UC Davis
UC Davis Electronic Theses and Dissertations

Title

Ultrafast Pulses Produced via Random Quasi-Phase-Matching Measured with Frequency-Resolved Optical Gating

Permalink

https://escholarship.org/uc/item/7kv139rc

Author

Glick, Noah Peter

Publication Date

2023

Peer reviewed|Thesis/dissertation

Ultrafast Pulses Produced via Random Quasi-Phase-Matching
Measured with Frequency-Resolved Optical Gating

By

Noah Glick
THESIS

Submitted in partial satisfaction of the requirements for the degree of

MASTER OF SCIENCE

in

Electrical and Computer Engineering

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

_____
William Putnam, Chair

_____
Weijian Yang

_____
Juan Sebastian Gomez-Diaz

Committee in Charge

2023

Ultrafast Pulses Produced via Random Quasi-Phase-Matching Measured with Frequency-Resolved Optical Gating

by

Noah Glick

Submitted to the Department of Electrical and Computer Engineering
on December 8, 2023, in partial fulfillment of the
requirements for the degree of
Master of Science in Electrical Engineering

Abstract

Random quasi-phase-matching (RQPM) is a nonlinear optical phenomenon known to occur naturally in polycrystalline materials. When driven by a pulsed source in the infrared, RQPM in ZnSe/S can generate second-harmonic light. This second-harmonic generation (SHG) can reach optical powers comparable to those of conventional ultrafast sources, on the order of hundreds of milliwatts. These pulses are not well characterized, nor has a complete analysis on them been done to determine factors such as compressibility. Previous measurements of the light generated by RQPM have focused on the power spectra and have said nothing about the spectral phase, which is required for making predictions about the temporal profile and coherence of the light. Here, we use frequency-resolved optical gating (FROG) to reconstruct a temporal description of a laser pulse produced by SHG via RQPM in ZnS. The experiment is based on a collinear SHG beam geometry that can be easily reconfigured for additional measurements, including linear and interferometric autocorrelations. After measuring the pulse, we perform an additional measurement in order to resolve an ambiguity inherent to SHG FROG. Finally, we will report on an experiment that suggests pulses produced by RQPM in ZnS might be compressible to durations of 35 fs or shorter.

Thesis Supervisor: William Putnam
Title: Assistant Professor

Acknowledgments

Contents

List of Figures

Chapter 1

Introduction

This thesis reports on the measurement of ultrafast laser pulses using the technique of frequency-resolved optical gating (FROG) [1]. The pulses under consideration are **ultrafast** in the sense that their instantaneous variations are on the order of femtoseconds (fs, *i.e.* 10^{-15} seconds) [2]. These variations are fantastically fast: there are a million billion femtoseconds in just a single second. Pulses whose durations are on the order of a few fs are among the shortest events created and recorded by human beings, although the boundary has been pushed even further in recent years toward shorter and shorter events. To measure an event so short in time seems daunting, but among our primary goals is that this thesis will help to elucidate some of the mystery.

The pulses studied in this paper are formed by a process not well characterized experimentally. Our objective will be to reconstruct a nearly-complete description of the these pulses. The pulses we study here have center wavelengths in the near-infrared (NIR), a region of the electromagnetic spectrum just outside the visible wavelengths corresponding to $\sim 700 - 3000$ nm.

The procedure for constructing the temporal description of an ultrafast pulse is called pulse **recovery** or **retrieval**. It might already be clear why retrieval is a nontrivial task, considering the rapid nature of the variations considered. All measurements in experimental science are limited in some way by the speed at which scientific instruments can physically operate. We say more about this in the introduction to Chapter 2, but the short version of the problem is that most of the measurement tools available to researchers are far too slow to

observe fs variations, such as those which characterize an ultrafast pulse. Clearly, methods other than direct measurement are required for retrieval. This thesis is devoted towards a description of those other methods.

To set the stage for describing our retrieval experiments, we will first outline a basic theory for pulse measurement. We will also unravel the signal processing required to characterize the pulse from a physical experiment. Our theoretical framework will detail the tools needed to perform such a measurement, whereas our study of signal processing will describe the steps necessary to carry out a retrieval. Our method is based on FROG, which is the first method invented to completely characterize a pulse. With the stage set, we will draw back the curtains by proceeding to report on a host of measurements, using several metrics to cross-reference our results. We will conclude by reporting on an exciting, preliminary result that speaks to the usefulness of these pulses for continued research.

1.1 Our Ultrafast Laser Source

In order to measure a pulse, we must first produce one, and here we examine a source capable of the task. Our system produces pulses in a Cr:ZnSe/S Kerr-lens modelocked (KLM) solid-state laser. We are measuring the output of a CLPF series laser manufactured by IPG Photonics Inc. Fig. 1-1 illustrates a diagram of our system. The Cr-doped ZnSe/S polycrystalline gain element is pumped using erbium-doped fiber lasers (EDFLs) operating at 1530-1570 nm. The gain element in our system is anti-reflection-coated (AR-coated) and the pump is normally incident to maximize pump and laser intensity in the gain element. This also has the effect of increasing the likelihood of producing a second-harmonic signal from the nonlinear optical effects in the crystal. The infrared signal exits the laser through an output coupler (OC).

The center wavelength of the Cr:ZnSe/S laser can vary, but a typical wavelength is around 2.4 μm. Our system supports pulse durations of approximately 30 fs, at a repetition rate of 80.97 MHz. The pulse energy is 40 nJ, and the time-bandwidth product (TBP) is less than 0.35. The optical power produced by the infrared output is 3.2 W.

The Cr:ZnSe/S laser system also produces a second-harmonic output via RQPM in the
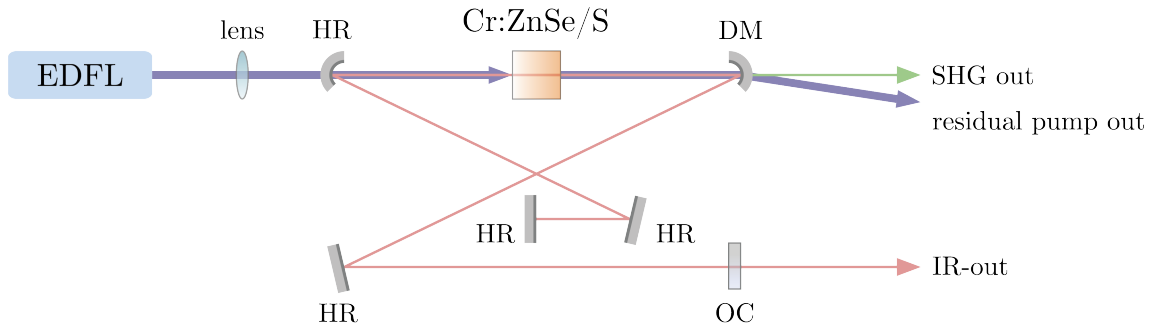
Figure 1-1: **Cr:ZnSe/S laser source:** HR is a highly reflective mirror, and other acronyms are defined in the main text.

gain element. Preliminary measurements have suggested that RQPM SHG might create pulses with a duration around 100-200 fs [3] [4]. Although additional predictions have been made about the spatial and temporal structure of RQPM pulses using simulations [5], to the best of our knowledge, an SHG laser pulse produced via RQPM has not been completely characterized to the degree provided by a FROG system. This thesis is devoted to a complete characterization of the SHG output of a pulse produced by RQPM in a Cr:ZnSe/S laser system such as the one illustrated in Fig. 1-1.

Putting it in plain terms: we will attempt to make a novel measurement of laser pulses that are produced in an unusual way. The pulses oscillate near 1.2 μm and their temporal structure is as-yet unknown.

1.2   Thesis Outline

This thesis is structured to provide a detailed theory for pulse measurement and retrieval, as well as a report of an RQPM SHG pulse measurement made using FROG. Chapter 2 will begin from first principles to outline the physical theory of interference and interferometry, with the ultimate goal of providing a formal description of the FROG measurement, called a **spectrogram** or **FROG trace**. In Chapter 3 we will study both the structure of the FROG trace as well as the algorithm that accompanies it, the latter of which produces a complete description of the pulse from the trace. Measurements will then be reported via experiments that we will detail in Chapter 4. Our closing remarks are made in Chapter 5.

Chapter 2

Interference and Interferometry

This chapter provides a background for the physics and signal processing that will be required to make an ultrafast pulse measurement. We will derive analytical expressions for the temporal and spectral signals produced by the experiments in Chapter 4. Beginning with a mathematical description of interference, we will proceed with an electromagnetic approach that requires little more than knowledge of plane wave solutions to Maxwell's equations.

Before proceeding it is crucial that we understand the effect our measurements will have on the electromagnetic signals. The measurements we make in this thesis will be taken by either a photodiode or a spectrometer. Without delving too deeply into the operation of these devices, we simply state their effects, and the signals that result from a measurement.

To begin, from electromagnetics we know that a real electric field $\boldsymbol{\mathcal{E}}(\mathbf{r}, t)$ has an intensity that is proportional to the magnitude-squared of its complex amplitude $\mathbf{E}(\mathbf{r}, t)$, defined by $\boldsymbol{\mathcal{E}}(\mathbf{r}, t) = \mathrm{Re}\{\mathbf{E}(\mathbf{r}, t)\}$. First, we will almost always prefer to work with the complex description of the field. Second, from the outset we will ignore the vector nature of the electric field, and describe it as a scalar quantity $E(\mathbf{r}, t)$. Third, we will additionally drop the spatial dependence in the course of our derivation by assuming that all measurements occur at the same point in space; our derivation will begin with the spatial dependence still explicit, but we drop it for this note. What remains is a description of an electric field that is a function only of time, denoted $E(t)$. This is the quantity we are measuring: the pulse. Finally, in this thesis we will ignore physical constants such as permittivity, and we will define the intensity of a laser pulse with the normalized expression:

$$I(t) \equiv |E(t)|^2 \ . \tag{2.1}$$

Photodiode

A photodiode (PD) converts an optical signal into an electric current by the generation of electron-hole pairs in the depletion region of a semiconductor *pn* junction. The electron-hole pairs are produced by an optical signal incident upon the device, which contains a certain number of photons per unit area per unit time, or photon flux. In short, the incident photons are destroyed by the interaction, and their energy is transferred to electrons in the valence band of the semiconductor, which are subsequently excited into the conduction band. In the conduction band they are free to move about and form a current. Therefore, the number of electron-hole pairs generated depends on the number of incident photons, which is a quantum mechanical description of the intensity contained in a classical electromagnetic wave. The upshot of these remarks is that a photodiode measures the intensity of an optical signal.

As alluded to in the introduction, any detector is characterized to some degree by its bandwidth, which in the case of a photodiode is a physical limitation on the speed at which charge carriers are free to move about. The signals in an ultrafast pulse have instantaneous variations at frequencies far outside the bandwidth of an ordinary photodiode (incidentally, this describes the problem of measuring an ultrafast pulse in the first place). Therefore, in the context of ultrafast photonics, a photodiode acts as a lowpass filter, in that it measures the time-integrated intensity. The notation we will adopt for this thesis uses an angled bracket notation to denote a time integral, so that the measurement at a photodiode of an ultrafast pulse $E(t)$ is given by:

$$\langle |E(t)|^2 \rangle \equiv \int |E(t)|^2 \, \mathrm{d}t \ , \tag{2.2}$$

where an integral without bounds will be implied to be taken over all values.

Spectrometer

A pulse can be described as a superposition of plane waves with distinct frequencies, or equivalently, by its Fourier transform. When a pulse is incident upon a spectrometer, a system of optics within the device serves to disperse the constituent plane waves throughout space. Since the individual wave components are dispersed spatially, individual intensities can be measured at an array of individual detectors. As an ensemble of measurements, the spectrometer provides a description of the pulse's power spectrum. We will denote the measurement of a pulse $E(t)$ at a spectrometer formally by the quantity:

$$|\tilde{E}(\omega)|^2 \equiv |\mathcal{F}\{E(t)\}|^2 = \left| \int E(t)\, e^{-i\omega t}\, \mathrm{d}t \right|^2 , \tag{2.3}$$

where the notation $\mathcal{F}\{\cdot\}$ will be used to denote a Fourier transform.

As a final word on notation, we will generally use the symbol $\tilde{S}(\omega)$ to denote the magnitude-squared of the pulse **spectrum** $\tilde{E}(\omega)$, that is, $\tilde{S}(\omega) = |\tilde{E}(\omega)|^2$. We will refer to $\tilde{S}$ as the **power spectrum**. We have taken particular care to make this distinction, as other conventions in the FROG literature refer to $\tilde{S}(\omega)$ as "the spectrum" [1] [6] [7]. In our convention, a spectrometer measures the power spectrum.

2.1 Interference

Interference arises in optics when considering the intensity of a sum of field quantities. The mathematical origins of interference can be understood by taking the magnitude-squared of a sum of complex numbers.

$$|z_1 + z_2 + \cdots + z_N|^2 = \left| \sum_{n=1}^{N} z_n \right|^2 = \left(\sum_{n=1}^{N} z_n \right) \left(\sum_{n=1}^{N} z_n \right)^* . \tag{2.4}$$

The conjugation distributes to each term in the latter sum, and both sums can be combined:

$$|z_1 + z_2 + \cdots + z_N|^2 = \sum_{n=1}^{N} \sum_{m=1}^{N} z_m z_n^* . \tag{2.5}$$

Eq. (2.5) is a generalized form of an equation usually called **the interference equation** [8].

We can write it in a more intuitive way by noting that there are two cases. When $m = n$, we get terms of the form $Z_n \equiv |z_n|^2$; when $m \neq n$ we get a pair of terms for each combination of nm, for example $z_1 z_2^* + z_2 z_1^*$. If we collect the individual magnitudes into a single term $Z_B = \sum_n^N Z_n$, the remaining pairs can be combined in one of two convenient ways depending on the problem at hand. Stating the results here for reference, we have:

$$|z_1 + z_2 + \cdots + z_N|^2 = Z_B + 2 \sum_{nm} \sqrt{Z_n Z_m} \, \cos \varphi_{mn} \, , \tag{2.6}$$

where $\varphi_{mn} = \angle z_m - \angle z_n$ is the phase difference between each interfering term, and

$$|z_1 + z_2 + \cdots + z_N|^2 = Z_B + 2 \, \mathrm{Re} \left\{ \sum_{nm} z_m z_n^* \right\} \, . \tag{2.7}$$

The notation $\sum_{nm}$ denotes a sum over every distinct of combination nm. Written this way, we can interpret the interference equation to consist of a **background term** and an **interference term**, the latter of which records the interference between each summed quantity. We will refer to these two forms as the **trigonometric interference equation** and the **complex interference equation**, respectively.

Both equations can be understood well using phasor diagrams like those in Fig. 2-1. Fig. 2-1a defines the geometry for a sum of two phasors using the tip-to-tail method, and Fig. 2-1b illustrates how the resulting magnitude of the sum changes (given by the length of the black arrow) as the phase difference is varied. We can generalize this idea to the sum of an arbitrary number of phasors (Fig. 2-1c).

(a)　　　　　　　　(b)　　　　　　　　(c)
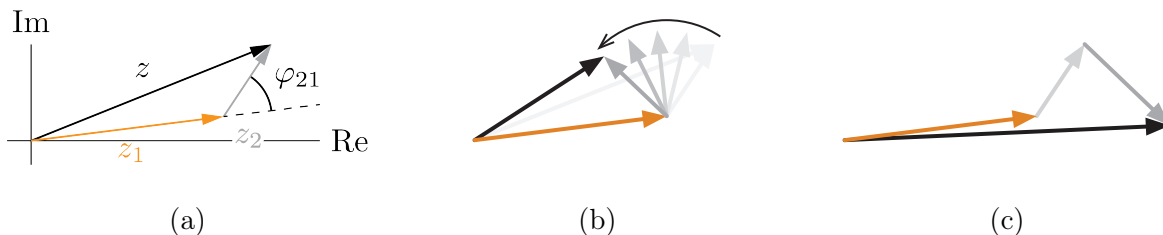
Figure 2-1: **Phasor geometry underlying the interference equation:** In 2-1a the sum of two phasors is given by a black arrow labelled $z = z_1 + z_2$. 2-1b shows how the magnitude of z depends on the phase difference φ_{21}. Note that the individual magnitudes $|z_1|$ and $|z_2|$ (the background) are unchanged. 2-1a illustrates that the idea of interference extends to any sum of phasors.

Eqs. (2.6) and (2.7) are not particularly nice expressions, but they are quite general. For small N, we can write out the terms explicitly. For example, in the case of linear interferometry, we will need to sum two field quantities and measure their resulting intensity. Choosing the trigonometric interference equation provides an intensity given by:

$$|z_1 + z_2|^2 = Z_1 + Z_2 + 2\sqrt{Z_1 Z_2} \cos \varphi_{21} \ . \tag{2.8}$$

We will soon find that in the course of FROG, which is a spectral form of nonlinear interferometry, we will need to sum three fields and measure the intensity. The result will take the form:

$$|z_1 + z_2 + z_3|^2 = Z_1 + Z_2 + Z_3 + 2\operatorname{Re}\{z_1 z_2^* + z_1 z_3^* + z_2 z_3^*\} \ . \tag{2.9}$$

We will prefer the complex form for its linear quality, which will simplify some of the analysis.

2.2 Interferometry

The interferometer is an optical system that exploits interference [2] [9]. The system works by producing two copies of an input wave at a beamsplitter, and then subjecting each wave to (possibly) different propagation lengths. The two waves are denoted as the **reference** and **signal**, and they each traverse a distinct **arm** of the interferometer. For our development, a beamsplitter is defined simply as an optical component that takes in an electromagnetic wave and splits half of its intensity into two output directions. The effect of propagation introduces a relative phase between the waves, which then form a superposition at a second beamsplitter. From the interference equation, this relative phase has an effect on the magnitude of the output. Note from Fig. 2-2 that the second beamsplitter produces two identical outputs, but most of our measurements will be based on observations at only one of the outputs. An aside: at the risk of using sloppy language, from here on we will freely exchange between the term *beam* and *wave* to describe an electric field or an electromagnetic wave.

One of the simplest optical layouts for recording interference is called a Michelson inter-
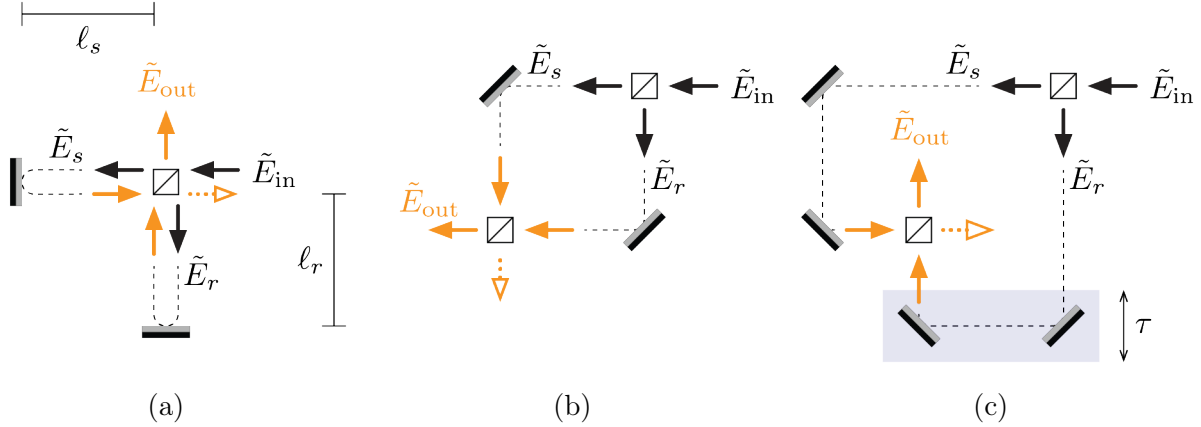
Figure 2-2: **Three configurations for an interferometer:** The Michelson interferometer (2-2a) is the simplest to construct. The Mach-Zehnder interferometer (2-2b, 2-2c) includes additional features for more complicated experiments, most notably the ability to reposition mirrors without altering the beam path. The quantities in this figure will be defined and utilized in the next section.

ferometer, wherein a single beamsplitter separates and recombines the constituent waves, as illustrated in Fig. 2-2a. By recombine we mean that at the second pass, the beamsplitter sees two inputs whose corresponding outputs are made to form a superposition. An alternative is the Mach-Zehnder interferometer which uses two beamsplitters, illustrated by two layouts in Figs. 2-2b and 2-2c. The latter scheme may be suitable for a number of reasons, including the ability for a beam to propagate through an optical component only once, a feature not easily permitted by the Michelson case. We prefer a Mach-Zehnder configuration for our experiments since the secondary output (illustrated with a dotted arrow) propagates in a direction separate from the input or primary output. This feature will allow us to use the secondary output for calibration, as we will examine in a later section. In particular, we use the configuration in Fig. 2-2c, since translating the mirrors in and out during an experiment will be necessary to produce a time delay τ. The configuration in 2-2b is not suitable, as shifting a mirror would alter the beam path, indicated by the dashed lines.

The remainder of this chapter is focused on examining signals produced by an interferometer subject to time-harmonic and ultrafast inputs. For the latter, we will observe the output with and without a nonlinear optical element, using both a photodiode and a spectrometer. We will compare the time domain and frequency domain content between all results.

9

2.2.1 Continuous Wave

In optics, a real, time-harmonic electric field is usually called a continuous wave (cw). If the interferometer input is cw, then we can express the electric field as a separable function:

$$E(\mathbf{r}, t) = \tilde{E}(\mathbf{r}) \, e^{i\omega_0 t} \, , \tag{2.10}$$

where the complex amplitude $\tilde{E}(\mathbf{r}) = \tilde{E}_{\text{in}} \, e^{-i\mathbf{k}\cdot\mathbf{r}}$ is time-invariant. Passage through a 50:50 beamsplitter produces two fields, which we call the reference and signal fields, with corresponding complex amplitudes $\tilde{E}_r(\mathbf{r}) = (\tilde{E}_{\text{in}}/\sqrt{2}) \, e^{-i\mathbf{k}_r\cdot\mathbf{r}}$ and $\tilde{E}_s(\mathbf{r}) = (\tilde{E}_{\text{in}}/\sqrt{2}) \, e^{-i\mathbf{k}_s\cdot\mathbf{r}}$. The factor $1/\sqrt{2}$ owes to the fact that the beamsplitters were defined to be lossless in the sense that $|\tilde{E}|^2 = |\tilde{E}_r|^2 + |\tilde{E}_s|^2$. Citing the geometry of the Michelson interferometer defined in Fig. 2-2a, note that propagation through the respective arms of the interferometer produces phase shifts $2k\ell_r$ and $2k\ell_s$, where $k = |\mathbf{k}_r| = |\mathbf{k}_s|$, so that $\tilde{E}_{\text{out}}(\mathbf{r}) = \tilde{E}_r(\mathbf{r}) \, e^{-i(2k\ell_r)} + \tilde{E}_s(\mathbf{r}) \, e^{-i(2k\ell_s)}$. The expression for the complete complex output field $E_{\text{out}}(\mathbf{r}, t) = \tilde{E}_{\text{out}}(\mathbf{r}) \, e^{i\omega_0 t}$ can be written[1]:

$$\tilde{E}_{\text{out}}(\mathbf{r}, t, t_r, t_s) = \frac{1}{\sqrt{2}} \left[\tilde{E}_r(\mathbf{r}) \, e^{i\omega_0(t-t_r)} + \tilde{E}_s(\mathbf{r}) \, e^{i\omega_0(t-t_s)} \right] \, , \tag{2.11}$$

where $t_j = 2k\ell_j/\omega_0$ for $j \in \{r, s\}$, and the additional factor of $1/\sqrt{2}$ owes to a second pass through another 50:50 beamsplitter (where the fields are recombined). Eq. (2.11) emphasizes that the phase accumulated by propagation through space can be interpreted as a shift in time for the fields in each arm of the interferometer. Making a change of frame to the origin of the reference field $t' = t - t_r$ and defining the relative delay between the signal and reference as $\tau = t_s - t_r$, Eq. (2.11) can be written more succinctly as:

$$E_{\text{out}}(\mathbf{r}, t', \tau) = \frac{1}{\sqrt{2}} \left[\tilde{E}_r(\mathbf{r}) \, e^{i\omega_0 t'} + \tilde{E}_s(\mathbf{r}) \, e^{i\omega_0(t'-\tau)} \right] = \frac{1}{\sqrt{2}} \left[E_r(\mathbf{r}, t') + E_s(\mathbf{r}, t' - \tau) \right] \, . \tag{2.12}$$

---

[1] Other effects can be accounted for by considering the phase accumulated by propagation through a beamsplitter, polarization effects from reflection, and others. In this analysis we neglect all of these other effects, in the approximation that they remain constant for the duration of an experiment. Our interest will simply be in the signal that is produced by changing the propagation length in one arm of the interferometer.

For convenience, we now assume that all of the observations in the remainder of this paper will take place at some position $\mathbf{r} = 0$. This frees us to drop the spatial dependence, and we also drop the prime on the shifted time variable. The result is a simplified expression for the output of the interferometer $E_{\text{out}}(t, \tau) = (1/\sqrt{2})\,[E_r(t) + E_s(t - \tau)]$ or or in terms of the interferometer input:

$$E_{\text{out}}(t, \tau) = \frac{1}{\sqrt{2}} \left[\frac{1}{\sqrt{2}} E(t) + \frac{1}{\sqrt{2}} E(t - \tau) \right] = \frac{1}{2} \left[E(t) + E(t - \tau) \right] . \tag{2.13}$$

At zero relative delay, the reference and signal fields recombine perfectly so that $E_{\text{out}}(t) = E(t)$.

Since Eq. (2.13) is expressed as the sum of two complex quantities we can invoke the interference equation directly to determine the intensity. Choosing the trigonometric form yields:

$$I_{\text{out}}(t, \tau) = |E_{\text{out}}(t, \tau)|^2 = \frac{1}{4} \left[I(t) + I(t - \tau) + 2\sqrt{I(t)I(t - \tau)}\,\cos\omega_0\tau \right] . \tag{2.14}$$

Noting that for time-harmonic inputs the time-dependence of $E(t) = \tilde{E}_{\text{in}}\,e^{i\omega_0 t}$ appears only in a complex phase factor, we can say that $I(t) = I(t - \tau) = |\tilde{E}_{\text{in}}|^2 = I_{\text{in}}$, a constant. Each term in Eq. (2.14) is thus constant for a given value of τ, and a time integral leaves the expression unchanged. Therefore, distributing a time integral to each term, we get:

$$U_{\text{out}}(\tau) = \langle I_{\text{out}}(t, \tau) \rangle = \frac{U_{\text{in}}}{2} \left(1 + \cos\omega_0\tau \right) , \tag{2.15}$$

where we will use the variable U to denote the time integrated intensity, which we will sometimes refer to as the *signal energy*. In Eq. (2.15) we have defined $U_{\text{in}} \equiv \langle I_{\text{in}} \rangle$, which in this case is just I_{in}. The factor of $1/2$ can be understood by noting that an interferometer has two output ports, and U_{out} corresponds to the signal energy observed at just one of the ports.

We should remark that in the case of cw signals, the trigonometric form of the interference equation is preferred since the term $\langle \sqrt{I(t)I(t - \tau)} \rangle = \langle |E(t)||E(t - \tau)| \rangle$ reduces to

11

a constant, which allowed us to combine it with the input term. This will not be the case with pulsed signals, for which a time-dependence exists separate from the complex oscillation factor. In any case, we will generally normalize our results (noting only proportionality), and we will take interest in the value of the background, in this case denoted $U_B = \langle I(t) \rangle$. Accordingly we will prefer the complex interference equation, which produces an alternative form of Eq. (2.15) given by:

$$U_{\text{out}}(\tau) \propto U_B + \text{Re}\left\{G_1(\tau)\right\} . \qquad (2.16)$$

The term $G_1(\tau) = \langle E(t)E^*(t - \tau) \rangle$ is called the field autocorrelation function [2] [10], and we will examine it further in the case of a pulsed input signal. For now, we simply note that this form neglects the factor of $1/2$, and emphasizes the distinction between the background term and the interference term (everything inside $\text{Re}\{\cdot\}$).

Eq. (2.15) is informative. The time-integrated intensity observed on a photodiode is an oscillating function of delay, which is produced in our case by manually changing the propagation distance in the signal arm. This underlies the rich theory of interferometry, and in basic terms, states that any change in the signal arm can be detected at a photodiode.

We have plotted a trace of Eq. (2.16) in Fig. 2-3 for future reference. Note that at zero delay, the interference term $G_1(0) = U_B$, which provides the trace a peak to background ratio of 2:1.
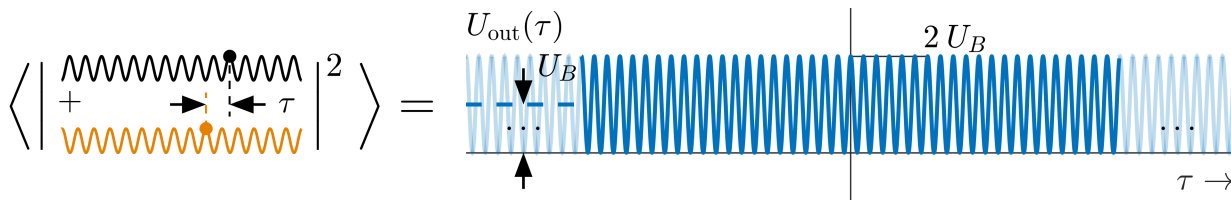
Figure 2-3: **Interferometer output energy for a CW input:** Notice the oscillatory nature of the trace, which arises from the interference between the fields in each arm of the interferometer. When the output power increases at one port, it decreases at the other, and vice versa.

In order to complete the forthcoming analogies for ultrafast optics, we also plot the spectrum of the interferometer output signal using a Fourier transform. This leads us to a crucial logistical detail: a function of delay τ has a spectrum that is a function of its Fourier

conjugate variable, which we call **delay frequency**, Ω. The Fourier transform of a function of delay will be denoted $\mathcal{F}_\tau\{\cdot\} \equiv \int \cdot \, e^{-i\Omega\tau} \, d\tau$. Upon quick examination of Eq. (2.15), we expect the spectrum to contain three delta functions, one at the origin for the background term, and two at $\pm\omega_0$ for the interference term:

$$\mathcal{F}_\tau\{U_{\text{out}}(\tau)\} = \pi U_{\text{in}} \left[\delta(\Omega) + \frac{1}{2}\,\delta(\Omega + \omega_0) + \frac{1}{2}\,\delta(\Omega - \omega_0) \right] . \tag{2.17}$$

This spectrum of U_{out} is plotted in Fig. 2-4.
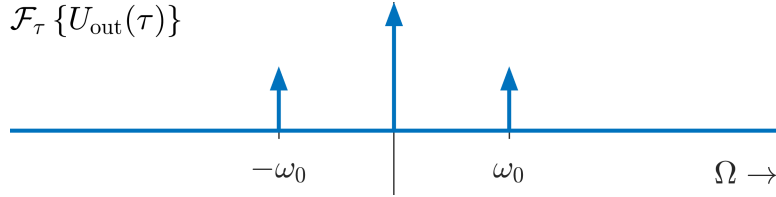
Figure 2-4: **The spectrum of the interferometer output energy for a cw input:** The presence of the delta functions is understood best from a signal processing perspective: from Fig. 2-3 the trace has infinite duration in time (since it is produced from everlasting sinusoids), and so each of its constituents has infinitessimal bandwidth [?].

2.2.2 Ultrafast

The equation that governs the system response of the interferometer is given by Eq. (2.15), and it was derived by assuming a time-harmonic input signal. An important feature to note is that it is a linear equation; therefore, the interferometer constitutes a linear *system*. Stated differently, Eq. (2.15) is valid even in the case that $E(t)$ is *not* time-harmonic. This follows from Fourier analysis, from which we know that $E(t)$ can be expressed as a superposition of time-harmonic functions. Since Eq. (2.15) holds for each term in the superposition, it holds for $E(t)$.

In this section, we indeed assume that $E(t)$ is not time-harmonic, and instead assume that it takes the conventional form of an ultrafast pulse [2], written:

$$E(t) = A(t)\, e^{i\omega_0 t} , \tag{2.18}$$

In this form, we see that the ultrafast pulse is conventionally described by the product of some **slowly-varying envelope** function $A(t)$ with a **carrier** $e^{i\omega_0 t}$. Although this is

fairly straightforward, we should address the conventions in order to be explicit about which quantities we are dealing with. First, recall that the electric field is a real quantity, and we prefer to work with its complex representation. Their relationship is repeated: $\mathcal{E}(t) = \text{Re}\{E(t)\} = [E(t) + E^*(t)]/2$. It is not always easy to visualize complex quantities, but we will use Fig. 2-5 as a reference.
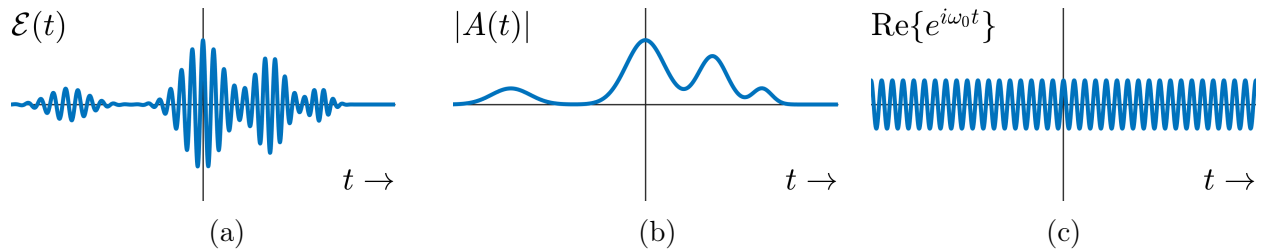
Figure 2-5: **Structure of an ultrafast pulse:** Plotted in 2-5a is the real field $\mathcal{E}(t)$. The real field is conventionally written in terms of an envelope, whose magnitude is plotted in 2-5b, and a carrier, whose real part is plotted in 2-5c.

Fig. 2-5a shows the real field $\mathcal{E}(t)$ corresponding to a pulse that we will use as a model for the remainder of this chapter. This pulse has more structure than commonly encountered pulses, in that it has more than one peak. We are modeling a complex pulse, because we will find that the pulse we examine in Chapter 4 has similar qualities.

The fluctuations shown in Fig. 2-5a are actual variations in the magnitude of the electric field as a function of time at some position in space. On the other hand, $A(t)$ illustrates the complex envelope of the complex field $E(t)$. In Fig. 2-5b we have plotted only the *magnitude* of $A(t)$, which of course is the same as the magnitude of $E(t)$—the carrier $e^{i\omega_0 t}$ has unity magnitude. The magnitude $|A(t)|$ gives us a sense of the overall shape of the pulse, ignoring the rapid fluctuations of the carrier term. Additionally, in Fig. 2-5c we have plotted the *real part* of the carrier, which is simply a cosine. The carrier gives us a sense of how quickly the real field $\mathcal{E}(t)$ fluctuates in magnitude. The complex representation $E(t)$, although not plotted, is simply a convenient way to store information about the pulse while making computations.

It is important to note that $A(t)$ also has a phase, that is $A(t) = |A(t)|\, e^{i\angle A(t)}$. This phase can be combined with the carrier, which produces a modulated cosine:

$$\mathcal{E}(t) = \text{Re}\{A(t)\, e^{i\omega_0 t}\} = \text{Re}\{|A(t)|\, e^{i\angle A(t)}\, e^{i\omega_0 t}\} = |A(t)|\, \cos(\omega_0 t + \angle A(t))\ . \qquad (2.19)$$

The phase $\angle A(t)$ contains information about the instantaneous frequency of the pulse, which is usually described by by a parameter called **chirp** [2]. There are conventions for dealing with chirp in both the temporal phase and the spectral phase of the pulse. In this paper, we are mostly concerned with experimental procedures, and we do not do any analysis of pulse chirp. To understand the distinction between the phase $\angle E(t)$ and the phase $\angle A(t)$, refer to Fig. 2-6.

$\cos(\omega_0 t)$

$\cos(\omega_0 t + \angle A(t))$

$t \rightarrow$ $t \rightarrow$

(a)   (b)

Figure 2-6: **Phase variations in an ultrafast pulse:** Figure 2-6a plots the same information as that in Fig. 2-5c. That is, we have plotted the fluctuations that arise from the carrier term alone. In 2-6b we see variations in the phase, caused by the phase of the envelope $A(t)$. These variations modulate the regular fluctuations of the carrier. Thus, 2-6b is a representation of the total phase of $E(t)$, whereas 2-6a is a representation of the phase of the carrier alone.

Just as important as the time-domain representation of the pulse is its spectral representation. There are other conventions in the ultrafast community for representing the spectra $\tilde{\mathcal{E}}(\omega)$, $\tilde{E}(\omega)$, and $\tilde{A}(\omega)$. In this thesis, we will deviate slightly from the convention, and we will be plotting the *full* spectrum of the real pulse, which is related to the spectrum of the complex representation according to $\tilde{\mathcal{E}}(\omega) = [\tilde{E}(\omega) + \tilde{E}^*(-\omega)]/2$. In turn, the spectrum $\tilde{E}(\omega)$ is related to the spectrum of the envelope by a shift property of the Fourier transform [10]:

$$\tilde{E}(\omega) = \mathcal{F}\{E(t)\} = \mathcal{F}\{A(t)\, e^{i\omega_0 t}\} = \tilde{A}(\omega - \omega_0)\ . \qquad (2.20)$$

It is conventional to work with the spectrum $\tilde{A}(\omega)$, because it is convenient to view plots that are centered about the origin. Fig. 2-7 is the reference for frequency-domain quantities.

15

$\tilde{E}(\omega)$ $\tilde{A}(\omega)$ $\delta(\omega - \omega_0)$

(a)   (b)   (c)
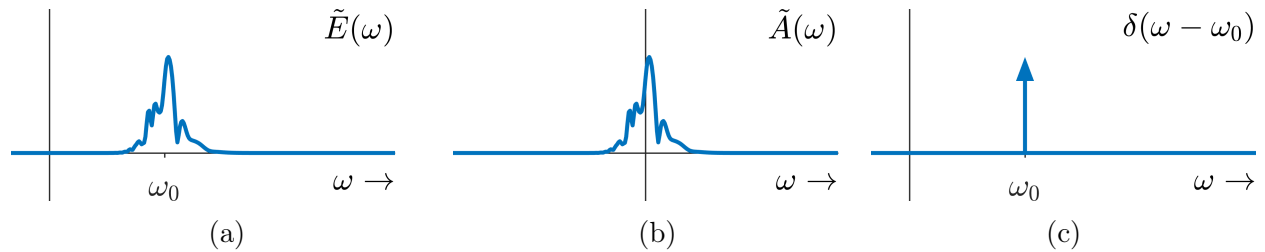
Figure 2-7: **Spectrum of an ultrafast pulse:** Plotted are the magnitude spectra of the complex field, the complex envelope, and the carrier. The spectrum in 2-7a is shifted from the origin, because it includes the carrier (see remarks leading to eq. (2.20)). Conversely, 2-7b is centered at the origin, being the spectrum of the envelope function alone. 2-7c is the spectrum of the carrier, which of course is a shifted delta function. Note: since in the time domain the complex field $E(t)$ is the product of the envelope and the carrier, in the frequency domain the corresponding spectrum is the convolution of the spectra of the envelope and carrier.

Our deviation from convention means that we will prefer to plot $\tilde{\mathcal{E}}(\omega)$, but that unless noted otherwise, the field we are working with is $E(t)$, whose spectrum is only *half* of the spectrum we are plotting. Our reasons for doing things this way are twofold. First, we believe it will provide a better progression from the autocorrelation to the spectrogram, which are quantities with which we will soon become intimately familiar. Second, in the lab we measure physically real quantities, and so the spectra we compute will be given by $\tilde{\mathcal{E}}(\omega)$. The notable exception to this is any measurement taken from a spectrometer, which provides spectral information directly, that is, information over positive frequency. For reference, the spectrum of the real field is plotted in Fig. 2-8.

$\tilde{\mathcal{E}}(\omega)$
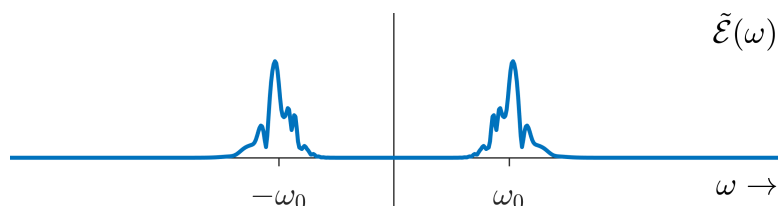
Figure 2-8: **Complete spectrum of an ultrafast pulse:** Plotted is the spectrum of $\mathcal{E}(t)$, which includes negative frequencies. The negative frequency region is a mirror image of the positive frequency region, with the phase inverted (not shown).

This spectrum has a negative frequency portion that is the complex conjugate of the positive frequency portion (hence the relationship we noted above). In summary, the reader should bear in mind that we will do our time-domain derivations in terms of $E(t)$, but our frequency-

domain plots will be of $\tilde{\mathcal{E}}(\omega)$.

We now examine the signals from three experiments that use an interferometer to measure an ultrafast pulse. These go by the names **Fourier-transform infrared** spectroscopy (FTIR[2]), **interferometric autocorrelation** (IAC), and **frequency-resolved optical gating** (FROG). We will soon understand these experiments to comprise a linear time-domain signal, a nonlinear time-domain signal, and a nonlinear frequency-domain signal, respectively. Each electric field signal will be written in terms of the interferometer output (Eq. (2.13)). Additionally, they will each be denoted by the subscript "sig", and—with the exception of FROG—a superscript labelling the experiment. The convention in FROG is to denote the signal field E_{sig}, leaving the superscript open to denote a particular beam geometry [1]. We will address these points more as needed.

Linear Signal

FTIR comprises a linear signal, wherein the signal field is linear in the output field. In our experiment, the signal field is identically the output of the interferometer,

$$E_{\text{sig}}^{\text{FTIR}}(t, \tau) = E_{\text{out}}(t, \tau) = \frac{1}{2} \left[E(t) + E(t - \tau) \right] , \tag{2.21}$$

whose intensity can be written:

$$I^{\text{FTIR}}(t, \tau) = |E_{\text{sig}}^{\text{FTIR}}(t, \tau)|^2 = \frac{1}{4} \left| E(t) + E(t - \tau) \right|^2 . \tag{2.22}$$

Choosing the complex form of the interference equation, we get:

$$I^{\text{FTIR}}(t, \tau) = \frac{1}{4} \left[I(t) + I(t - \tau) + 2 \operatorname{Re}\{E(t)E^*(t - \tau)\} \right] , \tag{2.23}$$

so that at a photodiode we record:

$$\langle I^{\text{FTIR}}(t, \tau) \rangle = \frac{1}{4} \left[\langle I(t) \rangle + \langle I(t - \tau) \rangle + 2 \operatorname{Re}\{\langle E(t)E^*(t - \tau) \rangle\} \right] , \tag{2.24}$$

---

[2]FTIR usually refers to a spectroscopic technique used in other scientific fields, one which is more or less identical to that which we describe here.

which can be written as:

$$U^{\text{FTIR}}(\tau) \propto U_B + \text{Re}\{G_1(\tau)\} \ . \tag{2.25}$$

The background term $U_B = \langle I(t)\rangle$ is again the input signal energy[3], and $G_1(\tau) = \langle E(t)E^*(t-\tau)\rangle$ is again the field autocorrelation. Eq. (2.25) is the ultrafast analog to Eq. (2.16) in the regime of linear optics; the two equations are identical, but we reiterate that Eq. (2.25) does not reduce to Eq. (2.15) in this case (see the discussion leading up to Eq. (2.16)). Furthermore, the autocorrelation is taken between functions that have finite duration, in contrast to the everlasting sinusoids we used previously. Accordingly, the autocorrelation trace is only nonzero for a finite region where the two copies of the pulse overlap, as shown in the figure below. Fig. 2-9 is the ultrafast linear analog to Fig. 2-3.
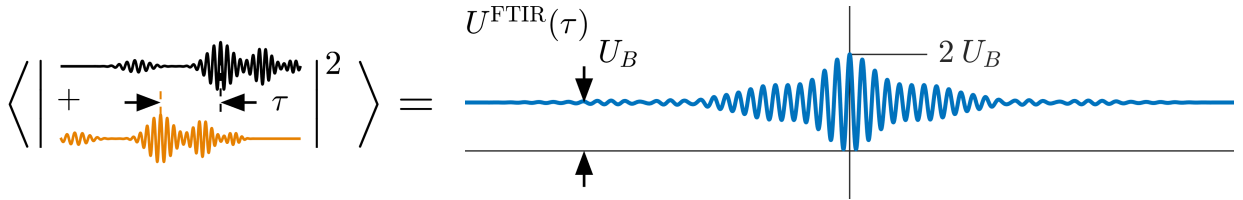
Figure 2-9: **FTIR trace:** When the delay is large ($\tau \to \infty$), the pulses do not interfere and half the pulse intensity is observed at each port, given by the delay-independent background term in eq. (2.25); when the delay is small ($\tau \to 0$) the two copies of the pulse interfere strongly, and the power oscillates between the output ports of the interferometer. Note that we have modelled a moderately complex sample pulse, and the structure of the autocorrelation echoes the multiple-peaked structure of the pulse.

Our present interest is in the spectral content of the FTIR. The first term is constant, and thus produces a delta function at the origin. In order to understand the spectrum of the second term, we simply invoke the autocorrelation theorem[4] [10] which tells us generally that $\mathcal{F}\{\langle x(t)x^*(t-\tau)\rangle\} = |\tilde{X}(\omega)|^2$, or in the present case:

---

[3] Note that even though $I(t)$ is not a constant in the case of a pulsed input, it still follows that $\langle I(t)\rangle = \langle I(t-\tau)\rangle$ so both terms can combine. This is easy to understand on physical grounds, since delaying the pulse does not change its optical energy; on mathematical grounds a change of variables $t' = t - \tau$ on the second integral yields the same result as the first, since the bounds of integration are infinite.

[4] The autocorrelation theorem is often called the Wiener-Khinchin theorem (when considering stochastic signals), and can be derived as a particular case of the convolution theorem.

$$\mathcal{F}_\tau\{\text{Re}\{G_1(\tau)\}\} = \frac{1}{2}\,\mathcal{F}_\tau\{G_1(\tau) + G_1^*(\tau)\} = \frac{1}{2}\,|\tilde{E}(\Omega)|^2 + \frac{1}{2}\,|\tilde{E}^*(-\Omega)|^2 = |\tilde{\mathcal{E}}(\Omega)|^2 \ . \quad (2.26)$$

Thus, Fourier transforming Eq. (2.25) tells us that the FTIR provides an alternative means of measuring the power spectrum. The right hand side of Eq. (2.26) is the same function we observed in Fig. 2-8. Including a delta function for the background term, we produce Fig. 2-10, which is the spectrum of an FTIR measurement.
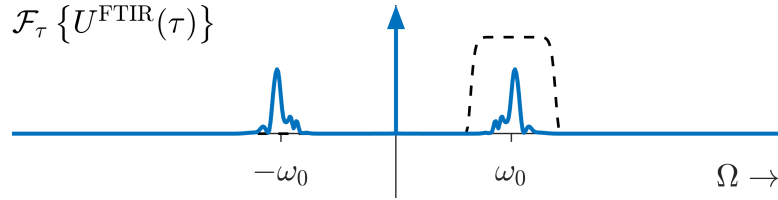
Figure 2-10: **FTIR spectrum:** The impulses in Fig. 2-4 have been smeared out, a feature again explained well by a signal processing perspective: from Fig. 2-9 the trace has finite duration in time, and so by the uncertainty relation of the Fourier transform, the corresponding spectral terms have finite bandwidth (the constant background still produces a delta function at the origin). The dashed line indicates a filter around the signal of interest, which is the pulse power spectrum.

The method of recovering the pulse power spectrum from an FTIR measurement proved to be crucial in our experimental studies, as we will discuss in Chapter 4. Briefly, we state that in practical situations, the spectrum can be recovered from a fast Fourier transform (FFT) of the recorded FTIR trace and by searching the spectrum over the region of interest. If the background term is not omitted, the spectrum will be dwarfed by the impulse, requiring the need to change the window size of the plot. Alteratively, a simpler approach is to normalize the trace to the background before taking an FFT; the resulting plots are much easier to interpret.

Nonlinear Signal

IAC comprises a nonlinear signal, wherein the signal field is the square of the interferometer output:

$$E_{\text{sig}}^{\text{IAC}}(t, \tau) = E_{\text{out}}^2(t, \tau) = \frac{1}{4}\left[E(t) + E(t - \tau)\right]^2 , \tag{2.27}$$

whose intensity is given by:

$$I^{\text{IAC}}(t, \tau) = |E_{\text{sig}}^{\text{IAC}}(t, \tau)|^2 = \frac{1}{16}\left|E^2(t) + E^2(t - \tau) + 2\,E(t)E(t - \tau)\right|^2 . \tag{2.28}$$

We can once again invoke the interference equation, in this case for a sum of three complex quantities (Eq. (2.9)), with the result

$$
\begin{aligned}
I^{\text{IAC}}(t, \tau) = \frac{1}{16}\Big[& I^2(t) + I^2(t - \tau) + 4\,I(t)I(t - \tau) \\
& + 2\,\text{Re}\left\{E^2(t)E^{*\,2}(t - \tau) + 2\,E^2(t)E^*(t)E^*(t - \tau) + 2\,E^2(t - \tau)E^*(t)E^*(t - \tau)\right\}\Big].
\end{aligned}
\tag{2.29}
$$

Taking the time integral, we find:

$$
\begin{aligned}
\langle I^{\text{IAC}}(t, \tau)\rangle = \frac{1}{16}\Big[& \langle I^2(t)\rangle + \langle I^2(t - \tau)\rangle + 4\,\langle I(t)I(t - \tau)\rangle \\
& + 2\,\text{Re}\left\{\langle E^2(t)E^{*\,2}(t - \tau)\rangle + 2\,\langle E^2(t)E^*(t)E^*(t - \tau)\rangle + 2\,\langle E^2(t - \tau)E^*(t)E^*(t - \tau)\rangle\right\}\Big].
\end{aligned}
\tag{2.30}
$$

Cancelling a factor of 2, the result can be reduced to:

$$U^{\text{IAC}}(\tau) \propto U_B + \text{Re}\left\{2\,\left[G_1^{\text{mod}}(\tau) + G_1^{\text{mod}\,*}(-\tau)\right] + G_1^{(2)}(\tau)\right\} + 2\,G_2(\tau) . \tag{2.31}$$

Examination of Eq. (2.31) reveals the usual background term and interference term (compare with Eqs. (2.16) and (2.25)), but the nonlinearity has produced additional complexity. First, we note the nonlinear background term is not simply in the input signal energy, $i.e.$ $U_B = \langle I^2(t)\rangle \neq \langle I(t)\rangle$. Additionally, the interference term for the nonlinear case contains interactions between the constituent fields at two frequencies: $G_1^{\text{mod}}(\tau) = \langle I(t){\cdot}E(t)E^*(t - \tau)\rangle$

represents a modified field autocorrelation, and $G_1^{(2)}(\tau) = \langle E^2(t)E^{*\,2}(t-\tau)\rangle$ represents a field autocorrelation of the second harmonic. Finally, the nonlinearity records an additional interaction in the term $G_2(\tau) = \langle I(t)I(t-\tau)\rangle$, a quantity usually called the **intensity autocorrelation**.

A key takeaway is that for large delay the autocorrelations disappear[5], and $U^{\mathrm{IAC}}(\tau) \to U_B$, whereas for small delays the constituent interference terms (and G_2) all approach U_B so that $U^{\mathrm{IAC}}(\tau) \to 8\,U_B$. In other words, the peak to background ratio of the IAC trace is 8:1. An example IAC trace was generated in Fig. 2-11 using the same pulse as that which produced the FTIR trace in Fig. 2-9.
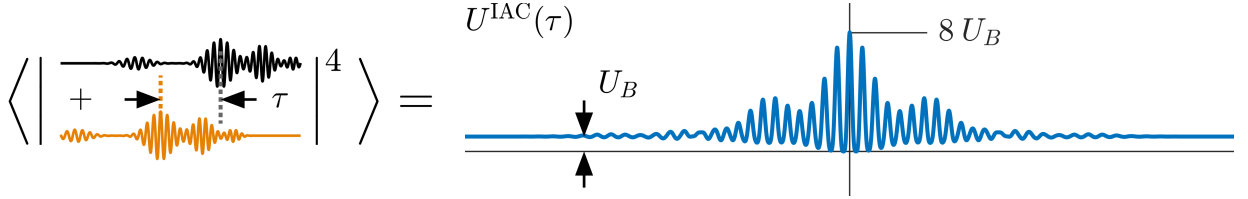
Figure 2-11: **IAC spectrum:** The nonlinearity used to produce this signal involves several interactions, and the corresponding signal energy includes interference terms beyond just the field autocorrelation. An important check on measurements is the peak to background ratio of the IAC trace, which should take the value 8:1 for a correctly measured trace.

In the nonlinear case the intensity autocorrelation produces spectral content at baseband ($I(t)$ does not contain a carrier term). On the other hand, the interference term produces spectral content in the neighborhood of $\pm\omega_0$ and in the neighborhood of $\pm 2\omega_0$, from G_1^{mod} and $G_1^{(2)}$ respectively. The latter result follows from our previous remarks about the autocorrelation theorem (*i.e.* $\langle E^2(t)E^{*\,2}(t-\tau)\rangle = |\mathcal{F}\{E^2(t)\}|^2$) and from the fact that $E^2(t) = A^2(t)\,e^{i2\omega_0 t}$ contains a carrier oscillating at $2\omega_0$. These remarks mostly explain the spectral features observed by Fourier transforming the IAC trace, as we have illustrated by the plot Fig. 2-12.

Nonlinear Spectral Signal

Abiding by the notation that is common to the FROG literature, we denote the signal field $E_{\mathrm{sig}}(t,\tau)$, and we define the **FROG field** as its frequency domain equivalent $\tilde{E}_{\mathrm{FROG}}(\omega,\tau) =$

---

[5]Since the signals in question have finite extent, sending τ to infinity basically amounts to saying that the zero regions of each term in $\langle x(t)x^*(t-\tau)\rangle$ cover the other term, so their product goes to zero.

$\mathcal{F}_\tau \left\{ U^{\mathrm{IAC}}(\tau) \right\}$

$-2\omega_0 \qquad -\omega_0 \qquad\qquad \omega_0 \qquad 2\omega_0 \qquad \Omega \to$
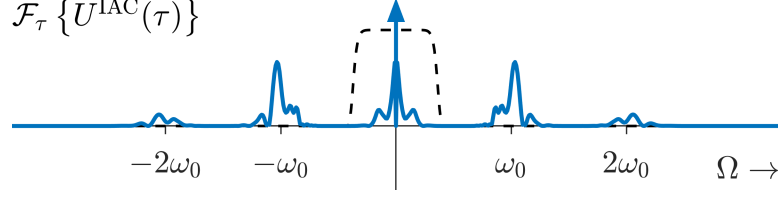
Figure 2-12: **IAC spectrum:** The intensity autocorrelation is contained in the baseband signal, and can be selectively measured by lowpass filtering (dashed lines). We will examine this term in a different context in the next section. The spectral features at $\pm\omega_0$ appear similar to the pulse spectrum recovered by an FTIR trace, but the terms are not exactly a field autocorrelation, and we have not ascribed to them any analytical description. On the other hand, the features at $\pm 2\omega_0$ appear from the term $\mathcal{F}\{\langle E^2(t)E^{*\,2}(t-\tau)\rangle\} = |\mathcal{F}\{E^2(t)\}|^2$, which is the second harmonic power spectrum.

$\mathcal{F}\{E_{\mathrm{sig}}(t,\tau)\}$. In our experiment, FROG comprises the same nonlinear system that is formed by an IAC, so the signal field is again the square of the output of the interferometer,

$$E_{\mathrm{sig}}(t,\tau) = E_{\mathrm{sig}}^{\mathrm{IAC}}(t,\tau) = E_{\mathrm{out}}^2(t,\tau) = \frac{1}{4}\left[E(t) + E(t-\tau)\right]^2 . \qquad (2.32)$$

In the case of FROG, the signal field is directed into a spectrometer, which amounts to taking the magnitude-squared of the Fourier transform of Eq. (2.32), which by linearity can be written:

$$\tilde{S}_{\mathrm{FROG}}(\omega,\tau) \equiv |\tilde{E}_{\mathrm{FROG}}(\omega,\tau)|^2 = \frac{1}{16}\left|\mathcal{F}\{E^2(t)\} + \mathcal{F}\{E^2(t-\tau)\} + 2\,\mathcal{F}\{E(t)E(t-\tau)\}\right|^2 . \tag{2.33}$$

Eq. (2.34) is the magnitude squared of the sum of three complex quantities—yet another application of the interference equation. Expanding the terms according to Eq. (2.7), we find:

$$\begin{aligned}
\tilde{S}_{\mathrm{FROG}}(\omega,\tau) = \frac{1}{16}\Big[& |\mathcal{F}\{E^2(t)\}|^2 + |\mathcal{F}\{E^2(t-\tau)\}|^2 + 4\,|\mathcal{F}\{E(t)E(t-\tau)\}|^2 \\
& + 2\,\mathrm{Re}\,\Big\{\mathcal{F}\{E^2(t)\}\mathcal{F}^*\{E^2(t-\tau)\} + 2\,\mathcal{F}\{E^2(t)\}\mathcal{F}^*\{E(t)E(t-\tau)\} \\
& + 2\,\mathcal{F}\{E^2(t-\tau)\}\mathcal{F}^*\{E(t)E(t-\tau)\}\Big\}\Big] .
\end{aligned} \tag{2.34}$$

As before, the first two terms are equivalent[6]: they are each the power spectrum of the second harmonic (as would be observed directly at a spectrometer), which we denote $\tilde{S}_B(\omega) \equiv \mathcal{F}\{E^2(t)\}$. We will sometimes find use to denote the second harmonic spectrum as $\tilde{\Gamma}_1(\omega) = \mathcal{F}\{E^2(t)\} = \tilde{E}(\omega) * \tilde{E}(\omega)$, which emphasizes that it is the *autoconvolution* of the fundamental spectrum. The remaining terms have analogous features to the terms in Eq. (2.30). An equation analogous to Eq. (2.31) can be written in similar terms:

$$\tilde{S}_{\text{FROG}}(\omega, \tau) \propto \tilde{S}_B(\omega) + \text{Re}\left\{2\,\tilde{\Gamma}_1(\omega)\left[1 + e^{-i\omega\tau}\right]\tilde{E}_{\text{FROG}}^{\text{SHG}*}(\omega, \tau) + |\tilde{\Gamma}_1(\omega)|^2\,e^{i\omega\tau}\right\} + 2\tilde{S}_{\text{FROG}}^{\text{SHG}}(\omega, \tau)\,. \tag{2.35}$$

Eq. (2.35) is not attractive, but with some care we can unpack it to discover some interesting features. First, a statement that is absolutely pivotal to the development of this thesis, is that we have discovered the quantities $\tilde{E}_{\text{FROG}}^{\text{SHG}}(\omega, \tau) = \mathcal{F}\{E(t)E(t - \tau)\}$, which is called the **SHG FROG field**, and $\tilde{S}_{\text{FROG}}^{\text{SHG}}(\omega, \tau) = |\tilde{E}_{\text{FROG}}^{\text{SHG}}(\omega, \tau)|^2$ which is the **SHG FROG spectrogram**. The SHG FROG spectrogram is the entire foundation upon which the pulse recovery process is built, as outlined in Chapter 3. Eq. (2.35) says that this important signal is contained within an interferometric autocorrelation measurement, since it was derived from $E_{\text{sig}}(t, \tau) = E_{\text{sig}}^{\text{IAC}}(t, \tau)$. We will say more about these signals in Section 2.2.3 and throughout Chapter 3. First, it is important to understand how to isolate this term from our measurement.

Since we are not performing a time integral, note that the interference equation applied to a spectral signal produces a 2-dimensional time-frequency distribution $\tilde{S}_{\text{FROG}}(\omega, \tau)$ called a spectrogram. The term spectrogram extends beyond FROG, but within the context of FROG it is usually modified in such a way that denotes which type of FROG geometry it describes (as is the case for the SHG FROG spectrogram). In this case, we have used the notation $\tilde{S}_{\text{FROG}}$ (without an identifier) to denote measurement of an interferometric autocorrelation signal. We will say a bit more about these topics in later sections.

---

[6]Again this can be understood on physical grounds, since delaying a pulse does not change its power spectrum. On mathematical grounds, the shift theorem tells us that a time delay amounts to a spectral phase shift, which leaves the spectral intensity unaffected. That is, $\mathcal{F}\{E^2(t-\tau)\} = \mathcal{F}\{E^2(t)\}\,e^{-i\omega\tau} = \tilde{\Gamma}_1(\omega)\,e^{-i\omega\tau}$, so $|\mathcal{F}\{E^2(t)\}|^2 = |\mathcal{F}\{E^2(t-\tau)\}|^2$.

We conventionally view the spectrogram with a two-dimensional color map. The added dimension arises from the fact that at each delay point we are now recording a power spectrum, which corresponds to a vertical slice over the spectrogram. Since each of the interfering terms in Eq. (2.35) are recorded in the same signal, we observe a fringe pattern over the spectrogram identical to the one that was observed in the interferometeric autocorrelation trace. We will study the relationship between the spectrogram and the autocorrelation in Section 3.1. For now we simply illustrate the analogy: Fig. 2-13 is the analog to Figs. 2-3, 2-9, and 2-11 for nonlinear spectral signals.
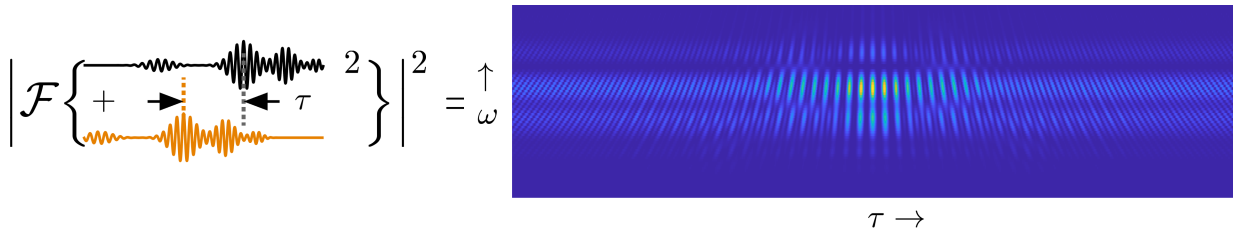
Figure 2-13: **Interferometric FROG spectrogram:** This image is a two-dimensional distribution that is formed by recording a spectrum for every delay point. This trace completely contains the autocorrelation trace recorded in Fig. 2-11, and then some. This is not surprising since the interferometric autocorrelation is produced by measuring the same signal at a photodiode. We will examine the various features of the spectrogram in Chapter 3.

In order to strengthen the analogy, we must examine the spectrum of Eq. (2.35). First, a note of caution. We will be looking at the spectrum of a 2-dimensional function over the delay variable, which involves a corresponding (1-dimensional) Fourier transform $\mathcal{F}_\tau\{\cdot\} \equiv \int \cdot\, e^{-i\Omega\tau}\, \mathrm{d}\tau$. However, our development will also require us to take the Fourier transform over the *regular coordinates* (t, ω), which we denote $\mathcal{F}\{\cdot\} = \int \cdot\, e^{-i\omega t}\, \mathrm{d}t$. To ease matters, the transforms $\mathcal{F}_\tau\{\cdot\}$ and $\mathcal{F}\{\cdot\}$ will be interpreted to be commuting linear operators in the sense that $\mathcal{F}\{\mathcal{F}_\tau\{\cdot\}\} \equiv \mathcal{F}_\tau\{\mathcal{F}\{\cdot\}\}$: the operators commute, and terms not affected by the operator can "push through".

In particular, we note that the resulting 2-dimensional spectrum will be unaltered along ω, that is, terms dependent upon only ω will push through the $\mathcal{F}_\tau\{\cdot\}$ operator. Breaking up the task into smaller pieces, we first check the spectrum of the SHG FROG field:

$$\mathcal{F}_\tau\{\tilde{E}^{\text{SHG}}_{\text{FROG}}(\omega,\tau)\} = \mathcal{F}_\tau\{\mathcal{F}\{E(t)E(t-\tau)\}\} = \mathcal{F}\{E(t)\,\mathcal{F}_\tau\{E(t-\tau)\}\}\;, \qquad (2.36)$$

where we have swapped $\mathcal{F}\{\cdot\}$ with $\mathcal{F}_\tau\{\cdot\}$. Applying the scaling and shift theorems to the integral over τ allows us to write:

$$\mathcal{F}_\tau\{\tilde{E}^{\text{SHG}}_{\text{FROG}}(\omega,\tau)\} = \mathcal{F}\{E(t)\,\mathcal{F}_\tau\{E(-(\tau+t))\}\} = \mathcal{F}\{E(t)\tilde{E}(-\Omega)\,e^{-i\Omega t}\}\;. \qquad (2.37)$$

Now, the term in Ω is invariant over the regular coordinates, and the complex exponential in t results in a spectral shift along ω in the remaining term, that is,

$$\mathcal{F}_\tau\{\tilde{E}^{\text{SHG}}_{\text{FROG}}(\omega,\tau)\} = \tilde{E}(-\Omega)\,\mathcal{F}\{E(t)\,e^{-i\Omega t}\} = \tilde{E}(-\Omega)\tilde{E}(\omega+\Omega)\;. \qquad (2.38)$$

Of course, we also need to account for both complex conjugation,

$$\mathcal{F}_\tau\{\tilde{E}^{\text{SHG}\,*}_{\text{FROG}}(\omega,\tau)\} = \tilde{E}^*(\Omega)\tilde{E}^*(\omega-\Omega)\;, \qquad (2.39)$$

as well as a complex exponential factor,

$$\mathcal{F}_\tau\{e^{-i\omega\tau}\,\tilde{E}^{\text{SHG}\,*}_{\text{FROG}}(\omega,\tau)\} = \tilde{E}^*(\Omega+\omega)\tilde{E}^*(\omega-(\Omega+\omega)) = \tilde{E}^*(-\Omega)\tilde{E}^*(\Omega+\omega)\;. \qquad (2.40)$$

Eqs. (2.39) and (2.40) tell us that the spectrum of the first interference term in Eq. (2.35) $\tilde{\Gamma}_1\,[1+e^{-i\omega\tau}]\,\tilde{E}^{\text{SHG}\,*}_{\text{FROG}}$ looks like the fundamental pulse spectrum multiplied by a copy of itself shifted left-and-right in delay frequency by an amount ω. Noting that a shift in ω *also* corresponds to a shift up-and-down in the 2-dimensional spectrum, we can expect to find spectral features in the neighborhood of the fundamental spectrum ($\pm\omega_0$) in both directions. If that is a brain twister, we can view things differently: Eqs. (2.39) and (2.40) are written in a way that seems similar to an autocorrelation, which must decay for large enough ω. Since there is no integral, features are spread out in two dimensions, but they are all in the

25

region of the carrier frequency ω_0.

The second interference term is easier to handle. A transform gives

$$\mathcal{F}_\tau\{|\tilde{\Gamma}_1(\omega)|^2\, e^{i\omega\tau}\} = |\tilde{\Gamma}_1(\omega)|^2\, \mathcal{F}_\tau\{e^{i\omega\tau}\} = |\tilde{\Gamma}_1(\omega)|^2\, \delta(\Omega - \omega)\ . \qquad (2.41)$$

In words, the second interference term produces a delta function that shifts linearly along ω as Ω is increased. In other words, for a given value of ω (a horizontal line across the 2-D spectrum), we expect a spectral feature with magnitude $|\tilde{\Gamma}_1(\omega)|^2$ at precisely the delay frequency $\Omega = \omega$ (a vertical line); over all ω this produces a sharp slanted line. Since this term contains the second harmonic spectrum $\tilde{\Gamma}_1(\omega) = \mathcal{F}\{E^2(t)\}$, these features are found in the region of $2\omega_0$.
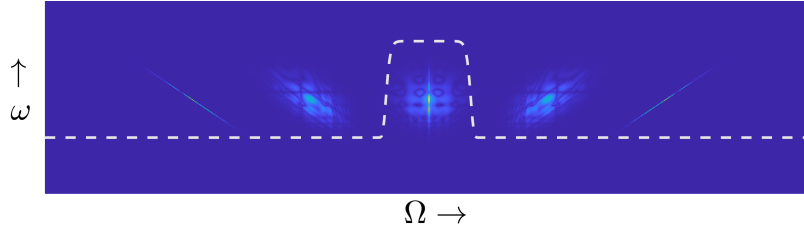
Figure 2-14: **2-dimensional spectrum:** This image was formed by taking a 1-dimensional Fourier transform of the FROG trace that was recorded by an IAC in Fig. 2-13. The features of this spectrum are analogous to the features contained in the 1-dimensional spectrum plotted in Fig. 2-12. Looking ahead, we will find that we can filter the baseband signal of the FROG spectrogram using a lowpass filter, indicated by the dashed line in this figure.

The most important takeaway from this chapter is stated here: the distinct terms in Eqs. (2.31) and (2.35) can be described by non-overlapping spectral features. This fact will enable us to recover the intensity autocorrelation from an IAC trace via a simple lowpass filter, as indicated by the dashed line in Fig. 2-12. Similarly (and more importantly), we can recover the SHG FROG trace from the *same* experimental setup as an IAC. All that is required is a lowpass filter applied to the 2-dimensional spectrum produced by a Fourier transform over delay of the spectrogram [11] [12]. The required filter is indicated by a dashed line in Fig. 2-14. The results of the filtering are plotted below in Fig. 2-15.

In short, this development explains our method for recovering an SHG FROG trace using *precisely* the same experimental layout as is used in an IAC, up to the detector. In the next

section we examine these signals in the context of different beam geometries, to provide geometrical intuition for the underlying physics.

$G_2(\tau)$

$\tau \rightarrow$

(a) Intensity autocorrelation

$\tau \rightarrow$
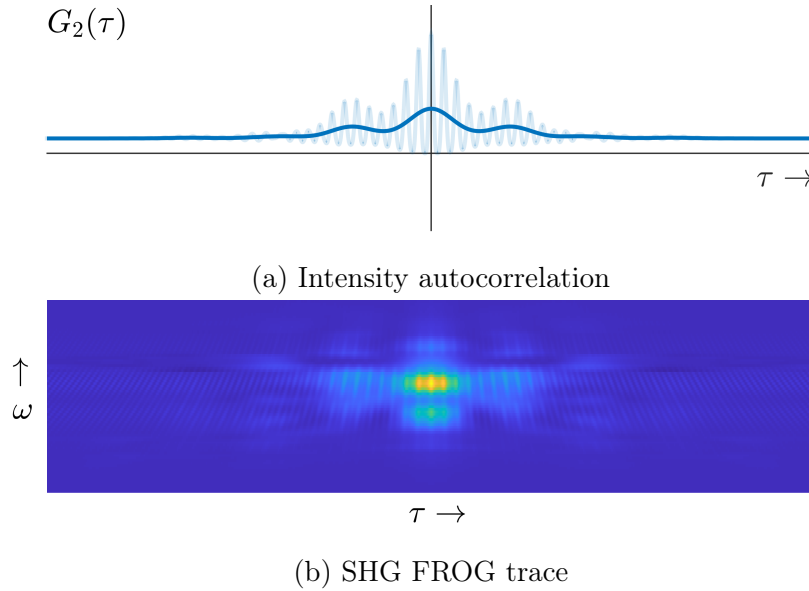
(b) SHG FROG trace

Figure 2-15: **Intensity autocorrelation and SHG FROG trace produced via lowpass filtering:** 2-15a is the intensity autocorrelation $G_2(\tau)$ produced from the IAC trace in Fig. 4-15 via lowpass filtering (dashed line in Fig. 2-12). 4-20 is the SHG FROG trace $\tilde{S}_{\mathrm{FROG}}^{\mathrm{SHG}}(\omega, \tau)$ produced from the collinear FROG trace in Fig. 2-13 via identical filtering (dashed line in Fig. 2-14). In both images we have left a ghost of the fringes for illustrative purposes, *i.e.* in order to compare them against the unfiltered signal. The true signals exhibit no such fringes.

2.2.3   SHG Beam Geometries

Our previous observations have led us to the discovery that an SHG FROG trace can be easily produced in one of two ways. The first method is the conventional method and is known as a **non-collinear geometry** [2]. The non-collinear geometry forms an intensity autocorrelation at a photodiode and an SHG FROG trace at a spectrometer. An example experimental layout illustrated in Fig. 2-16.

Producing an SHG FROG trace in the conventional way requires aligning the two outputs of an interferometer in such a way that they propagate in a parallel direction with no overlap. In the context of FROG, the reference and signal fields are called the **probe** and **gate**, denoted $E_p(t)$ and $E_g(t-\tau)$ respectively [1] [6]. The two parallel beam paths should be near to each other, so that they can both be focused by a single lens onto a nonlinear crystal with

a strong second order response (a $\chi^{(2)}$ crystal). The nonlinear response produces output signals, in addition to the propagated input signals, given by $E_p^2(t)$ and $E_g^2(t-\tau)$ which both propagate collinearly with their corresponding fundamental signal. An additional signal results from the nonlinear interaction of the two inputs; we will not address the formal nonlinear optics of the interaction, except to explain in a heuristic way that by symmetry, adding the **k**-vectors of the mixed inputs results in a new signal that propagates in a distinct spatial direction. A brief explanation of this idea is provided in the caption to Fig. 2-16. This distinct signal is the SHG FROG signal field $E_{\mathrm{sig}}^{\mathrm{SHG}}(t,\tau) = E_p(t)E_g(t-\tau)$, whose measurement at a spectrometer as a function of delay produces the SHG FROG spectrogram $\tilde{S}_{\mathrm{FROG}}^{\mathrm{SHG}}(\omega,\tau) = |\mathcal{F}\{E_{\mathrm{sig}}^{\mathrm{SHG}}(t,\tau)\}|^2$.
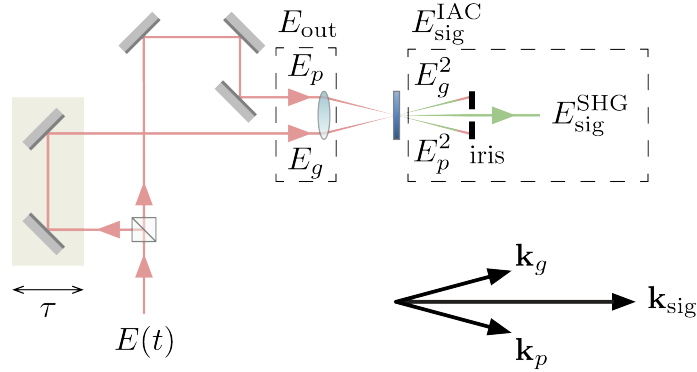
Figure 2-16: **Non-collinear beam geometry for SHG FROG:** The formation of a new signal can be understood in an intuitive way by considering the **k**-vectors of the probe and gate, denoted $\mathbf{k}_p$ and $\mathbf{k}_g$. Their geometric sum produces a new vector $\mathbf{k}_{\mathrm{sig}}$ whose direction is distinct from the others. This can be interpreted as a statement about conservation of momentum.

An alternative beam geometry overlaps the two interferometer outputs so that they propagate collinearly (Fig. 2-17). This is aptly called a collinear geometry, which forms an interferometric autocorrelation at a photodiode and a *collinear* SHG FROG trace at a spectrometer, sometimes called interferometric FROG, IFROG [11], or collinear SHG FROG [12]. This is the quantity we denoted by $\tilde{S}_{\mathrm{FROG}}(\omega,\tau)$ in the previous section. The collinear geometry can be built with a Mach-Zehnder inteferometer in the configuration we illustrated in Fig. 2-2c. Precisely the same signals are produced within the nonlinearity in the collinear case as in the non-collinear case, but addition of **k**-vectors does not produce any new beam

directions at the nonlinear element. Since the fundamental output is also an output of the crystal, and since all outputs propagate collinearly, an optical bandpass filter must be used to select the second harmonic signal $E_{\text{out}}^2(t, \tau)$, which is the same as $E_{\text{sig}}^{\text{IAC}}(t, \tau)$

The collinear signal as-measured cannot be written as the product of a probe and a variably-delayed gate, so the collinear SHG system does not constitute a conventional FROG beam geometry, which we will define in Section 3.1.1. In order to isolate the SHG FROG signal, a lowpass filter must be applied to the measurement after it has been recorded, using the methods described in Section 2.2.2.
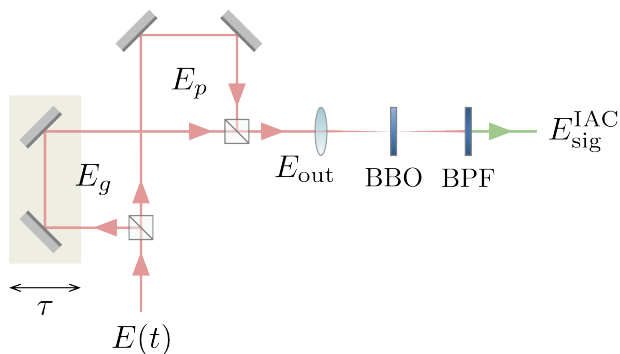
Figure 2-17: **Collinear beam geometry for SHG FROG**

We note that one of the merits of a collinear geometry is its support for pulses with complex spatial and/or temporal structure, such as the one we measure in this paper. We therefore utilize a collinear geometry in all of our experiments described in Chapter 4.

Chapter 3

FROG and the Pulse Retrieval Algorithm

Until the advent of FROG, methods of pulse characterization were limited to autocorrelation techniques, like IAC. These methods, while useful in their own right, do not completely characterize the temporal pulse profile, as they say cannot precisely determine the spectral phase. The merits of FROG were novel at the time: a complete characterization of a laser pulse up to a number of "trivial ambiguities" [6]. The ambiguities are minor, and they depend on the exact experimental layout used for the measurement. These days, a variety of options exist for pulse retrieval, which constitute a veritable zoo, with choices that include STRIPED-FISH [13], PENGUIN [14], SPIDER [2], and TADPOLE [15]. Many of these methods are based on FROG. Despite the arrival of newfound company, FROG has stood the test of time: FROG continues to be one of the most popular ultrafast pulse characterization techniques to this day, in large part due to its robustness and ease of implementation. In this chapter we investigate the method, which will expose us to a number of computational techniques and a brief study of linear algebra.

Let us first take a moment to outline the basics. FROG is an optical system accompanied by a computer algorithm, whose purpose is to measure an ultrafast pulse. As the name suggests, the technique involves the use of an interferometer to split an input pulse into two copies of itself, and to use one copy to optically *gate* the other. Mathematically, the gating amounts to an autocorrelation of the kind we examined in the previous chapter. The

measurements are recorded in the frequency domain, and the accompanying algorithm takes the result of the measurement as an input and spits out a description of the pulse. The measured quantity is called the spectrogram, or FROG trace. The retrieval algorithm has evolved and improved since FROG's conception, but the basic procedure for inverting a spectrogram into a pulse has remained mostly unchanged.

The following chapter is organized as follows. In Section 3.1 we will rebuild the spectrogram in the conventional way, but in a context different from that of Chapter 2. We will closely examine its various features. Section 3.2 includes a detailed description of the retrieval algorithm, the first steps of which are identical over its various incarnations; lastly, Section 3.2.2 is devoted to a particular implementation of the final step.

3.1   The Spectrogram

Our approach in Chapter 2 involved contextualizing the spectrogram by comparison to other interferometric signals. One reason for presenting topics in this way was to provide a comprehensive reference for the temporal and spectral signals produced by simple interferometric techniques. In particular, an analysis of this detail has not been found for collinear SHG FROG, even in the works cited. A more important reason is to provide the reader a simple physical framework for interpreting each of the examined signals.

The purpose of this section is to provide a detailed study of the spectrogram squarely in the context of FROG. We will first take note of the experimental factors required for producing other kinds of FROG traces, and we will find that the interferometric autocorrelation signal does not not qualify as a conventional FROG signal, but it can be filtered according the methods of Section 2.2.2. We will then conclude by examining additional signals that are contained in the trace, and we will state their corresponding uses.

3.1.1   Forming the Trace

SHG FROG is merely one type of geometry that can produce a spectrogram from an ultrafast pulse. The FROG literature is well stocked with resources for producing and examining other geometries. In this brief section, we simply state that in general, a FROG geometry

is structured in such a way to produce an optical signal that contains a probe $E_p(t)$ and a variably-delayed gate $E_g(t - \tau)$. Their product forms a signal field defined as $E_{\text{sig}}(t, \tau) = E_p(t)E_g(t - \tau)$. This expression reveals that all FROG geometries rely on some sort of nonlinearity to produce a signal. SHG FROG has become the most popular geometry, for its ease of implementation and the strength of the signal field (being a second-order technique). In the SHG geometry, the probe and the gate are the same function $E(t)$. In other geometries, this is not the case: one geometry, called polarization-gated (PG) FROG [1] [6], uses a gate that is given by $E_g(t - \tau) = |E(t - \tau)|^2$. Now, as we noted in Section 2.2.2, a general FROG spectrogram is formed by recording a power spectrum at each point in the sweep of an interferometer's delay stage. The resulting expression for a spectrogram formed by the most common geometries is given as:

$$\tilde{S}_{\text{FROG}}(\omega, \tau) = |\mathcal{F}\{E_{\text{sig}}(t, \tau)\}|^2 = \left| \int E_p(t)E_g(t - \tau)\, e^{-i\omega t}\, \mathrm{d}t \right|^2 \tag{3.1}$$

In the case of SHG FROG, the probe and the gate are identical. The resulting expressing is given by:

$$\tilde{S}_{\text{FROG}}^{\text{SHG}}(\omega, \tau) = |\mathcal{F}\{E_{\text{sig}}^{\text{SHG}}(t, \tau)\}|^2 = \left| \int E(t)E(t - \tau)\, e^{-i\omega t}\, \mathrm{d}t \right|^2 \tag{3.2}$$

To address a practical matter, the statement "each point in the sweep" is loose language. What is meant is that a delay stage is operated by some sort of electronic circuit that controls the position of the mirrors in the signal arm of an interferometer. The controller is usually driven electronically with a triangle wave whose amplitude corresponds to the total distance traversed by a mirror in a back and forth motion. A linear driving signal, *i.e.* half the period of a triangle wave (corresponding to the minimum/maximum displacement of the mirror), constitutes a sweep. Mirror displacement is imprinted onto an optical signal as a delay, as we saw in Chapter 2. In short, an electronic controller is usually responsible for the delay experienced by the gate in FROG.

For the remainder of Section 3.1 we will examine two additional signals that are hidden within a spectrogram. These signals are referred to as **marginals**, and they are formed by simple integration of a FROG trace over either of its dimensions. The delay marginal is

defined as:

$$M(\tau) \equiv \int \tilde{S}_{\text{FROG}}(\omega, \tau) \, \mathrm{d}\omega \; , \tag{3.3}$$

and the frequency marginal is defined as:

$$\tilde{M}(\omega) \equiv \int \tilde{S}_{\text{FROG}}(\omega, \tau) \, \mathrm{d}\tau \; . \tag{3.4}$$

3.1.2 The Delay Marginal as an Autocorrelation

The delay marginal is equivalent to an autocorrelation of the signal field that produces the spectrogram [2]. With some thought this fact becomes obvious. At a particular delay the signal at the detector carries a certain optical energy, depending on the mutual interference of the probe and gate. The spectrometer detects the energy of each frequency component; adding up the energy of each component (integrating) produces the total signal energy, which is precisely what is measured by a photodiode in an autocorrelation experiment.

The mathematical proof is as simple as they come. Examining the collinear signal first, by Parseval's theorem[1] we find:

$$\int \tilde{S}_{\text{FROG}}(\omega, \tau) \, \mathrm{d}\omega \equiv \int |\tilde{E}_{\text{FROG}}(\omega, \tau)|^2 \, \mathrm{d}\omega = \int |E^{\text{IAC}}_{\text{sig}}(t, \tau)|^2 \, \mathrm{d}t \equiv \int I^{\text{IAC}}(t, \tau) \, \mathrm{d}t \; . \tag{3.6}$$

The leftmost expression is the definition of the delay marginal, and the rightmost expression is that of the interferometric autocorrelation. They are equivalent:

$$M^{\text{IAC}}(\tau) = U^{\text{IAC}}(\tau) \tag{3.7}$$

---

[1]To be more precise, we should not that we have neglected a factor of 2π, since by convention we use angular frequency. Parseval's theorem in terms of regular frequency ν, reads

$$\int |\tilde{X}(\nu)|^2 \, \mathrm{d}\nu = \int |x(t)|^2 \, \mathrm{d}t \; . \tag{3.5}$$

Replacing $\mathrm{d}\omega = 2\pi \, \mathrm{d}\nu$ introduces the missing scaling factor. We simply neglect it, since we often normalize our results anyway.

In a non-collinear geometry, the proof is similarly simple:

$$\int \tilde{S}^{\text{SHG}}_{\text{FROG}}(\omega, \tau)\, \mathrm{d}\omega \equiv \int |\mathcal{F}\{E(t)E(t-\tau)\}|^2\, \mathrm{d}\omega = \int |E(t)E(t-\tau)|^2\, \mathrm{d}t \ . \qquad (3.8)$$

The delay marginal of the SHG FROG trace is the intensity autocorrelation:

$$M^{\text{SHG}}(\tau) = G_2(\tau) \ . \qquad (3.9)$$

Lastly, for any FROG geometry that exploits a probe $E_p(t)$ windowed by a gate $E_g(t-\tau)$, we have

$$\int \tilde{S}_{\text{FROG}}(\omega, \tau)\, \mathrm{d}\omega \equiv \int |\mathcal{F}\{E_p(t)E_g(t-\tau)\}|^2\, \mathrm{d}\omega = \int |E_g(t)E_g(t-\tau)|^2\, \mathrm{d}t \ , \qquad (3.10)$$

which can be written as an intensity cross-correlation:

$$M(\tau) = \langle I_p(t)I_g(t-\tau)\rangle \ . \qquad (3.11)$$

The standard use for the delay marginal is as an error check in FROG. In our experiments, we record both a IAC and a FROG, and we then compare the former to the delay marginal of the latter, as an assurance that both were recorded correctly. We will shortly find that the frequency marginal contains information that is even more useful than the quick sanity check provided by $M(\tau)$.

3.1.3 The Frequency Marginal as an Autoconvolution

The frequency marginal is equivalent to an autoconvolution [1]. The approach we take here to show this does not appeal to physical intuition in the same way as our note about the delay marginal, nor is this fact quite as obvious. We can derive the result by first interpreting the SHG FROG field as a convolution of the spectrum with a phase shifted copy of itself:

$$\tilde{E}_{\text{FROG}}^{\text{SHG}}(\omega,\tau) = \mathcal{F}\{E(t)E(t-\tau)\} = \frac{1}{2\pi}\mathcal{F}\{E(t)\} * \mathcal{F}\{E(t-\tau)\} = \frac{1}{2\pi}\left[\tilde{E}(\omega)\right] * \left[\tilde{E}(\omega)\,e^{-i\omega\tau}\right].$$

(3.12)

Writing out the convolution integral,

$$\tilde{E}_{\text{FROG}}^{\text{SHG}}(\omega,\tau) = \frac{1}{2\pi}\int \tilde{E}(\omega')\tilde{E}(\omega-\omega')\,e^{-i(\omega-\omega')\tau}\,\mathrm{d}\omega'$$

(3.13)

we find that we can express the SHG FROG trace as a double integral:

$$\begin{aligned}
\tilde{S}_{\text{FROG}}^{\text{SHG}}(\omega,\tau) &= |\tilde{E}_{\text{FROG}}^{\text{SHG}}(\omega,\tau)|^2 \\
&= \left[\frac{1}{2\pi}\int \tilde{E}(\omega')\tilde{E}(\omega-\omega')\,e^{-i(\omega-\omega')\tau}\,\mathrm{d}\omega'\right]\left[\frac{1}{2\pi}\int \tilde{E}(\omega'')\tilde{E}(\omega-\omega'')\,e^{-i(\omega-\omega'')\tau}\,\mathrm{d}\omega''\right]^* \\
&= \frac{1}{(2\pi)^2}\iint \tilde{E}(\omega')\tilde{E}^*(\omega'')\tilde{E}(\omega-\omega')\tilde{E}^*(\omega-\omega'')\,e^{-i(\omega''-\omega')\tau}\,\mathrm{d}\omega'\,\mathrm{d}\omega'',
\end{aligned}$$

(3.14)

where we have combined complex exponentials. The trick comes in taking the integral of Eq. (3.14) over τ:

$$\int \tilde{S}_{\text{FROG}}^{\text{SHG}}(\omega,\tau)\,\mathrm{d}\tau = \frac{1}{(2\pi)^2}\iiint \tilde{E}(\omega')\tilde{E}^*(\omega'')\tilde{E}(\omega-\omega')\tilde{E}^*(\omega-\omega'')\,e^{-i(\omega''-\omega')\tau}\,\mathrm{d}\omega'\,\mathrm{d}\omega''\,\mathrm{d}\tau.$$

(3.15)

Of course, the left hand side is the frequency marginal. The right hand side can be resolved by changing the order of integration and identifying that the exponential factor is the only τ-dependent term,

$$\tilde{M}^{\text{SHG}}(\omega) = \frac{1}{(2\pi)^2}\iint \tilde{E}(\omega')\tilde{E}^*(\omega'')\tilde{E}(\omega-\omega')\tilde{E}^*(\omega-\omega'')\int e^{-i(\omega''-\omega')\tau}\,\mathrm{d}\tau\,\mathrm{d}\omega'\,\mathrm{d}\omega'', \quad (3.16)$$

which delivers a delta function:

$$\tilde{M}^{\mathrm{SHG}}(\omega) = \frac{1}{2\pi} \iint \tilde{E}(\omega')\tilde{E}^*(\omega'')\tilde{E}(\omega - \omega')\tilde{E}^*(\omega - \omega'')\, \delta(\omega'' - \omega')\, \mathrm{d}\omega'\, \mathrm{d}\omega'' \ , \qquad (3.17)$$

that in turn collapses the inner integral (over ω'):

$$\tilde{M}^{\mathrm{SHG}}(\omega) = \frac{1}{2\pi} \int \tilde{E}(\omega'')\tilde{E}^*(\omega'')\tilde{E}(\omega - \omega'')\tilde{E}^*(\omega - \omega'')\, \mathrm{d}\omega'' \ . \qquad (3.18)$$

The spectral terms now depend on the same variable of integration ω'' and can thus be combined to give:

$$\tilde{M}^{\mathrm{SHG}}(\omega) = \frac{1}{2\pi} \int \tilde{S}(\omega'')\tilde{S}(\omega - \omega'')\, \mathrm{d}\omega'' \ , \qquad (3.19)$$

and what remains is identified as an autoconvolution of the power spectrum,

$$\tilde{M}^{\mathrm{SHG}}(\omega) = \frac{1}{2\pi}\tilde{S}(\omega) * \tilde{S}(\omega) \ . \qquad (3.20)$$

Again we note that for any FROG geometry which exploits a probe windowed by a gate, the above development holds, with the result

$$\tilde{M}(\omega) = \frac{1}{2\pi}\tilde{S}_p(\omega) * \tilde{S}_g(\omega) \ . \qquad (3.21)$$

As with the delay marginal, the frequency marginal can be used as an error check in FROG. Specifically, if a power spectrum of the pulse is available, a convolution can be done with the intention of comparing it against the recorded FROG trace's frequency marginal. The two traces should agree, and if they do not, certain error correction methods may be suitable to force their agreement.

The frequency marginal on its own has also been discovered to contain enough information to reconstruct the power spectrum directly from the FROG trace. We will examine this particular point in our discussion of the RANA approach as an aid to pulse recovery in Section 3.2.4.

3.2 Iterative Fourier Transform Algorithm

The procedure for inverting a spectrogram into a pulse is an example of a well known mathematical problem called phase retrieval [1] [16] [17]. A full treatment of the phase retrieval problem is beyond the scope of this thesis. Nonetheless, in order to motivate the retrieval algorithm, we take note here of some of the greatest hits:

- A 1-dimensional phase retrieval problem has no solution.

- A 2-dimensional phase retrieval problem has a (nearly) unique solution

- Both of the previous items can be derived from a discrete description of the spectrogram using the Fundamental Theorem of Algebra.

FROG constitutes a 2-dimensional phase retrieval problem, since it involves a two dimensional quantity (the spectrogram), for which only the amplitude is known. For our purposes, we take this all to mean: if we find *a* solution to our retrieval problem, then it is *the* solution.

Unfortunately, despite this strong uniqueness claim, the definition of a solution to our problem is qualitative at best. In brief, we will define a solution for an unknown pulse to be one which computes a spectrogram that *looks* very similar to the one we measure. The uniqueness statement above tells us that the pulse we use for the computation must be the one that produced our trace. Although the quality of a computed spectrogram is based on visual inspection, we will provide one quantitative metric for judging a solution when we get to our measurements in Chapter 4.

3.2.1 Retrieval Process

The statement of uniqueness can be misleading, if interpreted incorrectly. In order to understand it more clearly, consider the FROG field, rather than the spectrogram. The FROG field has a squared-magnitude that *is* the spectrogram. However, any number of FROG fields compute the spectrogram: just add to it an arbitrary phase, and it will still produce the correct spectrogram. Therefore, the uniqueness statement is really assurance that the FROG field has the correct phase. If it has the correct phase *and* the correct magnitude,

then it is uniquely the correct solution. Since the FROG field is computed directly from the probe and gate, a correct FROG field implies a correct probe and gate.

It is a trivial matter to compute a guess for the FROG field from a guess for the probe and gate. We will restate the steps below. On the other hand, it is a non-trivial task to do the reverse. In other words, it is trivial to multiply two functions together, but it is non-trivial to factor some arbitrary function into two functions which by definition must be related in some particular way.

That the two functions are constrained in some way will turn out to benefit the retrieval process. The basic steps are as follows:

- we will generate a guess for the pulse, either randomly or otherwise

- we will check the spectrogram computed from our guess against the trace that we measured experimentally

- if necessary, we will subject our guess to a sequence of computational steps known as the Iterative Fourier Transform Algorithm (IFTA) [1] [16]

- the algorithm will return a new guess, and we will iterate through the IFTA as needed.

We will now examine the complete process on a step by step basis.

First we need a guess to start the procedure, which we denote $E^{(1)}(t)$. In Section 3.2.4 we will figure out a way to come up with an educated guess, but here we assume that our guess is random.

Next, we form a signal field. The definition is repeated in Eq. (3.22):

$$E_{\text{sig}}(t, \tau) = E_p(t)E_g(t - \tau) . \tag{3.22}$$

Eq. (3.22) provides a direct computation for the signal field from our guess: $E_{\text{sig}}^{(1)}(t, \tau) = E_p^{(1)}(t)E_g^{(1)}(t - \tau)$. Interpreted differently, Eq. (3.22) *also* specifies that if we already have a signal field, as opposed to some other arbitrary function of t and τ, then it *must* factor into the right hand side of the above. If it does not factor, it is not a signal field. This requirement is called the **mathematical-form constraint**, and we will address it at a later step.

From the signal field, we form a FROG field via a Fourier transform $\tilde{E}^{(1)}_{\text{FROG}}(\omega, \tau) = \mathcal{F}\{E^{(1)}_{\text{sig}}(t, \tau)\}$. At this stage, we are able to check the spectrogram by taking the magnitude-squared, by definition. The quality of our guess is evaluated by visually comparing the computed spectrogram against the experimental trace. For all guesses except the correct one, the spectrogram will not be a match.

Assuming an incorrect match, we proceed to improve our guess. To do so, we can apply the **data constraint**. The statement of the data constraint is given by the definition of the spectrogram, which we repeat in Eq. (3.23):

$$\tilde{S}_{\text{FROG}}(\omega, \tau) = |\tilde{E}_{\text{FROG}}(\omega, \tau)|^2 \ . \tag{3.23}$$

Explicitly, the goal is to apply the data constraint onto the FROG field we previously computed. The expression for applying the data constraint is given by Eq. (3.24):

$$\tilde{E}^{'\,(1)}_{\text{FROG}}(\omega, \tau) = \tilde{E}^{`\,(1)}_{\text{FROG}}(\omega, \tau) \frac{\sqrt{\tilde{S}_{\text{FROG}}(\omega, \tau)}}{|\tilde{E}^{`\,(1)}_{\text{FROG}}(\omega, \tau)|} \ . \tag{3.24}$$

The prime symbol added to the left hand side implies that this quantity computes the correct spectrogram. In other words, it satisfies the data constraint given by Eq. (3.23). This can be seen by taking the magnitude squared of Eq. (3.24).

With the FROG field updated, it can be converted into a signal field with an inverse Fourier transform: $E^{'\,(1)}_{\text{sig}}(t, \tau) = \mathcal{F}^{-1}\{\tilde{E}^{'\,(1)}_{\text{FROG}}(\omega, \tau)\}$. What remains is a two dimensional function of t and τ, which satisfies the data constraint. We wish to know whether the function also satisfies the mathematical-form constraint.

Unfortunately, as mentioned above, this is a non-trivial task. Consider the case of a true signal field for SHG FROG:

$$E_{\text{sig}}(t, \tau) = E(t)E(t - \tau) \ . \tag{3.25}$$

It is simple to factor out of the left-hand side some function of t alone, call it $f(t)$, so that $E_{\text{sig}}(t, \tau) = f(t)g(t, \tau)$. The challenge is assuring that what remains satisfies the mathematical-form constraint, which for Eq. (3.25) would imply that $g(t, \tau)$ *must* take

the form $g(t, \tau) = f(t - \tau)$.

It is generally not possible to find an exact factorization that satisfies the mathematical-form constraint. In terms of our guess, then, this means that:

$$E_{\text{sig}}^{' \, (1)}(t, \tau) = E_p^{(2)}(t) E_g^{(2)}(t - \tau) , \qquad (3.26)$$

does not have a closed form solution for $E_p^{(2)}(t)$ and $E_g^{(2)}(t - \tau)$. The remainder of this section will be concerned with producing a good approximation to a solution.

In the process of attempting to satisfy the mathematical-form constraint, we will necessarily violate the data constraint. In reapplying the data constraint, we will necessarily violate the mathematical-form constraint. The goal of the IFTA is to iteratively apply these constraints until we converge on a solution that satisfies both, with as little error as possible.

In the next section, we will examine one scheme for applying the mathematical-form constraint. Here, we summarize the remainder of the algorithm in Eq. (3.27):

$$E_{\text{sig}}^{(k)}(t, \tau) = E_p^{(k)}(t) E_g^{(k)}(t - \tau)$$

$$\tilde{E}_{\text{CALC}}^{(k)}(\omega, \tau) = \mathcal{F}\{E_{\text{sig}}^{(k)}(t, \tau)\}$$

$$\tilde{E}_{\text{FROG}}^{' \, (k)}(\omega, \tau) = \tilde{E}_{\text{FROG}}^{(k)}(\omega, \tau) \cdot \frac{\sqrt{\tilde{S}_{\text{FROG}}^{(k)}(\omega, \tau)}}{|\tilde{E}_{\text{FROG}}^{(k)}(\omega, \tau)|} \qquad (3.27)$$

$$E_{\text{sig}}^{' \, (k)}(t, \tau) = \mathcal{F}^{-1}\{\tilde{E}_{\text{FROG}}^{' \, (k)}(\omega, \tau)\}$$

$$E^{(k+1)}(t) = E_{\text{sig}}^{' \, (t)}(t, \tau) \rightarrow \blacksquare .$$

Eq. (3.27) represents a single iteration of the IFTA. In the above, the superscript (k) corresponds to quantities that were computed from the kth guess, that is the kth iteration of the algorithm. An iteration ends with an updated guess $E^{(k+1)}(t)$ which satisfies the mathematical-form constraint, indicated by a black box in Eq. (3.27).

3.2.2 The Discrete Mathematical-Form Constraint

Early implementations of the IFTA produced solutions to the mathematical-form constraint using beam geometries other than SHG FROG. In the case of PG FROG—wherein the gate is given by $E_g(t - \tau) = |E(t - \tau)|^2$—the final step took the form of a simple integration over delay:

$$\int E(t)|E(t - \tau)|^2 \, \mathrm{d}\tau = E(t) \int |E(t - \tau)|^2 \, \mathrm{d}\tau \propto E(t) \, . \tag{3.28}$$

We have noted previously that a delay does not change intensity, so the integral evaluates to a constant. (This is not the case for SHG FROG, so we can not use this simple integral approach.) A later incarnation of the final step was based on minimizing the difference between $\tilde{S}_{\mathrm{FROG}}$ and $\tilde{S}_{\mathrm{CALC}}$, a procedure that has come to be known as a generalized-projections (GP) [18].

The Principal Component Generalized Projections Algorithm (PCGPA) [19] is a cousin of GP, but takes an altogether unique interpretation of the problem. It is predicated on the clever observation that a discrete FROG spectrogram is formed from the outer product of two discrete vectors. Emphasizing this point, we will find that in the ideal (noise-free) case, the spectrogram is formed from a *single* outer product, a fact which qualifies it as a rank-one matrix. More generally, we will show that any matrix of data can be formed as a superposition of outer products, in the worst case rank N, and it is a simple matter to determine which vectors form the constituent outer products. In other words, PCGPA frames the retrieval algorithm as an eigenvalue problem, asking the question: which two vectors are *mostly* responsible for forming a particular matrix? Equivalently it asks: what are the largest eigenvalues? We will find that by selecting eigenvectors corresponding to the largest eigenvalues, we can produce better and better guesses for the probe and gate (*i.e.* the pulse), until our guesses ultimately converge toward a solution.

To begin the derivation, we reiterate once more that in general, a FROG experiment produces an optical signal that is the product of a probe $E_p(t)$ and a gate $E_g(t - \tau)$. We examined two schemes for producing the nonlinear signal $E_{\mathrm{sig}}(t, \tau) = E_p(t)E_g(t - \tau)$ in physical terms in Section 2.2.3 (using collinear and non-collinear SHG). In this section, we

examine an alternative formulation of $E_{\text{sig}}(t, \tau)$ in strictly computational terms. It is this formulation that provides a basis for understanding PCGPA.

The conceptual transition from theory to computation requires a corresponding mathematical transition from continuous space to discrete space. In other words, we can only store a finite amount of information about a continuous function by means of a discrete complex vector, in our case of length N. In this paper we denote the discrete vector corresponding to the probe $E_p(t)$ using the notation:

$$
|p\rangle = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_N \end{bmatrix} .
\tag{3.29}
$$

The vector elements p_m are determined from the probe according to $p_m = E_p(t_m)$, with $m = 1, 2, \cdots, N$. Here $t_m = (-N/2 + m) \cdot \delta t$, and the time increment δt is determined by the resolution of the measurement. The discrete gate vector $|g\rangle$ corresponding to the un-delayed gate function $E_g(t)$ is defined similarly with elements $g_m = E_g(t_m)$.

To characterize the 2-dimensional quantity $E_g(t, \tau) = E_g(t - \tau)$ in a discrete fashion requires additional complexity. Since such a characterization must discretize both t and τ, we will need a distinct vector for each discrete value of delay. Matrix columns provide a particularly convenient bookkeeping strategy for storing these vectors. Before examining the form of the proposed matrix, we must address a subtle but crucial point: PCGPA requires that delaying the gate be executed with a *circular* shift of the elements of $|g\rangle$. MATLAB provides built-in functionality for a circular shift, but formally we can describe the discrete delay by $\tau_n = (1 - n) \cdot \delta t$, where some sort of modulo operator must be used to cycle through the indices. The formal details are mostly unimportant, except to say that this convention implies that if n denotes the column of our matrix, then the first column $n = 1$ should correspond to zero delay τ_1. Increasing the column index corresponds to a *more-negative* delay, that is an advance in time (left-shift on a delay axis), or an up-shift in the elements of a column vector.

Matters can be elucidated by writing out the matrix M_g corresponding to $E_g(t, \tau)$ as:

$$
M_g = \begin{matrix} & \overset{1}{\uparrow} & & \cdots & \overset{N-1}{\uparrow} & \\ \begin{bmatrix} g_1 & g_2 & \cdots & g_N \\ g_2 & g_3 & \cdots & g_1 \\ \vdots & \vdots & \ddots & \vdots \\ g_N & g_1 & \cdots & g_{N-1} \end{bmatrix} & \begin{matrix} t_1 \\ t_2 \\ \vdots \\ t_N \end{matrix} \\ \begin{matrix} \tau_1 & \tau_2 & \cdots & \tau_N \end{matrix} & \end{matrix}
\tag{3.30}
$$

Eq. (3.30) emphasizes that the mth row of M_g corresponds to a constant time $E_g(t_m - \tau)$, whereas the nth column corresponds to a constant delay $E_g(t - \tau_n)$. The arrows above the matrix indicate by how many elements each column vector formed from $|g\rangle$ must be circularly shifted to produce M_g. Finally, in order for M_g to more accurately describe $E_g(t, \tau)$ we require a re-ordering of the columns according to increasing delay. Although we perform this re-ordering in practice, we take the point for granted here, since it has no bearing on the development: we simply assume that we can reorder the columns at will[2].

Forming a matrix equivalent, M_{sig}, to the function $E_{\text{sig}}(t, \tau)$ is now a trivial task, involving element-wise multiplication of $|p\rangle$ with the columns of M_g. We can simply write this matrix as:

$$
M_{\text{sig}} = \begin{bmatrix} p_1 g_1 & p_1 g_2 & \cdots & p_1 g_N \\ p_2 g_2 & p_2 g_3 & \cdots & p_2 g_1 \\ \vdots & \vdots & \ddots & \vdots \\ p_N g_N & p_N g_1 & \cdots & p_N g_{N-1} \end{bmatrix}.
\tag{3.31}
$$

Eq. (3.31) says that M_{sig} stores every possible interaction between individual elements of the discrete vector $|p\rangle$ with those of $|g\rangle$. The rows of Eq. (3.31) are constant in time just as in Eq. (3.30), and the columns are similarly constant in delay.

There is another mathematical operation that stores each interaction between two vectors.

---

[2]The procedure is simple: in MATLAB a single line can be written, `fliplr(fftshift(Mg, 2))`. The inner function swaps the left and right halves of a matrix (the optional argument 2 indicates a swap over the 2$^{\text{nd}}$ dimension of the trace); this effectively places zero delay at the center of the trace. The outer function corrects the direction of time; this command is actually unnecessary for SHG FROG, since the trace is symmetric in delay.

The outer product of $|p\rangle$ and $|g^*\rangle$ is defined

$$|p\rangle\langle g^*| = \begin{bmatrix} p_1 g_1 & p_1 g_2 & \cdots & p_1 g_N \\ p_2 g_1 & p_2 g_2 & \cdots & p_2 g_N \\ \vdots & \vdots & \ddots & \vdots \\ p_N g_1 & p_N g_2 & \cdots & p_N g_N \end{bmatrix} \tag{3.32}$$

Note that for complex vectors $\langle v| \equiv |v\rangle^\dagger = |v\rangle^{T\,*}$ denotes the adjoint (transpose conjugate). Eq. (3.31) does not include any conjugated factors, a fact we have anticipated in advance by forming an outer product with the vector $\langle g^*| \equiv |g\rangle^T$. This additional conjugation is a feature that will need to be tracked closely in writing retrieval scripts.

Comparison of Eq. (3.32) with Eq. (3.31) reveals a correspondence between the outer product and the signal field. Being specific, if we already have two vectors $|p\rangle$ and $|g\rangle$, then we can form a discrete signal field by forming the outer product, and then shifting the rows of the resulting product to the left by the row number minus one. This is illustrated conceptually as:

$$|p\rangle\langle g^*| = \begin{bmatrix} p_1 g_1 & p_1 g_2 & \cdots & p_1 g_N \\ p_2 g_1 & p_2 g_2 & \cdots & p_2 g_N \\ \vdots & \vdots & \ddots & \vdots \\ p_N g_1 & p_N g_2 & \cdots & p_N g_N \end{bmatrix} \quad \begin{matrix} \Longleftrightarrow \\[6pt] 0 \leftarrow \\ 1 \leftarrow \\ \vdots \\ N-1 \leftarrow \\[6pt] \Longleftrightarrow \end{matrix} \begin{bmatrix} p_1 g_1 & p_1 g_2 & \cdots & p_1 g_N \\ p_2 g_2 & p_2 g_3 & \cdots & p_2 g_1 \\ \vdots & \vdots & \ddots & \vdots \\ p_N g_N & p_N g_1 & \cdots & p_N g_{N-1} \end{bmatrix} = M_{\mathrm{sig}}$$

$$\tag{3.33}$$

where the arrows indicate by how many elements the rows of $|p\rangle\langle g^*|$ must shift in order to produce M_{sig}. Note that a left/right shift leaves the factor p_m in each matrix element unaltered; the time t_m is invariant over the rows, but the delay is not, as only the gate is a function of delay. Since forming a FROG field from M_{sig} can be done simply by computing an FFT along each column, the salient result of this analysis is that up to some invertible algorithm (involving row shifting, column re-ordering, and FFT's), the outer product is

equivalent to a discrete description of the FROG field $\tilde{E}_{\mathrm{FROG}}(\omega, \tau)$.

In the context of PCGPA, then, the mathematical-form constraint can be written

$$M_{\mathrm{sig}} \sim |p\rangle\langle g^*| \ , \tag{3.34}$$

where the symbol $\sim$ in this context is used to denote equality up to some invertible algorithm. We will refer to (3.34) as the **discrete mathematical-form constraint**.

PCGPA turns this development on its head, by framing the pulse retrieval problem as follows: if we already have a FROG field, which two vectors form the corresponding outer product? The remainder of this section is devoted to answering this question. We will find that we can produce approximations to those two vectors, guesses for the probe and gate, via a simple matrix multiplication.

To summarize, in a single iteration of the IFTA we will transform the outer product of two guesses $|p\rangle$ and $|g\rangle$ into a FROG field, we will apply the data constraint, and then we will apply the inverse transformation. The resulting matrix will be called the "outer product form matrix" denoted O. Note that O produces a spectrogram with the correct image (*i.e.* magnitude, thanks to the data constraint), but itself has an arbitrary phase given by the initial guesses.Thus O is given by $|p\rangle$ and $|g\rangle$, and the program of PCGPA is a solution to the equation

$$O = |p'\rangle\langle g^{*\prime}| \tag{3.35}$$

for $|p'\rangle$ and $|g'\rangle$.

Unfortunately, by analogy with Eq. (3.26), it turns out that Eq. (3.35) also does not have a closed form solution. The next section is dedicated to finding a reasonable approximation to one.

3.2.3   Singular Value Decomposition

Here we state without derivation and for reference an important result from linear algebra:

Any $m \times n$ matrix O can be factored into the form $O = P\Sigma G^\dagger$, where P is $m \times m$ unitary, Σ is diagonal, and G is $n \times n$ unitary. This factorization is called a singular value decomposition (SVD).

We will say more about unitary and diagonal matrices in Appendix B. The SVD is completely equivalent to a linear combination of outer products. To quickly understand this equivalence, we explicitly show the case for a 2×2 matrix. Denoting the nth column of a general matrix X as $|x_n\rangle$, with elements ($x_{1n}, \quad x_{2n}, \quad \cdots, \quad x_{Nn}$), we can write:

$$O = P\Sigma G^\dagger = \begin{bmatrix} |p_1\rangle & |p_2\rangle \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \begin{bmatrix} |g_1\rangle & |g_2\rangle \end{bmatrix}^\dagger$$

$$= \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}^\dagger$$

$$= \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \begin{bmatrix} g_{11}^* & g_{21}^* \\ g_{12}^* & g_{22}^* \end{bmatrix}$$

$$= \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \begin{bmatrix} \sigma_1 g_{11}^* & \sigma_1 g_{21}^* \\ \sigma_2 g_{12}^* & \sigma_2 g_{22}^* \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_1 p_{11} g_{11}^* + \sigma_2 p_{12} g_{12}^* & \sigma_1 p_{11} g_{21}^* + \sigma_2 p_{12} g_{22}^* \\ \sigma_1 p_{21} g_{11}^* + \sigma_2 p_{22} g_{12}^* & \sigma_1 p_{21} g_{21}^* + \sigma_2 p_{22} g_{22}^* \end{bmatrix}$$

$$= \sigma_1 \begin{bmatrix} p_{11} g_{11}^* & p_{11} g_{21}^* \\ p_{21} g_{11}^* & p_{21} g_{21}^* \end{bmatrix} + \sigma_2 \begin{bmatrix} p_{12} g_{12}^* & p_{12} g_{22}^* \\ p_{22} g_{12}^* & p_{22} g_{22}^* \end{bmatrix}$$

$$= \sigma_1 |p_1\rangle\langle g_1| + \sigma_2 |p_2\rangle\langle g_2|$$

This result is easily generalized, such that any matrix O can be written as the linear combination:

$$O = \sum_n \sigma_n |p_n\rangle\langle g_n| \ . \tag{3.36}$$

where the scalars σ_n are called the singular values of O; they have a geometric significance

46

that we will not discuss.

In principle, the SVD provides a possible solution to the discrete mathmatical form constraint: simply take the outer product form matrix O, produced from arbitrary guesses for the probe and gate, and compute its SVD. If the guesses are correct, there will only be one nonzero term in the linear combination, insofar as we know that a FROG spectrogram is formed by a single outer product (via Eq. (3.34)).

Realistically, it is most likely that the arbitrary guesses will be incorrect, and the SVD will produce N nonzero terms in the linear combination. Nevertheless, if we interpret the singular values as weights in the linear combination, then we can select the term with the largest $|\sigma_n|$ and use the corresponding $|p_n\rangle$ and $|g_n\rangle$ as next guesses for the probe and gate. This amounts to saying that the outer product with the largest weight produces a spectrogram that *looks* most like the true spectrogram formed by O (compared to the other outer products in the sum). The vectors which formed the outer product with the greatest weight are called the principal components of O.

The approximate solutions provided by the principal components can be made into better solutions by projecting onto their outer product form matrix the data constraint. Then, we select the principal components of the updated matrix, and repeat. This is the description of the IFTA in the context of PCPGA.

MATLAB includes a routine for computing the SVD of an arbitrary matrix. The issue is one of computational cost: the SVD routine returns three $N \times N$ matrices where PCGPA requires only two $N \times 1$ vectors. For large N and a complicated spectrogram, this can result in very slow retrieval times that require many iterations of the SVD. Instead, an approximation for the principal components can be quickly computed with a simple matrix multiplication, in a routine called the "power method". We derive the method in Appendix B, but we state the result here, since it is the approach we will use in our retrieval algorithm. In the power method, the approximate solution to the discrete mathematical form constraint is provided by:

$$|p'\rangle \approx OO^\dagger |p\rangle \tag{3.37}$$

$$|g'\rangle \approx O^\dagger O |g\rangle \tag{3.38}$$

3.2.4   Generating a First Guess

For complicated pulses, as we expect to encounter for RQPM-based second-harmonic generation, convergence of the algorithm is not assured. Rather than making several attempts at a stubborn retrieval, we were able to improve our success rate by producing high-quality guesses to seed the first iteration of the IFTA. The approach we took was based on the Retrieved-Amplitude N-Grid Algorithmic (RANA) approach [7].

In order to produce a guess, we appeal to the frequency marginal. We showed in Section 3.1.1 that the frequency marginal of an SHG FROG trace is the autoconvolution of the power spectrum, up to a factor of 2π, which we neglect. That is, $\tilde{M}^{\text{SHG}}(\omega) = \tilde{S}(\omega) * S(\omega)$. Therefore in the time domain, we can describe the frequency marginal by:

$$\mathcal{F}^{-1}\{\tilde{M}^{\text{SHG}}(\omega)\} = \mathcal{F}^{-1}\{\tilde{S}(\omega) * \tilde{S}(\omega)\} = s^2(t) \ , \tag{3.39}$$

which means,

$$s(t) = \pm\sqrt{\mathcal{F}^{-1}\{\tilde{M}^{\text{SHG}}(\omega)\}} \ . \tag{3.40}$$

Here, $s(t)$ denotes the transform dual of the spectral *intensity*[3].

The RANA approach invokes the Paley-Wiener theorem, which for our purposes ensures that $s(t)$ and all of its derivatives are continuous. To understand why this is a powerful statement, note that computing $s(t)$ from Eq. (3.40) gives rise to an ambiguity about the sign of $s(t_0)$ at a particular time t_0: it is either the value corresponding to the positive square root or the negative one. By default, the square root operation in MATLAB returns vector elements that all have a positive real part, saying nothing about which elements correspond to which solution. Therefore, computing Eq. (3.39) alone does not produce a description

---

[3]Emphatically, not the spectrum, whose dual is simply $E(t)$.

of $s(t)$ and additional constraints are required to build one. The RANA approach forms a vector corresponding to $s(t)$ on an element-by-element basis by selecting from the two possibilities the one which produces the smoothest curve.

In discrete notation, we start with a vector of data $|p\rangle$ corresponding to the positive square-root in Eq. (3.40). Then we form a second vector $|m\rangle = -|p\rangle$ corresponding to the negative square-root. Lastly we form an additional vector $|s\rangle$ element-by-element; the elements of which are denoted by s_n.

The first few elements are set equal to the first elements of $|p\rangle$, denoted p_1, p_2, p_3. We could just as well have set them equal to the first elements of $|m\rangle$, as the first several elements of each vector are near zero anyways. Subsequent elements s_n are selected recursively from the $(n-1)^{\text{th}}$ element by checking which next element between p_n or m_n, is nearer in magnitude to the element s_{n-1}. As a measure of closeness, we mean that we are selecting the quantity with the smallest magnitude between the two options $\Delta_p = p_n - s_{n-1}$ and $\Delta_m = m_n - s_{n-1}$. This is our metric for ensuring continuity of $s(t)$ in a discrete space. If it happens that there is ambiguity in the next choice, i.e. if $\Delta_p \approx \Delta_m$, then we can impose continuity of the derivatives instead.

In the cited reference, the RANA approach was implemented by defining vectors computed from the difference between the above elements and checking for which vector the magnitude is smallest. This procedure was done up to a second derivative, with vectors whose elements are defined by:

$$
\begin{aligned}
\Delta_{p,n}^{(0)} &= p_n - s_{n-1} \\
\Delta_{p,n}^{(1)} &= [p_n - s_{n-1}] - [s_{n-1} - s_{n-2}] \\
\Delta_{p,n}^{(2)} &= \{[p_n - s_{n-1}] - [s_{n-1} - s_{n-2}]\} - \{[s_{n-1} - s_{n-2}] - [s_{n-2} - s_{n-3}]\} \\
\Delta_{m,n}^{(0)} &= m_n - s_{n-1} \\
\Delta_{m,n}^{(1)} &= [m_n - s_{n-1}] - [s_{n-1} - s_{n-2}] \\
\Delta_{m,n}^{(2)} &= \{[m_n - s_{n-1}] - [s_{n-1} - s_{n-2}]\} - \{[s_{n-1} - s_{n-2}] - [s_{n-2} - s_{n-3}]\} \ ,
\end{aligned}
\tag{3.41}
$$

Rather than checking each of these quantities individually, we followed an approach used

in the same reference, wherein the researchers selected s_n from the vector $|p\rangle$ or $|m\rangle$ that yielded the lesser value between the quantities:

$$\epsilon_p = c_0 \, |\Delta_{p,n}^{(0)}|^2 + c_1 \, |\Delta_{p,n}^{(1)}|^2 + c_2 \, |\Delta_{p,n}^{(2)}|^2$$
$$\epsilon_m = c_0 \, |\Delta_{m,n}^{(0)}|^2 + c_1 \, |\Delta_{m,n}^{(1)}|^2 + c_2 \, |\Delta_{m,n}^{(2)}|^2 \;, \tag{3.42}$$

where the weights c_0, c_1, c_2 were determined from a large number of empirical tests. The weights which produced convergence for 100% of trials corresponded to $c_0 = 0.09, c_1 = 0.425, c_2 = 1$. These are the same weights we used in our retrieval scripts, wherein we construct $s(t)$ using Eqs. (3.41) and (3.42).

Having constructed a discrete description of $s(t)$, we can take a Fourier transform to recover a description of $\tilde{S}(\omega)$. This means that we are able to compute the spectral intensity from the frequency marginal, that is, directly from the spectrogram. By adding a random phase to the computed intensity, we have a good first guess for the pulse spectrum, and thus a first guess for the pulse.

We can illustrate the RANA approach with the sample pulse we modeled in Chapter 2. Fig. 3-1 shows the spectrum of that pulse calculated from an FFT overlaid with the spectrum computed from the RANA approach. The agreement is not perfect, nor was there perfect agreement between the spectra that were presented in the original papers. However, the RANA approach clearly provides a reasonable estimate for the spectrum, which is convenient in that it is computed directly from the spectrogram with no additional measurements required.
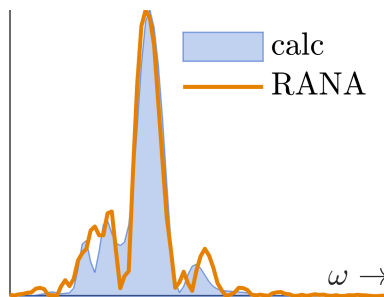

Figure 3-1: **Power spectrum computed from RANA approach:** The legend entry "calc" was computed by an FFT.

Strictly speaking, we did not find that the RANA approach was necessary for convergence, but it did improve our lot by producing a higher success rate. With a properly calibrated delay axis, convergence occurred on every attempt.

In Chapter 2 we discussed the physical theory that underlies our experiments, and in this chapter we outline the steps necessary to produce a description of the pulse from a measurement. In summary, we have provided a framework for performing an interferometric measurement and recovering a pulse. Chapter 4 reports on the results of measurements we made in this framework.

Chapter 4

Experiment

In this chapter we will describe the experimental implementation of the theoretical processes described in Chapters 2 and 3. We will state and analyze results from each of the measurements previously discussed, and at each step we will detail the associated difficulties. In Section 4.2 we will describe our initial experiments, and the issues faced in measuring the spectrum directly. Section 4.3 will be mostly oriented towards establishing a reliable description of the power spectrum. We will build an interferometric autocorrelation in Section 4.4, and from the trace we will make some predictions about the pulse's temporal intensity profile. The most comprehensive analysis will be provided by FROG in Section 4.5, where we will recover a mostly-complete description of the pulse in both time and frequency domains. The chapter concludes in Section 4.6 with an additional experiment to determine the sign of the spectral phase, as well as a brief catalog of the measurements we made over various input polarizations.

4.1   Optical Layout

The complete optical layout of our system is illustrated in Fig. 4-1. The system consists of an ultrafast laser source, the inner components of which were illustrated in Fig. 1-1. To review, the source produces pulses with a center wavelength of approximately 1.2 μm at a repetition rate of 80 MHz. Directly from the source, the pulses are directed into a two-lens telescope (Thorlabs LA1257-C and Thorlabs LA1805-AB) in order to narrow the beam

profile from a full-width-at-half-maximum (FWHM) of 6.8 mm to 2.6 mm in the transverse direction. The practical purpose of the telescope is to facilitate easier alignment, since the beam profile directly from the source is large and unwieldy for use with typical lab optics. In the complete experiment, the pulse also traverses an autocorrelator and a prism compressor. The latter two components are indicated to be separate blocks by the dashed lines in Fig. 4-1. The mirrors that direct the beam into each of these blocks are attached to flip-mounts, which allow for the beam to selectively bypass either block during alignment, and during other phases of the experiment. The flip-mounts are indicated by curved double arrows.
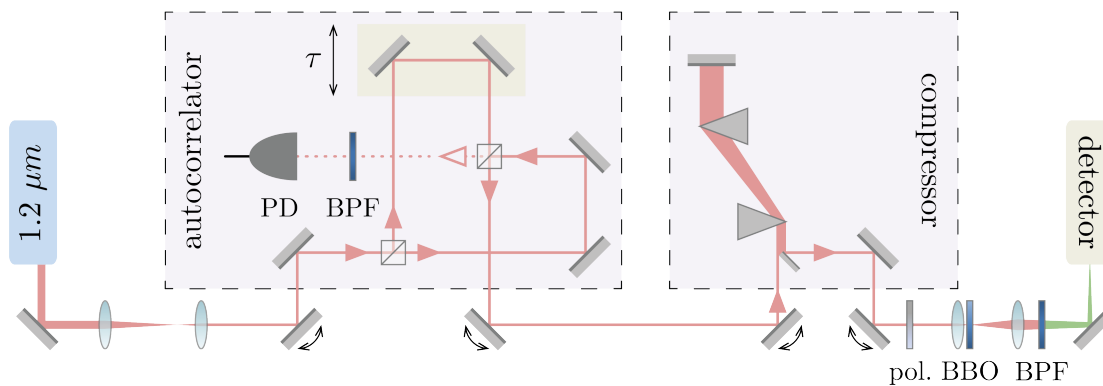

Figure 4-1: **Complete optical layout:** Through the use of flip mirrors (illustrated by double curved arrows), we were able to bypass the interferometer and compressor as needed. The detector is left vague in this figure, since it can be replaced by either a photodiode or a spectrometer. This has the nice feature that we can easily make the transition from one measurement to another, and it means that measurements were recorded at the same position.

The autocorrelator consists of an interferometer in a Mach-Zehnder configuration with a photodiode (PD) (Thorlabs PDA100A2) at the secondary output port. The beamsplitters in the interferometer are both designed to impart low group delay dispersion (GDD) over a spectral range from 1000-2000 nm (Thorlabs UFBS50502). The PD is used to dynamically calibrate the delay stage in a process that we will describe in Section 4.3. The mirrors in the signal arm of the interferometer are mounted onto a delay stage controlled by a nanopositioner with strain gauge, operated in "open-loop" mode (Newport NPX200SG and NPC3SG).

The stage operates by driving a piezoelectric (PZT) element inside the controller with an

electrical signal from a function generator. The PZT is a crystalline material that expands and contracts with applied voltage; the expansion and contraction has the effect of moving the mirrors in the signal arm of the interferometer, which shortens and lengthens the beam path periodically. The long duration of our pulse produces similarly long autocorrelation traces, which required that we drive the stage at the maximum supported voltage, in order to produce maximum displacement of the mirrors and maximum delay extent. For our autocorrelations we used a 10 Vpp triangle wave at a low frequency 15 mHz driving frequency. The low frequency sweep produced traces with high resolution. The sweep configuration for FROG was different, and we will address its operation in Section 4.5. The controller includes a monitor signal produced by a strain gauge, which acts as an indicator of the voltage across the PZT and thus of the position of the stage along its sweep. The monitor voltage was used in our first attempt at producing a delay axis, which was a troublesome procedure that we will discuss shortly.

The compressor consists of two fused silica prisms designed for dispersion compensation in ultrafast lasers (Thorlabs AFS-FS). The operation of a prism compressor is such that the angular dispersion produced by the first prism can be a source of negative GDD, which can compensate the positive GDD introduced by material dispersion. The second prism acts to realign all frequency components to a similar propagation direction. A mirror at the back end of the compressor makes the prism pair equivalent to a four prism sequence, the output of which is nearly collinear with the input beam path. The output signal of the compressor is slightly misaligned with respect to the input, so that a pick-off mirror can be used to redirect it toward a different path.

The remainder of the system consists of a polarizer, a 10 μm thick crystal of β-barium borate (BBO) crystal with a strong second-order nonlinear optical response[1], a colored-glass bandpass filter (BPF), and a detector. The bandpass filter (Thorlabs FGS600) is designed to pass signals in the wavelength range from 335 to 665 nm, thus enabling us to select the second-harmonic output of the nonlinear crystal, as is necessary in a collinear arrangement. We can easily remove the BPF to observe the fundamental component in an FTIR measurement.

---

[1]We do not say much about the theory of nonlinear optics in this paper. For our purposes, the BBO is a material that, when illuminated by $E(t)$, produces two outputs $E(t)$ and $E^2(t)$.

Additionally, the setup includes a 15-mm focal length N-BK7 lens (Thorlabs LA1540-C) in front of the BBO, which serves to tightly focus all of the optical power into the crystal. This tight focusing enables us to produce a maximum SHG signal. A second lens (Thorlabs LA1805-AB) is positioned shortly behind the crystal in order refocus the output onto a detector. A kinematic-mounted mirror is the final optical component, and is used as a steering mirror to guide the signal onto the detector. The system is configured in such a way that we can easily exchange detectors between a photodiode and a spectrometer, depending on the measurement.

4.2   Initial Measurements

Before launching our pulse into the autocorrelator, we made some simple measurements in a direct path from the source to a detector, in this case a spectrometer (Fig. 4-2). The only transmissive optical elements in this measurement are the two lenses that form the telescope, a polarizer, and the two lenses that refocus the optical power onto the detector.

Figure 4-2: **Direct-path layout:** Compared to the complete layout in Fig. 4-1, all mirror flip-mounts were in the down position, and we removed the nonlinear crystal and bandpass filter. The latter two elements were reinserted for the measurement of the SHG spectrum.

We began by making a simple measurement of the pulse spectrum. The spectrometer we used was a Blue-Wave Miniature Series model with measurement wavelengths in the 200-1150 nm range, manufactured by StellarNet Inc.. The spectrometer was calibrated to an integration time of 50 ms with smoothing disabled. The fundamental pulse spectrum from our initial measurement is plotted in Fig. 4-3a.

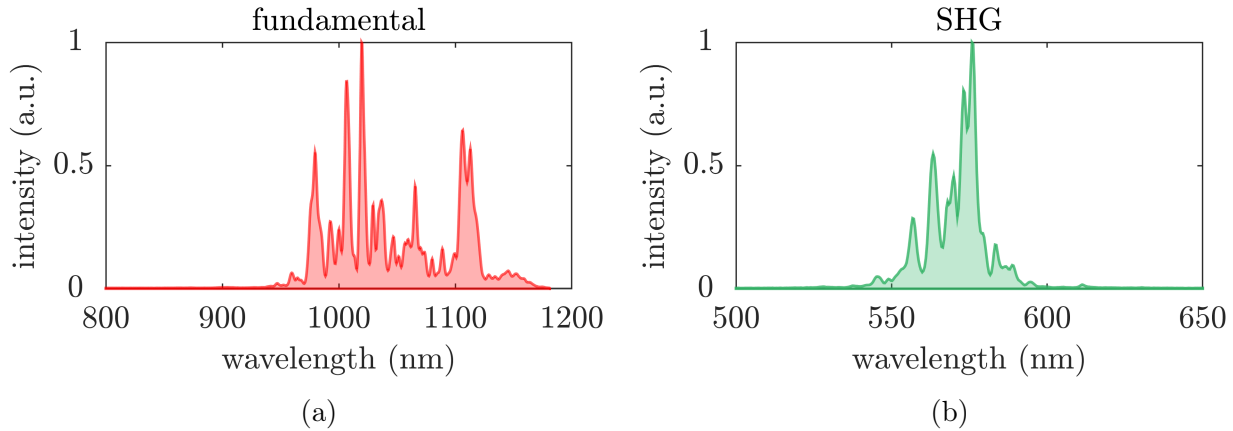Reinserting the BBO crystal and a BPF into the beam path gave us a direct measurement

Figure 4-3: **Directly measured fundamental and SHG power spectra:** The fundamental spectrum is shown in 4-3a. This signal appears finely structured, but we found that this measurement was unrepeatable, in that the spectrum we ultimately recovered with FROG extends past the bandwidth of the spectrometer. The SHG spectrum is shown in 4-3b. The latter signal was reliable and its measurement easily repeated.

of the second harmonic spectrum. The second harmonic signal was substantially weaker than the fundamental, but we were able to easily observe it on the spectrometer with a slight increase in integration time. This spectrum is plotted in Fig. 4-3b

Although we were able to observe a strong signal, we quickly faced a number of difficulties in recording a reliable fundamental spectrum. First, we observed a substantial change in the shape of the spectrum as we translated the entrance slit of the spectrometer in a direction transverse to the beam path. The effect is illustrated in Fig. 4-4, which shows variation in the spectral intensity as the alignment of the spectrometer is slightly altered over 6 random positions in the beam profile. We first interpreted this result to suggest that the beam exhibited a high degree of spatial chirp, a feature where certain frequency components were more prominent than others at different positions in the transverse direction.

Although we later determined that this issue was more likely a fault of the spectrometer bandwidth, we continued to attribute other measurement difficulties to spatial chirp, such as an inability to produce total destructive interference in our FTIR (Section 4.3). In any case, our final determination was that our spectrometer had insufficient bandwidth (with a cutoff of 1150 nm) to accurately produce the fundamental pulse spectrum. On the other hand, since the second-harmonic power spectrum is located near the center of the spectrometer measurement range, we remained confident that we produced the SHG spectrum accurately

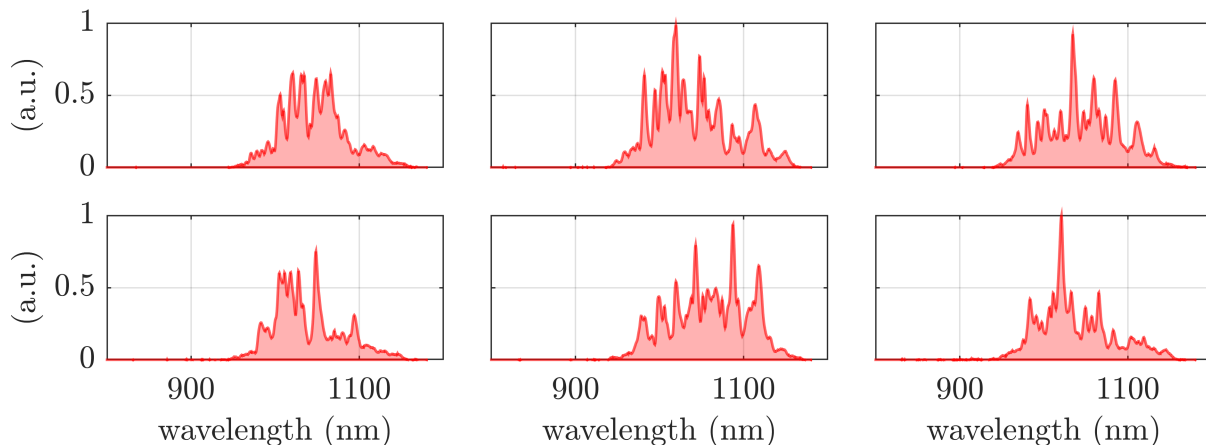and that our device was sufficient to record a FROG measurement.

Figure 4-4: **Spatially-varying power spectrum:** We observed a variation in the observed power spectrum for six random positions of the spectrometer entrance slit.

The second unusual effect we observed was a dramatic change in the spectral intensity with the input polarization. We will look more at the results of our polarization experiments in Section 4.6. Neither of these effects were predicted, nor have we been able to characterize them completely. Although the spatial effects mentioned above were not observed in the second-harmonic spectrum, the unusual polarization effects persisted in the SHG signal.

4.3   FTIR

FTIR is the simplest interferometric measurement we made, in that it comprises a linear system. However, despite the simplicity of the interferometer, our complex pulses introduced compilcations to these measurements. Our first conclusive observation of this complexity was the form of the FTIR trace.
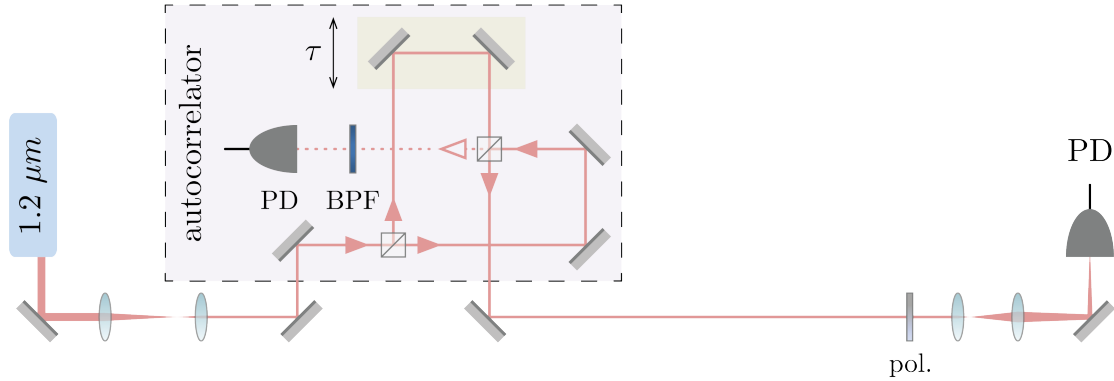
Figure 4-5: **FTIR layout**

Although we drove our controller with the maximum peak-to-peak voltage supported by the PZT, we observed that the signal was not fully contained by our sweep. In other words, the full autocorrelation trace could not be covered by the full range of our stage. This was not a catastrophic failure of our system however, since from signal processing we know that extending the range of a time domain signal has the effect only of increasing resolution in the spectral domain. This means that while clipping the trace may be a source of lost definition in our power spectrum, all of the most important structural features are recorded in the center of the trace. Therefore, the range of our stage still produced a good approximation to the power spectrum via an FFT.

Our interferometer was aligned so that the center of the autocorrelation corresponded to the center of the monitor voltage sweep. The FTIR trace and monitor voltage are plotted in Fig. 4-6. At a glance, the trace is even more structured than the multiple-peaked pulse we modeled in Section 2.2.2, implying fine detail in the pulse intensity profile.
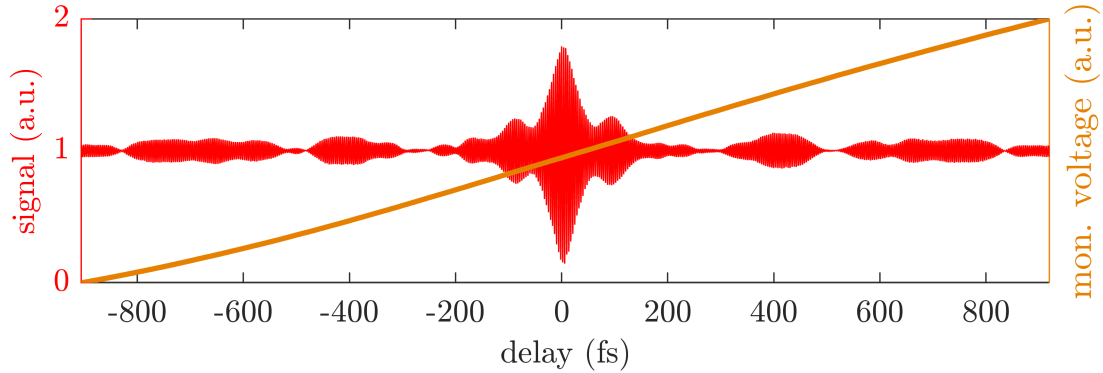
Figure 4-6: **FTIR signal**

The FTIR-recovered power spectrum is plotted in Fig. 4-7. These two images provide us our first insight into the complex structure of the pulse.
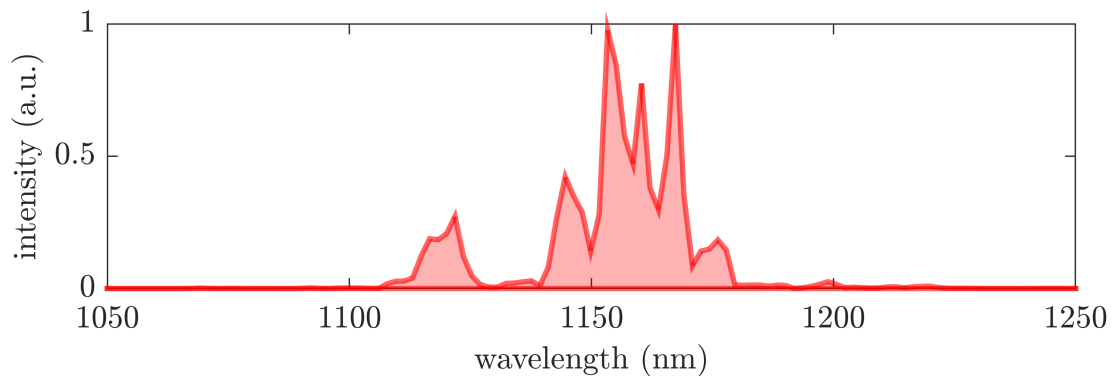

Figure 4-7: **FTIR-recovered power spectrum**

Before proceeding, we should note that initially we were not able to reliably produce the spectrum in Fig. 4-7. In hindsight, we have determined that this issue was a result of improper delay calibration. Being specific, we first modeled a very simple delay axis by fitting a line to the controller monitor voltage and doing a unit conversion (Fig. 4-8a). The conversion factor was estimated from the controller data sheet to be $L = 200 \ \mu m/10$ V. The result of the conversion gave us a displacement axis that approximately corresponded to the physical position of the signal mirrors. These were converted to delay via another unit conversion given by $2L/c$, where c is the speed of light and the factor of two accounts for propagation into and out of the signal arm (see Fig. 4-5).

The FTIR trace was plotted along this delay axis (uncalibrated) in Fig. 4-8b. Overlaid

59

is precisely the same data plotted on a delay axis that was formed with dynamic calibration (see below).
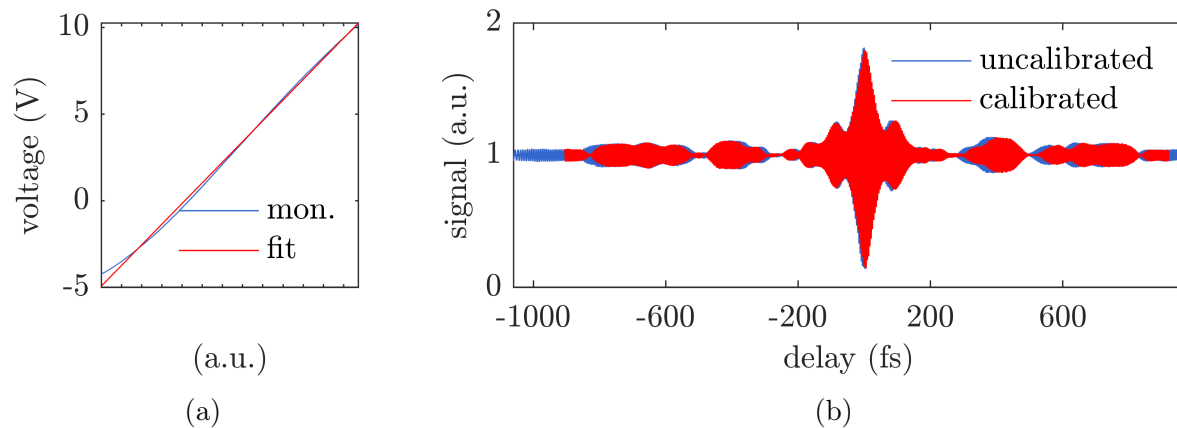

(a)

(b)

Figure 4-8: **The effects of delay axis calibration:** 4-8a shows a linear fit to the controller monitor voltage, which is easily changed into a delay axis via two conversion factors (see main text).

The two traces in Fig. 4-8 appear very similar. Nonetheless, we find significant discrepancies between the spectra recovered from the uncalibrated and calibrated delay axes, as shown in Fig. 4-9. In hindsight, we have determined that the monitor signal indicates non-uniform motion of the stage over its maximum and minimum displacement. We solved this problem using dynamic calibration.
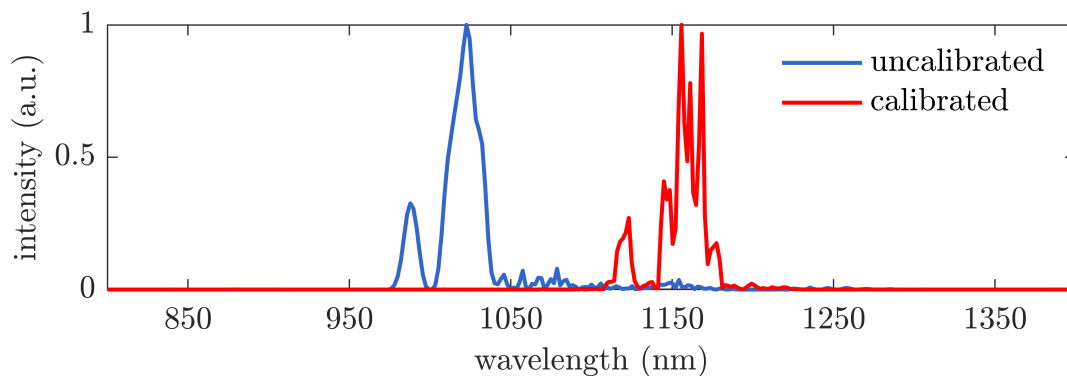

Figure 4-9: **Delay axis calibration and the power spectrum**

The procedure for dynamic calibration involves using the secondary output port of the interferometer to measure mirror displacement via an optical signal. By transmitting the secondary output signal through a narrow-bandpass filter, we can produce a "reference"

60

trace that approximates one that would be produced by cw inputs (Fig. 2.15): a narrower spectrum looks more like an impulse, and thus corresponds to a broader temporal extent. Furthermore, over a narrow spectral band the time-domain oscillations in the trace should be relatively unmodulated, that is, they should exhibit a stable wavelength λ corresponding to the center wavelength of the BPF.

The filtered signal can be used to calibrate a delay axis by subtracting the mean value of the signal and finding the zero crossings of the recorded data. Since the trace should cross zero every $\lambda/2$ units, this provides a precise measurement of the motion of the delay stage synchronized with the measurement of the primary output.

For our dynamic calibration we used a BPF with a center wavelength of 1100 nm and a FWHM estimated at 10 nm (Thorlabs FBH1100-10). The signal was recorded on a silicon photodetector (Thorlabs PDA100A2). The resulting trace is plotted in Fig. 4-10a.
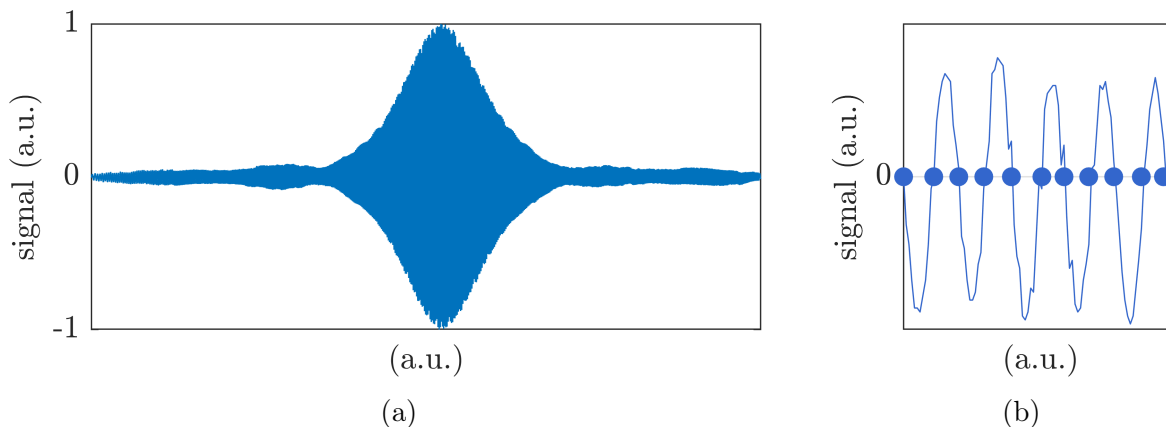

(a) (b)

Figure 4-10: **Dyanamic calibration signal:** 4-10a shows the signal observed at the secondary output of the interferometer. Despite the narrow BPF, the trace still decays to zero at the edges, motivating the need for an even narrower filter. Fig. 4-10b shows the computed zero crossings for a small portion of the signal. Each scatter point corresponds to a vector element where the signal crosses zero, which for a narrow band signal occurs at a regular spacing of $\lambda/2$ throughout the entire duration of the trace (unmodulated). These crossings give a direct measurement of the motion of the delay stage, and can be used to construct a delay axis.

Dynamic calibration improved the quality and repeatability of our measurements substantially, although further improvements could be made by using an even narrower BPF. Ideally, the BPF would be narrow enough to produce a trace that appears to be cw over the entire

extent of the delay sweep.

Another problem related to delay calibration was a determination of the correct center wavelength of the measured spectrum. This issue was first observed when varying parameters in the processing of the raw FTIR data. To understand this problem, note that when the control stage reaches a maximum or minimum in its sweep, distortions may arise from the sudden change of direction. As a solution, we clipped the trace at the edges to approximate a more linear sweep, at the cost of spectral resolution (see discussion above). Depending on the amount of clipping we chose, we observed an unexpected shift in the center frequency of the spectrum.

We resolved this ambiguity by computing a center frequency from the independently measured second-harmonic spectrum that we examined in section 4.2. By interpreting the spectrum as a probability density function and computing its mean via Eq. (A.2), we determined a second-harmonic center wavelength. This was easily converted to a fundamental wavelength, which we used in aligning the spectra for all subsequent analysis in this thesis. Fig. 4-11 illustrates the shift in center frequency as a function of trace clipping, with and without correction. In the corrected panel, all spectra share a common center wavelength, and a change in resolution is observed by increasing the extent of the trace, as expected.
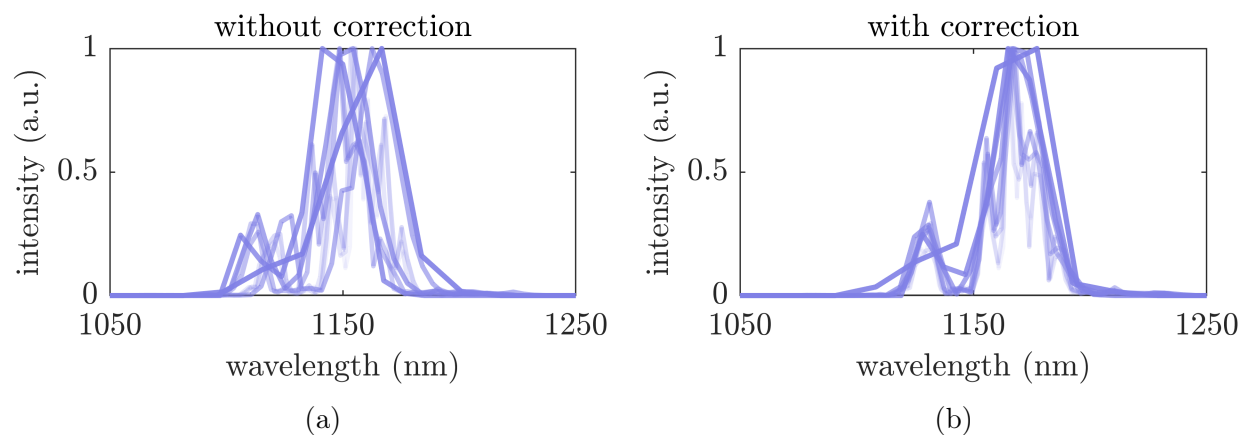

Figure 4-11: **FTIR-recovered power spectrum with changing resolution:** In 4-11a we observe an unexpected shift in the center frequency, depending on the exact window size our FTIR trace Also observed is a change in the trace resolution, which was expected. The change is resolution is repeated in 4-11b, but in that figure each spectrum was positioned to a common center frequency.

In order to further compensate the limited range of the delay stage, we used a micrometer in the reference arm of the interferometer to slightly adjust the position of the reference mirrors. This had the effect of altering the center of the autocorrelation trace with respect to the monitor voltage. We recorded two autocorrelations shifted to either side of center. In MATLAB we pieced together an extended trace with a wider temporal extent, using the peak of each as a reference point for stitching together the two signals. The extended trace is plotted in Fig. 4-12a, with the spectrum shown in Fig. 4-12b. The spectrum is overlaid with the original spectrum that was recovered with the clipped trace. While the temporal trace in Fig. 4-12a appears seamless, we can observe from the red shaded region in Fig. 4-12b a small disagreement between the two spectra. We attribute this disagreement to remaining calibration issues.
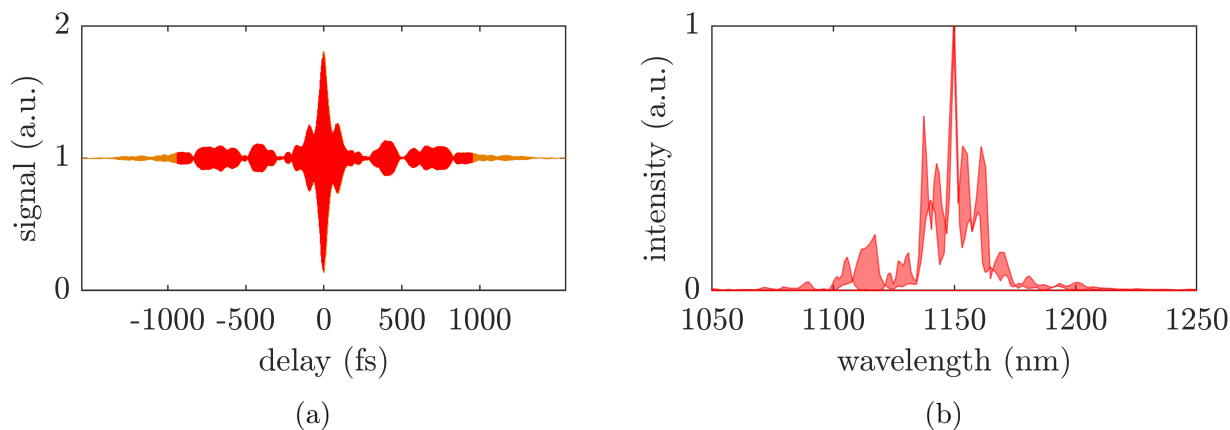

Figure 4-12: **Extended FTIR signal and power spectrum**

With further improvements to our calibration method, we may be able to produce better agreement between the spectra in Fig. 4-12b, and therefore a better overall measurement of the pulse spectrum. Nevertheless, the FTIR method provided a good sense of the spectral structure of our pulse. We will compare the spectrum recorded in this section to a power spectrum recovered by FROG in Section 4.5. The FTIR proved to the most reliable measurement of the fundamental pulse spectrum apart from FROG.

Our confidence in this technique was bolstered by comparing the spectra recovered by an FTIR measured on two different photodiodes. A photodiode is characterized by its responsivity—or the amount of photocurrent it can produce per unit of input optical power—

as a function of wavelength. The responsivity can be interpreted as a measure of the transfer function of the photodiode, or its filtering characteristics. The photodetectors used in the experiment described here are based on silicon (Si) (Thorlabs PDA10A2) and indium gallium arsenide (InGaAs) (Thorlabs PDA10D2). The silicon detector has a responsivity that decays in the neighborhood of 1100 nm, near the center of our pulse spectrum. On the other hand, the InGaAs detector has a fairly flat responsivity over the entire pulse spectral region.

The responsivity of these two detectors was resolved by recording two FTIR traces, and recovering the pulse spectrum from each. The responsivity raw data was taken directly from Thorlabs, and is plotted as patched shaded regions in Fig. 4-13. The InGaAs-recovered spectrum was shifted to a center wavelength that agreed with the independently measured SHG spectrum, in the same method that was previously described. Due to the persistent calibration issues, additional measurements were required in order to align the Si-recovered spectrum, including re-shifting and appropriate normalization. Note that the results of this experiment are not precise, but they are a good indicator of the quality of our FTIR retrievals.
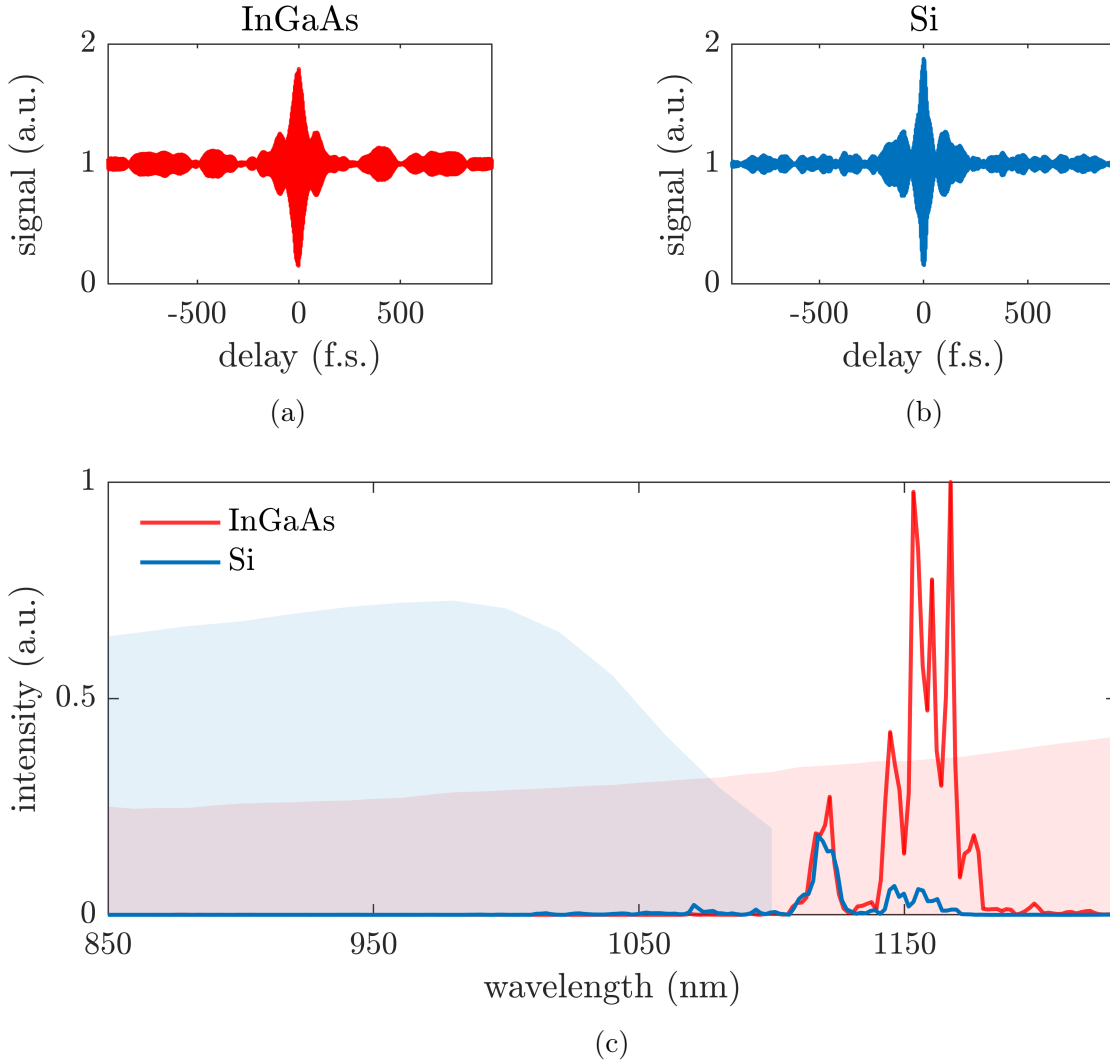
Figure 4-13: **Detector responsivity resolved with FTIR:** A trace was recorded on a Si detector (blue) and an InGaAs detector (red). The light shaded regions indicate the responsivity of each detector that was used to record the solid trace of the corresponding color. Notice the Si responsivity decays in the neighborhood of the pulse spectrum, and so the signal detected with Si becomes attenuated over those wavelengths. (The data provided by Thorlabs does not extend completely into the spectral region of interest, but the trend is clear.)

4.4 Interferometric Autocorrelation

In principle, the experimental transition from FTIR to interferometric autocorrelation involves simply reinserting the BBO and BPF. In practice however, we found that the SHG signal that was produced in the nonlinear element was too weak to overcome the electronic
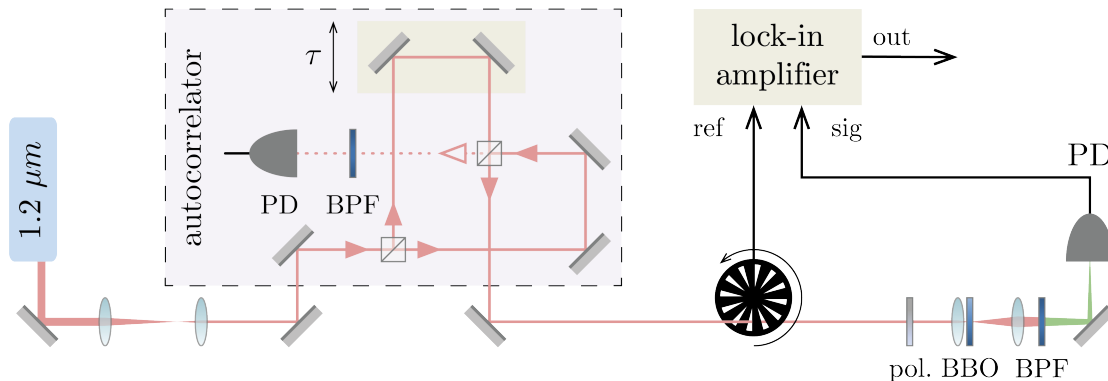
noise in the photodiode/oscilloscope circuit.

Figure 4-14: **IAC layout**

We overcame this problem by employing a technique common in optics involving a chopper and a lock-in amplifier (LIA), illustrated in Fig. 4-14. In short, the chopper modulates the beam so that the PD detects a square wave whose frequency f_R is related to the chopper speed, and whose amplitude is the signal of interest. An electrical signal oscillating at f_R produced by the chopper controller is used as the reference for a lock-in amplifier, and the PD signal is used as the input. The LIA mixes the two inputs. Since the optical signal was modulated by the chopper, it too has a fundamental component at f_R, and so it gets downconverted to baseband. With lowpass filtering and amplification, the signal overcomes the noise, and a measurement of the weak SHG signal can be made. The output of the amplifier forms our interferometric autocorrelation trace, which we plot in Fig. 4-15. This trace was produced by a Stanford Research Sytems SR830 set to a time-constant of 30 μs, a filter rolloff of 6 dB per decade, and a sensitivity of 200 mV. The amplifier was also set to low-noise mode, and notch filters at the line frequency and twice the line frequency were selected.
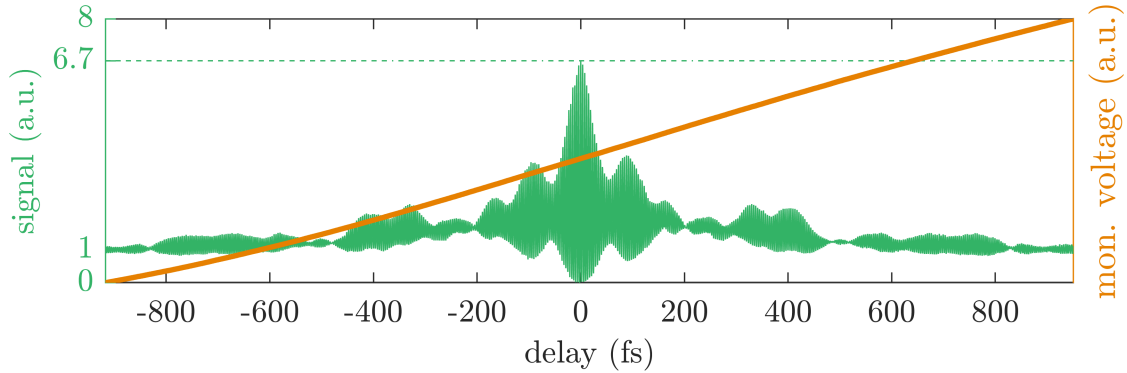
Figure 4-15: **Interferometric autocorrelation signal:** The signal shows an erroneous peak to background ratio of 6.7. From Section 2.2.2 we expected a ratio of 8. This immediately gave us concerns as to the quality of our measurement. Ultimately this proved to be a simple matter of clipping, where the extent of the trace was insufficient to allow it to decay to the true background.

In Fig. we extended the interferometric autocorrelation trace to a wider extent, in the same way that we did with FTIR. This revealed that the peak to background ratio of 6.7 that we observed in Fig. 4-15 was erroneous, since the true background was off scale. Increasing the extent improved the ratio to 7.5. Fig. 4-16b shows the SHG spectrum recovered from IAC against the same trace measured directly on a spectrometer. The IAC trace was shifted so that its center wavelength aligned with that of the spectrometer measurement, and it was normalized differently to highlight the areas of agreement. Although the main structure is in agreement between the two measurements, we observe that this was not the most accurate way to reproduce an SHG spectrum. We might have predicted this from our Section 2.2.2 analysis—specifically from Fig. 2-12 we see that the SHG component of the IAC spectrum is dwarfed by the baseband component and the modulated first harmonic component. That is, the signal to noise ration (SNR) is very low. FROG, by contrast, recovers the pulse from the baseband term (with a much higher SNR), and so we will find that the SHG spectrum *it* produces will agree much more closely with the spectrum produced by a spectrometer (Fig. 4-24).
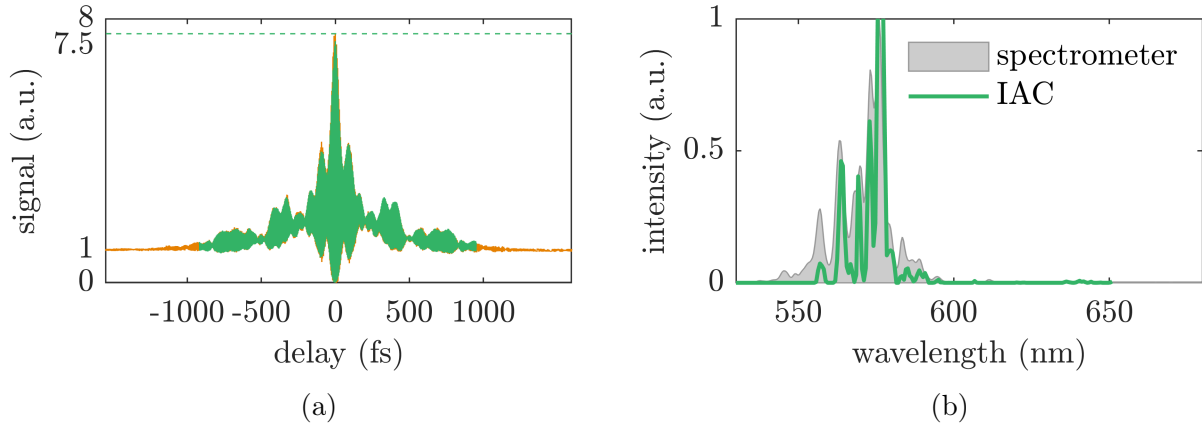
67

(a)

(b)

Figure 4-16: **Extended IAC trace with recovered SHG spectrum:** In 4-16a we see that the peak to background ratio improves to 7.5 (compared to value of 6.7 in Fig. 4-15) after we extended the trace to include the true background. In 4-16b we compare the SHG spectrum recovered from the second-harmonic band of the IAC trace with the spectrum measured on a spectrometer. The IAC recovery was shifted in wavelength and normalized to a peak of 1.5 (off scale) in order to improve the agreement. Better delay calibration may improve the need for these adjustments.

4.5 FROG

With the autocorrelator and nonlinear optics already positioned, the experimental transition to FROG involves simply replacing the photodiode with a spectrometer (and bypassing the lock-in circuit). The experiment is illustrated in Fig. 4-17.
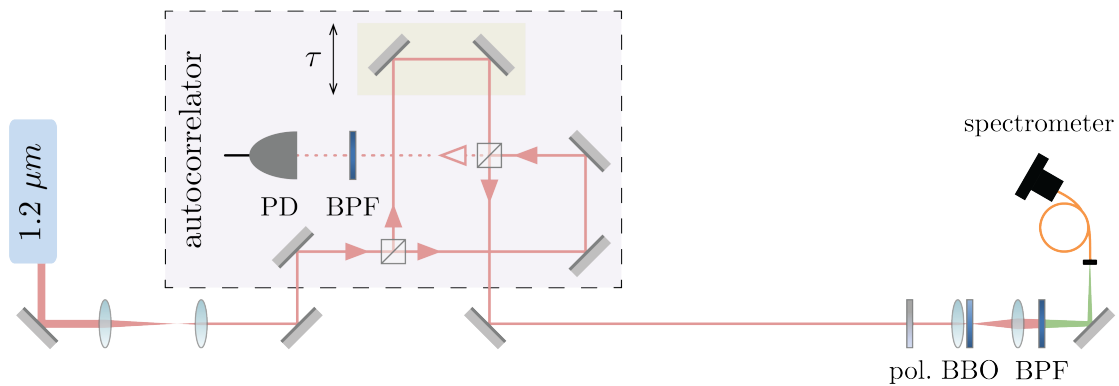

Figure 4-17: **FROG layout**

The ideal way to record a spectrogram involves driving the controller with the same

parameters as used in our other measurements, that is, with a function generator. The spectrometer should be automated to record the trace as the function generator produces a smooth sweep over the delay stage. Such methods of recording a spectrogram "on-the-fly" were studied in preliminary experiments. Unfortunately, the instruments in our setup required us to incrementally shift the delay stage, wait for the spectrometer to record a spectrum, and then repeat. The incremental motion of the delay stage can be unstable, and we observed ringing in the dynamic reference signal. This also caused recording times to be slowed dramatically, especially when recovering pulses from weak signals, which require an increased integration time. Common recording times in our setup were around 20 minutes. Finally, the incremental motion of the stage caused further problems in delay axis calibration for FROG. Ultimately this problem was resolved by finding a retrieval that converged successfully with a dynamically calibrated axis and using that axis as a calibration for all subsequent retrievals.

In order to build the system for recording a spectrogram, we first wrote a script in MATLAB to interface with both the controller and the spectrometer. Once communication was established, we set up a `for` loop to drive the controller with a linear sweep and to record a spectrum for each driving voltage. Several failures were logged in the first attempts, which were caused by user error due to improper calibration of the spectrometer. Calibration parameters were not many, but included options for varying the integration time as well as a default for "smoothing" the spectrum, which should be disabled.

With the spectrometer properly calibrated, we were able to successfully record a spectrogram. A good recording is plotted in Fig. 4-18a. This trace was recorded over 3500 delay points at an integration time of 150 ms per delay point, averaged over 2 spectra for every delay. In total, the raw trace had dimensions 2048×3500, which is very large for a FROG retrieval[2]. Fortunately, the SHG spectrum is fairly narrow compared to the bandwidth of the spectrometer (350 - 1150 nm), so there were many redundant zeros in the trace. This allowed us to crop the trace to a smaller size before attempting retrieval.

---

[2]Simple Gaussian pulses were reported to be retrieved successfully on traces as small as 64×64 in the early days of FROG!
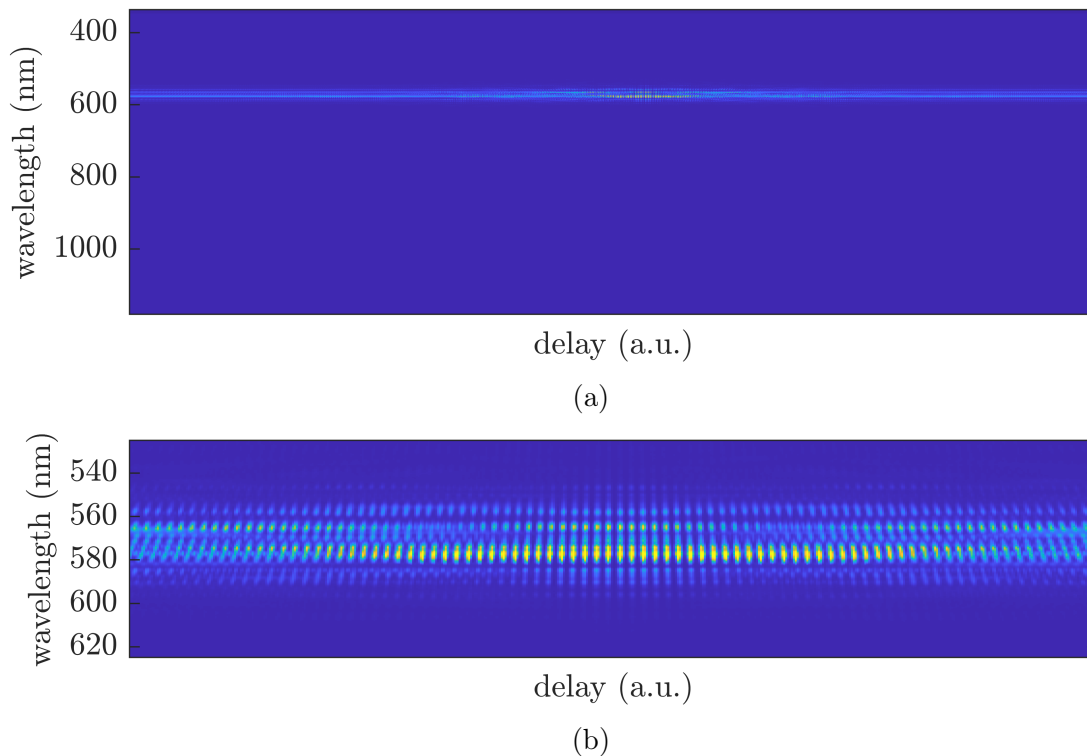
(a)

(b)

Figure 4-18: **Raw interferometric FROG spectrogram:** Interference fringes are observed in 4-18b, as expected from our analysis in Section 2.2.2.

A detailed view of the interesting portion of the raw spectrogram is shown in Fig. 4-18b. Interference fringes are clearly observed, and are in good agreement with the form of the theoretical trace we predicted from Fig. 2-14.

The first true sanity check on our measurement is provided by the delay marginal. Integrating the spectrogram over the vertical dimension, we recovered a close match for the IAC trace we recorded in Fig. 4-15. These results are plotted with a relative offset in Fig. 4-19. Their agreement gave us confidence that the trace was recorded successfully.
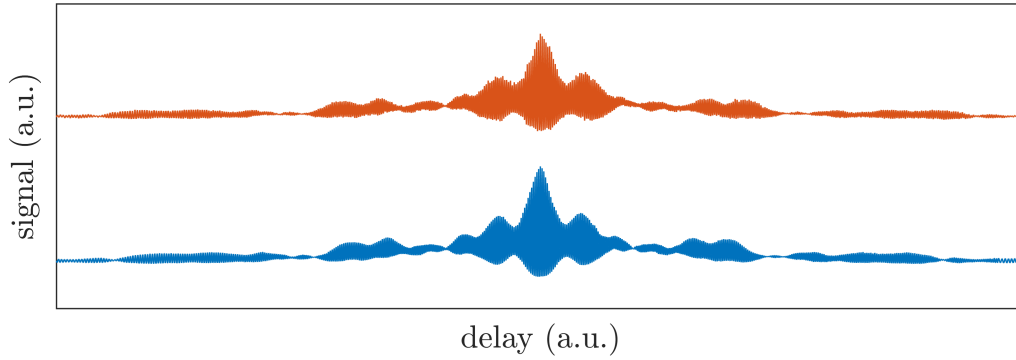
Figure 4-19: **Interferometric FROG delay marginal**

In order to recover the SHG FROG trace, we filtered the raw spectrogram according to the methods we discussed in Section 2.2.2. Taking an FFT over each row of the raw trace gave the 2-dimensional spectrum as a function of wavelength and delay-frequency shown in Fig. 4-20a. We plotted the square root of the spectrum in order to accentuate less intense regions of the spectrum. In this figure we clearly observe the baseband, fundamental, and second harmonic spectral bands. By selecting the baseband with a supergaussian filter applied over each row, and then inverse Fourier transforming back to the wavelength-delay domain, we recovered the image shown in Fig. 4-20b. Finally, we subtracted the background term which can be taken either from the first or last column of the trace; this is the component at large delay corresponding to $\tilde{S}_B(\omega)$ in Eq. (2.35). This gave us our first glimpse of the SHG FROG trace contained within our measurement (Fig. 4-20c).
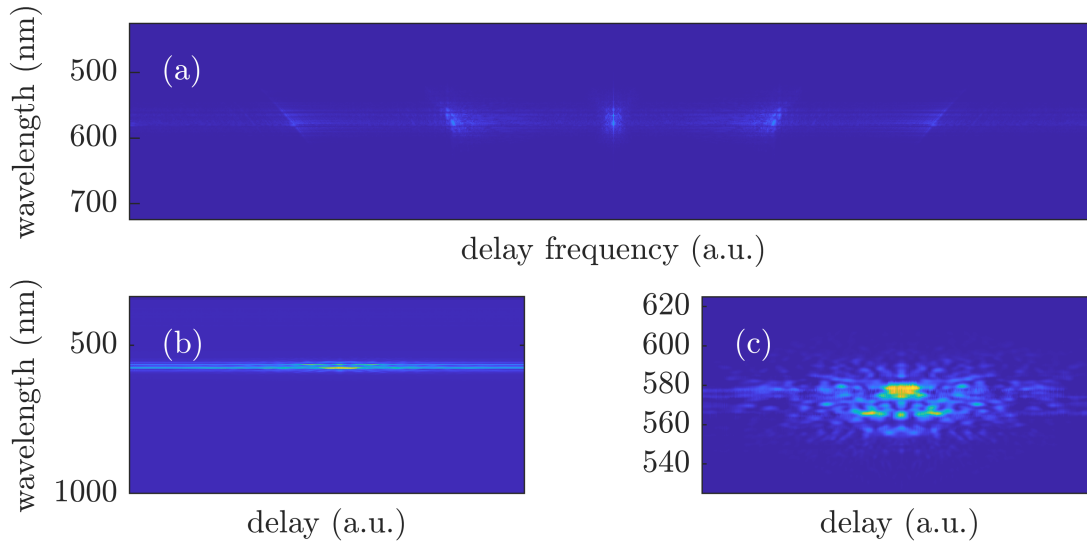
Figure 4-20: **SHG FROG trace recovered from raw spectrogram:** In 4-20a, the spectrogram was Fourier transformed over the delay direction. The baseband was isolated, and then an inverse Fourier transform over delay produced 4-20b. To recover the SHG FROG trace in 4-20c required subtraction of the final column of the resulting trace from each column. The orientation of the vertical axis was also reversed to produce the image in 4-20c, which has been zoomed in to focus on the interesting portion of the trace.

4.5.1   Pulse Measurement Results

With the spectrogram processed, we began our retrieval attempts. The final trace size was 1024×1024 pixels. The delay direction was padded with zeros so that the fractions of the trace in either direction filled by nonzero data were approximately equal. After zero padding, the trace was interpolated over delay to a square trace. In order to produce a good guess for the pulse, we implemented the RANA approach, discussed in Section 3.2.4. The algorithm ran for 100 iterations on the full size trace, and converged in ~ 20 seconds. The retrieval was considered successful since the guess for the pulse produced a FROG trace that was in close agreement with the measured trace. This is the most reliable method for evaluating a retrieval, though it is a qualitative measure. The quantitative metric of success is stated in Section 4.5.2.

The recovered pulse temporal and spectral intensity profiles are plotted in Fig. 4-21, along with the temporal and spectral phase. The FWHM of the pulse intensity and the root-mean-square (rms) width were in exact agreement at 125 fs. The rms width of the spectrum was

measured to be 28 nm about a center wavelength of 1141 nm. This center wavelength is shifted from the center of the main portion of the spectrum due to a significant second lobe centered around 1120 nm. The rms width in both cases was computed from Eq. (A.3). The center frequency was determined from the mean value of an independently measured SHG spectrum using Eq. (A.2).
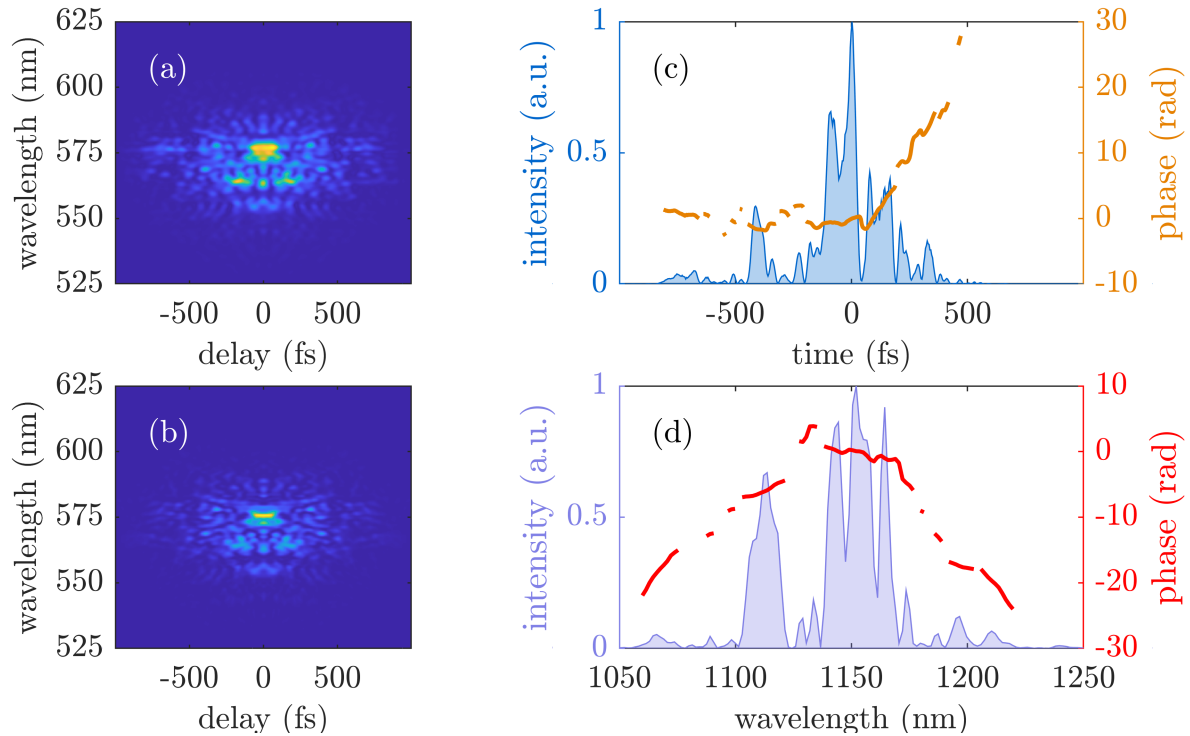
Figure 4-21: **Retrieval results:** 4-21a shows the measured SHG FROG trace, and the recovered trace is plotted in 4-21b. The pulse temporal intensity profile and phase are plotted in 4-21c. The spectral intensity and phase are illustrated in 4-21d.

The spectral intensity computed in the RANA approach was in fair agreement with the spectral intensity computed from the FTIR and from FROG, although we do note that the agreement was not perfect (Fig. 4-22). In the paper that introduced the approach, similar disagreements between their spectra were noted [6].

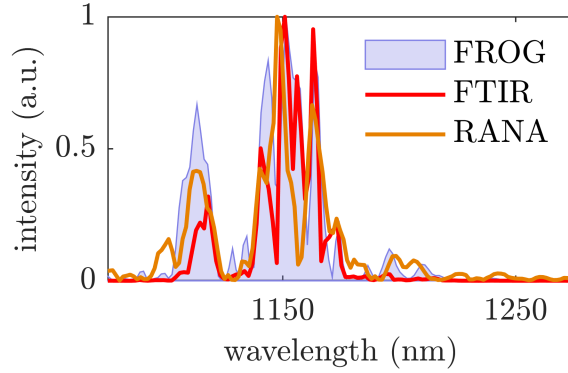Figure 4-22: **Comparison of FROG, FTIR, and RANA**

4.5.2 Error Checks

The success of any FROG algorithm to accurately recover a pulse is characterized by an error metric usually called the 'G error' [6]. Other names are the 'FROG error' and the 'rms error'. This metric, as the latter names suggests, is the rms difference between the spectrogram computed from the current guess $\tilde{S}_{\text{CALC}}^{(k)}(\omega, \tau)$ and the measured spectrogram $\tilde{S}_{\text{FROG}}(\omega, \tau)$. The difference is computed on a pixel by pixel basis for each iteration of the algorithm. It is described by:

$$G^{(k)} = \sqrt{\frac{1}{N^2} \sum_{m=1}^{N} \sum_{n=1}^{N} \left| \tilde{S}_{\text{FROG}}(\omega_m, \tau_n) - \tilde{S}_{\text{CALC}}(\omega_m, \tau_n) \right|^2}, \qquad (4.1)$$

where (ω_m, τ_n) denotes the value in the mth row and the nth column of the discrete trace.

The magnitude of the G error on its own is meaningless. However, compared to the value between iterations, the G error provides a quantitative metric for evaluating the success of the retrieval. The G error should always be reported with the size of the trace.

Our retrieval produced a minimum G error of 0.013 on a 1024×1024 matrix. A plot of the G error representative of this retrieval is shown in Fig. 4-23.
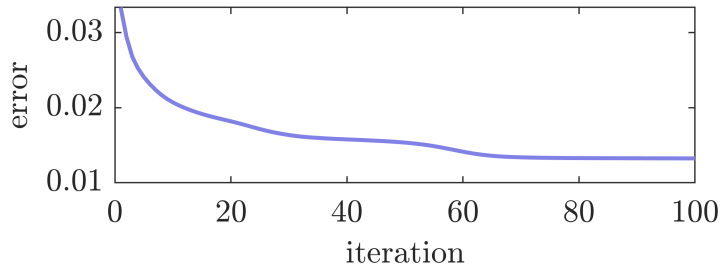
Figure 4-23: G **error**.

There are other checks we can perform to see if our recovered pulse is consistent with our other measurements. First, we find that the second-harmonic spectrum computed from the FROG recovery agrees very well with our spectrometer reading. This comparison is plotted in Fig. 4-24.
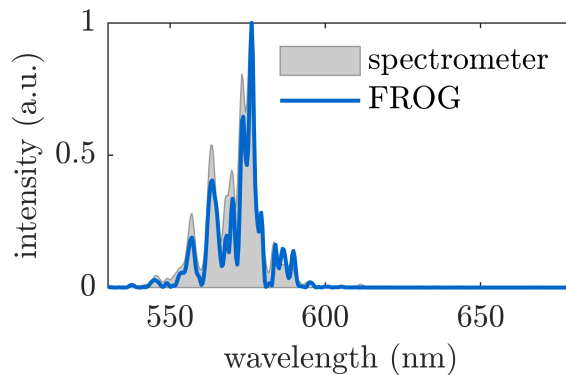
Figure 4-24: **SHG power spectrum computed from FROG:** A directly measured SHG power spectrum is overlaid with the same quantity computed from the FROG retrieval.

Next, we computed the autoconvolution of the power spectrum, and we compared the result with the frequency marginal of our SHG FROG trace $\tilde{M}(\lambda)$, in accordance with Eq. (3.20). This autoconvolution was performed on spectra recovered via 3 different methods, including the RANA approach, FROG, and FTIR. Agreement was good for the first two, and poor for the latter, as shown in Fig. 4-25. The good agreement is expected since each of the quantities under comparison were computed from the same data. The poor disagreement in the FTIR autoconvolution may be related to the delay calibration issues noted in Section 4.3.

good agreement — poor agreement

(a) (b)

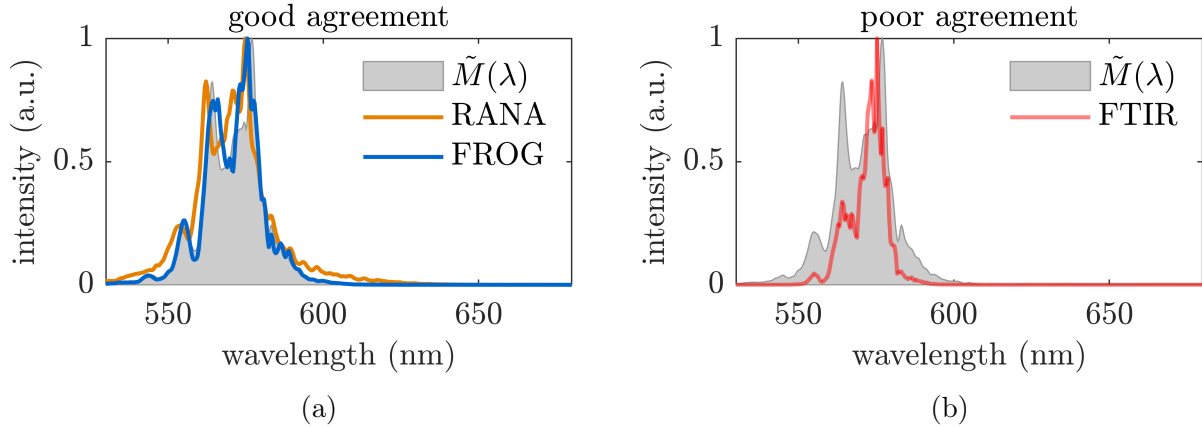Figure 4-25: **SHG FROG frequency marginal:** Overlaid is $\tilde{S}(\omega) * \tilde{S}(\omega)$ (plotted over wavelength), where $\tilde{S}(\omega)$ was computed both from the retrieved pulse, and from the RANA approach. In 4-25a we observe good agreement between the frequency marginal for $\tilde{S} * \tilde{S}$ and the result from FROG and RANA. This agreement is not particularly enthralling, however, since these measurements are self-referential in the sense that they were each computed from the raw spectrogram: the recovered pulse creates the correct spectrogram, so of course it creates the correct frequency marginal. The true test is a check against an independently measured spectrum. 4-25b shows the same computation using the power spectrum recovered from an FTIR, measured in an entirely different system. The disagreement is troubling; however we note that the main double-peak feature is observed in both measurements.

Lastly, we repeat the comparison in Fig. 4-22 for only FROG and FTIR. Agreement between these two quantities is arguably the best indicator of the overall quality of our system: FROG is nonlinear, FTIR is linear; FROG is recorded on a spectrometer, and FTIR is recorded on a photodiode. They form two distinct methods. In Fig. 4-26 we see good agreement between the power spectrum that each measurement produced. As previously noted, the most significant disagreement is in the relative size of the large "satellite" lobe in the spectral region around 1120 nm. The size of this lobe varied from measurement to measurement in the course of our experiments. This phenomenon remains unexplained.
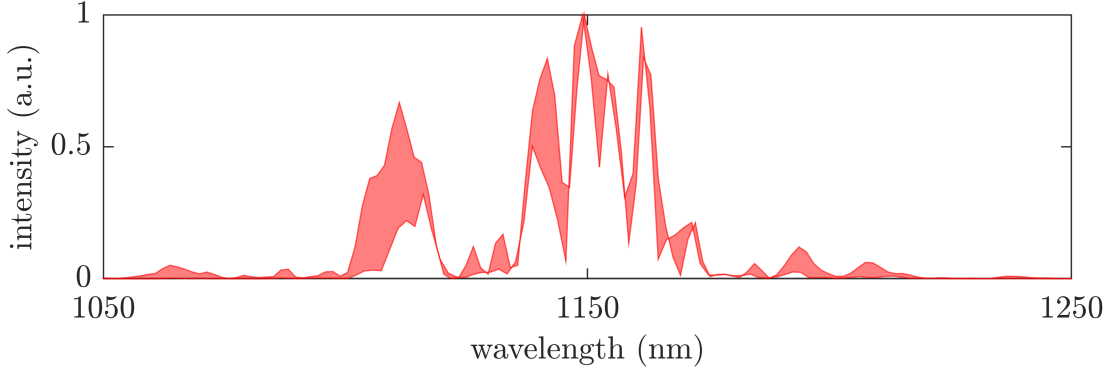
Figure 4-26: **FROG and FTIR comparison** - The red shaded region indicated the difference between the two distinct measurements we made, and thus acts a a measure of uncertainty in the precise shape of the spectrum.

4.6   Additional Measurements

4.6.1   SHG FROG Ambiguities

SHG FROG returns a complete characterization of an ultrafast pulse only up to a number of trivial ambiguities. Before addressing these ambiguities, we should address a subtle point about notation. In the FROG literature it is common to define the relationship between the real field and the complex field as $\mathcal{E}(t) = \text{Re}\{E(t)\,e^{i\omega_0 t}\}$, so that $E(t)$ is the envelope function [1]. We deviated from this convention in our previous analysis because it was a suitable way to proceed from pure electromagnetics. This may seem problematic to our formation of the spectrogram, until we note that in our notation the spectrogram can be written by inserting our convention for an ultrafast pulse Eq. (2.18) into the expression for the SHG FROG trace given by Eq. (3.2):

$$\tilde{S}^{\text{SHG}}_{\text{FROG}}(\omega, \tau) = \left| \int A(t)\, e^{i\omega_0 t} A(t - \tau)\, e^{i\omega_0(t-\tau)}\, e^{-i\omega t}\, \mathrm{d}t \right|^2 = \left| \int A(t) A(t - \tau)\, e^{-i\Delta\omega t}\, \mathrm{d}t \right|^2,$$

(4.2)

where $\Delta\omega = \omega - 2\omega_0$ and a factor of $e^{-i\omega_0 \tau}$ was pushed through the integral and erased by the magnitude-squared. The form of the spectrogram in Eq. (4.2) is equivalent to the definition provided by Eq. (3.2) up to some frequency offset determined by the pulse carrier frequency.

This means that FROG cannot resolve the carrier frequency in the first place, and it returns only the envelope function of the pulse. In other words, our original notation for the trace is equivalent to the conventional notation. For the remainder of this discussion we assume the conventional notation wherein $E(t)$ is the pulse envelope returned by the algorithm.

It is important that we noted this distinction in order to understand a FROG ambiguity that is inherent to SHG FROG, called the "direction-of-time" ambiguity, which results from the fact that the SHG FROG trace is symmetric in delay. We can derive this fact by noting:

$$\tilde{S}_{\text{FROG}}^{\text{SHG}}(\omega, \tau) = \left| \int E(t) E(t - \tau) \, e^{-i\omega t} \, dt \right|^2 = \left| \int E(t' + \tau) E(t') \, e^{-i\omega t'} \, dt' \right|^2 = \tilde{S}_{\text{FROG}}^{\text{SHG}}(\omega, -\tau)$$

(4.3)

A successful FROG retrieval should produce an accurate description of the real field $\mathcal{E}(t)$. The direction-of-time ambiguity tells us that even if a retrieval attempt converges, we cannot be certain whether our recovery describes $\mathcal{E}(t)$ or $\mathcal{E}(-t)$. In terms of the pulse envelope—which is the quantity returned by the algorithm—this is *also* an ambiguity in the sign of the temporal phase. This can be seen by expanding $E(t)$ (which is complex-valued) as $E(t) = |E(t)| \, e^{i\angle E(t)}$. Then by definition we have:

$$\mathcal{E}(t) = \text{Re}\{E(t) \, e^{i\omega_0 t}\} = \frac{|E(t)|}{2} \left[e^{i(\omega_0 t + \angle E(t))} + e^{-i(\omega_0 t + \angle E(t))} \right] \; ,$$

(4.4)

which implies:

$$\mathcal{E}(-t) = \frac{|E(-t)|}{2} \left[e^{-i(\omega_0 t - \angle E(-t))} + e^{i(\omega_0 t - \angle E(-t))} \right] = \text{Re}\{E^*(-t) \, e^{i\omega_0 t}\} \; .$$

(4.5)

Thus, the direction-of-time ambiguity between $\mathcal{E}(t)$ and $\mathcal{E}(-t)$ is equivalently written as an ambiguity between $E(t)$ and $E^*(-t)$. The additional conjugation is an important detail that must be tracked when performing computational tests on the pulse "as-recovered" or time-flipped. We also mention that in the frequency domain an ambiguous direction-of-time corresponds to an ambiguous sign-of-phase. This follows immediately from the basic properties of the Fourier transform, which tell us that $\mathcal{F}\{E^*(-t)\} = \tilde{E}^*(\omega)$.

In order to more completely characterize our pulse, and to make predictions about its

compressibility, we resolved the direction-of-time ambiguity with an additional measurement. Specifically, we were able to imprint onto our pulse the predicted phase shift from a glass sample. We chose for our sample 5 mm of zinc-selenide (ZnSe), which has a group delay dispersion of around 2500 fs^2 near our center wavelength of 1141 nm.

The procedure begins by adding to the spectrum of our recovered pulse $\tilde{E}_1(\omega)$ the phase shift from ZnSe denoted φ_{ZnSe}, which was computed from $\varphi(\omega) = \omega L \, n(\omega)/c$. Here the refractive index $n(\omega)$ was computed from the Sellmeier equation, c is the speed of light in vacuum, and L is the thickness of the sample. The spectrum of the broadened pulse is denoted by $\tilde{E}_b(\omega) = \tilde{E}_0(\omega) \, e^{-i\varphi(\omega)}$. Separately, we place a true 5-mm thick sample of ZnSe into our beam path, and we perform another FROG retrieval. The spectrum of the recovered pulse in this case is denoted $\tilde{E}_2(\omega)$. The temporal equivalents of these spectra are denoted $E_1(t)$, $E_b(t)$, and $E_2(t)$.

Individually, both $E_1(t)$ and $E_2(t)$ are ambiguous about their direction, for a total of four possible permutations. Together, the ambiguity can be resolved by comparing the profile of $E_b(t)$ to that of $E_2(t)$ for each permutation of $\mathcal{E}_1(\pm t)$ and $\mathcal{E}_2(\pm t)$. Agreement exists for only one permutation, as shown in Fig. 4-27. This agreement indicates the correct direction of time for both retrievals.
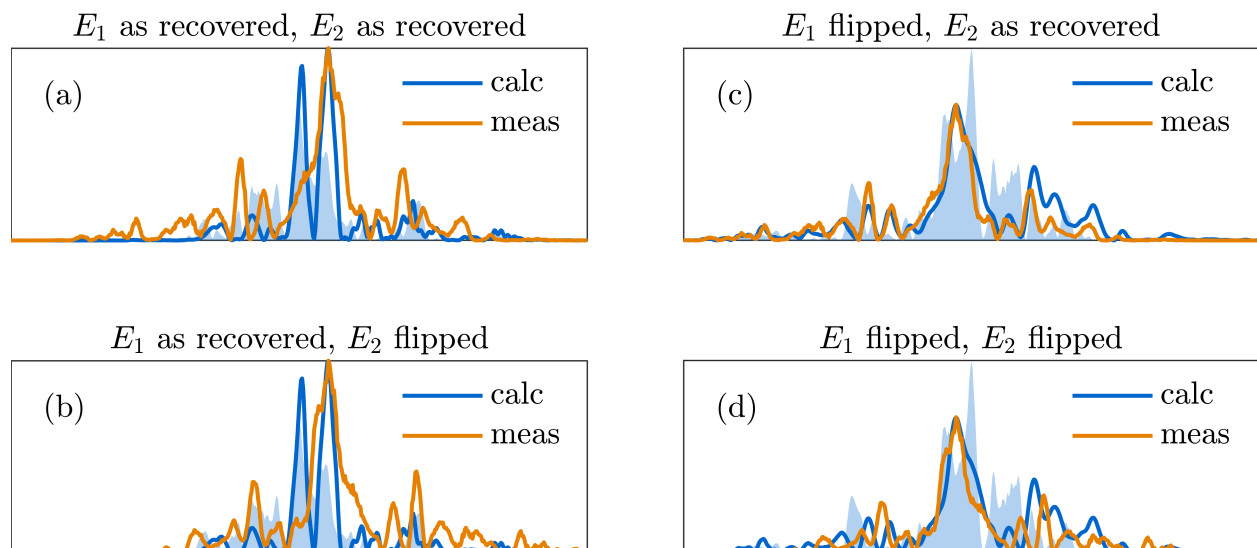

Figure 4-27: **Direction-of-time ambiguity:** Four permutations are plotted corresponding to possible directions for the recovered pulse and a separately recovered pulse after passage through 5 mm of ZnSe.

The resolved ambiguity is illustrated in the frequency domain in Fig. 4-28a. Note that our pulse has a complicated intensity profile where the intensity nears zero more than once. The phase is not defined when the intensity goes to zero, so FROG cannot resolve the value of the phase in those regions. This makes sense intuitively by noting that FROG cannot resolve *any* relative phase between the probe and gate (or a carrier-envelope offset), which follows from the fact that a global phase factor $e^{i\varphi_0}$ pushes through the Fourier transform operator, and gets deleted by recording the intensity, *i.e.* $|\mathcal{F}\{E(t)\,e^{i\varphi_0}\}|^2 = |\mathcal{F}\{E(t)\}e^{i\varphi_0}|^2 = |\mathcal{F}\{E(t)\}|^2$. Incidentally, this required us to zero all spectral phase plots in this thesis at the center of the spectrum, for better comparisons. The upshot here is that the shape of the spectral phase in each nonzero region is recovered accurately by FROG, but due to this so-called zeroth-order phase ambiguity, the relative values of the phases between distinct spectral regions are not known. Therefore the shape of $\angle\tilde{E}_2(\omega)$ agrees with $\angle\tilde{E}_b(\omega)$ only up to some offset which might vary between spectral regions. This same ambiguity is illustrated by plotting the phase shift from ZnSe (minus the linear phase) along with the *difference* in the phases between the two retrieved pulses.

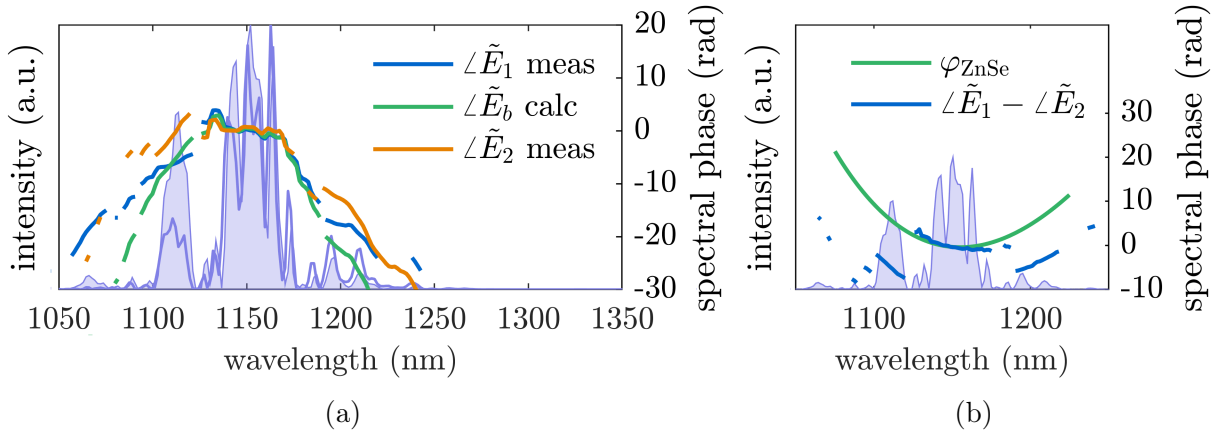Figure 4-28: **Sign-of-phase ambiguity:** This is a spectral representation of the direction-of-time ambiguity. "meas" denotes a quantity recovered by FROG, and "calc" denotes a prediction. See main text for definitions.

4.6.2 Polarization Effects

The RQPM pulse we measured had a complex spatial profile. Indications of this complexity were first noted by observing the beam profile on a near-infrared camera. We previously

noted in this thesis that a source of complexity in the beam profile might be related to spatial chirp.

We observed another unexpected effect in measurements related to the polarization of the beam. Specifically, we observed that changing the polarization direction of the beam significantly altered each of our measurements. We recorded FTIR, IAC, and FROG traces for beam polarizations between -90 and 90 degrees from vertical, where vertical was defined to be the direction normal to the optical table, and the positive direction was defined by a right hand rule. Measurements were taken at increments of 10 degrees.

The polarization effect was first observed in the FTIR trace. The fact that the trace is dependent on polarization suggests that the pulse spectrum varies with polarization. This in turn implies that the pulse temporal profile differs among the polarization components within the beam. Results of the FTIR measurements are plotted in Fig. 4-29.
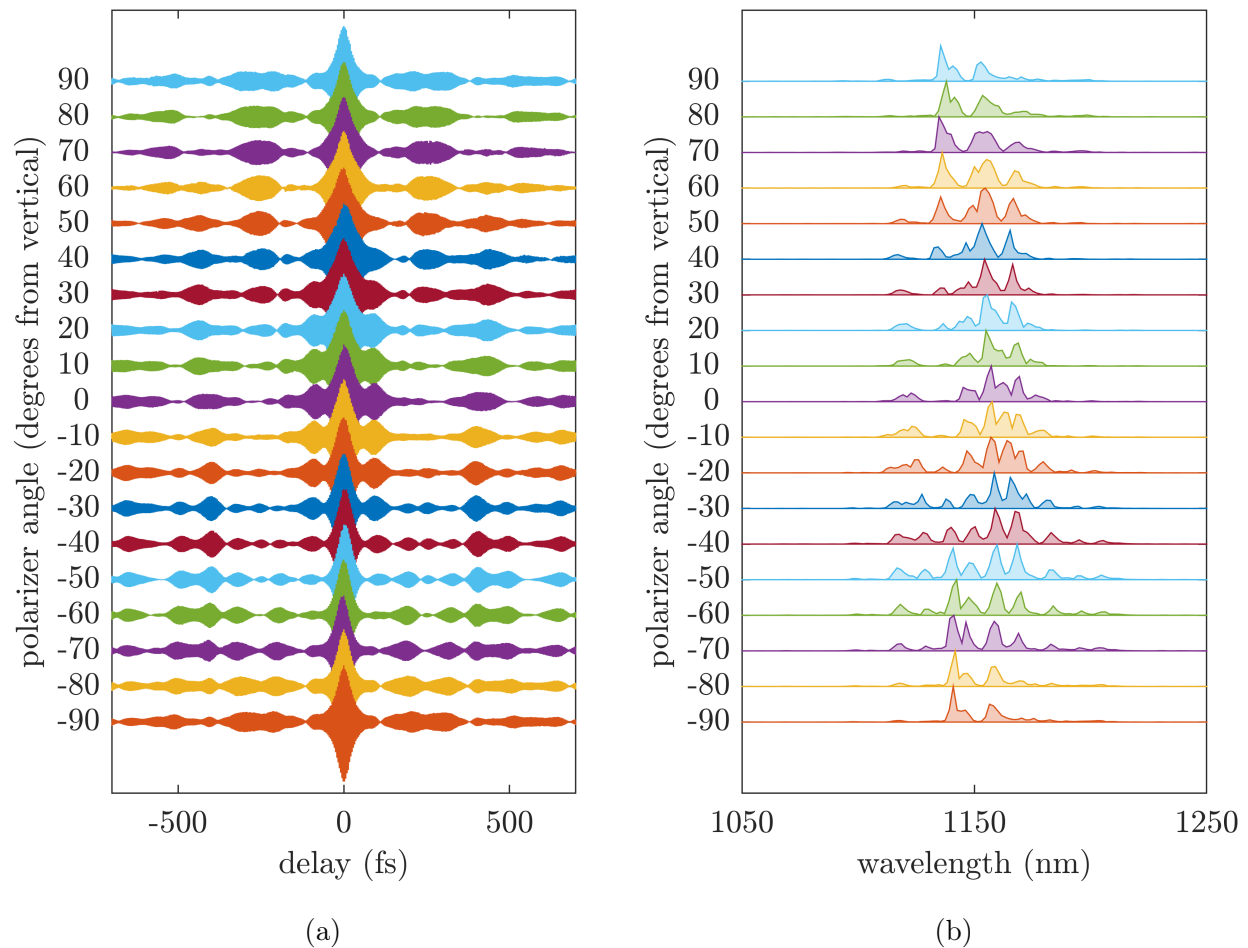
Figure 4-29: **Polarization effects observed in the FTIR:** In Fig. 4-30a we polarized the beam and observed changes in the FTIR. A change in the FTIR is accompanied by a change in the power spectrum (4-29b). Polarizer degrees are defined with 0 degrees corresponding to vertical, that is normal to the optical table (s polarization).

We then observed the polarization effects via changes in our IAC trace, as illustrated in Fig. 4-30.
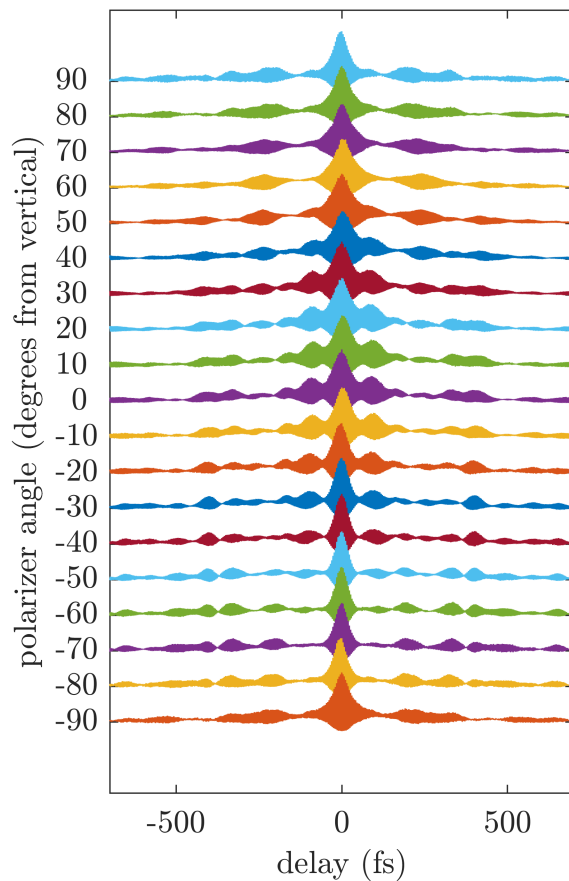
Figure 4-30: **Polarization effects observed in the IAC:** Similar polarization effects were observed in the IAC, as compared with the FTIR (4-29).

Lastly, we performed a FROG retrieval for the pulse observed at each polarization, with the direction of time corrected by a supplemental measurement with ZnSe, per Section 4.6.1. These recordings are cataloged in Fig. 4-31.
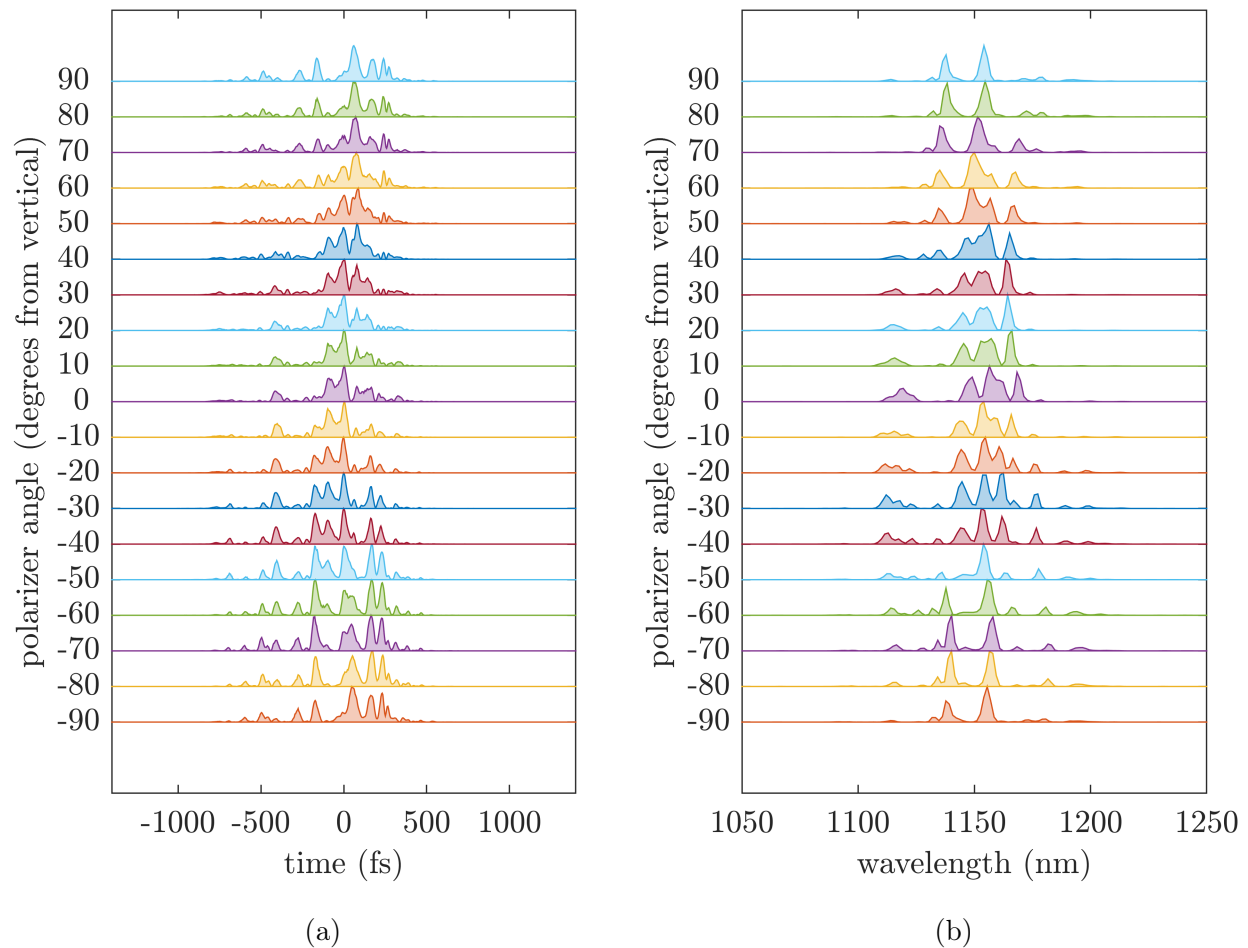
Figure 4-31: **Polarization effects observed in FROG:** A FROG retrieval was done for each beam polarization observed in the previous two measurements (FTIR and IAC). In 4-31a we observed the pulse temporal intensity profile changing with beam polarization. The corresponding power spectra are plotted in 4-31b.

Chapter 5

Outlook and Concluding Remarks

5.1 Compression

Of the many reasons for studying a pulse, one of the most important is determining whether it is compressible. In order to understand compressibility, we must understand that the spectral phase of a pulse is a major factor in determining the temporal shape. We will not go into the analysis here, but if a pulse has large variations in its spectral phase, then it is generally broader in time than the duration permitted by its bandwidth. In fact, a pulse with a completely flat spectral phase corresponds to the shortest possible pulse duration, given its particular power spectrum. A pulse with no phase variations is said to be **transform-limited**.

The question of compressibility is therefore equivalent to a question of whether there are any measures a researcher can take to manipulate the spectral phase of the pulse. There are a number of methods for exploring this question, but the one we take here is based on a simple sequence of prisms.

There is a straightforward derivation for the phase that a pulse would accumulate were it to propagate through a compressor, given certain parameters that define the compressor. If the phase on the pulse is suitable to manipulation, the possibility exists that the phase of the compressor, when added to the phase of the pulse, can alter the pulse temporal profile appreciably.

It can be shown that the phase introduced by propagation through a prism compressor

is given by:

$$\varphi(\omega) = \frac{2\omega L}{c} \cos\left(\theta_0 - \theta(\omega)\right) \ . \tag{5.1}$$

Here L is the distance between the apices of the prisms in the sequence, $\theta(\omega)$ is the exit-angle of the pulse with respect to the first prism, and θ_0 is the exit angle of the shortest frequency component in the power spectrum, that is $\theta_0 = \theta(\omega_{\min})$. The parameter $\omega_{\min}$ is determined by inspection of the power spectrum.

The exit angle $\theta(\omega)$ can be derived from geometry. The expression we used is:

$$\theta(\omega) = \sin^{-1}\left\{ n(\omega) \cdot \sin\left[\alpha - \sin^{-1}\left(\frac{\sin\theta_B}{n(\omega)}\right)\right] \right\} \ . \tag{5.2}$$

In Eq. (5.2), $n(\omega)$ is the refractive index of the prism glass, α is the apex angle of the prism (in our case provided by the manufacturer), and θ_B is the angle of incidence formed by the pulse and the surface of the first prism in the sequence. We have labelled it θ_B, because in our configuration, it corresponds to Brewster's angle, which is determined by the material.

First, we simulated the effects of a compressor with a apex separation of $L = 46.5$ cm. The prisms used for the simulation had an apex angle of 69.1° and were manufactured out of fused-silica (Thorlabs AFS-FS). The manufacturer also provided Brewster's angle, $\approx 55.6°$. The refractive index of fused-silica was computed from a Sellmeier equation.

Separately, we built a prism compressor with specifications identical to those used in our simulation. We performed another FROG retrieval, which ran for 100 iterations and produced a minimum G error of 0.004. The agreement between the recovered trace and the measured trace was excellent, and the recovered power spectrum agreed with the power spectrum before compression.
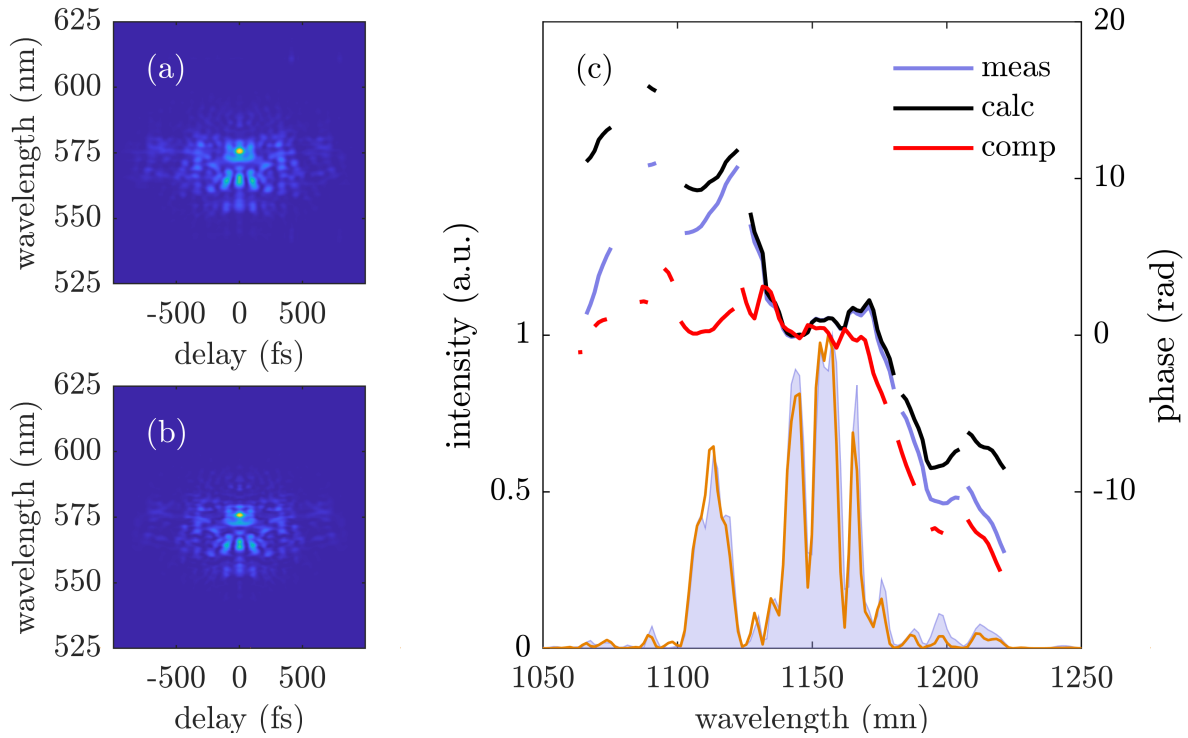
Figure 5-1: **FROG retrieval results for compressed pulse:** 5-1a shows the measured SHG FROG trace, and the recovered trace is plotted in 5-1b. In 5-1c the power spectrum for the orignal pulse is plotted as a shaded region, with an orange trace overlaid, which corresponds to the power spectrum of the compressed pulse. These two quantites should agree, since the compressor should only affect the phase. The legend entry "meas" corresponds to the original retrieval, "calc" is the calculated phase after adding the phase of the compressor, and "comp" is the retrieved phase of the compressed pulse.

Comparison of the temporal profiles is provided by Fig. 5-2. The figure shows excellent agreement between our computation and measurement, and indicates that we have achieved compression. The main double-peak with FWHM of 125 fs was compressed to a central peak with a FWHM around 37 fs, albeit with significant satellite peaks. Additional calculations suggest that other compressor configurations might reduce the intensity of the satellites. The magnitude of the recovered pulse was set to match that from our computation, which saw a factor of 1.95 increase with respect to the original pulse. The rms width decreased from 125 fs to 93 fs.
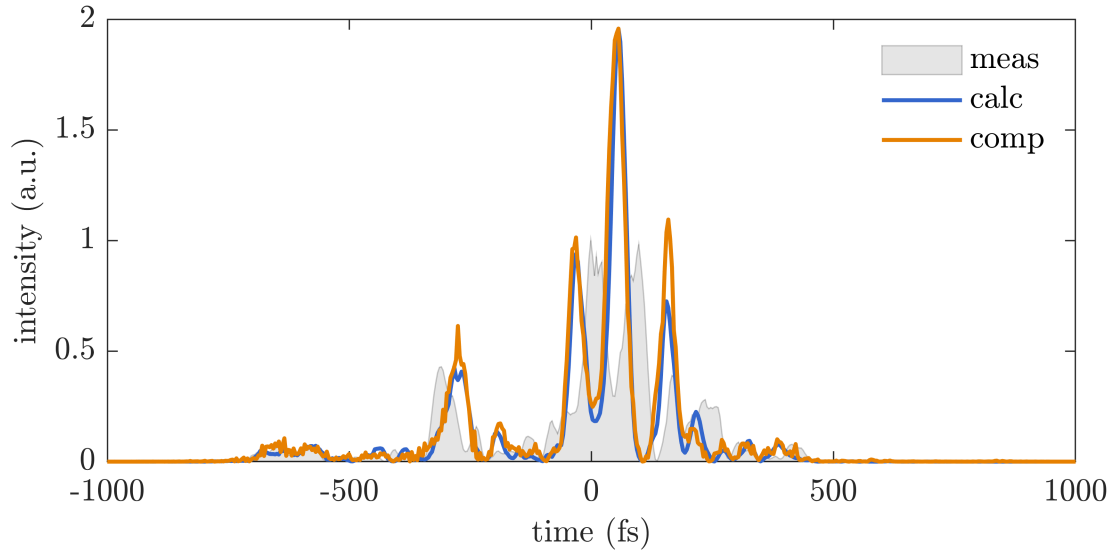
Figure 5-2: **Compressed pulse temporal profile.** Shown are the profiles of the original pulse, denoted by a shaded region with the label "meas", along with the profile predicted when the computed compressor phase was added, denoted by the blue trace with label "calc". The recovered compressed pulse is shown by the orange trace, with the label "comp".

5.2   Conclusion

In this thesis we reported on the first measurement of an ultrafast laser pulse produced by SHG via random quasi-phase-matching in polycrystalline ZnSe/S. We provided a theory of operation for a number of ultrafast experiments, and outlined the procedure for carrying out a pulse retrieval in a SHG FROG arrangement. We performed experiments using linear optics and nonlinear optics, and we discussed their measurement at both a photodiode and a spectrometer. We were successful at recovering a nearly complete description of the pulse temporal intensity profile and spectral phase. Cross-referencing the results of the retrieval with other measurements showed good agreement. We resolved an ambiguity inherent to our experiment in an additional measurement which amounted to resolving the phase shift from a known glass sample using FROG. We detailed the structure of the pulse, which has a center wavelength around 1141 nm and a temporal extent of around 125 fs. We cataloged a variety of pulse measurements corresponding to unusual and unexplained effects. Finally, we were successful in our initial compression attempts, and we plan to explore more options for compressing this pulse.

Appendix A

Ultrafast Metrics

For simple pulse shapes and spectral intensity profiles, the center frequency can be taken to be the frequency of maximum intensity. For complicated profiles, we measure the center frequency by interpreting the spectral intensity as a probability density function. This amounts to computing the mean value of the spectral intensity. The computation is a simple extension of the discrete mean to continuous functions. The mean of an arbitrary distribution F over the discrete index n is given by:

$$F_{\text{av}} = \frac{\displaystyle\sum_{n=-\infty}^{\infty} n \cdot F_n}{\displaystyle\sum_{n=-\infty}^{\infty} F_n} \tag{A.1}$$

For a distribution $f(x)$ over a continuous variable x, the mean is written as:

$$f_{\text{av}} = \frac{\displaystyle\int_{-\infty}^{\infty} x \cdot f(x)\,\mathrm{d}x}{\displaystyle\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x} \ . \tag{A.2}$$

We will use Eq. (A.2) to compute the center frequency of a complicated power spectrum $\tilde{S}(\nu)$.

Additionally, the standard deviation can be used to give a sense of pulse width and bandwidth. We will call this quantity the rms (root-mean-square) pulse width. For continuous functions, this takes the form:

$$\Delta f_{\text{rms}} = \frac{\int_{-\infty}^{\infty} (x - f_{\text{av}})^2 \cdot f(x) \, \mathrm{d}x}{\int_{-\infty}^{\infty} f(x) \, \mathrm{d}x} \quad . \tag{A.3}$$

We will report on our pulses using both the FWHM and the rms width.

Appendix B

The Power Method

In addition to a singular value decomposition (SVD), this derivation requires that we take for granted two additional results from linear algebra:

- By definition, a unitary matrix U has the property $U^\dagger = U^{-1} \Rightarrow UU^\dagger = U^\dagger U = I$, where I is the identity matrix. The columns of U also form an orthonormal basis for $\mathbb{C}^n$ (given that U is an $n \times n$ matrix).

- A diagonal matrix D with real entries has the property $D^\dagger = D$, and thus $DD^\dagger = D^\dagger D$.

B.1 Eigendecomposition

Since Σ is diagonal, it follows that $\Sigma\Sigma^\dagger = \Sigma^\dagger\Sigma \equiv \Lambda$, where Λ is diagonal with real nonnegative entries $\lambda_n = \sigma_n\sigma_n^* = \sigma_n^*\sigma_n$. We begin by showing that $\{\lambda_n\}$ are the eigenvalues of both $OO^\dagger$ and $O^\dagger O$, with corresponding eigenvectors $\{|p_n\rangle\}$ and $\{|g_n\rangle\}$.

Writing out the products explicitly, we have:

$$
\begin{aligned}
OO^\dagger &= (P\Sigma G^\dagger)(P\Sigma G^\dagger)^\dagger = (P\Sigma G^\dagger)(G^{\dagger\dagger}\Sigma^\dagger P^\dagger) \\
&= P\Sigma(G^\dagger G)\Sigma^\dagger P^\dagger = P\Sigma\Sigma^\dagger P^\dagger \\
&= P\Lambda P^\dagger
\end{aligned}
\tag{B.1}
$$

$$O^\dagger O = (P\Sigma G^\dagger)^\dagger (P\Sigma G^\dagger) = (G^{\dagger\dagger}\Sigma^\dagger P^\dagger)(P\Sigma G^\dagger)$$

$$= G\Sigma^\dagger (P^\dagger P)\Sigma G^\dagger = G\Sigma^\dagger \Sigma G^\dagger = \tag{B.2}$$

$$= G\Lambda G^\dagger$$

First addressing Eq. (B.1), we multiply on the right by an arbitrary column of P, and expand quantities as needed to make each point:

$$OO^\dagger \ket{p_n} = P\Lambda P^\dagger \ket{p_n} = P\Lambda \begin{bmatrix} \bra{p_1} \\ \vdots \\ \bra{p_N} \end{bmatrix} \ket{p_n} = P\Lambda \begin{bmatrix} \braket{p_1|p_n} \\ \vdots \\ \braket{p_N|p_n} \end{bmatrix}. \tag{B.3}$$

The set $\{\ket{p_n}\}$ is orthonormal, meaning $\braket{p_n|p_m} = \delta_{nm}$, so the vector of inner products on the right just returns the nth standard basis vector $\ket{e_n} = (0 \cdots 1 \cdots 0)$, where the nonzero entry occurs at the nth index. (For example, in $\mathbb{C}^3$, we have $\ket{e_2} = (0,1,0)$.) Plugging this result in, we find:

$$OO^\dagger \ket{p_n} = P\Lambda \ket{e_n} = P \begin{bmatrix} \lambda_1 \bra{e_1} \\ \vdots \\ \lambda_N \bra{e_N} \end{bmatrix} \ket{e_n} = P\lambda_n \ket{e_n} , \tag{B.4}$$

where the form $\Lambda = \begin{bmatrix} \lambda_1 \bra{e_1} & \cdots & \lambda_N \bra{e_N} \end{bmatrix}$ follows from the diagonality of Λ and the final step follows from the orthonormality of the standard basis. Thus, we have:

$$OO^\dagger \ket{p_n} = P\lambda_n \ket{e_n} = \lambda_n P \ket{e_n} = \lambda_n \begin{bmatrix} \ket{p_1} & \cdots & \ket{p_n} \end{bmatrix} \ket{e_n} = \lambda_n \ket{p_n} , \tag{B.5}$$

where again, $\ket{e_n}$ picks out the nth column. Eq. (B.5) is an eigenvalue equation. Therefore by definition the columns of P are eigenvectors of $OO^\dagger$, with corresponding eigenvalues λ_n. By inspection of Eq. (B.2), the columns of G are eigenvectors of $O^\dagger O$, with the same corresponding eigenvalues.

B.2 Principal Component Analysis

Any vector $|v\rangle \in \mathbb{R}^n$ can be written as a linear combination of basis vectors. If we choose our basis to be $\{|p_n\rangle\}$, then we can write:

$$|v\rangle = \sum_n c_n |p_n\rangle \ , \tag{B.6}$$

for some scalars c_n. Multiplying on the left by $OO^\dagger$ gives:

$$OO^\dagger |v\rangle = OO^\dagger \sum_n c_n |p_n\rangle = \sum_n c_n OO^\dagger |p_n\rangle = \sum_n c_n \lambda_n |p_n\rangle \ , \tag{B.7}$$

where the final step follows from Eq. (B.5) Multiplying on the left once more, we find:

$$(OO^\dagger)^2 |v\rangle = OO^\dagger \sum_n c_n \lambda_n |p_n\rangle = \sum_n c_n \lambda_n OO^\dagger |p_n\rangle = \sum_n c_n \lambda_n^2 |p_n\rangle \ . \tag{B.8}$$

This generalizes to arbitrary number m of multiplications:

$$\left(OO^\dagger\right)^m |v\rangle = \sum_n c_n \lambda_n^m |p_n\rangle \ . \tag{B.9}$$

Here's the rub: if λ_j is the largest of the eigenvalues ($\lambda_j > \lambda_n$ for each $j \neq n$), then for large m, λ_n^m overwhelms the other eigenvalues, and we can say $\left(OO^\dagger\right)^m |v\rangle \approx c_j \lambda_j^m |p_j\rangle$. Better yet, if we neglect the magnitude (we really only care about the magnitude *relative to the other eigenvectors*, and in practice we normalize our results anyway), then we can write:

$$\left(OO^\dagger\right)^m |v\rangle \approx |p_j\rangle \ . \tag{B.10}$$

An exactly analogous development shows that, we also can write:

$$\left(O^\dagger O\right)^m |v\rangle \approx |g_j\rangle \ . \tag{B.11}$$

B.3 PCGPA

Finally we are prepared to contextualize the results given by Eqs. (B.10) and (B.11). Since in our approximation $\lambda_j \gg \lambda_n$ for all other n, it follows that $|\sigma_j| \gg |\sigma_n|$. If we interpret σ_n as weights in the linear combination given by Eq. (3.36), then that equation reduces to a simple form:

$$O \approx \sigma_j \, |p_j\rangle\langle g_j| \ . \tag{B.12}$$

Eq. (B.12) is of the same form as the discrete mathematical form constraint given by Eq. (3.2.2), up to a factor of σ_j which we drop in normalization. Therefore, Eqs. (B.10) and (B.11) provide our approximate solutions. Explicitly,

$$|p'\rangle \approx OO^\dagger \, |p\rangle \tag{B.13}$$

$$|g'\rangle \approx O^\dagger O \, |g\rangle \tag{B.14}$$

Appendix C

Selected Computational Aspects

Besides the obvious experimental skills needed to perform the measurement of an ultrafast laser pulse, certain computational skills are also a necessity. The laboratory experiment itself produces an image of a spectrogram, or more fundamentally, a matrix of data. Obviously a number of numerical techniques must be deployed in order to properly manage this data and to transform it into the information we ultimately seek. Specifically, in order to recover the pulse completely, we have written a number of scripts which accomplish the tasks of: processing the raw data; performing the retrieval algorithm; and interpreting the recovered data. Each of these scripts of course includes a host of nested scripts and functions to perform individual tasks. Many of these are straightforward, but in this chapter we will discuss some of the more crucial aspects of computation that are perhaps less trivial (or obvious) than others.

C.1   Change of Coordinates

Ordinarily, spectrometers measure a spectrum against wavelength. The FROG retrieval algorithm involves iteratively transforming the spectrogram between the time and frequency domains, that is, *not* between the time and wavelength domains. Thus, we need a way to convert a spectrum between wavelength and frequency; mathematically this is accomplished by a change of coordinates. In addition to simply changing the independent variable, this involves scaling the spectrum by a factor called the Jacobian.

We can better understand this problem by noting that by conservation of energy, the integral of the power spectrum over frequency (the signal energy) must equal the integral over wavelength (also the signal energy), that is,

$$\int \tilde{S}(\nu)\, \mathrm{d}\nu = \int \tilde{S}(\lambda)\, \mathrm{d}\lambda \ . \tag{C.1}$$

This equation must hold over any wavelength (frequency) range, no matter how small, which requires

$$\tilde{S}(\nu)\, \mathrm{d}\nu = \tilde{S}(\lambda)\, \mathrm{d}\lambda \ , \tag{C.2}$$

or

$$\tilde{S}(\nu) = \tilde{S}(\lambda) \left| \frac{\mathrm{d}\lambda}{\mathrm{d}\nu} \right| \ . \tag{C.3}$$

where the absolute value brackets are required in the case of a negative derivative (the power spectrum is strictly nonnegative—see below). Rather than plugging in for this derivative, note again that the spectrometer provides wavelength data, meaning λ is the independent variable, and we need our derivative to reflect that fact. Therefore, we instead write:

$$\tilde{S}(\nu) = \tilde{S}(\lambda) \left| \frac{\mathrm{d}\nu}{\mathrm{d}\lambda} \right|^{-1} \ . \tag{C.4}$$

Noting that $\nu = c/\lambda$, and $\mathrm{d}\nu/\mathrm{d}\lambda = -c/\lambda^2$ (notice the need for the absolute value brackets) we end up with:

$$\tilde{S}(\nu) = \tilde{S}(\lambda) \frac{\lambda^2}{c} \ . \tag{C.5}$$

So in this case, the Jacobian scaling factor is given by λ^2/c. In pre-processing, the experimental FROG trace is processed for retrieval, and so we employ the change of coordinates described above. In post-processing, we wish to display our recovered data against wavelength, which involves the inverse change of coordinates. The Jacobian for the latter transformation is simply ν^2/c.

When it comes to computations involving data that are inevitably discrete, scaling the spectrogram is only half of the story. Notice that since ν and λ are not linearly related,

a discrete collection of data uniformly distributed in one domain will not be uniformly distributed in the other, even after scaling by the Jacobian. This problem is nicely illustrated by considering a double pulse, whose power spectrum is shown in Fig. C-1 below.
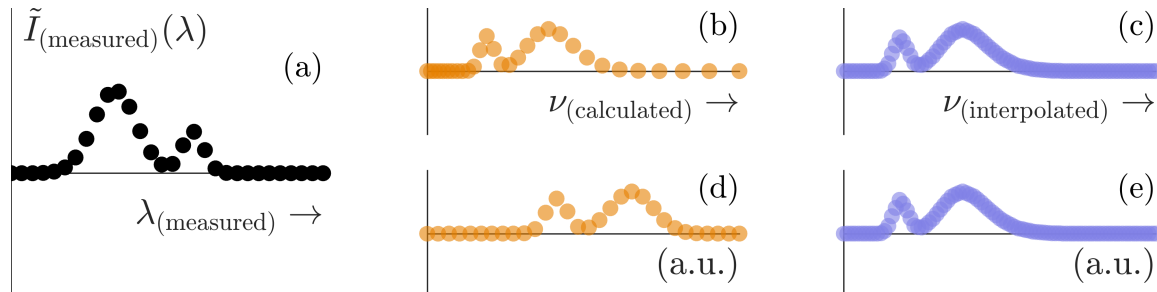

Figure C-1: **Computational aspects involved in a discrete change of coordinates:** A power spectrum against wavelength is simulated in (a). In (b), the Jacobian transformation is applied, but the resulting data is not uniformly distributed. Plots (b) and (d) convey the same data, but in (d) and abscissa is not specified, so the plot is skewed. Conversely, (c) shows a spectrum that was interpolated to a uniform frequency distribution, and (e) shows the same plot without an abscissa specified. The fact that they are identical emphasizes that (c) on its own indeed conveys uniformly distributed data.

In Fig. C-1a, we have plotted an example of a simulated power spectrum against wavelength, as might be recorded by a spectrometer. The plot was created in MATLAB using the command `plot(y, Sy)`, where `y` denotes discrete wavelength data provided by the detector. Fig. C-1b shows the same spectrum against frequency using the prompt `plot(v, Sv)`, where `Sv` is equivalent to `Sy` only scaled (vertically) by the Jacobian, the frequency vector was computed using `v = c/y`, and `c` was specified to be the speed of light in vacuum. Visually, this plot is scaled correctly, but this illusion is only accountable to the `plot()` command, which allows us to specify the abscissa (horizontal axis). Notice the problem: the data are not uniformly distributed, so the shorter frequencies (longer wavelengths) become much more tightly packed than the higher frequencies (shorter wavelengths). This is an artifact of the way we computed the frequency vector, which is an obvious result of the relationship $\lambda \propto 1/\nu$. This problem is further emphasized by comparing figure C-1b to C-1d, the latter of which plots the same exact data without specifying the abscissa, *i.e.* by commanding `plot(Sv)`. The data in this plot only *appear* to be uniformly distributed, but they are not scaled correctly in the horizontal direction.

This becomes a particularly salient problem when we think about the spectrogram. If we

did not provide a solution, then the "correct" image of the spectrogram would require some pixels to be larger than others, in the same way that some data spacings are larger than others in Fig. C-1b. This is of course not sensible. Moreover, note that the image is simply a way of viewing the spectrogram—the spectrogram itself is no more than a matrix of data. The algorithm does not include any features for determining how the data are distributed. All data are assumed by the algorithm to be uniformly distributed, so the data we provide to it need to be distributed as such.

To solve this problem, we need to interpolate the data. In brief, this involves creating a vector of frequency points that are uniformly distributed over the same frequency range `vUniform`, and then prompting `SUniform = interp1(v, spectrum_v, vUniform)`. The result is shown in Fig. C-1c. For comparison, Fig. C-1e was generated using the command `plot(SUniform)` without specifying the abscissa. Notice the two figures are identical, indicating that the data in Fig. C-1c are indeed uniformly distributed.

As a final word, we caution that the data plotted in Fig. C-1c were interpolated to a frequency spacing that matches the minimum frequency spacing in Fig. C-1b, hence the tightly packed data. This comes at the price of increasing the size of the data, which can be very costly to the retrieval algorithm. If the spectrum is not too finely structured then it is sufficient to interpolate over the same number of data points as the original spectrum. If the data *is* finely structured, care must be taken that the interpolation does not average-out any crucial spectral features.

Bibliography

[1] Rick Trebino. *Frequency Resolved Optical Gating: The Measurement of Ultrashort Laser Pulses.* Springer, 2000.

[2] Andrew M. Weiner. *Ultrafast Optics.* Wiley, 2009.

[3] Sergey Vasilyev, Igor Moskalev, Mike Mirov, Viktor Smolski, Sergey Mirov, and Valentin Gapontsev. Mid-ir kerr-lens mode-locked polycrystalline cr:zns and cr:znse lasers with intracavity frequency conversion via random quasi-phase-matching. *SPIE LASE*, page 97310B, 03 2016.

[4] Sergey Vasilyev, Igor Moskalev, Mike Mirov, Viktor Smolski, Sergey Mirov, and Valentin Gapontsev. Ultrafast middle-ir lasers and amplifiers based on polycrystalline cr:zns and cr:znse. *Opt. Mater. Express*, 7:2636–2650, Jul 2017.

[5] Jiahui Gu, Aaron Schweinsberg, Laura Vanderhoef, Michael Tripepi, Anthony Valenzuela, Christopher Wolfe, Trenton R. Ensley, Enam Chowdhury, and Miroslav Kolesik. Random quasi-phase-matching in polycrystalline media and its effects on pulse coherence properties. *Opt. Express*, pages 7479–7493, 03 2021.

[6] K.W. DeLong, Rick Trebino, and Daniel J. Kane. Comparison of ultrashort-pulse frequency-resolved-optical-gating traces for three common beam geometries. *J. Opt. Soc. Am. B*, page 97310B, 09 1994.

[7] Rana Jafari, Travis Jones, and Rick Trebino. 100% reliable algorithm for second-harmonic-generation frequency-resolved optical gating. *Opt. Express*, pages 2112–2124, 02 2019.

[8] B. E. A. Saleh and M.C. Teich. *Fundamentals of Photonics.* Wiley, 2019.

[9] Hermann A. Haus. *Waves and Fields in Optoelectronics.* Prentice Hall, 1984.

[10] B. P. Lathi. *Signal Processing and Linear Systems.* Berkeley-Cambridge Press, 1998.

[11] Gero Stibenz and Günter Steinmeyer. Interferometric frequency-resolved optical gating. *Opt. Express*, pages 2617–2626, 04 2005.

[12] Ivan Amat-Roldán, Iain G. Cormack, Pablo Loza-Alvarez, Emilio J. Gualda, and David Artigas. Ultrashort pulse characterisation with shg collinear-frog. *Opt. Express*, pages 1169–1178, 03 2004.

[13] Pablo Gabolde and Rick Trebino. Single-frame measurement of the complete spatiotemporal intensity and phase of ultrashort laser pulses using wavelength-multiplexed digital holography. *J. Opt. Soc. Am. B*, pages A25–A33, 06 2008.

[14] Alexander Gliserin, Soo Hoon Chew, Seungchul Kim, and Dong Eon Kim. Complete characterization of ultrafast optical fields by phase-preserving nonlinear autocorrelation. *Nature*, pages 884–886, 06 2022.

[15] David N. Fittinghoff, Jason L. Bowie, John N. Sweetser, Richard T. Jennings, Marco A. Krumbügel, Kenneth W. DeLong, Rick Trebino, and Ian A. Walmsley. Measurement of the intensity and phase of ultraweak, ultrashort laser pulses. *Opt. Lett.*, pages 884–886, 06 1996.

[16] J. R. Fienup. Phase retrieval algorithms: a comparison. *Applied Optics*, pages 2758–2769, 08 1982.

[17] J. R. Fienup. Reconstruction of a complex-valued object from the modulus of its fourier transform using a support constraint. *Applied Optics*, pages 118–123, 01 1987.

[18] Kenneth W. DeLong, David N. Fittinghoff, Rick Trebino, Bern Kohler, and Kent Wilson. Pulse retrieval in frequency-resolved optical gating based on the method of generalized projections. *Opt. Lett.*, pages 2152–2154, 12 1994.

[19] Daniel J. Kane. Recent progress toward real-time measurement of ultrashort laser pulses. *J. Quantum. Electron.*, pages 421–2154, 12 1994.