Ultrafast spectroscopy and control of correlated quantum materials

By

Bryan T. Fichera

B.S., University of Pennsylvania (2017)

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

Doctor of Philosophy

at the

MASSACHUSETTS INSTUTITE OF TECHNOLOGY

May 2024

Authored by: Bryan T. Fichera
               Department of Physics
               May?, 2024

Certified by: Nuh Gedik
               Donner Professor of Physics, Thesis supervisor

Accepted by: ?
               ? Professor of ?
               ? Chair of ?

Abstract

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Acknowledgements

Preface

The physics of solids is, to me, one of the most important and fundamental fields of modern science. This might seem, to some, a bit of a hot take. After all, by studying condensed matter physics, one learns next to nothing about, say, the formation of the stars and planets, or the origin of the universe. Nor does one learn about life, death, consciousness, disease, ethics, God, or any other question that perhaps puzzled humanity prior to about five hundred years ago. Certainly no one would argue that condensed matter physics is quite *useless*, given that nearly every device we interact with in modern life required some condensed matter physicist somewhere along the way to make one brilliant discovery or another—yet when the human mind starts to wander, and our thoughts turn to the metaphysical, we tend to look up, not down.

In my work I have taken a quite different view. Condensed matter physics, to me, is ultimately the study of how *truly boring* objects, when brought together in large quantites, *become* interesting, seemingly in spite of themselves. When electrons are put together in a lattice and allowed to interact slightly with the massive nuclei, at low enough temperatures they pair, the low-energy excitations become gapped, and current can flow for infinite times and with absolutely zero energy loss. Those same electrons, with some other set of interactions, may instead ionize (the opposite of pairing!) to create an electrically insulating state, whose low-energy excitation spectrum is nevertheless gapless and consisting of charge-neutral spin-$1/2$ particles. In all such cases, these systems exist in otherwise ordinary-looking rocks, fit in the palm of a hand[1], and are more or less indistinguishable from something you might find sticking into the bottom of your shoe.

While such systems may not tell us a lot[2] about the early universe, considering these and related problems lets us ask deep, fundamental questions about the world we live in—like, why is this thing a metal, but this thing is an insulator? What do those terms even mean?—that I don't think we would try to ask otherwise. To me, focusing our attention on

---

[1]Hopefully, gloved.

[2]This discussion is obviously intentionally reductive. In truth there is still quite a bit one can learn about, e.g. the early universe by studying condensed matter physics, see the Kibble and Pickett [15].

these problems, despite their obviously terrestrial nature, is not a waste of time; rather, I think they remind us that even the most mundane aspects of the human experience involve a level of complexity far beyond what we are capable of understanding absent the pursuit of science.

Throughout the seven years of my Ph.D., I hope to have made a few contributions to this pursuit. As the title of this work implies, I have mainly focused on the application of ultrafast techniques to the study of correlated quantum materials, which I loosely define as those materials in which the interaction between particles is large enough so as to compete with the kinetic energy of those particles. It is in these materials that I think lies the true frontier of condensed matter physics; here, much of our basic intuition about non- or weakly-interacting theory fails, and more complicated notions of phase competition, phase separation, disorder, pairing, coherence, etc. are needed to property describe the relevant physics.

In my own view, and in the view of many scientists in this field[2], the main question for strongly correlated physics amounts to: "Given a correlated system with some defined combination of different interaction strengths, is there a general theory which allows us to predict the phase diagram of this system *a priori*?" Related of course are questions about the origins of high-T_c superconductivity, strange metallicity, quantum spin liquids, and other exotic phases that we find emerging from strongly interacting systems. Since such a theory does not currently exist, at least with the level of predictive power that I think most would find satisfactory, new advances in this field typically come directly from experiment. Ultrafast optics plays a special role in this regard, for reasons that I will explain in chapter 1.

Progress thus happens in this field somewhat unsystematically, with small pieces of the puzzle added at random, but not infrequenct, intervals. Usually it is either new techniques or new materials that are the driving force here. To this end, I have tried to pursue both directions in my Ph.D. Appearing also in chapter 1 is thus a description of the materials I studied the most during my thesis, two of them, $CuBr_2$ and $CaMn_2Bi_2$ I consider criminally understudied. On the technique side, almost all of the work presented in this thesis was done using time-resolved second harmonic generation (tr-SHG), a relatively new, nonlinear optical technique which, at the most basic level, probes the point group assumed by the charge distribution function $\rho(\boldsymbol{x})$ at any given point in time. Second harmonic

generation (SHG) and tr-SHG are tricky techniques, with many pitfalls both practically and theoretically; chapters 2 and 3 are thus devoted to what I hope is a useful, if not fully compehensive, description of the technique. **??** is devoted to work that we did developing a new way to control the polarization of the light in a tr-SHG experiment using stepper motors. My hope is that these chapters are useful not only for the new student trying to build their own setup or analyze their own SHG data, but also for people for whom SHG is not a focus but nevertheless want to learn about it in slightly more detail than one would get from a typical paper or review article.

What follows, then, is a description of the three main research works I contributed during my Ph.D.. The first, which I describe in **??**, involves work that I did during my second and third years on $1T$-TaS_2, a very interesting charge density wave (CDW) material that, among other things, undergoes a mirror symmetry breaking CDW transition at 350 K that shows up in the SHG as a sudden distortion of the flower pattern at that temperature. Since this transition breaks mirror symmetry, two energetically degenerate domains should be present, corresponding to two opposite planar chiralities; in this work, we showed that SHG could differentiate between these two domains (i.e. the flower pattern in either domain looks different).

The second and third works, which I describe in **????**, in contrast to the $1T$-TaS_2 work, both involve taking the system out of equilibrium to study the dynamics. In $CaMn_2Bi_2$ (**??**), we discovered that photoexcitation causes the antiferromagnetic (AFM) order in that compound to reorient (relative to equilibrium) to a metastable state which is impossible to reach from the equilibrium state thermodynamically. Light is thus used to *control* the magnetic order in this material.

In $CuBr_2$ (**??**), light is not used to control the order parameter like in $CaMn_2Bi_2$, but it does excite coherent oscillations of the collective modes of the multiferroic order (electromagnons), whose frequency, amplitude, damping, etc. may be probed in tr-SHG as a function of temperature—a methodology referred to as ultrafast *spectroscopy*. In doing so, we found that one of these collective modes is actually quite special, as it is in fact the analogue of the Higgs mode of particle physics in the context of a multiferroic material.

I conclude with various remarks in **??**, as well as an appendix, in which I enumerate briefly all of the null-result experiments I performed during

my Ph.D., in the hopes that future scientists don't have to waste time on what we already know are fruitless pursuits. If you have any questions about this or any other section of this thesis, please do not hesitate to reach out via email.

Contents

List of Figures

List of Tables

Chapter One

Ultrafast optics in correlated electron systems

Interactions between electrons in solids can be treated broadly in two different ways, depending on the strength of the interaction. Let us start with the case where the interactions are weak compared to their kinetic energy. In this case, interactions may be treated as a perturbation

$$V = \sum_{\boldsymbol{pp'q},\sigma\sigma'} V(\boldsymbol{q}) c^{\dagger}_{\boldsymbol{p+q},\sigma} c^{\dagger}_{\boldsymbol{p'-q},\sigma'} c_{\boldsymbol{p'},\sigma'} c_{\boldsymbol{p},\sigma} \tag{1.1}$$

on top of the usual hamiltonian

$$H_0 = \epsilon_{\boldsymbol{p}} c^{\dagger}_{\boldsymbol{p},\sigma} c_{\boldsymbol{p},\sigma} \tag{1.2}$$

for the noninteracting electrons. The key insight, due conceptually to **?**] and later formalized by **?**], is that, as long as the interactions are sufficiently weak, this problem can be adiabatically connected to a similar problem with non-interacting quasiparticles. Since these quasiparticles are fermions, they obey the Pauli exclusion principle, and the phase space available for scattering low-energy excitations is thus goes to zero at small energies. By Fermi's golden rule, the lifetime of such particles thus approaches infinity as we get closer and closer to the Fermi surface; i.e., there are still well-defined quasiparticles there, even though we started with an interacting Hamiltonian. Such a system is hence referred as a "Fermi liquid," to signify the fact that we have made only a slight departure from the nominal model of a Fermi gas where the particles are treated as noninteracting.

Fermi liquid theory is almost unreasonably successful in describing the ground state of a large number of correlated electron systems. Truly *most* metals can be very adequately described as a simple Fermi gas with renormalized effective mass, specific heat, etc.; while these modifications may be large (e.g. $m^*/m \approx 10^3$ in $CeAl_3$), they otherwise look like normal metals. Nevertheless, we can discuss circumstances in which the theory fails. Clearly various electronic instabilities may gap out the Fermi surface, resulting in an insulator; this happens for arbitrarily weak interactions in quasi-1D systems due to the Peierls mechanism, and in higher dimensions due to, e.g., Fermi surface nesting. Whether the interaction is attractive or repulsive determines whether the instability occurs in the charge or the spin channel. In the other limit, where the kinetic energy hamiltonian 1.2 is treated as a perturbation to the interaction term 1.1, the Fermi liquid state gives way to an (Mott) insulating state where the electrons become localized so as to minimize the coulomb repulsion U.

Studying such departures from Fermi liquid theory has become one of the most important fields in condensed matter physics. It is usually referred to as "strongly correlated electron physics", but many of the instabilities mentioned above are present even in the limit of weak interactions. Much of this interest was spurred by the discovery in 1986 by **?**] of high-T_c superconductivity in $Ba_x La_{5-x} Cu_5 O_{5(3-y)}$, although the field has grown to include many other phenomena which occur in the strongly interacting limit that may or may not be related to superconductivity. Despite nearly four decades of research, however, there still exists no universal theory for strongly correlated electron systems, in the sense that, if someone hands you a random strongly correlated material, it is impossible from the outset to *predict* what will be the ground state, what are its low-energy excitations, etc. In fact, it is not even clear such a theory ought to exist[2].

The *field* of strongly correlated physics is thus, in some sense, still very much in its infancy, although that's not to say there haven't been huge advances, both theoretically and experimentally, especially since the discovery by **?**]. A lot of the effort experimentally has been focused on cataloguing the huge number of different exotic ordered phases—charge density wave, spin density wave, superconducting, strange metals, spin liquids, pseudogap, etc.—that are realized in these systems. Usually this is done by mapping out a phase diagram for a given material, which indicates which phases are present at various values of external parameters like magnetic field, pressure, strain, etc. Ultrafast optics has played an

important role in this respect, as it allows one (heuristically) to add a *nonequilibrium* axis to such phase diagrams. Such experiments, for example, have not only helped illustrate the extent to which different phases compete with one another in the cuprate phase diagram, but also the extent to which the action of the light pulse in strongly correlated materials can be used to tune the properties of those materials for practical purposes. This is paradigm is referred to as "ultrafast control," and I will discuss it in more detail in **??**.

A parallel effort in the field of ultrafast optics is to use the pump not to control the state of the material, but rather to excite coherent oscillations of the low-energy collective modes and study these oscillations in the time domain in a pump-probe scheme. This approach is advantageous for two reasons. For one, the frequencies accessible with this technique are bounded from below only by the length of one's delay stage, in contrast to conventional spectrometer-based methods which involve finite-frequency filters, gratings, etc. Second of all, the ability to select both the excitation mechanism (the pump) and the measurement apparatus (the probe) allows one to design experiments that target specific degrees of freedom of interest. Thus, for example, in multiferroics, one can use SHG to selecitively probe the collective modes which modulate the macroscopic polarization. This direction is referred to as "ultrafast spectroscopy," which I will explain in detail in section 1.1.

This chapter may thus be regarded as an introduction to strongly correlated electron systems, with a special emphasis on ultrafast experiments (of the two types explained above). A complete review of this material is beyond the scope of this thesis; instead, I will focus on a few seminal works that I think tell this story most pedagogically. We will also need a bit of machinery to understand **??????**, which focus on CDW, AFM, and multiferroic materials, respectively; hence I will primarily focus on these types of ground states, although of course this obviously misses e.g. strange metal phases, unconventional superconductivity, quantum spin liquids, etc. When appropriate, I point to pedagogical references that may be more useful than this thesis for these and other concepts.

1.1   Spectroscopy

1.1.1   Collective modes

The low-energy excitations of any many-body system are typically *collective*, in the sense that they involve motion of all of the particles in the system rather than just one. Let us consider, for example, the classical model consisting of two identical coupled harmonic oscillators

$$H = \sum_{i=1}^{2} \left(\frac{p_i^2}{2m} + \frac{m\omega_0^2}{2}x_i^2 \right) + gx_1x_2, \tag{1.3}$$

with mass m, natural frequency ω_0, and coupling constant g. When $g = 0$, the normal modes of this system are simply the independent oscillation modes of the two oscillators. However, for any nonzero g, the normal modes involve either symmetric or antisymmetric linear combinations of the two oscillator coordinates; that is, they involve collective motion of the two oscillators. The extension to an ensemble of harmonic oscillators is straightforward; there, too, the normal modes of the Hamiltonian involve collective motion of all of the coordinates at once.

Of course the solids that we are interested in are more complicated than a simple ensemble of harmonic oscillators. Usually, though, more complicated systems can be broken down into a small number of *subsystems* which may, as a first approximation, be treated separately. Thus, for example, it makes sense to refer to the phonon subsytem independently from the electronic subsystem, with different and independent collective mode spectra in the limit where the inter-subsystem coupling goes to zero.

When this coupling is finite, however, many interesting phenomena may occur. For example, spin-orbit coupling coupling—which implies a coupling between spin and orbital degrees of freedom of the electron—can, in the right circumstances, cause the statically ordered state of the spin subsystem to also induce a ferroelectric distortion of the electron orbitals. The normal modes of the system in the presence of this coupling are no longer pure magnon or orbital modes, but rather collective modes of the spin and orbital degrees of freedom together. Let us look at this phenomenon from a different perspective. Suppose we didn't know the ground state of the system, but we did know that the relevant low-energy

excitations involved both spin and orbital degrees of freedom (for example, we might see a response in both the time-resolved kerr rotation as well as the tr-SHG). This is good evidence that the ground state of the system involves this coupling to some extent. Thus, we have learned something quite important about our system without doing anything but look at the low-energy collective excitations.

1.1.2 Coherent oscillations

Clearly understanding the low-energy excitations corresponding to a given ground state is a useful way to understand its properties. So far, however, we have made no mention of how to *probe* these excitations in pump-probe spectroscopy. Let us consider the following Hamiltonian consisting of electrons $c_{k\sigma}$ (with dispersion $\epsilon_{k\sigma}$) and bosons b_q (with dispersion ω_q) interacting via some potential V_{kq}^σ

$$H = \sum_{k\sigma} \epsilon_{k\sigma} c_{k\sigma}^\dagger c_{k\sigma} + \sum_q \hbar\omega_q b_q^\dagger b_q + \sum_{\sigma kq} V_{kq}^\sigma \left(b_q + b_{-q}^\dagger \right) c_{k\sigma}^\dagger c_{k+q\sigma}. \quad (1.4)$$

The average lattice displacement is given by

$$\langle u(\boldsymbol{r}) \rangle \propto \sum_q \left(\langle b_q \rangle e^{i\boldsymbol{q}\cdot\boldsymbol{r}} + \langle b_q^\dagger \rangle e^{-i\boldsymbol{q}\cdot\boldsymbol{r}} \right). \quad (1.5)$$

Clearly, in order for us to have a macroscopic lattice displacement, we need to have a finite value for $\langle b_q \rangle$ and $\langle b_q^\dagger \rangle$, which is impossible if there are a definite number of phonons in the mode $\boldsymbol{q}$ (since $\langle n | b_q | m \rangle = 0$ for $n = m$). In contrast, if the wavefunction of the system consists of a coherent superposition of different phonon numbers, $\langle u(\boldsymbol{r}) \rangle$ may acquire a finite value. One thematic example of such a wavefunction is the so-called "coherent state" of the quantum harmonic oscillator

$$|\alpha_q\rangle = \sum_n \frac{\alpha^n e^{-z^2/2}}{n!} (b_q^\dagger)^n |0\rangle, \quad (1.6)$$

although the real wavefunction need not be fully coherent to have a nonzero average lattice displacement.

One can show (see **?**]) that the equation of motion for the operator $D_{\boldsymbol{q}} \equiv \left\langle b_{\boldsymbol{q}} \right\rangle + \left\langle b^{\dagger}_{-\boldsymbol{q}} \right\rangle$ due to Eq. 1.4 is

$$\frac{\partial^2}{\partial t^2} D_{\boldsymbol{q}} + \omega_{\boldsymbol{q}}^2 D_{\boldsymbol{q}} = -2\omega_{\boldsymbol{q}} \sum_{\boldsymbol{k}\sigma} V^{\sigma}_{\boldsymbol{k}\boldsymbol{q}} \left\langle c^{\dagger}_{\boldsymbol{k}\sigma} c_{\boldsymbol{k}+\boldsymbol{q}\sigma} \right\rangle , \qquad (1.7)$$

i.e., $D_{\boldsymbol{q}}$ obeys a *wave equation* with an inhomogenous part related (in this case) to the electronic subsystem.[1] Thus, if we manage to initialize a wavefunction with a finite $D_{\boldsymbol{q}}$, the frequency with which $D_{\boldsymbol{q}}$ oscillates in time is the frequency $\omega_{\boldsymbol{q}}$ of the boson $b_{\boldsymbol{q}}$.

The central idea in ultrafast spectroscopy is therefore to excite coherent modes like Eq. 1.6, and then measure the frequency, damping, etc. of these modes by measuring $D_{\boldsymbol{q}}$. This is in contrast to equilibrium spectroscopies, which measure, for example, the transfer of energy from the light field to states with a definite number of bosons (i.e., $|n\rangle \to |n+1\rangle$). In theory, of course, the information obtained is the same—the frequency and damping coefficient of the collective modes in question may readily be optained in the equilibrium spectroscopies as well as in the pump-probe scheme. However, as I argued at the start of this chapter, the pump-probe techniques offer a number of advantages, most notably (i) the ability to design the pump and the probe to specify exactly which excitations we would like to measure, and (ii) the ability to measure much lower frequencies than in conventional spectroscopy due to the energy being measured in the time domain, rather than the frequency domain.

1.1.3 Excitation mechanisms

The next question is how we typically excite these coherent collective oscillations in real materials. The truth is there are many such mechanisms; however, we can start by placing them into two generic categories. *Impulsive* mechanisms involve using the light pulse to apply an effective force to the relevant degrees of freedom in the material, which lasts for the duration of the light pulse; i.e., it is a delta function in time. *Displacive* mechanisms are, in contrast, typically a step function; i.e., the equilibrium position of the oscillator is different before and after the light pulse. One important experimental difference is that impulsive excitation results

---

[1]There will also be a damping term, which may be added phenomenologically but is otherwise not considered in this treatment.

in a coordinate $D_{\boldsymbol{q}}(t) \propto \sin(\omega t)$, whereas displacive excitation results in $D_{\boldsymbol{q}}(t) \propto \cos(\omega t)$; this can be seen simply by solving

$$\frac{\partial^2}{\partial t^2} D_{\boldsymbol{q}} + \omega_{\boldsymbol{q}}^2 D_{\boldsymbol{q}} = f(t) \tag{1.8}$$

for $f(t) \propto \delta(t)$ or $f(t) \propto \theta(t)$, respectively.

1.1.3.1 In absorption

Let us consider the Hamiltonian in Eq. 1.4. We saw that this Hamiltonian resulted in an equation of motion given by Eq. 1.7, which is a wave equation for $D_{\boldsymbol{q}}$ with a force term

$$f(t) = -2\omega_{\boldsymbol{q}} \sum_{\boldsymbol{k}\sigma} V_{\boldsymbol{k}\boldsymbol{q}}^{\sigma} \left\langle c_{\boldsymbol{k}\sigma}^{\dagger} c_{\boldsymbol{k}+\boldsymbol{q}\sigma} \right\rangle, \tag{1.9}$$

where the time-dependence of the right hand side is complicated but may be phonemnologically modeled. For example, let us suppose that we are in a semiconductor and that photon energy of our pump pulse is greater than the band gap of the material. Thus, the action of the pump is to excite electrons from the valence band into the conduction band. On a very fast timescale $(0.1\text{--}10\,\mathrm{fs})$ these electrons thermalize with themselves via electron-electron scattering, resulting in a quasi-equilibrium carrier distribution in which electrons and holes have settled at the bottom of the conduction band and the top of the valence band, respectively. Since further decay of these excitations is gapped, the relaxation of this state back to equilibrium may be quite long, especially if the gap is indirect; thus, together with the assumption that the lattice dynamics happen on much longer timescales than the electron thermalization time, it is appropriate to model the force term in Eq. 1.9 as a step function in time

$$f(t) \propto \begin{cases} 0 & t < 0 \\ 1 & t \geq 0 \end{cases}. \tag{1.10}$$

Thus we have the generic result that above-gap excitation typically excites coherent oscillations *displacively*. In the case that the bosons of Eq. 1.4 are phonons, this is known as displacive excitation of coherent phonons (DECP), and was studied by many authors, notably Zeiger et al.

[42]. An important insight which follows from Eq. 1.9 is that, in the limit where the quasi-equilibrium electron distribution doesn't break any symmetries of the original hamiltonian, the applied force also does not break any symmetries. Thus only totally-symmetric phonons may be excited via DECP.

1.1.3.2   In transparency

While the DECP-like mechanisms tend to dominate when the photon energy is above the band gap, they are forbidden in transparency. In this case, the dominant excitation mechanism is actually impulsive. Let us consider the case of a collective mode with energy $\hbar\omega_0$ which we wish to excite coherently with an ultrafast laser pulse, below the band gap. If the central frequency of the light pulse is resonant with the collective mode (i.e., we have $\nu_{\mathrm{photon}} = \hbar\omega_0$), then we may drive the coherent oscillation directly; i.e., the mode coordinate Q simply follows the electric field. This is of course only possible if the mode carries a finite dipole moment (i.e. it is odd under parity). Alternatively, we will see that, under the right conditions, we can also excite coherent oscillation of low-energy collective modes even if the photon energy is far away from any direct resonance.

To see this, recall that ultrafast laser pulses always have a nonzero bandwidth, proportional to the inverse of the pulse width. In a nonlinear optical effect known as difference frequency generation (DFG), pairs of these frequencies may *interfere* to produce electric field components at the difference frequency between the two members of the pair. Thus, as long as the collective mode in question has a frequency less than approximately the bandwidth $\delta\nu_{\mathrm{photon}}$ of the incident light pulse, there will exist pairs of frequencies (ω_1, ω_2) in that light pulse such that $|\omega_1 - \omega_2| \approx \omega_0$. This electric field will last for the duration of the light pulse, which may be treated as a delta function in time. This phenomenon is referred to as impulsive stimulated raman scattering (ISRS).

Since the force is a delta function in time, the result indeed is an impulsive excitation where the coordinate $D_{\boldsymbol{q}}(t) \sim \sin(\omega_0 t)$. Unlike the direct excitation mechanism mentioned above, ISRS requires two photons and the force is thus even under parity; i.e., we can only excite even-parity bosons.

It is instructive at this point to consider ISRS in a phenomenological Ginzburg-Landau (GL) model. We start by writing down an effective free

energy[?]

$$F = -\chi_{ij} E_i(t) E_j(t) Q \tag{1.11}$$

where $\boldsymbol{E}(t)$ is the incident electric field, Q the mode coordinate,[2] and χ_{ij} is some tensor of coefficients (see chapter 2). The effective force f due to Eq. 1.11 is

$$f = -\frac{\partial F}{\partial Q} \tag{1.12}$$

$$= \chi_{ij} E_i(t) E_j(t) \tag{1.13}$$

which appears on the right hand side of Eq. 1.8. Clearly, if the $\boldsymbol{E}(t)$ has Fourier components ω_1 and ω_2 such that $|\omega_1 - \omega_2| \approx \omega_0$, then the force will be *resonant* with the oscillator frequency.

Some interesting insights may be made in light of **??**. Let us consider for example the limit where the "mode" Q is a static magnetization M_k. The free energy reads[?]

$$F = -\chi_{ijk} E_i(t) E_j(t) M_k, \tag{1.14}$$

where we have expanded χ_{ij} to linear order in $\boldsymbol{M}$. Then, there is an *effective magnetic field*

$$H_k^{\text{eff}} = \frac{\partial F}{\partial M_k} \tag{1.15}$$

$$= \chi_{ijk} E_i(t) E_j(t) \tag{1.16}$$

which exists for the duration of the pump pulse. This is the inverse Faraday effect (IFE), which may thus be thought of as a particular limit of ISRS.[3] A similar effect (known as the inverse Cotton-Mouton effect (ICME)) is also present; expanding χ_{ij} to second order in $\boldsymbol{M}$, we have

$$F = -\chi_{ijkl} E_i(t) E_j(t) M_k M_l \tag{1.17}$$

---

[2]Here we treat the mode Q as nondegenerate, although the generalization to the degenerate case is straightforward.

[3]Despite the similarities presented here, the two effects are actually quite different. For example, the spectral content of the pump pulse in the IFE does not change before and after interacting with the sample[?].

and

$$H_l^{\text{eff}} = \frac{\partial F}{\partial M_l} \tag{1.18}$$

$$= \chi_{ijkl} E_i(t) E_j(t) M_k, \tag{1.19}$$

which, in contrast to the IFE, may occur when the pump is linearly polarized. The ICME is also available in AFMs; writing

$$F = -\chi_{ijkl} E_i(t) E_j(t) L_k L_l, \tag{1.20}$$

where $\boldsymbol{L}$ is the Néel vector (defined as the different in sublattice magnetizations), we have

$$H_l^{\text{eff}} = \frac{\partial F}{\partial L_l} \tag{1.21}$$

$$= \chi_{ijkl} E_i(t) E_j(t) L_k. \tag{1.22}$$

A full treatment of these and related effects is beyond the scope of this work;[4] the point is just to make a connection between the many different impulsive excitation mechanisms.

1.1.4 Collective modes in correlated materials

1.1.4.1 Phase transitions and Goldstone's theorem

In this section we illustrate one way we can understand the low-energy collective modes of a given material system without knowing precisely the form of the interaction Hamiltonian in Eq. 1.1. The important insight is that we can simply write down all of the *symmetry-allowed* terms in the Hamiltonian (or, at nonzero temperature, the free energy), and examine fluctuations near a phase transition where some order parameter takes on a nonzero value. In doing so, we find quite generally Goldstone's result that there is a gapless collective mode at the Γ point for each spontaneously broken symmetry. We also see that, in addition to the gapless Goldstone bosons, we also get a gapped *amplitude* mode which may be measurable in certain circumstances. The goal of this section is not to give a rigorous proof of Goldstone's theorem, but to illustrate how it comes about in a very simple model.

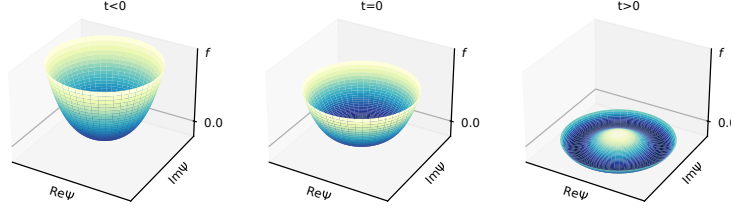[4]See Kirilyuk et al. [18] for a comprehensive review.

Figure 1.1: Schematic free energy density in the GL model.

Let us consider the GL free energy density in the case of superfluidity, which is a functional of the coarse-grained order parameter $\Psi(\boldsymbol{x}) \equiv \psi(\boldsymbol{x})e^{i\theta(\boldsymbol{x})}$ and whose lowest-order terms are

$$f[\Psi(\boldsymbol{x})] = -\frac{t}{2}|\Psi(\boldsymbol{x})|^2 + \frac{u}{4}|\Psi(\boldsymbol{x})|^4 + \frac{K}{2}|\nabla\Psi(\boldsymbol{x})|^2, \qquad (1.23)$$

where $u, K > 0$ and $t = (T_c - T)/T_c$. Clearly at all temperatures the minimum of Eq. 1.23 is spatially uniform, but for $T < T_c$ it occurs at a finite value of $\psi(\boldsymbol{x})$ (see Fig. 1.1)

$$\Psi(\boldsymbol{x}) = \overline{\psi} \equiv \sqrt{\frac{t}{u}}, \qquad (1.24)$$

where we have (without loss of generality) set $\overline{\theta} = 0$. Now consider fluctuations $\delta\theta(\boldsymbol{x})$ in the phase of the order parameter Ψ. We have

$$f[\overline{\psi}e^{i\delta\theta}] = f[\overline{\psi}] + \frac{K\overline{\psi}^2}{2}|\nabla\delta\theta(\boldsymbol{x})|^2, \qquad (1.25)$$

where we have used that $-t + u\overline{\psi}^2 = 0$. Let us write the fluctuation $\delta\theta(\boldsymbol{x})$ in terms of its Fourier transform

$$\delta\theta(\boldsymbol{x}) = \sum_{\boldsymbol{q}} e^{i\boldsymbol{q}\cdot\boldsymbol{x}}\delta\theta_{\boldsymbol{q}}. \qquad (1.26)$$

Then we have

$$F[\overline{\psi}e^{i\delta\theta(\boldsymbol{x})}] \equiv \int f[\overline{\psi}e^{i\delta\theta(\boldsymbol{x})}]\,d^3\boldsymbol{x} = F[\overline{\psi}] + \frac{KV\overline{\psi}^2}{2}\sum_{\boldsymbol{q}} q^2|\delta\theta_{\boldsymbol{q}}|^2. \qquad (1.27)$$

Thus, uniform ($\boldsymbol{q} = 0$) fluctuations in the phase of Ψ cost zero energy, and long-wavelength ($\boldsymbol{q} \approx 0$) fluctuations in the phase of Ψ cost very little energy. The collective modes corresponding to these long-wavelength fluctuations are exactly our Goldstone modes.

   Let us also examine the energy cost of fluctuations in the amplitude of Ψ, i.e. $\delta\psi(\boldsymbol{x}) \equiv \psi(\boldsymbol{x}) - \overline{\psi}$. To order $\delta\psi^2$, we have

$$f[\overline{\psi} + \delta\psi(\boldsymbol{x})] = f[\overline{\psi}] - \frac{t}{2}\delta\psi(\boldsymbol{x})^2 + \frac{3u}{2}\overline{\psi}^2\delta\psi(\boldsymbol{x})^2 + \frac{K}{2}(\nabla\delta\psi(\boldsymbol{x}))^2, \quad (1.28)$$

and, inserting $\overline{\psi} = \sqrt{\frac{t}{u}}$,

$$f[\overline{\psi} + \delta\psi(\boldsymbol{x})] = f[\overline{\psi}] + \frac{t}{2}\delta\psi(\boldsymbol{x})^2 + \frac{K}{2}(\nabla\delta\psi(\boldsymbol{x}))^2, \quad (1.29)$$

so that (for $T < T_c$), even *uniform* fluctuations in the amplitude of Ψ cost nonzero energy. The collective mode associated with such uniform fluctuations is the so-called "amplitude" or "Higgs" mode of our ordered superfluid, with the important result that the energy of this mode approaches zero for $T \to T_c$.

   Observing the amplitude mode in real systems is challenging because its decay into the lower-energy Goldstone modes is not typically forbidden, and so its lifetime is usually too low to observe it as a true quasiparticle[?]. Nevertheless, amplitude modes have been observed in a number of different circumstances; mainly, of course, in superconductors, where the amplitude mode may appear as a Raman peak in the A_{1g} channel[?], but also in charge density wave[?] and magnetic[? ?] systems (see ?] for a good review).

1.1.4.2   Example: electromagnons in multiferroics

In the above we have described how coherent oscillations of collective modes in solids may be excited with an ultrafast laser, and we have explored how we can understand what those modes are without having to solve the full many-body Hamiltonian. Let us start to apply this machinery to a relevant physical system. A full accounting of the different collective modes in correlated systems is beyond the scope of this work; instead. let us focus on a particular class of physical system—multiferroics—which encompasses a lot of the physics of correlated materials and may also be

described well in terms of the GL theory we outlined above. **??** is concerned with one member of this class ($CuBr_2$); in this material, we believe we are seeing the amplitude mode of the magnetic Hamiltonian imprinted on the charge degree of freedom in the form of an electromagnon. The understanding developed in this chapter will thus be quite helpful when we encounter that chapter.

Multiferroics are materials in which magnetic order exists in the same phase as ferroelectric order. They may typically be classified into two types. In the so-called type-II multiferroics, the transition to the ferroelectric state appears at the same temperature as a concomitant magnetic transition. This is because the ferroelectricity is *due to* some inversion symmetry breaking of the magnetic order, for instance by a helical spin density wave.[5] The microscopic mechanism will be discussed shortly, but let us contrast this with the case of type-I multiferroics—in type-I multiferroics, the magnetic and ferroelectric transition temperatures are different, and neither is "due" to the other; they just happen to exist in the same material. Intuitively this may happen when magnetic ions exist alongside different ions which participate in the ferroelectricity; this is the case in $BiFeO_3$[?]. Here we will focus on type-II multiferroics, not only because they tend to be more interesting but also because they usually have larger magnetoelectric effects (since the ferroelectric order is due to the magnetic order) and have thus been the central focus of the modern research on multiferroics.

The central question is then: can we describe (microscopically) how ferroelectricity may appear as a *result* of long-range magnetic order? A critical insight due to **?**] is that type-II multiferroicity can happen in quasi-1D Mott insulators when spin-orbit coupling is strong and the spins are ordered in a *magnetic spiral*. **?**] consider a 3-site 1D cluster model in which two transition metal ions with local spin-orbit coupled d electron orbitals interact via superexchange across an intermediate ligand atom (with an associated set of p orbitals). The tendency towards magnetic ordering is put in by hand, by placing a local spin at each transition metal site j in the direction $\hat{e}_j = (\cos\phi_j \sin\theta_j, \sin\phi_j \sin\theta_j, \cos\theta_j)$, and letting the d spins $\boldsymbol{S}_j$ interact with these local spins via a Hamiltonian $H_U = -U\sum_j \hat{e}_j \cdot \boldsymbol{S}_j$. They find quite generally the following geometrical relation between the

---

[5]The ferroelectricity is thus considered of the "improper" type.

spin directions $\hat{e}_j$ and an *induced* polarization $\boldsymbol{P}$:

$$\boldsymbol{P} \propto \hat{e}_{12} \times (\hat{e}_1 \times \hat{e}_2) \tag{1.30}$$

where $\hat{e}_{12}$ is the unit vector parallel to the bond connecting the two transition metal sites.

Clearly if $\hat{e}_1$ and $\hat{e}_2$ are either parallel or anti-parallel to one another, there is no induced polarization. However, if the magnetic order is *non-colinear*, and the spin plane is not perpendicular the chain direction, an induced polarization appears via Eq. 1.30. This explains very generally the multiferroicity of many noncollinear magnets. The multiferroic state thus described is also clearly "type-II", since the polarization is zero absent the magnetic ordering. Heuristically, such a spiral magnetic order may occur due to competition between a ferromagnetic nearest neighbor exchange interaction J_1 and an antiferromagnetic next-nearest neighbor interaction J_2 in a 1D spin chain; for $|J_2/J_1| < 4$ one can show that the ground state is such a noncollinear spiral.

Let us now consider the low-energy excitations of such a 1D chain. The Hamiltonian may be written as

$$H = H_1 + H_2 + H_3 + H_4, \tag{1.31}$$

where

$$H_1 = \sum_i J_1 \boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1} + J_2 \boldsymbol{S}_i \cdot \boldsymbol{S}_{i+2}, \tag{1.32}$$

$$H_2 = -\lambda \sum_i \boldsymbol{u}_i \cdot [\hat{x} \times (\boldsymbol{S}_i \times \boldsymbol{S}_{i+1})], \tag{1.33}$$

$$H_3 = \sum_i \frac{\kappa}{2} \boldsymbol{u}_i^2 + \frac{1}{2m} \boldsymbol{P}_i^2, \tag{1.34}$$

and

$$H_4 = \sum_i D(S_i^z)^2. \tag{1.35}$$

This Hamiltonian describes a 1D spin chain along $\hat{x}$ interacting with a charge coordinate u_i (with conjugate momentrum P_i) via Eq. 1.30. The anisotropy $D > 0$ pins the magnetic order to the xy plane. As discussed above, for $|J_2/J_1| < 4$ the ground state is a noncollinear spin spiral with spins lying in the xy plane, and the relative angle between adjacent spins is defined by some quantity $\phi \neq 0, \pi$ which is related to $|J_2/J_1|$.

?] discuss the low-lying excitations of this ground state in light of Eq. 1.31. First of all, clearly uniform rotations of all of the spins in the chain cost zero energy; this is the *Goldstone mode* of Eq. 1.31, and is in direct analogy with the Goldstone mode of the superfluid discussed in section 1.1.4.1.

If we temporarily set $D \to 0^+$ in Eq. 1.35, we may also observe that, in this limit, uniform rotations of the spin spiral *about the chain axis* also cost zero energy. This mode is also a Goldstone mode, although in this case it is due to the spontaneous breaking of the $\hat{x}$ rotational symmetry rather than the $\hat{z}$ rotational symmetry. For nonzero D, this mode of course acquires a finite energy; it is thus referred to as the *pseudo*-Goldstone mode, to highlight the fact that the symmetry is explicitly broken by the Hamiltonian rather than spontaneously broken.

The pseudo-Goldstone mode, however, is still quite interesting—note that, due to Eq. 1.33 and Eq. 1.30, a rotation of the spins about the chain axis *also* rotates the induced polarization about that axis. This mode is thus our first encounter with an electromagnon, which is any spin boson that acquires a finite electric dipole moment due to coupling to the charge degrees of freedom.

Electromagnons are in fact an intrinsic feature of multiferroics, and what we have discussed above is by no means the only mechanism by which such modes may be achieved.[6] While a complete description of electromagnons and multiferroics is beyond the scope of this work (see **?**] and **?**] for review), I hope to have nevertheless motivated the idea that elementary excitations in correlated systems may still be understood in the context of broken symmetries and the GL paradigm—a notion that we will encounter frequently in the course of this thesis.

1.2 Control

From a technological perspective, ultrafast spectroscopy is useful because it helps us understand the physics of materials in equilibrium and how those materials might function in practical applications. Ultrafast control is different in the sense that it seeks to understand the different ways we

---

[6]For example, in cases where magnetoelastic coupling is strong, a multi-particle excitation in which a spin boson couples to a phonon, and a phonon then couples to the electric polarization, is possible and common[**?**].

can *manipulate* materials, with the idea that the knowledge thus gained may be used as a control scheme in some future technology, whether that involves using light or some other stimulus. The material of study is thus of a different kind of importance in ultrafast control—we are more interested in the phenomenology of the light-matter interaction rather than the phenomenology of the material itself. Usually the phenomenology of interest is, for example, being able to control the *state* of the material (like the direction of a magnetization, or the presence of a charge density wave) using light.

In my view, the dynamics relevant for ultrafast control fall into two categories. In one case, the dynamics following the light pulse are *coherent*; that is, the light field, for instance, excites coherent oscillations (see section 1.1.2), and those coherent oscillations drive the system to a new state.[7] In contrast, *incoherent* dynamics do not involve coherent oscillations of any degrees of freedom; instead, the light is used to, for example, heat the material, or excite quasiparticles away from the Fermi surface. This can also drive the material into a different (nonequilibrium) state, for example by quenching some long-range order (LRO) of the electrons near the Fermi surface, but the mechanism is qualitatively different from the coherent case. We will discuss both cases below.

1.2.1   Incoherent control

1.2.1.1   Heating across a phase transition

The most basic level of ultrafast control is to use the light pulse simply to heat the material; if the initial and final temperatures lie below and above an equilibrium phase transition, then ultrafast control has been achieved. This is a somewhat trivial mechanism, but it played an important part in early research on ultrafast control, especially in materials where dynamical phase control *not* based on heating was rare or difficult (such as antiferromagnets[16]). In addition, interactions between the sample area melted by the pump pulse and the surrounding unpumped area may cause interesting effects, such as magnetization reversal in ferromagnetic thin films due to the stray field of the unpumped sample volume[**?**]. We

---

[7]Here I take the view that simply exciting coherent oscillations (i.e. without a dynamical phase transition) is *not* coherent control, but some may disagree with this point perspective, especially when the amplitude of such oscillations are large.

are more interested, however, in *nonthermal* ultrafast phenomena—that is, dynamical phenomena which cannot be replicated quasi-adiabatically.

1.2.1.2 Nonthermal melting

One example of such a phonemonon is the ultrafast quench (or "nonthermal melting") of some equilibrium order parameter via photoexcitation of electrons by the pump pulse. Let us consider, for example, the case discussed briefly in the beginning of this chapter of a quasi-1D Fermi liquid which becomes gapped by a periodic lattice distortion (i.e. via the Peierls mechanism). The basis for this mechanism is that quasiparticles near the Fermi vector in the undistorted Fermi liquid decrease their energy when the gap is opened via hybridization of folded bands, which happens at the expense of some energy gain associated with distorting the lattice away from its equilibrium configuration. When light is incident on this system in the distorted phase,[8] quasiparticles are excited *away* from the valence band top, and the energy savings we gained by distorting the lattice disappears. The system thus relaxes back to its previous configuration and we lose our periodic lattice distortion—the light pulse destroyed the LRO.

Note that this is referred to as a *nonthermal* effect, even though for typical metals the quasiparticles experience an effective temperature change in this picture on the order of 10^3 K[**?**]. Nevertheless, the dynamical phase transition we just described proceeds, by construction, via the creation of a *nonequilibrium* distribution of photoexcited electrons, and in this sense it is nonthermal. There is often difficulty encountered in *proving* that the action of the pump is truly nonthermal. Indeed, the existence of an ordered phase implies some finite temperature T_c where that order is destroyed by equilibrium thermal fluctuations—it is the job of the experimenter to show that the laser pulse is not simply heating the material above this critical temperature. This is obvious if the phenomenology following the quench (see the next section) is in any case not reproduced in equilibrium, but may also be shown (heuristically) by dividing the energy absorbed by the material from the pulse by the heat capacity.

---

[8]Let us say that the pump photon energy is above the Peierls gap.

1.2.1.3   Formal model: time-dependent Ginzburg-Landau theory

In describing many-body systems in the static limit, we have seen that it is useful to define a phenomenological model which captures the important physics (via a free energy density, an order parameter, etc.), rather than attempting to treat the inter-particle interactions explicitly. An important concept justifying this practice is the idea of *universality*, which states that, since the correlation length ξ diverges as we approach the critical point, the physics near the critical point is governed not by the microscopic details of the system but rather by a small set of characteristic quantities like the spatial dimensionality of the system and the symmetry of the order parameter[**?**]. This means that to understand any *complicated* system of a particular universality class, it suffices to study the *simplest* system in that universality class, as long as we are close to the critical point.[9]

In the dynamic limit, the idea of universality is still thought to apply (thus, for example, the wavevector- and frequency-dependent susceptibility $\chi(\boldsymbol{k}, \omega)$ is a homogenous function near the critical point), but the relevant parameters now include not just the spatial dimensionality and symmetry of the order parameter but also new characteristics, like the presence of conserved quantities, and Poisson-bracket relationships between those conserved quantities and the order parameter[**?**]. We are thus motivated, as in the static case, to consider *simple models* for each universality class to derive the important limiting behavior—this is the time-dependent Ginzburg-Landau (TDGL) theory. One such model, which is valid when there are no conserved quantities, is the so-called "Model A" equation of motion

$$\frac{\partial \psi(\boldsymbol{x}, t)}{\partial t} = -\Gamma \frac{\delta F}{\delta \psi(\boldsymbol{x}, t)} + \eta(\boldsymbol{x}, t), \tag{1.36}$$

where

$$F = \int \left[-\frac{t}{2}\psi^2 + \frac{u}{4}\psi^4 + \frac{K}{2}|\nabla \psi|^2 \right] d^d \boldsymbol{x} \tag{1.37}$$

and η is a Gaussian white noise source with zero mean and two-point correlation function

$$\langle \eta(\boldsymbol{x}, t)\eta(\boldsymbol{x}', t') \rangle = 2\Gamma \delta(\boldsymbol{x} - \boldsymbol{x}')\delta(t - t') \tag{1.38}$$

---

[9]The validity of the universality hypothesis away from criticality is not obvious, but is often heuristically assumed to be the case.

chosen to satisfy the fluctuation-dissipation theorem.

Different models are appropriate for systems with different conserved quantities. For example, the system which otherwise looks like the above but with where the order parameter is a conserved quantity may be modeled using what is referred to as "Model B"[?]; the equation of motion is

$$\frac{\partial \psi(\boldsymbol{x}, t)}{\partial t} = \lambda \nabla^2 \frac{\delta F}{\delta \psi(\boldsymbol{x}, t)} + \eta(\boldsymbol{x}, t). \tag{1.39}$$

A comprehensive review of all such models is given by ?].

1.2.1.4 Dynamics following an ultrafast quench

TDGL is thus used widely to study dynamical phase transitions in many physical systems. An interesting case, which we will treat in detail here due to its applicability to $CaMn_2Bi_2$ (see **??**), is when two order parameters compete with each other in the equilibrium free energy density, and we consider the TDGL dynamics after an *ultrafast quench* of the equilibrium order.[10] The free energy density is

$$f = f_1 + f_2 + f_c, \tag{1.40}$$

where

$$f_i = -\frac{t_i}{2}\psi_i^2 + \frac{u_i}{4}\psi_i^4 + \frac{K_i}{2}(\nabla \psi_i)^2 \tag{1.41}$$

and

$$f_c = c\psi_1^2\psi_2^2. \tag{1.42}$$

?? describes a system with two single-component order parameters ψ_1 and ψ_2 with a coupling term f_c. Without loss of generality, let us consider the case where $t_2 > 0$ and $t_1 < 0$, so that ψ_2 is favored in equilibrium, but let c take some value so that there is an additional local minimum for $\psi_1 \neq 0$.

We consider an ultrafast quench of the equilibrium order parameter ψ_2 which corresponds to a free energy density (**??**) with $t_1, t_2 < 0$. Since neither of the order parameters are conserved quantities, nor are there any other conserved quantities in the system, we are free to use Model A (Eq. 1.36) to describe the dynamics. Using Eq. 1.36, Sun and Millis [35]

---

[10]This discussion closely follows that of Sun and Millis [35].

showed that, as the relaxation proceeds towards (meta)stability, there is a
probability

$$p \sim 1 - \zeta^\delta \tag{1.43}$$

that the final state of the system has $\psi_1 \neq 0$ and $\psi_2 = 0$; i.e., that the order
has switched from ψ_2 to ψ_1. Here, ζ is the Ginzburg parameter which is
small when the mean-field theory is valid, and δ is related to the relaxation
rates γ_1 and γ_2 (see Eq. 1.36) of the two order parameters:[11]

$$\delta \sim \left| \frac{\gamma_1 t_1}{\gamma_2 t_2} \right|. \tag{1.44}$$

Thus, as long as mean-field theory holds and $\delta > 1$, the final state after
the quench is the metastable state with $\psi_1 \neq 0$.

The intuition associated with this result is that, following the ultrafast
quench, the dynamics are dominated not by whichever state has the lowest
free energy in equilibrium, but by the exponential amplification of long-
wavelength spatial fluctuations which occurs for *both* order parameters.
Thus, even a small relative difference in growth rates of these fluctuations
becomes exponentially amplified as the system relaxes. In light of Eq. 1.36,
this means that the dynamics are determined not by the minima of the free
energy, but by its gradient. At long times the equilibrium state of course
recovers, although this recovery involves slower processes like domain
nucleation so that the metastable state survives for long timescales.

Many examples of ultrafast trapping into metastable states following
an ultrafast quench can be understood in this paradigm. Sun and Millis
[35] consider, for instance, the enhancement of superconductivity that
occurs in $La_{1.675}Eu_{0.2}Sr_{0.125}CuO_4$ upon transient melting of the equilibrium
CDW state in that compound[11]. Since the CDW state—in contrast to
the superconducting state—involves collective motion of electrons and
the (heavy) lattice at the same time, recovery of this state after photoex-
citation is much slower than recovery of the superconducting state. The
superconducting state is thus favored after the quench.

The framework has also been applied to the case of $LaTe_3$, in which
an a-axis CDW appears in nonequilibrium following a quench of the
equilibrium c-axis CDW order.[12] In this case, both states are CDWs so the

---

[11]This is the value of δ when $\gamma_1 \gg \gamma_2$, but the same conclusions apply when $\gamma_1 \gtrsim \gamma_2$;
see Sun and Millis [35] for details.

[12]Strictly speaking the induced a-axis order is not truly long-ranged, see **?**].

relaxation rates γ_1 and γ_2 should be roughly equal. Eq. 1.44 should thus be replaced by a more stringent condition which also includes the fact that, if the cooling rate following the quench is too slow, the exponential amplification described above may not be effective in switching to the metastable state[35].

1.2.2 Coherent control

1.2.2.1 Via coherent oscillations

In the previous section we described how ultrafast phenomena may be understood phenomenologically using the GL or TDGL paradigm, where the important physics is captured by a generic a free energy gensity whose minima, gradients, etc. determine the relevant phases. In describing the action of the pump pulse, we treated the case where the light only quenches the equilibrium order via excitation of electrons from near the Fermi level, but of course, we know from section 1.1 that this is hardly the only thing the pump can do. For example, we studied in detail in section 1.1 that the pump may indeed excite coherent oscillations, and if these oscillations are strong, it turns out that they may also allow us to control the nonequilibrium or metastable phases of materials.

Let us consider for example the free energy density of **??**, and let us imagine that the constant c has some dependence on the coordinate Q of some collective mode, i.e.

$$f_c = \left(c(0) + \frac{\partial c(Q)}{\partial Q}\bigg|_{Q=0} Q + \dots \right) \psi_1^2 \psi_2^2. \tag{1.45}$$

Clearly if the pump pulse brings Q to some finite value, the free energy landscape may look quite different compared to equilibrium.

One example of this phenomenon occurs on intense resonant excitation of the $17\,\mathrm{THz}$ phonon mode in $DyFeO_3$, which is a magnetic material with a thermal metamagnetic transition at $51\,\mathrm{K}$ between a low-temperature AFM state and a high-temperature canted-AFM state with a small magnetic moment. Afanasiev et al. [1] showed the appearance of a long-lived offset in the Faraday effect signal following excitation which disappeared above $51\,\mathrm{K}$, signaling that the pump pulse triggers an ultrafast phase transition between the two magnetic states in this material. They claim that anharmonic phonon interactions rectify the coherent oscillation of the

$17\,\mathrm{THz}$ phonon mode and transfer a finite, time-independent amplitude to a *different* phonon mode associated with the displacement of Dy^{3+} ions. This displacement causes a long-lived change in the magnetic exchange interaction which modifies the free energy density and triggers a phase transition in a manner resembling Eq. 1.45. This is just one example of a whole field of research referred to as *nonlinear phononics*[?], which seeks to use anharmonic interactions like these to effect control over the crystal lattice with light.

1.2.2.2 Via Floquet engineering

A related scenario occurs when the light couples not to the constants t, u, c, etc. of the free energy, but rather acts as a *field h* for one of the order parameters, i.e. by the IFE (see section 1.1.3.2). Then we have, for example,

$$f(t) = f(0) + h(t)\psi_1(t). \tag{1.46}$$

and the field may thus provide an ultrafast "kick" to push the material into a new state.[13] In this case, the purpose of the pump is *not* to excite coherent oscillations—which go on to modify the free energy landscape—but rather to control the free energy landscape via the light-matter interaction *itself*. The study of mechanisms which fall under this umbrella is known as "Floquet engineering"[?], referring to the Floquet theorem[?] which gives the energy eigenstates of a Hamiltonian which is periodic in time.

The fact that an ultrafast light pulse may produce Floquet bands in solids and that those bands may hybridize with the eigenstates of the time-independent Hamiltonian was established by ?] and later ?] using time-resolved photoemission spectroscopy. Multiple control strategies using this effect, like the manipulation of band topologies[?], observation of a large optical stark effect[?], and manipulation of the magnetic exchange energy[?], were shown to occur as a result of this effect. A complete accounting of the progress experimentally and theoretically in this field is beyond the present scope, but the literature has been reviewed extensively by, e.g. ?] and ?].

---

[13]This is especially relevant in the case of magnets, where the magnetic field from the IFE can couple directly to the magnetic moment, although one needs intertial terms in the Lagrangian which are not present in the Landau-Lifshitz equation for a simple ferromagnet[17].

1.2.3 Control as spectroscopy

The above treatment of ultrafast control was motivated by a desire to manipulate quantum materials with light for some practical benefit, such as, for example, being able to control the direction of the spins in a magnetic memory device. However, there is an additional, related motivation which is quite important to note. Clearly, the mechanism by which the light manipulates the phases of materials depends quite sensitively on the material characteristics; for example, the discussion in section 1.2.1.4 depended critically on the fact that the two order parameters ψ_1 and ψ_2 *compete* in equilibrium—a which may not be obvious otherwise! Indeed, the fact that light may be used to suppress some phases and enhance others is sometimes the only information we have about how and whether those different phases interact with each other. The fact that CDW order competes with superconductivity in cuprates is, for example, strongly evidenced by the enhancement of superconductivity that occurs upon a photoinduced quench of the CDW order[11]. A similar observation about CDWs in CsV_3Sb_5 was recently reported by **?**]. The line between ultrafast spectroscopy and ultrafast control should not, then, be considered particularly sharp, and both of the motivations described here—to learn about materials, and to manipulate materials—are deeply related to one another.

Chapter Two

Second harmonic generation: theory

2.1 Space groups, point groups, and Neumann's principle

The utility of SHG in studying condensed matter systems is derived from the following simple statement, attributed to Franz Neumann[25] and later Pierre Curie[9]:

Theorem 2.1.1 (Neumann's principle) *Let P_G be the symmetry group of a crystal structure and P_H the symmetry group of some physical property of that crystal. Then, P_G is a subgroup of P_H.*

There are a few things to digest here. Let us start by understanding the meaning of the phrase "symmetry group". For any given crystal, there exists some infinitely large set of operations G under which the crystal structure is symmetric. Each of these operations may be decomposed into two parts: a "point-preserving operation" R, corresponding to either the identity, rotation, inversion, mirror, or the product of mirror and rotation, followed by a translation by some vector $\boldsymbol{\tau}$:

$$G = \{(R|\boldsymbol{\tau})\} \tag{2.1}$$

where $(R|\boldsymbol{\tau})$ means "Perform R, then translate by $\boldsymbol{\tau}$". Clearly, the set G forms a group, since if both g_1 and $g_2 \in G$ leave the crystal structure

invariant, so does the product $g_1 g_2$, and so $g_1 g_2 \in G$. Thus, G is called the *space group* of the crystal. In three dimensions, there are 230 crystallographic space groups, which are tabulated in a number of places, most usefully Wikipedia[7].

For 73 of these groups, the translation parts of the τ's in Eq. 2.1 are only ever linear combinations of integer multiples of the lattice vectors $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$; these are called *symmorphic* space groups. The remaining 157 groups involve translations that are not integer multiples of the lattice vectors; these are one's screw axes and glide planes, and so these groups are called *asymmorphic*.

Importantly, the "physical properties" of theorem 2.1.1 refer to the truly macroscopic properties of the crystal, like its conductivity, dielectric, or pyroelectric tensors. Consider, for example, that in SHG, we are typically studying the sample at optical wavelengths, where the wavelength of light is three or four orders of magnitude larger than the lattice spacing. Clearly, then, these properties do not care whether the correct symmetry is $(R|\tau)$ or $(R|\tau + {}^{a}\!/_{2})$. A more useful group, then, is the *point group* of the crystal

$$P_G = \{R \text{ s.t. } \exists\, \tau \text{ s.t. } (R|\tau) \in G\} \tag{2.2}$$

i.e., the point group is the set of point-preserving operations R for which R appears in G, regardless of whether you need to perform a translation with it. Once can show that P_G is also a group, and thus it is P_G which is involved in Neumann's principle for all intents and purposes[1].

The last ingredient that we need to understand theorem 2.1.1 is the concept of what is meant by "physical propery". The idea is that the response of the crystal $J_{i_1 i_2 \cdots i_n}$ (i.e., the current J_i, or the quadrupole moment Q_{ij}) is proportional to some field $F_{i'_1 i'_2 \cdots i'_m}$ via some tensor χ:

$$J_{i_1 i_2 \cdots i_n} = \chi_{i_1 i_2 \cdots i_n i'_1 i'_2 \cdots i'_m} F_{i'_1 i'_2 \cdots i'_m}. \tag{2.3}$$

For example, the conductivity σ_{ij} relates a current density J_i to an applied electric field E_j:

$$J_i = \sigma_{ij} E_j. \tag{2.4}$$

---

[1]Of course this breaks down when the wavelength of light is comparable to the lattice spacing; in that case you need to consider the full space group.

Likewise, the polarization P_i due to the pyroelectric effect is related to the a temperature difference ΔT by a tensor p_i:

$$P_i = p_i \Delta T. \tag{2.5}$$

The tensors σ_{ij}, p_i, and generally, $\chi_{i_1 i_2 \cdots i_n i'_1 i'_2 \cdots i'_m}$ are commonly referred to as *matter tensors*[30], to emphasize the fact that they are the only part of the response equations that depend on the material. It should be noted that matter tensors generically come in two types: those that transform like a vector under inversion and those that transform like a pseudovector under inversion. You can tell which is which by applying inversion to either side of the response equation. For example, the tensor ϵ_{ij} relating the displacement field to the electric field

$$D_i = \epsilon_{ij} E_j \tag{2.6}$$

is a polar tensor, whereas the tensor χ_{ij}^{me} describing the magnetoelectric effect

$$M_i = \chi_{ij}^{me} E_j \tag{2.7}$$

is an axial tensor.

We are now ready to restate theorem 2.1.1 in a slightly more useful form, using the terminology we have developed about point groups and matter tensors:

Theorem 2.1.2 (Neumann's principle, restated) *Let P_G be the point group of a given crystal, and let χ be a matter tensor describing some response function of that crystal. Then, for all $g \in P_G$, we have*

$$g(\chi) = \chi. \tag{2.8}$$

Eq. 2.8 can be more usefully expressed if we know the matrix R_{ij}^g corresponding to g. For example, if g is "threefold rotation about the z axis", we have

$$R_{ij}^g = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{2.9}$$

in which case one can show that Eq. 2.8 reads

$$(\det R^g)^t R_{i_1 i'_1}^g R_{i_2 i'_2}^g \cdots R_{i_n i'_n}^g \chi_{i'_1 i'_2 \cdots i'_n} = \chi_{i_1 i_2 \cdots i_n}, \tag{2.10}$$

where t is 0 if χ is a polar tensor and 1 if χ is an axial tensor. Theorem 2.1.2 tells us that there is one copy of Eq. 2.10 for each $g \in P_G$.

Apparently, each element $g \in P_G$ gives us a *constraint* on the numbers $\chi_{i_1 i_2 \cdots i_n}$, in that they have to satisfy Eq. 2.10. This is a remarkably useful fact. Since different point groups enforce different constraints on χ, that means the *form* of χ (e.g. when written as a list of numbers) depends quite sensitively on the point group of the crystal we are studying. As an example, here is the dielectric permittivity tensor for crystals with the point group (in Schoenflies notation) C_2:

$$\epsilon_{ij} = \begin{pmatrix} a & 0 & e \\ 0 & b & 0 \\ e & 0 & c \end{pmatrix}_{ij} \tag{2.11}$$

versus in the point group D_{3d}:

$$\epsilon_{ij} = \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & c \end{pmatrix}_{ij}. \tag{2.12}$$

Clearly, any *measurement* of ϵ_{ij} will be able to easily differentiate a crystal with point group C_2 from one with point group D_{3d}. This is the fundamental basis, then, for SHG. In SHG, we measure the tensor χ_{ijk} corresponding to the response equation[2]

$$P_i(2\omega) = \chi_{ijk} E_j(\omega) E_k(\omega); \tag{2.13}$$

the numbers χ_{ijk} thus tell us about the crystallographic point group we are measuring from.

There are a couple of advantages to measuring χ_{ijk} over any other matter tensor in a given system. For one thing, χ_{ijk} is a third rank tensor, which means it has a few more degrees of freedom to work with compared to ϵ_{ij}, and thus does a better job at uniquely specifying each point group. It also doesn't have *too many* degrees of freedom, so that most of the time your experiment will be able to tell you all of your tensor elements[3]. In addition, since we are typically doing SHG at optical wavelengths, the

---

[2]This discussion is a bit simplified in the sense that there are actually *many* response functions which will give you light at 2ω; for a more detailed discussion, see section 2.2.

[3]Quadrupole SHG has this problem, see section 2.2.

form of χ_{ijk} reflects the symmetry of the *charge distribution* $\rho(\boldsymbol{x})$, in contrast to e.g. x-ray diffraction, where the relevant tensors will tell instead you about the electron distribution, $n(\boldsymbol{x})$. This can be advantageous in cases where the long range order you are trying to study involves an ordering of the valence electrons but not the electrons in the cores of atoms. This is entirely the result of the fact that Neumann's principle, as expressed both in theorem 2.1.1 and theorem 2.1.2, tells us that the point group of our crystal is a *subgroup* of the point group we get from our measurement—the measurement can always be more symmetric than the crystal!

As another example of this fact, let us note that the response equation given by Eq. 2.13 clearly has an additional symmetry $j \leftrightarrow k$, since the two copies of the electric field on the right hand side are equivalent. Obviously this is not a result of the material we are studying, it is simply a fact of doing SHG. Thus, in addition to the constraints given by Neumann's principle and Eq. 2.10, we have the additional constraint

$$\chi_{ijk} = \chi_{ikj} \forall i, j, k. \tag{2.14}$$

This is known as *particularization*[4].

2.2 A classical understanding of SHG

In the last section we considered the SHG response function given by Eq. 2.13. Where does this relationship come from, and how is $P(2\omega)$ eventually measured? Our starting point in the classical treatment will be the inhomogenous electromagnetic wave equation

$$\left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right) E_i(\boldsymbol{x}, t) = S_i(\boldsymbol{x}, t), \tag{2.15}$$

which we understand as defining the field $E_i(\boldsymbol{x}, t)$ radiated by the source term $S_i(\boldsymbol{x}, t)$, which is induced by the incident field. To lowest order in a multipole expansion, $S_i(\boldsymbol{x}, t)$ is given by[14, 19]

$$\mu_0 \frac{\partial^2 P_i(\boldsymbol{x}, t)}{\partial t^2} + \mu_0 \left(\epsilon_{ijk}\nabla_j \frac{\partial M_k(\boldsymbol{x}, t)}{\partial t}\right) - \mu_0 \left(\nabla_j \frac{\partial^2 Q_{ij}(\boldsymbol{x}, t)}{\partial t^2}\right) \tag{2.16}$$

where $P_i(\boldsymbol{x}, t)$, $M_i(\boldsymbol{x}, t)$, and $Q_{ij}(\boldsymbol{x}, t)$ are the induced electric dipole, magnetic dipole, and electric quadrupole densities, and ϵ_{ijk} is the Levi-Civita tensor.

If the incident electric field is small, then the terms $P_i(\boldsymbol{x}, t)$, $M_i(\boldsymbol{x}, t)$, and $Q_{ij}(\boldsymbol{x}, t)$ are linear functions of that electric field. However, for larger incident fields (such as those generated by pulsed lasers), they may be more generally written as a taylor series:

$$P_i = \chi_{ij}^{ee} E_j + \chi_{ij}^{em} H_j + \chi_{ijk}^{eee} E_j E_k + \chi_{ijk}^{eem} E_j H_k + \cdots \tag{2.17}$$

$$M_i = \chi_{ij}^{me} E_j + \chi_{ij}^{mm} H_j + \chi_{ijk}^{mee} E_j E_k + \chi_{ijk}^{mem} E_j H_k + \cdots \tag{2.18}$$

$$Q_{ij} = \chi_{ijk}^{qe} E_k + \chi_{ijk}^{qm} H_k + \chi_{ijkl}^{qee} E_k E_l + \chi_{ijkl}^{qem} E_k H_l + \cdots \tag{2.19}$$

where we have suppressed the arguments $\boldsymbol{x}$ and t for brevity.

Assuming the incident field is monochromatic,

$$E_i(\boldsymbol{x}, t) = E_i(\omega) e^{i(\boldsymbol{k} \cdot \boldsymbol{x} - \omega t)} + \text{c.c.} \tag{2.20}$$

$$H_i(\boldsymbol{x}, t) = H_i(\omega) e^{i(\boldsymbol{k} \cdot \boldsymbol{x} - \omega t)} + \text{c.c.} \tag{2.21}$$

the induced sources are also monochromatic, and (keeping only terms proportional to $e^{i2\omega t}$) we thus get

$$P_i(2\omega) = \chi_{ijk}^{eee} E_j(\omega) E_k(\omega) + \chi_{ijk}^{eem} E_j(\omega) H_k(\omega) \tag{2.22}$$

$$M_i(2\omega) = \chi_{ijk}^{mee} E_j(\omega) E_k(\omega) + \chi_{ijk}^{mem} E_j(\omega) H_k(\omega) \tag{2.23}$$

$$Q_{ij}(2\omega) = \chi_{ijkl}^{qee} E_k(\omega) E_l(\omega) + \chi_{ijkl}^{qem} E_k(\omega) H_l(\omega). \tag{2.24}$$

Since Eq. 2.15 is linear, the electric field radiated by $S_i(\boldsymbol{x}, t)$ is simply proportional to it. In the limit where the first term of Eq. 2.22 dominates, the intensity measured at our detector thus satisfies

$$I(2\omega) \propto |\hat{e}_i^{\text{out}} \chi_{ijk}^{eee} \hat{e}_j^{\text{in}} \hat{e}_k^{\text{in}}|^2, \tag{2.25}$$

where $\hat{e}^{\text{in}}$ and $\hat{e}^{\text{out}}$ are unit vectors in the direction of the incoming and measured electric fields[4]. χ_{ijk}^{eee} does typically dominate when inversion symmetry is broken, but if not, you have to consider all of the terms in Eqs. 2.22–2.24. Actually, each of these terms needs to be considered twice, since there is both a surface contribution and a bulk contribution[5]. In my experience, the heirarchy of contributions (from most to least important, and assuming everything is allowed by symmetry) is typically:

---

[4]Usually there are polarizers in the experiment which define these directions.

[5]The space group which constrains the surface contributions is the bulk space group less the operations which involve some change in the z coordinate.

1. Bulk electric dipole

2. Surface electric dipole, bulk electric quadrupole, and bulk magnetic dipole, at the same order[6]

3. Everything else

I've never seen anything outside of items 1 and 2, but in rare cases an electronic resonance may cause an enhancement in one of the other contributions[12].

Let us take a moment now to emphasize the following extremely common misconception about SHG: just because you see SHG in your experiment, that does not mean that inversion symmetry is broken in your material! It also does not mean that your material is a ferroelectric, or really that there's anything special at all about your material, at least before you've done any further analysis. Similarly, if you *don't* see SHG, that doesn't mean inversion symmetry is preserved, either. I have repeatedly seen large electric quadrupole SHG show up in materials with inversion symmetry, while materials which definitely break inversion symmetry have absolutely zero SHG observable in the experiment. The reason for this is ultimately due to resonance, a topic which I will discuss in section 2.3, but I mention it here because it is truly quite common in the literature and it is surely a mistake worth avoiding. You are "allowed" to say your material breaks inversion symmetry only if there is no other contribution in Eqs. 2.22–2.24 which fits your data, and you are basically never allowed to say that your material preserves inversion symmetry when there is no SHG (a fact that should be obvious on a careful reading of theorem 2.1.1).

2.3   SHG in quantum mechanics

The description of SHG in the previous section is probably the most useful for understanding SHG from an "optics" perspective, but it gives little insight into the true microscopic origin of the SHG intensity. The

---

[6]Somehow the bulk electric quadrupole and magnetic dipole contributions have been labelled "exotic" by some in the community, but that has not been my experience. If I had to guess, almost half of the materials I have measured with inversion symmetry show electric quadrupole SHG.

quantum description, on the other hand, will tell you exactly where the SHG is coming from microscopically, but only if you have access to the eigenfunctions $|\psi\rangle$ of your hamiltonian—it is of little use otherwise. Nevertheless, we can still gain intuition about the dependence of our SHG intensity on the frequency of the light in the quantum picture, which will be useful for clearing up a whole other slew of misconceptions that have somehow made their way into the SHG literature. This treatment closely follows that of Boyd [5].

The starting point is to describe the system under study as a statistical ensemble specified by a Hamiltonian

$$H = H_0 + \lambda V \tag{2.26}$$

and a density matrix

$$\rho(t) = \sum_i p_i(t) \, |\psi_i(t)\rangle \, \langle\psi_i(t)| \tag{2.27}$$

where the $p_i(t)$'s specify the classical probability of the system being in state i at time t, and the $|\psi_i\rangle$'s are wavefunctions given by

$$|\psi_i(t)\rangle = \sum_n c_n^i(t) \, |n\rangle \tag{2.28}$$

for some $\{c_n^i(t)\}$, where

$$H_0 \, |n\rangle = E_n \, |n\rangle \tag{2.29}$$

for all n. In the presence of damping, the elements

$$\rho_{nm} = \langle\psi_n| \, \rho \, |\psi_m\rangle \tag{2.30}$$

of ρ satisfy the differential equation

$$\dot{\rho}_{nm} = \frac{1}{i\hbar}[H, \rho]_{nm} - \gamma_{nm}(\rho_{nm} - \rho_{nm}^{(\text{eq})}), \tag{2.31}$$

where γ_{nm} is a matrix of (phenomenological) damping parameters[7], and $\rho_{nm}^{(\text{eq})}$ is the density matrix corresponding to the equilibrium steady state of the system.

---

[7]This is just one choice of $p_i(t)$.

We consider the case where V may be treated as a perturbation on top of H_0, i.e. where λ is small. In this case, Eq. 2.31 can be written

$$\dot{\rho}_{nm} = -i\omega_{nm}\rho_{nm} + \frac{1}{i\hbar}\sum_k \lambda(V_{nk}\rho_{km} - \rho_{nk}V_{km}) - \gamma_{nm}(\rho_{nm} - \rho_{nm}^{(\text{eq})}), \quad (2.32)$$

where $\omega_{nm} = E_{nm}/\hbar$, and we seek a solution

$$\rho_{nm} = \rho_{nm}^{(0)} + \lambda\rho_{nm}^{(1)} + \lambda^2\rho_{nm}^{(2)} + \cdots. \quad (2.33)$$

Turning the crank (see Boyd [5] for details) gives us the solution

$$\rho_{nm}^{(N)}(t) = \int_{-\infty}^t \frac{1}{i\hbar}[\lambda V(t'), \rho^{(N-1)}(t')]_{nm} e^{(i\omega_{nm}+\gamma_{nm})(t'-t)}\, dt'. \quad (2.34)$$

Carrying out this series to second order in λ with the perturbation $V(t) = -\boldsymbol{\mu}\cdot\boldsymbol{E}(t)$, where $\boldsymbol{\mu}$ is the dipole moment and $\boldsymbol{E}(t) = \sum_q \boldsymbol{E}(\omega_q)e^{-i\omega_q t}$ is the incident electric field, we get an expression for the density matrix $\rho_{nm}^{(2)}(t)$ as a function of the dipole matrix elements

$$\boldsymbol{\mu}_{nm} = \langle n|\,\boldsymbol{\mu}\,|m\rangle, \quad (2.35)$$

the frequencies ω_q, the damping constants γ_{nm}, and $\rho_{nm}^{(0)}$.

Once we have $\rho_{nm}^{(2)}(t)$, we can compute the expectation value

$$\langle\boldsymbol{\mu}(t)\rangle = \sum_{nm}\rho_{nm}(t)\boldsymbol{\mu}_{nm}, \quad (2.36)$$

from which the susceptibility can be computed by taking two derivatives with respect to the electric field amplitudes[8]. Reproducing the final answer

---

[8]We are specializing here to the case of electric dipole SHG, although the calculation proceeds similarly for magnetic dipole and electric quadrupole.

here (again, from Boyd [5]):

$$\chi_{ijk}^{(2)}(\omega_p+\omega_q,\omega_q,\omega_p) = \frac{1}{2\epsilon_0\hbar^2}\sum_{lmn}(\rho_{ll}^{(0)} - \rho_{mm}^{(0)}) \times \Bigg\{$$

$$\frac{\mu_{ln}^i\mu_{nm}^j\mu_{ml}^k}{[(\omega_{nl} - \omega_p - \omega_q) - i\gamma_{nl}][(\omega_{ml} - \omega_p) - i\gamma_{ml}]}$$

$$+\frac{\mu_{ln}^i\mu_{nm}^k\mu_{ml}^j}{[(\omega_{nl} - \omega_p - \omega_q) - i\gamma_{nl}][(\omega_{ml} - \omega_q) - i\gamma_{ml}]}$$

$$+\frac{\mu_{ln}^j\mu_{nm}^i\mu_{ml}^k}{[(\omega_{nm} + \omega_p + \omega_q) + i\gamma_{nm}][(\omega_{ml} - \omega_p) - i\gamma_{ml}]}$$

$$+\frac{\mu_{ln}^k\mu_{nm}^i\mu_{ml}^j}{[(\omega_{nm} + \omega_p + \omega_q) + i\gamma_{nm}][(\omega_{ml} - \omega_q) - i\gamma_{ml}]}$$

$$\Bigg\},$$

$$(2.37)$$

where $\omega_{nm} = \omega_n - \omega_m$. The SHG susceptibility tensor is then obtained by taking the limit $\omega_p = \omega_q$.

We learned two things by doing the quantum calculation. First of all, clearly if we know all of the eigenfunctions $|n\rangle$ of our unperturbed Hamiltonian, we can calculate the susceptibility tensor *a priori*, although this is obviously difficult except in the simplest of cases. Secondly, we notice that there are two types of denominators in Eq. 2.37: those occuring at 2ω (remember we have set $\omega_q = \omega_p$) and those occuring at ω. Both can cause resonances in the SHG intensity and are observed abundantly in experiment[12]. The existence of resonances in the SHG spectrum makes comparison between different materials quite difficult if reference is made only to the SHG intensity at a single color. In one imfamous example, Wu et al. [40] incorrectly attributed the large SHG amplitude at optical wavelengths in TaAs to the presence of Weyl nodes near the Fermi level; later SHG spectroscopy measurements demonstrated that the enhancement was due to a simple band resonance at the excitation frequency used in that paper[27]. The emerging consensus is that the SHG intensity at optical frequencies has more or less nothing to do with the low-energy excitation spectrum or its topology.

2.4   SHG in the Ginzburg-Landau paradigm

While Eq. 2.37 is quite general and completely correct microscopically, it obviously lends very little intuition into what kind of phenomenology we can expect to find in the SHG signal, say, across a phase transition, where the dipole matrix elements which determine Eq. 2.37 certainly change but may not do so in a simple or straightforward way. An alternative approach is to treat SHG in a generalized Ginzburg-Landau paradigm, in which all of the physics of the system is cast in terms of an order parameter $\mathcal{O}_{i_1 i_2 \ldots i_r}$. This approach was mainly developed by Sa et al. [33], based off of early work by Pershan [28] in the 1960s.

The key insight of Pershan is that one can define a "time-averaged free energy" in nonlinear optics by considering the term (assuming the incident field $E^{\text{in}}(t)$ is monochromatic with frequency ω)

$$F_{SHG} = -P_i(2\omega)E_i^{*\text{out}}(2\omega) + \text{c.c.} \tag{2.38}$$

$$= -[\chi_{ijk}E_i^{*\text{out}}(2\omega)E_j^{\text{in}}(\omega)E_k^{\text{in}}(\omega) + \text{c.c.}]. \tag{2.39}$$

Since F is a free energy, it must be a real, totally symmetric scalar, and Eq. 2.39 thus gives us a way to *derive* the form of the tensor χ_{ijk} in a particular point group using similar arguments as in section 2.1.

In the context of a (spontaneous) symmetry-breaking transition at T_c, the Ginzburg Landau paradigm asserts that the free energy of Eq. 2.39 still obeys the symmetry group of the high temperature phase, even though the low temperature phase explicitly breaks some of those symmetries. This is because the spontaneous symmetry breaking is considered a property of the *solution* of the free energy minimization problem, rather than a property of the minimization problem itself. Thus, a valid expression for the free energy in the low temperature phase is given by

$$F = -[\chi_{ijkl_1 l_2 \cdots l_r}(T > T_c)E_i^{*\text{out}}(2\omega)E_j^{\text{in}}(\omega)E_k^{\text{in}}(\omega) + \text{c.c.}]\mathcal{O}_{l_1 l_2 \cdots l_r}. \tag{2.40}$$

Comparing Eqs. 2.39 and 2.40, we have an equation for the SHG susceptibility tensor in the low temperature phase

$$\chi_{ijk}(T < T_c) = \chi_{ijkl_1 l_2 \cdots l_r}(T > T_c)\mathcal{O}_{l_1 l_2 \cdots l_r}. \tag{2.41}$$

In light of Eq. 2.41, let us consider the consequences of the symmetry of $\mathcal{O}$ on the SHG signal. For simplicity, let us consider a situation that

resembles a ferroelectric phase transition; i.e., the high temperature phase preserves inversion symmetry, but the low temperature phase involves the emergence of a rank-1 order parameter $\mathscr{O}_l$ which is odd under inversion. In that case, we have

$$\chi_{ijk}(T < T_c) = \chi_{ijkl}(T > T_c)\mathscr{O}_l(T), \tag{2.42}$$

i.e. the SHG at low temperature is simply a product of a high-temperature tensor and the order parameter $\mathscr{O}$. We of course need to check that χ_{ijkl} is allowed in the high temperature phase; here, we need χ_{ijkl} to be an even rank polar tensor, which is allowed in the presence of inversion symmetry (see Eq. 2.10).

Let us now imagine that the polarization $\mathscr{O}_l$ is known to be directed along the y axis, i.e.

$$\mathscr{O}_l = (0, P_0, 0)_l, \tag{2.43}$$

for some P_0. Then, by Eq. 2.42, we have

$$\chi_{ijk}(T < T_c) = \chi_{ijky}(T > T_c)P_0. \tag{2.44}$$

Besides telling us that the susceptibility tensor is linear in the polarization magnitude P_0, what this formulation gets us is also that the elements of χ_{ijk} are just the elements of χ_{ijkl} with $l = y$. In many cases this is more information than you would have if you only knew the point group of the low temperature phase!

Eq. 2.41 is also quite useful in the case where the low temperature phase is heterogenous; i.e. the order parameter $\mathscr{O}$ varies spatially from one point to another. This is commonly the case in magnets, for example, where rotational symmetry is spontaneously broken at T_c and the low temperature free energy thus consists of multiple energetically degenerate states related to each other by elements R of the high temperature point group which are broken at low temperature. These states have the same free energy since, again, the symmetry-breaking is due to to the order parameter, not due to the free energy itself. The order parameter in the different domains are thus related to each other via

$$\mathscr{O}_{i_1 i_2 \cdots i_r}(\boldsymbol{x}_1) = R[\mathscr{O}_{i_1 i_2 \cdots i_r}(\boldsymbol{x}_2)], \tag{2.45}$$

where the right hand side is given by Eq. 2.10. Combining Eq. 2.41 and Eq. 2.45, the SHG tensor thus satisfies

$$\chi_{ijk}(\boldsymbol{x}_1) = R[\chi_{ijk}(\boldsymbol{x}_2)]. \tag{2.46}$$

Eq. 2.46 is the fundamental basis for our SHG works on $1T$-TaS_2 and $CaMn_2Bi_2$, discussed in chapters **??** and **??**.

Chapter Three

Second harmonic generation: practical

3.1 Description of the setup

In the last chapter we saw that the SHG intensity in a given crystal is related to the point group, order parameter, band structure, etc. of that crystal via the susceptibility tensor χ_{ijk}. It is also obvious from section 2.1 that the ideal scenario is to be able to measure as many of the numbers χ_{ijk} as possible; after all, if you only measured the xx component of Eqs. 2.11 and 2.12, you would have obtained essentially no information about your crystal whatsoever. The SHG setup that we built (whose design is mostly credited to Torchinsky and Hsieh[36], with some improvements by us which I will discuss below) was designed with exactly this goal in mind. There are two key insights which make this design work: one, the light is obliquely incident on the sample so there is some component of $\boldsymbol{E}^{\mathrm{in}}$ directed along the sample normal, and two, we rotate the plane of incidence so that the in-plane field direction sweeps an entire $360°$. The first point allows us to measure elements of χ_{ijk} with z indices[1], and the second point makes sure we get all of the x and y elements of χ_{ijk} too. All of the tensor elements are thus given a chance to contribute to the SHG intensity in a given experiment.

With those considerations in mind, let me proceed to give a schematic description of our SHG setup (see Figs. 3.1 and 3.2). Some of the choices

---

[1]Here and unless otherwise noted I define the sample normal to be the z axis.

51

P:Polarizer
HWP: Half waveplate
QWP: Quarter waveplate
DM: Dichroic Mirror
L: Lens
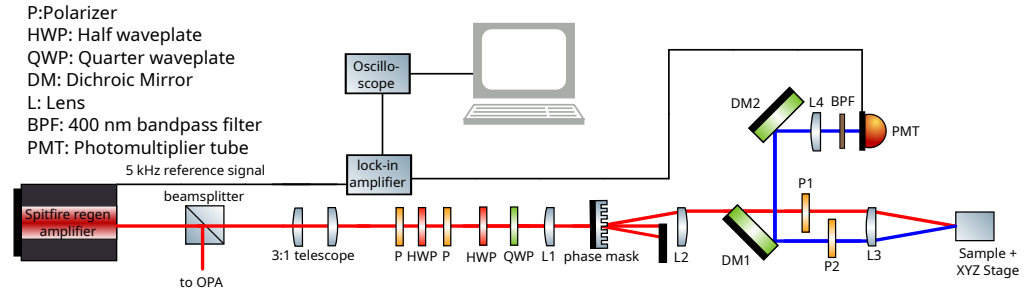BPF: 400 nm bandpass filter
PMT: Photomultiplier tube

Figure 3.1: Schematic drawing of the SHG setup used in this research. After Morey et al. [23].

we made may seem arbitrary right now, but I will go over them in detail in section 3.2. The starting point is our regenerative amplifier (Spectra-Physics Spitfire Sptf-100f-5k-xp), which produces 100 fs 800 nm pulses at a 5 kHz repetition rate from an 86 MHz seed laser (Spectra-Physics Tsunami 3941-M1S). 90% of the beam is split off to power an optical parametric amplifier (OPA), and the remaining 10% is used for the SHG probe beam. After passing through an optical telescope, which creates a collimated beam of width $1 - 2$ mm, this beam is attenuated with a polarizer - half-wave plate - polarizer triplet, and then elliptically polarized with a quarter-wave plate and half-wave plate in series. The ellipticity at this stage is set so the light is perfectly circularly polarized following transmission through a phase mask, as described below.

After passing through these polarization optics, the beam is focused with a lens onto the aforementioned phase mask, which acts as a transmissive diffraction grating and separates the beam into multiple different diffraction orders. The $+1$ order diffraction comes off at an angle of roughly $7°$, while the other orders are blocked with anodized aluminum foil. This (diverging) beam then propogates at $7°$ to the optical axis before meeting a lens set at the appropriate distance so as to both collimate the beam and rectify the $7°$ progation angle. Then, the light passes through a dichroic mirror (which transmits 800 nm and reflects 400 nm), becomes linearly polarized by a wire-grid polarizer, and is then focused onto the sample at a $10°$ angle of incidence by passing through the edge of a 1 in-diameter 50 mm achromatic focusing lens. The interaction between the light and the sample causes SHG to be radiated in reflection at the same
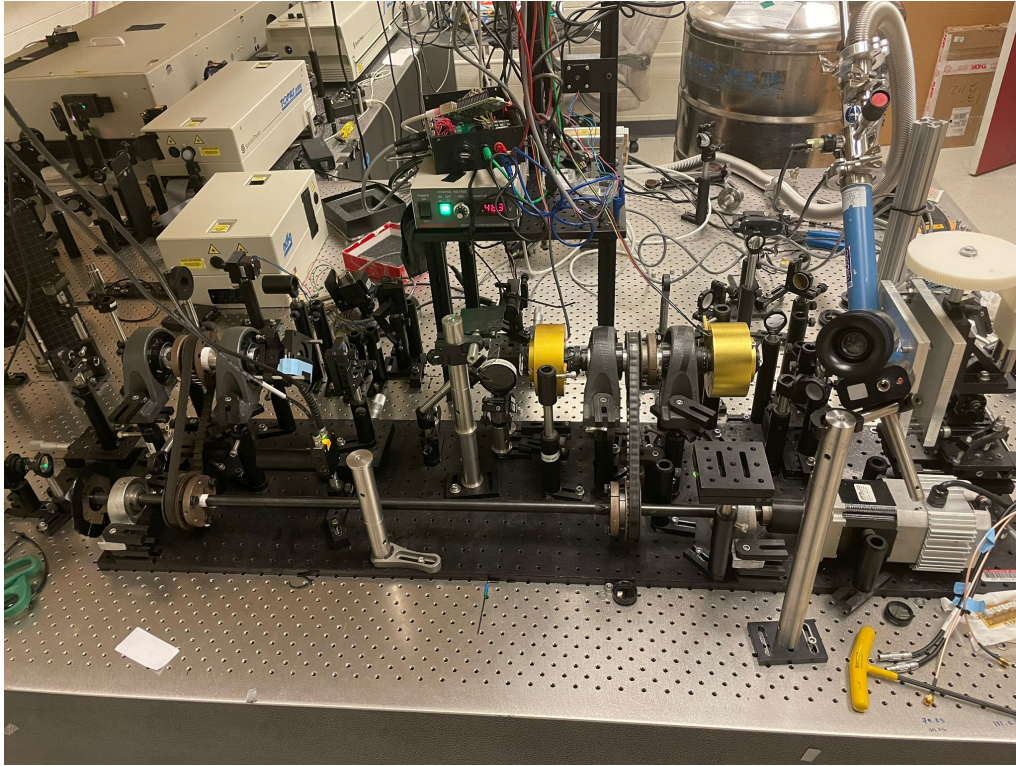
Figure 3.2: Photograph of the SHG setup used in this research. The sample lies in the blue cryostat on the right side of the image. The phase mask is mounted in a black holder with blue masking tape on the left side, and the polarizers are mounted in the two gold-colored slip rings (described in **??**).

$10°$ angle of incidence, so that the SHG beam passes through the opposite side of the 50 mm lens before passing through a second, independent polarizer which is used to control the polarization of the measured light.

Finally, the polarized output reflects off of two dichroic mirrors (oriented in such a way as to cancel the differing effect of the Fresnel equations on the reflectivity of S and P polarized light) and is focused through a 400 nm bandpass filter onto a photomultiplier tube (PMT) by a 400 mm lens. The current output of the PMT is filtered by a lock-in amplifier (for static SHG, set to the 5 kHz repetition rate of the laser) and read out on an oscilloscope. The phase mask, incoming polarizer, and outgoing polarizer

are mounted on rotating lens tubes which are connected via pulley to a common motor shaft driven at ~ 5 Hz by a brushless DC motor. The motor thus continuously rotates the plane of incidence of the experiment, since the latter is entirely defined by the phase mask and the polarizers. The rotation angle is tracked as a function of time by an optical rotary encoder, consisting of a laser pointer passed through a chopper wheel (with 100 slots) mounted at the end of the motor shaft and detected via photodiode. The encoder signal and the lock-in signal are both sent to a homemade oscilloscope (an Arduino Uno microcontroller which separates the lock-in signal into different individual rotations by looking for peaks in the encoder signal), the output of which is sent to a computer for further data processing.

3.2 Before you build

I list here a few essential aspects of SHG that should be considered before designing a new setup.

3.2.1 Spot size

One of the most important quantities in an SHG setup is the diameter of the probe spot. Ideally, this diameter is as small as possible, so as to measure the smallest samples or domain sizes. However, there is an important caveat: for constant fluence (i.e. supposing we are limited by the sample damage threshold), the SHG signal to noise ratio scales linearly in the area excited by the probe, and thus SHG microscopes have a difficult time measuring small SHG signals compared to traditional SHG setups with a larger excitation area. To see this, let us say that our detector measures the number of photons per SHG pulse, which is proportional to the pulse energy $U_p(2\omega)$. Assuming the input and output pulse intensity profiles have the shape of a square wave with width τ and height $I_p(\omega)$ and $I_p(2\omega)$, respectively, we have

$$U_p(2\omega) = AI_p(2\omega)\tau \tag{3.1}$$

and

$$U_p(\omega) = AI_p(\omega)\tau \tag{3.2}$$

where A is the area of the beam at the sample surface. The SHG intensity is proportional to the square of the input intensity

$$I_p(2\omega) \propto I_p^2(\omega) \tag{3.3}$$

so that

$$U_p(2\omega) \propto \frac{U_p^2(\omega)}{A\tau}. \tag{3.4}$$

Substituting for the fluence f

$$f(\omega) = \frac{U_p(\omega)}{A} \tag{3.5}$$

we have

$$U_p(2\omega) \propto \frac{f^2(\omega)A}{\tau} \tag{3.6}$$

i.e., if we hold the fluence constant at the sample damage threshold, the number of photons in the generated SHG pulse is proportional to the excitation area and inverse to the pulse width. The signal to noise ratio is then given by

$$\text{SNR} \propto \frac{U_p(2\omega)}{\sqrt{r}} \tag{3.7}$$

where r is the system repetition rate.

3.2.2 Oblique vs. normal incidence

While all of the results presented in this thesis utilized the setup in Fig. 3.1, where the incident beam makes a small angle with respect to the sample normal, plenty of groups use a different approach where that angle is set to $0°$. This has the obvious disadvantage of not specifying all of the tensor elements, since any element χ_{ijk} with $i, j, k = z$ is not accessible in this geometry. However, in some cases this can actually be something of an advantage. For example, sometimes unwanted SHG contributions (see section 2.2) may be avoided in the normal incidence geometry, assuming the numerical aperture (NA) of the focusing optic is small enough that longitudinal components of the electric field are nearly zero. Furthermore, in some materials the order parameter only couples to one or two elements of χ_{ijk}; if none of these elements have a z index, it is needless to complicate the analysis with oblique incidence.

In my experience, oblique incidence seems to be useful in two broad cases. For one thing, some order parameters only show up in the z components of χ_{ijk} (this is the case in $1T$-TaS$_2$, see **??**), in which case one obviously needs a nonzero angle of incidence to access these components. A more subtle point is that, even if in practice all of the phenomenology of a particular sample only shows up in the x and y indices of χ_{ijk}, still one must measure the full tensor to *rule out* unseen phenomenology in the other indices. Both the CaMn$_2$Bi$_2$ (**??**) and CuBr$_2$ (**??**) works presented in this thesis are examples of exactly this point, where the main scientific arguments involve either comparing SHG patterns in two domains or comparing oscillation amplitudes in different polarization channels. Clearly one needs to know all of the tensor elements to make those arguments exact.

3.2.3   Choice of detector

Our setup is somewhat unique in using a PMT for data collection rather than an electron multiplying charge coupled device (EM-CCD), which is the slightly more traditional method. Frankly, this decision was not based on the detection efficiency, but rather the fact that PMTs typically cost about two orders of magnitude less than EM-CCDs. However, the setup construction with a PMT is slightly different than with an EM-CCD, so you should probably decide which you want to use before you start building your setup. With an EM-CCD, the beam is sent into the device without focusing (L4 in Fig. 3.1) so that the beam traces a circle on the sensor as a function of time[13]. The rotational anisotropy signal is read off by performing a radial integration of the camera image (after masking the part of the image which is outside from the circle traced by the beam). In this way, the rotation angle of the motor is correlated with the SHG signal via the azimuthal degree of freedom on the camera image. In contrast, with a PMT the SHG signal is read out as a function of time on an oscilloscope (see section 3.1), and must be correlated with the rotation angle of the motor by some other method. We use an optical chopper wheel attached to the motor shaft; a beam from a laser pointer is directed through the chopper wheel and onto a photodiode, which produces a square wave signal that is used to trigger the oscilloscope. One must also be careful that the PMT is aligned as close to normal as possible to the axis

defined by DM2 and L4 in Fig. 3.1, as the PMT output is actually quite sensitive to the angle of incidence of the input radiation.

As for the detection efficiency of the two devies, an EM-CCD is basically an *array* of PMTs—thus, there is no fundamental difference in the detection efficiences of the two methods, although having never used an EM-CCD I cannot speak to any specific considerations that might favor one over the other.

3.3   tr-SHG: methodology

3.3.0.1   Polarization control

The main limitation for doing tr-SHG is simply that the experiment becomes longer. Each measurement of χ_{ijk} requires averaging four different polarization channels ($P_\text{in}P_\text{out}$, $P_\text{in}S_\text{out}$, $S_\text{in}P_\text{out}$, and $S_\text{in}S_\text{out}$), which, depending on the signal to noise ratio, can take as long as $2-3$ minutes each. In a time-resolved experiment, that procedure must be repeated at least once for each time delay; in fact, it us often useful to sweep the delay stage multiple times to reduce the extent to which systematic drifts in in the laser power, alignment, etc. affect the time trace. Not only is this time-consuming (typically taking $\sim 12\,\text{hours}$ to get a good dataset), but it also requires four rotations of the polarizers at each delay, which is simply not feasible if the polarizers are to be rotated manually. Unfortunately, automated polarizer rotation is difficult in the rotational anisotropy SHG (RA-SHG) experimental geometry because the polarizers must be rotated relative to a lens tube which is *itself* rotating at $5\,\text{Hz}$. The traditional method of rotating polarizers via stepper motor is thus not possible, unless one finds a method to transmit power between the stationary laboratory frame and the rotating frame that the polarizer lives in. Myself and Karna Morey designed such a method using an electronic device known as a hollow-bore electric slip ring, which uses ring-shaped conductive pads in combination with low-friction metallic brushes to conduct electricity between two rotating objects. In my opinion, such a design is essential to doing tr-SHG unless one is satisfied with only measuring a single polarization channel; thus, I have dedicated the entirety of **??** to our solution, which in my opinion represents the largest contribution we made to SHG methodology during this thesis.

3.3.0.2   The pump beam path

Having discussed the modifications needed for tr-SHG on the probe arm, let us now describe the pump beam path. 100 fs pulses from the regenerative amplifier (Spectra-Physics Spitfire Sptf-100f-5k-xp) are split by $90 : 10$ beamsplitter (with the 10% becoming the probe beam, see section 3.1) and are used as input to an OPA (Light Conversion TP8F1N3) which produces light of variable wavelength between 1100 and 2080 nm. The output of the OPA is directed through a wire-grid polarizer (to pick out the signal or idler, as needed) to an optical delay line (Newport DL125 with SMC100 motion controller) and a second polarizer to vary the beam power. The beam then passes through a NIR (near-infrared) longpass filter to remove unwanted visible wavelengths that are output by the OPA. A 400 nm lens focuses the beam past two mirrors (one mounted with epoxy onto the focusing lens L3 of Fig. 3.1[2]) and finally onto the sample. The diameter of the pump beam at the focus is set by moving the lens position along the beampath and monitoring the beam shape with a CCD camera. The pointing of the beam is aligned so that the light reflected off the sample is perfectly backscattered. Pump scatter is avoided by mounting a small black circular disk in the center of L3, and by subtracting "dark" scans without the probe beam from the final dataset.

An optical chopper wheel is placed somewhere in the pump beam path which blocks every other pulse from the OPA, so that the effective repetition rate of the pump pulsetrain is 2.5 kHz. The output of the PMT (see Fig. 3.1) is sent to a lock-in amplifier synced to this frequency, so that the output of the lock-in is proportional to

$$I^{\text{SHG}}_{\text{Pump+Probe}} - I^{\text{SHG}}_{\text{Probe}} \tag{3.8}$$

i.e. the output of the lock-in is the pump-induced change in the measured SHG intensity.

Finally, the pump-probe time delay is varied, and for each delay the rotational anisotropy in the SHG intensity is measured in each of the four polarization channels. Alternatively, the motor rotation may be parked at a fixed angle and the SHG intensity at this angle measured as a function of delay time.

---

[2]Remember that the probe beam is arriving at oblique incidence and thus passes through the edge, rather than the center, of L3.

3.4 Data analysis

Finally, we discuss the analysis of SHG data, first in the static limit and then in the context of tr-SHG.

3.4.1 Static RA-SHG patterns

The typical static RA-SHG dataset consists of a set of points $\{(\phi_n, I_n^p)\}$, $n \in \{0, 1, \ldots N-1\}$ and standard errors $\{\sigma_n^p\}$ for each of the four independent polarization channels

$$p \in \{P_{\text{in}} P_{\text{out}}, P_{\text{in}} S_{\text{out}}, S_{\text{in}} P_{\text{out}}, S_{\text{in}} S_{\text{out}}\}.$$

By Eq. 2.13, each of these patterns may be modeled by an equation[3]

$$I^p(\phi) \propto |\hat{e}_i^{p,\text{out}}(\phi) \chi_{ijk} \hat{e}_j^{p,\text{in}}(\phi) \hat{e}_k^{p,\text{in}}(\phi)|^2, \tag{3.9}$$

where $\hat{e}$ is a unit vector in the direction of the corresponding electric field. These unit vectors are given as a function of ϕ and the angle of incidence θ in table 3.1.

---

[3]I will focus on electric dipole SHG for now, although note that the equations for electric quadrupole and magnetic dipole SHG (Eqs. 2.23 and 2.24) are different—by Eq. 2.16, both involve a factor of the incident wavevector k, and the magnetic dipole term has a cross product.

Table 3.1: Vector definition of polarization channels. θ is the angle of incidence and ϕ is the angle of the plane of incidence with respect to the $\hat{x}$ axis.

Input	Output	$\hat{e}^{\text{in}}(\phi)$			$\hat{e}^{\text{out}}(\phi)$		
		$\hat{e}^{\text{in}}_x(\phi)$	$\hat{e}^{\text{in}}_y(\phi)$	$\hat{e}^{\text{in}}_z(\phi)$	$\hat{e}^{\text{out}}_x(\phi)$	$\hat{e}^{\text{out}}_y(\phi)$	$\hat{e}^{\text{out}}_z(\phi)$
P	P	$\cos\phi\cos\theta$	$\sin\phi\cos\theta$	$-\sin\theta$	$-\cos\phi\cos\theta$	$-\sin\phi\cos\theta$	$-\sin\theta$
P	S	$\cos\phi\cos\theta$	$\sin\phi\cos\theta$	$-\sin\theta$	$-\sin\phi$	$\cos\phi$	0
S	P	$-\sin\phi$	$\cos\phi$	0	$-\cos\phi\cos\theta$	$-\sin\phi\cos\theta$	$-\sin\theta$
S	S	$-\sin\phi$	$\cos\phi$	0	$-\sin\phi$	$\cos\phi$	0

Let us assume that we know the point group P_G of our material. Then, the first step is to constrain the values of χ_{ijk} using the arguments of section 2.1. One can do this manually by solving the system of equations defined by Eq. 2.10, or one can simply look up the answer in one of a number of tables (e.g. Boyd [5])[4]; in the end we are left with a tensor $\chi_{ijk}^{P_G}$ with M independent elements $\{\chi_i, i \in 0, 1, \ldots, M-1\}$.

The goal is to simultaneously fit the four patterns $\{(\phi_n, I_n^p)\}$ to Eq. 3.9 using the χ_i's. Our objective function is thus

$$f(\{\chi_i\}) = \sum_{\substack{p \in \{P_{\text{in}}P_{\text{out}}, P_{\text{in}}S_{\text{out}}, \\ S_{\text{in}}P_{\text{out}}, S_{\text{in}}S_{\text{out}}\}}} \sum_{n=0}^{N-1} \left(\frac{I^p(\phi_n) - I_n^p}{\sigma_n^p} \right)^2 \qquad (3.10)$$

with $I^p(\phi)$ defined in Eq. 3.9. Unfortunately, such a function is quite difficult to minimize. The reason is that $I^p(\phi)$ is ultimately a sum of products of trigonometric functions of ϕ (due to table 3.1) which can be difficult for typical minimization algorithms. An alternative is to cast the problem in terms of the Fourier transforms (FTs)

$$\hat{I}^p(k) \equiv \frac{1}{2\pi} \int_0^{2\pi} I^p(\phi)e^{-ik\phi} \, d\phi \qquad (3.11)$$

of $I^p(\phi)$ and the discrete Fourier transforms (DFTs)

$$\{(k, \hat{I}_k^p)\}, \qquad (3.12)$$

of $\{(\phi_n, I_n^p)\}$, where[5]

$$\hat{I}_k^p \equiv \frac{1}{N} \sum_{n=0}^{N-1} I_n^p e^{-ik\phi_n}. \qquad (3.13)$$

Since the DFT is unitary, the uncertainties $\hat{\sigma}_k^p$ are simply the DFT of the σ_n^p's. Thus, our new objective function is

$$f(\{\chi_i\}) = \sum_{\substack{p \in \{P_{\text{in}}P_{\text{out}}, P_{\text{in}}S_{\text{out}}, \\ S_{\text{in}}P_{\text{out}}, S_{\text{in}}S_{\text{out}}\}}} \sum_{k=0}^{N-1} \left(\frac{\hat{I}^p(k) - \hat{I}_k^p}{\hat{\sigma}_k^p} \right)^2. \qquad (3.14)$$

---

[4]Be aware that these tensors are given in just one of many choices of coordinates; a rotation of χ_{ijk} may be needed to agree with the orientation of the crystal in the experiment.

[5]I am assuming the ϕ_n's are equally spaced.

This is a much easier function to minimize for two reasons: one, the function is now just a quadratic function of the $\{\chi_i\}$'s, and two, since there are only six factors of trigonometric functions of ϕ in $I^p(\phi)$, $\hat{I}^p(k) = 0$ for $k \notin \{-6, -5, \ldots, 6\}$[6] (the other terms in Eq. 3.14 just contribute an overall constant to f).

What is left then is just the choice of minimization algorithm. Unfortunately, simple gradient descent algorithms are not ideal here unless you have a really good initial guess, since Eq. 3.11 is still quadratic in the fitting parameters and the global minimum is often surrounded by many local minima which can easily cause these algorithms to become trapped. Global minimization routines are thus optimal; I have used the simulated annealing algorithm of Xiang et al. [41] with success, as well as the related "basinhopping" algorithm of Wales and Doye [39], both of which are implemented in `scipy.optimize`[37]. Ultimately, though, it is up to the experimenter to ensure they have found the true global minimum of Eq. 3.14.

As part of this thesis, I wrote a software package in Python called ShgPy which uses `sympy`[22] to compute the FT model Eq. 3.11 and `scipy.optimize`[37] to fit this model to the data via Eq. 3.14. Installation instructions and a comprehensive documentation may be found at the project home page `https://bfichera.github.io/shgpy/`.

3.4.2 Modeling χ_{ijk}: the simplified bond hyperpolarizability model

Sometimes the symmetry-constrained χ_{ijk} of section 3.4.1 is difficult to analyze (perhaps it has too many fitting parameters), in which case it might be prudent try and justify an even more tightly constrained χ_{ijk} via some a microscopic model. We have already seen two such models via Eqs. 2.37 and 2.41, but the former is typically not tractable and the latter may only be valid in the presence of a spontaneous symmetry-breaking phase transition with a well-defined order parameter. One alternative is to model the microscopic charge degrees of freedom classically—suppose most of the motion of charges in our solid occurs in the bonds between atoms, and that that motion is constrained to occur along the bond direction only.

---

[6]This is the case for electric dipole SHG; for electric quadrupole and magnetic dipole, we can have as much as $k = \pm 8$.

Then, the solid can be modeled as an ensemble of oscillating dipoles (one for each bond in the unit cell) with corresponding hyperpolarizabilities $\{\alpha_n\}$. Symmetry may dictate how many independent α_n's there are in the unit cell, but otherwise these are considered unknown and to be fit to the data. The electric dipole SHG response of such an ensemble was analyzed by Powell et al. [29], and the electric quadrupole response was later considered by Bauer and Hingerl [3]; the answer in both cases is quite simple:

$$\chi_{ijk}^{eee} = \sum_n \alpha_n b_i^n b_j^n b_k^n \tag{3.15}$$

$$\chi_{ijkl}^{qee} = \sum_n \gamma_n b_i^n b_j^n b_k^n b_l^n \tag{3.16}$$

where $\boldsymbol{b}^n$ is a unit vector in the direction of bond n and the α_n's and γ_n's are the unknown parameters of the dipole and quadrupole models, respectively.

When valid, this simplified bond hyperpolarizability model (SBHM) is powerful for two reasons: for one, Eqs. 3.15 and 3.16 typically have fewer unknown parameters than the naive, symmetry-constrained tensor $\chi_{ijk}^{P_G}$ of section 3.4.1, and two, there is a direct connection between the unknown parameters α_n and γ_n and the microscopic degrees of freedom. One can thus draw conclusions *about* the microscopic degrees of freedom by fitting the SHG patterns to Eqs. 3.15 and 3.16 (see, e.g., Ron et al. [32]). Of course, one must be careful that the assumptions of this model—that the charges are localized to the bonds, and that the only relevant motion is along the bond direction—are actually correct before proceeding with this sort of analysis.

Myself and Karna Morey built a comprehensive interface to the SBHM in Python using `pymatgen`[26], which takes an arbitrary crystallographic information file (CIF) and calculates the relevant susceptibility tensors via Eqs. 3.15 and 3.16. The change to the susceptibility tensor induced by an elastic distortion mode ($\chi_{ijk}(q)$, where q is the mode amplitude) can be computed via integration with `ISODISTORT`[6, 34]. See the package documentation (`https://github.com/morey18k/sbhm.git`) for examples and further information.

3.4.3   Fitting time traces in tr-SHG

As the final section of this chapter, let me remark on a few aspects of fitting time-domain signals that are not necessarily specific to SHG, but are nevertheless essential for tr-SHG analysis.

Typically, in tr-SHG we plot the pump-induced change in the SHG intensity at a single angle ϕ_0 as a function of the delay time t; more complicated analysis of the tensor elements themselves is possible but often unnecessary. This intensity change $\Delta I_{2\omega}(t, \phi_0)$ is related to $P(2\omega)$ via

$$\Delta I_{2\omega}(t, \phi_0) = I_{2\omega}^{\text{Pump+Probe}}(t, \phi_0) - I_{2\omega}^{\text{probe}}(0, \phi_0) \tag{3.17}$$

$$\propto |P_{2\omega}(0, \phi_0) + \delta P_{2\omega}(t, \phi_0)|^2 - |P_{2\omega}(0, \phi_0)|^2 \tag{3.18}$$

$$\propto 2P_{2\omega}(0, \phi_0)\delta P_{2\omega}(t, \phi_0) + |\delta P_{2\omega}(t, \phi_0)|^2 \tag{3.19}$$

where we have used that the signal is presumably independent of time if there is no pump pulse. When $|\delta P_{2\omega}(t, \phi_0)| \ll |P_{2\omega}(0, \phi_0)|$, this reduces to

$$\Delta I_{2\omega}(t, \phi_0) \propto P_{2\omega}(0, \phi_0)\delta P_{2\omega}(t, \phi_0). \tag{3.20}$$

The basic problem is then that we have a dataset

$$\{t_n, I_n\}, n \in \{0, 1, \ldots, N-1\} \tag{3.21}$$

with uncertainties σ_n which we wish to fit to some model, typically consisting of some (e.g. polynomial) background

$$b(t, \{b_k\}), k \in \{0, 1, \ldots, K-1\} \tag{3.22}$$

plus 1 to M damped harmonic oscillators

$$I(t, \theta) = \begin{cases} 0 & t < 0 \\ b(t, \{b_k\}) + \sum_{m=1}^M A_m e^{-t/\tau_m} \cos(2\pi f_m t + \psi_m) & t > 0 \end{cases}, \tag{3.23}$$

where θ refers to the set of unknown parameters $\{b_k\}$, $\{A_m\}$, $\{\tau_m\}$, $\{f_m\}$, and $\{\psi_m\}$. We wish to determine the model coefficients θ by minimizing the objective function

$$f(\theta) = \sum_{n=0}^{N-1} \left(\frac{I(t_n, \theta) - I_n}{\sigma_n} \right)^2. \tag{3.24}$$

Various algorithms exist to find the minimum θ_0 of f; the Levenberg-Marquardt (LM) algorithm [20, 21] is known to have trouble with exponential functions, so it is recommended to do a Nelder-Mead search [24] first and then use LM to refine the solution. These methods also report the uncertainties $\{\sigma_i\}$ on each fitting parameter, basically by computing the curvature of the function f near θ_0 (which depends, among other things, on the data uncertainty $\{\sigma_n\}$). Importantly, however, these uncertainties are only valid when f is a linear function of the model parameters near the minimum![7] This is a point which is completely overlooked in the ultrafast condensed matter literature, even though there is really no reason that Eq. 3.24 should satisfy this criterion for typical uncertainty levels in pump-probe experiments.

The correct approach in this case is to note that the data set $D_0 \equiv \{t_n, I_n\}$ is just one of many possible datasets $D_0, D_1, D_2, \ldots$ that we *could* have measured during our experiment.[8] If we could run the minimization procedure described above for each of these hypothetical datasets, we would end up with a whole distribution of model coefficients $\theta_0, \theta_1, \theta_2, \ldots$ from which our measurement θ_0 was just a single draw. The spread of this distribution (or, more precisely, the spread of the distribution $\theta - \theta_{\text{true}}$) reflects exactly the quantitative uncertainty in our measurement of θ that we would like to report.

Our goal, then, is to approximate this distribution by some method. One approach (which works when we have many observations $\{t, I\}$ with the same t) is called the nonparametric bootstrap[8]. For each time t_n, let us say that we have a set of measured intensities $\{I_m^n\}$, with $m \in \{1, 2, \ldots, M_n\}$. Then, we generate new hypothetical datasets $\mathscr{D}_1, \mathscr{D}_2, \ldots$ by (for each t_n) drawing M_n items *with replacement* from the $\{I_m^n\}$'s. The fact that this new distribution of datasets is a faithful representation of the true distribution $\{D_0, D_1, D_2, \ldots\}$ is not obvious, but theoretical work by many authors (see Wackernagel [38] for a review) has established that it is nevertheless true.

If we are confident that our model $I(t, \theta)$ is, in fact, the true model (at least for some set of coefficients θ_{true}), then rather than using the dataset D_0 to generate our hypothetical datasets, we can instead choose to use the model *itself*. The idea is to take our estimate $I(t, \theta_0)$ and generate new

---

[7]Here, "near" means "in a neighborhood of size σ_i.

[8]The following discussion closely follows that of *Numerical Recipes*, chapter 15[31].

datasets $\{\mathscr{D}_m\}$ essentially by adding noise:

$$\mathscr{D}_m = \{(t_n, I(t_n, \theta_0) + N_n), n \in \{0, 1, \ldots, N-1\}\}, \qquad (3.25)$$

for some random variable N_n which we put in by hand. This is called the parametric bootstrap, referring to the fact that we are estimating our true distribution $\{D_0, D_1, D_2, \ldots\}$ via the parameters θ_0[10]. In order for the distribution of $\mathscr{D}_m$'s to faithfully represent the true distribution $\{D_0, D_1, D_2, \ldots\}$, obviously N_n needs to reflect the actual noise distribution of the experiment. Thankfully, in pump-probe experiments, we usually know this distribution—as long as we have enough points (t_n, I_n) with $t_n < 0$, we can estimate N_n with a normal distribution centered about 0 with variance $\langle (I_n)^2 \rangle_{<0} - \langle I_n \rangle_{<0}^2$, where $\langle \cdot \rangle_{<0}$ means "average over points with $t_n < 0$."

Once we have the hypothetical datasets $\{\mathscr{D}_m\}$ (regardless of how we generated them), we proceed by estimating a θ_m for each one using Eq. 3.24, and then plotting the distribution $\{\theta_m - \theta_0\}$; the spread of this distribution tells us the uncertainties σ_i in each of our model coefficients. The important part is that these uncertainties are much more reflective of the true uncertainty in our measurement of θ than the curvature of Eq. 3.24 near θ_0, which we have argued is strictly incorrect in many cases.

*

Bibliography

[1] D. Afanasiev, J. R. Hortensius, B. A. Ivanov, A. Sasani, E. Bousquet,
    Y. M. Blanter, R. V. Mikhaylovskiy, A. V. Kimel, and A. D. Caviglia.
    Ultrafast control of magnetic interactions via light-driven phonons.
    Nat. Mater., 20(5):607–611, May 2021.
    doi: 10.1038/s41563-021-00922-7.

[2] A. Alexandradinata, N. P. Armitage, Andrey Baydin, Wenli Bi, Yue
    Cao, Hitesh J. Changlani, Eli Chertkov, Eduardo H. da Silva Neto,
    Luca Delacretaz, Ismail El Baggari, G. M. Ferguson, William J.
    Gannon, Sayed Ali Akbar Ghorashi, Berit H. Goodge, Olga
    Goulko, G. Grissonnanche, Alannah Hallas, Ian M. Hayes,
    Yu He, Edwin W. Huang, Anshul Kogar, Divine Kumah,
    Jong Yeon Lee, A. Legros, Fahad Mahmood, Yulia Maximenko,
    Nick Pellatz, Hryhoriy Polshyn, Tarapada Sarkar, Allen Scheie,
    Kyle L. Seyler, Zhenzhong Shi, Brian Skinner, Lucia Steinke,
    K. Thirunavukkuarasu, Thaís Victa Trevisan, Michael Vogl,
    Pavel A. Volkov, Yao Wang, Yishu Wang, Di Wei, Kaya Wei, Shuo-
    long Yang, Xian Zhang, Ya-Hui Zhang, Liuyan Zhao, and Alfred
    Zong.
    The Future of the Correlated Electron Problem, July 2022.
    arXiv:2010.00584 [cond-mat].

[3] Klaus-Dieter Bauer and Kurt Hingerl.
    Bulk quadrupole contribution to second harmonic generation from
    classical oscillator model in silicon.
    Opt. Express, 25(22):26567, October 2017.
    doi: 10.1364/OE.25.026567.

[4] Robert R. Birss.

Symmetry and Magnetism.
North-Holland Pub. Co., 1964.

[5] Robert Boyd.
     Nonlinear Optics.
     Academic Press, third edition, mar 2008.

[6] Branton J. Campbell, Harold T. Stokes, David E. Tanner, and Do-
     rian M. Hatch.
     ISODISPLACE : a web-based tool for exploring structural distortions.
     J Appl Crystallogr, 39(4):607–614, August 2006.
     doi: 10.1107/S0021889806014075.

[7] Wikipedia contributors.
     List of space groups — Wikipedia, the free encyclopedia, 2024.
     [Online; accessed 06-March-2024].

[8] N.A.C. Cressie.
     Statistics for Spatial Data.
     A Wiley-interscience publication. J. Wiley, 1991.

[9] Pierre Curie.
     Sur la symétrie dans les phénomènes physiques, symétrie d'un
     champ électrique et d'un champ magnétiqueu.
     J. Phys. Theor. Appl., 3(1):393–415, 1894.

[10] F.M. Dekking, C. Kraaikamp, H.P. Lopuhaä, and L.E. Meester.
     *A Modern Introduction to Probability and Statistics: Understanding Why
     and How*.
     Springer Texts in Statistics. Springer London, 2006.

[11] 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(6014):189–191, January 2011.
     doi: 10.1126/science.1197294.

[12] M. Fiebig, D. Fröhlich, Th. Lottermoser, V. V. Pavlov, R. V. Pisarev,
     and H.-J Weber.
     Second Harmonic Generation in the Centrosymmetric Antiferromag-
     net NiO.

Phys. Rev. Lett., 87(13):137202, September 2001.
doi: 10.1103/PhysRevLett.87.137202.

[13] J. W. Harter, L. Niu, A. J. Woss, and D. Hsieh.
High-speed measurement of rotational anisotropy nonlinear optical
harmonic generation using position-sensitive detection.
Opt. Lett., 40(20):4671, October 2015.
doi: 10.1364/OL.40.004671.

[14] John David Jackson.
Classical electrodynamics.
Wiley, New York, NY, 3rd ed. edition, 1999.

[15] T.W.B Kibble and G.R Pickett.
Introduction. Cosmology meets condensed matter.
Phil. Trans. R. Soc. A., 366(1877):2793–2802, August 2008.
doi: 10.1098/rsta.2008.0098.

[16] A. V. Kimel, A. Kirilyuk, A. Tsvetkov, R. V. Pisarev, and Th. Rasing.
Laser-induced ultrafast spin reorientation in the antiferromagnet
$TmFeO_3$.
Nature, 429(6994):850–853, June 2004.
doi: 10.1038/nature02659.

[17] A. V. Kimel, B. A. Ivanov, R. V. Pisarev, P. A. Usachev, A. Kirilyuk,
and Th. Rasing.
Inertia-driven spin switching in antiferromagnets.
Nature Phys, 5(10):727–731, October 2009.
doi: 10.1038/nphys1369.

[18] Andrei Kirilyuk, Alexey V. Kimel, and Theo Rasing.
Ultrafast optical manipulation of magnetic order.
Rev. Mod. Phys., 82(3):2731–2784, September 2010.
doi: 10.1103/RevModPhys.82.2731.

[19] Challa S.S.R. Kumar, editor.
Magnetic Characterization Techniques for Nanomaterials.
Springer Berlin Heidelberg, Berlin, Heidelberg, 2017.
doi: 10.1007/978-3-662-52780-1.

[20] Kenneth Levenberg.
      A method for the solution of certain non-linear problems in least
          squares.
      Quart. Appl. Math., 2(2):164–168, 1944.
      doi: 10.1090/qam/10666.

[21] Donald W. Marquardt.
      An Algorithm for Least-Squares Estimation of Nonlinear Parameters.
      Journal of the Society for Industrial and Applied Mathematics, 11(2):431–
          441, June 1963.
      doi: 10.1137/0111030.

[22] Aaron Meurer, Christopher P. Smith, Mateusz Paprocki, Ondřej
          Čertík, Sergey B. Kirpichev, Matthew Rocklin, AMiT Kumar,
          Sergiu Ivanov, Jason K. Moore, Sartaj Singh, Thilina Rathnayake,
          Sean Vig, Brian E. Granger, Richard P. Muller, Francesco Bonazzi,
          Harsh Gupta, Shivam Vats, Fredrik Johansson, Fabian Pedregosa,
          Matthew J. Curry, Andy R. Terrel, Štěpán Roučka, Ashutosh Sa-
          boo, Isuru Fernando, Sumith Kulal, Robert Cimrman, and An-
          thony Scopatz.
      Sympy: symbolic computing in python.
      PeerJ Computer Science, 3:e103, January 2017.
      doi: 10.7717/peerj-cs.103.

[23] K Morey, Bryan T. Fichera, Baiqing Lv, Zongqi Shen, and Nuh Gedik.
      Automated polarization rotation for multi-axis rotational-anisotropy
          second harmonic generation experiments.
      In preparation.

[24] J. A. Nelder and R. Mead.
      A Simplex Method for Function Minimization.
      The Computer Journal, 7(4):308–313, January 1965.
      doi: 10.1093/comjnl/7.4.308.

[25] Franz Neumann.
      *Vorlesungen über die Theorie der Elasticität der festen Körper und des
          Lichtäthers, gehalten an der Universität Königsberg.*
      Leipzig, B. G. Teubner, 1885.

[26] Shyue Ping Ong, William Davidson Richards, Anubhav Jain, Geof-
 froy Hautier, Michael Kocher, Shreyas Cholia, Dan Gunter, Vin-
 cent L. Chevrier, Kristin A. Persson, and Gerbrand Ceder.
 Python Materials Genomics (pymatgen): A robust, open-source
 python library for materials analysis.
 Computational Materials Science, 68:314–319, February 2013.
 doi: 10.1016/j.commatsci.2012.10.028.

[27] Shreyas Patankar, Liang Wu, Baozhu Lu, Manita Rai, Jason D. Tran,
 T. Morimoto, Daniel E. Parker, Adolfo G. Grushin, N. L. Nair, J. G.
 Analytis, J. E. Moore, J. Orenstein, and D. H. Torchinsky.
 Resonance-enhanced optical nonlinearity in the Weyl semimetal
 TaAs.
 Phys. Rev. B, 98(16):165113, October 2018.
 doi: 10.1103/PhysRevB.98.165113.

[28] P. S. Pershan.
 Nonlinear Optical Properties of Solids: Energy Considerations.
 Phys. Rev., 130(3):919–929, May 1963.
 doi: 10.1103/PhysRev.130.919.

[29] G. D. Powell, J.-F. Wang, and D. E. Aspnes.
 Simplified bond-hyperpolarizability model of second harmonic gen-
 eration.
 Physical Review B, 65(20), May 2002.
 doi: 10.1103/PhysRevB.65.205320.

[30] Richard C. Powell.
 Symmetry, Group Theory, and the Physical Properties of Crystals.
 Springer New York, NY, 2010.

[31] William H. Press, Saul A. Teukolsky, William T. Vetterling, and
 Brian P. Flannery.
 Numerical Recipes 3rd Edition: The Art of Scientific Computing.
 Cambridge University Press, 3 edition, 2007.

[32] A. Ron, E. Zoghlin, L. Balents, S. D. Wilson, and D. Hsieh.
 Dimensional crossover in a layered ferromagnet detected by spin
 correlation driven distortions.
 Nat Commun, 10(1):1654, December 2019.

doi: 10.1038/s41467-019-09663-3.

[33] D. Sa, R. Valentí, and C. Gros.
A generalized Ginzburg-Landau approach to second harmonic generation.
Eur. Phys. J. B, 14(2):301–305, March 2000.
doi: 10.1007/s100510050133.

[34] H. T. Stokes, D. M. Hatch, and B. J. Campbell.
Isodistort.

[35] Zhiyuan Sun and Andrew J. Millis.
Transient Trapping into Metastable States in Systems with Competing Orders.
Phys. Rev. X, 10(2):021028, May 2020.
doi: 10.1103/PhysRevX.10.021028.

[36] Darius H. Torchinsky, Hao Chu, Tongfei Qi, Gang Cao, and David Hsieh.
A low temperature nonlinear optical rotational anisotropy spectrometer for the determination of crystallographic and electronic symmetries.
Review of Scientific Instruments, 85(8):083102, August 2014.
doi: 10.1063/1.4891417.

[37] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors.
SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python.
Nature Methods, 17:261–272, 2020.
doi: 10.1038/s41592-019-0686-2.

[38] H. Wackernagel.
Multivariate Geostatistics: An Introduction with Applications.
Springer Berlin Heidelberg, 2003.

[39] David J. Wales and Jonathan P. K. Doye.
Global Optimization by Basin-Hopping and the Lowest Energy Struc-
tures of Lennard-Jones Clusters Containing up to 110 Atoms.
J. Phys. Chem. A, 101(28):5111–5116, July 1997.
doi: 10.1021/jp970984n.

[40] Liang Wu, S. Patankar, T. Morimoto, N. L. Nair, E. Thewalt, A. Little,
J. G. Analytis, J. E. Moore, and J. Orenstein.
Giant anisotropic nonlinear optical response in transition metal
monopnictide Weyl semimetals.
Nature Phys, 13(4):350–355, April 2017.
doi: 10.1038/nphys3969.

[41] Y Xiang, D.Y Sun, W Fan, and X.G Gong.
Generalized simulated annealing algorithm and its application to the
Thomson model.
Physics Letters A, 233(3):216–220, August 1997.
doi: 10.1016/S0375-9601(97)00474-X.

[42] H. J. Zeiger, J. Vidal, T. K. Cheng, E. P. Ippen, G. Dresselhaus, and
M. S. Dresselhaus.
Theory for displacive excitation of coherent phonons.
Phys. Rev. B, 45:768–778, Jan 1992.
doi: 10.1103/PhysRevB.45.768.