1 - \omega_2$, whereas for the high pump and probe frequencies $\omega_{1,2}$ the index of refraction is nearly a constant: $n \approx \sqrt{\varepsilon_\infty}$. ρ_{is} can be related to the current density j in the graphene layer via the continuity equation, in the Fourier space yielding: $\rho_{is}(\omega_i, q_i) = (q_i/\omega_i) j_x(\omega_i, q_i)$. When depletion can be ignored, as is the case in Constant[14], the current densities at these frequencies can be approximated by their linear response results:

$$j_x(\omega_i, q_i) = \sigma^{(1)}(\omega_i) E_x(\omega_i, q_i). \tag{40}$$

Here, $E_x(\omega_i, q_i) = t_i E_{Ii} \sin\phi_i$ are the total parallel fields in the graphene layer.

The field generated at the difference frequency $\omega_3 = \omega_1 - \omega_2$ can be found similarly by using the boundary conditions at the graphene interface. We note in Fig. 1 now there is no incident field, but only reflected and transmitted fields $\vec{E}_{3R}$ and $\vec{E}_{3T}$ respectively. Using the boundary conditions for the electric fields and displacements at the graphene layer, we have:

$$E_{3R} \sin\theta_3 + E_{3T} \sin\phi_3 = 0, \tag{41}$$

$$E_{3R} \cos\theta_3 - n^2(\omega_3) E_{3T} \cos\phi_3 = \rho_{3s}/\varepsilon_0. \tag{42}$$

The surface charge density ρ_{3s} can be expressed in terms of the current density $j_x(\omega_3, q_3)$. We note now a source term j_s is included in j_x depending on the origin of DFG; this term describes the generated field independent from the graphene's linear response to the field at the frequency: $j_x(\omega_3, q_3) = \sigma^{(1)}(\omega_3) E_{3x} + j_s(\omega_3, q_3)$. j_s can arise from different nonlinear processes, e.g. in Constant[14] this is taken as a result of graphene nonlinearity, while in this contribution this originates from the Seebeck effect, and can then induce a plasmon field. Using Eqs (41) and (42) and the expression of j_x, one can then solve for E_{3x}.

Data Availability

All data created during this research are openly available from the University of Exeter's institutional repository located at https://ore.exeter.ac.uk/repository/handle/10871/35749.

References

1. Novoselov, K. S. *et al*. Electric field effect in atomically thin carbon films. *Sci.* **306**, 666–668 (2004).
2. Castro Neto, A. H. *et al*. The electronic properties of graphene. *Rev. Mod. Phys.* **81**, 109–162 (2009).
3. Mikhailov, S. A. Non-linear electromagnetic response of graphene. *Eur. Lett.* **79**, 27002 (2008).
4. Mikhailov, S. A. Quantum theory of the third-order nonlinear electrodynamic effects of graphene. *Phys. Rev. B* **93**, 085403 (2016).
5. Boyd, R. W. *Nonlinear Optics* (Academic Press, 2008).
6. Yao, X., Tokman, M. & Belyanin, A. Efficient nonlinear generation of THz plasmons in graphene and topological insulators. *Phys. Rev. Lett.* **112**, 055501 (2014).
7. Yao, B. *et al*. Broadband gate-tunable THz plasmons in graphene heterostructures. *Nat. Photonics* **11**, 272–278 (2017).
8. Tokman, M., Wang, Y., Oladyshkin, I., Kutayiah, A. R. & Belyanin, A. Laser-driven parametric instability and generation of entangled photon-plasmon states in graphene and topological insulators. *Phys. Rev. B* **93**, 235422 (2016).
9. Cheng, J. L., Vermeulen, N. & Sipe, J. E. Second order optical nonlinearity of graphene due to electric quadrupole and magnetic dipole effects. *Sci. Rep.* **7**, 43843 (2017).
10. Mikhailov, S. A. Theory of the giant plasmon-enhanced second-harmonic generation in graphene and semiconductor two-dimensional electron systems. *Phys. Rev. B* **84**, 045432 (2011).
11. Wang, Y., Tokman, M. & Belyanin, A. Second-order nonlinear optical response of graphene. *Phys. Rev. B* **94**, 195442 (2016).
12. Koppens, F. H. L., Chang, D. E. & García de Abajo, F. J. Graphene plasmonics: A platform for strong light-matter interactions. *Nano Lett.* **11**, 3370–3377 (2011).
13. Jablan, M., Buljan, H. & Soljačić, M. Plasmonics in graphene at infrared frequencies. *Phys. Rev. B* **80**, 245435 (2009).
14. Constant, T. J., Hornett, S. M., Chang, D. E. & Hendry, E. All-optical generation of surface plasmons in graphene. *Nat. Phys.* **12**, 124–127 (2016).
15. Graf, M. & Vogl, P. Electromagnetic fields and dielectric response in empirical tight-binding theory. *Phys. Rev. B* **51**, 4940 (1995).
16. Landau, L. & Lifshitz, E. *Quantum Electrodynamics: Volume 4*, 2nd edn. (Butterworth-Heinemann, 1982).
17. Rostami, H., Katsnelson, M. I. & Polini, M. Theory of plasmonic effects in nonlinear optics: The case of graphene. *Phys. Rev. B* **95**, 035416 (2017).
18. Shoji, I., Kondo, T. & Ito, R. Second-order nonlinear susceptibilities of various dielectric and semiconductor materials. *Opt. Quantum Electron.* **34**, 797–833 (2002).
19. Hale, P. J., Hornett, S. M., Moger, J., Horsell, D. W. & Hendry, E. Hot phonon decay in supported and suspended exfoliated graphene. *Phys. Rev. B* **83**, 121404 (2011).
20. Wei, P., Bao, W., Pu, Y., Lau, C. N. & Shi, J. Anomalous Thermoelectric Transport of Dirac Particles in Graphene. *Phys. Rev. Lett.* **102** (2009).
21. Zuev, Y. M., Chang, W. & Kim, P. Thermoelectric and magnetothermoelectric transport measurements of graphene. *Phys. Rev. Lett.* **102**, 1–4 (2009).
22. Tielrooij, K. J. *et al*. Generation of photovoltage in graphene on a femtosecond timescale through efficient carrier heating. *Nat. Nanotechnol.* **10**, 437–43 (2015).
23. Zhao, K. *et al*. Laser-induced ultrafast photovoltaic effect in $La_{0.67}Ca_{0.33}MnO_3$ films at room temperature. *Phys. B* **373**, 72–75 (2006).
24. Jackson, J. D. *Classical Electrodynamics*, 3rd edn. (John Wiley & Sons, 1998).
25. Constant, T. J., Hornett, S. M., Chang, D. E. & Hendry, E. Intensity dependences of the nonlinear optical excitation of plasmons in graphene. *Philos. Trans. R. Soc. A* **375** (2017).
26. Liu, C. H. *et al*. Ultrafast Lateral Photo-Dember Effect in Graphene Induced by Nonequilibrium Hot Carrier Dynamics. *Nano Lett.* **15** (2015).
27. Gusynin, V. P., Sharapov, S. G. & Carbotte, J. P. AC conductivity of graphene: from tight-binding model to 2 + 1-dimensional quantum electrodynamics. *Int. J. Mod. Phys. B* **21**, 4611–4658 (2007).
28. Gusynin, V. P., Sharapov, S. G. & Carbotte, J. P. Sum rules for the optical and Hall conductivity in graphene. *Phys. Rev. B* **75**, 165407 (2007).
29. Luxmoore, I. J. *et al*. Strong Coupling in the Far-Infrared between Graphene Plasmons and the Surface Optical Phonons of Silicon Dioxide. *ACS Photonics* **1**, 1151–1155 (2014).

Acknowledgements

The authors thank S.M. Hornett and W.L. Barnes for useful discussions. DEC acknowledges Fundacio Privada Cellex, Spanish MINECO Severo Ochoa Programme SEV-2015-0522 and CERCA Programme/Generalitat de Catalunya; CJT and JB acknowledge the EPSRC, via the EPSRC CDT in Metamaterials. EH acknowledges EPSRC (EP/K041215/1). SARH acknowledges the Royal Society and TATA (RPG-2016-186). All authors acknowledge the European Commission (FP7-ICT-2013-613024-GRASP).

Author Contributions

C.J.T. prepared the manuscript with assistance form J.B., C.J.T., D.E.C. and D.Z.L. derived and checked perturbation theory calculations; T.J.C. and D.E.C. modeled the substrate response; S.A.R.H., D.E.C., E.H., C.J.T. and D.Z.L. derived the photothermal model. All Authors contributed to reviewing the manuscript.

Additional Information

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Bibliography

[1] T. H. Maiman. "Stimulated Optical Radiation in Ruby". In: *Nature* 187.4736 (Aug. 1960), pp. 493–494. ISSN: 1476-4687. DOI: 10/c69k8z.

[2] A. L. Schawlow and C. H. Townes. "Infrared and Optical Masers". In: *Phys. Rev.* 112.6 (Dec. 1958), pp. 1940–1949. DOI: 10/dpmh2x.

[3] P. A. Franken et al. "Generation of Optical Harmonics". In: *Phys. Rev. Lett.* 7.4 (Aug. 1961), pp. 118–119. DOI: 10/bgxf8b.

[4] K. S. Novoselov et al. "Electric Field Effect in Atomically Thin Carbon Films". In: *Science* (Oct. 2004). DOI: 10/bcqksc.

[5] E. Hendry et al. "Coherent Nonlinear Optical Response of Graphene". In: *Phys. Rev. Lett.* 105.9 (Aug. 2010), p. 097401. ISSN: 00319007. DOI: 10/d95qc6. arXiv: 0912.5321.

[6] J.W. You et al. "Nonlinear Optical Properties and Applications of 2D Materials: Theoretical and Experimental Aspects". In: *Nanophotonics* 8.1 (Dec. 2018), pp. 63–97. ISSN: 2192-8614. DOI: 10/ghqcrs.

[7] Achim Woessner et al. "Highly Confined Low-Loss Plasmons in Graphene-Boron Nitride Heterostructures". In: *Nat. Mater.* 14.4 (Apr. 2015), pp. 421–425. ISSN: 14764660. DOI: 10/gfwst6. arXiv: 1409.5674.

[8] M. Zahirul Alam, Israel De Leon, and Robert W. Boyd. "Large Optical Nonlinearity of Indium Tin Oxide in Its Epsilon-near-Zero Region". In: *Science* 352.6287 (May 2016), pp. 795–797. ISSN: 10959203. DOI: 10/bhfn. arXiv: 1702.05883.

[9] Robert W Boyd. *Nonlinear Optics, Third Edition*. 3rd. Orlando, FL, USA: Academic Press, Inc., 2008. ISBN: 0-12-369470-1 978-0-12-369470-6.

[10] Orad Reshef et al. "Beyond the Perturbative Description of the Nonlinear Optical Response of Low-Index Materials". In: *Opt. Lett.* 42.16 (Aug. 2017), p. 3225. ISSN: 0146-9592. DOI: 10/gj7bdp. arXiv: 1702.04338.

[11] Piotr Kabaciński et al. "Nonlinear Refractive Index Measurement by SPM-induced Phase Regression". In: *Opt. Express, OE* 27.8 (Apr. 2019), pp. 11018–11028. ISSN: 1094-4087. DOI: 10/gk8bvs.

[12] Qianfan Xu, Vilson R. Almeida, and Michal Lipson. "Time-Resolved Study of Raman Gain in Highly Confined Silicon-on-Insulator Waveguides". In: *Opt. Express, OE* 12.19 (Sept. 2004), pp. 4437–4442. ISSN: 1094-4087. DOI: 10/drmj76.

[13] Yiyu Zhou et al. "Broadband Frequency Translation through Time Refraction in an Epsilon-near-Zero Material". In: *Nat. Commun.* 11.2020 (2020), p. 2180. ISSN: 2041-1723. DOI: 10/ggwrz8.

[14] Guang S. He. *Nonlinear Optics and Photonics*. Oxford: Oxford University Press, 2014. ISBN: 978-0-19-870276-4. DOI: 10.1093/acprof:oso/9780198702764.001.0001.

[15] Nathalie Vermeulen et al. "Graphene's Nonlinear-Optical Physics Revealed through Exponentially Growing Self-Phase Modulation". In: *Nat Commun* 9.1 (Dec. 2018), p. 2675. ISSN: 2041-1723. DOI: 10/gdsmmx.

[16] Tadatsugu Minami. "Transparent Conducting Oxide Semiconductors for Transparent Electrodes". In: *Semicond. Sci. Technol.* 20.4 (Apr. 2005), S35–S44. ISSN: 02681242. DOI: 10/bnk9br. arXiv: 0804.0063.

[17] Gururaj V. Naik, Vladimir M. Shalaev, and Alexandra Boltasseva. "Alternative Plasmonic Materials: Beyond Gold and Silver". In: *Adv. Mater.* 25.24 (2013), pp. 3264–3294. ISSN: 09359648. DOI: 10/f2c328. arXiv: 1109.5444.

[18] Alan X. Wang, Erwen Li, and Qian Gao. "Electrically-Tunable Subwavelength Grating Using Transparent Conductive Oxide". In: *Smart Photonic and Optoelectronic Integrated Circuits XX*. Ed. by El-Hang Lee and Sailing He. Vol. 1053626. SPIE, Feb. 2018, p. 80. ISBN: 978-1-5106-1557-1. DOI: 10/gj7bc4.

[19] Iñigo Liberal and Nader Engheta. "Near-Zero Refractive Index Photonics". In: *Nat. Photonics* 11.3 (Mar. 2017), pp. 149–158. ISSN: 17494893. DOI: 10/dfqn.

[20] P. R. Wallace. "The Band Theory of Graphite". In: *Phys. Rev.* 71.9 (May 1947), pp. 622–634. ISSN: 0031899X. DOI: 10/cs3tfn. arXiv: 1011.1669v3.

[21] K. S. Novoselov et al. "Electric Field in Atomically Thin Carbon Films". In: *Science* 306.5696 (Oct. 2004), pp. 666–669. ISSN: 00368075. DOI: 10/bcqksc. arXiv: cond-mat/0410550.

[22] A. H. Castro Neto et al. "The Electronic Properties of Graphene". In: *Rev. Mod. Phys.* 81.1 (Jan. 2009), pp. 109–162. ISSN: 00346861. DOI: 10.1103/RevModPhys.81.109. arXiv: 0709.1163.

[23] A. K. Geim. "Graphene: Status and Prospects". In: *Science* 324.5934 (2009), pp. 1530–1534. ISSN: 00368075. DOI: 10/bwk5x9. arXiv: 0906.3799.

[24] Jesus de La Fuente. *Properties of Graphene*. https://www.graphenea.com/pages/graphene-properties#.W9A7E80na70. 2013.

[25] Ricardo A Depine. "Electromagnetics of Graphene". In: *Graphene Optics: Electromagnetic Solution of Canonical Problems*. IOP Publishing, 2016. ISBN: 978-1-68174-309-7. DOI: 10.1088/978-1-6817-4309-7ch1.

[26] William L. Barnes, Alain Dereux, and Thomas W. Ebbesen. "Surface Plasmon Subwavelength Optics". In: *Nature* 424.6950 (2003), pp. 824–830. ISSN: 00280836. DOI: 10/c2w26g. arXiv: 1312.6806.

[27] Stefan A. Maier. *Plasmonics: Fundamentals and Applications*. 2007. ISBN: 0-387-33150-6. DOI: 10.1007/0-387-37825-1.

[28] D. N. Basov, M. M. Fogler, and F. J. García De Abajo. "Polaritons in van Der Waals Materials". In: *Science* 354.6309 (Oct. 2016), aag1992–aag1992. ISSN: 10959203. DOI: 10/f9b2w3. arXiv: 1610.04548.

[29] Tony Low et al. "Polaritons in Layered Two-Dimensional Materials". In: *Nat. Mater.* 16.2 (2017), pp. 182–194. ISSN: 14764660. DOI: 10/f9g6dv. arXiv: 1610.04548.

[30] Salvatore Campione, Igal Brener, and Francois Marquier. "Theory of Epsilon-near-Zero Modes in Ultrathin Films". In: *Phys. Rev. B* 91.12 (Mar. 2015), p. 121408. ISSN: 1098-0121. DOI: 10/dfqm.

[31] Stefan A. Maier. "Graphene Plasmonics: All Eyes on Flatland". In: *Nat. Phys.* 8.8 (2012), pp. 581–582. ISSN: 17452481. DOI: 10/gj7bcb.

[32] M. A. Noginov et al. "Transparent Conductive Oxides: Plasmonic Materials for Telecom Wavelengths". In: *Appl. Phys. Lett.* 99.2 (July 2011), p. 021101. ISSN: 0003-6951. DOI: 10/dh74kb.

[33] D. C. Look and K. D. Leedy. "ZnO Plasmonics for Telecommunications". In: *Appl. Phys. Lett.* 102.18 (May 2013), p. 182107. ISSN: 00036951. DOI: 10/gj7bc7.

[34] J. R. Mejía-Salazar and Osvaldo N. Oliveira. "Plasmonic Biosensing". In: *Chem. Rev.* 118.20 (Oct. 2018), pp. 10617–10625. ISSN: 15206890. DOI: 10/gfh4bt.

[35] Seungchul Kim et al. "High-Harmonic Generation by Resonant Plasmon Field Enhancement". In: *Nature* 453.7196 (2008), pp. 757–760. ISSN: 14764687. DOI: 10/bb8c5r. arXiv: 1106.1603.

[36] Nathaniel Kinsey et al. "Near-Zero-Index Materials for Photonics". In: *Nat. Rev. Mater.* 4.12 (2019), pp. 742–760. ISSN: 20588437. DOI: 10/gj7bm5.

[37] Ting S. Luk et al. "Directional Perfect Absorption Using Deep Subwavelength Low-Permittivity Films". In: *Phys. Rev. B - Condens. Matter Mater. Phys.* 90.8 (Aug. 2014), p. 085411. ISSN: 1550235X. DOI: 10/gj7bj8.

[38] Salvatore Campione et al. "Experimental Verification of Epsilon-near-Zero Plasmon Polariton Modes in Degenerately Doped Semiconductor Nanolayers". In: *Opt. Express* 24.16 (Aug. 2016), p. 18782. ISSN: 1094-4087. DOI: 10/gj7bk5.

[39] Marcus Freitag et al. "Photocurrent in Graphene Harnessed by Tunable Intrinsic Plasmons". In: *Nat. Commun.* 4 (2013), pp. 1951–1958. ISSN: 20411723. DOI: 10/gbdkth. arXiv: 1306.0593.

[40] Daniel Rodrigo et al. "Mid-Infrared Plasmonic Biosensing with Graphene". In: *Science* 349.6244 (2015), pp. 165–168. ISSN: 10959203. DOI: 10/f7jhs6.

[41] Hai Hu et al. "Far-Field Nanoscale Infrared Spectroscopy of Vibrational Fingerprints of Molecules with Graphene Plasmons". In: *Nat. Commun.* 7 (2016), pp. 1–8. ISSN: 20411723. DOI: 10/gj7bc6. arXiv: 1606.09327.

[42] E. M. Darmady. "Flexibility in Laboratory Design". In: *Ann. Clin. Biochem.* 2.2 (1962), pp. 33–34. ISSN: 17581001. DOI: 10/gj7bc5. arXiv: 1412.1061.

[43] David Alcaraz Iranzo et al. "Probing the Ultimate Plasmon Confinement Limits with a van Der Waals Heterostructure". In: *Science* 360.6386 (Apr. 2018), pp. 291–295. ISSN: 10959203. DOI: 10/gdfztm. arXiv: 1804.01061.

[44] Hugen Yan et al. "Damping Pathways of Mid-Infrared Plasmons in Graphene Nanostructures". In: *Nat. Photonics* 7.5 (May 2013), pp. 394–399. ISSN: 17494885. DOI: 10/f4v2qg.

[45] T. J. Constant et al. "All-Optical Generation of Surface Plasmons in Graphene". In: *Nat. Phys.* 12.2 (Feb. 2016), pp. 124–127. ISSN: 17452481. DOI: 10/gcph5k. arXiv: 1505.00127.

[46] Jianing Chen et al. "Optical Nano-Imaging of Gate-Tunable Graphene Plasmons". In: *Nature* 487.7405 (2012), pp. 77–81. ISSN: 00280836. DOI: 10/f997w9. arXiv: 1202.4996.

[47] F. Javier Garciía De Abajo. "Multiple Excitation of Confined Graphene Plasmons by Single Free Electrons". In: *ACS Nano* 7.12 (2013), pp. 11409–11419. ISSN: 19360851. DOI: 10/f5m8qz. arXiv: 1311.4690.

[48] A. Yu Nikitin et al. "Fields Radiated by a Nanoemitter in a Graphene Sheet". In: *Phys. Rev. B - Condens. Matter Mater. Phys.* 84.19 (2011), pp. 1–7. ISSN: 10980121. DOI: 10/bsb69m. arXiv: 1104.3558.

[49] Ryan Beams, Palash Bharadwaj, and Lukas Novotny. "Electroluminescence from Graphene Excited by Electron Tunneling". In: *Nanotechnology* 25.5 (2014). ISSN: 09574484. DOI: 10/gf69pr.

[50] Vladimir M. Shalaev. *ECE 695s Lecture 10: Surface Plasmon Excitation*. Nov. 2006.

[51] H. Ditlbacher et al. "Fluorescence Imaging of Surface Plasmon Fields". In: *Appl. Phys. Lett.* 80.3 (Jan. 2002), pp. 404–406. ISSN: 0003-6951. DOI: 10/dfbjhk.

[52] Craig F. Bohren and Donald R. Huffman. *Absorption and Scattering of Light by Small Particles*. John Wiley & Sons, Sept. 2008. ISBN: 978-3-527-61816-3.

[53] Lukas Novotny. "Effective Wavelength Scaling for Optical Antennas". In: *Phys. Rev. Lett.* 98.26 (June 2007), p. 266802. ISSN: 00319007. DOI: 10.1103/PhysRevLett.98.266802.

[54] N. C. Panoiu et al. "Nonlinear Optics in Plasmonic Nanostructures". In: *J. Opt. U. K.* 20.8 (Aug. 2018), p. 083001. ISSN: 20408986. DOI: 10/gj7bc3.

[55] Orad Reshef et al. "Nonlinear Optical Effects in Epsilon-near-Zero Media". In: *Nat. Rev. Mater.* 4.8 (Aug. 2019), pp. 535–551. ISSN: 20588437. DOI: 10.1038/s41578-019-0120-5.

[56] Baicheng Yao et al. "Broadband Gate-Tunable Terahertz Plasmons in Graphene Heterostructures". In: *Nature Photon* 12.1 (Jan. 2018), pp. 22–28. ISSN: 1749-4893. DOI: 10/gcwskh.

[57] Xianghan Yao, Mikhail Tokman, and Alexey Belyanin. "Efficient Nonlinear Generation of THz Plasmons in Graphene and Topological Insulators". In: *Phys. Rev. Lett.* 112.5 (2014), pp. 1–5. ISSN: 00319007. DOI: 10/gj7bc2. arXiv: 1308.2005.

[58] C.J. Tollerton et al. "Origins of All-Optical Generation of Plasmons in Graphene". In: *Sci. Rep.* 9.1 (2019). ISSN: 20452322. DOI: 10.1038/s41598-019-39961-1.

[59] C. Tollerton. "Nonlinear Optical Effects in Two Dimensional and Thin Film Materials". PhD thesis. University of Exeter, Dec. 2019.

[60] Tao Jiang et al. "Gate-Tunable Third-Order Nonlinear Optical Response of Massless Dirac Fermions in Graphene". In: *Nat. Photonics* 12.7 (July 2018), pp. 430–436. ISSN: 17494893. DOI: 10/gdjb79. arXiv: 1710.04758.

[61] Hassan A. Hafez et al. "Extremely Efficient Terahertz High-Harmonic Generation in Graphene by Hot Dirac Fermions". In: *Nature* 561.7724 (Sept. 2018), pp. 507–511. ISSN: 0028-0836. DOI: 10/gd56cq.

[62] Matthias Baudisch et al. "Ultrafast Nonlinear Optical Response of Dirac Fermions in Graphene". In: *Nat. Commun.* 9.1 (Dec. 2018), p. 1018. ISSN: 20411723. DOI: 10/gc7h2w.

[63] Joel D. Cox and F. Javier García De Abajo. "Plasmon-Enhanced Nonlinear Wave Mixing in Nanostructured Graphene". In: *ACS Photonics* 2.2 (Feb. 2015), pp. 306–312. ISSN: 23304022. DOI: 10/f63dmm. arXiv: 1502.01168.

[64] Irati Alonso Calafell et al. "Giant Enhancement of Third-Harmonic Generation in Graphene–Metal Heterostructures". In: *Nat. Nanotechnol.* 16.3 (Mar. 2021), pp. 318–324. ISSN: 1748-3395. DOI: 10/gjj2bs.

[65] N. Kinsey et al. "Epsilon-near-Zero Al-doped ZnO for Ultrafast Switching at Telecom Wavelengths". In: *Optica* 2.7 (July 2015), p. 616. ISSN: 2334-2536. DOI: 10/ghh84s.

[66] L. Caspani et al. "Enhanced Nonlinear Refractive Index in Epsilon-Near-Zero Materials". In: *Phys. Rev. Lett.* 116.23 (June 2016), p. 233901. ISSN: 10797114. DOI: 10/gj7bd3. arXiv: 1603.03581.

[67] Yuanmu Yang et al. "Femtosecond Optical Polarization Switching Using a Cadmium Oxide-Based Perfect Absorber". In: *Nat. Photonics* 11.6 (June 2017), pp. 390–395. ISSN: 17494893. DOI: 10.1038/nphoton.2017.64.

[68] Fengling Zhang et al. "Rationale behind Subpicosecond Optical Response of Transparent Conductive Oxides in Epsilon-near-Zero Region". In: *J. Appl. Phys.* 129.24 (June 2021), p. 243101. ISSN: 0021-8979. DOI: 10/gkzk78.

[69] M. Zahirul Alam et al. "Large Optical Nonlinearity of Nanoantennas Coupled to an Epsilon-near-Zero Material". In: *Nat. Photonics* 12.2 (Feb. 2018), pp. 79–83. ISSN: 17494893. DOI: 10.1038/s41566-017-0089-9.

[70] V. Bruno et al. "Negative Refraction in Time-Varying Strongly Coupled Plasmonic-Antenna-Epsilon-Near-Zero Systems". In: *Phys. Rev. Lett.* 124.4 (2020), p. 43902. ISSN: 10797114. DOI: 10/gj7bj6. arXiv: 1908.03908.

[71] Ting S. Luk et al. "Enhanced Third Harmonic Generation from the Epsilon-near-Zero Modes of Ultrathin Films". In: *Appl. Phys. Lett.* 106.15 (2015). ISSN: 00036951. DOI: 10.1063/1.4917457.

[72] Yuanmu Yang et al. "High-Harmonic Generation from an Epsilon-near-Zero Material". In: *Nat. Phys.* 15.10 (2019), pp. 1022–1026. ISSN: 17452481. DOI: 10/gf5d99. arXiv: 1902.03539.

[73] Vincenzo Bruno et al. "Broad Frequency Shift of Parametric Processes in Epsilon-near-Zero Time-Varying Media". In: *Appl. Sci. Switz.* 10.4 (Feb. 2020), p. 1318. ISSN: 20763417. DOI: 10/gj7bj4. arXiv: 1912.00052.

[74] Jacob B. Khurgin et al. "Adiabatic Frequency Shifting in Epsilon-near-Zero Materials: The Role of Group Velocity". In: *Optica* 7.3 (2020), p. 226. ISSN: 2334-2536. DOI: 10/gj7bm6.

[75] Cong Liu et al. "Photon Acceleration Using a Time-Varying Epsilon-near-Zero Metasurface". In: *ACS Photonics* (2021). ISSN: 2330-4022. DOI: 10/gjgq4g.

[76] H. Hora. "Y. R. Shen, The Principles of Nonlinear Optics, John Wiley & Sons, New York, 1984, 576 Pages." In: *Laser Part. Beams* 4.2 (May 1986), pp. 318–319. ISSN: 1469-803X, 0263-0346. DOI: 10/c2f4nk.

[77] H. G. Berry, G. Gabrielse, and A. E. Livingston. "Measurement of the Stokes Parameters of Light". In: *Appl. Opt.* 16.12 (1977), p. 3200. ISSN: 0003-6935. DOI: 10/bwpgg6.

[78] Beth Schaefer et al. "Measuring the Stokes Polarization Parameters". In: *Am. J. Phys.* 75.2 (2007), pp. 163–168. ISSN: 0002-9505. DOI: 10/c52jgf.

[79] Philip A. Thomas. *Narrow Plasmon Resonances in Hybrid Systems*. Springer Theses. Cham: Springer International Publishing, 2018. ISBN: 978-3-319-97525-2 978-3-319-97526-9. DOI: 10.1007/978-3-319-97526-9.

[80] Harland Tompkins and Eugene A. Irene. *Handbook of Ellipsometry*. William Andrew, Jan. 2005. ISBN: 978-0-8155-1747-4.

[81] Max Born and Emil Wolf. *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Seventh. Cambridge: Cambridge University Press, 1999. DOI: 10.1017/CBO9781139644181.

[82] Annemarie Berkhout and A. Femius Koenderink. "A Simple Transfer-Matrix Model for Metasurface Multilayer Systems". In: *Nanophotonics* 9.12 (Sept. 2020), pp. 3985–4007. ISSN: 2192-8614, 2192-8606. DOI: 10/gj9dr8.

[83] Steven J. Byrnes. "Multilayer Optical Calculations". In: *arXiv* (Mar. 2016). arXiv: 1603.02720.

[84] M Ferray et al. "Multiple-Harmonic Conversion of 1064 Nm Radiation in Rare Gases". In: *J. Phys. B: At. Mol. Opt. Phys.* 21.3 (Feb. 1988), pp. L31–L35. ISSN: 0953-4075, 1361-6455. DOI: 10/cr4sps.

[85] P. B. Corkum. "Plasma Perspective on Strong Field Multiphoton Ionization". In: *Phys. Rev. Lett.* 71.13 (Sept. 1993), pp. 1994–1997. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.71.1994.

[86] M. Lewenstein et al. "Theory of High-Harmonic Generation by Low-Frequency Laser Fields". In: *Phys. Rev. A* 49.3 (Mar. 1994), pp. 2117–2132. ISSN: 1050-2947, 1094-1622. DOI: 10/dqv5bw.

[87] Zenghu Chang et al. "Generation of Coherent Soft X Rays at 2.7 Nm Using High Harmonics". In: *Phys. Rev. Lett.* 79.16 (Oct. 1997), pp. 2967–2970. ISSN: 0031-9007, 1079-7114. DOI: 10/cgsjtt.

[88] Shambhu Ghimire et al. "Observation of High-Order Harmonic Generation in a Bulk Crystal". In: *Nature Phys* 7.2 (Feb. 2011), pp. 138–141. ISSN: 1745-2473, 1745-2481. DOI: 10/cmt5dw.

[89] O. Schubert et al. "Sub-Cycle Control of Terahertz High-Harmonic Generation by Dynamical Bloch Oscillations". In: *Nature Photon* 8.2 (Feb. 2014), pp. 119–123. ISSN: 1749-4885, 1749-4893. DOI: 10/f5skz6.

[90] T. T. Luu et al. "Extreme Ultraviolet High-Harmonic Spectroscopy of Solids". In: *Nature* 521.7553 (May 2015), pp. 498–502. ISSN: 0028-0836, 1476-4687. DOI: 10/f7d2f4.

[91] Hyunwoong Kim et al. "Generation of Coherent Extreme-Ultraviolet Radiation from Bulk Sapphire Crystal". In: *ACS Photonics* 4.7 (July 2017), pp. 1627–1632. ISSN: 2330-4022, 2330-4022. DOI: 10/gbqj2f.

[92] Nicolas Tancogne-Dejean and Angel Rubio. "Atomic-like High-Harmonic Generation from Two-Dimensional Materials". In: *Sci. Adv.* 4.2 (2018), eaao5207. ISSN: 23752548. DOI: 10/gc39mp.

[93] Shambhu Ghimire and David A. Reis. "High-Harmonic Generation from Solids". In: *Nature Phys* 15.1 (Jan. 2019), pp. 10–16. ISSN: 1745-2473, 1745-2481. DOI: 10/gftbcv.

[94] Georges Ndabashimiye et al. "Solid-State Harmonics beyond the Atomic Limit". In: *Nature* 534.7608 (June 2016), pp. 520–523. ISSN: 0028-0836, 1476-4687. DOI: 10/f8rt6g.

[95] Yong Sing You, David A. Reis, and Shambhu Ghimire. "Anisotropic High-Harmonic Generation in Bulk Crystals". In: *Nature Phys* 13.4 (Apr. 2017), pp. 345–349. ISSN: 1745-2473, 1745-2481. DOI: 10/gmkfxh.

[96] Stian Astad Sørngård, Sigrid Ina Simonsen, and Jan Petter Hansen. "High-Order Harmonic Generation from Graphene: Strong Attosecond Pulses with Arbitrary Polarization". In: *Phys. Rev. - At. Mol. Opt. Phys.* 87.5 (2013), pp. 1–5. ISSN: 10502947. DOI: 10/gj7bc9.

[97] Larisa A. Chizhova, Florian Libisch, and Joachim Burgdörfer. "High-Harmonic Generation in Graphene: Interband Response and the Harmonic Cutoff". In: *Phys. Rev. B* 95.8 (2017), pp. 1–9. ISSN: 24699969. DOI: 10/gj7bdb.

[98] Shunsuke A. Sato et al. "High-Order Harmonic Generation in Graphene: Nonlinear Coupling of Intraband and Interband Transitions". In: *Phys. Rev. B* 103.4 (Jan. 2021), p. L041408. ISSN: 2469-9950, 2469-9969. DOI: 10/gk3h2r.

[99] Sung-Young Hong et al. "Optical Third-Harmonic Generation in Graphene". In: *Phys. Rev. X* 3.2 (June 2013), p. 021014. ISSN: 2160-3308. DOI: 10/gdq72p.

[100] Nardeep Kumar et al. "Third Harmonic Generation in Graphene and Few-Layer Graphite Films". In: *Phys. Rev. B - Condens. Matter Mater. Phys.* 87.12 (2013), pp. 1–5. ISSN: 10980121. DOI: 10/gdq72q. arXiv: 1301.1042.

[101] Vl A. Margulis, E. E. Muryumin, and E. A. Gaiduk. "Frequency Dependence of Optical Third-Harmonic Generation from Doped Graphene". In: *Phys. Lett. Sect. Gen. At. Solid State Phys.* 380.1-2 (2016), pp. 304–310. ISSN: 03759601. DOI: 10/gj7bdg.

[102] Naotaka Yoshikawa, Tomohiro Tamaya, and Koichiro Tanaka. "High-Harmonic Generation in Graphene Enhanced by Elliptically Polarized Light Excitation". In: *Science* 356.6339 (May 2017), pp. 736–738. ISSN: 0036-8075, 1095-9203. DOI: 10/f98nv9.

[103] Hassan A. Hafez et al. "Terahertz Nonlinear Optics of Graphene: From Saturable Absorption to High-Harmonics Generation". In: *Adv. Optical Mater.* 8.3 (Feb. 2020), p. 1900771. ISSN: 2195-1071, 2195-1071. DOI: 10/ghgkwx.

[104] Jan-Christoph Deinert et al. "Grating-Graphene Metamaterial as a Platform for Terahertz Nonlinear Photonics". In: *ACS Nano* 15.1 (Jan. 2021), pp. 1145–1154. ISSN: 1936-0851, 1936-086X. DOI: 10/gk4pd7.

[105] Sergey Kovalev et al. "Electrical Tunability of Terahertz Nonlinearity in Graphene". In: *Sci. Adv.* 7.15 (Apr. 2021), eabf9809. ISSN: 2375-2548. DOI: 10/gmkd2w.

[106] Frank H.L. Koppens, Darrick E. Chang, and F. Javier García De Abajo. "Graphene Plasmonics: A Platform for Strong Light-Matter Interactions". In: *Nano Lett.* 11.8 (Aug. 2011), pp. 3370–3377. ISSN: 15306984. DOI: 10/frxnzt. arXiv: 1104.2068.

[107] Victor W. Brar et al. "Highly Confined Tunable Mid-Infrared Plasmonics in Graphene Nanoresonators". In: *Nano Lett.* 13.6 (2013), pp. 2541–2547. ISSN: 15306984. DOI: 10/f44d6k.

[108] Joel D. Cox and F. Javier García De Abajo. "Electrically Tunable Nonlinear Plasmonics in Graphene Nanoislands". In: *Nat. Commun.* 5 (Dec. 2014), p. 5725. ISSN: 20411723. DOI: 10/f6v39q. arXiv: 1410.2111.

[109] Chun Hung Lui et al. "Ultrafast Photoluminescence from Graphene". In: *Phys. Rev. Lett.* 105.12 (Sept. 2010), p. 127404. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.105.127404.

[110] Farhan Rana et al. "Carrier Recombination and Generation Rates for Intravalley and Intervalley Phonon Scattering in Graphene". In: *Phys. Rev. B - Condens. Matter Mater. Phys.* 79.11 (2009), pp. 1–5. ISSN: 10980121. DOI: 10/bgdhqp.

[111] Christoph Krüger et al. "Birefringence Measurements on Crystalline Silicon". In: *Class. Quantum Gravity* 33.1 (2015), p. 015012. DOI: 10/gmhj3r.

[112] Bahaa E. A. Saleh and Malvin Carl Teich. *Fundamentals of Photonics*. John Wiley & Sons, Mar. 2019. ISBN: 978-1-118-77006-1.

[113] Lukas Novotny and Bert Hecht. *Principles of Nano-Optics*. Cambridge: Cambridge University Press, 2006. DOI: 10.1017/CBO9780511813535.

[114] S. A. Mikhailov. "Quantum Theory of the Third-Order Nonlinear Electrodynamic Effects of Graphene". In: *Phys. Rev. B* 93.8 (2016), pp. 1–29. ISSN: 24699969. DOI: 10.1103/PhysRevB.93.085403. arXiv: 1506.00534.

[115] Justus Bohn et al. "All-Optical Switching of an Epsilon-near-Zero Plasmon Resonance in Indium Tin Oxide". In: *Nat. Commun.* 12.1 (Feb. 2021), p. 1017. ISSN: 2041-1723. DOI: 10/gj699k.

[116] E. N. Glezer et al. "Three-Dimensional Optical Storage inside Transparent Materials". In: *Opt. Lett.* 21.24 (Dec. 1996), p. 2023. ISSN: 0146-9592. DOI: 10/fnznhw.

[117] D. Cotter et al. "Nonlinear Optics for High-Speed Digital Information Processing". In: *Science* 286.5444 (Nov. 1999), pp. 1523–1528. ISSN: 00368075. DOI: 10/bzmnx3.

[118] John C. Howell et al. "Realization of the Einstein-Podolsky-Rosen Paradox Using Momentumand Position-Entangled Photons from Spontaneous Parametric down Conversion". In: *Phys. Rev. Lett.* 92.21 (May 2004), p. 210403. ISSN: 00319007. DOI: 10/dndkq4.

[119] Jonathan Leach et al. "Quantum Correlations in Optical Angle-Orbital Angular Momentum Variables". In: *Science* 329.5992 (Aug. 2010), pp. 662–665. ISSN: 00368075. DOI: 10/c6b9vf.

[120] Paolo Minzioni et al. "Roadmap on All-Optical Processing". In: *J. Opt. U. K.* 21.6 (June 2019), p. 063001. ISSN: 20408986. DOI: 10/gh84bb.

[121] Sergey V. Makarov et al. "Light-Induced Tuning and Reconfiguration of Nanophotonic Structures". In: *Laser Photonics Rev.* 11.5 (2017), pp. 1–25. ISSN: 18638899. DOI: 10/ghfsb9.

[122] Xinxiang Niu et al. "Epsilon-Near-Zero Photonics: A New Platform for Integrated Devices". In: *Adv. Opt. Mater.* 6.10 (2018), pp. 1–36. ISSN: 21951071. DOI: 10/dwjt.

[123] David A.B. Miller. "Are Optical Transistors the Logical next Step?" In: *Nat. Photonics* 4.1 (2010), pp. 3–5. ISSN: 17494885. DOI: 10/b7rjmm.

[124] Eyal Feigenbaum, Kenneth Diest, and Harry A. Atwater. "Unity-Order Index Change in Transparent Conducting Oxides at Visible Frequencies". In: *Nano Lett.* 10.6 (2010), pp. 2111–2116. ISSN: 15306984. DOI: 10/ckcd3r.

[125] Alessandro Ciattoni, Carlo Rizza, and Elia Palange. "All-Optical Active Plasmonic Devices with Memory and Power-Switching Functionalities Based on ϵ-near-Zero Nonlinear Metamaterials". In: *Phys. Rev. - At. Mol. Opt. Phys.* 83.4 (2011), pp. 1–8. ISSN: 10502947. DOI: 10/cvsj4p.

[126] Qiangbing Guo et al. "A Solution-Processed Ultrafast Optical Switch Based on a Nanostructured Epsilon-Near-Zero Medium". In: *Adv. Mater.* 29.27 (2017), pp. 1–7. ISSN: 15214095. DOI: 10/f95qjv.

[127] Georgios Sinatkas and Emmanouil E. Kriezis. "Silicon-Photonic Electro-Optic Phase Modulators Integrating Transparent Conducting Oxides". In: *IEEE J. Quantum Electron.* 54.4 (2018). ISSN: 00189197. DOI: 10/gj7bnj.

[128] Mohammad Taghinejad et al. "Hot-Electron-Assisted Femtosecond All-Optical Modulation in Plasmonics". In: *Adv. Mater.* 30.9 (2018), pp. 1–7. ISSN: 15214095. DOI: 10/gcvfvr.

[129] Xiaoming Qiu et al. "Multi-Layer MOS Capacitor Based Polarization Insensitive Electro-Optic Intensity Modulator". In: *Opt. Express* 26.11 (2018), p. 13902. ISSN: 1094-4087. DOI: 10/gj7bnb.

[130] Tobias Tyborski et al. "Ultrafast Nonlinear Response of Bulk Plasmons in Highly Doped ZnO Layers". In: *Phys. Rev. Lett.* 115.14 (2015), pp. 1–5. ISSN: 10797114. DOI: 10/f8bfq6.

[131] Ray Secondo, Jacob Khurgin, and Nathaniel Kinsey. "Absorptive Loss and Band Non-Parabolicity as a Physical Origin of Large Nonlinearity in Epsilon-near-Zero Materials". In: *Opt. Mater. Express* 10.7 (2020), p. 1545. ISSN: 2159-3930. DOI: 10/gj7bm8. arXiv: 2004.00134.

[132] Arrigo Calzolari, Alice Ruini, and Alessandra Catellani. "Transparent Conductive Oxides as Near-IR Plasmonic Materials: The Case of Al-Doped ZnO Derivatives". In: *ACS Photonics* 1.8 (Aug. 2014), pp. 703–709. ISSN: 2330-4022. DOI: 10/f6fhx4.

[133] Jacob B. Khurgin, Matteo Clerici, and Nathaniel Kinsey. "Fast and Slow Nonlinearities in Epsilon-Near-Zero Materials". In: *Laser Photonics Rev.* 2000291 (Dec. 2020), p. 2000291. ISSN: 1863-8880. DOI: 10/gj7bmw.

[134] Peijun Guo et al. "Large Optical Nonlinearity of ITO Nanorods for Sub-Picosecond All-Optical Modulation of the Full-Visible Spectrum". In: *Nat. Commun.* 7.1 (Nov. 2016), p. 12892. ISSN: 20411723. DOI: 10/f87vcj.

[135] Simon Vassant et al. "Berreman Mode and Epsilon near Zero Mode". In: *Opt. Express* 20.21 (2012), p. 23971. ISSN: 1094-4087. DOI: 10/dfqk.

[136] Ward D. Newman et al. "Ferrell-Berreman Modes in Plasmonic Epsilon-near-Zero Media". In: *ACS Photonics* 2.1 (2015), pp. 2–7. ISSN: 23304022. DOI: 10/gj7bnc. arXiv: 1505.06180.

[137] Jongbum Kim et al. "Dynamic Control of Nanocavities with Tunable Metal Oxides". In: *Nano Lett.* 18.2 (2018), pp. 740–746. ISSN: 15306992. DOI: 10/gcvv9r.

[138] Martin Van Exter and Ad Lagendijk. "Ultrashort Surface-Plasmon and Phonon Dynamics". In: *Phys. Rev.* 60.1 (1988), pp. 49–52.

[139] A. Devižis, V. Vaičikauskas, and V. Gulbinas. "Ultrafast Pump-Probe Surface Plasmon Resonance Spectroscopy of Thin Gold Films". In: *Appl. Opt.* 45.11 (2006), pp. 2535–2539. ISSN: 15394522. DOI: 10/ccxnwg.

[140] Israel De Leon et al. "Measurement of the Complex Nonlinear Optical Response of a Surface Plasmon-Polariton". In: *Opt. Lett.* 39.8 (2014), p. 2274. ISSN: 0146-9592. DOI: 10/gj7bkj.

[141] Klaus Ellmer and Rainald Mientus. "Carrier Transport in Polycrystalline ITO and ZnO:Al II: The Influence of Grain Barriers and Boundaries". In: *Thin Solid Films*. 5th International Symposium on Transparent Oxide Thin Films for Electronics and Optics 516.17 (July 2008), pp. 5829–5835. ISSN: 0040-6090. DOI: 10/ccdwf6.

[142] Fuzi Yang, J. R. Sambles, and G. W. Bradberry. "Long-Range Surface Modes Supported by Thin Films". In: *Phys. Rev. B* 44.11 (1991), pp. 5855–5872. ISSN: 01631829. DOI: 10/cxj72s.

[143] Michael Scalora et al. "Electrodynamics of Conductive Oxides: Intensity-dependent Anisotropy, Reconstruction of the Effective Dielectric Constant, and Harmonic Generation". In: *Phys. Rev. A* 101.5 (May 2020), p. 053828. ISSN: 2469-9926. DOI: 10/gj7bnq. arXiv: 2003.01059.

[144] Ll. G. Chambers, Carl M. Bender, and Steven A. Orszag. "Advanced Mathematical Methods for Scientists and Engineers". In: *Math. Gaz.* 63.424 (1979), p. 139. ISSN: 00255572. DOI: 10/b769nn.

[145] H. I. Elim, W. Ji, and F. Zhu. "Carrier Concentration Dependence of Optical Kerr Nonlinearity in Indium Tin Oxide Films". In: *Appl. Phys. B Lasers Opt.* 82.3 (2006), pp. 439–442. ISSN: 09462171. DOI: 10/d4sdjv.

[146] Baoli Yao, Liyong Ren, and Xun Hou. "Z-Scan Theory Based on a Diffraction Model". In: *J. Opt. Soc. Am. B* 20.6 (2003), p. 1290. ISSN: 0740-3224. DOI: 10/dzc8pb.

[147] Andrea Schirato et al. "Transient Optical Symmetry Breaking for Ultrafast Broadband Dichroism in Plasmonic Metasurfaces". In: *Nat. Photonics* 14.12 (Dec. 2020), pp. 723–727. ISSN: 1749-4885, 1749-4893. DOI: 10/gj7bm2.

[148] Kuidong Wang et al. "High Contrast, Femtosecond Light Polarization Manipulation in Epsilon-near-Zero Material Coupled to a Plasmonic Nanoantenna Array". In: *ACS Photonics* 8.9 (Sept. 2021), pp. 2791–2799. ISSN: 2330-4022, 2330-4022. DOI: 10/gmvhhj.

[149] Maxim R. Shcherbakov et al. "Ultrafast All-Optical Switching with Magnetic Resonances in Nonlinear Dielectric Nanostructures". In: *Nano Lett.* 15.10 (Oct. 2015), pp. 6985–6990. ISSN: 1530-6984. DOI: 10/752.

[150] Gustavo Grinblat et al. "Ultrafast Sub–30-Fs All-Optical Switching Based on Gallium Phosphide". In: *Sci. Adv.* 5.6 (June 2019), eaaw3262. ISSN: 2375-2548. DOI: 10/gkg2bt.

[151] Masaaki Ono et al. "Ultrafast and Energy-Efficient All-Optical Switching with Graphene-Loaded Deep-Subwavelength Plasmonic Waveguides". In: *Nat. Photonics* 14.1 (Jan. 2020), pp. 37–43. ISSN: 1749-4893. DOI: 10/ggjgpv.

[152] Wenlan Chen et al. "All-Optical Switch and Transistor Gated by One Stored Photon". In: *Science* 341.6147 (Aug. 2013), pp. 768–770. ISSN: 0036-8075, 1095-9203. DOI: 10/f466fc.

[153] Anton V. Zasedatelev et al. "A Room-Temperature Organic Polariton Transistor". In: *Nat. Photonics* 13.6 (2019), pp. 378–383. ISSN: 17494893. DOI: 10/gf256c.

[154] T. Gustafson, P. Kelly, and R. Fisher. "Subpicosecond Pulse Generation Using the Optical Kerr Effect". In: *IEEE J. Quantum Electron.* 5.6 (June 1969), pp. 325–325. ISSN: 1558-1713. DOI: 10/cqzs4m.

[155] S. A. Planas et al. "Spectral Narrowing in the Propagation of Chirped Pulses in Single-Mode Fibers". In: *Opt. Lett., OL* 18.9 (May 1993), pp. 699–701. ISSN: 1539-4794. DOI: 10/bcrmkf.

[156] C.A. Brackett. "Dense Wavelength Division Multiplexing Networks: Principles and Applications". In: *IEEE J. Sel. Areas Commun.* 8.6 (Aug. 1990), pp. 948–964. ISSN: 1558-0008. DOI: 10/bmf4wp.

[157] Dimitrios L. Sounas, Christophe Caloz, and Andrea Alù. "Giant Non-Reciprocity at the Subwavelength Scale Using Angular Momentum-Biased Metamaterials". In: *Nat Commun* 4.1 (Sept. 2013), p. 2407. ISSN: 2041-1723. DOI: 10/gbdxp9.

[158] Eran Lustig, Yonatan Sharabi, and Mordechai Segev. "Topological Aspects of Photonic Time Crystals". In: *Optica, OPTICA* 5.11 (Nov. 2018), pp. 1390–1395. ISSN: 2334-2536. DOI: 10/gmvbdg.

[159] S. A. Akhmanov, A. P. Sukhorukov, and A. S. Chirkin. "Nonstationary Phenomena and Space-Time Analogy in Nonlinear Optics". In: *Sov. Phys. JETP* 28.4 (1969), pp. 748–757.

[160] John C. AuYeung. "Phase-Conjugate Reflection from a Temporal Dielectric Boundary". In: *Opt. Lett., OL* 8.3 (Mar. 1983), pp. 148–150. ISSN: 1539-4794. DOI: 10/dmxcmm.

[161] J. T. Mendonca. *Theory of Photon Acceleration*. CRC Press, Dec. 2000. ISBN: 978-1-4200-3327-4.

[162] W. R. Donaldson, G. P. Agrawal, and B. W. Plansinis. "What Is the Temporal Analog of Reflection and Refraction of Optical Beams?" In: *Phys. Rev. Lett.* 115.18 (Oct. 2015), p. 183901. DOI: 10/ghgtzs.

[163] J. Paul et al. "Two-Beam Coupling by a Hot Electron Nonlinearity". In: *Opt. Lett., OL* 46.2 (Jan. 2021), pp. 428–431. ISSN: 1539-4794. DOI: 10/gj7bmr.