UC Santa Barbara

UC Santa Barbara Electronic Theses and Dissertations

Title

On-chip terahertz time domain spectroscopy for sub-diffraction van der Waals heterostructures

Permalink

https://escholarship.org/uc/item/4148d7qp

Author

Potts, Alexander Michael

Publication Date

2023

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

University of California
Santa Barbara

On-chip terahertz time domain spectroscopy for sub-diffraction van der Waals heterostructures

A dissertation submitted in partial satisfaction
of the requirements for the degree

Doctor of Philosophy
in
Physics

by

Alexander Michael Potts

Committee in charge:

Professor Andrea Young, Chair
Professor Mark Sherwin
Professor Leon Balents

September 2023

The Dissertation of Alexander Michael Potts is approved.

Professor Mark Sherwin

Professor Leon Balents

Professor Andrea Young, Committee Chair

September 2023

On-chip terahertz time domain spectroscopy for sub-diffraction van der Waals heterostructures

Copyright © 2023

by

Alexander Michael Potts

Dedicated to my father, Michael Joseph Potts,

for inspiring a lifetime of scientific ingenuity

Acknowledgements

I have been influenced by many excellent mentors before coming to Santa Barbara. The full listing would require tens of pages, so a short list is in order. Mrs. Biegler, Mrs. Zimmerman, Mrs. Koenig, Mrs. Eddy, Mr. DeFrancesco, Mrs. Hetherington, Mr. Heinmiller, Mr. Winner and Mr. Fill taught me many life lessons before my time in university. James Bartlett, Matt Nimon, and Rick Hruby of ViaSat mentored and shaped my career more than they probably know. Dr. Jeff Lindemuth, Dr. David Daughton, Patrick Hoffman, Rich Higgins and Scott Yano of Lake Shore Cryotronics are all masters of their crafts. They collectively kindled my interest in research, particularly in the areas of applied electromagnetics and quantum materials, and always pushed me to pursue my technical dreams. At Ohio State, Dr. Zane Jamal-Eddine proved the best class buddy I could have asked for. A special thanks to Michael Ganske, Darren Ng, Anthony Grano, Lewis Sandoval, Lincoln Ficek, Cassadie Baker, Nathan Randall, Dr. Jeevant Rampal, Nathan Hart, and Dr. Joe McGlone for helping me find my career. Dr. Matt Warren, Dr. Evan Jasper, Dr. Daniel Heligman, and Dr. Thuc Mai mentored me and showed me the beauty of THz spectroscopy. Dr. Rolando Valdés Aguilar always gave good technical feedback, pushed me to be a better scientist, and encouraged me to pursue my dreams. None of this would be possible without you.

I wouldn't be the same person without my childhood friends, now my adulthood friends. Dr. Allan Leishman, Richard Oxley, Paul Beck and Stasia Levakis, Paul and Emily O'Bryan, Alex and Krista Reynolds, Dan Sabo, Nathan Branam, Darrell Newman, Brian and Rachel Rice, Emily O'Bryan and Dr. Laura Collins - thank you all for making life fun.

I am grateful for the technical and friendly minds I've met and worked with along the way. My advisor, Dr. Andrea Young, has offered excellent technical feedback, developed

and enriched my perception of scientific research, and has made me a better writer. My fellow graduate student and postdoc labmates Dr. Fangyuan Yang, Dr. Gregory Polshyn, Dr. Yu Saito, Dr. Sasha Zibrov, Dr. Marec Serlin, Dr. Haoxin Zhou, Dr. Charlie Tschirhart, Liam Cohen, Noah Samuelson, Ludwig Holleis, Evgeny Redekop, Wenting Gong, Yi Guo, Owen Sheekey, and Caitlin Patterson have been an indelible part of my community at UCSB. I want to especially thank Dr. Joshua Island, Dr. Peter Kissin, Dr. Eric Spanton, Dr. Yiftach Frenkel, Dr. Richard Averitt and Dr. Abhay K. Nayak, Dr. Mark Sherwin, and Dr. Kelson Kaj for their contributions on all things THz. I aspire to have your patience, technical mastery, and physical insight.

I want to thank my classmates and friends: Dr. Matthew McEwen, Molly Wolfson, Jenny Smith, Sean McBride, Dr. Craig Pellegrino, and Dr. G. Mirek Brant. Your friendship has kept me going over the years, and I thank you all for it. In addition, my church friends including Diane Loewen, Joel Loewen, Rhett Austin, Ryler Austin, Nissi Vasa, Nico Para, Keith and Carissa Gagnon, Sam Biancone, Erin Etherton, Michaela Jorgenson, Dr. Lars Linton, Nate Wagner, Jay Campbell, Dr. Peter Ford, Jerry Stewart, Laurana Stewart, Max Moore, Dr. Randy Verhoef, John and Lauri Whalley, Jay Campbell, Bella Gregson, Mike Santarrosa, Danny Johnson, Dr. Rick (and Mary) Osgood, Peter Osgood, Max and Molly Moore, Paul Kovacevich, and all my students have greatly enriched my time and studies. Thank you all - you will all be missed.

My tenure with the Santa Barbara Yacht club has been unexpected, but immensely enriching and fun. All members are to be thanked, but especially Loren Toscher, Captain Dr. Rod Paragas, Tony Lee, Dr. Charlie Nusbaum, Jason Quisberth, Dr. Danny Klinenberg, Brian Casey, Captain Quincy Briscoe, Ellis Talton, Graham Guess, Chad Ress, Dan Oh, Dan (and Sue) Thompson, Jack Lorentson, Brian Lee, Erik Lobben, and Todd Dickey. I also want to thank my extended Toronto crew, specifically - Jonathan and Tabitha Davala and Hana 'Akka' Priya. Thank you all for taking care of me and

helping me heal from my cancer.

Liam, I couldn't have asked for a better friend. Your positivity and zeal for physics is infectious. You have seen me at my best and my worst and have single-handedly prevented from dropping out at least four times. You are my best friend and most trusted colleague. I will dearly miss you seeing you everyday.

I must also thank my parents, my brother, my aunt, and my extended family (Grandpa, Christy, Granny, Grampy, Greg, Janell, Lisa, Tim, Gary, Barbara, Carol, Rob, Sydney, Aidan, Linda, Inglis, Becky, Dean, Burke, Jen, Alex, Morgan, Lauren, Abby, Hannah, and Emily). Dad, you inspired me to take on a technical career and explore my own interests. Mom, you gave me the energy to explore and not give up. And Ben, your company remains fun and witty, and your insight remains razor sharp after all this time. Aunt Joycie, I admire your perspective and thirst for unconventional, creative, and otherwise shunned perspectives. Thank you all.

And, of course, there is the venerable Priscilla Pulletikurthi. What would I do without you? Thanks for taking care and loving me, my dearest Pris, through the most fun of times and through the darkness of my battle with cancer. I look forward to what lies next for us. Dhanyavaadalu, na bangaram.

Curriculum Vitæ

Alexander Michael Potts

Education

2023	Ph.D. in Physics (Expected), University of California, Santa Barbara.
2021	M.A. in Physics, University of California, Santa Barbara.
2017	M.S. in Electrical & Computer Engineering, The Ohio State University
2017	B.S. in Electrical & Computer Engineering, The Ohio State University, Honors, Magna Cum Laude
2016	B.S. in Physics, The Ohio State University, Magna Cum Laude

Publications

1. A. M. Potts A. K. Nayak, M. Nagel, K. D. Kaj, B. Stamenic, R.D. Averitt, and A. F. Young, "Cryogenic, On-chip, Time-domain Terahertz Spectrometer with Fast Sample Interchange Architecture." ACS Nano Letters 23, 9, pp. 3835–3841 (2023)
2. A. M. Potts, S. Bajaj, D. R. Daughton, A. A. Allerman, A. M. Armstrong, T. Razzak, S. H. Sohel, and S. Rajan, "$Al_{0.7}Ga_{0.3}N$ MESFET With All-Refractory Metal Process for High Temperature Operation." IEEE Transactions on Electron Devices Volume: 68, Issue: 9 (2021)
3. J. O. Island, P. Kissin, J. Schalch, X. Cui, S. Rubaiat U. Haque, A. M. Potts, T. Taniguchi, K. Watanabe, R. D. Averitt, and A. F. Young, "On-chip terahertz modulation and emission with integrated graphene junctions." Applied Physics Letters 116, 161104 (2020)
4. A. M. Potts, T. T. Mai, M. T. Warren, and R. Valdés Aguilar, "Corrective regridding techniques for non-uniform sampling in time-domain terahertz spectroscopy." Journal of the Optical Society of America B Vol. 36, Issue 4, pp. 1037-1043 (2019)
5. D. M. Heligman, A. M. Potts, N. Crescimanno, M. T. Warren, E. V. Jasper, T. T. Mai, R. Valdés Aguilar, "Size and Periodicity Effects on Terahertz Properties of Gammadion Metamaterials." arXiv:1811.06041 (2019)

Abstract

On-chip terahertz time domain spectroscopy for sub-diffraction van der Waals heterostructures

by

Alexander Michael Potts

Free-space time domain THz spectroscopy accesses the electrodynamic responses of quantum materials at frequencies ideally matched to the energy scales of interacting condensed matter systems. THz spectroscopy, however, is challenging when samples are physically smaller than the diffraction limit of $\sim$0.5 mm, as is typical, for example, in van der Waals materials and heterostructures. We examine foundational electrodynamics and spectroscopic principles relevant at THz frequencies. We present THz engineering challenges and overcome them by designing an on-chip, time-domain THz spectrometer with an interchangeable sample architecture and a bandwidth of 750 GHz. We use this spectrometer and extract the optical conductivity of a 7.5-um wide NbN film across the superconducting transition. We then extensively benchmark the spectrometer's operation in the context of THz engineering principles, and conclude by looking forward to the measurements this technology enables, including superconductivity, magnetism, and charge order in sub-diffraction materials and van der Waals heterostructures.

Contents

List of Figures

List of Tables

Chapter 1

Electrodynamics in linear media

> "Anyone who tells you that Maxwell's equations in media are easy is full of it"
>
> *F. Pincus*

Classical electrodynamics via Maxwell's equations describes the relationships between electric charges, electric fields, and magnetic fields. The application of Maxwell's equations to media is the foundational underpinning of engineering at THz frequencies and all THz spectroscopy.

The importance of Maxwell's equations to not only physics, but to human history, cannot be overstated. Maxwell's equations are one of the greatest scientific achievements of the human experience. Maxwell's equations include the wave equation, conservation of charge, Kirchkoff's voltage and current laws, the physics of radiation, the magnetosphere's screening of the sun's solar wind, and clues that would lead to Einstein's theory of special relativity. Maxwell's equations represent the first compact codification of a field theory, later inspiring quantum field theory and Einstein's quest for a grand unified field theory. Simply put: without Maxwell's equations, physicists would still be in the thralls of the industrial revolution pondering Atwood's machine.

Vector quantities are denoted by an overline (example: $\overline{E}$); dyadic tensor quantities are donated by two overlines (example: $\overline{\overline{\epsilon}}$). Time domain fields are indicated with lower case letters (example: $\overline{e}$); Frequency domain fields are indicated with upper case letters (example: $\overline{E}$). For more information on notation and sign conventions, see appendix E.

1.1 Time harmonic Maxwell's equations

Maxwell's equations is a set of four coupled, partial differential equations are individually called Gauss's law, no magnetic monopoles, Faraday's law, and Maxwell-modified Ampere's law. In differential form, Maxwell's equations in media are commonly[1] written as:

$$\begin{aligned} \nabla \cdot \overline{d} &= \rho_e \\ \nabla \cdot \overline{b} &= 0 \\ \nabla \times \overline{e} &= -\partial_t \overline{b} \\ \nabla \times \overline{h} &= \overline{j_e} + \partial_t \overline{d} \end{aligned} \tag{1.1}$$

Eqn. 1.1 places fields on the left side and source terms on the right side. Source, such as current, charge, and changing $\overline{e}$ and $\overline{h}$ fields, are responsible for producing the fields that appear on the left. The electric current density $\overline{j_e}$ includes both conduction currents and externally impressed currents. Application of the Fourier Transform to eqns. 1.1 converts Maxwell's equations into their time-harmonic form:

[1]Magnetic monopoles have never been observed, but it can be useful to define a fictitious magnetic monopole density, magnetic monopole current density, corresponding electric field vector potential, and magnetic loss tangents. Many antenna (and consequently, by Huygens principle, scattering) problems can be greatly simplified with the duality theorem and the dyadic Green's function derived from these extended Maxwell equations. Though somewhat useful for understanding some loss mechanisms introduced in Ch. 3, we ignore this entirely.

$$\begin{aligned} \nabla \cdot \overline{D} &= \rho_e \\ \nabla \cdot \overline{B} &= 0 \\ \nabla \times \overline{E} &= -j\omega\overline{B} \\ \nabla \times \overline{H} &= \overline{J_e} + j\omega\overline{D} \end{aligned} \tag{1.2}$$

Media are modelled as supporting a bound electric and magnetic polarizations in response to impinging fields. Bound electric polarization $\overline{P}$ is given by $\overline{P} \equiv \overline{\overline{\chi_E}}\overline{E}$. (scaled) imposed electric field intensity $\overline{E}$ plus the material response $\overline{P}$ produces electric flux intensity $\overline{D}$. Similarly, bound magnetic polarization $\overline{M}$ is given by $\overline{M} \equiv \overline{\overline{\chi_M}}\overline{H}$. Imposed magnetic field intensity $\overline{H}$ plus the bound magnetization $\overline{M}$ produces total magnetic flux density $\overline{B}$. We also assume that the electric field intensity $\overline{E}$ is related to and conduction electric current (J_e) by Ohm's Law) and write:

$$\begin{aligned} \overline{D} &= \overline{\overline{\epsilon}}\overline{E} = \epsilon_0(\overline{\overline{1}} + \overline{\overline{\chi_E}})\overline{E} \\ \overline{B} &= \overline{\overline{\mu}}\overline{H} = \mu_0(\overline{\overline{1}} + \overline{\overline{\chi_M}})\overline{H} \\ \overline{J_e} &= \overline{\overline{\sigma}}\overline{E} \end{aligned} \tag{1.3}$$

The relationships of eqn. 1.3 are called the constitutive relations. We completely define materials by their the dielectric permittivity $\overline{\overline{\epsilon}} = \overline{\overline{\epsilon}}_r\epsilon_0$, magnetic permeability $\overline{\overline{\mu}} = \overline{\overline{\mu}}_r\mu_0$, and electric monopole conductivity $\overline{\overline{\sigma}}$. These quantities are complex-valued tensors that, in general, are frequency-dependent, anisotropic, not diagonal, and spatially varying. For the purposes of this dissertation, all media will be assumed to be linear, which stipulates a linear relationship between $\overline{E}$ and $\overline{D}$, $\overline{H}$ and $\overline{B}$, and $\overline{J}$ and $\overline{E}$.

Maxwell's equations admit boundary conditions. These boundary conditions are derived by enforcing the integral forms of Maxwell's equations across a differentially small manifold containing a boundary with source terms [3]. At the boundary of two media (1

and 2), the $\overline{E}$ and $\overline{H}$ fields are related by the following conditions ($X|_{surf}$ refers to X at the media interface):

$$\begin{aligned} \hat{n} \times (\overline{H_1} - \overline{H_2})|_{interface} &= \overline{J_e}|_{surf} \\ \hat{n} \times (\overline{E_1} - \overline{E_2})|_{interface} &= 0 \\ \hat{n} \cdot (\overline{B_1} - \overline{B_2})|_{interface} &= 0 \\ \hat{n} \cdot (\overline{D_1} - \overline{D_2})|_{interface} &= \rho_e|_{surf} \end{aligned} \tag{1.4}$$

Eqn. 1.4 reveals the behavior of interfaces. If there are no surface electric currents, the tangential components of both $\overline{E}$ and $\overline{H}$ across the interface are continuous. Surface electric currents source $\overline{H}$; surface electric charges source electric field intensity $\overline{E}$.

The first material we will consider is the perfect electric conductor (PEC). A PEC is defined as medium with a purely tangential surface electric field and all internal fields ($\overline{E}$, $\overline{H}$, $\overline{B}$, and $\overline{D}$) equal to zero. A purely tangential surface electric field and $\overline{E_2} = 0$ (Eqn. 1.4) suggests that a PEC must support a surface electric current in response to any impinging electric fields. This provides intuition on the nature of a PEC - that a PEC's supports internal surface currents that cancel the tangential components of any incident electromagnetic fields.

1.1.1 Which is the "fundamental" magnetic field: B or H?

Physicists typically refer to $\overline{E}$ as the electric field and $\overline{B}$ as the magnetic field; engineers typically refer to $\overline{E}$ as the electric field and $\overline{H}$ as the magnetic field. One might ask: is B or H the fundamental magnetic field? To some extent, this is not worth serious consideration, as one can simply refer to a "B field" or "H field". Proper distinction, however, does play a key role in numerical calculation. Some physicists, including the venerable David Griffiths [4], have penned aggressive expostulations against anyone claiming that

$\overline{H}$ is the fundamental magnetic field. Griffiths is wrong. If $\overline{E}$ is the fundamental electric field, $\overline{H}$ *must* be the fundamental magnetic field. We now demonstrate this both theoretically and experimentally.

Examine Maxwell's equations carefully. $\overline{E}$ and $\overline{H}$ are only operated on by curl differential operators; $\overline{D}$ and $\overline{B}$ are acted upon by divergences and time derivatives. This is the basis for the Maxwellian duality theorem [3], which links $\overline{E}$ to $\overline{H}$ (not $\overline{B}$) as dual quantities. The duality theorem directly affects the boundary conditions in eqn. 1.4. The boundary conditions are derived from the integral form of Maxwell's equations. Thus, quantities involving a curl ($\overline{E}$ and $\overline{H}$) directly translate to being tangentially continuous across a source-free interface; quantities involving a divergence ($\overline{D}$ and $\overline{B}$) become normally continuous across a source-free interface.

This can be formalized in the language of differential topology. All relevant scalars and fields in Maxwell's equations must have finite energy, meaning they must belong to a square-integrable subspace L^2 on manifold Ω. Scalar charge density is related to fields $\overline{B}$ and $\overline{D}$ through the divergence operator. Not all $L^2(\Omega)$ functions will have a divergence still in $L^2(\Omega)$, so the subspaces of $\overline{B}$ and $\overline{D}$ are restricted to Sobolev space $\mathbb{H}(\nabla\cdot;\Omega)$ - the space of functions whose divergence lies in L^2. $\overline{E}$ and $\overline{H}$ (and vector potential $\overline{A}$, see appendix A) are related to $\overline{B}$ and $\overline{D}$ by the curl operator, and thus their function subspace is further reduced to $\mathbb{H}(\nabla\times,\nabla\cdot;\Omega)$: the subspace of functions whose curl is in $\mathbb{H}(\nabla\cdot;\Omega)$. The scalar potential Φ is related to the $\overline{E}$ via gradient, so it must reside in lower subspace, which we denote as $\mathbb{H}(\nabla,\nabla\times,\nabla\cdot;\Omega)$, which is the Sobolev space of functions whose gradient resides in $\mathbb{H}(\nabla\times,\nabla\cdot;\Omega)$. The resulting De Rham complex is summarized in Table 1.1.

A direct consequence of the duality theorem is that $\overline{E}$ and $\overline{H}$ are both 1-form objects; $\overline{B}$ and $\overline{D}$ are 2-form objects. This may seem a minor point, but is extremely important in both hand calculations and computational electromagnetics. If one were to expand

Form	Quantities	Minimum function space	Interface continuity
0-form	Φ	$\mathbb{H}(\nabla, \nabla\times, \nabla\cdot; \Omega)$	Continuous
1-form	$\overline{E}$, $\overline{H}$, $\overline{A}$	$\mathbb{H}(\nabla\times, \nabla\cdot; \Omega)$	Tangentially continuous
2-form	$\overline{D}$, $\overline{B}$	$\mathbb{H}(\nabla\cdot; \Omega)$	Normally continuous
3-form	ρ_e	$L^2(\Omega)$	Discontinuous

Table 1.1: Summary of electromagnetic cohomologies

the scalar potential $\Phi(\overline{r})$ as $\sum_{i=1}^{M} A_m \delta|\overline{r_i}|$ for discrete $\overline{r_i}$, it is obvious that $\overline{E} = -\nabla\overline{\phi}$ would be a discontinuous function and thus have non-physically infinite energy. It is imperative to expand fields in a basis appropriate for their cohomologies. Conflating the Sobolev spaces of $\overline{B}$ and $\overline{H}$ could lead to non-physical singularities and thus non-physical solutions. $\overline{E}$ and $\overline{H}$ *must* be expanded in the same Sobolev subspace; $\overline{E}$ and $\overline{B}$ cannot be expanded in the same Sobolev subspace. For more details of how expanding quantities in their appropriate subspaces manifests practically in simulation, aspects of finite element computation, see appendix B.

$\overline{E}$ and $\overline{H}$ are not just linked by theory. The line integral of a $\overline{E}$ produces a voltage, which is an experimentally accessible quantity. The line integral of $\overline{H}$, not $\overline{B}$, produces a current, which is also an experimentally accessible quantity. The line integral of $\overline{B}$ is not a directly measurable quantity without relating it to the line integral of $\overline{H}$. It is obvious, then, that $\overline{E}$ and $\overline{H}$ are also the *experimentally* fundamental quantities.

$\overline{E}$ and $\overline{H}$ are linked by the duality theorem, have the same interface continuity requirements, are cohomologous, and are the fundamentally measurable electric and magnetic fields. $\overline{E}$ and $\overline{H}$ are, therefore, the fundamental electric and magnetic fields, both experimentally and theoretically.

1.1.2 The Helmholtz equation

The Helmholtz wave equation may be derived from 1.2. We begin by taking the curl of Faraday's law (eqn. 1.2) and applying vector identities:

$$\begin{aligned} \nabla \times \nabla \times \overline{E} &= \nabla \times -j\omega\overline{B} \\ \nabla(\nabla \cdot \overline{E}) - (\nabla^2\overline{E}) &= \nabla \times -j\omega\overline{B} \end{aligned} \tag{1.5}$$

We have assumed a homogenous medium, as interfaces between media will be handled later. We apply Gauss's law to the divergence of $\overline{E}$, Maxwell-modified Ampere's law to the right hand side, and Ohm's law[2] to the resulting current:

$$\begin{aligned} \overline{\overline{\epsilon}}^{-1}\nabla(\rho_e) - (\nabla^2\overline{E}) &= -j\omega\nabla \times \overline{B} \\ \overline{\overline{\epsilon}}^{-1}\nabla(\rho_e) - (\nabla^2\overline{E}) &= -j\omega\overline{\overline{\mu}}(\overline{J_e} + j\omega\overline{\overline{\epsilon}}\overline{E}) \\ (\nabla^2\overline{E}) &= \overline{\overline{\epsilon}}^{-1}\nabla(\rho_e) + j\omega\overline{\overline{\mu}}(\overline{\overline{\sigma}}\overline{E} + j\omega\overline{\overline{\epsilon}}\overline{E}) \end{aligned} \tag{1.6}$$

The final line of eqn. 1.6 is an uncoupled AC-only differential equation for $\overline{E}$ in the form of a vector wave (Helmholtz) equation. An analogous wave equation for $\overline{H}$ can be also derived following the same logic. If we assume a source free-region, the Helmholtz equations for $\overline{E}$ and $\overline{H}$ reduce to:

$$\begin{aligned} \nabla^2\overline{E} &= j\omega\overline{\overline{\mu\sigma}}\overline{E} - \omega^2\overline{\overline{\epsilon\mu}}\overline{E} \equiv \gamma^2\overline{E} \\ \nabla^2\overline{H} &= j\omega\overline{\overline{\mu}}\ \overline{\overline{\sigma}}\overline{H} - \omega^2\overline{\overline{\epsilon}}\ \overline{\overline{\mu}}\overline{H} \equiv \gamma^2\overline{H} \end{aligned} \tag{1.7}$$

Eqn. 1.7 is a well-known form of the wave equation, whose time domain solution in one dimension (x) is a sum of plane waves, which individually take the form $e^{-\gamma x}$. All basic wave properties can be derived by deducing the propagation constant γ. In one dimension, the real and imaginary parts of γ can be simplified to:

[2]This assumes no externally impressed current (for example, a constant current source). Including such a term doesn't change the basic result.

$$\begin{aligned}\gamma \equiv \alpha + j\beta &= \sqrt{j\omega\mu(\sigma + j\omega\epsilon)} \\ \alpha &= \omega\sqrt{\mu\epsilon}\sqrt{\frac{1}{2}\left(\sqrt{1+(\sigma/\omega\epsilon)^2}-1\right)} \\ \beta &= \omega\sqrt{\mu\epsilon}\sqrt{\frac{1}{2}\left(\sqrt{1+(\sigma/\omega\epsilon)^2}+1\right)}\end{aligned} \tag{1.8}$$

α is the extinction coefficient, which quantifies a propagating wave's exponential attenuation. β is the phase coefficient, which describes the change of a wave's phase as it propagates. The basic electromagnetic wave properties of wave impedance, wavelength, phase velocity, index, and skin depth can be generated from eqn. 1.8. The importance of the quantities will be evident throughout the remaining chapters of this dissertation. The plane wave's wavelength is simply defined as $2\pi/\beta$. The impedance of a travelling electromagnetic wave is the ratio of $\overline{E}$ to $\overline{H}$ in orthogonal directions. For example, the impedance of a plane wave travelling along $\hat{z}$ with $\overline{E} = E_0\hat{x}$ and $\overline{H} = H_0\hat{y}$ would be $|E_0/H_0|$. The phase velocity of a plane wave (v_p), defined as ω/β, represents the translational velocity of a constant phase of a propagating plane wave. Phase velocity is often expressed in the form of optical index, which is defined as the vacuum speed of light ($1/\sqrt{\epsilon_0\mu_0} = 300\mu m/ps \equiv c_0$) divided by v_p. The skin depth is the distance after which a travelling plane wave is attenuated to amplitude $1/e \approx 37\%$. These quantities are summarized in 1.2, in exact form, in the limit of good dielectrics (where $(\sigma/\omega\epsilon)^2 \ll 1$), and in good metals (where $(\sigma/\omega\epsilon)^2 \gg 1$).

The careful examination and interpretation of the frequency dependence of electromagnetic wave parameters at THz frequencies defines THz spectroscopy, which will be the subject of the rest of this dissertation. Several trends are apparent from Table 1.2, which build our most naïve intuition for material parameters $\overline{\overline{\epsilon}}$, $\overline{\overline{\mu}}$, and $\overline{\overline{\sigma}}$. In good dielectrics, the impedance and phase velocity are frequency-independent; in good metals, the impedance and phase velocity have a frequency dependence. In good dielectrics,

Quantity	Exact	Good dielectric $(\sigma/\omega\epsilon)^2 \ll 1$	Good metal $(\sigma/\omega\epsilon)^2 \gg 1$
Extinction Coeff. α [1/m]	$\omega\sqrt{\mu\epsilon}\sqrt{\frac{1}{2}\left(\sqrt{1+(\sigma/\omega\epsilon)^2}-1\right)}$	$\frac{\sigma}{2}\sqrt{\mu/\epsilon}$	$\sqrt{\omega\mu\sigma/2}$
Phase Coeff. β [1/m]	$\omega\sqrt{\mu\epsilon}\sqrt{\frac{1}{2}\left(\sqrt{1+(\sigma/\omega\epsilon)^2}+1\right)}$	$\omega\sqrt{\mu\epsilon}$	$\sqrt{\omega\mu\sigma/2}$
Impedance η [Ohms]	$\sqrt{j\omega\mu/(\sigma+j\omega\epsilon)}$	$\sqrt{\mu/\epsilon}$	$(1+j)\sqrt{\omega\mu/2\sigma}$
Wavelength λ [m]	$2\pi/\beta$	$2\pi/(\omega\sqrt{\mu\epsilon})$	$2\pi\sqrt{2/\omega\mu\sigma}$
Phase velocity v_p [m/s]	ω/β	$1/\sqrt{\mu\epsilon}$	$\sqrt{2\omega/\mu\sigma}$
Optical index n [Unitless]	c_0/v_p	$\sqrt{\mu\epsilon}/\sqrt{\mu_o\epsilon_0}$	$\sqrt{\mu\sigma/2\omega\mu_0\epsilon_0}$
Skin depth δ [m]	$1/\alpha$	$2\sqrt{\epsilon/\mu\sigma^2}$	$\sqrt{2/\omega\mu\sigma}$

Table 1.2: Summary of important wave quantities.

the impedance purely real; in good metals, the impedance is equally real as it is imaginary. The skin depth is a key consideration in THz engineering. The skin depth of good dielectrics depends on $1/\sigma^2$, which is very large in dielectrics - meaning that the electromagnetic wave propagates deep into the dielectric medium without attenuation. The skin depth of metals is set by $1/\sqrt{\omega\sigma}$, which can be less than 100 nm at THz frequencies.

1.1.3 Interfaces between media

We have thus far considered only homogenous materials. We consider, now, the interface between two materials characterized by different functions of $\overline{\overline{\epsilon}}$, $\overline{\overline{\mu}}$, and $\overline{\overline{\sigma}}$. Imagine an interface between two semi-infinite media, with uniform isotropic constitutive parameters ϵ_1, μ_1, and σ_1 in medium 1 and ϵ_2, μ_2, and σ_2 in medium 2. By applying the boundary conditions (eqn. 1.4), one can show that a plane wave injected from medium 1 to medium 2 at normal incidence has reflection coefficient $\Gamma(\omega)$ and transmission coefficient $T(\omega)$

given by:

$$\begin{aligned} \Gamma(\omega) &= \frac{\eta_2(\omega) - \eta_1(\omega)}{\eta_2(\omega) + \eta_1(\omega)} \\ T(\omega) &= \frac{2\eta_2(\omega)}{\eta_2(\omega) + \eta_1(\omega)} \end{aligned} \tag{1.9}$$

Impedances are fully complex quantities, so reflection and transmission coefficients are also fully complex quantities that describe the phase of Γ and T. A useful rule of thumb implied by eqn. 1.9 is that transitions from (purely real) high to low impedances have reflection coefficients with an overall negative sign (π phase shift). Transitions from (purely real) low to high impedances have reflection coefficients with no overall phase shift. With pulsed measurements, where the incoming and reflected signals can be independently resolved in time, an inverted or non-inverted reflection can immediately answer whether the impedance transition was from $\eta_2 > \eta_1$ or $\eta_2 < \eta_1$.

Impedance is clearly the key parameter to understanding and engineering material interfaces, as it completely controls reflection and transmission coefficients. Finding the reflection coefficient of a few limiting cases gives vital information we will later use in circuit benchmarking. If material 2 is impedance matched to material 1 ($\eta_2 = \eta_1$), there is complete transmission and no reflection ($\Gamma = 0$, $T = 1$). If a "short (circuit)" is presented with $\eta_2/\eta_1 \ll 1$, there will be a complete reflection with a negative phase and no transmission ($\Gamma = -1$, $T = 0$). If an "open (circuit)" is presented with $\eta_2/\eta_1 \gg 1$, there will be complete reflection with no phase and a ($\Gamma = 1$, $T = 2$).

It curiously appears that the transmission coefficient is greater than unity when $\eta_2/\eta_1 > 1$. This does not, however, violate energy conservation. $\Gamma(\omega)$ and $T(\omega)$ in 1.9 are for the incident $\overline{E}$ fields. The reflection and transmission coefficients for incident and reflected power are found as $|\Gamma(\omega)|^2$ and $|T(\omega)|^2$ respectively. By the conservation of energy, we know that $|\Gamma(\omega)|^2 + |T(\omega)|^2 \leq 1$, with any sub-unity sum due to absorption

by surface currents. The power transmitted is $|\overline{E}|^2/\eta_2 \sim |T|^2/\eta_2$, which does not exceed unity for $\eta_2/\eta_1 > 1$. We see that the transmitted power drops to zero for $\eta_2/\eta_1 \to \infty$, meaning no power is transmitted across an open circuit.

Measuring the reflection and/or transmission coefficients from a sample can be used to extract a material's impedance, providing information on σ, ϵ, and μ. This is the foundational basis of THz and optical spectroscopy.

1.2 Transmission lines & the telegraphist's equations

We have explored unbound waves - propagating waves that are not guided by carefully engineering structures of conductors and dielectrics. We now explore waves bound in artificial guiding structures: transmission lines. A transmission line is defined a distributed structure capable of supporting propagating $\overline{E}$ and $\overline{H}$ waves. Transmission lines with total length L_0 differ from what are typically called "analog circuits" with total length L_0 in that transmission lines have $L_0\lambda \geq 1$ ("electrically large"), whereas analog circuits have $L_0/\lambda \leq 1$ ("electrically small"). That is: analog circuits are agnostic to the wave nature of electric and magnetic fields; transmission lines are not. Transmission lines are modelling as an infinite network of a differential element (Fig. 1.1) with the following parameters per unit length:

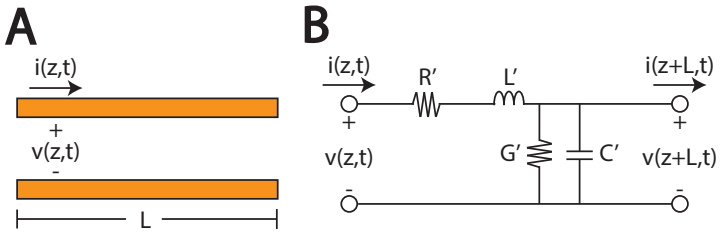

Figure 1.1: Transmission line schematics with (A) distributed model (B) lumped element model in differentially small element size L.

1. R': Resistance, determined by material conductivity and geometric effects

2. L': inductance, determined by the self-inductance of the two conductors

3. G': shunt conductance, determined by the transmission line's substrate

4. C': capacitance, determined by conductor geometry and local dielectric

Writing Kirchoff's voltage and current laws for Fig. 1.1, respectively, produce:

$$\begin{aligned} v(z,t) - R'i(z,t) - L'\partial_t i(z,t) - v(z+L,t) = 0 \\ i(z,t) - G'v(z+L,t) - C'\partial_t v(z+L,t) - i(z+L,t) = 0 \end{aligned} \tag{1.10}$$

To simplify, take the limit as $L/\lambda = nL\omega/c_0 \to 0$ and convert to the frequency domain. This produces two coupled equations known as the telegraphist's wave equations:

$$\begin{aligned} \partial_z V(z) = -(R' + j\omega L')I(z) \\ \partial_z I(z) = -(G' + j\omega C')V(z) \end{aligned} \tag{1.11}$$

These equations parallel Faraday's law and the Maxwell-corrected Ampere's law (eqn. 1.2) in a source-free, linear, and homogenous region, with voltage acting as a scalar stand-in for $\overline{E}$ and current acting as a scalar stand-in for $\overline{H}$. The solution to the telegraphist wave equations, much like Maxwell's equations in media, are current and voltage planes of the form:

$$\begin{aligned} V(z) = V_0^+ e^{-\gamma z} + V_0^- e^{\gamma z} \\ I(z) = V(z)/Z_0 \end{aligned} \tag{1.12}$$

Eqn. 1.12 describes plane waves bound to the transmission line that propagate through both space and time. The meaning of impedance to a plane wave in a transmission line is more familiar than it was for unguided waves - it simply relates the (complex) voltage to the (complex) current. Using our knowledge of eqn. 1.12 and the wave equation, we can duplicate Table 1.2 for transmission line circuits. The loss in transmission

lines that we use will be dominated by series resistance R' losses, so we will neglect losses associated with G'. It is very rare to design high-loss transmission lines (for all frequencies), so we list the approximate formulae for low-loss lines only.

Quantity	Exact	Low-loss
Extinction Coeff. α [1/m]	$\frac{\omega(L'G'+R'C')}{\sqrt{2(\omega^2L'C'-R'G')\sqrt{(G'^2+\omega^2C'^2)(R'^2-\omega^2L'^2)}}}$	$R'\sqrt{C'/L'} \approx R'/Z_0$
Phase Coeff. β [1/m]	$\frac{\sqrt{(\omega^2L'C'-R'G')\sqrt{(G'^2+\omega^2C'^2)(R'^2-\omega^2L'^2)}}}{\sqrt{2}}$	$\omega\sqrt{L'C'}$
Impedance Z [Ohms]	$\sqrt{\frac{R'+j\omega L'}{G'+j\omega C'}}$	$\sqrt{\frac{L'}{C'}}\left(1-j\frac{R'}{2\omega L'}\right)$
Wavelength λ [m]	$2\pi/\beta$	$2\pi/\omega\sqrt{L'C'}$
Phase velocity v_p [m/s]	ω/β	$1/\sqrt{L'C'}$
Effective mode index n [Unitless]	c_0/v_p	$\sqrt{L'C'/\mu_0\epsilon_0}$
Skin depth δ [m]	$1/\alpha$	$R'^{-1}\sqrt{L'/C'}$

Table 1.3: Summary of important wave quantities for transmission lines. The low loss limit is defined by $\omega L'/R' \gg 1, 1/\omega C'R' \gg 1$ and G' = 0.

Table 1.3 lists expressions for the same wave parameters we derived for unbound propagation (table 1.2). Here, $\overline{E}$ parallels V, $\overline{H}$ parallels I, L parallels μ, and C parallels ϵ. Shunt leakage conductance G' represents leakage in the capacitance, which is represented by a non-zero ϵ_2 in the unbound waves. Propagating waves become lossless ($\alpha \to 0$) in the case that R' and G' are zero, making R' and G' resistive (dissipative) parameters and L' and C' reactive parameters. Much like in free-space, impedance controls the reflection and transmission from an electrically small (physical length L $\ll \lambda$) interface, according to eqn. 1.9 (with η_i replaced by Z_i).

The loss of a low-loss transmission line is explicitly frequency-independent, but the devil is in the details. C' and L' are mostly determined by geometry, and thus only very

weakly frequency dependent. R', however, has a $\sqrt{\omega}$ frequency-dependence from the skin effect, in addition to other weak implicit dependences. Actual calculation of the R', L', G', and C' for a specific transmission line geometry is estimable for some geometries by hand calculation [3, 5], but is best calculated by finite element or method of moment computational software.

A (distributed) transmission line is a simple example of a 2-port network, which are defined as elements taking two inputs and having two output. Transmission lines take current and voltage inputs and produce current and voltage outputs. The input and output currents and voltages are related by admittance (Y), impedance (Z) or the transmission (ABCD) coefficient matrix. The ABCD matrix is defined as:

$$\begin{aligned} V_1 &= AV_2 + BI_2 \\ I_1 &= CV_2 + DI_2 \\ \begin{bmatrix} V_1 \\ I_1 \end{bmatrix} &= \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} V_2 \\ I_2 \end{bmatrix} \end{aligned} \tag{1.13}$$

The ABCD coefficients of more complicated circuits can also be written. One can identify the ABCD coefficients of a non-distributed shunt (series) circuit element in a transmission line with admittance $Y(\omega)$ (impedance $Z(\omega)$) by simply applying both the Kirchhoff current and voltage laws[5]. One can also use the simple definition of ABCD coefficients and the solutions to the telegraphist wave equation (eqn. 1.12) to write them for transmission lines. The ABCD coefficients are intuitive: with two 2-port networks are placed in cascade (series), the overall ABCD matrix is the product of the individual ABCD matrices. More complicated combinations of circuits, such as cascode and inter-connected cascode configurations, can be difficult time consuming to solve with ABCD matrices. In addition, A, B, C, and D have different interpretations and units, making

resulting formulae difficult to understand intuitively. To simplify, the s-parameters were introduced, which will be the preferred 2-port network parameters for this dissertation.

S-parameters are defined by incoming and outgoing waves from an N-port network with units of power and carrying a complex phase. Incident waves at port N are defined by $a_N = V_N^+/\sqrt{Z_0}$; outgoing waves are defined by $b_N = V_N^-/\sqrt{Z_0}$. The S-parameter $S_{xy} \equiv a_X/b_Y = V_X^- \sqrt{Z_Y}/V_Y^+ \sqrt{Z_X}$. Written as matrix, S-parameters take the form:

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} S_{11} = b_1/a_1 & S_{21} = b_2/a_1 \\ S_{12} = b_1/a_2 & S_{22} = b_2/a_2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \tag{1.14}$$

Intuitively, S_{xy} describes the transformation a signal's amplitude (gain) and phase at incident port x into output (measurement) port y. Parameters on the diagonal represent input and output at the same port, and thus represent reflection coefficients. Off-diagonal parameters represent coupling from one port to another, and can thus be thought of as transmission coefficients. The S-matrix, therefore, describes the gain and phase shift that a 2-port network imparts on incident current and voltage waves from any input port. A lossless, gainless device is described by a unitary matrix, as any incident power can only be transferred to another port (not absorbed by the device itself). Further, the S-matrix must be symmetric for reciprocal devices - ones which would cause the same gain and phase shift no matter where the incident waves arrive at port 1 or port 2.

1.2.1 Electrically small series and shunt elements

Many of our samples are electrically small, meaning $\lambda/L_0 \ll 1$ for feature size L_0. We will analyze lumped element devices using ABCD coefficients, convert them to S-parameters (using formulas derived in [5]), and interpret the S-parameters as reflection or transmission coefficients. We consider two electrically small circuit elements (the series

impedance and the shunt admittance) in a transmission line of impedance Z_0, which are pictured in 1.2.

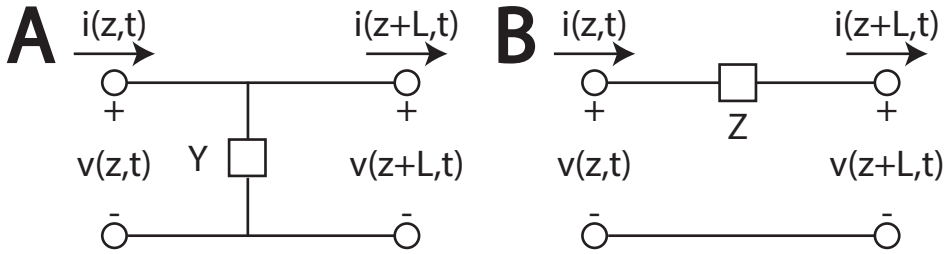

Figure 1.2: Electrically small, simple lumped element circuits for (A) a shunt admittance and (B) a series impedance.

We apply Kirchoff's voltage and current laws to the shunt admittance (Fig. 1.2A). This produces $v(z,t) = v(z+L,t)$ and $i(z,t) - i(z+L,t) - Yv(z,t) = 0$, which can be collapsed into an ABCD coefficient matrix. This is converted into an S-matrix using the formulas in [5]. The resulting S-matrix is:

$$\overline{\overline{S}} = \begin{bmatrix} \frac{-YZ_0}{2+YZ_0} & \frac{2}{2+YZ_0} \\ \frac{2}{2+YZ_0} & \frac{-YZ_0}{2+YZ_0} \end{bmatrix} \tag{1.15}$$

The S-matrix is symmetric along the diagonal and anti-diagonal, confirming that the lumped admittance is symmetric from the forward and reverse directions. The reflection coefficient ranges from -1 (infinite admittance - a short circuit) to 0 (zero admittance - a matched transmission line equivalent to placing no shunt element whatsoever). We also consider the total power transfer from the input, which is given by $\sum_n^N |S_{n1}|^2$. This is:

$$\sum_n^N |S_{n1}|^2 = \frac{4 + YY^*Z_0Z_0^*}{4 + YY^*Z_0Z_0^* + 2Z_0(Y + Y^*)} \tag{1.16}$$

The total power transfer is never greater than unity and can be unity if and only if $Y = -Y^*$. This can only occur if the admittance is purely imaginary. That is: a purely

reactive shunt circuit element will only redistribute the power between the reflected and transmitted waves. Any resistive component will dissipate some of the incident power.

Kirchoff's current and voltage laws can also be applied to the series element, producing an S-matrix of the form:

$$\overline{\overline{S}} = \begin{bmatrix} \frac{Z/Z_0}{2+Z/Z_0} & \frac{2}{2+Z/Z_0} \\ \frac{2}{2+Z/Z_0} & \frac{Z/Z_0}{2+Z/Z_0} \end{bmatrix} \tag{1.17}$$

Much like the shunt element S-matrix, the series element S-matrix is symmetric from both the forward and reverse directions. The reflection coefficient ranges from 0 (zero impedance - a matched transmission line equivalent to not placing a series element) to 1 (infinite impedance - an open circuit). The total power gain can also be shown to never exceed unity, and equal unity only in the case of a purely reactive circuit element.

Reactive circuit elements shift the signal to earlier times (leading) or later times (lagging phase). For both series and shunt geometries, resistive circuit elements decrease the magnitude of the reflected signal, but can only produce 0 or π phase shifts. Thus, the leading or lagging shift of a reflected or transmitted pulse's peak can instantly reveal the approximate magnitude and class of lumped element responsible, which is calculated from eqns. 1.15 and 1.17.

A summary of the analysis for electrically small series and shunt elements is in Table 1.4. The phase shift direction of an element placed in series is opposite that of the same element placed in shunt. In addition, the phase shift directions of an element's transmission and reflection coefficients must be opposite. The only Fresnel coefficients that can ever be negative are the reflection coefficients of shunt elements; all other coefficients are positive.

Quantity	Series C	Series L	Shunt L	Shunt L
Γ	$\frac{Z/Z_0}{2+Z/Z_0}$	$\frac{Z/Z_0}{2+Z/Z_0}$	$\frac{-YZ_0}{2+YZ_0}$	$\frac{-YZ_0}{2+YZ_0}$
Phase(Γ)	$atan(-2\omega CZ_0)$	$atan(2Z_0/\omega L)$	$atan(2/\omega CZ_0)$	$atan(-2\omega L/Z_0)$
Type	Lagging	Leading	Leading	Lagging
T	$\frac{2}{2+Z/Z_0}$	$\frac{2}{2+Z/Z_0}$	$\frac{2}{2+YZ_0}$	$\frac{2}{2+YZ_0}$
Phase(T)	$atan(\omega C/2Z_0)$	$atan(1/2Z_0\omega L)$	$atan(-Z_0/2\omega C)$	$atan(\omega LZ_0/2)$
Type	Leading	Lagging	Lagging	Leading

Table 1.4: Leading and lagging phase shift classification for simple lumped element capacitors (impedance $Z = 1/j\omega C$) and inductors (impedance $Z = j\omega L$).

1.2.2 Reflection and transmission sensitivity

The reflection and transmission coefficients of Table 1.4 can also be used to examine a spectrometer's sensitivities to small changes in admittance or impedance. As an example, Taylor expand the reflection coefficient in YZ_0 about point Y_SZ_0.

$$\Gamma(YZ_0) = \Gamma(Y_SZ_0) + \frac{\partial\Gamma(YZ_0)}{\partial(YZ_0)}|_{Y_SZ_0}(YZ_0 - Y_SZ_0) + \mathcal{O}(Y^2Z_0^3) \tag{1.18}$$

The first order term may be thought of as the sensitivity $(\partial\Gamma(YZ_0)/\partial(YZ_0)|_{Y_0Z_0})$ times the signal size (YZ_0 - Y_SZ_0). That is: extremize the following sensitivity (D_S) to produce the operating choice of YZ_0 around which small changes will be easiest to measure.

$$\frac{\partial D_S}{\partial(Y_SZ_0)} = \frac{\partial}{\partial(Y_SZ_0)}\left(\frac{\partial\Gamma(YZ_0)}{\partial(YZ_0)}|_{Y_SZ_0}(YZ_0 - Y_SZ_0)\right) \tag{1.19}$$

The choice of YZ_0 that maximizes the contrast between small changes is when $Y_S = 2/Z_0$. One can find the Y_SZ_0 maximizing contrast in transmission - which also occurs at $Y_S = 2/Z_0$. This is no accident: $Y_S = 2/Z_0$ is the point at which half of the signal reflect and half transmits from a lossless device. Small losses in reflection will amount to small increases in transmission, and vice versa. In the series geometry, contrast in both

reflection and transmission is optimized when $Z_S = 2Z_0$, which, once again, the operating point at which half of the signal transmits and half of the signal reflects from the sample.

This is a vital consideration in designing sample geometry. If a shunt sample's admittance is too large, small changes in impedance from quantum phenomena will not be measurable. It is *critical* to design samples geometries for their admittance and transmission lines with appropriate impedances to observe subtle spectroscopic features. The normalized sensitivity versus Y_SZ_0 for shunt geometries and Z_S/Z_0 for series geometries is plotted in Fig. 1.3. To get a sensitivity no smaller than half of the maximum, design YZ_0 or Z/Z_0 to be within 0.34 and 11.69.

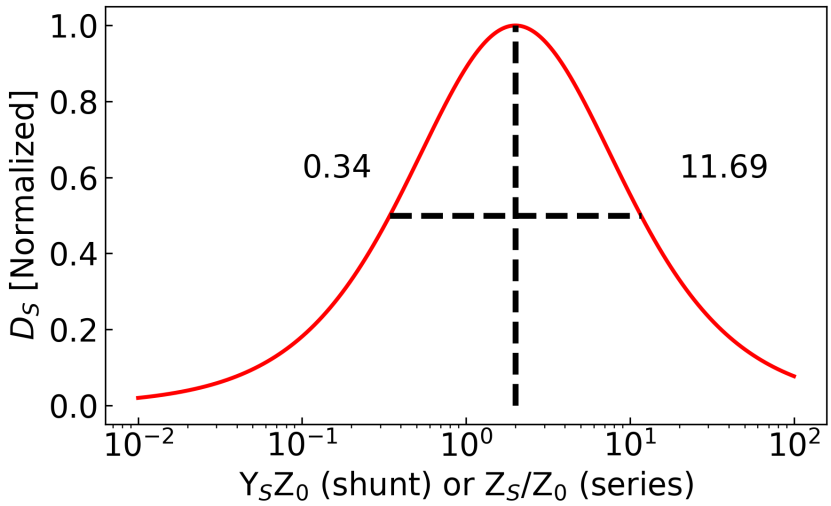

Figure 1.3: Signal contrast versus YZ_0 for shunt geometries / Z/Z_0 for series geometries, with unit normalization. Contrast is optimized at Y_SZ_0 or $Z_S/Z_0 = 2$, when a lossless sample reflects and transmits half of the incident signal.

1.2.3 Signal flow diagrams

Even the simplest realistic distributed circuits include the cascade, cascode, and/or interconnected cascode combinations of many 2-port devices, including transmission lines. For example, the most simplified form of a spectrometer contains a source, transmission

line, sample, and readout, plus leakage terms and radiative coupling. Describing these elements by solving cascade and cascoded combinations of distributed currents and voltages with ABCD matrices would be painful. Describing these elements by their effect on an incoming or outgoing wave, however, is more compact and intuitive. Each element or parasitic effect is quantified by a frequency-dependent linear, time-invariant transfer function. We are already familiar with reflection coefficients that link incoming wave a_1 to outgoing wave b_1, and forward transmission coefficients S_{21} that link incoming wave a_1 to outgoing wave b_1.

Additional transfer functions quantify additional parasitic elements and leakages in the circuit. Signal flow diagrams can be drawn to show the cascade and cascode network from one circuit point to the other. Such a diagram for the standard vector network analyzer measurement of a 2-port network are shown in Fig. 1.4. An input port at the left (1) is connected to an output port at the left (1) and an output port at the right (2), representing reflection and transmission, respectively. In addition to the usual S-matrix of reflection and transmission coefficients, there are parasitic transfer functions from inter-port isolation / crosstalk (E_x), leakage from input 1 to output 1 that sees no other transfer functions, called directivity (E_D), reflection tracking for systematic reflection errors (E_r), source-mismatch reflections (E_s), load-mismatch reflections (E_L) and load tracking for systematic transmission errors (E_t). The flow of these propagators is cumbersome to describe with words, but simple to describe diagrammatically.

Computing the transfer function relating input (1) to either output (1) or output (2) requires simplifying the diagram. This is done through a set of simplification rules reminiscent of Feynman diagram reduction [5] or through Mason's laws for signal diagrams. For details, see [5].

Signal flow diagrams are the basis for designing and analyzing all THz circuits. In order to measure the desired reflection (S_{11}) or transmission coefficient (S_{21}), we must

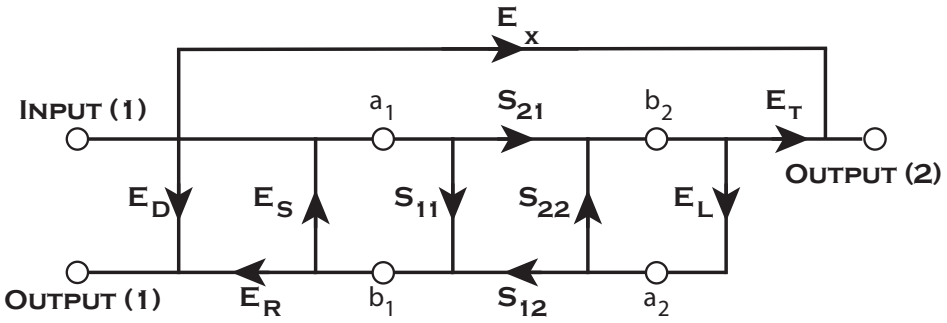

Figure 1.4: Signal flow diagram for measurement of a 2-port network, with systematic transfer functions. The transfer functions' complicated cascade and cascode connections would be difficult to represent without a diagram or with current / voltage descriptions.

know how they relate to additional propagators in the measurement circuit. The additional transfer functions must then be minimized or compensated for with a calibration. Removing systematic transfer functions is made easier by the time domain nature of on-chip THz. Since measurements at ports 1 and 2 are time-resolved, one can often identify the path of propagators that relates a measured signal to the input by is time delay. This greatly reduces the number of calibrations required to correctly de-embed a device's S-parameters from measurement systematics.

1.2.4 Coplanar striplines

The coplanar stripline (CPS) is a two-conductor transmission line geometry fabricated on the surface of a dielectric substrate. The transmission line is made of two metallic lines, often with a thickness of about 100 nm on substrate of thickness H. The metallic lines have width S and are separated by a gap of width W. A schematic is shown in Fig. 1.5. The electric field profile of a CPS are the primary reasons for its choice. At high frequency, current is restricted to only the innermost few skin depth on the inside of the metallic lines. This supports an electric field pointing from the innermost skin depths on the left metal to those on the right, forming an electric dipole-like field[6].

The resulting electric field points from one metal line to the other and is parallel to the substrate between the two conductors, with slight variations at larger heights.

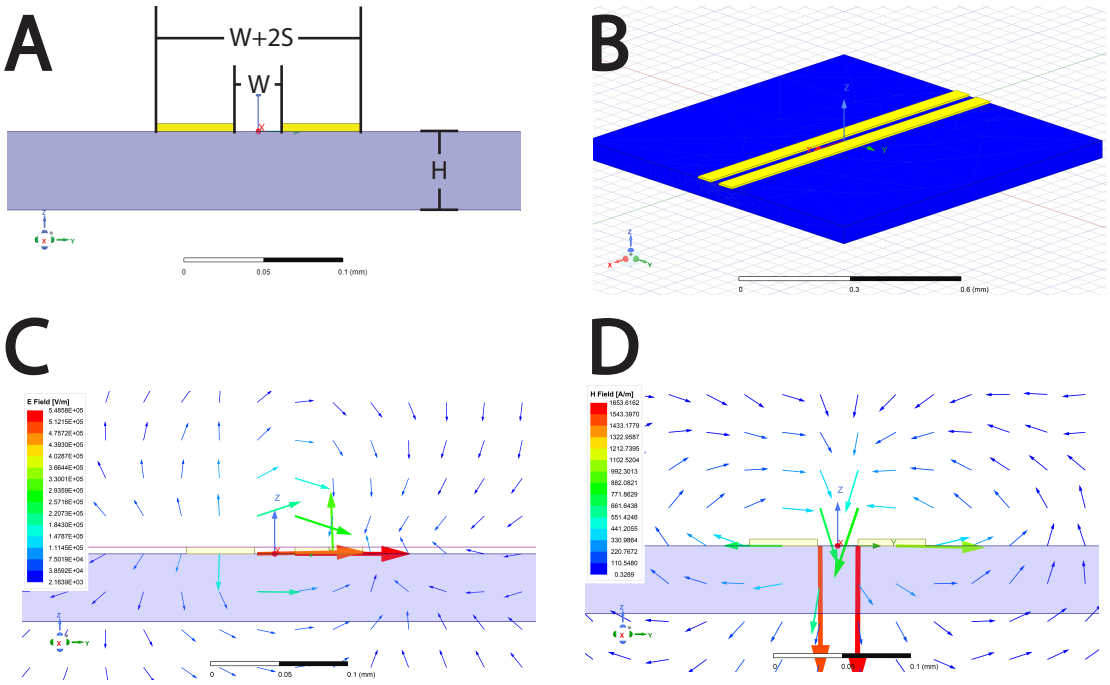

Figure 1.5: CPS geometry, from figures generated by finite element analysis in with Ansys HFSS. (A) Transverse geometry, showing two conductor strips with separation W and width S. (B) Dimetric view of the CPS geometry on a blue substrate. (C) $\overline{E}$ field lines. (D) $\overline{H}$ fields lines.

The effective relative permittivity can be shown to be $(1 + \sqrt{\epsilon_r})/2$ to zeroth order, which is the average dielectric constant in the substrate (dielectric) and superstrate (vacuum). The low-frequency phase velocity of signal in a CPS can be written as$\approx c_0/\sqrt{\epsilon_{eff}}$. The $\overline{E}$-field is evenly distributed in the substrate and superstrate, but the $\overline{D}$-field is concentrated in the dielectric. An analytic formula for the impedance can be derived as functions of W, S, and substrate ϵ, providing an impedance of about 165 Ω for $\epsilon_r = 2.26$, W $= 30$ μm, and S $= 50$ μm. These formulae cumbersome and only strictly true as $\omega \to 0$, these formulas give an estimate that can later be refined with numerical computation [7].

CPS-based devices have great advantage in that they make both series and shunt

device mounting possible. Series devices [8] replace the series transmission line with the sample; shunt devices [9] place the sample as a connection between the two metallic strips. In addition, the CPS is a balanced transmission line - meaning the same current magnitude flows on each conductor which allows for differential signaling.

The CPS does, however, have a great disadvantage in that the transmission line can support more than just the fundamental quasi-TEM mode. These additional modes are allowed field distributions that are often very lossy, make engineering impedance junctions more difficult, and obfuscate interpretation of reflection from devices. The CPS also supports TE_0 and TM_0 modes with at zero frequency[7]. The TE_0 mode has field lines parallel to the substrate surface, just like the quasi-TEM mode. Modes can easily mix between the two if one is not careful. The next highest order mode is the TE_1 mode for substrate of thickness H has lower cutoff[10]:

$$f_c^{TE_1} = c_0/2H\sqrt{\epsilon_r - 1} \tag{1.20}$$

Eqn. 1.20 highlights the importance of careful substrate choice. For 50 μ-thick cyclo-olefin polymer substrates, $f_c^{TE_1} = 2.6$ THz. For 50 μm thick quartz, $f_c^{TE_1} = 1.8$ THz. For standard 500 μm thick sapphire substrates, $f_c^{TE_1} = 0.09$ THz, which is much too low for most THz spectroscopy. Higher order modes of the CPS can be best avoided by using thin substrates with small dielectric constants.

Coplanar waveguides (CPW) transmission lines with three conductors (ground, signal, ground, or GSG) are generally considered to be of higher quality than the CPS. The CPS and CPW are complimentary structures[3], so their impedances and phase velocities (and, therefore, performance) are closely related through Babinet's Principle. CPWs are generally considered to be the superior coplanar transmission line at frequencies above

[3]Complimentary electromagnetic structure, here, means that superimposing the metal from an infinitely thin PEC CPS and infinitely thin PEC CPW produces a single infinite sheet of PEC

10 GHz, but will not be used in this dissertation. CPWs are more complicated than CPS, as CPWs support arbitrary superpositions of propagating waves in the left and right slots. This makes it more difficult to integrate both series and shunt samples, as well as to make inductive couplers. CPWs also require regular connection between the two ground planes. To avoid complication, we restrict our use to CPS transmission lines.

1.3 Material response functions

The present time response of any linear, time-invariant system to an external stimulus can be described by the impulse response of that system. A linear system output $y(t)$ depends on the past and present values of an external input $x(t)$ in the form:

$$y(t) = \int_{-\infty}^{t} h(t)x(t-t')\,dt' = \int_{-\infty}^{t} h(t-t')x(t)\,dt' \tag{1.21}$$

h(t) is the impulse response, which is a measure of the system's output if the input x(t) were a unit-weighted Dirac delta function. If the input could be written as two separated delta functions (namely: $x(t) = x_1\delta(t_1) + x_2\delta(t_2)$), the output would simply be the sum of the impulse response occurring at t_1 plus a second impulse response occurring at t_2 (namely: $y(t) = x_1h(t_1) + x_2h(t_2)$). One can imagine writing a continuous input as an infinitely dense sum of delta functions and performing the same exercise, only to find that y(t) becomes an infinitely dense sum of impulse responses. The limit of this process as the time between delta functions shrinks to zero, is eqn. 1.21. The integral's lower bound allows x(t) to contribute to the integral at $-\infty$; the upper bound stops x(t)'s contribution at the present time t. Future values of x, therefore, cannot influence the present value of y(t) and thus eqn. 1.21 assumes that the system response is causal.

The famous convolution theorem allows us to apply Fourier transforms to x(t), h(t),

and y(t) and write:

$$Y(\omega) = H(\omega)X(\omega) \tag{1.22}$$

By operating in the frequency domain, one can simply find $H(\omega) = Y(\omega)X(\omega)$. It is important to note that $Y(\omega)$, $H(\omega)$, and $X(\omega)$ are all complex-valued quantities. We notationally separate the real and imaginary parts as $H(\omega) = H_1(\omega) + jH_2(\omega)$. Different professions have different names for the quantities $h(t)$ and $H(\omega)$. "Impulse response" usually refers to the time domain quantity $h(t)$ and "Green's function", "transfer function function", or "response function" applies to the quantity $H(\omega)$. We maintain this convention.

The aim of linear THz spectroscopy will be to apply external stimuli in the form of THz-frequency $\overline{E}$ and $\overline{H}$ fields, extract the Green's functions $\overline{\overline{\sigma}}(\omega)$, $\overline{\overline{\epsilon}}$, or $\overline{\overline{\mu}}$ and infer the nature the of the underlying system.

1.3.1 Equivalence of $\overline{\overline{\epsilon}}$, $\overline{\overline{\mu}}$, and $\overline{\overline{\sigma}}$ at AC frequencies

The response functions $\overline{\overline{\sigma}}$, $\overline{\overline{\epsilon}}$, and $\overline{\overline{\mu}}$, no longer remain independent functions at non-zero frequencies. To see this, assume a uniform, isotropic medium and manipulate the time harmonic Maxwell-corrected Ampere's law (bottom, eqn. 1.1) by expanding $\overline{H}$ and $\overline{E}$ in terms of the electric polarization density and magnetization density.

$$\begin{aligned} \frac{1}{\mu_0}\nabla \times (\overline{B} - \overline{M}) &= \overline{J_e} + j\omega(\epsilon_0\overline{E} + \overline{P}) \\ \nabla \times \overline{B} &= \mu_0(j\omega\epsilon_0\overline{E} + \overline{J_e} + \nabla \times \overline{M} + j\omega\overline{P}) \end{aligned} \tag{1.23}$$

The terms on the right side of eqn. 1.23 are all currents responsible for sourcing a curl of $\overline{B}$. The $\mu_0(j\omega\epsilon_0\overline{E})$ term arises from a time-dependent $\overline{E}$ field. The $\mu_0\overline{J_e}$ term is from an injected or impressed current, via Ohm's law (eqn. 1.3). The $\mu_0(\nabla \times \overline{M})$ term is a current

that sources a magnetic field through the bound magnetization - the material response to an external magnetic field. The final term, given by $\mu_0(j\omega\overline{P})$, is a current associated with the electric polarization - the material response to an external electric field. The Ohmic current, bound magnetization current, and the bound polarization current, are related to the complex-valued material Green's functions that spectroscopists seek to measure and interpret.

At zero frequency, the Ohmic conductivity is a long-range phenomenon, wherein charge may be transported along very long length scales. Bound magnetization and polarization densities, however, are local, short range phenomena arising from ion-scale local electric and magnetic dipoles. They are associated with a length scale set by the inter-ionic spacing, and only become global phenomena when many ions are placed in close proximity. At high frequencies, the length scale which a charge travels per cycle reduces and the conductivity becomes a more local phenomenon. There is no experimental way to distinguish between the polarization, magnetization and Ohmic currents, as they all contribute to the curl of $\overline{B}$. As such, they are no longer independent parameters.

Measurement is capable of resolving α and β, meaning only two independent material parameters can be specified without overdetermining a material system. To see this, we revisit the definition of γ in eqn. 1.8. It is apparent that real values of σ or imaginary values of ϵ and μ would make extinction coefficient α non-zero, transferring energy from the propagating electric field to the medium. Imaginary values of σ or real values of ϵ and μ would make α zero, and thus let the plane wave propagate without loss.

It is also popular to use a hybrid description with the dielectric loss tangent δ_e or loss function. The dielectric loss tangent is defined by ignoring any externally injected currents and cleverly grouping the total current density from the Maxwell-corrected Ampere's law:

$$\begin{aligned} \nabla \times \overline{H} &= \sigma \overline{E} + j\omega\epsilon \overline{E} \\ \nabla \times \overline{H} &= j\omega\epsilon \left(1 - j\frac{\sigma}{\omega\epsilon}\right) \overline{E} = j\omega\epsilon(1 - j\tan\delta_e)\overline{E} \end{aligned} \tag{1.24}$$

The dielectric loss tangent $\tan\delta_e \equiv \sigma/(\omega\epsilon)$ is conventionally computed with purely real σ and ϵ. It compactly quantifies the ratio of the (resistive) Ohmic current to the (reactive) polarization density current. Small loss tangents correspond to low loss; large loss tangents correspond to high loss. Loss tangents are a convenient formulation because the extinction coefficient can be conveniently expanded for small loss tangents[4] as $\alpha \approx \frac{\omega}{2}\sqrt{\mu\epsilon}tan\delta_e$.

Though some THz and microwave engineers prefer to describe materials by specifying their real dielectric constant and dielectric loss tangent, this description is best avoided for calculations with time domain THz spectroscopy. The loss tangent is explicitly frequency-dependent, in addition to implicit frequency dependences of σ and ϵ. The explicit frequency dependence of $\tan\delta_e$ is more appropriate for narrowband design and not for the broadband THz spectrometer designs detailed in this dissertation. That said, specifying the loss tangent at a specific frequency is a quick way to quantify the relative lossiness of (substrate) materials in a single number. As long as the loss tangent is specified at the same frequency for every listed material, one can easily compare, say, sapphire, RT Duroid, silicon, cyclo-olefin polymer, and Pyrex. Anyone purchasing microwave and THz-compatible materials from manufacturers who specify $\tan\delta_e$ would be wise to the check the frequency at which δ_e is specified. Manufacturers with dubious morals are known to specify loss tangents at unrealistically small frequencies to hide excess loss from the customer.

The most popular choices of material parameters describe a material response using (a) real and imaginary conductivity, (b) real conductivity and real permittivity, or (c)

[4]Some literature will use $\tan\delta_e \equiv \epsilon_2/\epsilon_1$, this is fully equivalent

real and imaginary permittivity. Sometimes, the inverse permittivity is specified in form of a loss function. This is written as $-Im(1/\epsilon) = \epsilon_2/|\epsilon^2|$. Some spectroscopists prefer optical conductivities to be specified via units of Ohms, not Siemens. This loss function is the analogous quantity for dielectric descriptions - a loss specified in m/Farad instead of Farad/m. Physically, the loss function is proportional to the time derivative of a semiclassical electron's energy [11].

Since the above three descriptions are not independent, so one should be able to convert freely between the different descriptions. This is done by interpreting the material response as a purely bound (polarization) current and identifying the polarization current density $j\omega\epsilon(\omega)\overline{E}$ as the sum the material's Ohmic and material's polarization current densities $\sigma\overline{E}$ and $j\omega\overline{P}$ from eqn. 1.24. This produces:

$$\begin{aligned} j\omega\epsilon(\omega) &= \sigma_c + j\omega\epsilon_d \\ \epsilon(\omega) &= \epsilon_d - j\sigma_c/\omega \end{aligned} \tag{1.25}$$

The conductivity σ_c and permittivity ϵ_d, here, are nominally frequency-independent material parameters; the permittivity $\epsilon(\omega)$ is a purely dielectric description of the material response.

We can similarly describe the material's response purely with a conductivity $\sigma(\omega)$. By interpreting the material response purely as a free current, we identify conducted current $\sigma(\omega)$ as the sum of the material's Ohmic and material's polarization current densities $\sigma\overline{E}$ and $j\omega\overline{P}$ from eqn. 1.24. As for the dielectric, this produces:

$$\sigma(\omega) = \sigma_c + j\omega\epsilon_d \tag{1.26}$$

It is also possible to use real and imaginary μ, plus combinations of real and imaginary μ with other material parameters, but this is not common for free space THz measure-

ments and unwise for on-chip THz measurements. We will explore the reason in Ch. 3. Material response to an AC frequency is summarized in Table 1.5.

Quantity	Resistive part	Reactive Part	Commonly used at THz Freq.
$\sigma_1 + j\sigma_2$	σ_1	σ_2	Yes
$\epsilon_1 + j\epsilon_2$	ϵ_2	ϵ_1	Yes
$j\omega\epsilon_1 + \sigma_1$	σ_2	$j\omega\epsilon_1$	Yes
$\mu_1 + j\mu_2$	μ_2	μ_1	No
ϵ_1, $LF = -Im(1/\epsilon)$	LF	ϵ_1	No
ϵ_1, δ_e	δ_e	ϵ_1	No

Table 1.5: Summary of reactive and resistive non-zero frequency material descriptions.

1.3.2 Conjugate symmetry and the Kramers-Kronig relations

Response functions can be decomposed into their even and odd components. Consider the positive and negative frequency Fourier transforms $F(\omega)$ of the time domain function f(t) must be:

$$\begin{aligned} F(\omega) &= \int_{-\infty}^{\infty} f(t)e^{-j\omega t}\, d\omega \equiv F_1(\omega) + jF_2(\omega) \\ F(-\omega) &= \int_{-\infty}^{\infty} f(t)e^{j\omega t}\, (-d\omega) \equiv F_1(-\omega) + jF_2(-\omega) \end{aligned} \tag{1.27}$$

From eqn. 1.27 we see that if f(t) is a purely real function, then $F(\omega)$ must be conjugate-symmetric: $F(\omega) = F^*(-\omega)$. Suppose now that $F(\omega)$ is purely real (so $F_2(\omega) = F_2(-\omega) = 0$). This implies that $F_1(\omega) = F_1(-\omega)$, so we conclude that purely real $F(\omega)$ is even. Suppose now that $F(\omega)$ is purely imaginary (so $F_1(\omega) = F_1(\omega) = 0$). This implies that $F_2(\omega) = -F_2(-\omega)$, so we conclude that purely imaginary $F(\omega)$ is odd. When performing simulation calculations of the response of a theoretical model, it is tantamount to enforce this conjugate symmetry to the modelled material response function. Failure to do so can lead to non-physical results.

All response functions considered in this dissertation will be real-valued in the time domain, as imaginary-valued time domain response functions are not physically meaningful. Real-valued time domain response functions, therefore, can be decomposed into their even (real) and odd (imaginary) components. Such decomposition manifests in a grand way - as the basis for the Kramers-Kronig relations. THz spectroscopy will be used to the extract both the real and imaginary parts of material response functions to external electric and magnetic fields. At first inspection, one might think that the real and imaginary parts of response functions are not related, since the reactive and resistive (dissipative) parts of a response function seem to occur by different mechanisms. This is not so. The reactive and resistance parts of any causal response function are related via the Kramers-Kronig relations. The Kramers-Kronig relations may be derived arduously through the application of Cauchy's residue theorem. There is an alternate, more intuitive derivation in the time domain, using the conjugate symmetry of response functions we just derived.

As shown above, real-valued time domain response functions produce conjugate symmetric frequency domain response functions. If a time domain response function is causal, its value must be zero for all times $t < 0$. This can only be possible if the even and odd parts are equal and opposite for negative times and equal and bear the same sign at zero and positive times. This can be written, with the signum function $sgn(t)$, as:

$$\begin{aligned} f_{odd}(t) &= sgn(t)\ f_{even}(t) \\ f_{even}(t) &= sgn(t)\ f_{odd}(t) \end{aligned} \tag{1.28}$$

We compute the Fourier transform of 1.28. The time domain product becomes a frequency domain convolution. We have shown that $f_{even}(t) \Rightarrow F_1(\omega)$ and $f_{odd}(t) \Rightarrow F_2(\omega)$. Computing the Fourier transform of the $sgn(t)$ produces the frequency domain form of the well-known Kramers-Kronig relations:

$$F_2(\omega) = \frac{-1}{\pi} P \int_{-\infty}^{\infty} \frac{F_1(\omega)}{\omega' - \omega}\, d\omega'$$
$$F_1(\omega) = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{F_2(\omega)}{\omega' - \omega}\, d\omega' \tag{1.29}$$

P is the Cauchy principal value, coming from the Fourier Transform of the signum distribution. The mathematical operation performed to compute the Kramers-Kronig relations is better known as the Hilbert Transform, which is most often evaluated numerically. The Hilbert transform is a frequency-domain convolution, and thus obeys all the basic properties that typical convolution operations do: linearity, scaling, and argument shift, such that $\alpha F_1(\beta(\omega - \omega_0)) \Leftrightarrow \alpha F_2(\beta(\omega - \omega_0))$.

Eqn. 1.29 has profound implications. The response functions of all of the physical models we will derive and use are causal, and thus they must obey the Kramers-Kronig relations. Any frequency-dependence of the real part of any causal response function implies that there is a frequency-dependence of the imaginary part, and vice-versa. Consider, as an example, optical conductivity $\sigma(\omega)$. Any frequency dependence whatsoever in the *reactive* part σ_2 implies that there is a related frequency dependence in the *resistive* part. The reactive and resistive responses of a material are not independent.

To see this, we show that only structureless reactive parts of response functions produce structureless resistive parts of response functions. If we substitute $F_2(\omega) = 0$ at all frequencies into the bottom line of eqn. 1.29, the integral reduces to a constant. The only solution with a featureless resistive response requires a featureless reactive response, and vice versa. Put simply: change in material response function versus frequency *always* implies a change in the loss, and those changes can occur over different frequency bandwidths. Any frequency dependence in response function implies loss at some frequency. Since all materials are transparent in the $\omega \to \infty$ limit, all materials with non-unit parameters at zero frequency will be lossy at some frequency. This also suggests that

materials with larger low-frequency response functions will incur more loss than materials with smaller low-frequency response functions.

Furthermore, the nature of the convolution by the function $1/\omega$ means that narrowband changes in, say, the real part of the response function can have wideband consequences on the imaginary part, and vice versa. To see this, consider a causal response function $F(\omega)$ whose real part is a Gaussian $e^{-\omega^2/2}/\sqrt{2\pi}$. It's Hilbert transform, and therefore imaginary part, is $\sqrt{2}DawsonF(\omega/2)/\pi$, where $DawsonF()$ is the Dawson integral. The real and imaginary parts of $F(\omega)$ are plotted in 1.6, for $\omega_0 = 0$ (left) and ω_0 (right). The majority of the structure in the real part of the both Gaussians is localized within a few units around the function's mean. The imaginary part, however, has structure for 10 to 20 units on either side of the mean.

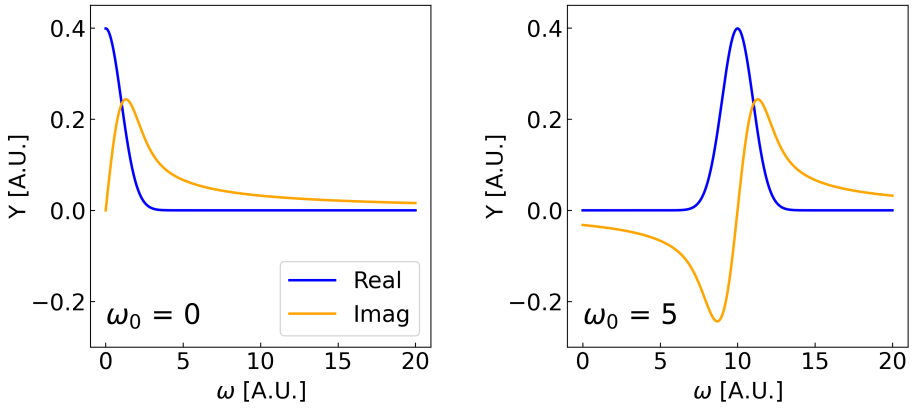

Figure 1.6: Real and imaginary parts of sample response functions, shifted by $\omega_0 = 0$ (left) and 5 (right). Despite somewhat narrowband structure in the real part, the imaginary part shows structure over a broader frequency range.

Experimental use of the Kramers-Kronig relations often takes a different form. If we multiple the numerator and denominator of eqn. 1.29 by $\omega' + \omega$ and use the conjugate symmetry ($F_1(\omega)$ is even, $F_2(\omega)$ is odd) to simplify, we find a form of the Kramers-Kronig more often used numerically.

$$F_2(\omega) = \frac{-2\omega}{\pi} P \int_0^\infty \frac{F_1(\omega')}{\omega'^2 - \omega^2}\, d\omega'$$
$$F_1(\omega) = \frac{2}{\pi} P \int_0^\infty \frac{\omega' F_2(\omega')}{\omega'^2 - \omega^2}\, d\omega' \tag{1.30}$$

Eqn. 1.30 is the form most often seen in theoretical treatments of spectroscopy. Many experiments can only measure the magnitude of a response function $F(\omega)$ without directly extracting the phase information (for example: Fourier transform infrared spectroscopy). One can still find the repsonse function phase by solving a system of equations: $|F(\omega)| = \sqrt{F_1^2 + F_2^2}$ and the Kramers-Kronig relations. The system of equations can be solved numerically for F_1 and F_2. This is exceptionally difficult to do in practice. It is much easier to measure the $|F(\omega)|$ and use the Bode gain-phase relationship, a form of the Kramers-Kronig relation that relates the magnitude and phase ($\phi(\omega)$) of a causal transfer function [12]:

$$\phi(\omega) = \frac{-1}{\pi} \int_{-\infty}^{\infty} \frac{\partial \ln |F(\omega')|}{\partial \nu} \ln\left(\coth\left(|\nu|/2\right)\right)\, d\nu \tag{1.31}$$

Here, $\nu \equiv \ln \omega'/\omega$. The $\ln \coth |\nu|/2$ term acts as a broadened delta function in ν. To demonstrate the power of the Bode gain-phase relation, consider a power law transfer function $F(\omega) = A\omega^{-n}$. The derivative $\partial \ln |F(\omega')|/\partial(-\ln(\omega/\omega')) = n$, the integral of $\ln \coth\left(|\nu|/2\right)$ evaluates to $\pi^2/2$, and the resulting phase is $-n\pi/2$. We note that the high frequency response of an n-pole low pass filter scales as such a power law at high frequency, and thus the filter phase *must* be $-n\pi/2$.

We are now positioned to answer the question: does the "ideal inductor", taught in sophomore engineering classes, actually exist? The ideal inductor has admittance $1/j\omega L$, with magnitude $|1/\omega L|$, phase[5] $-\pi/2$, and n = 1. Eqn. 1.31 suggests that phase must

[5]Phase of the conductivity (current to voltage relationship), not of the impedance (voltage to current relationship).

be $-\pi/2$, as taught in classrooms. However, we know from eqn. 1.30 that there must be some structure in the real part to produce a frequency dependence in the imaginary. We can deduce that this spectral weight must be at zero-frequency - as any signal at finite frequency would prevent the phase from being $-\pi/2$. One can substitute a zero frequency delta function into the Kramers-Kronig relation and produce the $1/\omega$ response. Thus, the concept of an ideal inductor having only an imaginary conductivity is only true at AC frequencies. We will explore consequences of this in our discussion of superconductors in Ch. 2.

The Bode gain-phase relationship implies that measuring the signal magnitude is enough to resolve the phase. This may lead one to declare direct measurements of field and field's phase overkill, and that one only ever need measure $|F(\omega)|$. Such declaration would be premature. Eqn. 1.30 and 1.31 require integrals over all frequencies. We have also demonstrated that narrowband transitions in F_1 (or F_2) can produce broadband structure in F_2 (or F_1). Failure to integrate over sufficient bandwidth can lure the unsuspecting careless spectroscopist into missing vital spectroscopic structure [13]. An conservative figure of merit for using and trusting the Kramer's Kronig is that trusting a Kramers-Kronig-produced feature at ω_0 requires measurement from $\omega_0/10$ to $10\omega_0$. Let's consider an example: suppose one measures $|F(\omega)|$ from ω_0 to $1000\omega_0$ and compute the real and imaginary parts through the Kramer's-Kronig relations. Without additional information, the phase of the computed response function should only be trusted between 10ω and 100ω. This is quite restrictive, but is not error prone.

Chapter 2

THz spectroscopy of model systems

> "To try to make a model of a condensed matter system by studying its spectrum is like trying to make a model of a grand piano by listening to the noise it makes when thrown down stairs"
>
> Adapted from *Francis Crick*

Having explored the basic response of electric and magnetic fields in matter, we turn our attention to simple theoretical models of quantum matter. The spectroscopy of quantum materials can produce abstruse and complicated real and imaginary response functions. Understanding such measurements begins with comparison to simple models of spectra of simple electronic systems. We develop simple models for uncorrelated electron gasses and oscillators and show that they may also be applied to quantum single- and many-body systems. We then derive the spectrum of superconductors with the Kramers-Kronig relations and examine additional features visible in the THz spectrum. This chapter closes with a brief examination of caveats that not all systems will respond

linearly to THz waves, using the amplitude and phase excitations of a charge density waves as an example.

Many sign conventions are used in literature. We henceforth ignore the dyadic nature of material response functions, unless explicitly noted otherwise. We emphasize that we maintain the following three conventions throughout this and following chapters:

- We always use SI units. $\epsilon(\omega) \equiv \epsilon_r(\omega)\epsilon_0$; $\epsilon(\omega)$ contains an implicit factor of ϵ_0; $\mu(\omega)$ contains an implicit factor of μ_0
- Complex quantities are always written as $Z(\omega) = Z_1(\omega) + jZ_2(\omega)$, even though some literature uses $\sigma(\omega) = \sigma_1(\omega) - j\sigma_2(\omega)$
- The Fourier transform of a time derivative is multiplication by $j\omega$ (see appendix E). This is equivalent to assuming monochromatic time-harmonic parameters of the form $e^{j\omega t}$.

Due to varying sign conventions, the following plots may have positive or negative imaginary parts compared to that seen in literature.

2.1 Semiclassical electron gasses

Semiclassical gasses of charge carriers form the basis of simple modelling. An electron gas is defined as a group of non-interacting electrons. We will show that the same formulations for electron gasses apply to single body quantum systems and generalize the results to interacting many body quantum systems.

2.1.1 Metals: the Drude model

The Drude model first stems from Paul Drude's 1900 extension of the kinetic theory of gasses to carriers in condensed matter systems. Drude assumed that free electrons

would respond to external electric and magnetic fields, but not interact with one another. Drude's model stipulates that electrons travel until they collide with something in the lattice capable of relaxing their momentum to zero. This average time to collision, τ, is called the momentum relaxation time. From the kinetic theory of gasses, we can write a differential equation describing the momentum ($\overline{p} = m\overline{v_d}$) evolution in the presence of an external electric field intensity $\overline{e(t)}$. This is written in the time and frequency domains as:

$$\begin{aligned} \partial_t \overline{p(t)} &= -e\overline{e(t)} - \overline{p(t)}/\tau \\ j\omega \overline{P(\omega)} &= -e\overline{E(\omega)} - \overline{P(\omega)}/\tau \end{aligned} \tag{2.1}$$

Substitute $\overline{P} = m\overline{V_d}$ and $\overline{J_e} = ne\overline{V_d}$ in eqn. 2.1. Simplifying:

$$\overline{J_e} = \frac{ne^2}{m}\frac{1}{1/\tau - j\omega}\overline{E} = \frac{\sigma_0}{1 - j\omega\tau}\overline{E} \tag{2.2}$$

Eqn. 2.2 is the AC extension of Ohm's law (eqn. 1.3), from which we identify the coefficient linking $\overline{J_e}$ to $\overline{E}$ as the optical conductivity. For convenience, define $\sigma_0 = ne^2\tau/m$ and plasma frequency $\omega_p = \sqrt{\sigma_0/\epsilon_d\tau} = \sqrt{ne^2/\epsilon_d m}$, using ϵ_r as the static dielectric constant of the medium holding the electron gas. The Drude model can also be expressed as a permittivity, using eqn. 1.26:

$$\begin{aligned} \sigma(\omega) &= \frac{\sigma_0}{1+\omega^2\tau^2} + j\omega\tau\frac{\sigma_0}{1+\omega^2\tau^2} \\ \epsilon(\omega) &= \epsilon_d(\omega) - \frac{\omega_p^2\tau^2}{(1+\omega^2\tau^2)} + j\frac{\omega_p^2\tau}{\omega(1+\omega^2\tau^2)} \end{aligned} \tag{2.3}$$

Using Table 1.2, we can approximate the optical index as $\sqrt{\epsilon}$. The magnitude of the frequency-dependent reflection coefficient from vacuum at normal incidence can be computed using the Fresnel equation: $R(\omega) \approx |(\sqrt{\epsilon}-1)/(\sqrt{\epsilon}+1)|$. The real and imaginary Drude conductivity and permittivity and resulting reflection coefficient magnitude are

plotted in Fig. 2.1.

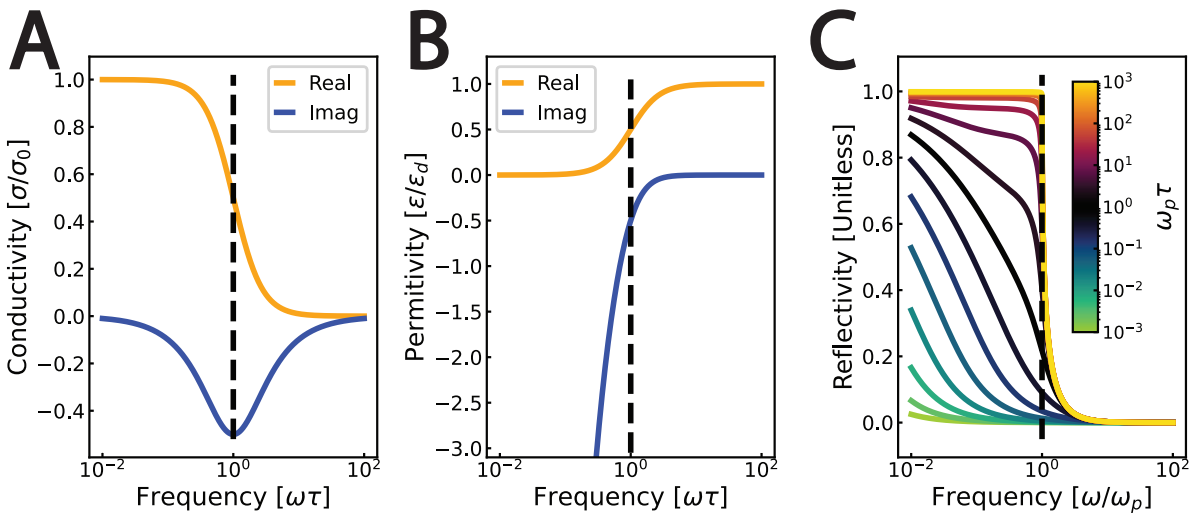

Figure 2.1: Response functions of the Drude model. (A) Real and imaginary optical conductivity. The imaginary conductivity peaks at $\sigma_0/2$ at $\omega = 1/\tau$, which is marked with a vertical line (B) Real and imaginary permittivity, with a vertical line indicating $|\epsilon_1| = |\epsilon_2| = |\epsilon_d/2|$ (C) Reflectivity magnitude for various $\omega_p\tau$, with a vertical limit marking $\omega = \omega_p$. When $\omega_p\tau \gg 1$, the reflectivity abruptly changes from 1 to zero at ω_p.

The Drude model real conductivity (Fig. 2.1A) is σ_0 for $\omega\tau \ll 1$ and approaches zero for $\omega\tau \gg 1$. The real conductivity is a resistive (dissipative) parameter, so we expect that the electron gas to attenuate incident electromagnetic radiation at $\omega\tau \ll 1$ and allow electromagnetic radiation to propagate at $\omega\tau \gg 1$. The imaginary conductivity, which is a reactive parameter, is zero in both the low and high frequency limits. Its magnitude grows to $\sigma_0/2$ at $\omega\tau = 1$, where its magnitude matches that of the real conductivity. Drude electron gasses have a simple signature: that $|\sigma_2(1/\tau)| = |\sigma_1(1/\tau)| = \sigma_0/2$. One can quickly examine a plot of the real and imaginary optical conductivity and seek this signature in a fast effort to identify the electron gas as Drude-conforming or Drude-non-conforming.

The Drude model's permittivity (Fig. 2.1B) contains the exact same intuition, but is

expressed differently. Understanding it improves intuition about the relationship between the pure conductivity and pure permittivity descriptions of a material. The low frequency permittivity is dominated by a $1/\omega$-like response in $\epsilon_2(\omega)$ (the resistive parameter) and near-zero $\epsilon_1(\omega)$ (the reactive parameter). The transition also occurs at $\omega\tau = 1$, with both real and imaginary permittivity taking on $|\epsilon_1(1/\tau)| = |\epsilon_2(1/\tau)| = \epsilon_d/2$.

The Drude model's physical intuition is made obvious by computing the reflectivity of a Drude electron gas. At low frequencies, the metal is reflective because the electrons can oscillate as frequently as the incident electric field, and thus absorb the incident energy. As the frequency increases, the carriers become increasingly unable to oscillate with the driving electric field. The crossover regime occurs at the plasma frequency $\omega_p \equiv \sqrt{ne^2/\epsilon_d m}$. The plasma frequency depends only on intrinsic properties of the electron gas and medium, and is independent of the scattering time τ. If $\omega_p\tau \ll 1$, the momentum relaxation rate is much larger than the plasma oscillation, meaning that carriers scatter their momentum before undergoing a full plasma oscillation. In this case, the electron gas's reflectivity decreases from unity over a broad bandwidth.

If $\omega_p\tau \gg 1$, the momentum relaxation rate is much smaller than the plasma oscillation, meaning that carriers undergo a plasma oscillation before scattering. This produces a sharp transition, wherein the electron gas goes from being completely reflective (opaque) to completely transmissive (transparent). In this limit, the plasma frequency is the maximum frequency at which electrons can oscillate. It is important to remember this interpretation: the plasma frequency constitutes a τ-independent frequency at which an electron gas goes from reflecting incident radiation to transmitting incident radiation (in the limit that $\omega_p\tau \gg 1$).

An illustrative example is the Earth's ionosphere. AM radio is assigned broadcast frequencies less than 2 MHz; FM radio is assigned broadcast frequencies greater than 88 MHz. AM radio is known to transmit long ranges, but FM radio is generally restricted

to line of sight communications. The reason for this lies in the Drude model. Ultraviolet radiation ionizes atmosphere gasses in the ionosphere, creating a free electron density of about n = 1E12 cm^{-3}. Using the electron mass and $\epsilon \approx \epsilon_0$, we estimate a plasma frequency of $\omega_p = 2\pi 9$ MHz. Incident AM radio broadcasts $\omega/\omega_p \ll 1$ see the ionosphere as a metal and are reflected. Incident FM broadcasts $\omega/\omega_p \gg 1$ see the ionosphere as an insulator and are transmitted. The long range transmission of AM broadcasts stems from ionospheric reflections. The lack of such reflections for FM broadcasts occurs because all waves incident on the ionosphere pass into outer space.

The reflectivity $R(\omega)$ can be expanded for $\omega/\omega_p \ll 1$ and $\omega\tau \ll 1$ as:

$$R(\omega) = 1 - 2\sqrt{\frac{2\epsilon_0\omega}{\sigma_0}} \tag{2.4}$$

This is the Hagen-Rubens relation. It may seem to be only a Taylor-series exercise in graduate-level, condensed matter textbooks. It is, but it is also has two practical uses. It is often used to fit the low-frequency limit of Drude systems when the plasma frequency is well *above* the measurement range [14]. This fit generates σ_0, which contains the momentum scattering time, implying that the plasma frequency can be computed without direct measurement. This assumes, however, that the system is truly Drude-like without directly measuring the plasma edge.

The Hagen-Rubens relation has also been used to extrapolate optical parameters when the low frequency limit is well *below* the measurement range. This has been done often in an efforts to increase the frequency range input to the numerical Kramers-Kronig relations [15], extending the frequency range over which the Kramers-Kronig output reliably operates. The downside, once again, is that this doesn't experimentally confirm that a system doesn't have other relevant low-energy physics and may, therefore, lead to fallacious interpretation.

2.1.2 Semiconductors: Lorentz oscillators

The Lorentz model is the second key model used in the analysis of THz spectra. The Lorentz model can be derived by considering the time-harmonic displacement of an electron subject to a damped, driven harmonic oscillator with mass m, damping rate $1/\tau$, spring constant k, and resonant frequency $\sqrt{k/m}$. We begin with the differential equation for the time domain displacement $x(t)$ and solve for the time-harmonic displacement $X(\omega)$.

$$\begin{aligned} m\partial_t^2 x(t) + \frac{m}{\tau}x(t) + m\omega_0^2 x(t) &= -ee(t) \\ X(\omega) &= -\frac{e}{m}\frac{1}{\omega_0^2 - \omega^2 + j\omega/\tau}E(\omega) \end{aligned} \tag{2.5}$$

We write the total polarization $P = -neX(\omega)$ and use the constitutive relations (eqn. 1.3) to write the permittivity of a Lorentz oscillator. This may also be converted into a conductivity with eqn. 1.25:

$$\begin{aligned} \epsilon_r(\omega) &= 1 + \frac{e^2 n}{m}\frac{1}{\omega_0^2 - \omega^2 + j\omega/\tau} \\ \sigma(\omega) &= j\omega\frac{e^2 n}{m}\frac{1}{\omega_0^2 - \omega^2 + j\omega/\tau} \end{aligned} \tag{2.6}$$

Eqn. 2.6 is the usual Lorentzian derived for the damped, driven harmonic oscillator from classical mechanics. Resonant frequency ω_0 is set by by the microscopic details of the model itself. A crude model pictures a purely classical atom, with a Lorentzian description of a spring connecting the large-mass ion to the valence electrons, with resonance frequency $\omega_0 \equiv \sqrt{k/m}$. The conductivity, permittivity, and reflection from a Lorentz model are shown in Fig. 2.2.

As the drive frequency ω approaches the resonance frequency ω_0, the reflectivity increases. Near the resonance frequency, electrons are able to oscillate at large amplitude (as the denominator of $X(\omega)$ is set only by the momentum relaxation rate). This allows

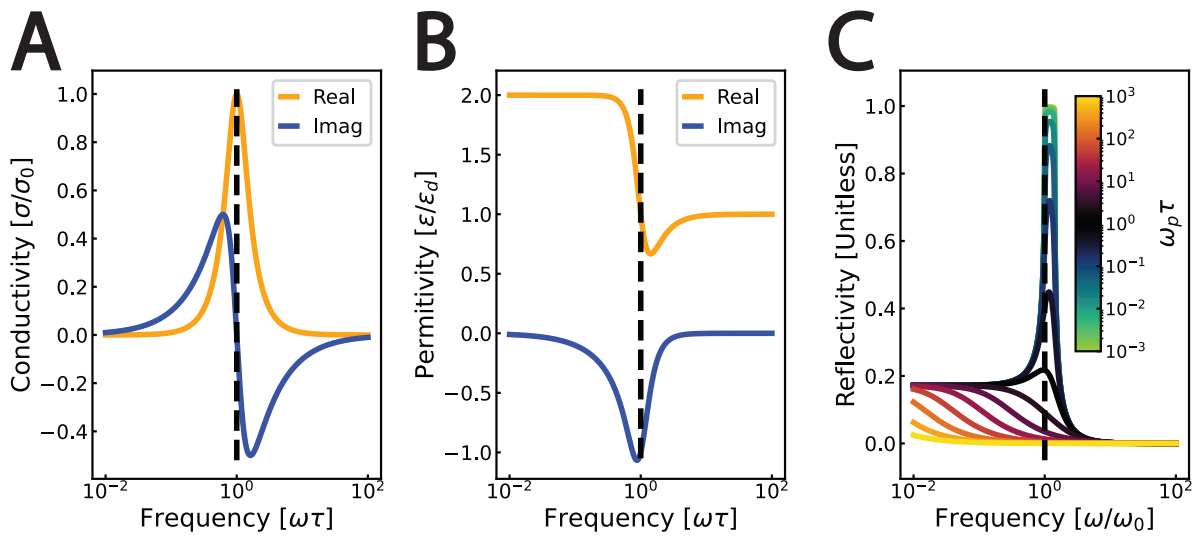

Figure 2.2: Response functions of the Lorentz model (with $\epsilon_0 = 1$ and $\epsilon_\infty = 1$, with $\omega = 1/\tau$ marked by vertical lines. (A) Real and imaginary optical conductivity. The real conductivity peaks at $\sigma_0 = \omega e^2 n/m = 1$. (B) Real and imaginary permittivity (C) Reflectivity magnitude for various $\omega_p\tau$. When $\omega_p\tau \ll 1$, the reflectivity abruptly increases near $\omega = \omega_0$

them to sustain large currents, which reflect the incident field. It, therefore, makes sense that the reflection coefficient R is large in the vicinity of ω_0 and small far from ω_0.

One can show that the Lorentz model reduces to the Drude model in the limit of $\omega_0 \to 0$. This implies that the Lorentz model describes the dynamics of a gas of free electrons, but with a restoring force.

2.1.3 Dielectrics: Debye model and the Jonscher universal dielectric response

Debye's model for dielectric relaxation assumes no free carriers, but a finite time for polarization to accrue in response to a step-function electric field. This time is the dielectric relaxation time, which mathematically behaves as the Drude momentum relaxation time τ, but applied to the dielectric susceptibility $\chi_e(\omega)$ (and not $\sigma(\omega)$). One can show

that this produces a simple dielectric coefficient of the form[16]:

$$\begin{aligned} \epsilon_1(\omega) + \epsilon_2(\omega) &= \epsilon_\infty + \frac{\epsilon_d - \epsilon_\infty}{1 - j\omega\tau} \\ \epsilon_1(\omega) = \epsilon_\infty + \frac{\epsilon_d - \epsilon_\infty}{1 + \omega^2\tau^2} \;\; ; \;\; \epsilon_2(\omega) &= \omega\tau\frac{\epsilon_d - \epsilon_\infty}{1 + \omega^2\tau^2} \end{aligned} \tag{2.7}$$

Here, ϵ_∞ is the high frequency permittivity, ϵ_d is the DC permittivity, and τ_d is the Debye dielectric relaxation time. The loss tangent is easily calculated:

$$tan\ \delta_e = \omega\tau\frac{\epsilon_d - \epsilon_\infty}{\epsilon_d + \epsilon_\infty\omega^2\tau^2} \tag{2.8}$$

Plots of the $\sigma(\omega)$, $\epsilon(\omega)$, and magnitude reflectivity $|R|$ are shown in Fig. 2.3. The permittivity resembles the conductivity of the Drude model, except that ϵ_1 reaches its maximum value at DC. The conductivity is low for $\omega/\tau \ll 1$, and increases at high frequencies. At low frequencies, Debye dielectrics are not lossy. This loss increases at high frequency, causing the reflection to decrease. The Debye model predicts a loss going as $1/\omega^2$ at $\omega\tau \gg 1$.

The Debye model is a perfect example to show that the loss tangent can be used to mislead, conceal, and otherwise nefariously muddle information. The maximum loss in the Debye model occurs when $\partial\epsilon_2(\omega)/\partial\omega = 0$, which occurs when $\omega = 1/\tau$ with $\epsilon(1/\tau) = (\epsilon_d + \epsilon_\infty)/2 + j(\epsilon_d - \epsilon_\infty)/2$. This is the frequency at which an incoming wave will be most attenuated - which is the main concern if one is using an insulator as a dielectric substrate.

The maximum of the loss tangent, however, occurs at a different frequency. The loss tangent is maximized when $\partial(tan\delta_e)/\partial\omega = 0$, which occurs when $\omega = 1/\tau\sqrt{\epsilon_d/\epsilon_\infty}$ and $\epsilon(\omega) = 2\epsilon_d\epsilon_\infty/(\epsilon_d + \epsilon_\infty) + j(\epsilon_d - \epsilon_\infty)/(\epsilon_d + \epsilon_\infty)\sqrt{\epsilon_d\epsilon_\infty}$. The loss tangent maximum frequency is different than that of the loss - as the loss tangent maximum only indicates where the resistive material response *relative to the reactive response* is maximized.

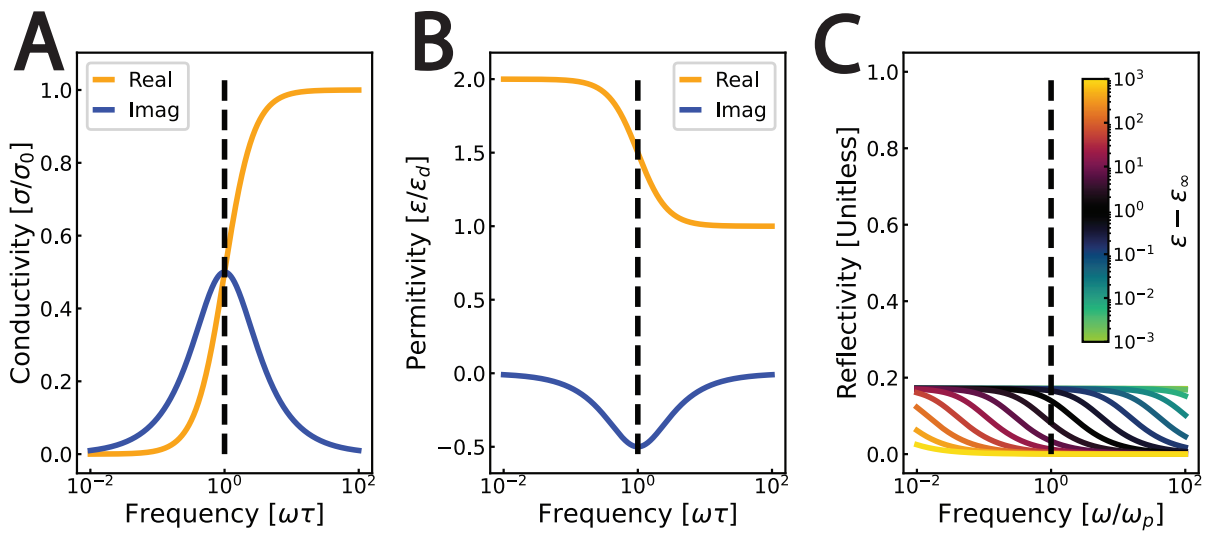

Figure 2.3: Response functions of the Debye model. (A) Real and imaginary optical conductivity. The imaginary conductivity peaks at $\sigma_0/2$ at $\omega = 1/\tau$, which is marked with a vertical line. (B) Real and imaginary permittivity, with a vertical line indicating $\epsilon_1 = |\epsilon_2| = (\epsilon_0 - \epsilon_\infty)/2$. (C) Reflectivity magnitude for various $\omega_p\tau$, with a vertical limit marking $\omega = \omega_p$. When $\omega_p\tau \gg 1$, the reflectivity abruptly increases near ...

Dielectric response, in practice, is more nuanced than the simple Debye model. One can account for distributions of dielectric relaxation processes, which leads to the Svensson model commonly used in high-frequency substrate benchmarking (which we explore in Ch. 3). Curiously, data taken from 16 orders of magnitude of frequency from vastly disparate samples, including single crystalline silicon, SiO_2, mica, amorphous silicon, β-carotene, and polymethyl methacrylate (PMMA), show similar not-quite-Debye dielectric relaxation. The loss peaks at different frequencies, but each of the vastly different samples shows high frequency loss that scales as ω^{-n-1} with $0 < n \leq 1$, meaning the loss often does not vanish as abruptly as the ω^-2 Debye model prediction. The ubiquity of this power law is remarkable. It occurs across 16 decades of frequency, in systems with covalent, ionic, and molecular bonding. The responding samples are both crystalline and polycrystalline insulators, glasses, and polymers whose geometry does not matter [17, 18, 16, 19]. The strikingly widespread applicability of this power law was identified

by Jonscher and has been subsequently been termed the "Jonscher universal dielectric response".

A full examination of the proposed mechanisms and significance of the Jonscher universal dielectric response is beyond the scope of this dissertation. For our purposes, we must merely remember that high-frequency dielectric losses do not vanish as abruptly as the Debye model predicts.

2.1.4 Application of classical gas models to single particle quantum systems

Drude's original model came before the development of quantum theory. Drude, Lorentz, and Debye physics, however, has been broadly applied in the spectroscopy of conductors and semiconductors. We show that the models apply to condensed matter systems and examine the broadness of the model in single-particle bandstructure systems.

The electronic properties of condensed matter systems are determined by band structure, which Drude's model also predated. A quantum treatment of electrons subjected to the periodic lattice potential produces electronic bandstructure: wherein electron waves constructively interfere and allow long-range propagation (in a band) or destructively interfere and prevent long-range propagation (in the bandgap) depending on the electron's energy. Bloch's theorem states that the propagating electrons have wavefunctions in the form of propagating plane waves. Propagating plane waves should never scatter - so one would naïvely expect that the Drude model, which describes momentum relaxation via scattering, would not apply to single particles in crystals.

The periodic potentials of real crystals, however, contain many anharmonicities and are not truly periodic. Structural defects like adatoms, vacancies, and stacking faults are capable of interacting with and relaxing the momentum of propagating electron waves.

Crystal lattices also contain phonons: quanta of lattice vibrations. Phonons, in real space, are ionic density waves that cause anharmonicities in the periodic potential. Drude electrons can relax their momentum by scattering via such disorder, phonons, or another source of crystalline anharmonicity.

Crystalline electrons are not the only charge carrier to follow the Drude model. The two key assumptions of the Drude model are that the constituents form a (a) gas, meaning negligible internal interactions and (b) contain charge carriers, which need not be electrons. Consider a heavily P-type doped semiconductors, for example. The holes are weakly interacting and are also electrically charged carriers. A gas of holes would exhibit Drude-like behavior.

The Drude model cannot account for transitions between bands, but it can be used for dynamics of charge carriers within a band. In quadratic bands, for example, semi-classical dynamics describe carrier behavior well if the carrier mass is replaced by the renormalized (effective) mass, which is computed from the band structure. Put crudely: carriers near the extrema of quadratic bands behave as a carrier gas with the regular mass replaced by the effective mass. The Drude model can simply be extended to dynamics of carriers in band by mapping the mass in eqn. 2.3 to the appropriate band effective mass m^*.

Transitions between bands cannot be represented by the Drude model. Consider $\chi_e(\omega)$ of a two-level quantum system with delta-function absorption given by $Im(\chi_e) = A_0\delta(\omega - \Omega_0)]$. We can extract $\epsilon(\omega)$ via the Kramers-Kronig relations (eqn. 1.30) and the dielectric constitutive relation (eqn. 1.3):

$$\begin{aligned} Re(\chi_e) &= \frac{2}{\pi} P \int_0^\infty \frac{\omega' A_0 \delta(\Omega_0)}{\omega'^2 - \omega^2}\, d\omega' = \frac{1}{\Omega^2 - \omega^2}\frac{2\Omega A_0}{\pi} \\ \epsilon_1(\omega) &= 1 + \frac{1}{\Omega^2 - \omega^2}\frac{2\Omega A_0}{\pi} \end{aligned} \tag{2.9}$$

The last line of eqn. 2.9 is a Lorentzian with zero broadening and a plasma frequency of $\sqrt{2\Omega A_0/\pi}$. The zero broadening is exactly in line with our assumption of the higher energy band having no width. It can be shown that adding broadening doesn't change the fundamental result - that Lorentzian material parameters describe transitions between quantum levels and energy bands [20]. The absorption spectra of three bands, for example, would be two Lorentzian responses.

The interplay between the two models can prove useful in practical measurements. Consider a metal to Mott-Hubbard insulator transition, with critical Mott-Hubbard temperature T_C and zero temperature gap of energy $\hbar\Omega_0$. In the high temperature metallic phase, the Fermi level crosses the conduction band and the system responds to incident THz radiation via the Drude model. This behavior continues as the material is cooled to temperatures just above T_C. Just below T_C, the Mott-Hubbard phase has a bandgap at the Fermi level, but some carriers are still thermally activated across the gap. The spectrum is best described by a Lorentzian term to describe excitation across the gap and a Drude term to describe the behavior of the thermally activated carriers. As the temperature continues the lower, the Drude term vanishes and the Lorentzian response dominates the THz response at the gap frequency Ω_0. As the condensed matter system is cooled through the critical temperature and a bandgap opens, spectral weight is transferred from zero frequency (Drude-like) response to the finite frequency (Lorentzian-like) response.

Spectral weight and its transfer is a key concept in the spectroscopy of quantum systems. Spectral weight can be understood by derivation from the Kramers-Kronig relation. No matter the low frequency behavior, all carriers will behave as if they were free (Drude-like) at extremely high frequency [21, 20]. Writing the Kramers-Kronig relation eqn. 1.30 for dielectric susceptibility $(\epsilon_1 - 1)$ and taking only the high-frequency behavior:

$$\begin{aligned}\epsilon_1 = 1 + \frac{2}{\pi} P \int_0^\infty \frac{\omega' \epsilon_2(\omega')}{\omega'^2 - \omega^2}\, d\omega' = 1 - \frac{2}{\pi\omega^2} P \int_0^\infty \omega' \epsilon_2(\omega')\, d\omega' \\ 1 - \omega_p^2/\omega^2 = 1 - \frac{2}{\pi\omega^2} P \int_0^\infty \omega' \epsilon_2(\omega')\, d\omega' \\ \frac{\pi\omega_p^2}{2} = P \int_0^\infty \omega' \epsilon_2(\omega')\, d\omega'\end{aligned} \tag{2.10}$$

Eqn. 2.10 is a sum rule, relying only on a causal dielectric susceptibility and Drude-like behavior at high frequency. An analogous sum rule can be derived in units of optical conductivity [22, 23]:

$$\frac{\pi\epsilon_0\omega_p^2}{2} = P \int_0^\infty \sigma(\omega'), d\omega' \tag{2.11}$$

Note that the squared plasma frequency in eqn. 2.11 contains a factor of the carrier density. As long as the charge carrier density is conserved, total spectral weight is also conserved. This can be used to estimate the weight transferred out of measurable frequency ranges into, say, narrow band absorptions or zero-frequency superconducting condensates. As long as the charge density is conserved, the spectral weight should be also be conserved.

The sum rule, however, does not tell us to what frequency band the spectral weight has transferred. The electrodynamic response of a solid contains contributions from valence and core electrons. If we could integrate that response over all frequencies, we could access the total density of core and valence electrons. Most optical spectroscopies are restricted to measuring intraband and some interband dynamics of the valence electrons, simply because transitions of core electrons occurs at frequencies beyond those considered in this dissertation. Spectral weight transfer is not always measurable.

In summary, the Drude model is best applied to intraband dynamics. The Lorentz model is best applied to interband dynamics. Appropriate sums and sum rules of Drude

and Lorentz models can be used to describe more complicated systems.

2.1.5 Warning: simple model extensions are often non-physical

There are many simple extensions of the Drude and Lorentz models, which include making m^* and τ frequency dependent, non-ideality parameters, and broadening from band tails. Though all of these models arise from some physical or phenomenological motivation, many produce non-causal, non-physical response functions. Be careful: non-physical models are strewn about the literature. When using a non-standard model or modifying one yourself, you should first whether the resulting response function obeys the Kramers-Kronig or not.

One physical extended model is the Cole-Davidson modification to the Drude model. The Cole-Davidson modification [24] adds modifies the denominator by ideality parameter β to get $\sigma(\omega) = \sigma_0/(1 - j\tau\omega)^{-\beta}$. It has been used successfully to fit the metallic response of lightly doped N- and P-type silicon. Inserting the real part of the model's optical conductivity into the Kramers-Kronig relations (eqn. 1.30) produces the published Cole-Davidson imaginary part. The Cole-Davidson modification is, therefore, a physically realizable system.

The Drude scattering time τ and carrier mass m^* (or plasma frequency ω_p) are frequency independent quantities. Some generalized Drude and Lorentz models assign these quantities an implicit frequency dependence, which have been successfully applied to describe the dynamics of heavy fermion, highly correlated, and Kondo systems [25, 26, 27, 28, 29]. Some frequency-dependent extended models, however, produce non-causal response functions. Theye et. al. [30] proposed a generalized Drude model with $\tau(\omega)$ chosen to match a measured resistive response. The reactive response is computed by inserting the calculated $\tau(\omega)$ into the Drude model. Even simple resonant forms of

Drude scattering time of the form $\tau = 1/(a + b\tau^2)$ fail to satisfy the Kramers-Kronig relations [31], and is thus non-causal.

A Kramers-Kronig-respecting frequency-dependent generalized model can be formulated if the real and imaginary response functions can be measured independently. For the Drude model, eqn. 2.3 suggests that the $\tau(\omega) = -\epsilon_1(\omega)\epsilon_2^{-1}(\omega)\omega^{-1}$. Using experimental data for $\epsilon_1(\omega)$ and $\epsilon_2(\omega)$ to calculate $\tau(\omega)$ ensures that all response functions remain causal, but may obfuscate poor fits to the Drude model [32]. Extended models using this philosophy have been applied to heavy Fermion systems to extract complicated effective masses and momentum relaxation times [25, 33]

There are also successful and failed extensions to the Lorentz and Debye models. The Brendel-Bormann oscillator, an extension to the Lorentz model, is another seemingly-simple model extension that convolves a simple Gaussian with the Lorentzian to account for absorption by the valence band tails of disordered systems [34]. It can be shown that the real and imaginary parts are not related by the Kramers-Kronig relations [35, 36], but has nevertheless been applied to the spectroscopy of glasses [37]. There are also Kramers-Kronig-respecting Gauss-Lorentz oscillator models, which have been used in the spectroscopy of disordered InGaAs [38].

In summary - be cautious when using generalized Drude, Lorentz, or Debye models that the resulting response functions are causal and still describe the physical systems in question.

2.1.6 Application to many particle quantum systems

We have shown that the spectroscopic models for a free electron gas or damped, driven oscillators can apply to quantum systems in which interparticulate interactions can be ignored. The carriers of real systems, of course, can interact with each other via

Coloumbic effects, exchange interactions, as mediated by the lattice, ... etc. Our simple models may also be extended to such general systems because these systems have low energy excitations that are long-lived, respond to electromagnetic waves, and only weakly interact with one another. This is the concept of a Fermi liquid [39, 40]. We summarize the key results.

A single-particle quantum system (Fermi gas) is a gas of non-interacting electrons and holes. A many particle system (Fermi liquid) is a gas[1] of quasiparticles and quasiholes. Both Fermi gas and Fermi liquid are described with spin and charge quantum numbers, and have an effective mass m^*. Fermi gas carriers (electrons and holes) are always eigenstates of the Hamiltonian, and are therefore always stable. Fermi liquid quasiparticles and quasiholes are, however, not eigenstates but interaction-dressed eigenstates. They are not always stable. The dissociation time, however, approaches infinity as $1/(E_1 - E_F)^{-2}$, where E_q is the quasiparticle energy and E_F is the Fermi energy. This implies that only the quasiparticles with energies near the Fermi surface are stable, and will actively respond to electromagnetic waves.

Excitations of many particle ground states are not necessarily fermions. There are also bosonic collective modes, such as magnons, excitons, and plasmons. These bosonic modes represent collective excitations of spin waves, bound electron and hole pairs, and collective charge oscillations. These modes can also couple to incident radiation.

Quasiparticles and other collective modes may react to incident radiation through the simple models we developed above. A Drude model, for example, would describe a gas of constituent quasiparticles. A Lorentz model would describe a gas of harmonic oscillators quasiparticles, with ω_0 representing a resonance frequency or an a quasiparticle transitioning between quasparticle bands. Some quasiparticles have more complicated in-

[1]The quasiparticles may interact, but they do so much more weakly than the underlying electrons and holes do

trinsic responses that cannot be described with simple modifications to Drude or Lorentz models, one of which we will consider in the next section.

Be weary of applying simple models entirely to the excitations of many body systems. Be aware that underlying electrons and holes or dielectric oscillators will also respond to incident radiation, so the total material response will be a sum of the quasiparticle, single-particle, and dielectric responses.

2.2 Superconductivity

Superconductivity is arguably condensed matter's most prolific many-body problem. The liquefaction of Helium and subsequent discovery of superconductivity by Onnes posed serious a challenge to callow quantum theories in 1911. Though many great discoveries have advanced the field, many aspects of superconductivity are not understood and haunt the physics community some 112 years later. Measuring $\sigma(\omega)$ has been a vital probe of superconducting systems for the last 80 years. Here, we introduce the BCS theory of superconductivity with the Kramers-Kronig relations, the THz spectroscopy of a superconductor, and provide a word of caution to the zealous Kramers-Kronig user.

Before the superconductor, we build intuition using the simple PEC from Ch. 1. Consider a zero-frequency PEC in a medium with $\sigma = A\delta(\omega)$. Electrical conductivity σ is a causal response function, so it must obey Kramers-Kronig relations. We insert $\sigma_1(\omega) = A\delta(\omega)$ into eqn. 1.29 and take advantage of the properties of the $\delta(\omega)$ to find:

$$\sigma_2(\omega) = -\frac{A}{\pi} P \frac{1}{\omega} \tag{2.12}$$

Thus, the imaginary conductivity goes as $-1/\omega$. In a circuit, a zero frequency PEC has imaginary conductivity proportional to $-1/\omega$, which defines an inductive reactance.

Thus, zero-frequency PECs are inductive at non-zero frequencies. This simple principle can be also be obtained by taking the limit as $\tau \to \infty$ in a Drude model, implying that a lack of momentum-losing collisions and inductance are completely equivalent.

A superconductor's inductance can be understood intuitively. Inductance L is proportional to $\partial\overline{j_e}/\partial t$. A voltage V applied to material generates current density $\overline{j_e} = ne\overline{v_d}$, for drift velocity $\overline{v_d}$. In a normal metal, the velocity of carriers accelerated by the fields repeatedly relaxes to zero, so the time-averaged drift velocity is time-independent. The inductance, therefore, remains zero. If the material is a PEC, however, $\partial\overline{v_d}/\partial t$ is a constant and the carriers acquire an inductance.

The superconductor and PEC are related, but not the same. Both have infinite real conductivity at zero frequency (and, thus, some inductance at finite frequency). At zero temperature, the ground state of a PEC is a gas of ballistic electrons. At zero temperature, the ground state of a superconductor is a gas of Cooper pairs. A Cooper pair is a Bosonic pairing of two carriers, which doesn't interact with other Cooper pairs [41]. Consider S-wave superconductors, which have an isotropic Cooper pair bandgap. At zero temperature, incident frequencies below the bandgap 2Δ cannot break a Cooper pair. As the incident frequency reaches 2Δ, the energy required to break a Cooper pair, normal carriers will be present in the sample. These normal state carriers respond resistively to the driving field, so one expects $\sigma_1(\omega)$ to rise and follow Drude (or more complicated) behavior. There is a corresponding Kramers-Kronig change in the imaginary part, which does little to divert the $-1/\omega$-like behavior at low frequencies. Collecting all contributions, one expects the real part of conductivity to have a zero-frequency delta function, be zero in the gap, and start increasing at $\omega = 2\Delta$. One expects the imaginary part to have a $-1/\omega$-like structure, with some slight modifications from the Kramers-Kronig relations.

The response is more complicated for real materials at non-zero temperatures. Non-

zero temperatures broaden the zero-frequency delta function. In addition, the finite momentum scattering time of the normal state carriers must be considered. If the normal state carriers scatter much faster than the gap frequency ($\tau \ll \hbar/2\Delta$), the normal state response is dominated by the usual carrier momentum relaxation. This applies to systems with lots of scattering, and thus is called the impure limit. In the opposite limit ($\tau \gg \hbar/2\Delta$) at frequencies above the 2Δ, the normal state carriers from the broken Cooper pairs travel ballistically. The elevated $\sigma_1(\omega > 2\Delta)$ is suppressed, as the excited carriers do not behave resistively. This applies to superconducting systems with little scattering, and which are thus said to be in the pure limit.

The optical conductivity $\sigma(\omega) = \sigma_1(\omega) + j\sigma_2(\omega)$ of samples with arbitrary purity $y = \hbar/2\Delta\tau$) and arbitrary temperature T/T_c is calculated using the Zimmerman model [42]. The result is plotted in Fig. 2.4. Note that 2Δ has a temperature dependence of the form $2\Delta(T) = 2\Delta(0)\tanh\left((\pi k_B T_c/\Delta(0))\sqrt{T_c/T-1}\right)$, where $2\Delta(0)$ is the zero-temperature gap and k_B is the Boltzmann constant.

As expected, the $T/T_c \ll 1$ and impure limit has a real part with a zero-frequency delta function, a zero value in the gap, and an absorptive feature that starts increasing at $\hbar\omega/2\Delta = 1$. The imaginary part has a $-1/\omega$-like dependence, with small corrections from the Kramers-Kronig relation. The pure limit has the non-zero frequency feature suppressed in the real part and a simpler $-1/\omega$-like relation. Higher temperatures show qualitatively similar features as low temperatures, but with thermal broadening features.

Real superconducting samples at non-zero temperatures also have spectral contributions from the normal state (uncondensed) carriers [41]. The modelling above can be improved with the two-fluid model, which contains a temperature dependent term with normal metal carrier density undergoing the Drude model and a temperature-dependent term with a superconducting carrier density undergoing the Zimmerman model. The sum of the spectral weight from the (thermally broadened) zero-frequency condensate,

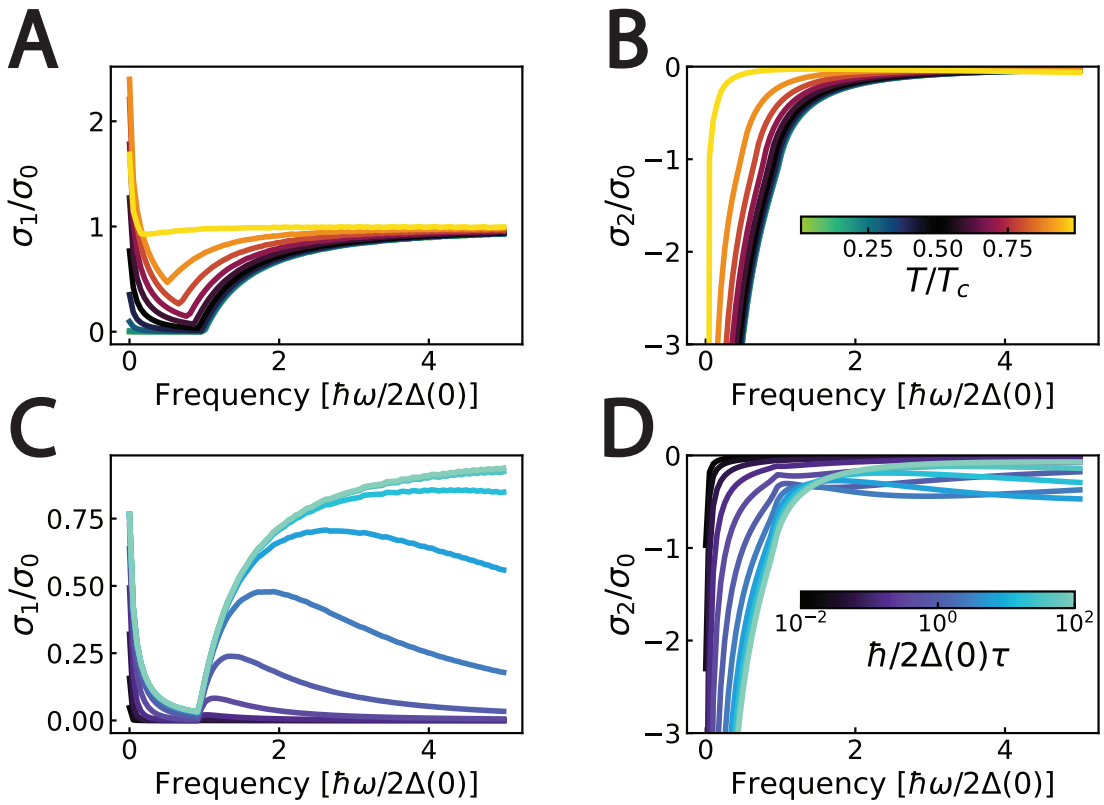

Figure 2.4: Response functions of the Zimmerman model, normalized to the normal metal ($T/T_c > 1$) optical conductivity. (A/B) $\sigma_1(\omega)$ and $\sigma_2(\omega)$ (respectively) of a Zimmerman superconductor versus normalized frequency with parametric steps in T/T_c, computed in the impure limit with purity $\hbar/2\Delta\tau = 500$. The real part develops a gap that increases with decreasing temperature; the imaginary part develops a strong -1/ω-like response. Thermal broadening is apparent at high temperatures. (C/D) $\sigma_1(\omega)$ and $\sigma_2(\omega)$ (respectively) of a Zimmerman superconductor versus normalized frequency with parametric steps in purity, computed with $T/T_c = 0.5$. The impure limit ($\hbar/2\Delta\tau \gg 1$) shows the same features as A/B; the pure limit ($\hbar/2\Delta\tau \ll 1$) shows a reduction in features in σ_1 and corresponding Kramers-Kronig change in $\sigma_2(\omega)$.

normal state carriers, and (thermally broadened) resistive carriers at $\omega(> 2\Delta)$ has been shown to follow the sum rule (eqn. 2.11)in overdoped $La_{2-x}Sr_xCuO_4$ [43]. This implies that charge is conserved as the sample is cooled from the normal metal into the superconducting phase.

2.2.1 Warning: Use caution with the Kramers-Kronig relations and superconductors

A word of caution must be issued regarding use of the Kramers-Kronig relations for superconductors. It is generally much easier to measure the magnitude of a sample's reflection coefficient $|R(\omega)|$ than to measure the magnitude and phase of $R(\omega)$. For this reason, it may be tempting to measure the reflection and compute σ_1 and σ_2 using the Kramers-Kronig relation. One should be cautious of doing so, and of any reports using this.

This criticism will focus on reports of light-induced superconductivity. The basic phenomenology is the claim that a normal metal at temperatures above the superconducting transition temperature can be driven into a transient superconducting state with a strong, ultrafast optical pulse. The dynamics are too fast for conventional DC or microwave electronics. THz pulses are, however, fast enough to monitor the change in electronic properties and the gap $2\Delta(T)$ lies within the resolvable bandwidth. Several experiments [15, 44, 45] detail a transient state with decreased spectral weight in $\sigma_1(\omega)$ and increased $1/\omega$-like response[2] in $\sigma_2(\omega)$ at $\omega < 2\Delta$. Real and imaginary conductivity are computed by measuring the reflectivity across 2.5 orders of magnitude in [15] and 1.5 orders of magnitude of frequency in [45] and using the Kramers-Kronig relations. Low frequency data is extrapolated using the Hagen-Rubens formula (2.4).

The Achilles heel of the Kramers-Kronig relation in experiments is that it requires an integral over all frequencies, as discussed in Ch. 1. No single experiment can acquire data over all frequencies. In an effort to increase the data's verisimilitude, the Hagen-Rubens formula was used to extrapolate the normal state data at lower frequencies from 0.75 THz to DC. This assumes that the low frequency conductivity contains no additional

[2]Their sign convention defines $\sigma(\omega) = \sigma_1(\omega) - j\sigma_2(\omega)$. We use their convention for section 2.2.1 only

features, and notedly cannot be done in the superconducting state.

In [15], the high-frequency cutoff is at a rather large frequency of 240 THz. Without questioning the use of the Hagen-Rubens relation, the (strict) criterion suggests that the use of Kramers-Kronig output can be trusted from DC to 24 THz in the normal metal and from 7.5 THz to 24 THz in the superconductor. This criterion is rather strict, but should remind all readers to be cautious in interpreting the results.

There is further issue in interpreting the imaginary conductivity. The Kramers-Kronig relations of a superconductor's response integrate the (thermally-broadened) zero-frequency delta-function. The transient state following the strong optical pump decays too quickly to be able to measure the zero-frequency features. Applying the Kramers-Kronig relations to a reflectivity measurement requires postulation of the low-frequency response. This can be motivated by observing some of the delta function's thermal broadening at non-zero but low frequencies, but is greatly complicated by the transient, non-equilibrium nature of the state.

Even if we establish that the $1/\omega$-like responsive can be trusted, it would be premature to definitively assign this response to incipient transient superconductivity. Recall that a perfect electric conductor (i.e. ballistic electronic) transport has this exact spectroscopic signature. Without observing the formation of a gap at $2\Delta(T)$, one *cannot* differentiate between a superconductor and a PEC.

The field of 3D superconductors typically requires three successful measurements to declare a sample a superconductor. These are (1) vanishing electrical resistance, (2) the Meisner effect, and (3) quantum coherence, in the form of an Andreev reflection, bound Andreev state, or a Josephson junction. Showing an imaginary conductivity of the form $1/\omega$ is equivalent to item 1 and item 1 only. A simple, precipitous drop in resistance at a certain temperature can occur by many mechanisms, including a sharp increase in mobility which could come from an injection of hot photodoped carriers by an optical

pulse.

2.3 THz vs. Raman active modes

It is important to note that not all condensed matter systems respond to incident electromagnetic radiation in the form of a linear response function. Systems can be (1) "THz" or infrared (IR) active[3] and/or (2) Raman active. IR (THz) active systems couple linearly to IR radiation; Raman active systems do not. As such, this dissertation will focus only on IR active systems.

IR spectroscopy directly excites a sample from one energy level to the other; Raman spectroscopy does so with an intermediate virtual state. IR activity refers to linear electromagnetic radiation - sample coupling, meaning that incident radiation at frequency ω will be absorbed and scattered at frequency ω. Raman spectroscopy, however, drives a system to a virtual energy level. The system will statistically not always return to the lowest energy level: it may return to an excited state. There will, therefore, be an energy difference between the incident and emitted photon, which is measured as scattered optical power at the incident and emitted photon frequency difference. Raman spectroscopy is, therefore, a nonlinear process.

Though both techniques are capable of measuring absorption at THz frequencies, they provide different information. IR spectroscopy provides phase information because it directly measures the reflected or transmitted electric field magnitude and phase; Raman spectroscopy cannot be used to directly extract phase because it measures only the scattered power. One could use the Kramers-Kronig relations developed in Ch. 1 to compute the phase via the Bode gain-phase relation (eqn. 1.31). This, however, requires an integral over all frequencies, which can lead to the errors. Accurate extraction of the

[3]We use the term "IR" active, though this applies to the THz band as well

phase requires Raman measurements over a very broad bandwidth, leading to the errors previously discussed.

Modes can be IR and Raman inactive, IR or Raman active, or both IR and Raman active. The IR and Raman activity for a mode can be deduced from the mode's symmetry. Determine the spatial location of the mode's charges and deduce the mode's irreducible representation via its point group. Then consider small deformations to that point group. Any deformation that results in a change to the dipole moment (change in magnitude of charge per site or change in the spacing between charges) will be IR active. Any deformation that results in a change in the polarizability will be Raman active. One can find the symmetries of the IR and Raman modes in documented tables [46], allowing one to identify which deformations can be measured by which technique.

Take a simple example from chemistry: a CO_2 molecule (group: D_{2H}). The molecule is depicted in Fig. 2.5. Consider two deformations: the symmetric stretch (B) and the asymmetric stretch (C). Imagine different charges localized to the carbon and oxygen atomic sites. The far field of the symmetrically stretched structure would be the same as that of the original structure, so the mode is not infrared active. The two atomic sites, however, do have different electronegativities and thus the symmetrically stretched structure will have a different polarizability. The symmetric structure, therefore, is Raman active. The asymmetric stretch, however, is not Raman active. The increase in polarizability of the left side is compensated by the decrease of the right side. The asymmetric stretch does, however, change the far-field dipole moment and is thus IR active.

The link between symmetry and optical activity can be extended to condensed matter systems. Consider a hypothetical material with a (simple) one dimensional standing charge density wave. That is: a one-dimensional periodic structure of alternating localized pockets of increased charge density and decreased charge density. Consider only one unit cell of the repeating charge density. The charge density wave could be excited by in-

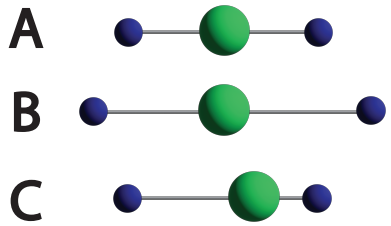

Figure 2.5: Illustrations of IR and Raman active deformations of CO_2 (oxygen as small, blue spheres; carbon as large green spheres) with (A) the bare molecule (B) a symmetric stretch and (C) an asymmetric stretch.

cident radiation in one of two ways. The incident radiation could increase the amplitude of the charge density wave by increasing (decreasing) the charge density in the statically increased (decreased) regions - called the amplitude mode. The amplitude excitation changes the charge polarization between the high and low charge density regions and is therefore Raman active. The total charge is conserved and the charging positions don't move, so the dipole moment is unaltered. The amplitude mode is therefore IR inactive.

The incident radiation could also shift some of the increased charge density into the decreased charge density, modulating the charge density wave's phase - called the phase mode. The phase excitation shifts the charges' local positions and thus alters the dipole moment. It also changes the charge density wave's polarizability and is also [47, 48] Raman active. THz spectroscopy is only capable of resolving the IR active modes and therefore, in this example, can identify the phase excitation but not the amplitude excitation. This is a word of caution: check and understand the symmetries in the spectroscopic models of a quantum systems before attempting to measure them.

.

Chapter 3

Active and Passive THz Engineering

"Never trust a materials grower
unless they are a data shower"

A. M. Potts

Different frequency bands of the electromagnetic spectrum have clearly delineated and standardized frequency / wavelength cutoffs. The precise cutoffs and distinctions start to fade in the mm-wave and THz bands, so we will define THz radiation as 0.1 - 10 THz. Spectroscopy, engineering, and even basic component design in the THz band has floundered and lagged the technological progress of adjacent frequency bands for decades. This makes work in THz bands particularly frustrating, but particularly exciting, as the reward of new discoveries and inventions await anyone bold and clever enough to undergo concentrated THz development efforts.

THz systems have three broad applications: (1) communication systems, (2) imaging, and (3) spectroscopy, plus mixes of these. The motivation for THz communication systems is simple: increasing the total signal bandwidth from microwave to THz bands increases the fractional bandwidth available to a communication network, per the Bode-Fano criterion[1][5]. Put simply: higher frequency carrier signals have higher data

[1] a Kramers-Kronig relation restricting the spectral weight of a communication channel's passband

throughput rates than low frequency carrier signals. The global demand for and subsequent overcrowding of cellular and wireless internet signals in and C band (2 - 4 GHz) and L band (4 - 8 GHz) frequencies would be greatly aided by pushes to higher frequencies. The extreme attenuation of THz signals by atmospheric gasses offers THz signals another advantage: covert, intercept-proof spacecraft-to-spacecraft communications. Signals at water atmospheric gas absorption lines can be safely transmitted from satellite to satellite without fear of ground interception, as any signal incident on the atmosphere will be sharply attenuated.

Imaging at THz frequencies also shows great scientific and technological promise. THz bands have a smaller wavelength than microwaves, giving THz waves better spatial resolution. In addition, THz is non-ionizing and is not harmful to cells for typical electric field strengths. THz has high enough energy to pass through fabrics (Debye model), but not enough to pass through metals (Drude model). THz scanners and imaging systems can be used to scan a briefcase or clothed human body for hidden metallic weapons and tools [49]. They can also see hidden compartments in, say, books that are used to smuggle weaponry. Such scanners are already in place in airports, for high-throughput, non-ionizing scanning for weapons [50, 51, 52, 53] and banned materials (such as methamphetamine [54, 55]). These scanners can also be used for pharmaceutical quality control, by monitoring the THz absorption of pill bottles without removing the medication from vacuum seals and packaging [56]. THz has also been known to show hidden internal defects in car parts, cracks and manufacturing faults in composites, and knots in lumber that would be difficult and time-consuming to detect by the naked eye and expensive/highly regulated to detect with X-rays [57, 58, 59, 60, 61].

Spectroscopy, the 3^{rd} main use of THz radiation, is the act of measuring and interpreting optical parameters $\sigma(\omega)$, $\epsilon(\omega)$, or $\mu(\omega)$ in physically meaningful systems. THz spectroscopy should be used when the features described in Ch. 2 lie within the THz

frequency band, meaning energy gaps, resonance frequencies, and the plasma frequencies of quasiparticles. This naturally occurs in a broad range of phenomena: ferromagnetic and antiferromagnetic resonances, intra-excitonic transitions, plasmons, superconducting energy gaps, localization effects in disordered systems, the Josephson plasmon resonance, and many more [62, 63]. In addition, THz spectroscopy has been shown to have biological applications, including the ability to screen for Alzheimer's disease[64] and breast, skin and colon cancers [65, 66, 67, 68, 69, 70, 71] . The sensitivity of THz to water vapor has made spectroscopy useful in assessing the depth of burn damage to human skin [72]. THz spectroscopy has even been applied to non-destructively analyzing paintings and even resolving high-profile art forgeries by comparing the THz absorption spectra of the paint, gold foil, and underlying canvas [73, 74] with that of known, historical materials (freely available online). This dissertation will apply THz spectroscopy to 2D quantum materials.

Historical progress and development of these applications has been slow and painful. Engineering at THz frequencies is much more difficult than for microwave and infrared/optical bands. Engineered electromagnetic systems have four primary subsystems: input, operation, propagation and output, which can be divided into active (capable of producing or measuring THz power) and passive (transfer function magnitude unity or less). Engineering both active and passive devices at THz frequencies are active areas of research.

To demonstrate the difficulty of THz engineering, we measure the attenuation coefficient of THz transmitted through a quasi-optical 1.5m path (Fig. 3.1). THz is more rapidly attenuated by atmospheric gasses than most other frequency bands are. We use an Agilent PNA with free space-coupled Virginia diode multipliers with WR1.5 configuration as a continuous wave THz source. Power is measured at the transmitter (TX), propagates through 1.5m of atmospheric gasses, and is measured at a receiver (RX). The difference in TX and RX power is the power absorbed by atmosphere gasses. The

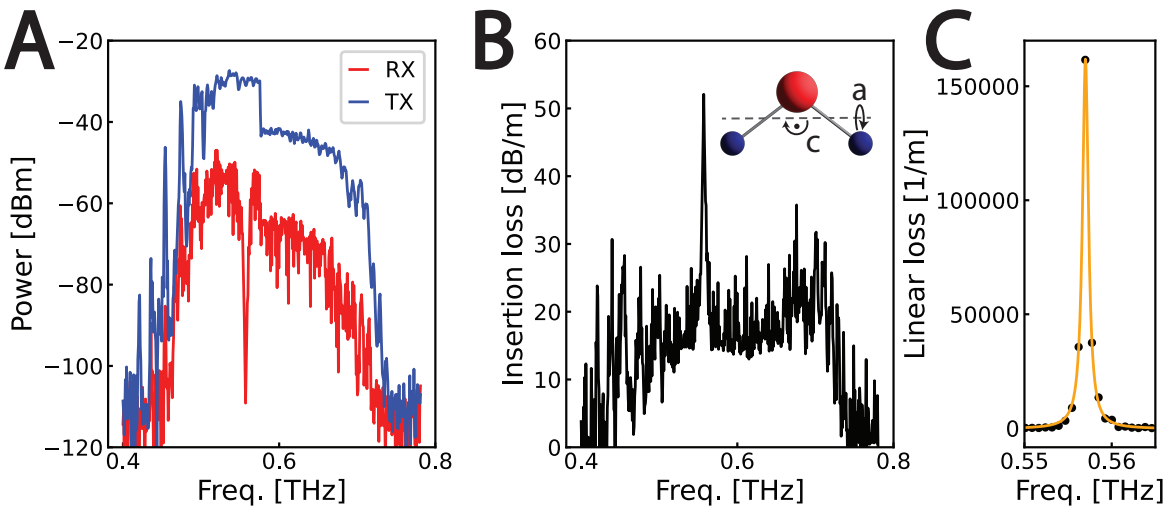

Figure 3.1: Continuous frequency THz propagation in atmospheric gasses. (A) Transmitted (TX) and received (RX) power, with all differences coming from gaseous absorption. (B) Insertion loss, computed as the difference between TX and RX power per meter. There is a broad continuum of absorption and a sharp resonance at 0.557 THz from water molecule resonance, whose lower bound extinction coefficient exceeds 50 dB/meter. This corresponds to >5 orders of power loss per meter of travel. Inset: schematic of water molecule, with the $1_{01} - 1_{10}$ rotational transition corresponding to resonant transition between a and c axis rotation (C) Lorentzian fitting (orange) of loss data points (black), with central frequency of 0.557 and a full width, half maximum of 800 MHz.

PNA noise floor is near -110 dBm, so attenuation can be accurately resolved only from 450 GHz to 750 GHz. The average attenuation is nearly 15 dB/m, meaning the transmitted THz power lessens by a factor of 31.6 for every travelled meter. The nearly 52 dB insertion loss corresponds to the power reduction by a factor of nearly 160,000 per meter of travel. This is a sharp Lorentzian resonance at 0.557 THz, corresponding to the $\delta J_{K_A K_C} = 1_{10} - 1_{01}$ rotational transition of gaseous water molecules [75, 76, 77]. This is an lower bound on extinction coefficient, as the RX signal reaches the PNA's noise floor.

Many other such resonances in water, carbon dioxide, and other atmospheric molecules occur at higher frequencies. The large attenuation at most frequencies and extreme attenuation at specific molecular resonance frequencies make atmospheric propagation of

THz nearly impossible. Though many use water vapor-deficient and nitrogen-rich environments to preserve signal integrity, this is unfeasible over several meter ranges.

Free space THz engineering suffers from severe attenuation. In order to make a spectrometer, this attenuation must carefully be controlled.

3.1 Free space and on-chip THz

Free-space THz time domain spectroscopy (TDS) is an experimental technique that has successfully probed a wealth of condensed matter phenomena in the last 25 years. The technique requires the use of a free space, pulsed laser with a pulse time < 200 fs. An ultrafast laser is partitioned into two beams. One beam strikes an emitter. By mechanisms we will explore later, this process causes the emission of a quasi-single cycle pulse with THz-frequency spectral content. This pulse is directed and focused by quasi-optical elements to a sample. The transmitted (or reflected) pulse is directed to a readout device. The second beam from the initial laser pulse is directed from retroreflector on a computer-controlled delay stage, which allows for in-situ tuning of the optical path length. The beam is then routed to the readout device, such that the second optical beam strikes the readout at about the same time as the propagating THz pulse. When both THz and laser (triggering) pulse are coincident on the readout device, the instantaneous electric field is recorded as a current. Sweeping the delay stage's position tunes the time delay between the triggering pulse and THz pulse, allowing full recording of the instantaneous electric field versus time. THz TDS is thus able to overcome electronic amplifier limitations by converting ultrafast timescales ($<$10 ps) into achievable changes in mirror positions mounted to a delay stage by taking advantage of the speed of light.

The optical beam incident on the emitter can be mechanically chopped at kHz frequencies, allowing one to extract the THz signal with a low frequency lock-in amplifier.

Laser pulses are emitted with a typical frequency of 100 MHz, so a mechanically chopped pulse can contain hundreds of thousands of THz pulses. Measuring each point with a typical 5 time constants (30 ms), a single data point contains the average of about 15 million THz pulses. The intrinsic averaging is key to free space THz TDS.

Free space THz techniques are best applied to the measuring samples of lateral length of 5 mm or greater. Quasi-optics setups cannot focus broadband radiation to a beam waist (paraxial diameter) smaller than the Rayleigh diffraction limit of 1.22λ. The lowest frequency components of a quasi-optical THz pulse have the largest projections on the sample plane (5 mm at 73 GHz). Many real samples are not at all uniform on this length scale and thus some artifacts can arise.

van der Waals (vdW) materials and heterostructures are made from atomically thin, exfoliated flakes of graphite, graphene, hexagonal boron nitride, WSe_2, WTe_2, $MoSe_2$, $MoTe_2$ and others. vdW materials are exceptionally clean, with mean free paths often larger than the physical samples themselves. In addition, the 2D nature of these materials makes them easy to engineer into devices, which can support a wide range of material physics. Tuning the voltage applied to (metallic) electrostatic gates allows one to change the carrier density of a vdW material in-situ. Unclean or problematic regions can also easily be removed by etching, allowing for careful design of very clean and very tunable quantum systems. Exfoliated flakes and resulting heterostructures are, however, on the order of 10 μ(m, which is much smaller than THz wavelengths (Fig. 3.2A). This is far too small to be measurable by free space THz pulses. Even if vdW heterostructures were 10 mm $\times$ 10 mm, free space THz would only pass through a few atomic layers of the sample. The pulse would barely interact with the sample, providing little contrast between different condensed matter system phases. What little contrast there is would be dominated by the metallic gates. Conventional free space THz TDS is, thus, completely unfeasible for 2D materials.

THz TDS can be adapted to vdW heterostructures, however, by changing the geometry. One can place the vdW heterostructure in a transmission line, making the vdW heterostructure a circuit element in a THz circuit. Since the vdW heterostructure is electrically small, it acts as a lumped element impedance / admittance in the transmission line, whose reflection and transmission coefficients are summarized in Table 1.4. Furthermore, the THz pulse travels in the plane of the vdW heterostructure. The relevant length for the THz-sample interaction is now the side length, which can be 10 μm instead of one atomic layer. Metallic gates above and/or below the vdW heterostructure will attenuate THz, but can be removed in post-processing through referencing because there is also contrast from the target vdW material itself [6, 78].

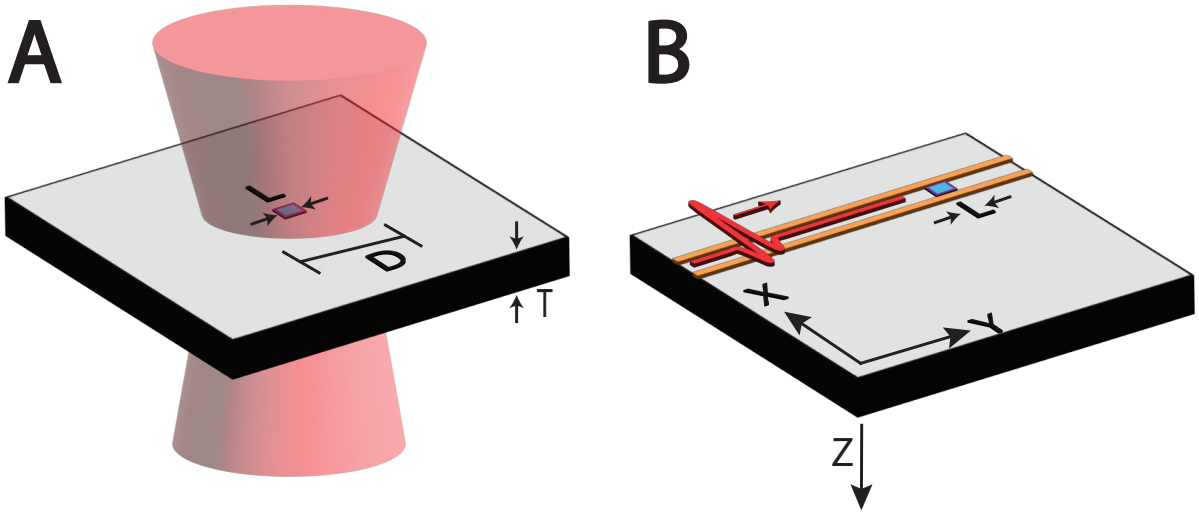

Figure 3.2: Comparison of geometries for the spectroscopy of a vdW heterostructure via (A) free space THz TDS (focused THz beam in red) and (B) on-chip THz TDS, showing a CPS-bound THz pulse in red

On-chip THz spectroscopy is also ideally suited for more complicated variants of free space THz TDS. Optical pump, THz probe (or THz pump, THz probe) spectroscopy is a variant by which an optical (or second THz) pulse excites the sample. Using a second optical delay stage, the instantaneous THz spectrum can be measured as a function of time after the excitation pulse, giving rise to a very exciting new field: the spectroscopy of light-

induced phases of matter [79, 15, 44]. This technique does, however, require complicated analysis due to mismatched skin depths ($\delta \sim \sqrt{\omega}$) of the pump and probe. The optical probe only penetrates the outermost few hundred nm in a metal, while THz can penetrate the entire sample. Resolving this requires postulating a time-dependent, frequency-dependent, depth-dependent index ("multilayer analysis") and fighting through cumbersome analysis. Worse, several claim that such depth-dependent analysis is susceptible to artifacts due to non-linearities from the strong pump [80] or transient thermal current [81]. These skin-depth mismatch-related issues are solved by performing optical pump, THz probe spectroscopy on-chip instead of in free space. The sample is much thinner than either the pump or the probe skin depth, so no multilayer analysis is required. This is an active area of research, and only one result has been submitted at the time of writing [82].

3.1.1 Warning: the interpretation of $\epsilon(\omega)$ and $\mu(\omega)$ in THz experiments is often non-physical

Classic electromagnetics texts [20, 21] suggest that one should always use $\epsilon(\omega) = \epsilon_1(\omega) + \epsilon_2(\omega)$ and rarely $\mu(\omega) = \mu_1 + j\mu_2$ at THz, IR, and optical frequencies. We will now demonstrate that this wisdom often applies for free space THz TDS, but does not apply to on-chip THz.

To see why $\mu(\omega)$ is considered to be physically meaningless in free space THz experiments, ignore Ohmic current densities and consider only the two bound current densities in eqn. 1.23: $\overline{J_{bound}} = \nabla \times \overline{M} + j\omega\overline{P}$. Interpreting the current density using only $\mu(\omega)$ requires us to neglect $\epsilon(\omega)$: namely, neglect contributions from $j\omega\overline{P}$. Interpreting the current density using only $\epsilon(\omega)$ requires us to neglect $\mu(\omega)$: namely, neglect contributions from $\nabla \times \overline{M}$. It is obvious that $\omega\overline{P}$ will dominate $\nabla \times \overline{M}$ when ω becomes large, making

pure-$\epsilon(\omega)$ descriptions more physical than pure $\mu(\omega)$ descriptions. The transition for this is usually taken in the THz frequency bands. Most textbooks conclude that using $\mu(\omega)$ is non-physical because bound polarization currents dominate bound magnetization currents. The conventional wisdom is imprecise: it doesn't clarify what is meant by "large" frequencies. We will show that "large frequencies" more properly means electrically large bodies. On-chip THz, notedly, is best applied to electrically small bodies. We must answer: does $\mu(\omega)$, $\epsilon(\omega)$, or neither have the most physical meanings at THz frequencies?

If the physical interpretation of $\mu(\omega)$ as purely induced magnetization is to hold, we must be in the limit given by:

$$|j\omega\overline{P}| \ll |\nabla \times \overline{M}| \tag{3.1}$$

Landau and Liftshitz envisioned a contrived scenario in which interpreting $\mu(\omega)$ as purely magnetization is most appropriate: one in which the current due to induced polarization current is small and bound magnetization current is large. This scenario is a coil of wire with length scale L and a ferrite core with susceptibility χ_m [21]. This is a classic textbook problem with with a famously strong $\overline{H}$ field along the center axis and as small an $\overline{E}$ as possible. The induced magnetization current density will scale as $\chi_m\overline{H}/L$. Assuming $\chi_e \approx 1$, the current from the induced polarization density is $|j\omega\overline{P}| \approx |j\omega\epsilon_0\overline{E}|$. We can approximate $\overline{E}$ using Faraday's law: $\nabla \times \overline{E} = -j\omega\overline{B}$. Approximating $|\nabla \times \overline{E}|$ as $|\overline{E}|/L$ gives the scaling relation $|\overline{E}| \approx j\omega L\mu_0(1+\chi_m)\overline{H}$. Combining the scaling relations for magnetization and polarization current densities, we derive the scaling relation for the polarization current $|j\omega\overline{P}| \approx |\omega^2 L\mu_0\epsilon_0(1+\chi_m)\overline{H}|$.

For the magnetization current to dominate the polarization current, we require eqn. 3.1 to hold. Substituting our above quantities, we require $|\omega^2 L\epsilon_0\mu_0(1+\chi_m)H| <\ll$

$\chi_m H/L$. This can be simplified to:

$$\frac{L}{\lambda} \ll \frac{1}{2\pi}\sqrt{\frac{\chi_m}{1+\chi_m}} \tag{3.2}$$

The right side of eqn. 3.2 is zero at zero χ_m and saturates at $(2\pi)^{-1}$ at large χ_m. The left size is the electrical size of the object, which is given by ratio of the relevant length scale to the wavelength. The most favorable condition for the physical interpretation of $\mu(\omega)$ is if the coil has a high permeability material and is electrically small (meaning $L/\lambda \ll 1$). At THz frequencies ($\lambda \approx 300$ μm), this scenario is possible with lithographically small (L ≈ 5 μm) coils with mumetal, purified iron, or permalloy cores [3]. Although no such wire coils will be used in this dissertation, eqn. 3.2 suggests that the electrodynamics of electrically small, highly magnetized systems are best-described by $\mu(\omega)$, not $\epsilon(\omega)$. This begs the question: what about free space and on-chip THz TDS?

Consider a plane wave propagating in a medium with length scale L and electric and magnetic susceptibilities χ_E and χ_H. To maintain some symmetry with the previous discussion, we explore the magnetization-pure interpretation $|j\omega\overline{P}| \ll |\nabla \times \overline{M}|$. We need only flip the final $\ll$ to a $\gg$ to switch to the polarization-pure inequality.

The induced magnetization current $|\nabla \times \overline{M}| = |\nabla \times \chi_m \overline{H}|$ will be computed by a curl on length scale L, so the induced magnetization current, once again, will scale as $\chi_m \overline{H}/L$. The polarization current density is $|j\omega\overline{P}| = j\omega\chi_e\epsilon_0\overline{E}$. The electric and magnetic field of planes wave are related by a material impedance η through $\overline{E} = -\eta(\hat{k} \times \overline{H})$, with unit wavevector $\hat{k}$. Examining the magnitudes only, $|\overline{E}| \approx \eta|\overline{H}|$, so the polarization current density scales as $|\omega\chi_e\epsilon_0\eta\overline{H}|$.

Inserting these quantities into the inequality in eqn. 3.1 and using the definitions of $\eta = \sqrt{\mu/\epsilon}$ and $c_0 = 1/\sqrt{\epsilon_0\mu_0}$, we deduce:

$$\begin{aligned} \omega\chi_e\epsilon_0\eta H &\ll \chi_m H/L \\ \frac{L}{\lambda_0} &\ll 2\pi\frac{\chi_m}{\chi_e}\frac{\eta}{\eta_0} \end{aligned} \tag{3.3}$$

The right side quantity of eqn. 3.3 contains the ratio of material impedance (η) to that of free space (η_0), but this is largely compensated by χ_m/χ_e. The right side of eqn. 3.3, therefore, is a number typically on the order of unity, but could be exceptionally large in the case of very high permeability materials (such as mumetal with $\chi_m \approx 1E5$ or supermalloy with $\chi_m \approx 1E6$) or very small in the case of extremely high dielectric materials (such as rutile, with $\chi_e \approx 180$). Most materials used in this dissertation are in neither such limit, so the term can be thought of as order unity.

The left side quantity of eqn. 3.3 is the body's electrical size. The right hand size is a quantity set to zeroth order by $2\pi\chi_m/\chi_e$. The right hand dictates how electrically small the body must be for the interpretation of $\mu(\omega)$ as magnetization to hold true. To make further examination more convenient, we simplify eqn. 3.3:

$$1 \ll 2\pi\frac{\lambda_0}{L}\frac{\chi_m}{\chi_e}\sqrt{\frac{(1+\chi_e)}{(1+\chi_m)}} \tag{3.4}$$

We plot the right side term of eqn. 3.4 versus both χ_m and χ_e when λ_0/L is large (1E2), intermediate (1), and small (1E-2) in the top, middle, and bottom panels of Fig. 3.3, respectively. This produces a map showing regions of pure $\epsilon(\omega)$-favorable interpretation (term is $\gg 1$), pure $\mu(\omega)$-favorable interpretation(term is $\ll 1$), or an intermediate region where neither pure description is physical (term is ~ 1). The pure ϵ region occurs where χ_e is large and χ_m is small. The pure μ region occurs, as one might expect, occurs at large χ_m and small χ_e. It is inappropriate to assign the standard physical interpretation to $\epsilon(\omega)$ or $\mu(\omega)$. The $\mu(\omega)$-favorable region is largest for electrically small systems; the $\epsilon(\omega)$-favorable region is largest for electrically large systems.

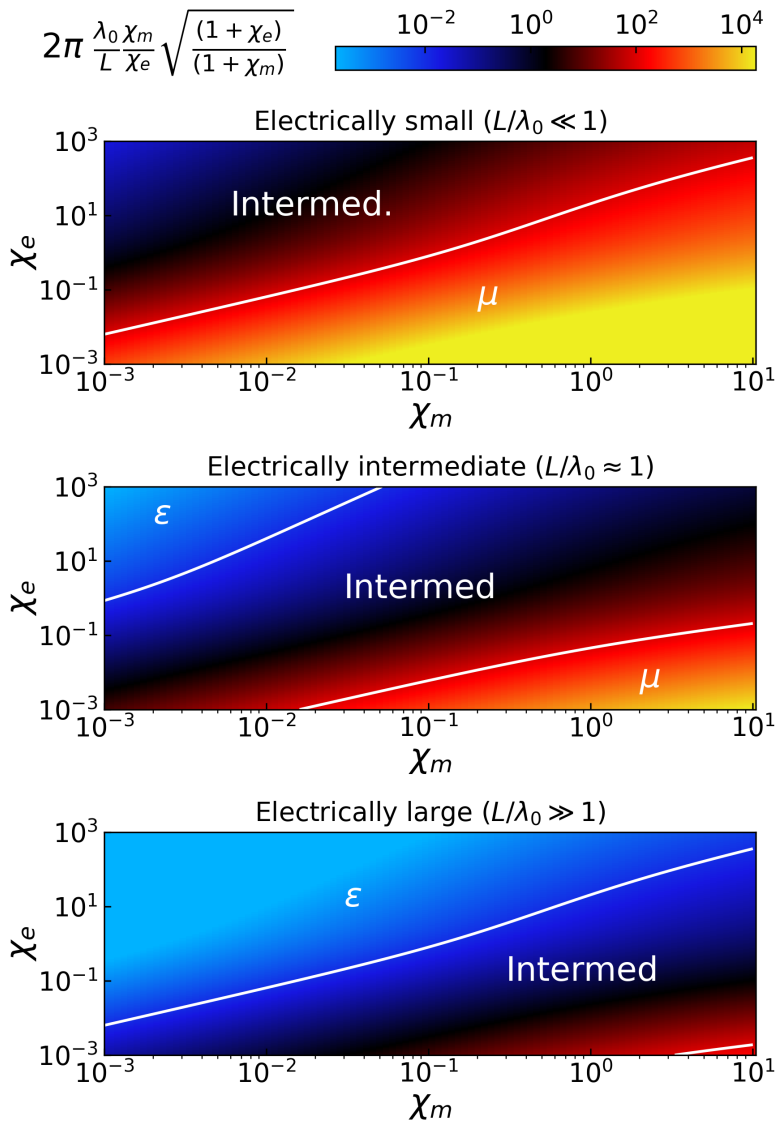

Figure 3.3: Regions of validity for 3.4 as a function of χ_m, χ_e, and electrical length. Contours for the much greater than ($\epsilon(\omega)$-favorable) and much lesser ($\mu(\omega)$-favorable) than regions are set when the right side of eqn. 3.4 takes values 1E2 and 1E-2.

It is clear that the electrical size of a body must be evaluated before choosing the most physical description. Some problems, however, have multiple length scales. The (correct) length scale L to consider is the length scale at which the operation $\nabla \times \overline{M}$ is evaluated. For example, consider the geometry shown in 3.2A. We take a monochromatic

free space optical beam (plane wave) used to illuminate a 3D sample that is electrically large in the transverse X-Y plane, but is electrically small in the Z-axis. $\overline{H}$ of the THz pulse lies in the sample plane, with the length scale of the Rayleigh limit. At 300 GHz, the free space wavelength is 1 mm (Rayleigh limit $\approx$ 1.2 mm), so a sample with 6 mm side length would be evaluated using $L/\lambda \approx 5$. An $\epsilon(\omega)$-pure/$\mu(\omega)$-pure interpretation map for such a free space THz system is shown in Fig. 3.4.

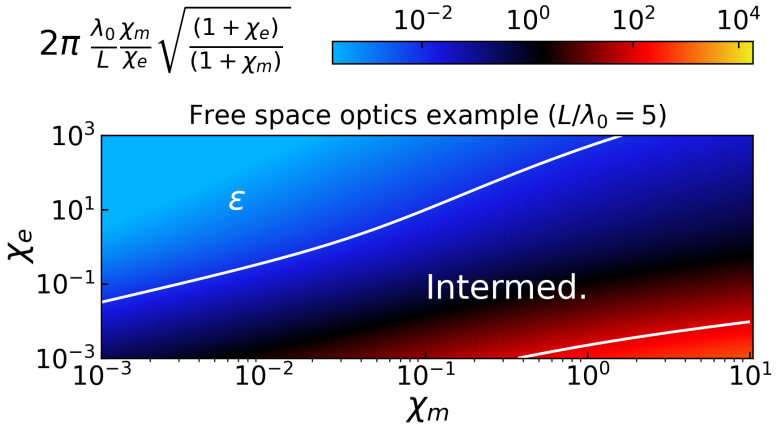

Figure 3.4: Validity of interpretation of $\epsilon(\omega)$ and $\mu(\omega)$ for free-space THz spectroscopy with $L/\lambda_0 = 5$.

Typical THz samples have values of χ_e between 1E-1 and 1E2 and χ_m between 0 and 1E-2 [3]. Almost all of these samples lie in the $\epsilon(\omega)$-favorable region, and thus it is appropriate to describe their behavior using $\epsilon(\omega)$ and not appropriate to describe their behavior using $\mu(\omega)$. Free space THz literature often reports parameters using $\epsilon(\omega)$, which is physically sensible.

Suppose now that the THz radiation were confined in a CPS transmission line separated by dimension T with sample of lateral length D. In the plane of the sample, the electric field $\overline{E} = E\hat{x}$ between the transmission lines. The $\overline{H}$-field circulates around the metallic strips in the XZ plane. The characteristic dimension of the $|\nabla \times \overline{M}|$ operation, therefore, is T. We will use such devices with a T of 30 μm or smaller. A $\epsilon(\omega)$ and

$\mu(\omega)$ validity map can be produced using 300 GHz ($\lambda_0 = 1$ mm) with electrical length $L/\lambda_0 = 0.03$. This map is shown in Fig. 3.5.

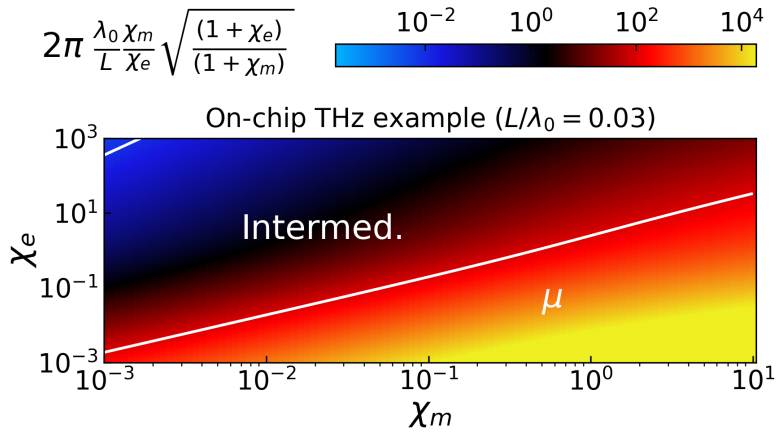

Figure 3.5: Validity of interpretation of $\epsilon(\omega)$ and $\mu(\omega)$ for free-space THz spectroscopy with $L/\lambda_0 = 0.03$.

The $\epsilon(\omega)$-valid region for on-chip THz is very small and may only be achieved when both χ_e and χ_m are very small. The samples we considered for free space THz (χ_e between 1E-1 and 1E2 and χ_m between 0 and 1E-2) would mostly lie within the intermediate interpretation region if measured with on-chip techniques. We therefore conclude that neither $\epsilon(\omega)$ nor $\mu(\omega)$ have strong physical interpretation for plane waves (or sums thereof) in on-chip THz measurements. This does not mean one cannot use $\mu(\omega)$ as a proper, well-defined material response function. It only means that interpreting $\mu(\omega)$ as originating from an AC magnetization density only is fallacious.

There is a less ambiguous approach. The four terms in the right hand side of eqn. 1.23 are all current densities, from external sources, Ohmic currents, bound magnetization, and bound polarization contributions. We can define a new conductivity, often called the optical conductivity, the infrared (IR) conductivity, or the THz conductivity $\sigma(\omega) = \sigma_1(\omega) + j\sigma(\omega)$. This conductivity describes all currents in response to THz excitation and therefore cannot break down in different limits. The preferred response function for

on-chip THz measurements is, therefore, intrinsic THz conductivity $\sigma(\omega)$ or $\sigma(\omega)$ times geometric parameters.

3.2 THz-frequency passive element engineering

The THz bands may be "just" another region of the electromagnetic spectrum, THz engineering is particularly difficult and less widely known than both microwave engineering and optical engineering for the visible spectrum. THz radiation in transmission lines behaves, to zeroth order, as RF and microwave radiation does. THz, however, is generally much lossier than RF and microwave radiation. There are three primary mechanisms for loss: metallic loss, dielectric loss, and radiative loss, all of which are exacerbated by operation at THz frequencies. This loss can be upwards of 30 dB/m [83]. This so severe that, if unmitigated, THz radiation cannot be measured after propagating for 5 mm or more. In addition, the mechanisms of loss by Cherenkov radiation and unbound leaky wave modes, which are irrelevant at lower frequencies, become important and even dominant at THz frequencies.

3.2.1 Very small classical skin depths

Metallic transmission lines are much lossier at THz frequencies than at microwave frequencies. This is modelled by a large resistance per unit length R' in the telegraphist's equations. R' for realistic dimensions, even at DC frequencies, can be surprisingly large. The metallic strips used for THz transmission lines have width 30 μm or smaller. A room temperature gold stripline with thickness 100 nm and width 30 μm has a surprisingly large DC resistance of 73 Ω/cm.

The resistance per unit length at THz frequencies is much higher than that at DC. THz frequency carriers do not occupy the conductor's entire width. Lateral field pene-

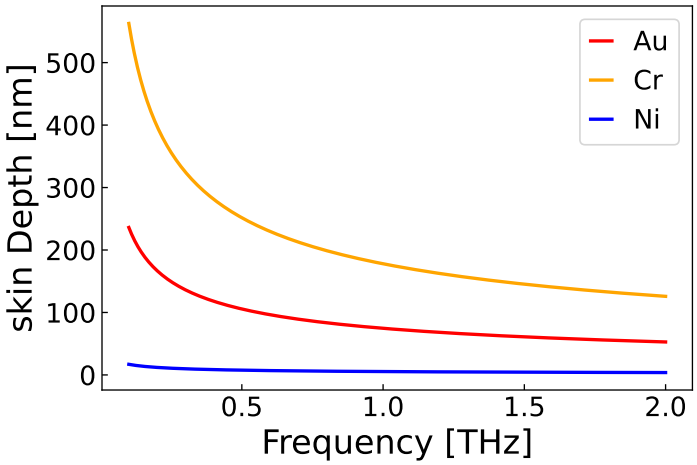

Figure 3.6: Classical skin depth of three common metals at THz frequencies. Fields only penetrate a few hundred nm into the metal. Nickel's skin depth is exceptionally small because of its large magnetic permeability.

tration into the metal is given by the skin depth, which goes $1/\sqrt{\mu\omega}$ for a good conductor (Table 1.2), resulting in the sheet resistance of $\sqrt{\omega\mu/2\sigma}$. A few skin depths of common materials at THz frequencies is shown in Table 3.6. Since the fields are restricted to only the few hundred nanometers near the conductor edge (and even less in magnetized metals), the THz resistance per unit length increases even though the DC resistivity stays constant. The loss at 1 THz ∼33 times that at 1 GHz and 1000 times that at 1 MHz.

The intrinsic THz-frequency conductor loss can only be combatted by depositing at least 1-2 skin depths of metal [84], decreasing the metal's permeability and increasing the metal's conductivity. Avoid magnetic metals to increase the skin depth. Cryogenically cooling metals increases their conductivity. This is quantified by the residual resistivity ratio (RRR) - the ratio of the room temperature resistivity to the cryogenic (typically 4K or 10K) resistivity. Cryogenic cooling improves the conductivity by a factor of RRR, so skin depth-related losses decrease by $\sqrt{1/RRR}$. For gold, the RRR from 300K to 10K is 101, so the conductor loss would naïvely improve by $\sqrt{101} \approx 10$. As we will show below, these expressions fail to capture the anomalous skin effect, which will limit the benefits

of cooling.

The skin effect also alters the transmission line's L' and C'. As frequency increases and the skin depth drops, the current bunching near the conductor edges changes the self-inductance and capacitance. This can change the phase velocity versus frequency, meaning that the 200 GHz Fourier component will travel faster than the 750 GHz Fourier component, resulting in dispersion and signal broadening. Such effects are best captured by FEM analysis. We explore the dispersion of phase velocity at THz frequencies later in this chapter.

In summary, Ohmic losses scale as $\sqrt{\omega\mu}$ and $\sqrt{1/RRR}$. They are best combatted by avoiding magnetic metals and operating at cryogenic temperatures, if possible. Cryogenic cooling, however, is limited to classical skin depths - where the classical expressions derived in Table 1.2 apply.

3.2.2 Surface roughness scattering

Surface roughness refers to wave scattering from both substrates and conductors with an irregular distribution of material. Surface roughness is characterized by an RMS height ζ_s. Most of a conductor's current is contained within the first few skin depths (δ) from the surface. One might expect roughness loss to increase when $\zeta_s \approx \delta$. Empirical relations have been proposed to approximate surface roughness loss, most famously as the Morgan-Hammerstad-Bekkadal model for microstrips [85, 86] with conductor skin depth δ_c:

$$\alpha^{'} = \alpha_0 \left(1 + \frac{2}{\pi} atan\left(\sqrt{2}\frac{\zeta_s}{\delta_c}\right)\right) \tag{3.5}$$

In equation 3.5, α_0 ($\alpha^{'}$) is the extinction coefficient before (after) accounting for surface roughness. The model strictly only applies to microstrip transmission lines but

has been shown to be approximate for other transmission line geometries [87]. Eqn. 3.5 suggests that surface roughness scattering only becomes a significant loss when the RMS surface roughness and skin-depth are comparable. A rule of thumb: when $\zeta_c = \delta_c$, the extinction coefficient will increase by 60%.

We can expand the Morgan-Hammerstad-Bekkadal model in small ζ_s/δ_c, which describes the increase in extinction for small surface roughnesses:

$$\alpha^{'} \approx \alpha_0 \left(1 + \frac{2}{\pi}\frac{\zeta}{\delta_c} - \frac{2}{3\pi}\left(\frac{\zeta_s}{\delta_c}\right)^3\right) \tag{3.6}$$

We can now estimate the surface roughness loss due to nanofabrication imperfections. The surface roughness that matters in a coplanar transmission line is the conductor roughness of the transmission line metal in the first few skin depths from the edge. The un-evenness of the deposited metal dominates to the total roughness. For dual-side polished wafers and a 100 nm Au metallization layer[2] deposited by electron-beam deposition processes, the total roughness is expected to be 1 nm or less, producing less than 1% increase in extinction coefficient at 1.0 THz.

Another source of surface roughness loss are nanofabrication dog ears, by which imperfect lift-off processes can produce bumps at the conductor edge ranging from zero thickness to that of the metal layer. Since dog ear size is comparable to relevant skin depths, dog ears can cause surface roughness loss. We model the dog ears as a uniform distribution, whose RMS value is $A\sqrt{1/12}$, where A is the metal layer thickness. For a 100 nm metallization layer, this increases the extinction coefficient by 23.5% at 1 THz, resulting in a substantial degradation in performance. Note that the same sized dog ears cause less than 1% increase in attenuation at frequencies less than 2 GHz.

The frequency dependence of eqn. 3.5 can be ascertained by using the skin depth

[2]We often use 10 nm of Cr or Ti as an intermediary to improve mechanical adhesion between the quartz and Au. This is ignored in these calculations.

formula from Table 1.2 for good conductors. Substituting and simplifying produces:

$$\alpha^{'} = \alpha_0 \left(1 + \frac{2}{\pi} atan(\zeta\sqrt{\omega\mu\sigma})\right) \tag{3.7}$$

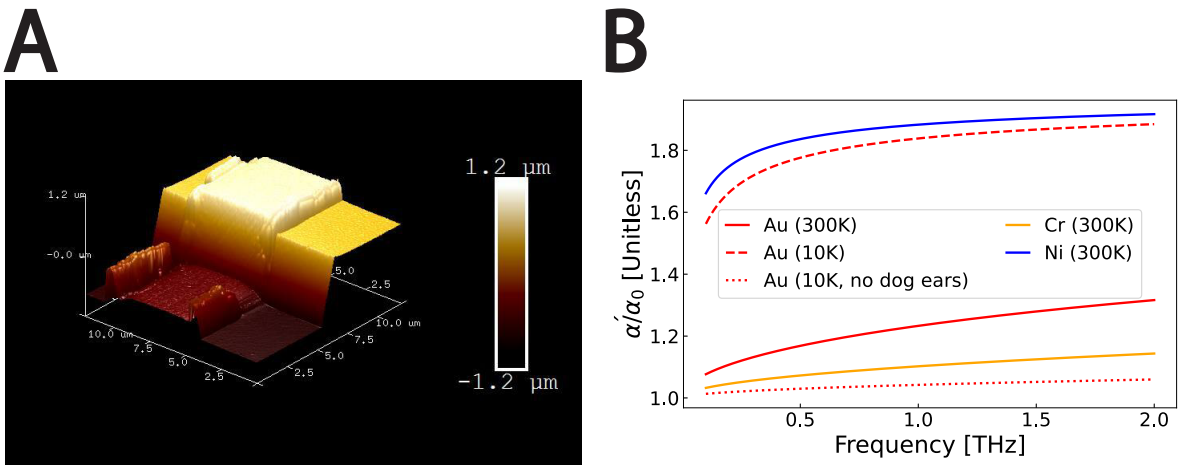

Figure 3.7: (A) Atomic force microscopy of nanofabrication dog ears on a conformal contact to a III-IV mesa (right). Stochastic bumps on the metal periphery are dog ears. (B) Dog ear-related increase in surface roughness loss for three metals at THz frequencies. Cryogenically cooled gold suffers because of the reduced skin depth from the increased conductivity.

It is clear that surface roughness loss will increase with frequency for the same surface roughness. We plot the surface roughness for gold, chromium, and nickel in Fig. 3.6. Nickel's inflated losses are worse than the other metals due to its large magnetic permeability. The surf roughness loss of Au is (counter-intuitively) worse than that of Cr, despite Au being a better metal than Cr. This stems from the σ in the atan in 3.7. The higher the conductivity, the smaller the skin depth and the more "relevant" the surface roughness becomes. For example, cryogenically cooling gold will increase the conductivity and decrease the skin depth by $\sqrt{RRR}$. This makes the ζ/δ_c increase, substantially worsening surface roughness loss.

Fig. 3.7 assumes the dog ear model proposed above. If there are no dog ears and the

surface roughness is set by the metal and underlying dielectric substrate roughness ($\zeta_c \approx$ 0.5 nm), the surface roughness loss increase is less than 6% at frequencies $\leq$1 THz. Even though surface roughness loss could be mitigated by using a lower conductivity metal, the conductor loss will increase. It is, therefore, not advisable to use poor metals.

In summary, surface roughness is most severe when surface roughness ζ is much smaller than the skin depth, which is very small at THz frequencies. The small skin depth in the terahertz bands makes such surface roughness for common fabrication problems much worse than in the microwave frequency bands. Surface roughness losses can be best mitigated by avoiding magnetic materials, using polished, optics grade substrates, and avoiding dog ears in nanofabrication by using subtractive processing or a bilayer resist. I recommend using electron beam deposition of titanium and gold deposted at $\leq$4 Å/s. Further details of such processes are in appendix A.

3.2.3 The anomalous skin effect

Conductivity-independent skin depths in good metals at high frequency was first reported in the late 1940s. Classical skin depths should scale as $\delta \sim \sqrt{1/\sigma}$ (Table 1.2), so this new phenomenology was dubbed the "anomalous" skin effect. The anomalous skin effect was shown to come from a microscopic violation of Ohm's law. This occurs when a carrier's mean free path L_e becomes comparable to the skin depth δ, so carriers do not experience an exponentially decaying electric field and macroscopic description by Ohm's law is microscopically violated. The effect was modelled numerically by locally solving the Boltzmann transport equation, which produced a non-exponential electric field decay profile [88, 89, 90, 91, 92]. Impinging electric and magnetic fields are, thus, not attenuated within the first mean free path, no matter the conductivity, resulting in the reported anomalous skin effect.

Detailed computation of the electric field and current profiles resulting from the anomalous skin effect requires solving a complicated integro-differential equation, most cases of which have no compact form. There is an alternative. Assume a semi-infinite vacuum adjacent to a semi-infinite good metal with unit normal vector $\hat{n}$ from the metal to the vacuum. The tangential fields $\overline{E_t}$ and $\overline{H_t}$ in the vacuum must satisfy the impedance relation: $\overline{E_t} = \eta_0 \overline{H_t} \times \hat{n}$. Since the fields must be continuous across the boundary, this impedance condition must also be true just inside the metal, but with vacuum impedance η_0 replaced by the material's surface impedance $\zeta(\omega) = \zeta_1(\omega) + j\zeta_2(\omega)$. This is the Leontovic boundary condition, which must be true even in the presence of the anomalous skin effect [21]. The approach to incorporating the anomalous skin effect in calculation is to define $\zeta(\omega)$

The relevance of the anomalous skin effect is set by $\Lambda_m = 3L_m^2/2\delta^2$ with classical skin depth δ. When incident reflected waves are much more specular than diffuse and $\omega\tau \ll 1$ (Drude momentum relaxation time τ), the surface impedance expressions simplify [91, 93]. In the non-anomalous limit ($\Lambda_m \gg 1$), the surface impedance is the classical expression $(1+j)\sqrt{\omega\mu/2\sigma}$ (Table 1.2). In the extreme anomalous limit ($\Lambda_m \ll 1$), the surface impedance is:

$$\zeta(\omega) = \sqrt[3]{\frac{\sqrt{3}\mu^2\omega^2 L_m}{16\pi\sigma}}(1 + j\sqrt{3}) \tag{3.8}$$

Note that both cases have a positive $\zeta_2(\omega)$, implying an inductive reactance. This is intuitively obvious - we know that ballistic electrons should have an inductance, from our discussion of perfect metals in Ch. 2. The real part of eqn. 3.8, which is the resistive part of the surface impedance, scales as $\omega^{2/3}$. The loss associated with the classical skin effect scales as $\omega^{1/2}$, meaning that the anomalous skin effect-related losses will outpace the classical skin effect losses at frequencies above critical frequency ω_c given by:

$$\omega_c = 32\pi^2/3\sigma\mu L_m^2 \tag{3.9}$$

In gold at 300K, this is on the order of 1E30 Hz, which is a ludicrously unfeasible frequency. The frequency-dependent loss of the anomalous skin effect is less than that of the classical skin effect at realistic frequencies, which stems from the ballistic transport of the carriers within one mean free path of the surface. It is worth nothing that the room-temperature mean free paths of gold, chrome, and nickel have been estimated or measured to be 38 nm, 15 nm, 5 nm, respectively [94]. This suggests that the 1.0 THz room temperature skin depths of 75 nm, 178 nm, and 5.4 nm are approaching the anomalous limit, and that significant cryogenic cooling will not reduce the loss by $\sqrt{RRR}$.

In summary, the anomalous skin effect prevents cryogenic cooling from reducing skin effect losses and can contribute surface inductance in the extreme anomalous limit.

3.2.4 Substrate dielectric loss

In addition to the metal, substrates become lossier at THz frequencies. As derived in Ch. 1, the extinction coefficient of a wave propagating in a good dielectric goes as $\omega\sqrt{\mu\epsilon}\, tan\delta_e$. At first glance, losses are frequency-independent, as the explicit frequency dependence of the loss tangent cancels the frequency dependence in extinction coefficient expression. The implicit frequency dependences of σ and ϵ, as well as the non-Debye power law of the Jonscher universal dielectric response, can lead to additional frequency dependence for commonly-used THz substrates.

The Debye model was extended by Svensson et. al. [95] and Djordjevic et. al. [96] for operation at high frequencies to a model now named the Svensson or Djordjevic-Sarkar model. The model assumes a continuous distribution of relaxation processes with different strengths and different dielectric relaxation times. The sum of individual Debye terms

becomes an integral between phenomenologically chosen smallest and largest relaxation times τ_1 and τ_2, resulting in:

$$\begin{aligned} \epsilon(\omega) &= \epsilon_d + \int_{\tau_1}^{\tau_2} \frac{1}{\tau'} \frac{1}{1+j\omega\tau'} d\tau' \\ \epsilon(\omega) &= \epsilon_d + a\left(\frac{1}{2}\ln\left(\frac{\tau_2^2(1+\omega^2\tau_1^2}{\tau_1^2(1+\omega^2\tau_2^2}\right) + j(atan(\omega\tau_1) - atan(\omega\tau_2))\right) \end{aligned} \tag{3.10}$$

The Svensson model in eqn. 3.10 describes a (Kramers-Kronig-respecting) transition between low frequency ($\omega \leq 1/\tau_1$) dielectric and high frequency ($\omega >\geq 1/\tau_2$) dielectric behaviors, with a loss tangent approximately constant in the transition regime. Many substrates follow this model with 1 THz $\approx 1/2\pi\tau_2$, so this model is often can be used to model dielectric materials for microwave and THz frequency substrates. It is, therefore, absolutely critical to choose substrates with small loss tangents, as losses will approximately scale as $\omega \tan(\delta_e) \sim \omega$ in the Svensson model. If the high-frequency Svensson model cutoff is less than the desired spectral range, other models are required. Loss beyond the transition regime is given crudely by the Debye model (ω^{-2}) or more accurately by the Jonscher dielectric exponent (ω^{-n-1} with $0 < n \leq 1$), but the basic loss by dielectric absorption doesn't change.

A rule of thumb: substrates with larger DC permittivities will be lossier at high frequencies. This is a direct consequence of the Kramers-Kronig relations. At (truly) infinite frequency, electromagnetic wave must experience a relative real permittivity of 1, no matter the substrate. An overall multiplicative factor (i.e. larger DC permittivity) in ϵ_1 must manifest as a larger ϵ_2, resulting in greater loss. One should note that this overall scale factor does not appear when calculating the loss tangent. This is another lesson in caution with using the loss tangent: high and low permittivity dielectrics may have the same loss tangent, but different absolute losses.

The approximately constant $\tan(\delta_e)$ only applies between dielectric relaxation times

τ_1 and τ_2. These often could correspond to specific material defect states or bands thereof, but could also come from low energy quasiparticles like phonons, charge density waves, spin waves... etc. One must be cautious in using the Svensson model across broad temperature ranges, as some defects and phonons do freeze out, but some low energy quasiparticles may emerge at low temperature (for example, strontium titanate [97]).

In summary, substrate dielectric losses can be best avoided by choosing substrates with small loss tangents and slightly helped by choosing substrates with small $\epsilon(\omega)$. Data from a few choice substrates will be shown in Ch. 4.

3.2.5 Radiative loss through parasitic antennas

Geometric radiation losses occur because circuit elements, such as transmission lines, couplers, and filters, act as parasitic antennas. Though difficult to calculate analytically, it is well-accepted that radiation loss increases with frequency, particularly increasing between 0.06 and 0.1 THz[98, 99]. Radiation and conduction loss can be independently estimated numerically by simulating a geometry twice: once with all perfect electric conductor (PEC) transmission lines and the other with finite conductivity metal lines. The loss in the pure-PEC computation are from radiation; the loss in the non-PEC case are from radiation and conduction.

There are a few relevant design rules to avoid geometric radiative loss:

1. **Avoid features with length** $n\lambda/2$

   Avoid designing circuit elements with $n\lambda/2$ for integer n. Such elements are (multiples of) parasitic half-wavelength dipoles that radiate energy. This is easy in a continuous wave spectrometer, where only a few lengths need to be avoided. Avoiding half-wavelength dipoles is much more difficult in broadband design, when operating frequency is specified by a range of wavelengths than a single number.

For the spectrometer described in Ch. 4 with operating wavelengths and quasi-TEM index of 1.46, circuit elements between lengths 1.0 mm and 2.7 mm (n = 1), 2.0 mm and 5.2 mm (n = 2)... etc. It quickly becomes apparent that this is impossible with any spectrometer with fractional bandwidth 1/2 or greater. It is therefore advisable to minimize and mollify radiating elements using numerical techniques.

2. **Minimize current loop size**

   Fields can also be thought of as radiating from loops between active signal-carrying transmission line metal and the current return (ground) metal[3] [3]. These loops are illustrated in Fig. 7 and 9 of [99], and can be reduced by shrinking transmission line widths or adding a ground plane. Using differential transmission line geometries and regular grounding with vias, particularly in coplanar waveguide transmission can drastically reduce radiation losses by reducing current loop sizes[99, 100]. This will not be vital consideration in the spectrometer in this dissertation, as our circuit sizes are dictated by sample size rather than loss considerations.

3. **Avoid impedance mismatches**

   Impedance mismatches reflect signals at THz frequencies, just like in microwave bands. As the frequency increases into the mm-wave and THz bands, the wavelength shrinks and the electrical size of the current loops described above increases. Impedance mismatches increasingly begin to radiate, increasing loss. Open circuits at low frequency are often designed by simple terminating transmission lines. The field lines "leaking" out of the transmission lines couple more strongly to free space at THz frequencies than at GHz frequencies.

[3]Such current loops can be shown to be equivalent to fictitious magnetic monopole current densities, subject to Maxwell's equations. Radiation loss can be thought of as an $n\lambda/2$ process for magnetic monopole currents.

Parallel plate waveguides don't suffer the same radiation losses that substrate-borne transmission lines do, as waveguides are completely encased by metal. Waveguides have notoriously narrow fractional bandwidths, and are thus not suitable to the broadband needs of spectroscopy. Even if their fractional bandwidth were increased, waveguides are not widely used in THz design, as the required width is on the order of a wavelength (300 μm at 1 THz). Suppressing higher order modes requires that the other dimension be half of the width or smaller, while avoiding surface roughness loss. Constructing such waveguides requires careful machining or many cm of precise ion milling to avoid the scattering losses.

In addition to geometric radiation losses, another form of radiation loss plagues and even dominates the above losses. This is loss in the form of Cherenkov radiation.

3.2.6 Cherenkov radiation

Cherenkov radiation is a general electromagnetic radiative effect that describes radiation produced when a charge moves through a medium faster than the phase velocity of light in that medium. Coplanar transmission lines on substrate with dielectric constant ϵ carry wavelets of propagating current with a phase velocity approximately $2c_0/(1+\sqrt{\epsilon})$, which is faster than light's phase velocity in the dielectric $c_0/\sqrt{\epsilon}$.

Cherenkov radiation is the electromagnetic cousin of acoustic shock waves, wherein a body moving faster than the speed of sound produces a "cone" of propagating acoustic waves. Cherenkov radiation can analogously be pictured as a spherical wavefront emitted by a bound current wavelet. When the wavelet phase velocity exceeds the speed of light in the dielectric, the locus of spherical wavefronts becomes a cone, whose angle is set by the ratio of relevant phase velocities. Such Cherenkov cones have been experimentally demonstrated by direct pump-probe imaging in the time domain in electro-optic $LiNbO_3$

[101].

Cherenkov radiation occurs for all pulses on all coplanar transmission lines, but is exacerbated by THz operation. The attenuation due to Cherenkov radiation α_C of a coplanar stripline with geometric parameters W and S can be shown to be:

$$\alpha_C = \frac{\pi^5}{32}\frac{\sqrt{2}(3-\sqrt{8})}{K(k)K'(k)}\frac{\omega^3}{c_0^3 W^2}\left(\epsilon_r^3(1-1/\epsilon_r)^2\sqrt{1+1/\epsilon_r}\right) \tag{3.11}$$

K(k) and K'(k) are the elliptic integrals of $k \equiv W/(W+2S)$, which are geometrically-determined constants. The cubic frequency dependence of eqn. 3.11 and realistic transmission line geometries implies a sharp increase of Cherenkov radiation loss at frequencies near 100 GHz [83]. Realistic parameter values and experiments [102, 103, 104, 105] show that all other mechanisms for transient attenuation are dominated by the Cherenkov radiation's cubic frequency dependence. Cryogenic cooling, even when limited by the anomalous skin effect, does little to reduce the ω^3 loss at frequencies above 100 GHz, implying that loss is Cherenkov-dominated [106].

Cherenkov radiation can, however, be mitigated by dielectric choice. The ϵ-dependence of eqn. 3.11 can be simply understood. Physical substrates must have a $\epsilon \geq 1$. The Cherenkov attenuation coefficient is meaningless for $\epsilon < 1$, goes to zero when $\epsilon = 1$, and then rises, with large ϵ behavior going as ϵ^3. Cherenkov radiation loss is zero for $\epsilon = 1$ because this reflects the same dielectric environment in the substrate and the superstrate. If a substrate could be made of vacuum, the current wavelet in the transmission line would be propagating at light's phase velocity in the substrate. As this phase velocity mismatch grows, Cherenkov radiation grows rapidly. To reduce Cherenkov radiation, the substrate $\epsilon(\omega)$ should be reduced as much as possible.

In addition, eqn. 3.11 assumes a semi-infinite substrate. Substrates whose vertical thickness is less than the characteristic mode length scale (W in a coplanar stripline)

present a reduced effective ϵ to the propagating mode, thereby reducing Cherenkov loss. Reducing a substrate thickness to be much less than the mode length scale, however, does not mitigate Cherenkov loss. For example: transmission lines with W = 50 μm are improved by using 50 μm-thick substrates. Transmission lines W = 5 μm are not improved by using 50 μm thick substrates. This had let to efforts in reducing the substrate thickness to micron-sized membranes, often of silicon nitride [107, 84, 106].

3.2.7 Unbound leaky wave modes

A complication arises in coplanar transmission lines on thin substrates used to combat Cherenkov attenuation. As the substrate thickness decreases, an effect arises by which Maxwell's equations admit an unbound mode. This is a propagating, leaky wave mode that is unbound to the transmission line - meaning that the mode is agnostic to the existence of any transmission line structure [108]. Though there remains a bound mode confined to the transmission line, the unbound mode propagates isotropically from the point of generation. This unbound mode can travel in free space with phase velocity c_0, along the substrate-superstrate interface as a surface wave, or freely in the substrate. Furthermore, the emitted fields can reflect from the substrate edges, further complicating the local electric field profile.

To see this, we examine data taken with a Protemics scanning THz system at Fraunhoffer HHI (Fig. 3.8). A CPS was fabricated on an InP substrate ($\epsilon_\infty \approx 9.6$) from the origin to the top of the scan window, aligned with the Y-axis. The CPS has a photoconductive switch at the origin, which is excited by an incident laser pulse near 0 ps delay. The photoconductive device generates a THz transient. The scanning tip contains a narrow CPS and photoconductive switch, which is rastered 20 μm above the from the substrate surface. The result is the time-resolved electric field versus X and Y position

at 50 fs time steps with a spatial resolution of about 15 μm.

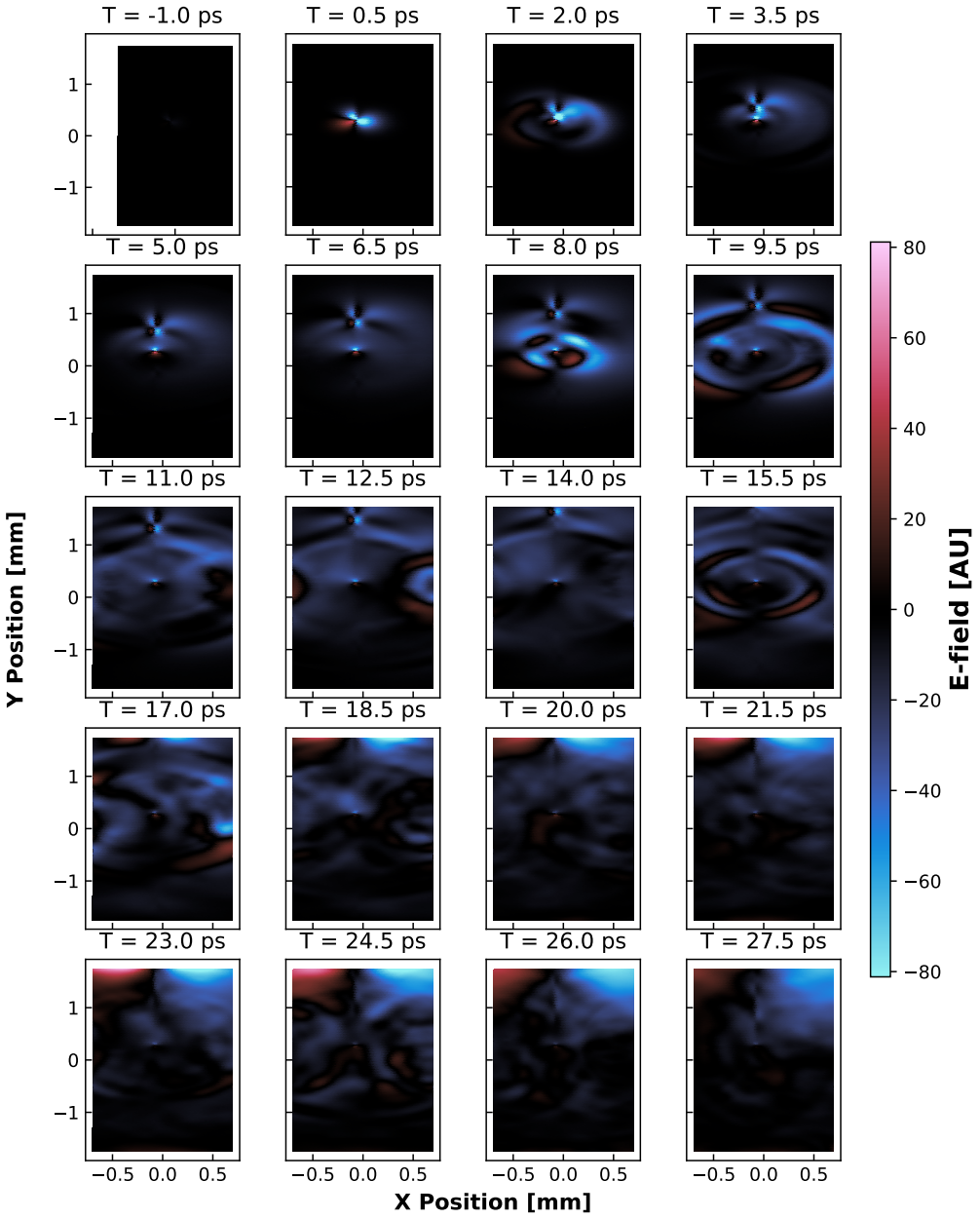

Figure 3.8: 2D spatiotemporal imaging of unbound, leaky wave substrate modes in 500 μm-thick InP using a Protemics scanning THz TDS system (20 μm above the sample surface) after photoexcitation at 0 ps delay. The bound transient propagates down the transmission, and is accompanied by a number of other features. The features not associated with the primary transient on the transmission line are due to unbound, leaky wave modes and their reflections. The pictured data is courtesy of Protemics GmbH and Fraunhofer HHI.

Fig. 3.8 shows several wavefronts emanating from the photoconductive device. At

delays less than 8.0 ps, the primary two wavefronts are a fast-propagating circular wavefront and slow-moving wave packet moving along the Y-axis. This circular wavefront has phase velocity speed 281 $\pm$ 21 μm/ps, which is the vacuum speed of light (300 μm/ps). This is a an unbound propagating vacuum wave whose wave properties are agnostic to transmission line. The slow-moving wave packet is a transient bound in the transmission line, which follows the familiar transmission line formulas from Table 1.3. This is the useful THz signal that can be engineered to interact with samples - meaning that all other THz electric fields are parasitic to interpretation and should be avoided. There is also a circular wavefront moving at the same phase velocity as the bound transient. This is from contributions by a Cherenkov shockwave and an unbound mode that confined to the dielectric (with the substrate acting as a parasitic dielectric waveguide). It is impossible to sensibly disentangle the electric fields of this substrate mode from the electric fields of the bound transient, since they move at the same phase velocity.

Additional electric fields appear at delays of 8.0 ps and larger. Wavefronts emitted vertically at zero delay propagate through the dielectric and reflect from the substrate undersides or edges. These modes can reflect from the substrate underside and manifest as additional unbound waves in the dielectric or above the substrate surface. These modes reflect from the substrates edges (11.0 ps and later) and create complicated field textures, which makes isolating contributions from only the bound mode impossible to isolate.

Unbound, leaky-modes must be suppressed by using a thick substrate (which would increase Cherenkov loss), reducing substrate dielectric coefficient, and/or careful choice of transmission line S and W. Analytic expressions to determine bound wave-only geometries and dielectric constants may be found in [108]. More simply: one can choose low dielectric constant substrates and geometries that have been shown in literature to not permit unbound, leaky wave mode propagation.

3.2.8 Geometric effects on dispersion and phase velocity

In Fig. 3.8, the bound wave packet propagates with $v_p = 91 \pm 8$ μm/ps. The high-frequency dielectric constant of InP is 9.7. One expects the quasi-TEM index to be the square root of the substrate and superstrate average dielectric: $\frac{1}{2}(1+\sqrt{\epsilon})$ and the corresponding phase velocity to be 146 μm/ps. The observed wave is however, much slower. Using only the substrate dielectric, one derives a phase velocity of $c_0/\sqrt{\epsilon_r} \approx 96\mu$m/ps, consistent with measurement. The slow wave near the CPS requires additional explanation (though this can be attributed to bound substrate modes away from the dielectric).

If one expands the transmission line propagation coefficient for $R/\omega L' \ll 1$ and $G/\omega C' \ll 1$ to second order, one can derive a propagation coefficient:

$$v_p = \frac{1}{\sqrt{L'C'(1+\frac{1}{\omega^2}\left(\frac{R'^2}{8L'^2}+\frac{G'^2}{8C'^2}-\frac{R'G'}{4L'C'}+\mathcal{O}(\omega^{-4})\right))}} \tag{3.12}$$

Eqn. 3.12 includes the skin effect ($R' \sim \sqrt{\omega}$), Svensson model for substrate leakage ($G \sim \sigma_1 \sim a_2 \omega atan(\omega\tau)$), and the anomalous skin effects on L'. If the term multiplying $1/\omega^2$ is positive, the phase velocity will saturate to $1/\sqrt{L'C'}$ at high frequencies. If the term multiplying $1/\omega^2$ is negative, the wave will slow down at high frequencies.

The complications of eqn. 3.12 suggest numerical modelling a better way to compute phase velocity. A phenomenological model [100, 109] based on numerical analyses of CPS geometries suggests THz pulses obey the effective high frequency $\epsilon_{eff}(\omega)$ of:

$$\sqrt{\epsilon_{eff}(\omega)} = \frac{\sqrt{\epsilon_r+1}}{2} + \frac{\sqrt{\epsilon_r}-\frac{\sqrt{\epsilon_r+1}}{2}}{1+e^{uln(S/W)+v}(4\omega H\sqrt{\epsilon_r-1}/c_0)^{-1.8}} \tag{3.13}$$

Here, u = 0.54 - 0.64 q + 0.015 q^2, v = 0.43 - 0.86 q + 0.54 q^2, and q = ln(S/H). This phenomenological model predicts nearly constant phase velocity from 1 GHz to 100

GHz, where the formula becomes strongly dependent on S, W, H, and ϵ. Ref. [109] shows that the phase velocity slightly increases around 300 GHz when S and W are both 2 μm (with $\epsilon_r = 13.1$), but decreases when S and W are both 20 μm. The transmission lines used in Fig. 3.8 have dimensions of 10s of microns, and thus eqn. 3.12 is one explanation for the slow phase velocities here. We will show in Ch. 5 that the bound transients in useful spectrometer designs also have slow phase velocities.

3.2.9 Passive THz Summary

Dominant THz-frequency loss mechanisms and their frequency dependences are shown in table 3.1. The strongest frequency dependence is that of Cherenkov radiation. There are additional geometric factors one must consider when designing passive THz elements. These are summarized in 3.2. One should generally avoid using high-ϵ dielectrics, and use thin substrates whenever possible.

Mechanism	ω-Dependence	Mitigated by
Atmospheric absorption	$\frac{1}{(\omega_0-\omega)^2+(1/2\tau)^2}$	Operation in inert gasses Operate far from ω_0
Classical skin effect	$\sqrt{\omega}$	Metal choice, cryogenic cooling
Anomalous skin effect	$\sqrt[3]{\omega^2}$	Not mollifiable
Svensson dielectric loss	ω	Choose substrates wisely
Surface roughness	$atan\sqrt{\omega}$	Avoid nanofabrication dog ears
Cherenkov radiation	ω^3	Thin substrates with small ϵ_r

Table 3.1: Summary of frequency dependence of passive element loss mechanisms relevant to THz engineering. The atmospheric absorption model is for a Lorentzian near resonance. Note that cryogenic cooling only helps the reduce classical skin effect losses up to the anomalous skin effect. The Svensson dielectric loss is only approximately true between low and high frequency cutoffs, which should be specified per substrate.

Mechanism	Problematic when	Mitigated by
CPS qTEM→TE_0 leakage	Impedance transitions	Avoid mismatches
CPS TE_1 mode	$\omega > c_0/2H\sqrt{\epsilon_r - 1}$	Thin substrates with small ϵ_r
Metal surf. roughness loss	Using subtractive Duroid substrates	Clean room nanofabrication
Parasitic $\lambda/2$ antennas	$L \approx n\lambda/2$	Careful length choices
Parasitic loop antennas	Large ground-signal separation	Miniaturizing S & W
Leaky wave modes	Thin substrates Poor choice of S & W	Substrates with small ϵ_r careful geometry
Geometric dispersion	Poor choice of S & W	Miniaturized dimensions

Table 3.2: Summary of geometric constraints on passive THz circuit elements.

3.3 THz-frequency active element engineering

We have seen great difficulty in engineering passive elements due to the excessive loss at THz frequencies, which is further complicated by geometric constraints on substrate thickness and permittivity, and transmission line geometry. These engineering considerations represent only one class of devices. We now consider active circuit elements: THz sources and detectors. Conventional, electronic amplifiers produce little to no gain in the THz bands. High frequency operation in heterojunction bipolar transistors and high electron mobility transitiors is limited by capacitive delay in the leads and gate, as well as the carrier transit time across the active channel itself. Commercial electronic amplifiers rarely oscillate at frequencies larger than a few hundred GHz [110].

In addition, conventional laser and related optical techniques fail to operate at THz frequencies. The binary, ternary, and quaternary alloyed semiconductors whose easily-tuned electronic bandgaps are responsible for laser operation cannot be easily controlled in the meV energy scale of THz photons, as meV-scale disorder and thermal activation across narrow bandgaps can prevent operation. Though there has been great progress in the design of quantum cascade lasers and backward wave oscillators and the use of free

electron lasers, low-power heterodyne microwave mixers, and two-color optically-pumped plasmas [111, 112, 113, 114, 115, 116, 117], they are not yet available on the chip-scale.

There has been greater success adapting photoconductive (PC) switch techniques to on-chip geometries. We will examine such PC switches, explore several candidate material systems, and compare their THz signals.

3.3.1 Photoconductive (Auston) switches

The Auston switch is a photoconductive (PC) element responsible for sourcing or measuring THz radiation when illuminated by ultrafast laser pulses. The PC switch is a photosensitive semiconducting device with a short photocarrier lifetime (<1 ps). A PC switch is contacted by two electrodes separated by the diameter of a focused laser beam. A DC bias is applied across the electrodes. In the dark condition (no illumination), the PC switch is insulating and no current flows across from one electrode to the other. When illuminated, photocarriers are rapidly populated, the switch becomes conducting, and current flows from one electrode to the other. Once the optical pulse has ended, the photodoped carriers relax to the valence band on the scale of the material's photocarrier lifetime, increasing the resistance and preventing current flow. All accelerating charges radiate - so some of transient current couples to transmission lines or antennas, from which it is directed to a sample. When the lifetime and laser pulse timescale are 1 ps or less, the emitted electric field pulse will have frequency components in the THz band.

PC switches used to emit THz require a laser pulse and a DC bias and output the THz transient. PC switches can also be used to measure THz pulses by taking a laser pulse and THz electric field and outputting a DC current. The optical pulse switches the PC switch to a conductor, and the instantaneous THz electric field at that time pushes carriers from one electrode to the other, which is captured and recorded as a current.

The sign of the current corresponds to the direction of the electric field in the PC switch. The allows for a direct measurement of the electric field's magnitude and phase (not power) which directly provides $\sigma(\omega) = \sigma_1(\omega) + j\omega_2$. Thus, the Kramers-Kronig relations are not required.

The name photoconductive switch describes the exact function of the device: it is a device switched from off to on by ultrafast illumination. It is therefore key to use a laser pulse that won't limit the transient bandwidth. Assuming the laser pulse is Gaussian (timescale τ_l) and the transient carrier density is a decaying exponential (timescale τ_s), the total THz bandwidth is set by $1/(\tau_l + \tau_s)$. A rule of thumb from Menlo systems is that laser pulses with Gaussian full-width, half maxima of 150 fs or less are sufficient for THz generation.

Engineering high-quality PC switches has several considerations. First: the material lifetime must be decreased from typical lifetimes of 100 ps to < 1 ps. This is usually done by intentionally introducing defects into the semiconductor, which allows photodoped carriers to quickly relax by scattering. Such defects do decrease the carrier lifetime, but they also lower the carrier mobility. This leads to a low off-state resistance and a higher on-state resistance, and decreases the transient current able to flow upon photoexcitation. The ideal PC would have carriers travelling in high mobility materials and only occasionally see high densities of lifetime-tuning defects.

There are other relevant effects in semiconductors that play a role in PC switch engineering. One is the photodember effect, wherein semiconductors with very different electron and hole mobilities develop different transient distributions for electrons and holes after photoexcitation. This produces internal fields, whose mobility is eventually limited to the ambipolar mobility. One should careful choose semiconductors and contact metallizations to avoid forming rectifying contacts to the PC switch material, which limits the measured readout current. The capacitance of such Schottky junctions can

also reduce measured transient bandwidth, and should be avoided all together.

Experimental tools to assess PC switch material performance can screen potential PC material for quality before development for usage in spectrometers. Many of these tools are not widely publicized. The most important quality test is the laser-IV curve (LIV), in which the current between two electrodes is measured in the dark and illuminated states as a function of voltage bias. The LIV curve provides a switch's on-off ratio and assesses the bias (electric field) range over which a PC switch responds linearly. A nonlinear LIV suggests that the measured current of a transient is nonlinearly related to the electric field, and that the PC switch readout cannot simply be interpreted as electric field.

Another diagnostic is the carrier lifetime, measured by optical pump, optical probe reflectometry or optical pump, THz probe spectroscopy. The optical pump photodopes the sample; the optical or THz probe monitors the subsequent carrier relaxation via the surface's reflectivity. Large lifetime materials ($>$ 1 ps) may require additional recombination tuning. In addition, one can monitor the optical spectrum for additional features (defect-driven physics, superlattice oscillations... etc.) that can reveal simple clues of the underlying physics. X-ray diffraction, atomic force microscopy, and additional electron diffraction experiments may also be used for for measuring aspects of III-V PC switch devices. The most important metric, of course, is the PC switch's performance when incorporated in a photoconductive antenna or transmission line.

The details of the nanofabrication processes used to build the following circuits are described in Appendix A.

3.3.2 PC switch testing circuit

PC switches were tested by incorporating them in a simple transmission circuit, shown in Fig. 3.9A. The structure incorporates 5 μm wide 10/100 nm thick Cr/Au CPS struc-

tures (W and S are 5 μm). Three sets of four alignment marks allow for precision device alignment and placement. PC switches are transferred to the leftmost and rightmost sets of alignment marks. A sample is then placed at the center set of alignment marks. Additional leads allow for current and voltage biases. The signal flow diagram, assuming that the leads do not scatter THz, is shown in Fig. 3.9B. Radiative propagators are shown in red, transmission line parameters are shown in orange, and scattering from both the first and second sample interfaces are in black.

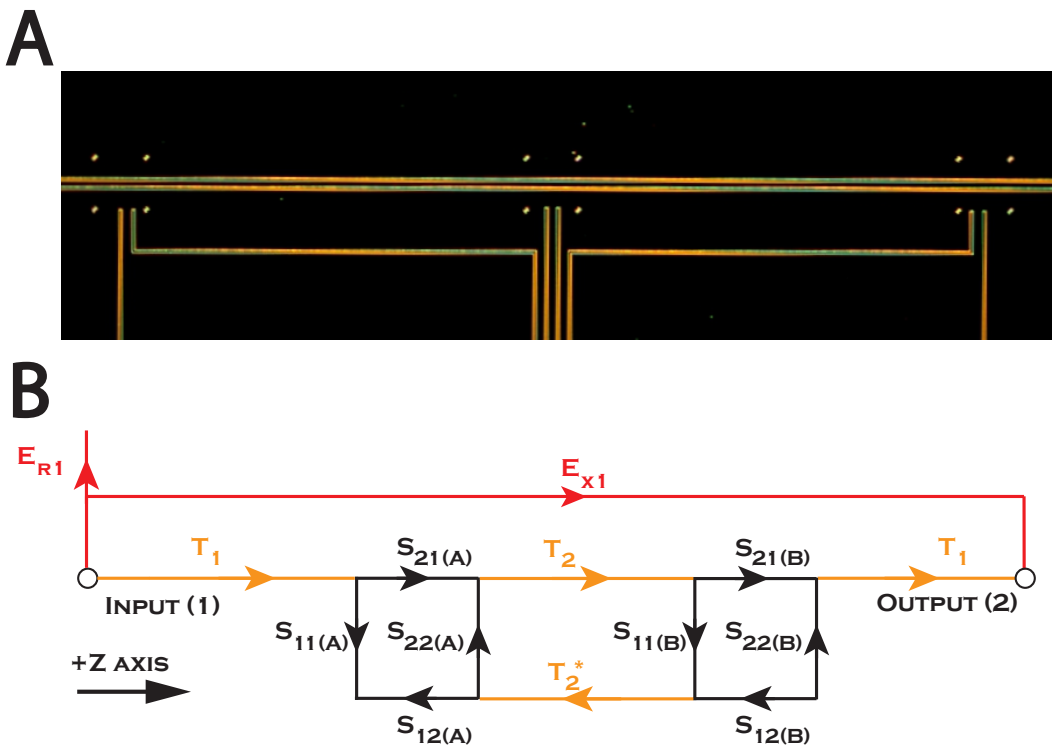

Figure 3.9: Testing circuit for PC switches. (A) Dark field optical micrograph, showing transmission lines, contacts for biasing and DC transport, and alignment markers (B) Signal flow diagram with transmission lines in orange, the sample black, and radiative coupling in red. We assume measurement in transmission.

As the electrical length of the sample decreases, the effects of the transmission line propagators T_2 and T_2^* cease. An electrically short sample, with no radiative coupling from input(1) to output(2), is transformed from the input to the output only by the

transmission line transmission and the sample transmission propagators. As a result, the analysis developed in [6, 78] accurately computes the transmission. That is: the transmitted or reflected signal is simply the series or shunt reflection or transmission coefficients from Table 1.4 times plane wave exponentials (appropriate products T_1 and T_1^*) describing the propagation along transmission lines.

3.3.3 Radiation-damaged silicon

Radiation-damaged silicon (rdSi) is popular because it is inexpensive compared to III-V PC switch material and the lithographic processing is comparatively straightforward (see appendix A). There are several forms of rdSi used as PC switches, including thermally evaporated and laser annealed [82, 118] Si and Si damaged by high-energy oxygen implantation [8]. It is on the latter we will focus.

We made PC switches with 300 nm silicon on sapphire. The silicon lifetime is lowered by implanting 100 keV O^+ ions or 185 keV O^{++} ions with density 1E15 cm^{-2}. Optical pump, THz probe spectroscopy was used to characterize the resulting carrier lifetime, which was found to be 560 fs [8]. The rdSi was etched, isolating a few hundred μm rdSi mesas. Metal transmission lines (10/100 nm Ti/Au) were evaporated to form the on-chip THz spectrometer test circuit.

The resulting LIV curves and transients are shown in Fig. 3.11. Panel A shows the LIV of switch 1 taken with 780 nm center wavelength, 10 mW, 75 fs optical pulses, with the corresponding transient in panel B. The achieved LIV currents are smaller than a nanoamp, and the transient itself is smaller than 100 fA: barely visible above the noise floor. The transient's 10 ps time scale suggests sub-100 GHz bandwidths. To achieve higher fluences, we switched to a laser system operating at UC San Diego with 200 kHz repetition rate, 800 nm center wavelength, and 25 fs pulse duration. For the same time-

averaged power, this system provides about 1000 times larger pulses due to the lower repetition rate. The LIV curves of switches 1 and 2 in the on- and off-states are shown in panel C, with the resulting transient in panel D. The on state LIV and THz signal have substantially larger amplitudes with the higher fluence optics. The Fourier transforms of the transients generated by the 250 MHz laser (purple) and 200 kHz laser (black) are shown in the inset of panel D, emphasizing the huge increase in signal quality by using a higher fluence laser. The usable[4] bandwidth increases from 79 GHz to nearly 220 GHz simply by switching from the 250 MHz laser to the 200 kHz laser.

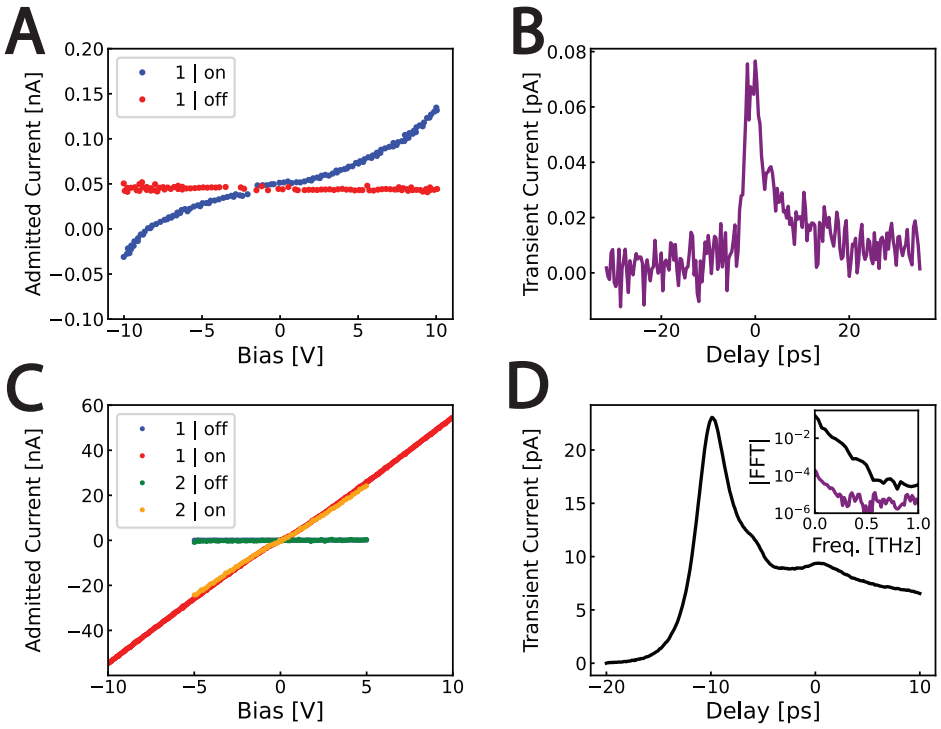

Figure 3.10: rdSi PC switch testing. (A) LIVs and (B) transient with 250 MHz laser, and (C) LIVs and (D) transient with the 200 kHz laser with (inset) magnitude of the Fourier transforms of the 250 MHz laser transient (purple) and 200 kHz laser transient (black).

The rdSi transient testing circuit was tested in measurement with an exfoliated flake

[4]defined as the lowest frequency at which the standard deviation of at least 5 identical scans becomes greater than 5%

of NbSe2, a vdW superconductor with transition temperature of 6.5K. Superconductivity is an ideal condensed matter system for testing because of the theoretically calculable response function, the clear difference between the real and imaginary parts of the optical conductivity, low-frequency spectral weight conservation (Ch. 2). The exfoliated few-layer flake was placed on one of the CPS conductors (Fig. 3.11). The 200 kHz, 25 fs, 800 nm wavelength laser from UC San Diego was used to excite THz transients. THz transients were sourced by positioning the laser over the left PC switch (emitter) between the bias line and outermost CPS conductor. The readout PC switch was excited in the space between the CPS conductors. The resulting THz transients in the time and frequency domains are shown in Fig. 3.11B and C. The analysis equations of [6, 8] apply, allowing us to extract real and imaginary optical admittance Y_1 and Y_2 (Fig. 3.11D and E).

There is a clear change in the transient as a function of temperature. As the sample is cooled below the superconducting transition temperature, the transmitted signal amplitude increases slightly. One can see, at the lowest temperatures, the low frequency spectral weight in the real part is transferred to the imaginary part. This is coarsely suggestive of the Zimmerman model introduced in Ch. 2, but bears little quantitative similarity to the model.

There are several other limitation in intrepretting the data. The contrast in signal above and below the superconducting transition temperature is quantitatively small. The usable bandwidth is only 219 GHz. This is too small to directly observe the superconducting gap ($2\Delta(0) = 520$ GHz) of $NbSe_2$. In addition, there is no simple way to estimate any drift in emitted transient amplitude and phase versus temperature after hours of measurement. It is possible that the change in signal occurred due to electronic drift, drift in the optical path length, temperature-dependent emission bandwidth... etc. The PC switches are so large that a THz transient could be emitted by optical incidence

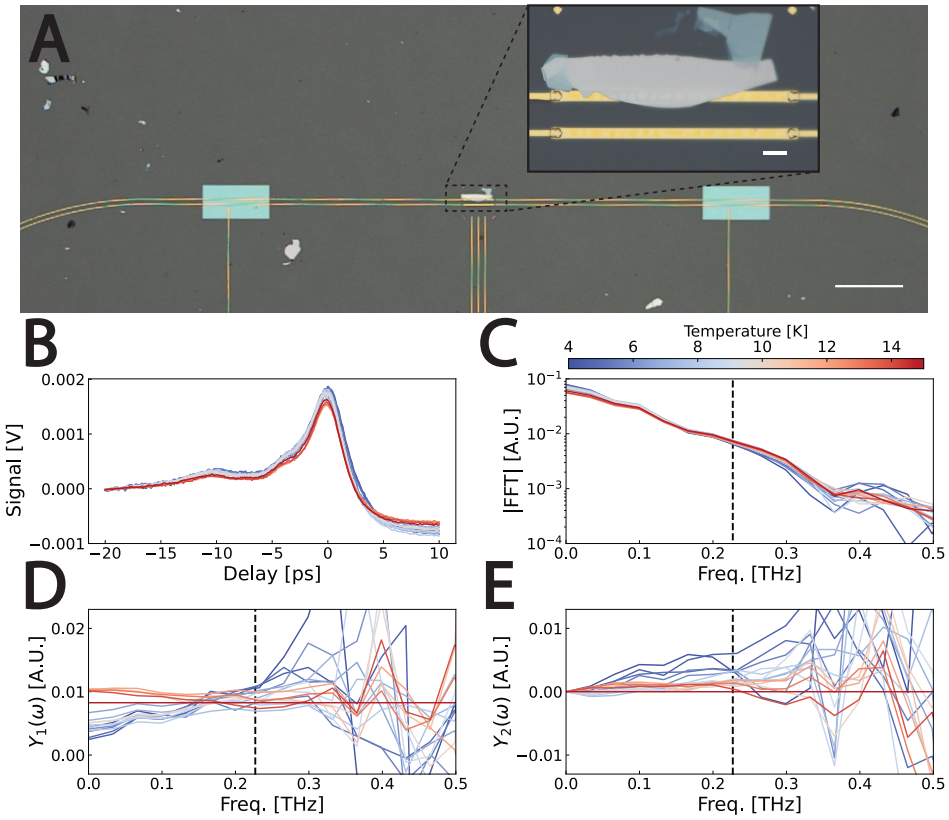

Figure 3.11: rdSi circuit with flakes of superconducting $NbSe_2$. (A) Optical micrograph (scale bar 400 μm) with inset (scale bar 5 μm) zoomed into $NbSe_2$. (B) Time and (C) frequency domain terahertz signals with different colors indicating different temperatures. The usable bandwidth (219 GHz) is demarcated by a vertical line on all frequency domain plots. The spectra were converted to (D) real and (E) imaginary admittance. There is contrast between the normal and superconducting states, but it is too weak to resemble the Zimmerman model of Ch. 2.

along a few hundred micron path along the PC switch. This is a limitation of the testing circuit, which does not provide an in-situ calibration of the emitted pulse. This physical distance could cause transients to be time delayed by over a picosecond, producing apparent drift in the signal phase. It is also likely that the minimal contrast between normal metal and superconductor stems from the poor impedance mismatch between the sample and the transmission line introduced in Ch. 1.

In summary, rdSi PC switches produce an insufficiently small bandwidth and suffer

from drift that cannot be removed in post-processing, despite having relatively simple nanofabrication processes and probing a material system that should have a well-understood response.

3.3.4 LT GaAs

Another candidate PC switch material system is low temperature-grown GaAs (LT GaAs). LT GaAs is grown like normal GaAs, except that LT GaAs is grown at temperatures near 250°C (pristine GaAs is grown at temperatures of 850°C or higher) in a <1% As overpressure. The As overpressure favors the formation of lifetime-tuning EL2 Arsenic anti-site defects [119, 120, 121], and the low temperature prevents the defects from being thermally annealed. The resulting LT GaAs's lifetime is lowered by the anti-site defects, whose density scales with the As overpressure. A brief post-growth anneal in forming gas substantially increases the dark resistance and the on-off ratio.

LT GaAs produces larger THz transients than rdSi because the mobility of carriers in LT GaAs exceeds that of rdSi. It is, however, about ten times more expensive and can only be purchased from speciality suppliers. In addition, LT GaAs cannot be grown on low-loss substrates. LT GaAs is usually prepared on ready-made unintentionally doped (UID) GaAs. UID GaAs wafers are cleaned and placed in a molecular beam epitaxial growth chamber. A thin insulating layer of stopetch material is grown (typically 300 nm $Al_xGa_{1-x}As$), followed by $\sim$ 1 μm of LT GaAs. The stopetch serves two benefits. Its chemistry enables a selective wet etch through the LT GaAs but not the stopetch layer, providing a reliable way to isolate LT GaAs mesas. It is also chosen to be insulating to incident laser photons, to prevent transient carriers from leaking from the active LT GaAs into the UID GaAs substrate.

All semiconductor bandgaps are weakly temperature dependent, per the Varshni equa-

tion [122], which describes the GaAs bandgap E_g versus temperature (T) as:

$$E_g(T) = 1.519 - \frac{(5.408E - 4)T^2}{T + 204} \tag{3.14}$$

PC switches operate best when every incident photon is capable of promoting electrons from the valence to conduction band. Incident photons that have sub-bandgap energy The zero temperature bandgap is equivalent to 816 nm wavelength light, whereas the room temperature bandgap corresponds to 872 nm light. Using an 800 nm optical pulse with a 50 nm full-width, half maximum would, for example, operate at full efficiency at room temperature but have decreased excitation efficiency at cryogenic temperatures. I recommend measuring the spectrum (or estimating it from datasheets / use an optical autocorrelator and computing the Fourier transform limit - see Appendix A).

We must characterize the photocarrier lifetime to ensure THz-frequency bandwidths are possible. We began by employing optical pump, reflection optical probe spectroscopy (OPOP) and optical pump, THz transmission probe spectroscopy (OPTP) on pieces larger than 5 mm x 5mm. In OPOP, the sample is excited with (strong) and measured by (weak, time-delayed) 25 fs near-infrared (NIR) pulses. The photoexcited carriers transiently display metallic properties, which reflect light until the photocarriers decay, per the Drude model (Ch. 2). This allows for direct extraction of the carrier lifetime. Since OPOP measures transient change in reflection ($\Delta R/R$ or ΔR) and the pump and probe have the same skin depth, OPOP probes the photoexcited surface carrier density at the surface only. In OPTP, the sample is excited with an NIR pulse at zero delay and measures the sample by free-space quasi single cycle THz that reflects from or passes through the sample. The recorded electric field is that of the THz pulse's time domain peak following photoexcitation. The contrast mechanism is reflection by transient photocarriers, most typically those which photoexcited at the sample's surface

due to the NIR pump pulse's shallow skin depth ($\sim$ 100 nm).

The shape of the OPOP or OPTP decay can be fit to one of two models. The first is a simple exponential decay with delay τ_0 and decay time τ_1 of the form:

$$\Delta R/R = A_1 e^{-(t-\tau_0)/\tau_1} + C \tag{3.15}$$

Fitting to a simple decaying exponential is complicated by the fact that the decaying exponential is the impulse response, which assumes an input of $\delta(t)$. The input is, however, a Gaussian optical pulse with finite temporal width. As a result, it is up to the spectroscopist to chose which points should be excluded from the fit, leading to some subjectivity. This can be remedied by fitting to the more complicated form, which we term the "Kissin model", which is derived by convolving a Gaussian input with width τ_2 with eqn. 3.15.

$$\Delta R/R = A_1 \left(\frac{1}{2} + \frac{1}{2} Erf\left(\frac{\sqrt{2}(t-\tau_0)}{\tau_2} \right) \right) e^{(t-\tau_0)/\tau_1} + C \tag{3.16}$$

Our NIR pulses were provided by the 200 kHz repetition rate laser from UC San Diego. THz probe pulses for OPTP were generated by optical rectification in birefringent ZnTe [123], with usable bandwidth from 0.3 to 2.2 THz. The results of OPTP and OPOP measurement, with fits to the Kissin model, are shown in Fig. 3.12A-B respectively.

OPTP taken with photoexcitation of the front (LT GaAs) or back (UID GaAs) reveals a photocarrier with a lifetime of about 1 ns. This is far too long for incorporation as a PC switch for THz TDS experiments. The long times can be explained because the probe operates in transmission - meaning that it passes through the UID GaAs. UID GaAs carrier Drude response (Ch. 2) overwhelms the transient reflectivity from the active layer. OPTP taken in reflection mode from the active surface may reveal the LT GaAs carrier dynamics only. Sadly, that data is not available.

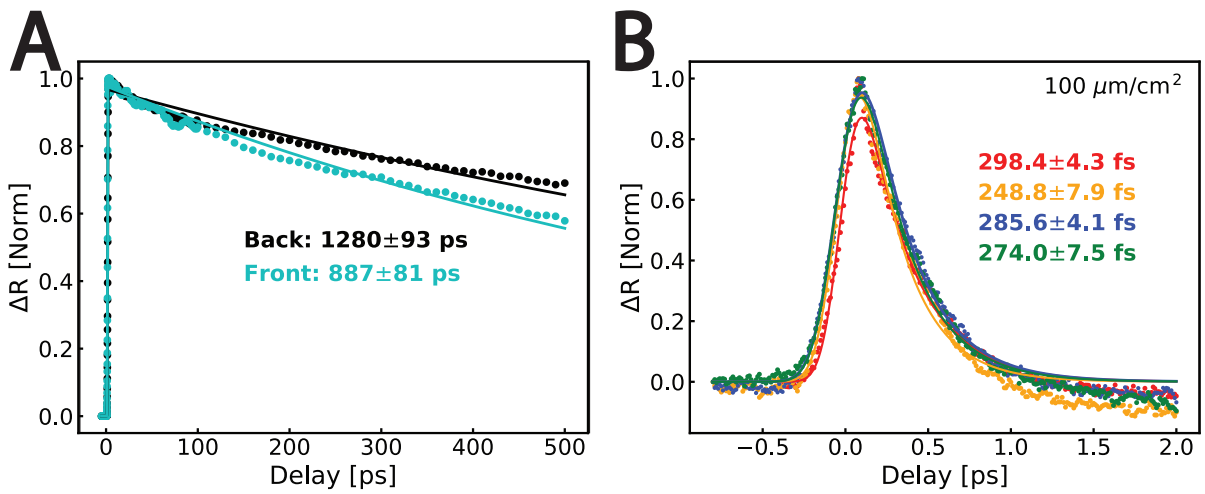

Figure 3.12: (A) OPTP and (B) OPOP measurement of photocarrier lifetime in LT GaAs, with fits to the Kissin model (eqn. 3.16). OPTP measurements were taken on the active and back of the LT GaAs wafer; OPOP measurements were taken on four separate pieces of the active side to check for consistency.

OPOP, however, shows a response from only the active surface because it is taken in reflection (and the electric fields do not pass through the substrate). From OPOP, we identify a lifetime between 250 and 300 fs, which is sufficiently short to produce THz-bandwidth transients. Switches could be fabricated in one of three ways. One could evaporate a transmission line directly on the active layer ("sliding contact"). One could also perform mesa isolation: where the active layer is removed except where PC switches will be, electrically isolating PC switches from one another (as was done for rdSi). One could perform epitaxial lift off (ELO) and transfer patterned PC switches from the growth substrate to externally fabricated spectrometers.

Sliding contact spectrometers were made by evaporating a Ti/Au test spectrometer on LT GaAs. LIVs with 15 mW laser power (780 nm, 75 fs, 250 MHz) were taken, showing an on-off ratio of about 1E3. The excitation beam was mechanically chopped at 1 kHz, allowing the use of a lock-in amplifier to record the THz transient. Instead of finding a sharp, THz-bandwidth transient, we measured a 10 nA background independent of delay. Adding a high pass filter (100 Hz cutoff) didn't change the background, suggesting that

the signal was not an artifact from measurement electronics.

This large background could come from one of two sources: non-trivial photocarrier diffusion length or coupling to carriers across the stopetch layer. That is: photoexcitation of UID GaAs carriers through the stopetch layer could capacitively couple to carriers in the LT GaAs layer. Photocarriers excited by a tightly-focused laser spot could diffuse radially outwards with an average length (l_d) given by:

$$l_d = \mu V \tau_1 / W \tag{3.17}$$

Here, μ is the carrier mobility[5], V is the applied bias, τ_1 is the photocarrier lifetime from Kissin model fitting (Fig. 3.12), and W is the CPS gap spacing from Ch. 1. Using mobilities of 8500 cm^2/Vs (UID GaAs) and 300 cm^2/Vs (LT GaAs) [124], 30V bias and a 10 μm gap, l_d in LT GaAs is 100 nm; l_d in UID GaAs is 3 mm. It is, therefore, possible for carriers to diffuse from one laser spot to another laser spot within the laser's finite repetition time (4 ns). This is an upper bound (as the 30V electric field is the maximum maintained between the conductors, but not radially outward), but nevertheless illustrates that coupling between long-lived UID GaAs carriers and short-lived LT GaAs carriers could cause this high background. We overcame this challenge by purchasing a new[6] LT GaAs wafer (with stopetch $Al_{0.9}Ga_{0.1}As$, to increase the stopetch bandgap and avoid resistively coupling carriers from active to UID GaAs layers) and performed mesa isolation - which removes excess LT GaAs near the PC switches. Mesa isolation ensures that carriers from adjacent PC switches cannot couple to one another. The results of mesa isolation and a new wafer on a spectrometer are shown in Fig. 3.13.

LIVs show an on-off ratio between 1E2 and 1E3 and a broad transient at biases near

[5]This should be the ambipolar mobility, but the carrier density is unknown, so we use the electron mobility.

[6]Commercial supplier previously mistakenly grew $Al_{0.75}Ga_{0.25}As$, which has a bandgap not insulating to 780 nm photons

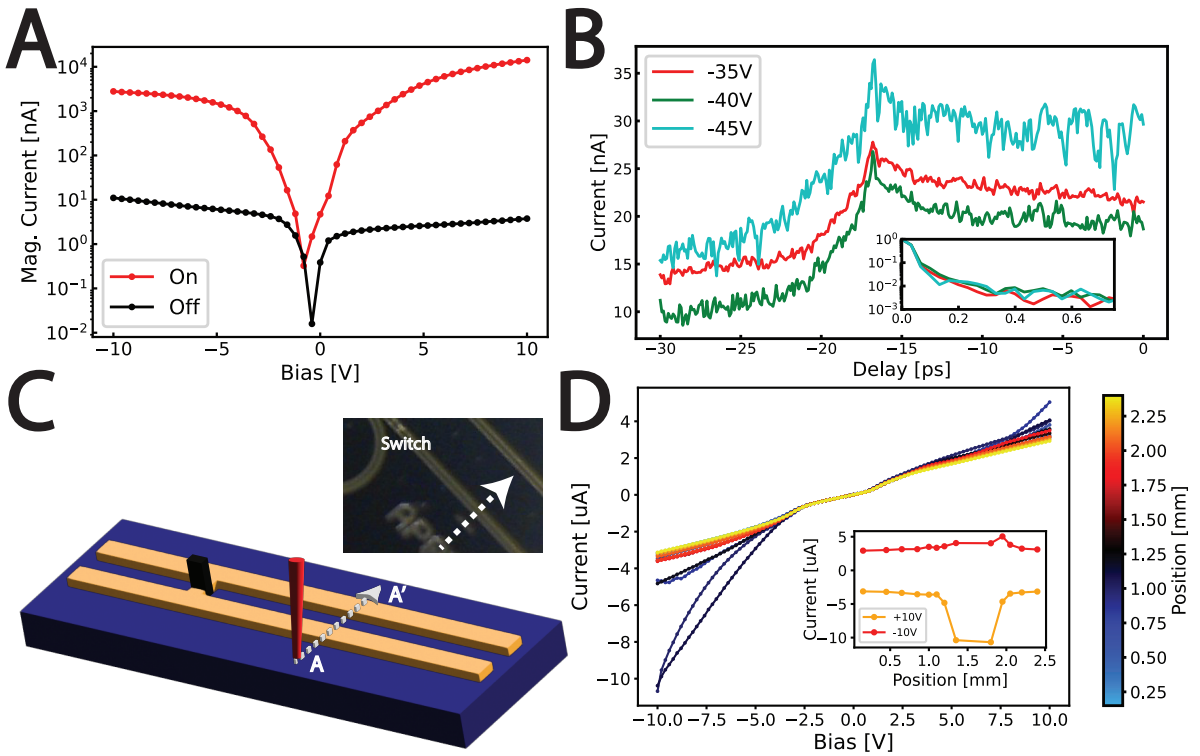

Figure 3.13: LT GaAs spectrometer with mesa isolation. (A) LIVs and (B) transients taken at large negative bias with (inset) corresponding Fourier transforms magnitudes. (C) Cartoon and image, showing focused laser beam on AlGaAs (purple) near PC switch (black) with transmission lines (gold). The laser spot is swept from A to A', producing (D) LIVs versus position along AA'. The inset shows the current at ±10V versus position.

the breakdown voltage (-47V). The signal increases with increasing bias magnitude, and has a usable bandwidth of 64 GHz. The backgrounds, however, are very large. Mesa isolation rules out direct travel from PC switch to PC switch, suggesting that the large backgrounds and poor bandwidths come from coupling through the substrate. To confirm this, the laser was positioned 2 mm from the PC switch (black) and swept about 2 mm from position A to A' (Fig. 3.13C). An LIV was measured at each position, with the resulting current at ±10V shown in the inset of Fig. 3.13D. Here, the transmission line is grazed by the broadly-focused laser spot only at 1.4 - 1.8 mm. The large change in reported current suggests that the photodoped AlGaAs is conducting current between

the two metal conductors. The large background is thus likely due to the long-lived AlGaAs photocarriers. Though intrinsic $Al_{0.9}Ga_{0.1}As$ should be insulating to 708 nm photons, residual and unintentional doping likely adds midgap states and reduces the photon energy required for problematic photodoping. Though the laser could, in principle, be aligned to the LT GaAs mesa only, the LT GaAs mesa height is only about 2 skin depths. Some radiation would inevitably reach the AlGaAs, and produce large background signals. This could be solved by ELO: a process which we will explore for ErAs:GaAs superlattice PC switches.

In summary, LT GaAs switches have been limited by supplier errors and large background signals in sliding contact samples. I firmly believe LT GaAs spectrometers made with high quality materials and non-sliding contact designs will have THz bandwidth signals, as seen in [6, 78].

3.3.5 ErAs:GaAs Superlattices

ErAs:GaAs superlattices (abbrev. ErAs) are PC switches made by epitaxially growing pristine GaAs with interspersed planes of ErAs quantum dots. These quantum dots are distributed randomly with a large average density in-plane. The interspersed ErAs planes form a c-axis superlattice. In the dark state, no carriers propagate through the pristine GaAs. Upon illumination, photocarriers are excited at the surface and diffuse along the c-axis away from the surface. These carriers mostly travel in pristine, high mobility GaAs. Some fraction of the photocarrier current recombines upon interaction with each plane of quantum dots. Decreasing the superlattice period decreases the average time to recombination, but it also decreases the signal size (as carriers spend less time travelling in pristine GaAs). Increasing the superlattice period size increases the signal size (more time spent in pristine GaAs), but decreases the transient's bandwidth [125, 126, 127,

128, 129, 130].

ErAs was epitaxially grown on 300 nm $Al_{0.75}Ga_{0.25}As$ on 500 μm UID GaAs. ErAs was grown with a 25 nm superlattice period. Growth was monitored by resonant, high energy electron diffraction (RHEED), confirming the nucleation of individual atomic layers. Samples (individually identified by labels of the form: TMC###) were grown with 500 nm to 2 μm thick active layers. Individual wafers could be screened by melting Indium on the surface for contact via probe station. Epitaxial liftoff (ELO, see appendix A) was then performed (details in Appendix A), creating 50 μm to 100 μm by 25 μm mesas.

The test spectrometers were fabricated with Ti/Au metallization on a 625 μm quartz substrate. Mesas were individually transferred with dry polymer transfer [131] to test spectrometers, wherein their LIVs could be extracted. Two example optical micrographs are shown in Fig. 3.14A/B. The planes of ErAs quantum dots cause significant strain, which can cause switches to crack upon transfer. Transferred switches are often broken (A) but can transferred with enough attempts (B). The LIVs of individual wafers can have wildly different behaviors. TMC417, for example, has a high-slope linear in V current at low bias. This suggests that small electric fields with zero mean value are capable of producing large currents. Other growths, however, do not show large slopes. Other growths not only have much lower currents and conductances, they have different shapes, suggesting each growth is essentially a different underlying material system[7].

We explore the material nature of these wafers by performing X-ray diffraction (XRD). The (004) peak contains a main peak at $2\theta \approx 66°$, associated with pristine GaAs's lattice constant. Additional superlattice peaks result from Fabry-Pérot reflections between the ErAs layers. The superlattice peak spacing gives the superlattice period per growth,

[7]Note that a detailed explanation of the transport shown here is beyond the scope of this dissertation, and is not well-understood in the literature.

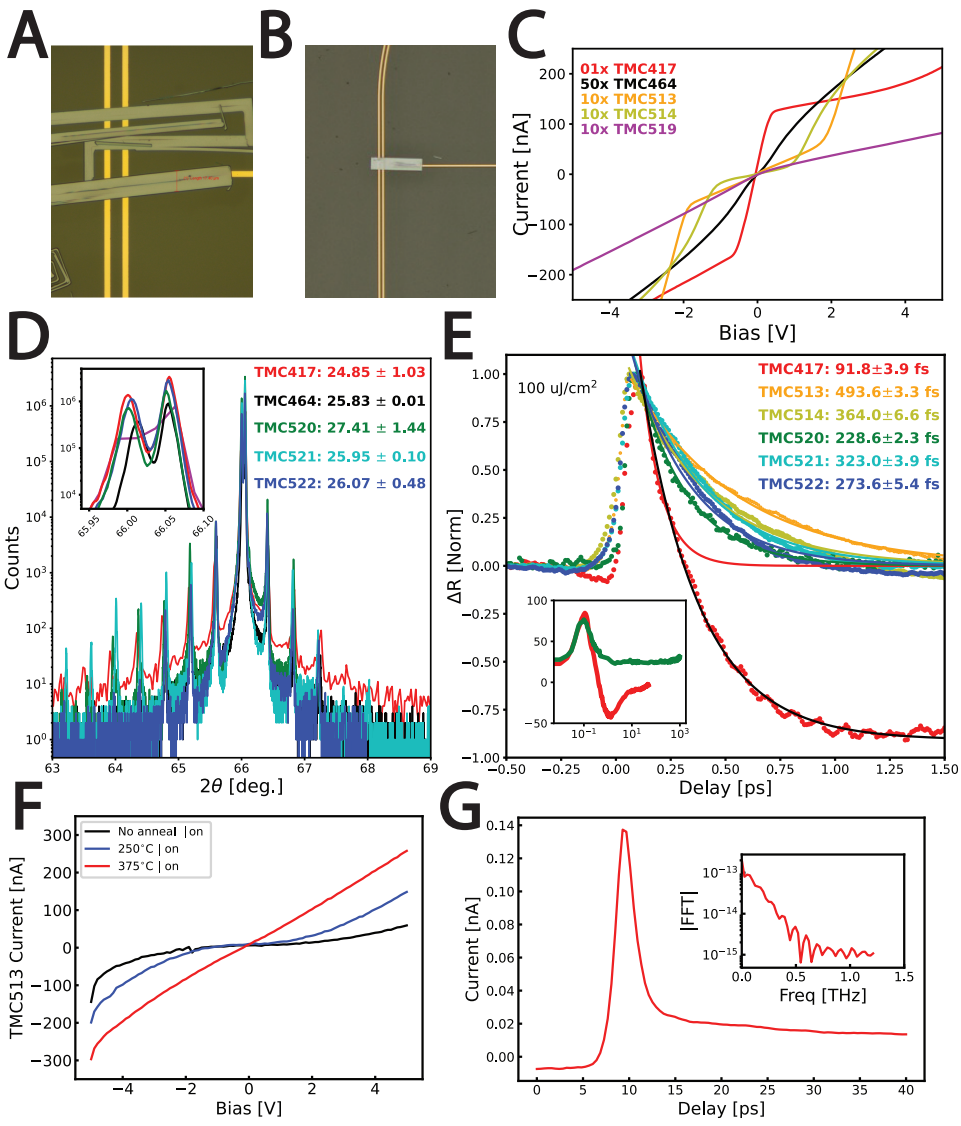

Figure 3.14: ErAs superlattice PC switches. Note that colors representing specific growths are consistent across all plots. (A) Cracked and (B) non-cracked PC switches transferred to a test spectrometer on 625 μm thick quartz. (C) LIV curves from a variety of wafers (note the multipliers for some wafers). (D) ErAs X-ray diffraction, (004) peak, with (inset) showing a zoom of the central peak and per-sample labels of the superlattice spacing. (E) OPOP of several ErAs wafers with (inset) showing long delay behavior (F) Effects of variable temperature annealing on Ohmic contact of ErAs mounted on a probe station with a bright illuminating light. (G) THz transient of ELO PC switches on quartz.

indicated on Fig. 3.14D, which are all consistent with the target 25 nm period. The peak width, shown in inset, is broadened by strain. High ErAs quantum dot densities

increase the broadening - so broader peaks suggest better THz performance. TMC417 has the most broad (004) XRD peak and has the most desirable LIV; TMC464 is the least broad and has the least desirable LIV.

Sample structure can be further probed by OPOP (Fig. 3.14E). For simplicity, curves were fit by simple decaying exponentials (eqn. 3.15), not the Kissin model. Most samples show a simple decay structure with photocarrier lifetimes between 228 and 364 fs. The exception is TMC417, whose transient reflectance change drops below zero after a 92 fs decay. The transient reflectance contains visible 89 fs oscillations. The signal recovers to zero after about 100 ps. This is likely caused by superradiant damping - a cooperative phenomenon cause by nearest neighbor coupling between adjacent ErAs [132]. This only occurs when the ErAs has nucleated quantum dots, as opposed to simply doping the local GaAs. Together, the superradiant oscillations in OPOP and broadening in XRD suggest that TMC417 has the requisite quantum dot structure, whereas all other wafers are merely Er-doped, As-rich GaAs.

Only devices made with TMC417 will be capable of producing THz transients - no other material was found capable of producing THz-bandwidth signals. A test spectrometer on 625 μm quartz was made from ELO PC switches from TMC417. The spectrometer was annealed to improve Ohmic contact. The improvement with simple 1 hour anneals at 375°C in vacuum can be seen in Fig. 3.14F under a fixed, not-focused white light source (no laser required). The resulting THz transient (Fig. 3.14G) and Fourier transform show a signal with spectral weight at frequencies less than 289 GHz. Though this is an improvement from rdSi switches, it is far from providing true THz-bandwidth signals. In addition, the difficulty in consistently producing ErAs quantum dots, as opposed to Er-doped GaAs, suggest this not to be reliable way to consistently fabricate THz spectrometers.

We had great difficulty ensuring that high quality LT GaA and ErAs was provided,

even by academic and well-reputed commercial suppliers, in part because standardized quality control protocols for these materials *do not exist*. If you are reading this with intent to build a THz setup of your own, I highly advise that you: (1) never assume that the provided material is of high quality without provided data, (2) never assume future samples will work simply because past samples worked, (3) request data sheets and growth logs with every growth, so you can diagnose simple mistakes in layer chemistry, (4) mark the underside of the substrate with a small diamond scribe mark, so you can identify it in processing, and (5) use pieces from the wafer center to make your spectrometers, as the temperature during growth is monitored there and therefore their performance better matches the data in the growth logs.

In summary, ErAs switches are limited fundamentally by inconsistent materials growth and require substantially more research and development.

3.3.6 2D Dirac materials

The thermallization of excited photocarriers in prototypical Dirac semimetal graphene has been shown to follow a time-dependent Fermi temperature after the initial ∼100 fs process. This cooling is limited by disorder via supercollisions, a direct result of graphene's famous Dirac cone bandstructure. Electron thermalization requires electrons to scatter from point to point in the Fermi surface. Graphene's small, circular Fermi surface at non-zero carrier density has very few available allowed momenta transfer from one point on the circle to the others. Scattering from the large available phonon space (at high temperature) or from disorder (at low temperature) in spite of the ultralow disorder can limit the thermalization time [133, 134]. These times, which can be one to several tens of picoseconds, can be exploited to employ Dirac semimetals as PC switches [135, 136].

To demonstrate graphene's use as a PC switch, we performed on-chip OPTP. We fabricated a single-conductor Goubau transmission lines that link two rdSi PC switches, with a platinum-gated, hexagonal Boron Nitride-encapsulated monolayer graphene linking the left and right sides of the circuit. We illuminate each PC switch with the 200 kHz repetition rate laser from UC San Diego. An additional beam is directed to the monolayer graphene heterostructure, which optically excites the graphene. This second beam is delayed and mechanically chopped independently of the PC switch beams. The log of the measured current as a function of both the graphene and PC switch delay time, as well as a linecut at sampling delay of -25 ps and the Fourier transform thereof, are shown in Fig. 3.15A-C respectively.

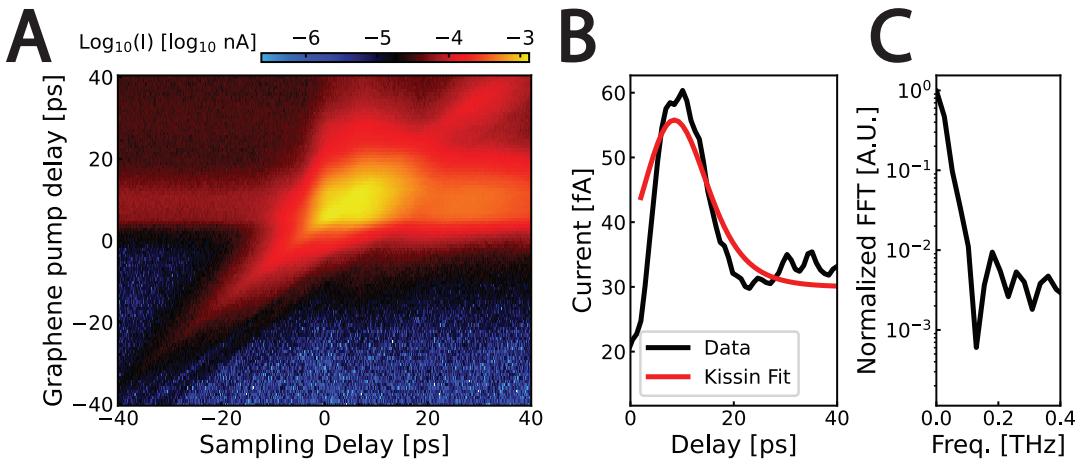

Figure 3.15: Photocurrent showing graphene operation as a PC switch. (A) $\log_{10}$ of the magnitude of measured photocurrent versus the PC switch ("sampling") delay and graphene delay (B) A linecut taken with the sampling delay set to -25 ps, with an indicated Kissin-style fit (C) Fourier transform of the data from panel B.

The measured signal for most negative graphene pump delays is at the spectrometer's noise floor. At all sampling delays and graphene pump delays greater than zero, a signal is measured corresponding to THz emission by the graphene stack. To isolate the graphene-only component, we take a linecut of the measured electric field before the PC switch transient (feature at nearly zero sampling delay and zero graphene pump delay). The

linecut shown in Fig. 3.15B, with a Kissin-style fit (eqn. 3.16). The resulting fit has time scale 6.12 ± 3.61 ps. The Fourier transform (Fig. 3.15C) shows a bandwidth less than 100 GHz. The useable bandwidth may be estimated from several vertical linecuts - which is 81 GHz. This is far too small to make a true THz device.

New geometries and better engineering may, however, lead to higher bandwidth spectroscopy. Other groups have shown graphene to be an effective PC switch [136, 137] and that graphene can be engineered to operate at several hundred GHz [135]. Using 2D materials as PC switches is a very exciting research direction, whose exploration could one day lead to an all vdW spectrometer.

3.3.7 Professionally fabricated LT GaAs

The primary failures of the previous PC switch materials are: rdsi - low bandwidth, LT GaAs - large backgrounds associated with poor supply and substrate limitations, ErAs superlattices - inconsistent growths and graphene - low bandwidth. It is obvious that another source of PC switch material is required. We will show in Ch. 4 that this is a professional supplier - Protemics GmbH. LT GaAs supplied by Protemics will undergo ELO and be transferred to low loss, thin, low permittivity substrates to avoid severe Cherenkov radiation and to avoid unbound modes.

3.3.8 Active THz engineering summary

Table 3.3 summarizes the primary metrics achieved in PC switch design in this chapter. Note that this excludes professionally fabricated LT GaAs PC switches, which we will explore in the subsequent chapter. None of the following PC switches are currently viable for incorporation into THz-bandwidth spectrometers.

Metric	Unit	rdSi	LT GaAs	ErAs:GaAs	Graphene
Cost	n/a	$	$$	$$$	$
Reliability	n/a	High	Moderate	Very poor	Moderate
Mesa isolation	n/a	Easy	Hard	Hard	Easy
Achieved lifetime	fs	560	248 ± 7.9	91.8 ± 3.2	6120 ± 3610
Achieved bandwidth	GHz	219	25	289	81
Recommended DC Bias	V	10	10	0.1	—
Achieved on/off	Unitless	552	3767	3890	—
Achieved Dynamic range	n/a	1243	2.3	71.3	161
Achieved peak current	pA	23.2	3.5E4	137	0.060

Table 3.3: Summary of PC switch material systems. Entries marked with "—" are not available. LT GaAs and ErAs figures assume 780 nm, 75 fs Gaussian laser pulses; rdSi and graphene figures assume 800 nm, 25 fs Gaussian laser pulses. We linearly extrapolate data to obtain the rdSi off current at 9V. The graphene data assumes zero gate voltage and no applied bias, due to the geometry. The dynamic range quantifies the largest frequency domain signal to the average of the noise floor at measured frequencies greater than 2.0 THz.

3.4 Use cases for THz spectroscopies

On-chip THz TDS is superior to free space THz TDS if at least one of the following is met:

1. The sample is electrically small and/or is thinner than several skin depths.
2. The sample is a device and has electrostatic gates.
3. The sample is not uniform for the entire beam waist (further exacerbated if low-frequency information is critical).
4. The sample is being interrogated with optical pump, THz probe spectroscopy and the optical pump/THz probe have mismatched skin depths.

On-chip THz TDS is inferior to free space THz TDS if at least one of the following is met:

1. High frequency information (>1.0 THz, with current state of the art) is desired.

2. One wishes to use single-shot THz spectroscopy based on a spatial distribution of pulses [138, 139, 140, 141, 142].

3. Spectroscopic information is needed at temperatures which exceed the glass transition temperature of substrates. For the spectrometer that will be presented shortly, this transition occurs at T> 150 °C.

A spectroscopist's most effective tools in PC switch material quality control are:

1. DC transport (I-V curves) in the dark and under illumination.

2. Optical pump, optical probe spectroscopy or optical pump, THz reflection probe. Optical pump, THz transmission probe can contain irrelevant information from substrate carriers.

3. RHEED oscillations during epitaxial growth.

4. X-ray diffraction.

5. Atomic force microscopy of the surface, though surface reconstruction in III-Vs [143] can make this difficult to interpret.

6. Rutherford backscattering analysis (RBS) or particle-induced X-ray scattering (PIXE). They are better suited to identifying the atoms of homogeneous thin films. We used both techniques to screen for accidental Fe doping, not identify the formation of ErAs quantum dots. The absorption signatures of ErAs quantum dots as compared to Er-doped GaAs by RBS and PIXE measurements are not known at the time of writing.

7. THz test circuits with and without mesa isolation.

8. Wet etches sensitive to epi-layer chemistry. This allows you to verify correct chemistry. See appendix A for details on citric acid / hydrogen peroxide etches.

THz engineering differs substantially from microwave and optical frequency engineering. In addition to difficulties in passive element engineering, THz is more difficult to generate and detect. We summarize PC switch engineering considerations in Table 3.3. If you are reading this with the intent to design or improve free space THz path links, there can be additional free space propagation losses. Such losses are modelled by Friis formula, the Lee model, and Hata models for signal fading. The loss in these models does becomes more severe at THz frequencies than at microwave frequencies. For further reading, see [144].

Type	Free space	On-chip
Minimum sample size	Diameter of 1.22λ	Nanofabrication-limited
Maximum sample size	No	Yes
Low-frequency limits	Rayleigh limit	Fabry-Pérot reflections
High-frequency limit	Switch-limited	Transmission-limited
Can measure gated devices?	No	Yes

Table 3.4: Summary of differences and uses cases for free space and on-chip THz spectroscopy. Note that low frequency measurement free space THz TDS demands sample uniformity over the entire Rayleigh-limited area. In on-chip THz TDS, non-uniform sample regions can be removed in nanofabrication without adversely affecting spectroscopy.

Chapter 4

On-chip Cryogenic, Fast Sample Interchange Architecture

"Thoroughly conscious ignorance is the prelude to every real advance in science."

J. C. Maxwell

The contents of this section are adapted from ACS Nano Lett. 2023, 23, 9, 3835–3841: [9].

4.1 On-chip time-domain terahertz spectroscopy of superconducting films below the diffraction limit

THz TDS can be used to extract the complex optical conductivity of quantum materials without Kramers-Kronig relations. The few-meV energy scale and the picosecond time scale of a typical THz pulse makes THz spectroscopy well suited to experiments in many fields of research, particularly strongly correlated phenomena [62, 145, 146, 147], superconductivity [148, 44, 149, 43] and non-equilibrium dynamics[44, 15]. However,

free space THz TDS, which has been used extensively to study bulk materials and thin films, cannot be used directly to study correlated phenomena in vdW heterostructures, an increasingly active area for probing these same phenomena [150, 151, 152]. Free space THz TDS inherently suffers from the inability to focus THz beams on samples smaller than the Rayleigh diffraction limit of 488 μm at 0.75 THz. vdW heterostructures, often 10 μm or smaller, have a reduced cross-section with free-space THz beams, leading to a drastically diminished signal-to-noise ratio that precludes most experiments. Though THz near-field imaging techniques [153, 154, 155, 156, 157, 158] can achieve extreme sub-wavelength resolution, they are often difficult to integrate into cryostats, have low excitation efficiency if standard far-field excitation techniques are used, and are limited to narrow bands of THz frequencies.

The diffraction limit can also be circumvented by confining THz pulses within transmission lines patterned on a chip. Early on-chip THz experiments incorporated photoconductive (PC) switches into transmission line circuits to emit and detect THz radiation[118, 159, 106]. Recent on-chip THz efforts [160, 161, 162, 163, 164, 165, 107, 166, 135, 82] have explored new geometries, leading to increases in the THz amplitude and spectral bandwidth. On-chip THz TDS has been recently used to study electronic phases at cryogenic temperatures in carbon nanotubes [167], topological insulators [147], and two-dimensional electronic systems such as GaAs [168] and graphene [79, 6, 8, 136, 169]. Most of the existing designs for on-chip THz spectroscopy require monolithic fabrication of the THz emitter/receiver and the vdW device. The combined complexity of fabricating both the THz spectrometer and a vdW device to perform a single experiment, however, reduces measurement throughput.

Higher throughput may be achieved by disentangling the fabrication of the THz spectrometer (switch board) and the device (on the sample board) as introduced recently by Lee *et al.* [170] at room temperature. Here, we demonstrate cryogenic operation of

an on-chip THz spectrometer with a fast sample exchange architecture, and use it to extract the complex optical conductivity of a lithographically-defined NbN device below its superconducting transition temperature. Our experimental geometry is shown in Fig. 4.1, and consists of a 'switch board' containing transmission lines and photoconductive switches for both THz generation and readout and an interchangeable 'sample board' containing a sample of interest, transmission line segments, and additional deposited electrical contacts.

The switch board is a permanent fixture containing LT GaAs PC switches coupled to lithographically defined coplanar stripline transmission lines (Fig. 4.1A). The transmission lines have a conductor width of 50 μm, a separation of 30 μm, and are terminated by photonic bandgap filters that reduce THz reflections from the transmission line terminations at long times. The switch board is fabricated on a 50 μm thick cyclo-olefin polymer (COP) substrate. High transparency, low permeability and a nearly temperature and frequency-independent optical index of $\approx$1.52 at THz frequencies ensures minimal pulse broadening [171, 172, 173]. The thin substrate and comparatively small dielectric constant substrate mitigates the ω^3-dependent Cherenkov radiation loss that often dominates on-chip THz propagation [102, 83, 174, 101, 107, 106]. In addition, the local dielectric environment of thin COP substrates does not support unbound leaky-wave modes (from Ch. 3).

We use separate transmission lines for THz emission and detection. The transmission line for the THz detection is gapped between the detector A-switch and B-switch to allow for the placement of the sample board. Sample boards are fabricated on 50 μm thick z-cut dual-side polished quartz. Nanofabrication of fragile 50 μm quartz is accomplished by crystal-bonding quartz to a silicon carrier wafer, patterning the CPS and performing any additional sample-related nanofabrication steps, and then dissolving the crystal bond to release the quartz film. The sample board is then aligned to the switch board under an

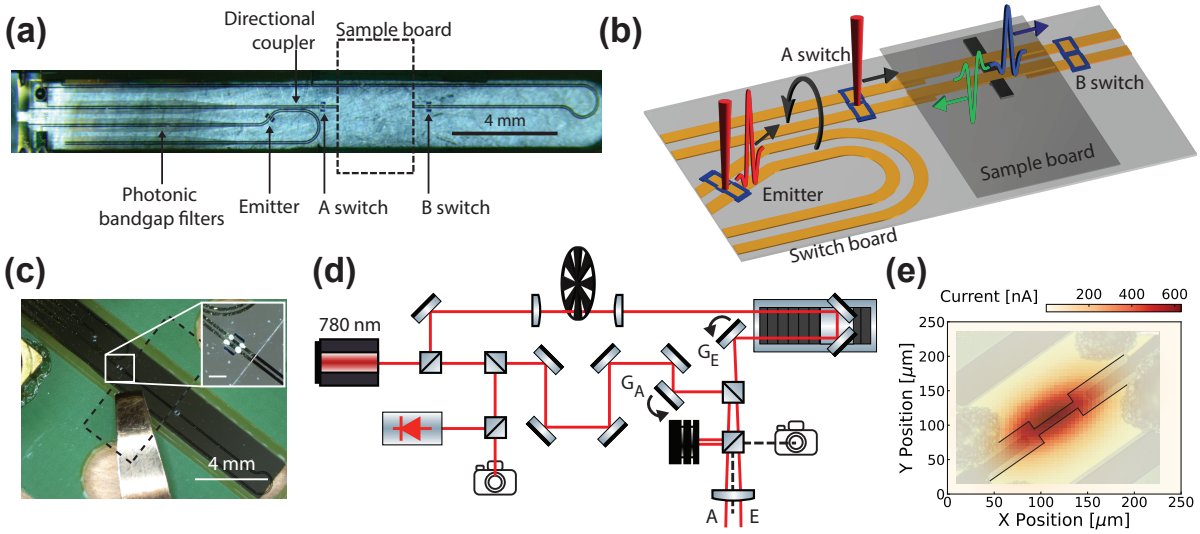

Figure 4.1: On-chip THz spectrometer with fast sample exchange architecture. (A) Micrograph of the switch board, showing key circuit components and three LT GaAs PC switches (emitter, A-switch, and B-switch) with the location of the replaceable sample board indicated. (b) Schematic illustration of the on-chip THz spectrometer, the switch-sample board architecture, the emitted propagating transient (red), reflected transient (green) and transmitted transient (blue). (C) Micrograph of a thru sample board (outlined by the black dashed box) mounted on the switch board. The sample board transmission lines are precisely aligned to the switch board transmission lines, as shown in the inset (scale bar: 250 μm). A metal clip holds the sample board firmly in place to ensure excellent coupling of the switch board and sample board transmission lines. (D) Block optics diagram, showing a pulsed laser split into two paths. One beam is chopped, delayed, and directed by scanning galvanometer G_E to the emitter (E) switch. The other beam is statically delayed and steered by scanning galvanometer G_A to the readout (A) switch. The reflection from the sample is directed to a camera through the beamsplitter, used for coarse alignment. Some power from the A-switch beam is siphoned for power and Poynting stability monitoring. (E) Dynamic alignment scan map, showing recorded DC current versus galvanometer-controlled beam position with overlaid micrograph of the emitter photoconductive switch. Black lines denote the edges of gold transmission lines.

optical microscope, resulting in a reliable placement precision of $\leq$10 μm. The resulting combined structure is shown schematically in Fig. 4.1B while a micrograph of an aligned thru sample board is shown in Fig. 4.1C.

We illuminate both the emitter and detector (A/B) PC switches with 3 mW of 780 nm, 75 fs Gaussian laser pulses with a repetition rate of 250 MHz. The optical path is illustrated in Fig. 4.1D. The beam to the emitter switch is chopped at $\sim$3.2 kHz, routed through an optical delay stage, directed on the emitter switch using a scanning galvanometer (G_E), and focused to diameter $\approx 25\ \mu m$. A beam-splitter extracts $\sim$500 μW for power and spot stability monitoring on the A-switch beam path. We apply a 10V DC bias across the emitter switch, providing the energy that is converted into a THz transient in the emission line when the emitter switch is illuminated. This THz pulse is subsequently coupled to the detection transmission line via an inductive directional coupler, which provides a DC block between the emission and detection transmission lines (Fig. 4.1B). To improve the stability of our measurements, we spatially map the DC photocurrent across the emitter switch using scanning galvanometer G_E (Fig. 4.1R), allowing us to stabilize switch operation at the position of maximum photocurrent through real-time feedback.

The beam for the detector switch is routed through another scanning galvanometer (G_A) prior to being focused on the switch, thus allowing for automatic and precise alignment of the laser beam on the PC switches. The THz pulse traveling along the detection transmission line is then measured at the detector A-switch. The detector switch is momentarily triggered by an optical pulse and thus allows us to measure the current across the detector switch, which samples the local electric field across the detection PC switch. Since the instantaneous current across the detector A-switch is only hundreds of picoamps, we use a transimpedance amplifier with gain 1E-7 A/V, the output of which is then demodulated using a lock-in amplifier synchronized to the chopper. Sweeping the

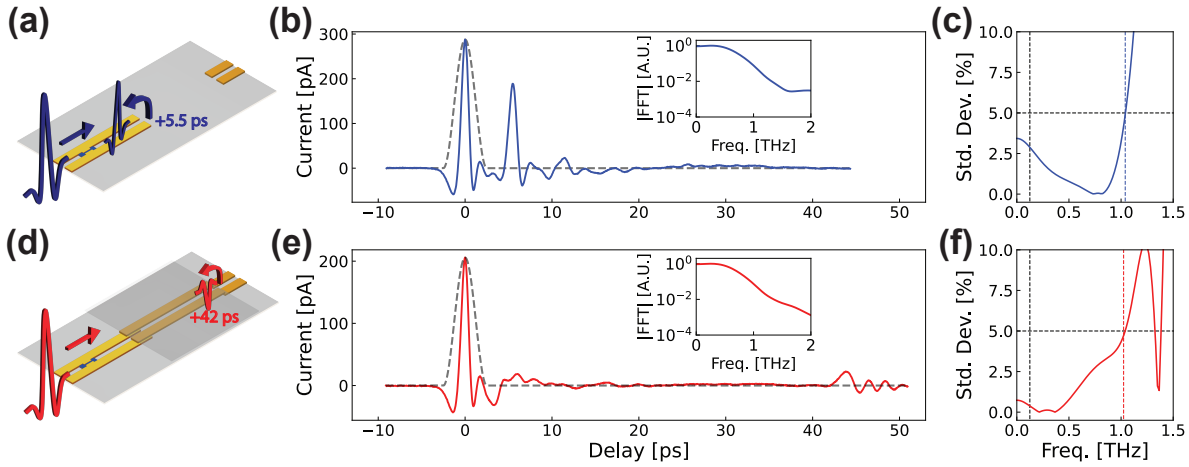

Figure 4.2: THz reflection transients for two simple samples at room temperature. (A) Schematic showing incident transient on bare switch board and reflection from open circuit termination. (B) Time domain electric field with no sample board, showing the directivity pulse, corresponding isolating window (in dashed lines), and delayed reflections corresponding to the switch-sample board interface and a round trip from the second switch-sample board interface (inset: Fourier amplitude of system directivity $D(\omega)$). (C) 5% magnitude repeatability, showing a bandwidth of 1.05 THz. (d) Schematic showing incident transient with a mounted thru sample board. (E) Time domain electric field with thru sample board, showing the directivity pulse and corresponding window (in dashed lines), a minimal first switch-sample board interface reflection, and a delayed reflection from the second switch-sample board interface (inset: Fourier amplitude of system directivity). (F) 5% magnitude repeatability, showing a bandwidth of 1 THz.

optical delay stage continuously allows us to vary the time delay of the optical pulses exciting the emitter and the detector switches. This allows us to extract the time-domain electric field of the incident, reflected, and transmitted THz pulse from the transients recorded at the PC switch A and B. We move the optical delay stage continuously, measuring the time-domain THz signal with a Nyquist frequency of nearly 350 THz. We, then, apply standard THz TDS signal processing [175, 176] and smooth the time-domain data with a 101-point, 3rd order Savitsky-Golay filter.

Fig. 4.2A-C show characterization data of the spectrometer at room temperature with no sample board mounted. In this configuration, the THz pulse traverses the A-switch (producing a 'directivity' peak at zero delay) and is subsequently reflected at the

open termination of the transmission line, producing a peak 5.5 ps after the directivity. The directivity and reflected signal have the same polarity, correctly indicating that the transmission line termination behaves as an open circuit. Additional structure, including a weaker peak at 11 ps, results from Fabry-Pérot reflections between the open circuit and A-switch. The spectral amplitude and phase of both the directivity pulse D(ω) and the sample-reflected pulse R(ω) are isolated using appropriate Hann windows and computed by Fourier transform. We define usable bandwidth as frequencies for which the standard deviation is less than 5% of the signal amplitude as shown in Fig. 4.2C, yielding a bandwidth of about 1 THz.

Characterization data for a sample board containing a bare transmission line (thru) are shown Fig. 4.2D-F. Accurate alignment of the transmission line allows the strength of the THz pulse reflected from the sample board interface (around 5.5 ps) to be smaller than 10% of the directivity, indicating good coupling between switch and sample boards. The transmitted THz pulse propagates across the sample board transmission line until the second interface, where it undergoes another partial reflection. This reflected pulse is observed at the A-switch about 42 ps after the directivity. The reflected signal is featureless between $\sim$20 and $\sim$42 ps. This window corresponds to a significant fraction of the transmission line length on the sample board, and sets the bounds for placement of the device to be tested.

For both configurations, $D(\omega)$ is repeatable to within 5% out to $\approx$1 THz (Fig. 4.2C and F). Recording $D(\omega)$ and $R(\omega)$ from the sample board interface (in addition to reflections from the device to be tested) allows us to characterize the emitted THz pulse and the sample board alignment. We consequently focus on THz reflection rather than transmission in what follows, even though our circuit permits the measurement of both THz reflection and transmission.

Having characterized the on-chip THz spectrometer at room temperature, we now

demonstrate its cryogenic capability by extracting the complex optical conductivity ($\sigma(\omega)$) of a superconductor as it is cooled below its transition temperature. Conventional superconductivity exhibits distinct features such as the superconducting gap (2Δ) in the real optical conductivity ($\sigma_1(\omega)$) and $-1/\omega$ like response in the imaginary optical conductivity ($\sigma_2(\omega)$) [42]. To inform the design of the sample, we model a sample on the transmission lines as complex admittance connecting one metal line to the other. The admittance takes the form $\mathrm{Y}(\omega) = \mathrm{Y}_1(\omega) + \mathrm{jY}_2(\omega)$ for radial frequency ω and imaginary unit j.The reflected signal $\mathrm{R}(\mathrm{YZ}_0,\omega)$ is proportional to the directivity $\mathrm{D}(\omega)$ and sample reflection coefficient $\Gamma(YZ_0, \omega)$, which is given by:

$$\Gamma(YZ_0, \omega) = \frac{-Y(\omega)Z_0}{2 + Y(\omega)Z_0} \tag{4.1}$$

where Z_0 is the characteristic impedance of the transmission line. We are interested in maximizing the response to small changes in sample admittance δYZ_0, which can be expressed as $\delta\Gamma = \delta YZ_0(\partial\Gamma/\partial YZ_0)$. Using eqn. 4.1 for Γ, we find that this quantity is maximized when $\mathrm{YZ}_0 = 2$ (see Ch. 1). The lateral width and thickness of the superconducting film are thus chosen to satisfy the maximal contrast condition at temperatures just above the superconducting transition.

For the superconducting film, we use NbN, chosen for its moderate superconducting transition temperature (T_C) and well-documented electrodynamics[177, 178, 179, 180, 181]. NbN films were deposited by magnetron sputtering (AJA ATC 2200-V sputtering system) with power 300W, pressure 1.8 mTorr, and 45/3.8 sccm $\mathrm{Ar/N_2}$ gas flow. The four-terminal DC transport curve is shown in Fig. 4.3A, with the data scaled by a geometric factor to accurately represent the resistance shorting the transmission line.

The sample has a T_C of 11.1K, where the T_C is defined as 90% of the normal state

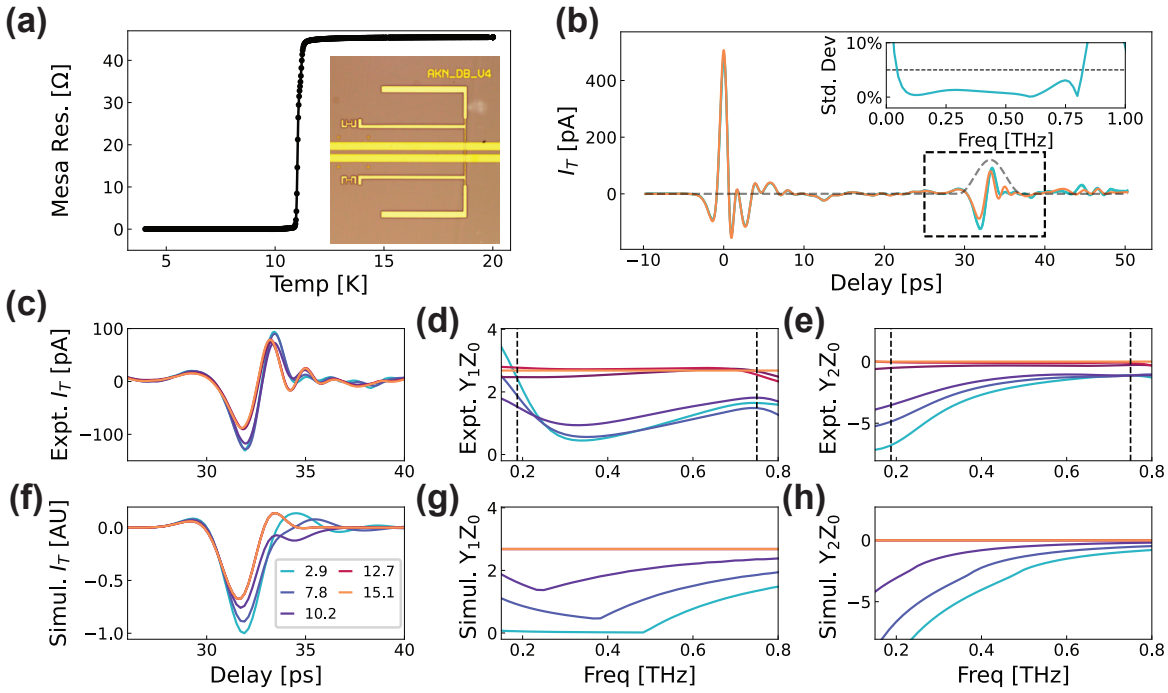

Figure 4.3: On-chip THz spectroscopy of superconducting NbN. (A) DC, four-terminal measurement of 7.5 μm NbN mesa with inset showing sputtered NbN device on a sample board (B) On-chip THz TDS of NbN mesa, with dashed lines drawn around NbN reflection (inset: representative 5% magnitude of 810 GHz). Red and orange colors indicate temperatures above T_c; purple and blue colors indicate temperatures below T_c. (C) Windowed time-domain NbN reflection (D) experimentally-measured real admittance and (E) experimentally measured imaginary admittance (F) Zimmermann-simulated time domain transient (G) Zimmermann-simulated real admittance and (H) Zimmermann-simulated imaginary admittance.

resistance at 15K. With lateral width 7.5 μm, the sample is 3.2% of one wavelength at 750 GHz, justifying the lumped element approximation. The sample thickness of 151 nm is chosen to bring the resistance in the normal state as close as possible to the maximum contrast condition. Experimentally, we find that the sample presented here has $\mathrm{Y}Z_0 = 2.58$ at room temperature, which, given the residual resistivity ratio of 0.94 for NbN, ensures good matching at low temperatures. The sample board containing the NbN film is aligned to the switch board and cooled down to $T \approx 3$K in an exchange gas environment. The exchange gas ensures thermalization of the sample, despite the poor thermal conductivity of the COP switch board.

To extract the complex admittance of the superconducting state, we assume a frequency-independent admittance in the normal state, Y_N, which we take to be equivalent to the conductance measured by DC transport. The ratio of the reflected signal in the normal and superconducting states can then be used to extract the superconductor's admittance $\mathrm{Y}_S(\omega)$ via the relation:

$$Y_S(\omega)Z_0 = \frac{-2}{1 - \frac{R_N(\omega)}{R_S(\omega)}\frac{D_S(\omega)}{D_N(\omega)}\frac{2+Y_N Z_0}{Y_N Z_0}} \tag{4.2}$$

The THz time-domain reflection signal of the NbN sample measured above and below the T_C is presented in Fig. 4.3B, with the dashed box outline highlighting the signal reflected from the NbN device. The repeatability of the THz signal reflected from the sample at a temperature of 3K is shown in the inset, showing a usable bandwidth extending to $f \approx 750$ GHz. The dynamic range, defined as the ratio of the peak Fourier amplitude and the background noise level exceeds 2000 for the NbN reflection.

Fig. 4.3C shows the time-domain reflection from the NbN, after application of the window shown in gray in Fig. 4.3B. The measured reflection is temperature indepen-

dent at high temperatures but exhibits strong temperature dependence after the onset of superconductivity; qualitatively, the peak signal increases while simultaneously also shifting to later times by $\delta t \leq 200$ fs as the temperature is lowered below T_C. These behaviors are consistent with expectations. Specifically, the increase in the peak reflection signal results from decreased THz absorption by normal carriers, while the lagging phase arises from the inductive response of the incipient Cooper pairs[41].

The real ($Y_1(\omega)$) and imaginary ($Y_2(\omega)$) admittance, expressed in units of transmission line impedance (Z_0), are extracted using eqn. 4.2 and are shown in Fig. 4.3D and E, respectively. Free space THz conductivity measurements of NbN in the normal state report a featureless real conductivity equal to the measured DC value and a negligible imaginary conductivity [177, 179] from DC to 750 GHz. We therefore use the DC value of Y_N for all measurable frequencies.

The upper and the lower frequency cutoffs are delineated with black dashed lines in Fig. 4.3d and e. The upper frequency cutoff is defined as the frequency at which the normalized standard deviation of the THz signal magnitude is at least 5% (inset), while the lower frequency cutoff is defined as the product of the minimum resolvable frequency of the windowed directivity or windowed NbN reflection and the equivalent noise bandwidth (ENBW) of the isolating Hann window of 1.5. The minimum resolvable frequency is given by the maximum size of the isolating window used for the reflection, which is set by the distance between the signal of interest and features arising from unwanted reflections such as those that occur at the switch board/sample board interface. Practically, in our current geometry the low frequency limit is 200 GHz. This could be be extended by increasing the physical spacing between the detector switch A and the switch-sample board interface.

We simulate the electromagnetic response of the superconductor using a model developed for disordered superconductors by Zimmermann and collaborators [42]. We take

the NbN to be in impure limit ($y = 500$), use $T_C = 11.1$K, and use a superconducting gap $2\Delta(T)$ of the following form:

$$2\Delta(T) = 2\Delta(0)\tanh\left(\frac{\pi k_B T_c}{\Delta(0)}\sqrt{T_c/T - 1}\right). \tag{4.3}$$

where, $2\Delta(0)$ is 500 GHz and k_B is the Boltzmann constant. The simulated time-domain response of the superconductor and the corresponding real and imaginary admittances are plotted in Fig. 4.3F, G, and H, respectively. We use the complex admittance obtained from the BCS model[42] to simulate the reflected signal in the frequency domain, as shown in Fig. 4.3G and H. The time domain signal shown in Fig. 4.3F was then reconstructed from the complex admittance using the inverse Fourier transform.

The measured and the simulated time-domain profiles show similar qualitative features such as the increase in signal peak value and the phase lag below the superconducting transition temperature. The real admittance ($\propto \sigma_1(\omega)$) of the superconducting state reduces as we go from the high frequency ($\omega > 2\Delta$) response of the normal carriers into the gap ($\omega < 2\Delta$), followed by an increase due to thermal broadening of the superfluid peak at $\omega = 0$. The imaginary admittance ($\propto -\sigma_2(\omega)$) of the superconducting state shows the characteristic $1/\omega$-like response indicative of the superfluid density, whose spectral weight increases as the temperature decreases.

While our results demonstrate the qualitative electrodynamic response of a superconducting thin film, further improvements in phase accuracy and bandwidth are needed to get a better quantitative understanding of, say, unconventional superconductors. Imperfect alignment of the laser spot and the PC switches, for instance, introduces a random variation in the THz emission and the propagation length of the THz pulses. We use the directivity ratio $D_S(\omega)/D_N(\omega)$ in eqn. 4.2 to calibrate the system, however, the THz

path error of length δL subsequently accumulates to a phase error of $\omega\delta L/c$. Better phase sensitivity can likely be achieved by miniaturizing the photoconductive switches to match the laser spot.

In summary, we have demonstrated a cryogenic-compatible on-chip THz spectrometer with a fast sample interchange architecture, and used it to extract the complex optical admittance of a deeply sub-wavelength superconducting NbN sample between 200 and 750 GHz. The confinement of the THz field within the transmission line and ease of sample interchange makes this spectrometer architecture ideal to investigate gate-tunable vdW heterostructures, with near-term applications to the physics of correlated electronic states and ultra-high frequency devices.

4.2 COP and unbound substrate modes

The generation of unbound leaky wave modes [108] greatly affects the ability of an on-chip THz spectrometer to measure the reflection and transmission from the device. Ideally, the THz pulse would propagate in only the CPS's bound quasi-TEM mode [6]. If any coplanar transmission line and substrate dielectric environment supports leaky-wave modes, such modes are radiated isotropically upon emitter excitation. Some radiation goes into free space, some radiation is coupled to the transmission line as a bound mode, and some radiation is coupled as a leaky-wave mode in the substrate. All three types of modes are propagating THz electric fields that can be sampled by readout PC switches [182]. Unbound free space modes will propagate to the readout PC switch many picoseconds before the bound transient. The bound transmission line mode, then the unbound substrate modes, will reach the readout PC switch. Unbound substrate modes may propagate longitudinally in the substrate or may reflect from the substrate underside and edges, delaying their arrival at the readout PC switch by up to hundreds of

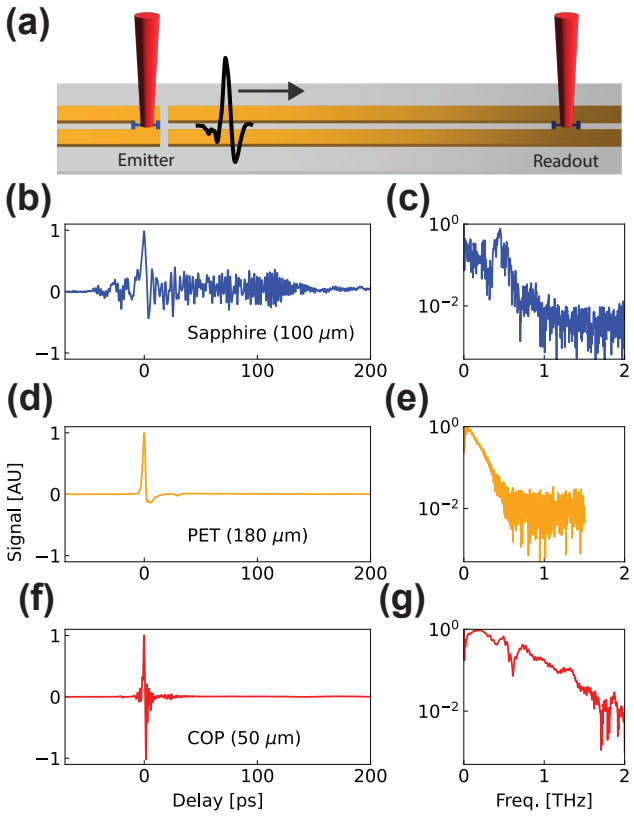

Figure 4.4: Presence of unbound leaky-wave modes in alternate substrates. (A) Illustration of measurement schematic, showing a bound quasi-TEM transient propagating from the emitter to the readout, and time domain/frequency domain THz transients measured after 10 mm propagation with substrates: (B/C) 100 μm thick sapphire, (D/E) 180 μm thick PET and (F/G) 50 μm thick COP.

picoseconds (Ch. 3). As the substrate thickness is reduced, the substrate begins acting as a dielectric waveguide for unbound modes, producing measurable electric fields at long time delays.

To demonstrate the effects of unbound leaky-wave modes, a thru circuit was fabricated with two LT GaAs switches placed 10 mm apart (Fig. 4.4A). The emitter is connected to the thru CPS of the same dimensions used in the main text via an capacitive coupler. The circuit was fabricated on three substrates: 50 μm thick COP, 180 μm thick polyethylene terephthalate (PET) and 100 μm thick sapphire. The resulting transients and their normalized Fourier transforms are shown in Fig. 4.4B-G.

The measured transmission with a sapphire substrate shows a signal 35 ps before the main transient from unbound free space modes. The bound transient, mixed with unbound modes, appears at zero delay. Strong oscillations continue for 130 ps, due to the leaky wave substrate modes and their reflections from the substrate underside and periphery. After 130 ps, the oscillations reduce in amplitude, but do not vanish. The bound transient cannot be isolated with a narrow window because that peak contains bound and unbound contributions. Even if the bound transient could be isolated, the window would have to be so narrow that the minimum resolvable frequency would be reduced to about 1 THz.

The unbound mode cannot propagate for thin PET and COP substrates as a result of decreased substrate permittivity. The moderate loss tangent and dispersive index in PET, however, results in a bound transient with only 300 GHz bandwidth. COP's frequency-independent index and minimal loss tangent result in a large signal, large bandwidth transient, even after 10 mm of propagation. COP is therefore the best of the three thin substrates for THz applications.

In the case of short (< 1 mm) transmission lines fabricated on thick (> 1 mm) substrates, the reflected leaky-wave modes do not have sufficient time to propagate to the readout PC switch and are thus not conflated with the desirable bound mode signal. Both short transmission lines and thick substrates are undesirable. Short transmission lines increase the minimum resolvable frequency; thick substrates suffer exacerbated Cherenkov radiation loss. The use of COP substrates mitigates both factors, although at a cost in difficulty of switch fabrication.

Chapter 5

Additional Spectrometer Characterization

"Experiments must be designed by geniuses so they can be run by idiots."

A. F. Young

We present additional characterization information, including numerical modelling, loss characterization, and a systematic review of best operating conditions, for the fast sample exchange spectrometer introduced in Ch. 4.

5.1 Finite Element Simulation

Modal finite element analysis in Ansys HFSS was used to compute basic transmission line parameters. The switch board was simulated using COP with $\epsilon_r = 2.26$, quartz sample boards with $\epsilon_r = 3.78$, and gold transmission lines with the 300K value of conductivity 2.2 $\mu\Omega$ cm. Finite element simulations (1 and appendix B) were performed with a sample-board / switch spacing δz of 0, 5, 10 15, and 20 μm, as well as with the sample board placed infinitely far away ($\delta z \to \infty$). The impedance and complex

propagation coefficients were extracted directly from simulation. The phase velocity was derived using Table 1.3.

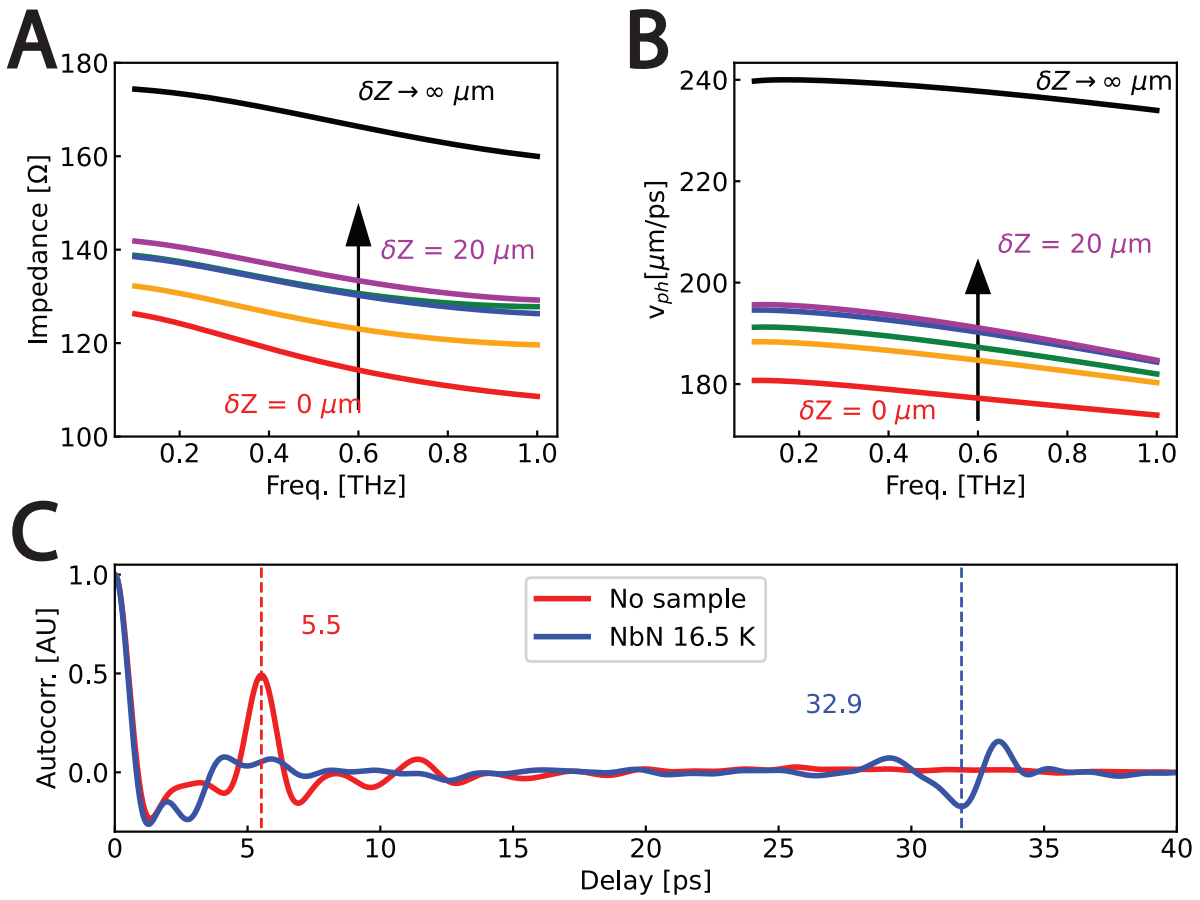

Figure 5.1: Finite element simulation of spectrometer architecture, with sample boards placed infinitely far away (black), 0 μm away (red), and in steps of 5 μm. (A) Impedance and (B) phase velocity. (C) Autocorrelogram for no sample and with NbN at 16.5K, whose timings are used in time flight analysis to extract phase velocities.

The results, plotted in Fig. 5.1, show a weakly frequency-dependent impedance that decreases as frequency increases, but increases as board separation increases. From Table 1.3, a transmission line's characteristic impedance to zeroth order in R' is $\sqrt{L/C} \sim \sqrt{\mu/\epsilon}$. As the two boards are increasingly separated, the dielectric volume through which the most intense electric field lines go decreases, lowering the impedance.

To test the simulation, the phase velocity was estimated through time of flight calcu-

lations. The pulse's autocorrelation (Fig. 5.1C) was computed with no sample mounted (300K) and with NbN (16.5K). The autocorrelation, which measures waveform self-similarity at all time delays, shows a peak from the board interface at 5.52 ± 0.01 ps. The pulse travels 565 ± 15 μm, depending on laser alignment, yielding a phase velocity of 204.7 ± 10.8 μm/ps and index of n = 1.48 ± 0.08. The minimum of the NbN reflection occurs at 32.9 ± 0.01 ps. The sample is placed 2496 ±1 μm from the board interface. Using the sample board phase velocity derived above, the speed light in the sample board section is 175 ± 2.2 μm/ps, with index 1.72 ± 0.02.

The simulation, therefore, overestimates the phase velocity in the switch board. The measured index is that of pure COP (1.52), within the uncertainty due to the automatic realignment procedure. HFSS's computation matches the zeroth order quasi-TEM, static index derived from conformal mapping: $n \approx \frac{1}{2}(\sqrt{\epsilon} + 1)$. The slow speed of this mode could likely comes from the geometric dispersion effects described in Ch. 3, as S and W are several 10s of μm. The phase velocity in the sample board suggests $\delta z \leq 5\mu$m. This is consistent with the observation that placing the clip near the transmission lines increases coupling and, therefore, decreases δz.

5.1.1 Charge per pulse / Electric field strength

Most THz TDS articles list measured electric field as the current or voltage in arbitrary units, which obfuscates the exceptionally small pA-level signal amplitudes from the common eye. Each THz transient can be thought of as a propagating wavelet of charge on the transmission lines. At first inspection, one might conclude that the total charge is less than one electron. The $\sim$ 100 pA current is the time-averaged current from many pulses. The THz pulses themselves have a the same duty cycle as the laser (250 MHz). The THz pulses last only 1 ps per every 4 ns. One should multiply the charge per pulse

by the duty cycle[1], which suggests that the transients have an instantaneous current of $Q_{pulse} = I_{avg} t_{rep} \approx 4E-14C = 2.5E5e^-$. The actual total charge of the directivity and NbN reflection pulses, which have total charge 6.64E-14 C (4.13E5 e^-) and 2.10E-14 C (1.31E5 e^-), respectively.

One should also note that the reported current in Ch. 4 is the current that flows across the PC switch, which is not (necessarily) the current that supports the propagating electric field transient. An illuminated PC switch has a linear relationship between the DC biasing voltage (DC electric field) and the admitted current, which describes the conversion an electric field across the PC switch into a current. If this slope were infinite (optically induced zero-resistance PC switches), the transmission lines' resistance to the chopped 3.2 kHz signal on the order of hundreds of Ohms limits the registered current. Even an (impossible) infinite current-field slope would still be limited by the transmission line resistance.

A lower bound on the electric field strength at the A-switch can be computed from the LIV curves. If the LIVs are linear and taken under the same illumination, they represent the long time, averaged current that flows under a fixed DC bias. The transient electric field (E_N) that produces current (I_N) for LIV bias V_{bias} and LIV current I_{LIV} is given by:

$$E_N = \frac{V_{LIV}}{dI_{LIV}} I_N \tag{5.1}$$

This yields electric fields of ∼10 kV/cm and ∼1 kV/cm (through the bulk of the A-switch) for the directivity and reflection respectively. These are weak THz fields. For comparison, the electric field intensities used for the nonlinear THz spectroscopy of a marginal Fermi glass in doped silicon [183] are a few of MV/cm. Tapering the

[1] and account for the difference between RMS and peak values reported by lock in amplifiers, see [79]

transmission line to a small widths of less than 1 μm may increase the field to strengths sufficient for nonlinear on-chip THz TDS.

5.2 Signal flow diagram

A full signal flow diagram for the fast sample exchange spectrometer is shown in Fig. 5.2. The photoconductive switches are modelled as input(1) (the emitter), output(1) (A-switch) and output(2) (B-switch). Transmission line propagators are indicated in orange; the sample board / switch board interchange mechanism is indicated in blue, radiative propagators are indicated in red, and the device under test in black. As before, each arrow represents a complex, frequency-dependent propagator. For convenience, we define the z-axis as moving from left to right.

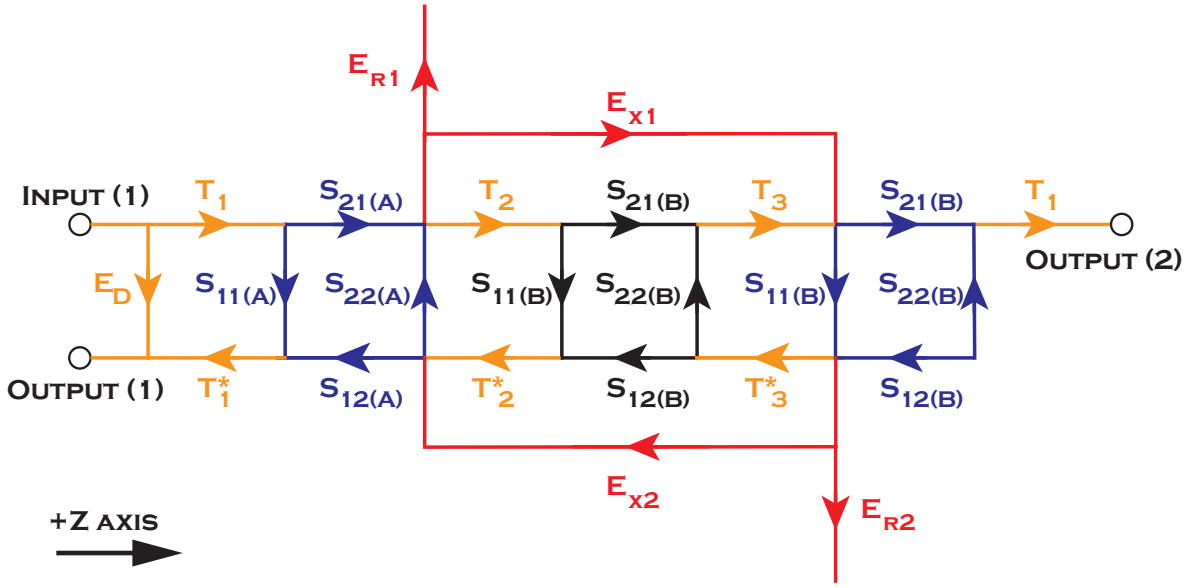

Figure 5.2: Signal flow diagram for the fast sample interchange spectrometers with color-coded transmission lines (orange), interchange mechanism (blue), radiative coupling (red) and device under test (black).

Signal is sourced at input(1) and propagates to the A-switch. The signal flows directly from the input to the output by the directivity (E_D) propagator. The directivity

propagator is excluded from all loops, and thus its existence never detracts from measuring additional propagators. We used this property in Ch. 4 to calibrate emitted pulse amplitude and phase. The signal propagates down the transmission line (T_1) until it reaches the switch board - sample board interface (blue). The sample board has a full two-port network S-matrix. The signal transmitted through the interface continues propagating down the transmission line (T_3), where it reaches another full S-matrix for the second switch board - sample board interface. The transmitted signal propagates down another transmission line, which is assumed to be identical to the first transmission line segment (T_1). If the sample is placed in the center of the sample board, then $T_2 = T_3$. The two sample board - switch board interfaces cannot be assumed to be identical, as the mounting can be different. Note that there is also radiative coupling to free space (E_{R1}) or the second switch board - sample board interface (E_{X1}).

A full calibration would require the use of Mason's laws to simplify the signal flow diagram's complicated loops and interconnected structure. This is intractably complex, particularly considering the radiation to free space couplings E_{R1} and E_{R2}. One must make approximations, take advantage of the time domain, and simplify the signal flow piece by piece in order to obtain a simple calibration. We will now consider the relevance of each set of propagators.

5.2.1 Interchangeable board interfacial reflections

THz signals incident on the switch board / sample board from the switch board transition from $\sim 160\Omega$ to $\sim 110\Omega$, and should thus incur a reflection coefficient of -0.19. Instead, they reflect with positive polarity, suggesting that the interfacial impedance transition is dominated by a capacitive gap between the two boards. The horizontal gap can be reduced by careful alignment (see appendix C). The vertical gap can be reduced

by squeezing the two boards closer together.

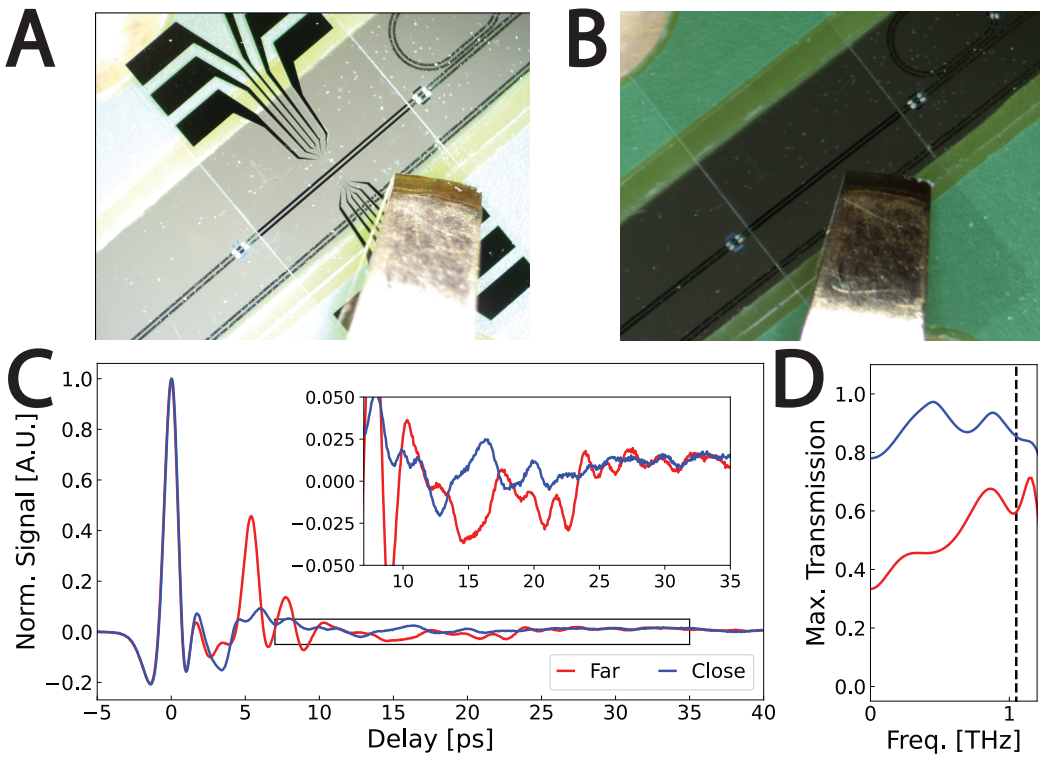

Figure 5.3: Importance of clip placement in interfacial coupling. Optical micrographs of the clip affixing thru connectors to the sample board, in the (A) "far" and (B) "close" positions. (C) Time domain data, showing that placing the clip close to the transmission line suppresses the interfacial reflections (inset: zoom of the boxed region, showing that the close clip has fewer spurious signals at delays longer than the first interfacial reflection). (D) Computed upper bound on interfacial transmission coefficient.

In Fig. 5.3, the two optical micrographs show the clip far from (A) and close to (B) the transmission line. Time domain data taken from the A-switch shows that placing the clip far from the transmission line results in a large interfacial reflection: 45% of the directivity amplitude. Placing the clip within 200 μm of the transmission line, however, reduces this reflection signal size to 9%. The benefits of reduced interfacial reflection last to later time delays. the signal is reduced even at long delays when the clip is close to the transmission line because the severity of Fabry-Pérot reflection points is reduced

when the interfacial impedance mismatch is lessened. Reducing such spurious reflections is key to simplifying the analysis, as we will show shortly.

Using the computed upper bounds on extinction coefficient, one can solve for the interfacial transmission. We see that this transmission coefficient is near unity and frequency-independent when the clip is close. The lack of frequency dependence suggests that the interface coupling capacitance is acting as a short. The transmission coefficient is smaller and frequency-dependent when the clip is far. The low transmission at low frequency suggests capacitive impedance, consistent with moderate vertical separation between the two boards.

In summary, place the clip near the sample board transmission lines to keep the coupling large.

5.2.2 Transmission line attenuation

Consider the transmission line couplings (orange), which simply describes the propagation of a polychromatic wave packet along the z-axis and are therefore of the form $T = Ae^{-(\alpha - j\beta)z}$. α and β are given in Table 1.3, and A is some constant. Waves propagating in -z have a conjugated exponent, meaning they attenuate as much as the forward moving wave, but the phase evolves in the opposite direction. Assuming a very weakly temperature-dependent substrate dielectric coefficient, these propagators should be the same for the reference and the relevant condition, so they will be unimportant.

The signal flow diagram has too many propagators to fully resolve the frequency-dependent extinction and propagation coefficients. Their upper bounds, however, can be estimated. We define the switch board only section as region 1 (length L_1) and the sample board / switch board hybrid as region 2 (length L_2). Using the propagation length of 565 $\pm$ 15 μm, the interfacial reflection is $I(\omega) = T_1 T_1^* S_{11(A)}$. The interface is

not a perfect open due to radiative loss, but assuming it is ($S_{11(A)} = 1$) provides an upper bound on the extinction coefficient in region 1:

$$\alpha_1 = \frac{1}{2L_1} ln \left(\frac{D(\omega)_{open}}{I(\omega)_{open}} \right) \tag{5.2}$$

We then insert a sample board with a short circuit. The interfacial reflection $I_{short}(\omega)$ can be estimated using the extinction coefficient just derived. Thus: $S_{11(A)_{short}} = I_{short}(\omega)e^{2(\alpha_1 L_1)}/D_{short}(\omega)$. Assume that the interface is lossless, so $S_{21(A)_{short}} = 1 - S_{11(A)_{short}}$. Assuming also that the short circuit is lossless and reflects all incident signal, an upper bound on the extinction coefficient in region 2 can be solved from the measured reflected signal $R(\omega)$:

$$\alpha_2 = \frac{ln \left(\frac{-R(\omega)}{D(1-S^2_{11(A)_{short}})} \right) - 2\alpha_1 L_1}{2L_2} \tag{5.3}$$

The frequency-dependent extinction coefficient upper bounds in the two regions are plotted in Fig. 5.4. Both have superlinear frequency dependences, and were fit to a power law with constant offset. It is clear that wave attenuation is much more severe in region 2 than in region 1. The extinction coefficient in region 1 scales with exponent 1.734 ± 0.048; the extinction coefficient in region 2 scales with exponent 4.258 ± 0.301. Cherenkov radiation should scale as ω^3, in addition to weaker, implicit frequency dependences in ϵ_r.

From eqn. 3.11, one expects the Cherenkov extinction coefficient to be 4.7 times more severe in region 2 than region 1[2]. Cherenkov radiation semi-quantitatively fits the ratio of extinction coefficients, but fails to capture the different power laws in frequency. The sub-cubic power law in region 1 suggests that Cherenkov loss is not the dominant loss mechanism. The super-cubic power law in region 2 may stem from some implicit

[2]The sharp ϵ^3 power law in extinction by Cherenkov radiation would imply that the extinction coefficient in sapphire is 160 times that of COP - and the total attenuation is exponentially sensitive to this ratio.

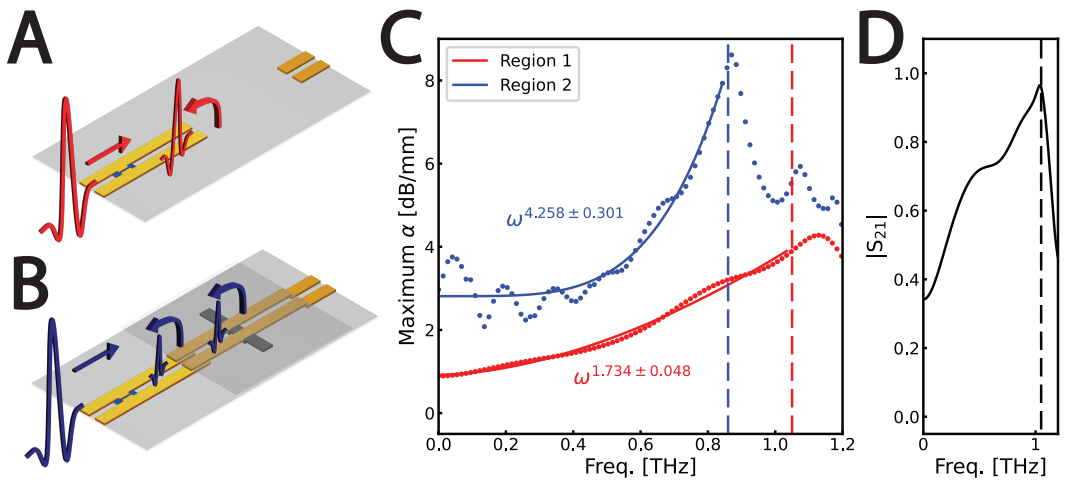

Figure 5.4: Upper bounds on the extinction coefficient in regions 1 and 2. (A) "Open" circuit geometry used to compute upper bound on α_1. (B) Mounted short circuit used to compute upper bound on α_2 and interfacial reflection coefficient. (C) Computed extinction coefficients in regions 1 and 2 with raw data in dots and cubic fits in solid lines. Vertical lines demarcate the upper bandwidth limits derived in Ch. 4. (D) Computed interfacial transmission coefficient, which increases with frequency to the indicated bandwidth limit. The increased high-frequency coupling stems from better capacitive coupling between the boards at higher frequencies.

frequency dependence in eqn. 3.11 or from trapped surface wave between the two boards. The increased DC loss in region 2 makes it obvious that the low-frequency dominant loss mechanisms are also different. It is difficult to precisely discern power laws without more precise or wider ranges of data. As a check of calculation, the interfacial transmission coefficient was also computed. It should mediocre coupling (<50% near DC) that increases to near unity in the upper range of the calculable bandwidth. The clip was placed far from the transmission line, so we are confident in our calculation.

In summary, the extinction coefficients in Fig. 5.4 are upper bounds for the CPS on pure COP and for the fast sample interchange region. They follow power laws, and aren't obviously limited by pure Cherenkov radiation. Thin COP outperforms thin quartz at THz frequencies.

5.2.3 Thick substrates are lossier than thin ones

The formalism of the previous section allows us to compare the extinction coefficients of thin (50 μm thick) quartz substrates to that of thick (625 μm thick) quartz substrates. Transmission lines of the same metallization and dimensions as the previous section were fabricated on sample boards made from 625 μm thick Z-cut quartz ("thick quartz"), 50 μm thick Z-cut quartz ("thin quartz"), and 50 μm thick COP. Identical short circuits were fabricated on the thick and thin quartz sample boards. The time and frequency domain results, and cutoff frequencies for measurement, are shown in Fig. 5.5.

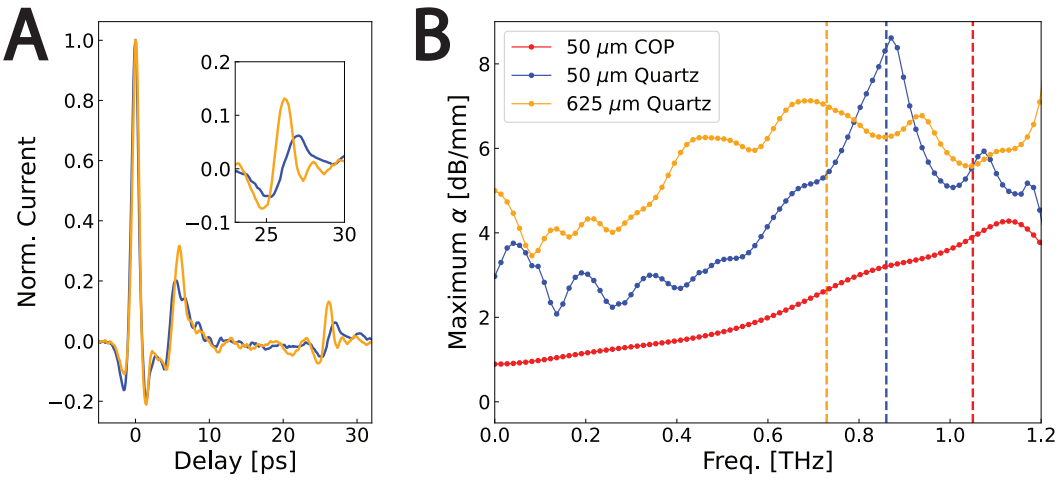

Figure 5.5: Substrate-thickness dependent loss. (A) Time-domain transients from short circuits fabricated on thick and thin quartz, with an inset showing the short circuit reflection. (B) Extinction coefficients for the three substrates, with vertical lines indicating the usable bandwidth limits. These are upper bounds on α, as the analysis is not capable of resolving the difference in switch board / sample board interfaces for the 50 μm and 625 μm thick quartz substrates.

The usable bandwidth difference between the short circuit reflection is easily visible in the time domain. The short circuit reflection on thick quartz has a smaller amplitude and longer wavelength features than the reflection from its thin quartz counterpart. The extinction coefficients make it obvious: thin COP greatly outperforms thin quartz, which greatly outperforms thick quartz.

The extinction coefficients follow a simple pattern. The extinction coefficient of COP at ~ 1 THz is nearly that of thin quartz at low (< 100 GHz) frequencies. Thin quartz's extinction coefficient sharply increases around 600 GHz to ~ 4 dB/mm, which is nearly the extinction coefficient of thick quartz at low frequencies. The usable bandwidth of the thin quartz is smallest, simply because high frequency components are more rapidly attenuated than in thin quartz or COP.

The propagating signal is exponentially suppressed by the extinction coefficient. To illustrate, take a typical distance (2 mm) as an example. After propagating, a 100 GHz signal would be 65.1% of its original size in COP, 47.7% of its original size in thin quartz, and 41.8% of its original size in thick quartz. The loss at high frequencies is much worse. A 700 GHz component would be 57.1% of its original size in COP, 30.1% of its original size in thin quartz, and only 19.8% of its original size in thick quartz.

Thick and thin quartz should have the same loss tangent, so the loss difference must be attributed to another mechanism. We established Cherenkov radiation as the leading loss mechanism which, critically, depends on the relative permittivity presented to the transient. Quartz has a permittivity of 3.78, larger than COP's permittivity of 2.26. Cherenkov loss scales as the permittivity cubed (eqn. 3.11), suggesting that quartz should be 4.7 times lossier than COP. The transient's fringe electric fields penetrate the thin substate into free space, but are confined in the thick substrate. That is: the additional dielectric volume of the substrate is responsible for the additional loss with thick quartz.

In conclusion: use COP when possible. Thin quartz can also work. Thick quartz introduces extra loss. High-ϵ substrates are dramatically worse.

5.2.4 Radiative coupling

The spectrometer signal flow diagram includes radiative coupling between the first interface and the output. This is not as insignificant as one might expect - the high propensity for radiation at THz frequencies (Ch. 3) leads to increased radiative coupling and loss.

To demonstrate, consider the A-switch and B-switch signals when no sample board is loaded and when a thru sample board, with a bare transmission line, is loaded (Fig. 5.6). One might expect the unloaded case to behave as an open circuit, wherein all incident signal at the transmission line edge would reflect with $\Gamma = 1$ and there would be no transmitted signal. This is not the case - there is a clearly transmitted signal. Time of flight analysis produces velocity between transmission lines of $302 \pm 16\ \mu m/ps$, which is the vacuum speed of light ($300\ \mu m/ps$), confirming that this is a free space radiative coupling between the two transmission lines. The signal transmitted through the thru connector arrives much later, and propagates with phase velocity $150 \pm 4\ \mu m/ps$. The slower speed is due to the increased dielectric volume with the sample board.

Neglecting radiation coupling on this basis, however, would be premature. The transmitted signals for the "open" circuit and thru connectors, shown in real units in Fig. 5.6, are the same order of magnitude - suggesting that the radiative coupling cannot be neglected. Although there is a substantial arrival time difference for the free space and coupled pulses, one cannot assume that a terminated transmission line will behave as a true open circuit.

Note that the radiative coupling propagators do not flow to output(1) in the signal flow diagram (Fig. 5.2), except after reflection from the second interface, which would arrive at a later time than the sample reflection. Since all signals are exponentially suppressed in signal flow path length, radiative effects are much small in reflection geometry

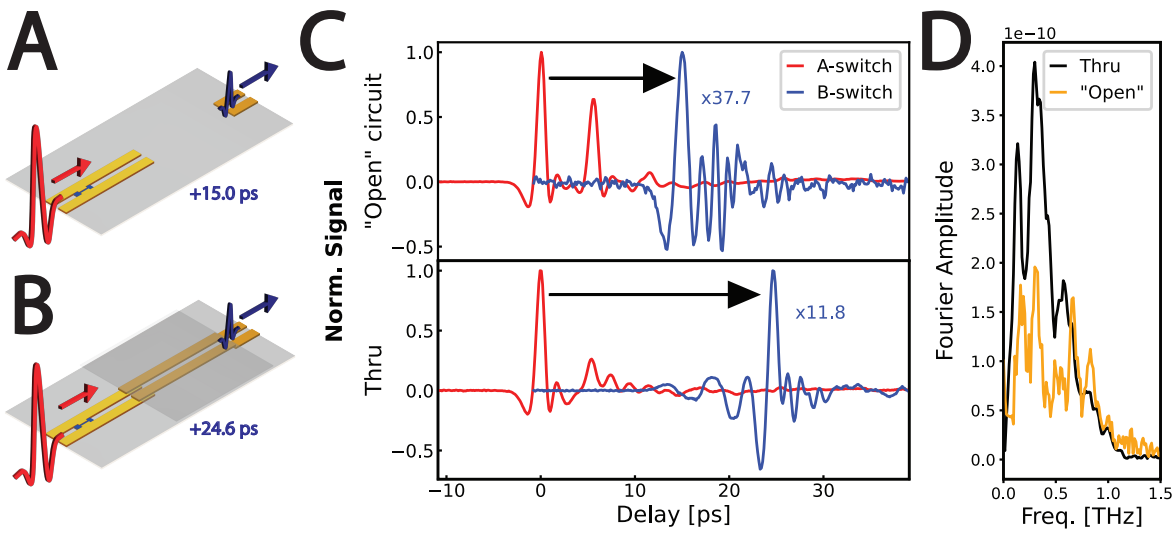

Figure 5.6: Radiative coupling. Schematics for reflection and transmission with (A) an "open" circuit (no mounted sample board) and (B) with a thru sample board. (C) Normalized time-domain reflection (ref) and transmission (blue) THz signals for an "open" circuit (top) and with a mounted thru (bottom). The transmitted signal in the open circuit comes from free-space radiative coupling. (D) Non-normalized Fourier amplitudes for transmitted THz signals, showing their ratio to be order unity. Note the multipliers in panel C.

than in transmission. Even though radiative couplings should be removed by referencing operators, it is preferable to measure in only reflection than only transmission.

5.2.5 Calibration

Calibration requires removing all systematic error propagators, which are all of the parameters in Fig. 5.2 except for $S_{11(B)}$ in reflection or $S_{21(B)}$ in transmission from a measurement at one or both of the outputs. Computing the signal flows in reflection mode, to output(1), involves E_D, first sample board - switch board reflection, sample reflection, second sample board - switch board reflection, internal Fabry-Pérot reflections, and Fabry-Pérot reflections associated with the radiative coupling. Assuming lossless board interfaces and incorporating radiative loss E_{R1} in T_1, one must remove 7 independent

systematic error propagators before measuring the sample. This requires 7 independent calibrants, which are both frequency and temperature dependent.

In frequency domain network analyzer calibration, this is done with short circuits, open circuits, well-calibrated loads, match, reflect, and thru standards (Ch. 1). These standards are designed to have well-known, frequency-independent impedances. Measuring with several standards produces a system of equations that can be solved for the error propagators, which can be mathematically removed from subsequent measurements. Such calibration is considerably more difficult at THz frequencies, as it is difficult to make true standard calibrants. For example, using a terminated transmission line as an open circuits radiates at THz frequencies, meaning the open circuit calibrant does not have a well-known impedance. The small skin depths and anomalous skin effect at THz frequencies and cryogenic temperatures mean that short standard calibrants have considerable kinetic inductance with a significant frequency dependence. This can be approximated by a lateral metal strip connecting CPS conductors, but would be best with a vertical metal wall. Frequency-independent load and match standards are nearly impossible to design. These standards may be simulable in finite element software, but at the expense of convenience. The slightly different switch board - sample board interface for each calibrant would require additional assumptions to remove, greatly lowering the accuracy of standard calibration.

It should go without saying that full frequency-dependent and temperature-dependent calibration is impossible. Ad hoc short and open calibrants may work as a system check, but are insufficient to perform full 7-parameter calibration. In addition, the fast sample interchange architecture means that each of the board interface terms will change with every new sample or calibrant.

Calibration is greatly simplified by the spectrometer's time domain nature. The temporal separation between the directivity, the board interfacial reflections, and sample

reflection allow one to isolate each pulse with the windows introduced in Ch. 4. Unless one wants to precisely measure the error propagators, one no longer needs to resolve all error terms. Increasing the transmission line lengths associated with T_1 and T_2 allow for longer windowing functions (and, therefore, lower frequency cutoffs), but lead to smaller signals.

The directivity, first interfacial reflection (which we will ignore), and sample reflection can easily be identified. The simplest calibration makes use of the directivity. Following Fig. 5.2, only E_D contributes to the directivity. The sample reflection, without Fabry-Pérot oscillations, is composed of T_1, $S_{21(A)}$, T_2, $S_{11(B)}$, and related propagators. One can use the shunt reflection coefficient derived in Ch. 1 and write the reflected signal $R(\omega)$ using the directivity $D(\omega)$ as:

$$\begin{aligned} R(\omega) &= D(\omega)T_1S_{21(A)}T_2S_{11(B)}T_2^*S_{12(A)}T_1^* + D(\omega)\xi \\ R(\omega) &= D(\omega)e^{-\gamma_1 l_1}S_{21(A)}e^{-\gamma_2 l_2}\frac{-Y(\omega)Z_0}{2+Y(\omega)Z_0}e^{\gamma_2^* l_2}S_{12(A)}e^{\gamma_1^* l_1} + D(\omega)\xi \\ R(\omega) &= D(\omega)e^{-2(\alpha_1 l_1-\alpha_2 l_2)}S_{21(A)}S_{12(A)}\frac{-Y(\omega)Z_0}{2+Y(\omega)Z_0} + D(\omega)\xi \end{aligned} \tag{5.4}$$

We have allowed an additional spurious reflection term $D(\omega)\xi$, which can come from scattering from any impedance mismatch or internal Fabry-Pérot oscillation. For example, the 12 ps time domain signal in Fig. 4.3, common to all temperatures, is a Fabry-Pérot reflection between the first interface and the directional coupler. such $D(\omega)\xi$ can be avoided by choosing making a thru connector, identifying time domain features, and placing your sample at a place to avoid such features.

We simplify by $P \equiv S_{21(A)}S_{12(A)}$ and $e^{-2\alpha l} \equiv e^{-2(\alpha_1 l_1+\alpha_2 l_2)}$. Writing eqn. 5.4 for the normal (N) and superconducting (S) states and assuming that the interfacial and transmission line propagators are temperature independent, one can solve the ratio of reflected signals (R_N/R_S) for the sample admittance as:

$$Y_s(\omega) = \frac{-2D_nPY_nZ_0 + 2e^{2\alpha l}(2 + Y_nZ_0)(D_s - D_nR_s/R_n)\xi}{PZ_0(D_nY_nZ_0 - D_s(2 + Y_nZ_0)R_n/R_s) - e^{2\alpha l}Z_0(2 + Y_nZ_0)(D_n - D_sR_n/R_s)\xi} \tag{5.5}$$

Eqn. 5.5 simplifies to eqn. 4.2 if $\xi \to 0$. The moral: choose your sample location carefully to avoid spurious reflections!

Similar expressions can be derived for transmission, using the lumped shunt transmission coefficient. If one simultaneously measures reflection (R) and transmission (T) in the normal and superconducting states, one can show:

$$(R_n/T_n)/(R_s/T_s) = Y_n/Y_s \tag{5.6}$$

This is a key result: all error propagators, even the directivity, are irrelevant if one simultaneously measures R and T, and references the reflection to the transmission. Note that different sensitivities (different LIVs) for the A and B switches would manifest as a scaling between R and T, which will also be cancelled out in eqn. 5.6. The beautiful cancellation suggests referencing the reflection to the transmission is the most elegant and robust method for identifying the sample admittance, if one's optics can easily accommodate it.

5.3 Best operating conditions

5.3.1 Pre-pulses

Wide-area photocurrent scans of the A-switch shows two local maxima (Fig. 5.7A). One spot (1) has small area and large photocurrents; one spot (2) has a large area and ~ 5 times smaller photocurrents. The transients corresponding to these two spots are

shown in Fig. 5.7B. The transient from spot 2 occurs 11 ps before the transient from spot 1. In addition, the spot 2 transient is about x10 smaller than the spot 2 transient. The spot 2 transient is a pre-pulse - which has been known to occur in other THz TDS systems. THz pulses can shift to smaller delays if the A-switch beam traverses a longer optical path. The free space optical path length corresponding to the shift is 3.3 mm, which has round trip length 1.65 mm. The COP's transparency allows a nearly-aligned beam to reflect from the cold finger and excite the underside of the A-switch. The COP switch board is mounted about 1.6 mm above a reflective cold finger. The laser is focused reasonably well at spot 1. Poor focus from the additional 3.3 mm of propagation cause the decreased peak current and broadened pattern at spot 2.

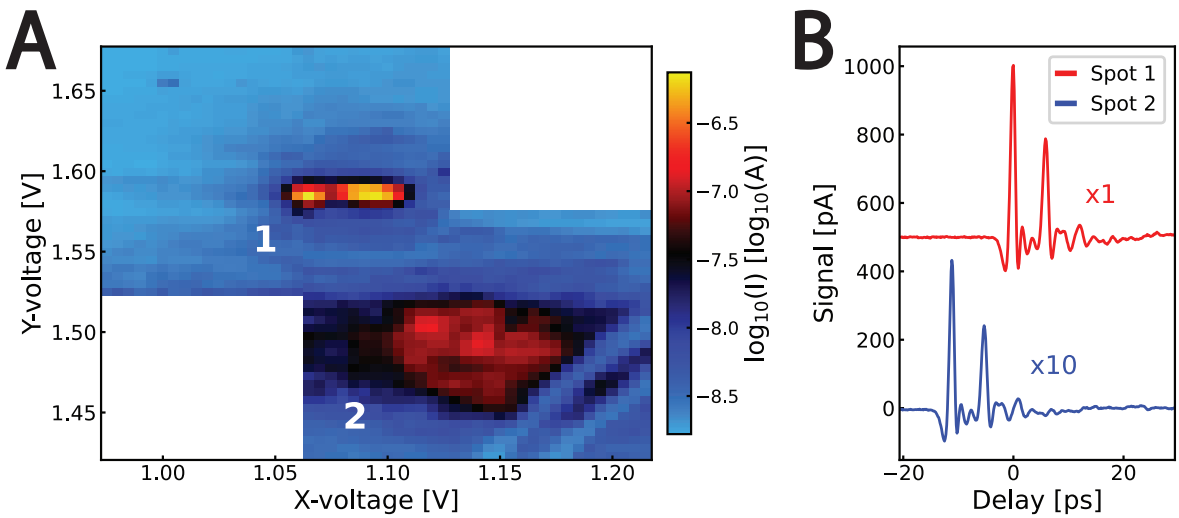

Figure 5.7: Origin of pre-pulses. (A) Log-scaled photocurrent scan, showing two local maxima (1: localized, 2: spread-out) (B) A-switch transients associated with aligning the laser to spots 1 and 2. No sample board is mounted to the switch board. Note the x10 multiplier on the transient from spot 2.

In summary, align (at first) by taking large area scans and identifying the primary and second alignment locations. Adjust optics so that the main switch illumination doesn't also excite a pre-pulse - which complicates the analysis by adding a second directivity pulse.

5.3.2 Optical power and focus

We illuminate the emitter and A/B switches with 3 mW. Higher power does produce larger signals, but also risks damage to the spectrometer. High fluence can cause irreversible breakdown in the LT GaAs; high powers can locally melt the COP switch board. Long term degradation can deform the substrate and lead to lower LIV and THz currents. We regularly exposed one set of emitter and A switches to powers exceeding 5 mW, which were damaged by over the course of a few months. Though THz signal amplitude temporarily increased, the switch DC and THz currents decreased to 20% of their original operating levels. Local damage to switches from long term, high power use can be seen in micrographs in Fig. 5.8.

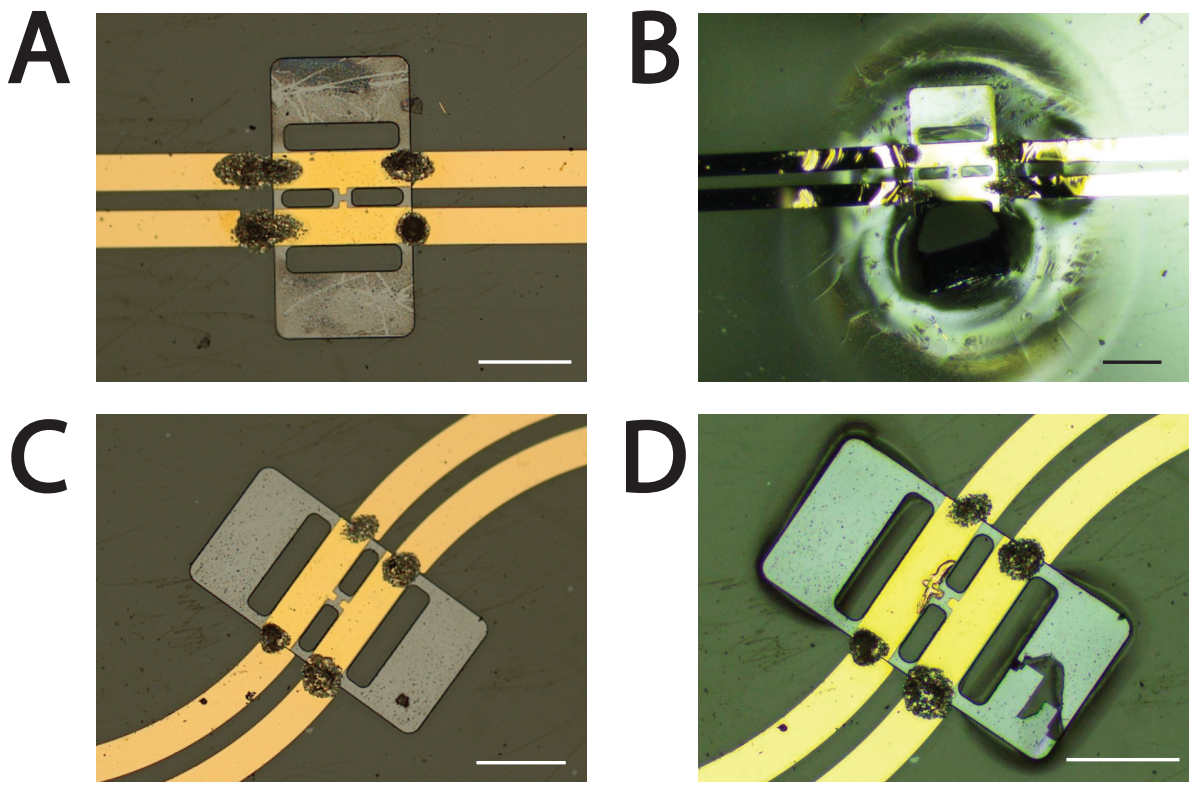

Figure 5.8: Photoconductive switch degradation from long-term expose to high power. (A) before and (B) after images of the B-switch; (C) before and (D) after images of the emitter switch. All scale bars correspond to 130 μm.

High power operation melted and deformed the thin COP. Deformations in the substrate can put the switches under compressive or tensile strain and eventually cause their long term failure. The switches in panels B and D are barely useable.

In summary, use MHz repetition-rate lasers to keep the fluence low and use less than a few mW to avoid locally deforming the polymer substrate.

5.3.3 Best biasing conditions

The spectrometer operates with a DC bias applied across the emitter and a DC bias applied across the A/B switches for automated alignment and monitoring. The applied DC bias and measured LIV slope determines the amplitude of the emitted THz signal. The manufacturer does not recommend operating these PC switches at biases larger than 10V. The bias conditions applied to the A/B switches, however, could be arbitrary. The biases (and the corresponding current) is supplied by a Keithley sourcemeters, which have easily tunable biases and bias ranges. Here, we show the that the dynamic range is maximized for small readout biases with large sourcemeter ranges.

The emitter was fixed to constant bias (10V) with constant range. The A/B/ sourcemeter bias (with range 10 μA) was varied in half orders of magnitude from 1 mV to 10V, and the resulting THz transient was measured (Fig. 5.9). One can see, in both the time domain and the frequency domain, that the signal is large and the noise is small for biases $\leq$0.1 V. As the bias increases to 1V and above, the signal amplitude decreases and the noise floor simultaneously rises, causing the THz (AC) dynamic range to degrade. It is likely that the decreased signal amplitude at high bias comes from nonlinearity in the LIV.

One can also measure dynamic range of the DC feedback signal, which is given by the ratio of the measured DC feedback signal to its fluctuation. This dynamic range

drops to nearly zero at 1 mV, as this is approximately the photoconductive switch's built-in potential and thus the feedback is almost entirely fluctuation about zero. The experiment should be operated at the bias where the dynamic ranges of both the AC and DC dynamic ranges are maximized. This occurs when their product is maximized, which occurs between 100 mV and 1V.

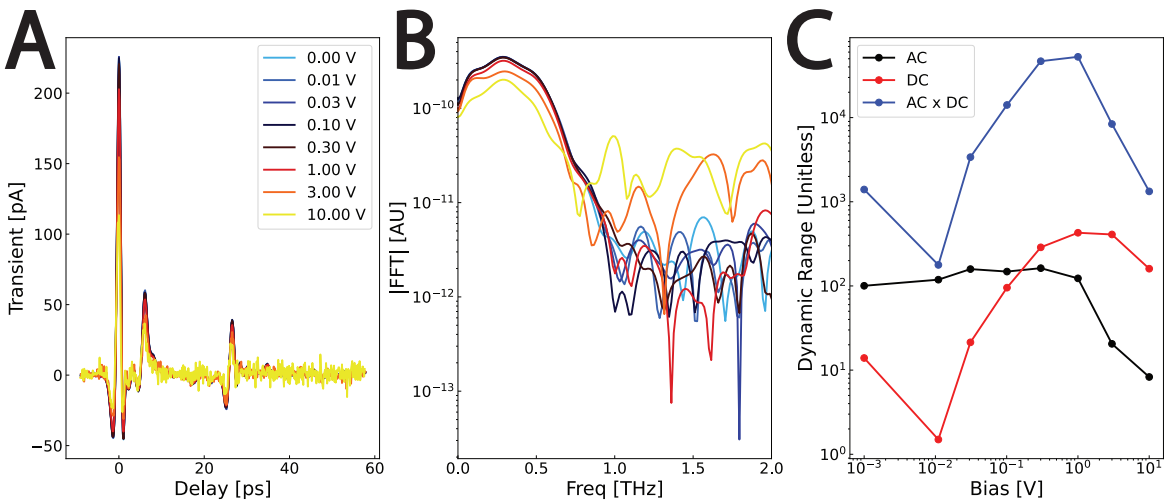

Figure 5.9: THz signal quality for varying A switch bias at constant range (10 μA), with the emitter at constant bias (10V) and range (10 μA). (A) Time domain transients (B) Fourier transform (C) Dynamic range for AC (THz) transient, DC sourcemeter readout, and their product.

Different sourcemeter ranges have different internal circuitry, and a different input impedance to ground. The emitter bias (10V) and range (10 μA) was subsequently fixed, the A/B bias was fixed (100 mV) and the bias range was swept from 0.01 μA to 100 μA (Fig. 5.10). The THz dynamic range is small with the 10 mV range, which monotonically increases with voltage range. The dynamic range of the feedback, however, decreases nearly monotonically with range. The product of the DC and AC ranges is maximized at 10 μA, which, therefore, is the best operating range.

In summary, the A/B line should be biased with 100 mV at 10 μA range.

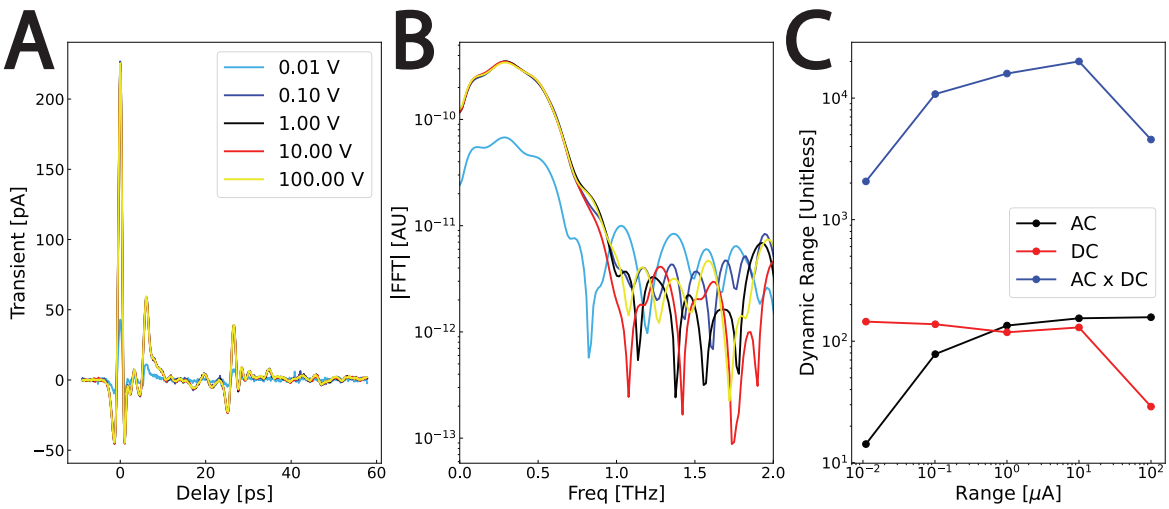

Figure 5.10: THz signal quality for varying A switch sourcemeter bias range, with the emitter at constant bias (10V) and range (10 μA). (A) Time domain transients (B) Fourier transform (C) Dynamic range for AC (THz) transient, DC sourcemeter readout, and their product. This analysis follows the same structure as Fig. 5.9.

Chapter 6

Outlook

"When you think you know everything, you get a Bachelor's degree. When you realize you know nothing, you get a Master's degree. When you realize that no one else knows anything either, you get a doctorate."

R. Verhoef

The fast sample exchange spectrometer detailed in this dissertation is capable of resolving low energy excitations in materials with energy scales between 200 and 750 GHz. A few minor engineering improvements, including adding robust DC transport contacts and increasing the distance between A-switch and switch board / sample board interface to improve low frequency operation could easily enable the 100 - 800 GHz spectroscopy of gated vdW heterostructures. An obvious candidate for measurement are vdW heterostructures of gated graphene monolayers and bilayers. One may expect to see THz-frequency plasmon physics and monopolar hydrodynamic plasmons in hydrodynamic regimes [78]. One could also probe quantum criticality in the linear-in-temperature

resistivity [184, 185] in the superconductor's parent phase in magic angle twisted bilayer graphene [186], which can reveal a quantum-limited scattering rate [187].

The future of on-chip THz extends to 2D materials beyond simply graphene. THz frequency magnons in magnetically-ordered chromium trihalides ($CrBr_3$, CrI_3) could be directly probed [188, 189, 190]. The energy spacing between energy levels in many Hubbard excitons lies in THz bandwidths - one could imagine pursuing the details of the intraexcitonic structure of transition metal dichalcogenide heterostructures [191, 192, 193].

The capacitive coupling between transmission lines and sample in on-chip THz also suggests that DC transport contacts are not required. At the time of writing, the community struggles to make Ohmic contacts to TMDs with electron doping due to a lack of contacts. In addition, making contact to many materials is stymied by a fast-growing, self-limiting oxide or a self-limited layer of unintentional oxygen doping. Making DC electrical contact to, for example, the famous high temperature superconductor $Bi_2Sr_2CaCu_2O_{8+\delta}$ can require expensive and complicated setups involving exfoliation at liquid nitrogen temperatures and cold welding [194]. These issues can by bypassed by THz measurement. Assuming that the self-limited oxygen-doped layer is THz inactive, THz excitation simply passes through it, allowing for low frequency (100 GHz $\ll 2\Delta(0)$) transport measurement. Superconductivity has clear THz-frequency real and imaginary response functions, so why stop with $Bi_2Sr_2CaCu_2O_{8+\delta}$? DC contact-free measurements targeting the exotic pairing in magic angle twisted $Bi_2Sr_2CaCu_2O_{8+\delta}$ [195] or transition metal dichalogenides [196] are achievable in the near term.

It should also be noted that on-chip THz spectroscopy need not be restricted to vdW materials. The THz spectrum of any material or device fabricated to be smaller than the diffraction limit is best captured by on-chip techniques. To illustrate this, consider Weyl semimetal Cd_2As_3, where both topological edge states and non-topological bulk carrier physics contribute to THz absorption [197]. The conductivity contribution of carriers in

the bulk could be mitigated by reducing sample thickness, but this makes free space THz measurement difficult (Ch. 3). Using on-chip THz spectroscopy could be used to study only the edge physics of samples with a minimal density of bulk (< 10 nm) carriers.

Another exciting avenue is ultrafast transport measurement of Floquet states induced by a strong laser pump. Pioneered by McIver et. al. [79], examining optically driven out-of-equilibrium superconductivity [82], charge order [198], and flatband physics [199] is a nascent field ripe for high-performing on-chip THz spectrometers. With only a few atomic layers, there can be no pump / probe depth mismatch. This greatly simplifies the analysis and may make Floquet features more clear. In addition, increasing electric field strengths by, say, tapering the CPS's W to 30 nm (instead of 30 μm) will increase the field strengths by a factor of 1000. Though this will complicate THz engineering principles, it may allow access to nonlinear on-chip THz TDS useful to spin liquid research [200].

Despite challenging engineering considerations at THz frequency operations, the high-throughput spectrometer fabricated, tested, and used here is a convenient and robust platform to chase many important scientific and technological goals. I, personally, am thrilled to be contributing member in a field that has much to offer.

Appendix A

Standard Nanofabrication Processes and Travelers

> "If something is hard, don't do it. (Find a more clever way to solve the problem.)"
>
> *L. A. Cohen*

The following are processes executed in the UCSB nanofabrication facility. They may be adapted for use in other facilities, but all process parameters listed here are known to work with tools at UCSB. For this section, "s" means seconds and "m" means minutes.

We assume that all wafers have been cleaned by maximum power, maximum frequency sonication in acetone (5m), methanol (2m) and isopropanol (2m) and dried immediately with dry nitrogen gas. This step will be referred to as a solvent clean for all subsequent process listings. Clean solvent glassware with unused solvent only (not DI H_2O - never use a polar solvent to clean residue from a nonpolar solvent or vice versa). Clean developer, acid, and stopdevelop/stopetch DI H_2O beakers with three fills and flushes of clean DI H_2O. Clean HF glassware with five fills and flushes of clean DI H_2O.

Many processes here can be done with alternative tools. A complete list of substitutions is listed at the end of this section.

A.1 rdSi circuit fabrication

Required tools	Wet bench, solvent glassware, developer glassware, ICP2, EB3
Required supplies	300 nm Si on sapphire substrates
Difficulty	Simple
Common pitfalls	Incorrect photolithography dose / focus
Suggested scheduling	4 hours + overnight liftoff

Step	Operation	Tool	Parameter 1	Parameter 2
01	**Ion implantation**	Leonard Kroko Inc	100 keV or 185 keV	O^+ or O^{++} ions
02	**Mesa lith.**			
02A	SPIN	Spin coater	SPR955 0.9	3000 RPM [5]
02B	BAKE	Hotplate	105°C	90s
02C	EXPOSE	Stepper 2	0.35s	—
02D	BAKE	Hotplate	105°C	60s
02E	DEVELOP	AZMiF300	60s	–
03	**Mesa isolation**	ICP2	Etch 174	1.7 Pa
	RF (top): 450W	RF (bot): 50W	Cl_2: 10 sccm	SF_6: 50 sccm
04	**Metal lith.**			
04A	SPIN	Spin coater	SPR955 0.9	3000 RPM [5]
04B	BAKE	Hotplate	105°C	90s
04C	EXPOSE	Stepper 2	0.35s	—
04D	BAKE	Hotplate	105°C	90s
04E	DEVELOP	AZMiF300	60s	–
05	**Metal deposition**			
05A	DEPOSIT	EB3	10 nm Ti	100 nm Au
05B	LIFTOFF	NMP chemistry	4+ hours	80°C

Table A.1: Clean room processing for rdSi circuits. Any UCSB metal evaporation tool can be used for metal deposition. One can also use the Heidelberg MLA150 for optical lithography, although exposure fluence and defocus are not known for rdSi on sapphire.

A.2 Spectrometer on growth substrate

Required tools	Wet bench, solvent glassware, developer glassware, Acid/HF bench, Heidelberg MLA 150, EB3
Required supplies	LT GaAs or ErAs:GaAs on $Al_xGa_{1-x}As$ with x$\geq$ 0.75
Difficulty	Moderate
Common pitfalls	Temperature/stoichiometry-related etch rate drift Mesa isolation etch selectivity $\approx$ 100:1 for x = 0.75 Wafer "upside-down" errors
Suggested scheduling	4 hours (session 1) + liftoff + 2 hours (session 2) + liftoff

Step	Operation	Tool	Parameter 1	Parameter 2
01	**Mesa lith.**	—	—	—
01A	SPIN	Spin coater	SPR955 0.9	3000 RPM [5]
01B	BAKE	Hotplate	105°C	90s
01C	EXPOSE	MLA150	220 mJ	-3 defoc.
01D	BAKE	Hotplate	105°C	60s
01E	DEVELOP	AZMiF300	60s	–
03	**Mesa isolation**	—	—	
03A	ETCH	Acid bench	CA:H_2O:H_2O_2 6:6:1	246 nm/m
03B	STOPETCH	DI H_20	30s	N_2 blow dry
03C	RESIST STRIP	NMP chemistry	4+ hours	80°
04	**Line lith.**	—	—	—
04A	SPIN	Spin coater	SPR955 0.9	3000 RPM [5]
04B	BAKE	Hotplate	105°C	90s
04C	EXPOSE	MLA150	220 mJ	-3 defoc.
04D	BAKE	Hotplate	105°C	60s
04E	DEVELOP	AZMiF300	60s	–
05	**Metal deposition**	—	—	—
05A	DEPOSIT	EB3	10 nm Ti	100 nm Au
05B	LIFTOFF	NMP chemistry	4+ hours	80°
06	**Contact lith.**	—	—	—
06A	SPIN	Spin coater	SF8	2000 RPM [3]
06B	BAKE	Hotplate	200°C	180s
06C	SPIN	Spin coater	SPR955 0.9	3000 RPM [5]
06D	BAKE	Hotplate	105°C	60s
06E	EXPOSE	MLA150	240 mJ	-4 defoc.
06F	BAKE	Hotplate	105°C	90s
06G	DEVELOP	AZMiF300	60s	–
07	**Conformal metal**	Sputter 4	—	—
07A	SPUTTER	tilt 10	Ti 200W	180s ($\sim$ 8 nm)
	3 mTorr	Ar: 45 sccm	300W	V.FDBK $\approx$ 350 V
07A	SPUTTER	tilt 10	Potts Au 200W	300s ($\sim$ 150nm)
	5 mTorr	Ar: 45 sccm	200W	V.FDBK $\approx$ 525 V
07C	LIFTOFF	NMP chemistry	4+ hours	80°C

Table A.2: Clean room processing for III-V on growth substrates circuits. Note that ErAs etches at 246 nm/min, but LT GaAs etches faster: 296 nm/min.

A.3 Epitaxial lift off

Required tools	Wet bench, HF bench, solvent glassware, developer glassware, RTA Acid/HF bench, Heidelberg MLA 150, EB3, vdW stacking station
Required supplies	LT GaAs or ErAs:GaAs on $Al_xGa_{1-x}As$ with x$\sim$ 0.75 HF-safe gloves, face shield, apron, and quartz wafers
Difficulty	Hard
Common pitfalls	Difficulty handling HF PC switch movement following ELO
Suggested scheduling	5 hours (ELO) + 2 days (spectrometer)

Step	Operation	Tool	Parameter 1	Parameter 2
01	**Dicing**			
01A	SPIN	Spin coater	SPR955 0.9	3000 RPM [5]
01B	BAKE	Hotplate	105°C	90s
01C	DICE	Dicing saw	53 mm chuck	2.187-4C-9RU-3
		1 cm x 1cm	35 KRPM	<2 mm/s
01D	RESIST STRIP	NMP chemistry	4+ hours	80°
02	**Mesa lith.**			
02A	SPIN	Spin coater	AZ5214 (N)	6000 RPM [9]
02B	BAKE	Hotplate	95°C	60s
02C	EXPOSE	MJB3 (v8)	5s	Z = 13.6
02D	BAKE	Hotplate	110°C	60s
02E	FLOOD EXPOSE	MJB3	60s	—
02E	DEVELOP	AZMiF300	45s	—
03	**Mesa isolation**			
03A	ETCH	Acid bench	CA:H_2O:H_2O_2 6:6:1	250 nm/m
03A	STOPETCH	DI H_20	30s	N_2 blow dry
04	**Epitaxial liftoff**			
04A	ETCH	HF bench	HF 5%	210 min / 300 nm
04B	COLLECT	Silicon	Scoop from HF	—
04C	STOPETCH	DI H_20	DI 1 (1m)	DI 2 (1 min)
04D	CLEAN	Kimwipe dab	Gentle dry N_2	—
04D	ANNEAL	RTA	300°C	20 m

Table A.3: Clean room processing for preparing epitaxially-lifted off (ELO) III-V PC switches. Many spin coaters shatter wafers at rotational rates > 3000 RPM - clean the chuck in acetone (remove the O-ring first) and IPA before use. Mix citric acid (CA) / hydrogen peroxide brew by carefully dispensing two 5mL pumps of H_2O_2. Measure 60 mL of citric acid:H_2O 1-1 (labelled "CAW 1-1") and slowly add to H_2O_2. This reaction is slightly exothermic, so go slowly. Pour the brew from one beaker to another at least four times before using. Dispose with a plenum flush. Use extreme caution with HF or BOE solutions (and polypropylene beakers). The rapid thermal anneal (RTA) is used to remove α-As redeposited during stopetch removal.

Step	Operation	Tool	Parameter 1	Parameter 2
01	**Metal dep.**			
01A	DEPOSIT	Therm. Evp. 1	20 nm Cr	
02	**Line lith.**			
02A	SPIN	Spin coater	SPR955 0.9	3000 RPM [5]
02B	BAKE	Hotplate	105°C	90s
02C	EXPOSE	MLA150	220 mJ	-3 defoc.
02D	BAKE	Hotplate	105°C	90s
02E	DEVELOP	AZMiF300	60s	–
03	**Metal. dep.**			
03A	DEPOSIT	EB3	10 nm Ti	100 nm Au
05B	LIFTOFF	NMP chemistry	4+ hours	80°C
04	**Contact lith.**	—	—	—
04A	SPIN	Spin coater	SF8	2000 RPM [3]
04B	BAKE	Hotplate	200°C	180s
04C	SPIN	Spin coater	SPR955 0.9	3000 RPM [5]
04D	BAKE	Hotplate	105°C	90s
04E	EXPOSE	MLA150	240 mJ	-4 defoc.
04F	BAKE	Hotplate	105°C	90s
04G	DEVELOP	AZMiF300	60s	–
04B	LIFTOFF	NMP chemistry	4+ hours	80°C
05	**Conformal metal**	Sputter 4	—	—
05A	SPUTTER	tilt 10	Ti 200W	180s ($\sim$ 8 nm)
	3 mTorr	Ar: 45 sccm	300W	V.FDBK $\approx$ 350 V
05A	SPUTTER	tilt 10	Potts Au 200W	300s ($\sim$ 150nm)
	5 mTorr	Ar: 45 sccm	200W	V.FDBK $\approx$ 525 V
05C	LIFTOFF	NMP chemistry	4+ hours	80°C
06	**Metal etch**			
06Ai	WET ETCH	Acid bench	1020 Cr Etchant	15s
06Bi	STOPETCH	Acid bench	Di H_2O	>30s
06Ci	PLEN. FLUSH	NMP chemistry	4+ hours	80°C

Table A.4: Clean room processing fabricating ELO spectrometers on low-loss substrates. The sputter4 Ti recipe uses 300W, but is permanently mislabeled "Ti 200W". See traveler A.5 for alternative dry metal etching processing.

Step	Operation	Tool	Parameter 1	Parameter 2
06ii	*Alternative*			
06Aii	DEPOSIT	EB3	10 nm Ti	100 nm Au
06Bii	LIFTOFF	NMP chemistry	4+ hours	80°
06Cii	DRY ETCH	ICP2	Etch 137 (2 min)	?? Pa
	RF (top): ??	RF (bot): ??	Cl_2: ?? sscm	SF6: 50 sccm

Table A.5: Alternative clean room processing for Cr etch. The selectivity of the dry Cr etch is high, but the exact value unknown. Dry Cr etch is superior to wet etching to avoid undercutting, but takes longer (as ICP2 is more difficult to book) and the selectivity/etch rate are not precisely known. The listed parameters on ICP2 are unknown at the time of writing (indicated by "??").

A.4 Thin quartz wafer bonding

The following was largely developed by Biljana Stamenic, Demis John, and Abhay Nayak.

Thin substrates ($\leq 100\mu$m quartz substrates) are very fragile and will not survive most nanofabrication processes without cracking. Their structural integrity in processing can be immensely improved by bonding them with carrier silicon wafers. I use 275 μm thick Si carrier wafers. There are three requisite processes: (1) crystal bonding a thin quartz wafer to a carrier wafer, (2) crystal bonding a diced quartz piece (several mm by several mm), and (3) removing the carrier wafer. Note that handling small, fragile samples is greatly simplified by using vacuum tweezers instead of normal, soft-tipped tweezers[1]. Closely monitor the crystal bond during the heating, as the quartz can easily slide off and become very difficult to reposition.

Wafer-bonded 50 μm quartz on 275 μm Si should be treated as having an average thickness of 400 μm in other tools, including the dicing saw. Dice wafer-bonded quartz with silicon recipes (Thermocarbon 2.187-4C-9RU-3, 35 KRPM, 1-2 mm/s).

Required tools	Wet bench, solvent glassware
Required supplies	Thin quartz, carrier wafer, crystal bond, mass balance
Difficulty	Moderate
Common pitfalls	Insufficiently flat crystal bond for lithography, cracked quartz boards inability to see thin, transparent quartz in NMP chemistry
Suggested scheduling	3 hours for bond, 6 hours for bond removal

[1]Though vacuum tweezers can easily become dirty and are hard to clean.

Step	Operation	Tool	Parameter 1	Parameter 2
01i	**Wafer bonding**			
01Ai	ASH	300 mTorr O_2	100W	60s [5]
01Bi	COAT	Hotplate	Al foil	
01Ci	BAKE SI	Hotplate	120°C	100s
01Di	CRYS. BOND 509	Place on Si	0.14g	5m
01Ei	RAMP TEMP	Hotplate	$120°C \rightarrow 220°C$	30m
01Fi	RAMP TEMP	Hotplate	$220°C \rightarrow 60°C$	
01Gi	ANNEAL	Hotplate	220∘C, bake 3m, then 60∘C	As needed
01ii	**Piece bonding**			
01Aii	PREPARE	Hotplate	Si carrier piece	in weighboat
01Bii	BAKE	Hotplate	120°C	100s
01Cii	CUT	Crystal bond 509	Few flakes on Si	distributed
01Dii	BAKE	Hotplate	120°C	100s
01Eii	RAMP TEMP	Hotplate	$120° \rightarrow 220°$	30m
01Fii	RAMP TEMP	Hotplate	$220° \rightarrow 60°$	
01Gii	ANNEAL	Hotplate	220∘, bake 3m, then 60∘	As needed
02	**Bond dissolution**			
02A	LIFTOFF	NMP chemistry	4+ hours	$80°C$
02B	REMOVE	Vacuum tweezers	Acetone rinse	IPA rinse

Table A.6: Fragile wafer/piece bonding and bond removal traveler.

A.5 Additional information

A.5.1 Best lithographic practices

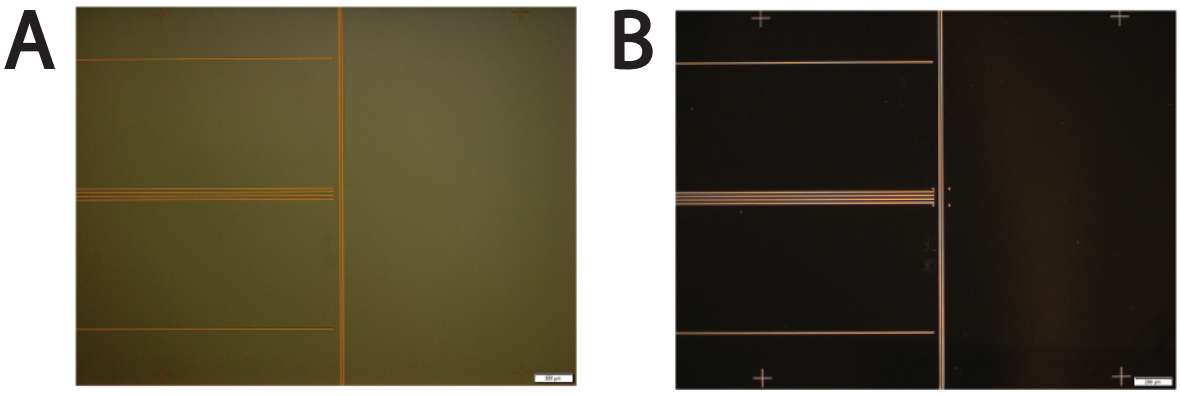

Figure A.1: Comparison of one transmission line device taken as standard and DFIC optical micrographs. (A) Standard microscopy and (B) DFIC microscopy. Features and contamination barely visible in optical microscopy are clear with DFIC microscopy.

Dose testing should always be done if precise (feature size $\leq$1.5 μm) are required. Dose testing should be done with cleanroom-provided patterns, which are used to measure the horizontal/vertical skew, drift, and feature resolution limits. It is wise to metallize dose tests, as the proper metallic features, not photoresist features, are the process target. At the least, inspect dose tests under differential (field) interference contrast microscopy (DFIC microscopy). DFIC microscopy produces an image with contrast from the gradient of the optical path length (feature height), making it very sensitive to patches of undeveloped photoresist. An example of a device micrograph taken with and without DFIC microscopy is shown in Fig. A.5.1.

A.5.2 Photolithography on transparent substrates

Photolithography on transparent substrates can be done in one of four ways. Each has advantages and disadvantages. The following list sorts simplest / least precise to

most complicated / most precise.

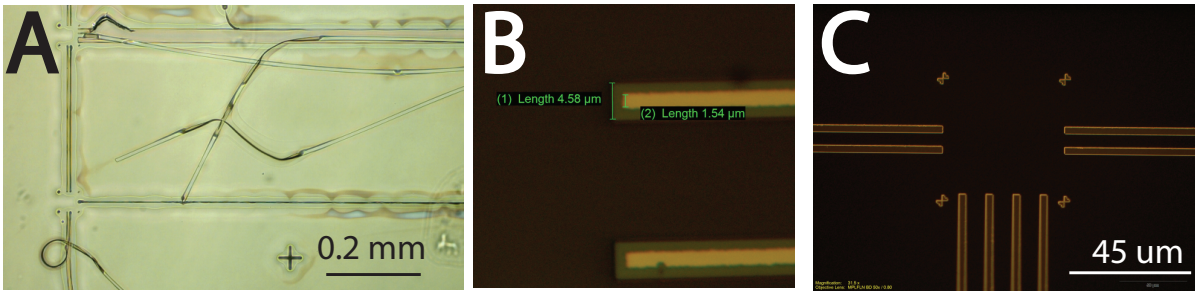

Figure A.2: Nanofab on transparent substrates (A) developed optical lithography on 625 μm thick quartz, performed on stepper 2. Features are so broadened that strips of undercut photoresist are lifting off. This substantially improves by using the Heidelberg MLA150. (B) Purely wet-etched Cr/Au with a target feature size of 4.5 μm. Feature size has been reduced to 1.5 μm by uncontrolled interfacial wet etching (C) dark field image of hybrid wet etching / liftoff processing, showing pristine features.

1. **Straightforward liftoff**

   Treat the transparent substrate as if it were an opaque substrate. Perform positive photolithography, deposit metal, and perform liftoff. Fabry-Pérot oscillations from the sample underside / substrate-holding chuck can cause pattern broadening. This is not advised on stepper 2 and not advised for features more precise than 3 μm in the MLA150. A sample exposure from stepper 2 is shown in Fig. A.5.2A.

2. **Pure wet subtractive processing**

   Deposit metal uniformly on the substrate surface. Perform negative photolithography (suggested negative photoresist: AZ5214). Use wet gold etchant (TFA, 2.8 nm/s at 25°) to etch the gold, and wet Ti etchant (Transene TFTN, 35 nm/s) or Cr etchant (Transene 1020, 4.0 nm/s) to etch the sticking layer. Etches generally proceed faster at interfaces. The lateral gold etch and chrome undercut are difficult to manage when the feature size becomes approximately 5μm (see Fig. A.5.2B), which is why this technique was abandoned by the semiconductor industry [201].

Note: all etch rates here are specified at 25° and double with every 10°. I recommend taking the temperature before using wet etchants.

3. **Hybrid wet subtractive / liftoff process**

   A compromise between the two previous processes is to first deposit uniform sticking layer on metal. A typical exposure wavelength is 365 nm, where the anomalous skin effect controls the frequency-independent, non-exponential electric field versus depth (Ch. 3). A good rule of thumb is that the metallized substrate should be more opaque than transparent under visible light ($\approx$ 20 - 30 nm Cr). Perform photolithography on the sticking layer, deposit another sticking layer and the top metallization (Au), and perform liftoff. Then, wet etch the sticking layer. The very thin sticking layer etches before the gold features can be undercut. I have consistently used this hybrid technique to produce 1 μm features with high yield. An example is shown in Fig. A.5.2C.

   One can also remove chrome sticking layers by using ICP2 with etch #137. There remain some unknown parameters, as I have never measured the selectivity to Ti and Au and the undercut rates. I have never found a use case in which wet etching Cr is insufficient and dry etching is required instead.

4. **Pure dry subtractive processing**

   Deposit metal uniformly on the substrate surface. Perform negative photolithography on the metal. Etch using an ion mill. This process is most precise and completely avoids the use of wet solvents, but I have never personally tested it, measured the selectivity, or optimized the parameters. I have never found a use case that could not be completed by one of the three aforementioned techniques.

A.5.3 Epitaxial liftoff and transfer

The epitaxial liftoff (ELO) of LT GaAs or ErAs:GaAs from UID GaAs substrates requires dissolving $Al_xGa_{1-x}As$ in concentrated HF or buffered oxide etch (BOE), which is acutely toxic in miniscule doses. The majority of the lifted-off mesas float or sink in HF, making their retrieval difficult. There are several approaches to collecting and cleaning the mesas.

One approach is to simply ignore the >90% of mesas in solution and collect only the mesas that vdW bonded to the UID GaAs substrate during ELO. Be sure to hold the substrate parallel to floor and move slowly when lifting it from the HF beaker to ensure that the maximum number of switches stay on the substrate. Dry with nitrogen with the gun held at a far distance. Many of these mesas are strongly bonded to the UID GaAs, making their transfer impossible and furthering lower yields. This method has the advantages of simplicity and is the safest, but has the lowest yield.

Another approach is to clean and ash (O_2 plasma, 300 mTorr, 100W, 30s) a chip of bare silicon. *Very* carefully, scoop the floating PC switches out of the HF beaker. Clean them with a very gentle rinse of DI H_2O, patting them dry with a lint-free wipe or drying with low pressure dry nitrogen. The active side of the switches will be cleaned in solvents before incorporation as spectrometer. This method is advantageous in that I find that it has the highest yield, but requires the most caution and skill in handling HF.

There are additional approaches. One can melt a thin layer of black wax onto mesas before placing them in HF. Large clumps of wax should float to the top of the HF beaker, where they can be scooped out with cleaned silicon. The resulting wax clumps are then placed in toluene, which dissolves the wax. I find that the mesas tend to float away in toluene, and that the dark colored wax makes them harder to collect than the HF. This method is comparatively safe, but difficult to execute and, in my experience, has the

lowest yield. In addition, the large strain in ErAs can close the etch channel when the mesas are undercut, making ELO infeasible.

ELO can deposit amorphous arsenic on the PC switch mesas. Amorphous As have a distinct Raman signature, and are visually identifiable as charcoal color. Amorphous As will be removed after rapid thermal annealing at 300°C for 20m. The mesas then need to be transferred to patterned transmission lines. This can mostly easily be done by using a vdW stacking setup. Polymer dry transfer is performed with polycarbonate (PC) films [131] and can align switches with few micron precision. I have also successfully transferred mesas using only the vdW attraction between a mesa and the tip of a probe station probe. The probe station has more uncontrolled variables than the polymer transfer, but can be done very quickly.

A.5.4 Auxiliary mask structures

Optical masks contain the basic transmission lines and alignment markers, but also should include a number of other structures. Such structures are useful for simplifying nanofabrication and troubleshooting material problems. Representative images of the useful diagnostic structures and dicing marks are shown in Fig. A.5.4.

Alignment crosses should be placed in key mask locations, in the event that one needs to add or repair features at a later time. Sets of four coarse crosses (2 μm wide, 45 μm long) for coarse alignment (rectangular pattern, side length of 2 mm) and sets of four fine crosses (2 μm wide, 8 μm long) for fine alignment (rectangular pattern, side length of 100 μm) should be placed so their centers are at the target feature center. To avoid reflection and transmission effects from alignment marks, they should be placed at least three times the CPS gap W. One should also add four "L" shaped dicing marks at the edge of the (rectangularly-distributed) features, which later makes aligning dicing saw cuts straight-

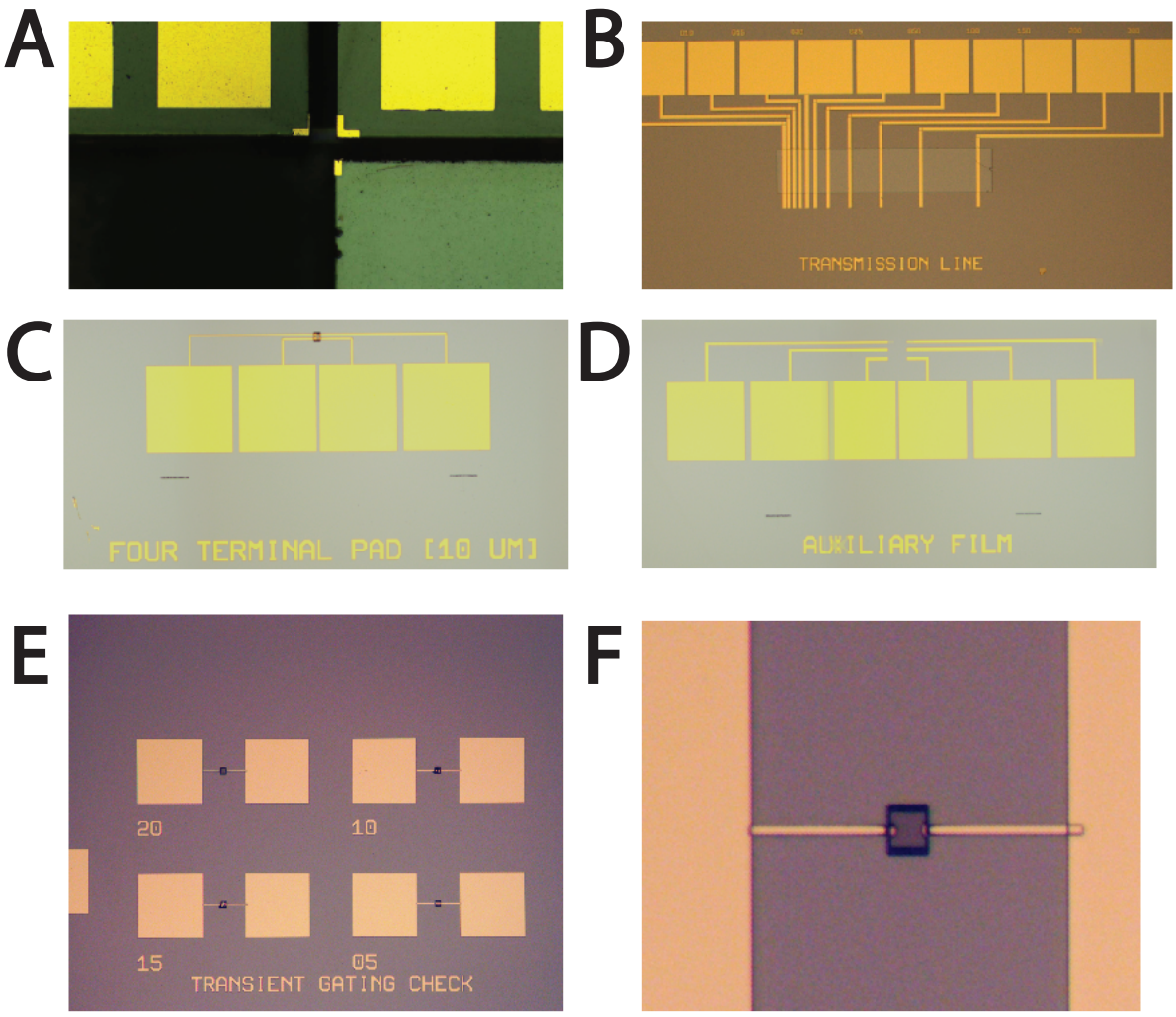

Figure A.3: Optical micrographs of useful mask structure. (A) Dicing marks on the top left, top right, and bottom right, showing cut trenches immediately after dicing. (B) Transmission line structure, used for estimating contact resistance and mesa resistivity. The structure is the standard used in the semiconductor industry (without the Schottky contacts) [1] and consists of increasingly separated electrodes. (C) Four terminal contact to a III-V PC switch mesa. (D) four terminal contacts for auxiliary (sample) films. (E) PC switches separated by decreasing distances, labelled in microns. These are used to measure the ideal PC switch size. (F) Zoom in of the 20 μm structure from panel E.

forward. Expect dicing saw cuts made with quartz dicing blades to cut 230 μm wide dicing lanes, removing any material contained thereon. Write down the XY coordinates of all alignment and dicing structure to easily access them while in the cleanroom. I also recommend a metallized label for each sample, for quick sample identification and sample orientation identification. If your samples are symmetric, include a macroscopic (no microscope required), symmetry-breaking feature in a corner for easy identification without a microscope.

Diagnostic structures should also be included on masks. They are used to assess material and contact quality. One should include an ungated transmission line structure to extract the contact and mesa resistivities. A set of four contacts to the mesa can be used to measure contact the contact Schottky barrier and estimate mesa resistance. If the sample has a low resistance, make the length of each electrode the same. A include four square PC switch mesas of with side lengths from 5 - 20 μm. Each PC switch mesa is contacted by two electrodes, which are connected to wirebonding pads. These structures can be used to estimate the two terminal resistance of an PC switch in the dark or illuminated conditions as a function of PC switch size. I also include a structure for measuring the four terminal resistance of an auxiliary structure (such as a NbN thin film), though this is not critical to assessing performance.

A.5.5 Suitable alternate tools

Metal electron beam evaporation tools (EB) 1, 3, and 4 at UCSB can all be used for simple metal deposition. Thermal evaporator 1 can also be used, except for the deposition of chrome. EB1 and EB4 provide the best liftoff. EB1 is manually operated but rarely booked; EB4 is well automated but always booked. EB3 is sufficient for most metal layers.

If there is a shortage of AZ300MiF, CD26 can be used without process parameter modification. Dilute (<1% TMAH) can also be used, but timing needs to be tested before use. In addition AZ NMP (1165) can be used interchangeably with PG remover for the processes listed here.

Stepper 2 and the Heidelberg MLA150 can both be used for photolithographic processes. The doses and defocus values listed here are subject to change with the tool cycling, but should offer a baseline. Note that stepper2 is faster, but operation requires a mask and is more difficult; the MLA150 is slower, but requires no mask and is very simple to use. I recommend using the MLA150 unless devices need to be produced en masse.

Most deposited metals contain and adhesion layer (Ti, Cr, or Ge) and the top metal (Au or Pt). The adhesion layer improves adhesion between the substrate and subsequent metals. I have never observed differences in THz performance between Ti- and Cr-sticking layer samples. There is a nanofab reason to prefer Cr as the first sticking layer. If ones deposits Ti, removes the sample from the evaporator, places the sample back in the evaporator, and deposits metal a second time, additional sticking material should be deposited the second time. Removing the sample from the evaporator exposes the metal to air, oxidizing the surface. Top metal layers do not stick to oxidized surfaces. If removed from the evaporator, Ti develops "less sticky" oxides than Cr. That is: Ti adheres better to processed Cr than Cr adheres to processed Ti.

Flood exposure for negative photoresist can be done in the Karl Suss MA6, dedicated flood exposure tool, or with an MJB3 aligner with no mask. I prefer the MJB3, as it is always available and is very easy to use.

Appendix B

Circuit Simulation and Finite Element Computation

"You should always know the answer to a problem before you solve it"

B. Munk and *J. F. Lee*

Though one can write down Maxwell's equations and boundary conditions for realistic problem geometries, solving four coupled partial differential equations subject to the Silver-Müller radiation condition for millions of spatial coordinates is intractable to hand calculation. Such calculations should be done with complicated and often obtuse numerical electromagnetics computer programs. These tools are invaluable to crafting and understanding solutions, but should *never* be blindly trusted.

A common complaint in antenna engineering firms is that early career engineers blindly trust the outputs of finite element or method of moments computations because "it came from a computer". One should *never* simply accept the answer delivered by such software as true. This epidemic is so widespread that some engineers jokingly report that their colleagues would say that the distance from Los Angeles to Cleveland is "500 kg, because that is what my software said."

It is the user's job, as a physicist and/or an engineer, to verify the correctness of solution to the problem, to validate that the solution answers the correct problem, and to communicate the results in the context of nanofabrication, practical limitations, and funding. To do, we must first agglutinate our intuition from elementary electrodynamics from Ch. 1 and from elementary numerical principles, which we now develop.

B.1 Additional Electromagnetic principles

We must develop the vector potential and the radiation condition, both of which will be instrumental for computational electrodynamic calculations.

B.1.1 Potentials and the Helmholtz theorem

Two potentials are defined for convenient use. Electrostatics defines the scalar potential as $\overline{E} = -\nabla\phi(\overline{r})$. This breaks down at non-zero frequency electrodynamics. The vector potential is defined by $\overline{A}$ such that $\overline{B} = \nabla \times \overline{A}$. Maxwell's equations must still be obeyed. Substituting the definition of $\overline{A}$ into divergence of $\nabla \cdot \overline{B} = 0$ law still holds, as the divergence of any curl is zero. Substituting the definition of $\overline{A}$ into Faraday's law, however, produces $\nabla \times \overline{E} = -j\omega(\nabla \times \overline{A})$, which implies that $\nabla \times (\overline{E} + j\omega\overline{A}) = 0$. Since the curl of $(\overline{E} + j\omega\overline{A})$ must vanish, we can write this as the gradient of some potential. Identifying this potential as the scalar potential $-\phi(\overline{r})$, we can now identify $\overline{E}$ and $\overline{B}$ using:

$$\begin{aligned} \overline{E} &= -\nabla\phi(\overline{r}) - j\omega\overline{A} \\ \overline{B} &= -\nabla \times \overline{A} \end{aligned} \tag{B.1}$$

The Helmholtz theorem states that a vector function can be uniquely specified as

the sum of its curl and divergence[1], up to an additive constant [4]. This is the basis for Maxwell's equations, which specify the electric and magnetic fields in terms of their divergences and curls. The Helmholtz theorem has another consequence: we can perform gauge transformations. Neither $\phi(\omega, \overline{r})$ nor $\overline{A}$ are physical quantities: $\overline{E}$ and $\overline{B}$ are. We can perform any self-consistent transformation on $\phi(\omega, \overline{r})$ or $\overline{A}$ that preserve the proper values of $\overline{E}$ and $\overline{B}$.

To see this, add an irrotational field to the vector potential, which, by construction, cannot change the curl. We choose the gradient of some scalar function Λ for convenience. Since the curl of any gradient is zero, adding $\nabla(\Lambda(\omega, \overline{r}))$ to $\overline{A}$ does not affect $\overline{B}$. The inclusion of $\nabla\Lambda$ cannot change $\overline{E}$, so, inserting the new vector potential $\overline{A}' = \overline{A} + \nabla\Lambda$ into eqn. B.1, we find that the new scalar potential is $\phi'(\overline{r}) = \phi(\overline{r}) - j\omega\Lambda$.

The act of gauge transformation is choosing the function Λ to best simplify the problem. The Lorenz gauge specifies $\nabla \cdot \overline{A} = -j\omega\mu\epsilon\phi$. Though we can change our choice of gauge at any time, using the Lorenz will prove most convenient in later problems[2].

The Lorenz gauge can be used to rederive the wave equations of Ch. 1. It is convenient to use a description of sources and potentials. By defining the D'Alembertian operator $\Box$ as $\nabla^2 + \omega^2/c_0^2$, we can write:

$$\begin{aligned} \Box V &= -\epsilon^{-1}\rho \\ \Box A &= -\mu\overline{J_e} \end{aligned} \tag{B.2}$$

The physical meaning of eqn. B.2 is clear. In the Lorenz gauge, the scalar potential is sourced by charges; the vector potential is sourced by currents. As expected, both equations can be shown to support the wave behavior developed in Ch. 1 and can be used in

[1] assuming that the divergence and curl go to zero faster than $1/r^2$ and that the function itself goes zero faster than $1/r$ as $r \to \infty$

[2] Note that if we had specified the existence of fictitious magnetic monopoles in eqn. 1.2, we would also have to define electric vector potential $\overline{F}$ and repeat the gauge derivations shown above. Though this is beyond the scope of this dissertation, one would find the same essential results.

computational electromagnetic calculations. We will shortly show that the computation of potentials is the primary means by which simple finite element calculations subsist.

B.1.2 Uniqueness and the Silver-Müller radiation condition

One can show that the solutions to the time-harmonic Maxwell's equations on finite or infinite sized domains Ω are not generally unique without additional information[3]. It can be shown that the time-harmonic Maxwell's equations require either a boundary condition at $|\overline{r}| \to \infty$ for an infinitely-sized Ω and a lossy medium or an impedance boundary condition on $\partial\Omega$ for their solutions on a finite-sized domain to be unique. For infinite Ω, the time-harmonic Maxwell's equations are often accompanied by the Silver-Müller radiation condition, usually written in one of two ways:

$$\begin{gathered} \lim_{|r|\to\infty} r(\nabla \times \overline{E} \times \hat{r} + jk\hat{r} \times \overline{E} \times \hat{r}) = 0 \\ \hat{r} \times \overline{E} \times \hat{r} = \eta_0 \overline{H} \times \hat{r} + \mathcal{O}(1/r^n) \end{gathered} \tag{B.3}$$

In eqn. B.3, n must be greater than unity to ensure that the contributions vanish sufficiently fast as $|\overline{r}| \to \infty$. The radiation condition, intuitively, specifies the vacuum impedance η_0 at $|r| \to \infty$ and therefore is merely an impedance boundary condition specified at $|r| \to \infty$. We have also asserted that the time harmonic Maxwell equations are only unique in finite domains if the medium is lossy or an impedance boundary condition is specified on $\partial\Omega$ (much like the Silver-Müller condition did in the infinite domain problem).

The intuition is rather interesting: Maxwell's equations are only guaranteed to produce unique solutions if a non-zero resistance is specified in the bulk or a reactance is specified on the periphery. This fact has a well-known application to circuits. AC voltage

[3]Full, rigorous proofs may be read at [202]

and current sources, the circuit equivalent of oscillating charges and current densities, placed alone in circuits do not have a unique specify their outputs. The output current of only an AC voltage source is not unique; the voltage output of only an AC current source is not unique. The circuit is only solvable when one of two types of circuit elements is placed: either a lossy resistor is placed (the Thevenin configuration for voltage sources or the Norton configuration for current sources) or an impedance boundary condition relating the current to the voltage (a reactive capacitor or inductor).

We will not explicitly apply the Silver-Müller radiation condition to solve any problems. It is, however, important to remember in numerical problems, which we will use to analyze spectrometer designs. Failure to assign proper boundary conditions, even at $|\overline{r}| \rightarrow \infty$, may result in internal resonances in finite element and method of moments calculations, causing non-physical results.

B.2 The finite element method

Several numerical techniques are commonly used to solve for electric and magnetic fields in realistic structures, including the finite difference time domain (FDTD), Method of moments (MoM), and finite element method (FEM). Each technique has different strengths and weaknesses. The full study of these techniques can exceed a career, so we only summarize the phenomenology of solutions from each technique.

FDTD solutions define electric and magnetic fields at every point in given volume Ω and enforce Maxwell's equations in sequential time domain steps. The technique is relatively simple to understand and doesn't require matrix diagonalization. FDTD, however, suffers from dispersion error at high frequencies and large time windows, stemming from error in the computed phase velocity accruing over distance with each time step [202]. Computational complexity scales as $f_{max}^6 L^5$ (with length scale L) and is therefore

exacerbated in electrically large problems at high frequencies, particularly if there are multiple length scales. On-chip THz TDS systems have electrically long transmission lines and photonic bandgap filters to avoid time domain reflections and require solution over many length scales (substrate thickness, conductor width, conductor thickness, and transmission line length), making FDTD infeasible for calculation.

MoM computations rely on solving a variational formulation of the electric field integral equation (EFIE), magnetic field integral equation (MFIE), and/or the combined field integral equation (CFIE). The EFIE, MFIE, and CFIE are derived from the representation formulae of Maxwell's equations and are only solved on domain periphery $\partial\Omega$, reducing the computational complexity. Since MoM computations reside on boundaries, they are very natural for antenna and scattering problems. Though they reduce solution dimensionality and are computationally simpler than finite element methods, the Lax-Milgram convergence theorem does not apply and the solutions may not converge with increasing mesh density. The sparse matrices utilized by MoM solvers require intricately engineered preconditioners, which we avoid altogether.

We, instead, use the finite element method. Understanding best practices first requires understanding a model calculation. The following example is adapted from the Autumn 2017 ECE 5510 notes of Dr. J. F. Lee and Dr. K. Sertel at The Ohio State University.

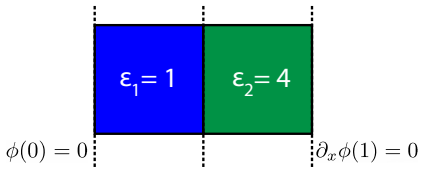

Figure B.1: Geometry for illustrative boundary value problem, with one Dirichlet (left) and one Neumann (right) boundary value.

Consider the one dimensional geometry pictured in Fig. B.2 with electrostatic potential $\phi(x)$ from 0 to 1. Two adjacent dielectrics of width 1/2 are subjected to a grounded

potential at x = 0 and a potential with a vanishing derivative at x = 1. The electrostatic potential $\phi(x)$ of this system is solved by manipulating Gauss's law (eqn. 1.2) with the constitutive relation and the definition of electrostatic potential ($\overline{E} = -\nabla\phi(x)$):

$$-\partial_x \cdot \epsilon \partial_x \phi(x) = \rho_e \tag{B.4}$$

Eqn. B.4 is the Sturm-Liouville equation in one dimension. For simplicity, we set ϵ_0 and ρ_e to 1. The analytic solution, which must be the only solution by the uniqueness theorem, is obtained by simply integrating twice and applying boundary conditions.

$$\phi = \left\{ \begin{array}{l} \frac{x^2}{2} + x, 0 \leq x < \frac{1}{2} \\ \frac{-x^3}{8} + \frac{x}{4} + \frac{9}{32}, \frac{1}{2} <= x < 1 \end{array} \right\} \tag{B.5}$$

This analytic expression can be compared to the "best" numerical solution obtained by creating a mesh. A mesh refers to the division of the geometry into individual elements. We expand the potential in trial basis functions ψ_m as the linear combination $\phi(x) \approx \sum_{m=1}^{M} a_m \psi_m$ with weighting coefficients a_m. We must now chose the form and domain of the basis functions.

The most obvious and simple mesh requires one element for each dielectric region. Using the Sturm-Liouville equation, we define a residual error R(x). We define a functional $P : \mathbb{C}^n \mapsto \mathbb{C}^1$ of the integrated residual error, subjected to arbitrary weighting function $w_m(x)$.

$$\begin{aligned} 1 + \partial_x \epsilon \partial_x \phi(x) &= R(x) \approx 0 \\ P &= \int_\Omega w_m(x)(1 + \partial_x \epsilon \partial_x \phi(x))\, dx \\ P &= \sum_{m=1}^{M} \int_\Omega w_m(x)(1 + \partial_x \epsilon_n \partial_x (a_m \psi_m))\, dx \end{aligned} \tag{B.6}$$

We apply Galerkin's method, for simplicity, which sets the trial and testing weighting functions equal to one another ($\psi_n = w_m$). The Euler-Lagrange theorem is used to determine a stationary point in the integrated residual P. The result is called the Galerkin weak solution. The Galerkin weak solution will stably converge to the analytic answer if the Galerkin solution is bounded and coercive, which is guaranteed for the Poisson equation (and fully time-harmonic Maxwell's equations) by a form of the generalized Lax-Milgram theorem [202]. Finite element solutions work by reducing the residual error as the density of meshed elements M/L (for length scale L) is increased. All that remains in our simple example is to define a mesh, choose basis functions, and variationally minimize the residue.

The mesh is simple. We divide the domain Ω into discrete segments of length $\delta L = L/M$. The finite element method will guarantee convergence as $\lim_{M\to\infty}$, though this process may be slow.

To use the finite element, we define a new functional $I : \mathbb{C}^n \mapsto \mathbb{R}^1$ for the linear differential operator $\mathcal{L}$ acting on function $\nu(x) = \nu$.

$$I(x) = \langle \mathcal{L}(\nu(x)) | \nu(x) \rangle - 2 \langle \mathcal{L}(\nu_0(x)) | \nu(x) \rangle \tag{B.7}$$

The inner product $\langle \alpha(x) | \beta(x) \rangle$ is the usual $\int_\Omega \alpha(x)\beta(x)dx$. The function $\nu \in \mathcal{H}^1(\Omega)$ has admissible Hilbert space is the space of functions that can be differentiated once and still have finite energy (be L^2 integrable)[4]. I is a measure of the difference in residual between *some* numerical solution $\nu(x)$ and *the best* numerical solution $\nu_0(x)$. Deriving the factor of two preceding the $\langle \mathcal{L}(\nu_0(x)) | \nu(x) \rangle$ is beyond the scope of this dissertation, details can be found in [202]. Applying eqn. B.7 to the Sturm-Liouville differential operator and integrating by parts produces:

[4]The Sobolev space $\mathcal{H}^n(\Omega)$ is the normed Hilbert space of functions whose n^{th} derivative is L^2 integrable on Ω.

$$
\begin{aligned}
I(x) &= \int_{\Omega} \left[(-\partial_x(\epsilon(x)\partial_x\nu(x)) + \rho(x)\nu(x))\nu(x) - 2f(x)\nu(x)\right] dx \\
I(x) &= \int_{\Omega} \left[\epsilon(x)\partial_x\nu(x)^2 + \rho(x)\nu(x)^2 - 2f(x)\nu(x)\right] dx - \epsilon(x)\partial_x\nu(x)\nu(x)|_{\partial\Omega}
\end{aligned}
\tag{B.8}
$$

The one Dirichlet and one Neumann boundary condition make the boundary ($\partial\Omega$) term vanish. Deriving the best numerical answer stipulates that $I(\nu_0) \leq I(\nu_0(x) + \delta u(x))\ \forall \delta u(x)\mathcal{H}^1(x)$. We expand eqn. B.8 for small δ, collect all terms of order δ, and set their sum to zero to find a stationary point in I. The resulting expression, after integrating by parts, is:

$$
\begin{aligned}
0 &= \int_{\Omega} \left[\epsilon(x)\partial_x\nu_0(x)\partial_x u(x) + \rho(x)\nu_0(x)u(x) - f(x)u(x)\right] dx \\
0 &= \int_{\Omega} \left[\partial_x(\epsilon(x)\partial_x\nu_0(x)) + \rho(x)\nu(x))\nu_0(x) - f(x)\right] \nu(x)dx - \epsilon(x)\nu_0(x)u(x)|_{\partial\Omega}
\end{aligned}
\tag{B.9}
$$

We need to numerically solve eqn. B.9 by expanding in basis functions. These basis functions reside in $\mathcal{S}^1 \subset \mathcal{H}^1(\Omega)$, the subspace of discrete functions whose first derivatives have finite energy. We chose the first order interpolatory Ritz basis for ψ_m. The ith individual basis element for mesh size δl is defined as $y = 1+1/\delta l$ from $x \in [-\delta(i-1), \delta li]$, $y = 1 - 1/\delta l$ for $x \in [\delta li, \delta l(i+1)]$, and zero elsewhere. Assigning one such function per mesh element produces a first order hold interpolation between adjacent elements. As before, we must find the numerical statement $I(\nu_0) \leq I(u)\ \forall \nu_0 \in \mathcal{H}^1$ that by setting the first variation to zero.

Plugging the expansion $\nu(x) = \sum_i^M a_m\psi_m = \langle a_m|psi_m\rangle$ into eqn. B.8 and simplifying for purely real operators and functions produces:

$$
\begin{aligned}
I(\nu(\psi_m)) &= \int_\Omega \left[\epsilon(x)(\partial_x \nu)^2 + \rho(x)\nu(x)^2 - 2f(x)\nu(x)\right] dx \\
I(\nu(\psi_m)) &= \int_\Omega \left[\epsilon(x)(\partial_x \langle a_m|\psi_m\rangle)^2 + \rho(x)\langle a_m|\psi_m\rangle^2 - 2f(x)\langle a_m|\psi_m\rangle\right] dx \\
I(\nu(\psi_m)) &= \int_\Omega \left[\epsilon(x)\langle a_m|\partial_x|\psi_m\rangle\langle\psi_m|\partial_x^\dagger|a_m\rangle + \rho(x)\langle a_m|\psi_m\rangle\langle\psi_m|a_m\rangle^2 - 2f(x)\langle\psi_m|a_m\rangle\right] dx \\
I(\nu(\psi_m)) &= \langle a_m| \int_\Omega \epsilon(x)\left[\partial_x|\psi_m\rangle\langle\psi_m|\partial_x^\dagger dx\right]|a_m\rangle + \langle a_m| \int_\Omega \rho(x)|\psi_m\rangle\langle\psi_m|a_m\rangle \\
&\quad -2\int_\Omega f(x)\langle\psi_m|a_m\rangle dx
\end{aligned}
\tag{B.10}
$$

One can show that the setting the first variation of eqn. B.10 to zero, as we did with the non-expanded form in eqn. B.9. This produces a final matrix equation that needs to be solved for the best interpolatory basis expansion coefficients:

$$
\begin{aligned}
|a_{m0}\rangle &= \left[\int_\Omega \epsilon(x)\partial_x|\psi_m\rangle\langle\psi_m|\partial_x^\dagger dx + \int_\Omega \rho(x)|\psi_m\rangle\langle\psi_m| dx\right]^{-1} \int_\Omega |\psi_m\rangle dx \\
|a_{m0}\rangle &= \mathcal{A}^{-1}\,|\int_\Omega \psi_m dx\rangle
\end{aligned}
\tag{B.11}
$$

For the Ritz method with the simple interpolatory basis, the matrix elements of $\mathcal{A}$ follow a regular pattern. The matrices in eqn. B.11 reduce to a tridiagonal matrix. We label the left and right sides of individual triangular elements as "l" and "r", respectively. The dielectric constant and charge density on the left leg and right leg of the triangular basis elements become ϵ_l and ϵ_r, and ρ_1 and ρ_2 respectively. The ϵ matrix has diagonal elements $(\epsilon_L + \epsilon_R)/\delta L$; the ρ matrix has diagonal elements $\delta L/2(\rho_L + \rho_R)$. Diagonal adjacent elements of the ϵ matrix are $-\epsilon_R/\delta L$; diagonal adjacent elements of the ρ matrix are $-\delta L\rho_R/2$. Note that higher order basis functions may produce matrices less sparse than the tridiagonal one shown here for the simple interpolatory basis.

The finite element solution for our Sturm-Liouville boundary value problem is plotted in Fig. B.2. Note that the solution automatically obeys the lower order (Dirichlet) boundary condition because the interpolatory basis obeys the lower order boundary condition. As the mesh density increases, so does the accuracy with which the higher order (Neumann) boundary condition is enforced.

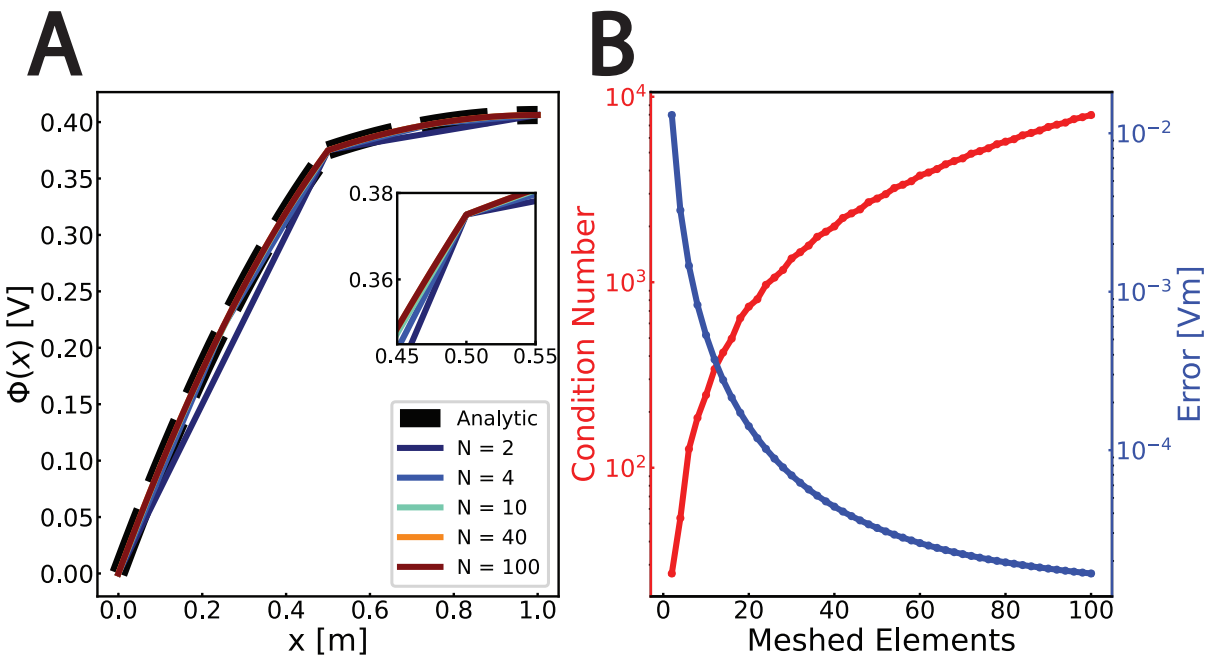

Figure B.2: Finite element computation for the Sturm-Liouville electrostatics problem presented in Fig. B.2. (A) Finite element solutions to the potential versus mesh density, with the analytic solution in large dashed black line for comparison and (inset) zoom showing high error in regions where $\partial_x \phi(x)$ is large (B) condition number and integrated error versus mesh density

Two additional quantities are plotted in Fig. B.2. The condition number of a matrix is the ratio of the largest eigenvalue to the smallest eigenvalue and quantifies the range of numerical values involved in a calculation. Generally, the condition number of the involved matrices increases with the mesh density. As this number becomes larger and larger, repeated matrix operations begin to lose precision. If the condition number were, say, 1E-16, one would expect operations like Gauss-Jordan elimination to lose all precision

for floating point numbers. A rule of thumb - sudden jumps in the condition number versus number of meshed elements are good reason to be skeptical of an answer. Our condition number smoothly increases, as one might expect.

We also plot the integrated absolute error in the answer, which we can easily calculated by plugging the finite element computation result back into $\mathcal{L}(x)$. The error monotonically converges to zero, as we would expect. One may ask: how accurate is "accurate enough"? One typically defines an acceptable error budget (1E-2 Vm, for example). The flow of work is thus: perform a simple finite element computation, measure the error, increase the mesh density, and repeat.

We learn another simple lesson from our toy calculation: increasing the number of meshed elements uniformly about Ω is wasteful. Most of the error is concentrated in regions where the potential derivative is large. Since we knew the analytic solution before starting, we anticipated needing a mesh element centered at the x = 1/2. This luxury does not exist in generalized finite element problems. Instead, the computer program must "guess" a multidimensional mesh, measure the error, and increase the mesh density in the high derivative (error-prone) regions. This is the basic intuition underlying the complicated adaptive mesh used by most professional software packages.

We can take a few lessons from our simple finite element Sturm-Liouville exercise:

1. The finite element method produces a best solution by functionally minimizing residuals, subject to (for us) the Galerkin weak statement. Computations are all matrix equations and subject to the challenges of numerical analysis.

2. The condition number is a simple benchmark that monitors the numerical precision of repeated computations. It is always wise to monitor it and watch for sudden jumps.

3. Increasing the mesh density increases the solution accuracy, but also increases the

problem complexity. Since the solution depends on matrix inversion, naïvely Gauss-Jordan will scale as $\mathcal{O}(N^3)$. Increasing computation speed requires faster matrix inversion algorithms, numerical preconditioners, or taking advantage of the tridiagonal matrices. Increasing mesh density can be done intelligently to strategically reduce the integrated error, usually through adaptive meshing packages.

4. The basis functions *must* remain in the appropriate Hilbert spaces, or the calculation risks non-physical results. Higher order basis functions may improve accuracy but greatly increase computation time because they don't generate the simple tridiagonal matrices that can be exploited by a pre-conditioner.

The analytical solution is guaranteed by the uniqueness theorem to be the correct and only solution. The numerical solution is guaranteed to be the best solution for a given form of basis functions and mesh size. Analytic solutions also offer greater intuitive descriptions than numerical solution, as one can algebraically manipulate them without having to re-run an entire calculation, but...

B.3 Best practices

... analytic solutions cannot be written for most realistic THz problems due to the complex geometries. We will resort to calculation using the finite element method, but can still monitor three main parameters to assess the veracity of our model. The following questions can help in assessing whether the finite element or method of moments code solved the desired problem accurately or not.

1. **Hand calculations of simplified systems**

   Simplify the geometry (infinite conductivity metals, non-tapered waveguides, small pieces of the geometry only... etc.) and perform simple hand calculations. Does

the simulation predict approximately the same as the hand calculations? Are differences between simplified hand calculations and the full model explicable?

2. **Passivity**

   Passivity is a measure of the energy output by a system divided by the energy input to a model. Numerical simulators input a finite amount of energy through wave, lumped, and transient ports, which may be absorbed by the model elements (metals, dielectrics, etc.) or by a boundary (another wave, lumped, or transient plot, or a radiation absorbing boundary). If the passivity exceeds unity, the solution is non-physical. If the passivity as a function of frequency rises abruptly, one should also be suspicious and re-run the simulation with a more dense mesh, higher-order basis function, and/or lower error tolerance.

3. **Condition numbers**

   The matrix equation in eqn. B.11 requires exact or approximate diagonalization to solve. The condition number - the ratio of the largest to the smallest eigenvalue - measures the ability of matrix operations to maintain precision over extended computations. For example, a condition number of 1E16 should immediately raise one's suspicion (as subsequent numerical operations with floating point numbers will rapidly loose precision), whereas a condition number of 100 is very reasonable.

I use Ansys electronics (HFSS) for finite element electrodynamic computations. In my experience, the adaptive mesher produces meshes such that the condition numbers need not be carefully monitored. One should, however, check with hand calculations and *always* scrutinize the passivity in frequency sweeps. I have found a number of best practices in Ansys electronics that have improved my workflow, decreased my time to getting an answer, and allowed me to spot errors in THz system engineering. These are:

1. **Follow the toolbar**

   Modern versions (2020 and newer) of Ansys electronics were designed with greater user friendliness than previous versions. Once an HFSS design has been inserted into the workspace, the workflow to creating a solution is detailed under the project file header in the project manager toolbar. Populate components of the solution sequentially from top to bottom, as shown in Fig 1.

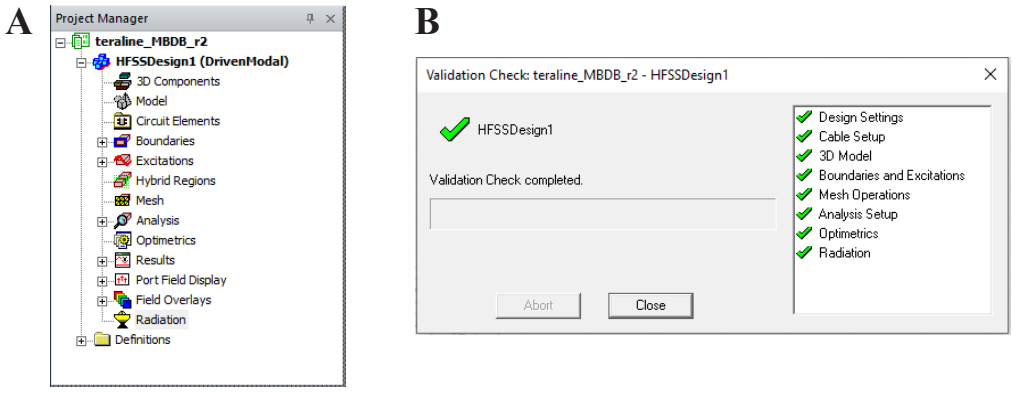

Figure B.3: (A) HFSS project manager menu and (B) Validation check for solution readiness.

   (a) Populate the geometry, which will appear under 3D components

   (b) Add a radiation bounding box with a radiation absorbing boundary condition. Per the Silver-Müller radiation condition, some boundary condition or loss is needed to guarantee uniqueness. Failure to do may produce non-physical results.

   (c) Add excitations, in the form of wave, transient, Floquet, or terminal ports

   (d) Add a solution setup. Chose basic meshing options and the swept frequency range.

   (e) Add parametric geometry steps and/or optimization, as needed.

   (f) Plot the results once simulated.

Once you think your solution is ready to run, go to HFSS → Validation Check. A menu (Fig. 1B) will appear, showing aspects of the design ready for simulation (green checks) and not yet ready for simulation (red stop sign). When all checks are green, the system is prepared to run the simulation. This does *not* guarantee successful computation.

2. **Clear, simple geometries**

   HFSS's powerful CAD system has great strength in use of variables. I recommend producing a hand drawn or otherwise computer-aided diagram clearly detailing all dimensions. This makes adding or removing circuit elements straightforward and will make parametric design steps simple. I recommend a simple design convention: name all variables describing lengths in the x-axis as "x_{name}"... etc.

   Note that external models can be imported into HFSS or exported from HFSS. Use this feature when producing diagrams for nanofabrication, photolithography masks, or dicing saw diagrams.

   I find that errors are most easily identified when the top of all dielectric surfaces rest at Z = 0, so all transmission line metallization occurs at positive heights. I also center all of my designs at (X, Y) = (0, 0) and maintain coordinate symmetry as best as possible. Before simulating any design, I change intentional change dimension variables to ensure that the design scales as desired. This is vital if you will be using any parametric or optimization in your design.

3. **Do not renormalize ports**

   Upon creation, wave and terminal ports request several (optional) inputs, including number of modes, integration lines... etc. Under Post Processing, HFSS gives the option to perform port renormalization. Renormalizing ports means that all ports

perform as ideal 50Ω, which can fix the simulated device's impedance or cause reflections between undesired components in the circuit. This is useful in microwave engineering problems, where the impedance of a network analyzer needs to be modelled. This is not meaningful, however, when only the characteristic impedance and propagation characteristics of a transmission line need to be measured. In simple THz TDS simulations, ports should not be renormalized.

4. **Monitor the solution convergence**

   FEM solution quality is guaranteed to improve with increased mesh density, barring numerical problems. One might wonder: how dense a mesh is dense enough to produce a reliable answer? HFSS addresses this by meshing the simulation geometry with N tetrahedra, solving the full S matrix (N x N, for N ports), and mapping the S-matrix to its the determinant ΔS_m. HFSS's adaptive mesher then increases the number of tetrahedra and computes the new ΔS_{m+1}. If the $\Delta S_{m+1} - \Delta S_m$ exceeds a threshold (default: 2%), the adaptive mesher increases mesh density and repeats the process. More accurate solutions may be obtained by decreasing the tolerated $\Delta S_{m+1} - \Delta S_m$ to 0.01 or 0.005. The number of allowed mesh refinements can also be increased, for particularly complicated geometries.

   I recommend checking the mesh convergence by in the project manager → Analysis → Setup → Convergence. If the mesh was far from converging and suddenly dropped below the tolerated threshold in one refinement, one should be suspicious of the results.

5. **Always check passivity**

   As detailed above, passivity is an important tool to benchmark solution quality. No solution should be considered valid until the operator first checks the passivity.

In Ansys HFSS, this is done by inserting a rectangular modal solution result and selecting passivity.

6. **Check wavevectors with multiple modes**

   If you are simulating multiple modes in a transmission line, check the imaginary wavevector of the modes to be sure they are propagating and not evanescent solutions. Once computation is done, insert a modal solution report of Im(gamma(1:1,1:1)) and Im(gamma(1:1,1:2)), or equivalent. This plots the imaginary wavevector of the wave injected at port 1, mode 1 and measured at port 1, mode 1. If the imaginary wavevector is zero, the wave is evanescent. If the imaginary wavevector is nonzero, the wave is propagating and is a valid part of the solution. This is often not obvious from field or simple S-parameter displays. An example is further developed in the following best practice.

7. **Ensure higher order modes aren't waveport effects**

   The feed geometry of any device can be done with a lumped element port (which is best if the feed physically comes from a point) or a waveport (if the input does not correspond to a point excitation). In general, wave behavior is best captured by waveports and we will use them. Waveports seem innocuous, but their finite geometry can lead to insidious and non-physical effects in simulation. Waveports must be large enough to capture "all" of the fields surrounding a transmission line. I suggest using 10x the characteristic transverse length, S+2W for a CPS.

   It is wise to parametrically vary the waveport geometry of your final design to ensure that your results do not stem from boundary effects. For example, we examine the 3^{rd} mode of a simulated CPS with a waveport of finite width in Fig. 7A. At low frequency, β_3 is zero or near zero, indicating that low frequency waves are in a

bandgap and do not propagate. β_3 is non-zero at higher frequencies, so a wave is allowed to propagate. The transition, however, is waveport-width dependent, suggesting that the higher-order mode is related to the waveport and not the CPS.

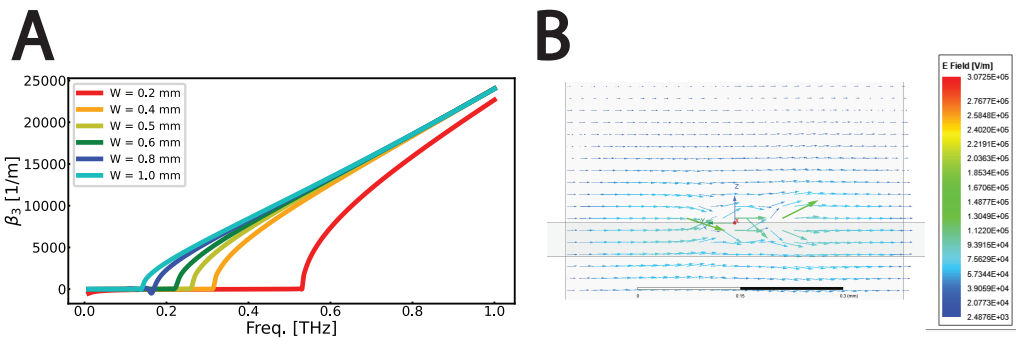

Figure B.4: 3^{rd}-order mode of a CPS with varied waveport geometry. (A) Imaginary wavevector of the 3^{rd}-order mode (β_3) versus frequency for varied waveport width. (B) Field diagram exported from Ansys (HFSS), showing field lines pointing laterally from one edge of the waveport (vertical lnes at the left and right edges of the diagram) to the other.

We examine the electric field of the real space mode structure (Fig. 7B). The CPS metallization is near the origin, sitting atop the 50 μm substrate and is fed by waveports, shown as vertical blue rectangles. The field lines point from the left to the right in the substrate and vacuum superstrate, with small deviations near the conductors. The strength of the field in the vacuum confirmed our suspicions: this mode is some sort of waveguide mode whose existence relies on the waveport size.

The final example example beautifully illustrates the grander point of this appendix. Use your intuition and carefully check the results of the computation before reporting them as physically meaningful.

Appendix C

Daughterboard alignment

"This is so easy that a gorilla could do it"

A. F. Young

The interchangeable sample architecture developed in Ch. 4 requires the consistent and precise alignment of a quartz sample board to a COP switch board. This can be done consistently in less than ten minutes with a dedicated setup. The setup requires three components: a probe holder, a switch board movement assembly, and simple optics.

The probe holder is a key component to holding the cryostat probe steady while samples are being installed / removed. The probe holder contains several sections of PTFE machined to match the probe profile. The PTFE blocks have tapped screwholes populated by soft-tipped set screws. These set screws are tightened into the probe, preventing the probe from moving while the sample is being installed.

The switch board movement assembly is a micrometer with a metallic wire. The wire is small enough to allow for careful translation of the sample board. Metallic (not carbon-tipped) wire dissipates any built-up electrostatic charge and prevents the sample boards from jumping on the switch boards. We observe the alignment through a dedicated, long focal length microscope with continuously tunable zoom and a monitor. We use

an AM Scope 3.5X-180X Trinocular Stereo Microscope (Amazon B0719B2S61), which costs about 1,000 USD and provides zoom near and far zoom alignment images. One could equivalently use a USB camera (such as the SVPRO B096V84M57) which costs only 120.99 USD and offers the same level of performance.

To install a sample board to a mounted switch board, tighten the set screws to ensure the probe is tight in the sample holder. Ensure that at least two of the switch board screws are populated. Place the sample board on the switch board, with the active side facing the switch board. I find using vacuum tweezers to be the least error-prone. Then, secure the micrometer to another surface and being gently pushing the sample board into position. Use the microscope to align the sample board as precisely as possible to the switch board. When the sample board is placed accurately, use tweezers to move the SEM clip to pin the sample board to the switch board. Lower the SEM clip directly down, avoid lateral movements. Place the clip within a few hundred microns of the switch board transmission line. THz coupling quality, as measured from the switch board - sample board reflection, becomes measurably worse as the SEM clip touchdown location moves further from the switch board transmission line (Ch. 5).

Appendix D

Practical THz TDS optics for beginners

"Optics is the geometry of light as music is the arithmetic of sounds."

adapted from *Claude Debussy*

The ultrafast laser is the heart of a THz TDS setup. THz data integrity can only be as good as the laser, the optical elements, and the combination thereof permits. This appendix details basic principle of using and handling, and debugging each of the main optical elements depicted schematically in Fig. 4.1D.

This appendix is not meant to be read sequentially - rather, it is to be used as a reference when building a setup for the first time or to debug issues that later arise in well-established setups. The intended use of every section is labelled as either "background" (theory), "advice" (suggestions on installation and usage) or "debugging" (what to do when issues arise). Included are basic principles of fiber laser operation and simple diagnostics to monitor system health, basic table optics principles (lenses, mirrors, telescopes, irises, delay stages, fast photodiodes, etc.) and list suggested components when designing a system.

The names and prices of all components described here are current as of the filing of this document - they may chance over time. The majority of work described in this dissertation was performed with a Menlo Systems M-fiber 1560 nm Er-doped fiber laser with a second harmonic stage. This was later replaced by a Menlo ELMO 780 Er fiber laser. Unless otherwise specified, assume that described diagnostics refer to the M-fiber laser.

D.1 Ultrafast fiber laser pulses

The modelocked laser must output a sharp (<150 fs) ultrafast pulse centered at an infrared wavelength of 780 nm. The laser must be modelocked so that it only lases one mode, whose repetition rate is carefully controlled and monitored. Details of the modelocking mechanism in rare earth fiber lasers may be found elsewhere [203].

Laser performance is generally temperature and humidity sensitive. I recommend installing an array of temperature and humidity sensors to continually log the lab environmental parameters. We use a temperature watchdog sensor (array) and the Vertiv humidity monitor to do so, and have been able to monitor the effects of few degree Fahrenheit swings in the lab. Part details are included at the end of this section.

D.1.1 Background: S-, P-, and circular polarization

The $\overline{E}$ field of a propagating plane wave can be written as $\overline{E} = E_x\hat{x}e^{j\phi_x} + E_y\hat{y}e^{j\phi_y} + E_z\hat{z}e^{j\phi_z}$. Many optical components require certain incoming polarization or output only certain polarizations. Literature refers to different polarizations by different names. It is important to remember the different names used by different sub-fields:

One usually pictures each spatial electric field component as having the same phase, but this need not be. One can create circularly polarized light by introducing a $\pi/2$ phase

$\overline{E}$ is...	Notation 1	Mnemonic	Notation 2
Parallel to table	S	"Sideways"	TE
Perpendicular to table	P	"Perpendicular"	TM

Table D.1: Polarization notation for free space optics

difference between two components, wherein the locus of $\overline{e(t)}$ is a circle. Introducing $E_x \neq E_y$ causes $\overline{e(t)}$ to trace an ellipse, which is referred to as elliptical polarization. Optical fibers, dielectric structures that guide plane waves, can be engineered to either maintain or not maintain the field polarization.

Note that some components (autocorrelators, beamsplitters, polarization-maintaining fibers, rhombs... etc.) require purely S-, P-, or circularly polarized light. Some components (lasers, beamsplitters, waveplates, Wollaston prisms...) output purely S- or P-polarized light. One should keep track of the polarization at various positions on the table. If one needs to rotate a linear polarization by an arbitrary angle, install a half-waveplate on a rotating mount. I recommend installing using a ThorLabs WPH10M-780 on a continuous rotation mount (RSP1X15) in each beamline to allow arbitrary polarization rotations.

D.1.2 Background: simple laser diagnostics

Three tests are used to characterize a modelocked 780 nm laser output: the optical spectrum, the optical autocorrelation, and RF pulse spectrum from a fast photodiode. The optical spectrum is a measurement of power per lasing wavelength, and is important to ensure that the majority of the spectral content is at energies greater than the PC switch bandgap. The autocorrelation is a measurement of the ultrafast pulse time and shape. The bandwidth of the emitter THz pulse is fundamentally limited by the sum of the ultrafast pulse and PC switch response times, so it is important to be sure that the

ultrafast pulse is sufficiently short (rule of thumb: < 150 fs). The RF pulse spectrum ensures that the laser's repetition rate is as specified (250 MHz). A proper RF pulse spectrum is merely a broadened delta function at 250 MHz. Additional spectral features can be an early indication of modelock failure.

The Menlo M-fiber laser's optical spectrum was measured with a ThunderOptics spectrometer with Spectragryph software. The optical spectrum analyzer is nonlinear at high fluence, which can be circumvented by placing the sensor and a white paper card on optics mounts. Place the card in the beam path and bring the supported sensor near the card. Backscattered light from the card will enter the sensor, and the power level can easily be tuned with an external neutral density filter wheel or by simply moving the sensor closer to / further from the card. Sample the spectrum with several (low, high... etc.) signal levels. The most accurate spectrum is that with components that do not saturate the detector readout. Three spectra fitting this criterion are shown in D.1.2A, showing a bandwidth of about 27 nm.

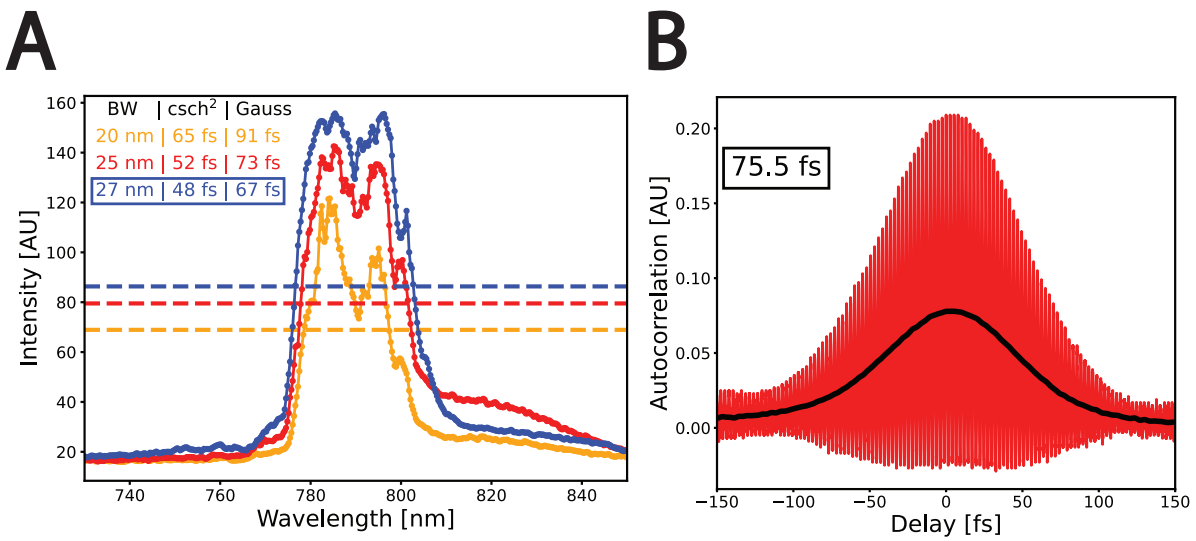

Figure D.1: M-fiber laser diagnostics. (A) Optical spectrum, taken with three incident powers. The effects of the detector non-linearity are difficult to quantify, but blue is likely the most accurate spectrum. (B) Pulse autocorrelogram (red), with fitted Gaussian pulse (black) width of 75.5 fs.

This is a lesson for on-chip THz different than that taught by free space THz. Free space THz setups can use 800 nm lasers with short pulse times because of larger 300K GaAs bandgap. On-chip THz systems with GaAs PC switches should avoid 800 nm lasers and extremely short (<40 fs) optical pulses at 780 nm to avoid losing emission efficiency at low temperatures, per Varshni [122] and Ch. 3.

The bandwidth, defined as the full-width, half-maximum, is indicated by dashed lines. The transform limit, which is the minimum pulse duration set by the uncertainty principle for a given pulse shape and bandwidth. The blue spectrum has a bandwidth of 27 nm and suggests a minimum Gaussian pulse time of 67 fs. Menlo M-fiber lasers are specified to produce transform-limited +10% Gaussian pulses, so one naïvely expects a 74 fs Gaussian pulse from the optical spectrum alone.

The pulse width, however, could suffer from the second harmonic stage and fiber couplings, many of which had to be customized in the M-fiber laser. To ensure a sufficiently small pulse time, we also measure the pulse autocorrelation with an interferometric Mach-Zenhder autocorrelator (ThorLabs FSAC). The autocorrelogram is measured with a Mach-Zehnder interferometer, wherein an incoming pulse is divided into two arms (with one of variable length). The two pulses are recombined and their interference pattern is measured. The result is shown in Fig. D.1.2B. Individual oscillations have period 2.6 fs, which correspond to the individual electric field oscillations of a 780 nm wavelength optical pulse. The overall shape can be fit to simple theoretical modelling [2] and can be used to extract a Gaussian or csch^2 pulse width. The extracted Gaussian pulse width is 75.5 fs, matching remarkably well with our crude estimate from the optical spectrum analyzer. Note that this pulse autocorrelogram was measured near the laser output, not near the sample. The pulse does broaden as it goes through optical elements like mirrors, lenses, and beamsplitters. Simple calculations from component datasheets suggests that this broadening is less than 30 fs, so pulse arriving on the sample still has < 150 fs

lifetime.

D.1.3 Debugging: optical base station

I recommend building a base station. The term "base station" refers to a handful of optical elements used to continuously record fiber laser diagnostics (beam power and Poynting stability) during operation. We insert a 90:10 beamsplitter into the main optical path, which diverts 10% of the power into another beamsplitter. One output of the second beamsplitter is fed to a power meter, which continually records the spot power. Note that the noise floor can be substantially reduced by including a 1" lens tube on the power meter. The other beamsplitter output is passed through a ND 6.0 or greater filter and directed to a CMOS camera with a small pixel size. Continuously recording the spot profile on this camera allows one to measure drift in the peak, and can be used to quickly diagnose problems. This camera will also be useful for precise delay stage alignment, as we will later see.

I do not recommend including a ThorLabs FSAC autocorrelator in the base station, as it operates best with several mW of optical power. Instead, I suggest inserting a mirror on a flip mount (ThorLabs TRF90) early in the beamline, so one can easily divert the beam to record the autocorrelation or optical spectrum.

D.2 Simple optical elements

We review simple techniques in assembling an optics table for on-chip THz TDS measurement. A simple list of required tools, pitfalls, and scheduling estimates is provided for each process. It is important, before beginning any of the following processes, to know your laser mount and output height. All optical components will be mounted in a plane parallel to the optics table, with a few exceptions. Chose this height (usually the

height of the output port, if convenient) before performing too much work. Ensure you have sufficiently many posts and bases of the correct height. Set all components to this height before placing them in the layout: this will save a lot of time.

D.2.1 Advice: miscellaneous component handling

The most general and popular type of base we use is the clamping fork with studded bases. Clamping forks and studded bases have a small profile and allow one to greatly increase the component density. Note that students in adjacent labs have used these successfully on a vertically mounted breadboard without any problems. I recommend using ThorLabs R2 collars as much as possible to ease coarse alignment

One can always build periscopes, but the ThorLabs pre-built periscope mounts (RS99) are rumored to have less drift on yearly time scales.

Always wear gloves when handling optics components. Components should always be handled by their edges and never by their bulks. Components' holders should be labelled with the components' part number. A convention should be developed to determine the direction of a lens or filter, such that they "forward" direction through a filter or the convex side of a lens is always on the same side of the lens holder.

I am fond of ThorLabs LMR1 or equivalent for holding lenses. I find it best to install components by lowering them with a gloved hand into the lens holder, placing the retaining ring, and screwing it in with a spanner wrench (ThorLabs SPW602 or equivalent). One can also use a narrow flathead screwdriver, but risks scratching antireflective coatings if your hand slips.

Keep the boxes that filters, lenses, mirrors... etc. ship in – they are useful for storing components later. This is particularly useful for micrometers and drives, where electrostatic discharge sensitive or padded foam packaging is customized to the parts

themselves.

If you need to clean optics, try blowing clean with small canisters of compressed air usually used for cleaning keyboards. If that is insufficient, apply IPA to a lint-free cloth. Place the lint free cloth onto the component. Carefully drag the cloth across the optic's surface and blow dry with compressed air. Use the smallest amount of IPA possible. You might find it helpful to use a pipette to carefully dispense IPA, particularly when cleaning sensitive components (such as beamsplitters). ThorLabs lens cleaning notes may be found here. A Newport optics cleaning guide is found here.

D.2.2 Advice: affixing components to an optics table

Be sure to tightly secure components to the table. Few things in an optics lab are as annoying as aligning a delay stage and later finding it misaligned because a component wasn't securely fastened. Use washers, when possible, to avoid damaging bases.

Note that the lamination on all optics tables slowly wears out over time. To find valid machined holes, use the back end of a table driver to tap on the table. Near the center of the table, where the table is still sound, the tap will make a deep resonant "thud". Near the edge of the table, likely near the outermost hole, the "thud" will become nasal and shallow – avoid putting components here. They will drift much more over time that those nearer the center.

D.2.3 Background / advice: mirrors

Besides posts and mounts, the most ubiquitous component is the mirror. Mirrors function best when the incident spot is located in the vertical and lateral innermost 1/3 of the mirror. For 780 nm light, one can use silver or dielectric mirrors. Silver mirrors attenuate about 3% of incident light and slightly temporally broaden the pulse. Broad-

band dielectric mirrors attenuate about 1% of the incident light and barely broadens incident pulses. Consider the total mirror attenuation when estimating power requirements. Twenty silver mirrors attenuate 46% of incident power; twenty dielectric mirrors attenuate 18% of incident power. Also consider the effects of pulse broadening. A general rule of thumb is that silver mirrors are sufficient for pulses with widths greater than 40 fs.

D.2.4 Background / advice: two mirrors and beam levelling

A single mirror merely reflects light. Combinations of two or more mirrors exist into two basic configurations: two mirrors in the XY plane and two mirrors that translate beam height (the "periscope"). Aligning either case will require walking the beam. We first note, however, that periscopes can change the polarization of light, whereas purely in-plane mirror setups cannot. This can simply be understood from the electric-field boundary conditions (eqn. 1.4) applied to a PEC. It is key to remember this, as one needs to be aware of the beamline polarization with polarizing beamsplitters, autocorrelators, and half-waveplates.

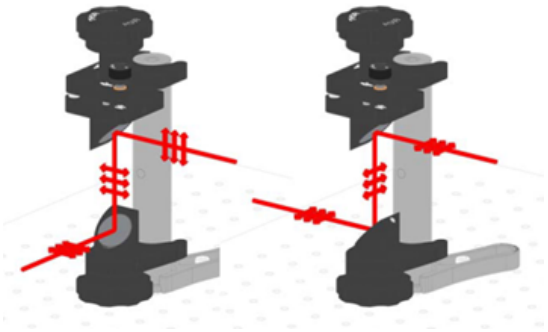

Figure D.2: (Left) Polarization change from lower beam (S-) to higher beam (P-) in a 90° periscope. (Right) No polarization change from lower beam to higher beam (S-) in a 180° periscope. This graphic was adapted from the ThorLabs FSAC autocorrelator manual [2].

The process of levelling / walking a beam is that used to fix a beamline to a particular optical axis. Beams are levelled to prepare a parallel and well-controlled input to subsequent optical elements, which could be beamsplitters, telescopes, the sample itself, etc. In addition, one should level a beam coming out of a laser or oscillator, as output beams are often not as level as one might assume. Excellent tutorials may be found on the ThorLabs YouTube channel and the Edmunds Optics YouTube channel. I summarize below the method I use to level a beam - meaning align it to be parallel to the table and parallel to a line of tapped holes on the table.

Beams are often levelled in one of three configurations, which are depicted in Fig. D.2.4. Beams whose Poynting vector needs to reversed are levelled in the U-configuration (panel A). Beams needing to exit 90° from their input are routed through the 4-configuration (panel B). Beams needing to maintain their input Poynting vector are routed through the Z-configuration (panel C). The process of levelling a beam will not depend on the chosen configuration. Choose the configuration most needed for your application and table layout.

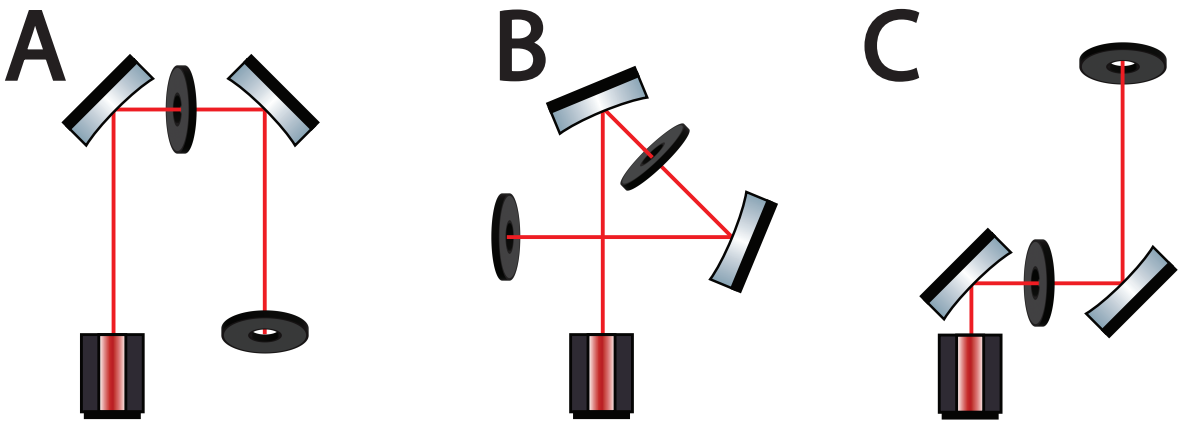

Figure D.3: Beam routing, mirrors, and irises for (A) the U-configuration (B) the 4-configuration, and (C) the Z-configuration.

Measure the incoming beam height with a ruler. We assume this need not change.

If it does, build a periscope. Indicate this position on the ruler with tape or a sharpie, or otherwise remember this height. Get two mirrors on kinematic mounts. Using the kinematic mounts, set them to be as level as possible so you will have the largest possible travel range for adjustments. Place the first mirror, keep the laser spot on the inner 1/3 of the mirror. Move the ruler as far away as possible. Tune the kinematic mirror until the beam is at the correct height.

Place the second mirror in the beam path such that the beam hits the mirror's inner 1/3. The further the second mirror is placed from the first, the easier alignment will be. Move the ruler as far away from the mirrors as possible and, once again, level the beam to the appropriate height by tuning the second mirror's kinematic mount.

Prepare two irises on posts. Take time to make sure they are both at the chosen height: the "levelness" of your outgoing beam will be limited by the heights. I suggest putting the irises as close as possible to each other, open the first one / close the back one, and slowly close the first one until you can identify how the one in the back is misaligned. Repeat until you are content. Place one iris close to the second mirror; place one iris as far from the two mirrors on the optical table as is possible. Adjust the first mirror in the optical path to ensure the laser goes through the center of the first iris. I recommend holding the ThorLabs spot-viewing card behind the iris. The beam should pass through the center of the viewing card.

Then, fully open the first iris and see where the laser lands on the second iris. Adjust the second mirror mount to make the laser go through the second iris. Go back to the first iris and see how far off the laser has become misaligned. Adjust the first mirror and continue iterating until you are content.

Note that you can adjust two mirrors at the same time (as Bill from ThorLabs does in the ThorLabs YouTube tutorial), especially if you can easily the spot. He watches the second iris and adjust the first mirror and second mirrors in opposite directions along the

same axis. This bypasses a lot of repetitive alignment steps. Note that you may have to redo the coarse alignment without the iris if you run out travel range.

Walking a beam can also be done to align two optical axes. It can be helpful to inject a secondary beamline from a visible laser into the beamline. If the optical axis are well matched, one can avoid cumbersome use of IR viewing scopes and beam viewing cards.

D.2.5 Background / advice: lenses and telescopes

The simplest form of optical lens we use is the plano-convex achromatic doublet. Our laser is pulsed and has a bandwidth of about 27 nm. Simple, monochromatic components are often too dispersive, so using achromatic doublets tends to be safer. In general, lenses with smaller radii and shorter focal lengths will perform better. We often use ThorLabs LA1131-B or LA1509-B. Make sure to use the wavelength-appropriate anti-reflective coating.

Note that the focal length is measured from the flat of the plano-side, not the center of the glass. One should always record the contents of mounted optical components with a label, but this is often not the case. A quick way to determine the focal length of a lens is to hold the lens with one hand and bring a planar object a measure able distance away. When the image of the building lights (whose light is semi-collimated) comes into focus, record the distance as the focal length.

Effort should be made to ensure that the beam passes through the center of the lens (of focal length F). As the beam moves from near the center (F/∞) to the edge ($\sim$F/2), the beam increasingly suffers from spherical aberration, where the different incoming rays focus to different lengths. A general rule of thumb: plano-convex lenses perform best when the incoming beam goes through a point no further than F/10.

Single lenses are used to focus light. The simplest combinations of two lenses is the

telescope, wherein two lenses of focal length F are separated by distance 2F. The first lens focuses incoming collimated light to the midpoint of the two lenses, after which the second lens collimates the outgoing beam. Telescopes in our setup are used for one of three purposes: (1) Adjusting the collimation. If the incident light to a telescope is not perfectly collimated, one can introduce small deviation δ to the distance between the two lenses (now $2F + \delta$) to correct for the collimation. (2) Telescopes are also used to focus light between the blades of a dense mechanical chopper wheel. The better focused the laser spot is, the higher frequency that the beam can be chopped, and the less the effect of prevalent 1/f noise becomes. (3) One can focus the light with a telescope and use an iris to remove non-circular beam components. Non-circular components of the beam can complicate PC switch scanning kernels, making alignment difficult.

Build a telescope by installing the first lens. Some incident light on his lens will be scattered back to the previous component (likely an iris). This is called the backreflection, and is key to installing lenses and beamsplitters. Position the lens such that the beam enters through the center of the convex side and that the backreflection is aligned to the incoming beam from the previous component. You may need to use the IR viewer, and alignment laser, or a beam viewer card. Place the second lens on a translation stage. Place this component approximately 2F from the first. Tune the position of the translation stage to produce a satisfactorily well-collimated beam at a long distance. This can be monitored by a CMOS camera (which may require $\geq$ND 6.0 attenuation). Place an iris at the center of the telescope, unless it is intended for a chopper. It is important to have the planar faces of the lenses facing the focus (not the collimated beams) to minimize spherical and chromatic aberration. This is depicted in Fig. D.2.5.

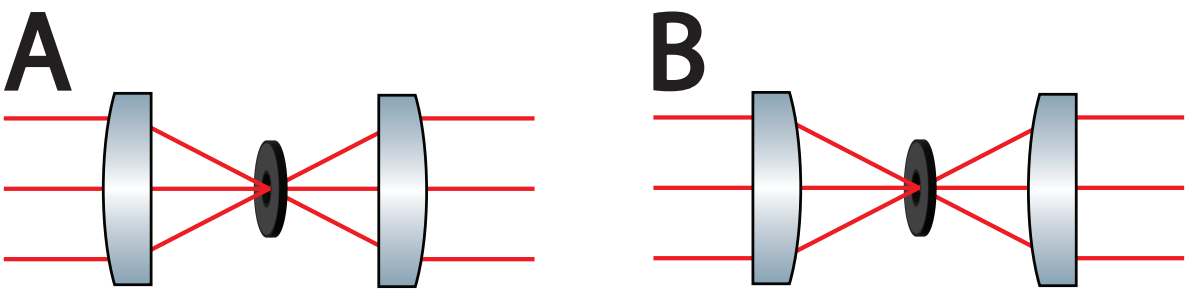

Figure D.4: Using plano-convex achromatic doublets in (A) a properly-built telescope and (B) an improperly-built telescope.

D.2.6 Advice: irises

As demonstrated above, irises with diaphragms are key components in building an aligning optical setups. Irises should be used liberally. I recommend buying many Thor-Labs D15S-P5 or Newport M-ID-1.0-PK iris packs, with tunable diameters from 1.5 mm to 25mm. Irises serve three primary functions in THz TDS setups, namely:

1. **Levelling and aligning beams**

   We have seen that irises are vital components in levelling / walking a beam. I recommend you to leave the "first" iris in from beam levelling, even after the beam has been levelled. This allows for easy corrections to misalignments.

2. **Correcting setup perturbations**

   Adding irises after each and every manually tunable optical mount allows one to easily correct for long term drift, mistakes made while tuning element positions, and component harvesting by members from other labs. This includes irises used when walking the beam, aligning beamsplitters, and aligning galvanometers. It is key to know which kinematic mount is used to tune misalignments appearing throughout the beamline. Assume there are N elements and N irises, with one iris installed per component on kinematic mount. Locate the earliest element in the

beamline where the misalignment error occurs (denoted as the N^{th} element). Use the N-1^{th} kinematic mount to tune for misalignment errors on the N^{th}. This is schematically shown in Fig. D.5, where misalignments in the colored irises can be removed by tuning the mirror mounts of the same color.

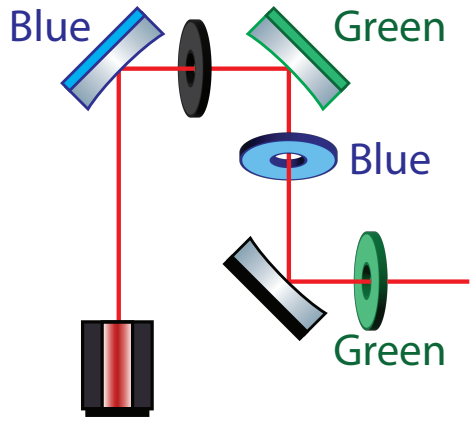

Figure D.5: Appropriate irises to adjust in to remove alignment errors. The mirror on the blue mount tunes the position of the laser beam passing through the blue iris, the mirror on the green mounts tunes the position of the laser beam passing through the green iris... etc.

3. **Spatially filtering a focused beam**

When placed at, say, the focus of a telescope, an iris removes non-ideal beam components that may have been caused by couplers or fiber collimators. If a circular beam shape is vital, an iris in the center of a telescope can be used to remove non-idealities.

It is wise to include one iris per tunable kinematic mount. In this way, the element(s) responsible for drift can easily be identified as the first in the beamline whose irises (per D.5) have beams not passing through the center. Iris effectively compartmentalize optical realignment: so one needs realign only individual components and not the whole table.

D.2.7 Background / advice: scanning galvanometers

A scanning galvanometer (galvo) is a 2-axis axis voltage-controlled 90°-periscope. The two periscope mirrors are controlled by two external DC voltage sources. I use ThorLabs GVS012, which has large (10 mm) silver mirrors and requires ±10V DC control. Installing the GVS012 requires several considerations that are not immediately obvious. The default cables are very short and not suitable for realistic setups. Long cables can sustain parasitic resonant currents that can drive measurable oscillations in the galvo Poynting. Mitigate this by ordering special made cables from ThorLabs, which have a long lead time. Note that GVS012 redirects the beam by 90° in the XY plane, rotates the polarization from S- to P- (or vice versa), and translates the beam height by about 1/2 inch. This is important when planning sample mounts and the beamline, as failure to do so will result in a misplaced beam. I recommend carefully measuring the travel range before finalizing the setup.

GVS012 has an internal jumper that allows one to set the voltage sensitivity S_V). The system defaults to 1°/V with a 20° scan range per axis. With careful installation, one can use the 0.5°/V scan range to increase scanning resolution. When a voltage bias V is applied, the beam displacement x at distance D is found by simple trigonometry: $x = D \tan S_v V$. Taylor-expanding the expression for small bias changes δV allows one to compute the scanning pixel size δx as $\delta x \approx D S_V \delta V$. Our DC voltage sources have a 20V range with 20 bits, and therefore 20 μV resolution. With D of 33 cm, scans have pixels that are approximately 3.3 μm x 3.3 μm, with 1 mV bias increments, corresponding to 16 fs error for the design described in Ch. 4. This is much smaller than the phase error associated with the automated realignment process. This error can be mitigated by filtering the input to the galvo controller, shortening the galvo - sample distance, and using ThorLabs-made cables.

I recommend machining a simple control box for the GVS012 control and power modules with two input BNC cables (X and Y voltage control) and three outputs (cable to the galvo unit, BNC X and Y feedback outputs). The default length cable provided by ThorLabs is very short. One can order longer (3m) cables, but may risk internal cable resonances (ThorLabs can build custom cables devoid of such resonances, but the lead time is long). For PC switch realignment scans, which need not be fast, I recommend adding simple low-pass RC filters with a 1 kHz pole to reduce noise from capacitive ringing.

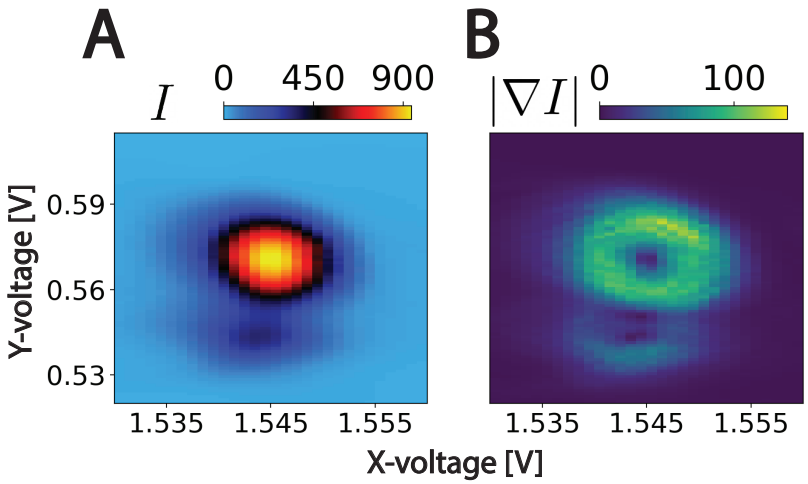

Figure D.6: Galvanometer scans of the A-switch of the spectrometer used in Ch. 4. (A) Raw current with 10V bias versus position (current in nA). (B) Magnitude of the current gradient (nA/V), showing regions of rapid change. The gradient ascent algorithm locates the PC switch at the center of the green circle, where the gradient is minimal. There is a small, secondary feature to the bottom left that is likely from a secondary reflection (Ch. 5).

Automated realignment is a key feature enabled by using a scanning galvanometer. Temperature-dependent length contraction of both the cold finger and switchboard shift the PC switch position by several mm over the course of 300K to 4K cooldown. Incorporating automated realignment ensures continuous data with better-than-manual alignment and a higher density of temperature points. We take a coarse map of the

PC switch current versus spot position, move the beam to the center, and use the galvanometer feedback voltages and simple gradient ascent algorithm to find the location of maximum current. The simplicity of PC switch design ensures that the maximum can be reliably found. This is best shown in a photocurrent and photocurrent gradient scan map, shown in Fig. D.2.7. This process can be run simultaneously for the emitter and A or B switch, as the lines are decoupled.

D.2.8 Advice: delay stage alignment

Delay stage alignment is the process which ensures that the PC switch spot position will not change as a function of delay stage position. This is vital to ensure the same operation conditions as a function of delay, so that the THz amplitudes at +0 ps and +100 ps are the same. The process essentially requires several iterations of beam levelling. A diagram is shown in Fig. D.2.8, with each stage of process indicated by different-colored irises.

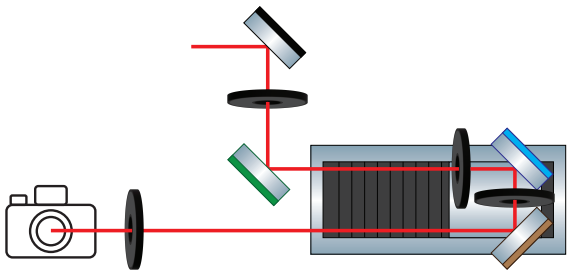

Figure D.7: Delay stage alignment setup. The stage moves continuously from the left (near) to the right (far) positions.

Begin by levelling the incoming beam using the black and green mirrors, with the delay stage in the far position. Add the cyan and brown mirrors, with an iris between them. Add an iris at a far position, the further the better. Behind this iris, place a

CMOS camera[1] with 5-10 μm pixel resolution.

Levelling the beam with the cyan and brown mirrors constitutes the majority of delay stage alignment. Move the delay stage to the far position and center the beam on the iris with the cyan mirror. Then, move the delay stage to the near position and center the beam on the iris with the brown mirror. Move the delay stage back to the far position and use the cyan mirror to align the beam once again. Repeat until the beam is centered on the iris in both the near and far positions. Open the iris and further refine the process using the CMOS camera. I recommend using ThorCam software and enabling thresholding, which highlights the spot maximum, allowing one to easily see the peak intensity as the delay stage moves.

To confirm that the delay stage is truly properly aligned, one should measure the current across a PC switch versus delay for the entire delay stage. If the alignment is proper, there should be minimal change in current versus delay. PC switch emitted THz signal / sensitivity to THz are linear in the photocurrent, so, no matter whether the emitter or readout PC switch beamline is delayed, errors are linear in photocurrent change versus delay. Alignment error can be reduced by restricting delay stage range in measurement.

This process can take several hours. It can be greatly sped up by: (1) enabling maximum delay stage acceleration and travel speed, (2) coarsely aligning the beam before using the CMOS camera, and (3) pre-emptively guessing the required alignment changes. After a few iterations of moving the delay stage, you should know the proper knob adjustment direction. Keep the stage in one position and adjust both mirrors simultaneously along their proper axes, keeping the spot fixed on the CMOS camera readout. Make adjustments 5 - 20 times more than you normally would. Move the delay stage and measure the results of your adjustment. This will save hours of time.

[1]We use a ThorLabs DCC1545M with an ND 6.0 filter to avoid pixel damage

It is very easy to confuse oneself of the proper adjustment direction per mirror. I recommend drawing arrows on the optics table with a dry-erase marker to avoid setting yourself further behind.

D.2.9 Advice / Debugging: true time zero

All forms of ultrafast pump-probe setups require the temporal overlap of the two beams. This can be time consuming and prone to error. Here, I outline three sequential steps for finding "true time zero" - the delay stage position at which the two beams temporally overlap on the sample.

1. **Coarse alignment - a tape measure**

   Build the delayed and non-delayed beamlines, recombining the two beams into a single optical axis with a beamsplitter[2]. Estimate the optical path length of the delayed beam path with a tape measure. I suggest positioning the delay stage at one of its extremes, zeroing the position, and moving the stage to its midpoint. This ensures maximal travel range in positioning. Measure the optical path length length of the non-delayed optical path. Add mirrors and other static elements to the non-delayed optical path to get the two optical path lengths to match within 1 - 2 cm.

2. **Intermediate alignment - fast photodiodes**

   Place the beamviewing camera next to the beamsplitter such that backreflections can be seen. Place a lens at the final beamsplitter output. Connect a fast Si PIN photodiode (Photonic solutions EOT 2000) to a spectrum analyzer capable of resolving the laser repetition frequency. Place this Si photodiode on a translation

[2]In our THz TDS setup, both of these beams should be routed through their respective scanning galvanometers.

stage and approximately align the sensor to the focused beams. This is greatly aided by the image from the beam camera (Fig. 2A), which clearly shows the beam and sensor positions. Block one beam and use the galvanometer to align one beam. Maximize the signal on the spectrum analyzer. Once maximized, block the other beam and repeat. When both beams are aligned, route the photodiode output to a fast oscilloscope capable of logging statistics. After waiting for upwards of 1000 averages, record the peak to peak voltage and the pulse width. The absolute value of these quantities is not meaningful, but their dependence on delay stage position is. As the two optical pulses arrive closer and closer to one another, the measured response will shrink in time and grow in amplitude. The width should reach a minimum near true time zero; the peak to peak ("pkpk") should reach a maximum near true time zero (Fig. 2B). If either parameter has a monotonic dependence on delay stage position, remeasure with a tape measure and restart the process.

3. **Fine alignment - Measure an optics signal**

   Once one confirms that true time zero is near the center of the optical delay stage, one should measure a THz transient (or other pump/probe signal). This will confirm the exact time of time zero. If you cannot find such a signal, confirm that the coarse and intermediate alignment steps were done correctly.

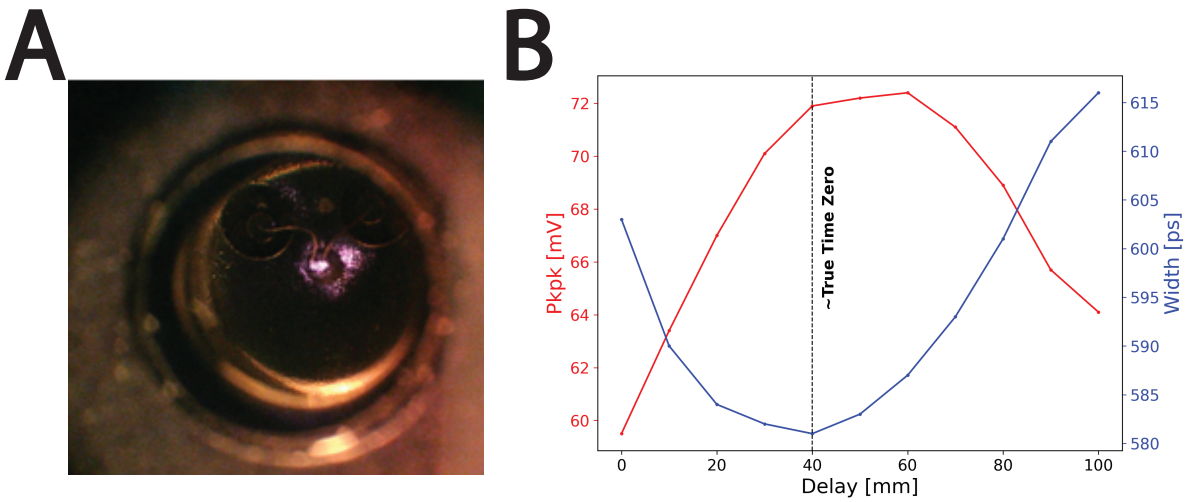

Figure D.8: Intermediate true time zero alignment with a fast photodiode. (A) Image of a Si PIN sensor, taken through a beamviewing camera through a beamsplitter. The sensor is illuminated with 8 mW from only one beamline. (B) Measurement of peak to peak voltage (pkpk) and pulse width versus delay. The pkpk voltage and width are both extremized near 40 mm, with a pkpk/width mismatch in reported time zero likely coming from insufficient data collection times. This is sufficient to progress to the next stage of alignment.

D.3 Recommended components

Table D.2 lists the set of key components we use. Additional laser goggles, posts, bases, shakers, mounts, mirrors, small-range delay stages, translation stages, cables, electronics ... etc. are not included. All prices are per unit and current as of 2023. Note that ThorLabs VRC2 is strictly not for 780 nm operation, but works reasonably well. VRC4 does not perform well for our laser, and ThorLabs VRC5 is laminated and risks poorly controlled reflections.

Name	Supplier	Part	Comment	Cost [USD]
Ultrafast laser	Menlo	ELMO780	Bare model	≈ 45,000
Flow cryostat	Lake Shore Cryotronics	SVP100	need LHe	31,690.00
Alignment Laser	ThorLabs	CPS532	few mW	176.33
Align. Laser Power	ThorLabs	CPS11K	n/a	222.95
Power Meter	ThorLabs	PM121D	Si sensor	1,625.14
Chopper controller	ThorLabs	MC2000B	n/a	1,391.73
Chopper blades	ThorLabs	MC2F53600	n/a	53.57
Galvo.	ThorLabs	GVS012	x2	3,468.21
Galvo. Heatstink	ThorLabs	GHS003	x2	31.45
Galvo. power	ThorLabs	GPS011-US	x2	571.52
Alignment camera	Amazon SVPRO	B096V84M57	x2	120.99
Optics camera	ThorLabs	DCC1545M	Obsolete	355.00
Delay stage	Newport	ESP302-1N	n/a	7,015
Delay stage	Newport	UTS100CC	n/a	5,364
Half-waveplate	ThorLabs	WPH10M-780	x2	610.27
Waveplate Mount	ThorLabs	RSP1X15	x2	152.57
Temp. sensing array	Amazon	B00WOSF82G	n/a	199.00
Humidity sensor	Amazon	Vertiv	n/a	137.71
Cryostat Z-stage	Newport	MVN80	Needs drives	1,205.00
Z-stage drive	Newport	BM17.25	n/a	243.00
Z-stage baseplate	Newport	M-PBN8	n/a	111.00
XY stages	Newport	426	x2	620.00
XY stage drives	Newport	SM-25	x2	153.00
Longpass filter	ThorLabs	FELH0600	Metric	140.99
Shortpass filter	ThorLabs	FESH1000	Metric	140.99
Lenses	ThorLabs	LA1131B	F = 50 mm	37.28
Lenses	ThorLabs	LA1509-B	F = 100 mm	36.11
Infrared viewers	Edmunds	36-084	n/a	2,335.00
Beam viewcard	ThorLabs	VRC2	See caption	94.05
Periscope assembly	ThorLabs	RS99	No mirrors	331.92
Autocorrelator	ThorLabs	FSAC	Needs oscilloscope	7,113.75
Better autocorrelator	FemtoEasy	ROC-PS1 FS10	Standalone	≈15,000
Si PIN photodiode	Photonic Solutions	ET2000	Standalone	338.00

Table D.2: Component suggestions

Appendix E

Symbols and units

> "By relieving the brain of all unnecessary work, a good notation sets it free to concentrate on more advanced problems, and in effect increases the mental power of the race."
>
> *Alfred North Whitehead*

Real and complex-valued scalar quantities are listed with no special notation (example: ρ_e). Vector quantities are indicated with an overline (example: $\overline{E}$) and dyadic tensors are indicated with two overlines (example: $\overline{\overline{\epsilon}}$). Complex quantites are often split into real and imaginary parts, with imaginary unit j. Real parts acquire a (1) superscript; imaginary parts acquire a (2) superscript. For example, consider $\sigma(\omega) = \sigma_1(\omega) + j\sigma_2(\omega)$. σ_1 is the real part of σ; σ_2 is the imaginary part of σ. Complex conjugation is denoted by the operator *.

We use signal processing Fourier transform conventions. These define the Fourier transform for functions $f(t) \Leftrightarrow F(\omega)$ as:

$$F(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(t)e^{-j\omega t}$$
$$f(t) = \int_{-\infty}^{\infty} F(\omega)e^{j\omega t} \tag{E.1}$$

A derivative with respect to parameter x is indicated by ∂_x or $\partial/\partial x$. Per eqn. E.1, differentiation in the time domain domains corresponds to multiplying by $j\omega$ in the frequency domain. This can be thought of as assuming time-harmonic quantities take monochromatic time-dependent form $e^{j\omega t}$. If a quantity has both time and frequency domain quantities, time domain values are in lower case (example: $\overline{e}$) and frequency domain quantities are in upper case (example: $\overline{E}$).

Bibliography

[1] A. M. Potts, S. Bajaj, D. R. Daughton, A. A. Allerman, A. M. Armstrong, T. Razzak, S. H. Sohel, and S. Rajan, *$Al_{0.7}ga_{0.3}n$ mesfet with all-refractory metal process for high temperature operation*, *IEEE Transactions on Electron Devices* **68** (9, 2021) 4278–4282.

[2] ThorLabs, *Fsac interferometric autocorrelator user guide*, tech. rep., ThorLabs, 2017.

[3] C. A. Balanis, *Advanced Engineering Electrodynamics*. John Wiley & Sons, 2 ed., 2012.

[4] D. Griffiths, *Introduction to Electrodynamics*. Pearson, 3 ed., 2013.

[5] D. Pozar, *Microwave Engineering*. John Wiley & Sons, 3 ed., 2012.

[6] P. Gallagher, C. S. Yang, T. Lyu, F. Tian, R. Kou, H. Zhang, K. Watanabe, T. Taniguchi, and F. Wang, *Quantum-critical conductivity of the dirac fluid in graphene*, *Science* **364** (4, 2019) 158–162.

[7] R. Simons, *Coplanar Waveguide Circuits, Components, and Systems*. John Wiley & Sons, 1 ed., 2001.

[8] J. O. Island, P. Kissin, J. Schalch, X. Cui, S. R. U. Haque, A. Potts, T. Taniguchi, K. Watanabe, R. D. Averitt, and A. F. Young, *On-chip terahertz modulation and emission with integrated graphene junctions*, *Applied Physics Letters* **116** (4, 2020) 161104.

[9] A. M. Potts, A. K. Nayak, M. Nagel, K. Kaj, B. Stamenic, D. D. John, R. D. Averitt, and A. F. Young, *On-chip time-domain terahertz spectroscopy of superconducting films below the diffraction limit*, *Nano Letters* (5, 2023).

[10] L. Smith, W. Gomma, H. Esmaeilsabzali, T. Darcie, and T. Darcie, *Tapered transmission lines for terahertz systems*, *Optics Express, Vol. 29, Issue 11, pp. 17295-17303* **29** (5, 2021) 17295–17303.

[11] F. Wooten, *Optical Properties of Solids*. Academic Press, Inc., 1972.

[12] J. Bechhoefer, *Kramers–kronig, bode, and the meaning of zero*, *American Journal of Physics* **79** (9, 2011) 1053.

[13] R. K. Ahrenkiel, *Modified kramers-kronig analysis of optical spectra*, *Journal of the Optical Society of America* **61** (12, 1971).

[14] E. Hagen and H. Rubens, *Über beziehungen des reflexions- und emissionsvermögens der metalle zu ihrem elektrischen leitvermögen*, *Annalen der Physik* **316** (1903) 873–901.

[15] M. Mitrano, A. Cantaluppi, D. Nicoletti, S. Kaiser, A. Perucchi, S. Lupi, P. D. Pietro, D. Pontiroli, M. Riccò, S. R. Clark, D. Jaksch, and . A. Cavalleri, *Possible light-induced superconductivity in k 3 c 60 at high temperature*, *Nature* (2016).

[16] G. G. Raju, *Dielectrics in Electric Fields : Tables, Atoms, and Molecules*. CRC Press, 7, 2017.

[17] A. K. Jonscher, *The 'universal' dielectric response*, *Nature 1977 267:5613* **267** (6, 1977) 673–679.

[18] A. K. Jonscher and R. Holloway, *Dielectric relaxation in solids*, *Journal of Physics D: Applied Physics* **32** (7, 1999) R57.

[19] A. K. Jonscher, *The universal dielectric response and its physical significance*, *IEEE Transactions on Electrical Insulation* **27** (1992) 407–423.

[20] A. Zangwill, *Modern Electrodynamics*. Cambridge University Press, 1 ed., 2012.

[21] L. Landau and E. Lifshitz, *Electrodynamics of Continuous Media*, vol. 8. Pergamon Press, 2 ed., 1960.

[22] R. A. Ferrell and R. E. Glover, *Conductivity of superconducting films: A sum rule*, *Physical Review* **109** (2, 1958) 1398.

[23] F. W. King, *Sum rules for the optical constants*, *Journal of Mathematical Physics* **17** (8, 1976) 1509.

[24] K. S. Cole and R. H. Cole, *Dispersion and absorption in dielectrics i. alternating current characteristics*, *The Journal of Chemical Physics* **9** (12, 1941) 341.

[25] G. Bossé, L. S. Bilbro, R. V. Aguilar, L. Pan, W. Liu, A. V. Stier, Y. Li, L. H. Greene, J. Eckstein, and N. P. Armitage, *Low energy electrodynamics of the kondo-lattice antiferromagnet cecu 2 ge 2*, *Physical Review B* **85** (2012) 155105.

[26] P. Zheng, C. Wang, Y. Xu, L. Wang, W. Wu, Y. G. Shi, Y.-F. Yang, and J. L. Luo, *Uniaxial hybridization in the quasi-one-dimensional kondo lattice* $ceco_2ga_8$, *Physical Review B* **105** (2022) 35112.

[27] A. Pustogow, Y. Saito, A. Löhle, M. S. Alonso, A. Kawamoto, V. Dobrosavljević, M. Dressel, and S. Fratini, *Rise and fall of landau's quasiparticles while approaching the mott transition*, *Nature Communications 2021 12:1* **12** (3, 2021) 1–8.

[28] Y. Wang, H. P. Nair, N. J. Schreiber, J. P. Ruf, B. Cheng, D. G. Schlom, K. M. Shen, and N. P. Armitage, *Separated transport relaxation scales and interband scattering in thin films of srruo$_3$, caruo$_3$, and sr$_2$ruo$_4$*, *Physical Review B* **103** (2021) 205109.

[29] Y. Wang, G. Bossé, H. P. Nair, N. J. Schreiber, J. P. Ruf, B. Cheng, C. Adamo, D. E. Shai, Y. Lubashevsky, D. G. Schlom, K. M. Shen, and N. P. Armitage, *Subterahertz momentum drag and violation of matthiessen's rule in an ultraclean ferromagnetic srruo3 metallic thin film*, *Physical Review Letters* **125** (11, 2020) 217401.

[30] M. L. Thèye, *Investigation of the optical properties of au by means of thin semitransparent films*, *Physical Review B* **2** (10, 1970) 3060.

[31] J. W. Allen and J. C. Mikkelsen, *Optical properties of crsb, mnsb, nisb, and nias*, *Physical review B* **15** (1977) 2952–2960.

[32] J. J. Hopfield, H. C. Wolfe, and D. H. Douglass, *Infrared properties of transition metals*, *AIP Conference Proceedings* **4** (6, 1972) 358–366.

[33] G. Bossé, L. Pan, Y. S. Li, L. H. Greene, J. Eckstein, and N. P. Armitage, *Anomalous frequency and temperature-dependent scattering and hund's coupling in the almost quantum critical heavy-fermion system cefe2ge2*, *Physical Review B* **93** (2, 2016).

[34] R. Brendel and D. Bormann, *An infrared dielectric function model for amorphous solids*, *Journal of Applied Physics* **71** (8, 1998) 1.

[35] J. Orosco and C. F. Coimbra, *Optical response of thin amorphous films to infrared radiation*, *Physical Review B* **97** (3, 2018) 094301.

[36] C. F. M. Coimbra and J. Orosco, *On a causal dispersion model for the optical properties of metals*, *Applied Optics, Vol. 57, Issue 19, pp. 5333-5347* **57** (7, 2018) 5333–5347.

[37] C. R. Gautam, A. K. Yadav, C. R. Gautam, and A. K. Yadav, *Synthesis and optical investigations on (ba,sr)tio3 borosilicate glasses doped with la2o3*, *Optics and Photonics Journal* **3** (8, 2013) 1–7.

[38] T. N. Le, J.-L. Pelouard, F. Charra, and S. Vassant, *Determination of the far-infrared dielectric function of a thin ingaas layer using a detuned salisbury screen*, *Optical Materials Express* **12** (2022) 2711.

[39] M. Mardar, *Condensed Matter Physics*. John Wiley & Sons, 1 ed., 2010.

[40] L. Landau and E. Lifshitz, *Statistical Physics, Part 2: Theory of the Condensed*, vol. 9. Pergamon Press, 2 ed., 1969.

[41] M. Tinkham, *Introduction to Superconductivity*. McGraw-Hill, New York, 2nd ed., 1975.

[42] W. Zimmermann, E. H. Brandt, M. Bauer, E. Seider, and L. Genzel, *Optical conductivity of bcs superconductors with arbitrary purity*, *Physica C: Superconductivity* **183** (11, 1991) 99–104.

[43] F. Mahmood, X. He, I. BoŽović, and N. P. Armitage, *Locating the missing superconducting electrons in the overdoped cuprates la2-xsrxcuo4*, *Physical Review Letters* **122** (1, 2019) 027003.

[44] D. Fausti, R. I. Tobey, N. Dean, S. Kaiser, A. Dienst, M. C. Hoffmann, S. Pyon, T. Takayama, H. Takagi, and A. Cavalleri, *Light-induced superconductivity in a stripe-ordered cuprate*, *Science* **331** (1, 2011) 189–191.

[45] A. Cantaluppi, M. Buzzi, G. Jotzu, D. Nicoletti, M. Mitrano, D. Pontiroli, M. Riccò, A. Perucchi, P. D. Pietro, and A. Cavalleri, *Pressure tuning of light-induced superconductivity in k3c60*, *Nature Physics 2018 14:8* **14** (5, 2018) 837–841.

[46] E. B. W. Jr., J. C. Decius, and P. C. Cross, *Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra*. Dover Publications, 3, 1980.

[47] S. Sugai, Y. Takayanagi, and N. Hayamizu, *Phason and amplitudon in the charge-density-wave phase of one-dimensional charge stripes in la2-xsrxcuo4*, *Physical Review Letters* **96** (2006).

[48] D. H. Torchinsky, F. Mahmood, A. T. Bollinger, I. Božović, and N. Gedik, *Fluctuating charge-density waves in a cuprate superconductor*, *Nature Materials* **12** (5, 2013) 387–391.

[49] A. V. Shchepetilnikov, P. A. Gusikhin, V. M. Muravev, G. E. Tsydynzhapov, Y. A. Nefyodov, A. A. Dremin, and I. V. Kukushkin, *New ultra-fast sub-terahertz linear scanner for postal security screening*, *Journal of Infrared, Millimeter, and Terahertz Waves* **41** (6, 2020) 655–664.

[50] A. Luukanen, L. Grönberg, T. Haarnoja, P. Helistö, K. Kataja, M. Leivo, A. Rautiainen, J. Penttilä, J. E. Bjarnason, C. R. Dietlein, M. D. Ramirez, and E. N. Grossman, *Passive thz imaging system for stand-off identification of concealed objects: results from a turn-key 16 pixel imager*, in *SPIE Defense and Security Symposium*, vol. 6948, p. 69480O, SPIE, 4, 2008.

[51] C. A. Martin, J. A. Lovberg, W. H. Dean, and E. Ibrahim, *High-resolution passive millimeter-wave security screening using few amplifiers*, in *Defense and Security Symposium*, vol. 6548, pp. 6548061–654806–10, SPIE, 4, 2007.

[52] T. May, G. Zieger, S. Anders, V. Zakosarenko, H.-G. Meyer, M. Schubert, M. Starkloff, M. Rößler, G. Thorwirth, and U. Krause, *Safe visitor: visible, infrared, and terahertz object recognition for security screening application*, in *SPIE Defense, Security, and Sensing*, vol. 7309, p. 73090E, SPIE, 5, 2009.

[53] D. T. Petkie, C. Casto, F. C. D. Lucia, S. R. Murrill, B. Redman, R. L. Espinola, C. C. Franck, E. L. Jacobs, S. T. Griffin, C. E. Halford, J. Reynolds, S. O'brien, and D. Tofsted, *Active and passive imaging in the thz spectral region: phenomenology, dynamic range, modes, and illumination*, *Journal of the Optical Society of America B* **25** (2008) 1523–1531.

[54] K. Kawase, Y. Ogawa, Y. Watanabe, H. Inoue, P. H. Siegel, J. E. Parmeter, D. W. Murray, and D. W. Hannum, *Non-destructive terahertz imaging of illicit drugs using spectral fingerprints*, *Optics Express* **11** (2003) 2549–2554.

[55] Y. C. Shen, T. Lo, P. F. Taday, B. E. Cole, W. R. Tribe, and M. C. Kemp, *Detection and identification of explosives using terahertz pulsed spectroscopic imaging*, *Applied Physics Letters* **86** (6, 2005) 1–3.

[56] A. J. Fitzgerald, B. E. Cole, and P. F. Taday, *Nondestructive analysis of tablet coating thicknesses using terahertz pulsed imaging*, *Journal of Pharmaceutical Sciences* **94** (2005) 177–183.

[57] D. M. Mittleman, R. H. Jacobsen, and M. C. Nuss, *T-ray imaging*, *IEEE Journal of selected topics in quantum electronics* **2** (1996).

[58] M. Strag and W. Swiderski, *Defect detection in aramid fiber-reinforced composites via terahertz radiation*, *Journal of Nondestructive Evaluation* **42** (3, 2023).

[59] M. Zhong, B. Liu, C. Li, Z. Wang, D. Wei, B. Zhou, X. Dai, and Y. Xu, *Terahertz spectroscopy and imaging detection of defects in civil aircraft composites*, *Journal of Spectroscopy* **2020** (2020).

[60] M. Bauer, R. Hussung, C. Matheis, H. Reichert, P. Weichenberger, J. Beck, U. Matuschczyk, J. Jonuscheit, and F. Friederich, *Fast fmcw terahertz imaging for in-process defect detection in press sleeves for the paper industry and image evaluation with a machine learning approach*, *Sensors* **21** (10, 2021).

[61] X. Yang, P. Liu, S. Wang, B. Wu, K. Zhang, B. Yang, and X. Wu, *Defect detection of composite material terahertz image based on faster region-convolutional neural networks*, *Materials* **16** (1, 2023).

[62] D. N. Basov, R. D. Averitt, D. V. D. Marel, M. Dressel, and K. Haule, *Electrodynamics of correlated electron materials*, *Reviews of Modern Physics* **83** (6, 2011) 471–541.

[63] D. N. Basov, R. D. Averitt, and D. Hsieh, *Towards properties on demand in quantum materials*, *Nature Materials 2017 16:11* **16** (10, 2017) 1077–1088.

[64] L. Shi, P. Shumyatsky, A. Rodríguez-Contreras, and R. R. Alfano, *Terahertz spectroscopy of brain tissue from a mouse model of alzheimer's disease*, *https://doi.org/10.1117/1.JBO.21.1.015014* **21** (1, 2016) 015014.

[65] Y. Peng, C. Shi, X. Wu, Y. Zhu, and S. Zhuang, *Terahertz imaging and spectroscopy in cancer diagnostics: A technical review*, *BME Frontiers* **2020** (9, 2020).

[66] P. H. Siegel, *Terahertz technology in biology and medicine*, *IEEE Transactions on Microwave Theory and Techniques* **52** (2004).

[67] R. Woodward, V. Wallace, B. Cole, R. Pye, D. Arnone, al M Ruth Woodward, V. P. Wallace, B. E. Cole, R. J. Pye, D. D. Arnone, and E. H. Linfield, *Terahertz pulse imaging in reflection geometry of skin tissue using time-domain analysis techniques*, *https://doi.org/10.1117/12.469785* **4625** (6, 2002) 160–169.

[68] A. J. Fitzgerald, V. P. Wallace, M. Jimenez-Linan, L. Bobrow, R. J. Pye, A. D. Purushotham, and D. D. Arnone, *Terahertz pulsed imaging of human breast tumors*, *https://doi.org/10.1148/radiol.2392041315* **239** (5, 2006) 533–540.

[69] P. C. Ashworth, E. Pickwell-MacPherson, E. Provenzano, S. E. Pinder, A. D. Purushotham, M. Pepper, V. P. Wallace, P. Y. Han, G. C. Cho, X. C. Zhang, P. H. Bolívar, M. Nagel, F. Richter, M. Brucherseifer, H. Kurz, A. Bosserhoff, and R. Büttner, *Terahertz pulsed spectroscopy of freshly excised human breast cancer*, *Optics Express, Vol. 17, Issue 15, pp. 12444-12454* **17** (7, 2009) 12444–12454.

[70] V. P. Wallace, A. J. Fitzgerald, S. Shankar, N. Flanagan, R. Pye, J. Cluff, and D. D. Arnone, *Terahertz pulsed imaging of basal cell carcinoma ex vivo and in vivo*, *British Journal of Dermatology* **151** (8, 2004) 424–432.

[71] V. P. Wallace, A. J. Fitzgerald, E. Pickwell, R. J. Pye, P. F. Taday, N. Flanagan, and H. A. Thomas, *Terahertz pulsed spectroscopy of human basal cell carcinoma*, *http://dx.doi.org/10.1366/000370206778664635* **60** (8, 2016) 1127–1133.

[72] D. Mittleman, *Terahertz Imaging*. Springer, Berlin, Heidelberg, 2003.

[73] K. Fukunaga and M. Picollo, *Terahertz spectroscopy applied to the analysis of artists' materials*, *Applied Physics A: Materials Science and Processing* **100** (9, 2010) 591–597.

[74] J. Labaune, J. B. Jackson, S. Pagès-Camagna, I. N. Duling, M. Menu, and G. A. Mourou, *Papyrus imaging with terahertz time domain spectroscopy*, *Applied Physics A: Materials Science and Processing* **100** (9, 2010) 607–612.

[75] J. Tennyson and O. L. Polyansky, *Water on the sun: The sun yields more secrets to spectroscopy*, *http://dx.doi.org/10.1080/001075198181955* **39** (2010) 283–294.

[76] L. Wiesenfeld and A. Faure, *Ab initio computation of the broadening of water rotational lines by molecular hydrogen*, *RAPID COMMUNICATIONS PHYSICAL REVIEW A* **82** (2010) 40702.

[77] A. R. Conrad, *Rotational spectroscopy of biomolecules*, tech. rep., Kent State University, 2011.

[78] W. Zhao, S. Wang, S. Chen, Z. Zhang, K. Watanabe, T. Taniguchi, A. Zettl, and F. Wang, *Observation of hydrodynamic plasmons and energy waves in graphene*, *Nature 2023 614:7949* **614** (2, 2023) 688–693.

[79] J. W. McIver, B. Schulte, F. U. Stein, T. Matsuyama, G. Jotzu, G. Meier, and A. Cavalleri, *Light-induced anomalous hall effect in graphene*, *Nature Physics 2019 16:1* **16** (11, 2019) 38–41.

[80] J. S. Dodge, L. Lopez, and D. G. Sahota, *Photoinduced superconductivity reconsidered: The role of photoconductivity profile distortion*, *arXiv* (2022).

[81] H. Niwa, N. Yoshikawa, K. Tomari, R. Matsunaga, D. Song, H. Eisaki, and R. Shimano, *Light-induced nonequilibrium response of the superconducting cuprate* $la_{2-x}sr_xcuo_4$, *Physical Review B* **100** (2019) 104507.

[82] E. Wang, J. D. Adelinia, M. Chavez-Cervantes, T. Matsuyama, M. Fechner, M. Buzzi, G. Meier, and A. Cavalleri, *Nonlinear transport in a photo-induced superconductor*, 2023.

[83] D. Grischkowsky, I. N. Duling, J. C. Chen, and C. C. Chi, *Electromagnetic shock waves from transmission lines*, *Physical Review Letters* **59** (10, 1987) 1663.

[84] H. J. Cheng, J. F. Whitaker, T. M. Weller, and L. P. Katehi, *Terahertz-bandwidth characteristics of coplanar transmission lines on low permittivity substrates*, *IEEE Transactions on Microwave Theory and Techniques* **42** (1994) 2399–2406.

[85] S. P. Morgan, *Mode conversion losses in transmission of circular electric waves through slightly non-cylindrical guides*, *Journal of Applied Physics* **21** (1950) 329–338.

[86] E. O. Hammerstad and F. Bekkadal, *Microstrip Handbook*. Electronics Research Laboratory, University of Trondheim, Norwegian Institute of Technology, 1 ed., 1975.

[87] Z. Wu and L. E. Davis, *Surface roughness effect on surface impedance of superconductors*, *Journal of Applied Physics* **76** (1994) 3669–3672.

[88] A. B. Pippard, *The surface impedance of superconductors and normal metals at high frequencies iii. the relation between impedance and superconducting penetration depth*, *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **191** (11, 1947) 399–415.

[89] A. B. Pippard, *The surface impedance of superconductors and normal metals at high frequencies i. resistance of superconducting tin and mercury at 1200 mcyc. /sec*, *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **191** (11, 1947) 370–384.

[90] A. Pippard, *The surface impedance of superconductors and normal metals at high frequencies ii. the anomalous skin effect in normal metals*, *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **191** (11, 1947) 385–399.

[91] G. E. Reuter and E. H. Sondheimer, *Theory of the anomalous skin effect in metals*, *Nature 1948 161:4089* **161** (1948) 394–395.

[92] R. G. Chambers, J. Reuter, and P. R. Soc, *Anomalous skin effect in metals*, *Nature 1950 165:4189* **165** (1950) 239–240.

[93] S. Lucyszyn, *Evaluating surface impedance models for terahertz frequencies at room temperature*, *Piers Online* **3** (2007).

[94] D. Gall, *Electron mean free path in elemental metals*, *Journal of Applied Physics* **119** (2, 2016) 085101.

[95] C. Svensson and G. E. Dermer, *Time domain modeling of lossy interconnects*, *IEEE Transactions on advanced packaging* **24** (2001).

[96] A. R. Djordjević, R. M. Biljić, V. D. Likar-Smiljanić, and T. K. Sarkar, *Wideband frequency-domain characterization of fr-4 and time-domain causality*, *IEEE Transactions on Electromagnetic Compatibility* **43** (11, 2001) 662–667.

[97] J. Hemberger, P. Lunkenheimer, R. Viana, R. Bohmer, and A. Loidl, *Electric-field-dependent dielectric constant and nonlinear susceptibility in srtio3*, .

[98] P. Heydari, *Terahertz integrated circuits and systems for high-speed wireless communications: Challenges and design perspectives*, *IEEE Open Journal of the Solid-State Circuits Society* **1** (10, 2021) 18–36.

[99] A. Tsuchiya and H. Onodera, *Impact of radiation loss in on-chip transmission-line for terahertz applications*, *2012 IEEE 16th Workshop on Signal and Power Integrity, SPI 2012 - Proceedings* (2012) 125–128.

[100] M. Y. Frankel, S. Gupta, J. A. Valdmanis, and G. A. Mourou, *Terahertz attenuation and dispersion characteristics of coplanar transmission lines*, *IEEE Transactions on Microwave Theory and Techniques* **39** (1991) 910–916.

[101] Z. Wang, F. Su, F. A. Hegmann, L. Razzari, X. Ropagnol, H. c Bandulet, F. Vidal, R. Morandotti, J. c Kieffer, T. Ozaki, H. Tiedje, H. Haugen, M. Reid, and F. A. Hegmann, *Ultrafast imaging of terahertz cherenkov waves and transition-like radiation in linbo¡sub¿3¡/sub¿*, *Optics Express, Vol. 23, Issue 6, pp. 8073-8086* **23** (3, 2015) 8073–8086.

[102] D. H. Auston, *Subpicosecond electro-optic shock waves*, *Applied Physics Letters* **43** (1983) 713–715.

[103] D. H. Auston, K. P. Cheung, J. A. Valdmanis, and D. A. Kleinman, *Cherenkov radiation from femtosecond optical pulses in electro-optic media*, *Physical Review Letters* **53** (10, 1984) 1555.

[104] D. A. Kleinman and D. H. Auston, *Theory of electrooptic shock radiation in nonlinear optical media*, *IEEE Journal of Quantum Electronics* **20** (1984) 964–970.

[105] B. B. Hu, X. C. Zhang, D. H. Auston, and P. R. Smith, *Free-space radiation from electro-optic crystals*, *Applied Physics Letters* **56** (6, 1998) 506.

[106] D. R. Grischkowsky, *Optoelectronic characterization of transmission lines and waveguides by terahertz time-domain spectroscopy*, *IEEE Journal on Selected Topics in Quantum Electronics* **6** (11, 2000) 1122–1135.

[107] R. Smith and T. Darcie, *Demonstration of a low-distortion terahertz system-on-chip using a cps waveguide on a thin membrane substrate*, *Optics Express, Vol. 27, Issue 10, pp. 13653-13663* **27** (5, 2019) 13653–13663.

[108] H. Shigesawa, M. Tsuji, and A. A. Oliner, *Simultaneous propagation of bound and leaky dominant modes on printed-circuit lines: A new general effect*, *IEEE Transactions on Microwave Theory and Techniques* **43** (1995) 3007–3019.

[109] U. D. Keil, D. R. Dykaar, A. F. Levi, R. F. Kopf, L. N. Pfeiffer, S. B. Darack, and K. W. West, *High-speed coplanar transmission lines*, *IEEE Journal of Quantum Electronics* **28** (1992) 2333–2342.

[110] J. W. Chung, W. E. Hoke, E. M. Chumbes, and T. Palacios, *Algan/gan hemt with 300-ghz fmax*, *IEEE Electron Device Letters* **31** (3, 2010) 195–197.

[111] C. Yoo, M. Huang, J. H. Kawamura, K. W. West, L. N. Pfeiffer, B. S. Karasik, and M. S. Sherwin, *Demonstration of a tunable antenna-coupled intersubband terahertz (tacit) mixer*, *Applied Physics Letters* **116** (1, 2020) 013504.

[112] D. A. Bandurin, D. Svintsov, I. Gayduchenko, S. G. Xu, A. Principi, M. Moskotin, I. Tretyakov, D. Yagodkin, S. Zhukov, T. Taniguchi, K. Watanabe, I. V. Grigorieva, M. Polini, G. N. Goltsman, A. K. Geim, and G. Fedorov, *Resonant terahertz detection using graphene plasmons*, *Nature Communications 2018 9:1* **9** (12, 2018) 1–8.

[113] J. B. Costello, S. D. O'Hara, Q. Wu, D. C. Valovcin, L. N. Pfeiffer, K. W. West, and M. S. Sherwin, *Reconstruction of bloch wavefunctions of holes in a semiconductor*, *Nature 2021 599:7883* **599** (11, 2021) 57–61.

[114] S. Ariyoshi, C. Otani, A. Dobroiu, H. Sato, K. Kawase, H. M. Shimizu, T. Taino, and H. Matsuo, *Terahertz imaging with a direct detector based on superconducting tunnel junctions*, *Applied Physics Letters* **88** (5, 2006) 203503.

[115] N. V. Vvedenskii, A. I. Korytin, V. A. Kostin, A. A. Murzanev, A. A. Silaev, and A. N. Stepanov, *Two-color laser-plasma generation of terahertz radiation using a frequency-tunable half harmonic of a femtosecond pulse*, *Physical Review Letters* **112** (2, 2014) 055004.

[116] D. R. Daughton, R. Higgins, and S. Yano, *Field and temperature-dependent thin film characterization with a continuous-wave terahertz magneto-spectrometer*, *International Conference on Infrared, Millimeter, and Terahertz Waves, IRMMW-THz* (11, 2014).

[117] S. Saxena, S. Bagchi, M. Tayyab, J. A. Chakera, S. Kumar, and D. N. Gupta, *Enhanced broadband terahertz radiation from two-colour laser pulse interaction with thin dielectric solid target in air*, *Journal of Infrared, Millimeter, and Terahertz Waves* **42** (7, 2021) 747–760.

[118] D. H. Auston, *Picosecond optoelectronic switching and gating in silicon*, *Applied Physics Letters* **26** (1975) 101–103.

[119] M. C. Beard, G. M. Turner, and C. A. Schmuttenmaer, *Subpicosecond carrier dynamics in low-temperature grown gaas as measured by time-resolved terahertz spectroscopy*, *Journal of Applied Physics* **90** (12, 2001) 5915–5923.

[120] M. R. Melloch, J. M. Woodall, E. S. Harmon, N. Otsuka, F. H. Pollak, D. D. Nolte, R. M. Feenstra, and M. A. Lutz, *Low-temperature grown iii-v materials*, *Annu. Rev. Mater. Sci* **25** (1995) 547–600.

[121] R. Jiang, S. Cheng, and Q. Li, *Near-band-edge photoluminescence of gaas epitaxial layers grown at low temperature*, *Japanese Journal of Applied Physics To* **36** (1997) 623.

[122] Y. P. Varshni, *Temperature dependence of the energy gap in semiconductors*, *Physica* **34** (1, 1967) 149–154.

[123] K. Wiesauer and C. Jördens, *Recent advances in birefringence studies at thz frequencies*, *Journal of Infrared, Millimeter, and Terahertz Waves* **34** (11, 2013) 663–681.

[124] P. Arifin, E. M. Goldys, and T. L. Tansley, *Electron mobility in low temperature grown gallium arsenide*, *Materials Science and Engineering: B* **35** (1995) 330–333.

[125] M. Griebel, J. H. Smet, D. C. Driscoll, J. Kuhl, C. A. Diez, N. Freytag, C. Kadow, A. C. Gossard, and K. V. Klitzing, *Tunable subpicosecond optoelectronic transduction in superlattices of self-assembled eras nanoislands*, *Nature Materials* **2** (2003) 122–126.

[126] J. F. O'Hara, J. M. Zide, A. C. Gossard, A. J. Taylor, and R. D. Averitt, *Enhanced terahertz detection via eras:gaas nanoisland superlattices*, *Applied Physics Letters* **88** (6, 2006).

[127] A. Mingardi, W. D. Zhang, and E. R. Brown, *Thz performance of 1550-nm-driven photoconductive switches made from gaas:er with eras quantum dots*, *2017 IEEE National Aerospace and Electronics Conference (NAECON)* (2017).

[128] C. Kadow, S. B. Fleischer, J. P. Ibbetson, J. E. Bowers, A. C. Gossard, J. W. Dong, and C. J. Palmstrøm, *Self-assembled eras islands in gaas: Growth and subpicosecond carrier dynamics*, *Applied Physics Letters* **75** (11, 1999) 3548–3550.

[129] E. R. Brown, A. Mingardi, W. D. Zhang, A. D. Feldman, T. E. Harvey, and R. P. Mirin, *Abrupt dependence of ultrafast extrinsic photoconductivity on er fraction in gaas:er*, *Applied Physics Letters* **111** (7, 2017).

[130] J. R. Middendorf and E. R. Brown, *A study of ultrafast extrinsic photoconductivity vs wavelength in eras:gaas photoconductive switches*, *2013 38th International Conference on Infrared, Millimeter, and Terahertz Waves (IRMMW-THz)* (2013).

[131] J. Y. Kim, X. Ju, K. W. Ang, and D. Chi, *Van der waals layer transfer of 2d materials for monolithic 3d electronic system integration: Review and outlook*, *ACS Nano* **17** (2, 2023) 1831–1844.

[132] W. Zhang, E. R. Brown, A. Mingardi, R. P. Mirin, N. Jahed, and D. Saeedkia, *Thz superradiance from a gaas: Eras quantum dot array at room temperature*, *Applied Sciences (Switzerland)* **9** (2019).

[133] J. C. Song and L. S. Levitov, *Energy flows in graphene: Hot carrier dynamics and cooling*, *Journal of Physics Condensed Matter* **27** (4, 2015).

[134] J. C. Song, M. Y. Reizer, and L. S. Levitov, *Disorder-assisted electron-phonon scattering and cooling pathways in graphene*, *Physical Review Letters* **109** (9, 2012).

[135] K. Yoshioka, T. Wakamura, M. Hashisaka, K. Watanabe, T. Taniguchi, and N. Kumada, *Tracking ultrafast photocurrent generation and transport in hbn-encapsulated graphene using on-chip thz spectroscopy*, *2022 47th International Conference on Infrared, Millimeter and Terahertz Waves (IRMMW-THz)* (9, 2022) 1–2.

[136] N. Hunter, A. S. Mayorov, C. D. Wood, C. Russell, L. Li, E. H. Linfield, A. G. Davies, and J. E. Cunningham, *On-chip picosecond pulse detection and generation using graphene photoconductive switches*, *Nano Letters* **15** (3, 2015) 1591–1596.

[137] T. J. Echtermeyer, P. S. Nene, M. Trushin, R. V. Gorbachev, A. L. Eiden, S. Milana, Z. Sun, J. Schliemann, E. Lidorikis, K. S. Novoselov, and A. C. Ferrari, *Photothermoelectric and photoelectric contributions to light detection in metal-graphene-metal photodetectors*, *Nano Letters* **14** (7, 2014) 3733–3742.

[138] F. Y. Gao, Z. Zhang, Z. Sun, L. Ye, Y.-H. Cheng, Z.-J. Liu, J. G. Checkelsky, E. Baldini, and K. A. Nelson, *Snapshots of a light-induced metastable hidden phase driven by the collapse of charge order*, *Sci. Adv* **8** (2022) 9076.

[139] F. Y. Gao, Z. Zhang, Z.-J. Liu, and K. A. Nelson, *High-speed two-dimensional terahertz spectroscopy with echelon-based shot-to-shot balanced detection*, *Optics Letters* **47** (7, 2022) 3479.

[140] T. Shin, J. W. Wolfson, S. W. Teitelbaum, M. Kandyla, and K. A. Nelson, *Dual echelon femtosecond single-shot spectroscopy*, *Review of Scientific Instruments* **85** (2014).

[141] L. Dhar, J. T. Fourkas, and K. A. Nelson, *Pulse-length-limited ultrafast pump-probe spectroscopy in a single laser shot*, *Optics Letters* **19** (1994).

[142] J. T. Fourkas, L. Dhar, K. A. Nelson, and R. Trebino, *Spatially encoded, single-shot ultrafast spectroscopies*, *Journal of the Optical Society of America B* **12** (1995).

[143] B. Z. Nosho, W. H. Weinberg, W. Barvosa-Carter, B. R. Bennett, B. V. Shanabrook, and L. J. Whitman, *Effects of surface reconstruction on iii-v semiconductor interface formation: The role of iii/v composition*, *Applied Physics Letters* **74** (3, 1999) 1704–1706.

[144] C. Levis, J. T. Johnson, and F. L. Teixeira, *Radiowave Propagation: Physics and Applications*. John Wiley & Sons, 1 ed., 2010.

[145] R. V. Aguilar, A. V. Stier, W. Liu, L. S. Bilbro, D. K. George, N. Bansal, L. Wu, J. Cerne, A. G. Markelz, S. Oh, and N. P. Armitage, *Terahertz response and colossal kerr rotation from the surface states of the topological insulator bi2se3*, *Physical Review Letters* **108** (2, 2012) 087403.

[146] L. Wu, M. Salehi, N. Koirala, J. Moon, S. Oh, and N. P. Armitage, *Quantized faraday and kerr rotation and axion electrodynamics of a 3d topological insulator*, *Science* **354** (12, 2016) 1124–1127.

[147] C. Kastl, C. Karnetzky, H. Karl, and A. W. Holleitner, *Ultrafast helicity control of surface currents in topological insulators with near-unity fidelity*, *Nature Communications* **6** (3, 2015) 1–6.

[148] L. S. Bilbro, R. V. Aguilar, G. Logvenov, O. Pelleg, I. Božović, and N. P. Armitage, *Temporal correlations of superconductivity above the transition temperature in la2-xsrxcuo4 probed by terahertz spectroscopy*, *Nature Physics* **7** (2, 2011) 298–302.

[149] W. Hu, S. Kaiser, D. Nicoletti, C. R. Hunt, I. Gierz, M. C. Hoffmann, M. L. Tacon, T. Loew, B. Keimer, and A. Cavalleri, *Optically enhanced coherent transport in yba2cu3o6.5 by ultrafast redistribution of interlayer coupling*, *Nature Materials* **13** (5, 2014) 705–711.

[150] D. M. Kennes, M. Claassen, L. Xian, A. Georges, A. J. Millis, J. Hone, C. R. Dean, D. N. Basov, A. N. Pasupathy, and A. Rubio, *Moiré heterostructures as a condensed-matter quantum simulator*, *Nature Physics* **17** (2, 2021) 155–163.

[151] L. Balents, C. R. Dean, D. K. Efetov, and A. F. Young, *Superconductivity and strong correlations in moiré flat bands*, *Nature Physics* **16** (5, 2020) 725–733.

[152] E. Y. Andrei, D. K. Efetov, P. Jarillo-Herrero, A. H. MacDonald, K. F. Mak, T. Senthil, E. Tutuc, A. Yazdani, and A. F. Young, *The marvels of moiré materials*, *Nature Reviews Materials* **6** (3, 2021) 201–206.

[153] S. Schäffer, Y. Loth, M. Nagel, S. Sawallich, A. Wigger, and P. H. Bolivar, *A terahertz s-snom tip with integrated photoconductive emitter switch evaluated on an afm-system platform*, *International Conference on Infrared, Millimeter, and Terahertz Waves, IRMMW-THz* **2021-August** (2021).

[154] M. Eisele, T. L. Cocker, M. A. Huber, M. Plankl, L. Viti, D. Ercolani, L. Sorba, M. S. Vitiello, and R. Huber, *Ultrafast multi-terahertz nano-spectroscopy with sub-cycle temporal resolution*, *Nature Photonics 2014 8:11* **8** (10, 2014) 841–845.

[155] T. L. Cocker, V. Jelic, M. Gupta, S. J. Molesky, J. A. Burgess, G. D. L. Reyes, L. V. Titova, Y. Y. Tsui, M. R. Freeman, and F. A. Hegmann, *An ultrafast terahertz scanning tunnelling microscope*, *Nature Photonics 2013 7:8* **7** (7, 2013) 620–625.

[156] V. Jelic, K. Iwaszczuk, P. H. Nguyen, C. Rathje, G. J. Hornig, H. M. Sharum, J. R. Hoffman, M. R. Freeman, and F. A. Hegmann, *Ultrafast terahertz control of*

extreme tunnel currents through single atoms on a silicon surface, *Nature Physics 2017 13:6* **13** (2, 2017) 591–598.

[157] D. M. Mittleman, P. Dean, A. Valavanis, J. Keeley, K. Bertling, Y. L. Lim, R. Alhathlool, A. D. Burnett, L. H. Li, S. P. Khanna, D. Indjin, T. Taimre, A. D. Rakic, E. H. Linfield, and A. G. Davies, *Twenty years of terahertz imaging [invited]*, *Optics Express, Vol. 26, Issue 8, pp. 9417-9431* **26** (4, 2018) 9417–9431.

[158] Z. Fei, G. O. Andreev, W. Bao, L. M. Zhang, A. S. McLeod, C. Wang, M. K. Stewart, Z. Zhao, G. Dominguez, M. Thiemens, M. M. Fogler, M. J. Tauber, A. H. Castro-Neto, C. N. Lau, F. Keilmann, and D. N. Basov, *Infrared nanoscopy of dirac plasmons at the graphene-sio2 interface*, *Nano Letters* **11** (11, 2011) 4701–4705.

[159] D. H. Auston and M. C. Nuss, *Electrooptic generation and detection of femtosecond electrical transients*, *IEEE Journal of Quantum Electronics* **24** (1988) 184–197.

[160] C. D. Wood, D. Mistry, L. H. Li, J. E. Cunningham, E. H. Linfield, and A. G. Davies, *On-chip terahertz spectroscopic techniques for measuring mesoscopic quantum systems*, *Review of Scientific Instruments* **84** (8, 2013) 085101.

[161] J. Cunningham, M. B. Byrne, C. D. Wood, and L. Dazhang, *On-chip terahertz systems for spectroscopy and imaging*, *Electronics Letters* **46** (12, 2010) 34–37.

[162] L. Dazhang, J. Cunningham, and M. B. Byrne, *On-chip terahertz goubau-line waveguides with integrated photoconductive emitters and mode-discriminating detectors*, *Appl. Phys. Lett* **95** (2009) 92903.

[163] D. M. Heligman, *Terahertz Time Domain Spectroscopy Techniques for Antiferromagnets and Metamaterials*. PhD thesis, The Ohio State University, Columbus, OH, 2021.

[164] C. Karnetzky, P. Zimmermann, C. Trummer, C. D. Sierra, M. Wörle, R. Kienberger, and A. Holleitner, *Towards femtosecond on-chip electronics based on plasmonic hot electron nano-emitters*, *Nature Communications* **9** (6, 2018) 1–7.

[165] Z. Shi, H. Zhang, K. Khan, R. Cao, Y. Zhang, C. Ma, A. K. Tareen, Y. Jiang, M. Jin, and H. Zhang, *Two-dimensional materials toward terahertz optoelectronic device applications*, *Journal of Photochemistry and Photobiology C: Photochemistry Reviews* **51** (6, 2022) 100473.

[166] C. Wood, J. Cunningham, P. C. Upadhya, E. H. Linfield, I. C. Hunter, A. G. Davies, and M. Missous, *On-chip photoconductive excitation and detection of pulsed terahertz radiation at cryogenic temperatures*, *Applied Physics Letters* **88** (4, 2006) 142103.

[167] Z. Zhong, N. M. Gabor, J. E. Sharping, A. L. Gaeta, and P. L. McEuen, *Terahertz time-domain measurement of ballistic electron resonance in a single-walled carbon nanotube*, *Nature Nanotechnology* **3** (3, 2008) 201–205.

[168] C. D. Wood, D. Mistry, L. H. Li, J. E. Cunningham, E. H. Linfield, and A. G. Davies, *On-chip thz generation and detection at milli-kelvin temperatures for the study of ultrafast phenomena in confined semiconductor systems*, *International Conference on Infrared, Millimeter, and Terahertz Waves, IRMMW-THz* (2012) 1–3.

[169] L. Prechtel, L. Song, D. Schuh, P. Ajayan, W. Wegscheider, and A. W. Holleitner, *Time-resolved ultrafast photocurrents and terahertz generation in freely suspended graphene*, *Nature Communications* **3** (1, 2012) 1–7.

[170] J. Lee, S. Sawallich, M. C. Lemme, and M. Nagel, *On-chip terahertz sensor based on low-loss coplanar strip lines for the analysis of microscale two-dimensional materials*, in *2022 47th International Conference on Infrared, Millimeter and Terahertz Waves (IRMMW-THz)*, pp. 1–2, 2022.

[171] S. Yanagi, M. Onuma, J. Kitagawa, and Y. Kadoya, *Propagation of terahertz pulses on coplanar strip-lines on low permittivity substrates and a spectroscopy application*, *Applied Physics Express* **1** (1, 2008) 012009.

[172] P. D. Cunningham, N. N. Valdes, F. A. Vallejo, L. M. Hayden, B. Polishak, X. H. Zhou, J. Luo, A. K. Jen, J. C. Williams, and R. J. Twieg, *Broadband terahertz characterization of the refractive index and absorption of some important polymeric and organic electro-optic materials*, *Journal of Applied Physics* **109** (2, 2011) 043505–1 – 043505–5.

[173] T. Kaji, Y. Tominari, T. Yamada, S. Saito, I. Morohashi, and A. Otomo, *Terahertz-wave generation devices using electro-optic polymer slab waveguides and cyclo-olefin polymer clads*, *Optics Express* **26** (2018) 30466–30475.

[174] S. A. Sychugin, E. A. Mashkovich, A. V. Maslov, and M. I. Bakunov, *Terahertz cherenkov radiation from a tightly focused ultrashort laser pulse in an electro-optic medium*, *J. Opt. Soc. Am. B* **36** (Apr, 2019) 1101–1107.

[175] J. Neu and C. A. Schmuttenmaer, *Tutorial: An introduction to terahertz time domain spectroscopy (thz-tds)*, *Journal of Applied Physics* **124** (12, 2018) 231101.

[176] A. M. Potts, M. T. Warren, R. V. Aguilar, and T. T. Mai, *Corrective re-gridding techniques for non-uniform sampling in time-domain terahertz spectroscopy*, *J. Opt. Soc. Am. B* **36** (4, 2019) 1037–1043.

[177] B. Cheng, L. Wu, N. J. Laurita, H. Singh, M. Chand, P. Raychaudhuri, and N. P. Armitage, *Anomalous gap-edge dissipation in disordered superconductors on the brink of localization*, *Physical Review B* **93** (5, 2016) 180511.

[178] T. Hong, K. Choi, K. I. Sim, T. Ha, B. C. Park, H. Yamamori, and J. H. Kim, *Terahertz electrodynamics and superconducting energy gap of nbtin*, *Journal of Applied Physics* **114** (12, 2013) 243905–1 – 243905–4.

[179] K. I. Sim, Y. C. Jo, T. Ha, J. H. Kim, J. H. Kim, and H. Yamamori, *Terahertz electrodynamics and superconducting energy gap of nbn*, *Journal of the Korean Physical Society* **71** (10, 2017) 571–574.

[180] M. C. Nuss, K. W. Goossen, J. P. Gordon, P. M. Mankiewich, M. L. O'Malley, and M. Bhushan, *Terahertz time-domain measurement of the conductivity and superconducting band gap in niobium*, *Journal of Applied Physics* **70** (1991) 2238–2241.

[181] M. C. Nuss and J. Orenstein, *Terahertz time-domain spectroscopy*, *Millimeter and Submillimeter Wave Spectroscopy of Solids* **74** (1998) 7–50.

[182] S. Lauck, S. Sawallich, R. B. Kohlhaas, A. Michalski, L. Liebermeister, M. Nagel, M. Schell, and B. Globisch, *Photonic thz near-field imaging: Characterizing high-frequency components from 100 ghz to 4 thz*, *International Conference on Infrared, Millimeter, and Terahertz Waves, IRMMW-THz* (2022).

[183] F. Mahmood, D. Chaudhuri, S. Gopalakrishnan, R. Nandkishore, and N. P. Armitage, *Observation of a marginal fermi glass*, *Nature Physics* **17** (5, 2021) 627–631.

[184] H. Polshyn, M. Yankowitz, S. Chen, Y. Zhang, K. Watanabe, T. Taniguchi, C. R. Dean, and A. F. Young, *Large linear-in-temperature resistivity in twisted bilayer graphene*, *Nature Physics* **15** (10, 2019) 1011–1016.

[185] Y. Cao, D. Chowdhury, D. Rodan-Legrain, O. Rubies-Bigorda, K. Watanabe, T. Taniguchi, T. Senthil, and P. Jarillo-Herrero, *Strange metal in magic-angle graphene with near planckian dissipation*, *Physical Review Letters* **124** (2, 2020).

[186] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, *Unconventional superconductivity in magic-angle graphene superlattices*, *Nature* **556** (4, 2018) 43–50.

[187] L. Prochaska, X. Li, D. C. MacFarland, A. M. Andrews, M. Bonta, E. F. Bianco, S. Yazdi, W. Schrenk, H. Detz, A. Limbeck, Q. Si, E. Ringe, G. Strasser, J. Kono, and S. Paschen, *Singular charge fluctuations at a magnetic quantum critical point*, *Science367* (1, 2020) 2023.

[188] B. Huang, G. Clark, E. Navarro-Moratalla, D. R. Klein, R. Cheng, K. L. Seyler, D. Zhong, E. Schmidgall, M. A. McGuire, D. H. Cobden, W. Yao, D. Xiao, P. Jarillo-Herrero, and X. Xu, *Layer-dependent ferromagnetism in a van der waals crystal down to the monolayer limit*, *Nature* **546** (6, 2017) 270–273.

[189] B. Huang, G. Clark, D. R. Klein, D. MacNeill, E. Navarro-Moratalla, K. L. Seyler, N. Wilson, M. A. McGuire, D. H. Cobden, D. Xiao, W. Yao, P. Jarillo-Herrero, and X. Xu, *Electrical control of 2d magnetism in bilayer cri3*, *Nature Nanotechnology* **13** (7, 2018) 544–548.

[190] D. R. Klein, D. Macneill, J. L. Lado, D. Soriano, E. Navarro-Moratalla, K. Watanabe, T. Taniguchi, S. Manni, P. Canfield, J. Fernández-Rossier, and P. Jarillo-Herrero, *Probing magnetism in 2d van der waals crystalline insulators via electron tunneling*, *Science* **360** (2018).

[191] K. L. Seyler, P. Rivera, H. Yu, N. P. Wilson, E. L. Ray, D. G. Mandrus, J. Yan, W. Yao, and X. Xu, *Signatures of moiré-trapped valley excitons in mose2/wse2 heterobilayers*, *Nature* **567** (3, 2019) 66–70.

[192] P. Rivera, J. R. Schaibley, A. M. Jones, J. S. Ross, S. Wu, G. Aivazian, P. Klement, K. Seyler, G. Clark, N. J. Ghimire, J. Yan, D. G. Mandrus, W. Yao, and X. Xu, *Observation of long-lived interlayer excitons in monolayer mose 2-wse 2 heterostructures*, *Nature Communications* **6** (2, 2015).

[193] C. Jin, E. C. Regan, A. Yan, M. I. B. Utama, D. Wang, S. Zhao, Y. Qin, S. Yang, Z. Zheng, S. Shi, K. Watanabe, T. Taniguchi, S. Tongay, A. Zettl, and F. Wang, *Observation of moiré excitons in wse2/ws2 heterostructure superlattices*, *Nature* **567** (3, 2019) 76–80.

[194] Y. Yu, L. Ma, P. Cai, R. Zhong, C. Ye, J. Shen, G. D. Gu, X. H. Chen, and Y. Zhang, *High-temperature superconductivity in monolayer bi2sr2cacu2o8* $+ \delta$, *Nature* **575** (11, 2019) 156–163.

[195] O. Can, T. Tummuru, R. P. Day, I. Elfimov, A. Damascelli, and M. Franz, *High-temperature topological superconductivity in twisted double-layer copper oxides*, *Nature Physics* **17** (4, 2021) 519–524.

[196] Y. M. Wu, Z. Wu, and H. Yao, *Pair-density-wave and chiral superconductivity in twisted bilayer transition metal dichalcogenides*, *Physical Review Letters* **130** (3, 2023).

[197] B. Cheng, T. Schumann, S. Stemmer, and N. P. Armitage, *Probing charge pumping and relaxation of the chiral anomaly in a dirac semimetal*, *Science Advances* **7** (4, 2021) 914–930.

[198] A. Kogar, A. Zong, P. E. Dolgirev, X. Shen, J. Straquadine, Y. Q. Bie, X. Wang, T. Rohwer, I. C. Tung, Y. Yang, R. Li, J. Yang, S. Weathersby, S. Park, M. E. Kozina, E. J. Sie, H. Wen, P. Jarillo-Herrero, I. R. Fisher, X. Wang, and N. Gedik, *Light-induced charge density wave in late3*, *Nature Physics* **16** (2, 2020) 159–163.

[199] O. Katz, G. Refael, and N. H. Lindner, *Optically induced flat bands in twisted bilayer graphene*, *Physical Review B* **102** (10, 2020).

[200] W. Choi, K. H. Lee, and Y. B. Kim, *Theory of two-dimensional nonlinear spectroscopy for the kitaev spin liquid*, *Physical Review Letters* **124** (3, 2020).

[201] T. A. Green, *Gold etching for microfabrication*, *Gold Bulletin* **47** (5, 2014) 205–216.

[202] J. M. Jin, *Theory and Computation of Electromagnetic Field*. John Wiley and Sons, 9, 2010.

[203] R. Marjoribanks, *A primer on ultrafast fiber lasers*, tech. rep., The University of Toronto, 2003.