THz polaritonics:
Optical THz generation, detection, and control on a chip

by

Christopher A. Werley

B.A. Chemistry, Stanford University (2005)

Submitted to the Department of Chemistry
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2012

Signature of Author ____
Department of Chemistry
April 27, 2012

Certified by ____
Keith A. Nelson
Professor of Chemistry
Thesis Supervisor

Accepted by____
Robert W. Field
Professor of Chemistry
Chairman, Department Committee on Graduate Students

This doctoral thesis has been examined by a committee of the Department of Chemistry as follows:

Professor Moungi G. Bawendi ____
Chairperson

Professor Keith A. Nelson ____
Thesis Supervisor

Professor Sylvia T. Ceyer ____

THz polaritonics:
Optical THz generation, detection, and control on a chip

by
Christopher A. Werley

Submitted to the Department of Chemistry on April 27, 2012 in partial fulfillment
of the requirements for the degree of Doctor of Philosophy in Chemistry

ABSTRACT

The THz polaritonics system is an on-chip platform for THz generation, detection, and control. THz-frequency electromagnetic waves are generated directly in a thin slab of lithium niobate crystal where they can be amplified and guided. Time-resolved phase-sensitive imaging lets us capture movies of THz waves as they propagate at the light-like speeds. I developed polaritonics methodologies and used the platform to study various microstructures interacting with THz waves.

I began technique development by deriving a quantitative model explaining THz wave propagation in an anisotropic slab waveguide. From this model, I extracted the frequency-dependent wave velocity and used this knowledge to design an optical pumping geometry that phase-matches and coherently amplifies a selected THz frequency. This geometry can generate high-amplitude THz waves with a tunable center frequency and bandwidth. Much like the generation, the detection was also revamped. New optical designs, acquisition procedures, and hardware let us quantitatively measure THz field strengths. The image resolution was improved from ~50 μm to 1.5 μm, and measurement noise was reduced by 50-fold.

Using the improved generation and detection methods, we studied two classes of microstructures: laser-machined air gaps and deposited metal antennas. Air gaps cut into the lithium niobate slab effectively reflect, waveguide, and scatter THz waves. We fabricated structures that demonstrate wave phenomena such as diffraction and interference and captured movies of THz waves interacting with these structures. The movies can be useful tools in lectures on electromagnetism because they beautifully illustrate the fundamental effects and bring cutting-edge research into the classroom. In addition to air structures, we studied metal antennas, which are interesting because of their ability to enhance optical fields and localize electromagnetic waves well below the diffraction limit. The polaritonics platform enabled incisive study of fundamental antenna behavior and scaling because we could map the antenna's near-field with $\lambda/100$ spatial resolution and we could quantify large THz electric field amplitudes and enhancements in a deeply sub-wavelength gap between antennas. Antenna field enhancement is already facilitating nonlinear THz research, and the polaritonics platform will enable improved study of photonic systems such as metamaterials and photonic crystals.

Thesis Supervisor: Keith A. Nelson
Title: Professor of Chemistry

Acknowledgements

Many people have helped me during my thesis work in many different ways. Some contributed directly to the work that is presented, some helped me grow as a scientist, and some were friends and companions. Many of the people I was closest to filled more than one of these roles. I want to extend a heart-felt thank you to all these people; you have made my PhD productive, educational, and fun.

To start, I want to acknowledge the people who contributed directly to the work. First and foremost is my advisor Keith Nelson. He is a font of information and ideas, and his excitement about science in infectious and inspiring. Equally important to those of us in the trenches, his group is happy, cohesive and collaborative, and I am extremely grateful to have Keith as a role model for my future endeavors.

Next I want to thank the older students who showed me the ropes when I first joined the lab: Eric Statz, Kung-Hsuan Lin, and Taeho Shin. They were helpful and patient and made me happy to go to work each day. Without their instruction and advice, I would no doubt still be bumbling around the lab today and would likely have managed to put out an eye with the laser. Generating narrowband THz radiation in the waveguide with a tilted pulse front was Kung-Hsuan's idea, and he built the first generation of the instrument.

The next people who deserve recognition are the younger students who worked under me. The first was Ryan Tait, an undergrad who resurrected the laser-machining setup and fabricated many structures. The work in chapter 5 was done closely with Ryan. He was a quick study, and when I was teaching Ryan I truly understood many of the concepts for the first time. Next is Steph Teo, who was the first graduate student to join my project. Steph was a life saver. Without her help the quality of the work would have suffered, the scope of the projects would have been limited, my graduation would likely have been delayed, and work just wouldn't have been as fun. Steph helped with the DAQ card work, and was intimately involved with the antenna study, the most exciting and important section of my thesis. After Steph, three more students joined the polaritonics project: Ben Ofori-Okai, Prasahnt Sivarajah, and LongFang Ye. All three picked up on the projects quickly and are great to work with. As of the publication of this thesis, Steph, Ben, Prasahnt, and LongFang and I are working on a broad set of exciting new experiments.

The final group who contributed directly to the work are collaborators. We got to share knowledge and expertise and combine skill sets to do experiments that would not otherwise be possible. First was Qiang Wu, a professor from Nankai University who visited the group for a year and continued the collaboration after he returned to his lab. Qiang is a great guy, and we worked hard to develop improved THz imaging techniques. After he returned to China, he and his student Chengliang Yang worked with me to understand THz wave propagation in the anisotropic lithium niobate slab waveguide. I also collaborated with August Dorn from the Bawendi Lab, who made the phase mask for phase contrast imaging. August was good to work with because he always carefully considered projects before charging down a path that could be time consuming and fruitless. The final group of collaborators was Rick Averitt from Boston University and his students Drew Strikwerda and Kebin Fan. They were all great to work with. As experts on THz metamaterials, they taught me a lot about metal microstructures and their interaction with electromagnetic waves. Drew did the simulations for the antenna work and Kebin did the fabrication.

Moving on from people who contributed directly to the work, there are many people who contributed to my academic development. I can't name everyone, as there were many friends and classmates with whom I had valuable discussions. In particular I would like to thank the Nelson group for being such a good place to work. Everyone was always willing to share their equipment, expertise, insights, and time. I can't come close to counting the number of discussions on scientific topics that we engaged in (there were some every day), and it was often during these discussions that I really advanced my understanding. Thank you all.

The last major group of people I want to thank are friends and family. There are too many to name everyone individually, so instead I will name categories of people. The first are friends from my pre-graduate life, both Los Alamos and Stanford, who were always excited to hear about my latest grad school adventures. Next, my friends from mechanical engineering (and the other hangers-on like myself). We played some great games and had amazing ski trip adventures. Next are friends from running and working out, movie night, and volleyball. These recurring activities and the people I did them with formed the backbone of my social life outside of school, and it was a wonderful backbone. Finally, there are friends from the department and the Nelson lab whom I saw every day. It is these people, even more than the work, who make grad school such a fun place to be.

The next group of obvious importance is my family. My parents, step parents, brother and sister, grandparents, aunts, uncles, and cousins are all an important part of my life. My parents in particular did so much for me through the years. They supported me in whatever I found interesting, and above all encouraged my curiosity in the world around me. I think this curiosity has been one of my most important assets during graduate school.

Finally, I want to thank Ashley. She has been my best friend and companion for many years both during and before my PhD. I am grateful for all she's done for me and for her patience enduring the long hours and busy schedule of a graduate student.

To all those I mentioned, and any I may have forgotten, my sincerest thank you.
-Kit

Outline

IV DEVELOPMENTS IN POLARITONICS METHODOLOGY: THZ GENERATION, DAQ CARD DETECTION, AND THZ IMAGING 75

V THZ WAVES IN CLASSIC EXPERIMENTAL GEOMETRIES 121

Chapter I

Introduction to THz Spectroscopy and Polaritonics

Terahertz (THz) radiation is the part of the electromagnetic spectrum that lies between the infrared and microwave, and it is typically associated with frequencies between 0.1 – 10 THz (wavelength: 30 – 3000 μm, energy: 3 – 300 cm^{-1} = 0.4 – 40 meV). THz radiation is an important tool in basic science because it can be used to interrogate many THz-frequency phenomena including molecular rotations in a gas, vibrations in a molecular crystal (like sugar), and electronic transitions in nanostructures such as quantum wells or quantum dots. It also can be used to probe a variety of more exotic condensed phase phenomena including Cooper pairs, polarons, and magnons. In addition, THz has practical applications as it is at the frontier in high-speed electronics and optical communications, and may prove useful as a replacement for x-ray scanners in airports.

Because THz frequencies lie above what is easily accessible with electronics and below what is easily accessible with tunable lasers, THz technology is less advanced than technology in other regions of the electromagnetic spectrum. In the late 1980s, Auston and coworkers developed two ultrafast methods to generate and detect THz radiation: photoconductive switches and optical rectification. In a photoconductive switch, free carriers (electrons and holes) excited by an ultrafast optical pulse are accelerated by a DC electric field, and the accelerated charges emit THz radiation [Auston 1984]. In optical rectification, an ultrafast optical pulse is sent into a nonlinear crystal and a difference-frequency process generates THz [Auston 1988]. These two ultrafast methods spurred rapid development in THz research and technology. Today, THz spectrometers are available commercially and there are multiple international research conferences attended by hundreds of groups. Modern THz research includes such topics as THz waveguide spectroscopy [Laman 2008], tomography [Abraham 2011], and nonlinear spectroscopy [Hebling 2010].

In spite of all these advances, THz methods still lag behind their microwave and optical analogs. The generation of broadband THz pulses still requires ultrafast lasers, and the generation of intense THz radiation requires an expensive amplified ultrafast laser system and significant technical expertise. Although this is starting to change, detection usually requires the time-consuming process of scanning a delay stage to record the full THz field temporal profile, and Fourier transforming to retrieve the spectral information. In addition to the challenges of generation and detection, control and guidance of THz radiation are more difficult than in other spectral regions. Free-space THz radiation can be guided by optics (flat and curved mirrors, lenses, etc.) much like visible light, but the THz is much harder to work with because of its longer wavelength. Electromagnetic beams are described by the Gaussian beam formalism [Saleh 2007]. In this formalism, the most important parameter is the ratio of the beam diameter d to the wavelength λ. For visible beams, a "well-collimated" beam is typically at least 1 mm in diameter, so $d/\lambda \sim 2000$. For a THz beam ($\lambda = 300$ μm at 1 THz) to have this same ratio, the beam diameter must be 60 cm, which is impractical for tabletop experiments. As a result it is very difficult both to send THz beams long distances and to focus them tightly. As an alternative to optics and free-space beams, one might try to control and guide THz waves using metal wires, resistors, capacitors, etc. as is done for electronics, but this is also is challenging. Most electrical components cannot respond at THz frequencies and high damping in metals severely limits many applications.

To address some of these issues, the Nelson group introduced the "polaritonics" platform, in which the THz wave is confined within a dielectric material for the entire experiment [Feurer 2007]. This avoids the damping that plagues metallic components and circumvents the problem of beam divergence by waveguiding the THz. The dielectrics of choice have been the ferroelectric crystals lithium niobate ($LiNbO_3$, often abbreviated LN) and lithium tantalate ($LiTaO_3$, or LT) whose high indices of refraction at THz frequencies, 5.11 and 6.15 respectively, make strong waveguiding possible. Performing optical experiments "on chip" in this way enables compact experimental design, requires only small amounts of sample material, and facilitates the interfacing between optical components and electronics. Typical on-chip experiments also present challenges, because it is usually difficult to couple light onto the chip and couple it back out for detection. The polaritonics platform avoids these problems by directly generating, amplifying, and detecting the THz wave on chip. This diverse set of capabilities is made possible by the unique optical properties of $LiNbO_3$ and $LiTaO_3$. These crystals are transparent at both optical and THz frequencies, and they have large nonlinear optical coefficients which make interconversion between optical and THz frequencies possible. In addition both crystals have strong electro-optic responses, which means that the THz wave changes the crystal's index of refraction. This change can then be detected by an optical probe pulse.

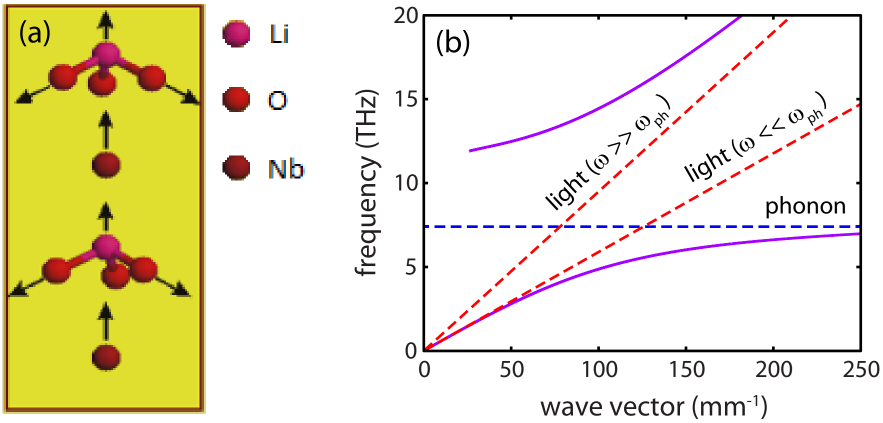

Figure 1.1 | Phonon character of THz waves. (a) The atomic motions for the relevant phonon mode in lithium niobate. (b) Dispersion curves for a phonon polariton wave in $LiNbO_3$ (purple). Dashed red lines are the dispersion curves for light at frequencies well above the phonon resonance (slope $= c/n_\infty$) and well below the phonon resonance (slope $= c/n_0$). Dashed blue is the phonon dispersion curve (slope ~ 0).

The Nelson group first became interested in polaritonics through its study of phonon-polaritons waves, which are mixtures of vibrational waves (phonons) and electromagnetic waves (polaritons). In $LiNbO_3$ and $LiTaO_3$, all THz frequency waves have some at least some phonon character and some electromagnetic character. In $LiNbO_3$, the relevant phonon resonance is at 7.6 THz; the atomic motions associated with this mode are shown in Fig. 1.1(a). This resonance is the ferroelectric "soft mode" and is largely responsible for the strong nonlinear response of this material [Brennan 1997]. It is this strong nonlinear response that enables efficient THz generation and detection and makes $LiNbO_3$ such a good substrate for the polaritonics platform. The dispersion curve of bulk $LiNbO_3$ is shown in Fig. 1.1(b) [Feurer 2007]. In all the measurements in my thesis, we are concerned with the lower polariton branch, the purple curve below 7.6 THz. At low frequencies, it lies along the dashed red line, which shows the behavior of light in the absence of a phonon resonance. At higher frequencies, it lies along the dashed blue line, which shows the behavior of the phonon mode if it did not cross the light line. In this thesis, all experiments are concerned with frequencies below 2.5 THz where the behavior is very light-like. Even at these frequencies, however, much of the energy of the phonon-polariton wave is stored in the ionic displacement.

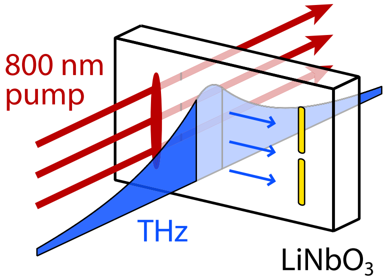

Figure 1.2 | The experimental pumping geometry. A THz wave is generated by an ultrafast optical pump pulse and guided down the $LiNbO_3$ slab until its evanescent field interacts with a sample deposited on the surface (a gold antenna is shown here as an example). Alternatively, the wave could interact with a structure machined into the crystal slab, or a sample deposited in a hole machined into the slab.

The phonon character of the propagating wave underlies the microscopic mechanism for THz generation. When an intense, ultrafast optical pump pulse enters the $LiNbO_3$ crystal, it drives a THz response via the nonlinear process of impulsive stimulated Raman scattering (ISRS) [Dougherty 1992]. We use this to generate THz directly in the polaritonics chip. The experimental geometry for the polaritonics platform is illustrated in Fig. 1.2. The pump pulse passes through a $LiNbO_3$ slab, a 30 – 50 µm thick, free-standing single crystal that is thin enough to act as a waveguide for THz radiation. The generated THz wave propagates orthogonal to the pump pulse, and it is guided down the slab where it can interact with a machined air-gap, metallic microstructure, or sample deposited on the surface (see Fig. 1.2). During my thesis work, I improved understanding of THz behavior on-chip and introduced and improved many on-chip techniques. This included an improved understanding of how waves propagate in the crystal slab [Yang 2010], tunable THz generation and THz amplification [Lin 2009], improved THz imaging and detection [Wu 2009, Werley 2010, Werley 2011], diffractive and waveguiding elements in the slab [Werley AJP 2012], and deposition of metallic components on the slab surface with large field enhancements and sub-wavelength localization [Werley OE 2012]. The following paragraphs offer brief summaries of each project.

To understand how a THz wave evolves after it is generated, I developed a simple analytical theory for wave propagation in an anisotropic slab waveguide and found that it fully matched experimentally measured dispersion curves [Yang 2010] (see chapter 2). Figure 1.3 shows the dispersion curves (data underneath, theory plotted on top) when the optic axis of the waveguide is rotated 20° relative to the THz propagation direction. The excellent agreement at this and all other angles demonstrates that the theory fully explains waveguiding in the anisotropic slab and enables an intuitive understanding of the behavior. In bulk anisotropic material, waves are divided into two normal modes, ordinary waves and extraordinary waves. In an isotropic waveguide there are also two uncoupled eigenmodes, the transverse electric (TE) and transverse magnetic (TM) modes. In the anisotropic waveguide, however, all the modes couple together and the new eigenmodes of the system are neither purely TE nor TM and also not purely ordinary or extraordinary. The coupling affects the mode profiles, dispersion curves, and effective refractive indices in a fundamental and significant way, which can all be predicted by the theory.

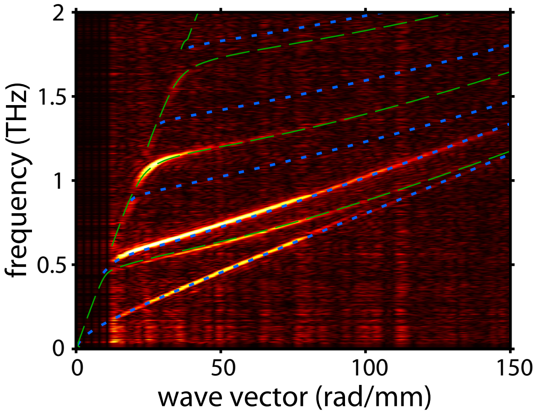

Figure 1.3 | Anisotropic waveguide modes. Analytical theory (overlaid, dashed green lines are TM-like modes and dotted blued lines are TE-like modes) is in excellent agreement with experimentally measured dispersion curves (yellow continuous lines) in the anisotropic $LiNbO_3$ slab. In this case the optic axis of the crystal is rotated 20° away from the THz propagation direction. This theory gives us full knowledge of the group and phase velocities of the THz waves.

One of the outputs from the theory of waveguiding is the frequency-dependent phase velocity of the THz waves. The phase velocity in the waveguide has a very strong frequency dependence, transitioning from the speed of light, c, at low frequencies to c/n_{LN} at high frequencies, with n_{LN} the index of refraction in bulk lithium niobate. Because of this large variation in phase velocities, it is possible to velocity match, and thus coherently amplify, a single THz frequency using an optical pulse whose intensity front has been tilted by a diffraction grating [Lin 2009] (see section 4A). Generating pulses with high spectral brightness in this way enables strong and selective pumping of a single THz resonance without exciting neighboring modes. By tuning the tilt angle, it is possible to tune the frequency of the generated THz wave (see Fig. 1.4).

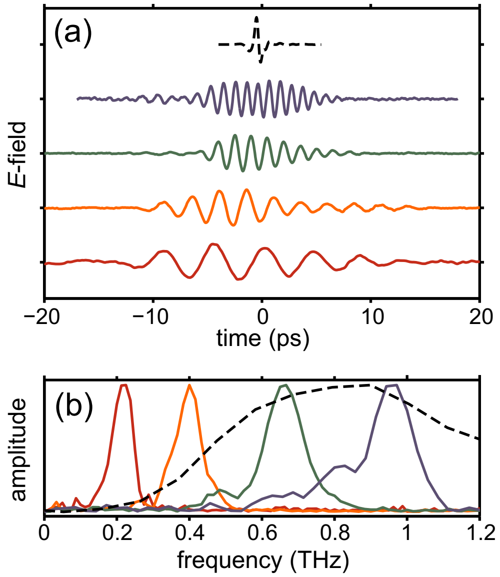

Figure 1.4 | Tunable narrowband generation. (a) The electric field time traces of a broadband (dashed black) and various narrowband THz pulses made in the $LiNbO_3$ waveguide with a tilted optical intensity front. (b) The corresponding spectra, spanning about a decade in frequency.

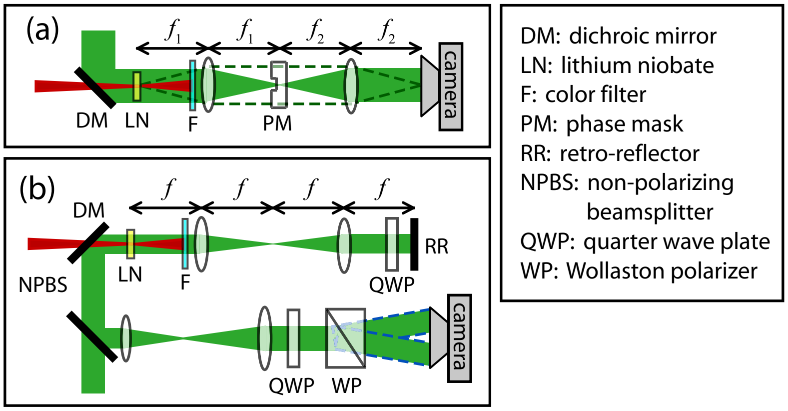

Figure 1.5 | THz wave imaging setups. The experimental setups used to image THz waves. (a) Phase contrast imaging. The diffracted light is phase-shifted by 90º relative to the 0^{th} order beam by a phase mask in the Fourier plane, leading to interference and thus phase-to-amplitude conversion in the image plane. The resolution is 1.5 μm, ~λ/100 for the THz frequencies we typically use, about 50x better than the previous imaging design. (b) Polarization gating imaging, a technique complementary to phase contrast, uses changes in the polarization state of the probe beam to detect the THz wave, which enables balanced detection. The resolution is somewhat coarser (~5 μm), but the signal-to-noise ratio is improved by more that 10x relative to the previous method.

Another major thrust in my thesis was to improve detection of THz waves in the $LiNbO_3$ slab [Wu 2009, Werley 2010, Werley 2011] (see sections 4C & D). A key capability of the polaritonics chip is the ability to measure the full E-field profile of the THz wave at each point in time as the wave propagates. This information can be played back as a video showing interactions between the wave and structures in or on the chip, providing exceptional insight into the behavior of photonic components. This is possible because $LiNbO_3$ is an electro-optic crystal, so the THz field $E(x,y)$ induces a change in the index of refraction $\Delta n(x,y)$, which shifts the phase $\Delta\varphi(x,y)$ of the expanded optical probe beam used to detect the THz wave: $E(x,y) \rightarrow \Delta n(x,y) \rightarrow \Delta\varphi(x,y)$. We then use a phase sensitive imaging technique to record the induced shift, and step the time delay between pump and probe pulses to build up the full evolution of the wave. I developed two complementary, phase sensitive imaging methods: phase contrast imaging and polarization gating imaging. Overall, the resolution was improved from ~25 μm to 1.5 μm, better than $\lambda/100$ resolution for many of the frequencies we use, and the signal-to-noise ratio was increased by more than 10-fold, which greatly expands the set of phenomena that can be observed. Figure 1.5 shows the two experimental setups.

The next stage in my thesis was to add components to the chip, and use the generation and detection techniques already developed to study them. Using laser machining, I was able to cut holes and other features in the $LiNbO_3$ slab. Because $LiNbO_3$ has a very high index of refraction for THz radiation (n = 5.1), the air gaps strongly scatter the THz. Launching a multicycle wave using the tilted optical pulse and imaging the wave as it interacts with the structures yields a very complete picture of both wave and photonic element. These videos are excellent illustrations of the principles of electromagnetism, and as such we realized that many of the results had significant educational value. Direct visualizations of propagating electromagnetic waves are more modern versions of classic water wave demonstrations, and these short videos can easily be shown in a lecture. I demonstrated a number of classic experimental geometries including two-slit interference, diffraction off a grating, focusing of a wave, and waveguiding in a dielectric slab [Werley AJP 2012] (see chapter 5). Figure 1.6 shows a frame from such a movie showing 5-slit diffraction. In addition to their educational value, some of the elements could be useful in future devices or on-chip experiments.

The final major thrust of my thesis work was to deposit metallic microstructures onto the surface of the $LiNbO_3$ slab [Werley EO 2012] (see chapter 6). I focused on pairs of half-wave antennas aligned end-to-end and separated by a small gap [see Fig. 1.7(a)]. Antennas for optical and infrared frequencies have received attention recently because of their ability to provide very large field enhancements in regions much smaller than a diffraction-limited spot. Our antenna work had three purposes: to develop a component that can interconvert between propagating electromagnetic waves and subwavelength electrical signals, to harness the antenna's field enhancement to generate very high amplitude electric fields for future nonlinear THz experiments, and to improve fundamental understanding of antenna behavior and gain intuition that can be applied at any frequency range.

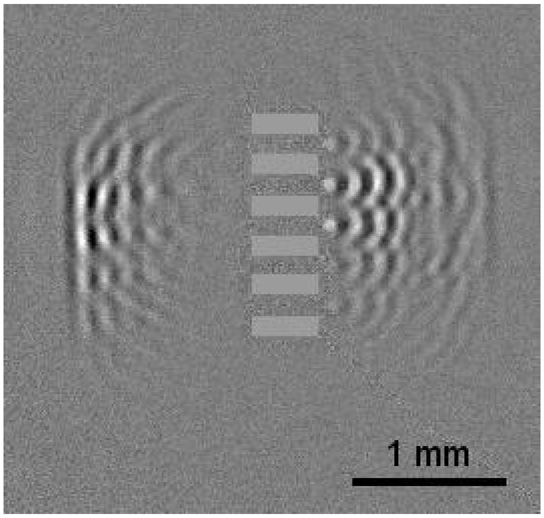

Figure 1.6 | THz wave diffracting through 5 slits. An experimental image of a THz wave soon after interacting with 5 slits. The transmitted wave (right) and reflected wave (left) display diffraction and interference. The light gray regions are air gaps machined into the crystal slab. This is an example of a photonic element that can be made through laser machining; other elements include mirrors, waveguides, and gratings.

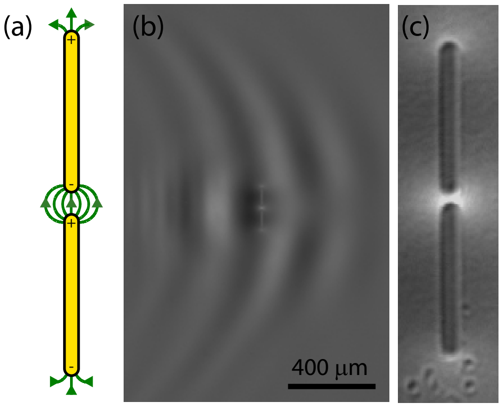

Figure 1.7 | THz fields in an antenna. (a) A diagram of the antenna geometry and how the fields localize at the gaps and antenna ends. (b) An experimental picture of a rightward propagating, resonant THz wave interacting with an antenna pair. The antenna response is 90° out of phase with the driving field. (c) A magnified view of (b) showing field enhancement in the antenna gap.

Our ability to quantitatively record the THz *E*-field with sub-cycle temporal and λ/100 spatial resolution enables particularly incisive study of antenna behavior, allowing us to refine the current understanding. We non-invasively mapped the *E*-field in the antenna's near-field [see Fig. 1.7(b) & (c)] and directly measured field enhancements (up to 40-fold). In addition, we determined the spectral response across more than a decade in bandwidth spanning from DC across multiple resonances, and observed distinct behavior in the near- and far-field. By modeling the antenna as a simple, damped harmonic oscillator we explained the full spectral response. Finally, we measured the field enhancement and resonant frequency as a function of gap size and antenna length and developed intuitive models to predict the trends. These insights are applicable at all frequency ranges and will aid the design of antennas for various applications including single-molecule fluorescence, surface enhanced Raman spectroscopy, near-field scanning optical microscopy, photonic devices, and nonlinear THz spectroscopy.

We are currently using the polaritonics toolkit to study a new set of complex and interesting structures and elements. We have recently begun work on photonic crystals fabricated by cutting holes into the slab, and plan to demonstrate components such as waveguides, filters, and splitters. We have also started depositing metamaterial structures, such as split ring resonators, onto the slab surface. Initial calculations indicate that it will be possible to demonstrate many interesting effects including negative index behavior and cloaking. We expect that our capabilities will yield deep insight into photonic crystals and metamaterials, much as they did with the antennas. After developing a complete understanding of these structures, we plan to pursue active structures that can control the THz waves using electrical biases (e.g. with graphene) or optical pulses (e.g. semiconductors or superconductors). Finally, the spectral intensity in the antenna gap, which determines how strongly a resonance can be driven, is among the largest ever demonstrated at THz frequencies because of the high peak field strengths and multiple cycles covering a relatively narrow frequency range. We are launching a study which will use these enhanced fields to drive nonlinear THz responses. The driven nonlinear responses of various degrees of freedom (vibrational, rotational, and electronic) can be probed with THz, infrared, visible, or even x-ray light in order to build a complete picture of energy and coherence flow within these systems.

References

[Abraham 2011] E. Abraham, Y. Ohgi, M. Minami, M. Jewariya, M. Nagai, T. Araki, and T. Yasui, “Real-time line projection for fast terahertz spectral computed tomography,” *Opt. Lett.* **36**, 2119-2121 (2011).

[Auston 1984] D. H. Auston, K. P. Cheung, and P. R. Smith, “Picosecond photoconducting Hertzian dipoles,” *Appl. Phys. Lett.* **45** 284-286 (1984).

[Auston 1988] D. H. Auston and M. C. Nuss, “Electrooptic generation and detection of femtosecond electrical transients,” *IEEE J. Quant. Electron.* **24** 184-197 (1988).

[Brennan 1997] C. J. Brennan, Thesis title: *Femtosecond wavevector overtone spectroscopy of anharmonic lattice dynamics in ferroelectric crystals*. (Massachusetts Institute of Technology, Cambridge, MA, 1997).

[Dougherty 1992] T. P. Dougherty, G. P. Wiederrecht, and K. A. Nelson, “Impulsive stimulated Raman scattering experiments in the polariton regime,” *J. Opt. Soc. Am.* **9**, 2179-2189 (1992).

[Feurer 2007] T. Feurer, N. S. Stoyanov, D. W. Ward, J. C. Vaughan, E. R. Statz, and K. A. Nelson, “Terahertz polaritonics,” *Annu. Rev. Mater. Res.* **37**, 317-350 (2007).

[Hebling 2010] J. Hebling, M.C. Hoffmann, H.Y. Hwang, K.-L. Yeh, and K.A. Nelson, "Observation of nonequilibrium carrier distribution in Ge, Si, and GaAs by terahertz pump-terahertz probe measurements,” *Phys. Rev. B* **81**, 035201 (2010).

[Laman 2008] N. Laman, S. S. Harsha, D. Grischkowsky, and J. S. Melinger, "7 GHz resolution waveguide THz spectroscopy of explosives related solids showing new features," *Opt. Express* **16**, 4094-4105 (2008).

[Lin 2009] K.-H. Lin, C. A. Werley, K. A. Nelson, "Generation of multicycle terahertz phonon-polariton waves in a planar waveguide by tilted optical pulse fronts," *Appl. Phys. Lett.* **95**, 103304 (2009).

[Saleh 2007] Saleh, B. E. A. and Teich, M. C. *Fundamentals of Photonics*, 2^{nd} ed. (Wiley, Hoboken, NJ, 2007).

[Werley 2010] C. A. Werley, Q. Wu, K.-H. Lin, C. R. Tait, A. Dorn, and K. A. Nelson, "Comparison of phase-sensitive imaging techniques for studying terahertz waves in structured $LiNbO_3$," *J. Opt. Soc. Am. B* **27**, 2350-2359 (2010).

[Werley 2011] C. A. Werley, S. M. Teo, and K. A. Nelson, "Pulsed laser noise analysis and pump-probe signal detection with a DAQ card," *Rev. Sci. Inst.* **82**, 123108 (2011).

[Werley AJP 2012] C. A. Werley, C. R. Tait, and K. A. Nelson, "Direct visualization of terahertz electromagnetic waves in classic experimental geometries," *Am. J. Phys.* **80**, 72-81 (2012).

[Werley OE 2012] C. A. Werley, K. Fan, A. C. Strikwerda, S. M. Teo, X. Zhang, R. D. Averitt, and K. A. Nelson, "Time-resolved imaging of near-fields in THz antennas and direct quantitative measurement of field enhancement," *Opt. Express* **20**, 8551-8567 (2012).

[Wu 2009] Q. Wu, C. A. Werley, K.-H. Lin, A. Dorn, M. G. Bawendi, K. A. Nelson, "Quantitative phase contrast imaging of THz electric fields in a dielectric waveguide," *Opt. Express* **17**, 9219-9225 (2009).

[Yang 2010] C. Yang, Q. Wu, J. Xu, K. A. Nelson, and C. A. Werley, "Experimental and theoretical analysis of THz-frequency, direction-dependent, phonon polariton modes in a subwavelength, anisotropic slab waveguide," *Opt. Express* **18** 26351-26364 (2010).

Chapter II

Dielectric slab waveguides

A. ISOTROPIC SLAB WAVEGUIDES

A dielectric slab waveguide, also commonly called a planar dielectric waveguide, is a slab of dielectric material that extends infinitely in two dimensions. The slab, also called the core, has a higher dielectric constant than the surrounding material, also called the cladding. Electromagnetic waves can be bound inside the slab, so they never escape. These bound eigenmodes propagate along one direction parallel to the slab surface, extend infinitely in the second direction parallel to the surface, and have a transverse profile along the direction orthogonal to the surface that does not evolve as the wave propagates. Because of the boundary conditions at the dielectric interface, bound modes are allowed only at quantized energy levels. That is, for a specific frequency, only a discrete set of wavelengths are allowed. These allowed modes are called the waveguide modes, and the derivation for determining the allowed wavelengths and the transverse profile follows. In this thesis, we will primarily be concerned with waveguides made from lithium niobate which guide THz-frequency waves.

1. Waveguide mode derivation

The derivation below will treat the simple situation of a symmetric, dielectric slab waveguide with both the slab (the core) and the surrounding material (the cladding) isotropic. This derivation procedure can be easily generalized to more complex situations where the geometry is not symmetric [Burnes 1974], some materials are anisotropic [Yang 2010; Marcuse 1978; Marcuse 1979; Burnes 1974], or the index of refraction is negative [Wu 2003; Shadrivov 2003] . The derivation here employs a significantly different strategy than in many books [e.g. Cronin 1995; Saleh 2007], and both ways of thinking can be valuable for intuition.

The slab is assumed to extend infinitely along x and z, both core and cladding have no magnetic response, and all three materials are isotropic. See Fig 2.1 for the geometry. Note that the index of refraction in the cladding has a lower value (n_l) than the high index in the core (n_h). The wave is assumed to propagate along the x-direction. Finally, to simplify the analysis we assume that the waves are harmonic in space and along the propagation direction: $\bar{E}(x,y,z,t)=\bar{E}(y)\exp[i(\beta x-\omega t)]$, where $\beta = k_x$ is the propagation constant.

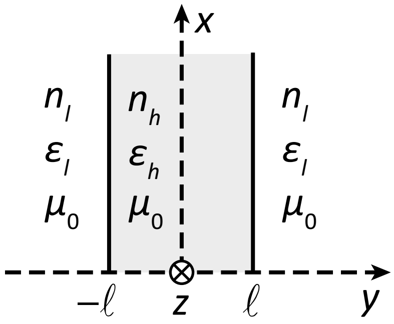

Figure 2.1 | The geometry for a symmetric dielectric slab waveguide. The surfaces of the slab are perpendicular to *y*, and the slab extends infinitely along *x* and *z*. The slab itself, called the core, has a higher index, n_h, than the lower index, n_l, of the surroundings, also called the cladding. For this analysis, we will assume that both core and cladding have no magnetic response. The slab is 2ℓ thick.

The first step in the derivation is to determine the characteristics of waves in bulk material, i.e. the dispersion curves (the relationships between wave vector and frequency) and polarizations, in both core and cladding. Linear combinations of these bulk waves, constrained by system symmetry, are used to build the waveguide modes and lay out the general functional form of the solutions. The boundary conditions at the waveguide surface generate a homogeneous system of equations which can be used to solve for the coefficients in the linear combination. The system of equations can be recast in matrix notation, and a solution exists when the determinant of the matrix is zero. This occurs only for certain pairs of frequency and propagation constant, and these allowed solutions correspond to the waveguide dispersion curve

To simplify notation we define some variables. The propagation constant, $\beta \equiv k_x$ was defined above, and the wave vector component orthogonal to the slab surface is defined both outside the crystal, $i\alpha \equiv k_y^{\text{out}}$, and inside the crystal, $\kappa \equiv k_y^{\text{in}}$. α is defined as imaginary because bound modes will have evanescent, decaying fields in the cladding. Because we assumed that the field extends infinitely along the *z* direction, $k_z = 0$. The bulk dispersion curves are given by the relationship $\vec{k} \cdot \vec{k} = \omega^2 / v^2$, with ω the frequency, *v* the velocity, and $\vec{k}$ the wave vector. This relationship can be easily recognized as the standard relationship between wavelength and frequency:

$$v = \lambda f = \omega / k = \omega / \left|\vec{k}\right| = \omega / \sqrt{k_x^2 + k_y^2 + k_z^2} \quad \Rightarrow \quad k_x^2 + k_y^2 + k_z^2 = \omega^2 / v^2 = \omega^2 n^2 / c^2 \tag{2.1}$$

The dispersion curves result from combining Maxwell's equations to get the wave equation and solving [see i.e. Marcuse 1979]. There are two relevant bulk dispersion curves, one for the cladding and one for the core. They are:

$$\text{Cladding:} \quad \beta^2 - \alpha^2 = \omega^2 n_l^2 / c^2 = k_0^2 n_l^2 \tag{2.2a}$$

$$\text{Ordinary:} \quad \beta^2 + \kappa^2 = \omega^2 n_h^2 / c^2 = k_0^2 n_h^2 \tag{2.2b}$$

where *c* is the speed of light and k_0 is the wave vector in free space. These relations are used to eliminate *α* and *κ* from the equations which follow so everything is expressed in terms of *β* and *ω*.

For a specific pair of *β* and *ω*, there are four possible plane waves in each region, two signs for k_y and two polarizations. Because the materials are isotropic, we can choose any pair of orthogonal polarizations, so for convenience we choose the first polarization component to be $\hat{z}$, along the *z*-axis. The second polarization component must be orthogonal to the *z*-axis and the wave vector. In the low-index cladding we have:

$$\vec{l}^{\,\pm} = \frac{\vec{k}_l^{\pm} \times \hat{z}}{\left\| \vec{k}_l^{\pm} \times \hat{z} \right\|} = \frac{1}{\sqrt{\beta^2 - \alpha^2}} \begin{bmatrix} \pm i\alpha \\ -\beta \\ 0 \end{bmatrix} = \frac{1}{k_0 n_l} \begin{bmatrix} \pm i\alpha \\ -\beta \\ 0 \end{bmatrix}, \tag{2.3a}$$

and similarly for the high index core we have:

$$\vec{h}^{\pm} = \frac{\vec{k}_h^{\pm} \times \hat{z}}{\left\| \vec{k}_h^{\pm} \times \hat{z} \right\|} = \frac{1}{\sqrt{\beta^2 + \kappa^2}} \begin{bmatrix} \pm \kappa \\ -\beta \\ 0 \end{bmatrix} = \frac{1}{k_0 n_h} \begin{bmatrix} \pm \kappa \\ -\beta \\ 0 \end{bmatrix}. \tag{2.3b}$$

Now that we have the polarizations and dispersion curves of the bulk plane waves, we can write the most general form for the waveguide solutions:

$$\begin{aligned}
&\text{cladding, } y < -\ell: \ \vec{E}(y) = A_1 \hat{z} \exp[\alpha y] + A_3 \hat{z} \exp[-\alpha y] + A_2 \vec{l}^{\,-} \exp[\alpha y] + A_4 \vec{l}^{\,+} \exp[-\alpha y] \\
&\text{core: } \vec{E}(y) = B_1 \hat{z} \exp[i\kappa y] + B_3 \hat{z} \exp[-i\kappa y] + B_2 \vec{h}^{+} \exp[i\kappa y] + B_4 \vec{h}^{-} \exp[-i\kappa y] \\
&\text{cladding, } y > \ell: \ \vec{E}(y) = C_1 \hat{z} \exp[\alpha y] + C_3 \hat{z} \exp[-\alpha y] + C_2 \vec{l}^{\,-} \exp[\alpha y] + C_4 \vec{l}^{\,+} \exp[-\alpha y]
\end{aligned} \tag{2.4}$$

where A_i, B_i, and C_i are scalar constants and the +/- superscripts correspond to the sign of *α* or *κ*. This general solution can be immediately simplified. For bound solutions, we require that the electric field decays to zero as $y \to \pm\infty$, so the terms in the cladding that are exponentially growing can be discarded. We now apply the symmetry condition that there is a reflection plane down the center of the sample, which eliminates half of the constants. In this situation, the solution must be made of symmetric and antisymmetric modes. Absorbing some constant factors into the coefficients and making use of Euler's formula, we have:

Symmetric:

$$\text{cladding, } y < -\ell: \ \vec{E}(y) = \begin{bmatrix} \alpha A_2 \\ -i\beta A_2 \\ A_1 \end{bmatrix} \exp[\alpha(y + \ell)]$$

$$\text{core: } \vec{E}(y) = \begin{bmatrix} B_2 \kappa \cos(\kappa y) \\ -iB_2 \beta \sin(\kappa y) \\ B_1 \cos(\kappa y) \end{bmatrix} \tag{2.5a}$$

$$\text{cladding, } y > \ell: \ \vec{E}(y) = \begin{bmatrix} \alpha A_2 \\ i\beta A_2 \\ A_1 \end{bmatrix} \exp[-\alpha(y-\ell)]$$

<u>Antisymmetric:</u>

$$\text{cladding, } y < -\ell: \ \vec{E}(y) = \begin{bmatrix} \alpha A_2 \\ -i\beta A_2 \\ A_1 \end{bmatrix} \exp[\alpha(y+\ell)]$$

$$\text{core: } \vec{E}(y) = \begin{bmatrix} B_2 \kappa \sin(\kappa y) \\ iB_2 \beta \cos(\kappa y) \\ B_1 \sin(\kappa y) \end{bmatrix} \tag{2.5b}$$

$$\text{cladding, } y > \ell: \ \vec{E}(y) = \begin{bmatrix} -\alpha A_2 \\ -i\beta A_2 \\ -A_1 \end{bmatrix} \exp[-\alpha(y-\ell)]$$

Note that for the symmetric case reflection across the symmetry plane, which lies in the center of the slab, should not change the sign of a vector pointing along *x* or *z*, but should change the sign of a vector pointing along *y*. This condition is met in the symmetric mode solutions, and the opposite is met for the antisymmetric solutions. Applying the symmetry conditions eliminates half the unknowns, so now we need only apply boundary conditions at one interface to solve for the unknown coefficients. The boundary condition is that the tangential E and H-fields must be continuous across the boundary [Born 1999]. Because all the calculations thus far have been performed using the electric field, we use Faraday's law to recast the boundary conditions in terms of electric field. Faraday's law is:

$$\vec{\nabla} \times \vec{E} = -\frac{d\vec{B}}{dt} = i\omega \vec{B} = i\omega\mu_0 \vec{H} \ \rightarrow \ -\frac{i}{\omega\mu_0} \vec{\nabla} \times \vec{E} \tag{2.6}$$

Remembering that our functional form is $\vec{E}(x,y,z,t) = \vec{E}(y)\exp[i(\beta x - \omega t)]$, some derivatives simplify to $\partial/\partial z = 0$, $\partial/\partial x = i\beta$, and $\partial/\partial t = -i\omega$. The boundary conditions all in terms of E-field are:

$$\begin{gathered} E_{z,\text{clad}} = E_{z,\text{core}} \\ \frac{\partial E_{z,\text{clad}}}{\partial y} = \frac{\partial E_{z,\text{core}}}{\partial y} \\ E_{x,\text{clad}} = E_{x,\text{core}} \\ i\beta E_{y,\text{clad}} - \frac{\partial E_{x,\text{clad}}}{\partial y} = i\beta E_{y,\text{core}} - \frac{\partial E_{x,\text{core}}}{\partial y} \end{gathered} \tag{2.7}$$

All the above boundary conditions should be evaluated at the interface ($y = \ell$), and they must be solved independently for the symmetric and antisymmetric modes. The four expressions above yield a set of homogeneous equations which can be recast in matrix notation.

Symmetric:

$$\begin{bmatrix} 1 & -\cos(\kappa\ell) & 0 & 0 \\ \alpha & -\kappa\sin(\kappa\ell) & 0 & 0 \\ 0 & 0 & \alpha & -\kappa\cos(\kappa\ell) \\ 0 & 0 & n_l^2 & n_h^2\sin(\kappa\ell) \end{bmatrix} \begin{bmatrix} A_1 \\ B_1 \\ A_2 \\ B_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{2.8a}$$

Antisymmetric:

$$\begin{bmatrix} 1 & \sin(\kappa\ell) & 0 & 0 \\ \alpha & -\kappa\cos(\kappa\ell) & 0 & 0 \\ 0 & 0 & \alpha & \kappa\sin(\kappa\ell) \\ 0 & 0 & n_l^2 & n_h^2\cos(\kappa\ell) \end{bmatrix} \begin{bmatrix} A_1 \\ B_1 \\ A_2 \\ B_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{2.8b}$$

Because of our choice of polarizations, these equations are uncoupled for both the symmetric and antisymmetric cases. The modes associated with A_1 and B_1 have pure *z*-polarization and are the transverse electric or TE modes, which means that the electric field (along *z*) is perpendicular to the propagation direction (along *x*). The modes associated with A_2 and B_2 are the transverse magnetic or TM modes, whose magnetic field is perpendicular to the propagation direction.

As is true for any homogeneous system of equations, there is a solution when the determinant is zero. Because these break up into 2x2 matrices, the determinants can be easily calculated analytically. They yield a set of transcendental equations whose solutions are the waveguide dispersion curves:

$$\text{Symmetric, TE: } \frac{\alpha}{\kappa} = \tan(\kappa\ell) \tag{2.9a}$$

$$\text{Antisymmetric, TE: } \frac{\alpha}{\kappa} = -\cot(\kappa\ell) \tag{2.9b}$$

$$\text{Symmetric, TM: } \frac{\alpha}{\kappa} = -\frac{n_l^2}{n_h^2}\cot(\kappa\ell) \tag{2.9c}$$

$$\text{Antisymmetric, TM: } \frac{\alpha}{\kappa} = \frac{n_l^2}{n_h^2}\tan(\kappa\ell) \tag{2.9d}$$

Finally, for the allowed solutions which fall on the dispersion curve, we can write the field profiles using Eq. 2.8 to solve for A_i and B_i. We get:

TE, symmetric:

$$\vec{E}(y) = \hat{z}E_0 \left| \begin{array}{ll} \cos(\kappa\ell)\exp[\alpha(y+\ell)] & y < -\ell \\ \cos(\kappa y) & -\ell < y < \ell \\ \cos(\kappa\ell)\exp[-\alpha(y-\ell)] & y > \ell \end{array} \right. \tag{2.10a}$$

TE, antisymmetric:

$$\vec{E}(y) = \hat{z}E_0 \left| \begin{array}{ll} -\sin(\kappa\ell)\exp[\alpha(y+\ell)] & y < -\ell \\ \sin(\kappa y) & -\ell < y < \ell \\ \sin(\kappa\ell)\exp[-\alpha(y-\ell)] & y > \ell \end{array} \right. \tag{2.10b}$$

TM, symmetric:

$$\vec{E}(y) = E_0 \left\{ \hat{x} \begin{bmatrix} \cos(\kappa\ell)\exp[\alpha(y+\ell)] \\ \cos(\kappa y) \\ \cos(\kappa\ell)\exp[-\alpha(y-\ell)] \end{bmatrix} - i\hat{y} \begin{bmatrix} -(\beta/\alpha)\cos(\kappa\ell)\exp[\alpha(y+\ell)] \\ (\beta/\kappa)\sin(\kappa y) \\ (\beta/\alpha)\cos(\kappa\ell)\exp[-\alpha(y-\ell)] \end{bmatrix} \right\} \begin{array}{c} y < -\ell \\ -\ell < y < \ell \\ y > \ell \end{array} \tag{2.10c}$$

TM, antisymmetric:

$$\vec{E}(y) = E_0 \left\{ \hat{x} \begin{bmatrix} -\sin(\kappa\ell)\exp[\alpha(y+\ell)] \\ \sin(\kappa y) \\ \sin(\kappa\ell)\exp[-\alpha(y-\ell)] \end{bmatrix} + i\hat{y} \begin{bmatrix} (\beta/\alpha)\sin(\kappa\ell)\exp[\alpha(y+\ell)] \\ (\beta/\kappa)\cos(\kappa y) \\ (\beta/\alpha)\sin(\kappa\ell)\exp[-\alpha(y-\ell)] \end{bmatrix} \right\} \begin{array}{c} y < -\ell \\ -\ell < y < \ell \\ y > \ell \end{array} \tag{2.10d}$$

2. Numerical solutions

Finally, to get all the solutions we must find the wave vectors, which means solving the transcendental equations (Eq. 2.9). The bulk dispersion curves (Eq. 2.2) can be used remove all dependence on *α*. Defining two new unitless variables to facilitate numerical solution, we get:

$$\text{Symmetric, TE: } \sqrt{\frac{A^2(n_h^2-n_l^2)}{b^2}-1} = \tan(b) \tag{2.11a}$$

$$\text{Antisymmetric, TE: } \sqrt{\frac{A^2(n_h^2-n_l^2)}{b^2}-1} = -\cot(b) \tag{2.11b}$$

$$\text{Symmetric, TM: } \frac{n_h^2}{n_l^2}\sqrt{\frac{A^2(n_h^2-n_l^2)}{b^2}-1} = -\cot(b) \tag{2.11c}$$

$$\text{Antisymmetric, TM: } \frac{n_h^2}{n_l^2}\sqrt{\frac{A^2(n_h^2-n_l^2)}{b^2}-1} = \tan(b) \tag{2.11d}$$

with $\omega = cA/\ell$, $\kappa = b/\ell$, $\beta = \sqrt{A^2 n_h^2 - b^2}\,/\ell$, $\alpha = \sqrt{A^2(n_h^2-n_l^2)-b^2}\,/\ell$.

It is important to remember that tangent and cotangent are periodic, so there are many solutions. Now we take advantage of the trig identities $\tan(\phi+\pi/2)=-\cot(\phi)$ and $\tan(\phi)=\tan(\phi+m\pi)$ with m an integer. This conveniently lets us compress the four equations into two:

$$\text{TE:}\quad \sqrt{\frac{A^2(n_h^2-n_l^2)}{b^2}-1}=\tan(b+m\pi/2) \tag{2.12a}$$

$$\text{TM:}\quad \frac{n_h^2}{n_l^2}\sqrt{\frac{A^2(n_h^2-n_l^2)}{b^2}-1}=\tan(b+m\pi/2) \tag{2.12b}$$

with m a non-negative integer. For the TE case, the m even are the symmetric solutions and m odd are the antisymmetric solutions, while the opposite holds for the TM case. Each m corresponds to a different waveguide mode, and for TE modes m corresponds to the number of nodes. Note that the lowest mode (m = 0) exists for all frequencies, but the higher modes have cutoff frequencies which depend on the slab thickness and indices. See [Saleh 2007] for details. The transcendental equations (2.12a and b) can be solved numerically, for instance by bisection. To easily find the numerical solutions, good guesses for the upper and lower bounds must be made. These can be accurately chosen by understanding the structure of the solutions for the right hand side (RHS) and left hand side (LHS) of Eq. 2.12a and b. Figure 2.2 below shows the structure of these functions for the TE case.

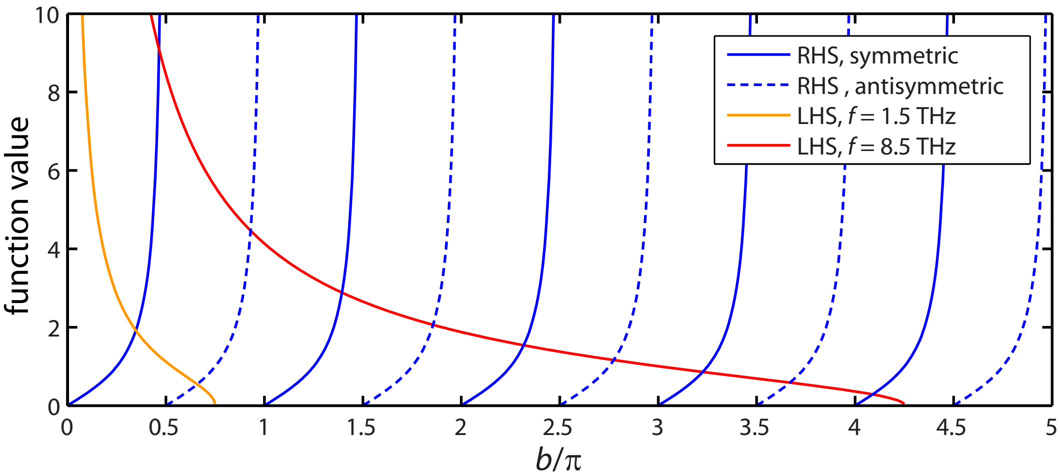

Figure 2.2 | Trancendental equation plots. The curves for TE modes in a high-index waveguide in air (n_h = 5.1, n_l = 1, $\ell = 15\,\mu\text{m}$). The solid and dashed blue lines are the RHS plots for symmetric and antisymmetric modes, respectively. The orange and red lines are the LHS plots for f = 1.5 and 8.5 THz. At 1.5 THz, there are two allowed modes (the RHS and LHS are equal when the plots cross), while there are 9 allowed modes at 8.5 THz.

The solution for the first mode will always lie between 0 and $\pi/2$, the second (if it is allowed) between $\pi/2$ and π, and so on. The final allowed mode has an upper bound where the LHS of Eq. 2.12 is equal to zero, or $b_{\max} = A\sqrt{n_h^2 - n_l^2}$.

Figure 2.3 shows the dispersions curves, or the allowed values for the wave vectors, for the first three TE (solid) and TM (dashed) modes. The cutoff frequencies below which the second and third mode do not exist are clearly visible. α extends all the way to zero, indicating that the evanescent field extends to infinity at the cutoff frequency for each mode. κ has clear cutoff wave vectors corresponding to integer numbers of half-periods within the slab. The propagation constant β is bounded by the dispersion curves for the bulk materials. At low frequencies the waves behave primarily like a wave in bulk cladding material, and at high frequencies, the waves behave primarily like a wave in bulk core material.

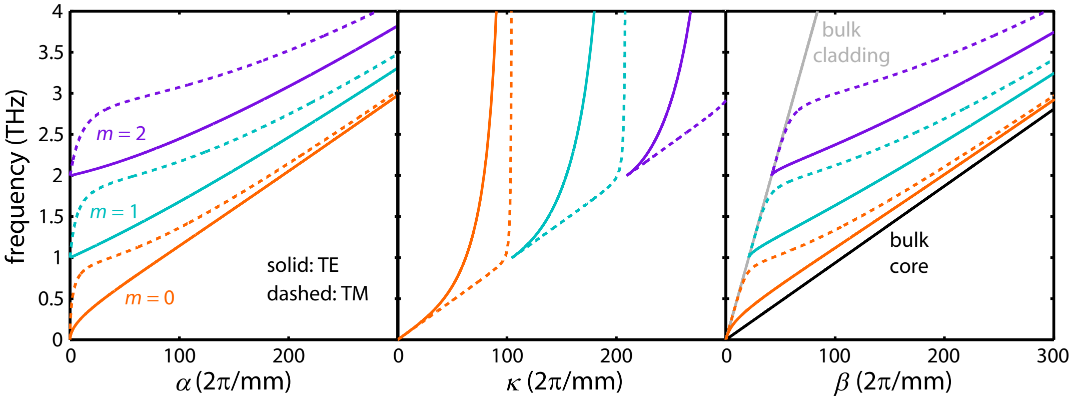

Figure 2.3 | Example dispersion curves for the first three TE and TM modes. This example is for a high-index waveguide in air designed for THz frequencies. The slab is 30 μm thick, $n_h = 5.1$, and $n_l = 1$. The first frame shows f vs. α, the second f vs. κ, and the third f vs. β. The lowest mode is in orange, the second mode is in light blue, and the third mode is in purple. TM modes are shown as dashed lines and TE modes are shown as solid lines. In the third frame, the black and gray line show the bulk dispersion curve for the core and cladding respectively.

Another valuable quantity which builds intuition for wave behavior is the effective index. The effective phase index, the ratio between the speed of light and the phase velocity and defined as $n_{\text{ph}} = c\beta/\omega$, can be used to determine the velocity at which the phase fronts propagate in the waveguide. The group index, the ratio between the speed of light and the group velocity and defined as $n_{\text{gr}} = c\dfrac{d\beta}{d\omega}$, can be used to determine the velocity at which wave packets move in the waveguide. These two values are plotted in Fig. 2.4. The phase index transitions from the cladding index to the core index, while the group index greatly overshoots the bulk value, leading to very slowly propagating wavepackets, before it asymptotes to the core group index at high frequencies.

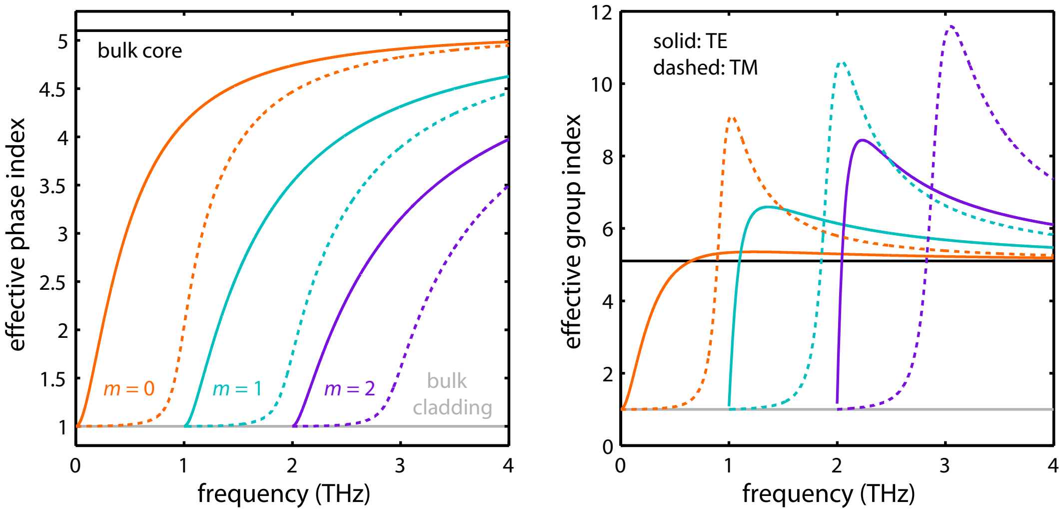

Figure 2.4 | Phase and group index for the first three TE and TM modes. This example is for a high-index waveguide in air designed for THz frequencies. The slab is 30 µm thick, $n_h = 5.1$, and $n_l = 1$. The first frame shows n_{ph} vs. f while the second shows n_{gr} vs. f. The lowest mode is in orange, the second mode is in light blue, and the third mode is in purple. TM modes are shown as dashed lines and TE modes are shown as solid lines. The black and gray lines show the bulk index of the core and cladding respectively.

With the dispersion curves in hand, it is now possible to plot some example field profiles using Eq. 2.10a-d. Figure 2.5 shows TE (solid) and TM (dashed) field profiles for the first three modes at different frequencies. The boundary conditions at the interface (Eq. 2.7) say that E_z must be continuous and its first derivative must be continuous, leading to smooth solutions with profiles identical to the quantum mechanical problem of a particle in a box with finite potential wall height. The boundary conditions for the TM modes are more complicated (Eq. 2.7), resulting in more complicated field profiles. There are two polarization components. E_x must be continuous at the boundary, although its derivative need not be, and neither E_y nor its derivative need be continuous. For both solutions, there are exponentially decaying evanescent waves in the cladding and sinusoidal solutions in the core. For each mode and both TE and TM waves, the decay length of the evanescent field is longer near the cutoff frequency so that more energy is in the cladding. The decay length becomes shorter as the frequency increases so that at high frequencies most energy is in the core. For a given frequency, the evanescent decay length is always longer for TM waves. Finally, we can predict the number of nodes by the mode number, m. For E_y and E_z the number of nodes is equal to m, while for E_x, the number of nodes is $m + 1$.

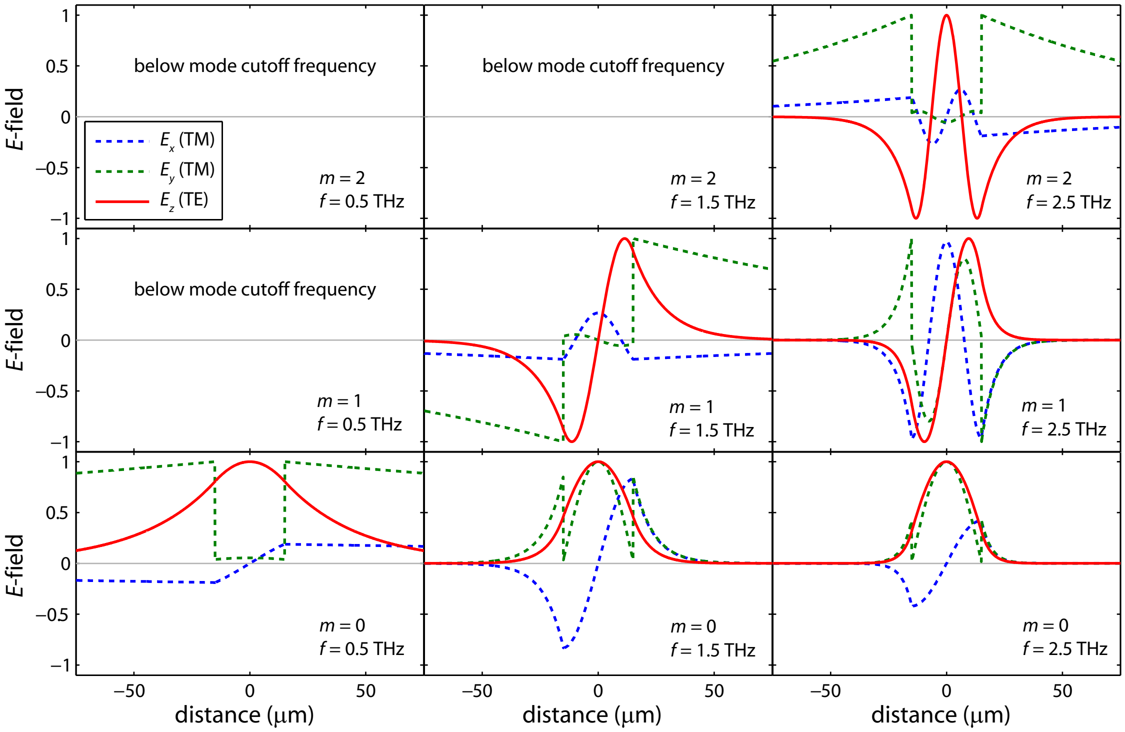

Figure 2.5 | Example E-field profiles for different modes and frequencies. This example is for a high-index waveguide in air designed for THz frequencies. The slab is 30 μm thick, $n_h = 5.1$, and $n_l = 1$. The solid, red line is for TE modes where the field is *z*-polarized. Dashed blue and dashed green correspond to the *x* and *y*-polarized components of the TM wave. The top three frames are for the third mode, the middle three for the second mode, and the bottom three for the lowest mode. The left three frames are at 0.5 THz, the middle three at 1.5 THz, and the right three at 2.5 THz. The three frames to the upper left have no fields shown because the higher modes cannot propagate cannot propagate at frequencies below their cutoff.

B. ANISOTROPIC SLAB WAVEGUIDE, EXPERIMENT & THEORY

Content from: C. Yang, Q. Wu, J. Xu, K. A. Nelson, and C. A. Werley. “Experimental and theoretical analysis of THz-frequency, direction-dependent, phonon polariton modes in a subwavelength, anisotropic slab waveguide,” *Opt. Express* **18**, 26351-26354 (2010).

1. Abstract

Femtosecond optical pulses were used to generate THz-frequency phonon polariton waves in a 50 micrometer lithium niobate slab, which acts as a subwavelength, anisotropic planar waveguide. The spatial and temporal electric field profiles of the THz waves were recorded for different propagation directions using a polarization gating imaging system, and experimental dispersion curves were determined via a two-dimensional Fourier transform. Dispersion relations for an anisotropic slab waveguide were derived via analytical analysis and found to be in excellent agreement with all observed experimental modes. From the dispersion relations, we analyze the propagation-direction-dependent behavior, effective refractive index values, and generation efficiencies for THz-frequency modes in the subwavelength, anisotropic slab waveguide.

2. Introduction

Terahertz-frequency phonon polariton generation, control and detection have received extensive attention in recent years due to their outstanding capabilities in terahertz (THz) spectroscopy, imaging and advanced signal processing [Lee 2009, Koehl 1999, Stoyanov 2002, Feurer 2003, Feurer 2007]. Phonon polariton waves result from the coupling of lattice vibrational waves and electromagnetic waves, and can be generated in ferroelectric crystals such as $LiNbO_3$ (LN) via impulsive stimulated Raman scattering (ISRS) using femtosecond optical pulses [Dougherty 1992,Yan 1985]. The electromagnetic component of the phonon polariton wave can be coupled into free space and is a source for intense THz pulses [Auston 1988, Lee 2000, Hebling 2004, Yeh 2007, Lin 2009]. THz waves generated in the sample do not propagate collinearly with the pump beam due to the large index-mismatch between optical and THz frequencies. Instead they generate a Cherenkov radiation pattern and propagate primarily in the lateral direction [Auston 1984, Wahlstrand 2003]. This lateral propagation facilitates coherent control of the THz wave, which can easily be made to interact with subsequent optical pulses, other THz waves, or patterned structures all in the same small crystal of LN. As a result, a LN slab can serve as a platform for THz processing because generation, propagation, detection, and control can be fully integrated in one sample [Feurer 2007, Stoyanov 2003]. Furthermore, when the sample thickness becomes comparable to or less than the THz wavelength, the strong evanescent field of the THz wave can interact with material deposited on the crystal surface. This opens the door for spectroscopic analysis and interfacing of the LN slab with other optical or photoelectric devices.

Because the THz wave propagates almost perpendicular to the optical pump beam, it is possible to obtain time-resolved images of the electric field in the LN slab. As the THz wave propagates through the crystal, its electric field changes the refractive index through the electro-optic effect. The time-delayed probe pulses, which can be expanded to illuminate the whole crystal, experience a spatially dependent phase shift proportional to the refractive index change. Four methods have been introduced to convert this phase pattern to an amplitude image: Talbot imaging [Koehl 1999], Sagnac interferometry [Peier 2008, Werley 2010], polarization gating [Peier 2008], and phase contrast imaging [Wu 2009]. In a recent comparison [Werley 2010], an improved geometry for polarization gating was found to offer the best sensitivity and most reliable field quantification, while phase contrast imaging was best in situations requiring high spatial resolution. In this paper, we used the polarization gating system similar to that shown in [Werley 2010] to record a sequence of images. The full spatio-temporal evolution was extracted from the image sequence and double Fourier transformed to obtain the wave vector vs. frequency dispersion curves [e.g. Wu 2009]. The data collection and analysis were performed as a function of wave propagation direction to study the complex mode structure present in an anisotropic slab waveguide, which was found to be in excellent agreement with theory. From the dispersion relations we extract the mode and propagation-angle dependent effective refractive index (ERI) and discuss pumping efficiencies for THz phonon-polariton waves in a LN waveguide.

3. Experimental

The experiments were performed with a Ti:sapphire regenerative amplifier whose pulse duration was 120 fs, central wavelength was 800 nm, and repetition rate was 1 KHz. The laser pulses were divided into a pump beam (370 μJ per pulse) and probe beam (35 μJ per pulse). The vertically polarized pump beam was routed through a mechanical delay stage and then focused to a line on the sample by a 200 mm focal length cylindrical lens. The probe was frequency-doubled to 400 nm in a BBO crystal and expanded to be larger than the sample. Figure 2.6(a) shows a sketch of the experimental setup and the coordinate system. A quarter-wave plate (QW1) and a retroreflective mirror were used in a 4-f system. The mirror and lenses imaged the sample precisely back onto itself without magnification or inversion. The axis of QW1, which was the same as the first Glan-Taylor polarizer (GTP1), was at +45° so it exchanged the ordinary and the extraordinary polarization components of the probe. In this way the spatially varying phase shift between the vertical and horizontal polarization components accumulated from the probe's first pass through the sample was compensated after the second pass. The phase shift after the first pass resulted from the intrinsic birefringence of the LN slab, and self-compensation was necessary to correct for spatial inhomogeneities in the phase shift due to thickness variation, strain, or other imperfections in the slab. The phase shift electro-optically induced by the THz wave, however, was not compensated because the THz wave was launched only after the probe pulse had passed through the sample the first time. The THz-induced phase information was converted to amplitude information prior to detection with the camera by QW2 (oriented vertically) and GTP2 (oriented at -45°). In this geometry a positive (negative) field results in a positive (negative) amplitude change [Werley 2010].

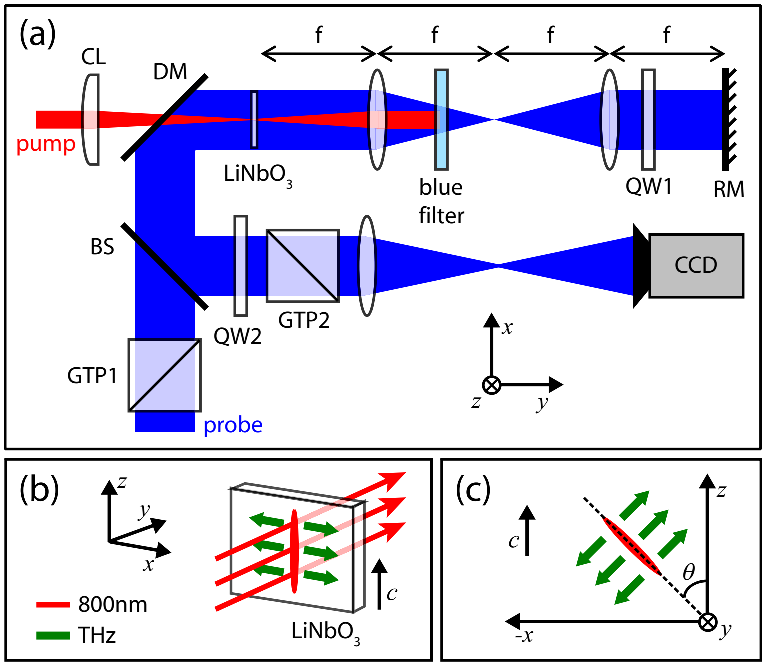

Figure 2.6 | The experimental geometry. (a) Overview diagram of the experimental setup. GTP1 and GTP2 are Glan-Taylor prisms, whose polarizations are at +45° and -45° to *z*-axis respectively. BS: 400 nm beam splitter; CL: cylindrical lens; DM: dichroic mirror; RM: retroreflective mirror. QW1 and QW2 are zero order 400 nm quarter-wave plates with optic axes at +45° and parallel to z-axis respectively. The 800 nm pump (red) and 400 nm probe (blue) are nearly collinear when they arrive at the sample, a 50 µm thick $LiNbO_3$ slab. (b) The pump geometry and coordinate system. The 800 nm pump beam (red) propagates through the crystal, orthogonal to the crystal surface, while the THz (green) is guided down the slab. (c) The cylindrical lens can be rotated by θ relative to the *z*-axis (the *c* crystallographic axis of the LN sample) in order to launch the THz wave in a 90°-θ direction.

The pump geometry is shown in Fig. 2.6(b). Red lines represent the 800 nm pump beam and green the broadband THz waves generated when the pump is focused into the 50 µm thick $LiNbO_3$ crystal slab. Because the center wavelength of the THz phonon polariton wave is about 100 µm, the slab acts as a sub-wavelength waveguide. As Fig. 2.6(c) shows, the THz wave propagation direction was changed by rotating the cylindrical lens. Because of the strong anisotropy of LN at THz frequencies ($n_e \sim 5.1$, $n_o \sim 6.5$ [Feurer 2007]), the nature and behavior of the waveguide modes change drastically as the propagation direction rotates relative to the optic axis.

4. Results

By changing the delay between the pump and probe pulses, a series of images can be obtained. The image sequence can be compiled to form a movie showing THz propagation [Koehl 1999, Stoyanov 2002, Feurer 2003, Feurer 2007, Wu 2009, Werley 2010]. Frames from such a movie (video 9 in [Werley 2012]) are shown in Fig. 2.7. The optical pump pulse used to launch the THz wave is focused to a line, so the generated THz wave is uniform along that line (the vertical dimension in Fig. 2.7). Figure 2.7(a) shows the moment of generation, when the pump is still in the crystal. Shortly later [Fig. 2.7(b)], two equal-amplitude, counter-propagating waves are observed. As they propagate further, the waveguide modes begin to separate, and frequencies within each waveguide mode also separate. The first three waveguide modes are clearly visible in Fig. 2.7(c) & (d). In (d), the chirp of the lowest waveguide mode (which has propagated farthest from the generation region) is clearly visible. Figure 2.7(e) shows the waves after they have reflected off the crystal edges, and (f) shows the standing wave generated by the counter-propagating waves as they cross and interfere.

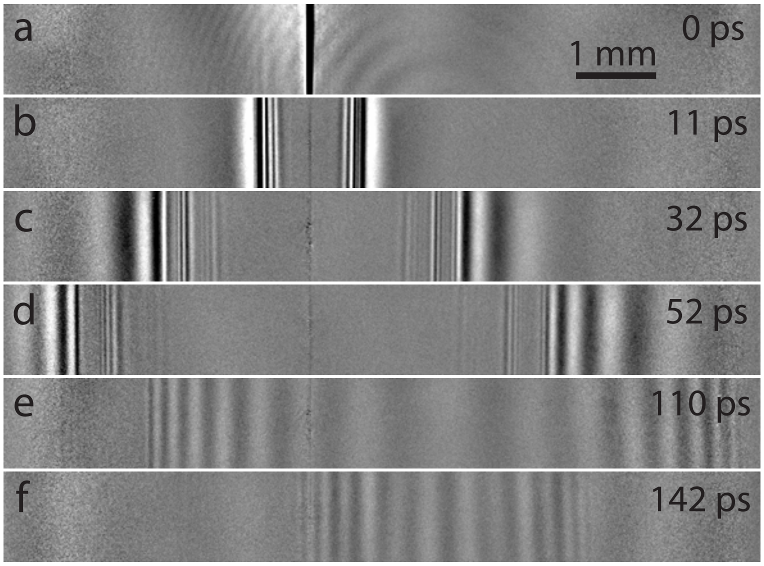

Figure 2.7 | Images of a broadband, guided THz wave. Panels (a)-(f) are frames from video 9 in [Werley 2012] depicting broadband THz waves propagating in an unstructured, 50 μm thick $LiNbO_3$ crystal. Counter-propagating waves are launched by a cylindrically focused "line" of pump pulse light in (a). The first three waveguide modes are clearly visible in (c) and (d). After frame (d) the wave reflects off the crystal edges so the waves are propagating toward one another in (e), and they have begun to overlap and interfere in (f).

Videos of the propagating THz waves, like the one represented by the time-incremented sequence of images in Fig. 2.7, can be used to determine the waveguide dispersion curve. To do this, we first calculate a space-time plot showing wave propagation. Looking at the images in Fig. 2.7, it is clear that the signal is uniform along the vertical dimension. Averaging over this direction collapses each 2D matrix of values (the image) into a 1D vector. The 1D vectors for all time delays are placed in rows of the space-time matrix, one above the other in time order. The result, shown in Fig. 2.8(a), depicts the full temporal and spatial evolution of the wave. In Fig. 2.8(a) we can clearly see all the effects mentioned in the discussion of Fig. 2.7: dispersion, reflection, and different waveguide modes. A 2-dimensional Fourier transformation of Fig. 2.8(a) yields the THz dispersion curves [Fig. 2.8(b)]. Along the vertical axis, time is transformed to frequency, and along the horizontal axis, space is transformed to the wave vector, k_x, which is often called the waveguide propagation constant, β. In Fig. 2.8, $\theta = 0$, so the crystal's c-axis is parallel to the 800 nm pump polarization. In this geometry, which has been most used in previous work [Koehl 1999, Stoyanov 2002, Feurer 2003, Feurer 2007, Stoyanov 2003, Peier 2008, Wu 2009, Werley 2010], only z-polarized THz is generated, and true transverse electric (TE) modes are launched in the slab. Overlaid on the experimental data are the dispersion curves for air (white line), bulk LN (magenta line), and the calculated TE mode dispersion curves (see e. g. [Saleh 2007]) up to a frequency of 2 THz for an isotropic slab waveguide with $n = n_e$ (dotted blue lines). The curves show four TE waveguide modes, which propagate at different group velocities, $v_g = d\omega / dk$, and phase velocities, $v_p = \omega / k$. Cutoff frequencies can be seen for the all but the first mode as expected. Although the isotropic waveguide analysis is predictive in this simple geometry, a more complete analysis is required when $\theta \neq 0$, as will be shown below.

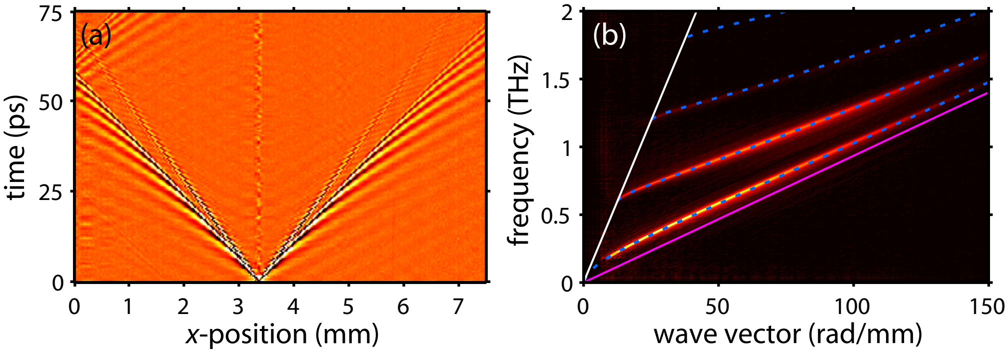

Figure 2.8 | 2-dimensional THz wave plots. (a) Space-time plot of a propagating THz wave. We can see waveguide dispersion (the frequencies separate as time progresses), reflection from the crystal edge, and the first two waveguide modes (the second mode has a higher frequency and a steeper slope because of its lower group velocity) in this picture. The horizontal axis is the x-axis of the coordinate system in Fig. 2.6 and the vertical axis is the delay time between the probe and pump. (b) Dispersion curves of the THz wave in the LN slab waveguide computed by 2D Fourier transformation of (a). The horizontal axis is the wave vector, k_x (also called the propagation constant, β), and the vertical axis is frequency of THz wave in the sample. Theoretical dispersion curves in air (white), bulk LN (magenta), and a 50 μm slab waveguide (dotted blue) are overlaid on the experimental data where the first three modes are visible.

In an anisotropic waveguide, constraints relating to propagation in bulk anisotropic material and constraints relating to propagation in a waveguide both come into play. In bulk anisotropic material waves are divided into two normal modes, ordinary waves and extraordinary waves, which propagate through the material at different velocities [Born 1999]. In an isotropic waveguide there are also two uncoupled eigenmodes, the transverse electric (TE) and transverse magnetic (TM) modes, which propagate through the waveguide at different velocities [Saleh 1991]. When $\theta = 0$ or 90°, these modes map directly onto one another. For $\theta = 0°$ the TE mode is an extraordinary wave and the TM mode is an ordinary wave while for $\theta = 90°$ the opposite pairing holds. When $\theta \neq 0$, however, the high-symmetry configuration is broken and all the modes couple together. The new eigenmodes of the system are neither purely TE nor TM and also not purely ordinary or extraordinary. The coupling effects the mode profiles, dispersion curves, and effective refractive indices in a fundamental and significant way, as will be demonstrated experimentally (presented immediately below) and theoretically (the full analysis can be found in Sec. 2.B.5 below) in the remainder of this paper.

With the experimental system mentioned above, we measured the dispersion curves for different propagation directions by rotating the cylindrical lens and CCD camera together, which kept the THz wavefront aligned vertically in the images. In this manner the THz wave propagation direction was varied from 0 to 90 degrees relative to the *c*-axis as shown in Fig. 2.6(c). The polarization of the 800 nm pump light was not rotated and thus was parallel to the *c*-axis in all measurements. Because of the strong r_{33} electro-optic coefficient in LN, this ensured efficient pumping of THz waves with a large component polarized along the optic axis [Auston 1988, Barker 1967, Stoyanov 2004]. Using the same data collection and analysis procedure as was used to generate Fig. 2.8(b), the dispersion curves were measured for different angles θ, some examples of which are shown in Fig. 2.9. Overlaid on the experimental data are the theoretical solutions for a bound mode propagating in an anisotropic, dielectric slab waveguide with a thickness of 50 μm, an extraordinary index of 5.11, and an ordinary index of 6.8. The full derivation is presented in Sec. 2.B.5. For all modes and all angles, the data agree very well with theoretical predictions.

In Fig. 2.9(a) where $\theta = 20°$, one set of modes is very TE-like, and one is strongly TM-like. Because the TE-like modes have their primary polarization component along the *c*-axis, they were pumped much more strongly than the TM-like modes, which were too weak to be observed clearly. Blue dotted lines are calculated TE-like modes and green dashed lines are TM-like modes. An interesting effect resulting from propagation in the anisotropic waveguide when $\theta \neq 0$ or 90° is visible in the region containing the white lines, a magnified view of which is shown in the lower right corner. Although the TM-like modes are not visible, we still see an avoided crossing when two modes with the same symmetry (symmetric or antisymmetric) cross. The avoided crossing, visible in experiment and predicted by theory, results from coupling between TE- and TM-like modes in the anisotropic waveguide. Here the two lowest symmetric modes, the lowest TE-like mode and the second TM-like mode, avoid each other.

Figure 2.9(b) shows dispersion curves for the case of $\theta = 50°$. The three TE-like modes predominate, although their strength is reduced, and TM modes still cannot be observed. No avoided crossings occur between modes of the same symmetry within the bandwidth of the experiment. As θ increases, the velocity of the extraordinary wave approaches that of the ordinary wave [Marcuse 1979], which means that the slopes of TE-like and TM-like modes tend to be more similar at higher frequencies. Figure 2.9(c) shows the results for $\theta = 70°$, where both TM-like and TE-like modes can be seen clearly. The strength of the TE-like modes continues to decrease with increased θ and TM-like modes are finally pumped strongly enough to detect. Although some of the modes are weak, the first seven modes can be observed in the experiment, and all the dispersion curves agree with theoretical predictions. Continuing the trend, at high frequencies the slopes of the TM-like and TE-like modes become even more similar. Finally, Fig. 2.9(d) shows results from for $\theta = 90°$, where only TM modes can be observed.

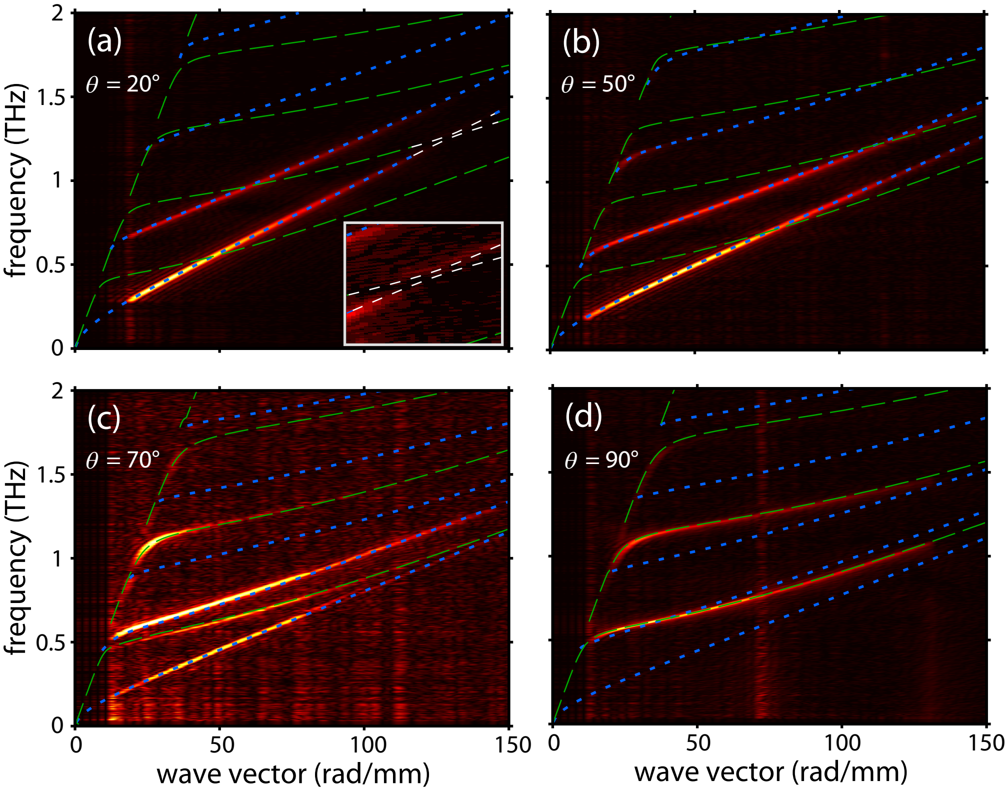

Figure 2.9 | Experimental and theoretical dispersion curves for various propagation directions in a 50 μm thick LN slab waveguide. Angles give the optical pump cylindrical focus orientation relative to the LN optic axis as shown in Fig. 2.6(c). Blue dotted lines are calculated TE-like mode dispersion curves and green dashed lines are TM-like modes. (a) Dispersion curves for θ = 20°. Experimentally we see three TE-like modes and no TM-like modes. The white box in the lower right shows a blow-up of the region around an avoided crossing between the two lowest symmetrical modes. (b) Dispersion curves for θ = 50°. In this case, TE-like modes still predominate and TM-like modes are too weak to be observed. (c) Dispersion curves for θ = 70°. We can clearly see both the first two TE-like modes and the first two TM-like modes, in addition to hints of several others. (d) Dispersion curves for θ = 90°, in which only the TM modes are excited. All the experimental data agree well with the calculated curves.

Using the derivation in Sec. 2.B.5, we can calculate the E-field profile of THz waves as shown in Fig. 2.10. Figures 2.10(a) and (b) show field profiles for TE and TM modes respectively at θ = 0°. The coordinate system in Fig. 2.10 is the same as in Sec. 3.B.5 (see Fig. 2.14), where the axes are defined by the propagation direction of the wave and not by the lab frame as in Fig. 2.6. Blue, green and red lines represent electric field along x-axis, y-axis and z-axis respectively. The electric field along the y-axis, whose polarization is perpendicular to the surface of the slab, changes drastically at the slab surface (±25 μm). As mentioned above, pure TE and TM modes only exist at 0 and 90 degrees. At any other angle the eigenmodes are superpositions of TE and TM modes and contain all three polarization components, as shown in Fig. 2.10(c) and (d) where θ = 50°.

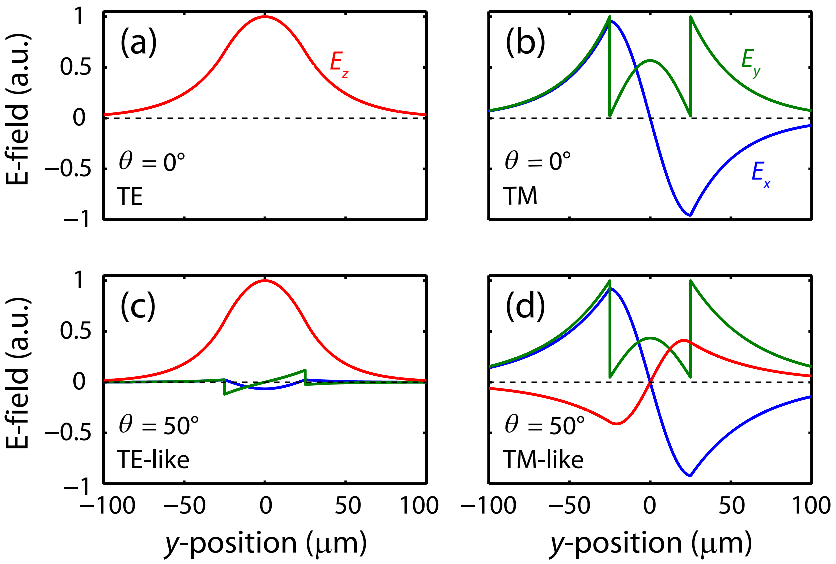

Figure 2.10 | Electric field profiles. Electric field profiles for the lowest symmetric and antisymmetric modes at 0.5 THz. E_x, E_y and E_z are represented by blue, green and red lines respectively. The discontinuities in E_y located at ±25 μm occur because of the slab surfaces. (a)-(b) TE and TM profiles when θ is 0°. (c) and (d) The electric field profile when θ is 50°.

One can extract the group and phase effective refractive index (ERI) from dispersion curves like those shown in Fig. 2.9. The phase ERI can be retrieved directly from the dispersion curve, $n_p = c / v_p = ck/(2\pi f)$, and the group ERI can be retrieved from the slope of the dispersion curve, $n_g = c / v_g = c\Delta k/(2\pi\Delta f)$. Here v_p and v_g are the phase and group velocities, and the wave vector, k, and frequency, f, correspond to the axes in Fig. 2.9. From the derivation in Sec. 2.B.5, we can calculate both phase and group ERI for all propagation directions of TE-like and TM-like modes in the LN slab waveguide. Based on the agreement between experimental data and theoretical predictions shown in Fig. 2.9, Fig. 2.11 gives the theoretically calculated phase and group ERI for different angles, modes and frequencies. The ERI values are important for phase-matching in THz generation and for many nonlinear as well as linear optical processes.

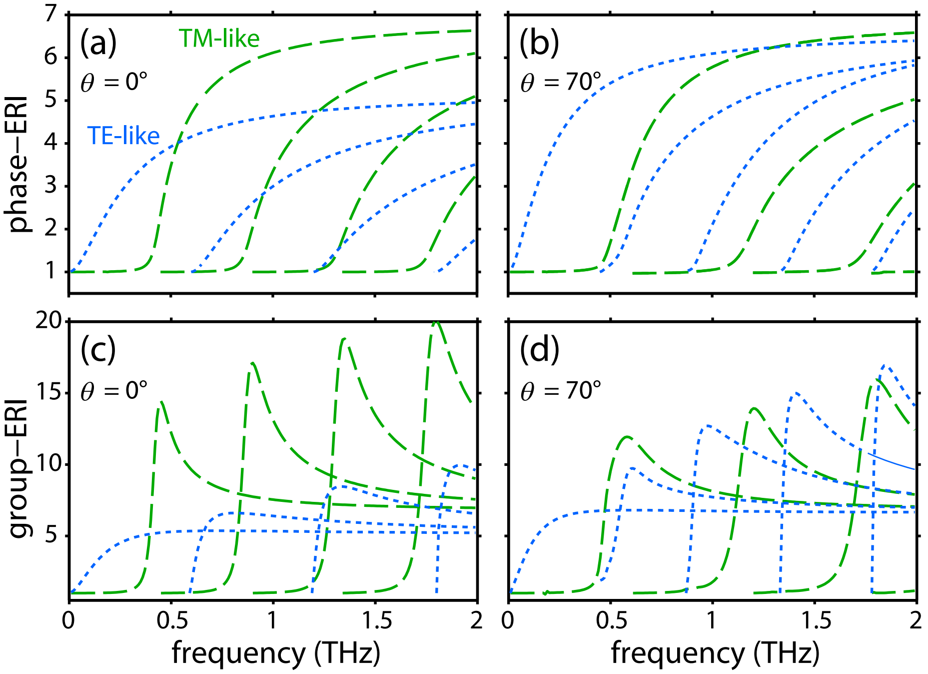

Figure 2.11 | The group and phase index of refraction. (a) The frequency- and mode-dependent phase ERI for TE-like (dotted blue) and TM-like (dashed green) modes when $\theta = 0°$. (b) The phase ERI when $\theta = 70°$. (c) and (d) are the same as (a) and (b), but for the group ERI.

In Fig. 2.11 dotted blue lines are the calculated ERI for TE-like modes and dashed green lines are the ERI for TM-like modes. From Fig. 2.11(a) & (b), we can see that the phase ERI for both TE-like and TM-like modes transitions from 1 (the index of air) to the bulk effective index. For the TM-like waves the bulk index is always the ordinary index of refraction, $n_o \approx 6.8$, while for the TE-like waves the bulk index is that for the extraordinary wave in the anisotropic material, and changes from ~5.1 at 0° to ~6.8 at 90°. At all angles, the low-frequency TM-like modes have most of their energy in the evanescent field in the air and have ERI values near unity. The ERI then transitions rapidly to bulk-like values at higher frequencies. Much like the phase ERI, the group ERI transitions from 1 to the bulk effective index [see Fig. 2.11 (c) & (d)]. In contrast to the phase ERI, however, the group ERI rises well above the bulk values before approaching them asymptotically at high frequencies. In contrast to the phase index, where higher modes always have lower ERIs, the group ERI is usually higher for higher modes. Another difference is that the peak group index changes drastically with θ for both TE-like and TM-like modes, while the peak phase ERI for the TM-like modes is just the bulk value and insensitive to angle.

A useful way to display the ERI is with an index ellipse, which highlights the angle-dependant behavior. Figure 2.12 follows the phase ERI for the first three TE-like modes at wave vector magnitude $\beta = 50$ rad/mm as a function of angle, tracing out the phase ERI ellipse. We measured data in the first quadrant, and because of the symmetry these results can also be used for 90$^{\mathrm{o}}$ to 360$^{\mathrm{o}}$. Values over 70$^{\mathrm{o}}$ were not recorded because the TE-like modes were too weak to be observed. Fig. 2.12 shows the measured values for the first three TE-like modes as open symbols and the calculated values predicted by the derivation in Sec. 2.B.5 as solid lines. The experimental data can be fit to an ellipse, where the long and short axes are 5.44 and 4.18 for the first mode, 3.36 and 2.59 for the second mode and 2.01 and 1.74 for the third mode. The value of the long axis represents the ERI for an ordinary wave (the TE mode is purely ordinary at 90°) and the short axis represents the ERI for an extraordinary wave (the TE mode is purely extraordinary at 0°).

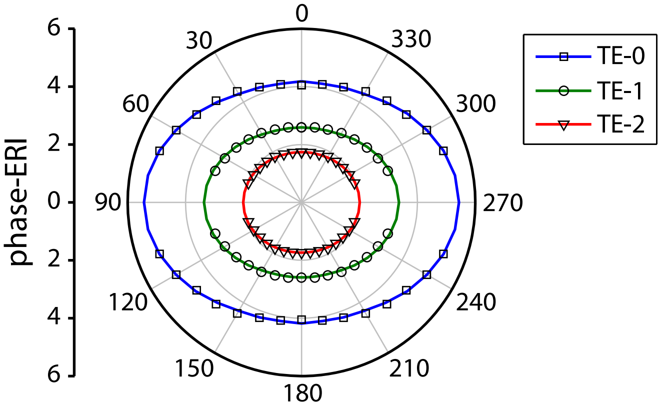

Figure 2.12 | The index ellipsoid. Effective refractive index (phase ERI) ellipse for three TE modes at a wave vector β = 50 rad/mm in a 50 µm LN slab waveguide. The open symbols are experimental data and the solid lines are calculated results. The scale along the x-axis is the same as that along y.

Because the 800 nm light was always polarized along the c-axis, we only pumped through the r_{33} electro-optic coefficient, which generated THz polarized along the optic axis [Feurer 2007]. The THz generated by the pump can be represented by a linear combination of waveguide modes, and the magnitude of the contribution from a given mode is related to the projection of its polarization along the c-axis. Thus, for a given mode and frequency, one can make a rough estimate of the relative pumping efficiency $\eta(\theta)$ by looking at the fraction of the mode energy corresponding to a field inside the crystal oriented along the optic axis:

$$\eta(\theta) = \frac{\int_{-\ell}^{\ell} \left[E_x^2(y)\sin^2\theta + E_z^2(y)\cos^2\theta\right]dy}{\int_{-\infty}^{\infty} \left[E_x^2(y) + E_y^2(y) + E_z^2(y)\right]dy} \tag{2.13}$$

Through the integration limits, the expression also takes into account the degree to which the mode is localized within the slab, which improves the efficiency since generation only occurs in the crystal, or is extended into the (air) cladding where no generation occurs.

Figure 2.13 shows η as a function of θ. When θ = 0°, the TE mode is polarized purely along the optic axis and is pumped most efficiently. As θ increases, the component of the TE wave along the optic axis slowly decreases. In contrast, the component of the TM wave along the optic axis increases, especially after 60°, and at 90°only the TM mode is pumped. At 70°, both modes are pumped with similar efficiencies. The qualitative trends in η explain the mode amplitudes observed in Fig. 2.9. For θ less than about 60°, TM-like modes are too weak to be observed, while for θ more than about 85°, the TE-like modes are not visible. As predicted, both kinds of modes are visible at 70° as shown in Fig. 2.9(c). Using different pump polarizations and reflective elements integrated into the waveguides, it will be possible to generate modes not observed in this study.

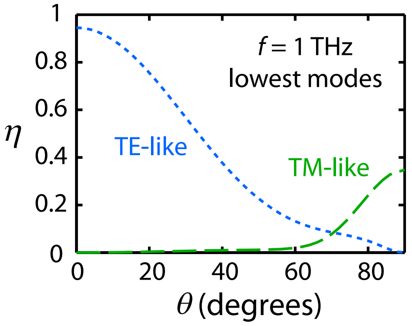

Figure 2.13 | Mode pumping efficiency. The fraction of total mode energy corresponding to a field inside the crystal polarized along the optic axis, η, which gives a rough prediction for pumping efficiency. The dotted blue line corresponds to the first symmetric, TE-like mode at 1 THz and the dashed green dashed line corresponds to the first antisymmetric, TM-like mode at the same frequency. As the angle increases, the TE-like mode becomes weaker while the TM-like mode grows in.

5. Conclusions

We have measured the propagation properties of THz waves in a 50 μm $LiNbO_3$ anisotropic slab waveguide using a self-compensating polarization gating imaging system. This system can detect the THz electric fields both temporally and spatially over a wide wavelength range. Using the system, we studied the propagation-direction-dependent behavior of waveguide modes and determined the dispersion curves and effective refractive index for THz waves. A general solution for waveguide modes in a uniaxial slab waveguide was derived and found to agree with the experimental data.

Dispersion is integral to many processes in THz science and generally in linear and nonlinear optics, including broadening of ultrashort pulses, walk-off between pump and probe pulses, phase-matching of parametric processes, and generation of optical solitons. Because dispersion in a waveguide is determined by both the intrinsic material dispersion and geometric dispersion, it is essential to understand waveguiding effects. The results presented here will facilitate the design of functional devices with new capabilities in the $LiNbO_3$ platform for integrated THz experiments and processing.

6. Solution to a uniaxial slab waveguide with isotropic cladding

Anisotropic slab waveguides were extensively studied in the 1970's [Wang 1973, Russo 1973, Burns 1974, Ramaswamy 1974, Nemoto 1977, Marcuse 1978, Marcuse 1979]. In many cases, attention was focused on anisotropic films deposited on a substrate that was itself anisotropic because mode converters, polarization mode filters, and other devices of that time had such a geometry [Burns 1974]. The case in this paper is somewhat simpler because the geometry is symmetric (see Fig. 2.14). An additional simplification is that the anisotropic core (the slab) is embedded within an isotropic cladding (air in our experiment). In the derivation of the waveguide dispersion curves and mode profiles presented below, we assume the experimentally relevant conditions that the crystal is uniaxial (like $LiNbO_3$) and its optic axis is parallel to the slab surface. The slab is assumed to extend infinitely along x and z and both core and cladding have no magnetic response. The wave is assumed to propagate along the x-direction and extend infinitely along the z-direction. In the experiment the cylindrical lens generating the wave was rotated instead of the sample, so the derivation here is performed in the coordinate frame of the lens. Finally, to simplify the analysis we assume that the waves are harmonic in space and along the propagation direction: $\vec{E}(x,y,z,t)=\vec{E}(y)\exp[i(\beta x-\omega t)]$, where $\beta = k_x$ is the propagation constant.

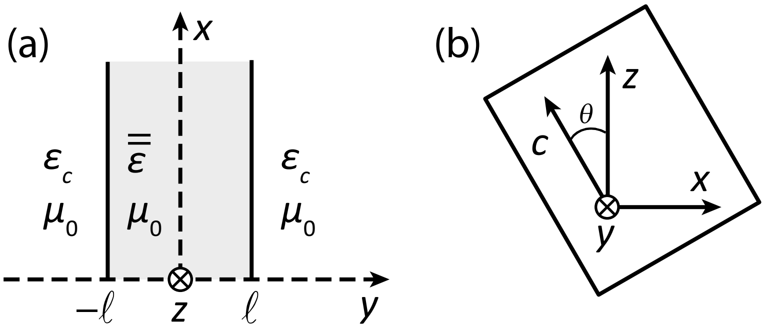

Figure 2.14 | The geometry for the waveguide mode derivation. (a) An anisotropic slab of width 2ℓ centered at $y = 0$ embedded in an isotropic cladding which extends to infinity. The bound wave propagates along x and extends infinitely along z. ε and μ are the permittivity and permeability in the different regions. (b) The coordinate system for the derivation is defined by the slab surface normal and the propagation direction, which differs from Fig. 2.6 where the coordinates are defined in the lab frame. θ is the angle between the z-axis and the optic axis of the crystal.

The derivation presented below will loosely follow the analysis of Marcuse and Kaminow [Marcuse 1979] where the more complicated symmetric geometry of an anisotropic slab with anisotropic cladding is studied. For the sake of brevity our analysis will skip some intermediate steps, many of which can be found in [Marcuse 1979] and are explained more fully in the simple case in Sec. 2.A. The first step in the derivation is to determine the characteristics of waves in bulk material, i.e. the dispersion curves and polarizations, in both core and cladding. Linear combinations of these bulk waves, constrained by system symmetry, are used to build the waveguide modes and lay out the general functional form of the solution. The boundary conditions at the waveguide surface generate a homogeneous system of equations which can be used to solve for the coefficients in the linear combination. Solutions exist for this system of equations, i.e. the determinant of the corresponding matrix is zero, only for certain pairs of frequency and propagation constant. These allowed solutions correspond to the waveguide dispersion curves.

To simplify notation we define several important variables. The propagation constant, $\beta \equiv k_x$, was defined above, and the wave vector orthogonal to the slab surface is defined both outside the crystal, $i\alpha \equiv k_y^{\text{out}}$, and for the ordinary and extraordinary waves inside the crystal, $\kappa_o, \kappa_e \equiv k_{yo}^{\text{in}}, k_{ye}^{\text{in}}$. α is defined as imaginary because bound modes will have evanescent, decaying fields in the cladding. There are three relevant bulk dispersion curves which define the relationships between wave vector, frequency, and index, one for the cladding and one each for the ordinary and extraordinary waves in the uniaxial core. They are:

$$\text{cladding:} \quad \alpha^2 = \beta^2 - k^2 n_c^2 \tag{2.14a}$$

$$\text{ordinary:} \quad \kappa_o^2 = k^2 n_o^2 - \beta^2 \tag{2.14b}$$

$$\text{extraordinary:} \quad \kappa_e^2 = k^2 n_e^2 - \beta^2 \left(\cos^2\theta + \frac{n_e^2}{n_o^2} \sin^2\theta \right) \tag{2.14c}$$

where $k = \omega / c$ is the wave vector in free space and $n_c, n_o, \text{and}\, n_e$ are the cladding index, ordinary index in the slab, and extraordinary index in the slab respectively. These relations are used to eliminate $\alpha, \kappa_o, \text{and}\, \kappa_e$ from the equations which follow, so everything is expressed in terms of β and k.

For a specific pair of β and k, there are four possible plane waves in each region, two signs for k_y and two polarizations. In the anisotropic medium, the polarizations correspond to the ordinary (later represented by $\vec{o}$) and extraordinary (later represented by $\vec{e}$) waves. In the cladding, any pair of orthogonal polarizations can be chosen, so for convenience we choose the TE polarization (represented below by $\vec{v}$ for vertical polarization) and the TM polarization (represented later by $\vec{h}$ for horizontal polarization). We can write out the most general form of the waveguide mode solution as:

$$\text{cladding:} \quad \vec{E}(y) = A_1 \vec{v}^- \exp[\alpha y] + A_2 \vec{h}^- \exp[\alpha y] + A_3 \vec{v}^+ \exp[-\alpha y] + A_4 \vec{h}^+ \exp[-\alpha y] \tag{2.15a}$$

$$\text{core:} \quad \vec{E}(y) = B_1 \vec{e}^+ \exp[i\kappa_e y] + B_2 \vec{o}^+ \exp[i\kappa_o y] + B_3 \vec{e}^- \exp[-i\kappa_e y] + B_4 \vec{o}^- \exp[-i\kappa_o y] \tag{2.15b}$$

$$\text{cladding,:} \quad \vec{E}(y) = C_1 \vec{v}^- \exp[\alpha y] + C_2 \vec{h}^- \exp[\alpha y] + C_3 \vec{v}^+ \exp[-\alpha y] + C_4 \vec{h}^+ \exp[-\alpha y] \tag{2.15c}$$

where A_i, B_i, and C_i are scalar constants and the +/- superscripts correspond to the sign of k_y.

The polarizations in the expression above can be determined from the appropriate vector constraints. In the cladding:

$$\vec{v}^\pm = \vec{v} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \vec{k}^\pm \times \vec{v} \propto \vec{h}^\pm \equiv \begin{bmatrix} \pm i h_x \\ -h_y \\ 0 \end{bmatrix} \propto \begin{bmatrix} \pm i\alpha \\ -\beta \\ 0 \end{bmatrix} \tag{2.16}$$

where h_x and h_y are the magnitudes of the components of the normalized polarization vector. In contrast to the isotropic cladding, where any orthogonal polarizations could be chosen, in the slab the polarizations are uniquely determined as the ordinary and extraordinary wave polarizations in bulk material. The ordinary wave will be orthogonal to the plane containing the crystal axis and the wave vector: $\vec{o} \propto \vec{k} \times \vec{c}$. The extraordinary wave will be located in the plane of $\vec{k}$ and $\vec{c}$. The displacement field will be given by $\vec{D} \propto \vec{k} \times (\vec{k} \times \vec{c})$, and the electric field is given through the constitutive relation: $\vec{E} = \overline{\overline{R}}(-\theta)\overline{\overline{\varepsilon}}^{-1}\overline{\overline{R}}(\theta)\vec{D}$ where $\overline{\overline{R}}$ is the rotation matrix for rotation around the y-axis. This yields:

$$\vec{o}^{\,\pm} \equiv \begin{bmatrix} \pm o_x \\ -o_y \\ \mp o_z \end{bmatrix} \propto \begin{bmatrix} \pm \kappa_o \cos\theta \\ -\beta\cos\theta \\ \mp \kappa_o \sin\theta \end{bmatrix} \tag{2.17a}$$

$$\vec{e}^{\,\pm} \equiv \begin{bmatrix} e_x \\ \mp e_y \\ e_z \end{bmatrix} \propto \begin{bmatrix} \left[\dfrac{1}{\varepsilon_o} - \dfrac{1}{\varepsilon_e}\right]\left(\beta^2 + \kappa_e^2\right)\cos^2\theta\sin\theta + \kappa_e^2\sin\theta\left[\dfrac{\cos^2\theta}{\varepsilon_o} + \dfrac{\sin^2\theta}{\varepsilon_e}\right] \\ \mp\dfrac{\beta\kappa_e\sin\theta}{\varepsilon_o} \\ \left[\dfrac{1}{\varepsilon_o} - \dfrac{1}{\varepsilon_e}\right]\kappa_e^2\sin^2\theta\cos\theta + \left(\beta^2 + \kappa_e^2\right)\cos\theta\left[\dfrac{\sin^2\theta}{\varepsilon_o} + \dfrac{\cos^2\theta}{\varepsilon_e}\right] \end{bmatrix} \tag{2.17b}$$

where $o_x, o_y, o_z, e_x, e_y, \text{and } e_z$ are the magnitudes of the components of the normalized polarization vectors.

With the dispersion curves (Eq. 2.14) and polarizations (Eqs. 2.16 & 2.17) of waves in the bulk material in hand, we can simplify the expressions in Eq. 2.15 for $\vec{E}(y)$. For bound solutions, we require that the electric field decays to zero as $y \to \pm\infty$, so the terms in the cladding that are exponentially growing can be discarded. We now apply the symmetry condition that there is a reflection plane down the center of the sample, which eliminates half of the coefficients. In this situation, the solution must be made of symmetric and antisymmetric modes. Absorbing some constant factors into the coefficients, we have:

Symmetric:

$$\text{cladding, } y < -\ell: \ \vec{E}(y) = \begin{bmatrix} h_x A_2 \\ -ih_y A_2 \\ A_1 \end{bmatrix} \exp[\alpha(y+\ell)] \tag{2.18a}$$

$$\text{core: } \vec{E}(y) = B_1 \begin{bmatrix} e_x\cos(\kappa_e y) \\ -ie_y\sin(\kappa_e y) \\ e_z\cos(\kappa_e y) \end{bmatrix} + B_2 \begin{bmatrix} -o_x\cos(\kappa_o y) \\ io_y\sin(\kappa_o y) \\ o_z\cos(\kappa_o y) \end{bmatrix} \tag{2.18b}$$

$$\text{cladding, } y > \ell: \ \vec{E}(y) = \begin{bmatrix} h_x A_2 \\ ih_y A_2 \\ A_1 \end{bmatrix} \exp[-\alpha(y-\ell)] \tag{2.18c}$$

Antisymmetric:

$$\text{cladding, } y < -\ell: \ \vec{E}(y) = \begin{bmatrix} h_x A_2 \\ -ih_y A_2 \\ A_1 \end{bmatrix} \exp[\alpha(y+\ell)] \tag{2.19a}$$

$$\text{core: } \vec{E}(y) = B_1 \begin{bmatrix} e_x\sin(\kappa_e y) \\ ie_y\cos(\kappa_e y) \\ e_z\sin(\kappa_e y) \end{bmatrix} + B_2 \begin{bmatrix} o_x\sin(\kappa_o y) \\ io_y\cos(\kappa_o y) \\ -o_z\sin(\kappa_o y) \end{bmatrix} \tag{2.19b}$$

$$\text{cladding, } y > \ell: \ \vec{E}(y) = \begin{bmatrix} -h_x A_2 \\ -ih_y A_2 \\ -A_1 \end{bmatrix} \exp[-\alpha(y-\ell)] \tag{2.19c}$$

Applying the symmetry conditions eliminated half the unknowns, so now we need only apply boundary conditions at one interface to solve for the coefficients. The boundary condition is that the tangential E and H fields must be continuous across the boundary [Born 1999]. Using Faraday's law and the fact that $\partial/\partial z = 0$, $\partial/\partial x = i\beta$, and $\partial/\partial t = -i\omega$ for our functional form, $\vec{E}(x,y,z,t) = \vec{E}(y)\exp[i(\beta x - \omega t)]$, we can express all the boundary conditions in terms of the electric field components:

$$E_{z,\text{clad}} = E_{z,\text{core}} \tag{2.20a}$$

$$\frac{\partial E_{z,\text{clad}}}{\partial y} = \frac{\partial E_{z,\text{core}}}{\partial y} \tag{2.20b}$$

$$E_{x,\text{clad}} = E_{x,\text{core}} \tag{2.20c}$$

$$i\beta E_{y,\text{clad}} - \frac{\partial E_{x,\text{clad}}}{\partial y} = i\beta E_{y,\text{core}} - \frac{\partial E_{x,\text{core}}}{\partial y} \tag{2.20d}$$

The constant coefficients in functional form of the solutions (Eqs. 2.18 & 2.19) must be chosen so the above boundary conditions are satisfied at the interface ($y = \ell$). They must be solved independently for the symmetric and antisymmetric modes. The four expressions above yield a set of homogeneous equations which can be recast in matrix notation.

Symmetric:

$$\begin{bmatrix} -1 & 0 & e_z \cos(\kappa_e \ell) & o_z \cos(\kappa_o \ell) \\ -\alpha & 0 & e_z \kappa_e \sin(\kappa_e \ell) & o_z \kappa_o \sin(\kappa_o \ell) \\ 0 & h_x & -e_x \cos(\kappa_e \ell) & o_x \cos(\kappa_o \ell) \\ 0 & h_y \beta - h_x \alpha & (e_y \beta + e_x \kappa_e)\sin(\kappa_e \ell) & -(o_y \beta + o_x \kappa_o)\sin(\kappa_o \ell) \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \\ B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{2.21}$$

Antisymmetric:

$$\begin{bmatrix} 1 & 0 & e_z \sin(\kappa_e \ell) & -o_z \sin(\kappa_o \ell) \\ \alpha & 0 & -e_z \kappa_e \cos(\kappa_e \ell) & o_z \kappa_o \cos(\kappa_o \ell) \\ 0 & h_x & e_x \sin(\kappa_e \ell) & o_x \sin(\kappa_o \ell) \\ 0 & h_y \beta - h_x \alpha & (e_y \beta + e_x \kappa_e)\cos(\kappa_e \ell) & (o_y \beta + o_x \kappa_o)\cos(\kappa_o \ell) \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \\ B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{2.22}$$

The polarizations (Eqs. 2.16 & 2.17) and bulk dispersion curves (Eq. 2.14) can be used to remove all dependence on $\alpha, \kappa_o, \text{and } \kappa_e$, so for a given angle θ, the only variables are β and k. The determinant will be zero, i.e. the set of equations has a solution, only for β and k pairs that are on the waveguide dispersion curve, and finding all allowed pairs traces out these curves. Using the allowed pairs, the bulk dispersion curves, and one additional "normalization condition" such as $B_1 + B_2 = 1$, all wave vectors and coefficients can be completely determined. The theoretical dispersion curves for several angles are plotted along with the experimental data in Fig. 2.9, selected electric field profiles are shown in Fig. 2.10, and the effective indices of refraction for two angles are shown in Fig. 2.11.

C. REFERENCES

[Auston 1984] D. H. Auston, K. P. Cheung, J. A. Valdmains, and D. A. Kleinman, “Cherenkov Radiation from Femtosecond Optical Pulses in Electro-Optic Media,” *Phys. Rev. Lett.* **53**, 1555-1558 (1984).

[Auston 1988] D. Auston and M. Nuss, “Electrooptic Generation and Detection of Femtosecond Electrical Transients,” *IEEE J. Quantum Electron.* **24**, 184-197 (1988).

[Barker 1967] A. S. Barker, JR, and R. Loudon, “Dielectric properties and optical phonons in $LiNbO_3$,” *Phys. Rev.* **158**, 433-445 (1967).

[Born 1999] M. Born and E. Wolf, *Principles of Optics,* 2nd ed. (Cambridge University Press, Cambridge, 1999).

[Born 1999] M. Born and E. Wolf, *Principles of Optics,* 2nd ed. (Cambridge University Press, Cambridge, 1999).

[Burns 1974] W . K. Burns and J. Warner. “Mode dispersion in uniaxial optical waveguides,” *J. Opt. Soc., Amer.,* **64**, 441-446, (1974).

[Cronin 1995] N. J. Cronin, *Microwave and Optical Waveguides* (Institute of Physics, 1995).

[Dougherty 1992] T. P. Dougherty, G. P. Wiederrecht, and K. A. Nelson, “Impulsive Stimulated Raman Scattering Experiments in The Polariton Regime,” *J. Opt. Soc. Am.* **9**. 2179-2189 (1992).

[Feurer 2003] T. Feurer, J. C. Vaughan, and K. A. Nelson, “Spatiotemporal Coherent Control of Lattice Vibrational Waves,” *Science* 299, 374-377 (2003).

[Feurer 2007] T. Feurer, N. S. Stoyanov, D. W. Ward, J. C. Vaughan, E. R. Statz, and K. A. Nelson, “Terahertz Polaritonics,” *Annu. Rev. Mater. Res.* **37**, 317-350 (2007).

[Hebling 2004] J. Hebling, A. G. Stepanov, G. Almási, B. Bartal, and J. Kuhl, "Tunable THz pulse generation by optical rectification of ultrashort laser pulses with tilted pulse fronts," *Appl. Phys. B* **78,** 593-599 (2004).

[Koehl 1999] R. M. Koehl, S. Adachi, and K. A. Nelson, “Real-Space Polariton Wave Packet Imaging,” *J. Chem. Phys.* 110, 1317-1320, (1999).

[Lee 2000] Y. S. Lee, T. Meade, V. Perlin, H. Winful, T. B. Norris, and A. Galvanauskas, "Generation of narrow-band terahertz radiation via optical rectification of femtosecond pulses in periodically poled lithium niobate," *Appl. Phys. Lett.* **76**, 2505-2507 (2000).

[Lee 2009] Yun-Shik Lee, *Principles of Terahertz Science and Technology* (Springer, New York, 2009).

[Lin 2009] K.-H. Lin, C. A. Werley, K. A. Nelson, “Generation of multicycle THz phonon-polariton waves in a planar waveguide by tilted optical pulse fronts,” *Appl. Phys. Lett.* **95**, 103304 (2009).

[Marcuse 1978] D. Marcuse, “Modes of a symmetric slab optical waveguide in birefringent media-Part I: Optical axis not in plane of slab,” *IEEE J. Quantum Electron.* **QE-14**, 736-741, (1978).

[Marcuse 1979] D. Marcuse and I. P. Kaminow, “Modes of a Symmetric Slab Optical Waveguide in Birefringent Media Part II: Slab with Coplanar Optical Axis,” *IEEE J. Quantum Electron.* **QE-15**, 92-101, (1979).

[Nemoto 1977] S. Nemoto and T. Makimoto, “Further discussion of the relationship between phase and group indices in anisotropic inhomogeneous guiding media,” *J. Opt. Soc. Am.* **67**, 1281-1283 (1977).

[Peier 2008] P. Peier, S. Pilz, F. Müller, K. A. Nelson, and T. Feurer, "Analysis of phase contrast imaging of terahertz phonon-polaritons," *J. Opt. Soc. Am. B* **25**, B70-B75 (2008).

[Ramaswamy 1974] V. Ramaswamy, "Propagation in asymmetric anisotropic film waveguides," *Appl. Opt.* **13**, 1363-1371 (1974).

[Russo 1973] D. P. Gia Russo and J. H. Harris, "Wave propagation in anisotropic thin-film optical waveguides," *J. Opt. Soc. Am.* **63**, 138-145 (1973).

[Saleh 2007] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics, 2nd Ed.* (Wiley, 2007).

[Shadrivov 2003] I. V. Shadrivov, A. A. Sukhorukov, and Y. S. Kivshar. "Guided modes in negative-refractive-index waveguides," *Phys. Rev. E* **67**, 057602 (2003).

[Stoyanov 2002] N. S. Stoyanov, D. W. Ward, T. Feurer, and K. A. Nelson, "Terahertz Polariton Propagation in Patterned Materials," *Nat. Mat.* **1**, 95-98 (2002).

[Stoyanov 2003] N. S. Stoyanov, T. Feurer, D. W. Ward, and K. A. Nelson, "Integrated diffractive terahertz elements," *Appl. Phys. Lett.* **82**, 674-676 (2003).

[Stoyanov 2004] N. S. Stoyanov, T. Feurer and K. A. Nelson, "Direct visualization of a polariton resonator in the THz regime," *Opt. Express* **12**, 2387-2396 (2004).

[Wahlstrand 2003] J. K. Wahlstrand and R. Merlin, "Cherenkov radiation emitted by ultrafast laser pulses and the generation of coherent polaritons," *Phys. Rev. B* **68**, 054301 (2003).

[Wang 1972] S. Wang, M. L. Shah, and J. D. Crow, "Wave propagation in thin film optical waveguides using gyrotropic and anisotropic materials as substrates," *IEEE J. Quantum Electron.,* **QE-8**, 212-216 (1972).

[Werley 2010] C. A. Werley, Q. Wu, K. H. Lin, C. R. Tait, A. Dorn, and K. A. Nelson, "A comparison of phase sensitive imaging techniques for studying THz waves in structured $LiNbO_3$," *J. Opt. Soc. Am. B* **27**, 2350-2359 (2010).

[Werley 2012] C. A. Werley, C. R. Tait, and K. A. Nelson, "Direct visualization of THz-frequency electromagnetic waves in classic experimental geometries," *Am. J. Phys.* **80**, 72-81 (2012).

[Wu 2003] B.-I. Wu, T. M. Grzegorczyk, Y. Zhang, and J. A. Kong. "Guided modes with imaginary transverse wave number in a slab waveguide with negative permittivity and permeability," *J. Appl. Phys.* **93**, 9386-9388 (2003).

[Wu 2009] Q. Wu, C. A. Werley, K. H. Lin, A. Dorn, M. G. Bawendi, and K. A. Nelson, "Quantitative phase contrast imaging of THz electric fields in a dielectric waveguide," *Opt. Express* **l**(17), 9219-9225 (2009).

[Yan 1985] Y. X. Yan, E. B. Gamble, and K. A. Nelson, "Impulsive Stimulated Scattering: General Importance in Femtosecond Laser Pulse Interactions with Matter, and Spectroscopic Applications," *J. Chem. Phys.* **83**. 5391-5399 (1985).

[Yang 2010] C. Yang, Q. Wu, J. Xu, K. A. Nelson, and C. A. Werley. "Experimental and theoretical analysis of THz-frequency, direction-dependent, phonon polariton modes in a subwavelength, anisotropic slab waveguide," *Opt. Express* **18**, 26351-26354 (2010).

[Yeh 2007] K. L. Yeh, M. C. Hoffmann, J. Hebling, and Keith A. Nelson, "Generation of 10 μJ ultrashort terahertz pulses by optical rectification," *Appl. Phys. Lett.* **90**, 171121 (2007).

Chapter III

The electro-optic effect in lithium niobate, electro-optic detection, and ion displacements in lithium niobate

A. ELECTRO-OPTIC EFFECT IN LITHIUM NIOBATE

The electro-optic effect is the mechanism through which a DC or low-frequency field changes the optical index of refraction in a crystal. The applied field deforms the electron cloud in the crystal, which modifies its response to optical light. A treatment can be found in nonlinear optics or photonics textbooks such as [Boyd 2008] or [Yariv 2007]. The electro-optic effect is of particular interest to us because a THz field is low-frequency relative to the optical probe we use to detect THz waves, so the formalism is appropriate. As a result, the electro-optic effect determines the magnitudes of the signals we see, and can be used to quantify THz field amplitudes based on the recorded pump-probe signals.

1. The electro-optic tensor

In a crystal, the displacement field is related to the electric field via the dielectric permittivity tensor:

$$\vec{D}_{\text{opt}} = \varepsilon_0 \overline{\overline{\varepsilon}} \vec{E}_{\text{opt}} \tag{3.1}$$

The subscript "opt" indicates the that these fields are optical fields in our experiments, not to be confused with the DC or THz fields discussed later. Equation 3.1 is one of two constitutive relations, the other being $\vec{B}_{\text{opt}} = \mu_0 \overline{\overline{\mu}} \vec{H}_{\text{opt}}$, which along with Maxwell's relations govern the propagation of light in materials. In lossless materials, $\overset{=}{\varepsilon}$ is a real, symmetric matrix, and in the appropriate coordinate system, it is diagonal. For lithium niobate or any uniaxial crystal:

$$\overline{\overline{\varepsilon}} = \begin{bmatrix} \varepsilon_o & 0 & 0 \\ 0 & \varepsilon_o & 0 \\ 0 & 0 & \varepsilon_{eo} \end{bmatrix} = \begin{bmatrix} n_o^2 & 0 & 0 \\ 0 & n_o^2 & 0 \\ 0 & 0 & n_{eo}^2 \end{bmatrix} \tag{3.2}$$

with n_o the ordinary index of refraction and n_{eo} the extraordinary index.

The effect of an applied field is most conveniently expressed in terms of the impermeability tensor, so I will introduce this tensor here. The impermittivity tensor is found by inverting Eq. 3.1:

$$\vec{E}_{\text{opt}} = \frac{1}{\varepsilon_0} \overline{\overline{\eta}} \vec{D}_{\text{opt}}. \tag{3.3}$$

$\overline{\overline{\eta}}$ is the matrix inverse of $\overline{\overline{\varepsilon}}$ and is real and symmetric for a lossless material. In the coordinate system where the permittivity tensor is diagonal, the impermittivity tensor is simply:

$$\overline{\overline{\eta}} = \begin{bmatrix} 1/\varepsilon_o & 0 & 0 \\ 0 & 1/\varepsilon_o & 0 \\ 0 & 0 & 1/\varepsilon_{eo} \end{bmatrix} = \begin{bmatrix} 1/n_o^2 & 0 & 0 \\ 0 & 1/n_o^2 & 0 \\ 0 & 0 & 1/n_{eo}^2 \end{bmatrix}. \tag{3.4}$$

Now that we have the basic structure of the matrices, we consider what happens when a field is applied to the crystal. $\overline{\overline{\eta}}$ will be modified, and to lowest order in the applied field (i.e. the linear electro-optic effect) it will be:

$$\overline{\overline{\eta}} = \overline{\overline{\eta}}_0 + \overline{\overline{\Delta\eta}} = \overline{\overline{\eta}}_0 + \mathbf{r}\vec{\mathrm{E}}. \tag{3.5}$$

Here $\overline{\overline{\eta}}_0$ is the impermittivity tensor in the absence of an applied field. $\vec{\mathrm{E}}$ is the THz electric field, not to be confused with the optical electric field, E_{opt}, in Eqs. 3.1 and 3.3. **r** is the electro-optic tensor, a rank three tensor, which can be represented by a 3x3x3 matrix. This means that every element in $\overline{\overline{\Delta\eta}}$ can in principle be modified by each polarization of the electric field. Written out completely, we have:

$$\overline{\overline{\Delta\eta}} = \begin{bmatrix} r_{111} & r_{121} & r_{131} \\ r_{211} & r_{221} & r_{231} \\ r_{311} & r_{321} & r_{331} \end{bmatrix} \mathrm{E}_x + \begin{bmatrix} r_{112} & r_{122} & r_{132} \\ r_{212} & r_{222} & r_{232} \\ r_{312} & r_{322} & r_{332} \end{bmatrix} \mathrm{E}_y + \begin{bmatrix} r_{113} & r_{123} & r_{133} \\ r_{213} & r_{223} & r_{233} \\ r_{313} & r_{323} & r_{333} \end{bmatrix} \mathrm{E}_z. \tag{3.6}$$

The third index in r_{ijk} corresponds to the direction along which the DC or THz field is applied ($1 \rightarrow x$, $2 \rightarrow y$, $3 \rightarrow z$), and the first two indices correspond to the elements in the impermittivity tensor. Because $\overline{\overline{\eta}}$ is symmetric, r_{ijk} must be symmetric across its first two indices: $r_{ijk} = r_{jik}$. Said another way, each of the three matrices in Eq. 3.6 is symmetric, and there are only 6 potentially independent values in the electro-optic tensor for each 9 elements. Standard notation in the field takes advantage of this symmetry to reassign the first two indices to a single index in a fairly arbitrary way. The mapping is:

$$\begin{aligned} r_{11k} &\equiv r_{1k} \\ r_{22k} &\equiv r_{2k} \\ r_{33k} &\equiv r_{3k} \\ r_{23k} &= r_{32k} \equiv r_{4k} \\ r_{13k} &= r_{31k} \equiv r_{5k} \\ r_{12k} &= r_{21k} \equiv r_{6k} \end{aligned}. \tag{3.7}$$

Plugging the reassigned values from Eq. 3.7 into Eq. 3.6, we get:

$$\overline{\overline{\Delta\eta}} = \begin{bmatrix} r_{11} & r_{61} & r_{51} \\ r_{61} & r_{21} & r_{41} \\ r_{51} & r_{41} & r_{31} \end{bmatrix} \mathrm{E}_x + \begin{bmatrix} r_{12} & r_{62} & r_{52} \\ r_{62} & r_{22} & r_{42} \\ r_{52} & r_{42} & r_{32} \end{bmatrix} \mathrm{E}_y + \begin{bmatrix} r_{13} & r_{63} & r_{53} \\ r_{63} & r_{23} & r_{43} \\ r_{53} & r_{43} & r_{33} \end{bmatrix} \mathrm{E}_z. \tag{3.8}$$

There are potentially 18 independent electro-optic values, and to save space they are always written as a 6x3 matrix:

$$\mathbf{r} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \\ r_{41} & r_{42} & r_{43} \\ r_{51} & r_{52} & r_{53} \\ r_{61} & r_{62} & r_{63} \end{bmatrix}. \tag{3.9}$$

The number of unique and non-zero components in **r** is determined by the crystal symmetry. In a triclinic crystal, the lowest-symmetry type, all 18 coefficients are non-zero and they are all different, but most crystals have higher symmetry and simpler electro-optic tensors. See [Yariv 2007] for the nonzero elements in the electro-optic tensor and the relationships between these elements for all crystal symmetry groups. Lithium niobate is a 3*m* trigonal crystal, and its electro-optic tensor is (there is a sign error in [Boyd 2008]):

$$\mathbf{r} = \begin{bmatrix} 0 & -r_{22} & r_{13} \\ 0 & r_{22} & r_{13} \\ 0 & 0 & r_{33} \\ 0 & r_{51} & 0 \\ r_{51} & 0 & 0 \\ -r_{22} & 0 & 0 \end{bmatrix}. \tag{3.10}$$

Note that most of the nonzero elements are equal to others. Using the definition of the $\overline{\overline{\eta}}_0$ (Eq. 3.4), the expression for $\overline{\overline{\eta}}$ in the presence of an external field (Eq. 3.5 & 3.8), and the electro-optic tensor for lithium niobate (Eq. 3.10), we get the full impermittivity tensor in $LiNbO_3$ in the presence of an applied DC or THz field:

$$\overline{\overline{\eta}} = \begin{bmatrix} 1/n_o^2 - r_{22}\mathrm{E}_y + r_{13}\mathrm{E}_z & -r_{22}\mathrm{E}_x & r_{51}\mathrm{E}_x \\ -r_{22}\mathrm{E}_x & 1/n_o^2 + r_{22}\mathrm{E}_y + r_{13}\mathrm{E}_z & r_{51}\mathrm{E}_y \\ r_{51}\mathrm{E}_x & r_{51}\mathrm{E}_y & 1/n_{eo}^2 + r_{33}\mathrm{E}_z \end{bmatrix}. \tag{3.11}$$

Multiple experiments have measured the values of the electro-optic tensor components for $LiNbO_3$, and there is significant spread in the measured values [Wong 2002]. The preparation of the crystal can matter, for example whether it is grown from a congruent melt or stoichiometric melt and what dopants are present in the crystal. For THz experiments, we typically use ~5% MgO doped stoichiometric melt $LiNbO_3$. In addition, the frequency of all three light components matter [Wong 2002]. In particular, when the applied field is at frequencies below the mechanical resonance of the crystal being tested (unclamped measurement), the entire crystal can expand or contract via the piezo-electric response, which changes the index. At frequencies above the mechanical resonance (clamped conditions), only the electronic response matters. THz frequencies are well above the mechanical resonance condition, so clamped values should be used. Table 3.1 contains my best choices for the appropriate refractive indices and electro-optic coefficients.

	400 nm	532 nm	800 nm		cm/kV
n_o	2.43	2.325	2.26	r_{51}	25e-7
n_{eo}	2.32	2.23	2.175	r_{22}	3.4e-7
				r_{13}	9e-7
				r_{33}	31e-7

Table 3.1 | LiNbO$_3$ indices and electro-optic coefficients. The ordinary and extraordinary refractive indices at three optical wavelengths and the electro-optic coefficients for visible light [Wong 2002]. The electro-optic tensor values are for undoped LiNbO$_3$ at room temperature grown from a stoichiometric melt; values did not appear to change upon doping with MgO. Refractive index values are for an MgO-doped crystal.

2. The optical indicatrix

Equation 3.11 gives the full impermittivity tensor, but what is relevant experimentally (particularly for THz detection) is the change in the index of refraction. Propagation in an anisotropic medium is often described by the optical indicatrix, also commonly referred to as the index ellipsoid or the ellipsoid of wave normals. The definitive description can be found in [Born 1999]. It corresponds to a surface of constant energy density, and it is a 3D ellipsoid described by:

$$\eta_1 x^2 + \eta_2 y^2 + \eta_3 z^2 + 2\eta_4 yz + 2\eta_5 xz + 2\eta_6 xy = 1 \tag{3.12}$$

When no field is applied and $\overline{\overline{\varepsilon}}$ and $\overline{\overline{\eta}}$ are diagonal, the ellipsoid has its principal axes along x, y, and z. In LiNbO$_3$, the ellipsoid in the absence of an applied field is an oblate spheroid (shaped like an M&M candy) because the extraordinary index is smaller than the two ordinary indices. When a field is applied, the lengths of the axes change (due to the on-diagonal terms), and the principal axes rotate slightly away from the axes of the coordinate system (because of the off-diagonal terms). For light propagating through a material along a vector $\vec{v}$, one slices the indicatrix with a plane perpendicular to $\vec{v}$ passing through the origin. The intersection between the indicatrix and the plane is an ellipse. The directions of the major and minor axes lie along the new primary axes of the crystal, and their magnitudes indicate the impermittivity, which can be used to retrieve the refractive indices.

3. Electro-optic index changes in LiNbO$_3$

The electro-optic effect in lithium niobate is discussed in some detail in [Lenzo 1966]. The indicatrix is found easily by substituting the matrix elements of $\overline{\overline{\eta}}$ (Eq. 3.11) into Eq. 3.12. This gives:

$$\begin{aligned}(1/n_o^2 - r_{22}\mathrm{E}_y + r_{13}\mathrm{E}_z)x^2 + (1/n_o^2 + r_{22}\mathrm{E}_y + r_{13}\mathrm{E}_z)y^2 + (1/n_{eo}^2 + r_{33}\mathrm{E}_z)z^2 \\ + 2r_{51}\mathrm{E}_y yz + 2r_{51}\mathrm{E}_x xz - 2r_{22}\mathrm{E}_x xy = 1.\end{aligned} \tag{3.13}$$

In all the measurements described in my thesis, the samples are thin slabs of x-cut LiNbO$_3$, diagrammed in Fig. 3.1. In experiments, the optic axis (along z) is typically aligned vertically in the lab frame. The probe beam propagates perpendicular to the slab surface, along x.

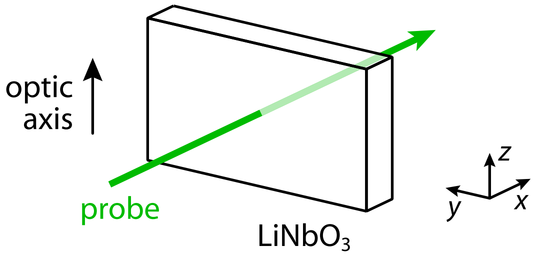

Figure 3.1 | Crystal orientation in experiments. In the experiments in this thesis, the $LiNbO_3$ slabs are *x*-cut, and the probe propagates along *x*, perpendicular to the slab surface. The optic axis is aligned vertically, along *z*. In a typical experiment, the THz wave propagates along *y* and is polarized along *z*.

Because the optical light propagates along *x*, we take a plane perpendicular to *x* passing through the origin, and find the intersection between this plane and the ellipsoid in Eq. 3.13. In this geometry, we simply set $x = 0$ in Eq. 3.13, which yields:

$$(1/n_o^2 + r_{22}\mathrm{E}_y + r_{13}\mathrm{E}_z)y^2 + 2r_{51}\mathrm{E}_y yz + (1/n_{eo}^2 + r_{33}\mathrm{E}_z)z^2 = 1. \qquad (3.14)$$

This equation describes an ellipse with its major and minor axes rotated slightly away from *y* and *z* respectively. The *z*-component of the THz field shows up only in diagonal terms, so it changes the length of the axes but does not rotate the coordinate system. In contrast, the *y*-component affects both on- and off-diagonal terms, so it changes the axis length and rotates the coordinate system.

An introductory textbook on calculus which describes the conic sections [Larson 1990] shows how to transform an ellipse described by $Ay^2 + Byz + Cz^2 = 1$ (axes NOT parallel to the coordinate axes) to an ellipse described by $A'y^2 + C'z^2 = 1$ (axes parallel to the coordinate axes). The system should be rotated by an angle

$$\theta = \frac{1}{2}\arctan\left(\frac{B}{A-C}\right). \qquad (3.15)$$

The new coefficients are:

$$\begin{aligned} A' &= A\cos^2\theta + B\cos\theta\sin\theta + C\sin^2\theta \\ C' &= A\sin^2\theta - B\cos\theta\sin\theta + C\cos^2\theta \end{aligned} \qquad (3.16)$$

To determine *θ*, *A*', and *C*', we must calculate *A*, *B*, and *C*. Using the values from Table 3.1 and assuming a 532 nm optical probe and THz field strength of E_y = 20 kV/cm, which is the largest peak field strength we have observed for any THz polarization without enhancement from a metal microstructure, $\theta = 0.2°$. This is really too small to be observed experimentally and is within the error of the initial alignment of the slab's axes, so we can approximate the rotation angle to be zero. $\sin(\theta)$ = 3e-3, so the induced change in impermittivity resulting from the off-diagonal component $[\Delta\eta_{\text{off diagonal}} = B\sin(\theta)\cos(\theta) = 2r_{51}E_y\sin(\theta)\cos(\theta)]$ is much smaller than the on-diagonal modifications to the index $(\Delta\eta_{\text{on diagonal}} = r_{22}E_y)$ and can be neglected. As a result, we have:

$$\begin{aligned}\Delta\eta_2 &= \Delta\left(\frac{1}{n_o^2}\right) = r_{22}\mathrm{E}_y + r_{13}\mathrm{E}_z \\ \Delta\eta_3 &= \Delta\left(\frac{1}{n_{eo}^2}\right) = r_{33}\mathrm{E}_z.\end{aligned} \tag{3.17}$$

Finally, we use differentiation by parts to get Δn:

$$\Delta n = \left(\frac{d\eta}{dn}\right)^{-1}\Delta\eta = -\frac{n^3}{2}\Delta\eta. \tag{3.18}$$

It seems common in the literature to neglect the negative sign in Eq. 3.18, and in many cases the sign of the electro-optic coefficient is not known so the sign is somewhat irrelevant. I will keep the negative sign here for accuracy. Using Eqs. 3.17 and 3.18 we get the final result for *x*-cut LiNbO$_3$:

$$\begin{aligned}\Delta n_o &= -\frac{n_o^3}{2}(r_{22}\mathrm{E}_y + r_{13}\mathrm{E}_z) \\ \Delta n_e &= -\frac{n_{eo}^3}{2}r_{33}\mathrm{E}_z.\end{aligned} \tag{3.19}$$

It is useful for develop a feeling for the order of magnitude of the effects that this will introduce. As will be discussed more fully in the next section and the next chapter, the signal magnitude in our measurements is related to the induced phase shift. For phase contrast imaging:

$$\left(\frac{\Delta I}{I_0}\right)_{\mathrm{PC}} = 2\Delta\varphi_{\mathrm{PC}} = 4\pi\frac{\ell}{\lambda_{\mathrm{opt}}}\Delta n_{eo}, \tag{3.20}$$

with $\Delta\varphi_{\mathrm{PC}}$ the induced phase shift for a particular probe polarization and ℓ the slab thickness. For polarization gating imaging with balancing:

$$\left(\frac{\Delta I}{I_0}\right)_{\mathrm{PG}} = 2\Delta\varphi_{\mathrm{PG}} = 4\pi\frac{\ell}{\lambda_{\mathrm{opt}}}(\Delta n_{eo} - \Delta n_o), \tag{3.21}$$

with $\Delta\varphi_{\mathrm{PG}}$ the phase shift between vertically and horizontally polarized components of the optical light. If we assume the slab is 50 μm thick, the probe wavelength is 532 nm, and $\mathrm{E}_y = 0$, we get:

$$\begin{aligned}\left(\frac{\Delta I}{I_0}\right)_{\mathrm{PC}} &= -0.020 * \mathrm{E}_z \\ \left(\frac{\Delta I}{I_0}\right)_{\mathrm{PG}} &= -0.014 * \mathrm{E}_z\end{aligned} \tag{3.22}$$

with E_z the THz field strength in kV/cm.

B. ELECTRO-OPTIC AND OPTICAL KERR EFFECT DETECTION

Two mechanisms through which light or an electric field can change a material's index of refraction are the electro-optic effect and the optical Kerr effect (OKE). The electro-optic effect is where an applied electric field changes the index of refraction of a material. The index change is related to the applied field by the electro-optic tensor (see a nonlinear optics text such as [Boyd 2008] for details). The optical Kerr effect is where an optical pulse changes the index of refraction, but the index change depends on the light intensity, not the field strength. In both of these effects, the anisotropy of the material changes. Thus to detect the induced change in index, it is natural to place the sample (in which the E-field or pump pulse is present) between crossed polarizers. This section explains how to perform this measurement carefully in order to quantitatively extract the induced change in index. This is particularly relevant to our work because the E-field of the THz pulse can be the applied field, and using the electro-optic effect is important for recording THz time traces and quantifying THz field strengths. To track the polarization state of light through an optical system the Jones matrix or Jones calculus formalism, described in the next section, is very convenient.

1. Jones Matrices

A good description of the Jones calculus formalism can be found in [Yariv 2007] and in Wikipedia, but I will summarize here and include some extra matrices that are particularly relevant to our studies. In the Jones calculus formalism, the electric field is represented by a 2x1 complex vector. The magnitudes of the first and second components describe the amplitude of the field polarized horizontally and vertically, respectively. The phases of the two components describe the phases of the electric fields. In the most general case the Jones vector is:

$$\mathbf{v} = \begin{bmatrix} E_x \exp(i(kz - \omega t + \phi_x)) \\ E_y \exp(i(kz - \omega t + \phi_y)) \end{bmatrix}. \tag{3.23}$$

Often the absolute amplitude and phase are not important, so these are factored out (or discarded) and the first component is set equal to 1:

$$\mathbf{v} = E_x \begin{bmatrix} 1 \\ (E_y / E_x)\exp(i(\phi_y - \phi_x)) \end{bmatrix}. \tag{3.24}$$

The intensity that would be detected, for instance by a photodiode, is calculated by multiplying the Jones vector by its conjugate transpose $I = \mathbf{v}^{*T}\mathbf{v} = E_x^2 + E_y^2$. Some common, important polarization states (with the amplitude dropped) are summarized in the table below.

Vertical:	$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$	Horizontal:	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$
+45°:	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	-45°:	$\begin{bmatrix} 1 \\ -1 \end{bmatrix}$
Right circular:	$\begin{bmatrix} 1 \\ -i \end{bmatrix}$	Left circular:	$\begin{bmatrix} 1 \\ i \end{bmatrix}$

Table 3.2 | Jones vectors for common polarization states

The polarization states are intuitive: vertical polarization has the horizontal polarization set equal to zero. Circularly polarized light has the same amplitude for vertical and horizontal polarizations, but there is a 90° phase shift between them.

When light propagates through an optic the polarization state is modified, and each optic can be described by a 2x2 complex matrix. Below is a set of matrices for common optics (the amplitudes have been dropped).

Vertical polarizer	$\mathbf{P_v} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$	Horizontal polarizer	$\mathbf{P_h} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$
Half wave plate (HWP)	$\mathbf{H} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	Quarter wave plate (QWP)	$\mathbf{Q} = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}$
Mirror not at 0°	$\mathbf{M} = \begin{bmatrix} R_p e^{i\varphi_p} & 0 \\ 0 & R_s e^{i\varphi_s} \end{bmatrix} \equiv \begin{bmatrix} 1 & 0 \\ 0 & Ae^{i\alpha} \end{bmatrix}$	Lens, mirror at 0°, isotropic substrate at 0°	$\mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
Birefringent slab	$\mathbf{B} = \begin{bmatrix} e^{i\varphi_x} & 0 \\ 0 & e^{i\varphi_y} \end{bmatrix} \equiv \begin{bmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{bmatrix}$	Isotropic slab with induced birefringence	$\mathbf{S} = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\Delta\phi} \end{bmatrix}$

Table 3.3 | Jones matrices for common optics. In all cases it is assumed that the principal axes of the optic are aligned along the *x*- and *y*-axes in the lab frame. For instance, this means the fast axis of the QWP, HWP, and birefringent slab are aligned vertically or horizontally. The mirror is in a vertical plane so that the *p*- and *s*-polarizations are along x and y. For a good mirror A is close to 1 and α is small. Typically for the optical pulse or field induced birefringence, $\Delta\phi$ is small.

The last important concept in Jones calculus is what happens when an optic is rotated so its principal axes are not along x and y. In this case we must first rotate the frame to the principal axes of the optic, apply the matrix for that optic, and then rotate back to the lab frame:

$$\mathbf{O}(\theta) = \mathbf{R}(\theta)\mathbf{O}\mathbf{R}(-\theta) \tag{3.25}$$

with

$$\mathbf{R}(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}. \tag{3.26}$$

Below is a table with polarizers and wave plates at 45°, the most common angle other than 0 and 90°.

Polarizer, +45°	$\mathbf{P}_{+45} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$	Polarizer, -45°	$\mathbf{P}_{-45} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$
Half wave plate, +45°	$\mathbf{H}_{45} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$	Quarter wave plate, +45°	$\mathbf{Q}_{45} = \begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix}$

Table 3.4 | Jones matrices for common, rotated optics. Here the optics have been rotated 45° relative to the lab frame. As before, the amplitude in front of the matrix has been omitted.

2. Electro-optic detection: the ideal case

To develop intuition for how detection of induced birefringence works, we will study the simplest case. A representative experimental setup is shown in Fig. 3.2. Here, a vertically polarized optical pump pulse is sent into glass (fused silica). The vertical index is changed by the optical Kerr effect (OKE), while the horizontal index is unmodified. The phase shift in matrix **S** from table 3.3 is related to the induced change in index Δn by $\Delta\phi = 2\pi\ell\Delta n/\lambda$, with ℓ the thickness of the glass slab and λ the optical wavelength in free space. This geometry is commonly used for doing a cross-correlation to determine pulse durations. Because the response of glass results purely from deforming electron clouds, it is effectively instantaneous on the 100 fs timescale.

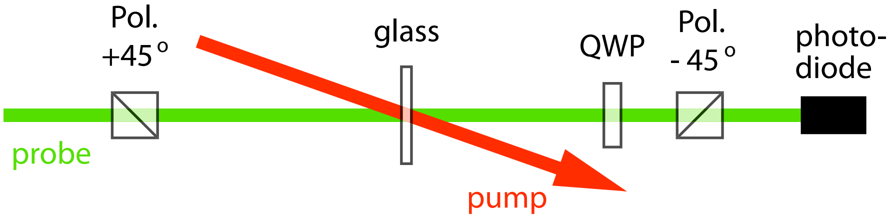

Figure 3.2 | Simple optical Kerr effect geometry. The simplest geometry for detecting an induced change in birefringence. QWP is quarter wave plate.

After the initial polarizer the polarization state is set to [1; 1], which means the vertical and horizontal polarizations have equal amplitudes with no relative phase shift. When the pump is present, a small phase shift is introduced between the two components by the birefringence induced in the glass, and an additional $\pi/2$ phase shift is introduced by the quarter wave plate (QWP). The final polarizer collapses the polarization onto a -45° linear polarization, and the horizontal and vertical polarizations interfere. Mathematically, this can be expressed with Jones calculus by:

$$\mathbf{v} = \mathbf{P}_{-45}\mathbf{QSP}_{+45}\mathbf{v_0} = E_0\left(1 - ie^{i\Delta\phi}\right)\begin{bmatrix} 1 \\ -1 \end{bmatrix}. \tag{3.27}$$

Note that the matrices are applied from right to left (the rightmost matrix represents the first optic). We can now calculate the intensity detected at the photodiode:

$$I = \mathbf{v}^{*T}\mathbf{v} = I_0(1 + \sin\Delta\phi)\,. \tag{3.28}$$

In general, we calculate the signal in a pump-probe type measurement by calculating $\Delta I / I_0 = (I - I_0)/I_0$, as this is typically proportional to the physical quantity we are trying to measure (see for example Eq. 3.29). We measure I_0 with the pump blocked, when $\Delta\phi = 0$, and we obtain:

$$\frac{\Delta I}{I_0} = \sin\Delta\phi \approx \Delta\phi = 2\pi\frac{\ell}{\lambda}\Delta n\,, \tag{3.29}$$

where the approximately equal sign comes from a Taylor expansion for small phase shifts. From this expression we can quantitatively determine Δn, the physically relevant quantity. The complete relationship between Δn and the THz electric field in $LiNbO_3$ is given by Eq. 3.19.

Notice that in this experimental geometry, the signal is directly proportional to the phase shift as shown in Eq. 3.29. This is because we added the quarter wave plate, which effectively heterodyned the beam.

If the QWP were not used, the result would be:

$$\mathbf{v} = \mathbf{P}_{-45}\mathbf{S}\mathbf{P}_{+45}\mathbf{v}_0 = E_0\left(1 - e^{i\Delta\phi}\right)\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \tag{3.30}$$

where the only difference from above is the *i* in front of the exponential. Now the intensity is:

$$I = \mathbf{v}^{*T}\mathbf{v} = I_0(1 - \cos\Delta\phi) \approx I_0\Delta\phi^2/2\ . \tag{3.31}$$

This is effectively a homodyne measurement. Notice that the intensity is proportional to $\Delta\phi^2$, so any information relating to the sign of the signal has been lost. In addition, *I* is very small for small phase shifts. Another problem is that it can be very difficult to quantify $\Delta\phi$ because calculating $\Delta I / I_0$ to normalize for I_0 involves dividing by zero. The advantage of the homodyne measurement is that it is a dark background measurement, which means that it is potentially possible to detect very small signals. Typically, however, the case with the QWP is preferred.

An additional advantage of the quarter wave plate comes from balancing, where the signal from two photodiodes can be used to suppress noise. The advantages will be discussed more fully and quantified in Sec. 3.C on DAQ card detection, but the principle can be discussed here. If the final polarizer is replaced by a Wollaston prism at 45°, two polarization states will come out, one at +45° and the other at -45°. The -45° case has already been discussed above. Running through the same calculation with the +45° case, we get: $I_+ = I_0(1 - \sin(\Delta\phi))$, which has a sign flip relative to the above. Balancing is implemented by subtracting the two signals:

$$\frac{\Delta I_-}{I_0} - \frac{\Delta I_+}{I_0} \approx 2\Delta\phi\ , \tag{3.32}$$

so the signal is doubled while laser intensity fluctuations, which are the same sign for both photodiodes, are subtracted so noise is suppressed.

3. Correcting for real optics with a polarizer and wave plate

Now I introduce an experimental situation which we have encountered where non-ideal optics complicate quantitative measurement of $\Delta\phi$. As shown in Fig. 3.3, a pair of dichroic mirrors is used to combine the pump and probe in the sample and separate them after it. The probe is transmitted through the first and reflected off the second. In addition, it develops a static phase shift from the $LiNbO_3$ sample which is ideally corrected fully by the compensating slab. In reality the compensation is imperfect due to variation in slab thickness.

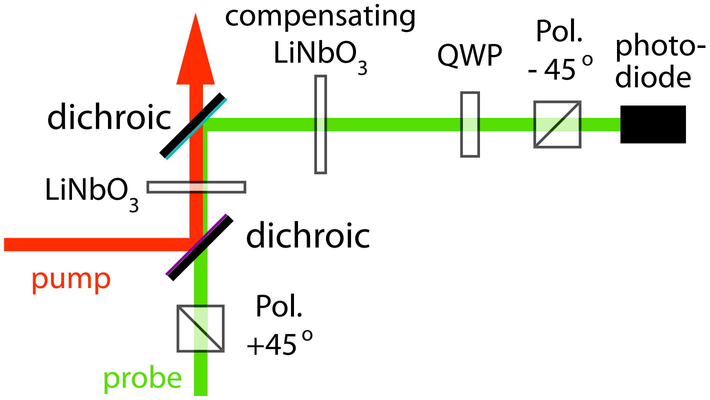

Figure 3.3 | Typical induced birefringence geometry with real optics. A pair of dichroic mirrors is used to combine the pump and probe in the sample, and later to separate them. The first reflects the pump and transmits the probe, while the second does the opposite. In addition, the sample is lithium niobate, which is strongly birefringent. This makes it necessary to use a compensating crystal that has been rotated by 90° relative to the sample which roughly compensates for the static birefringence in the sample crystal.

The behavior of this system can be modeled again using Jones calculus. The Jones vector just before the final polarizer is

$$\mathbf{v} = \mathbf{QR}(90)\mathbf{B}_2\mathbf{R}(-90)\mathbf{M}_2\mathbf{SB}_1\mathbf{M}_1\mathbf{P}_{+45}\mathbf{v}_0 = \begin{bmatrix} 1 \\ iA_1A_2\exp[i(\phi_1-\phi_2+\alpha_1+\alpha_2+\Delta\phi)] \end{bmatrix} \equiv \begin{bmatrix} 1 \\ iA\exp[i(\phi+\Delta\phi)] \end{bmatrix}. \tag{3.33}$$

Here, $\mathbf{M_1}$ is for transmission through the first dichroic, and $\mathbf{M_2}$ is for reflection off the second dichroic. $A_1 = T_s/T_p$ for the first dichroic and $A_2 = R_s/R_p$ are the relative transmissivity and reflectivity for *s*- and *p*-polarized light from the two dichroic mirrors. α_1 and α_2 are the relative phase shifts from the mirrors. $\phi_1 - \phi_2$ is the difference in phase shift introduced by the sample and compensating slab, and should be 0 for perfect compensation. $\Delta\phi$ is the THz-induced (via the electro-optic effect) phase shift that we want to measure. Following the same procedure as in the section above, the signal at the photodiode is

$$\frac{\Delta I}{I_0} = \frac{2A}{1+A^2}\sin(\phi+\Delta\phi), \tag{3.34}$$

and it is impossible to quantitatively extract $\Delta\phi$ unless both A and ϕ are known very accurately.

It is, however, possible to experimentally correct for the non-ideal behavior of the optics. Looking at the expression for the Jones vector just before the last polarizer, it can be seen that the relative amplitude and phase of vertical and horizontal polarizations are different from the ideal case. The amplitude can be corrected by rotating the first polarizer (around the axis parallel to beam propagation) to control the relative amounts of vertical and horizontal light at the outset. With proper alignment, after transmission through the first dichroic and reflection off the second dichroic, the vertical and horizontal polarizations will have exactly the same amplitude. This will set $A = 1$. To control the phase, we tilt the quarter wave plate. When the wave plate is aligned so its surface is perpendicular to the direction of beam propagation, it induces a phase shift of $\pi/2$.

If it is tilted, the path length through the birefringent material is increased and the induced phase shift is $\pi/2+\psi$, where the sign of ψ depends on whether the fast axis of the wave plate is aligned vertically or horizontally. The tilt can be chosen such that $\psi=-\phi$, returning the polarization state to the ideal.

An effective procedure to correctly align the polarizer and wave plate is not trivial to develop, so I will outline one here.

1. Set the second polarizer to +45° so it is parallel to the first polarizer (all the light should be transmitted). It is important that this angle is set accurately. It will be easiest if a Wollaston prism is used.
2. Replace the QWP with a HWP that has its fast (or slow) axis aligned vertically. Again, this alignment should be done carefully, probably by placing the HWP between crossed vertical and horizontal polarizers and rotating the HWP until no light passes through the second polarizer. After placing the HWP between the polarizers parallel at +45°, if optics are ideal, no light should be transmitted through the second polarizer. With non-ideal optics, some light will make it through.
3. Tilt the HWP back and forth by rotating the post (which has a collar) in the post holder. This will scan $\psi+\phi$ through 0, and there should be visible minima (although not yet zero) in the transmission through the second polarizer. As the HWP is being tilted back and forth, very slowly rotate the first polarizer until the minimum in transmission is an actual null. When this happens, $A = 1$ and the first polarizer is set accurately.
4. Replace the HWP with a QWP that has been accurately aligned to vertical. Slowly tilt the QWP until the two outputs of the Wollaston are balanced. If the tilt is too great, rotate the QWP by 90° (switching the fast and slow axes) and tilt again until balancing is achieved.

After this procedure, both the amplitude and phase should be corrected.

4. Correcting for real optics in the self-compensating geometry

The procedure described in the previous section assumes spatially homogeneous optics, meaning that the relative phase shift introduced between the two polarization components after passing through the sample and compensating crystal is uniform across the entire beam. If the probe is focused through the sample and compensating crystal, this may easily be true. For the imaging of THz waves, however, an expanded optical beam passes through the sample. Experience shows that the thickness and/or birefringence of the LN slab has significant spatial variation. If two orthogonal LN slabs are placed between crossed polarizers, it is impossible to fully extinguish the entire beam even if the slabs are tilted and rotated optimally.

To correct for the spatial inhomogeneity, as well as non-ideal behavior of other optics, a self-compensating geometry is used (See Fig. 3.4). An expanded probe beam passes through the LN sample. It is imaged onto the retro-reflector, and then imaged back onto the sample, mapping the image of the sample exactly back onto itself. The beam passes through a QWP at 45° travelling both to and from the retroreflector, which flips the vertical and horizontal polarizations. On the second pass through the sample, the static, spatially dependent phase shift is fully corrected, as are any problems introduced by the dichroic mirror.

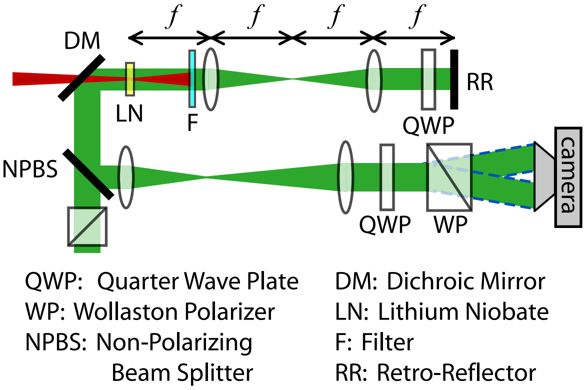

Figure 3.4 | Self-compensating induced birefringence geometry with real optics. An expanded probe beam is initially transmitted through the non-polarizing beam splitter, after which it passes through the LN sample. It is imaged onto the retro-reflector, and then imaged back onto the sample, mapping the image of the sample exactly back onto itself. The sample is then imaged onto the camera, reflecting off the NPBS on the way. The initial polarization is set to 45°.

To analyze the polarization behavior carefully, again we use Jones calculus. Here we assume that the non-polarizing beam splitter behaves ideally, since it is specially designed to not affect the polarization state. The Jones vector after the Wollaston is

$$\mathbf{v} = \mathbf{P}_{-45}\mathbf{QMSBQ}_{45}\mathbf{Q}_{45}\mathbf{BMP}_{+45}\mathbf{v_0} = E_0\left(1 - ie^{i\Delta\phi}\right)\begin{bmatrix} 1 \\ -1 \end{bmatrix}. \tag{3.35}$$

This is identical to the ideal case above, demonstrating that the self-compensating geometry can correct for non-ideal optics such as mirrors and dichroics as well as spatially dependent inhomogeneity in the sample.

C. ION DISPLACEMENT AMPLITUDES IN $LiNbO_3$

As discussed in the introduction (Fig. 1.1), the THz wave in $LiNbO_3$ is a phonon-polariton wave, so it has both electromagnetic and vibrational components. In section 3A, I used the electro-optic effect to determine the relationship between experimental signals and the absolute *E*-field amplitudes with real units. Here I derive the ionic displacements (in Angstroms) of the associated phonon mode as a function of frequency and *E*-field amplitude. As will be discussed more fully below, knowing the magnitude of the ionic displacement enables estimates of the nonlinear response and approximations of the field strengths required to switch the ferroelectric orientation of the crystal. In the analysis I present, everything is treated classically, which will help with intuition and should be fairly accurate for the high THz fields and large vibrational amplitudes we hope to achieve. I will first go through the derivation of a generic, damped, driven harmonic oscillator. I will then explain all the terms (driving force, mass, frequency, and damping) in the simple case of an aligned diatomic molecule such as HCl. Finally, I will introduce the important phonon mode in $LiNbO_3$ and estimate actual atomic displacements in our experiments.

1. The damped, driven harmonic oscillator

Discussions of the damped, driven harmonic oscillator can be found in many textbooks. The derivation starts with Newton's second law: $F = ma$. There are three contributions to the force term on the left hand side. The first is the restoring force of the spring itself, $F_{\text{restore}} = -kx$. If this were the only force term, the solution would be simple harmonic motion $x(t) = \sin(\omega_0 t)$ with the resonant frequency $\omega_0 = \sqrt{k/m}$. The second force term is the damping term, $F_{\text{damp}} = -c(dx/dt)$. With both the restoring and damping term, the solution (in the underdamped case) is sinusoidal oscillations multiplied by an exponential decay term. The final force term is the driving term. Collecting all these together, we get the differential equation:

$$F_{\text{drive}} + F_{\text{restore}} + F_{\text{damp}} = F_{\text{drive}} - kx - c\frac{dx}{dt} = ma \tag{3.36}$$

For convenience when deriving the solutions, the terms are typically rearranged and new constants are defined to give:

$$\frac{d^2x}{dt^2} + 2\gamma\omega_0\frac{dx}{dt} + \omega_0^2 x = \frac{F(t)}{m}. \tag{3.37}$$

Here ω_0 is the resonant frequency of the undamped harmonic oscillator as described above, $F(t)$ is the externally applied, time-dependent driving force, and γ is the dimensionless damping ratio. When $\gamma < 1$, the system is underdamped and experiences oscillations. γ should be significantly less than 1 for most molecular and crystalline vibrations, so a delta-function input will cause the system to oscillate many times before damping away. The generalized form of the differential equation shown in Eq. 3.37 represents many different systems including torsional systems and *RLC* circuits.

There are two damped harmonic oscillator responses that are particularly relevant for understanding atomic displacements in a molecule or crystal driven by an electromagnetic wave:

the impulse response and the response to a sinusoidal driving force. The impulse response to a unit delta function is a damped exponential:

$$g(t) = \frac{1}{m\omega_0\sqrt{1-\gamma^2}} \exp(-\gamma\omega_0 t)\sin\left(\sqrt{1-\gamma^2}\,\omega_0 t\right) \tag{3.38}$$

The oscillation frequency $\omega_1 = \omega_0\sqrt{1-\gamma^2}$ is slightly lower than the undamped frequency ω_0. $g(t)$ can be used as a Green's function to determine the system response to an arbitrary force $F(t)$ (such as a single cycle THz pulse) via a convolution:

$$x(t) = F(t) \otimes g(t) = \int_{-\infty}^{\infty} F(\tau)g(t-\tau)d\tau. \tag{3.39}$$

Note that the units of $g(t)$ are sec/kg, so that after convolution with the force, we retrieve the displacement $x(t)$. A plot of the impulse response for $\gamma = 0.07$ is shown in Fig. 3.5(a).

The second particularly important response for a dipole interacting with an electromagnetic wave is when the dipole is driven by a sinusoidal force $F(t) = F_0\sin(\omega t)$ (e.g. when the $LiNbO_3$ is excited by a narrowband THz wave). The solution can be derived, for example, by convolving the driving force with $g(t)$. It is

$$x(t) = A(\omega)\sin[\omega t + \phi(\omega)] \tag{3.40a}$$

$$A(\omega) = \frac{F_0}{m\omega_0^2}\frac{1}{\sqrt{(2\gamma q)^2 + (q^2-1)^2}} \tag{3.40b}$$

$$\phi(\omega) = \arctan\left(\frac{2\gamma q}{q^2-1}\right) \tag{3.40c}$$

$$q \equiv \frac{\omega}{\omega_0}. \tag{3.40d}$$

The amplitude and phase for $\gamma = 0.07$ is plotted in Fig. 3.5(b). There is a DC response at zero-frequency, and its magnitude is simply $A_{\mathrm{DC}} = A(0) = \frac{F_0}{m\omega_0^2} = \frac{F_0}{k}$. For significantly underdamped systems, the maximum response is given by $A_{\max} = \frac{F_0}{m\omega_0^2}\frac{1}{2\gamma\sqrt{1-\gamma^2}}$ and occurs at a frequency $\omega_{\max} = \omega_0\sqrt{1-2\gamma^2}$. $A_{\max}$ and $\omega_{\max}$ are plotted as a function of γ in Fig. 3.5(c).

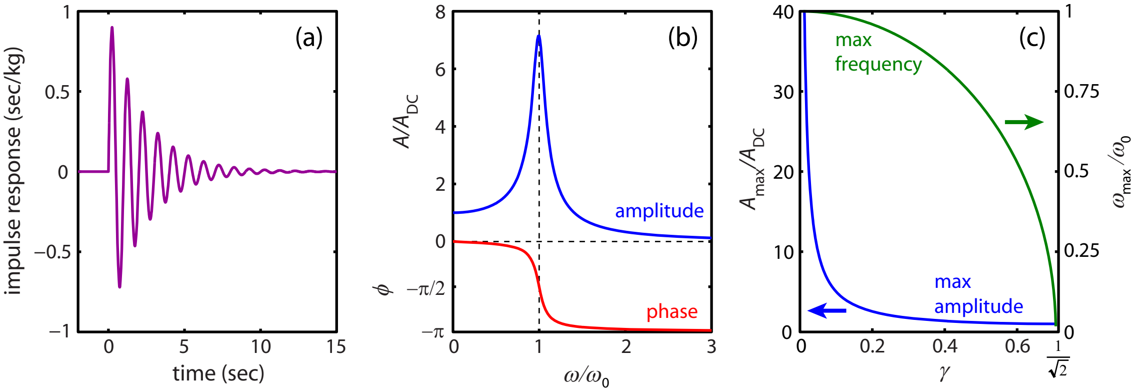

Figure 3.5 | Generic damped, driven harmonic oscillator. (a) The impulse response for a damped harmonic oscillator with $\gamma = 0.07$. The mass is set to $m = 1$ kg and the resonant frequency to $\omega_0 = 2\pi$. (b) The amplitude and phase of the response when a damped harmonic oscillator is driven by sinusoidal force at a frequency ω. Again, $\gamma = 0.07$. (c) The left axis shows the peak amplitude [maximum of blue curve in (b)] and the right axis shows the frequency at which the peak occurs as a function of γ.

2. Ion displacement amplitudes in the simple diatomic HCl

Although we are actually interested in atomic displacements in lithium niobate, as an example I will first calculate displacements in hydrogen chloride. This simpler example lets us build intuition without dealing with a complicated unit cell and vibrational mode. We first assume that the HCl molecule is isolated in a gas and the dipole moment is aligned parallel to the electric field of the incident light wave, as shown in Fig. 3.6(a). Because HCl is a polar molecule, there is a partial positive charge $+\delta$ on the hydrogen and a partial negative charge $-\delta$ on the chlorine. The field pulls the charges in opposite directions, causing the molecule to vibrate. We assume a sinusoidal driving field of the form $E(t) = E_0 \sin(\omega t)$. The force on atoms is simply the field multiplied by the charge, $F(t) = \delta E(t) = \delta E_0 \sin(\omega t)$, and for this periodic driving force the atomic motion is described by Eq. 3.40. To calculate the amplitude, the only further modification to Eq. 3.40b is that the mass must be replaced by the reduced mass: $m_{eff} = m_{\mathrm{H}} m_{\mathrm{Cl}} / (m_{\mathrm{H}} + m_{\mathrm{Cl}})$. Thus for HCl, we have the maximum amplitude of the atomic displacement given by:

$$A(q = \omega / \omega_0) = \frac{\delta E_0}{m_{eff}\,\omega_0^2} \frac{1}{\sqrt{(2\gamma q)^2 + (q^2 - 1)^2}} \tag{3.41}$$

The damping ratio γ can be estimated from the collision time. By setting the relaxation time of the impulse response (Eq. 3.38) equal to the collision time τ, we get $\gamma \sim 1/(\omega_0 \tau)$. At STP in air, the collision time for HCl molecules is τ ~ 0.3 ns and so γ ~ 6e-6. The values for the other constants of $\mathrm{H}^1\mathrm{Cl}^{35}$ are shown in Fig. 3.6(b), and the frequency-dependent amplitude for a driving field of E_0 = 100 kV/cm is shown in Fig. 3.6(c), revealing a very sharp resonance with large motion when driven resonantly.

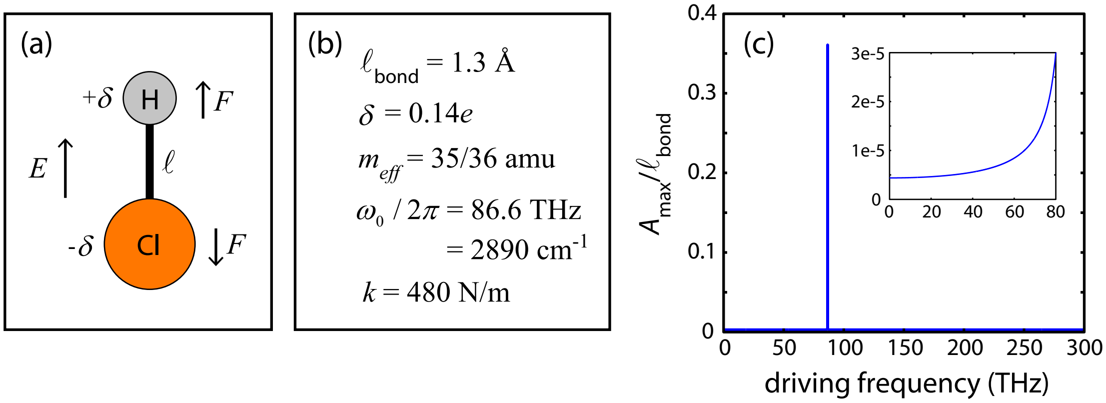

Figure 3.5 | HCl as a harmonic oscillator. (a) HCl in a field. (b) Relevant constants for HCl [HCl]. (c) Atomic displacement normalized by the bond length when driven by a CW THz field with an amplitude of 100 kV/cm. Inset is a magnified view of the response below the resonance at 86.6 THz.

3. Ion displacements in $LiNbO_3$

Now I address the primary purpose of this section, estimating atomic displacements in $LiNbO_3$ under excitation with a THz field. I will only consider displacement along a single normal mode (a single optic phonon mode), the soft mode of $LiNbO_3$ [see Fig. 3.6(a)]. The mode is called "soft" because its resonant frequency becomes very low as the crystal temperature approaches the ferroelectric phase transition temperature. The soft mode is particularly interesting because motions of atoms in the unit cell along the directions of this normal mode are in large part responsible for taking $LiNbO_3$ from one ferroelectric orientation to another. By driving the phonon mode to high amplitude with a THz field, it should be possible to switch the ferroelectric orientation. Figure 3.6(b) shows the potential for this mode. In the ferroelectric phase the crystal lies in one of the potential minima, which corresponds to Li and Nb atoms shifted up and O atoms shifted down relative to their high-symmetry locations [see Fig. 3.6(a) for direction of motion]. [Weis 1985] gives a clear description of the crystal structure and the effects of the ferroelectric phase transition. In the sections below, I will explain how I estimated the value of each parameter necessary for calculating atomic displacements. Some of the estimates are fairly coarse, and it should be possible to determine values more accurately. I hope by explaining the overall method and approximations clearly here, it will be fairly straightforward to substitute the improved values as they are calculated or found in references I missed.

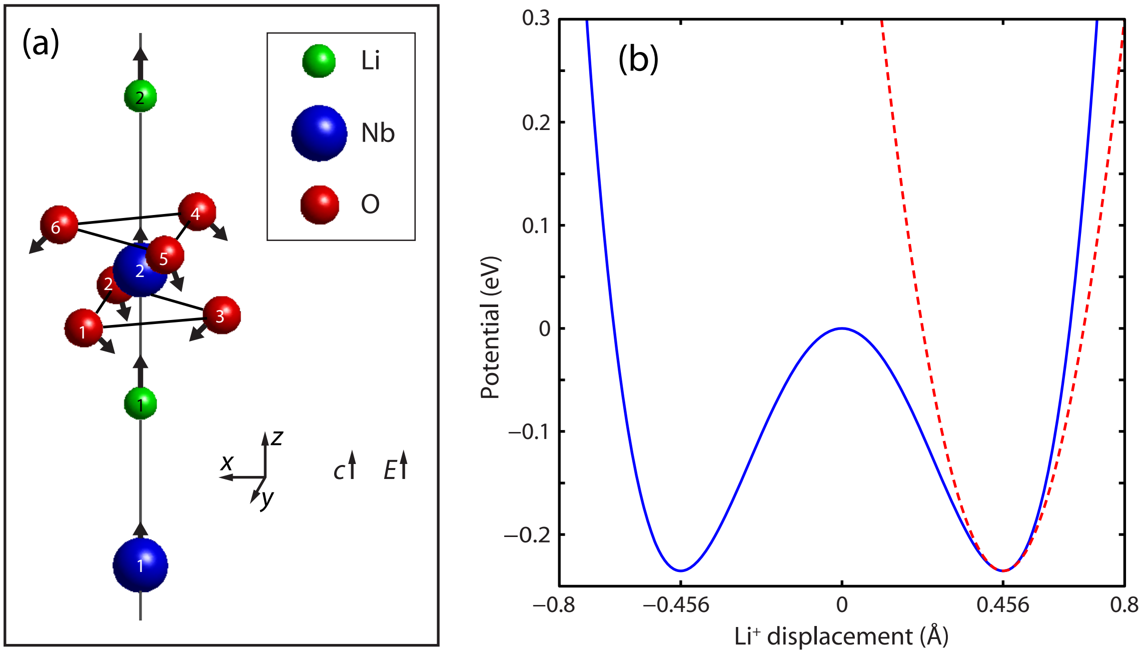

Figure 3.6 | $LiNbO_3$ normal mode and potential. (a) The soft mode in $LiNbO_3$, resonant at 7.4 THz, which is largely responsible for the ferroelectric phase transition. (b) The quartic double-well potential of the normal mode as a function of Li ion displacement is shown in solid blue. Displacement amplitudes for O and Nb are proportional to the Li atom displacement (see Table 3.5). The dashed red line is a quadratic fit around one minimum, used for modeling small displacements in the ferroelectric phase.

a. Determining the normal mode

Finding the exact normal mode of lithium niobate associated with the 7.4 THz resonance was difficult; I could not find a clear and unambiguous description in the literature. As an approximation, I will follow [Barker 1967], and approximate the normal mode as lying along the displacement vectors of $LiNbO_3$ when it goes through the ferroelectric phase transition. Since the 7.4 THz resonance is thought to be primarily responsible for driving the transition, this should be a passable approximation. As in [Barker 1967], the motion of the oxygen atoms along x and y is unknown, and we assume it to be at 45° relative to the z-axis.

Ion	$\vec{r}$ **Location**[a,b] **(Å)**	a **Normal mode relative amplitude**[c]	$\vec{n}$ **Normal mode unit vector**[c]	δ **partial charge**[d] **(units of** e**)**
Li_1	(0, 0, 3.92)	1	(0, 0, 1)	+1
Li_2	(0, 0, 10.86)	1	(0, 0, 1)	+1
Nb_1	(0, 0, 0)	0.14	(0, 0, 1)	+1.4
Nb_2	(0, 0, 7.20)	0.14	(0, 0, 1)	+1.4
O_1	(1.41, 1.10, 6.31)	0.59	(-0.56, -0.43, -0.71)	-0.8
O_2	(0.25, -1.77, 6.31)	0.59	(-0.10, 0.70, -0.71)	-0.8
O_3	(-1.65, 0.67, 6.31)	0.59	(0.66, -0.27, -0.71)	-0.8
O_4	(-1.41, -1.10, 8.09)	0.59	(-0.56, -0.43, -0.71)	-0.8
O_5	(-0.25, 1.77, 8.09)	0.59	(-0.10, 0.70, -0.71)	-0.8
O_6	(1.65, -0.67, 8.09)	0.59	(0.66, -0.27, -0.71)	-0.8
	[a][NRL] [b][Boysen 1994]	[c][Barker 1967]	[c][Barker 1967]	[d][Lines 1970]

Table 3.5 | Atomic locations, normal mode displacement vectors, and partial charges in $LiNbO_3$. Listed here are two Li atoms, 2 Nb atoms, and 6 O atoms. The octahedron of O atoms surrounding the bottom Nb is rotated slightly relative to the one that is shown, but this is omitted because it is not needed to calculate normal mode amplitudes. Data are drawn from [NRL], and the original source is [Boysen 1994].

b. Calculating spring constant and effective mass

The spring constant and effective mass can be easily determined once the parabolic potential is known. In this section, I estimate the parabolic potential by Taylor expanding the full theoretical curve around the potential minimum in the ferroelectric phase. The full potential in $LiNbO_3$ can be modeled by a quartic function of the form [Romero-Rochin 1999]:

$$V = V_0\left[\left(\frac{Q}{Q_{\min}}\right)^4 - 2\left(\frac{Q}{Q_{\min}}\right)^2\right] \tag{3.42}$$

with V_0 the well depth, Q the displacement along the normal mode, and $Q_{\min}$ the displacement corresponding to the well minimum. When plotting the solid blue curve in Fig. 3.6(b), I used the theoretically calculated well depth (V_0 = 0.24 eV from [Inbar 1996]) and the experimentally measured displacement of the Li atom from the high-symmetry point ($Q_{\min} = 0.456$ Å from [Weis 1985]). A displacement of $Q = 0$ corresponds to the Li in the high-symmetry point where it is found in the paraelectric phase.

I estimate the harmonic potential by Taylor expanding the quartic potential around $Q_{\min}$. The potential is given by

$$V = \frac{4V_0}{Q_{\min}^2} R^2 - V_0 \tag{3.43}$$

with $R \equiv Q - Q_{\min}$. R represents the amplitude of the normal mode, and it is proportional to the amplitude of the actual ion displacements (see Table 3.5 for proportionality constants and displacement directions). The harmonic potential in Eq. 3.43 is plotted as the dashed red line in Fig. 4.6(b). It is apparent that the quadratic approximation starts to break down even for modest vibrational amplitudes where $R/Q_{\min}$ is relatively small. There is nonetheless significant insight that can be gained because displacements in this anharmonic potential will exceed those in the harmonic potential, so the results give the approximate magnitude and serve as a lower bound. In addition, by modeling the mode as a harmonic oscillator, we can use the intuition developed previously. Eq. 3.43 can be used, along with the resonant frequency, to determine constants we seek. The spring constant, retrieved from the derivative of the potential, is $k = 8V_0 / Q_{\min}^2 = 148\,\text{N/m}$. The effective mass is $m_{eff} = k / \omega_0^2 = 41\,\text{amu}$.

c. Calculating the force

The final important parameter to determine atomic displacements is the force generated by the applied field. In HCl the force was simply the field multiplied by the partial charge, but the more complex normal mode in $LiNbO_3$ makes life slightly more complicated. To start, we calculate the energy potential from the applied field as a function of normal mode displacement $U(R)$. The force is then given by $F = dU/dR$. The potential from the applied field is

$$U(R) = \vec{E} \cdot \vec{\mu}(R). \tag{3.44}$$

The dipole moment is determined by the atomic charges, atomic locations, and the normal mode amplitude and unit vectors from Table 3.5.

$$\vec{\mu}(R) = \sum_i \delta_i \left(\vec{r}_i + a_i \vec{n}_i R\right) = \sum_i \delta_i \vec{r}_i + R \sum_i \delta_i a_i \vec{n}_i = \vec{\mu}_0 + R \sum_i \delta_i a_i \vec{n}_i. \tag{3.45}$$

Here the sums run over all the atoms in one unit cell: Li_1, Nb_1, and O_1, O_2, and O_3. The simple potential is just a constant plus a term with a linear dependence on R. Taking the derivative with respect to R to get the force, the constant term vanishes and we get

$$F = \vec{E} \cdot \sum_i \delta_i a_i \vec{n}_i. \tag{3.46}$$

For this mode, the change in dipole lies entirely along z (the optic axis). Said another way, the components along x and y in the vector sum in Eq. 3.46 cancel. In addition, we are interested in an applied field polarized along z, so we can take the z-component of the normal mode vector to convert the vectors to scalars:

$$F(t) = E(t) \sum_i \delta_i a_i n_{z,i} \equiv E(t) \delta_{eff} \tag{3.47}$$

where I have defined an effective charge in analogy to the simple diatomic case. Using the values from Table 3.5, we get $\delta_{eff} = 2.4e$.

d. THz electromagnetic wave driven responses in LiNbO₃

Resonant frequency	$\omega_0/2\pi = 7.4$ THz
Damping ratio	$\gamma = 0.043$
Spring constant	$k = 148$ N/m
Effective mass	$m_{eff} = 41$ amu
Effective charge	$\delta_{eff} = 2.4e$
Li$^+$ displacement from high symmetry point	$Q_{\min} = 46$ pm
DC amplitude constant	$\dfrac{\delta_{eff}}{m_{eff}\omega_0^2} = 0.26\dfrac{\text{pm}\cdot\text{cm}}{\text{MV}}$

Table 3.6 | Collected normal mode parameters in LiNbO$_3$. Most parameters were calculated in the previous sections. The damping ration is from [Feurer 2007]. The final row facilitates quick calculation of displacement amplitudes when the narrowband THz field strength is known (in MV/cm).

The relevant parameters for calculating vibrational amplitudes in LiNbO$_3$ are compiled in Table 3.6. With these values in hand, we can calculate the impulse response and the response to a CW driving field, all with real units. Following Eq. 3.38, the response to a delta function driving *E*-field is:

$$g(t) = \frac{\delta_{eff}}{m_{eff}\omega_0\sqrt{1-\gamma^2}}\exp(-\gamma\omega_0 t)\sin\left(\sqrt{1-\gamma^2}\,\omega_0 t\right). \tag{3.48}$$

The normal mode amplitude can be extracted by a convolution: $R(t) = E(t) \otimes g(t)$. Following Eq. 3.40, the amplitude of the response to a sinusoidal driving field is:

$$R(q) = \frac{\delta_{eff}E_0}{m_{eff}\omega_0^2}\frac{1}{\sqrt{(2\gamma q)^2 + (q^2-1)^2}}. \tag{3.49a}$$

$$q \equiv \frac{\omega}{\omega_0}. \tag{3.49b}$$

The results are summarized in Fig. 3.7. Figure 3.7(a) shows the impulse response, $g(t)$ in LiNbO$_3$. Figure 3.7(b) shows a single-cycle THz driving field with a peak amplitude of 1 MV/cm (orange curve, left axis) and the resulting normal mode amplitude $R(t)$ (green curve, right axis), which is equal to the Li ion displacement. For frequencies below about 5 THz, the ionic displacement directly follows the input field. Displacements for Nb and O atoms can be determined from Table 3.5. Finally, Fig. 3.7(c) shows $R(\omega)$ for a CW driving field of 1 MV/cm. The peak ionic displacement is only about 1/15th of the distance required (46 pm) to switch the ferroelectric orientation, indicating that a 15 MV/cm field would be required on resonance to switch the phase. The reality would be about two times easier because the real quartic potential is much softer than the quadratic potential used in this model.

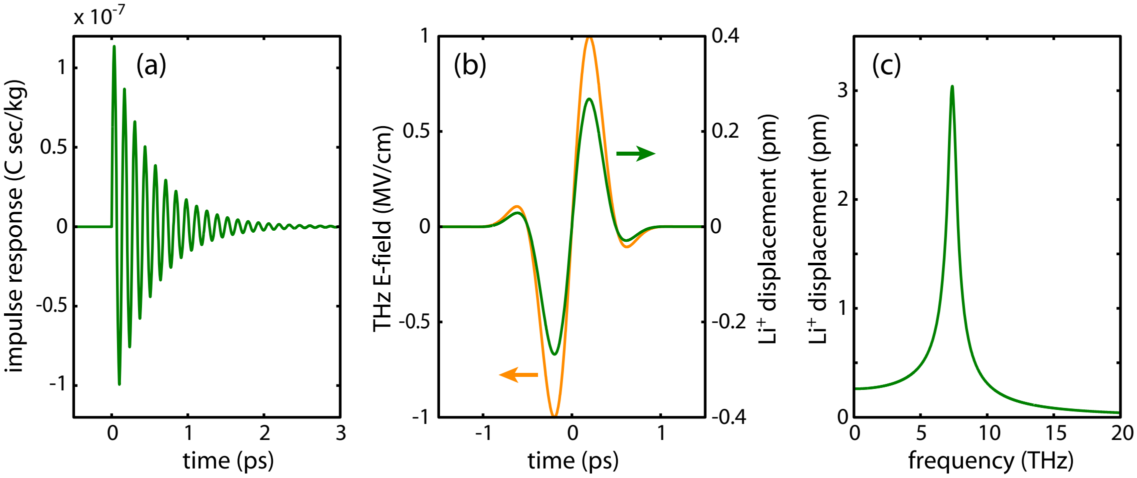

Figure 3.7 | Atomic displacements in driven $LiNbO_3$. (a) The impulse response of the soft mode (7.4 THz at room temperature). (b) The ionic displacement (in green, right axis) when the $LiNbO_3$ is driven by a single-cycle THz field (orange, right axis) with a peak field strength of 1 MV/cm. (c) The amplitude response for different, CW driving frequencies with an amplitude of 1 MV/cm. As shown in Fig. 3.6(b), the maximum of the quartic potential is 46 pm from the minimum.

Finally, we seek some confirmation that the calculated values in Table 3.6 and the results in Fig. 3.7 are reasonable. One such test is the coercive field. In a ferroelectric crystal, the coercive field amplitude is the external DC field strength necessary to switch the ferroelectric orientation. Using a completely different method than what we use here, it is theoretically predicted to be 5.4 MV/cm [Kim 2002]. The final row of Table 3.6 is the constant which determines the DC amplitude (see Eq. 3.49a). Its units are chosen so the displacement in pm can be calculated directly when the field amplitude is known in MV/cm. We estimate the coercive field by finding the field necessary to raise the potential energy by V_0, the height of the maximum in the double well potential [see Fig. 3.6(b)]. In the harmonic approximation this corresponds to a displacement of $Q_{\min}/2$ = 23 pm (see Eq. 3.43), and we find the coercive field strength to be ~100 MV/cm. Obviously there is a huge, ~20-fold discrepancy between this and the literature value. In reality, there are many phonons that contribute to the phase transition, so it is expected that the calculated value for driving directly over the hump in the double-well potential would be an overestimate, but a factor of 20 seems too large. I would say the results presented in this section are called into question and results should be tested against another calculation methodology such as a rigorous DFT simulation.

D. REFERENCES

[Barker 1967] A. S. Barker and R. Loudon, “Dielectric properties and optical phonons in $LiNbO_3$,” *Phys. Rev.* **158**, 433-445 (1967).

[Born 1999] M. Born and E. Wolf, *Principles of Optics*, 7th ed. (Cambridge Univ. Press, Cambridge, 1999).

[Boyd 2008] R. W. Boyd, *Nonlinear Optics*, 3rd ed. (Elsevier, Burlington, 2008).

[Boysen 1994] H. Boysen and F. Altorfer, “A neutron powder investigation of the high-temperature structure and phase transition in $LiNb0_3$,” *Acta Cryst.* **BS0**, 405-414 (1994).

[Feurer 2007] T. Feurer, N. S. Stoyanov, D. W. Ward, J. C. Vaughan, E. R. Statz, and K. A. Nelson, “Terahertz polaritonics,” *Annu. Rev. Mater. Res.* **37**, 317-350 (2007).

[HCl] Various sources on the internet that did not cite their source. Quantitative values should be regarded with suspicion.

[Inbar 1996] I. Inbar and R. E. Cohen, “Comparison of the electronic structures and energetics of ferroelectric $LiNbO_3$ and $LiTaO_3$,” *Phys. Rev. B* **53**, 1193-1204 (1996).

[Kim 2002] S. Kim, V. Gopalana, and A. Gruverman, “Coercive fields in ferroelectrics: A case study in lithium niobate and lithium tantalate,” *Appl. Phys. Lett.* **80**, 2740-2742 (2002).

[Larson 1990] R. E. Larson, R. P. Hostetler, and B. H. Edwards, *Calculus*, 4th ed. (D. C. Heath and Co., Lexington, 1990).

[Lenzo 1966] P. V. Lenzo, E. G. Spencer, and K. Nassau, “Electro-optic coefficients in single-domain ferroelectric lithium niobate,” *J. Opt. Soc. Am.* **56**, 633-635 (1966).

[NRL] http://cst-www.nrl.navy.mil/lattice/struk/LiNbO3f.html

[Romero-Rochin 1999] V. Romero-Rochín, R. M. Koehl, C. J. Brennan, and K. A. Nelson, “Anharmonic phonon-polariton excitation through impulsive stimulated Raman scattering and detection through wave vector overtone spectroscopy: Theory and comparison to experiments on lithium tantalite,” *J. Chem. Phys.* **111**, 3559-3571 (1999).

[Weis 1985] R. S. Weis and T. K. Gaylord, “Lithium niobate: Summary of physical properties and crystal structure,” *Appl. Phys. A* **37**, 191-203 (1985).

[Wong 2002] D. H. Jundt, C. J. G. Kirkby, and C. Florea, "Optical, dielectric and electro-optical properties," in *Properties of Lithium Niobate*, K. K. Wong ed. (Short Run Press, Exeter, 2002), pp. 113–175.

[Yariv 2007] Yariv, A. & Yeh, P. *Photonics: Optical Electronics in Modern Communications*, 6th ed. (Oxford Univ. Press, New York, 2007).

Chapter IV

Developments in polaritonics methodology: THz generation, DAQ card detection, and THz imaging

PREFACE

The polaritonics platform had been under development for more than 7 years when I joined the lab, and an array of capabilities had been demonstrated [Feurer 2007]. These included THz imaging [Koehl 1999], programmable control over THz generation [Feurer 2003], and an array techniques for THz manipulation implemented by laser machining air structures into the host material [Ward 2004; Danielson 2006; Stoyanov 2004; Stoyanov 2002; Weiss 2001]. To press forward, however, improved experimental techniques were required for both THz generation and THz detection.

THz fields generated in the waveguide were relatively weak, and waveguide dispersion (see chapter 2) quickly chirped the pulse. In addition, there were so many generated frequency components that it was difficult to follow what was happening in machined structures. We used phasematching and coherent amplification to generate high-amplitude, multicycle waves in the polaritonics chip. Both the center frequency and bandwidth could be tuned experimentally (see [Lin 2009] and section 4A below). The high-amplitude narrowband THz pulses were useful for studying machined microstructures (see [Werley AJP 2012] and chapter 5) and resonant metallic antennas (see [Werley OE 2012] and chapter 6). In the future we plan to use these high-amplitude waves to strongly drive resonances in nonlinear THz studies.

The THz detection techniques were also limiting. We developed a method for point-source detection that used a data acquisition (DAQ) card instead of a lock-in amplifier. The DAQ card system offered quantitative monitoring of laser noise, simple, quantitative measurement of THz field amplitudes, an improved signal-to-noise ratio, and more versatility in terms of rapid scanning and differential chopping (see [Werley 2011] and section 4B). The DAQ card method can be implemented in almost any ultrafast experiment. The THz imaging methodology had even more room for improvement. The technique in use when I joined the lab, Talbot imaging, is an out-of-focus technique so the resolution is very poor. I introduced phase contrast imaging (see [Wu 2009, Werley 2010] and sections 4C & 4D), an in-focus technique which improved image resolution from ~100 μm to ~1 μm. This enabled mapping the near-fields around a metal antenna with λ/100 image resolution [Werley OE 2012]. In addition to resolution problems, neither pump modulation nor balancing was used for noise suppression, so only very large signals could be detected. I reworked the detection electronics (see section 4C) and introduced polarization gating imaging (see [Werley 2010] and section 4C & D) to implement balanced detection. These improvements reduced the minimum signal detection threshold by more than a factor of 100. The improved signal-to-noise greatly facilitated all subsequent measurements [Werley AJP 2012, Werley OE 2012]. These improved detection methods will continue to be useful in future experiments, where we aim to study photonic crystals and metamaterial structures.

A. GENERATION OF MULTICYCLE THz PHONON-POLARITON WAVES IN A PLANAR WAVEGUIDE BY TILTED OPTICAL PULSE FRONTS

Content from: K.-H. Lin, C. A. Werley, and K. A. Nelson, "Generation of multicycle THz phonon-polariton waves in a planar waveguide by tilted optical pulse fronts," *Appl. Phys. Lett.* **95**, 103304 (2009).

1. Abstract

We demonstrate generation of frequency-tunable, multicycle THz phonon-polariton waves in a $LiNbO_3$ slab waveguide. Because the waveguide modes show considerable phase-velocity dispersion, we are able to enhance frequency-selected narrowband THz waves by using femtosecond optical pulses whose intensity fronts are tilted at angles that meet the appropriate noncollinear phase-matching conditions. The pump light is spread across a large area of the crystalline waveguide within which coherent THz wave generation occurs, averting material damage while yielding peak-to-peak THz field amplitudes in the waveguide of 50 kV/cm.

2. Introduction

A phonon-polariton mode, which results from coupled photon and optic phonon modes, can be excited by optical pulses in noncentrosymmetric crystals such as $LiNbO_3$ (LN) through impulsive stimulated Raman scattering [Crimmins 2002]. Phonon-polariton waves generated in bulk crystals can be coupled into free space and are a good source for intense, THz electromagnetic radiation [Hebling 2002; Hoffmann 2007; Yeh 2007]. In LN, the spatial and temporal profiles of the THz waveform are determined by those of the optical pump. This makes possible waveform synthesis, manipulation, and guidance through the use of spatially and temporally shaped optical excitation pulses [Feurer 2003; Ward 2004] and fabricated structures in the host crystals [Ward 2004; Danielson 2006; Stoyanov 2004; Stoyanov 2002; Weiss 2001]. These capabilities, coupled with real-space imaging of the full THz fields [Koehl 1999; Feurer 2007; Peier 2008; Wu 2009; Werley 2010], make LN slabs and waveguides promising polaritonics platforms for integrated THz photonics, optoelectronics, and spectroscopy [Feurer 2007].

For many applications in coherent control and generally for studying resonant phenomena, it would be useful to generate narrowband, high-amplitude THz waves with a tunable center frequency in slab-waveguide structures. Tunable multiple-cycle waves can be generated through the use of crossed pump pulses that produce an optical interference pattern in the crystal [Ward 2005; Katayama 2008; Stepanov 2004] to select the THz wave vector or through the use of multiple optical pulses that are timed to select the THz frequency [Ward 2004]. The maximum THz amplitudes that these techniques can achieve, however, are inherently limited by the LN optical damage threshold. Here we take advantage of waveguide dispersion and noncollinearly velocity-matched generation by using an optical pump pulse with a tilted intensity front, which substantially circumvents this problem. The tilted pulse-front geometry has been used to produce high-power THz waves in bulk LN [Hebling 2002], where the tilt is set so that the peak intensity of the optical pump moves laterally across crystal at a speed that matches the lateral THz phase velocity, thereby continuously intersecting the peak of the THz wave and coherently enhancing its amplitude. For frequencies near the transverse optical phonon frequency (7.6 THz), material dispersion enables phase-matched generation of multiple-cycle waves [Hebling 2004]. Well below the phonon frequency (less then ~4 THz), however, the dispersion curve in bulk LN is nearly linear, so a wide range of frequencies is noncollinearly velocity-matched and broad-band, single-cycle THz pulses are generated. In this letter, we demonstrate that the unique dispersion properties of THz waveguide modes in LN can be exploited for facile generation of intense, frequency-tunable multiple-cycle THz waves with high spectral brightness in the 0.1 – 1.5 THz frequency range.

3. Experimental, results, and discussion

Our sample was a 33-μm thick, x-cut $LiNbO_3$ slab. The polarization of the generated THz wave was parallel to the optic axis of LN. Figure 4.1(a) shows the calculated dispersion relation for bulk LN (dotted lined), air (dashed line), and the first three transverse electric waveguide modes in the slab. Figure 4.1(b) shows the effective index for these modes, $n_{\text{eff}}(\omega) = c\beta/\omega$, with β the propagation constant. Note that n_{eff} for the waveguide modes has a much steeper frequency dependence than the bulk index. Although we are able to generate THz waves in several waveguide modes, for the remainder of this paper we will discuss only the lowest waveguide mode because it shows the highest generation and detection efficiencies.

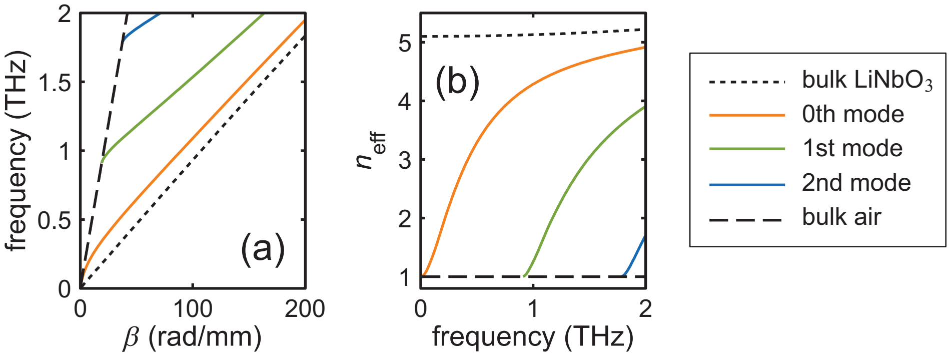

Figure 4.1 | Dispersion curves and effective index. (a) Calculated dispersion relations for electromagnetic waves in air (dashed line) and for bulk (dotted line) and waveguide (solid lines) phonon-polariton modes in $LiNbO_3$. β is the propagation constant, or wave vector. (b) The frequency-dependent effective index corresponding to each dispersion curve.

Because the THz phase velocity varies as a function of frequency, we can manipulate the tilt angle of the optical pump's intensity front, γ, to selectively enhance the amplitude of a desired frequency. Figure 4.2(a)-(c) shows the sample geometry as viewed from above at three different interaction times. As the tilted pump pulse propagates through the slab, its point of intersection with the slab moves laterally at a speed $v_{\text{int}}(\gamma) = c\cot(\gamma)$. Selective enhancement of a THz wave of frequency ω_0 is achieved through the velocity-matching condition:

$$v_{\text{int}}(\gamma) = v_{\text{THz}}(\omega_0) = c / n_{\text{eff}}(\omega_0), \tag{4.1}$$

which is met when

$$n_{\text{eff}}(\omega_0) = \tan\gamma. \tag{4.2}$$

It is not only possible to select the center frequency of the THz wave by setting γ, but also to control the bandwidth by specifying the velocity-matching length: the longer the interaction length, the smaller the bandwidth as shown in Fig. 4.2(d)-(f). One way to visualize the dispersion and phase-matching process is through the mismatch between the THz phase and group velocities. Due to the slower group velocity, the THz wave envelope lags behind the single field cycle that is being continuously amplified by the optical pump. Although phase cycles grow into the trailing edge of the THz pulse as they would in a freely propagating wave, pumping prevents them from dying out at the leading edge. This leads to a gradual increase in the number of cycles at the selected frequency ω_0 and a narrowing of the bandwidth.

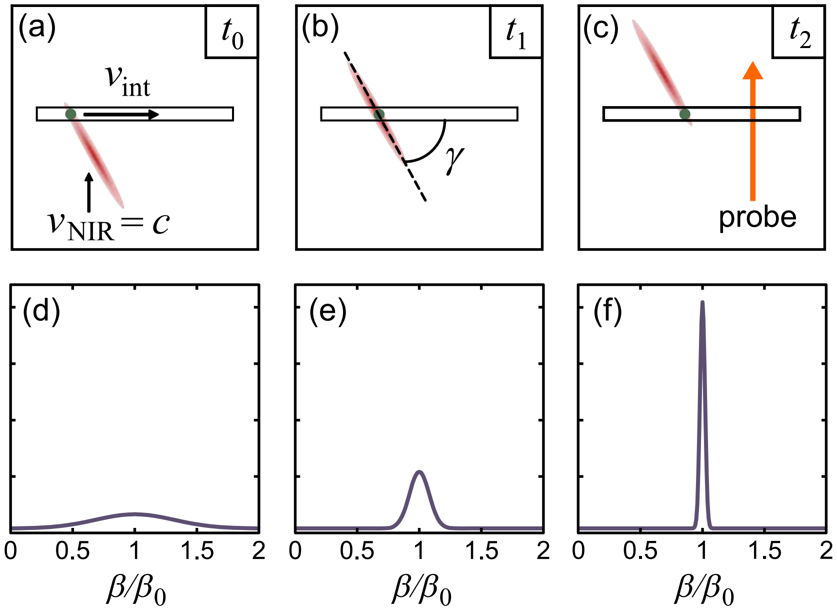

Figure 4.2 | Phase matching and bandwidth narrowing. (a) – (c) The intersection of the tilted optical pulse with the LN slab, shown from above, at times t_0, t_1, and t_2. The optical pulse travels perpendicular to the crystal surface at the speed of light in air, while the intersection of the optical pulse with the crystal (depicted as a filled circle) travels to the right at the intersection velocity $v_{\text{int}} = c\cot(\gamma)$, with γ shown in (b), and c the speed of light in vacuum. The generated THz wave propagates in the waveguide and is probed at a later time in a spatially separate location as shown in (c). (d) – (f) The wave vector content of the THz pulse at times t_0, t_1, and t_2.

To demonstrate the generation and coherent amplification of tunable, narrowband THz waves in a LN waveguide, we used optical pulses that were generated from a Ti:sapphire regenerative amplifier system with central wavelength of 800 nm, repetition rate of 1 kHz, pulse duration of 70 fs, and energy of 600 μJ. A 2000 lines/mm grating was used to tilt the pulse front, and the grating surface was imaged onto the sample by directing the -1 diffraction order through a single 10-cm cylindrical lens. We detected the THz wave at a localized region distinct from the generation region, as shown in Fig. 4.2(c).

The pulse tilt angle γ at the LN crystal depends on both the angle of incidence onto of the grating and the magnification of the imaging system, M [Palfalvi 2008]. We held the angle of incidence constant at 46°, yielding for the -1 diffraction order a tilt angle $\gamma_1 = 73°$, and we varied the magnification of the one-lens imaging system by moving the grating and lens while keeping the image plane location fixed at the LN crystal. The resulting tilt angle at the crystal was given by the relation $\tan\gamma = \tan(\gamma_1)/M$. The interaction length, determined by the pump spot size on the grating as well as the magnification, ranged from roughly 2 mm ($M = 0.8$) to 5 mm ($M = 2$).

For detecting the THz waves, we focused the probe (output of a homebuilt noncollinear optical parametric amplifier; 650 nm central wavelength; 50 fs pulse duration) onto the LN waveguide with a 10-cm spherical lens for point detection. Following the polarization gating setup described by [Peier 2008], crossed polarizers oriented at ±45° relative to the LN optic axis were placed in front of and behind the 33 μm LN crystal, whose static birefringence was compensated by a second, identical crystal whose optic axis was perpendicular to that of the sample. A quarter wave plate was placed immediately before the second polarizer. The THz field modulates the optical refractive indices through the electro-optic effect, and the induced phase shift between the vertical and horizontal polarization components is given by:

$$\Delta\phi = -2\pi\frac{\ell}{\lambda_{\text{opt}}}(\Delta n_{eo} - \Delta n_o) = -2\pi\frac{\ell}{\lambda_{\text{opt}}}\frac{r_{33}n_{eo}^3 - r_{13}n_o^3}{2}E_{\text{THz}} \tag{4.3}$$

with ℓ the sample thickness, λ_{opt} the probe wavelength, n_o and n_{eo} respectively the ordinary and extraordinary refractive indices for the probe, and r_{13} and r_{33} the appropriate electro-optic coefficients. The transmitted probe light after the second polarizer was sent into a PIN silicon photodiode. Bandpass filters were placed in front of the photodiode to eliminate scattered pump light, the pump beam was chopped at 500 Hz, and the signal was detected using a lock-in amplifier. By measuring the transmission as a function of time delay between pump and probe, we recorded the time-dependent THz field, $E(t)$.

Figure 4.3(a) shows the time-dependent field profiles of multi-cycle THz waves with central frequencies of 1.0, 0.6, 0.4, and 0.2 THz and a typical single-cycle THz wave generated in a bulk crystal under comparable conditions (dashed black) for comparison. The corresponding Fourier spectra in Fig. 4.3(b) have full-width at half maximum (FWHM) values ranging from 70 GHz to 120 GHz, much narrower than the single-cycle pulse. Despite the fact that dispersion is less, the percent bandwidth $(\Delta f_{\text{FWHM}}/f_0)$ tends to be smaller at higher frequencies because the velocity-matching length is longer relative to the THz wavelength.

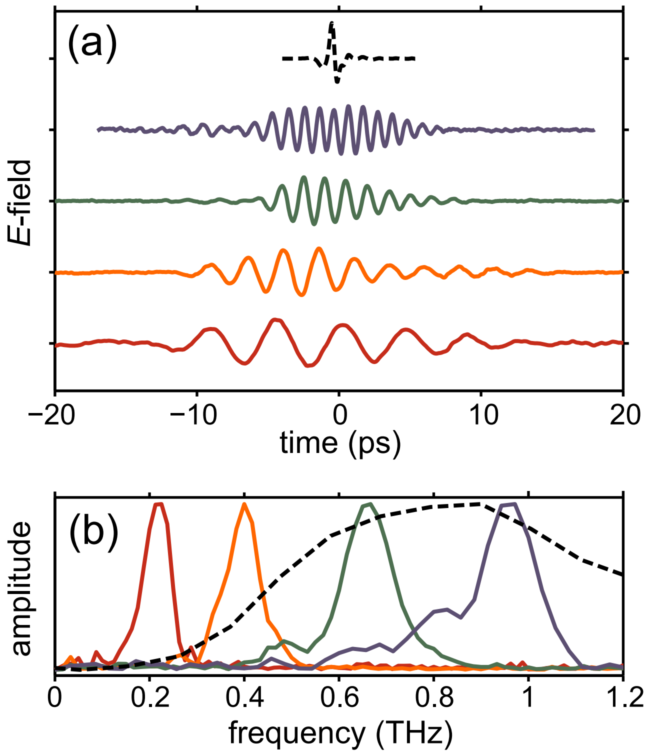

Figure 4.3 | The time traces and spectra of selected THz pulses. The measured THz waveforms when certain frequencies were selectively enhanced due to appropriate velocity-matching. (a) Time domain field profiles. (b) Corresponding Fourier spectra. The dashed black curve is a typical single-cycle THz pulse generated in a bulk LN crystal, shown for comparison.

Because phonon-polariton generation through ISRS is a second-order process [Yan 1985], for a given optical pump pulse energy the generation of THz is most efficient when the pump intensity is as high as possible within the limit imposed by the optical damage threshold of LN. In order to maximize THz field strengths, we decreased the spot size to below 1 mm in both dimensions at the LN waveguide to increase the optical intensity. We tuned to ~600 GHz where generation is most efficient, and measured peak-to-peak electric field amplitudes of ~50 kV/cm (data not shown). This value is 2-3 times larger than the maximum electric field that can be generated in the same LN slab with a non-tilted optical pulse of comparable intensity. Although the enhancement in peak field amplitude is somewhat modest, enhancements in spectral intensity at the center frequency can exceed 30 fold.

The values for THz waveform characteristics such as bandwidth, central wavelength, and amplitude presented in this paper do not span the entire range that is possible with this technique, particularly if LN slabs of different thicknesses are used. The bandwidth is determined by the waveguide dispersion curve, which is itself determined by the waveguide thickness. Thinner waveguides have more dispersion, so narrower bandwidths can be achieved with the same generation length. Thinner waveguides also have dispersion over a broader frequency range, so higher frequencies can be selectively enhanced, and ultimately the intrinsic dispersion of bulk LN will dictate the phase-matching conditions [Hebling 2004]. The peak THz amplitude (especially at higher frequencies where damping is strong at room temperature) can be increased considerably by reducing the sample temperature [Stepanov 2003].

4. Conclusions

In summary, we have demonstrated generation of narrowband THz waves in a planar waveguide. Due to the dispersion of the waveguide, it is possible to selectively phase-match femtosecond pulses with tilted intensity fronts to a desired THz frequency. The frequency can be tuned by changing the tilt angle of the pulse front, and the bandwidth can be set by changing the phase-matched interaction length. This technique is capable of generating high amplitude phonon-polariton waves in the waveguide because coherent amplification of the THz pulse as it propagates circumvents the limit imposed by the optical damage threshold.

B. PULSED LASER NOISE ANALYSIS AND PUMP-PROBE SIGNAL DETECTION WITH A DAQ CARD

Content from: C. A. Werley, S. M. Teo, and K. A. Nelson, “Pulsed laser noise analysis and pump-probe signal detection with a data acquisition card,” *Rev. Sci. Instrum.* **82**, 123108 (2011).

1. Abstract

A photodiode and data acquisition card whose sampling clock is synchronized to the repetition rate of a laser are used to measure the energy of each laser pulse. Simple analysis of the data yields the noise spectrum from very low frequencies up to half the repetition rate and quantifies the pulse energy distribution. When two photodiodes for balanced detection are used in combination with an optical modulator, the technique is capable of detecting very weak pump-probe signals ($\Delta I/I_0 \sim 10^{-5}$ at 1 kHz), with a sensitivity that is competitive with a lock-in amplifier. Detection with the data acquisition card is versatile and offers many advantages including full quantification of noise during each stage of signal processing, arbitrary digital filtering *in silico* after data collection is complete, direct readout of percent signal modulation, and easy adaptation for fast scanning of delay between pump and probe.

2. Introduction

Pulsed lasers are ubiquitous in modern laboratories, and they make it possible to study ultrafast dynamics through pump-probe techniques. This technique has enabled an important and diverse set of experiments including short pulse characterization [Ippen 1975], ultrafast motion of non-thermal electrons [Brorson 1987], resolution of terahertz (THz) electric fields [Auston 1988], and nuclear motion in proteins during electron transfer [Vos 1993], among many others. The pulsed lasers used in these studies experience shot-to-shot and long-term fluctuations in pulse energy, which complicates the study of phenomena with small experimental signals. Here we present a straightforward method that uses a data acquisition (DAQ) card to record the energy of every laser pulse for repetition rates less than ~500 kHz. Because the energy of every pulse is recorded, the method is very versatile and enables quantification of laser noise and arbitrary digital signal processing to measure weak pump-probe signals. Although DAQ card acquisition has been used by multiple labs in pump-probe experiments (e.g. [Ruhman 1988]; [Lim 1993]), we could find no paper detailing the implementation. We prefer this method to using a lock-in amplifier or boxcar integrator, and hope this demonstration will make it easier for others to adopt DAQ cards as a major tool for pump-probe signal acquisition.

In pump-probe measurements, the time-dependent response is recorded by stepping the delay between pump and probe pulses, and information is encoded as a pump-induced change in the probe pulse. The pump affects the real and/or imaginary parts of the sample’s index of refraction, which in turn modifies the amplitude, phase, or polarization of the reflected or transmitted probe. When the phase or polarization are modified, interferometry or polarization sensitive detection is generally used to convert changes in phase or polarization to changes in probe intensity, which are sensed by the detector. Consequently, regardless of how the pump influences the sample, the information is ultimately encoded in the energy of the probe pulse that reaches one or more detectors. When the pump is modulated, typically with every alternate pulse blocked, the probe light intensity I_0 reaching a detector with the pump blocked and the intensity I with the pump present can be measured and the fractional signal modulation $\Delta I / I_0$ can be retrieved by taking the ratio of consecutive probe pulse energies. It is very common that the desired signal is directly proportional to $\Delta I / I_0$ [e.g. Valk 2004; Wright 2008; Lessing 1976; Goodno 1998], and acquisition with the DAQ card gives rapid and straightforward access to this value.

3. Experimental methods

A photodiode detects the laser output, and the photodiode signal (a charge) is converted to a time-dependent voltage and amplified by a current preamplifier. The analog voltage output from the current preamplifier is digitized and recorded with the DAQ card [see Fig. 4.4(a)]. If the DAQ card is set to sample much faster than the repetition rate of the laser, the card behaves like a digital oscilloscope and records the time-evolution of the voltage. Figure 4.4(b) shows some example traces. In this demonstration, we used a titanium-sapphire laser with a regenerative amplifier to generate pulses at 1 kHz with a center wavelength of 800 nm, a pulse energy of 1.5 mJ, and a pulse duration of 80 fs. The photodiode was a low-noise, large area, unbiased silicon photodiode from Hamamatsu (model: S2281). We chose the current preamplifier (Stanford Research Systems, model: SR570) because of its low noise performance and filtering capabilities, whose use will be described below. The DAQ card, a National Instruments model: PCI-6281, had a maximum sampling rate of 625 kHz and an analog input with 18 bit resolution. At this bit depth, the discrete step size due to digitization, 76 μV, is very small compared to laser noise (see Fig. 4.4(d)) and digitization effects can be ignored.

To record the energy of every laser pulse, it was important to measure some parameter that was directly proportional to the pulse energy. The peak voltage and integrated voltage of an electrical pulse both satisfy this criterion. We chose to measure peak voltage (which we confirmed was linearly proportional to pulse energy through all signal processing steps), as it was easier to acquire. The dashed traces in Fig. 4.4(b) show the input to the DAQ card, where a 3 kHz lowpass filter was applied by the preamplifier. The photodiode signal was filtered so that the voltage evolved on a slower timescale than the maximum sampling rate of the DAQ card, but still decayed to zero before the next laser pulse arrived at the photodiode. Alternative analog signal processing devices, such as a boxcar integrator or a sample and hold circuit, could be as effective as the current preamplifier with proper implementation. We considered these devices but chose the current preamplifier because it carefully handles the current signal from the photodiode and has excellent noise performance. It is also possible to implement this detection system without any of these analog devices. Pulse energy measurement works for a 1 kHz laser if the input impedance of the DAQ card is manually set to ~100 kΩ and the DAQ card's on-board amplifier is used prior to signal digitization. Although eliminating these analog devices reduces the expense of the system, it loses the versatility of filtering and amplification. These capabilities enable the detection of lower probe pulse energies and higher laser repetition rates, so including the current preamplifier, boxcar integrator, or other analog device is preferred.

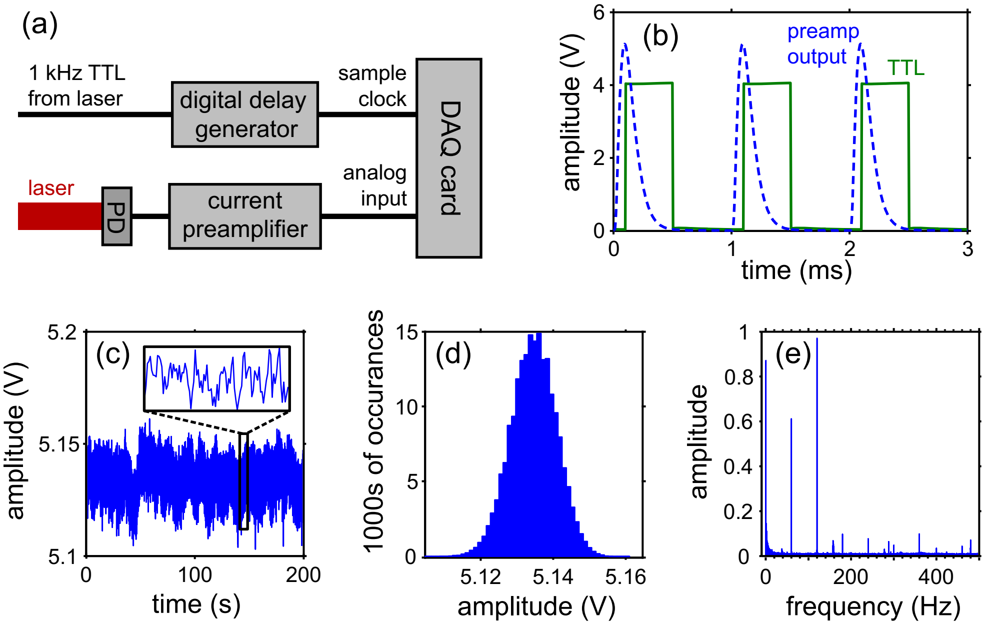

Figure 4.4 | Laser noise detection schematic and example. (a) The experimental setup for laser noise detection. The laser pulses are directed into a photodiode (PD) whose output is converted from current to voltage, amplified, and filtered to slow down the response. The measured voltage output from the current preamplifier is shown by the dashed line in (b). The sample clock of the DAQ card is given by an external TTL signal, which is synchronized to the laser. The TTL signal (shown as the solid line in (b)) is delayed such that the photodiode trace is sampled once at the peak of each pulse, effectively measuring the laser pulse energy. (c) The result of sampling 200,000 consecutive pulses, and the inset shows an expanded view of a short time period. (d) The distribution of the measured pulse energies. (e) The amplitude spectrum calculated by taking the Fourier transform of (c).

One feature of the DAQ card is that an external timing signal can be used as the sampling clock. In this configuration, whenever the card detects the rising edge of a logic signal, it immediately (within 1.6 μs, the delay is given approximately by the inverse of the sampling rate) records the voltage at the input of the analog-to-digital (A/D) converter. The TTL logic signal used as the sample clock, shown as a solid line in Fig. 4.4(b), was synchronized to the repetition rate of the laser. The TTL pulse train was delayed relative to the laser output with a digital delay generator (Stanford Research Systems, model: DG535) such that the rising edge of the TTL pulse corresponded to the peak voltage of the preamplifier output [see Fig. 4.4(b)]. In this way, the phase between the detected pulses and the sampling clock was stable. The peak voltage was measured for every pulse, and this voltage was directly proportional to the laser pulse energy. The voltage corresponding to 200,000 consecutive laser pulses is shown in Fig. 4.4(c) and a histogram of the measured values is shown in Fig. 4.4(d). In this system, laser noise is much greater than noise from the electronics (measured as described later), and essentially all the observed fluctuations in pulse energy come from the laser itself. In this data set, the laser was optimized for stability and displayed RMS fluctuations of only 0.12%. This value, as well as pulse duration and pulse energy, is an important indicator of laser performance.

Figure 4.4(e) shows the noise spectrum of the laser calculated by taking the Fourier transform of the time trace in Fig. 4.4(c). The noise spectrum is very valuable because it reveals the frequencies of strong noise components, which should be avoided when modulating the pump for lock-in detection. Many features of the noise spectrum shown here are typical for the three Ti:Saph lasers we measured, including a constant background of white noise, the $1/f$ noise at low frequencies [Bak 1987], and the noise at the AC line frequency (60 Hz here) and its harmonics. The strong signal at 120 Hz, in particular, results from the rectifying power supply for the diodes in the pump laser for the regenerative amplifier. In addition, there were typically some noise components, like the peak at ~160 Hz and the cluster around 280 Hz, that were unique to the laser, day, or room.

To adapt the technique to measure pump-probe signals, a conceptually similar but more complicated setup is required [see Fig. 4.5]. We employed standard strategies to achieve optimal noise reduction: balanced detection was implemented to suppress laser intensity fluctuations and the pump laser was modulated to shift the signal to higher frequencies and avoid $1/f$ noise. The digital delay generator, which was synchronized to the laser, provided the sample clock for the DAQ card and the synchronization frequency for the optical chopper. There were two photodiodes; one detected the probe before it interacted with the sample and one detected after interaction. Each photodiode output was filtered and amplified as described above with a current preamplifier, and both preamplifier outputs were recorded with the analog inputs of the DAQ card. We chopped the pump beam at 500 Hz, half the laser repetition rate, so that every other pump pulse was blocked. Because the digital delay generator set the sample clock, we measured the energy of every probe pulse arriving at each photodiode. Data acquisition by the DAQ card was triggered by the chopper controller, so that for every acquisition run, all the even numbered pulses had the pump present and odd numbered pulses had the pump blocked.

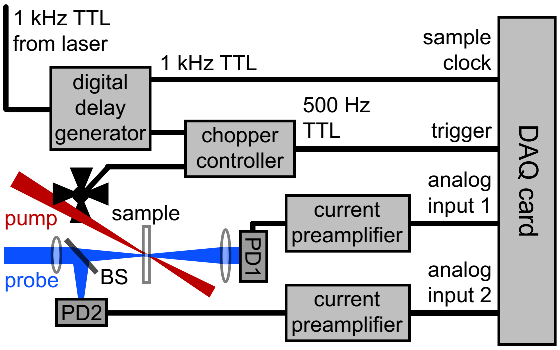

Figure 4.5 | Experimental setup for taking pump-probe traces with the DAQ card. The pump, which can be variably delayed relative to the probe, is modulated at half the laser repetition rate with a chopper, which is synchronized to the laser. A beamsplitter (BS) placed before the sample splits off part of the probe beam to implement balanced detection. The current outputs of both photodiodes are converted to voltage, filtered, and amplified by the current preamplifiers prior to digitization by the DAQ card. The DAQ card sampling clock is synched to the laser so that every laser pulse energy is measured at each photodiode. The trigger signal from the chopper controller assures that the pump is always blocked when the first pulse energy is acquired.

After all pulse energies are recorded, data processing in the computer is easily implemented. The signal from each photodiode for a pulse when the pump is present (pulse *a*) is normalized to the subsequent pulse when the pump is blocked (pulse *b*), correcting for differences in the optical paths. The normalized value from the diode after the sample (PD1) is divided by the normalized value from the diode preceding the sample (PD2), correcting for temporal fluctuations in the laser power. Finally we subtract 1 to retrieve the modulation:

$$\frac{\Delta I}{I_0} = \frac{\mathrm{PD1}_a}{\mathrm{PD1}_b} \Bigg/ \frac{\mathrm{PD2}_a}{\mathrm{PD2}_b} - 1 \tag{4.4}$$

It is convenient that the output of this processing is $\Delta I/I_0$, the physically relevant quantity in many experiments. Different detection schemes, such as balancing with a Wollaston prism or dark background detection, require different mathematical processing that can easily be implemented in software.

The maximum repetition rate laser to which this method can be applied depends on the photodiode, current preamplifier, and DAQ card used. For the hardware described above, the rate is limited by the DAQ card, which has a single analog-to-digital converter. The limit is set by the time it takes to sample analog input 1 followed by analog input 2, ~7 μs in practice. Because performance is best when the signal detected by each analog input channel is similar, the voltage trace should evolve on a significantly slower timescale than the sampling rate, and we found 20 kHz to be a practical upper limit. With a biased PIN photodiode, enough probe input power, and a simultaneous-sampling DAQ card with a sufficiently high sampling rate, it should be possible to push this method up to 500 kHz repetition rate lasers or beyond.

4. Results and discussion

a. Signal processing and noise

An advantage of the DAQ card detection system is that the noise can be quantified after each noise-suppression technique (i.e. chopping, balancing, and averaging) is implemented. Figure 4.6 shows the noise spectrum and histogram of values for different schemes, where each value has been normalized to the mean. To generate these plots, 200,000 pulses of 400 nm light were collected from a pair of photodiodes as shown in Fig. 4.5, and the values were processed in different ways. The number in the upper right of each histogram is the standard deviation of the distribution, which gives the RMS noise. The boxed number in the lower right of each spectrum is the factor by which the spectral amplitude was multiplied in that subplot. Figure 4.6(a) and 4.6(b) show the spectrum and histogram for the unprocessed laser output measured at a single photodiode. When these data were measured, the laser was much noisier than for the data shown in Fig. 4.6, and the distribution of values extends well beyond the bounds on the axes. Figure 4.6(c) & (d) show the output when every laser pulse from a single photodiode is divided by the subsequent pulse, simulating chopping. The frequency spectrum only extends to 250 Hz because pairs of pulses are combined to make a single value. Chopping can be thought of as a high-pass filter: it suppresses low-frequency noise but high-frequency noise is only slightly reduced. Figures 4.6(e) & (f) show the output if pulse energies are recorded by a pair of photodiodes, and one photodiode output is normalized by the other. This greatly suppresses noise at all frequencies, as can be easily seen by the much narrower distribution of values and the greater than 10-fold

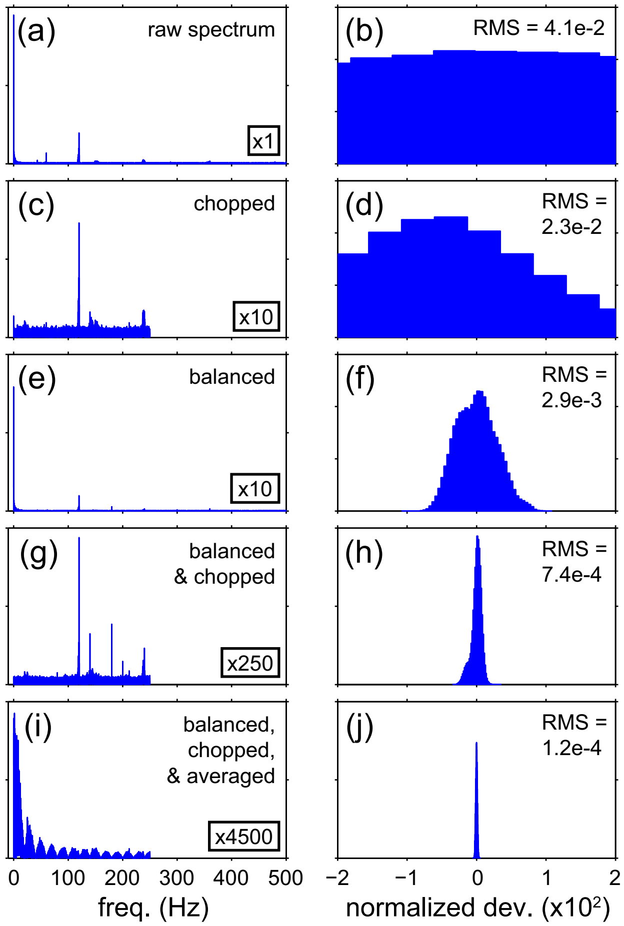

Figure 4.6 | Effects of signal processing on niose. The noise amplitude spectrum and histogram of measured values normalized to the mean for 200,000 pulses of 400 nm light processed in different ways. The amplitude for each spectrum was multiplied by the boxed factor in the lower right of the plot. The RMS fluctuation value in the upper right of the histogram plots is the standard deviation of the distribution, which reflects the efficacy of each processing step. (a) and (b) are from the raw pulse energies with no processing. For (c) and (d), each pulse from one photodiode was divided by the subsequent pulse to mimic chopping. In (e) and (f), the outputs from two photodiodes, as shown in Fig. 4.5, were divided to mimic balancing. (g) and (h) show the result of dividing one photodiode output by the other and then dividing ratios from successive time points to mimic chopping and balancing. (i) and (j) have the same processing as in (g) and (h), but the entire trace is convolved with a rectangle function 25 values long to mimic averaging.

reduction in noise from the unprocessed data. Figures 4.6(g) & (h) show the result of the operations in Eq. 4.4, which incorporates balancing and chopping. This displays good overall noise suppression and excellent suppression at low frequencies, which is why these noise suppression techniques are so widely implemented. It is possible to further reduce the noise by digital filtering of the data. Figure 4.6(i) & (j) show the result of the simplest digital filter: unweighted averaging. This can be thought of as convolving the pulse time sequence with a rectangular function, which is equivalent to multiplying by a sinc function in the frequency domain, as can be clearly seen in Fig. 4.6(i). We chose the number of pulses to average such that an integer number of cycles of the largest noise components were averaged over, almost completely suppressing those noise components. In other words, we chose the width of the convolving rectangular function so that the nodes of the sinc function in the frequency domain fell on top of the largest noise components. Here we averaged 25 pulse pairs, with a total averaging time of 50 ms or 3 cycles of the line frequency, which suppressed noise at 60 Hz and its harmonics. The 5.5-fold noise suppression between Figs. 4.6(h) & (j) is roughly consistent with $\sqrt{n}$ scaling predicted for averaging. Figure 4.6(j) has an RMS noise of only 1.2×10^{-4}, which highlights the capability to detect weak signals even with short averaging times when using this measurement system.

Even though the laser noise is significantly reduced by balancing and chopping, laser fluctuations are still the dominant noise source. Shot noise resulting from Poisson statistics in the probe has a magnitude of $1/\sqrt{N}$ [Yariv 1971], where N is the number of photons in the laser pulse. If we use pulse energies at the high end of the linear response regime of our photodiode, roughly 0.5 nJ/pulse at 400 nm or $\sim10^9$ photons/pulse, shot noise is around 3×10^{-5} for a single pulse. Because the shot noise for each pulse measured is uncorrelated, the noise variances add when performing the mathematical operations in Eq. 4.4, leading to a total contribution from shot noise of 6×10^{-5}, which is more than an order of magnitude smaller than fluctuations not fully suppressed by balancing and chopping. To quantify the noise arising from the measurement electronics, we blocked all light from hitting the photodiodes and used a setting on the current preamplifiers that generates a small DC offset current. This current was amplified and detected at 1 kHz by the DAQ card in the same way as the real data, then processed as if balancing and chopping were implemented. With well shielded cabling and good connections between the photodiode and the preamplifier, the normalized RMS deviation was 9×10^{-5}, significantly smaller than the laser noise from the balanced and chopped signal. When a very low-noise laser is used, such as that shown in Fig. 4.4 where the normalized RMS deviation is only 0.13%, it is possible for the balanced and chopped output to have an RMS deviation of $\sim3\times10^{-4}$ (data not shown), still significantly larger than noise from the electronics or shot noise. Thus the performance of the detection electronics is sufficiently good that measurements are limited by pointing and shot-to-shot energy fluctuations in the laser.

b. Comparison with a lock-in amplifier

Lock-in amplifiers [Temple 1975] are often used to measure pump-probe signals. Figure 4.7 shows a direct comparison of terahertz electric field time traces measured using a lock-in amplifier and the DAQ card system. In these measurements the terahertz field induced birefringence in an electro-optic crystal, changing the polarization of a transmitted optical pulse whose orthogonal polarization components were directed to two photodiodes for balanced detection [Valdmanis 1983]. For both measurements, we used the 1 kHz TiSaph laser system described above and implemented balanced detection and chopping at 500 Hz. For the lock-in, balancing was implemented by sending both signals into the instrument, a Stanford Research Systems SR810, which subtracted the inputs before performing additional signal processing. In this scheme, careful balancing is required for optimal noise suppression. Slow drift of laser pointing unbalances the diodes, which can adversely affect performance. For the DAQ card system, the mathematical divisions in Eq. 4.4 make careful balancing unnecessary and significantly reduce the sensitivity to drift in the balancing ratio.

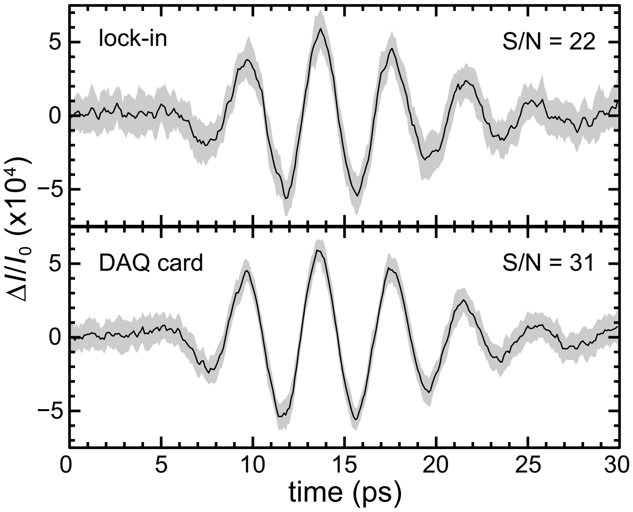

Figure 4.7 | DAQ vs. lock-in comparison. The time trace of a multi-cycle THz pulse with a center-frequency of 250 GHz measured using a lock-in amplifier (top panel) and the DAQ card measurement system (bottom panel). The total measurement time was the same for both methods. For each method, the trace was measured 25 times. The black is the average and the gray is $\pm\sigma$ for each time delay, so that 68% of measurements fall within the gray region. The S/N for each method compares the average uncertainty (which was roughly independent of time delay) to the peak signal.

In a lock-in amplifier, typically both the time constant τ and the filter roll-off are programmable. The roll-off reflects the number n of sequential *RC*-type filters the lock-in signal passes through (n = 1 for 6 dB/octave, n = 2 for 12 dB/octave, n = 3 for 18 dB/octave, etc.). Passing though n filters can be described mathematically as convolving the lock-in signal with an exponential function n times, or convolving the signal with the function:

$$f(t) = \frac{t^{n-1}}{(n-1)!} e^{-t/\tau} u(t) \tag{4.5}$$

with $u(t)$ the Heaviside step function [Oppenheim 1997]. The lock-in output at a time t most strongly reflects the peak of this function, which occurred earlier at a time $t-(n-1)\tau$. It takes more time constants for the majority of the signal to decay out of the buffer when changing conditions (e.g. the time delay in a pump-probe measurement). In our measurement, we chose τ = 30 ms and a fourth order, 24 dB/octave filter. We waited 8 time constants, 240 ms, so that there was a roughly 4% contribution to the signal at each time delay from the previous time step. 250 pulses were averaged for the DAQ card, which never has parasitic contributions from previous time delays. Using these settings, the number of pulses averaged at each delay and the total scan time were roughly the same for both methods. The time trace was collected 25 times for each method to build up statistics for signal to noise ratio (S/N) comparisons. In Fig. 4.7, the black line shows the average of the traces and the gray area shows one standard deviation (σ) on either side of the average for each time delay, giving the spread in the measurements. As can be seen from the figure, the traces look very similar and the data have a similar spread. The standard deviation was basically independent of time delay for both methods, so the S/N was calculated with the maximum signal and the standard deviation averaged over the delay time as: $\mathrm{S/N} = \max\left(\Delta I / I_0\right) / \left(\sqrt{n}\langle\sigma\rangle\right)$. The S/N is about 40% better for the DAQ card here, but in our experience the relative performance of the two techniques depends on the nature of the laser fluctuations, and overall both methods have similar minimum-signal detection thresholds.

c. Rapid scanning

In spite of the similarity in detection thresholds for long averaging times, the DAQ card system has some advantages at short averaging times. Because the lock-in multiplies the input voltage by a sine wave and convolves with the expression in Eq. 4.5, it is necessary for the lock-in time constant to be significantly longer than the modulation period. If this is not the case and the expression in Eq. 4.5 evolves significantly during a single modulation period, quantification of the signal amplitude will be impaired. This is not the case in the DAQ card system because it uses equally weighted, rectangular function averaging. The minimum averaging time is one modulation period, or two laser pulses if modulation occurs at half the repetition rate. Another disadvantage of the exponential averaging implemented by the lock-in is that when stepping to a new time delay in a pump-probe measurement, it is necessary to wait many time constants for the signal from the previous time delay to decay away. This is why we had the 8τ waiting time when we took the lock-in trace in Fig. 4.7. No waiting is necessary with rectangular function averaging, so a time trace can be acquired more quickly with the DAQ card.

A capability made possible by the short averaging time required by the DAQ card system is rapid scanning. To implement this, we used an ILS150PP delay stage and ESP300 stage controller from Newport. The stage was accelerated to a target velocity, at which point the DAQ card acquired signal continuously until the desired time delay was reached. Data acquisition was then stopped, the stage was quickly returned to the starting position, and the whole process was performed repeatedly. Knowing the stage velocity and laser repetition rate, it is straightforward to measure femtosecond or picosecond time traces with accurate time dynamics, assuming the signal is large enough to measure with minimal averaging. We were able to display 10 ps time scans and their corresponding frequency spectra at a refresh rate of ~1 Hz (data not shown). This rapid scanning and real-time feedback was very useful for improving optical alignment and optimizing pump-probe overlap, pulse compression, and pump pulse spatiotemporal shaping. The delay stage and controller we used to implement fast scanning were not chosen with this application in mind, and a delay stage with a faster acceleration will enable even shorter acquisition times.

d. Differential chopping

When chopping low repetition rate lasers (i.e. ~10 kHz or below), generally one requires that the chopping is implemented such that pump pulses either pass completely through the chopper wheel without clipping or are fully blocked. This ensures that the pump intensity profile at the sample is never altered. To implement this cleanly, the pump is chopped at an even integer fraction of the repetition rate *R* (e.g. *R*/2: one passed, one blocked; *R*/4: two passed, two blocked; *R*/6; etc.). When two pump lasers are directed into the sample, the coupling between their induced effects can be sensitively detected by implementing differential chopping. In this scheme, the pumps are chopped at different frequencies, and the signal is detected with a lock-in amplifier at the difference frequency. If one requires chopping at even integer fractions, the fastest difference frequency that is not degenerate with one of the fundamental modulation frequencies is achieved by chopping one pump at *R*/4, the other at *R*/6, and detecting at *R*/12, which extracts only the signal resulting from the coupling between the pumps. The noise can be a problem because of the low detection frequency, and it takes additional work to extract signals from just one pump or the other. Although we have not tested it, we believe that the DAQ card method offers an elegant alternative. In the DAQ card scheme, one pump would be chopped at *R*/2 and the other at *R*/4. Looking at four sequential probe pulses, the state would be S_{++}, S_{+-}, S_{-+}, and S_{--}, where "+" means the pump is present and "−" means the pump is absent. From these four values all the relevant information can be extracted. The signal resulting from both pumps, the first pump only, the second pump only, and the coupling between the pumps is given by $S_{++}/S_{--}-1$, $S_{+-}/S_{--}-1$, $S_{-+}/S_{--}-1$, and $\left(S_{++}-S_{+-}-S_{-+}+S_{--}\right)/S_{--}$, respectively. In addition, all four signals are detected at *R*/4, a relatively high frequency, leading to effective noise suppression. As a result, the DAQ card system offers multiple advantages in a differential chopping scheme.

5. Conclusions

The detection system described here, which consists of a photodiode, a current preamplifier, and a data acquisition card clocked by the laser, is a versatile and useful tool for pump-probe experiments. When the DAQ sampling rate is driven by its internal clock, the system acts like an oscilloscope. When the DAQ card's sampling clock is a TTL pulse train synchronized to the laser repetition rate, the intensity of every pulse is measured. This enables laser noise quantification and measurement of the noise spectrum, much like a spectrum analyzer. When the DAQ card system is used with pump modulation and balanced detection, it is able to measure very small signals in much the same way as a lock-in amplifier. Because the energy of every laser pulse is measured, it is possible to perform arbitrary linear or nonlinear filtering on the data to further reduce the noise. For example, with a more unstable laser it might be beneficial to set a threshold so that abnormally large or small pulses would be discarded and not included when calculating average behavior, as has been done previously (e.g. [Ruhman 1988]). A final advantage is fast scanning, which facilitates the optimization of optical alignment and real-time feedback on the fast dynamics in the system.

In addition to these diverse capabilities, data processing in the DAQ system is intuitive. Signal and reference are measured both through the sample and avoiding the sample, and everything is divided to cancel out temporal fluctuations and path differences. This processing scheme, which is implemented in many common instruments, is simple to understand for experimenters new to this kind of technique. In addition, the direct output is $\Delta I/I_0$, which is useful for comparing signal magnitude across different experiments and is often proportional to some relevant physical quantity of the system. Combined with the quantification of laser noise, this helps build intuition about the system under study and the experimental apparatus. DAQ cards have been and are used for pump-probe measurements by various labs, but they are not as ubiquitous as lock-in amplifiers. Because of its conceptual simplicity and general versatility, we believe that the DAQ card measurement system merits wider adoption in the pump-probe community.

C. A COMPARISON OF PHASE SENSITIVE IMAGING TECHNIQUES FOR STUDYING THz WAVES IN STRUCTURED LiNbO$_3$

Content from: C. A. Werley, Q. Wu, K.-H. Lin, C. R. Tait, A. Dorn, and K. A. Nelson, "A comparison of phase sensitive imaging techniques for studying THz waves in structured LiNbO$_3$," *J. Opt. Soc. Am. B* **27** 2350-2359 (2010).

1. Abstract

Four phase-sensitive imaging methods (Talbot, phase contrast, Sagnac, and polarization gating) used for detecting terahertz-frequency waves in structured lithium niobate slabs are compared analytically and experimentally. It is demonstrated that both phase contrast and a self-compensating polarization gating geometry can generate in-focus images of the sample and quantitatively measure the terahertz electric field. Of these two methods polarization gating has better signal-to-noise ratio and so is preferred for most situations, while phase contrast imaging has better spatial resolution and so is preferred for measurements involving fine structures or near-field effects.

2. Introduction

Technology relating to terahertz (THz) radiation has historically lagged behind technology relating to the other regions of the electromagnetic spectrum, although this has been rapidly changing due to an international effort to develop THz methods [Lee 2009]. THz techniques employ a very diverse set of tools, and in many cases great care and effort are required to use these tools in concert. It has been suggested that the nonlinear crystal lithium niobate (LN) could serve as an integrated platform for THz generation, control, and detection [Feurer 2007]. THz "on a chip" or "polaritonics" has the potential benefits of being compact, efficient, versatile, and simple, and broad set of capabilities has already been demonstrated in LN. Using femtosecond optical pulses, it is possible to generate THz phonon-polariton waves with high field strengths and a tunable center frequency and bandwidth [Lin 2009], a single cycle, or an arbitrary field profile imprinted through reconfigurable spatial, temporal, or spatiotemporal shaping of the optical intensity profile [Feurer 2003]. The generated waves can be focused [Hornung 2007], diffracted [Stoyanov APL 2002], waveguided [Stoyanov NM 2002], or filtered [Statz 2007, Peier 2009] by structures which have been cut into the crystal through laser micromachining [Ward 2007]. A final advantage of the LN system is that it is possible to follow the spatial and temporal evolution of the THz waves by using variably delayed optical probe pulses for time-resolved real-space imaging [Koehl 1999], which permits determination of the full THz field under a wide range of conditions [Feurer 2007, Feurer 2003, Hornung 2007, Stoyanov APL 2002, Stoyanov NM 2002, Statz 2007, Peier 2009, Ward 2007, Koehl 1999, Wahlstrand 2003].

LN (with or without THz polariton waves) is transparent to probe light in and near the visible spectral region so no information is encoded in the amplitude of the probe. However, because LN is an electro-optic crystal the phase can be modulated through a THz-induced change in the index of refraction: $\Delta\phi(x,z) \propto \Delta n(x,z) \propto E_{\mathrm{THz}}(x,z)$. If a camera, an intensity detector, were placed immediately after the sample the phase pattern would not be visible. Instead, various methods are needed to convert the phase information into amplitude information. In this paper we compare four phase-sensitive imaging techniques which can be used to detect THz waves: Talbot imaging, Sagnac interferometry, phase contrast imaging, and polarization gating. Talbot imaging has been used the longest for detecting waves in the LN platform [Koehl 1999]. Improved techniques were more recently introduced: Sagnac and polarization gating imaging in 2008 [Peier 2008] and phase contrast imaging in 2009 [Wu 2009]. A comparison of several methods has been presented for imaging THz waves in unstructured LN [Peier 2008]. Here we show that the need to image THz waves *and* sample structures with which they are interacting introduces new considerations in the comparison among imaging methods. This has significant consequences for assessing the advantages and disadvantages of each. Here we introduce an improved, self-compensating geometry for polarization gating imaging and compare it with the other methods for imaging THz in structured samples.

The experimental setups for the four methods are illustrated in Fig. 4.9(c)-(f). All of these methods use interference to convert phase to amplitude information, but the interference is implemented in very different ways. In Sagnac imaging, interference occurs between two beams that have been separated by a beamsplitter. In polarization gating imaging, the vertically and horizontally polarized components of the beam interfere. In phase contrast and Talbot imaging, the light diffracted by the THz wave pattern interferes with the 0^{th} order (undiffracted) beam. These methods have different advantages and disadvantages, but all are executed using common-path optics, which ensures passive relative optical phase stability between the two interfering light fields throughout the course of the experiment.

The goal for all four methods is to quantitatively extract the two-dimensional spatial profile of the THz electric field as the wave undergoes propagation, dispersion, reflection, interference, etc. in the host crystal. The data are compiled in the form of a video clip in which each frame shows the THz electric field at a different time after the THz wave has been launched. Figure 4.8 shows frames from such a video at a few selected time points. In this example, which is reminiscent of the classic single-slit diffraction experiment, a rightward propagating wave interacts with a LN bridge structure. Air gaps, masked in light grey in Fig. 4.8(a), have been carved into a thin LN slab. Most of the wave is transmitted through the bridge, which is about three wavelengths wide. After the wave emerges from the bridge, it diffracts and evolves into the familiar far-field pattern. As an example of the information that can be extracted from any time or position within the movie, Fig. 4.8(f) shows the E-field measured by averaging over the vertical dimension within the dashed box in Fig. 4.8(e). It is clear that if such data sets can be recorded accurately, they provide a wealth of qualitative and quantitative insight into THz wave behavior in structured and unstructured media.

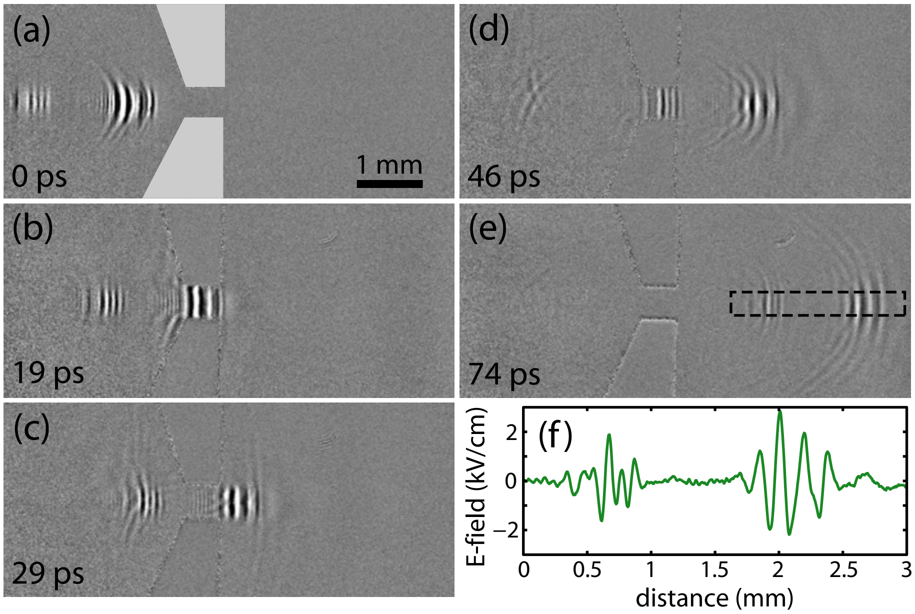

Figure 4.8 | Sequential images of a THz wave diffracting through a slit. (a)-(e) Polarization gating imaging frames from a video clip of a rightward propagating THz wave. The wave consists of two several-cycle wavepackets separated by 27 ps. Two air-gaps [light grey in (a)] were generated by removing these regions from the LN slab using laser machining. A thin LN bridge, ~3 wavelengths tall, remains. The THz wave passes through the gap and diffracts as it is emitted. (f) An *E*-field trace extracted by averaging over the vertical dimension of the dashed box in (e).

For the data in Fig. 4.8 and the rest of the paper, we will discuss THz waves propagating in 53 μm thick LN samples. For frequencies less than 2 THz (i.e. wavelengths longer than about 30 μm in the crystal) these thin samples are best described as dielectric slab waveguides. In this frequency range, material dispersion and absorption in lithium niobate are small, and we can neglect these effects for the millimeter propagation lengths of interest. Thus the THz field evolves as a set of transverse electric (TE) waveguide modes propagating in a dielectric slab of constant index n = 5.11 [Feurer 2007]. One advantage to working in the planar waveguiding regime is that the thin sample is entirely in the image plane, so structures will be clear and in focus if the appropriate imaging method is used. Another advantage is that the geometrical dispersion experienced by the guided THz wave enables high-amplitude, multiple-cycle THz wave generation [Lin 2009], which is useful for study of a variety of systems. A third advantage is that the strong evanescent field of the waveguide mode can interact with any material deposited on the crystal surface. These capabilities make it possible to generate, control, and detect the THz field and to study samples of interest in the integrated platform of the LN waveguide.

3. Experimental approaches and methods

For all the imaging measurements described below, a THz wave was generated through impulsive stimulated Raman scattering [Dougherty 1992] by focusing an intense pump pulse (800 nm, 70 fs, 400 μJ at 1kHz) through an *x*-cut LN slab using a cylindrical lens. The width of the cylindrically focused "line" of pump light was 20 μm. The pump pulse, which was variably delayed relative to the probe pulse and polarized parallel to the crystal's optic (*z*) axis, was directed at normal incidence through the crystal slab [Fig. 4.9(a)-(b)]. The generated THz wave, which was also polarized along the optic axis, traveled in the crystal plane orthogonal to the direction of pump propagation [Fig. 4.9(b)].

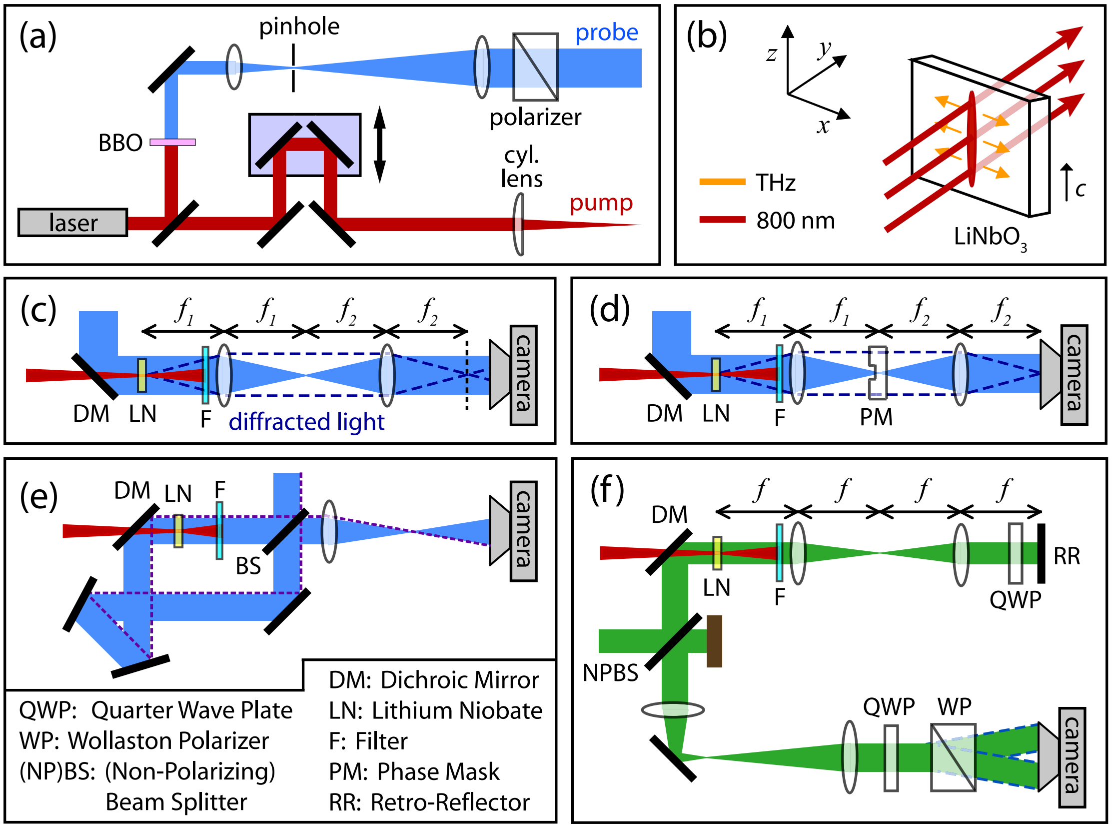

Figure 4.9 | Experimental setups. (a) Optical components common to all four imaging techniques. (b) The pumping geometry showing that the THz wave propagates in the plane of the LN slab, orthogonal to the generating pump light. (c) Talbot imaging (d) Phase contrast imaging (e) Sagnac imaging; the dashed line follows one edge of the beam through the interferometer. (f) Polarization gating. Instead of 400 nm light from second harmonic generation in a BBO crystal, the probe was 532 nm light generated from a non-collinear optical parametric amplifier (NOPA).

In all the experiments, we probed at a different wavelength from that of the optical pump so a filter could be used to block the pump light. For the first three techniques the probe was at 400 nm, the second harmonic of the pump. Because LN is a nonlinear crystal, some frequency doubling of the pump light occurs. The resulting 400 nm light could not be removed by a spectral filter and was blocked by a spatial filter in the form of a razor blade. For the polarization gating measurement, instead of probing at 400 nm we used the output of a home-built, non-collinear optical parametric amplifier (NOPA) tuned to 532 nm. This allowed us to use a single interference filter to remove both the 800 and 400 nm light from the pump. Regardless of its wavelength, the probe beam was spatially filtered and expanded so it was larger than the LN crystal [Fig. 4.9(a)], and was made nearly collinear with the pump beam using a dichroic mirror immediately before the sample.

A Princeton Instruments PIXIS 2048x2048 pixel, 16-bit, thermoelectrically cooled CCD camera was used to capture the images. This minimized electrical noise, so that the dominant sources of noise were shot-to-shot laser intensity fluctuations and pointing instability. As will be demonstrated below more rigorously for each imaging method, the THz *E*-field is proportional to the signal modulation: $E_{\mathrm{THz}}(x,z) \propto \Delta\phi(x,z) \propto P(x,z) \equiv \Delta I(x,z)/I_0(x,z)$ [Magnification]. Thus to determine the THz *E*-field, it is necessary to measure the probe intensity both in the presence (I) and absence (I_0) of the THz wave. Because the laser noise was strongest on slow time scales, it was beneficial to measure the signal image *S* with the THz wave present and a reference image *R* with the THz wave absent as close together in time as possible. The use of a shutter in the pump beam, which opened and closed on a 500 ms time scale, was ineffective because of mild sample and filter heating and thermal expansion. The thermal expansion produced a phase shift in the probe, and because it was modulated in the same way as the THz generation, it obscured the signal. To eliminate slow thermal effects but still maintain a short time between the acquisition of the signal and reference images, we placed optical choppers (New Focus model 3501) in both the pump and probe beams, and chopped both beams at 500 Hz with the chopper phases under computer control. The signal and reference images were acquired with the choppers in and out of phase, respectively. The phase cycling sequence was repeated as needed for signal averaging at a repetition rate of ~1 Hz. In this scheme the pump always irradiated the sample at 500 Hz, and since the timescale for steady-state heating or cooling was much slower than 2 ms, there were no detectable thermal effects in the signal. Small amounts of non-probe light reaching the detector, such as room light or scattered pump light, were accounted for by recording a background image *B* with the probe blocked. The signal modulation due to the THz field for each image was calculated as $\Delta I/I_0 = (S-B)/(R-B)-1$, which was then used for quantitative determination of $\Delta\phi$.

a. Talbot imaging

The Talbot effect, in which a periodic phase or amplitude pattern in a beam is reproduced again and again as the beam propagates through space, was first reported in 1836 by H. F. Talbot [Talbot 1836]. In this classic experiment Talbot observed a ray of sunlight after it had been transmitted through a grating provided by Joseph von Fraunhofer. This self-imaging phenomenon has been studied extensively since then [Patorski 1989], and was first applied to imaging of THz waves in 1999 [Koehl 1999]. Talbot imaging has been the most commonly used technique for THz imaging up to now [Feurer 2003, Hornung 2007, Stoyanov APL 2002, Stoyanov NM 2002, Statz 2007] because of the ease of experimental setup. In this method, the camera is moved out of the image plane of the sample [Fig. 4.9(c)] so that spatially varying phase information interferes with the 0^{th} order beam, bringing about phase to amplitude conversion. As has been discussed [11], for a particular displacement y of the camera from the image plane, a free-space probe wavelength λ, and n a positive, odd integer, the part of the phase image with a wave vector $k_x = 2\pi\sqrt{n/2\lambda y}$ will be completely converted to an amplitude image while all other wave vectors will be converted to amplitude information with lower efficiency. In addition to the wave vector-dependent response, which hampers quantitative measurement of the THz electric field, moving the camera out of the image plane blurs any fabricated structure in the LN crystal and obscures interactions between wave and structure. These drawbacks, which have been carefully analyzed by Peier et al [Peier 2008], have motivated development of in-focus imaging methods that enable quantitative field measurement. It should be noted that Talbot imaging can still be useful for some qualitative applications because it is very easy to align and requires no specialized optics.

b. Phase contrast imaging

Phase contrast imaging, an important technique for imaging of transparent samples in biology, was first introduced by Frits Zernike in the 1930s [Zernike 1942], and a good overview of the development of this technique is given in his Nobel speech [Zernike 1955]. The phase contrast method was recently applied to the imaging of THz fields [Wu 2009]. As in Talbot imaging, phase-to-amplitude conversion occurs by interference between the diffracted light and the 0^{th} order beam, but in this case the camera is located in the image plane, so both fabricated structures and THz fields are in focus. In this geometry the sample is imaged onto the camera using a two-lens, 4*f* imaging system [see Fig. 4.9(d)]. In the back focal plane (the Fourier plane) of the first lens, the main, undiffracted beam, also called the 0^{th} order beam, is focused to a small spot. In contrast, the high-wave-vector components of the light, the light that diffracted from structures in the sample or the THz wave, are spread throughout the plane and are spatially separated from the 0^{th} order beam. A phase plate with a 30 μm square, recessed region is placed in this plane such that only the 0^{th} order beam passes through this recessed square, which introduces a π/2 phase shift between the 0^{th} order beam and the diffracted light. When these two fields are recombined in the image plane (the back focal plane of the second lens) they interfere, bringing about phase to amplitude conversion. Using the small phase shift approximation, the measured intensity can be shown to be $I(x,z) \cong I_0(x,z)[1-2\Delta\phi(x,z)]$ [Goodman 2005]. Thus the signal modulation for phase contrast imaging is:

$$P_{\mathrm{PC}}(x,z) \equiv -\frac{\Delta I(x,z)}{I_0(x,z)} \cong 2\Delta\phi_{\mathrm{PC}}(x,z) \tag{4.6}$$

and the phase can be quantitatively extracted.

c. Sagnac interferometer imaging

The Sagnac interferometer was first developed by Georges Sagnac in 1913 to detect the relative motion of ether [Sagnac 1913]. The Sagnac geometry has been used extensively since then to detect various ultrafast phenomena (i.e. [Gabriel 1991]). In addition to point source detection, it has been used for imaging by scanning of the detection point. This has been used to study, for instance, propagating surface acoustic waves [Tachizake 2006]. Use of a Sagnac interferometer for imaging of THz waves was first demonstrated in 2008 [Peier 2008]. In this method the light is divided by a beam splitter, and one beam travels clockwise and the other counterclockwise around a ring [Fig. 4.9(e)]. Every time a beam reflects off a mirror, it is inverted through the vertical plane. For the beams to overlap properly after they are coupled out of the interferometer there must be an odd number of mirrors in the ring [see the dashed line along beam edge in Fig. 4.9(e)]. This ensures that the image of any structure in the crystal will be oriented the same way. The sample is placed so that the counterclockwise traveling wave passes through the sample before the pump pulse while the clockwise propagating wave passes through the sample after the pump pulse. The lenses after the Sagnac ring image the sample and the THz waves onto the camera for the clockwise ring. For the counterclockwise arm, the distance from the sample to the first lens is much longer, so this shows up as a blurred image on the camera. The *E*-field at the camera is:

$$E(x,z) = E_{\mathrm{CW}}(x,z)\exp\left[i\Delta\phi(x,z) + i\alpha(x,z)\right] + E_{\mathrm{CCW}}(x,z), \tag{4.7}$$

with $E_{\mathrm{CW}}(x,z)$ the in-focus image from the clockwise beam, $\Delta\phi(x,z)$ the THz-induced phase shift, $\alpha(x,z)$ the phase shift from the Sagnac interferometer, and $E_{\mathrm{CCW}}(x,z)$ the out-of-focus image from the counterclockwise propagating beam. The image detected by the camera is:

$$I(x,z) = EE^{*} = I_{\mathrm{CW}}(x,z) + I_{\mathrm{CCW}}(x,z) + 2E_{\mathrm{CW}}(x,z)E_{\mathrm{CCW}}(x,z)\cos\left[\Delta\phi(x,z) + \alpha(x,z)\right]. \tag{4.8}$$

For simplicity, we can assume that $\alpha(x,z) = \pi/2$ and Taylor expand in $\Delta\phi$. After dividing by the image measured when the pump is blocked ($\Delta\phi = 0$, $I_0 \equiv I_{\mathrm{CW}} + I_{\mathrm{CCW}}$) and subtracting 1, the measured signal modulation for Sagnac interferometry is:

$$P_{\mathrm{SI}}(x,z) \equiv -\frac{\Delta I(x,z)}{I_0(x,z)} \cong \frac{2E_{\mathrm{CW}}(x,z)E_{\mathrm{CCW}}(x,z)}{I_{\mathrm{CW}}(x,z) + I_{\mathrm{CCW}}(x,z)}\Delta\phi_{\mathrm{SI}}(x,z). \tag{4.9}$$

To quantitatively extract the phase from this measurement, we need to independently measure $E_{\mathrm{CW}}(x,z)$ and $E_{\mathrm{CCW}}(x,z)$, which is not possible because the beams are co-propagating. In the experiment, this means that the diffracted light from the out-of-focus, counterclockwise beam shows up in the image and obscures the signal.

In principal this problem could be fixed by introducing imaging optics *inside* the Sagnac ring. One setup that would do this is a 4-lens, 8-*f*, non-inverting imaging system which would use the CCW beam to create an image of the sample inside the Sagnac ring which is an equal distance from the beam-splitter as the sample itself. In this way images from both beams would be overlapping and in-focus. The significant difficulty involved with aligning this system coupled with the difficulty of controlling $\alpha(x,z)$, which we conveniently assumed to be π/2 in the analysis above, makes Sagnac imaging an inconvenient method. Because phase contrast imaging described above and polarization gating described below do not have these issues, we discard Sagnac imaging as a competitive method for quantitative imaging of THz waves in patterned LN samples. We note that in clean, unpatterned samples Sagnac imaging is more viable because the unfocused CCW image does not have spatially varying features [Peier 2008].

d. Polarization gating imaging

Polarization gating, where the pump pulse introduces a transient birefringence in the sample, has long been a tool in ultrafast optics. It has been used extensively for studying relaxation dynamics in liquids through the optical Kerr effect (OKE) [Greene 1982, Kinoshita 1995] and for measuring THz *E*-field time traces through electro-optic sampling [Auston 1988, Valk 2004]. This method was first used to image THz fields propagating collinearly with the optical probe in ZnTe [Wu 1996, Jiang 1999] and more recently discussed in the non-collinear geometry used here [Peier 2008]. In our setup, the probe light is polarized at 45 degrees and the vertically and horizontally polarized beams are phase shifted by different amounts as they pass through the sample. The small electric field-induced phase shift rides on top of a much larger phase shift between the two polarizations caused by the inherent birefringence of the LN crystal. We initially tried to compensate the inherent birefringence using a second LN slab of equal thickness with its optic axis rotated by 90 degrees, as demonstrated earlier [Peier 2008]. However, sample imperfections and non-uniformity in the slab thickness made it impossible to sufficiently extinguish the beam across the entire sample when two orthogonal crystals were placed between crossed polarizers. The sensitivity of the polarization state to small crystal variations was heightened by the large birefringence of LN ($n_o = 2.44$, $n_{eo} = 2.31$) [Feurer 2007].

To relax the requirements on sample quality, we modified the setup to be self-compensating, as shown in Fig. 4.9(f). The crystal is imaged onto a retro-reflecting mirror by a two-lens telescope with a magnification of 1. Immediately before and after reflecting off the mirror, the beam passes through a quarter wave plate with its fast axis rotated to 45°. The double pass through the QWP flips the vertical and horizontal polarizations. The inverted image at the retroreflector is imaged back onto the sample by the same lens pair, forming a correctly oriented image with the polarizations reversed. In this way the crystal self compensates any phase shift introduced by its inherent birefringence or by any static inhomogeneities in its thickness. The pump launches the THz wave in the sample shortly before the first pass of the probe. By the time the probe passes through the sample a second time the THz wave has long since decayed away entirely, so while the inherent birefringence is fully compensated, the detection technique is sensitive to the THz-induced birefringence. After passing through the sample the second time, the retroreflected beam is imaged onto the camera. *En route*, it passes through the non-polarizing beamsplitter, a QWP at 45°, and a Wollaston prism at 0° which splits the beam into vertical and horizontal polarizations, both of which are imaged onto different regions of a single CCD chip, enabling balanced imaging. The 45° difference between the orientation of the QWP and Wollaston places the interfering fields $\pi/2$ out of phase, which gives the strongest, most linear response and assures that the induced signals in the two images have opposite signs.

Using the Jones calculus formalism [Yariv 2007] and assuming a perfect beamsplitter and mirrors (equal reflectivities for *s* and *p* polarizations and no relative phase shift between the polarizations, a reasonable assumption if optics are chosen carefully), the intensity detected at the camera is:

$$I^{\pm}(x,z) = I_0^{\pm}(x,z)[1 \pm \sin(\Delta\phi(x,z))] \tag{4.10}$$

with $I_0^{\pm}(x,z)$ the intensity in the absence of the THz field, and "+" and "-" used to indicate vertical and horizontal polarizations respectively. Because the signal modulation is equal in magnitude but opposite in sign for the two polarizations, subtracting doubles the signal strength while cancelling noise, which has the same sign in both sub-images. To correctly overlap the two images for subtraction *in silico*, an affine transformation [Goshtasby 2005], which consists of translation, rotation, scaling, and shear, was applied in Matlab to one of the images which maps a set of reference points onto corresponding points in the other image. The transform allowed subtraction to be performed without any noticeable blurring of structure edges or other sharp features. Taylor expanding in $\Delta\phi$, the percent signal modulation for polarization gating is:

$$P_{\mathrm{PG}}(x,z) \equiv \frac{I^{+}(x,z)}{I_{0}^{+}(x,z)} - \frac{I^{-}(x,z)}{I_{0}^{-}(x,z)} \cong 2\Delta\phi_{\mathrm{PG}}(x,z). \tag{4.11}$$

Thus it is possible to quantitatively extract the phase shift and additionally cancel low wave vector spatial noise due to the balanced imaging.

4. Results and discussion

The methods described above can be compared experimentally by analyzing both the real space (position/time) and Fourier space (wave vector/frequency) responses. For each technique, a THz wave was generated by a line focus in an unstructured LN crystal (see video clip in [Wu 2009]). Because of the line focus, there was uniformity along the vertical (z) dimension. For each time frame in the movie, averaging over the vertical dimension, as was done to calculate Fig. 4.8(f), yields the spatial field profile in the lateral (x) dimension. For each time delay, the resulting vector was placed in a different row of a 2-D matrix. This matrix, a color-map of which is shown for each imaging method in Fig. 4.10, contains the full spatial and temporal evolution of the propagating THz waves. Two counter-propagating waves were launched at $t = 0$, and as they propagated, the different waveguide modes separated and the THz pulses became highly chirped due to waveguide dispersion. At intermediate times the waves reflected off the crystal edges, and at late times the two pulses again overlapped and interfered with one another. For the Sagnac method [Fig. 4.10(c)], the propagating signal is cut by vertical nodes. These result from diffraction off the edge of the crystal which showed up as fringes in the CCW propagating beam. Their appearance was predicted in the Sagnac analysis above, and this spatially non-uniform response (exacerbated in a patterned LN crystal) makes Sagnac imaging a poor method. Fringes can also be seen in Talbot data [Fig. 4.10(a)], although it is not obvious from this particular figure. For a clearer example of out-of-focus diffraction fringes obscuring the signal, see for instance [Stoyanov NM 2002]. The phase contrast method [Fig. 4.10(b)] also shows some weak artifacts near the edge of the crystal due to the large change in the static phase shift, but this is a fairly small effect. The polarization gating method [Fig. 4.10(d)] is basically artifact free.

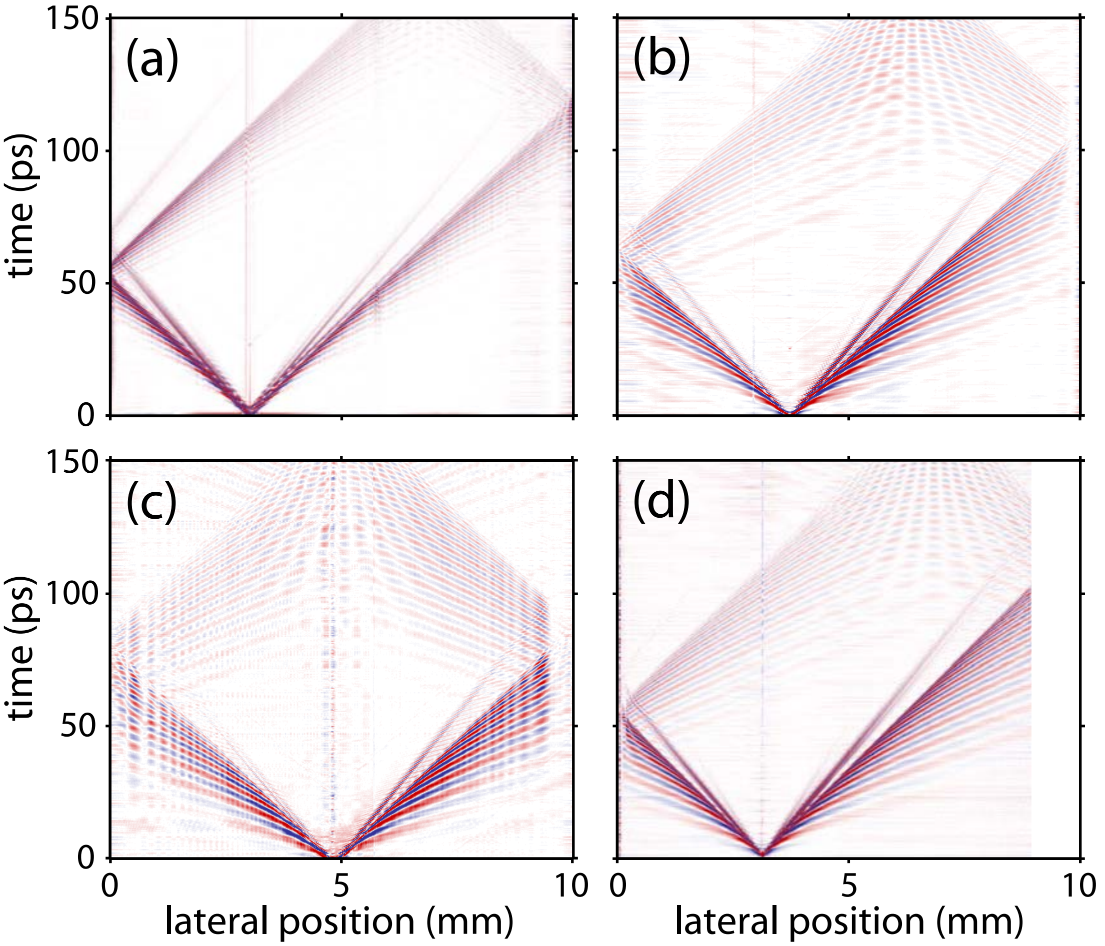

Figure 4.10 | Space-time plot. Plots from each imaging method showing a broadband THz wave propagating in a 53 um thick, 1 cm wide, unstructured LN slab. The color scale indicates the relative value of the electric field. (a) Talbot imaging. (b) Phase contrast imaging. (c) Sagnac imaging. (d) Polarization gating imaging.

We can analyze the wave vector and frequency response by taking the absolute value of the two-dimensional Fourier transform of the matrices shown in Fig. 4.10. The time axis transforms as frequency and the spatial axis, which is along the direction of waveguide mode propagation, transforms as k_x, which is often referred to as the waveguide propagation constant, β. The results of this transform, shown in Fig. 4.11 for the polarization gating data, are the waveguide dispersion curves. The three visible lines correspond to the first three waveguide modes, and their locations and shapes are in excellent quantitative agreement with theoretical predictions [Wu 2009].

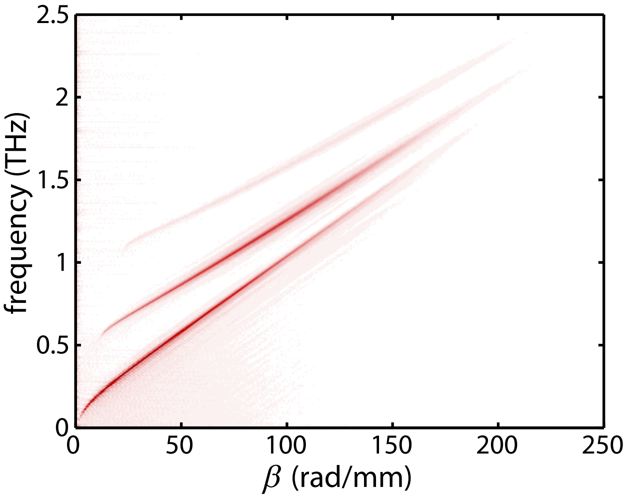

Figure 4.11 | Experimental dispersion curve. Taking the 2-dimensional Fourier transforms of the space-time plots in Fig. 4.10 (position transforms to wave vector and time transforms to frequency) yields the waveguide dispersion curves. Here the transform of the polarization gating data from Fig. 4.10(d) is shown. The dispersion curves for the first three waveguide modes are clearly visible.

Figure 4.12 shows the average wave vector response for each method calculated by averaging over the frequency axis. The percent signal modulation, $P = \Delta I / I_0$, depends on $\Delta\phi$, the actual wave vector content of the induced phase shift immediately after the sample, and *R*, the instrument response: $P(k_x) \propto \Delta\phi(k_x) R(k_x)$. Although the four methods have some variation in the induced phase shift because the details of pumping may have changed slightly when optics were rearranged during the probe reconfiguration, relative differences between the traces in Fig. 4.12 are due primarily to a varying instrument response. The modulation in the Talbot response (dashed line), in particular the weakened response at low wave vectors, ~60 rad/mm, and ~120 rad/mm, was predicted in section 3a above. This non-uniform, wave vector-dependent response, coupled with out of focus imaging of structures, makes Talbot imaging a poor method. Note that both phase contrast (solid line) and polarization gating (dotted line) have smooth and un-modulated wave vector responses and large detection bandwidths. In this case the bandwidth is limited by THz generation and not the detection technique.

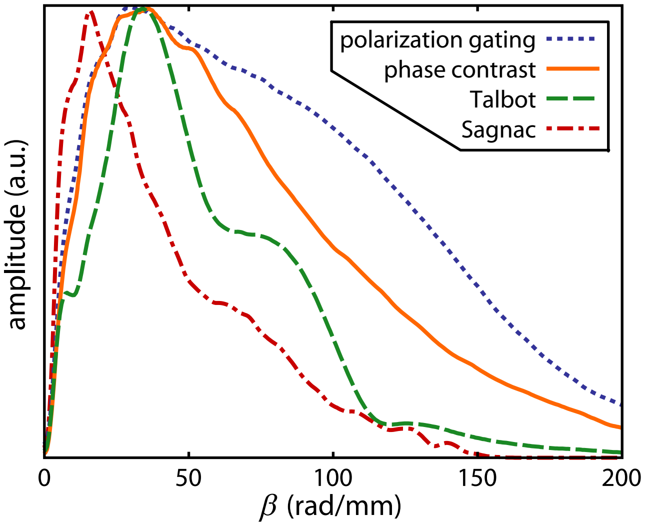

Figure 4.12 | Measured wave vector content. The measured wave vector bandwidth of the images, calculated by taking the dispersion curves like the one shown in Fig. 4.11 and integrating over the frequency axis. The amplitude is proportional to the actual wave vector content multiplied by the instrument response.

Both the analysis and experimental results indicate that the phase contrast and polarization gating techniques show the most promise, so a more careful comparison is required. For phase contrast imaging, $\Delta\phi$ represents the THz-induced phase shift in vertically polarized light. The phase shift is:

$$\Delta\phi_{\text{PC}}(x,z) = -2\pi\frac{\ell}{\lambda_{\text{opt}}}\Delta n_{eo}(x,z) = -2\pi\frac{\ell}{\lambda_{\text{opt}}}\frac{r_{33}n_{eo}^3}{2}E_{\text{THz}}(x,z), \tag{4.12}$$

where ℓ is the slab thickness, λ_{opt} is the free space optical wavelength, Δn_{eo} is the THz-field induced change in index for the extraordinary axis, and r_{33} is the appropriate electro-optic coefficient. Equation 4.12 assumes that the THz field is uniform along y through the depth of the crystal, which is not strictly correct. See Section 6 below for a rigorous derivation of the THz-induced phase shift as a function of mode number and frequency. For polarization gating, $\Delta\phi$ represents the difference in THz-induced phase shift between vertically (along the extraordinary axis) and horizontally (along the ordinary axis) polarized light. It is given by:

$$\Delta\phi_{\text{PG}}(x,z) = -2\pi\frac{\ell}{\lambda_{\text{opt}}}\left[\Delta n_{eo}(x,z) - \Delta n_o(x,z)\right] = -2\pi\frac{\ell}{\lambda_{\text{opt}}}\frac{r_{33}n_{eo}^3 - r_{13}n_o^3}{2}E_{\text{THz}}(x,z). \tag{4.13}$$

Inserting Eqs. 4.11 & 4.12 into the expressions for percent signal modulation (Eqs. 4.6 & 4.11), we have:

$$\frac{P_{\text{PC}}(x,z)}{P_{\text{PG}}(x,z)} = \frac{2\Delta\phi_{\text{PC}}(x,z)}{2\Delta\phi_{\text{PG}}(x,z)} = \frac{r_{33}n_{eo}^3}{r_{33}n_{eo}^3 - r_{13}n_o^3} \approx 1.6, \tag{4.14}$$

so the signal modulation is larger for phase contrast imaging. The significant reduction in noise resulting from balanced imaging, however, more than compensates for the slightly weaker signal. Movies captured using the polarization gating method look qualitatively better because low wave vector noise resulting from pointing instability in the laser is almost completely eliminated. In addition, the self-compensating geometry makes quantitative measurement of the phase shift more reliable because the relative phase between vertical and horizontal polarizations is uniform and carefully controlled. The phase difference between the diffracted orders and the 0th order beam is less well controlled in phase contrast imaging because of inhomogeneities in sample thickness. In addition, light diffracted at low wave vectors off relatively large static structures may not be sufficiently separated from 0th order light to avoid the sharp edges of the depression in the phase plate, further modulating the background. These effects are visible in Fig. 4.13(a), a raw image from the phase contrast setup which shows significant variation in background intensity. Figure 4.13(b) shows a raw image from the polarization gating setup, where the two bright regions correspond to the two polarizations emerging from the Wollaston prism. The intensity in Fig. 4.13(b) is much more uniform due to the self-compensation. The more reliable quantification of electric field and superior signal-to-noise ratio makes polarization gating the preferred method in most situations.

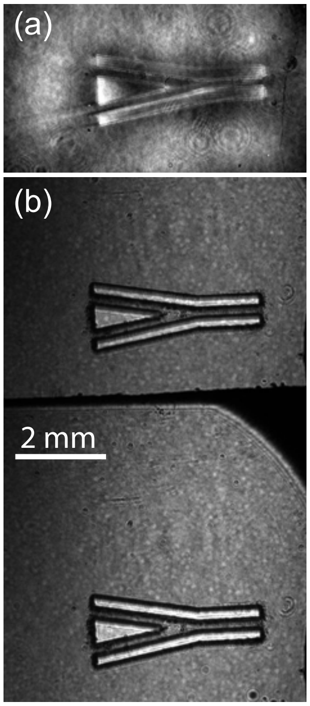

Figure 4.13 | Raw images. The unprocessed images of a Y-coupler structure imaged through (a) the phase contrast setup and (b) the polarization gating setup. Modulation in the background intensity of the phase contrast image is apparent, while it is smooth and unmodulated for the polarization gating image. The two images in (b) result from the two polarizations that have been imaged onto the same CCD chip.

Polarization gating is not the best method, however, when high spatial resolution is required. The retro-reflecting geometry necessitates that there be a large number of optics between the sample and first lens that images the sample onto the camera. In practice, we found that the shortest focal length that could be conveniently used was 20 cm, which limits the numerical aperture (NA), and thus the resolution, of the imaging system. In addition, the sample must be very accurately imaged back onto itself [see Fig. 4.9(f)] to obtain the NA-limited resolution. In practice the resolution is on the order of 10 μm, which is significantly sub-wavelength for 1 THz waves ($\lambda \sim 60$ μm in bulk LN) and so not an issue in most situations. However, if measurement of much higher wave vector waves or near-field effects around the edges of microstructures is desired, then the limited resolution could be an issue. The phase contrast geometry is well suited for high-NA imaging and is preferred in these situations.

5. Conclusions

Because structured LN slabs provide a versatile platform for THz generation and control, it is important to have versatile detection techniques to study these systems. Phase sensitive imaging is of significant interest because it enables the recording of the full spatiotemporal evolution of the fields. If quantitative measurement of THz *E*-field strengths is not necessary and out-of-focus imaging of structures can be tolerated, then Talbot imaging is preferred because of its simplicity and ease of alignment. For quantitative THz *E*-field measurement and in-focus imaging, the phase contrast or polarization gating methods are preferred. Polarization gating provides the most reliable *E*-field measurements because the self-compensating geometry corrects for sample imperfections and variations in thickness. It also provides the best signal to noise ratio because balanced imaging can be implemented. These advantages make polarization gating the method of choice for most situations. Phase contrast imaging, however, lends itself to high numerical aperture imaging and so is preferred when high spatial resolution is required. Thus Talbot, phase contrast, and polarization gating form a complementary set of techniques for imaging THz waves.

6. Appendix: Quantitative *E*-field measurement

The experimental setups described above measure the induced phase shift in the optical probe, but generally the amplitude of the THz electric field is desired. If the applied field is DC, or perfectly phase-matched detection is used, the *E*-field can be determined from the simple electro-optic effect given by Eqs. 4.12 and 4.13 above. However, in all of our setups we detect AC THz waves propagating perpendicular to the direction of probe propagation, a geometry which is necessarily *not* phase-matched. This results in a detection sensitivity that depends on both the waveguide mode and frequency. By understanding frequency-dependent sensitivity, it is possible to correct for the phase matching error. As a *z*-polarized probe pulse travels a distance *dy* along the *y*-direction [see Fig. 4.9(b) for axes], it accumulates a THz-induced phase shift, $\Delta\phi(x,z)$ through the electro-optic effect:

$$d\Delta\phi(x,z) = \frac{\omega_{\text{opt}}}{c}\Delta n(x,y,z)dy = \frac{\omega_{\text{opt}}}{c}\frac{n_{eo}^3 r_{33}}{2}E_{\text{THz}}(x,y,z)dy \tag{4.15}$$

Here E_{THz} is the THz electric field, ω_{opt} is the probe frequency, c is the speed of light in vacuum, n_{eo} is the extraordinary index of refraction for the probe, and r_{33} is the appropriate electro-optic coefficient. We can approximate the optical probe pulse by a δ-function in space and time when the optical pulse duration is significantly shorter than the THz temporal oscillation period and the diffraction-limited spot that can be resolved by the imaging system is significantly smaller than the THz wavelength. This lets us calculate the induced phase shift as a function of waveguide mode, m, and frequency, ω_{THz}, by inserting the theoretical description for the waveguide mode field profile into Eq. 4.15 and integrating. During the integration we must account for the time dependent THz phase: $\psi = \omega_{\mathrm{THz}} t + \psi_0 = \omega_{\mathrm{THz}} n_{\mathrm{opt}} y / c + \psi_0$, which changes as the optical probe propagates through the waveguide. Putting this all together yields:

$$\Delta\phi(\omega_{\mathrm{THz}}, m) = \frac{\omega_{\mathrm{opt}}}{c} \frac{n_{eo}^3 r_{33}}{2} \int_{-\ell/2}^{\ell/2} E_T(y, \omega_{\mathrm{THz}}, m) \cos\left(\frac{\omega_{\mathrm{THz}} n_{eo}}{c} y + \psi_0 \right) dy \tag{4.16}$$

with ℓ the slab thickness. The transverse field profile of the THz wave in the lithium niobate given by the well known eigenstates of the dielectric slab waveguide [Cronin 1995]:

$$E_T(y, \omega_{\mathrm{THz}}, m) = E_0 \cos\left[k_y(\omega_{\mathrm{THz}}, m) y + m\pi/2 \right] \tag{4.17}$$

Here the transverse wave vector, k_y, depends on the mode number and the THz frequency. Even modes are cosine functions and odd modes are sine functions, with $y = 0$ defined as the center of the crystal. The sensitivities for the first three waveguide modes are plotted in Fig. 4.14, with the sensitivity defined as the phase shift induced by the THz wave divided by the phase shift that would be induced by a DC field of amplitude E_0. The system is most sensitive to the 0th order mode for frequencies below ~1.5 THz because it is the most uniform along y, and generally performs better for lower frequencies since these THz waves travel only a small fraction of their wavelength during the transit time of the optical probe pulse through the LN slab. For situations where the waveguide mode and frequency are known, as is often the case for a narrowband THz wave [Lin 2009], the E-field can be quantitatively measured using Eq. (4.12) or (4.13) and dividing by the sensitivity to correct for the imperfect phase matching.

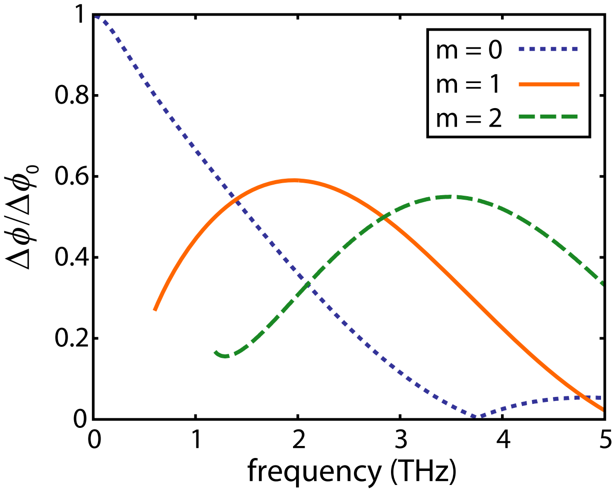

Figure 4.14 | Detection sensitivity. The mode- and frequency-dependent sensitivity for a 53 µm LN slab. Sensitivity is defined as the actual phase shift induced in the probe as it propagates through the LN slab normalized by the phase shift that would be expected for a DC electric field.

D. ADVANCES IN THz IMAGING METHODS

1. Introduction

Section 4C and the paper it was drawn from [Werley 2010] made a careful comparison of phase sensitive imaging techniques for visualizing THz waves in a $LiNbO_3$ slab. The primary conclusion was that phase contrast is the best choice for high-resolution imaging, while self-compensating polarization gating offers the best signal-to-noise ratio and most quantitative extraction of THz electric field strengths. Since the publication of [Werley 2010], there have been a number of advances in the optical design of both phase-contrast and polarization gating imaging, primarily targeted at increasing resolution. In addition, there have been advances in the cameras and hardware to improve the signal to noise. My aims in this section are to show the different iterations for both techniques, and to use that to explain the design decisions and hopefully point out some pitfalls to avoid in future work.

2. Phase contrast imaging

There are two primary decisions when designing the optical setup for phase contrast imaging. The first is how to combine the pump and probe in the sample and subsequently re-separate them without damaging any of the optical components. For example, the pump will likely destroy any objective lens it passes through. The second major consideration is what lens to use after the sample (this lens sets the numerical aperture (NA)), and the design of the phase mask paired with this lens. Figure 4.15 shows the evolution of the imaging setup.

In Fig. 4.15(a), the phase mask was made by depositing a λ/4-thick layer of poly(methyl methacrylate) (PMMA) photoresist onto a flat fused silica substrate. The photoresist in the center was then exposed and removed, resulting in a 30 x 30 μm recessed square in the center of the substrate. The size of the square was designed so that the main beam would focus through the depression without clipping the edges, while little diffracted light would pass through the depression. Gold arrows deposited on the phase mask facilitated alignment. In this design, the pump light was separated from the probe light using a dichroic mirror, but it turns out that a tilted optic in any imaging system will introduce significant aberration and skew the image. A simple solution to this is shown in Fig. 4.15(b), where the dichroic mirror is replaced by an absorptive filter.

The design in Fig. 4.15(b) was a huge step forward over Talbot imaging in terms of image resolution. Talbot imaging is an out-of-focus technique and the resolution is roughly 50μm. It is useful in some situations because the resolution is finer than the wavelength of lower-frequency waves, but the Talbot method is obviously somewhat limited. The iteration of phase contrast in Fig. 4.15(b) increased the resolution to ~ 5 μm, but there were still some challenges and limitations. First, the absorptive filter and simple achromatic lens introduced some spherical aberration, so the image was not completely diffraction limited. Second, the NA of the imaging system was fairly modest. In addition to these standard imaging considerations, the phase mask had some problems. The PMMA was delicate and was easily damaged by the tightly focused probe. In addition, the gold arrows deposited on either side of the recessed region to aid in alignment [Wu 2009] blocked some of the diffracted light. Finally, the pump generated significant second harmonic in the $LiNbO_3$ slab, which could not be removed with a color filter because it was the same frequency as the probe light. As a result, the pump had to be sent into the sample at an angle, so the generated second harmonic would be spatially separated and could be blocked by a razor blade or other thin metal element. Blocking the pump with a thin element has several drawbacks, including the fact that it is difficult to remove the pump light entirely and the element obscures some diffracted information in the probe.

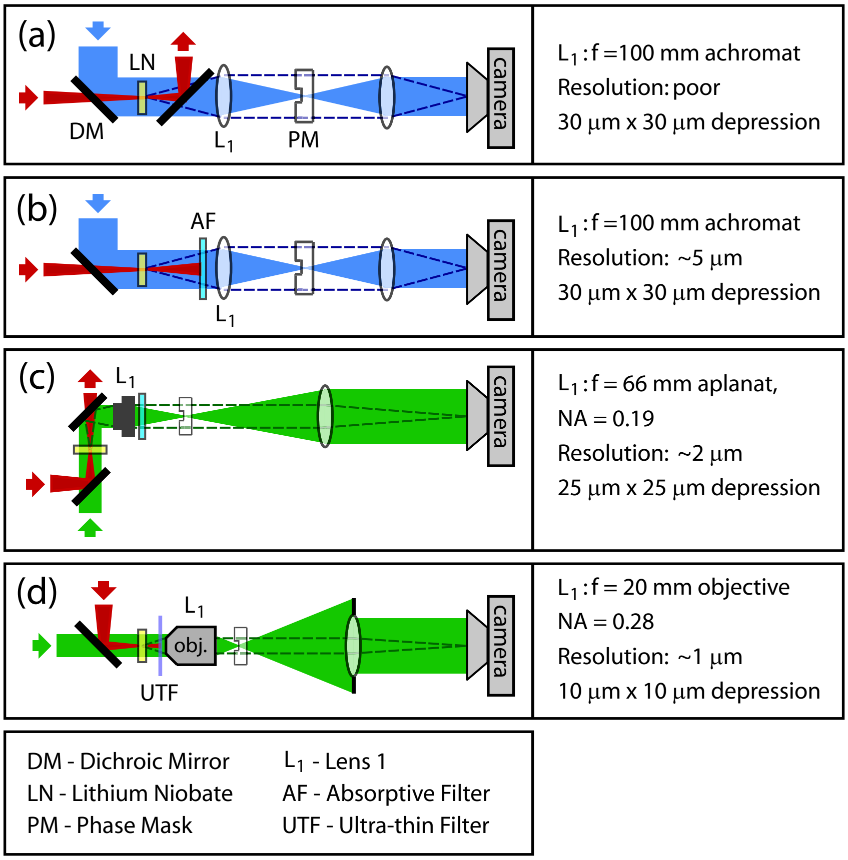

Figure 4.15 | Evolution of phase-contrast design. This figure shows the optical design to the left and indicates key design characteristics to the right. Abbreviations are defined at the bottom. In the first two designs, probing is performed with 400 nm light, and SHG from the pump in the $LiNbO_3$ must be blocked spatially. In (c) and (d), probing is performed with 532 nm light from a NOPA, so pump and SHG can be blocked with a spectral filter. (a) The initial design. The pump light was reflected off a dichroic mirror after the sample. The diffracted light was diverging when it passed through this angled optic and thus degraded the image. (b) The dichroic was replaced by an absorptive filter at 0°, enabling imaging with moderate resolution. (c) This design was used in the antenna studies and featured a higher NA and reduced spherical aberration. (d) A proposed design, tested for regular imaging in the antenna work, which pushes the NA higher and further improves resolution.

The design in Fig. 4.15(c) addressed many of the problems with the design in 4.15(b). The primary filter was a reflective dichroic mirror, which is aberration free, and the first lens was an aplanat which largely removed spherical aberration (CVI part #: LAP & APM-100-20, WD =

45 mm, NA = 0.19, *f* = 66 mm). Thus the system yielded a nearly diffraction limited image with a resolution of about 2 μm [λ/(2NA) = 1.5 μm]. Furthermore, instead of depositing photoresist on the fused silica, the glass was etched away to a quarter-wave depth in a 25 x 25 μm square, significantly increasing the damage threshold. The gold arrows for alignment were also moved 1.5 cm away from the recessed region, far enough off center that they did not block any diffracted light. Finally, the probe light was switched from 400 nm (SHG of the 800 nm light) to 532 nm light from the output of a non-collinear optical parametric amplifier (NOPA). Because the NOPA output was well separated in wavelength from both the 800 nm pump and the pump-generated 400 nm light, it was possible to reject all pump light with a single dielectric filter (pass ~ 512 nm to 552 nm). 532 nm light was chosen in particular because of the availability of diverse, high-quality optics at this wavelength. The spectral filter let us remove all spatial blocks from the imaging system so the diffracted light could pass unimpeded. This design was used to collect very beautiful movies of the antenna (Chapter 6, [Werley 2012]).

Although the design in Fig 4.15(c) is quite flexible and solved most of the problems with phase contrast imaging, there still remain situations where an even higher resolution would be desirable. When studying the antennas, for example, the imaging resolution was already 100 times finer than the THz wavelength, but there was structure in the near-fields that we could not resolve. To improve the resolution of the imaging system further, a higher NA objective is required [see Fig. 4.15(d)]. Not any objective can be used, however, because the pump light must still be removed upstream of the objective. I propose using an experimental geometry we developed for aligning the probe spot into the antenna gap in point-source detection. This method uses a long working distance objective and rejects pump light with an ultra-thin reflective filter (CVI part #: SWP-0-R800-T400-ET-25.4-0.30-UV). The filter is thin enough that it introduces minimal aberration. I suggest Mitutoyo's long working distance objectives; either the 10x (CVI part #: 04OVI010, working distance (WD) = 33mm, NA = 0.28, *f* = 20 mm) or the 20x (CVI part #: 04OVI020, WD = 20 mm, NA = 0.42, *f* = 10 mm) seem like good candidates. The 10x gives 1 μm resolution, and has already been tested with the ultra-thin filter. I believe the filter would also be able to protect the 20x objective without being damaged, and this would offer 700 nm resolution, enabling λ/300 resolution for lower THz frequencies. While even finer resolution may be desirable, objectives with higher NA's have shorter working distances and will be difficult to protect from the high pump fluence. Replacing the current lens with a higher NA objective would require a new phase mask to be incorporated as well. The recessed region of the new phase mask would have reduced dimensions (the size of the square should scale with the focal length of the lens). It would also be beneficial to use a thinner substrate for the phase mask because it seemed the 3/8-inch substrate in the current design [Fig. 4.15(c)] introduced significant aberration.

3. Polarization gating imaging

Phase contrast imaging offers the most promise for high-resolution imaging, which is critical for characterizing the deeply sub-wavelength near-field structure of metal components. For measurement of purely dielectric structures such as photonic crystals, however, a more modest resolution is acceptable. For these measurements, polarization gating imaging provides the most accurate quantification of THz field strengths and greatly improved noise suppression via balanced imaging (section 4C and [Werley 2010]). In polarization gating, 45°-polarized probe light passes through the sample and the THz is detected via the induced change in sample birefringence. However, working with 45°-polarized light is challenging because each mirror has a different response for *s*- and *p*-polarization states and will depolarize the beam. Another challenge is that the sample itself is strongly birefringent even in the absence of a THz field. When strategies to control the polarization state of the light are implemented, careful optical design is required to maintain good image resolution.

Figure 4.16(a) shows the first generation design for polarization gating imaging [Peier 2008]. When the probe passes through the highly anisotropic $LiNbO_3$ slab, the vertically polarized component develops a large ($>>2\pi$), wavelength dependent phase shift relative to the horizontally polarized component. To compensate for this static birefringence, a second LN slab with its optic axis rotated by 90° was placed in the imaging path, which should in principle correct for the induced phase shift. However, we found in practice that slight inhomogeneities in sample thickness could not be compensated and led to large spatial variations in the polarization state of the beam. This prevented quantitative measurement of the THz field strength and gave the system a spatially varying sensitivity to THz fields. In addition, it was difficult to separate the 400 nm probe from the pump second harmonic generated in the $LiNbO_3$ slab, much like in the early generations of phase contrast imaging.

The solution to these problems is shown in Fig. 4.16(b). To address the difficulty of separating the probe light from the pump light, we probed with the 532 nm output of a NOPA and used a single dielectric filter to block all pump light. The next advance was self-compensation. By imaging the sample back onto itself, after flipping the vertical and horizontal polarization components, it was possible to have a nearly pure polarization state after the second pass through the sample (see 4C and [Werley 2010]). This also enabled balancing, which suppressed the noise by a factor of ~5 (see Fig. 4.6 and [Werley 2011]) and was critical for removing the low wave vector background flicker that appears in old polaritonics movies. The downside of this design, however, was its resolution. The sample was imaged back onto itself, and any imperfections in alignment hurt the resolution. In addition, the main probe beam and the diverging diffracted light were transmitted through the non-polarizing beam splitter, which introduced significant aberration. Perhaps the worst offender was the Wollaston prism. The different polarization states refract in different directions, and that introduces significant smearing in the image. The technique is "in-focus", so the resolution is better than for Talbot imaging, but the resolution is only 15 – 20 μm. For many measurements this is sufficiently fine because it is still significantly smaller than the THz wavelength. However, this can be limiting when dealing with structures, even those without plasmonic near fields. In a photonic crystal, for example, the bandgap occurs when the hole diameter is roughly 1/4 of the THz wavelength ($\lambda/4$), and fields can vary on even shorter length scales. Since a 1-THz phonon-polariton has a wavelength of 60 μm in bulk $LiNbO_3$, $\lambda/4$ corresponds to 15 μm. It is desirable to use polarization gating because of its quantitative field measurement and low noise, but the resolution in Fig. 4.16(b) is clearly limiting.

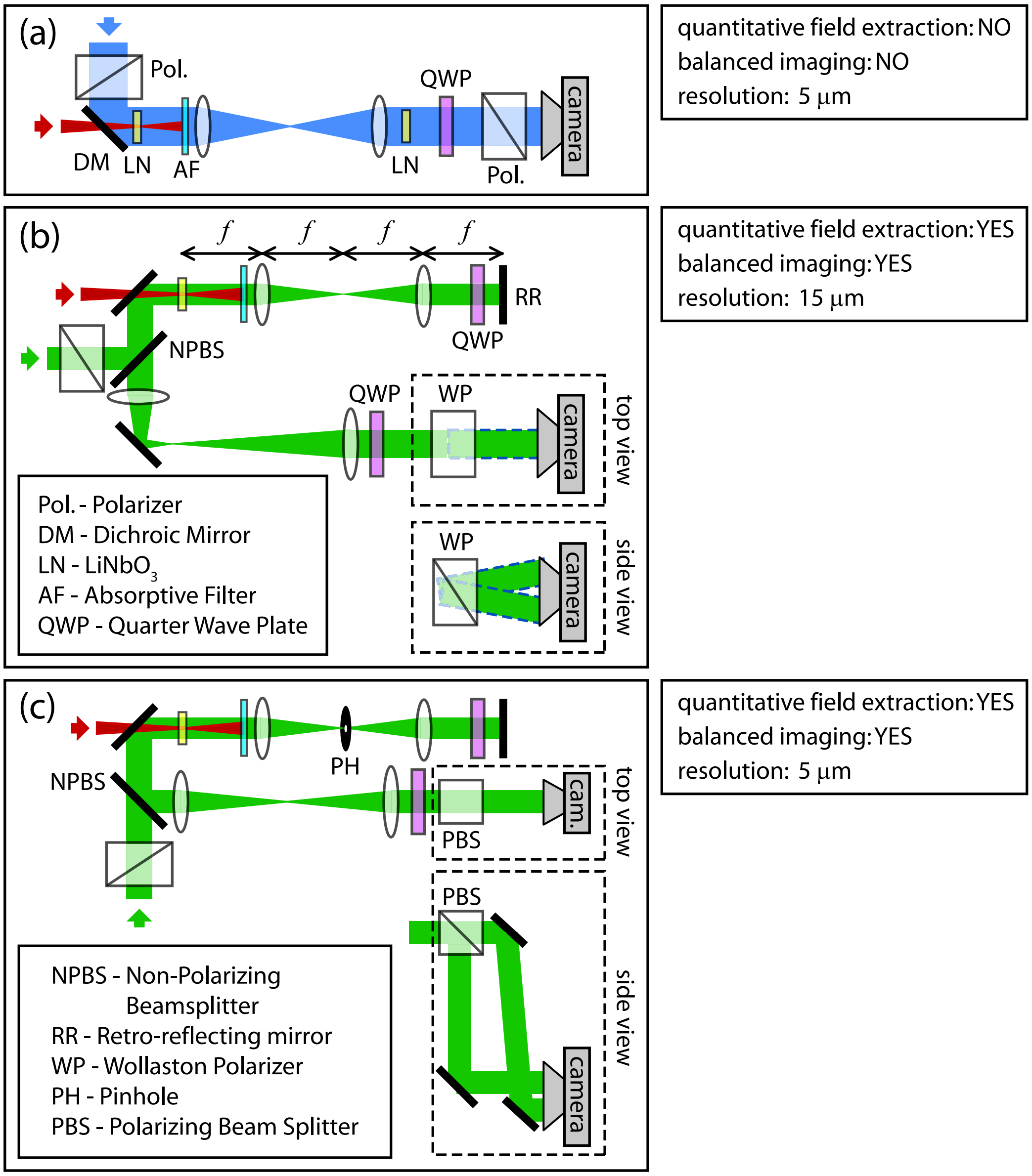

Figure 4.16 | Evolution of polarization gating optical design. Abbreviations are defined in the insets and performance characteristics are indicated to the right. In the first design (a), probing was performed with 400 nm light, and SHG from the pump in the $LiNbO_3$ had to be blocked spatially. In (b) and (c), probing is performed with 532 nm light from a NOPA, so pump and SHG can be blocked with a spectral filter. (a) In this initial design, a second slab of $LiNbO_3$ with its optic axis rotated by 90° was placed in the beam to compensate for the static birefringence of the first slab. The strategy proved ineffective due to inhomogeneities in slab thickness. (b) The sample slab is imaged back onto itself (with its polarization rotated by the QWP), self-compensating its static birefringence. (c) A design with improved image resolution. The image is reflected off the non-polarizing beamsplitter to minimize aberration, and a polarizing beam splitter replaces the Wollaston prism, which smeared the image when it refracted the two polarization states in opposite directions.

To address the resolution limitations, we redesigned the optics as shown in Fig. 4.16(c). The simple first correction was to reflect the image off the non-polarizing beamsplitter instead of transmitting through it, eliminating it as source of aberration. Next, we placed a pinhole in the Fourier plane (the focal plane) of the retro-reflecting imaging system. The pinhole size must be large enough to transmit the information about thickness variation in the $LiNbO_3$ slab for self-compensation, which tends to be slowly varying, but small enough to block the high wave vector information diffracted off sample structures. Although this is inexact, we found in practice that it does indeed improve resolution because two images (one from each pass through the sample) need not be overlaid on top of each other. Perhaps with very careful imaging in the retro-reflecting setup using low-aberration, high NA lenses, the pinhole would have no effect on final image resolution. The final step to improve image resolution was to replace the Wollaston prism with a polarizing beamsplitter. Although the optical setup is more inconvenient, the design in Fig. 4.16(c) lets us see the whole width of the $LiNbO_3$ slab in the two balanced images on top and bottom of the CCD chip. For both images to be in-focus, the path lengths of both beams must be the same, and this can be tuned by changing the angle of the beam that was transmitted through the polarizing beamsplitter. The advantage of the polarizing beamsplitter is that it does not refract the beams as the Wollaston does, and therefore introduces minimal aberration into the imaging system. The above modifications improved the resolution to 5 μm, sufficiently fine for measurements that don't have metal structures with strongly localized near fields.

In fact, for photonic crystals and other machined microstructures, we want the depth of focus to be equal to the slab thickness. The depth of focus, also known as the confocal parameter b, is related to the numerical aperture (NA) by $b = \pi\lambda/[8(NA)^2]$. The resolution is given by $d = \lambda/(2NA)$. For a 50 μm thick slab and 532 nm probe, the optimal NA is 0.06, which yields a resolution of 4 μm. For small divergence angles (long focal length lenses) $NA = D/(2f)$, so the lens aperture should be $D \sim 0.12f$. In principle, higher NA lenses (e.g. CVI part #: LAP and APM-300.0-60.0, WD = 163 mm, NA = 0.19) can be used to further improve the resolution for studying metallic structures on the surface, but this will be difficult in practice, so phase contrast imaging is preferred for high-resolution imaging. Another possible improvement to the polarization gating setup could come from using a Wollaston with a very low separation angle. This would simplify the alignment considerably over the polarizing beamsplitter and may introduce sufficiently low smearing to make it a better option in many situations.

4. Cameras and hardware

The previous two sections tracked the evolution of the optical design, which generally focuses on increasing image quality and resolution, filtering of pump light, and improved quantification of the THz-induced phase shifts in the optical probe. In parallel with the evolution of optical design, the electronics, hardware, and referencing procedure have also improved. In this case, the primary focus has been improving the signal to noise ratio (SNR) and reducing acquisition time. To measure the THz field, it is always necessary to measure two images: a signal image when the pump is present and a reference image when the pump is blocked. The signal image is divided by the reference image to extract $\Delta I/I_0$, and this is used to measure the THz field strength.

To design the optimal detection setup we need to minimize total noise, which comes both from the laser and the detection electronics. Although electronic noise can be significant for high-noise detectors such as a CMOS camera, laser noise usually dominates ultrafast measurements (see Ch. 4 Sec. B and [Werley 2011]). It is of paramount importance to suppress this laser noise, which requires understanding its characteristics. The largest component results from long-timescale fluctuations of laser power and pointing, the laser drift. The magnitude of this drift is inversely proportional to the frequency, and so is called $1/f$ noise, a type of noise common in many physical systems [Bak 1987]. It can be clearly observed at low frequencies in our laser until it drops below the flat, white-noise floor at ~50 Hz [see Fig. 4.4(e)]. To combat $1/f$ noise, it is beneficial to record signal and reference images as close together in time as possible.

Table 4.1 shows the different methods for switching between signal and reference images, and indicates the time between them. In the first method (first row), the reference image was recorded once at the beginning of data acquisition. Then the delay stage was slowly scanned and a signal image was recorded at each time delay, typically taking ~30 minutes to record an entire trace. During this 30-minute interval, the laser would inevitably drift significantly so that the signal and reference traces were quite different, leading to bad noise in the THz movie. The solution to this problem was to collect both signal and reference at each time delay. The first implementation of this (second row) was to place a computer-controlled shutter in the pump. At each time delay the computer would close the shutter, capture a reference image, open the shutter, and capture a signal image. This was very helpful for suppressing the effects of slow laser drift, but there was an unanticipated problem. When the pump shutter was opened, the sample would heat and expand, and our phase-sensitive imaging methods would detect this change. The movies showed the phase shift introduced both by the THz wave and by the temperature change.

To minimize the effects of sample heating and cooling we used choppers instead of a shutter (third row). Both pump and probe were chopped at 500 Hz with choppers synched to the laser. The signal image was collected when the choppers were in-phase: pump and probe pulses derived from the same laser pulse would pass through the choppers. When the choppers were out-of-phase, pump and probe pulses from alternating laser shots would pass through the choppers and the reference was detected. In the out-of-phase configuration, the probe pulse arrived ~ 1 ms after the pump, so the THz was long gone and a clean reference could be recorded. The pump laser hit the sample at 500 Hz when collected both signal and reference images, so the thermal load was constant and a temperature difference was not detected. This acquisition method was slightly slower than using shutters because changing the chopper from in-phase to out-of-phase took approximately 1 second, but it was still much better than acquiring the reference only once at the beginning of the scan.

Pump modulation method	Detector Type	Time between signal and reference images
Take single reference image, record signal images at each time delay in the entire trace.	Slow, low noise CCD	30 min.
Place a shutter in pump, close for signal, open for reference at each time delay.	Slow, low noise CCD	1 sec.
Chop both pump and probe at 500 Hz. Choppers in phase for signal and out of phase for reference. Change phase at each time delay.	Slow, low noise CCD	2 sec.
Chop pump at 500 Hz, detect every probe pulse on a camera synched to the laser.	fast, high noise CMOS	1 msec.

Table 4.1 | Evolution of imaging hardware. The table gives a brief description of how signal and reference images were collected in the first column. The detector is indicated in the second column; the CCD was a Pixis camera from Princeton Instruments and the CMOS camera was a Fastcam 1024PCI from Photron. The last column shows the time between signal and reference images. Shorter times more effectively suppress laser noise.

Although using the choppers to capture signal and reference images suppressed some of the $1/f$ noise, it was still significantly slower than the ideal. It would be preferable to reduce the time between signal and reference acquisition even further. A very common noise suppression strategy is to modulate the pump and detect the signal at the modulation frequency, as is done for a lock-in amplifier or for our DAQ card detection system (Ch. 4 Sec. B, [Werley 2011]). Generally the modulation frequency should be as high as possible, and it should be chosen to avoid systematic noise, such as the AC line frequency at 60 Hz and its harmonics. For a 1 kHz laser, the optimal choice is to chop at 500 Hz (every other pump pulse blocked). To implement this, we purchased a fast CMOS camera from Photron that could be synched to the laser, thus acquiring an image for every laser pulse. We chopped the pump at 500 Hz, so that every even numbered probe pulse (when the pump was unblocked) detected the signal image, and every odd numbered probe pulse detected the reference. There are still some challenges associated with the fast camera, including difficulties in rapidly processing the massive amounts of raw data (data rate = 2 Gbyte/sec) and a CMOS detector that is noisier than its slower CCD counterpart. In spite of these challenges, the CMOS detector is greatly preferred because of its high detection frequency and excellent laser noise suppression. When combined with balancing as described in Fig. 4.16, movies of THz waves can be detected with sensitivity approaching that of point probe lock-in detection. The exceptional sensitivity and improved resolution from these advances will enable a broad set of future experiments.

E. REFERENCES

[Auston 1988] D. H. Auston and M. C. Nuss, “Electrooptic generation and detection of femtosecond electrical transients.” *IEEE J. Quantum Electron.* **24**, 184-197 (1988).

[Bak 1987] P. Bak, C. Tang, and K. Wiesenfeld, “Self-organized criticality: an explanation of 1/*f* noise,” *Phys. Rev. Lett.***59**, 381-384 (1987).

[Brorson 1987] S. D. Brorson, J. G. Fujimoto, and E. P. Ippen, “Femtosecond electronic heat-transport dynamics in thin gold films,” *Phys. Rev. Lett.* **59**, 1962-1965 (1987).

[Crimmins 2002] T. F. Crimmins, N. S. Stoyanov, and K. A. Nelson, “Heterodyned impulsive stimulated Raman scattering of phonon–polaritons in $LiTaO_3$ and $LiNbO_3$,” *J. Chem. Phys.* **117,** 2882-2896 (2002).

[Cronin 1995] N. J. Cronin, *Microwave and Optical Waveguides* (Institute of Physics Publisher, 1995).

[Danielson 2006] J. R. Danielson, N. Amer, and Y. S. Lee, “Generation of arbitrary terahertz wave forms in fanned-out periodically poled lithium niobate,” *Appl. Phys. Lett.* **89,** 211118 (2006).

[Dougherty 1992] T. P. Dougherty, G. P. Wiederrecht, and K. A. Nelson, “Impulsive stimulated Raman scattering experiments in the polariton regime,” *J. Opt. Soc. Am.* **9**, 2179-2189 (1992).

[Feurer 2003] T. Feurer, J. C. Vaughan, and K. A. Nelson, “Spatiotemporal coherent control of lattice vibrational waves,” *Science* **299**, 374-377 (2003).

[Feurer 2007] T. Feurer, N. S. Stoyanov, D. W. Ward, J. C. Vaughan, E. R. Statz, and K. A. Nelson, “Terahertz polaritonics,” *Annu. Rev. Mater. Res.* **37**, 317-350 (2007).

[Gabriel 1991] M. C. Gabriel, N. A. Whitaker, Jr., C. W Dirk, M. G. Kuzyk, M. Thakur “Measurement of ultrafast optical nonlinearities using a modified Sagnac interferometer,” *Opt. Lett.* **16**, 1334-1336 (1991).

[Goodman 2005] J. W. Goodman, *Introduction to Fourier Optics, 3rd edition* (Roberts & Company Publishers, 2005), Chap. 8.

[Goodno 1998] G. D. Goodno, G. Dadusc, and R. J. D. Miller, “Ultrafast heterodyne-detected transient-grating spectroscopy using diffractive optics,” *J. Opt. Soc. Am. B* **15**, 1791-1794 (1998).

[Goshtasby 2005] A. A. Goshtasby, *2-D and 3-D Image Registration* (John Wiley & Sons, 2005), Chap. 5.

[Greene 1982] B. I. Greene, R. C. Farrow, “Direct measurement of a subpicosecond birefringent response in CS_2,” *J. Chem. Phys.* **77**, 4779-4780 (1982).

[Hebling 2002] J. Hebling, G. Almasi, I. Z. Kozma, and J. Kuhl, “Velocity matching by pulse front tilting for large area THz-pulse generation,” *Opt. Express* **10,** 1161-1166 (2002).

[Hebling 2004] J. Hebling, A. G. Stepanov, G. Almaasi, B. Bartal, and J. Kuhl, “Tunable THz pulse generation by optical rectification of ultrashort laser pulses with tilted pulse fronts,” *Appl. Phys. B* **78,** 593-599 (2004).

[Hoffmann 2007] M. C. Hoffmann, K. L. Yeh, J. Hebling, and K. A. Nelson, “Efficient terahertz generation by optical rectification at 1035 nm,” *Opt. Express* **15,** 11706-11713 (2007).

[Hornung 2007] T. Hornung, K. L. Yeh, K. A. Nelson, “Terahertz nonlinear response in lithium niobate,” In Ultrafast Phenomena XV, ed. P. Corkum, D. M. Jonas, R. J. D. Miller, A. M. Weiner. Berlin: Springer, (2007).

[Ippen 1975] E. P. Ippen and C. V. Shank, "Dynamic spectroscopy and subpicosecond pulse compression," *Appl. Phys. Lett.* **27**, 488-490 (1975).

[Jiang 1999] Z. Jiang and X.-C. Zhang, "Terahertz Imaging via Electrooptic Effect," *IEEE Trans. Microwave Theory Tech.* 47, 2644-2650 (1999).

[Katayama 2008] K. Katayama, H. Inoue, H. Sugiya, Q. Shen, T. Toyoda, and K. A. Nelson, : Generation and detection of tunable phonon polaritons using a single transmission grating," *Appl. Phys. Lett.* **92,** 031906 (2008).

[Kinoshita 1995] S. Kinoshita, Y. Kai, M Yamaguchi, T. Yagi, "Direct Comparison between Ultrafast Optical Kerr Effect and High-Resolution Light Scattering Spectroscopy," *Phys. Rev. Lett.* **75**, 148-151 (1995).

[Koehl 1999] R. M. Koehl, S. Adachi, and K. A. Nelson, "Direct Visualization of Collective Wavepacket Dynamics," *J. Phys. Chem. A* **103**, 10260-10267 (1999).

[Lee 2009] Y.-S. Lee, *Principles of Terahertz Science and Technology* (Springer, 2009).

[Lessing 1976] H. E. Lessing and A. von Jena, "Separation of rotational diffusion and level kinetics in transient absorption spectroscopy," *Chem. Phys. Lett.* **42**, 213-217 (1976).

[Lim 1993] M. Lim, T. A. Jackson, and P. A. Anfinrud, "Nonexponential protein relaxation: Dynamics of conformational change in myoglobin," *Proc. Natl. Acad. Sci.* **90**, 5801-5804 (1993).

[Lin 2009] K.-H. Lin, C. A. Werley, K. A. Nelson, "Generation of multicycle THz phonon-polariton waves in a planar waveguide by tilted optical pulse fronts," *Appl. Phys. Lett.* **95**, 103304, (2009).

[Magnification] In reality the signal at the camera is a magnified, inverted image of the phase pattern after the sample: $P(-mx,-mz) \propto \Delta\phi(x,z)$. This scaling and inversion of the field is not important for understanding the concepts in this paper, so the magnification factor will be omitted from equations in the text. Said another way, the analysis assumes a non-inverting imaging system with a magnification of one.

[Oppenheim 1997] A. V. Oppenheim, A. S. Willsky, and S. H. Nawab, *Signals and Systems, 2^{nd} ed.* (Prentice Hall, Upper Saddle River, New Jersey, 1997).

[Palfalvi 2008] L. Palfalvi, J. A. Fulop, G. Almasi, and J. Hebling, "Novel setups for extremely high power single-cycle terahertz pulse generation by optical rectification," *Appl. Phys. Lett.* **92,** 171107 (2008).

[Patorski 1989] Patorski, K. "The self-imaging phenomenon and its applications," in *Progress in Optics*, E. Wolf, ed. (Elsevier, 1989), Vol. 27, pp. 1-108.

[Peier 2008] P. Peier, S. Pilz, F. Müller, K. A. Nelson, and T. Feurer, "Analysis of phase contrast imaging of terahertz phonon-polaritons," *J. Opt. Soc. Am. B* **25**, B70-B75 (2008).

[Peier 2009] P. Peier, S. Pilz, T. Feurer, "Time-resolved coherent imaging of a THz multilayer response," *J. Opt. Soc. Am.* B **26**, 1649-1655 (2009).

[Ruhman 1988] S. Ruhman, A. G. Joly, and K. A. Nelson, "Coherent molecular vibrational motion observed in the time domain through impulsive stimulated Raman scattering," *IEEE J. Quantum Electron.* **24**, 460-469 (1988).

[Sagnac 1913] G. Sagnac, "L'ether lumineux demontre par l'effet du vent relatif d'ether dans un interferometre en rotation uniforme," *Compt. Rend.* **157**, 708–710 (1913).

[Statz 2007] E. R. Statz, D. W. Ward, and K. A. Nelson, “Phonon-polariton excitation in ferroelectric slab waveguides and photonic crystals,” in *Ultrafast Phenomena XV*, P. Corkum, D. M. Jonas, R. J. D. Miller, A. M. Weiner eds. (Springer, 2007), pp. 784-786.

[Stepanov 2003] A. G. Stepanov, J. Hebling, and J. Kuhl, “Efficient generation of subpicosecond terahertz radiation by phase-matched optical rectification using ultrashort laser pulses with tilted pulse fronts,” *Appl. Phys. Lett.* **83**, 3000 (2003).

[Stepanov 2004] A. G. Stepanov, J. Hebling, and J. Kuhl, “Generation, tuning, and shaping of narrow-band, picosecond THz pulses by two-beam excitation,” *Opt. Express* **12,** 4650-4658 (2004).

[Stoyanov APL 2002] N. S. Stoyanov, T. Feurer, D. W. Ward, K. A. Nelson, “Integrated diffractive THz elements,” *Appl. Phys. Lett.* **82**, 674–676 (2002).

[Stoyanov NM 2002] N. S. Stoyanov, D. W. Ward, T. Feurer, and K.A. Nelson, “Terahertz polariton propagation in patterned materials,” *Nat. Mater.* **1**, 95-98 (2002).

[Stoyanov 2004] N. S. Stoyanov, T. Feurer, D. W. Ward, E. R. Statz, and K. A. Nelson, “Direct visualization of a polariton resonator in the THz regime,” *Opt. Express* **12,** 2387-2396 (2004).

[Tachizaki 2006] T. Tachizaki, T. Muroya, O. Matsudaa, Y. Sugawara, D. H. Hurley, O. B. Wright, “Scanning ultrafast Sagnac interferometry for imaging two-dimensional surface wave propagation,” *Rev. Sci. Inst.* **77**, 043713, (2006).

[Talbot 1836] H. F. Talbot, “Facts relating to optical science no. IV,” *Philos. Mag.* **9**, 401-407 (1836).

[Temple 1975] P. A. Temple, “An introduction to phase-sensitive amplifiers: An inexpensive student instrument,” *Am. J. Phys.* **43**, 801-807 (1975).

[Valdmanis 1983] J. A. Valdmanis, G. A. Mourou, and C. W. Gabel, “Subpicosecond electrical sampling,” *IEEE J. Quantum Electron.* **QE-19**, 664-667 (1983).

[Valk 2004] N. C. J. van der Valk, T. Wenckebach, and P. C. M. Planken, “Full mathematical description of electro-optic detection in optically isotropic crystals,” *J. Opt. Soc. Am. B* **21**, 622-631 (2004).

[Vos 1993] M. H. Vos, F. Rappaport, J-C Lambry, J. Breton, and J-L Martin, “Visualization of coherent nuclear motion in a membrane protein by femtosecond spectroscopy,” *Nature* **363**, 320-325 (1993).

[Wahlstrand 2003] J. K. Wahlstrand and R. Merlin, “Cherenkov radiation emitted by ultrafast laser pulses and the generation of coherent polaritons” *Phys. Rev. B* **68**, 054301 (2003)

[Ward 2004] D. W. Ward, J. D. Beers, T. Feurer, E. R. Statz, N. S. Stoyanov, and K. A. Nelson, “Coherent control of phonon-polaritons in a terahertz resonator fabricated with femtosecond laser machining,” *Opt. Lett.* **29,** 2671-2673 (2004).

[Ward 2005] D. W. Ward, E. R. Statz, K. A. Nelson, R. M. Roth, and R. M. Osgood, “Terahertz wave generation and propagation in thin-film lithium niobate produced by crystal ion slicing,” *Appl. Phys. Lett.* **86,** 022908 (2005).

[Ward 2007] D. W. Ward, E. R. Statz, and K. A. Nelson, “Fabrication of polaritonic structures in $LiNbO_3$ and $LiTaO_3$ using femtosecond laser machining,” *Appl. Phys. A* **86**, 49-54 (2007).

[Weiss 2001] C. Weiss, G. Torosyan, Y. Avetisyan, and R. Beigang, “Generation of tunable narrow-band surface-emitted terahertz radiation in periodically poled lithium niobate,” *Opt. Lett.* **26,** 563-565 (2001).

[Werley 2010] C. A. Werley, Q. Wu, K.-H. Lin, C. R. Tait, A. Dorn, and K. A. Nelson, "A comparison of phase sensitive imaging techniques for studying THz waves in structured $LiNbO_3$," *J. Opt. Soc. Am. B* **27** 2350-2359 (2010).

[Werley 2011] C. A. Werley, S. M. Teo, and K. A. Nelson, "Pulsed laser noise analysis and pump-probe signal detection with a data acquisition card," *Rev. Sci. Instrum.* **82**, 123108 (2011).

[Werley 2012] C. A. Werley, K. Fan, A. C. Strikwerda, S. M. Teo, X. Zhang, R. D. Averitt, and K. A. Nelson, "Time-resolved imaging of near-fields in THz antennas and direct quantitative measurement of field enhancements," *Opt. Express* **20**, 8551-8567 (2012).

[Werley AJP 2012] C. A. Werley, C. R. Tait, and K. A. Nelson, "Direct visualization of terahertz electromagnetic waves in classic experimental geometries," *Am. J. Phys.* **80**, 72-81 (2012).

[Werley OE 2012] C. A. Werley, K. Fan, A. C. Strikwerda, S. M. Teo, X. Zhang, R. D. Averitt, and K. A. Nelson, "Time-resolved imaging of near-fields in THz antennas and direct quantitative measurement of field enhancement," *Opt. Express* **20**, 8551-8567 (2012).

[Wright 2008] O. B. Wright, B. Perrin, O. Matsuda, and V. E. Gusev, "Optical excitation and detection of picosecond acoustic pulses in liquid mercury," *Phys. Rev. B* **78**, 024303 (2008).

[Wu 1996] Q. Wu, T. D. Hewitt, and X.-C. Zhang, "Two-dimensional electro-optic imaging of THz beams," *Appl. Phys. Lett.* 69, 1026-1028 (1996).

[Wu 2009] Q. Wu, C. A. Werley, K.-H. Lin, A. Dorn, M. G. Bawendi, and K. A. Nelson, "Quantitative phase contrast imaging of THz electric fields in a dielectric waveguide," *Opt. Express* **17**, 9219-9225 (2009).

[Yan 1985] Y. X. Yan, E. B. Gamble, and K. A. Nelson, "Impulsive stimulated scattering: General importance in femtosecond laser pulse interactions with matter, and spectroscopic applications," *J. Chem. Phys.* **83,** 5391-5399 (1985).

[Yariv 1971] A. Yariv, *Introduction to Optical Electronics* (Holt, Rinehart and Winston, New York, 1971).

[Yariv 2007] A. Yariv and P. Yeh, *Photonics, Optical Electronics in Modern Communications*, 6th ed., (Oxford University Press 2007).

[Yeh 2007] K.-L.Yeh, M. C. Hoffmann, J. Hebling, and K. A. Nelson. "Generation of 10 μJ ultrashort THz pulses by optical rectification ," *Appl. Phys. Lett.* **90**, 171121 (2007).

[Zernike 1942] F. Zernike, "Phase contrast: a new method for the microscopic observation of transparent objects," *Physica* **9** 686-698 (1942).

[Zernike 1955] F. Zernike, "How I discovered phase contrast," *Science* **121**, 345-349 (1955).

Chapter V

Direct visualization of THz waves in classic experimental geometries

Content from: C. A. Werley, C. R. Tait, and K. A. Nelson, “Direct visualization of THz-frequency electromagnetic waves in classic experimental geometries,” *Am. J. Phys.* **80**, 72-81 (2012).

PREFACE

In recent years, educators in science have been striving to bring modern research into their classrooms and build connections between current applications and more traditional coursework. As we were developing the experimental methods discussed in chapter 4, in particular the imaging of THz waves, we realized that many of the results had significant educational value. Direct visualizations of propagating electromagnetic waves are more modern versions of classic water wave demonstrations. We determined to fabricate a set of structures using laser machining that would demonstrate classic experimental geometries, capture videos of THz waves interacting with these structures, and compile and present them for use in classes and lectures. This chapter is drawn from the resulting paper [Werley 2012].

A. ABSTRACT

In this study, we used cutting-edge research methods to collect educational video clips of electromagnetic waves propagating at the speed of light. The electromagnetic waves were of terahertz frequencies and were generated and detected in $LiNbO_3$ crystals structured to reproduce classic optical geometries and experiments, such as two-slit interference and diffraction from a grating. Direct visualization of the phase fronts as the electromagnetic pulses propagate, reflect, diffract, and interfere helps develop intuition and insight about the fundamental behavior of light and waves. We believe these videos will be a valuable addition to lectures on introductory optics and physics, as they will bring modern research to the classroom and provide clear and direct experimental demonstrations of light and wave behavior.

B. INTRODUCTION

In both optics and everyday life, the behavior of light is determined by its wave-like nature, but for students of optics, this can be difficult to visualize. Simple lab experiments that demonstrate wave behavior, such as diffraction or interference, generally do not enable direct visualization of the processes as they unfold or directly follow the electromagnetic wave as it interacts with objects. In Young's two-slit experiment, for example, a screen is usually placed far from the slits to observe an intensity pattern. This is a beautiful experiment, but it does not highlight effects like phase front curvature or interference effects close to the slits. In this work, we explain an experimental technique that enables the direct imaging of electromagnetic waves and the capture of short video clips of waves propagating in a nonlinear optical crystal *at the speed of light*. We then present videos of some classic optical experiments collected using this technique which clearly show wave behavior such as reflection, focusing, interference, diffraction, and waveguiding.

It is our hope that this direct experimental record of propagating electromagnetic waves will provide readers of this article and viewers of the associated videos with valuable physical intuition about wave behavior. We envision that the videos can most effectively be used in lectures or as supplements to lectures. For instance, after deriving the far-field diffraction pattern from two slits, a lecturer could very briefly explain our experimental technique (or at a simpler level, could just indicate that these are genuine, direct experimental measurements) and then show the appropriate video. Figure 5.2 and Sec. 5.C.3 present the key ideas required for understanding how the videos are collected. It should be possible to explain the technique at a basic level and show the selected videos in under 5 minutes, thus not using too much valuable lecture time. We think that students should be encouraged to seek experimental proof of theoretical concepts they are taught, and the videos here provide such proof convincingly. Ultimately we hope that, with the help of other labs, an online library of videos demonstrating a broad set of important optical phenomena will be posted.

It is also important to consider the relative roles of these experimental videos and theoretical simulations. The experimentally compiled video library provides students with opportunities to compare classroom discussions of wave phenomena with real results including non-ideal behavior resulting from stochastic noise fluctuations, sample imperfections, etc. This sort of comparison is a key element of the scientific method. The ability to critically compare expectations with experimental results is critical for physicists and important for all informed citizens when weighing any issues at the intersection of science and society such as climate change, genetically modified foods, and energy resources. Theoretical simulations play a complementary and equally important role. Simulations can build intuition, especially in connection with interactive programs in which users can vary parameters and observe changes in system behavior. In most cases simulations are inexpensive and their results are effectively free of noise, yielding clear illustrations the behavior that results from a particular model. A majority of the results in this paper can be treated with analytical theory, as discussed in Sec. 5.D below. Many simple geometries from introductory electricity and magnetism can also be accurately modeled with finite difference time domain (FDTD) simulations (e.g. [Taflove 2005; Sipos 2008]), and the experiments in which the wave propagates in lithium niobate can be reproduced with a slightly more complicated formulation of the FDTD simulation [Ward 2003]. Using conceptual formulation, analytical theory and simulation, and experimental results in concert can build competence in the specific subject being taught and in scientific thinking in general.

The waves displayed in the videos are terahertz (THz) frequency electromagnetic waves. THz radiation, roughly classified as having a frequency between 0.1 and 10 THz, is nestled between the infrared and microwave regions of the electromagnetic spectrum. The wavelength of these THz waves, $\lambda = c/f \approx 30 - 3000$ μm in free space, is large enough to be easily resolved using optical imaging but small enough that many wave cycles fit within a 1 cm square crystal. Insight gained at THz frequency can be generalized to any other frequency range because the behavior of light has no fundamental length scale [Joannopoulos 2008]. Since length and frequency are scaled by a constant factor $(x' \rightarrow ax, y' \rightarrow ay, z' \rightarrow az, \omega' \rightarrow \omega / a)$, the behavior is identical at any length scale. Many of the phenomena discussed below, such as diffraction, interference, reflection, and focusing, are universal for all wave types, including acoustic (sound or pressure) and water waves.

The field of THz research has been particularly active since around 1990 when new tools for generating and detecting THz radiation were developed [Auston 1988; Grischkowsky 1990]. The first experiments on THz imaging took place in 1996, where THz radiation was imaged after diffracting off an object of interest [Wu 1996; Jiang 1999]. In these measurements the THz field is used for imaging of objects, but the THz itself is not visualized. In contrast, our purpose is to observe the THz field itself. The first experiments of this type were presented moer than ten years ago [Koehl 1999]. Since then numerous experiments have employed imaging to study THz waves [Feurer 2007]. These techniques have been used to study THz generation [Feurer 2002; Feurer 2003; Lin 2009] and the interaction of THz with various structures [Stoyanov APL 2002; Stoyanov 2002; Stoyanov 2004; Peier 2009; Peier 2010] Recent advancements in the methodology of imaging [Peier 2008; Wu 2009; Werley 2010] have made it possible to collect high resolution, in-focus videos of THz waves interacting with a variety of fabricated structures. These recent developments in THz imaging and generation were incorporated in this study to make the images and videos as clear and understandable as possible.

The experimental section that follows explains how THz waves are generated and detected and how videos of propagating waves are captured. It also explains how we produce samples with structures such as slits and waveguides. This background will help the interested reader to better understand the experiments but is not necessary for understanding the physical principles demonstrated by the videos that are presented in the results section. The first experiment highlights interference by crossing two beams, while the second emphasizes diffraction resulting from a wave passing through a single slit. The next three experiments (Young's double-slit experiment, a transmission grating, and a reflection grating) display the combined effects of diffraction and interference. The sixth experiment shows focusing of a curved phase front. The final two experiments display the capabilities and effects of waveguiding. The variety of geometries in these experiments demonstrate a number of important optical behaviors and phenomena and opens the door to many further instructive examples.

C. EXPERIMENTAL

1. Generation of THz electromagnetic waves

THz waves were generated by focusing short, intense laser pulses into a nonlinear crystal. The laser pulses were produced by an "ultrafast" titanium:sapphire laser and amplifier [Strikland 1985], with a center wavelength of 800 nm, a duration of 100 femtoseconds (10^{-13} seconds), and about 1 mJ of energy. Such high-energy, short duration pulses have extremely large peak intensities around 3 x 10^{11} W/cm^2, which allow them to exert significant forces on the ions in a crystal lattice, displacing them from their equilibrium positions. This class of phenomena falls under the umbrella of nonlinear optics, and there are many textbooks written on the subject (e.g. [Boyd 2008; Shen 2003]). For understanding this paper, the key result is that a small part of the energy from the short, high-frequency (~375 THz) optical "pump" pulse is converted into a short, low-frequency (~0.5 THz) THz pulse which is typically several picoseconds in duration.

In the present experiments, the nonlinear crystal used was lithium niobate, $LiNbO_3$. The nonlinear physics of THz generation in $LiNbO_3$ has been well studied [Auston 1988; Dougherty 1992], and in this crystal impulsive stimulated Raman scattering is responsible for converting some of the optical frequency light to THz frequencies. A THz wave generated in $LiNbO_3$ propagates at a large angle relative to the optical pump pulse that produces it [Auston 1988] because the speed of the optical pulse is far higher than the speed of the THz pulse. The effect, labeled the “Cherenkov” effect, is very similar to that of a boat rushing through water faster than the speed of water waves, where the wake left by the boat propagates mainly sideways relative to the boat speed in a typical inverted “V” pattern. In the experiments described below, we restrict the THz propagation angle to be not just mainly but fully perpendicular to the direction of the optical pulse by guiding the THz wave with a very thin slab of $LiNbO_3$. The $LiNbO_3$ slab, which has a high index of refraction and is only 50 μm thick, acts as a waveguide that traps the THz wave within it through total internal reflection. (Optical fibers are also examples of waveguides; even when they are bent, light propagates through them without escaping into the air.) The optical pump propagates perpendicular to the slab surface, and the generated THz travels in the plane of the slab [see Fig. 5.1(a)]. Many characteristics of the THz wave can be controlled by changing the pump pulse characteristics [Feurer2003; Lin2009]. Here we choose the center frequency and duration of the THz pulse which most clearly demonstrate the optical principles involved.

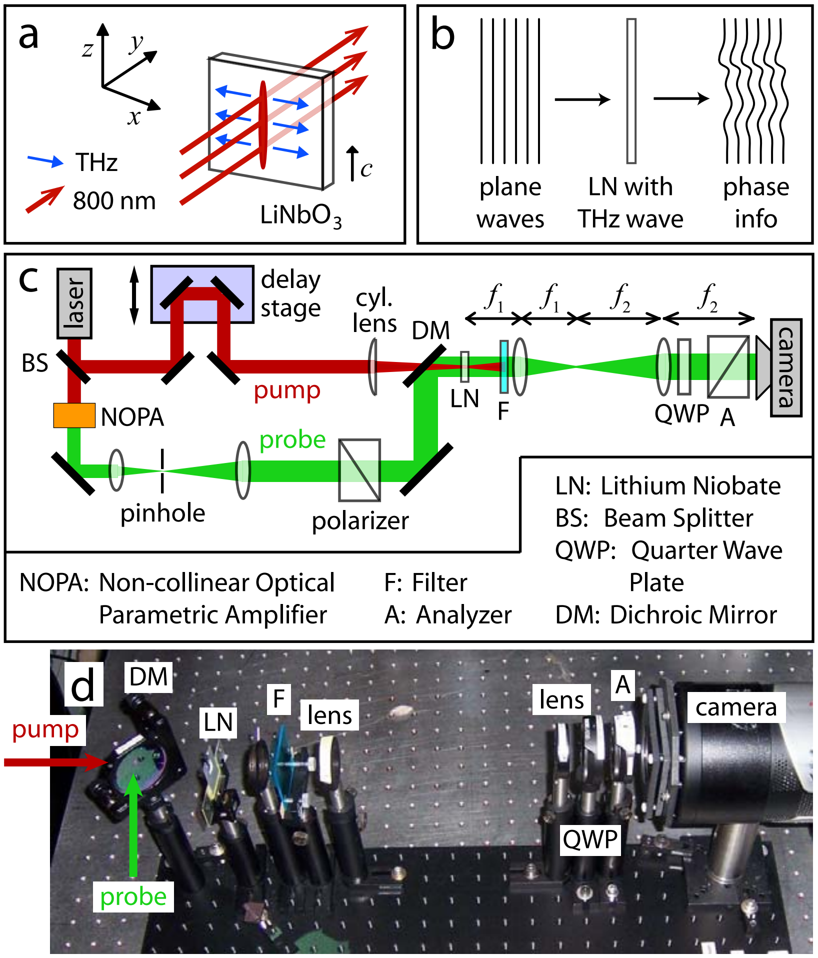

Figure 5.1 | Experimental setup. (a) The generation geometry: the pump laser pulse propagates orthogonal to the slab surface, and the generated THz wave propagates perpendicular to the pump pulse, in the plane of the slab. (b) The vertical lines to the left of the sample represent the optical plane wave phase fronts of the incoming probe pulse. After the sample, the phase fronts are distorted due to the presence of the THz wave. (c) A simplified layout for the entire experiment. The short pulses coming out of the laser are split by the 90% reflecting beam splitter. The pump beam is routed through a delay stage so it can be delayed relative to the probe. The sample is placed between crossed polarizers so any THz-induced birefringence in the sample can be recorded on the camera, providing an image of THz field profile. (d) A photograph of the detection part of the experimental setup corresponding the diagram in (c). The drilled holes on the breadboard and optical table are spaced 1 inch apart.

2. Imaging of THz electromagnetic waves

A THz wave generated as described above propagates at the speed of light in the sample, and so must be detected using an extremely fast method. Here, a second ultrafast optical pulse is used, called a probe pulse. In these experiments, we converted 800 nm light to 532 nm light for the probe, so it could be separated from the pump after the sample with a single spectral filter. The probe light was spatially filtered to improve the mode quality and expanded [Fig. 5.1(c)] so it would illuminate the whole sample. The probe is sensitive to a THz-induced change in the sample. Since $LiNbO_3$ is an electro-optic crystal because of its particular nonlinear properties, an applied electric field changes its index of refraction: $n = n_0 + aE$, with n_0 the index of refraction in the absence of an electric field, E the applied field, and a a crystal-dependent constant. In reality, index and field are connected through a tensor relationship which depends on the angle of the field, crystal, and probing light [Feurer 2007], but these details are not essential for understanding this work. The THz wave can be thought of as a spatially dependent electric field, $E_{\mathrm{THz}}(x,z)$, which can be approximated as constant through the depth of the crystal due to the waveguiding geometry [Werley 2010]. Thus the THz field induces a spatially dependent index of refraction in the sample

$$n(x,z) = n_0 + aE_{\mathrm{THz}}(x,z) \tag{5.1}$$

The probe pulse is an expanded laser beam that can be approximated as a plane wave. As the plane waves propagate through the sample, they develop a phase shift proportional to the index of refraction [see Fig. 5.1(b)]. Where the refractive index has been increased, the probe light speed is reduced and the probe phase lags behind that of the unperturbed plane wave. Where the refractive index has been decreased, probe light speed is increased and the probe phase runs ahead. Thus, immediately after the sample, a phase pattern is imposed on the beam, and the induced phase shift at any point in the (x,z) plane is directly proportional to the THz electric field at that point.

The sample is imaged onto the camera using a 2-lens, 4-focal length imaging system. If nothing more were done, the $LiNbO_3$ would be imaged perfectly onto the camera but no THz wave would be detected. The THz would not be visible because it modulates the phase of the probe light, and the camera is only sensitive to the intensity. Thus, some optical technique must be used to perform phase-to-amplitude conversion. A number of methods can do this conversion [Werley 2010], but here we will use the polarization gating method.

A simplified experimental geometry containing the key conceptual features of the polarization gating setup is shown in Fig. 5.1(c). Polarization gating detects a change in the birefringence of the $LiNbO_3$ sample induced by the THz field, which changes the index of refraction of different crystal axes by different amounts. To perform polarization gating, the sample must be placed between two crossed polarizers. The first polarizer sets the probe polarization to 45 degrees, and the THz-induced change in the polarization state of the probe changes the transmission through the second analyzing polarizer. The quarter-wave plate sets the unperturbed transmission to 50% and ensures that a positive phase shift will be detected as a positive amplitude change and visa-versa. The detected intensity for the diagram in Fig. 5.1(c) is given by [Werley 2010]:

$$I(x,z) = I_0(x,z)[1 + \sin(\Delta\varphi(x,z))] \tag{5.2}$$

I_0 is the unperturbed intensity measured when there is no THz wave present (e.g. when the pump pulse is blocked) and $\Delta\varphi$ is the difference in induced phase shift experienced by vertically and horizontally polarized light in the presence of the THz wave. This is given by:

$$\Delta\varphi(x,z) = -2\pi\frac{d}{\lambda_{\mathrm{pr}}}\left[\Delta n_v(x,z) - \Delta n_h(x,z)\right] = 2\pi\frac{d}{\lambda_{\mathrm{pr}}}a' E_{\mathrm{THz}}(x,z) \tag{5.3}$$

with d the slab thickness, λ_{pr} the probe wavelength, Δn_v and Δn_h the THz-induced change in index for vertically and horizontally polarized light, and a' a material dependent constant. By measuring both $I(x,y)$ and $I_0(x,y)$, the spatially dependent THz field can be determined.

3. Capturing videos of THz waves: the pump-probe method

The imaging method in the previous section explains how to capture the spatial profile of a THz wave at one moment in time. To collect a video of the wave as it propagates and evolves, the pump-probe method is required. The short laser pulse coming out of the laser is split by a beamsplitter [see Fig. 5.1(c)]. Most of the energy goes into the pump, which is used to launch the THz wave, and some goes into the probe, which is used to image the THz wave that the pump launched. Both pulses propagate in air at the speed of light, and the time it takes for each pulse to reach the $LiNbO_3$ sample is given simply by the path length divided by c. Thus, the amount of time the probe arrives after the pump is:

$$\tau = (d_{\text{pr}} - d_{\text{pu}})/c \tag{5.4}$$

where d_{pr} and d_{pu} are the respective optical path lengths of the probe and pump pulses. When the path lengths of both the pump and probe arms are exactly the same, as shown in the first frame of Fig. 5.2, the probe measures the moment of THz wave generation. The image in the lower right shows the signal that would be collected on the camera for point source generation (pump focused into the crystal). As the pump path length is shortened (by moving the motorized delay stage in Fig. 5.2 toward the probe beam), τ is increased and the THz wave has longer to evolve before its spatial profile is captured by the probe. To make a video, an image is captured as described above for a series of time delays. Each delay corresponds to one frame in the video.

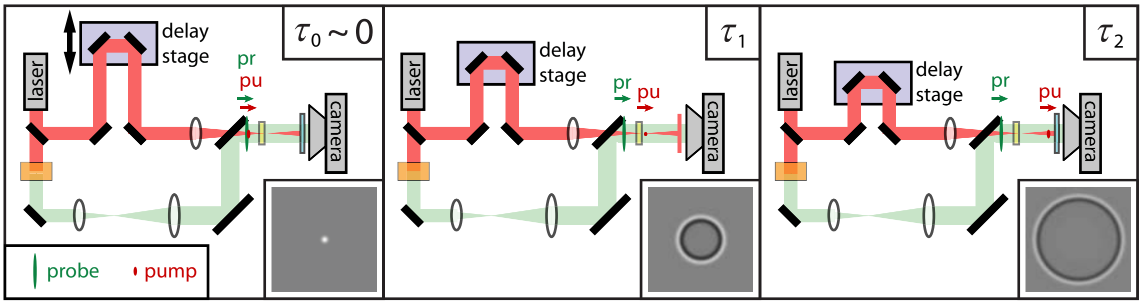

Figure 5.2 | Pump-probe measurements. Three simplified pictures of the experimental setup showing that as the delay stage moves, the time delay between pump and probe pulses is changed. The images inset in the lower right of each frame show the THz wave that would be captured on the camera if the wave were launched by a point source pump, expanding outward much like the water wave that is launched when a stone is thrown into a pond.

The general technique of increasing the delay between two pulses is called the pump-probe method, and it is widely used to study phenomena that happen on timescales faster than ~1 nanosecond, that is, the attosecond, femtosecond, and picosecond regimes. In order for this method to work, the same experiment must be performed hundreds of times (at least once for each time delay), and the results must be reproducible to build a consistent picture. Here, we rely on the fact that nearly identical laser pulses launch the THz waves the same way every time without damaging or otherwise modifying the sample. Because we have to make the measurement at many time delays, and we average the results of many pulses at each delay to reduce noise, acquiring these movies often takes many hours.

4. Structuring the sample using laser machining

In order to recreate classic experiments in the thin $LiNbO_3$ slab, it is necessary to introduce structures such as mirrors, slits, waveguides, etc. directly into the slab. These structures are introduced by cutting air gaps into the crystal slab. Because $LiNbO_3$ has a very high index of refraction ($n \sim 5.1$) for THz-frequency light, reflection off air interfaces is very efficient. The total internal reflection (TIR) angle is only 11 degrees in bulk $LiNbO_3$, so THz waves are reflected off even gently tilted surfaces with 100% efficiency. We used femtosecond laser machining to cut air gaps in the crystal slab [Ward 2007]. In this method, a short-pulsed laser is tightly focused into the sample using a microscope objective. The sample is moved along a specified trajectory using computer controlled motors, and the laser cuts through the crystal slab as it moves. If the sample is moved in a circle, a circular hole is cut out of the slab. If the sample is moved along the edges of a square, a square hole is cut. Any pattern can be programmed into the computer and cut into the sample, making it relatively easy to generate interesting structures.

D. RESULTS: CLASSIC EXPERIMENTS VISUALIZED

Each of the following sections 1-8 describes a different experimental geometry. Each geometry has one or two video clips in the supporting online material, and a selection of video frames are shown in the associated figure. In all the videos and figures, the darker-gray region is $LiNbO_3$ crystal, and the uniform, lighter-gray region is a machined air gap cut as described in Sec. 5.C.4 above. In Figs. 5.3-5.6, there is a theoretical prediction for the field pattern in addition to the experimental data. The theoretical prediction assumes a continuous wave (CW) input of frequency ν that is truly periodic in time. It shows the root mean square electric field: $E_{rms}(x,z) = \left[\nu \int_{\nu t=0}^{1} E^2(x,z,t)\,dt \right]^{1/2}$, or the square root of the time averaged intensity, which is what would typically be observed by eye in a simple experiment with a HeNe laser and diffractive optic in an introductory physics lab.

1. Crossing two beams

In this simple experiment, two beams,which are well approximated by plane waves, are crossed to make an interference pattern [see Fig. 5.3(a)]. As with many of the experiments in this paper, the behavior is equally applicable to any kind of wave, including electromagnetic, acoustics (sound), and water waves. The many applications of this geometry include crossing beams to generate a periodic pattern in optical lithography [Mack 2007] or a transient grating for a wide array of experiments [Eichler 1986]. The spatial period of the intensity pattern, Λ, depends only on the wavelength in the material, λ, and the crossing angle, θ:

$$\Lambda = \frac{\lambda}{2\sin(\theta/2)} \tag{5.5}$$

Figure 5.3(a) shows the amplitude pattern predicted for crossing two beams, one propagating from upper left to lower right in the figure, and the other propagating from lower left to upper right. Figures 5.3(b) – 5.3(e) contain frames from video 1 in [Werley 2012] material showing two beams crossing. Initially, a THz wave with several optical cycles reflects off two tilted surfaces. The reflected waves cross, forming an interference pattern with horizontal nodes as predicted by the theory. As the waves propagate, the phase within the intensity maxima evolves but the horizontal nodes in the interference pattern remain unchanged. By the end of the video, the two beams have separated and the interference pattern has mostly disappeared. The angle between the beams is 60 degrees, so Eq. 5.5 above predicts that the period of the *intensity* pattern should be equal to the THz wavelength (~ 160 μm), which agrees with the experiment. Note that Fig. 5.3(a) shows the *time-averaged rms value* of the field, so the time-dependent field oscillations between positive and negative values that are apparent in (b)-(e) do not appear in (a).

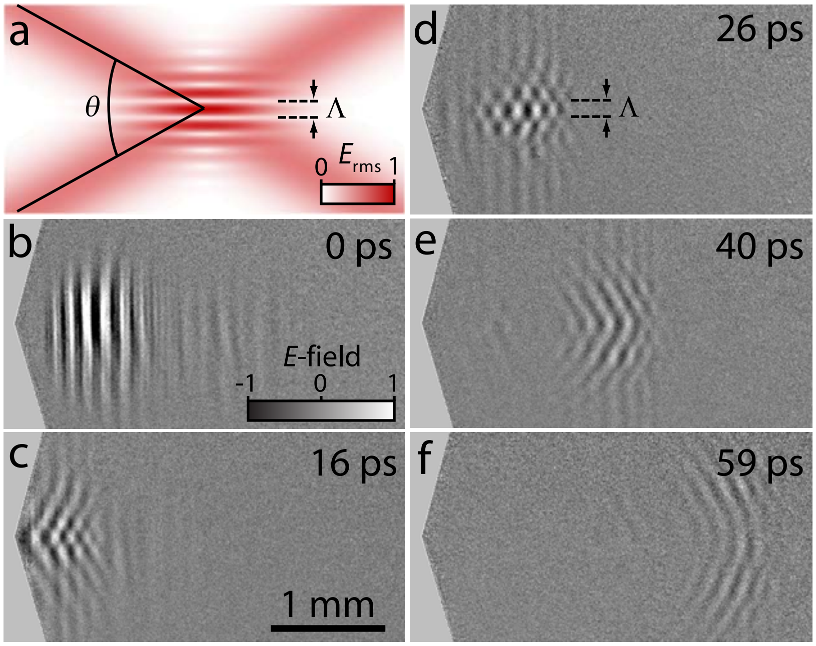

Figure 5.3 | Beams crossing at an angle of 60 degrees. (a) shows the theoretical time-averaged E-field amplitude ($|E| = \sqrt{I}$) which results from crossing two continuous-wave beams. (b) – (f) show frames from video 1 in [Werley 2012] at a sequence of time delays. The uniform, light gray region on the left of each frame is an air gap that has been cut into the thin LiNbO3 slab. The THz wave is clearly visible in each frame of the video, with white corresponding to a positive electric field and black corresponding to a negative electric field. The wave is propagating to the left in (b). In (c) the wave has just reflected off the air interface and the upper half and lower halves of the initial pulse have formed two beams that are crossing at an angle of 60 degrees. In (d) the two beams are fully overlapped, and by (f) they have almost completely separated.

2. Diffraction through one slit

Diffraction through a single slit is a classic wave mechanics phenomenon that appears in many freshman level physics texts. In this experiment, a plane wave is incident on a small slit, and the diffraction pattern of the transmitted light is observed. Although it is possible to determine the behavior near the slit (a complicated interference pattern) using the Fresnel Kirchhoff formula [Fowles 1989], the treatment is mathematically too complex for an introductory text. It is possible to determine behavior far from the slit, the Fraunhoffer diffraction limit, using fairly simple methods [Resnik 2002]. The far-field electric field pattern as a function of angle is:

$$I = I_0 \left[\frac{\sin \gamma}{\gamma} \right]^2 \tag{5.6}$$

with $\gamma = \pi a \sin(\theta)/\lambda$ and λ the wavelength. The slit width a and the emission angle θ are defined in Fig. 5.4(a). If the slit width is less than the wavelength, then the slit acts like a point source, which yields the dipole emission pattern, and only the main central lobe is observed. This is the case in Fig. 5.4(a) – (f) and video 2, where the slit width is $a = 4\lambda/5$. Figure 5.4(a) and (g) show the RMS E-field patterns calculated exactly using the Fresnel Kirchhoff formula [Fresnel Kirchhoff], a method which will also be used in future Figs. 5.5(a) and 5.6(a). In Fig. 5.4(g) – (l) and video 3, the slit width is $a = 3\lambda$. As predicted by Eq. 5.6, there are nodes in the emission pattern for this larger gap spacing. The central lobe contains most of the energy and diverges more slowly than in the smaller gap spacing.

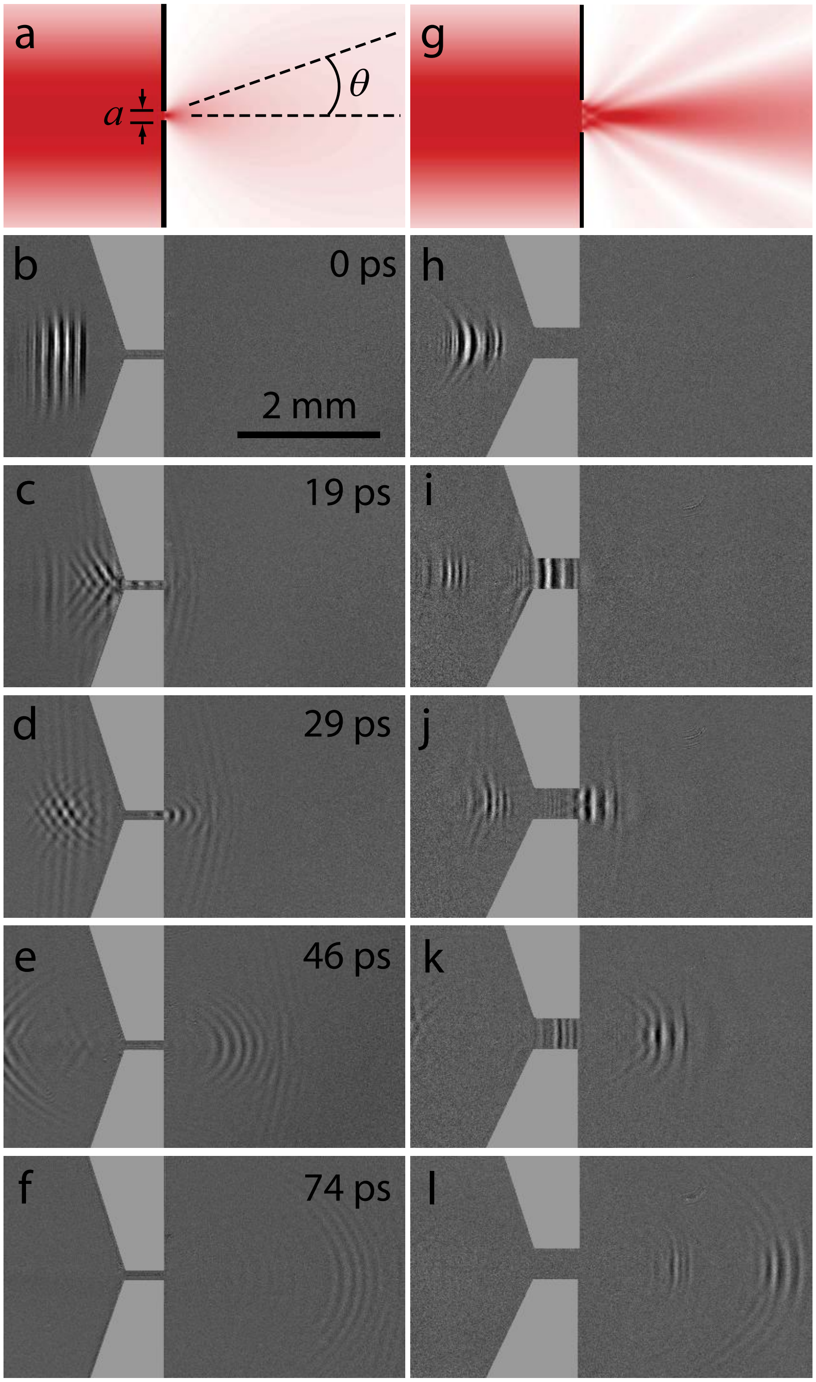

Figure 5.4 | Diffraction through a single slit. (a) – (f) are for a wave propagating through a slit whose width is $4\lambda/5$. (g) – (l) are for a slit whose width is 3λ. (a) and (g) are theoretical predictions for the time-averaged electric field amplitudes for a continuous wave beam entering from the left. (b) – (f) and (h) – (l) are frames from videos 2 and 3 respectively in [Werley 2012] showing experimentally measured electric field profiles for a short THz pulse entering from the left. Adjacent images have the same time delays. The uniform light gray regions are air gaps that have been cut into the crystal.

3. Diffraction through two slits

Two-slit diffraction is another classic experiment discussed in many introductory texts. A plane wave is incident on a pair of slits of width a separated by a distance d [see Fig. 5.5(a)]. Light diffracts from each slit as described in Sec. 2 above, and the light from the two sources interfere with each other. Again, it is complicated to analytically describe the behavior near the slits, but the far-field behavior can be calculated with relative ease [Resnik 2002]:

$$I = I_0 \left[\frac{\sin\gamma}{\gamma} \right]^2 \cos^2[\delta] \tag{5.7}$$

where $\delta = \pi d\sin(\theta)/\lambda$ has a similar definition to γ in Sec. 2 above. Equation 5.7 is simply a periodic interference pattern multiplied by the single slit diffraction pattern from Sec. 2. In Fig. 5.5 and video 4 in [Werley 2012], the slit spacing is $d = 2.5\lambda$ and the slit width is $a = \lambda$. Because a is close to the wavelength, the diffraction part of the equation is smooth and slowly decaying as one moves away from the center. The modulation results from interference between the outputs from the two slits. The video shows the interference evolving close to the slits and eventually developing into the far-field pattern described by Eq. 5.7 above.

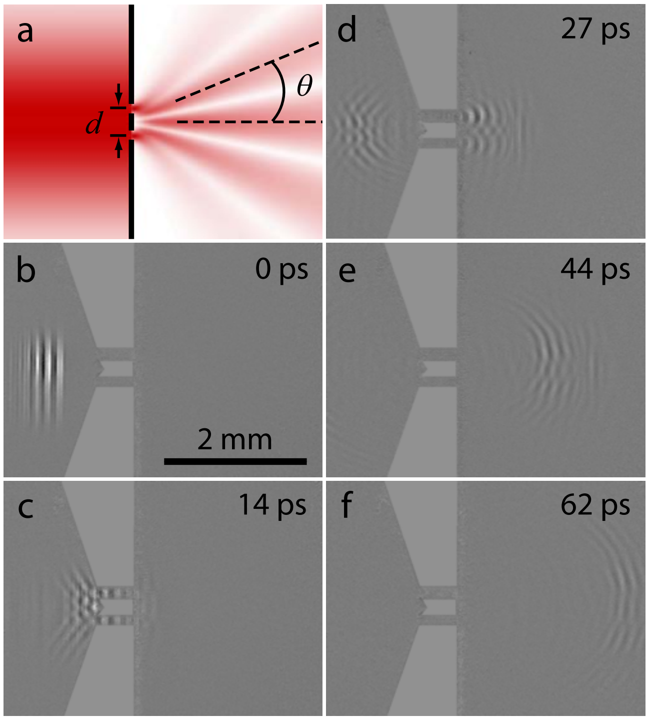

Figure 5.5 | Two-slit diffraction. The figure shows diffraction through a pair of slits with a width of lambda and a slit center to slit center spacing of 2.5λ. (a) shows the theoretical, time-averaged electric field amplitude. (b) – (f) show frames from video 4 in [Werley 2012]. The uniform, slightly lighter gray region is an air gap. The wave is initially propagating to the right, and a significant fraction is reflected off the air gap while part propagates through the slits. The THz waves reflected or transmitted from the two slits interfere with one another as the waves propagate away from the slits.

4. Diffraction through *N* slits

As more slits are added, the intensity maxima become narrower and brighter. Between the bright peaks are a set of nulls and significantly weaker maxima [see Fig. 5.6(a)]. The exact expression in the far-field is given by [Fowles 1989]:

$$I = I_0 \left[\frac{\sin \gamma}{\gamma} \right]^2 \left[\frac{\sin(N\delta)}{N \sin(\delta)} \right]^2 \tag{5.8}$$

with $\gamma = \pi a \sin(\theta)/\lambda$ and $\delta = \pi d \sin(\theta)/\lambda$. As in Sec. 5.D.3 above, the single-slit envelope function is multiplied by a factor describing the interference. The limit of an infinite number of slits is a transmission grating, in which the weak peaks disappear completely and there are a set of narrow, well defined diffraction directions. The experiment here, where roughly 5 slits are illuminated, shows an intermediate case between the double slit and a transmission grating. Although it is difficult to see all the minor maxima in the data {Fig. 5.6(b) – (f) and video 5 in [Werley 2012]}, it is interesting to see how a complicated interference pattern close to the slit evolves into a simpler diffraction pattern in the far-field.

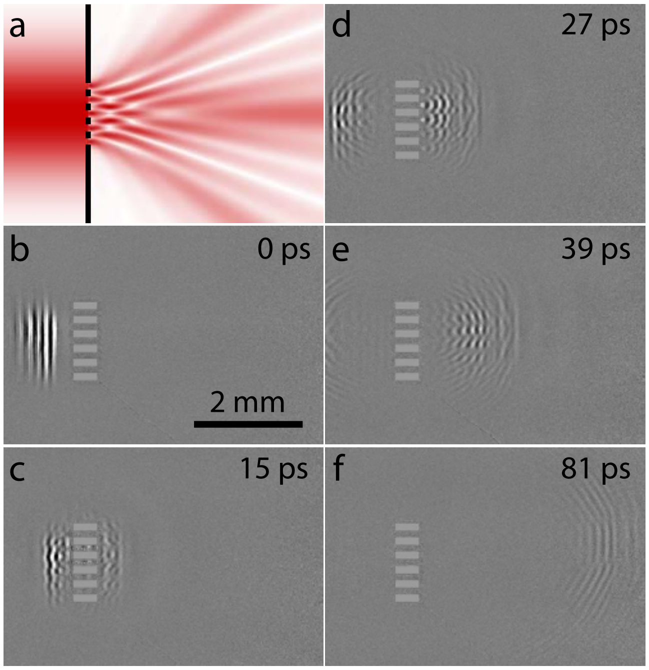

Figure 5.6 – Five-slit diffraction. Panel a shows the RMS *E*-field for a CW laser beam. Frames (b)-(f) are from video 5 in [Werley 2012]. The uniform light gray regions are air gaps. Part of the initially rightward propagating wave reflects off the air gaps and part is transmitted through the slits. Wave propagation reveals the mechanism through which interference close to the slits evolves into the far-field emission pattern. Slit period $d = 5\lambda/3$, slit width $a = 5\lambda/6$.

5. Reflection Grating

Of all the geometries described so far, the reflection grating is the most used as a component in typical optical experiments. Because the points of diffraction are so close together, a monochromatic source generates very well-defined diffraction orders. The well-known grating equation relating the diffraction directions and groove spacing is [Palmer 2005]:

$$m\lambda = d[\sin(\alpha) + \sin(\beta_m)] \tag{5.9}$$

where m is an integer, λ is the wavelength, d is the groove period on the grating, α is the angle of incidence and β_m is the angle of reflection as defined in Fig. 5.7(a). The convention used to determine the signs of the angles is indicated by the + and - in that figure. Any periodic surface with period d will follow Eq. 5.9. Propagating diffraction orders will exist so long as $|m\lambda/d| < 2$, when the sum of the sines does not exceed 2. Although the orders exist according to Eq. 5.9, the distribution of intensity into the different orders is determined by the detailed profile of the periodic structure.

Many gratings are created with a saw-tooth surface pattern. This allows one to selectively send most of the energy into a specific diffraction order and efficiencies in excess of 90% are routinely achieved [Palmer 2005]. In a grating with such a saw-tooth pattern, the blaze angle, or the angle of the tooth surface, is set so that if the surface of each tooth were extended, the incident and diffracted beam would obey the law of reflection. In effect, each facet or tooth behaves like a sub-wavelength mirror. For a overview of different types of gratings and grating designs, see Newport's Diffraction Grating Handbook [Palmer 2005]. The Littrow configuration, where the angle of incidence is equal to the angle of reflection ($\beta_1 = \alpha$), is very common due to its high efficiency and is shown here.

To prepare a grating in the Littrow configuration, the period of the steps are designed and the grating is angled such that the $m = 1$ diffraction order propagates directly back along the incoming direction. The grating is blazed such that the surface of each step is vertical and the sub-wavelength mirrors reflect directly backwards, improving the efficiency. Figure 5.7(b) - (f) shows frames from video 6 in [Werley 2012]. The wave initially propagates to the right and diffracts off the grating. The zero order reflection, visible in Fig. 5.7(d) - (f), propagates up and to the left. The first order diffracted beam propagates directly back along the incoming direction. The intensity front of the diffracted wave is tilted, a feature that would not be visible if the light source were continuous instead of a short pulse. In fact, a grating was used to tilt the visible pump pulse that generated the THz wave in this video [Lin 2009].

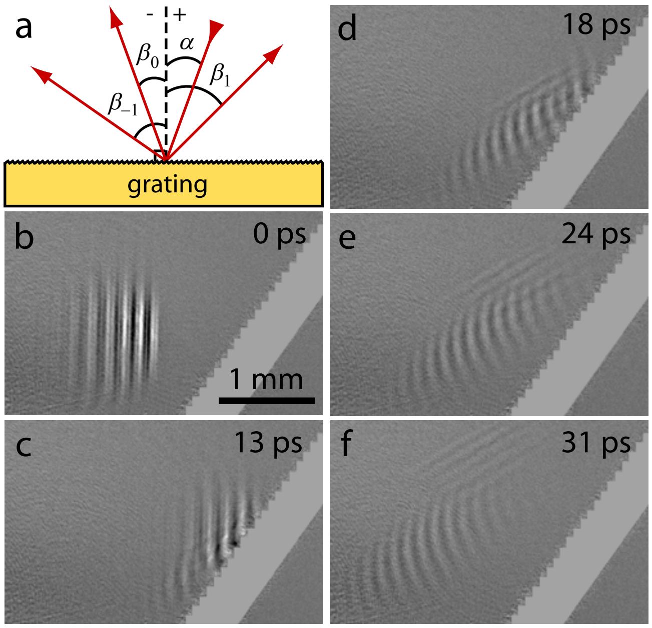

Figure 5.7 | A reflection grating. Panel (a) defines the angles for the grating geometry. Panels (b)-(f) are frames from video 6 in [Werley 2012]. The uniform light gray region is an air gap that has been carved into the crystal. The grating period, angle, and blaze have been optimized for the Littrow configuration, shown here. The initially rightward propagating wave is split into the 0th order beam (the reflection that would be produced by a smooth surface) which propagates up to the left and the 1st order beam which is diffracted directly back along the input path.

6. Wave Focusing

The simple geometry here follows a focusing electromagnetic wave. A curved THz wave is generated using an axicon, a glass cone, in combination with a lens to focus the pump beam to a ring on the sample [Stoyanov2002]. The curvature of the generated phase front is similar to the phase front resulting from transmission through a lens or reflection off a curved mirror. When collecting this data set, the right side of the ring of light was blocked to form a semicircular THz wave pattern. Figure 5.8 and video 7 in [Werley 2012] show the focusing THz wave. The wave collapses to a tight spot, enhancing and localizing the electric field, and then quickly diverges. The spot size at the focus is given by:

$$d \approx \frac{\lambda_0}{2\mathrm{NA}} = \frac{\lambda_0}{2n\sin\theta} \tag{9}$$

where λ_0 is the wavelength in free space, NA is the numerical aperture, n is the index of refraction, and θ is the angle defined in Fig. 5.8. Lenses and mirrors with large acceptance angles and thus large numerical apertures, such as microscope objectives, can focus to (or resolve) very small spots. In the experiment shown in Fig. 5.8, the index is 6.15 and θ is almost 90°, yielding an NA of about 6, much larger than that of common optical lenses and mirrors. The wavelength observed in the experiment is λ_0/n, so the diameter of the focused spot is about half the observed wavelength. One other interesting phenomenon noticeable in video 7 is the Gouy phase shift ─ the change in wave polarity as it passes through the focus [Feurer 2002; Guoy 1980]. Direct visualization of this effect is important for understanding how fields change in the focus of an electromagnetic wave.

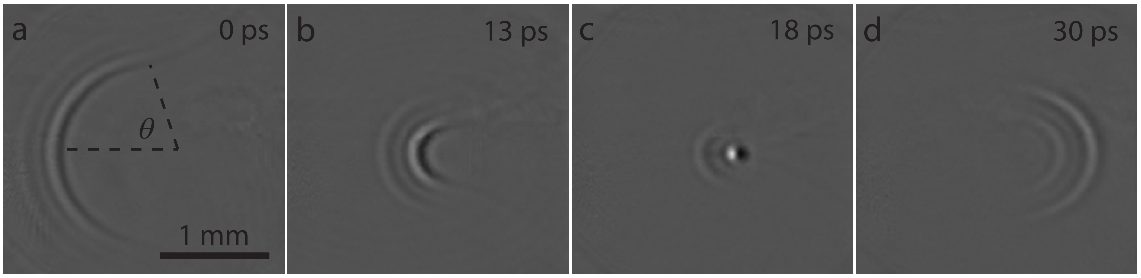

Figure 5.8 | Focusing. Panels (a)-(d) are frames from video 7 in [Werley 2012]. An axicon focuses the pump pulse into a 5 mm thick slab of $LiTaO_3$ to generate a curved phase front. The rightward propagating wave is focused to a diffraction-limited spot in panel (c), before it starts diverging.

7. Y-coupler

A Y-coupler is a simple structure that demonstrates waveguiding and interference. Air gaps (light gray in Fig. 5.9 and video 8 in [Werley 2012]) have been cut into the slab leaving $LiNbO_3$ in the shape of a "Y." A tall, uniform THz wave is launched initially and a small section of it is incident on the top and bottom arms of the Y. These sections of the THz wave are waveguided, trapped inside the high index of refraction material due to total internal reflection. (Recall that THz waves inside a lithium niobate slab are already in a planar waveguide, but since the confinement is along the same direction as the optical pump and probe beam propagation, preventing the THz waves from escaping the plane of the lithium niobate, the images do not directly illustrate waveguiding. In the present case, waveguiding is apparent in the images because confinement prevents the THz waves from propagating along some directions within the image plane.) The THz waves travel down each arm of the Y until they meet in the center and interfere. The wave resulting from their superposition is further guided down the stem of the Y until it is emitted. The idea of using a high-index material to trap and guide light is an important one because it is responsible for the functioning of optical fibers and other photonic components.

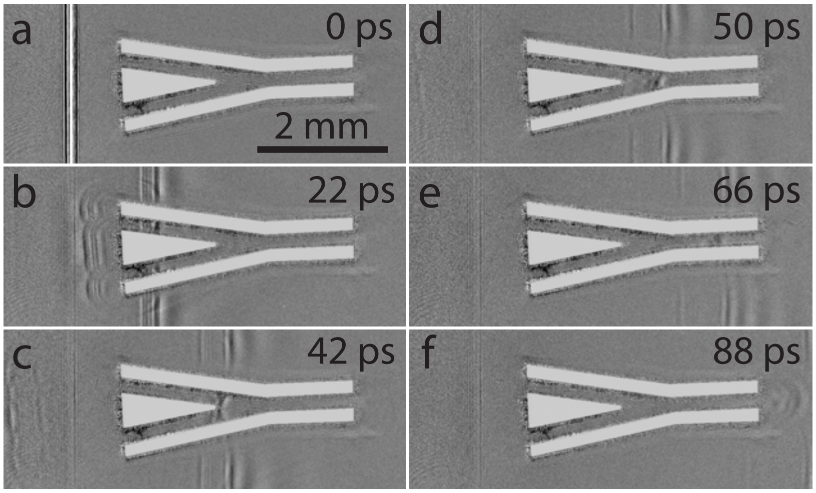

Figure 5.9 | Y-coupler. Panels (a)-(f) are frames from video 8 in [Werley 2012]. The uniform light gray region is an air-gap carved into a 50 μm slab of $LiNbO_3$. The rightward-propagating THz wave is initially guided down each arm of a "Y". When the waves meet in the middle they interfere, and are further guided down the stem of the Y. When the THz wave is emitted at the end it starts to diverge due to diffraction.

8. Dielectric slab waveguide propagation

As mentioned briefly in the experimental section, THz waves were guided in a slab of $LiNbO_3$ for most above measurements. The waveguiding was useful because it confined the THz wave to a thin region that could be imaged. The final experiment demonstrates how waveguiding behavior can be visualized and studied in the imaging setup. Before showing the experimental results, a quick review of the dielectric slab waveguide, also called a planar waveguide, will be presented. Unlike the material from the previous sections which would generally be covered in an introductory physics class, the material discussed in this section is slightly more advanced and would usually be covered in a graduate or advanced undergraduate optics class. In spite of the disparity in complexity, we discuss waveguiding here because the experiment provides a clear qualitative and quantitative visualization of waveguiding behavior which could be very helpful in a more advanced class. Below, we discuss the results and solutions to the dielectric waveguide but not their derivation, which can be found in many textbooks, e.g. [Cronin 1995].

A dielectric slab waveguide consists of a high-index core, such as $LiNbO_3$, surrounded by a low-index material, such as air. In a waveguide, the profile of the electric and magnetic field down the depth of the crystal [along y in Fig. 5.1(a)] does not change as the wave propagates. This is similar to an eigenmode in quantum mechanics, with time t from quantum mechanics replaced by the distance along the propagation direction x, and the solutions for the transverse electric (TE) mode of the dielectric slab waveguide are exactly the same as those for a particle in a box with finite potential walls. Discussion here will be limited to the TE modes, which have the electric field parallel to the slab surface and perpendicular to propagation direction, because they are the only modes generated in the experiment. For a specific propagation constant, $\beta = 2\pi/\lambda$, with λ the wavelength along the direction of propagation, there are multiple modes with different numbers of nodes. The 0^{th} mode has no nodes, the 1^{st} mode has 1 node, etc. Profiles of the first three modes in a 50 μm thick $LiNbO_3$ slab in air with $\beta = 100$ rad/mm (i.e. a wavelength $\lambda = 63$ μm) are shown in Fig. 5.10(a). The solutions are harmonic inside the high-index material (cosine for the even, symmetric modes and sine for the odd, antisymmetric modes). In contrast, the solutions are evanescent and exponentially decaying in the low-index material outside.

The wavelength inside the high-index material and the decay length in the low-index material depend on the frequency. Focusing only on one waveguide mode, one can see how the profile changes with frequency [see Fig. 5.10(b)]. At low frequencies, i.e. long wavelengths, the evanescent decay length is very long and most of the wave's energy is in the air, while at higher frequencies, i.e. shorter wavelengths, the evanescent field decays very quickly and most of the energy is localized in the $LiNbO_3$.

One common way to plot this frequency- and mode-dependent behavior is with a dispersion curve [see Fig. 5.10(c)]. A dispersion curve shows the frequency (or energy) as a function of wave vector. The two thin dashed lines show the dispersion curve for air, $\omega = c\beta$, and for bulk $LiNbO_3$, $\omega = c\beta/n$. The dispersion curve for the waveguide shows that at low frequencies, the wave behaves more like a wave in air because so much of the energy is in the evanescent field. At high frequencies, the wave behaves more like a wave in bulk $LiNbO_3$ because most of the energy is in the $LiNbO_3$ slab. The lowest mode extends all the way to zero frequency and wave vector. The higher modes, however, have cutoff frequencies below which there are no modes with the appropriate number of nodes bound in the slab.

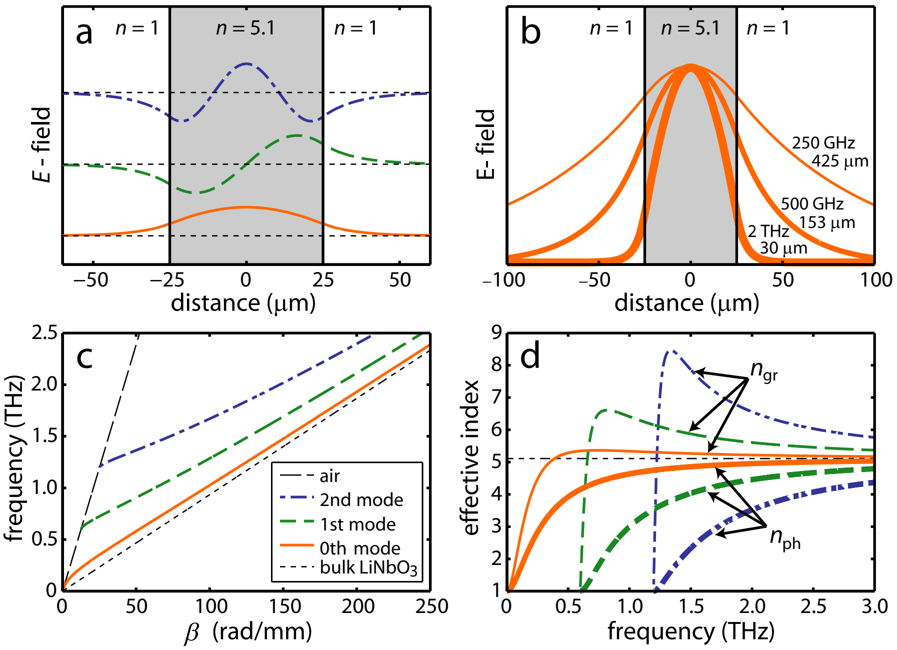

Figure 5.10 | Analytical waveguide solutions for TE modes in a dielectric slab waveguide. (a) shows the E-field profile of the first three modes at a specific propagation constant β = 100 rad/mm. (b) shows the E-field profiles for the lowest mode at three different frequencies. The wavelength along the propagation direction is indicated below the frequency. (c) shows the dispersion curves for the first three modes. (d) shows the effective group index (upper three, thinner curves) and effective phase index (lower three, thicker curves) for the waveguide.

Another way to display the same information is with the effective indices of refraction, both the phase index, n_{ph}, and the group index, n_{gr}. The phase index is defined as $n_{ph}(\omega) = c\beta(\omega)/\omega$. The velocity at which a single phase front moves, the phase velocity, is given simply by $v_{ph} = c/n_{ph}$, and the wavelength along the direction of propagation is given by $\lambda = 2\pi c/(\omega n_{ph})$. The lower three, thinner curves in Fig. 5.10(d) show the phase index as a function of frequency for the lowest three modes. n_{ph} transitions smoothly from the index of air to the index of bulk $LiNbO_3$. The velocity at which a pulse or wave-packet travels, the group velocity, is given by $v_{gr} = c/n_{gr} = d\omega/d\beta$. Although the group index starts close to 1 for each mode, it quickly increases until it is larger than the group index for the bulk material [see the upper three curves in Fig. 5.10(d)]. The large group index means that wave-packets of the appropriate frequency actually travel more slowly than in the bulk material. As the frequency increases and the slab becomes thick compared to the wavelength, both n_{gr} and n_{ph} approach the value for the bulk index of the dielectric — the group index from above and the phase index from below.

In the experiment described here, THz waves are excited in a plain, unstructured waveguide by focusing the pump beam to a line in the $LiNbO_3$ slab. This geometry excites a single-cycle THz wavepacket that contains many frequencies and the first three waveguide modes. Figure 5.11 shows frames from video 9 in [Werley 2012]. Initially two counter-propagating single-cycle waves are launched. As the waves propagate, the first three waveguide modes separate due to their differences in group velocity. The three modes are clearly visible as distinct wavepackets in Fig. 5.11(c) & (d). Looking just at the 0^{th} waveguide mode, it is easy to see the lower frequency, longer wavelength components travel faster than the high-frequency components, a consequence of the fact that the low-frequency, long-wavelength components are propagating mostly in air. This leads to the pulse broadening and becoming "chirped", i.e. having a time-dependent frequency. The waves reflect off the crystal edge after frame (d) in Fig. 5.11 and eventually cross, at which point the two counter-propagating waves form a standing wave [Fig. 5.11(f)]. Counter-propagation is a limiting case of Eq. 5.5 in Sec. 5.D.1 with $\theta = \pi$. The interference pattern that is formed by the intersecting beams is striking because the fringe spacing starts out relatively large and gradually becomes smaller as the shorter-wavelength components arrive at the region of intersection between the two beams.

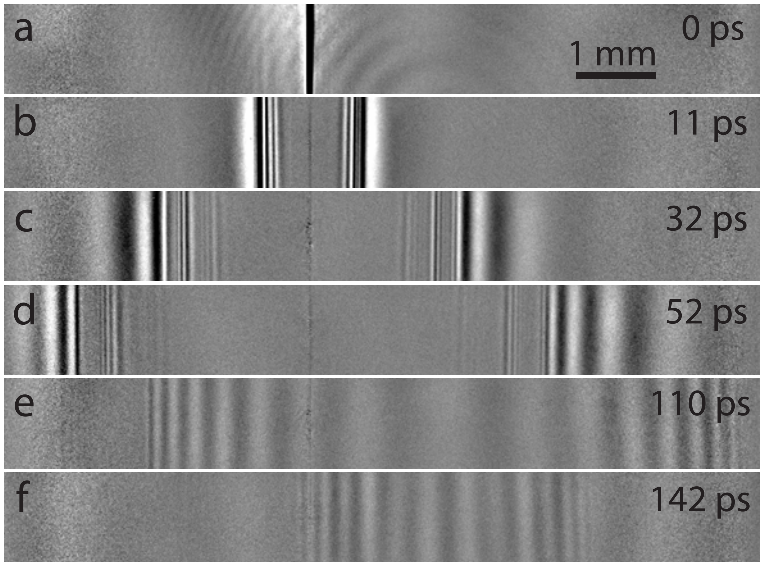

Figure 5.11 | Images of a broadband, guided THz wave. Panels (a)-(f) are frames from video 9 in [Werley 2012] depicting broadband THz waves propagating in an unstructured $LiNbO_3$ crystal. Counterpropagating waves are launched by a cylindrically focused “line” of pump pulse light in (a). The first three waveguide modes are clearly visible in (c) and (d). After frame (d) the wave reflects off of the crystal edges so the waves are propagating toward one another in (e), and they have begun to overlap and interfere in (f).

In every frame in video 9, the E-field is uniform along the vertical (z) direction. As a result, it is possible to compress each m x n picture into a 1 x n array by averaging over the vertical dimension with no information loss. The 1 x n array is the electric field profile as a function of horizontal position at one time delay. The compression is performed for each frame in the video, and each 1 x n array is placed into one row of a matrix, one above the other, in time order so that from bottom to top the field profiles at successively later times are stored. The resulting space-time plot shows the full evolution of the wave as it propagates [Fig. 5.12(a)]. In this single figure, it is possible to see the modes separate, the pulses begin to become chirped, and the reflection and interference of the waves. In this plot, the slope is directly proportional to the index: $m = \Delta t/\Delta x = 1/v = n/c$. The slope of the wavepacket is given by the group index, and the slope of the phase front is given by the phase index. It is clear from examining the figure that the group index of the second mode is higher than that of the first mode, but the phase index is lower. Taking the 2D Fourier transform of the matrix in Fig. 5.12(a) enables a quantitative comparison with analytical theory. The time dimension transforms to frequency and the spatial dimension transforms to the propagation constant, so the result of the 2D Fourier transform is a direct experimental determination of the dispersion curve. Figure 5.12(b) shows this matrix, which is in good agreement with the analytical theory shown in Fig. 5.10(c) [Wu 2009]. The direct observation of guided electromagnetic waves helps develop intuition for the formation of discrete modes and waveguide dispersion, effects common to optical fibers and many other waveguide geometries.

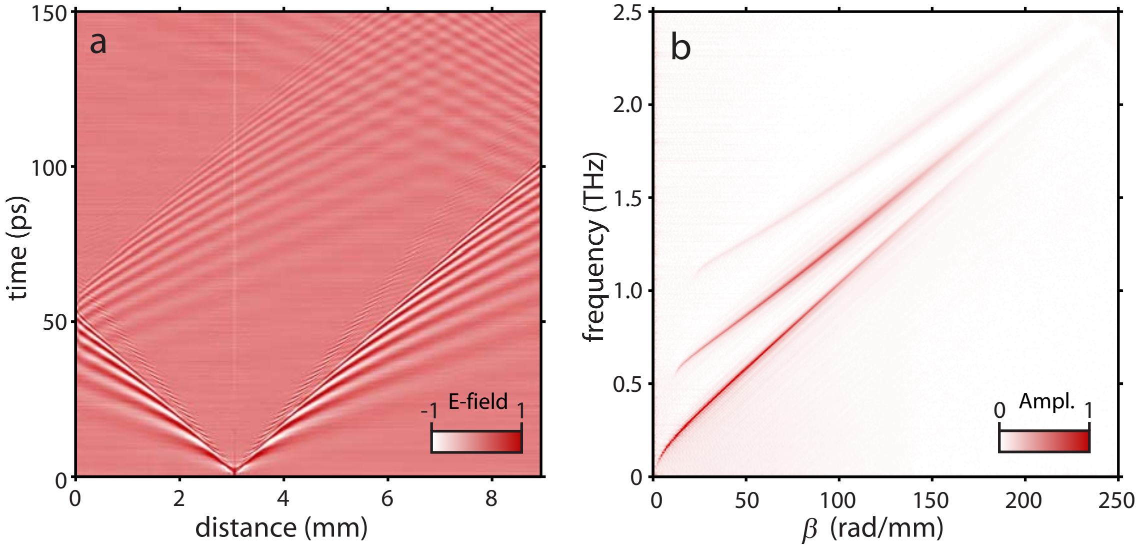

Figure 5.12 | 2D plots of a waveguided THz wave. (a) is a space-time plot showing the evolution of a broadband THz pulse in an unstructured slab waveguide. The 2D Fourier transform of (a) yields the experimentally determined dispersion curves shown in (b).

E. CONCLUSIONS

The experimental technique described above enables visualization of electromagnetic waves propagating at the speed of light. Videos of classic experiments help build intuition for wave propagation and fundamental effects like waveguiding, interference, and diffraction. Directly observing the phase fronts highlights the wave nature of light, and the time-resolved measurement solidifies builds understanding of group and phase velocities and their ramifications. We hope that this paper and its accompanying videos can help bring modern research into the classroom and provide a valuable tool in an introductory and intermediate physics curriculum. We also hope that other researchers in THz optics and spectroscopy will contribute additional figures and videos to a growing library of illustrations of wave mechanics principles. To facilitate the growth of such an online library, we would gladly add videos from other researchers (with appropriate attribution) to the collection already on our website.

F. REFERENCES

[Auston 1988] D. Auston and M. Nuss, “Electrooptic generation and detection of femtosecond electrical transients,” *IEEE J. Quantum Electron.* **24**, 184-197 (1988).

[Boyd 2008] R. W. Boyd, *Nonlinear Optics*, 3rd ed. (Elsevier, Burlington, 2008).

[Cronin 1995] N. J. Cronin, *Microwave and Optical Waveguides* (Institute of Physics Publisher, Philadelphia, 1995).

[Dougherty1992] T. P. Dougherty, G. P. Wiederrecht, and K. A. Nelson, “Impulsive stimulated Raman scattering experiments in the polariton regime,” *J. Opt. Soc. Am.* **9**, 2179-2189 (1992).

[Eichler 1986] H. J. Eichler, P. Günter, and D. W. Pohl, *Laser-Induced Dynamic Gratings* (Springer-Verlag, Berlin, 1986).

[Feurer 2002] T. Feurer, N. S. Stoyanov, D. W. Ward, and K. A. Nelson, "Direct visualization of the Gouy phase by focusing phonon polaritons" *Phys. Rev. Lett.* **88**, 257402 (2002).

[Feurer 2003] T. Feurer, J. C. Vaughan, and K. A. Nelson, “Spatiotemporal coherent control of lattice vibrational waves,” *Science* **299**, 374-377 (2003).

[Feurer 2007] T. Feurer, N. S. Stoyanov, D. W. Ward, J. C. Vaughan, E. R. Statz, and K. A. Nelson, “Terahertz polaritonics,” *Annu. Rev. Mater. Res.* **37**, 317-350 (2007).

[Fowles 1989] G. R. Fowles, *Introduction to Modern Optics*, 2nd ed. (Dover Publications, Mineola, 1989), Chap. 5.

[Fresnel Kirchhoff] We re-derived the Fresnel Kirchhoff formula following the derivation in [Fowles 1989], but in two dimensions instead of three because wave behavior in the waveguide follows 2-dimensional behavior. Making the additional assumption that plane waves are incident on the aperture, the Green’s function is $G(x,z) = -Y_0(2\pi\rho/\lambda)[x/\rho+1]$ with Y_0 a Bessel function of the second kind and $\rho^2 = x^2 + z^2$ the radius in two dimensions. The *E*-field can then be calculated by performing a numerical convolution of the aperture function with the Green’s function: $E(x,z,t) \propto [A(z) \otimes G(x,z)]\exp(-i\omega t)$.

[Grischkowsky 1990] D. Grischkowsky, S. Keiding, M. van Exter, and C. Fattinger, "Far-infrared time-domain spectroscopy with terahertz beams of dielectrics and semiconductors," *J. Opt. Soc. Am. B* **7**, 2006-2015 (1990).

[Guoy 1890] L. G. Gouy, “Sur une propriete nouvelle des ondes lumineuses” C. R. Acad. Sci. Paris **110**, 1251 (1890)

[Jiang 1999] Z. Jiang and X.-C. Zhang, “Terahertz Imaging via Electrooptic Effect,” *IEEE Trans. Microwave Theory Tech.* **47**, 2644-2650 (1999).

[Joannopoulos 2008] J. D. Joannopoulos, S. G. Johnson, J. N. Winn, and R. D. Meade, *Photonic Crystals, Molding the Flow of Light*, 2nd ed. (Princeton University Press, Princeton, 2008), Chap. 2.

[Koehl 1999] R. M. Koehl, S. Adachi, and K. A. Nelson, “Direct visualization of collective wavepacket dynamics,” *J. Phys. Chem. A* **103**, 10260-10267 (1999).

[Lin 2009] K.-H. Lin, C. A. Werley, and K. A. Nelson, “Generation of multicycle THz phonon-polariton waves in a planar waveguide by tilted optical pulse fronts,” *Appl. Phys. Lett.* **95**, 103304 (2009).

[Mack 2007] C. Mack, *Fundamental principles of optical lithography : the science of microfabrication* (John Wiley & Sons, Hoboken 2007).

[Palmer2005] C. Palmer, *Diffraction Grating Handbook*, 5th ed. (Newport, Rochester 2005).

[Peier 2008] P. Peier, S. Pilz, F. Müller, K. A. Nelson, and T. Feurer, "Analysis of phase contrast imaging of terahertz phonon-polaritons," *J. Opt. Soc. Am. B* **25**, B70-B75 (2008).

[Peier 2009] P. Peier, S. Pilz, T. Feurer, "Time-resolved coherent imaging of a THz multilayer response," *J. Opt. Soc. Am. B* **26**, 1649-1655 (2009).

[Peier 2010] P. Peier, H. Merbold, V. Pahinin, K. A. Nelson, and T. Feurer, "Imaging of THz waves in 2D photonic crystal structures embedded in a slab waveguide," *New J. Phys.* **12**, 013014 (2010).

[Resnik 2002] R. Resnik, D. Halliday, and K. S. Krane, *Phyiscs*, 5th ed. (John Wiley & Sons, New York, 2002).

[Shen 2003] Y. R. Shen, *The Principles of Nonlinear Optics* (John Wiley & Sons, Hoboken 2003).

[Sipos 2008] M. Sipos and B. G. Thompson, "Electrodynamics on a grid: The finite-difference time-domain method applied to optics and cloaking" *Am. J. Phys.* **76**, 464-469 (2008).

[Stoyanov 2002] N. S. Stoyanov, D. W. Ward, T. Feurer, and K. A. Nelson, "Terahertz polariton propagation in patterned materials," *Nat. Mater.* **1**, 95-98 (2002).

[Stoyanov 2004] N. S. Stoyanov, T. Feurer, D. W.Ward, E. R. Statz and K. A. Nelson, "Direct visualization of a polariton resonator in the THz regime," *Opt. Express* **12**, 2387-2396 (2004).

[Stoyanov APL 2002] N. S. Stoyanov, T. Feurer, D. W. Ward, K. A. Nelson, "Integrated diffractive THz elements," *Appl. Phys. Lett.* **82**, 674–676 (2002).

[Strikland 1985] D. Strickland and G. Mourou, "Compression of amplified chirped optical pulses," *Opt. Commun.* **56**, 219-221 (1985).

[Taflove 2005] *Computational Electrodynamics: The Finite-Difference Time-Domain Method*, 3rd ed., edited by A. Taflove and S. C. Hagness (Artech House, Boston, 2005).

[Ward 2003] D. W. Ward, E. R. Statz, N. S. Stoyanov, and K.A. Nelson, "Simulation of phonon-polariton propagation in ferroelectric $LiNbO_3$ crystals," in Engineered Porosity for Microphotonics and Plasmonics, MRS Symposium Proc. **762**, C11.60.1-6 (2003).

[Ward 2007] D. W. Ward, E. R. Statz, and K. A. Nelson, "Fabrication of polaritonic structures in $LiNbO_3$ and $LiTaO_3$ using femtosecond laser machining," *Appl. Phys. A* **86**, 49-54 (2007).

[Werley 2010] C. A. Werley, Q. Wu, K.-H. Lin, C. R. Tait, A. Dorn, and K. A. Nelson, "Comparison of phase-sensitive imaging techniques for studying terahertz waves in structured $LiNbO_3$," *J. Opt. Soc. Am. B* **27**, 2350-2359 (2010).

[Werley 2012] C. A. Werley, C. R. Tait, and K. A. Nelson, "Direct visualization of THz-frequency electromagnetic waves in classic experimental geometries," *Am. J. Phys.* **80**, 72-81 (2012).

[Wu 1996] Q. Wu, T. D. Hewitt, and X.-C. Zhang, "Two-dimensional electro-optic imaging of THz beams," *Appl. Phys. Lett.* **69**, 1026-1028 (1996).

[Wu 2009] Q. Wu, C. A. Werley, K.-H. Lin, A. Dorn, M. G. Bawendi, K. A. Nelson, "Quantitative phase contrast imaging of THz electric fields in a dielectric waveguide," *Opt. Express* **17**, 9219-9225 (2009).

Chapter VI

Time-resolved imaging of near-fields in THz antennas and direct quantitative measurement of field enhancements

Content from: C. A. Werley, K. Fan, A. C. Strikwerda, S. M. Teo, X. Zhang, R. D. Averitt, and K. A. Nelson, "Time-resolved imaging of near-fields in THz antennas and direct quantitative measurement of field enhancements," *Opt. Express* **20**, 8551-8567 (2012).

A. ABSTRACT

We investigate the interaction between terahertz waves and resonant antennas with sub-cycle temporal and $\lambda/100$ spatial resolution. Depositing antennas on a $LiNbO_3$ waveguide enables non-invasive electro-optic imaging, quantitative field characterization, and direct measurement of field enhancement (up to 40-fold). The spectral response is determined over a bandwidth spanning from DC across multiple resonances, and distinct behavior is observed in the near- and far-field. The scaling of enhancement and resonant frequency with gap size and antenna length agrees well with simulations.

B. INTRODUCTION

Antennas, which convert electromagnetic radiation into electrical currents and vice versa, have important applications spanning the electromagnetic spectrum from radio to visible frequencies. The technology is mature at radio and microwave frequencies where antennas are used in mobile telephones, television broadcasting, and many other applications. At these frequencies, study has focused on sending and receiving information [Balanis 2005] and therefore has primarily considered signals far from the internal antenna elements, i.e. in the far-field. More recently, studies at infrared and optical frequencies have focused instead on the antenna's ability to provide field enhancement and subwavelength field localization [Kinkhabwala 2009; Fromm 2006; Willets 2007; Cubukcu 2008; Muhlschlegel 2005; Ghenuche 2008; Ward 2010; Crozier 2003; Barnard 2011; Schnell 2010]. Antennas have been used in single-molecule fluorescence [Kinkhabwala 2009], surface enhanced Raman spectroscopy [Fromm 2006; Willets 2007], near-field scanning optical microscopy [Höppener 2008], photonic devices [Cubukcu 2008] and novel nonlinear optics [Muhlschlegel 2005; Ghenuche 2008; Ward 2010]. Although near-field signals assume primary interest in these studies, it has proved difficult to quantify electric field amplitudes and enhancements and characterize field profiles around the antenna because of the small length scales and a lack of methods for direct field visualization. It has also been difficult to determine the full spectral response because of the ultra-broad bandwidth required: an antenna resonant in the visible displays important behavior spanning from DC to ultraviolet frequencies. Because of the experimental challenges, most of the understanding of near-field behavior in antennas has relied on simulations that have only been experimentally verified in the far-field. Comparing near-field simulations with experiments is complicated by sample inhomogeneity, surface roughness, and chemical adsorbates; even the ideal case presents challenges because modeling tightly localized fields requires a very fine mesh, while capturing the far-field behavior needs a large simulation volume. In short, quantitative measurement of near-field behavior is critical for validating theoretical methods that will be used in the design of future devices and experiments.

Many experimental studies at THz, infrared, and visible frequencies have tried to characterize fields around resonant antennas [Cubukcu 2008; Muhlschlegel 2005; Ghenuche 2008; Ward 2010; Crozier 2003; Fischer 2008; Bitzer 2010; Blanchard 2011; Barnard 2011; Schnell 2010], but results have not been definitive. Scanning near-field probes yielded high-resolution field maps [Cubukcu 2008; Imura 2006; Schnell 2010], but there is strong evidence that scanning tips significantly perturb the fields being measured [Imura 2006; Mujumdar 2007; Wang 2004]. Nonlinear microscopy at visible frequencies [Ghenuche 2008] and electro-optic probing at THz frequencies with sub-wavelength resolution [Bitzer 2010; Blanchard 2011] made it possible to map antenna mode structure without seriously perturbing the system, but the tightly localized, strongly enhanced fields near antenna ends and gaps were not spatially resolved or directly quantified. Recently, photocurrents in an optical antenna were measured [Barnard 2011], but volume-averaging of fields prevents the quantification of near-fields with subwavelength resolution. Another recent study determined field strengths in angstrom-sized gaps between electrodes using tunneling currents [Ward 2010], but it is difficult to infer antenna behavior from the atomic-scale gaps in these irregularly-shaped structures. We use a THz polaritonics platform [Feurer 2007] to study antennas without perturbing the system or compromising on spatial resolution.

The polaritonics platform allows versatile THz generation [Feurer 2003; Lin 2009], control [Stoyanov 2002; Peier 2009], and time-resolved imaging of THz electric fields [Koehl 1999; Peier 2008; Wu 2009; Werley 2010]. Several favorable considerations enable particularly incisive study of antennas and other metallic microstructures using this system. First, because the THz wave propagates in an electro-optic medium, it is possible to directly record its electric field (*E*-field) and thus quantify field enhancement, a capability that has proved difficult in previous experiments. Second, the *E*-field is imaged with an optical probe without the need for a tip or other intrusive element. The spatial resolution is ultimately set by the optical diffraction limit [Wu 2009; Werley 2010], which is hundreds of times smaller than the THz wavelength. The temporal resolution is set by the duration of the femtosecond optical probe pulse, which is a small fraction of the THz period. Third, the THz antennas can be deposited directly onto the THz propagation medium with high precision using optical lithography. Finally, facile control over the THz field allows us to initially probe the antennas with broadband THz waves in order to characterize the frequency-dependent response from 0.1-2 THz and survey for resonances. We subsequently drive the antennas with multi-cycle THz waves tuned to the antenna resonance frequencies in order to study the resonant responses in detail and maximize field enhancement. Because of the versatility of the THz polaritonics system and its ability to spatially, temporally, and spectrally resolve the antenna response, it is possible to develop a detailed and scalable understanding of antenna behavior.

C. RESULTS AND DISCUSSION

The experimental geometry is shown in Fig. 6.1. An 800 nm, 80-fs pump pulse passes through a thin $LiNbO_3$ slab where it generates a THz response via impulsive stimulated Raman scattering [Dougherty 1992]. The THz field is wave-guided within the high-index $LiNbO_3$ slab, propagating from the generation region to a gold micro-antenna deposited on the crystal surface. The antenna was a pair of half-wave strips aligned end-to-end with a small gap between them, as shown in Fig. 6.2(a). Strips were chosen because they generate larger enhancements than other geometries [Crozier 2003; Fischer 2008], and the ends were given a smooth, 5 μm radius of curvature so that fields in the gap would be relatively uniform and undistorted by sharp tip effects. The antennas were 150 nm thick gold deposited with optical lithography (see Sec. E). All antennas studied were 10 μm wide, with lengths ranging from 30 to 110 μm and gaps ranging from 2 to 20 μm. Figure 6.2(b) shows an image of a typical antenna. The white spot in the gap is the reflection of a tightly focused optical beam (see Sec. G, Fig. 6.8), which was used for recording THz transients in the gap.

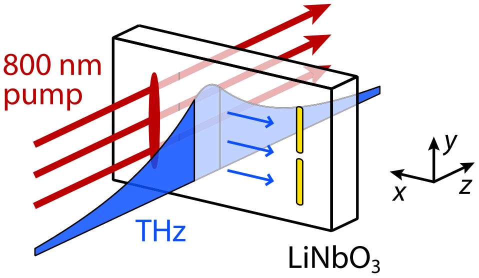

Figure 6.1 | The experimental pumping geometry. A THz wave is generated by an ultrafast optical pump pulse and guided down the 30 μm thick $LiNbO_3$ slab until its evanescent field interacts with the gold antenna deposited on the surface.

When driven on resonance, current flow in a half-wave antenna results in a standing wave charge distribution with maxima at the ends of each antenna arm. The antenna pair acts like a small capacitor [Kang 2009], and the *E*-field, represented by green lines in Fig. 6.2(a), is confined and enhanced in the subwavelength gap. To study field localization and mode structure, we excited the antenna on resonance with a multi-cycle THz wave [Lin 2009] and recorded movies of the wave as it interacted with the antenna using time-resolved, phase-sensitive imaging techniques [Wu 2009; Werley 2010] as illustrated in Sec. F, Fig. 6.7. Figure 6.2(c)-(e) are frames from such movies, which show the *E*-field of the propagating THz wave. Figure 6.2(c) (see Media 1 [Werley 2012]) shows a zoomed-out view of the THz wave tuned to 255 GHz, the frequency of maximum enhancement (see below), as it interacts with the antenna at image center. Figure 6.2(d) is zoomed in on the antenna. Field enhancement at the antenna ends and even greater enhancement in the gap center is visible, in agreement with the qualitative depiction in Fig. 6.2(a). Figure 6.2(e) (from Media 2 [Werley 2012]) shows the more complex $\ell = 3\lambda/2$ antenna mode. The left side of the image shows the field emitted from the antenna just after the rightward propagating THz wave has passed. Figure 6.2(f) shows the predicted wave pattern emitted from a pair of

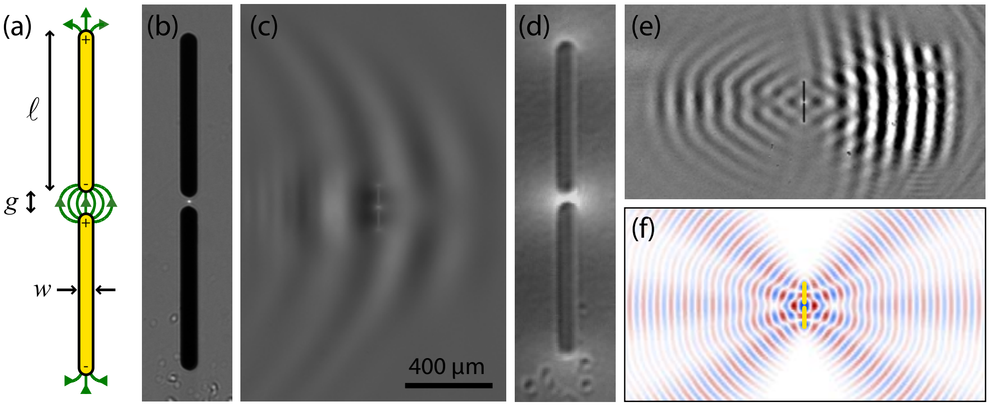

Figure 6.2 | Antenna design and images of THz fields. (a)-(f) w = 10 µm and g = 5 µm. In (b)-(d) $\ell = 94\,\mu\text{m}$ and in (e) and (f) $\ell = 110\,\mu\text{m}$. (a) The charge distribution and electric field lines (in green) for the lowest antenna mode. (b) An optical image of the antenna. (c) An experimental image of the full E-field profile at one moment in time as a rightward-propagating, resonant THz wave interacts with the antenna in (b). (d) A magnified view of (c) with better than 2 µm resolution. (e) The field profile soon after a rightward-propagating THz wave at 760 GHz (the second antenna resonance) has passed the antenna. (f) The theoretical wave pattern emitted from a pair of $\ell = 3\lambda/2$ wires in free space.

$3\lambda/2$ wire antennas in a homogeneous dielectric oriented and separated to match experiment [Balanis 2005]. The radiation pattern emitted from each antenna arm has two horizontal nodes, and interference generates the observed E-field profile.

In order to perform spectroscopy on the antennas, we switched from imaging the THz waves to localized detection, which provides a higher signal-to-noise ratio (S/N). We focused the probe beam to a 1 µm diameter spot which could be accurately positioned relative to the sample (see Sec. G) to enable direct probing of the field in the antenna gap. Spectroscopy was performed both in transmission and reflection geometries. Transmission yielded excellent S/N but measured the THz E-field averaged through the 30-µm crystal thickness. In contrast, reflection yielded weaker signals but quantitatively measured the E-field amplitude directly at the $LiNbO_3$ surface on which the antenna was deposited (see Sec. G).

To find the resonances of the antenna, the structure was interrogated with a broadband THz wave and detected in the optical transmission geometry. We recorded the time-resolved THz transients [Fig. 6.3(a)] at three locations as shown in the inset of Fig. 6.3(b): a reference trace far removed from the antenna, a point soon after the THz wave had propagated past the antenna (similar to a typical THz transmission measurement), and directly in the antenna gap as seen in Fig. 6.2(b). The THz amplitude spectra, calculated by taking the Fourier transforms of the time domain signals, are shown in Fig. 6.3(b), and the transmission spectra, the ratios of signal and reference, are shown in Fig. 6.3(c). The spectrum measured after the antenna (orange) shows clear dips, corresponding to transmission minima resulting from resonant modes of the antenna. The two lowest modes ($\ell = \lambda/2$ at 350 GHz and $\ell = 3\lambda/2$ at 860 GHz) are clearly visible. The ratio between these frequencies would be exactly three for an extremely thin wire in a

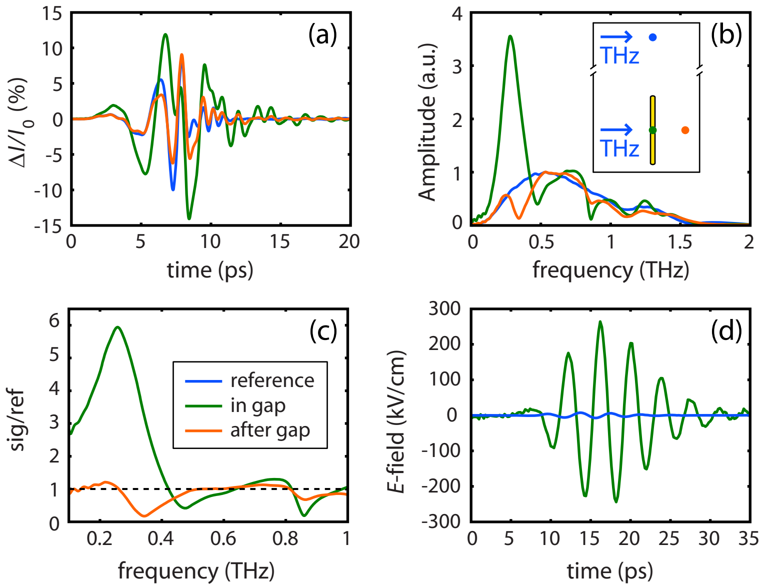

Figure 6.3 | Time traces and spectra. (a) Time traces of a broadband THz wave measured far above the antenna (blue spot in the inset of (b), the reference field location); ~100 μm after the antenna (orange spot); and directly in the 5 μm antenna gap (green spot). (b) The spectral amplitude of each trace calculated from the Fourier transform of data in (a). (c) The ratio of signal and reference traces shown in (b). (d) The time trace for a multi-cycle reference THz wave (blue) and the enhanced field measured in a 2 μm antenna gap (green).

homogeneous dielectric [Balanis 2005], but it is significantly lower here because of the complex geometry (a finite width antenna deposited on a high-index planar waveguide substrate [Lin 2009; Werley 2010]).

Behavior in the antenna gap [green traces in Fig. 6.3] is very different from that measured after the antenna. The amplitude spectrum in Fig. 6.3(c) reveals a clear and strong enhancement at 255 GHz. The frequency shift between the in-gap enhancement maximum and the maximum scattering frequency (the transmission minimum) is not unexpected, as such a shift has been previously predicted [Messinger 1981] and observed [Ghenuche 2008] in similar structures. Other notable features of the in-gap spectrum are enhancement down to DC and an amplitude reduction at a higher frequency than the transmission minimum (~490 GHz). The DC enhancement, amplitude reduction, and in-gap enhancement maximum are also present in other antennas with different arm lengths and gap sizes (see Fig. 6.4 & 6.5), showing that these are general behaviors for half-wave antennas. They can be understood qualitatively by thinking of the antenna as a damped driven harmonic oscillator, which also has an amplitude maximum that occurs at a lower frequency than the resonant frequency, a DC response that depends only on the resonant frequency and oscillator strength, and a frequency-dependent phase shift (see Sec. H). Above the resonant frequency, the phase shift leads to destructive interference between the driving field and induced response, which is evident in our results as an enhancement minimum with a value less than unity.

To determine the effects of antenna geometry on the antenna response (enhancement, resonant frequency, etc.), we fabricated antennas with different arm lengths and gap sizes. The amplitude ratios between reference traces and scans recorded in and after the gap, like those shown in Fig. 6.3(c), are shown for each antenna in Fig. 6.4. In Fig. 6.4(a) & (c), which vary the gap size but keep the antenna arm length constant, the DC offset and the maxima, minima are clearly visible for the first two modes of each antenna. Other than the magnitude of the near-field response, the spectra are nearly independent of gap size. In contrast, the resonant frequencies, linewidths, and peak response depend strongly on antenna length [Figs. 6.4(b) & (d)]. In the 30 μm long antenna [dark green in Fig. 6.4(b)], for instance, the enhancement is weak and nearly constant for frequencies below ~750 GHz. The 110 μm long antenna [dark purple in Fig. 6.4(b)], in contrast, is sharply peaked at ~200 GHz and shows much larger enhancements.

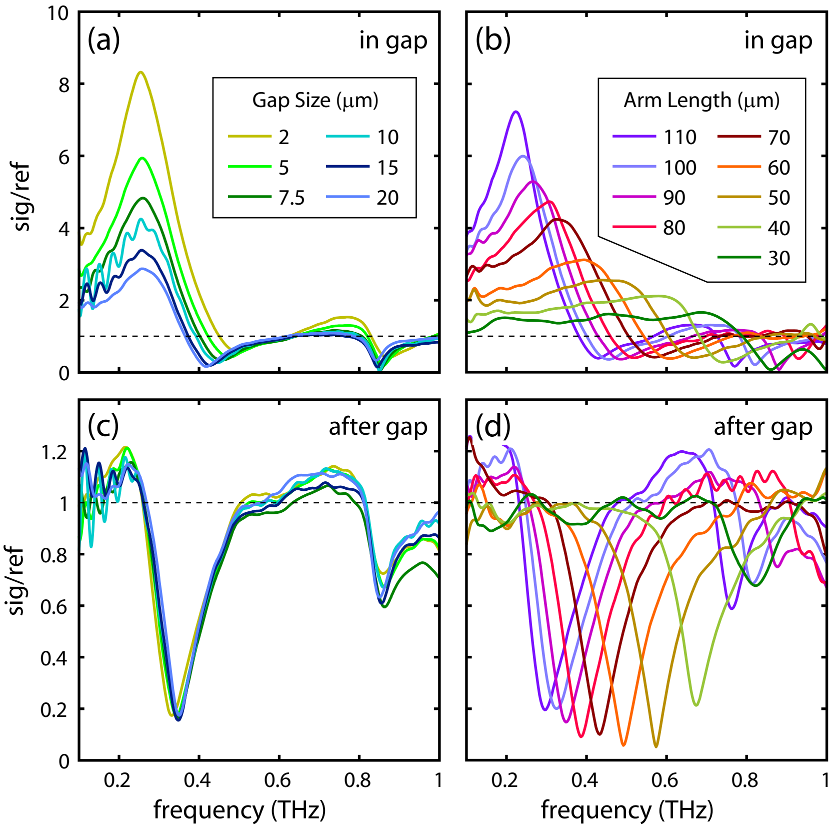

Figure 6.4 | Transmission and enhancement spectra. In-gap enhancement spectra [green dot in Fig. 6.3(b)] for various antennas are shown in (a) and (b), and transmission spectra [orange dot in Fig. 6.3(b)] are shown in (c) and (d). (a) and (c) Spectra for a fixed arm length $\ell = 94\,\mu\text{m}$ for different gap sizes. (b) and (d) Spectra for a fixed gap size $g = 5$ μm for different arm lengths.

To understand the scaling of antenna properties with arm length and gap size, we extracted the frequency at which minimum transmission occurred and the frequency at which maximum enhancement occurred from the traces in Fig. 6.4. In addition, traces like those in Fig. 6.3(d) were recorded in reflection for each antenna to quantify the peak field enhancement. Finite difference time domain (FDTD) simulations were performed using the commercial software CST Microwave Studio to corroborate experimental results. Experimental results (solid symbols), simulation results (dashed orange and solid green lines), and simple trend lines (dotted and dot-dashed black) are collected in Fig. 6.5.

Figure 6.5(a) shows the frequency of minimum transmission and maximum enhancement for the $\lambda/2$ (lower set of data) and the $3\lambda/2$ (upper set of data) modes. Experiment and simulation show these frequencies are nearly independent of gap size. This lack of dependence indicates that the length of a single antenna arm determines the resonant frequency, and coupling between antenna arms is weak at these gap sizes. Figure 6.5(b) shows the frequencies as a function of arm length ℓ for both the $\lambda/2$ and $3\lambda/2$ modes. Large frequency shifts between near- and far-field resonances are clearly visible for all gap sizes and antenna lengths. Experiment and simulation agree that the resonant frequency increases sharply for shorter lengths. Microwave theory predicts that antennas in free space will have maximum scattering at a frequency where the arm length is an integer number of half the wavelengths, *m*, with *m* odd [Balanis 2005]. We fit the data with a simple one-parameter model: $f_{\mathrm{res}} = c/(n_{\mathrm{eff}}\lambda_{\mathrm{res}}) = mc/(2n_{\mathrm{eff}}\ell)$, where an effective index, n_{eff}, is used to account for the complex sample geometry of an antenna deposited on a subwavelength, high-index substrate. Although it was not possible to fit the peak scattering frequency with a single n_{eff} as was expected, the simple $1/\ell$ fit agrees reasonably well with the frequency of peak enhancement for both the $\lambda/2$ and $3\lambda/2$ modes; the dotted black line in Fig. 6.5(a) & (b) is the resonant frequency predicted for $n_{\mathrm{eff}} = 6.4$, an index higher than the substrate. Even though the theory of microwave antennas is quite mature, it is still important to do careful experiments and modeling to fully understand the behavior of samples with non-ideal geometries.

Figure 6.5(c) shows the peak amplitude enhancement (measured in the frequency domain) as a function of gap size. As the gap is decreased, the enhancement increases rapidly. As was discussed previously, at a given instant the antenna can be thought of as a capacitor, where the charge driven onto the ends of the antenna arms by the incoming THz wave is independent of gap size. For smaller gaps, the same voltage drop occurs over a smaller distance, resulting in an increased *E*-field. Experiment, simulation, and intuitive modeling all indicate that even larger field enhancements would be present in gaps smaller than the ones measured here, consistent with THz enhancements estimated in nano-slits [Seo 2008; Seo 2009]. The black dot-dashed line shows the arm length divided by the gap size, ℓ/g, with no fit parameters. This can serve as a rule of thumb for predicting antenna enhancement.

Figure 6.5(d) shows the enhancement in the gap as a function of arm length when each antenna was driven by a multi-cycle THz pulse tuned to its frequency of maximum enhancement. Small imperfections in the antenna structures and variations in the multicycle THz waveforms at the different frequencies as well as limited S/N in the reflection measurements all contributed to scatter in the plot, but the results and those of simulations and transmission measurements [Fig. 6.4(b)] all indicate that, for a given gap size, the peak enhancement increases approximately linearly with length. This results from the fact that electrons from the entire length of the antenna are concentrated at the ends when excited on resonance, leading to more charge, and thus a larger field, in longer antennas. Again, the black dot-dashed line shows ℓ/g.

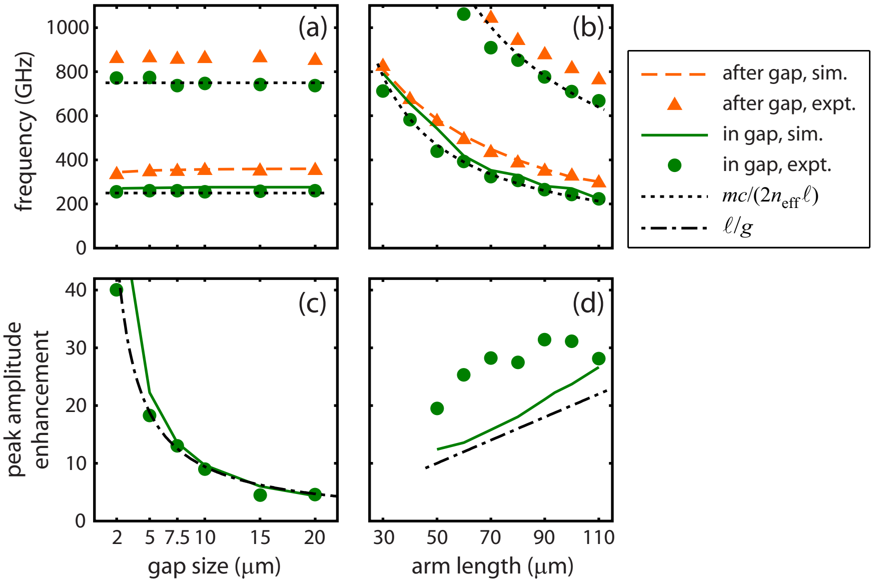

Figure 6.5 | Resonant frequency and enhancement trends. Trends as a function of gap size for a fixed arm length of 94 μm [(a) and (c)] and as a function of arm length for a fixed gap size of 5 μm [(b) and (d)]. (a) and (b) show the frequency at which the peak enhancement occurred (green) and the frequency of the antenna for which the transmission was minimized (orange).The upper set of data in (a) and (b) (circles and triangles) corresponds to the $3\lambda/2$ antenna mode and the lower data corresponds to the $\lambda/2$ mode. (c) and (d) show the maximum amplitude enhancement in the antenna gap. Green circles are experimental measurements in the gap and orange triangles are experimental measurements after the gap. The solid green (in gap) and dashed orange (after gap) lines are the results of FDTD simulations with the same gap sizes and arm lengths as in the experiments.

D. CONCLUSIONS

Our measurements have revealed key antenna features including the near-field spectral response, magnitude of the maximum field enhancement, and the time-dependent near-field profile in unprecedented detail. This will facilitate antenna design at all frequency ranges. Our methodology can also be applied to complex structures including metamaterials whose unique capabilities are based on near-field profiles and enhancements within antenna-like elements. In addition, enhancements of almost 40-fold and greater enhancements in future antennas with smaller gaps will enable new nonlinear THz phenomena to be explored. In spite of the fact that the enhanced field is tightly localized, the ability to probe exclusively in the gap region lets us take full advantage of the enhancement. High THz peak fields have been generated by other tabletop, ultrafast methods [Bartel 2005; Yeh 2007; Hirori 2011], but the THz output is typically a single-cycle pulse that has its energy distributed over a broad spectral range. The THz transients measured in our smallest antenna gaps have greater than 500 kV/cm peak-to-peak field strengths, comparable to those in strong single-cycle pulses. Our antenna-enhanced multicycle transients, however, have their energy concentrated in a much narrower spectral range, with spectral brightness exceeding that of the most intense broadband THz pulses. This spectral brightness makes it possible to strongly drive selected resonant transitions (electronic, vibrational, or rotational) for waveguide-based nonlinear spectroscopy and coherent control, while avoiding unwanted responses driven by extraneous frequency components. In the future, substantially larger *E*-fields could be produced by driving the antennas with stronger THz waves produced simply with additional optical pump pulse energy or with intensity-modulated optical pump waveforms tailored for efficient generation of multi-cycle THz pulses.

E. SAMPLE PREPARATION

The antennas, patterned by direct laser writing with a DWL 66 (Heidelberg Instruments), were deposited on a 30-μm $LiNbO_3$ substrate. First, a 1.4 μm thick layer of photoresist (AZ5214 E from Clariant Co.) was spin-coated onto the $LiNbO_3$. The sample was then scanned in the plane of the laser focus with the photoresist exposure calibrated appropriately for resist thickness and substrate. This allowed programmable patterning with ~1 μm resolution. After development and removal of the exposed photoresist [Fig. 6.6(a)], a 10 nm adhesion layer of chromium followed by a 150 nm layer of gold was evaporated on the photoresist and the patterned areas [Fig. 6.6(b)]. The final structure was achieved through a lift-off process [Fig. 6.6(c)]. Due to the negative wall profiles of patterned photoresist, the gold on AZ5214 E can be gently rinsed with acetone and removed without an ultrasonic process that could damage or crack the thin slab.

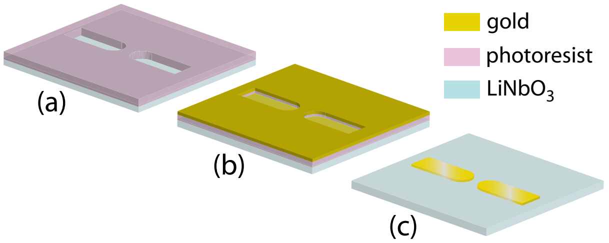

Figure 6.6 | Sample fabrication procedure. The optical lithography process for depositing 150 nm thick gold antennas on a 30 μm thick $LiNbO_3$ slab.

F. ULTRAFAST PHASE SENSITIVE IMAGING

All experiments studying the interaction between THz waves and the gold antennas used ultrafast methods. Both the imaging experiments (Figs. 6.2 & 6.7) and single-point spectroscopy measurements (Figs. 6.3, 6.4, 6.5, and 6.8) were conducted with an amplified titanium-sapphire laser with a repetition rate of 1 kHz and a pulse duration of 100 fs. A beamsplitter separated pump and probe pulses, and the pump was delayed relative to the probe using a mechanical delay stage.

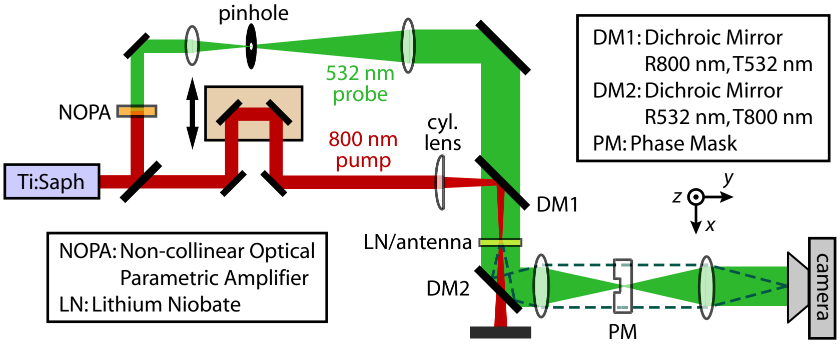

Figure 6.7 | Experimental geometry for imaging. The experimental geometry for time-resolved, phase contrast imaging of THz waves in LN sample.

To study the antenna mode profile, we used time-resolved, phase-sensitive imaging [Koehl 1999; Peier 2008; Wu 2009; Werley 2010] as discussed more fully in sections 4C & 4D. The pump generates a THz wave that is wave-guided down the thin $LiNbO_3$ slab. The THz electric field changes the $LiNbO_3$ index of refraction via the electro-optic effect. A spatially expanded probe beam is routed through the sample (see Fig. 6.7), where it accumulates a spatially varying phase shift due to the THz-induced change in index: $\Delta\phi(x,y) \propto \Delta n(x,y) \propto E_{\mathrm{THz}}(x,y)$. Because the camera detects intensity, a phase-sensitive imaging technique is required to observe the phase shift. We used phase contrast imaging [Wu 2009; Werley 2010]. In the back focal plane of the first lens after the sample, i.e. the Fourier plane, the light diffracted off the THz wave is spatially separated from the zero-order beam. In that plane, a mask (a flat plate with a 25 μm x 25 μm recessed region) introduces a $\pi/2$ phase shift between the zero-order and diffracted light, and when the light recombines at the image plane it interferes to make an amplitude image which is detected by the camera. By changing the time delay between pump and probe, a series of images were recorded and compiled into a video clip showing the propagation of the THz wave (see Media 1 and 2 in [Werley 2012]).

We used a non-collinear optical parametric amplifier (NOPA) to generate 100 fs probe pulses at 532 nm. A spectral bandpass filter before the camera transmitted the 532 nm probe light but blocked the 800 nm pump light and 400 nm light resulting from second harmonic generation in the $LiNbO_3$ sample. To facilitate tight focusing of the zero-order probe beam at the phase mask and high resolution imaging, the first lens after the $LiNbO_3$ sample was an aplanatic lens with a numerical aperture of 0.15, resulting in an image resolution better than 2 μm.

G. SPECTROSCOPY IN REFLECTION AND TRANSMISSION

In order to perform careful spectroscopy of the antennas, we switched from imaging to point source detection, which greatly enhanced the signal to noise ratio (S/N) and enabled reflection measurements. To generate broadband THz waves for spectroscopy, such as those shown in Figs. 6.3(a)-(c), the pump was focused to a line (~1 cm tall, 30 μm wide, 100 μJ pulse energy) in the lithium niobate sample [Wu 2009]. For quantitative measurement of *E*-field and antenna enhancement, multi-cycle, "narrowband" THz waves tuned to the antenna resonance were preferred over broadband THz waves [see Fig. 6.3(d)]. Pumping the antennas on resonance increased the peak *E*-field enhancement and the spectral amplitude of the signal for improved S/N in reflection mode measurements. Because of waveguide dispersion, different THz frequencies propagate at very different velocities, so a tilted intensity front in the optical pump pulse can be used to velocity match, and thus coherently amplify, a selected THz frequency [Lin 2009]. The tilted pump pulse for narrowband THz waves was prepared by imaging a diffraction grating onto the sample (see Fig. 6.8), and the tilt angle was controlled by changing the magnification of the imaging system. We used 1 mJ pump pulse energy with a spot size at the sample of ~5 mm wide and 300 μm tall. The two pumping geometries complement each other in that the broadband THz pulse can be used to scan a wide frequency range and determine the resonant frequencies of the first several antenna modes in a single measurement. The narrowband pulse can then be tuned to the resonant frequency of the antenna to measure field enhancement and characterize the resonant response.

To fully characterize the THz fields, it was necessary to focus the probe tightly enough that it could fit into a 2 μm antenna gap and to position the probe relative to the antenna with sub-micrometer accuracy. It also was beneficial to make both optical transmission measurements, which have the excellent S/N required for broadband spectroscopy, and optical reflection measurements, which enable quantitative determination of THz *E*-fields at the surface of the lithium niobate crystal where the gold antenna was deposited. This was important because the THz field is not uniform throughout the LN thickness for two reasons. First, the LN slab acts like a planar THz waveguide, and while the lowest-order waveguide mode that predominates far from any antenna structure is fairly uniform as a function of depth in the crystal, the higher-order modes that are weakly excited are not. Second, in the antenna regions where THz field features vary sharply on micron length scales, the variation is sharp as a function of depth as well as lateral position, so averaging over the LN crystal thickness is particularly inaccurate for quantitative characterization of the field profile and enhancement in the antenna gap. The experimental setup in Fig. 6.8 enables tight focusing and high-precision positioning of the probe beam as well as switching between transmission and reflection modes without adjusting the sample or any beam-steering optics.

In this setup, 800 nm light in the probe arm is frequency-doubled to 400 nm in a β-BBO crystal. The 400 nm light is spatially filtered and expanded to make a Gaussian beam with a diameter designed to back-fill the objective used to focus the probe onto the sample. To prevent sample damage by the tightly focused probe, the beam is attenuated to less than 1 nJ/pulse. After a polarizer, which is used to control the polarization state in transmission and reflection measurements, 20% of the probe beam is reflected off the beamsplitter (BS2). This light is then transmitted through a dichroic mirror and focused onto the sample with an 0.28 numerical aperture, 3.3 cm working distance objective lens, resulting in a 1 μm probe spot at the lithium niobate surface.

In addition to the 400 nm probe light, incoherent, 530 nm light from a light emitting diode (LED) is incident on the sample from the back side. The sample is imaged onto the camera, and the objective used to focus the probe is the first lens in this imaging system, yielding 1 μm imaging resolution. In addition to the incoherent light being imaged onto the camera, a small fraction of 400 nm probe light reflected from the sample surface (~10%) is also reflected from the 532 nm reflecting dichroic mirror and imaged onto the camera. The image on the camera shows the antenna structure and the 1 μm probe spot. The sample is mounted on a 3-dimensional translation stage with differential micrometers, allowing positioning relative to the probe spot with ~200 nm accuracy. Figure 6.2(b) shows the probe accurately positioned in the middle of a 5 μm antenna gap. After the sample was correctly positioned, the LED was shut off to prevent extraneous light from hitting the photodiodes during pump-probe measurements.

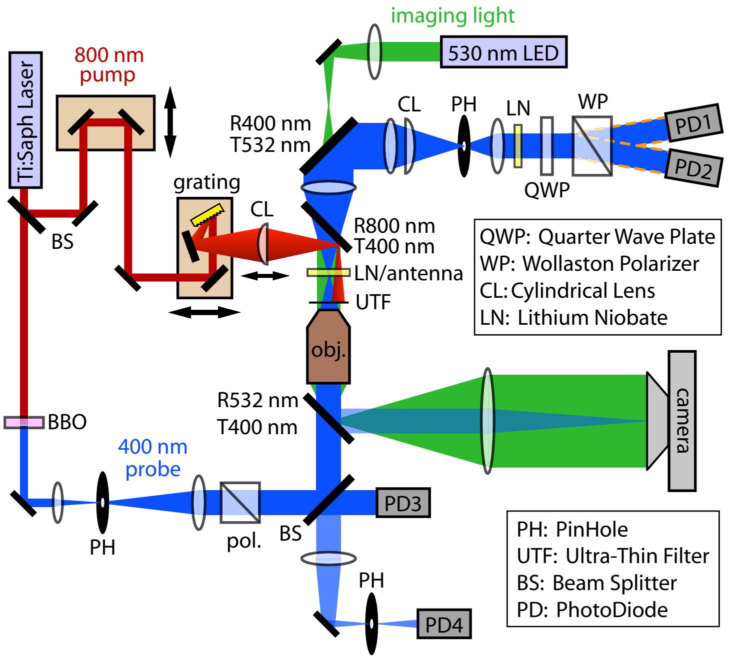

Figure 6.8 | The experimental geometry for point-source detection. The sample is imaged onto the camera using incoherent 530 nm light from an LED source. A small fraction of the 400 nm probe light also reaches the camera, allowing simultaneous visualization of antenna and probe spot. The sample is mounted on a 3D translation stage with differential micrometers, making it possible to position the sample with ~200 nm accuracy relative to the probe spot which is focused to 1 μm with a 0.28 numerical aperture microscope objective [see Fig. 6.2(b)]. After the sample is correctly positioned, the THz transients can be measured in transmission or reflection mode. The intensity front of the optical pump pulse is tilted by imaging a diffraction grating onto the sample.

After correctly positioning the pump, probe, and sample, optical transmission or reflection mode measurements could be made without changing alignment. For both transmission and reflection measurements, the pump was chopped at 500 Hz and the signal was detected at the modulation frequency using a data acquisition (DAQ) card [Werley 2011]. Balancing was implemented between photodiodes PD1 and PD2 for transmission and between PD3 and PD4 for reflection (see Fig. 6.8). The overall minimum signal detection threshold was $\Delta I / I_0 \sim 10^{-4}$.

Transmission was used for broadband THz measurements because of the excellent S/N. To correctly implement transmission measurements, it is necessary to control the beam collimation. After passing through the sample, the transmitted probe beam is rapidly diverging. Before being collimated, it passes through a R800 nm/T400 nm dichroic mirror angled at 45°, which changes the apparent location of the focus for divergence along the horizontal and vertical directions. A cylindrical lens (see Fig. 6.8) corrects for this, and makes it possible to focus the probe to ~100 μm and later to collimate the beam. The pinhole was necessary to block 400 nm light that was generated through second harmonic generation of the pump light in the $LiNbO_3$ sample, which could not be rejected by a spectral filter like the 800 nm pump light.

In transmission mode, the THz signal was encoded as a change in the polarization state of the probe. The probe polarization was initially set to 45º, and the THz induced a change in the relative index between vertically and horizontally polarized light in the $LiNbO_3$. The change in relative index was encoded as a phase shift, $\Delta\varphi$, between the horizontal and vertical components. The final quarter wave plate (which changes the 45º polarized light to circularly polarized light) and Wollaston polarizer convert the phase shift to an amplitude shift. The signal detected at the photodiodes is (for ideal optics):

$$I = I_0 (1 \pm \sin \Delta\phi) \tag{6.1}$$

where the "+" corresponds to PD1 and the "–" to PD2. Subtracting the output of the two diodes doubles the signal and suppresses common mode noise. The second slab of $LiNbO_3$, located just before the quarter wave plate, is rotated 90º relative to the sample to correct for the inherent, static birefringence of the $LiNbO_3$ sample. Additional adjustment is needed to correct for the depolarizing effects of the beamsplitter and three dichroic mirrors between the two polarizers that reflect or transmit the probe. These optics affect the relative amplitudes of vertical and horizontal polarization components and introduce a static phase shift between these components. Jones matrix analysis [Yariv 2007] shows that the amplitude change can be corrected by rotating the input polarizer away from 45º, and the static phase shift can be corrected by tilting the quarter wave plate so that it induces a phase shift greater than 90º. After making these corrections, we observe a nearly ideal polarization state after the quarter wave plate, and can use supplemental Eq. 6.1 to quantitatively extract the THz-induced phase shift.

The THz-induced phase shift is directly proportional to the THz electric field, with the proportionality constant determined by the electro-optic effect. Taylor expanding Eq. 6.1 in $\Delta\phi$ gives [Werley 2010]:

$$\frac{\Delta I}{I_0} \approx 2\Delta\phi = 4\pi \frac{\ell}{\lambda_{\text{opt}}} \langle \Delta n_{eo} - \Delta n_o \rangle = 2\pi \frac{\ell}{\lambda_{\text{opt}}} \left(r_{33} n_{eo}^3 - r_{13} n_o^3 \right) \langle E_{\text{THz}} \rangle \tag{6.2}$$

where ℓ is the $LiNbO_3$ slab thickness, λ_{opt} is the free space optical wavelength, Δn_{eo} and Δn_o are the THz-field-induced changes in refractive index for the extraordinary and ordinary axis, respectively, r_{33} and r_{13} are the appropriate electro-optic coefficients, and $\langle E_{\mathrm{THz}}\rangle$ is the average THz electric field experienced by the probe as it propagates through the crystal. Transmission mode has very strong signals (typical peak $\Delta I / I_0$ is 20%), and in an unmodified dielectric slab waveguide the effect of averaging the THz *E*-field through the crystal depth can be calculated analytically [Werley 2010] and used to quantitatively determine the peak *E*-field. The antenna, however, strongly localizes the electric field, and averaging through the crystal depth prevents accurate quantitative measurement of enhancements in the near-field region of the gap.

To measure the THz electric field at the surface where the antenna is deposited, reflection mode measurements are required. For reflection measurements, the 400 nm probe is polarized vertically, parallel to the THz polarization, the optic axis of the $\mathrm{LiNbO_3}$ crystal, and the long axis of the antenna. Of the probe light that is incident on the lithium niobate crystal, about 14% is reflected off the front surface, and this is collected and collimated by the same objective lens used to focus the probe. The majority of this reflected light is transmitted through the dichroic mirror (R532 nm/T400 nm) and beamsplitter (BS2) and detected by photodiode PD4.

At normal incidence, the reflection off a dielectric interface is given simply by $R=(n_2-n_1)^2/(n_1+n_2)^2$. Given that the vertically polarized optical probe is incident from free space, then, the signal modulation after Taylor expansion in Δn_{eo} is

$$\frac{\Delta R}{R_0} \approx \frac{4\Delta n_{eo}^{\mathrm{surf}}}{n_{eo}^2-1} = \frac{2r_{33}n_{eo}^3}{n_{eo}^2-1} E_{\mathrm{THz}}^{\mathrm{surf}} \quad (6.3)$$

Here $E_{\mathrm{THz}}^{\mathrm{surf}}$ is the THz electric field at the surface of the $\mathrm{LiNbO_3}$ slab. Because fields parallel to a dielectric interface must be continuous, this measurement reports directly on the *E*-field in the antenna gap and can be used to quantify field enhancement.

Equation 6.3 assumes that Δn_{eo} is uniform through the depth, while in fact the THz field weakens as one moves away from the surface and gap. To confirm that the non-uniform THz field profile has a negligible effect on ΔR, we calculated the reflectivity for a material with a depth-dependent induced change in index [Thomsen 1986]. The amplitude reflection coefficient (Eq. 32 from [Thomsen 1986] after substitution of variables) is given by

$$r = r_0 + \frac{i4k}{(n_{eo}+1)^2}\int_0^\infty \exp(2ikz)\Delta n_{eo}(z)dz \equiv r_0 + \Delta r \quad (6.4)$$

with *k* the optical wave vector in the $\mathrm{LiNbO_3}$ and $r_0=(1-n_{eo})/(1+n_{eo})$ the reflectivity from the dielectric interface in the absence of a THz field. We assume the field decays exponentially away from the surface: $\Delta n_{eo}(z)=\Delta n_{eo}^{\mathrm{surf}}\exp(-z/l)$. Evaluating the integral in Eq. 6.4, calculating $\Delta R/R_0$ using $R=|r|^2$, and keeping only terms that are first order in Δn yields

$$\frac{\Delta R}{R_0} = \frac{1}{1+[\lambda_{\mathrm{opt}}/(4\pi n_{eo}l)]^2}\frac{4\Delta n_{eo}^{\mathrm{surf}}}{n_{eo}^2-1} \quad (6.5)$$

which can be compared directly with Eq. 6.3. The signal strength is not attenuated appreciably relative to Eq. 6.3 so long as $l \gg \lambda_{\mathrm{opt}}/(4\pi n_{eo}) = 14\ \mathrm{nm}$. Because we expect the decay length of the

THz field to be on the order of the gap size, for the gaps in this paper l is far larger than 14 nm and $\Delta R/R_0$ is accurately predicted by Eq. 6.3.

A drawback of reflection measurements is that reflection mode signals ($\Delta R/R_0$) are roughly 700-fold weaker than transmission mode signals ($\Delta I/I_0$), as can be seen by comparing Eqs. 6.2 and 6.3. This reduction in S/N makes it difficult to perform broadband spectroscopy in reflection mode. An additional drawback of reflection mode measurements is that the beam has contributions from reflections off both the front and back surfaces of the LiNbO$_3$ slab, but the enhanced electric fields are localized at the front surface where the antenna is deposited. To eliminate the reflection off the back surface of the slab, the reflected probe beam was focused through a 15 μm pinhole with a 30 cm focal-length lens. In addition to excluding the reflection off the back surface, the pinhole also blocked 400 nm pump light. When the pinhole was translated along the direction of beam propagation, the reflections off of both surfaces could be clearly observed 5 mm apart, as predicted by Gaussian beam analysis [Saleh 2007]. When the pinhole was correctly positioned to transmit the reflection off the front surface, the back surface reflection was attenuated by ~15-fold, effectively eliminating its contribution to the enhanced signal in the antenna gap. As a control, the probe polarization was rotated 90° to horizontal. The reflection signal was reduced by a factor of ~3 as predicted by the electro-optic effect ($\Delta n_{eo}/\Delta n_0 = r_{33}n_{eo}^3/r_{13}n_o^3 = 3.1$), indicating that the signal was accurately described by Eq. 6.3.

When capturing reference scans, in which there was no field enhancement at the front surface, the pinhole's 15-fold attenuation was not sufficient to completely suppress the contribution from the back surface reflection, which was strongly modulated by the THz during its round-trip through the sample. To measure the reference scan, we recorded the THz time trace in transmission mode. Rearranging Eq. 6.2, the electric field at the surface in a plain, unstructured slab waveguide is given by:

$$E_{\mathrm{THz}}^{\mathrm{surf}} = \frac{\xi_{\mathrm{THz}}^{\mathrm{surf}}}{\langle \xi_{\mathrm{THz}} \rangle} \langle E_{\mathrm{THz}} \rangle = -\frac{1}{2\pi\left(r_{33}n_{eo}^3 - r_{13}n_o^3\right)} \frac{\lambda_{\mathrm{opt}}}{\ell} \frac{\xi_{\mathrm{THz}}^{\mathrm{surf}}}{\langle \xi_{\mathrm{THz}} \rangle} \frac{\Delta I}{I_0} \tag{6.6}$$

where $\xi_{\mathrm{THz}}^{\mathrm{surf}} / \langle \xi_{\mathrm{THz}} \rangle$ is the ratio between the E-field amplitude at the crystal surface and the average E-field experienced by the probe as it passes through the crystal. Because the mode structure in a planar waveguide is well known, $\xi_{\mathrm{THz}}^{\mathrm{surf}} / \langle \xi_{\mathrm{THz}} \rangle$, which is on the order of 1, can be determined analytically as a function of mode and frequency [Werley 2010]. This correction can be applied in the Fourier domain to quantitatively determine spectral amplitudes at the surface of an unstructured waveguide from a transmission mode setup. To confirm the ability to quantitatively measure surface fields in transmission mode, we recorded reflection and transmission scans in an unstructured waveguide and used Eqs. 6.3 and 6.6 respectively to calculate $E_{\mathrm{THz}}^{\mathrm{surf}}$. To eliminate the contribution from the back surface in the reflection measurement, we used a longer focal length lens before the 15 μm pinhole to increase the attenuation of the back surface reflection. In addition, we coated the back surface of the LiNbO$_3$ slab with nail polish, which served as a simple antireflection coating since its index was intermediate between LiNbO$_3$ and air. A scan of the pinhole confirmed that a negligible amount of light from the back surface made it through the pinhole. Measurements of peak $E_{\mathrm{THz}}^{\mathrm{surf}}$ calculated with the two methods differed by ~10%, confirming the ability to use transmission scans for reference measurements. Transmission mode reference measurements were used to record the blue trace in Fig. 6.3(d) and for determining field enhancements in Fig. 6.5(c) & (d).

H. MODELING THE ANTENNA AS A DAMPED DRIVEN HARMONIC OSCILLATOR

As mentioned in the main text of the paper, the line shape can be understood qualitatively by modeling the antenna as a damped driven harmonic oscillator. The amplitude of the *E*-field in the antenna gap should be proportional to the amount of charge built up at the antenna ends. The differential equation describing the charge is:

$$\frac{d^2Q}{dt}+2(2\pi\nu_0)\gamma\frac{dQ}{dt}+(2\pi\nu_0)^2Q=(2\pi)^2A_0\sin(2\pi\nu t)\,. \tag{6.7}$$

Here ν is the driving frequency, ν_0 is the undamped resonant frequency, and γ is the damping ratio. The well known solution (ignoring transients) is a harmonic response at the driving frequency with a driving frequency-dependent amplitude and phase:

$$Q(t)=A(\nu)\sin[2\pi\nu t+\phi(\nu)] \tag{6.8a}$$

$$A(\nu)=\frac{A_0}{\sqrt{(2\nu\nu_0\gamma)^2+(\nu^2-\nu_0^2)^2}} \tag{6.8b}$$

$$\varphi(\nu)=\arctan\left(\frac{2\nu\nu_0\gamma}{\nu^2-\nu_0^2}\right). \tag{6.8c}$$

The amplitude and phase are plotted in Fig. 6.9(a) with γ = 0.18. The amplitude has a DC response, peaks at $\nu_{\max}=\nu_0\sqrt{1-2\gamma^2}$ (slightly lower in frequency than the undamped resonance), and decays to zero at high frequencies. The response is at very low frequencies, is phase shifted by -π/2 at ν_0, and is out of phase at high frequencies.

Because the signal strengths were too weak to measure the broadband response in reflection mode, we recorded these spectra in transmission mode as described above. As a result, the measured near-field signal interferes with a background accumulated as the probe travels through the depth of the crystal. The measured signal is approximately the near-field signal plus the reference signal: $E_{\text{sig}}(\nu)\approx E_{\text{nf}}(\nu)+E_{\text{ref}}(\nu)$. The traces plotted in Figs. 6.3 and 6.4 are $\left|\frac{E_{\text{sig}}(\nu)}{E_{\text{ref}}(\nu)}\right|\approx\left|\frac{E_{\text{nf}}(\nu)}{E_{\text{ref}}(\nu)}+1\right|$, where $\frac{E_{\text{nf}}(\nu)}{E_{\text{ref}}(\nu)}$ is a measure of the antenna response. Figure 6.9(b) shows the in-gap signal for a 110 μm arm-length antenna from Fig. 6.4(b) and $|Q(\nu)+1|$ with ν_0 = 0.24 THz, γ = 0.18, and A_0 = 0.14. The characteristics of the spectrum are predicted fairly well, including the DC response, the enhancement peak, and a minimum above the resonance. The minimum results from destructive interference between the near-field signal, which is somewhat out-of-phase with the driving field, and the background signal. The structure in the data at higher frequencies than the minimum (above ~0.45 THz) is complicated by the higher order antenna modes, so it is probable that the response from the lowest mode decays away as predicted. It is possible to capture many of the high-frequency features in the spectrum by including additional resonances in the model. One feature not completely explained is the observed frequency difference between the enhancement maximum and the transmission minimum, which is significantly greater than the red shift from the undamped resonance shown in Fig. 6.9(b). In spite of this difficulty, it seems that thinking of an antenna as a damped harmonic oscillator with charge driven by the input wave provides excellent intuition for observed antenna behavior.

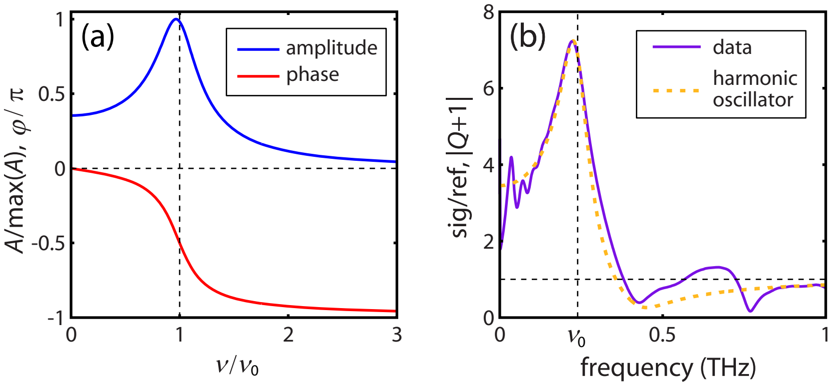

Figure 6.9 | Modeling the antenna as a damped harmonic oscillator. (a) The amplitude and phase of a damped harmonic oscillator ($\gamma = 0.18$) as a function of the normalized driving frequency, ν/ν_0, with ν_0 the resonant frequency. (b) The experimentally measured enhancement in the gap of a 110 μm arm-length antenna (purple) and $|Q+1|$, the interference between the harmonic oscillator response and a constant background (orange).

I. REFERENCES

[Balanis 2005] C. A. Balanis, *Antenna theory: analysis and design*, 3rd ed. (Wiley, Hoboken, NJ, 2005).

[Barnard 2011] E. S. Barnard, R. A. Pala, and M. L. Brongersma, “Photocurrent mapping of near-field optical antenna resonances,” *Nat. Nanotechnol.* **6**, 588-593 (2011).

[Bartel 2005] T. Bartel, P. Gaal, K. Reimann, M. Woerner, and T. Elsaesser, “Generation of single-cycle THz transients with high electric-field amplitudes,” *Opt. Lett.* **30**, 2805 (2005).

[Bitzer 2010] A. Bitzer, A. Ortner, and M. Walther, “Terahertz near-field microscopy with subwavelength spatial resolution based on photoconductive antennas,” *Appl. Opt.* **49**, E1-E6 (2010).

[Blanchard 2011] F. Blanchard, A. Doi, T. Tanaka, H. Hirori, H. Tanaka, Y. Kadoya, and K. Tanaka, “Real-time terahertz near-field microscope,” *Opt. Express* **19**, 8277-8284 (2011).

[Crozier 2003] K. B. Crozier, A. Sundaramurthy, G. S. Kino, C. F. Quate, “Optical antennas: Resonators for local field enhancement,” *J. Appl. Phys.* **94**, 4632-4642 (2003).

[Cubukcu 2008] E. Cubukcu, N. Yu, E. J. Smythe, L. Diehl, K. B. Crozier, and F. Capasso, “Plasmonic laser antennas and related devices,” *IEEE J. Sel. Top. Quant.* **14**, 1448-1461 (2008).

[Dougherty 1992] T. P. Dougherty, G. P. Wiederrecht, and K. A. Nelson, “Impulsive stimulated Raman scattering experiments in the polariton regime,” *J. Opt. Soc. Am.* **9**, 2179-2189 (1992).

[Feurer 2003] T. Feurer, J. C. Vaughan, and K. A. Nelson, “Spatiotemporal coherent control of lattice vibrational waves,” *Science* **299**, 374-377 (2003).

[Feurer 2007] T. Feurer, N. S. Stoyanov, D. W. Ward, J. C. Vaughan, E. R. Statz, and K. A. Nelson, "Terahertz polaritonics," *Annu. Rev. Mater. Res.* **37**, 317-350 (2007).

[Fischer 2008] H. Fischer and O. J. F. Martin, "Engineering the optical response of plasmonic nanoantennas," *Opt. Express* **16**, 9144-9154 (2008).

[Fromm 2006] D. P. Fromm, A. Sundaramurthy, A. Kinkhabwala, P. J. Schuck, G. S. Kino, and W. E. Moerner, "Exploring the chemical enhancement for surface-enhanced Raman scattering with Au bowtie nanoantennas," *J. Chem. Phys.* **124**, 061101 (2006).

[Ghenuche 2008] P. Ghenuche, S. Cherukulappurath, T. H. Taminiau, N. F. van Hulst, R. Quidant, "Spectroscopic mode mapping of resonant plasmon nanoantennas," *Phys. Rev. Lett.* **101**, 116805 (2008).

[Hirori 2011] H. Hirori, A. Doi, F. Blanchard, and K. Tanaka, "Single-cycle terahertz pulses with amplitudes exceeding 1 MV/cm generated by optical rectification in $LiNbO_3$," *Appl. Phys. Lett.* **98**, 091106 (2011).

[Höppener 2008] C. Höppener and L. Novotny, "Antenna-based optical imaging of single Ca^{2+} transmembrane proteins in liquids," *Nano Lett.* **8**, 642–646 (2008).

[Imura 2006] K. Imura and H. Okamoto, "Reciprocity in scanning near-field optical microscopy: illumination and collection modes of transmission measurements," *Opt. Lett.* **31**, 1474-1476 (2006).

[Kang 2009] J. H. Kang, D.S. Kim, and Q.-H.Park, "Local capacitor model for plasmonic electric field enhancement," *Phys. Rev. Lett.* **102**, 093906 (2009).

[Kinkhabwala 2009] A. Kinkhabwala, Z. Yu, S. Fan, Y. Avlasevich, K. Müllen, W. E. Moerner, "Large single-molecule fluorescence enhancements produced by a bowtie nanoantenna," *Nat. Photoics* **3**, 654-657 (2009).

[Koehl 1999] R. M. Koehl, S. Adachi, and K.A. Nelson, "Direct visualization of collective wavepacket dynamics," *J. Phys. Chem. A* **103**, 10260-19267 (1999).

[Lin 2009] K.-H. Lin, C. A. Werley, and K. A. Nelson, "Generation of multicycle THz phonon-polariton waves in a planar waveguide by tilted optical pulse fronts," *Appl. Phys. Lett.* **95**, 103304 (2009).

[Messinger 1981] B. J. Messinger, K. U. von Raben, R. K. Chang, and P. W. Barber, "Local fields at the surface of noble-metal microspheres ," *Phys. Rev. B* **24**, 649-657 (1981).

[Mühlschlegel 2005] P. Mühlschlegel, H.-J. Eisler, O. J. F. Martin, B. Hecht, D. W. Pohl, "Resonant optical antennas," *Science* **308**, 1607-1609 (2005).

[Mujumdar 2007] S. Mujumdar, A. F. Koenderink, R. Wüest, and V. Sandoghdar, "Nano-optomechanical characterization and manipulation of photonic crystals," *IEEE J. Sel. Top. Quantum Electron.* **13**, 253-261 (2007).

[Peier 2008] P. Peier, S. Pilz, F. Müller, K. A. Nelson, and T. Feurer, "Analysis of phase contrast imaging of terahertz phonon-polaritons," *J. Opt. Soc. Am. B* **25**, B70-B75 (2008).

[Peier 2009] P. Peier, S. Pilz, and T. Feurer, "Time-resolved coherent imaging of a THz multilayer response," *J. Opt. Soc. Am. B* **26**, 1649-1655 (2009).

[Saleh 2007] Saleh, B. E. A. and Teich, M. C. *Fundamentals of Photonics*, 2nd ed. (Wiley, Hoboken, NJ, 2007).

[Schnell 2010] M. Schnell, A. Garcia-Etxarri, J. Alkorta, J. Aizpurua, and R. Hillenbrand, "Phase-resolved mapping of the near-field vector and polarization state in nanoscale antenna gaps," *Nano Lett.* **10**, 3524-3528 (2010).

[Seo 2008] M. A. Seo, A. J. L. Adam, J. H. Kang, J. W. Lee, K. J. Ahn, Q. H. Park, P. C. M. Planken, and D. S. Kim, "Near field imaging of terahertz focusing onto rectangular apertures ," *Opt. Express* **16**, 20484-20489 (2008).

[Seo 2009] M. A. Seo, H. R. Park, S. M. Koo, D. J. Park, J. H. Kang, O. K. Suwal, S. S. Choi, P. C. M. Planken, G. S. Park, N. K. Park, Q. H. Park and D. S. Kim, "Terahertz field enhancement by a metallic nano slit operating beyond the skin-depth limit," *Nature Photon.* **3**, 152-156 (2009).

[Stoyanov 2002] N. S. Stoyanov, D. W. Ward, T. Feurer, and K.A. Nelson, "Terahertz polariton propagation in patterned materials," *Nat. Mater.* **1**, 95-98 (2002).

[Thomsen 1986] C. Thomsen, H. T. Grahn, H. J. Maris, and J. Tauc, "Surface generation and detection of phonons by picosecond light pulses," *Phys. Rev. B* **34**, 4129-4138 (1986).

[Wang 2004] K. Wang, D. M. Mittleman, N. C. J. van der Valk, and P. C. M. Planken, "Antenna effects in terahertz apertureless near-field optical microscopy," *Appl. Phys. Lett.* **85**, 2715-2717 (2004).

[Ward 2010] D. R. Ward, F. Hüser, F. Pauly, J. C. Cuevas, D. Natelson, "Optical rectification and field enhancement in a plasmonic nanogap," *Nat. Nanotechnol.* **5**, 732-736 (2010).

[Werley 2010] C. A. Werley, Q. Wu, K.-H. Lin, C. R. Tait, A. Dorn, and K. A. Nelson, "A comparison of phase sensitive imaging techniques for studying THz waves in structured $LiNbO_3$," *J. Opt. Soc. Am. B* **27** 2350-2359 (2010).

[Werley 2011] C. A. Werley, S. M. Teo, and K. A. Nelson, "Pulsed laser noise analysis and pump-probe signal detection with a data acquisition card," *Rev. Sci. Instrum.* **82**, 123108 (2011).

[Werley 2012] C. A. Werley, K. Fan, A. C. Strikwerda, S. M. Teo, X. Zhang, R. D. Averitt, and K. A. Nelson, "Time-resolved imaging of near-fields in THz antennas and direct quantitative measurement of field enhancements," *Opt. Express* **20**, 8551-8567 (2012).

[Willets 2007] K. A. Willets, and R. P. Van Duyne, "Localized surface plasmon resonance spectroscopy and sensing," *Annu. Rev. Phys. Chem.* **58**, 267-297 (2007).

[Wu 2009] Q. Wu, C. A. Werley, K.-H. Lin, A. Dorn, M. G. Bawendi, and K. A. Nelson, "Quantitative phase contrast imaging of THz electric fields in a dielectric waveguide," *Opt. Express* **17**, 9219-9225 (2009).

[Yariv 2007] Yariv, A. and Yeh, P. *Photonics: Optical Electronics in Modern Communications*, 6th ed. (Oxford Univ. Press, New York, 2007).

[Yeh 2007] K.-L.Yeh, M. C. Hoffmann, J. Hebling, and K. A. Nelson. "Generation of 10 μJ ultrashort THz pulses by optical rectification ," *Appl. Phys. Lett.* **90**, 171121 (2007).