

# Ultrafast Electronic and Structural Dynamics in Solids

Eli Nathan Glezer



Harvard University  
Cambridge, Massachusetts  
September, 1996

# **Ultrafast Electronic and Structural Dynamics in Solids**

A thesis presented

by

Eli Nathan Glezer

to

The Division of Engineering and Applied Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Applied Physics

Harvard University

Cambridge, Massachusetts

September, 1996

# Abstract

This thesis investigates the dynamics of electrons and atoms in solids driven by intense, ultrashort laser pulses. Results of two series of experiments are presented. In the first set, the changes in the electronic properties of the semiconductor GaAs are determined by measuring the changes in its optical properties in response to 70-fs laser pulses. A fluence range of up to, and above, the damage threshold is examined. The experiments differ from previous work in the field, in that they are direct time-resolved measurements of the dielectric function and second-order optical susceptibility — fundamental quantities that characterize the optical state of the material. The dielectric function is measured from 1.5 to 3.5 eV, and at 4.4 eV, while the second-order susceptibility is measured at a single frequency of 2.2 eV. The results suggest a new view of the underlying electronic and structural changes. Three regimes of behavior are observed: at low excitation, rapid bandstructure changes are followed by lattice heating for about 10 ps; at medium excitation, stronger bandstructure changes are followed by a loss of long-range order in the crystal within several picoseconds; and at high excitation, an increasingly rapid transition to a metallic state is seen.

In the second set of experiments, the effect of ultrafast excitation inside the bulk of a solid is studied. It is shown that submicron-diameter voxels can be produced inside many transparent materials by tightly focusing 100-fs laser pulses. The use of such voxels for high-density 3-D optical data storage is demonstrated. Scanning electron microscopy and atomic force microscopy are used to examine 200-nm diameter voxels. The results suggest that extreme temperatures and pressures create a micro-explosion, leading to the formation of a void surrounded by densified material. Permanent structural changes are produced even in such hard materials as quartz and sapphire.

# Acknowledgments

Graduate school has been a rewarding and enjoyable experience for me. In large part this is thanks to the friends and colleagues who have enriched my life over the last five years.

Prof. Eric Mazur has been my advisor since my first day at Harvard. I could not have made a better choice. I thank Eric for being a constant source of inspiration and guidance, while giving me the freedom to pursue my ideas. I appreciate Eric's treatment of his students as younger colleagues — it is both challenging and rewarding. I've had the opportunity to be involved in important aspects of research, in addition to the core elements of carrying out and analyzing experiments: I've gained experience in coming up with ideas for experiments, writing grant proposals, and presenting results. Eric's insistence on clarity in communicating ideas and focusing on essentials have set an example for me. Finally, on a personal level, I've always been able to rely on Eric for support and encouragement.

The collaborative approach to research in our group has given me a chance to work with and learn from some very bright and capable fellow students. Chapters 3 and 4 of this thesis are based on work carried out together with Yakir Siegal. Yakir and I worked side by side for the first three years of my graduate school, and along the way he taught me a lot of physics and optics. We spent many long days running experiments, and many more days getting to that point. Yakir and I jointly came up with the idea and the approach to measuring the dielectric constant instead of just reflectivity, which led to a re-examination of common assumptions and a new view of electron and lattice dynamics. From designing experiments to interpreting results, I enjoyed very much working with Yakir.

Another major contributor to the work in this thesis, and especially to Chapter 5, is Li Huang.

Li joined Yakir and me as an enthusiastic beginning graduate student, participated in the work in Chapters 3 and 4, and in the process became a very capable and confident experimentalist. Li deserves much of the credit for pushing through the final phase of the difficult experiments described in Chapter 5. It has been a real pleasure for me to work with Li. Both Li and I benefitted greatly from the contribution of the newest member of our project, Paul Callan. Paul's dedication and strong computational background allowed him to make a significant impact on the success of the project in his first year with our group.

The work described in Chapters 6 and 7 is unique in that I was able to draw on the help of several other group members — Rich Finlay, Tsing-Hua Her, and Milos Milosavljevic. What started out as a quick test of some ideas quickly grew into a very exciting and extensive research topic. Although this is the most 'solo' work that I've done in designing experiments and analyzing results, it was great to have the help of Milos, Li, Rich, Tsing-Hua, and Paul in carrying out various parts of the experiments. I look forward to continuing this work together with Chris Schaffer, who, within weeks of joining our group, has already demonstrated impressive experimental skills and unusual dedication.

Several older students in the group played an important part in my graduate school experience. I've learned a lot from and enjoyed many insightful conversations with Shrenik Deliwala on physics, optics, and many topics far outside of lab. Shrenik's enthusiasm for science and flow of creative ideas are truly captivating. I've also learned from Jay Goldman, especially his skillful work in the lab and realistic perspective. Jay's cheerful attitude always made work more fun. Shrenik and Jay designed and built the Ti:Sapphire laser I later used for the experiments in Chapters 6 and 7. Juen-Kai Wang and Peter Saeta built the CPM laser and laid the groundwork for the semiconductor experiments (Chapters 3 – 5). Ka Yee Lee gave me a warm and cheerful welcome

on my first visit to Harvard, long before I even joined Eric's group.

Although Aryeh Feder and I have yet to work on a common project, we've always discussed new ideas and results. Aryeh's thoughtful approach and critical thinking have been very helpful to me, and I've enjoyed our many conversations over the last four years. Rich has made valuable suggestions on experiments and provided lots of help with computer control. I thank Rich, Tsing-Hua and Claudia Wu for letting me use their lab and showing me how to run the Ti:Sapphire laser. I've also benefitted from working with Walter Mieher, Qing-Yu Wang, and Wendy Zhang.

I thank my thesis committee, Profs. Henry Ehrenreich, Tim Kaxiras, and Mike Aziz, for many valuable discussions and their interest in my work. Their input has made me question assumptions and has clarified the interpretation of our experimental results. I'm grateful to Prof. Nico Bloembergen who, although not formally on my committee, has been enthusiastic about my work, often pointing out critical issues and making valuable suggestions throughout many discussions. I'm lucky to have met and become friends with Prof. Rino Di Bartolo, whose enthusiasm and warmth of spirit is contagious. During his sabbatical in our group, we spent many afternoons discussing physics and occasionally drifting off to politics, religion, arts and literature. I thank him for giving me a unique opportunity to give a set of lectures on the techniques of ultrafast spectroscopy at the NATO summer school in Erice, on which Chapter 2 of this thesis is based. I thank Prof. Claus Klingshirm for valuable discussions during his summer sabbatical in our group. I thank Prof. Andrew Lange for guidance in my undergraduate research in experimental astrophysics at Berkeley, and giving me a unique opportunity to participate in everything from designing and building an instrument to taking it to the top of Mauna Kea, Hawaii to do astronomical observations. I thank my high school teacher, Mr. Owen, for his remarkable dedication to teaching, and his emphasis on experiments as the foundation physics.

Many close friends aided me throughout graduate school. In particular, for memorable experiences and moral support I want to thank Eddie Kreisberg, Debby Lu, Tom Steinbach, Jay Goldman, Deanna Griffin, Viviana Waisman, Mark Fish, Lisa Morgan, Yakir Siegal, Shrenik Deliwala and Phalguni Shah, Catherine Jeannin, Ron Kim, John Baskey, Dave Morin, the regulars on the Ultimate field in the middle of the Harvard track and my teammates on the leaky basketball court of Hemenway gym.

Finally, I want to thank my family — my parents, my grandparents, and my brother Jacob — for their unconditional love, encouragement, and unlimited generosity. I'm extremely lucky to have such a close family, and they have been my greatest support in graduate school and throughout my life.

## **Acknowledgment of Financial Support**

I gratefully acknowledge the support a graduate fellowship from the Hertz Foundation.

# Contents

<b>Abstract</b>	<b>iii</b>
<b>Acknowledgments</b>	<b>iv</b>
<b>Table of Contents</b>	<b>viii</b>
<b>List of Figures</b>	<b>xi</b>
<b>Chapter 1: Introduction</b>	<b>1</b>
Introduction	54
Optical Properties and Electronic Structure	55
<b>Chapter 2: Techniques of Ultrafast Spectroscopy</b>	<b>6</b>
Introduction	6
Introduction to Ultrafast Spectroscopy	6
Limits of Time and Frequency Resolution	10
Propagation of Light Pulses	15
Producing Femtosecond Pulses	20
Femtosecond Lasers	20
Amplification	33
Measuring and Shaping Femtosecond Pulses	37
Pulse Characterization	38
Pulse Shaping	47

<b>Chapter 3: Laser-Induced Band-Gap Collapse in GaAs</b>	<b>53</b>
Introduction	54
Optical Properties and Electronic Structure	55
Experimental Setup	56
Experimental Results	57
Analysis and Discussion	62
Conclusion	64
<b>Chapter 4: Behavior of <math>\chi^{(2)}</math> During a Laser-Induced                     Phase Transition in GaAs</b>	<b>66</b>
Introduction	67
$\chi^{(1)}$ and $\chi^{(2)}$ Dependence of Second-Harmonic Signal	67
Experimental Setup	69
Experimental Results	70
Discussion	72
Conclusion	73
<b>Chapter 5: GaAs Under Ultrafast Excitation: Response                     of the Dielectric Function</b>	<b>75</b>
Introduction	76
Background	76
Experimental Setup	78
Measuring the Dielectric Function	82
Results	87

Discussion	91
Conclusion	97
<b>Chapter 6: 3-D Optical Storage Inside Transparent Materials</b>	<b>100</b>
<b>Chapter 7: Ultrafast-Laser Driven Micro-Explosions</b>	
<b>Inside Transparent Materials</b>	<b>104</b>
Introduction and Background	104
Description of Experiments	107
Results	108
Discussion	118
Applications	131
Conclusion	132

# List of Figures

## Chapter 2

Fig. 1	7-fs pulse of 800-nm wavelength light	7
Fig. 2	Pump-probe technique	8
Fig. 3	Example of time-domain spectroscopy	9
Fig. 4	Example of spectroscopy with femtosecond time resolution	9
Fig. 5	Time representation of two pulses	11
Fig. 6	Corresponding frequency representations	11
Fig. 7	Limits of time and frequency resolution: thought experiments	12
Fig. 8	Pump-probe thought experiments: case 2 and 3 in more detail	13
Fig. 9	Very fast saturable absorber	14
Fig. 10	Hypothetical measurements on a very fast saturable absorber	14
Fig. 11	Self-phase-modulation	18
Fig. 12	Self-focusing	19
Fig. 13	Wigner representation of a Gaussian pulse	20
Fig. 14	Wigner representation of a chirped pulse	21
Fig. 15	Basic elements of a femtosecond laser cavity	21
Fig. 16	Two examples of a femtosecond laser: dye-based and solid-state	22
Fig. 17	Gain and loss spectra in a CW laser	23
Fig. 18	Active mode locking	24
Fig. 19	Passive mode locking with ‘slow’ saturable absorption	25
Fig. 20	Fast saturable absorber	26
Fig. 21	Examples of KLM with self-focusing and an aperture	27
Fig. 22	Pulse-shortening rate for different mode-locking mechanisms	28
Fig. 23	Group velocity in fused silica glass	28
Fig. 24	GVD compensation with gratings and prisms	29
Fig. 25	Schematic for calculating grating pair dispersion	30
Fig. 26	GVD, GDD, and optical path derivatives	31
Fig. 27	Gain narrowing in an amplifier	34
Fig. 28	Two-lens correlator	36

Fig. 29	Pulse stretcher	36
Fig. 30	Lasers and amplifiers: state of the art for sub-100-fs pulses	37
Fig. 31	Typical spectrum of an ultrashort pulse from a Ti:Sapphire laser	38
Fig. 32	Non-collinear autocorrelation	39
Fig. 33	7-fs pulse and the corresponding SH intensity autocorrelation	40
Fig. 34	Interferometric autocorrelation	41
Fig. 35	7-fs pulse and the corresponding SH interferometric autocorrelation	41
Fig. 36	Interferometric time-delay setup without a nonlinear element	42
Fig. 37	7-fs transform-limited pulse and its linear interference trace	43
Fig. 38	14-fs transform-limited pulse and its linear interference trace	43
Fig. 39	7-fs pulse, with autocorrelation and linear interference traces	44
Fig. 40	7-fs pulse after propagating through a dispersive medium	44
Fig. 41	General approach to joint spectral-temporal pulse characterization	45
Fig. 42	Frequency-resolved optical gating (FROG)	45
Fig. 43	Spectrally-resolved autocorrelation	46
Fig. 44	Example of a spectrally-resolved autocorrelation	47
Fig. 45	Pulse shaper	48
Fig. 46	Example of simple pulse-shaping	49
Fig. 47	Examples of shaped pulses	50

### Chapter 3

Fig. 1	Dielectric function of GaAs	55
Fig. 2	Electronic band structure of GaAs	55
Fig. 3	Probing geometry for 22- and 44-eV measurements	57
Fig. 4	Dielectric constant at 22 eV vs pump-probe time delay	58
Fig. 5	Dielectric constant at 22 eV vs pump fluence	58
Fig. 6	Pump-probe time delays at which $\text{Re}(\epsilon) = 0$ at 22 eV	58
Fig. 7	Schematic representation of the band-gap collapse	59
Fig. 8	Dielectric constant at 44 eV vs pump-probe time delay	60
Fig. 9	Dielectric constant at 44 eV vs pump fluence	60
Fig. 10	Pump-probe time delays at which $\text{Re}(\epsilon) = 0$ at both 22 and 44 eV	60

Fig. 11	Comparison of observed dielectric constant to the Drude model	61
Fig. 12	Time evolution of double-oscillator model dielectric function	62

## Chapter 4

Fig. 1	Effect of the dielectric constant on the electric field orientation	68
Fig. 2	Probing geometry for second-harmonic-generation measurements	68
Fig. 3	Normalized second-harmonic signal vs pump fluence	70
Fig. 4	Square of the second-order susceptibility vs pump fluence	70
Fig. 5	Square of the second-order susceptibility vs pump-probe time delay	71
Fig. 6	Second-harmonic and the square of second-order susceptibility	71
Fig. 7	Second-harmonic and $45^\circ$ reflectivity after (a) few seconds, (b) 100 ps	71

## Chapter 5

Fig. 1	Broadband probe generation and spatial filtering	78
Fig. 2	Original and filtered spectra averaged over 100 pulses	80
Fig. 3	Experimental setup	80
Fig. 4	Temporal dispersion of the spectrum of the white-light probe	81
Fig. 5	Mapping of the complex dielectric-function plane to reflectivity pairs	83
Fig. 6	Reflected and reference spectra from a single pulse	84
Fig. 7	Reflectivity of GaAs at 58 and 75 degrees, p-polarized	85
Fig. 8	Dielectric function at negative time delay	86
Fig. 9	Dielectric function for a pump fluence of $0.45 F_{th}$ after 500 fs and 4 ps	87
Fig. 10	Lattice heating in the low fluence range	88
Fig. 11	Dielectric function for a pump fluence of $0.7 F_{th}$ after 500 fs and 4 ps	89
Fig. 12	Dielectric function for a pump fluence of $1.6 F_{th}$ after 500-fs and 4-ps	91

## Chapter 6

Fig. 1	Schematic diagram of 3-D optical data storage	101
Fig. 2	Binary data pattern stored inside fused silica with $2\text{-}\mu\text{m}$ bit spacing	102
Fig. 3	SEM view of a cross-section through a $5\text{-}\mu\text{m}$ spaced array of bits	102
Fig. 4	Damage produced inside fused silica by 200-ps pulses.	103

## Chapter 7

Fig. 1	Binary data pattern stored inside fused silica with 2- $\mu\text{m}$ bit spacing	109
Fig. 2	2- $\mu\text{m}$ spaced array of voxels photographed in reflection	110
Fig. 3	SEM view of a cross-section through a 5- $\mu\text{m}$ spaced array of bits	111
Fig. 4	AFM scans of three pits and three bumps	112
Fig. 5	Diffraction from a 5- $\mu\text{m}$ spaced array of voxels	113
Fig. 6	Damage produced inside fused silica by 200-ps pulses	114
Fig. 7	SEM view of damage produced inside fused silica by a 10-ns pulse	115
Fig. 8	Schematic diagram of the setup for time-resolved imaging	116
Fig. 9	Time-resolved images 1 ps, 5 ps, 100 ps, and 1 second after excitation	117
Fig. 10	Transmission of light in the plasma cone at 1 ps, 5 ps, and 100 ps	118
Fig. 11	Diffraction efficiency of a voxel array as a function of $\Delta n$	119
Fig. 12	Effect of self-focusing of the beam profile	123
Fig. 13	Beam waist as a function of self-focusing strength	123
Fig. 14	Defocusing due to the effect of the plasma on the index of refraction	126
Fig. 15	Plasma absorption along the beam axis	127
Fig. 16	Ratio of critical power for defocusing to critical power for absorption	127

To my parents, Mila and Boris

# Chapter 1

## Introduction

Light plays a special role in the physical world as a courier of information. It makes possible our dominant sense of perception — vision, and with the help of instruments allows us to extend our ability to see beyond our natural range, out to distant galaxies, down to the sub-cellular level, and across the entire electromagnetic spectrum, from microwaves to x-rays.

The ability to produce extremely short pulses of light has opened up a new realm for optical exploration, that of ultrafast dynamics of electrons and atoms. Not only do these light pulses allow us to observe dynamics on an unprecedented timescale, but they can also be used to create unique conditions in matter.

### General overview

This thesis investigates the dynamics of electrons and atoms in solids driven by intense, ultrashort laser pulses. The short duration and high intensity of the excitation in these experiments, typically 100 fs ( $10^{-13}$  sec) duration and up to  $10^{15}$  W/cm<sup>-2</sup> intensity, create highly non-equilibrium conditions, such as a plasma of energetic electrons in an atomic lattice that is still cold. This provides an opportunity to observe the behavior of the hot electrons, their interaction with the atomic lattice, and the structural changes that result.

In the first set of experiments presented in this thesis, (Chapters 3, 4 and 5) we determined the changes in the electronic properties of the semiconductor GaAs by measuring the changes in its optical properties. When the optical pulse promotes valence electrons to higher energy states, it immediately alters the electronic properties of the material. Over the course of about 10 ps, the

excited electrons transfer their energy to the ions by inducing vibrations in the lattice. When the laser excitation is strong enough, it has the more drastic effect of destabilizing the bonding in the crystal. Within several picoseconds the instability in the bonding leads to a loss of crystalline order. Even stronger excitation causes a structural phase transition to occur in under a picosecond.

The second set of experiments presented in this thesis (Chapters 6 and 7) examines the effect of ultrafast excitation inside the bulk of a solid, rather than on its surface. When an intense ultrashort laser pulse is absorbed on the surface of a solid, material is vaporized and ablated, leaving a cratered surface. Internal optical excitation is possible if the solid is transparent to the laser beam. By focusing the ultrashort pulse inside the transparent solid, energy is deposited only at the focus, where the laser intensity is high enough for nonlinear absorption. The expansion that normally accompanies heating, melting, and vaporization at the surface, is restrained by the surrounding lattice. The tremendous pressures cause an explosive expansion that forces material into the surrounding volume. Permanent structural changes are produced even in such hard materials as quartz and sapphire.

## **Content and organization**

The next chapter is a general introduction to the techniques of ultrafast optics. It begins with a description of ultrafast spectroscopy, considers the limits of time and frequency resolution, and reviews the linear and nonlinear propagation of light pulses in a dispersive medium. Next, the basic elements of ultrashort laser pulse generation are described, including gain medium requirements, mode-locking mechanisms, compensation for group velocity dispersion, and pulse amplification. The last part of the chapter deals with the measurement of ultrashort pulses, including joint time-frequency techniques, and also describes pulse-shaping in the frequency domain.

Chapters 3, 4 and 5 present the results of linear and nonlinear optical measurements of the

response of GaAs to intense 70-fs excitation. The data in Chapter 3 are the first time-resolved measurements of the dielectric constant of a semiconductor following ultrafast excitation. Both the real and imaginary parts of the dielectric constant are measured at two photon energies of 2.2 eV and 4.4 eV. The experiments differ from previous work in the field, in that they are direct measurements of the fundamental quantity that characterizes the optical state of the material. Our data show that previous assumptions of how electrons respond to the intense excitation are incorrect, and the results suggest a new view of the underlying electronic and structural changes.

In Chapter 4, the second-order nonlinear susceptibility,  $\chi^{(2)}$ , is determined and used to monitor the changes in the long-range order in the GaAs crystal following the ultrafast excitation. In previous studies, subtle changes in  $\chi^{(2)}$  were masked by changes in the dielectric constant. By combining new measurements of second-harmonic generation with the measurements of the dielectric constant from Chapter 3 we extract the time evolution of  $\chi^{(2)}$  below and above the damage threshold. Several new effects are observed.

In Chapter 5 we move beyond single-frequency optical measurements by using a new broadband spectroscopic technique to simultaneously monitor the response of the dielectric function,  $\epsilon(\omega)$ , of GaAs across the spectral range from 1.5 to 3.5 eV. By measuring the evolution of  $\epsilon(\omega)$  with 100-fs time resolution we obtain the most detailed view thus far of the electron and lattice dynamics over a wide range of excitation strengths. The data support the conclusions reached in Chapters 3 and 4, and provide further insight into bandstructure changes, lattice heating, loss of crystalline order, and the semiconductor-metal transition.

Chapters 6 and 7 present the results of a new project on femtosecond-laser-induced structural changes inside transparent materials. In Chapter 6, a novel method for 3-D optical data storage is demonstrated. The method is based on our discovery that tight focusing of ultrashort pulses can

produce submicron-diameter structures, or voxels, inside many transparent materials. Binary information is recorded by writing such voxels as bits in multiple planes, which are read out using a microscope objective with a short depth of field. Data storage and retrieval with 2- $\mu\text{m}$  in-plane bit spacing and 15- $\mu\text{m}$  inter-plane spacing (17 Gbits/cm<sup>3</sup>) is demonstrated.

Chapter 7 presents the results of further optical and structural examination of these voxels. The results suggest that extreme temperatures and pressures create a ‘micro-explosion’, resulting in the formation of a void surrounded by densified material. Scanning electron microscopy and atomic force microscopy show structural changes confined to an area 200 nm in diameter. The roles of nonlinear optical processes, including multi-photon absorption, self-focusing, plasma absorption and defocusing are discussed.

## **Chapter 2**

### **Techniques of Ultrafast Spectroscopy**

The following paper appeared in *Spectroscopy and Dynamics of Collective Excitations in Solids*, Ed. B. Di Bartolo (Plenum, NATO ASI series, 1996).

# TECHNIQUES OF ULTRAFAST SPECTROSCOPY

Eli N. Glezer

Gordon McKay Laboratory  
Division of Applied Sciences  
Harvard University  
9 Oxford Street  
Cambridge, MA 02138

## ABSTRACT

This chapter begins with a general introduction to ultrafast spectroscopy, considers the limits of time and frequency resolution, and reviews the linear and nonlinear propagation of light pulses in a dispersive medium. Next, the basic elements of ultrashort laser pulse generation are described, including gain medium requirements, mode-locking mechanisms, compensation for group velocity dispersion, and pulse amplification. The last section deals with the measurement of ultrashort pulses, including joint time-frequency techniques, and also describes pulse-shaping in the frequency domain.

## I. INTRODUCTION

### I. A. Introduction to Ultrafast Spectroscopy

The development of lasers over the last thirty five years has greatly expanded the range of physical phenomena accessible to optical measurements. Even before the invention of the first laser, classical spectroscopy was a well-developed science and an important experimental tool. Classical spectroscopy involves measuring the transmission, reflection or emission of light by some medium, as a function of the frequency of the light. Laser spectroscopy has made possible measurements which could not be made with other light sources. The advantages of using a laser source are many: narrow linewidth, tunability, spatial coherence, temporal coherence, and high intensity. The tunability and narrow linewidth, (today as narrow as 1 KHz or even 1 Hz for some frequency-stabilized lasers), has greatly increased the achievable spectral resolution and has allowed, for example, such measurements as the Lamb shift, observed in the fine structure of the Hydrogen absorption spectrum [1]. But it is not the narrow linewidth alone that has made many of the new measurements possible: the high intensity of laser light, especially when the laser is operated in a pulsed mode, has made possible nonlinear spectroscopy, with such techniques as saturation spectroscopy, coherent Raman spectroscopy, photon-echo, multiphoton absorption, and transient-grating, just to name a few. For example, in the Hydrogen fine structure measurements, saturation spectroscopy was used as a clever tool to circumvent Doppler broadening by probing only atoms at rest, because only they would

be simultaneously resonant with narrow-frequency light from counter-propagating beams.

Ultrafast spectroscopy is a relatively recent branch of laser spectroscopy. Lasers have produced pulsed outputs since the first laser (Ruby pumped by a Xenon flashlamp), and production of much shorter pulses by Q-switching was demonstrated shortly after that. Since then, there has been a progression to shorter and shorter pulses, from nanoseconds, to picoseconds, to femtoseconds (fs). The term ‘ultrafast’ is generally applied to phenomena that occur on a femtosecond- or picosecond- time scale. In calling anything ‘ultra’ there is a risk that soon something ‘more ultra’ will be discovered. However, in the case of ultrafast spectroscopy the term may be justified, at least for visible and infrared light. Today's shortest optical pulses – less than 10 fs in duration – consist of only a few cycles of the electromagnetic field and thus require a bandwidth that is a sizable fraction of the central frequency. Figure 1 shows the electric field of a 7-fs pulse of 800-nm wavelength light. Because the duration of a single cycle of visible light is about 2 fs, producing pulses much shorter than a few femtoseconds will require going beyond the visible spectrum into the ultraviolet or X-ray region.

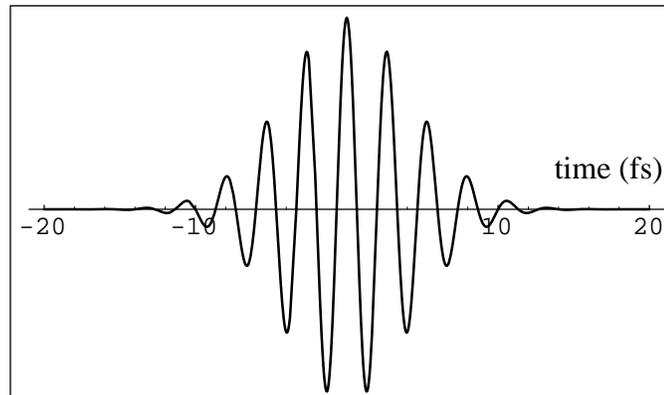


Figure 1. 7-fs pulse of 800-nm wavelength light.

Two questions fundamental to ultrafast optics are: how do we measure such short optical pulses when there are no electronic detectors with fast enough response times? and how do we use ultrashort pulses for studying ultrafast processes? The answer to the first question will be discussed in section III. A. The second question has been answered in a great variety of ways in clever experiments, but there is one general technique that includes a wide range of linear and nonlinear optical measurements: the pump-probe technique, illustrated in Fig. 2. In the simplest version, two ultrashort pulses are employed: a strong ‘pump’ pulse is used to excite the system of interest, and a weak ‘probe’ pulse that arrives at some later time is used to observe the resulting dynamics. By changing the time delay between the two pulses, the time evolution of the excited system can be mapped out. A more complicated pump may consist of multiple pulses, pulses of different frequencies, or specially tailored waveforms, and the probe may involve a variety of linear or nonlinear optical interactions, but the concept of separately inducing and monitoring dynamics still applies.

Ultrafast spectroscopy comes in two basic versions: 1. Spectroscopy in the time domain, and 2. Spectroscopy in frequency with time resolution. In the first case, the

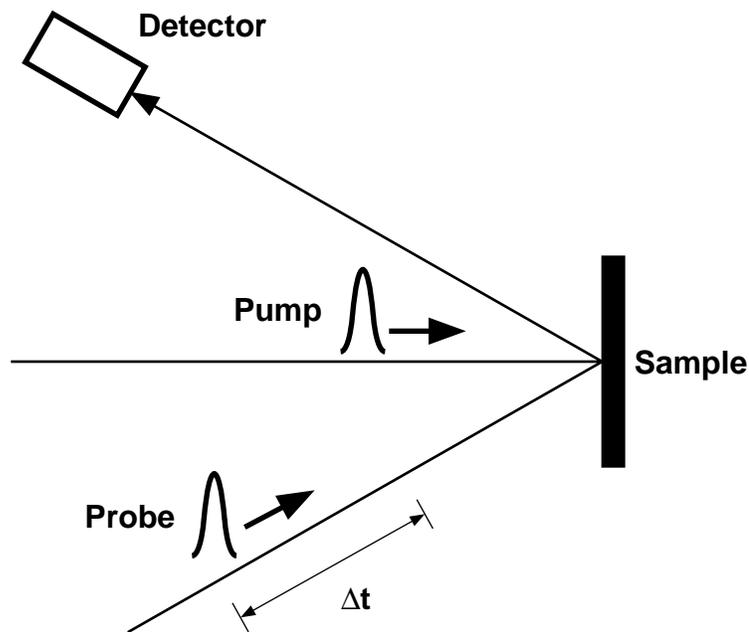


Figure 2. Pump-probe technique.

frequency information comes entirely from the temporal response of the system. An excellent example of this are impulsively driven lattice vibrations observed in time. Figure 3 shows oscillations in the reflectivity of an Antimony crystal as it responds to a 70-fs optical excitation, as measured by varying the time-delay of a 70-fs probe [2]. By analyzing the frequency content of the response through a Fourier transform, shown in the inset, the authors identify which of the known phonon modes are driven by the excitation. In this case, only the  $A_{1g}$  symmetric breathing mode is excited, while other Raman-active modes are not. The authors are able to determine that the excitation is not through a Raman process, but instead is due to a near-instantaneous displacement of the potential surface seen by the atoms when electrons are transferred to higher energy states through direct absorption. As the lattice vibrates, the electronic bandstructure oscillates with it, causing the modulation in the reflectivity. In this example the frequency information in the THz range is obtained by simply Fourier transforming the temporal response.

The second kind of ultrafast spectroscopy involves obtaining frequency information in the optical range by using a spectrometer to directly resolve the probe pulse into its spectral components. An impressive early example of this type of spectroscopy is the measurement of the vibration of an organic dye molecule in response to a 6-fs laser pulse excitation. Fig.4 shows the change in the transmittance of the dye solution as a function of both the pump-probe time-delay and the probe wavelength [3]. In fact, the entire wavelength range is covered by a single 6-fs probe pulse, which is spectrally dispersed after it is transmitted through the sample.

A 6-fs pulse in the visible range has a minimum bandwidth limit of over 50 nm. However, as we see in Fig. 4, it is certainly possible to do a measurement with higher frequency resolution than the minimum bandwidth of the probe pulse. This example may be a good backdrop for considering the question of fundamental limits of frequency and

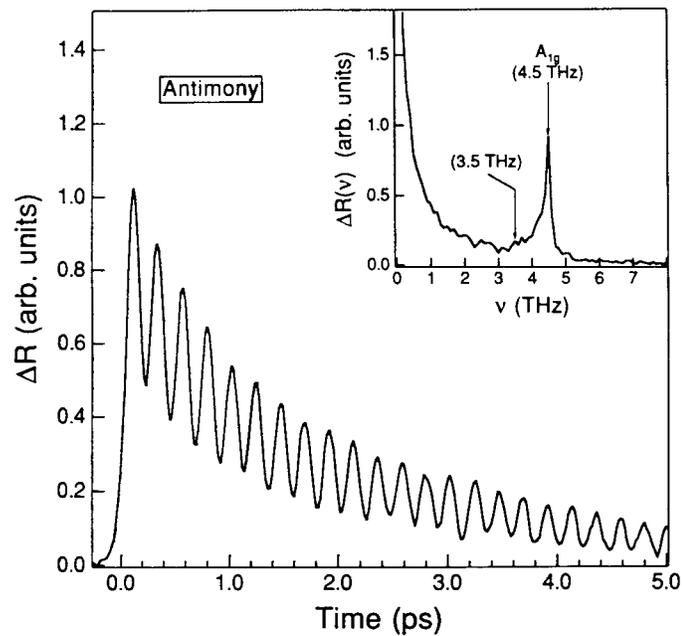


Figure 3. Example of time-domain spectroscopy (from [2]).

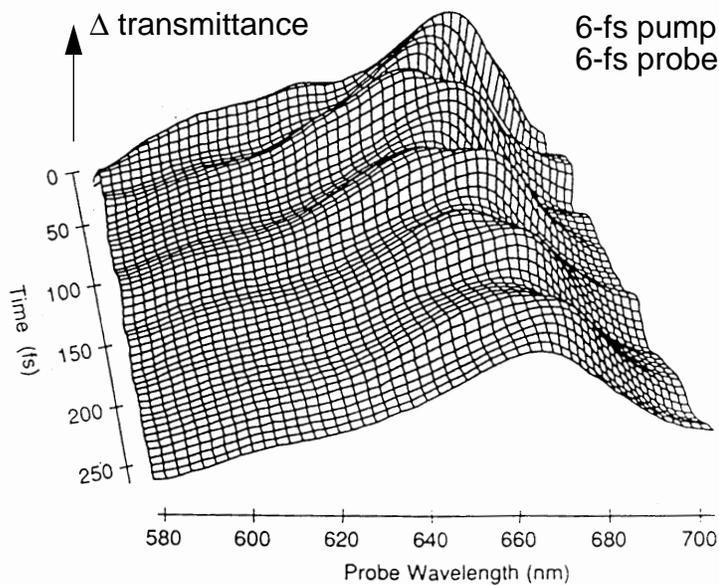


Figure 4. Example of spectroscopy with femtosecond time resolution (from [3]).

time resolution in ultrafast spectroscopy. What are these limits and how does the uncertainty principle enter into such measurements? A related, and perhaps more experimentally concrete question can be asked: if you use a high-resolution spectrometer, do you lose temporal resolution?

## I. B. Limits of Time and Frequency Resolution

Before considering the question of the limits of time and frequency resolution, it may be useful to review the relationship between the time and frequency domains. Any pulse waveform can be represented equivalently in time and frequency domains. A very useful pulse shape is the Gaussian, which closely approximates many real laser pulses and also has the property that its Fourier transform is another Gaussian. A transform-limited pulse is one that has minimal phase variation over its spectrum (in the frequency domain) and has a minimal time-bandwidth product — i.e. it has the shortest possible time duration for a given spectral bandwidth. The electric field of such a pulse can be represented in the time domain as

$$E(t) \sim e^{-\frac{t^2}{2\sigma_t^2}} e^{i\omega_0 t}, \quad (1)$$

and in the frequency domain as

$$E(\omega) \sim e^{-\frac{(\omega - \omega_0)^2}{2\sigma_\omega^2}}, \quad (2)$$

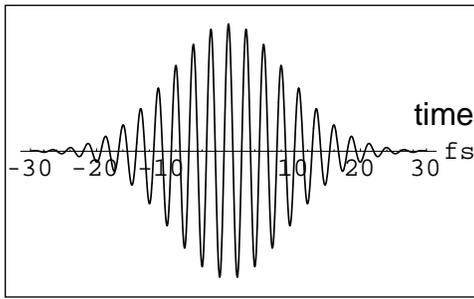
where  $\sigma_t$  characterizes the duration of the pulse,  $\omega_0$  is the center frequency, and  $\sigma_\omega$  characterizes the width of the pulse spectrum. The pulse duration and spectral width are inversely related:  $\sigma_t = 1/\sigma_\omega$ . (Note that while the complex notation for the field is useful for representing amplitude and phase, only the real part of the complex quantity has physical significance.) The time and frequency domain representations are related by Fourier transforms:

$$E(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} E(t) e^{-i\omega t} dt \quad (3)$$

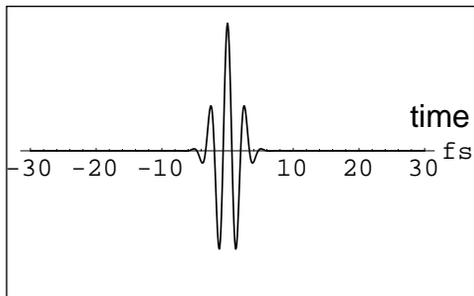
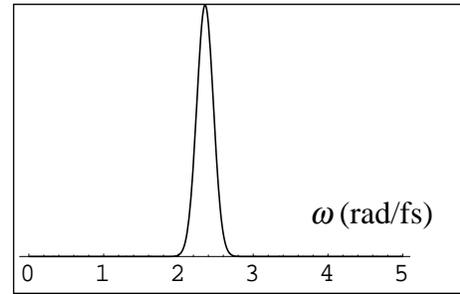
$$E(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} E(\omega) e^{i\omega t} d\omega. \quad (4)$$

These relationships hold generally and are not limited to Gaussian pulses. For illustration, two transform-limited Gaussian pulses are shown in both time and frequency in Fig. 5 and 6. Note that the shorter pulse has a broader spectral width.

Let's now conduct a thought experiment to provide a concrete example for the discussion of time and frequency resolution limits. Consider a pump-probe experiment performed on a material with an instantaneous response, where the medium is totally absorbing except when the presence of the pump pulse opens a window for the probe pulse. (Although bleaching and recovery of absorption in real materials cannot be instantaneous, it is possible to achieve a nearly instantaneous response for example by combining a nonlinear polarization rotation element between two crossed polarizers.) The detector in this experiment can be used to measure either the time-integrated probe intensity as a function of pump-probe time delay, or simply the spectrum of the transmitted probe. Case 1 of Fig. 7 shows the pump-probe setup and both the spectrum of the detected signal and the time delay curve. Now we add a spectral filter to the setup, case 2, which narrows the



**Fourier  
Transform**



**Fourier  
Transform**

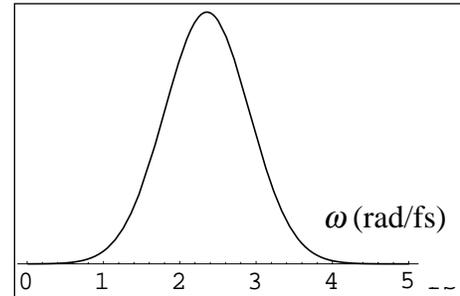


Figure 5. Time representation of two pulses.

Figure 6. Corresponding frequency representations.

spectrum of the probe pulse and correspondingly stretches the probe pulse in time before it is incident on the sample. The stretched probe pulse will produce a longer time-delay curve, as shown in case 2. In case 3, the same filter is placed after the sample and thus after the pump-probe experiment. Clearly the detected probe spectrum will now be as narrow as the transmission spectrum of the filter since it is the last element before the detector. But what about the time-delay curve? Will it be short, as in case 1, or long, as in case 2? And if it is short, have we somehow circumvented the uncertainty principle in getting the original time resolution for a narrower frequency band? It may be tempting to assume that the order of the sample and filter should not matter, and that the results of cases 2 and 3 should be the same. However, that assumption only holds for linear, time-invariant systems; the arrival of the pump pulse at the sample clearly disqualifies the time-invariance assumption.

To answer these questions let's consider in more detail what happens to the probe pulse in cases 2 and 3. Fig. 8 shows the temporal envelope of the probe pulse at several points in the experiment, as well as the pump-probe time-delay curve that would be measured. In case 2, the probe pulse is first stretched by the filter, but then only a short part of it is transmitted through the sample during the 'time window' opened by the pump pulse. (Note that this will actually make the spectral width of the probe broader than the filter transmission width). In case 3, the short pulse is transmitted through the sample and then stretched by the filter. However, the duration of the final probe pulse is not what is being measured. In fact, the duration of the transmitted pulse does not in any way affect the time-delay curve since the intensity of the pulse gets integrated by the detector anyway. Recall that the experimental time resolution is not due to a fast response time of the detector, but instead comes from scanning the pump-probe time-delay  $\Delta t$ . What matters is the duration of the probe pulse at the sample, and thus the time-delay curve in case 2 is long, but in case

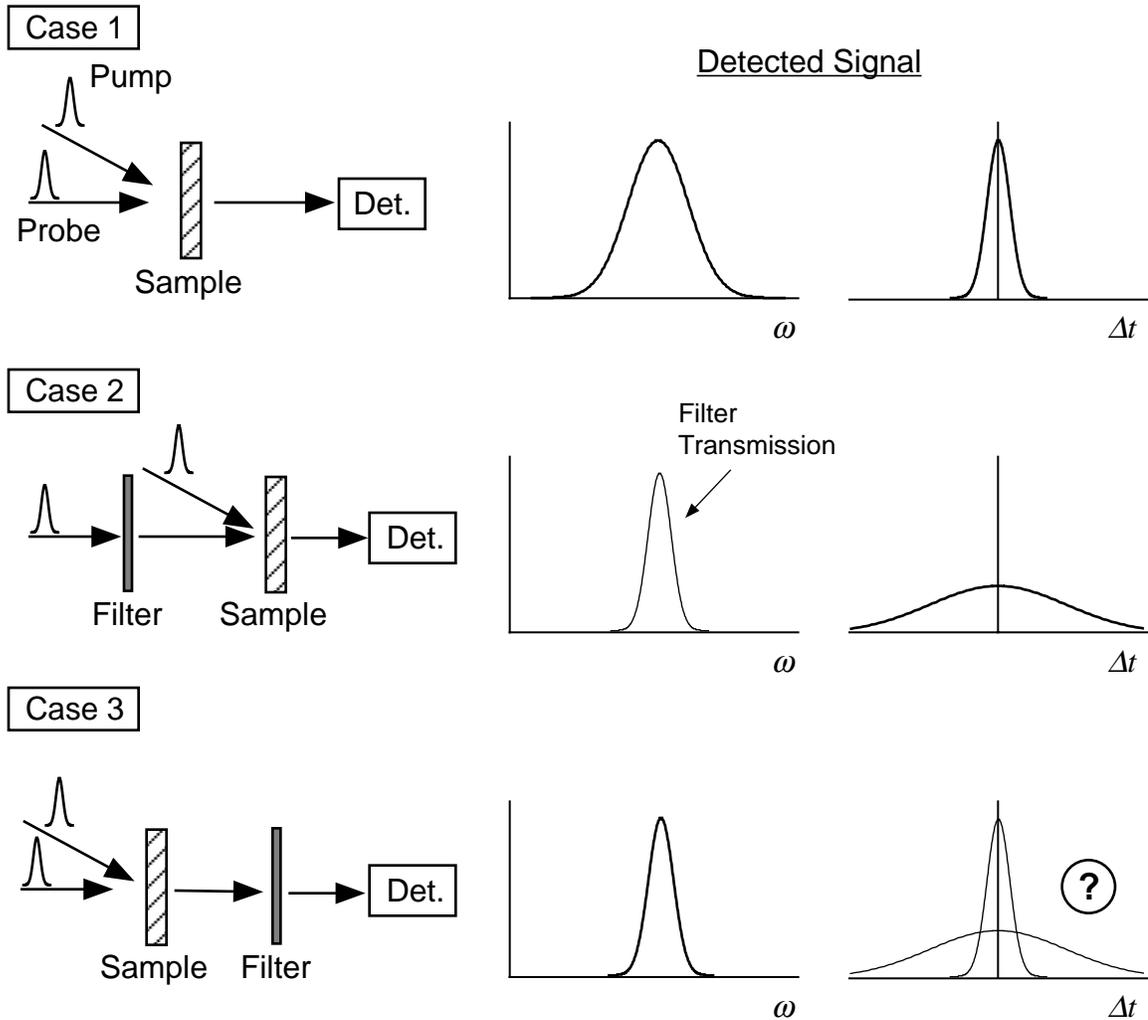


Figure 7. Limits of time and frequency resolution: pump-probe thought experiments.

3 is short. There is no loss in measured time resolution in case 3:  $\Delta\omega_{\text{filter}}$  does not constrain the time resolution, no matter how narrow the filter bandwidth may be! We can replace the filter in case 3 with a spectrometer and an array of detectors, thus obtaining the same time resolution across the entire spectral range of the probe simultaneously. So what really limits the time-frequency resolution?

There is no fundamental limit to the time-frequency resolution in this measurement. Answering the posed question, we can say that using a high-resolution spectrometer does not destroy the temporal resolution in the pump-probe experiment. Instead, we need to focus our attention on the sample being studied, because it is the physical system itself that dictates the appropriate time and frequency resolution in the measurement.

Again let's consider a specific example: a fast, but not instantaneous, real saturable absorber as the sample in our thought experiment. The action of the saturable absorber is illustrated in Fig. 9. Before the arrival of the pump pulse the electrons are in their ground state, during the pulse they are promoted to higher lying states saturating the ability of the

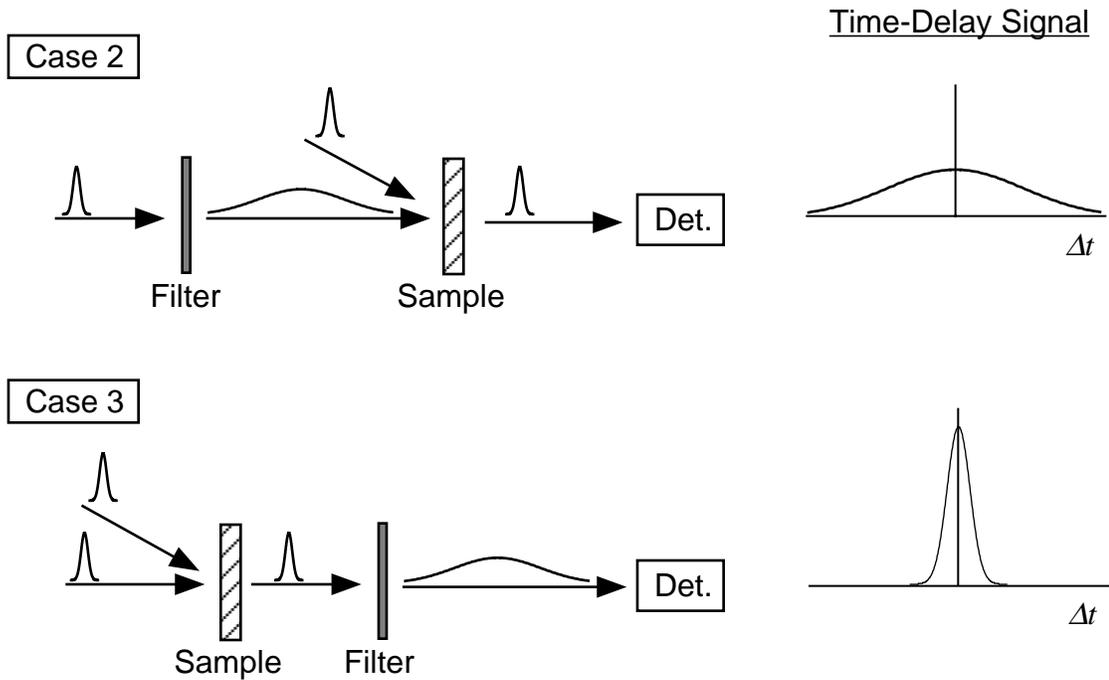


Figure 8. Pump-probe thought experiments: case 2 and 3 in more detail.

medium to absorb photons, and shortly after the pump pulse they return to the ground state. Fig. 10 shows two sets of hypothetical measurements performed with 5-fs pump and probe pulses. At ‘negative time-delay’ (–10 fs) the probe pulse arrives before the pump and thus sees the full absorption spectrum of the medium. At zero time-delay, the probe sees the full effect of the pump, which in the first case is a broad spectral hole in the absorption spectrum, whereas in the second case it is a set of spectrally narrow lines. At 10 fs after the pump pulse, the absorption has returned to its original state. It would appear that the second situation is an ideal case for employing a high resolution spectrometer in an ultrafast measurement. However, the second situation is impossible. A system cannot have sharp spectral features if it has such fast relaxation. This is why it is the system itself that dictates the appropriate time and frequency resolution in the measurement. A medium with a fast response must have a broad spectrum and thus a measurement would not profit from high frequency-resolution; similarly, sharp spectral features necessarily evolve slowly, and thus do not benefit from ultrafast time-resolution.

According to the above discussion, there is no fundamental limit on the simultaneous frequency and time resolution of an optical measurement. The final issue that was raised in Section I. A is how does the uncertainty principle enter into such measurements? In quantum mechanics, the ‘uncertainty’ in the uncertainty principle is a result of the inherently probabilistic nature of the theory, i.e. the fundamental variation in the measurement of physical observables represented by operators that do not commute. While there are mathematical parallels between quantum mechanics and time-frequency signal analysis, the analogy cannot be taken too far. In a deterministic theory, such as the classical theory of electromagnetism, there really is no ‘uncertainty principle’, but rather just a time-bandwidth relation which says that a waveform and its Fourier transform cannot be made

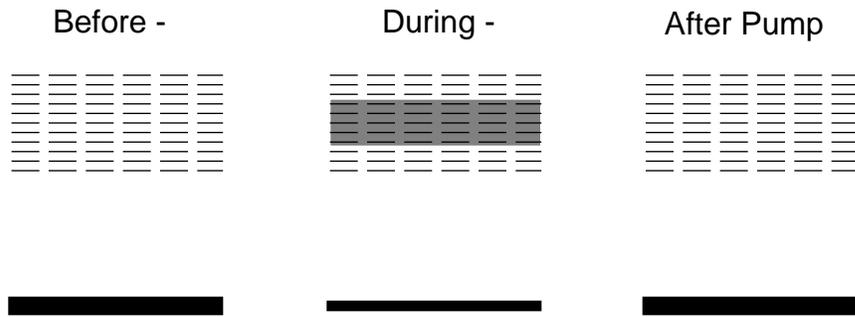


Figure 9. Very fast saturable absorber.

'Measurement': 5-fs pump/ 5-fs probe

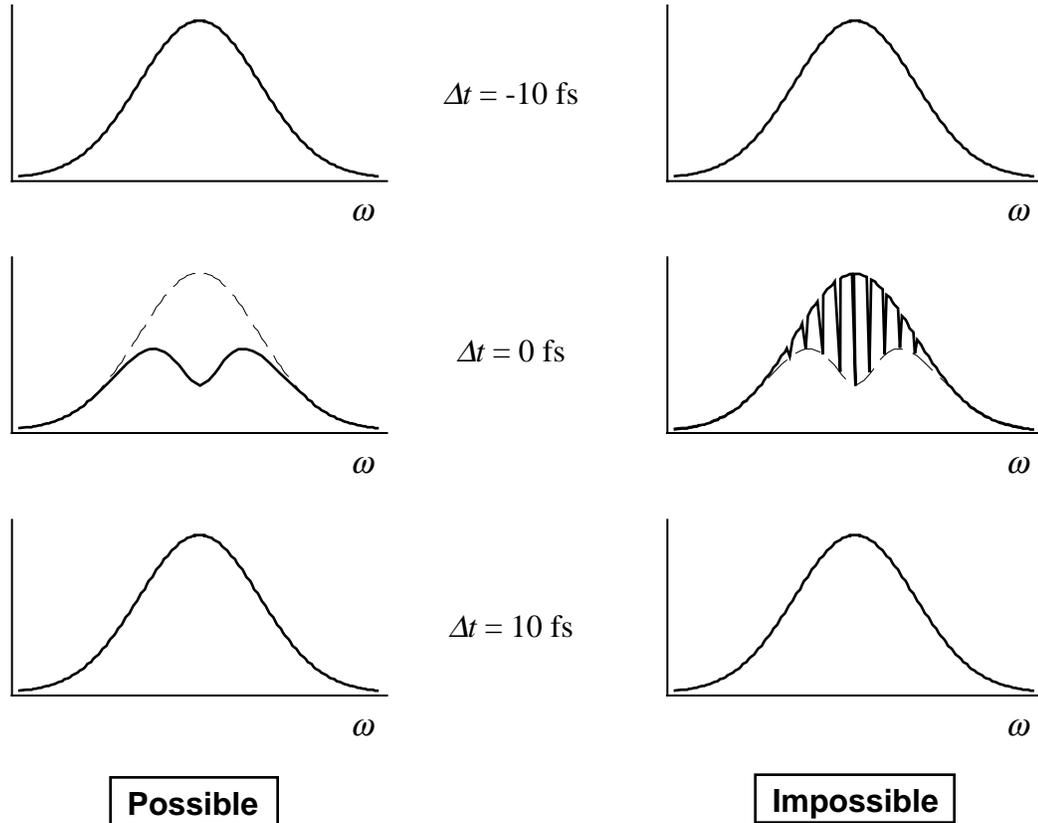


Figure 10. Hypothetical measurements performed with 5-fs pump and probe pulses on a very fast saturable absorber.

arbitrarily narrow simultaneously. But the waveform itself can be known exactly, and the resolution in a joint time-frequency representation (such as the one shown in Fig. 4) is unlimited. For a detailed and insightful review of joint time-frequency analysis and the connection to quantum mechanics see the review by Cohen [4].

### I. C. Propagation of Light Pulses

The propagation of ultrashort pulses is governed by the same equations as the propagation of any light. However, some features of nonlinear and even linear propagation are more evident or more extreme for ultrashort pulses. In this section we will look at the effects of dispersion on a pulse, the meaning of group velocity in ultrashort pulse propagation, and the nonlinear effects of self-phase-modulation and self-focusing. We will also review the Wigner representation, which is useful in describing a pulse simultaneously in time and frequency.

**1. Dispersion.** Propagation of electromagnetic radiation is governed by the wave equation which is derived by combining Maxwell's equations. The wave equation for the electric field is

$$\nabla^2 \mathbf{E} - \frac{\mu \varepsilon \partial^2 \mathbf{E}}{c^2 \partial t^2} = 0, \quad (5)$$

and has the solutions

$$\mathbf{E} = \mathbf{E}_0 e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}, \text{ where } k = \frac{\sqrt{\mu \varepsilon}}{c} \omega. \quad (6)$$

The electric field of the wave induces a polarization in the medium, proportional to the field,  $P = \chi E$ , where  $\chi$  is the linear susceptibility and describes both the strength and phase of the polarization response. The electric field in the medium consists of the displacement field  $D$ , minus the response of the bound charges:  $E = D - 4\pi P$ . Writing  $D$  in terms of  $E$ ,  $D = E + 4\pi P = (1 + 4\pi\chi) E = \varepsilon E$ , where  $\varepsilon$  is the dielectric constant. The dielectric constant is really not a constant, but a function of frequency, and characterizes the optical response of the medium. The complex index of refraction is defined as

$$\tilde{n} \equiv \sqrt{\mu \varepsilon}, \quad (7)$$

and if the magnetic permeability is  $\mu = 1$ ,

$$\tilde{n} \equiv \sqrt{\varepsilon}. \quad (8)$$

The real and imaginary parts of  $k$  correspond to dispersion and absorption,

$$k = \frac{\tilde{n} \omega}{c} = \beta + i \frac{\alpha}{2}, \quad (9)$$

where  $\alpha$  is the absorption coefficient, and is defined so that it gives the exponent of the decay of the intensity with propagation distance:

$$I = I_0 e^{-\alpha z}. \quad (10)$$

The dispersion of the medium is determined by  $\beta$ . If we consider propagation through a transparent medium of length  $L$ , then the field (in the frequency representation) will acquire the following frequency dependent phase:

$$E_{\text{out}}(\omega) = e^{i\phi(\omega)} E_{\text{in}}(\omega) = e^{i\beta(\omega)L} E_{\text{in}}(\omega). \quad (11)$$

It is useful to Taylor expand  $\beta$  near  $\omega_0$  to examine the effects of the coefficients on the pulse shape:

$$\beta(\omega) = \beta_0 + \beta' \Big|_{\omega_0} (\omega - \omega_0) + \frac{1}{2} \beta'' \Big|_{\omega_0} (\omega - \omega_0)^2 + \dots \quad (12)$$

The zero-order coefficient,  $\beta_0$ , is just a fixed phase shift that has no effect on the pulse shape. Notably,  $\beta'$  also has no effect on pulse shape, and is equivalent to propagation through vacuum. So  $\beta''$  is the first term that alters the pulse shape; its effect on a transform-limited pulse is to introduce a ‘chirp’ — a spreading in time of the spectral components of the pulse. An example of the effect of chirp on a pulse is shown later in the top plot of Fig. 40.

**2. Group Velocity.** The group velocity determines the speed of propagation of the energy in a wave of a given frequency, and is thus an important physical quantity. It is defined as

$$V_g = \frac{\delta\omega}{\delta k} = [\beta']^{-1}. \quad (13)$$

We can put this into more useful form by differentiating  $k = n\omega/c$

$$\frac{\delta k}{\delta\omega} = \frac{1}{c} \left[ n + \frac{\delta n}{\delta\omega} \cdot \omega \right], \quad (14)$$

giving

$$V_g = \frac{c}{n + \frac{\delta n}{\delta\omega} \omega}, \quad (15)$$

or in terms of wavelength,

$$V_g = \frac{c}{n - \frac{\delta n}{\delta\lambda} \lambda}. \quad (16)$$

The two dispersion regimes,  $\delta n/\delta\omega > 0$  and  $\delta n/\delta\omega < 0$  are called normal and anomalous dispersion, respectively. In the anomalous regime,  $V_g$  can become infinite. This occurs near a sharp resonance, and brings into question the physical meaning of an infinite group velocity. Of course energy does not propagate at an infinite velocity, so we must examine the assumptions behind the definition of  $V_g$ . In fact, the definition is only an approximation, and it breaks down when  $k(\omega)$  varies too rapidly. The approximation requires that

$$\frac{\delta^2 k}{\delta\omega^2} \omega \ll \frac{\delta k}{\delta\omega}. \quad (17)$$

An explicit derivation of the group velocity of a waveform based on a plane wave

decomposition can be found in [5].

**3. Self-Phase-Modulation.** Because of the high intensity of ultrashort pulses, nonlinear effects often play an important role in their propagation. Expanding the induced polarization in terms of field strength,

$$P = \chi^{(1)} E + \chi^{(2)} EE + \chi^{(3)} EEE + \dots, \quad (18)$$

the first term corresponds to ordinary linear propagation, and the rest describe the nonlinear response which introduces intensity dependent propagation and also couples together different frequencies. In general, the fields are vectors and the susceptibility coefficients are tensors. For example, the third order susceptibility,  $\chi^{(3)}$ , is a fourth order tensor. The range of nonlinear phenomena is very extensive, but here we will focus on just one particular third-order nonlinearity that leads to an intensity dependent modulation of the index of refraction.

Consider the following element of the third-order polarization:

$$\begin{aligned} P^{(3)}(\omega) &= 3\chi^{(3)} E(\omega) E^*(\omega) E(\omega) \\ &= 3\chi^{(3)} I(\omega) E(\omega) \end{aligned} \quad (19)$$

where  $I \equiv |E|^2$ . Now combining this with the linear response, gives

$$\begin{aligned} P &= P^{(1)} + P^{(3)} \\ &= (\chi^{(1)} + 3\chi^{(3)} I) E \\ &= \chi_{\text{eff}} E \end{aligned} \quad (20)$$

where  $\chi_{\text{eff}}$  is an intensity-dependent ‘effective susceptibility’. The index of refraction correspondingly becomes

$$n = \sqrt{\epsilon} = \sqrt{1 + 4\pi\chi_{\text{eff}}}, \quad (21)$$

and can be expanded about its ordinary value if the nonlinear contribution is relatively small:

$$\begin{aligned} n &\approx \sqrt{1 + 4\pi\chi^{(1)}} + \frac{1}{2} \frac{4\pi 3\chi^{(3)} I}{\sqrt{1 + 4\pi\chi^{(1)}}} \\ &= n_0 + 6\pi \frac{\chi^{(3)}}{n_0} I \\ &= n_0 + n_2 I \end{aligned} \quad (22)$$

This phenomenon is called the Kerr effect, and  $n_2$  is referred to as the nonlinear index. The intensity dependence of the index of refraction produces a phase modulation throughout the duration of the pulse. We can calculate the phase delay,  $-\phi$ , due to propagation through a medium of length  $L$ , by relating it to the effective optical path,  $nL$ ,

$$\frac{-\phi}{2\pi} = \frac{nL}{\lambda} \quad (23)$$

$$\phi = \frac{-2\pi}{\lambda} L (n_0 + n_2 I) . \quad (24)$$

The intensity-dependent phase modulation produces a time-dependent frequency shift

$$\Delta\omega = \frac{d\phi}{dt} = \frac{-2\pi}{\lambda} L n_2 \frac{dI}{dt} . \quad (25)$$

Figure 11 illustrates the frequency shift for a Gaussian pulse. For positive  $n_2$ , which is the

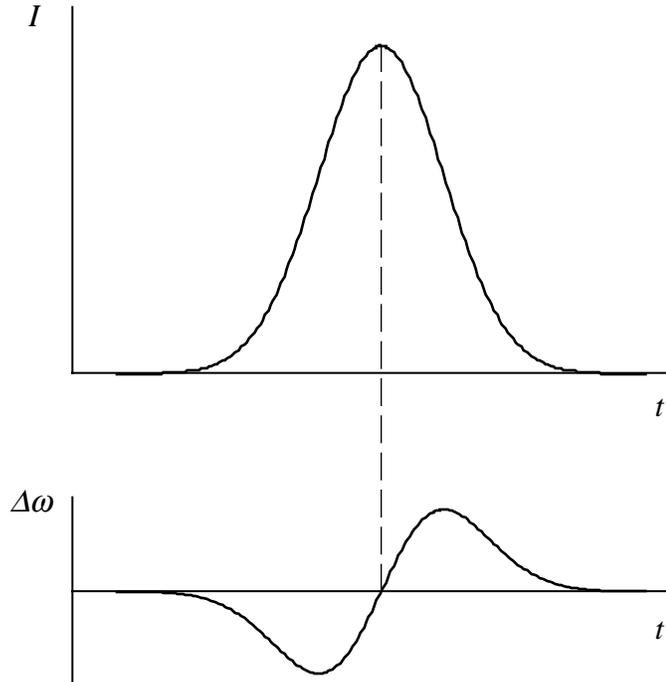


Figure 11. Self-phase-modulation.

case for most materials, the front part of the pulse is red-shifted (to lower frequencies), while the back part is blue-shifted. This results in a nearly linear frequency chirp across the central part of the pulse. Self-phase-modulation plays an important role in mode-locking dynamics in ultrafast lasers, in propagation of solitons in optical fibers, and in white-light generation.

**4. Self-Focusing.** A closely related phenomenon to self-phase-modulation is self-focusing. Its source is the same third-order nonlinearity, i.e. the Kerr effect, except that it is the *spatial* variation in intensity, rather than temporal, that produces self-focusing. If the beam profile is given by  $I(r)$ , then the index of refraction is simply

$$n = n_0 + n_2 I(r) . \quad (26)$$

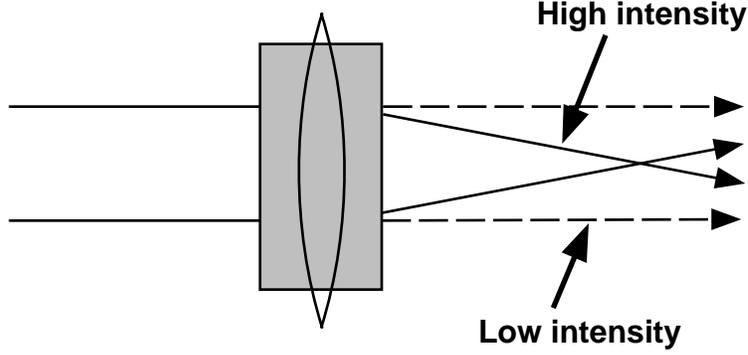


Figure 12. Self-focusing.

A Gaussian beam propagating through this medium creates a lens for itself, as illustrated schematically in Fig. 12. Near its center, a Gaussian can be approximated by

$$I(r) = I_0 e^{-\frac{r^2}{a^2}} \approx I_0 \left( 1 - \frac{r^2}{a^2} \right), \quad (27)$$

where  $a$  characterizes the beam width. This intensity distribution creates a quadratic variation in the index of refraction, resulting in a lens-like effect:

$$n(r) = n_0 + n_2 I_0 \left( 1 - \frac{r^2}{a^2} \right). \quad (28)$$

Self-focusing is used in Kerr-lens mode locking, described in Section II. A.2.

**5. Joint Time-Frequency Representation.** While either time or frequency representation of a pulse offers a complete description, in some cases it can be insightful to use a joint time-frequency representation. Several such representations exist, (see for example [4]), but here we'll describe one that's referred to as the Wigner representation [6] or more specifically, the chronocyclic representation when it is applied to electromagnetic pulses (for a comprehensive review see [7]). The Wigner function is calculated from the field as represented in either frequency or time:

$$\begin{aligned} W(t, \omega) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} E\left(\omega + \frac{\omega'}{2}\right) E^*\left(\omega - \frac{\omega'}{2}\right) e^{-i\omega't} d\omega' \\ &= \int_{-\infty}^{\infty} E\left(t + \frac{t'}{2}\right) E^*\left(t - \frac{t'}{2}\right) e^{-i\omega t'} dt' \end{aligned} \quad (29)$$

The Wigner function attempts to represent the spectral and temporal distribution of the energy, or the 'chronocyclic intensity', as a two-dimensional real function (although it can take on negative values), while treating time and frequency on an equal footing. Some of its properties include: a time integral gives the intensity spectrum,

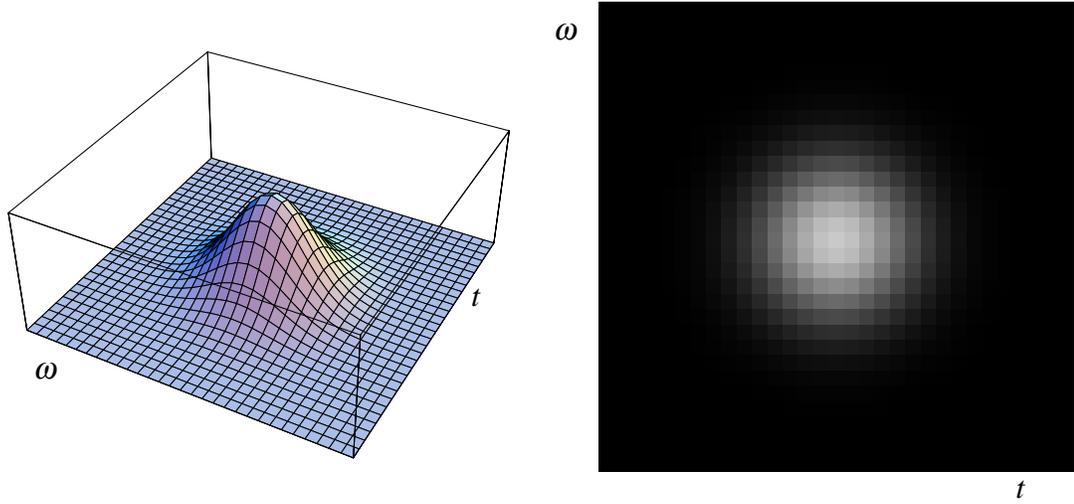


Figure 13. Wigner representation of a Gaussian pulse as a surface in 3-D and as a density plot.

$$\int_{-\infty}^{\infty} W(t, \omega) dt = |E(\omega)|^2 = I(\omega), \quad (30)$$

integrating in frequency gives the intensity profile,

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} W(t, \omega) d\omega = |E(t)|^2 = I(t), \quad (31)$$

and the area integral in both time and frequency gives the energy in the pulse,

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(t, \omega) dt d\omega = \text{Energy} \quad . \quad (32)$$

In order to satisfy the limits imposed by Fourier relations, the Wigner function must be non-zero in a phase-space area larger than  $\pi$ . A transform-limited Gaussian pulse is shown in the Wigner representation in Fig. 13 as a 3-D plot and as a density plot. Figure 14 shows a chirped pulse, and demonstrates the immediate graphic appreciation of the pulse properties when seen in this representation. Finally, the Wigner representation provides a close connection to the joint time-frequency measurements discussed in Section III. A.

## II. PRODUCING FEMTOSECOND PULSES

### II. A. Femtosecond Lasers

Today there are many types of lasers that produce femtosecond pulses. They have some common principles of operation that appear in a variety of real embodiments. A femtosecond laser cavity, shown schematically in Fig. 15 contains three essential elements: a gain medium, a mode-locking element, and an element that compensates for group velocity dispersion (GVD) in the rest of the cavity. The gain medium has a broad gain bandwidth to allow a large number of modes (frequencies) to lase simultaneously. When

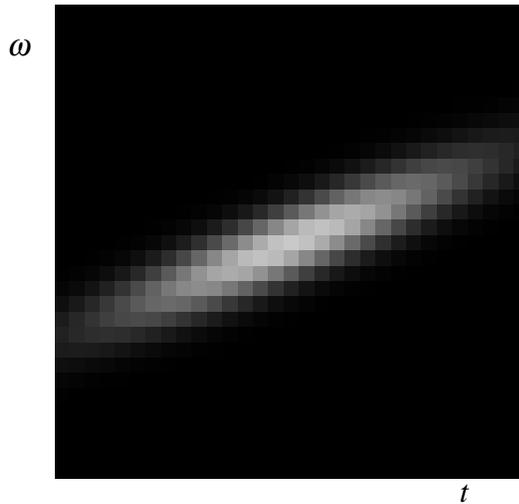


Figure 14. Wigner representation of a chirped pulse.

properly locked in phase, these modes add up to a short pulse that circulates in the cavity. A fraction of this pulse is ejected out of the cavity through an output coupler, providing a train of pulses spaced by the round-trip time in the cavity.

Two examples of real femtosecond laser systems are shown in Fig. 16a & b. Fig. 16a

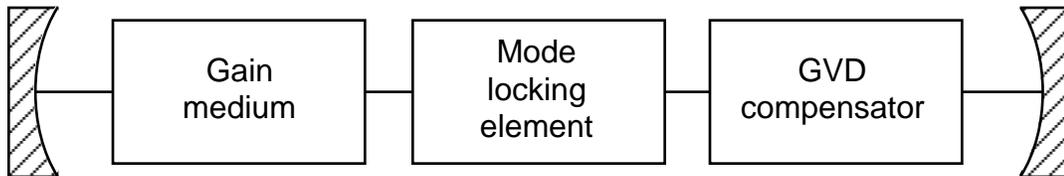


Figure 15. Basic elements of a femtosecond laser cavity.

is a dye-based laser system, which was the first type of laser to produce ultrashort pulses [8]. The gain medium is a jet of an organic dye in solution, most commonly Rhodamine 6G in ethylene glycol, pumped continuously by an Argon-Ion laser. The mode locking is provided by another jet of a different dye which acts as a saturable absorber. Finally, a sequence of prisms provides the GVD compensation. Fig. 16b is a solid-state based laser, where the gain medium is a doped crystal, in this case Titanium-doped Sapphire ( $\text{Ti:Al}_2\text{O}_3$ ). This type of laser was first found to produce mode-locked pulses in 1991 [9], and since then there has been tremendous growth in solid-state ultrashort-pulse lasers based on a wide variety of materials, cavity types and geometries, and mode-locking methods. In the example shown in Fig. 16b, the  $\text{Ti:Al}_2\text{O}_3$  crystal provides both the gain and mode locking in the laser. The prisms again provide the GVD compensation. In the following sections we will examine in more detail gain, mode locking, and GVD compensation in an ultrashort-pulse laser.

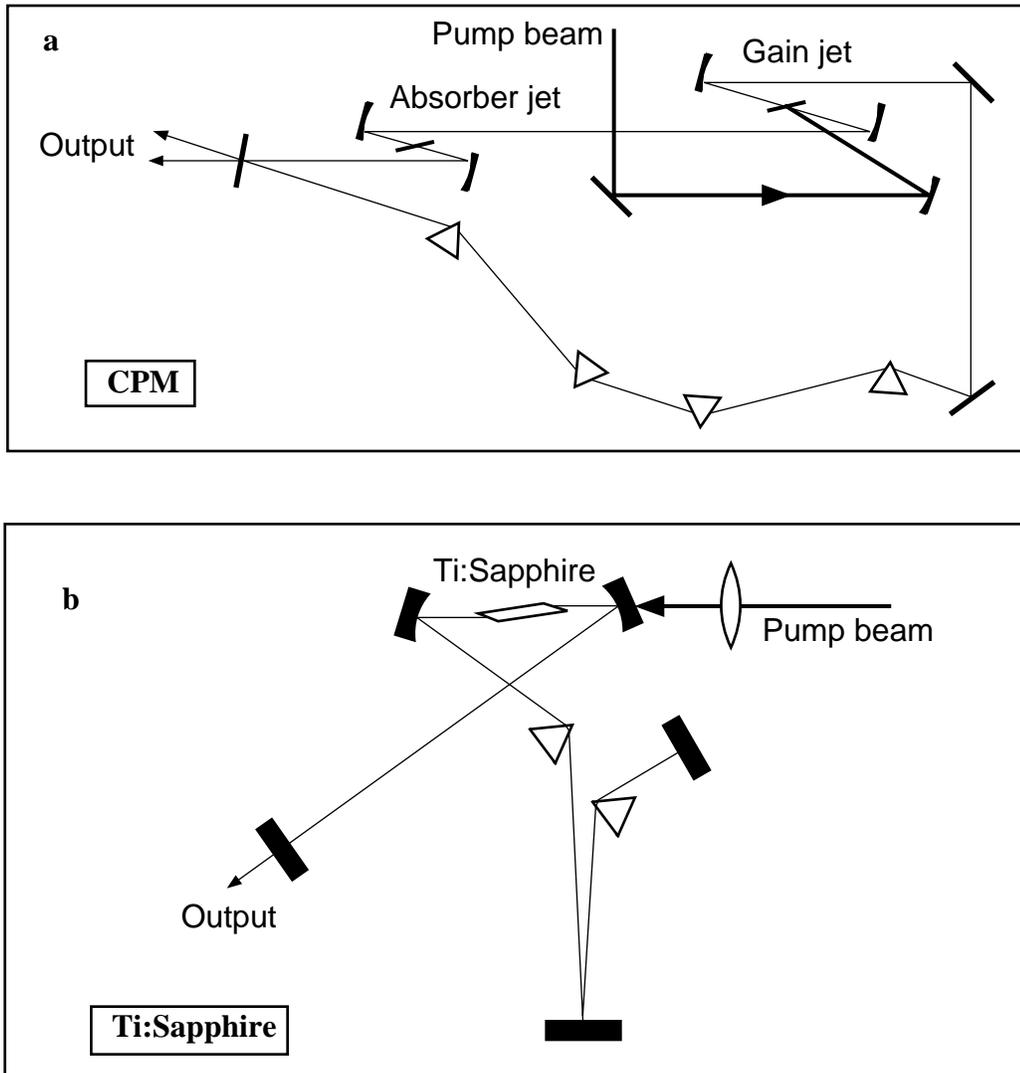


Figure 16. Two examples of a femtosecond laser: **a.** dye-based laser; **b.** solid-state laser.

**1. Gain Medium.** The first requirement for the gain medium is that it have a broad gain spectrum. A narrow-band gain medium could not possibly produce ultrashort pulses, because a wide range of Fourier components is required to add up to a short pulse. A figure of merit has been suggested for comparing different gain media [10]:

$$M = \sigma\tau\Delta\nu, \quad (33)$$

where  $\sigma$  is the peak stimulated emission cross-section,  $\tau$  is the lifetime of the inverted population and  $\Delta\nu$  is the width of the fluorescence line. The larger the product  $\sigma\tau$ , the lower the pumping threshold. Note that the lifetime of the inverted population does not need to be short in order to produce ultrashort pulses. The larger the  $\Delta\nu$ , the shorter the pulse duration limit or alternatively, if the bandwidth is broader than necessary for the desired pulse duration, it can be used to provide tunability in frequency. Other desirable features in a gain medium include: convenient absorption band for pumping, photochemical stability,

and of course mechanical stability.

As a contrast to ultrashort-pulse lasers, it may be useful to review gain in a CW (continuous wave) laser. The gain and loss spectra in an ideal CW laser are shown in Fig. 17. A number of cavity modes are shown within the gain spectrum, spaced by  $\Delta\omega = 2\pi(c/2L)$ , where  $2L$  is the round-trip length of the cavity. The shown gain spectrum assumes that the lasing line is homogeneously broadened, so that spectral hole burning is not possible, and the line shape is fixed. A further assumption is that there is no spatial hole burning, so that the gain spectrum is constant throughout the gain medium. The main observation is that in steady state, the highest-gain mode cannot exceed the loss, and thus there will only be a single lasing mode! The other modes will see a gain smaller than the loss, and will not lase.

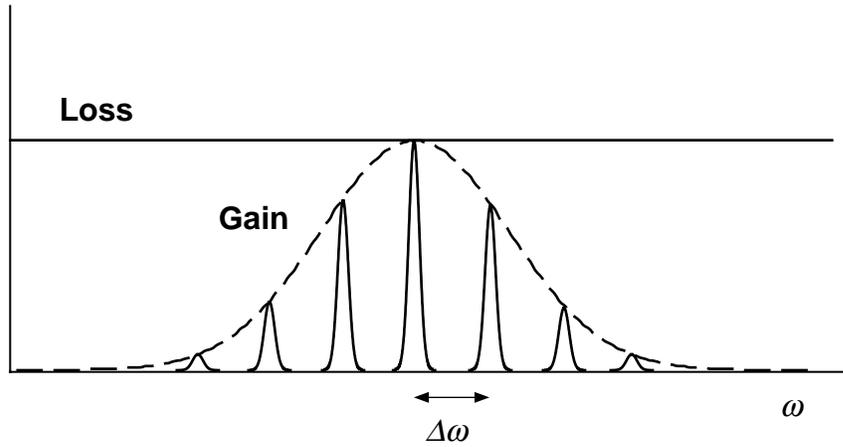


Figure 17. Gain and loss spectra in a CW laser.

Single frequency operation is desirable for narrow-linewidth CW laser, but is quite the opposite to what is required for producing ultrashort pulses. A femtosecond laser must have many lasing modes (typically on the order of  $10^4$ - $10^5$ ), and these modes must all be locked in phase.

**2. Mode Locking.** The term ‘mode locking’ refers to establishing a desired fixed phase relationship between the different frequency modes in the laser cavity. The field can be written in the frequency domain as a sum of  $n$  cavity modes,

$$E(\omega) = \sum_n E_n(\omega_n) e^{i\phi_n}. \quad (34)$$

If the  $\phi_n$  are random, then the laser output will just be a continuous fluctuating signal. If the  $\phi_n$  are all zero on the other hand, then the modes will add up constructively to form a short pulse oscillating in the cavity. An additional effect of mode locking is to broaden the lasing spectrum to include a much larger number of modes.

Although ‘mode locking’ suggests a frequency domain perspective, the frequency description is not convenient for dealing with the propagation of ultrashort pulses inside the cavity because a very large number of modes is involved and they are nonlinearly coupled. For this reason a time-domain description is more useful and insightful.

The mode locking discussion in this section follows an excellent review article on the subject by Ippen [11]. The essential principle in mode locking is providing a mechanism for time dependent modulation of loss and/or gain. This modulation may be active, for example by including an externally driven modulator in the cavity, or it can be passive, by having an element in the cavity that lets the pulse itself vary the gain or loss in time so that the laser prefers to operate in a pulsed mode.

Figure 18 shows the gain and loss in an actively modulated cavity as a function of time. In the time window where the gain exceeds the loss a pulse will form as a result of the modulation. However, active mode locking is not used to produce ultrashort pulses because the speed of modulation is essentially limited by the response time of the modulator.

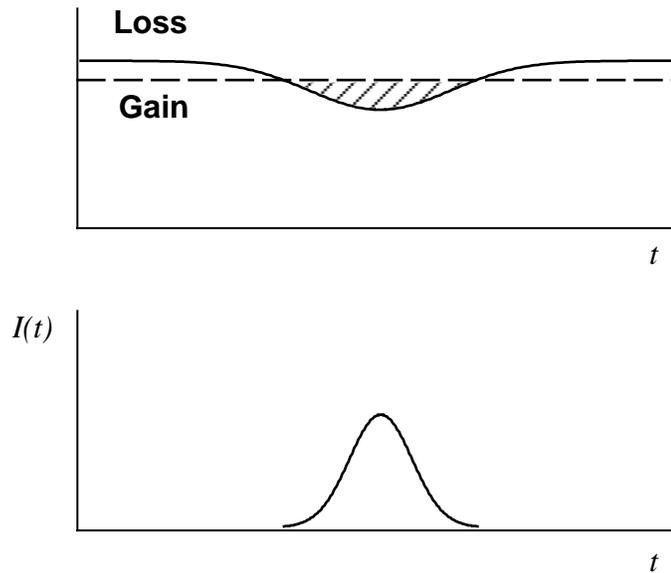


Figure 18. Active mode locking.

The basic approach in passive mode locking is to use a saturable absorber — an element whose ability to absorb is reduced at high intensity. The effect of the saturable absorber is to favor the propagation of a short burst of light over continuous lasing in the cavity. There are two classes of saturable absorbers, ‘slow’ and ‘fast’, characterized by their response time relative to the duration of the ultrashort pulse that is produced. In both cases, the response time is fast relative to the cavity round-trip time.

In the case of a slow saturable absorber, its action needs to be supplemented by the depletion of gain in the gain medium. Let's assume that a pulse somehow forms in the cavity, and examine the resulting gain and loss in the cavity. (Initial pulse formation will be discussed later). Figure 19 shows the pulse intensity envelope together with the gain and as a function of ‘local time’,  $t-x/v_g$ , so that the combined effect of two elements can be shown even though they are spatially and temporally separated. Despite the ‘slow’ recovery of both the absorption and the gain, the combined action leads to a fast modulation of the difference, as shown in the third graph.

The saturation of absorption results from exciting a large fraction of the valence electrons before they have a chance to relax to the ground state. Similarly, the depletion of

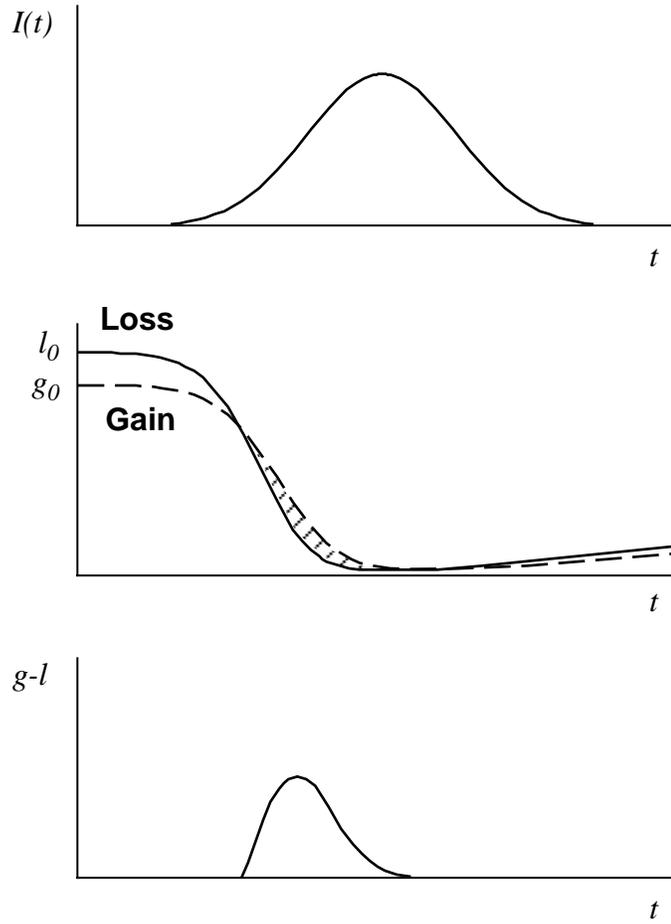


Figure 19. Passive mode locking with the combined action of ‘slow’ saturable absorption and gain depletion.

the gain results from stimulating a large fraction of the inverted population to radiatively recombine, before the inversion can be reestablished by the pump source. Assuming the recovery is slow, the saturation of absorption and the depletion of the gain can be modeled as

$$l(t) = l_0 e^{-\sigma_a \int_{-\infty}^t |E(t')|^2 dt'} \quad (35)$$

$$g(t) = g_0 e^{-\sigma_g \int_{-\infty}^t |E(t')|^2 dt'} \quad (36)$$

where  $l_0$  and  $g_0$  are the initial (or ‘low intensity’) loss and gain, and  $\sigma_a$  and  $\sigma_g$  characterize the saturation and depletion. The requirements for mode locking are that  $l_0 > g_0$ , so that there is no continuous lasing, and  $\sigma_a > \sigma_g$ , so that the absorption is saturated before the gain is depleted.

In the case of a fast saturable absorber, the gain medium does not need to play a role in the mode locking. This greatly expands the range of gain media that can be used in ultrafast lasers. Figure 20 shows a pulse and the resulting gain and loss in a laser with a fast saturable absorber and no depletion of gain.

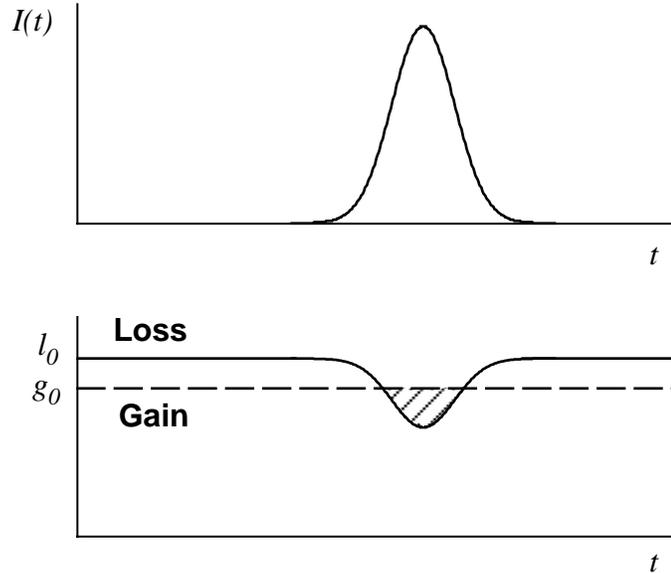


Figure 20. Fast saturable absorber.

In the limit of instantaneous response time of the absorber, the decrease in the loss simply follows the pulse intensity:

$$l(t) = l_0 - \gamma |E(t)|^2. \quad (37)$$

However, no real absorber has an instantaneous response. Absorption consists of electrons making transitions to excited states, and it takes time for them to relax to the ground state (at least a few picoseconds). Thus it would seem that the recovery timescale precludes the possibility of a nearly instantaneous saturable absorber, like the one in Fig 20. However, a clever new approach has been invented to get around the recovery time limit. It was first discovered experimentally [9], and even called ‘magic mode locking’ before the mode locking mechanism was understood.

The trick is to use nonlinear phase modulation and convert it into amplitude modulation. Phase modulation can be nearly instantaneous because it does not require absorption; by using a fast, nonresonant nonlinearity it is possible to modulate the real part of the susceptibility without affecting the imaginary part. For example, self-focusing and self-phase-modulation in a transparent material do not induce absorption — the imaginary part of the susceptibility remains zero. But how can we convert the induced changes in the index of refraction into amplitude modulation to produce an artificial ‘instantaneous saturable absorber’? One approach is to combine self-focusing with a physical aperture in such a way as to select the self-focused propagation mode of a short intense pulse over the non-self-focused mode of continuous lasing. This type of mode locking is called Kerr-lens mode locking (KLM) and is the mechanism responsible for ‘magic mode locking’.

Figure 21 shows two possible arrangements combining intensity-dependent focusing with an aperture [12]. The ‘lens’ is not a real lens, but the effect of self-focusing, and thus is only seen by the short, intense pulse; the low intensity continuous radiation is not affected. Of the two modes shown in each of the two cases, the mode for the pulse is more focused than the mode for continuous lasing. Both cases produce the smallest loss for the ultrashort-pulse mode and mimic the action of an almost instantaneous saturable absorber. Note that the nonlinear effect itself is not dissipative — amplitude modulation is achieved by coupling with other elements such as an aperture. An alternative configuration uses the pump laser beam as a gain aperture, avoiding the need for a loss aperture.

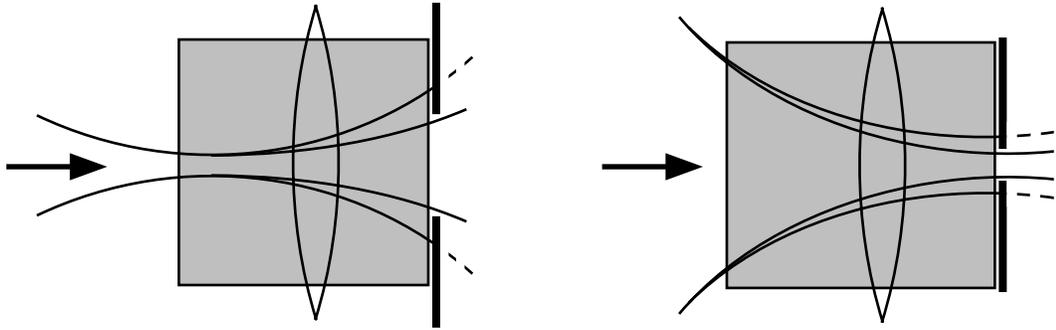


Figure 21. Examples of KLM with combinations of a self-focusing element and an aperture.

Self-phase-modulation is used in other mode-locking schemes with coupled cavities or cross-polarized modes in the same cavity, and is called additive pulse mode locking (APM) or coupled cavity mode locking (CCM). The most significant application of these mode-locking techniques has been in fiber-based lasers because self-focusing cannot be used in fibers, and also because the long fiber length, cavity stability, and good polarization control makes fiber lasers good candidates for APM. For more information on various types of mode locking see [11].

In our discussion of passive mode locking we have considered how a short pulse affects the loss and gain in the laser cavity. The discussed conditions are necessary for the ultrashort pulse to be a stable solution, but we have not considered whether the pulse will evolve from initial fluctuations. The ability of a laser to mode lock itself is called ‘self-starting’. Whether a laser is self-starting depends on whether an initial fluctuation will shorten within a cavity coherence time. If this does not happen, the fluctuation will be dispersed through random phase fluctuations, and the laser will not be self-starting. Pulse-shortening rate, a non-dimensional quantity, can be used to illustrate pulse duration limits and also to characterize the ability of a mode-locking mechanism to self-start. Figure 22 shows the pulse-shortening rate,  $-\Delta\tau/\tau$ , vs.  $1/\tau$  for different mode-locking mechanisms (from [11]). Active mode locking is most effective for long pulses (small  $1/\tau$ ), whereas a fast saturable absorber is most effective for short pulses (large  $1/\tau$ ). So while a fast saturable absorber is ideal for sustaining ultrashort pulses, it usually requires an external perturbation to initiate the mode locking. In practice, this perturbation may be provided by the vibration resulting from a light tap on the laser platform, or a more sophisticated mechanism with feedback control that will automatically restart the mode locking if for some reason it is interrupted.

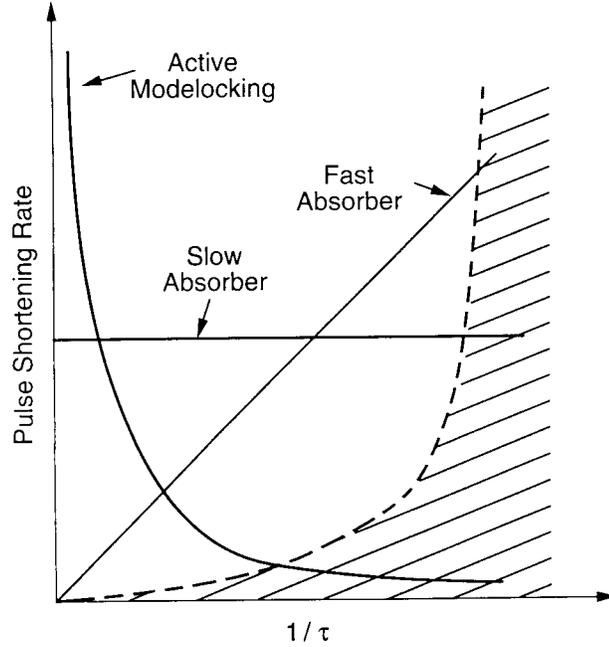


Figure 22. Pulse-shortening rate for different mode-locking mechanisms (from [11]).

**3. GVD Compensation.** In order to produce ultrashort pulses it is necessary to minimize the group velocity dispersion (GVD) in the laser cavity, so that the broad range of spectral components in the ultrashort pulse can all propagate at the same velocity (averaged over a cavity round-trip). There are two effects that require compensation: the material dispersion in the cavity (a linear effect), and self-phase-modulation (a nonlinear effect).

Material dispersion in a solid state laser is usually dominated by the gain medium, but also includes contributions from all other elements in the cavity. Figure 23 shows the group

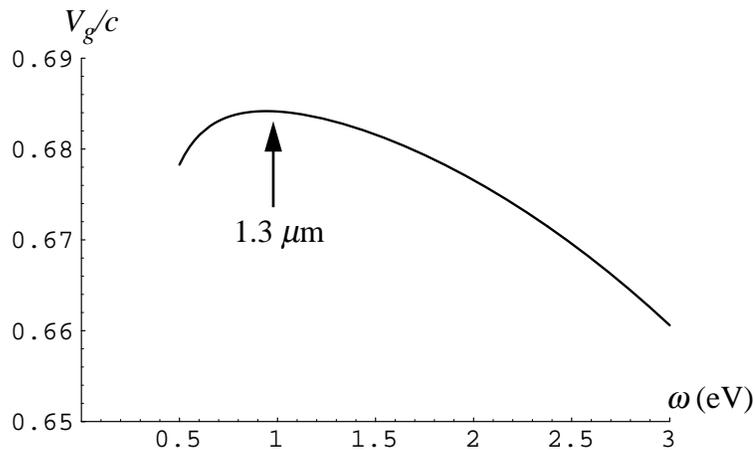


Figure 23. Group velocity in fused silica glass.

velocity in fused silica glass as a function of frequency (expressed in units of photon energy). Other transparent materials produce very similar curves. GVD is just the derivative of the group velocity,  $dV_g/d\omega$ . The dispersion results from resonances that lie outside of the transparent range; the negative slope at visible frequencies is caused by electronic transitions further out in the ultraviolet, while the positive slope on the infrared side is due to optically active vibrational transitions in the far infrared. A balance between these effects occurs at the peak of the group velocity, where GVD is zero, at  $1.3 \mu\text{m}$  (0.9 eV) in fused silica. For lasers operating in the visible or in the very near infrared, the GVD causes the ‘blue’ (higher frequency) part of the spectrum to travel more slowly than the ‘red’ (lower frequency) part.

Self-phase-modulation introduces a frequency depended phase shift, which has an effect similar to group velocity dispersion (in addition to producing new spectral components). Because it is a nonlinear effect, it cannot just be generally represented as a simple curve in frequency, but depends on the duration, intensity, and spectrum of the pulse itself.

Two methods of compensating for GVD have been devised involving arrangements of gratings or prisms. Different colors take different paths as shown in Fig. 24. The basic idea is to set up a geometric arrangement to allow the blue part of the spectrum to ‘catch up’ to the red part. With grating pairs, this is easy to see intuitively, simply by comparing the optical path taken by different frequency components. With prisms the situation is much more tricky, as we’ll see below.

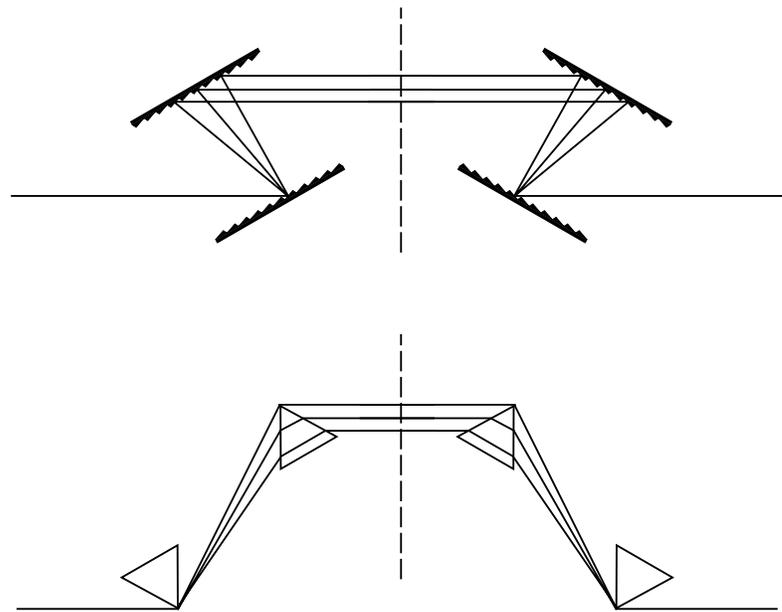


Figure 24. GVD compensation with gratings and prisms.

To calculate the optical path through a pair of anti-parallel gratings, we can use simple geometry. Because the gratings are anti-parallel, and we use the first order diffracted light, all spectral components will be propagating in the same direction after the first pair of gratings. The second pair serves to recombine the spectral components back into a single

beam, as well as doubling the temporal effect of the first pair. (Alternatively, a mirror can be used to send the spectral components back through the first pair of gratings.)

To compare path lengths for different frequencies, we can calculate the path length to a plane perpendicular to all the spectral components, and then (by symmetry) just double that to get the full path length. Figure 25 shows the path  $l$ , consisting of segments  $l_1$  plus  $l_2$ , where

$$l_1 = \frac{L}{\cos \theta_r}, \quad (38)$$

$$l_2 = l_1 \cos (\theta_i - \theta_r). \quad (39)$$

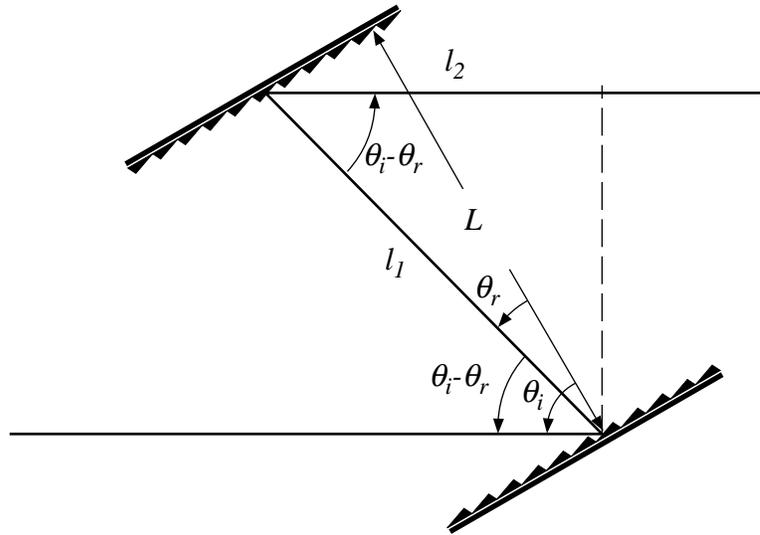


Figure 25. Schematic for calculating grating pair dispersion..

The reflected and incident angles  $\theta_i$  and  $\theta_r$  are related by

$$\sin \theta_r = -\sin \theta_i + \frac{m\lambda}{\Lambda}, \quad (40)$$

where  $m$  is the order of the diffraction (in this case  $m=1$ ),  $\lambda$  is the wavelength, and  $\Lambda$  is the groove spacing. For a proper choice of angles,  $l$  decreases almost linearly with  $\omega$ , about a central frequency  $\omega_0$ , giving a shorter path for the higher frequency components.

We now turn to prisms, to examine their effect on the dispersion, following the treatment of Fork, Martinez, and Gordon [13]. In Fig. 26 we assume that the prisms are used at the angle of minimum deviation for the central frequency. (At this angle the optical path inside the prism is perpendicular to the angle bisector). Point A is the location of the apex of the first prism, where the beam is incident, and L is the distance to the apex of the second prism. To calculate the actual optical beam path ABC, we can use an equivalent fictitious path through the vertex of the second prism, A'B'C'. The two optical path lengths are

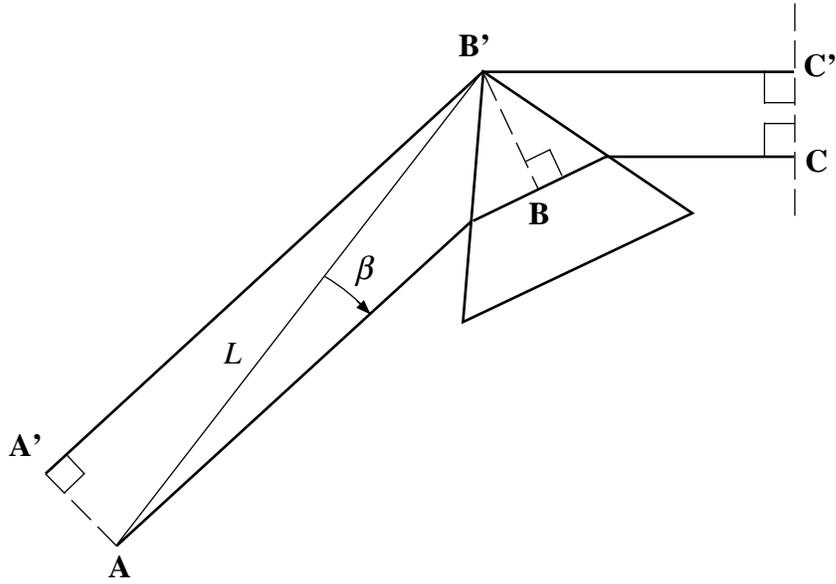


Figure 26. Schematic for calculating prism pair dispersion.

equivalent, because they start and end at a common phase front. We now note that two *different* frequencies will have the same optical path length from the plane through BB' to the plane through CC'. So the frequency dependent part of the optical path can just be taken as the distance from A to B, or equivalently from A' to B'. This 'prism-pair' path is simply

$$l_{p-p} = L \cos [\beta(\omega)] \quad (41)$$

where  $\beta$  is the angle between the apex-to-apex connector and the actual beam path AB. The frequency dependence of the index of refraction affects the optical path length through  $\beta$ . The second pair of prisms simply doubles the effect of the first pair, while also recombining the spectral components into a single beam.

Without doing any calculations, we can see that because the index of refraction increases with frequency, a greater  $\omega$  produces a smaller  $\beta$ , thus a larger  $\cos[\beta]$  and hence a larger  $l_{p-p}$ . So the optical path *increases* with frequency! In other words, the derivative of the path length with frequency is positive. But this is the opposite of what we were interested in, and the opposite of the effect of gratings! So how is it possible for the higher frequencies to 'catch up' to the lower frequencies if the optical path length increases with frequency?

The short answer is that it is the second derivative of the phase that determines the dispersion, and not the first derivative of the optical path. The phase acquired by traveling through the four prisms is

$$\phi(\omega) = \frac{\omega}{\epsilon} 2l_{p-p} = \frac{\omega}{\epsilon} 2L \cos [\beta(\omega)] . \quad (42)$$

The condition for the desired dispersion (for the higher frequencies to 'catch up' to the lower frequencies) is

$$\frac{d^2\phi}{d\omega^2} < 0. \quad (43)$$

Note that it is possible to meet this condition despite having a path length that increases with frequency.

This short answer may seem unsatisfying, because the optical path length may be a more intuitive concept than phase. Let's examine in more detail the relationship between the effective optical path  $l_{\text{eff}}$ , phase  $\phi$ , and group velocity  $V_g$ . For concreteness, we can consider the specific example of a pulse traveling through a block of transparent material of length  $l_0$ , with a frequency-dependent index of refraction  $n$ , so that  $l_{\text{eff}} = nl_0$ . The input field will acquire a frequency dependent phase,

$$E_{\text{out}}(\omega) = e^{i\phi(\omega)} E_{\text{in}}(\omega). \quad (44)$$

The phase is related to the optical path by

$$\phi = kl_0 = \frac{2\pi}{\lambda} l_0 = \frac{2\pi}{\lambda_0} l_{\text{eff}} = \frac{\omega}{c} l_{\text{eff}}, \quad (45)$$

where  $\lambda$  is the wavelength in the material and  $\lambda_0$  is the wavelength in vacuum. The first and second derivatives of the phase are:

$$\frac{d\phi}{d\omega} = \frac{1}{c} [l_{\text{eff}} + l'_{\text{eff}}\omega], \quad (46)$$

$$\frac{d^2\phi}{d\omega^2} = \frac{1}{c} [2l'_{\text{eff}} + l''_{\text{eff}}\omega]. \quad (47)$$

Now we turn to the group velocity:

$$\begin{aligned} V_g &= \frac{d\omega}{dk} = \left[ \frac{dk}{d\omega} \right]^{-1} = \left[ \frac{d\phi}{d\omega} \right]^{-1} l_0 \\ &= \frac{cl_0}{[l_{\text{eff}} + l'_{\text{eff}}\omega]} \end{aligned} \quad (48)$$

The group velocity dispersion (GVD) is

$$\begin{aligned} \frac{dV_g}{d\omega} &= \frac{-cl_0}{[l_{\text{eff}} + l'_{\text{eff}}\omega]^2} [2l'_{\text{eff}} + l''_{\text{eff}}\omega] \\ &= \frac{-l_0}{\left( \frac{d\phi}{d\omega} \right)^2} \cdot \frac{d^2\phi}{d\omega^2} \end{aligned} \quad (49)$$

The second derivative of the phase is often referred to as group delay dispersion (GDD). Thus we see that GVD and GDD are related by

$$\text{GVD} = \frac{-l_0}{\left(\frac{d\phi}{d\omega}\right)^2} \cdot \text{GDD}. \quad (50)$$

Table 1 shows the sign of the GVD, GDD, and first and second derivatives of the optical path for both prism and grating pairs. From eq. (49) it is apparent that the sign of the GVD depends on both the first and second derivatives of the optical path. In the case of gratings, the dominant term in the GVD is  $l'_{\text{eff}}$ , and thus the sign of the GVD is determined by the sign of  $l'_{\text{eff}}$ . In the case of prisms, on the other hand, at the appropriate angle,  $l''_{\text{eff}}\omega$  can have a larger magnitude than  $2l'_{\text{eff}}$ , thus giving the counter-intuitive result of higher frequencies arriving earlier, despite traveling a longer optical path.

	$\frac{dV_g}{d\omega}$	$\frac{d^2\phi}{d\omega^2}$	$l'_{\text{eff}}$	$l''_{\text{eff}}$
Gratings	+	-	-	small
Prisms	+	-	+	-

Table 1. GVD, GDD, and optical path derivatives.

In practice, grating pairs are used when large dispersion is required, such as in stretching a pulse by a large factor before sending it into an amplifier, or compressing it afterwards, as will be discussed in the next section. Inside a laser cavity however, prisms are usually used because they produce much smaller loss and the small dispersion compensation they provide is sufficient.

An interesting new method for GVD compensation that does not use gratings or prisms was demonstrated by Stingl et al. [14]. Recognizing that reflection off dielectric mirrors can produce negative GDD (or positive GVD), they have designed specially coated multilayer mirrors that provide a nearly constant negative GDD across a very broad wavelength range ( $\sim 200$  nm). A Ti:Sapphire laser with mirror-controlled dispersion has produced pulses of only 8-fs duration. An advantage of this type of dispersion compensation is that unlike the prisms pairs, it is largely independent of the cavity alignment.

## II. B. Amplification

Femtosecond pulses can be amplified outside of the laser cavity. Solid-state femtosecond lasers produce pulses that are typically a few nJ in energy, while dye-based femtosecond lasers produce pulses of even lower energy. The pulse repetition rate is determined by the cavity round-trip time, and is typically on the order of 100 MHz. Often, higher pulse energy at a lower repetition rate is desired. Before describing specific types of amplifiers, we review the concepts of gain narrowing and gain saturation, and how they apply to ultrashort pulses.

The gain of an amplifier is the ratio of the output intensity to the input intensity, and has a spectrum determined by the line shape of the amplifying medium. If the spontaneous emission line shape is given by  $\alpha(\omega)$ , then the gain over a distance  $z$  is

$$G(\omega) = \frac{I_{\text{out}}(\omega)}{I_{\text{in}}(\omega)} = e^{\alpha(\omega)z}. \quad (51)$$

If the line shape is Lorentzian (homogeneous broadening), then

$$\alpha(\omega) = \frac{N_2 - N_1}{1 + \frac{(\omega - \omega_0)^2}{\left(\frac{\Delta\omega}{2}\right)^2}}, \quad (52)$$

where  $N_2 - N_1$  is the population inversion,  $\omega_0$  is the center frequency, and  $\Delta\omega$  is the line width.

Amplification often reduces the spectral width of the input signal. This is called gain narrowing. To illustrate this effect, in Fig. 27 we plot the normalized gain spectrum for amplification by a factor of 10, 100, and 10,000. At high amplification, the gain narrowing can be significant. Because ultrashort pulses necessarily have a broad spectrum, the gain medium in the amplifier should have a gain bandwidth that is as broad as possible, otherwise the amplified pulses will have a longer duration than the input pulses.

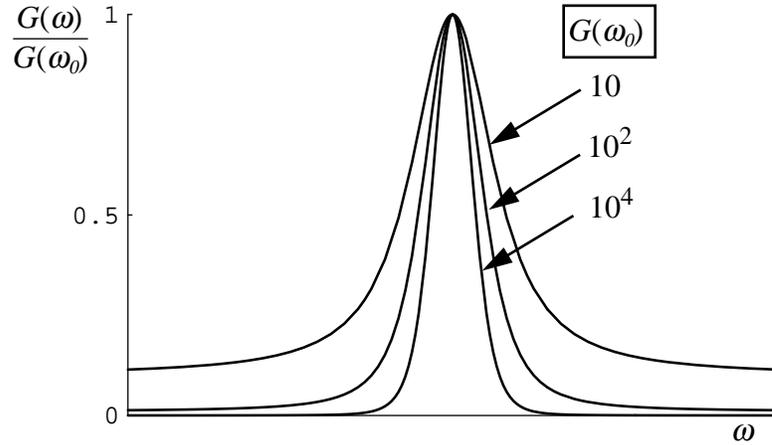


Figure 27. Gain narrowing in an amplifier.

Perhaps even more important than gain narrowing is the effect of gain saturation in an amplifier. For a homogeneously broadened medium, the following formula for saturation can be found in most laser textbooks:

$$\alpha = \alpha_0 \frac{1}{1 + I/I_{\text{sat}}} \quad (53)$$

where  $\alpha_0$  is the gain coefficient for a small signal, and  $I_{\text{sat}}$  is the saturation intensity. The gain coefficient decreases at higher intensity, producing smaller and smaller gain. Can this formula be used for ultrashort pulses? For example, is it possible to avoid saturation by stretching a pulse in time (thus reducing the intensity), then amplifying it, and then compressing it?

To answer this question we should examine what causes the gain saturation. The saturation occurs because stimulated emission in the amplifier reduces the population inversion and thus reduces the gain. With long pulses or continuous radiation the maximum intensity is limited by the pump rate, i.e. the rate at which the population inversion is being replenished. The underlying assumption here is that the system is in equilibrium: the rate of emission cannot exceed the pumping rate. This assumption holds as long as the pulse duration is much longer than the gain recovery time. (Note that in steady state, when the pump rate balances the decay rate, the gain recovery time equals the lifetime of the inverted population.)

With femtosecond and even picosecond pulses, the assumption of equilibrium is certainly false for almost all gain media. For example, in Ti:Sapphire the gain recovery time is three microseconds, so the pump rate is so slow that on the timescale of an ultrashort pulse it has no effect on the population inversion. In this opposite limit, where the pulse is much shorter than the population recovery time, the emission can certainly exceed the pump rate, and the only limit is that the total number of photons in the pulse cannot exceed the number of electrons in the inverted state. Thus, in amplifying ultrashort pulses, the proper measure of saturation is a saturation fluence, rather than a saturation intensity:

$$\alpha = \alpha_0 \frac{1}{1 + F/F_{\text{sat}}} \quad (54)$$

Thus we see that stretching an ultrashort pulse, even say by a factor of 1000, will not avoid saturation, because it is the fluence rather than the intensity that determines the saturation. (One can of course produce greater energy pulses by magnifying the beam and using an amplifier with a larger cross-section, thus reducing the fluence.) The saturation fluence for dyes in solution is several  $\text{mJ}/\text{cm}^2$ , whereas for doped solids, such as Ti:Sapphire, it is much greater, several  $\text{J}/\text{cm}^2$  [15].

Yet there is a different reason why stretching the input pulse to an amplifier is useful: nonlinear effects, such as self-focusing, are not desirable in the amplifier, and can be avoided if a pulse is stretched in time. A pulse can be stretched reversibly by imposing a ‘chirp’ on it. The term ‘chirp’ refers to a linear spread in time of the frequencies in the pulse, and results from constant (non-zero) GDD. A negative chirp, where the higher frequency components are ahead of the lower frequency components, can be achieved with a grating pair, as discussed in the previous section. To recompress the pulse after the amplifier we would need to provide positive chirp. How can we impose a positive chirp on a pulse? One approach is to pass the pulse through a dispersive material. But if a large positive chirp is desired, the required material length may be too long.

An inventive approach to creating positive chirp uses a pair of gratings in a two-lens system. Figure 28 shows two lenses in a so called ‘4-F’ or ‘two-lens correlator’ arrangement. This type of system thus finds great use in optical spatial information processing, for such operations as Fourier filtering, convolution, and correlation. Its main feature is that it performs a Fourier transform from the input plane to the second plane, and then again from the second plane to the output plane. For our purposes, we note that it is an inverting imaging system (from input to output planes), that all rays take equal length optical paths, and that the slope of the rays is ‘inverted’ from input to output.

We now modify the two-lens system by adding two gratings and a mirror, as shown

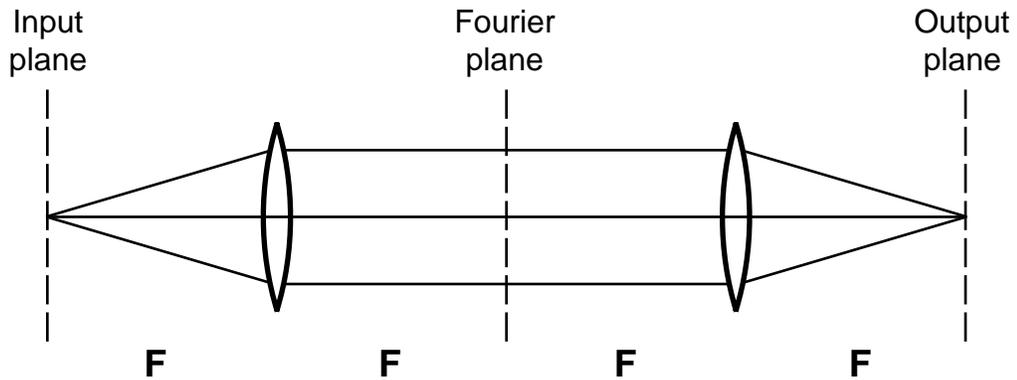


Figure 28. Two-lens correlator.

in Fig. 29. In this arrangement higher frequencies are made to take a longer path. The positions at which two spectral components ('red' and 'blue') have traveled equal paths are marked in the drawing as  $P_r$  and  $P_b$ . From those points to the mirror the 'blue' beam has a longer path than the 'red'. Note that if the second grating is placed at the output plane, then no dispersion is introduced. Thus by adjusting the position of the second grating one can control the amount of chirp. Positive chirp is obtained by placing the grating in front of the output plane, and negative chirp is obtained by placing the grating behind the output plane. This arrangement was originally designed to compensate for the dispersion in an optical fiber in the  $1.3 - 1.6 \mu\text{m}$  wavelength range, where the material dispersion is negative [16].

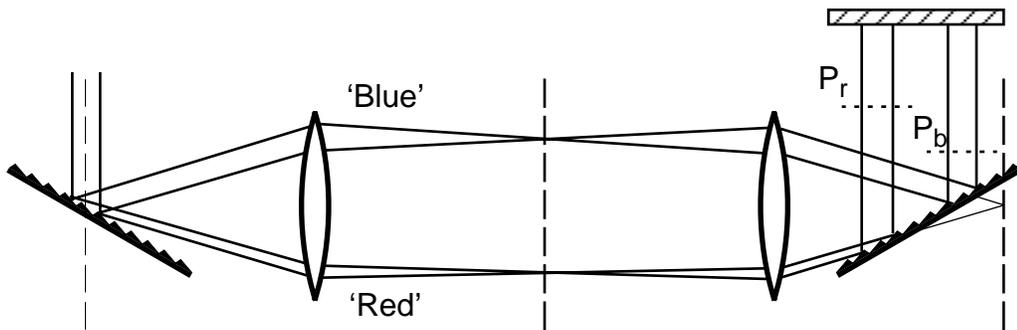


Figure 29. Pulse stretcher.

In chirped pulse amplification (CPA) [17], such a method for introducing positive chirp is used to stretch an ultrashort pulse by more than 1000 times, the pulse is then amplified, and a grating pair is then used to recompress the pulse. Curved mirrors can be used instead of lenses in the stretcher to avoid spectral aberrations, and cylindrical rather than spherical optics can be used. A carefully designed stretcher-amplifier-compressor system that is dispersion compensated to fourth order (i.e.  $d^2\phi/d\omega^2 = d^3\phi/d\omega^3 = d^4\phi/d\omega^4 = 0$ ) can stretch an input pulse of 25 fs to 300 ps, amplify it, and recompress it back to about 27 fs [18].

Different types of amplifiers are used depending on the desired energy and pulse repetition rate. Figure 30 shows a representative selection of current state-of-the-art laser/amplifier systems producing sub-100-fs pulses. The 100 MHz and 1 GHz points are the direct (unamplified) output from the laser cavity of two Ti:Sapphire lasers, one optimized for maximum energy [19], the other optimized for compactness and high repetition rate [20]. The point at 400 KHz represents the output of CPA system with a regenerative amplifier pumped continuously by an Argon ion laser [21]. Similar amplifiers, pumped by Q-switched lasers, produce higher energy pulses in the low kHz range [22]. The regenerative amplifier is essentially a second laser cavity into which laser pulses are injected and later ejected after the gain is saturated after some number of round-trips. The pulse injection and ejection are accomplished by polarization rotation with a Pockels cell, which provides fast electro-optic switching. Regenerative amplifiers have the advantages of excellent beam mode quality, high efficiency, and compactness. For even higher energies, large linear amplifiers (single- or multi-pass) are used following the regenerative amplification, as for example in [23].

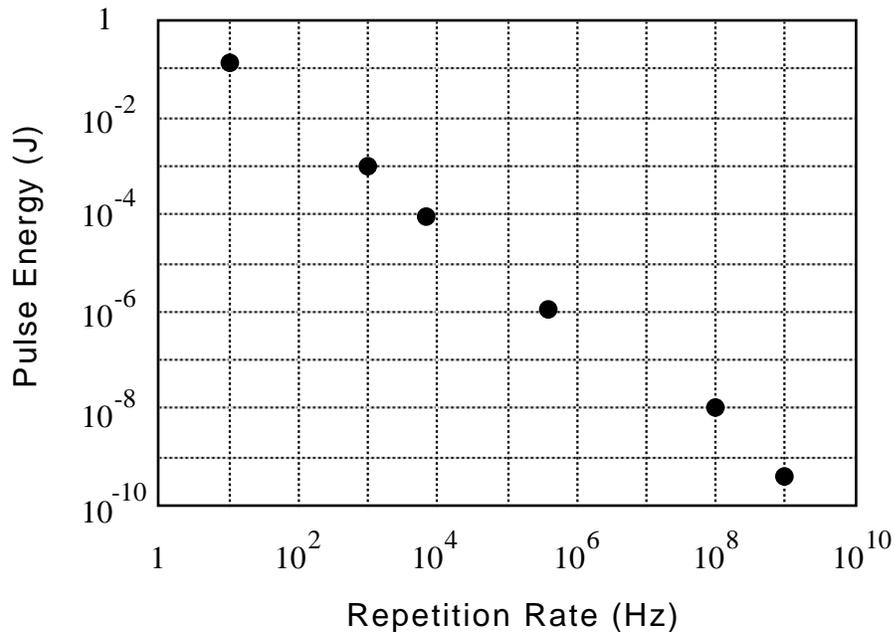


Figure 30. Lasers and amplifiers: state of the art for sub-100-fs pulses

### III. MEASURING AND SHAPING FEMTOSECOND PULSES

We now return to a question posed in the introduction: how do we measure ultrashort optical pulses when there are no electronic detectors with fast enough response times? Since the development of ultrashort-pulse sources, a great deal has been learned about how such pulses can be characterized, so that recently it has even become possible to measure the exact electric field of an ultrashort pulse. In addition to detailed characterization, it is now possible to shape the pulses in frequency and time, producing, for example, controlled multiple-pulse sequences out of a single pulse, or imposing a specific phase variation on an originally unchirped pulse.

### III. A. Pulse Characterization

**1. Pulse Spectrum.** The simplest characterization of a pulse is its spectrum. An example of a spectrum typical of the shortest pulses (about 10 fs) from Ti:Sapphire lasers is shown in Fig. 31. The width of the spectrum places a limit on the shortest possible pulse duration, as discussed in Section I.B. For a Gaussian pulse the limit is  $\sigma_t=1/\sigma_\omega$ . At this point it may be useful to mention that pulse durations and spectral widths are often given as ‘FWHM’ — full width at half maximum of the intensity. For a Gaussian pulse

$$E(t) \sim e^{-\frac{t^2}{2\sigma_t^2}} \quad \text{and} \quad I(t) \sim e^{-\frac{t^2}{\sigma_t^2}}$$

$$\frac{I\left(\frac{\Delta t_{\text{FWHM}}}{2}\right)}{I(0)} \equiv \frac{1}{2} \tag{55}$$

$$\Delta t_{\text{FWHM}} = 2\sqrt{\ln 2}\sigma_t$$

Similarly, in frequency

$$\Delta\omega_{\text{FWHM}} = 2\sqrt{\ln 2}\sigma_\omega. \tag{56}$$

While it does set a limit on the shortest possible pulse duration, the pulse spectrum cannot be used to measure pulse duration because the spectrum contains no phase information.

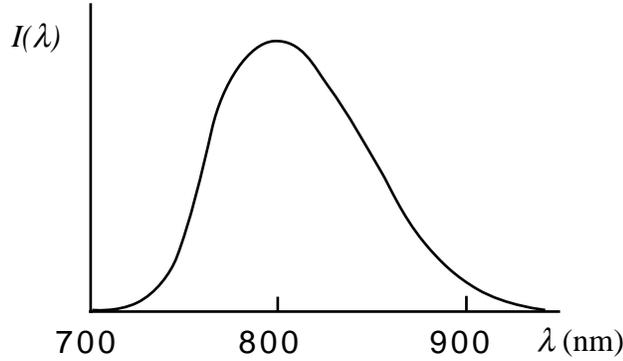


Figure 31. Typical spectrum of an ultrashort pulse from a Ti:Sapphire laser.

**2. Temporal Characterization.** To measure the pulse in time we can use the pulse itself. Figure 32 shows one simple technique that measures the temporal autocorrelation of a pulse. The pulse is split into two parts, one of which passes through an adjustable delay line. The two parts are then focused together onto a second-harmonic crystal, producing three beams at the second-harmonic frequency.

This arrangement is the so-called ‘non-collinear’ or ‘background-free’ autocorrelation. By measuring the energy in beam #2 as a function of the time-delay ( $\Delta t$ ) we obtain an autocorrelation of the intensity profile of the pulse. To show how this occurs, we consider the field incident on the SH crystal:

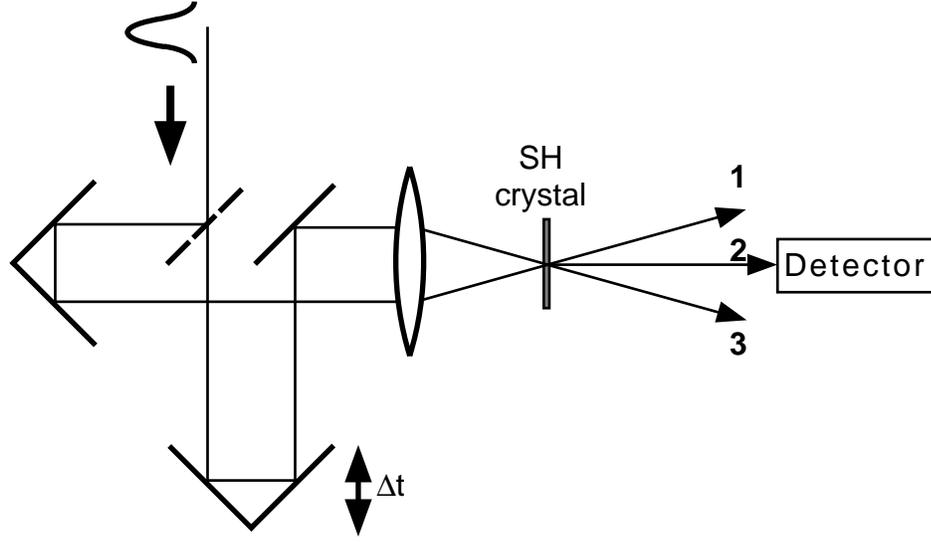


Figure 32. Non-collinear autocorrelation.

$$E_{\text{tot}}(t, \Delta t) = \frac{1}{\sqrt{2}}E(t) + \frac{1}{\sqrt{2}}E(t + \Delta t). \quad (57)$$

The second harmonic field is proportional to the second-order susceptibility times the square of the incident field. (This is true only in the limit where only a small fraction of the incident energy is converted into the second-harmonic).

$$E_{2\omega} \sim \chi^{(2)} E_{\text{tot}}^2. \quad (58)$$

The second-harmonic intensity is proportional to the square of the second-harmonic field:

$$I_{2\omega}(t, \Delta t) \sim |\chi^{(2)}|^2 |E_{\text{tot}}^2|^2 = |\chi^{(2)}|^2 |E^2(t) + 2E(t)E(t + \Delta t) + E^2(t + \Delta t)|^2. \quad (59)$$

So far we have not made the spatial dependence explicit. By measuring the energy in beam #2, we effectively select only the second term in eq. (59). The detector response is much longer than the duration of an individual pulse, so the detector integrates the intensity in time:

$$F_{2\omega}(\Delta t) = \int I_{2\omega}(t, \Delta t) dt \sim \int |\chi^{(2)}|^2 4|E(t)|^2 |E(t + \Delta t)|^2 dt \quad (60)$$

$$F_{2\omega}(\Delta t) \sim \int I(t) I(t + \Delta t) dt. \quad (61)$$

Thus we see that the measured signal is proportional to an autocorrelation of the intensity envelope of the pulse. Figure 33 shows the electric field of a 7-fs pulse, and the corresponding intensity autocorrelation.

An alternative pulse-duration measurement based on second-harmonic generation employs a collinear arrangement in which all 3 terms of eq. (59) are spatially coincident. As

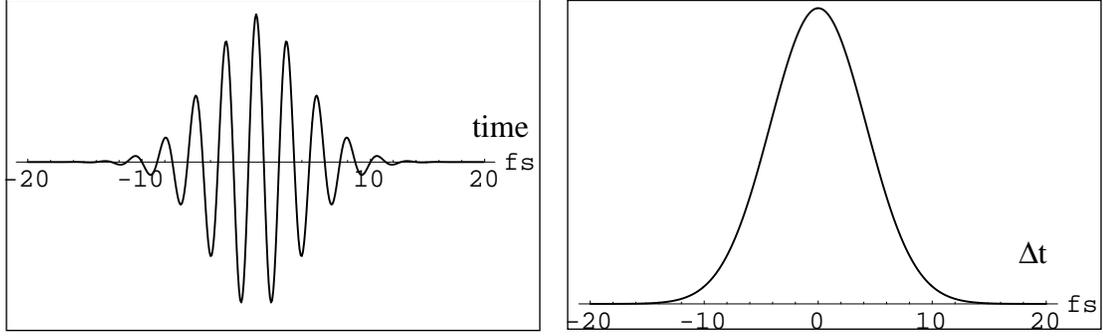


Figure 33. 7-fs pulse and the corresponding SH intensity autocorrelation

shown in Fig. 34, a Michelson interferometer is used to split the pulse into two, delay one part relative to the other, and then colinearly recombine the two. After the SH crystal, a filter is used to block the beam at the fundamental frequency and transmit the SH beam. In this arrangement the 3 SH beams are colinear, and thus add coherently. The SH intensity is given by

$$\begin{aligned}
 I_{2\omega}(t, \Delta t) &\sim |\chi^{(2)}|^2 |E^2(t) + 2E(t)E(t + \Delta t) + E^2(t + \Delta t)|^2 \\
 I_{2\omega}(t, \Delta t) &\sim |\chi^{(2)}|^2 \{ |E(t)|^4 + |E(t + \Delta t)|^4 + 4|E(t)|^2|E(t + \Delta t)|^2 \\
 &\quad + E^2(t) [E^2(t + \Delta t)]^* + \text{c. c.} \\
 &\quad + E^2(t) 2E^*(t) E^*(t + \Delta t) + \text{c. c.} \\
 &\quad + 2E(t) E(t + \Delta t) [E^2(t + \Delta t)]^* + \text{c. c.} \}
 \end{aligned} \tag{62}$$

where c.c. stands for complex conjugate. The detected integrated intensity is again a function of the time-delay,  $\Delta t$ ,

$$F_{2\omega}(\Delta t) = \int I_{2\omega}(t, \Delta t) dt. \tag{63}$$

Figure 35 again shows the electric field of a 7-fs pulse, and the corresponding ‘interferometric autocorrelation’, also referred to as a ‘fringe-resolved autocorrelation’. The time-delay curve and even its envelope are not simple autocorrelations, but the width of the envelope is related to the width of the pulse. At  $\Delta t = 0$  all the terms in equation (62) give non-zero contributions to the detected signal,

$$F_{2\omega}(\Delta t = 0) \sim 16 \int |E(t)|^4 dt, \tag{64}$$

while at  $\Delta t = \pm\infty$  only the first two terms in equation (62) contribute, giving

$$F_{2\omega}(\Delta t = \pm\infty) \sim 2 \int |E(t)|^4 dt. \tag{65}$$

This gives a ratio of 8 between the peak at  $\Delta t = 0$  and the background signal at long time-

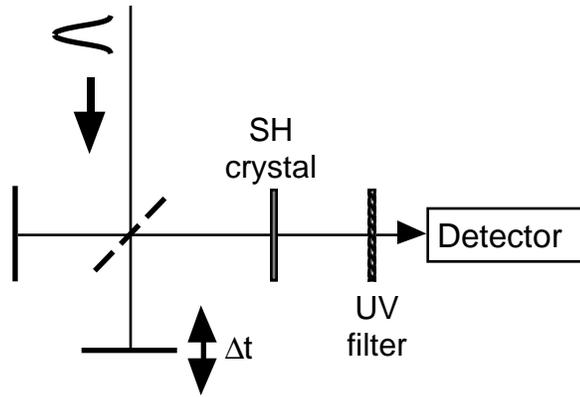


Figure 34. Interferometric autocorrelation.

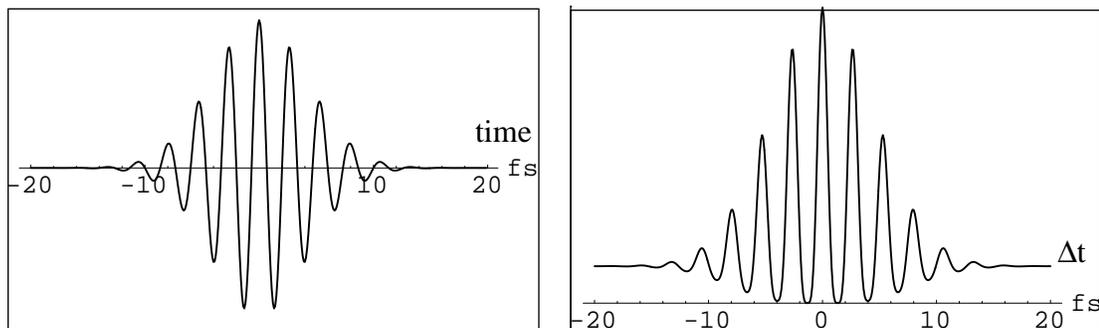


Figure 35. 7-fs pulse and the corresponding SH interferometric autocorrelation

delay when the two pulses are well-separated in time; for this reason the interferometric autocorrelation is sometimes referred to as the ‘8-to-1 curve’.

One of the two autocorrelations described above is sometimes more appropriate than the other, depending on what properties of the pulse are of interest. The advantages of the interferometric autocorrelation include: the self-calibrating nature of the curve due to the presence of the fringes, the lack of temporal distortion due to the zero relative angle between the two beams (only significant for the very shortest pulses), and the 8-to-1 ratio, which serves as a check that the second-harmonic signal is not affected by other factors (such as saturation or self-focusing in the crystal). The main advantage of the intensity autocorrelation is that it is ‘background-free’, due to the non-colinear geometry. This is particularly useful in checking for small pre- or post-pulses, which can be more easily detected without a background which unavoidably has some noise.

Both autocorrelation techniques described above use a nonlinear crystal to produce a second-harmonic signal. We now ask the question: Is the nonlinear crystal necessary?

Figure 36 shows the same Michelson interferometer setup, except this time the SH crystal and the filter have been removed, so the detector now measures the integrated intensity at the fundamental frequency, again as a function of time-delay ( $\Delta t$ ). The intensity is proportional to the square of the sum of the fields,

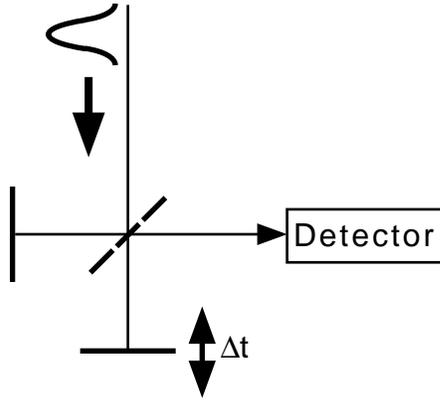


Figure 36. Interferometric time-delay setup without a nonlinear element.

$$I_{\omega} \sim |E(t) + E(t + \Delta t)|^2, \quad (66)$$

and the detected signal is thus:

$$F_{\omega}(\Delta t) = \int I_{\omega}(t, \Delta t) dt \quad (67)$$

$$F_{\omega}(\Delta t) \sim \int \{ |E(t)|^2 + |E(t + \Delta t)|^2 + E(t) E^*(t + \Delta t) + E^*(t) E(t + \Delta t) \} dt. \quad (68)$$

The first two terms are just a constant ‘background’ independent of  $(\Delta t)$ , while the last two are interference terms. Figure 37 shows a 7-fs pulse and the corresponding ‘interference trace’ as a function of time-delay. Figure 38 shows a transform-limited 14-fs pulse (of narrower spectral width), and its interference trace. The broader envelope of the time-delay trace corresponds to the longer pulse duration. From this example it is tempting to conclude that the linear interference trace can be used to measure pulse duration, and that the nonlinear crystal is unnecessary.

But now we consider a different case. The original 7-fs pulse is passed through a dispersive medium, thereby stretching it in time while leaving its spectral width unaffected. The two pulses appear in Figs. 39 and 40, which also show the background-free SH, fringe-resolved SH, and the linear interference time-delay traces for both the 7-fs transform-limited pulse, and its ‘stretched’ or ‘chirped’ version. While both of the second-harmonic autocorrelations are broader for the longer pulse, the linear interference traces in the two cases are identical! This may seem counter-intuitive, because after all, why shouldn’t a longer pulse produce a ‘longer’ interference trace?

A qualitative explanation for this is that the ‘red’ part of the pulse ‘does not interfere’ with the ‘blue’ part. If the front end of the pulse contains the lower frequency components, when it overlaps in time with the back end of its copy which contains the higher frequency components, there will be no significant constructive or destructive interference, i.e. the time-integrated intensity will not be significantly changed. While this explanation may give some insight, it gives only a rough justification since it mixes frequency and time

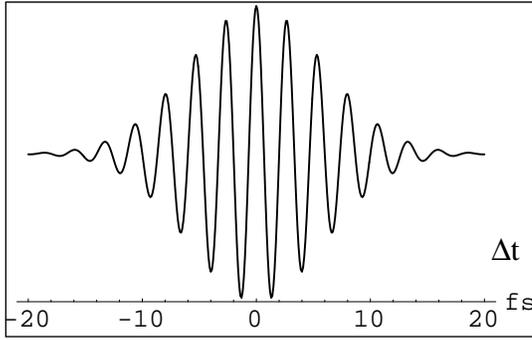
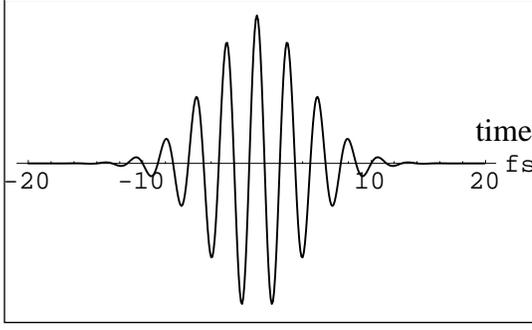


Figure 37. 7-fs transform-limited pulse and its linear interference trace.

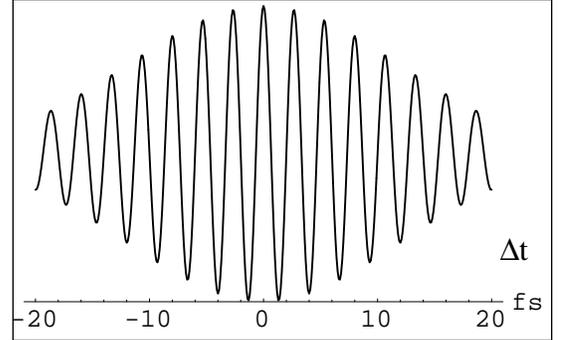
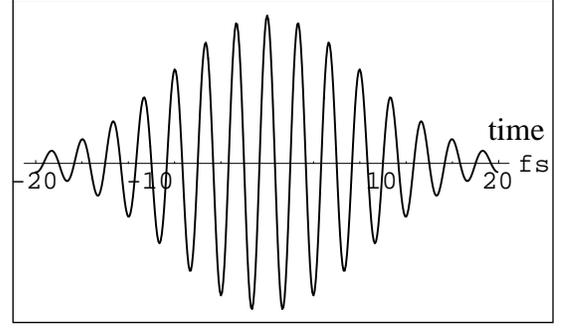


Figure 38. 14-fs transform-limited pulse and its linear interference trace

descriptions of the pulse.

To prove that chirp does not affect the linear interference trace, we treat the pulse in its frequency representation. As the original pulse,  $\tilde{E}(\omega)$ , propagates through a dispersive medium, it acquires a frequency-dependent phase factor, producing  $\tilde{E}(\omega) e^{i\phi(\omega)}$ . We now consider the interference terms in eq. (68),

$$\int \{ E^*(t) E(t + \Delta t) + \text{c. c.} \} dt. \quad (69)$$

Using the correlation theorem,

$$g_1(t) \otimes g_2(t) \equiv \int g_1(t + \Delta t) g_2^*(t) dt = \mathbf{F}^{-1} \{ \tilde{g}_1(\omega) \tilde{g}_2^*(\omega) \}, \quad (70)$$

(where the  $\tilde{g}$  designates the Fourier transform of  $g$ ), we rewrite the interference terms for the chirped pulse as

$$\begin{aligned} \int E_{\text{chirped}}(t + \Delta t) E_{\text{chirped}}^*(t) dt &= \mathbf{F}^{-1} \{ \tilde{E}(\omega) e^{i\phi(\omega)} \tilde{E}^*(\omega) e^{-i\phi(\omega)} \} \\ &= \mathbf{F}^{-1} \{ |\tilde{E}(\omega)|^2 \} \\ &= \int E(t + \Delta t) E^*(t) dt \end{aligned} \quad (71)$$

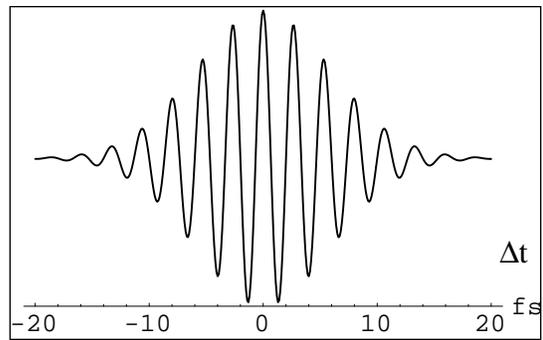
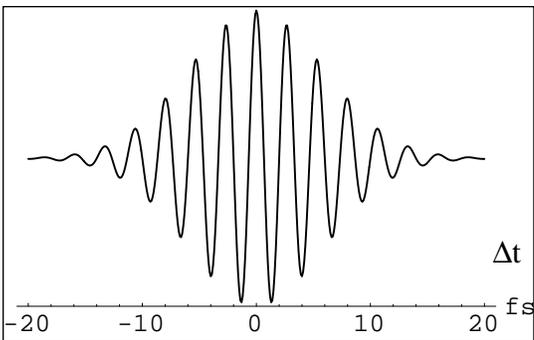
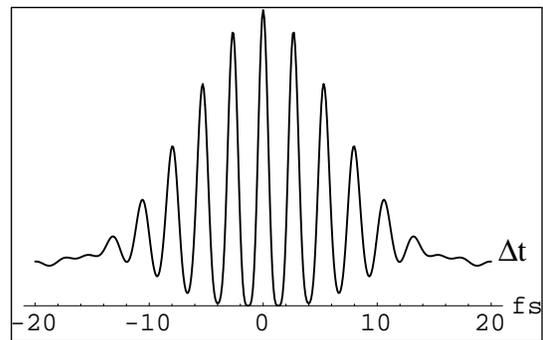
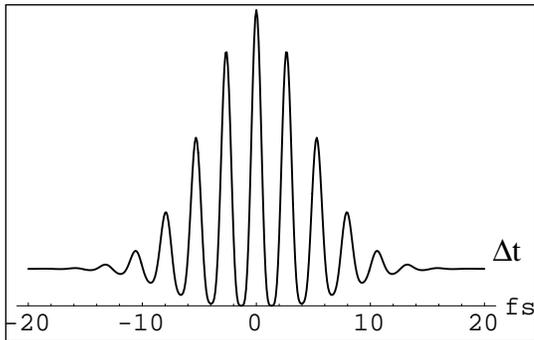
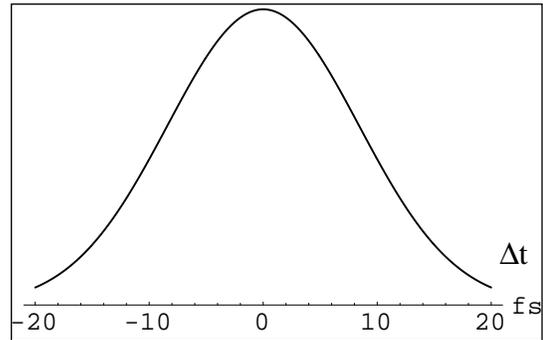
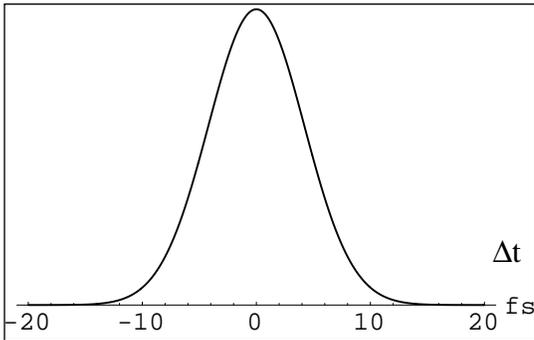
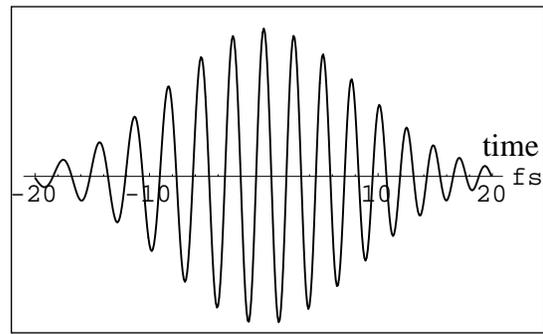
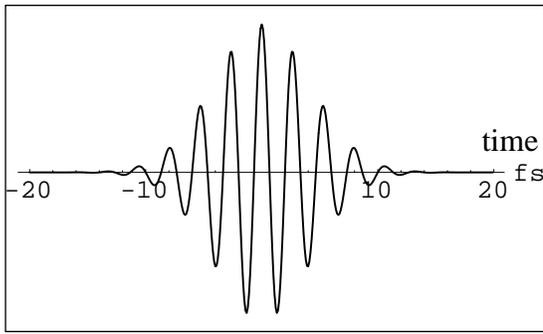


Figure 39. 7-fs pulse, with its SH intensity autocorrelation (AC), SH interferometric AC, and linear interference trace.

Figure 40. Same pulse after propagating through a dispersive medium, with its SH intensity AC, SH interferometric AC, and linear interference trace.

Thus we have proved that the linear interference time-delay trace is completely determined by the intensity spectrum — and unaffected by the phase.

**3. Joint Time-Frequency Characterization.** We have seen that a nonlinear element can be used to obtain an autocorrelation of an ultrashort pulse, and thus provide a measure of its duration. When more detailed information about the electric field of a pulse as a function of time is desired, measurements that combine temporal and spectral resolution can be employed. The general approach is to use a pulse split into two in a time-delay arrangement, just as in the autocorrelations described above, but this time to measure the spectrum of the nonlinear signal rather than just its energy. This is illustrated schematically in Fig. 41.

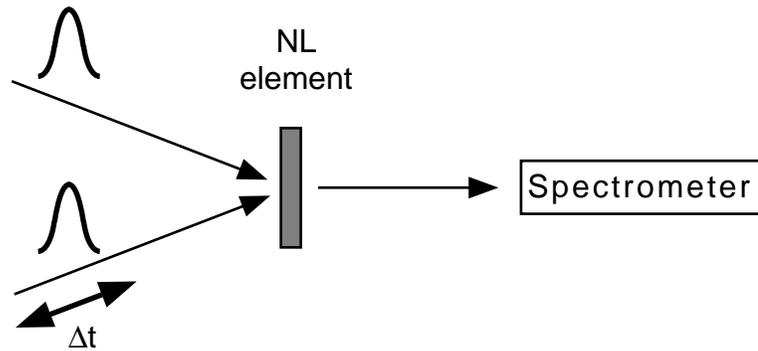


Figure 41. General approach to joint spectral-temporal pulse characterization.

One specific example of joint time-frequency analysis is called ‘frequency-resolved optical gating’ (FROG), [24] shown schematically in Fig. 42. The ‘optical gating’ refers to using one part of the pulse as a gate for the other. This is accomplished with a third-order nonlinear element, with the gate pulse (#2) used to rotate the polarization of the ‘direct’ pulse (#1). As can be seen in Fig. 42, the sequence of crossed polarizers will block pulse #1, unless its polarization is slightly rotated by the gate pulse. If the response of third-order element is essentially instantaneous, i.e. non-resonant, (a simple piece of glass is often used), then the field in the transmitted beam will be proportional to the square of the magnitude of the electric field of the gate pulse. Note that although the transmitted signal

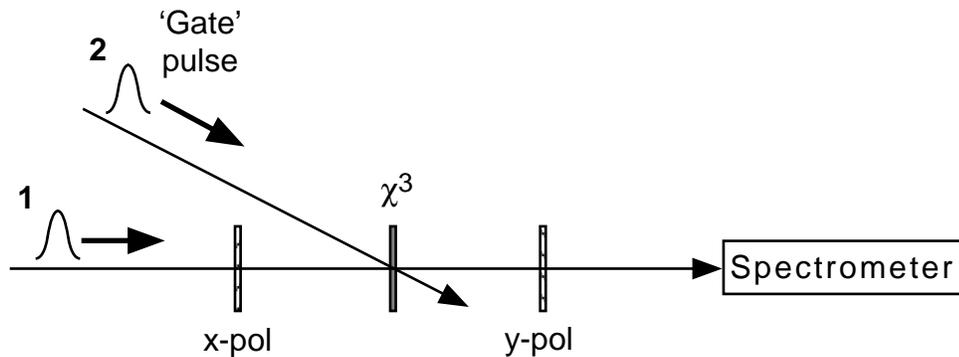


Figure 42. Frequency-resolved optical gating (FROG).

is of the same frequency as the original pulses, it does involve a nonlinear process. Assuming the gate pulse is just a time-delayed version of the ‘direct’ pulse, the transmitted signal is

$$E_{\text{signal}}(t, \Delta t) \sim \chi^{(3)} E(t) |E(t + \Delta t)|^2. \quad (72)$$

The spectrometer performs a Fourier transform on the signal and the array of detectors in the spectrometer measures the time-integrated spectral intensity:

$$\begin{aligned} F_{\text{FROG}}(\omega, \Delta t) &\sim |\chi^{(3)}|^2 |E_{\text{signal}}(\omega, \Delta t)|^2 \\ &\sim |\chi^{(3)}|^2 \left| \int E_{\text{signal}}(t, \Delta t) e^{-i\omega t} dt \right|^2 \end{aligned} \quad (73)$$

One of the advantages of joint time-frequency characterization is that the signal (measured as a function of both frequency and time-delay) uniquely determines the electric field,  $E(t)$ . Although it is not possible to invert eqs. (72) and (73) to directly calculate the field from the FROG signal, the field can be obtained iteratively by successive approximation. This procedure is described in detail in [25]. Another advantage is that a plot of the FROG signal visually conveys information about the chirp of the pulse, and provides a close connection to the Wigner representation.

Joint time-frequency characterization is not restricted to using the specific third-order nonlinearity described above. For example, sum-frequency generation can be used in a ‘spectrally-resolved autocorrelation’, as shown in Fig. 43. The sum-frequency field is

$$E_{\text{signal}}^{2\omega}(t, \Delta t) \sim \chi^{(2)} E(t) E(t + \Delta t), \quad (74)$$

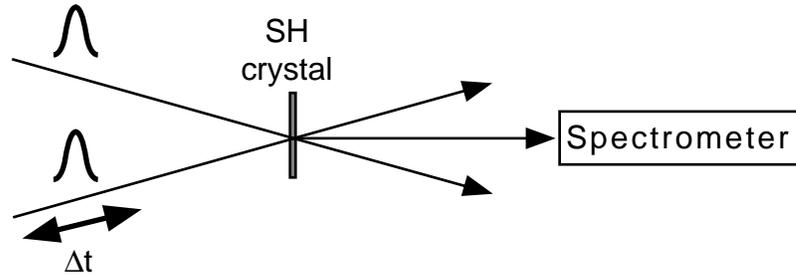


Figure 43. Spectrally-resolved autocorrelation.

and the measured spectrum is

$$\begin{aligned} F^{2\omega}(\omega, \Delta t) &\sim |\chi^{(2)}|^2 |E_{\text{signal}}^{2\omega}(\omega, \Delta t)|^2 \\ &\sim |\chi^{(2)}|^2 \left| \int E_{\text{signal}}^{2\omega}(t, \Delta t) e^{-i\omega t} dt \right|^2 \end{aligned} \quad (75)$$

Once again, the spectrally-resolved signal uniquely determines  $E(t)$ . An example of the spectrally resolved autocorrelation trace is shown in Fig. 43 together with the retrieved electric field of the 13-fs pulse (from [26]).

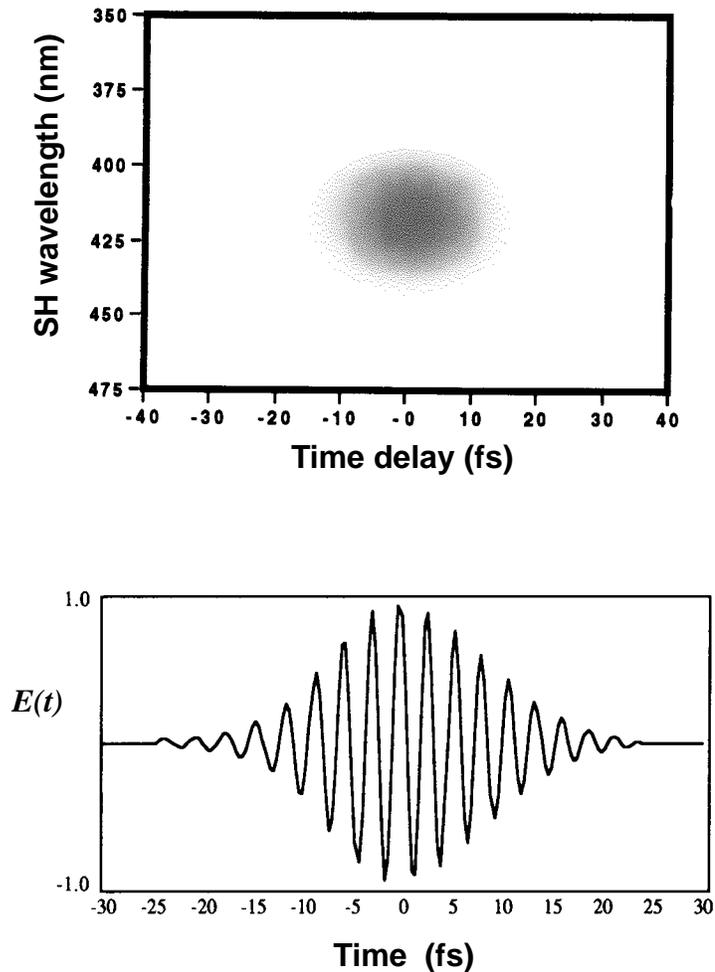


Figure 44. Example of a spectrally-resolved autocorrelation measurement and the electric field obtained from it (from [26])

### III. B. Pulse Shaping

Now that we've answered the question of how ultrashort pulses can be measured, we are ready to examine how these pulses can be actively shaped or tailored. As a motivation for wanting to shape ultrashort pulses, consider two very different applications: coherent control and temporal information processing. 'Coherent control' refers to the optical manipulation of vibrational and electronic states of molecules and solids. Through precise control of the phase and amplitude of the electric field of an optical pulse or sequence of pulses, it becomes possible to achieve such goals as selectively breaking molecular bonds, controlling chemical reactions, and driving phase transitions in solids. For an introduction to coherent control see [27]. Temporal information processing takes advantage of the similarity between spatial and temporal domains to extend the techniques of spatial signal processing (such as correlation, convolution, matched filtering) to processing temporal waveforms [28]. Furthermore, conversion of information between spatial and temporal domains makes possible such operations as encoding spatial information directly into a

pulse sequence, retrieving time-encoded information into a spatial signal distribution, and identifying specific signatures in an optical data stream by correlating it with a spatial pattern [29]. Some very clever experimental implementations of these operations have recently been demonstrated [30], opening up an exciting new field of ultrafast spatio-temporal information processing.

How can we shape ultrashort optical pulses? Certainly no modulators can be driven fast enough to directly modulate the temporal envelope of the pulse. However, there is a very clever approach that does not require fast modulation at all — it involves shaping a pulse in its Fourier domain (i.e. in frequency). The trick is to somehow map the frequency domain onto a spatial coordinate, thus obtaining direct access to the frequency domain and making it possible to apply a spatial phase and amplitude mask to the pulse spectrum. Then, if the spectral components can be recombined into a single beam, the pulse will have been shaped by the Fourier transform of the applied mask. All this is accomplished by the optical arrangement shown in Fig. 45, first demonstrated by Weiner, Heritage and Kirschner [31]. (A somewhat different arrangement was used earlier for shaping picosecond pulses [32])

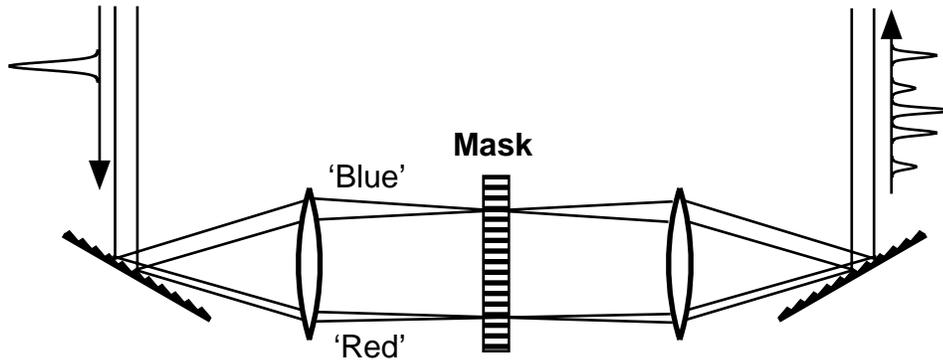


Figure 45. Pulse shaper.

Note that this is the same basic arrangement as in the pulse-stretcher discussed in Section II.B, in a zero-chirp configuration with the gratings positioned exactly in the input and output planes of the two-lens correlator. The role of the first grating is to disperse spectral frequencies into different directions (or ‘spatial frequencies’), which are then mapped onto the transverse spatial coordinate by the Fourier transform action of the first lens. A phase and amplitude mask is applied in the Fourier plane. Finally, the shaped spectrum is recombined back into a single beam: the second lens transforms positions in the Fourier plane into spatial frequencies in the output plane, and the second grating brings the spectral components back into a single beam.

With such a pulse-shaping system almost any waveform can be produced that is within the bandwidth of the original pulse. If  $E_{\text{out}}(t)$  is the desired waveform, then the required mask is

$$M(\omega) = \frac{E_{\text{out}}(\omega)}{E_{\text{in}}(\omega)} = \frac{F\{E_{\text{out}}(t)\}}{F\{E_{\text{in}}(t)\}}. \quad (76)$$

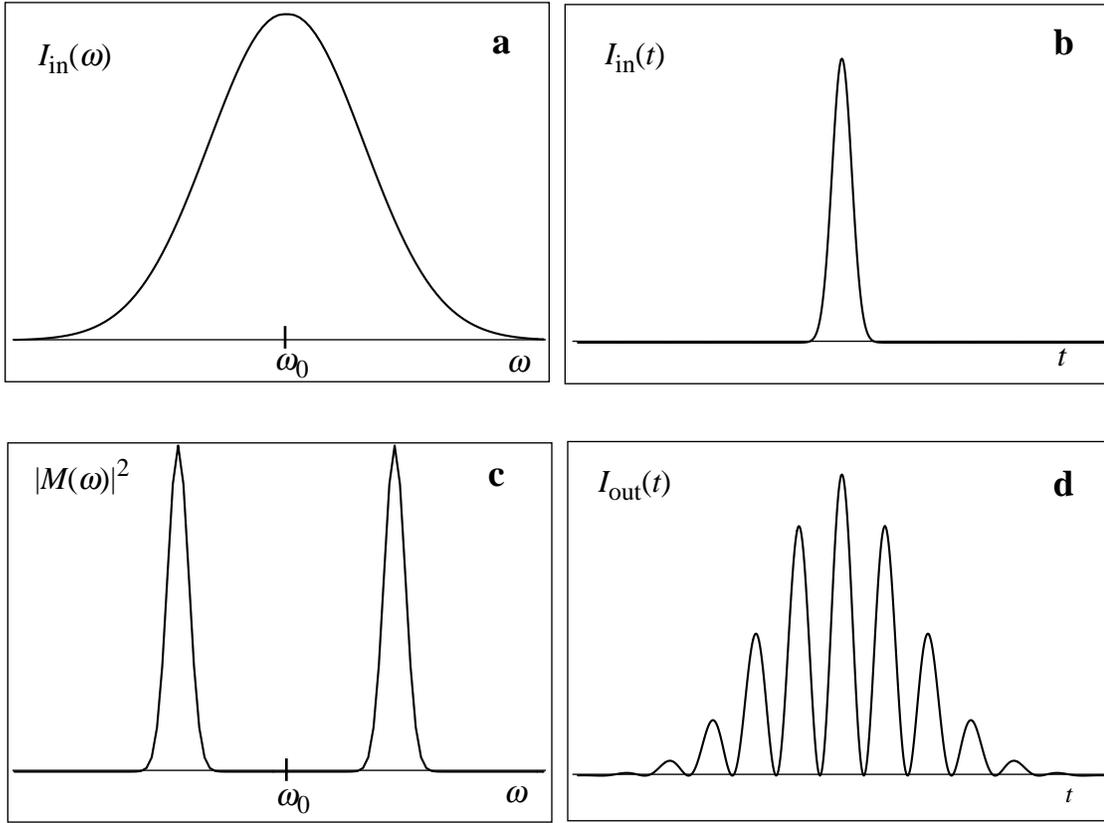


Figure 46. Example of pulse shaping. **a.** Spectrum of input pulse; **b.** Temporal intensity envelope of input pulse; **c.** Amplitude mask; **d.** Temporal intensity envelope of output pulse.

The constraint on the mask is  $|M(\omega)| \leq 1$ , since it is not an amplifier. Figure 46 shows an example of a Gaussian input pulse, shaped by a simple amplitude mask consisting of two Gaussian transmission peaks. The frequency filtering significantly lengthens the duration of pulse, and imposes a temporal modulation corresponding to the spectral spacing between the two transmission peaks. The result is a sequence of pulses inside a Gaussian envelope, with a total duration inversely proportional to the width of the transmission peaks.

The first masks were permanent phase and/or amplitude masks fabricated on glass substrates [31]. Subsequently, programmable masks have been designed which provide a great deal of flexibility in dynamically changing the waveform from pulse to pulse [33]. These masks are based on liquid crystal spatial light modulators (LC SLM's), in which the applied electric field controls the index of refraction of each pixel and thus provides control over phase delay across the spectrum. By combining a LC SLM with polarizers an amplitude mask can be assembled. In fact, both amplitude and phase control can be achieved in one mask by combining two SLM's and two crossed polarizers [34]. In this arrangement, two SLM's, one oriented in the x-plane, the other in the y-plane, are sandwiched between two polarizers at  $+45$  and  $-45$  degrees. Expressing the polarization of the electric field as a vector,

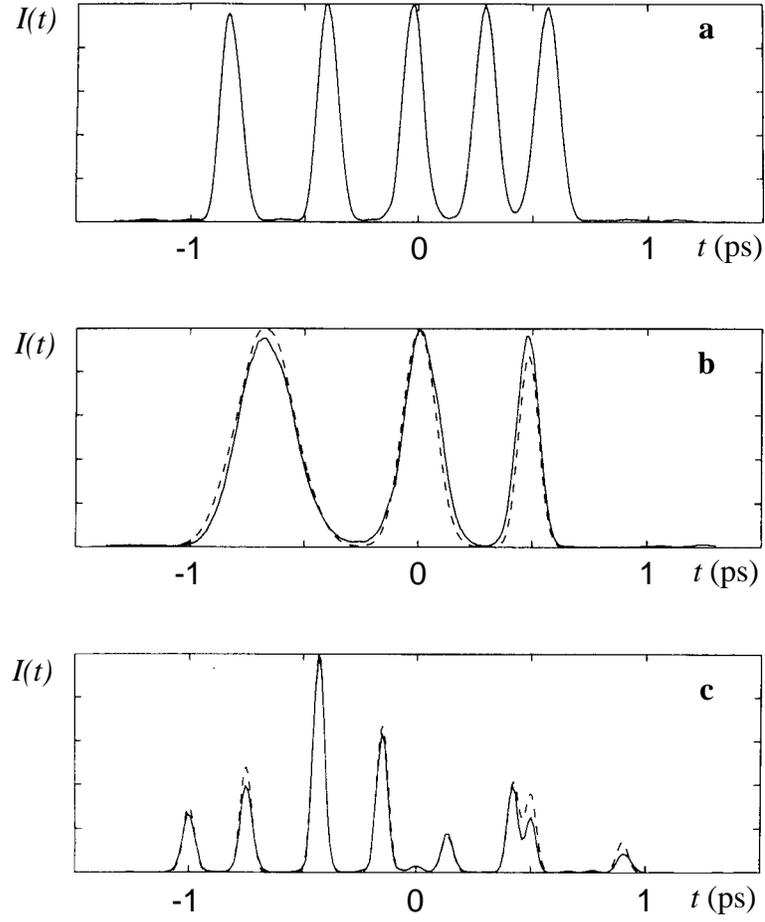


Figure 47. Examples of shaped pulses (from [34]).

$$\mathbf{E} = \begin{bmatrix} E_x \\ E_y \end{bmatrix}, \quad (77)$$

we can use a matrix notation for the  $\pm 45^\circ$  polarizers,

$$P_{45^\circ} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad P_{-45^\circ} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \quad (78)$$

and for  $x$ - and  $y$ - axis oriented masks,

$$M_x = \begin{bmatrix} e^{i\theta_x} & 0 \\ 0 & 1 \end{bmatrix} \quad M_y = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\theta_y} \end{bmatrix}. \quad (79)$$

Combining these four elements into a single mask matrix, we obtain

$$M = \begin{bmatrix} m & m \\ -m & -m \end{bmatrix}, \text{ where } m = \frac{1}{4} (e^{i\theta_x} - e^{i\theta_y}). \quad (80)$$

The output field is

$$\mathbf{E}_{\text{out}} = \begin{bmatrix} m & m \\ -m & -m \end{bmatrix} \begin{bmatrix} E_x \\ E_y \end{bmatrix} = 2m \begin{bmatrix} E_x + E_y \\ -(E_x + E_y) \end{bmatrix}, \quad (81)$$

polarized at  $-45^\circ$  as expected. By defining  $\delta$  and  $\sigma$  as

$$\delta \equiv \frac{1}{2} (\theta_x - \theta_y); \quad \sigma \equiv \frac{1}{2} (\theta_x + \theta_y), \quad (82)$$

we see that we have independent control over both the amplitude and phase of the output field:

$$m = \sin(\delta) e^{i\sigma}. \quad (83)$$

Each pixel of the mask, and thus each spectral component of the pulse, can be individually controlled in both amplitude and phase. The finite number of pixels places a limit on the spectral resolution of the mask.

Some examples of the kind of pulse shaping that is possible with programmable masks are shown in Fig. 32. The desired waveform is shown as a dashed line, and the measured waveform obtained with a 128-pixel phase and amplitude mask is shown as a solid line. The three waveforms shown are: a pulse sequence of decreasing separations, a pulse sequence with varying degrees of chirp, and an arbitrary pulse sequence.

The LC SLM is not the only available means for dynamic pulse shaping. An acousto-optic modulator (AOM) driven by microsecond radio-frequency electrical pulses can be used as a diffractive element in the Fourier plane of the pulse shaper [35]. This potentially offers the advantages of faster update times and no pixel gaps.

In addition to shaping pulses, the arrangement of Fig. 45 can be used for storage, recall, and processing of femtosecond waveforms. This is done by replacing the mask in the Fourier plane with a recordable hologram. The possibility of using holographic techniques for temporal pulse processing was first recognized and demonstrated by Weiner et al. [28]. These ideas have recently been extended to connect the spatial and temporal domains, by realizing that an ordinary hologram and one recorded with temporal signals inside a pulse-shaper are equivalent and interchangeable [29]. This is making possible such operations as encoding spatial information directly into a pulse sequence, retrieving time-encoded information as a spatial signal distribution, identifying specific signatures in an optical data stream, and high-speed all-optical data processing [30].

## Acknowledgments

I thank Prof. Eric Mazur, Prof. Baldassare Di Bartolo, Dr. Shrenik Deliwala, and Aryeh Feder for insightful discussions during the preparation of these lectures, and Prof.

Nicolaas Bloembergen, Paul Callan, Rich Finlay, Li Huang and Prof. Claus Klingshirn for their helpful comments on the manuscript.

I'm grateful to the Hertz Foundation for financial support throughout my graduate work.

## REFERENCES

1. T. W. Hansch, I. S. Shahin, and A. L. Schawlow, *Nature (Physical Science)* **235**, 63 (1972).
2. T. K. Cheng et al., *Appl. Phys. Lett.* **57**, 1004 (1990).
3. H. L. Fragnito et al., *Chem. Phys. Lett.* **160**, 101 (1989).
4. L. Cohen, *IEEE Proceedings* **77**, 941 (1989).
5. J. D. Jackson, *Classical Electrodynamics*, John Wiley and Sons, New York (1975), Section 7.8.
6. E. Wigner, *Phys. Rev.* **40**, 749 (1932).
7. J. Paye, *IEEE JQE* **28**, 2262 (1992).
8. R. L. Fork, B. I. Greene, C.V. Shank, *Appl. Phys. Lett.* **38**, 671 (1981).
9. D. E. Spence, P. N. Kean, and W. Sibbett, *Opt. Lett.* **16**, 42 (1991).
10. F. Krausz et al., *IEEE JQE* **28**, 2097 (1992).
11. E. P. Ippen, *Appl. Phys. B* **58**, 159 (1994).
12. D. Huang et al., *Opt. Lett.* **17**, 511 (1992).
13. R. L. Fork, O. E. Martinez, and J. P. Gordon, *Opt. Lett.* **9**, 150 (1984).
14. A. Stingl et al., *Opt. Lett.* **19**, 204 (1994); *Opt. Lett.* **20**, 602 (1995).
15. P. Maine et al., *IEEE JQE* **24**, 398 (1988).
16. O. E. Martinez, *IEEE JQE* **23**, 59 (1987).
17. D. Strickland and G. Mourou, *Opt. Comm.* **56**, 219 (1985).
18. B. E. Lemoff and C. P. J. Barty, *Opt. Lett.* **18**, 1651 (1993).
19. B. E. Lemoff et al., *Opt. Lett.* **17**, 1367 (1992).
20. M. Ramaswamy-Paye et al., *Opt. Lett.* **19**, 1756 (1994).
21. T. B. Norris, *Opt. Lett.* **17**, 1009 (1992).
22. F. Salin et al., *Opt. Lett.* **16**, 1964 (1991).
23. C. P. J. Barty et al., *Ultrafast Phenomena* **7** (1994).
24. R. Trebino and D. J. Kane, *J. Opt. Soc. Am. A* **10**, 1101 (1993).
25. K. W. DeLong et al., *Opt. Lett.* **19**, 2152 (1994).
26. G. Taft et al., *Opt. Lett.* **20**, 743 (1995).
27. W. S. Warren, H. Rabitz, and M. Dahleh, *Science* **259**, 1581 (1993); S. Rice, *Science* **258**, 412 (1992).
28. A. M. Weiner et al., *IEEE JQE* **28**, 2251 (1992).
29. M. C. Nuss et al., *Opt. Lett.* **19**, 664 (1994).
30. M. C. Nuss and R. L. Morrison, *Opt. Lett.* **20**, 740 (1995).
31. A. M. Weiner et al., *JOSA B* **5**, 1563 (1988).
32. J. P. Heritage, A. M. Weiner, and R. N. Thurston, *Opt. Lett.* **10**, 609 (1985).
33. A. M. Weiner et al., *Opt. Lett.* **15**, 326, (1990).
34. M. M. Wefers and K. A. Nelson, *Opt. Lett.* **20**, 1047, (1995).
35. C. W. Hillegas et al., *Opt. Lett.* **19**, 737 (1994).

## **Chapter 3**

### **Laser-Induced Band-Gap Collapse in GaAs**

The following paper appeared in Phys. Rev. B **51**, pp. 6959-6970 (1995).

# Laser-induced bandgap collapse in GaAs

E. N. Glezer, Y. Siegal, L. Huang, and E. Mazur

*Division of Applied Sciences and Department of Physics, Harvard University, Cambridge, MA 02138*

We present experimentally determined values of the dielectric constant of GaAs at photon energies of 2.2 eV and 4.4 eV following excitation of the sample with 1.9-eV, 70-fs laser pulses spanning a fluence range from 0 to 2.5 kJ/m<sup>2</sup>. The data show that the response of the dielectric constant to the excitation is dominated by changes in the electronic band structure and not by the optical susceptibility of the excited free carriers. The behavior of the dielectric constant indicates a drop in the average bonding-antibonding splitting of GaAs following the laser pulse excitation. This drop in the average splitting leads to a collapse of the bandgap on a picosecond time scale for excitation at fluences near the damage threshold of 1.0 kJ/m<sup>2</sup> and on a subpicosecond time scale at higher excitation fluences. The changes in the electronic band structure result from a combination of electronic screening of the ionic potential as well as structural deformation of the lattice caused by the destabilization of the covalent bonds.

## 1. Introduction

Reflectivity and second-harmonic generation studies of femtosecond laser-excited semiconductors show evidence of rapid changes in the material (within a few hundred femtoseconds) following the excitation.<sup>1, 2, 3, 4, 5, 6</sup> However, the nature of these material changes cannot be extracted from the results of these experiments. Interpretation of reflectivity and second-harmonic generation results is difficult because these quantities do not directly yield the behavior of intrinsic material properties. In particular, the reflectivity at a specific wavelength, polarization, and incident angle depends on both the real and imaginary parts of the dielectric constant at that wavelength. Furthermore, the measured second-harmonic radiation depends on the dielectric constant at both the fundamental and second-harmonic wavelengths as well as on the second-order susceptibility. The amount of information in the linear reflectivity and second-harmonic generation measurements, therefore, is not sufficient to uniquely determine the behavior of the linear or nonlinear optical material properties.

Without direct determination of the time-evolution of the dielectric constant, interpretation of reflectivity and second-harmonic data has relied to date on making assumptions about the functional form of the dielectric constant. Specifically, it has been assumed that the changes in the dielectric constant induced by the excitation are dominated by the free carrier contribution to the optical susceptibility. Under this assumption, the changes in the dielectric constant have been modeled using a Drude-model formalism.<sup>6, 7, 8</sup> While this type of assumption is legitimate at lower excitation regimes, it is misleading in the case of laser-induced disordering experiments. Misinterpretation of the data arises because a single-incident-angle reflectivity value does not correspond to a unique value of the dielectric constant. Thus, one can reproduce single-incident-angle reflectivity data using dielectric constant values that are completely different from the actual ones.

To avoid relying on an assumed functional form for the dielectric constant in interpreting the results of femtosecond pump-probe experiments on GaAs, we directly determined the time evolution of the real and imaginary parts of the dielectric constant. Specifically, we experimentally determined the behavior of the complex dielectric constant at photon energies of 2.2 eV and 4.4 eV following excitation with an intense, 70-fs pump pulse at 1.9 eV. To uniquely extract the real and imaginary parts of the dielectric constant, two independent measured quantities are necessary. Therefore, at each probe frequency we measured the  $p$ -polarized reflectivity at two carefully chosen angles of incidence using two simultaneous 70-fs probe beams.<sup>9, 10, 11</sup> We then converted each measured pair of reflectivities to the corresponding complex value of the dielectric constant as a function of pump-probe time delay. We verified that our two-angle technique yielded dielectric constant values consistent with reflectivity measurements at a third angle of incidence. The general approach of employing multiparameter optical probing to distinguish Drude and interband contributions to the dielectric constant has been used in other recent work, using ellipsometric techniques.<sup>12, 13, 14, 15</sup> The data we present in this paper show that for excitation fluences greater than 0.5 kJ/m<sup>2</sup>, the Drude model does not adequately describe the induced changes to the dielectric constant. The results indicate that changes in the interband transition contribution to the optical susceptibility dominate the behavior of the dielectric constant, as opposed to changes in the free carrier contribution as has been generally assumed. This finding implies that the previously observed initial reflectivity rise following high-intensity femtosecond laser-pulse excitation of semiconductors<sup>3, 4, 5, 6, 7</sup> is due to large changes in the electronic band structure which result from the excitation.

## **2. Optical Properties and Electronic Structure**

### *2.1 Interband Transition Contribution to the Dielectric Function*

The interband transition contribution  $\chi_{\text{interband}}(\omega)$  to the dielectric function arises from the coupling of states in different bands through the applied electric field. In a direct-gap semiconductor like GaAs,  $\chi_{\text{interband}}(\omega)$  is dominated by direct, or vertical, transitions and is therefore closely related to the joint density of states  $J(\omega)$  of the material. The joint density of states, in turn, is directly determined by the electronic band structure:<sup>16, 17</sup>

$$J(\omega) = \sum_{v,c} \frac{2}{(2\pi)^3} \int d^3k \delta[\omega_{vc}(\mathbf{k}) - \omega], \quad (1)$$

where  $v$  and  $c$  are respectively valence and conduction band indices,  $\hbar\omega_{vc}$  is the energy separation between band  $v$  and band  $c$  at a crystal momentum value of  $\mathbf{k}$ , and the integral is over the Brillouin zone. In the ground state (valence band filled and conduction band empty), the imaginary part of  $\chi_{\text{interband}}(\omega)$  approximately satisfies the relation<sup>16, 17</sup>

$$\text{Im}[\chi_{\text{interband}}(\omega)] \propto \frac{1}{\omega^2} J(\omega), \quad (2)$$

illustrating the connection between the electronic band structure and the interband transition contribution to the dielectric constant.

Most Group IV and III-V semiconductors have qualitatively similar joint densities of states, leading to similar dielectric functions in the ground state.<sup>16, 17, 18</sup> The solid curves in Fig. 1 show both the real and imaginary parts of the dielectric function of undoped GaAs at room temperature.<sup>19</sup> The three main features labeled  $E_0$ ,  $E_1$ , and  $E_2$  are common to the Group IV and III-V semiconductors although their locations and the relative sizes of these features vary from material to material. The point  $E_0$  (at 1.4 eV for GaAs) marks the fundamental band edge below which  $\text{Im}(\epsilon)$  is zero;  $E_1$  and  $E_2$ , (located at 3.0 eV and 4.75 eV, respectively, for GaAs<sup>20</sup>) label the two main absorption peaks in the spectrum. These peaks arise from regions in the band structure (the region around the L-point for the  $E_1$  peak and the region around the X-point for the  $E_2$  peak — see GaAs band structure diagram in Fig. 2) in which the valence band is roughly parallel to the conduction band, resulting in a large joint density of states for direct interband transitions.<sup>16</sup> The  $E_2$  peak, which is roughly coincident with the zero-crossing in  $\text{Re}(\epsilon)$ , is

the stronger of the two absorption peaks, and its location approximately gives the value of the average bonding-antibonding splitting of GaAs.<sup>17</sup>

In the simplest approximation, the overall shape of the GaAs dielectric function (the solid curves in Fig. 1) resembles that of a charged particle in a damped harmonic oscillator potential. We can model the GaAs dielectric function as a single average harmonic oscillator with a resonant frequency that corresponds to the average bonding-antibonding splitting,<sup>16</sup> and a width that is related to the spectral range over which the joint density of states is large. The dashed curves in Fig. 1 represent a fit of the single-oscillator dielectric function<sup>21</sup> to the ground-state dielectric function of GaAs. While this type of picture does not describe the structure in the semiconductor dielectric function in detail, it provides a simple physical interpretation for the overall shape and highlights some important characteristics. In particular, it illustrates that the main interband absorption peak in the semiconductor dielectric function has the same features as an oscillator resonance: a zero-crossing in the real part coinciding with a peak in the imaginary part. This simplified picture will be helpful in the interpretation of the data presented in Section 4.

## *2.2 Free-Carrier Contribution to the Dielectric Function — The Drude Model*

The free carrier contribution to the dielectric constant of a semiconductor is generally treated in the framework of the Drude model for the ac conductivity of free electrons.<sup>8, 22, 23</sup> For a semiconductor with a conduction-band electron density of  $N_e$  and a valence-band hole density of  $N_h$ , the Drude contribution of the free carriers to the optical susceptibility is given by

$$\chi_{\text{Drude}}(\omega) = \frac{ie^2}{\omega} \left[ \frac{N_e \tau_e}{m_e^* (1 - i\omega\tau_e)} + \frac{N_h \tau_h}{m_h^* (1 - i\omega\tau_h)} \right], \quad (3)$$

where  $e$  is the charge of the electron, and  $m_{e,h}^*$  and  $\tau_{e,h}$  are the effective mass and mean collision time of the free electrons and holes. If the free carriers are produced by optical

excitation, then  $N_e = N_h = N$ , where  $N$  is the number of electron-hole pairs created by the excitation.

The general shape of the Drude optical susceptibility is very different from that of the optical susceptibility near an interband transition resonance. Defining a reduced effective mass  $m^*$  such that  $1/m^* = 1/m_e^* + 1/m_h^*$  and assuming that  $\tau_e \approx \tau_h \approx \tau$ , the real and imaginary parts of  $\chi_{\text{Drude}}(\omega)$  are given by

$$\text{Re}[\chi_{\text{Drude}}(\omega)] = -\frac{Ne^2\tau^2}{m^*(1 + \omega^2\tau^2)} \quad (4a)$$

and

$$\text{Im}[\chi_{\text{Drude}}(\omega)] = \left(\frac{1}{\omega\tau}\right)\frac{Ne^2\tau^2}{m^*(1 + \omega^2\tau^2)}. \quad (4b)$$

Equations (4a-b) show that  $\text{Re}[\chi_{\text{Drude}}(\omega)]$  is a featureless, monotonically-increasing function of  $\omega$  while  $\text{Im}[\chi_{\text{Drude}}(\omega)]$  is a featureless, monotonically-decreasing function of  $\omega$ . If  $\omega\tau \gg 1$ , which is generally the case for semiconductors in the visible frequency range, the magnitude of the imaginary part is much smaller than that of the real part, and we can write  $\chi_{\text{Drude}}(\omega) \approx -\omega_p^2/4\pi\omega^2$ , where  $\omega_p \equiv [4\pi Ne^2/m^*]^{1/2}$  is the plasma frequency. The Drude dielectric function  $\epsilon(\omega) = 1 + 4\pi\chi_{\text{Drude}}(\omega)$  goes through a resonance at  $\omega = \omega_p$  as  $\text{Re}[\epsilon(\omega)]$  passes through zero. This plasma resonance corresponds to the excitation of plasmons, or collective oscillations of the free-carrier charge density.<sup>24</sup> Note that  $\text{Im}[\epsilon(\omega)]$  continues to decrease as a function of  $\omega$  at a plasma resonance, in sharp contrast to the peak shape of  $\text{Im}[\epsilon(\omega)]$  at an interband transition resonance. At fixed  $\omega$  in a semiconductor, one can in principle go through a plasma resonance by increasing the free carrier density  $N$ . In this case,  $\text{Re}[\epsilon(N, \omega)]$  decreases through zero as the carrier density increases while  $\text{Im}[\epsilon(N, \omega)]$  increases monotonically with carrier density. The contrast between a plasma resonance and an interband transition resonance will further aid in the interpretation of the experimental data in Section 4.

### *2.3 Standard Assumption for $\epsilon(\omega)$ Following Laser Excitation*

As evidenced by changes observed in the reflectivity, excitation of a semiconductor with a laser pulse modifies the dielectric constant of the material.<sup>1, 2, 3, 4, 8, 25, 26, 27</sup> The model generally used for the time-dependent dielectric function  $\varepsilon(\omega, t)$  of a semiconductor following laser-pulse excitation involves the assumption that the interband contribution to the susceptibility does not change significantly as a result of the excitation:  $\chi_{\text{interband}}(\omega, t) \approx \chi_{\text{interband}}(\omega)$ .<sup>7, 28</sup> Instead, changes in the dielectric function are attributed mainly to the time dependence of  $\chi_{\text{Drude}}(\omega, t)$  arising from the time-dependent free-carrier density  $N(t)$  produced by the excitation. This model can be expressed as

$$\varepsilon(\omega, t) = 1 + 4\pi \left[ \chi_{\text{interband}}(\omega) + \frac{iN(t)e^2\tau}{m^* \omega(1 - i\omega\tau)} \right]. \quad (5)$$

The free-carrier density  $N(t)$  created by the laser-pulse excitation depends on the excitation fluence  $\phi$ . At a fixed time  $t$  after the excitation,  $N(\phi, t)$  increases monotonically with excitation fluence. So according to Eq. (5) one would expect  $\text{Re}[\varepsilon(\omega, \phi, t)]$  to decrease monotonically and  $\text{Im}[\varepsilon(\omega, \phi, t)]$  to increase monotonically with increasing excitation fluence. Furthermore, with  $\omega\tau \gg 1$ , the increase in  $\text{Im}[\varepsilon(\omega, \phi, t)]$  should be only slight.

While this Drude-like model for the dielectric constant of a laser-excited semiconductor adequately describes data at low and moderate carrier excitation levels ( $N \leq 10^{20} \text{ cm}^{-3}$ ),<sup>12, 29</sup> the data we present below show that it cannot be used to analyze reflectivity data at high carrier excitation levels ( $N \geq 10^{21} \text{ cm}^{-3}$ ). The experimental results in Section 4 indicate that, in this strong excitation regime, the response of the dielectric function to the laser pulse excitation is dominated by changes in the interband transition contribution to the dielectric constant resulting from a major alteration of the electronic band structure. The difference between the expected behavior of  $\varepsilon(\omega, \phi, t)$  based on previous assumptions and the experimentally observed behavior will be highlighted in Section 4.5.

### **3. Experimental Setup**

#### *3.1 Two-Color Amplified Femtosecond Laser System*

The results presented in this chapter involve two sets of measurements made using a two-color pump-probe technique. One set consists of data taken using a 70-fs, 1.9-eV pump beam and a simultaneous pair of 70-fs, 2.2-eV probe beams while the other set consists of data taken using the same pump beam conditions but a doubled probe photon energy of 4.4 eV. To generate pump and probe beams at different frequencies, we pass the amplified output of a colliding pulse modelocked laser through a 20-mm, single-mode, polarization-preserving optical fiber.<sup>30</sup> Self-phase modulation in the fiber broadens the spectrum of the input pulse from 5 to 200 nm. By splitting this continuum beam with a broadband beamsplitter, we can independently amplify different spectral regions within the 200-nm bandwidth using two separate amplifier chains.<sup>31</sup> A three-stage amplifier using the dye DCM produces a 300- $\mu$ J pump beam centered at 635 nm with a 20-nm bandwidth; a two-stage amplifier using the dye Rhodamine 6G produces a 30- $\mu$ J probe beam centered at 570 nm with a 10-nm bandwidth. Both amplifiers are pumped by a frequency-doubled, 10-Hz Nd:YAG laser. Following amplification, each beam is compressed by a separate grating pair to a 70-fs pulse width (FWHM).

#### *3.2 Time-Resolved Determination of the Dielectric Constant*

To determine both the real and imaginary parts of the dielectric constant of GaAs with femtosecond time resolution, we simultaneously measure the reflectivity at two different angles of incidence as a function of pump-probe time delay.<sup>31</sup> In the 2.2-eV experiment, the pair of probe beams is produced simply by splitting the 570-nm beam in two. In the 4.4-eV experiment, we first double the 570-nm beam in a 100- $\mu$ m thick BBO crystal before splitting the beam. The probing geometries for both experiments are summarized in Fig. 3. In both cases, the incident beams are polarized in the plane of incidence and are focused onto the same spot on an insulating (110) GaAs wafer (Cr

doped,  $\rho > 7 \times 10^7 \Omega \text{ cm}$ ), which is exposed to air. To monitor a uniformly excited region, we focus the probe beams more tightly than the pump beam: the probed surface area is about 16 times smaller than the  $0.01\text{-mm}^2$  focal area of the pump beam on the sample. Uniform excitation in the probed region is further assured by the smaller penetration depth of the probe beams (between 5 and 170 nm at 2.2 eV and 5 and 10 nm at 4.4 eV)<sup>32</sup> compared to that of the pump beam (270 nm). The pump pulse fluence at each pump-probe time delay spans a range from 0 to  $2.5 \text{ kJ/m}^2$ . The probe beam fluence never exceeds  $0.1 \text{ kJ/m}^2$  so as not to produce any detectable changes in the dielectric constant. To avoid cumulative damage effects, we translate the sample during data collection so that each data point is obtained at a new spot on the sample.

We convert each pair of reflectivity measurements to the corresponding real and imaginary parts of the dielectric constant by numerically inverting the Fresnel formula for reflectivity as a function of incident angle. Setting one of the probe beam angles of incidence to the Brewster angle provides good sensitivity in distinguishing changes in  $\text{Re}(\epsilon)$  from changes in  $\text{Im}(\epsilon)$  because the  $p$ -polarized reflectivity at this angle is determined mainly by  $\text{Im}(\epsilon)$ .<sup>33</sup> We base our choice of the second incident angle, which is not as critical, on constraints in the experimental setup. The two angles of incidence for the 2.2-eV measurements were  $75.8^\circ$  and  $70.9^\circ$ ; in the 4.4-eV measurements we used  $76.0^\circ$  and  $58.5^\circ$ . To keep the angular separation between pump and probes small, we used a  $63^\circ$  incident angle for the  $p$ -polarized pump beam in the 2.2-eV experiment and a  $68^\circ$  incident angle in the 4.4-eV experiment.

Because the GaAs surface oxidizes in air, we use a three-phase model (air-oxide-GaAs) in converting the reflectivity data to values for the dielectric constant. Surface roughness effects can be accounted for in this model by adjusting the effective thickness of the oxide layer.<sup>34</sup> We calibrated the effective thickness of the oxide layer by measuring reflectivity as a function of incident angle in the absence of excitation by a pump pulse. Using the ground-state dielectric constant of GaAs<sup>19</sup> and a value of  $\epsilon = 4$  for the dielectric

constant of the oxide layer,<sup>34</sup> we fit the three-phase model to the measured angle dependence with the effective oxide layer thickness as a fit parameter. This procedure consistently yielded an effective thickness for the oxide layer of about 4 nm. In the 2.2-eV experiment the obtained value was  $4.2 \pm 0.4$  nm while in the 4.4-eV experiment the value was  $4.4 \pm 0.4$  nm.

As a consistency check of our determination of the dielectric constant, we measured the time-evolution of the reflectivity at a third, completely different, angle of incidence ( $45^\circ$  for the 2.2-eV case and  $35.5^\circ$  for the 4.4-eV case) under similar pump pulse excitation conditions at both probe frequencies. We then calculated the expected reflectivity at that third incident angle using our experimentally determined time- and fluence-dependent values for the dielectric constant. The measured reflectivity showed excellent agreement with the calculated reflectivity at this third angle of incidence for both the 2.2-eV and 4.4-eV data. An example of this consistency check is shown in Section 4.5 (see Fig. 11).

## 4. Experimental Results

### 4.1 Dielectric Constant at 2.2 eV

Figures 4 and 5 summarize the experimental data on the dielectric constant at 2.2 eV. In Fig. 4,  $\text{Re}(\epsilon)$  (filled circles) and  $\text{Im}(\epsilon)$  (open circles) are plotted vs. pump-probe time delay for four different excitation fluences; in Fig. 5,  $\text{Re}(\epsilon)$  and  $\text{Im}(\epsilon)$  are plotted vs. pump fluence at four different time delays. The change induced in the dielectric constant by the pump pulse excitation is completely different from that expected from the free carrier contribution to the optical susceptibility. At pump fluences near  $1 \text{ kJ/m}^2$ ,  $\text{Im}(\epsilon)$  starts at an initial value of about 2, rises to a peak near 60, and then drops to somewhere between 10 and 15 — a strong contrast to the slight, featureless increase predicted by the Drude model.  $\text{Re}(\epsilon)$ , meanwhile, initially decreases slightly but then sharply increases before

dropping through zero. Note that the zero-crossing of  $\text{Re}(\epsilon)$  roughly coincides with the peak in  $\text{Im}(\epsilon)$ .

The results shown in Figs. 4 and 5 indicate that a strong absorption peak comes into resonance with the probe frequency as a result of the excitation. The resonance behavior is most striking in Fig. 5 because the features are particularly clear when plotted versus pump fluence. This behavior must result from an interband absorption peak and not from a free carrier plasma resonance because the zero-crossing in  $\text{Re}(\epsilon)$  is accompanied by a peak in  $\text{Im}(\epsilon)$  rather than by a steady increase. From the behavior of  $\text{Re}(\epsilon)$ , we can infer the time evolution of this interband absorption peak. Because  $\text{Re}(\epsilon)$  is initially positive, the resonant frequency of the observed absorption peak evidently starts out higher than the probe frequency; it then sweeps down through the probe frequency as  $\text{Re}(\epsilon)$  drops through zero.

The rate at which the resonant frequency of the absorption peak drops through the probe frequency depends on the strength of the excitation: the higher the pump fluence, the faster  $\text{Re}(\epsilon)$  drops through zero. Figure 6 illustrates this dependence by showing the time delay at which  $\text{Re}(\epsilon)$  crosses through zero plotted vs. pump fluence. For fluences around  $2.0 \text{ kJ/m}^2$ , the absorption peak comes into resonance with the probe frequency within a few hundred femtoseconds; at fluences just above  $0.8 \text{ kJ/m}^2$ , on the other hand, the absorption peak takes on the order of 10 picoseconds to come into resonance. For fluences below  $0.8 \text{ kJ/m}^2$ ,  $\text{Re}(\epsilon)$  never goes through zero, indicating that the excitation is not strong enough to bring the resonant frequency of the peak down to the probe frequency.

The dashed line in Fig. 6 at  $1.0 \text{ kJ/m}^2$  indicates the threshold fluence for permanent damage to the sample. We determined this threshold by correlating pump pulse fluence with the size of damage spots on the sample measured through a microscope. Above the damage threshold the pump pulse induces irreversible changes in the sample while below the damage threshold the induced changes are reversible. Measurements taken

several seconds after the excitation confirm that the dielectric constant eventually returns to its initial value for fluences below the damage threshold. Note, however, that the absorption peak comes into resonance with the probe frequency even for pump fluences below this damage threshold.

The use of the Fresnel equations to extract the dielectric constant from reflectivity measurements presumes the existence of a sharp boundary between two different media. While this assumption is clearly justified below the damage threshold, at higher fluences density gradients can develop due to hydrodynamic expansion at the surface.<sup>13</sup> Furthermore, the hydrodynamic expansion can lead to the formation of an absorptive cloud, accompanied by non-specular scattering of the probe light. The formation of such an absorptive cloud has been seen in silicon, beginning at about 10 ps after excitation by a pulse with fluence about five times greater than the damage threshold.<sup>2</sup> In GaAs, a decrease in the specular reflectivity and the onset of non-specular scattering of the probe light, also beginning about 10 ps after the excitation, has been observed above the damage threshold.<sup>27</sup> To remain within the validity range of the sharp boundary assumption we limit the range of time delays over which we extract the dielectric constant to 8 ps. By measuring the reflectivity at a third angle for both 2.2 eV and 4.4 eV experiments we verified that the Fresnel equations are indeed valid over our entire experimental range.

#### *4.2 Qualitative Picture — Bandgap Collapse*

It is useful at this point to develop a qualitative interpretation of the data presented in the preceding section; a more quantitative analysis of the results will appear in Section 4.4. The first step in interpreting the data is to attach a physical significance to the interband absorption peak. As discussed in Section 2.1, the dielectric function of GaAs can be approximated by that of a damped single harmonic oscillator with a resonant frequency equal to the average bonding-antibonding splitting.<sup>16</sup> Identifying the interband

absorption peak in the data with the harmonic oscillator absorption peak, we can think of the shift in this peak as a drop in the average bonding-antibonding splitting. In GaAs, the average bonding-antibonding splitting in the ground state is about 4.75 eV.<sup>17, 20</sup> The 2.2-eV data then indicate that the laser-pulse excitation induces a drop in the average bonding-antibonding splitting from 4.75 eV to below 2.2 eV. Note that this drop in the average splitting by more than a factor of two occurs even for fluences below the damage threshold, an excitation regime in which the induced changes are *reversible*.

Figure 7 illustrates schematically the qualitative picture that emerges from the 2.2-eV data. The average bonding-antibonding splitting  $\Delta E_{b-a}$  starts out far above 2.2 eV, so the probe photon energy lies at the foot of the single-oscillator absorption peak where  $\text{Im}(\epsilon)$  is small (step 1 in Fig. 7). As a result of the excitation in step 1,  $\Delta E_{b-a}$  begins to decrease, leading to a downward shift in the resonant frequency of the single-oscillator absorption peak and therefore a rise in  $\text{Im}(\epsilon)$  at 2.2 eV (step 2 in Fig. 7).<sup>35</sup> As  $\Delta E_{b-a}$  drops past 2.2 eV (step 3 in Fig. 7),  $\text{Im}(\epsilon)$  goes through a peak. If  $\Delta E_{b-a}$  drops far enough, the minimum in the conduction band will drop below the maximum in the valence band and the semiconductor will take on metallic properties (step 4 in Fig. 7). The 2.2-eV dielectric constant at high fluences and long time delays, after the collapse of the bandgap, is consistent with the characteristics of a poor metal and is similar to that of liquid silicon<sup>22</sup> and liquid carbon<sup>36</sup> produced by laser pulse excitation.

This interpretation of the 2.2-eV data in terms of a drop in the average bonding-antibonding splitting allows us to predict qualitatively the behavior of the dielectric constant at different photon energies under similar excitation conditions. In particular, for a given excitation strength, the dielectric constant at a probe photon energy between 2.2 eV and 4.75 eV should exhibit resonance features at an earlier pump-probe time delay than the dielectric constant at 2.2 eV. Equivalently, for a fixed pump-probe time delay, the dielectric constant at a probe photon energy in the above range should exhibit resonance features at a lower pump fluence than the dielectric constant at 2.2 eV.

### *4.3 Dielectric Constant at 4.4 eV*

To verify the interpretation described in the preceding section, we measured the behavior of the dielectric constant at 4.4 eV, which is only slightly below the initial value of the average bonding-antibonding splitting of GaAs. Figures 8 and 9 summarize the data at 4.4 eV. Figure 8 presents the time-dependence of the 4.4-eV dielectric constant at the same four pump fluences shown in Fig. 4, and Fig. 9 shows the fluence dependence of the 4.4-eV dielectric constant for fixed pump-probe time delay. Note that at equal pump fluence the peak in  $\text{Im}(\epsilon)$  and the zero-crossing in  $\text{Re}(\epsilon)$  occur at earlier time delays in the 4.4-eV case than in the 2.2-eV case. Correspondingly, these features occur at lower fluences in the 4.4-eV data than in the 2.2-eV data for equivalent time delays. This behavior stands out in Fig. 10, which adds to Fig. 6 the corresponding points for the 4.4-eV results.  $\text{Re}(\epsilon)$  at 4.4 eV crosses zero for fluences as low as  $0.5 \text{ kJ/m}^2$  compared to the minimum fluence of  $0.8 \text{ kJ/m}^2$  required for a zero-crossing at 2.2 eV.

The behavior of the dielectric constant at 4.4 eV is indeed consistent with the picture described in Section 4.2 of a drop in the average bonding-antibonding splitting. Following the pump pulse excitation,  $\Delta E_{\text{b-a}}$  drops from its initial value of about 4.75 eV first past 4.4-eV and then continues down past 2.2-eV. A stronger excitation causes a faster drop through both probe frequencies. At pump fluences between  $0.5 \text{ kJ/m}^2$  and  $0.8 \text{ kJ/m}^2$ , the excitation is strong enough to bring the resonant frequency of the absorption peak below 4.4 eV but not all the way down to 2.2 eV. Note that since 4.4 eV is close to the initial value of the average bonding-antibonding splitting,  $\text{Im}[\epsilon(4.4 \text{ eV})]$  does not rise much above its initial value before coming down.

### *4.4 The Role of Free-Carrier Contributions*

The striking resonance behavior we observe indicates that the response of the dielectric constant to the excitation is dominated by changes in the electronic band

structure rather than by free-carrier contributions to the optical susceptibility. However, this does not mean that the free-carrier contributions are absent. One can get a sense for the size of these contributions to the optical susceptibility by looking at low fluence data where the band structure changes are not strong enough to overwhelm the free-carrier effects. At  $0.5 \text{ kJ/m}^2$ , for example,  $\text{Re}(\epsilon)$  at  $2.2 \text{ eV}$  shows an initial decrease within a few hundred femtoseconds and then recovers on a picosecond time scale (see top graph in Fig. 4). This behavior is consistent with the expected effect of the optical susceptibility of the excited free carriers on  $\text{Re}(\epsilon)$ .<sup>37</sup> According to Eq. (4a), the initial creation of free carriers by the pump pulse should be accompanied by a negative contribution to  $\text{Re}(\epsilon)$ . As the free carrier population decreases due to strong Auger recombination, the magnitude of the free-carrier contribution decreases and  $\text{Re}(\epsilon)$  recovers. The observed recovery time of a few picoseconds agrees with the predicted high carrier-density Auger recombination time predicted by Yoffa.<sup>38</sup> However, even at  $0.5 \text{ kJ/m}^2$ , the changes in the interband transition contribution are already large enough to affect the dielectric constant at  $2.2 \text{ eV}$ , as evidenced by the behavior of  $\text{Im}(\epsilon)$ , which shows no recovery on a picosecond time scale. At fluences exceeding  $0.6 \text{ kJ/m}^2$ , the interband transition resonance behavior at  $2.2 \text{ eV}$  is so strong that it masks the free-carrier contributions to the dielectric constant.

#### *4.5 Comparison of Observed Behavior with Expected Behavior*

Figure 11 highlights the unexpected nature of the experimental results by comparing the observed behavior of the dielectric constant with that expected based on the standard assumption expressed in Eq. (5). This figure also illustrates the ambiguity inherent in the interpretation of single angle-of-incidence reflectivity measurements. The top graph in Fig. 11a shows the experimentally determined dielectric constant at  $2.2 \text{ eV}$  plotted against pump fluence at a time delay of  $2 \text{ ps}$  (this corresponds to the third graph in Fig. 5). The dielectric constant values in this graph are used to calculate the

corresponding  $p$ -polarized reflectivity at an incident angle of  $45^\circ$ , shown by the curve in the bottom graph in Fig. 11a. This curve agrees with the experimental values of the  $45^\circ$   $p$ -reflectivity also measured at a time delay of 2 ps and represented by the open squares in the graph. Fig. 11a therefore provides a consistency check for our two-angle technique. For comparison, Fig. 11b presents in the top graph a Drude-like dielectric constant based on Eq. (5), without the restriction  $\omega\tau \gg 1$ , that is chosen to also reproduce the measured  $45^\circ$ -reflectivity values, as seen in the bottom graph. It is important to emphasize that the distinctive resonance features of the actual, experimentally determined dielectric constant in Fig. 11a produce the same reflectivity values at  $45^\circ$  incidence as the Drude-like behavior of the computer-generated, but incorrect, dielectric constant in Fig. 11b. However, the two dielectric constants reflect very different underlying physics: changes in  $\chi_{\text{interband}}(\omega)$  dominate the former while the latter would indicate the prevalence of changes in  $\chi_{\text{Drude}}(\omega)$ . Thus, experimental determination of the complex dielectric constant is essential to understanding laser-induced phase transitions in semiconductors.

## 5. Analysis and Discussion

### 5.1 Quantitative Analysis — Double-Oscillator Model

To refine the qualitative picture we have so far developed of a laser-induced drop in the average bonding-antibonding splitting,<sup>9</sup> we can use a damped double-oscillator model that better captures the features of the ground-state dielectric function of GaAs. The double-oscillator dielectric function is given by the expression

$$\varepsilon(\omega) = 1 + \omega_p^2 \left[ \frac{f_1}{(\omega_1^2 - \omega^2 + i\omega\Gamma_1)} + \frac{f_2}{(\omega_2^2 - \omega^2 + i\omega\Gamma_2)} \right], \quad (6)$$

where  $\omega_i$  and  $\Gamma_i$  are the resonant frequency and the spectral width of the  $i^{\text{th}}$  oscillator. The oscillator strength  $f_i$ , which obeys the sum rule  $f_1 + f_2 = 1$ , corresponds to the fraction of the total number of electrons  $N$  (appearing in  $\omega_p^2$ ) that contributes to the

strength of the  $i^{\text{th}}$  oscillator. Fitting Eq. (6) to the GaAs ground-state dielectric function<sup>19</sup> using  $\hbar\omega_1 = 3.10$  eV and  $\hbar\omega_2 = 4.75$  eV to match the location of the  $E_1$  and  $E_2$  absorption peaks yields  $\hbar\Gamma_1 = 1.6$  eV,  $\hbar\Gamma_2 = 0.68$  eV,  $\hbar\omega_p = 13.7$  eV,  $f_1 = 0.16$ , and  $f_2 = 0.84$ . The higher value of  $f_2$  compared to  $f_1$  indicates that a larger region of the band structure contributes to the  $E_2$  peak than to the  $E_1$  peak.

To model the response of the dielectric function to the laser pulse excitation using the double-oscillator model, we fit Eq. (6) simultaneously to the real and imaginary parts of the dielectric constant at *both* 2.2 eV and 4.4 eV for each pump fluence and time delay. The fit parameters here are  $\omega_1$ ,  $\omega_2$ ,  $\Gamma_1$ , and  $\Gamma_2$  ( $\omega_p$ ,  $f_1$  and  $f_2$  are held constant), implying that both the location and the width of each of the two absorption peaks can vary as a result of the excitation. To ensure reasonable fluence and time evolution of the parameters, we impose the additional constraint when fitting Eq. (6) to the data that each parameter not change too much for each subsequent fluence or time step. For measurements taken at a pump-probe time delay of  $-0.33$  ps, before the pump pulse arrives at the sample, the double-oscillator fit to the data produces values of  $\hbar\omega_1 = 3.10$  eV,  $\hbar\omega_2 = 4.77$  eV,  $\hbar\Gamma_1 = 1.9$  eV,  $\hbar\Gamma_2 = 0.73$  eV, closely reproducing the ground-state dielectric function (see top graphs in Fig. 12). This result for negative pump-probe time delay provides a consistency check for the double-oscillator fitting routine.

Figure 12 shows the time-evolution of the full double-oscillator dielectric function (plotted versus photon energy) for excitation with a pump fluence of  $1.0$  kJ/m<sup>2</sup>. As seen in the figure, the double-oscillator model fits the observed values of the complex dielectric constant at both 2.2 and 4.4 eV very well. The fitting results at other pump fluences are in similar agreement with the data. The most interesting aspect of the time-evolution of the dielectric function shown in Fig. 12 is that it initially exhibits primarily a broadening of the absorption peaks followed by a shift to lower photon energy in the location of the resonance features.

We can interpret the plots in Fig. 12 in terms of the electronic band structure of GaAs (see Fig. 2). The initial broadening and shift of the absorption peaks may result from the valence and conduction bands becoming less parallel in the regions around the L- and X-points, which contribute respectively to the  $E_1$  and  $E_2$  peaks. Such an effect would occur, for example, if the direct gap at the X-point became significantly smaller while the direct gap at the  $\Gamma$ -point did not. The subsequent narrowing of the absorption peaks as they continue to slide to even lower photon energies suggests that as parts of the valence and conduction bands get closer together, they also become more parallel. This type of behavior could arise, for instance, from the region between the L-point and the  $\Gamma$ -point if the direct gap at the L-point decreased from its initial value of about 3 eV while the direct gap at the  $\Gamma$ -point continued to remain roughly the same at about 1.4 eV. The behavior of  $\epsilon(\omega)$  near  $\omega = 0$  (infrared frequencies) at a pump-probe time delay of 4.0 ps, reminiscent of the infrared spectrum of metals, is consistent with a collapse of the minimum band gap. The large, positive values of  $\text{Im}(\epsilon)$  and large, negative values of  $\text{Re}(\epsilon)$  in this frequency range develop as the resonant frequency of one of the absorption peaks approaches zero. Once the minimum band gap goes to zero, one would indeed expect a resonance at  $\omega = 0$ .

It is important to note that we fit the four-parameter double-oscillator model to just four independent values ( $\text{Re}(\epsilon)$  and  $\text{Im}(\epsilon)$  at 2.2 and 4.4 eV) at each pump fluence and time delay. Thus, while the experimental data presented in Section 4 strongly indicate significant band structure changes marked by a drop in the average bonding-antibonding splitting, the details of these changes are not uniquely determined by the double-oscillator model analysis. Further refinement of our picture of the laser-induced changes in the electronic band structure requires an experimental determination of the dielectric function over a wide range of probe photon energies rather than at a few discrete points. To this end, we are currently developing a broadband probe as well as incorporating broadband light detection capabilities into the two-angle reflectivity technique. We

should stress, however, that although we do not currently know all of the details, the data presented in this paper conclusively show that intense femtosecond laser-pulse excitation of GaAs leads to a major change in the electronic band structure marked by a drop in the average separation between the valence band and the conduction band.

### *5.2 Underlying Physics — Electronic Screening and Structural Change*

What underlying physical effects are responsible for this alteration of the band structure? To answer this question, we should examine two main sources of band structure modification: electronic screening and structural change. Through electron-hole pair generation, the pump pulse creates a large population of mobile charge carriers that can partially screen the ionic potential in the material. Since the average bonding-antibonding splitting increases with the strength of the ionic potential,<sup>17</sup> electronic screening reduces the average bonding-antibonding splitting. Note that electronic screening, through its modification of the band structure, changes the interband transition contribution to the dielectric constant and is therefore completely different from the direct intraband contribution of the free carriers to the dielectric constant through a Drude term. A recent calculation shows that when 10% of the valence electrons are excited to the conduction band, the direct gap at the X-point in the band structure of GaAs will decrease by roughly 2 eV due to electronic screening and many-body bandgap renormalization while the direct gap at the  $\Gamma$ -point changes only slightly.<sup>39</sup> Such an effect would lead to an immediate broadening and shift of the absorption peaks in the dielectric function as discussed in Section 5.1.

Because the strength of electronic screening increases with the free-carrier density, the effects due to screening should follow the carrier density instantaneously. Therefore, the effect of screening on the dielectric constant should be largest when the free carrier density is highest—*i.e.* immediately following the excitation. As Auger recombination and diffusion reduce the free-carrier density, the influence of electronic screening on the

band structure should correspondingly decrease. The time scale for electronic screening is consistent with that for the initial broadening of the absorption peaks, seen in Fig. 12 at the earliest time delays. However, the subsequent changes in the dielectric constant, marked primarily by a shift in the resonance features, continue to grow with time delay during the picoseconds following the excitation, when the carrier density is already decreasing from its peak value. Electronic screening by itself, therefore, cannot explain all of the observed behavior.

To account for the time evolution of the shift in the resonance features of the dielectric constant, we must examine the effect of lattice structural change on the band structure. The electronic band structure is determined by the crystal structure. The semiconducting behavior of Group IV and III-V materials such as GaAs arises from the tetrahedrally-coordinated diamond or zincblende arrangement of the constituent atoms. If this arrangement is disturbed, the electronic properties will change accordingly. In general, deformation of the diamond or zincblende structure leads to a collapse of the band gap and a semiconductor-metal transition.<sup>17, 40, 41, 42</sup> Just a 10% change in average bond length is enough to cause a semiconductor-metal transition.<sup>42</sup> Note that an ionic velocity as small as 25 m/s is already sufficient to achieve a 10% change in the GaAs bond length within 1 ps.

Because the covalent bonds of semiconductors like GaAs are stabilized by the valence electrons, excitation of a sufficient number of electrons from bonding valence states to antibonding conduction states can lead directly to lattice instability.<sup>43, 44, 45</sup> If the femtosecond pump pulse is intense enough to excite this critical density of electrons, the resulting instability in the lattice will cause the atoms to move towards a new minimum potential energy configuration. This deformation of the lattice begins immediately following the excitation but continues long after the incidence of the pump pulse. The change in the dielectric constant accompanying the lattice deformation should

therefore continue to increase in the picoseconds following the excitation, in agreement with the observed behavior of the dielectric constant.

Note that even below the damage threshold, in the fluence range of 0.8–1.0 kJ/m<sup>2</sup> the laser-pulse excitation induces a drop in the average bonding-antibonding splitting from 4.75 eV down to 2.2 eV within 8 ps. This suggests that reversible structural change takes place below the permanent damage threshold. The changes in the optical properties are considerably greater than those that could result from lattice heating, even up to the melting temperature of 1511 K. The effect of lattice heating on the dielectric constant has been observed at much lower excitation fluences (10<sup>-3</sup> kJ/m<sup>2</sup>) in Ge and Si<sub>1-x</sub>Ge<sub>x</sub> alloys<sup>12</sup>. Based on the temperature dependence of the optical properties of silicon<sup>46</sup>, heating from 10 K up to 1000 K results in about 0.4-eV decrease in the average optical gap. This decrease is linear with temperature above roughly 200 K. Similar temperature dependence has been measured in GaAs, although only up to 300 K, by Walter *et al.*<sup>47</sup> It is quite unlikely that even near the melting temperature the average bonding-antibonding splitting would drop down to 2.2 eV. Finally, the observed resonance in the 0.8–1.0 kJ/m<sup>2</sup> fluence range is very similar to that observed above 1.0 kJ/m<sup>2</sup>, only slower, indicating that even below the damage threshold the underlying cause for the drop in the splitting is structural change. The value of 0.8 kJ/m<sup>2</sup> does not represent a fundamental threshold, but is simply the lowest fluence that brings the bonding-antibonding splitting into resonance with the 2.2 eV probe frequency within 8 ps.

Below an excitation fluence of 0.8 kJ/m<sup>2</sup> it is more difficult to identify the source of the observed changes in optical properties. For example, at 0.5 kJ/m<sup>2</sup> excitation (Figure 4, top graph), we see that the imaginary part of the 2.2 eV dielectric constant has not recovered after 8 ps. Even more pronounced is the drop in the real part of the dielectric constant at 4.4 eV, which also does not recover by 8 ps (Figure 8, top graph). These persistent effects are due to changes in the bandstructure, and correspond to a

downward shift and broadening of the resonant features in the dielectric spectrum. The source of these changes may be lattice heating (thermal motion of the ions), structural change (a deformation of the lattice structure), or a combination of the two.

## **6. Conclusion**

Our data show that a Drude model cannot be used to analyze high-intensity femtosecond pump-probe reflectivity measurements. At pump fluences greater than 0.5 kJ/m<sup>2</sup>, the excitation causes major changes in the electronic band structure. Because of these changes, the behavior of  $\chi_{\text{interband}}(\omega)$  rather than  $\chi_{\text{Drude}}(\omega)$  dominates the response of the dielectric constant to the excitation.

The data indicate that the femtosecond laser pulse excitation induces a drop in the average bonding-antibonding splitting. This drop manifests itself in a decrease in frequency of the main absorption resonances in the GaAs dielectric function. The rate and extent of the drop in average splitting increases with pump fluence. At fluences between 0.8 and 1.0 kJ/m<sup>2</sup>, the average splitting drops *reversibly* by more than a factor of two. If the average splitting decreases enough, the minimum in the conduction band drops below the maximum in the valence band, resulting in metallic behavior. Electronic screening may account for the initial changes in the dielectric function, consisting primarily of a broadening of the absorption peaks, while the subsequent shift in the resonance features most likely results from lattice deformation caused by destabilization of the covalent bonds.

## **Acknowledgments**

We appreciate many useful discussions with Professors N. Bloembergen, H. Ehrenreich, and E. Kaxiras. E. N. G. gratefully acknowledges a Fannie and John Hertz Fellowship. This work was supported by contract numbers ONR N00014-89-J-1023 and NSF DMR 89-20490.

## References

- <sup>1</sup>C. V. Shank, R. Yen, and C. Hirlimann, Phys. Rev. Lett. **51**, 900 (1983).
- <sup>2</sup>M. C. Downer, R. L. Fork, and C. V. Shank, J. Opt. Soc. Am. B **2**, 595 (1985).
- <sup>3</sup>H. W. K. Tom, G. D. Aumiller, and C. H. Brito-Cruz, Phys. Rev. Lett. **60**, 1438 (1988).
- <sup>4</sup>S. V. Govorkov, I. L. Shumay, W. Rudolph, and T. Schroeder, Opt. Lett. **16**, 1013 (1991).
- <sup>5</sup>K. Sokolowski-Tinten, H. Schulz, J. Bialkowski, and D. von der Linde, Appl Phys A **53**, 227 (1991).
- <sup>6</sup>P. N. Saeta, J. Wang, Y. Siegal, N. Bloembergen, and E. Mazur, Phys. Rev. Lett. **67**, 1023 (1991).
- <sup>7</sup>C. V. Shank, R. Yen, and C. Hirlimann, Phys. Rev. Lett. **50**, 454 (1983).
- <sup>8</sup>H. Kurz, and N. Bloembergen, in *Energy Beam-Solid Interactions and Transient Thermal Processing*, edited by D. K. Biegelsen, G. A. Rozgonyi and C. V. Shank (Materials Research Society, Pittsburgh, 1985).
- <sup>9</sup>Y. Siegal, E. N. Glezer, and E. Mazur, Phys. Rev. B **49**, 16403 (1994).
- <sup>10</sup>Y. Siegal, E. N. Glezer, and E. Mazur, in *Femtosecond Chemistry*, edited by J. Manz and L. Wöste (Verlag Chemie, Berlin, 1994).
- <sup>11</sup>Y. Siegal, E. N. Glezer, L. Huang, and E. Mazur, in *Ultrafast Phenomena in Semiconductors*, edited by D. K. Ferry and H. M. van Driel (SPIE, Bellingham, Washington, 1994).
- <sup>12</sup>H. R. Choo, X. F. Hu, and M. C. Downer, Appl. Phys. Lett. **63**, 1507 (1993).
- <sup>13</sup>X. Y. Wang, and M. C. Downer, Opt. Lett. **17**, 1450 (1992).
- <sup>14</sup>M. C. Downer, H. Ahn, D. H. Reitze, and X. Y. Wang, Int. J. Thermophysics **14**, 361 (1993).
- <sup>15</sup>Ahn, X. Y. Wang, and M. C. Downer, in *Short Wavelength V: Physics with Intense Laser Pulses*, edited by M. D. Perry and P. B. Corkum (1993).

- <sup>16</sup>M. L. Cohen, and J. R. Chelikowsky, *Electronic Structure and Optical Properties of Semiconductors*; (Springer-Verlag, Berlin, 1988).
- <sup>17</sup>W. A. Harrison, *Electronic Structure and the Properties of Solids: The Physics of the Chemical Bond*; (Dover, New York, 1989).
- <sup>18</sup>H. R. Philipp, and H. Ehrenreich, Phys. Rev. **129**, 1550 (1963).
- <sup>19</sup>E. D. Palik, in *Handbook of Optical Constants of Solids*, edited by E. D. Palik (Academic Press, Inc., New York, 1985).
- <sup>20</sup>D. E. Aspnes, G. P. Schwartz, G. J. Gualtieri, A. A. Studna, and B. Schwartz, J. Electrochem. Soc. **128**, 590 (1981).
- <sup>21</sup>J. D. Jackson, *Classical Electrodynamics*; (Wiley, New York, 1975).
- <sup>22</sup>P. M. Fauchet, and K. D. Li, J. Non-Crystalline Solids **97&98**, 1267 (1987).
- <sup>23</sup>H. M. van Driel, Phys Rev B **35**, 8166 (1987).
- <sup>24</sup>N. W. Ashcroft, and N. D. Mermin, *Solid State Physics*; (Saunders College, Philadelphia, 1976).
- <sup>25</sup>J. M. Liu, A. M. Malvezzi, and N. Bloembergen, in *Energy Beam-Solid Interactions and Transient Thermal Processing*, edited by D. K. Biegelsen, G. A. Rozgonyi and C. V. Shank (Materials Research Society, Pittsburgh, 1985).
- <sup>26</sup>A. M. Malvezzi, H. Kurz, and N. Bloembergen, in *Energy Beam-Solid Interactions and Transient Thermal Processing*, edited by D. K. Biegelsen, G. A. Rozgonyi and C. V. Shank (Materials Research Society, Pittsburgh, 1985).
- <sup>27</sup>P. N. Saeta, Ph.D. Thesis, Harvard University, 1991.
- <sup>28</sup>A. Othonos, H. M. van Driel, J. F. Young, and P. J. Kelly, Solid-State Electron. **32**, 1573 (1989).
- <sup>29</sup>A. Othonos, H. M. van Driel, J. F. Young, and P. J. Kelly, Phys. Rev. B **43**, 6682 (1991).
- <sup>30</sup>J. K. Wang, Y. Siegal, C. Z. Lü, and E. Mazur, Opt. Comm. **91**, 77 (1992).
- <sup>31</sup>Y. Siegal, Ph.D. Thesis, Harvard University, 1994.

- <sup>32</sup>The experimental results presented here show major changes in the dielectric constant following excitation by the pump beam. Since the penetration depth is determined by the dielectric constant, its value when the probe beam arrives at the sample depends on the strength of the excitation and on the pump-probe time delay.
- <sup>33</sup>D. L. Greenaway, and G. Harbeke, *Optical Properties and Band Structure of Semiconductors*; (Pergamon Press, Oxford, 1968).
- <sup>34</sup>R. F. Potter, in *Optical Properties of Solids*, edited by S. Nudelman and S. S. Mitra (Plenum Press, New York, 1969).
- <sup>35</sup>Note that if the oscillator strength and the spectral width of an absorption peak are held fixed, then the height of the absorption peak increases as its resonant frequency decreases.
- <sup>36</sup>D. H. Reitze, H. Ahn, and M. C. Downer, *Phys. Rev. B* **45**, 2677 (1992).
- <sup>37</sup>The effect of the Drude contributions is not noticeable in the behavior of  $\text{Re}(\epsilon)$  at 4.4 eV because the Drude term has a  $1/\omega^2$  dependence and is therefore smaller by a factor of 4 at the higher photon energy. Also, the band structure changes at low fluences have a larger effect on  $\text{Re}(\epsilon)$  at 4.4 eV than at 2.2 eV because of the larger slope at 4.4 eV of  $\text{Re}(\epsilon)$  as a function of photon energy.
- <sup>38</sup>E. J. Yoffa, *Phys. Rev. B* **21**, 2415 (1980).
- <sup>39</sup>D. H. Kim, H. Ehrenreich, and E. Runge, *Solid State Commun.* **89**, 119 (1994).
- <sup>40</sup>V. M. Glazov, S. N. Chizhevskaya, and N. N. Glagoleva, *Liquid Semiconductors*; (Plenum Press, New York, 1969).
- <sup>41</sup>W. Jank, and J. Hafner, *J. Non-Crystalline Solids* **114**, 16 (1989).
- <sup>42</sup>S. Froyen, and M. L. Cohen, *Phys. Rev. B* **28**, 3258 (1983).
- <sup>43</sup>J. A. Van Vechten, R. Tsu, and F. W. Saris, *Phys. Lett.* **74A**, 422 (1979).
- <sup>44</sup>R. Biswas, and V. Ambegoakar, *Phys. Rev. B* **26**, 1980 (1982).
- <sup>45</sup>P. Stampfli, and K. H. Bennemann, *Phys. Rev. B* **42**, 7163 (1990).
- <sup>46</sup>G. E. Jellison, and F. A. Modine, *Phys. Rev. B* **27**, 7466 (1983).

<sup>47</sup>J. P. Walter, R. R. L. Zucca, M. L. Cohen, and Y. R. Shen, Phys. Rev. Lett. **24**, 102 (1970).

## Figure Captions

- Fig. 1** Dielectric function of GaAs. The solid curves show the dielectric function taken from the literature (see Ref. 19).  $E_0$  labels the fundamental absorption edge, corresponding to the minimum band gap while  $E_1$  and  $E_2$  label the two main absorption peaks. The dashed curves show a fit of the single-oscillator dielectric function to the solid curves.
- Fig. 2** A section of the electronic band structure of GaAs (see Ref. 17) showing some of the regions that contribute to the  $E_1$  and  $E_2$  absorption peaks in the dielectric function.
- Fig. 3** Probing geometry and incident angles for the 2.2-eV and 4.4-eV measurements. All beams are  $p$ -polarized.
- Fig. 4** Dielectric constant at 2.2 eV vs. pump-probe time delay for four different pump fluences. ●:  $\text{Re}(\epsilon)$ , ○:  $\text{Im}(\epsilon)$ .
- Fig. 5** Dielectric constant at 2.2 eV vs. pump fluence for four different pump-probe time delays. ●:  $\text{Re}(\epsilon)$ , ○:  $\text{Im}(\epsilon)$ .
- Fig. 6** Pump-probe time delays at which  $\text{Re}(\epsilon) = 0$  for different pump fluences. The solid curve is drawn to guide the eye, and the dashed line corresponds to the damage threshold of  $1.0 \text{ kJ/m}^2$ . At fluences above this value, the induced changes in the material are irreversible while at fluences below the damage threshold, the induced changes are reversible.

**Fig. 7** Schematic representation of the bandgap collapse. The pump pulse leads to a drop in the average bonding-antibonding splitting from its initial value of about 4.75 eV to below the probe photon energy of 2.2 eV. If the minimum in the conduction band drops below the maximum in the valence band, the material takes on metallic properties. The drop in average bonding-antibonding splitting appears as a shift in the main absorption peak in the dielectric function, as illustrated in the figure. cb: conduction band; vb: valence band;  $\Delta E_{b-a}$ : average bonding-antibonding splitting;  $\hbar\omega_{\text{probe}}$ : probe photon energy.

**Fig. 8** Dielectric constant at 4.4 eV vs. pump-probe time delay for four different pump fluences. ●:  $\text{Re}(\epsilon)$ , ○:  $\text{Im}(\epsilon)$ .

**Fig. 9** Dielectric constant at 4.4 eV vs. pump fluence for four different pump-probe time delays. ●:  $\text{Re}(\epsilon)$ , ○:  $\text{Im}(\epsilon)$ .

**Fig. 10** Pump-probe time delays at which  $\text{Re}(\epsilon) = 0$  for different pump fluences at both 2.2 eV and 4.4 eV. The curves are drawn to guide the eye. ●: 4.4 eV, ○: 2.2 eV.

**Fig. 11** Comparison of observed behavior with expected behavior at a time delay of 2 ps. (a) Top graph shows experimentally determined values for the 2.2-eV dielectric constant. ●:  $\text{Re}(\epsilon)$ , ○:  $\text{Im}(\epsilon)$ . (b) Top graph shows computer generated, Drude-like dielectric constant chosen to reproduce the measured reflectivity. ———:  $\text{Re}(\epsilon)$ , - - - -:  $\text{Im}(\epsilon)$ . In both (a) and (b), the bottom graph shows the  $p$ -polarized reflectivity at 45° incident angle. ———: reflectivity calculated from dielectric constant, □: measured reflectivity.

**Fig. 12** Time evolution of double-oscillator model dielectric function vs. photon energy for excitation with a pump fluence of  $1.0 \text{ kJ/m}^2$ . Thicker curves in the top graphs show the dielectric function of GaAs from the literature (see Ref. 19). The data points are experimental values at 2.2 and 4.4 eV.  $\circ$ :  $\text{Im}(\epsilon)$  at 2.2 eV,  $\square$ :  $\text{Im}(\epsilon)$  at 4.4 eV,  $\bullet$ :  $\text{Re}(\epsilon)$  at 2.2 eV,  $\blacksquare$ :  $\text{Re}(\epsilon)$  at 4.4 eV.

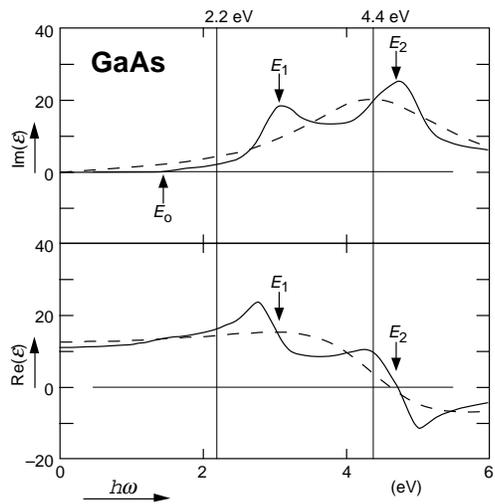


Fig. 1

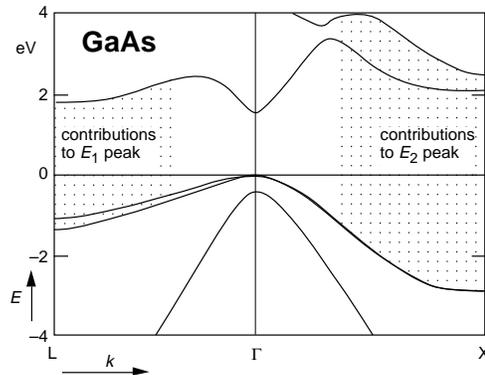


Fig. 2

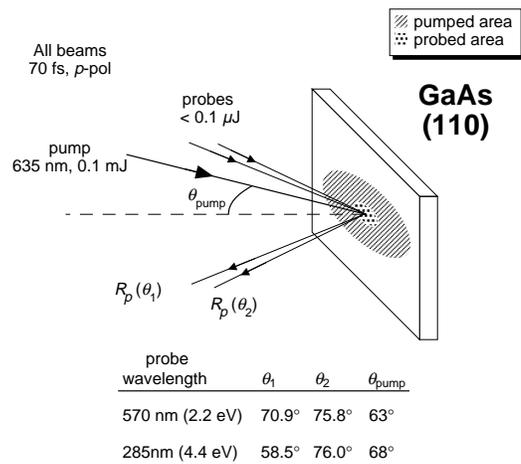


Fig. 3

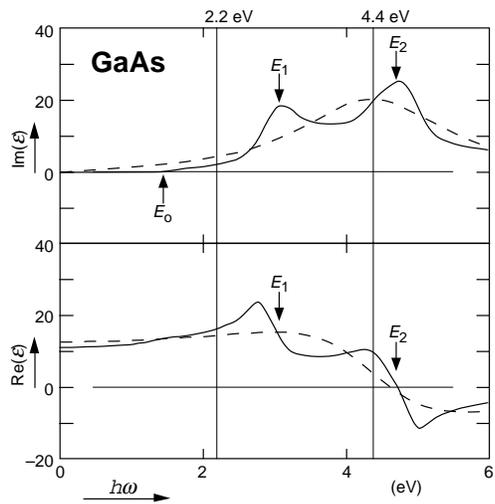


Fig. 1

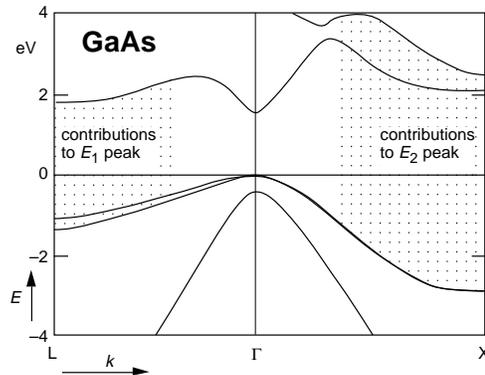


Fig. 2

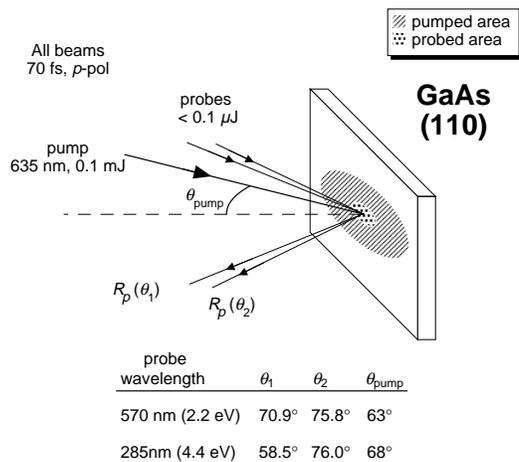


Fig. 3

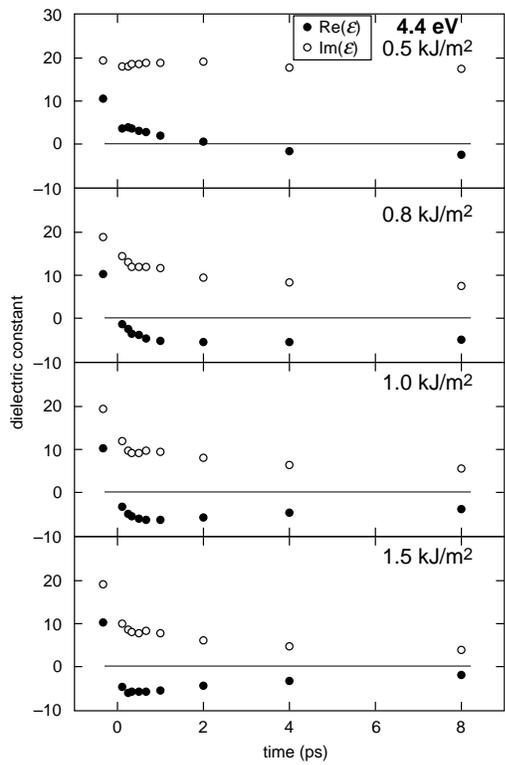


Fig. 8

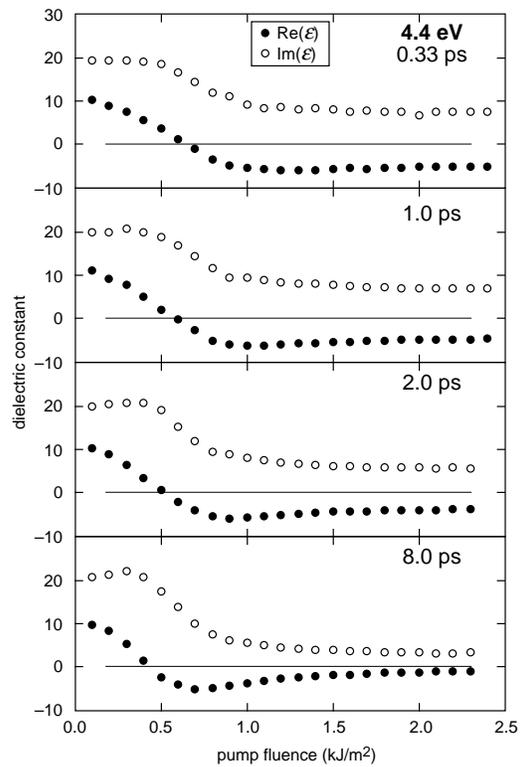


Fig. 9

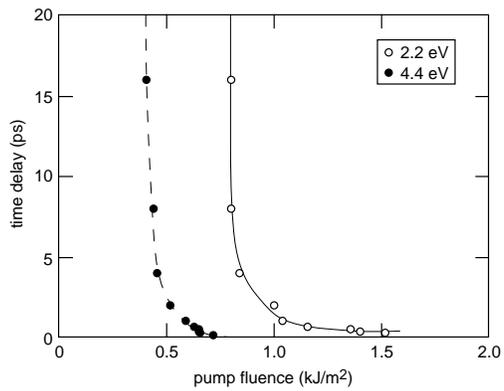


Fig.10

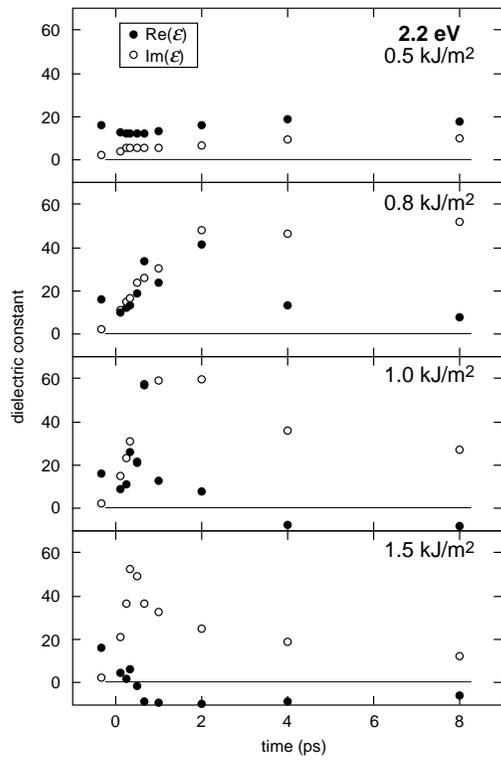


Fig. 4

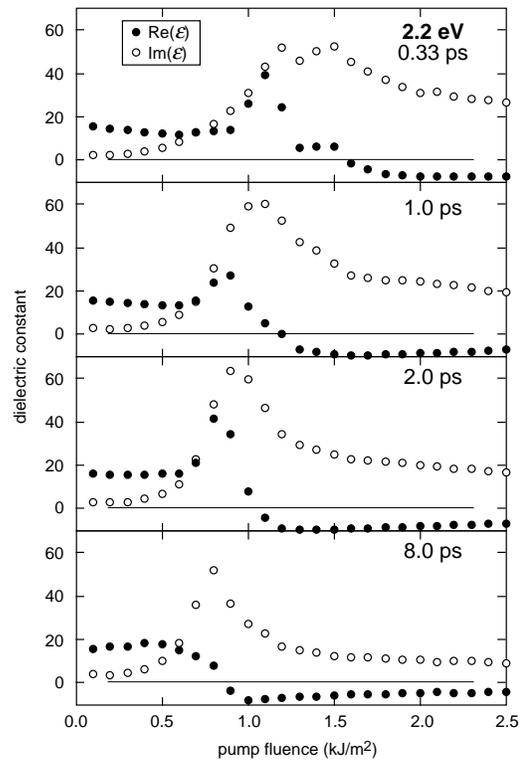


Fig. 5

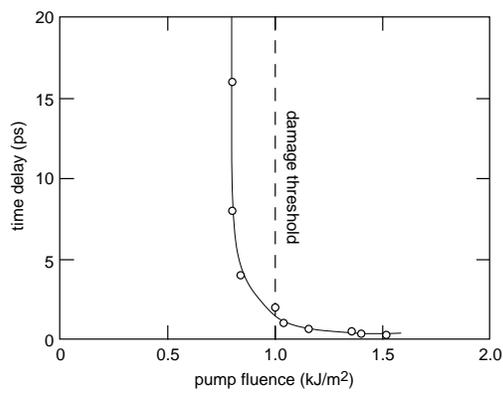


Fig. 6

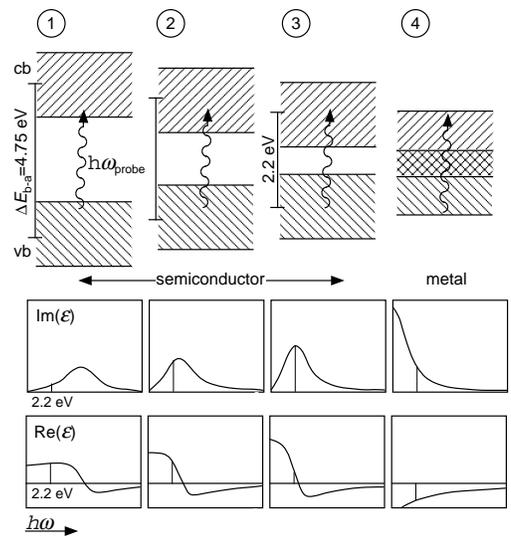


Fig. 7

## **Chapter 4**

### **Behavior of $\chi^{(2)}$ During a Laser-Induced Phase Transition in GaAs**

The following paper appeared in Phys. Rev. B **51**, pp. 9589-9596 (1995).

# **The behavior of $\chi^{(2)}$ during a laser-induced phase transition in GaAs**

E. N. Glezer, Y. Siegal, L. Huang, and E. Mazur

*Division of Applied Sciences and Department of Physics, Harvard University, Cambridge, MA 02138*

We explicitly determine the second-order optical susceptibility of GaAs following intense femtosecond laser-pulse excitation from second-harmonic generation measurements. To separate the dependence of the 4.4-eV second-harmonic signal on the second-order susceptibility from its dependence on the linear dielectric constant, we use experimentally determined values for the dielectric constant of GaAs at 2.2 and 4.4 eV. The results show that the excitation of electrons and the resulting changes in the lattice affect the behavior of the second-order susceptibility. At pump fluences of 0.6 kJ/m<sup>2</sup> and higher, the material loses long-range order on a time scale ranging from 100 femtoseconds to tens of picoseconds, depending on the pump fluence. A recovery of the second-order susceptibility to its initial value at pump fluences between 0.6 and 1.0 kJ/m<sup>2</sup> shows that the loss of long-range order is reversible in this fluence regime.

Intense, femtosecond laser-pulse excitation of semiconductors provides a unique opportunity for observing the dynamics of a phase transition. The semiconductor-metal transition that can result from such excitation<sup>1-11</sup> is particularly interesting because it illustrates the critical role high free-carrier densities can play in modifying the electronic and structural properties of semiconductors. Understanding the complex dynamics involved in laser-induced phase transitions requires an explicit determination of the behavior of intrinsic material properties during these transitions.

Because of its sensitivity to crystal symmetry, second-harmonic generation has been used by a number of researchers to study laser-induced phase transitions in semiconductors.<sup>3-7, 12-14</sup> The sensitivity of second-harmonic generation to the symmetry properties of a nonlinear crystal arises from the dependence of second-harmonic generation on the material's second-order optical susceptibility  $\chi^{(2)}$ , which reflects the symmetry group of the crystal.<sup>15</sup> A change in the material's symmetry properties, such as may occur in a phase transition, affects  $\chi^{(2)}$  and results in a change in the detected second-harmonic signal. However, the detected second-harmonic signal depends on more material properties than just  $\chi^{(2)}$ . In particular, it depends also on the values of the linear optical susceptibility  $\chi^{(1)}$  (or, equivalently, the linear dielectric constant  $\epsilon$ ) at both the fundamental frequency  $\omega$  and the second-harmonic frequency  $2\omega$  of the probe beam used for second-harmonic generation.<sup>15</sup> Thus, to extract the behavior of  $\chi^{(2)}$  from second-harmonic generation measurements, one must first know the behavior of  $\epsilon(\omega)$  and  $\epsilon(2\omega)$ .

We combined experimental measurements of  $\epsilon(\omega)$  and  $\epsilon(2\omega)$  following intense femtosecond laser-pulse excitation of GaAs<sup>8-11</sup> with second-harmonic generation measurements under identical excitation conditions to unambiguously determine the response of  $\chi^{(2)}$  to this type of excitation. This experiment is unique because the response of  $\chi^{(2)}$  to femtosecond laser-pulse excitation is extracted from second-harmonic generation measurements by explicitly taking into account the experimentally

determined response of  $\chi^{(1)}$  to the same excitation rather than by assuming that the effect of  $\chi^{(1)}$  on the second-harmonic generation measurements is small. In fact, the experimentally-determined changes in  $\chi^{(1)}$  are much larger than expected,<sup>8-11</sup> and our results show that these changes have a significant effect on the detected second-harmonic signal, contrary to earlier assumptions.<sup>3, 5, 7</sup> This effect masks the behavior of  $\chi^{(2)}$  at fluences below 0.6 kJ/m<sup>2</sup>, as will be discussed in Section 3.2. Expressly incorporating the changes in the dielectric constant in this experiment has uncovered previously unobserved behavior in this fluence range marked by recovery of  $\chi^{(2)}$  to its initial value on a picosecond time scale.

## 1. $\chi^{(1)}$ - and $\chi^{(2)}$ -Dependence of Second-Harmonic Signal

While second-harmonic generation requires a nonzero  $\chi^{(2)}$  in the dipole approximation, the dielectric constant  $\epsilon$  also affects the strength of this process. The role of  $\chi^{(2)}$  in second-harmonic generation is direct and straightforward: it determines the nonlinear polarization  $\mathbf{P}^{(2)}(2\omega)$ , induced by an electric field  $\mathbf{E}(\omega)$ , which acts as a source for the second-harmonic field  $\mathbf{E}(2\omega)$ . The role of  $\chi^{(1)}$  in second-harmonic generation is less direct but still very important. Changes in the dielectric constant can affect both  $\mathbf{P}^{(2)}(2\omega)$  itself as well as the amount of second-harmonic radiation produced by a given  $\mathbf{P}^{(2)}(2\omega)$ .

The dielectric constant influences  $\mathbf{P}^{(2)}(2\omega)$  by affecting the orientation and magnitude of  $\mathbf{E}(\omega)$  relative to the crystallographic axes. As an example, if  $\mathbf{E}(\omega)$  is oriented in the  $y$ - $z$  plane of a crystal and at an angle  $\theta$  to the  $y$ -axis, then the nonlinear polarization produced through  $\chi_{xyz}^{(2)}$ , the  $xyz$  element of the  $\chi^{(2)}$  tensor, is given by

$$P_x^{(2)}(2\omega) = \chi_{xyz}^{(2)} E(\omega) \cos\theta E(\omega) \sin\theta = \chi_{xyz}^{(2)} E^2(\omega) \sin(2\theta)/2. \quad (1)$$

If the beam at frequency  $\omega$  is initially incident on the crystal in the  $y$ - $z$  plane at some fixed angle  $\theta_i$  to the  $y$ -axis, then, through Snell's Law,  $\epsilon(\omega)$  determines the angle  $\theta$  inside the material that appears in Eq. (1), as shown in Fig. 1. Thus, changes in the dielectric constant

affect  $\mathbf{P}^{(2)}(2\omega)$  through the dependence of  $\mathbf{P}^{(2)}(2\omega)$  on  $\theta$ . In addition, through the Fresnel formulas for reflection and refraction,  $\epsilon(\omega)$  also determines the magnitude of the field in the material for a given incident field. Note that the field magnitude  $E(\omega)$  appearing in Eq. (1) is the field magnitude inside the material, so changes in reflectivity can have a significant effect on  $\mathbf{P}^{(2)}(2\omega)$ .

Besides its effects on the strength of  $\mathbf{P}^{(2)}(2\omega)$  itself, the dielectric constant also affects the amount of second-harmonic radiation generated from a given  $\mathbf{P}^{(2)}(2\omega)$  in a number of ways. Firstly, the index of refraction, given by  $n(\omega) = \text{Re}[\epsilon(\omega)]^{-1/2}$ , determines the phase velocity for light at frequency  $\omega$ . The efficiency of second-harmonic generation depends on  $n(\omega) - n(2\omega)$ , which determines the phase mismatch between the induced nonlinear polarization and the resulting second-harmonic field: the larger the phase mismatch, the less efficient is second-harmonic generation.<sup>15</sup> The phase mismatch is important because it determines the length scale over which second-harmonic radiation generated by the propagating fundamental (frequency  $\omega$ ) beam adds with the proper phase to the propagating second-harmonic beam generated earlier along the beam path. Another way in which  $\epsilon(\omega)$  and  $\epsilon(2\omega)$  affect second-harmonic generation is by determining the absorption depth at  $\omega$  and at  $2\omega$ . The absorption depth can affect second-harmonic generation by limiting the interaction length over which this process takes place.<sup>15</sup> If the energy depleted from the fundamental beam by second-harmonic generation is negligible, as it is in our experiment, then the total amount of second-harmonic generation produced is proportional to the square of the interaction length. So if the absorption depth at one or both of the frequencies becomes smaller than the interaction length, production of second-harmonic radiation will decrease accordingly. Finally, by modifying the reflection of the second-harmonic radiation at the surface of a material, changes in  $\epsilon(2\omega)$  affect the amount of second-harmonic radiation that gets out of the material.

Because GaAs absorbs in the visible and the ultraviolet, we carried out our second-harmonic generation measurements in a reflection geometry, as illustrated in Fig. 2. One can derive an expression for the intensity of the reflected second-harmonic radiation as a function of  $\varepsilon(\omega)$ ,  $\varepsilon(2\omega)$ ,  $\chi^{(2)}$ , and the incident field magnitude  $E_i(\omega)$  and angle  $\theta_i$  of the fundamental beam. To do this, one must first solve the Maxwell equations for the electric field amplitude inside the material, including the effects of the nonlinear polarization. Then, the amplitude of the reflected second-harmonic field is determined by imposing the proper boundary conditions at the surface and using a nonlinear generalization of Snell's Law.<sup>15, 16</sup> For GaAs in the probing geometry illustrated by Fig. 2,<sup>17</sup> the reflected second-harmonic field amplitude is given by<sup>15</sup>

$$E_r(2\omega) = -4\pi P^{(2)}(2\omega) \left\{ \frac{\sqrt{\varepsilon(\omega) - \sin^2 \theta_i} - \sqrt{\varepsilon(2\omega) - \sin^2 \theta_i}}{[\varepsilon(\omega) - \varepsilon(2\omega)] \left[ \sqrt{\varepsilon(2\omega) - \sin^2 \theta_i} - \cos \theta_i \right]} \right\}, \quad (2)$$

where

$$P^{(2)}(2\omega) = 2\chi^{(2)} E_i^2(\omega) \left\{ \frac{4 \sin \theta_i \cos^2 \theta_i \sqrt{\varepsilon(\omega) - \sin^2 \theta_i}}{\left[ \varepsilon(\omega) \cos \theta_i + \sqrt{\varepsilon(\omega) - \sin^2 \theta_i} \right]^2} \right\}. \quad (3)$$

The intensity of the reflected second-harmonic signal, and therefore the detected signal  $S$ , is proportional to  $|E_r(2\omega)|^2$ .

In the experiment presented here, we want to extract the behavior of  $|\chi^{(2)}|^2$  following femtosecond laser-pulse excitation from measurements of the second-harmonic signal. In other words, we want to determine  $|\chi^{(2)}(\phi, t)|^2$ , where  $\phi$  is the fluence of the excitation pulse and  $t$  is the time delay between the excitation pulse and the probe pulse. Note from Eqs. (2) and (3) that we can separate the dependence of the detected signal on the dielectric constant from its dependence on the second-order susceptibility as follows:

$$\frac{S}{|E_i^2(\omega)|^2} = F[\theta_i, \varepsilon(\omega), \varepsilon(2\omega)] \times |\chi^{(2)}|^2, \quad (4)$$

where the function  $F[\theta, \varepsilon(\omega), \varepsilon(2\omega)]$  depends only on the dielectric constant and incident angle and not on the second-order susceptibility. We can define the normalized second-harmonic signal  $S_{norm}(\phi, t) \equiv S(\phi, t)/S(0,0)$ , where  $S(0,0)$  is the second-harmonic signal detected in the absence of any excitation. Similarly, we define  $F_{norm}(\phi, t) \equiv F(\phi, t)/F(0,0)$  and  $\chi_{norm}^{(2)}(\phi, t) \equiv \chi^{(2)}(\phi, t) / \chi^{(2)}(0,0)$ . We can calculate  $F_{norm}(\phi, t)$  with values for  $\varepsilon(\omega, \phi, t)$  and  $\varepsilon(2\omega, \phi, t)$  determined experimentally using a two-angle reflectivity technique under similar laser-excitation conditions as in the measurements of  $S_{norm}(\phi, t)$  presented here.<sup>8-11, 18</sup> Then, with the measured values of  $S_{norm}(\phi, t)$ , we get from Eq. (4) values for  $|\chi_{norm}^{(2)}(\phi, t)|^2$ :

$$|\chi_{norm}^{(2)}(\phi, t)|^2 = \frac{S_{norm}(\phi, t)}{F_{norm}(\phi, t)}. \quad (5)$$

Eq. (5) allows us to extract the desired information on the behavior of the second-order susceptibility from our second-harmonic generation measurements.

## 2. Experimental Setup

The experiment presented in this paper consists of measurements carried out on an insulating (100) GaAs wafer (Cr doped,  $\rho > 7 \times 10^7 \Omega \text{ cm}$ ) in air. In this experiment, 70-fs, 1.9-eV (635 nm) pump pulses excite the sample and 70-fs, 2.2-eV (570 nm) probe pulses monitor second-harmonic generation and reflectivity at various time delays with respect to the excitation. To generate pump and probe pulses at different frequencies, we pass the amplified output of a colliding-pulse modelocked laser through a 20-mm, single-mode, polarization-preserving optical fiber.<sup>19</sup> Self-phase modulation in the fiber broadens the spectrum of the input pulse from 5 nm to 200 nm. By splitting this continuum beam with a broad band beamsplitter, we can independently amplify different spectral regions within this 200-nm band width using two separate amplifier chains.<sup>18</sup> A three-stage amplifier using the dye DCM produces a 300- $\mu\text{J}$  pump beam centered at 635 nm with a 20-nm band width; a two-stage amplifier using the dye Rhodamine 6G produces a 30- $\mu\text{J}$  probe beam centered at 570 nm with a 10-nm band

width. Both amplifiers are pumped by a frequency-doubled, 10-Hz Nd:YAG laser. Separate grating pairs compress each beam temporally to a final pulse width of 70 fs (FWHM).

Figure 2 shows the probing geometry for the second-harmonic generation measurements. The pump beam arrives at an incident angle of  $63^\circ$  with respect to the surface normal while the probe beam comes in with an incident angle of  $45^\circ$ . Both beams are polarized in the plane of incidence and are focused to the same spot on the sample. The (100) sample surface is set to a position with two of the crystallographic axes in the plane of incidence of the beams and the third one perpendicular to it. To monitor a uniformly excited region, we focus the probe beams more tightly than the pump beam: the probed surface area is about 25 times smaller than the  $0.01\text{-mm}^2$  focal area of the pump beam on the sample. Uniform excitation in the probed region is further assured by the small absorption depth of the second-harmonic radiation generated in the sample (about 20 nm) compared to that of the pump beam (270 nm). In the measurements of  $\varepsilon(\omega)$  and  $\varepsilon(2\omega)$  the absorption depth of the probe beams is also significantly shorter than the absorption depth of the pump beam.<sup>11</sup> The pump pulse fluence at each pump-probe time delay is varied over a range from 0 to  $2.0\text{ kJ/m}^2$ . The probe beam fluence does not exceed  $0.1\text{ kJ/m}^2$  so as not to produce any detectable changes in the optical properties to within our experimental resolution. To avoid cumulative damage effects, we translate the sample during data collection so that each data point is obtained at a new spot on the sample.

To extract the behavior of the second-order susceptibility from the second-harmonic generation measurements, we need to combine these measurements of  $S_{norm}(\phi, t)$  with the previous measurements of  $\varepsilon(\omega, \phi, t)$  and  $\varepsilon(2\omega, \phi, t)$ , as described in Section 1. We checked the consistency between the present data set and these previous measurements by simultaneously measuring the linear reflectivity along with the second-harmonic signal in the present experiment. To this end, we used a dichroic beamsplitter

to separate the second-harmonic radiation generated in reflection from the reflected fundamental radiation, measuring the second-harmonic signal using a photomultiplier tube and the fundamental signal with a calibrated phototube. The reflectivity values measured in this way agree with the reflectivity values we calculate for  $45^\circ$  incident angle and polarization in the plane of incidence using the measured dielectric constant from the previous set of data, verifying that the excitation conditions and fluence calibrations for both sets of measurements are indeed identical.

### 3. Experimental Results

#### 3.1 Effect of the Linear Optical Susceptibility on the Second-Harmonic Signal

Figure 3 summarizes the second-harmonic generation measurements and highlights the importance of taking into account the effects of changes in the linear optical susceptibility on the second-harmonic signal. The filled circles in Fig. 3 represent the measured values of the normalized second-harmonic signal  $S_{norm}(\phi, t)$  plotted versus pump fluence  $\phi$  at four different pump-probe time delays  $t$ . The open circles in the figure represent the values of  $F_{norm}(\phi, t)$  calculated using Eqs (2)-(4) and the experimentally determined behavior of the dielectric constant at photon energies of 2.2 and 4.4 eV.<sup>8-10</sup> <sup>18</sup> Note from Eq. (5) that the function  $F_{norm}(\phi, t)$  is identical to  $S_{norm}(\phi, t)$  if  $\chi_{norm}^{(2)}(\phi, t)$  is held constant at its initial value of 1. In other words,  $F_{norm}(\phi, t)$  (the open circles in Fig. 3) shows the changes that result in the second-harmonic signal *solely* from the behavior of  $\chi^{(1)}$  following the excitation; it does not include any effects from changes in  $\chi^{(2)}$ .

As Fig. 3 shows, the changes induced in  $\chi^{(1)}$  by the excitation have a significant impact on the measured second-harmonic signal, contrary to previous assumptions.<sup>3, 5, 7</sup> The locations of the minimum value of  $F_{norm}(\phi, t)$  in each of the plots in Fig. 3 coincide with a large peak in the imaginary part of the dielectric constant at 2.2 eV described in previous papers.<sup>8-11</sup> The order of magnitude decrease in the absorption depth of the fundamental beam accompanying this peak in  $\text{Im}[\epsilon(2.2 \text{ eV}, \phi, t)]$  plays a major role in the large drop in

$F_{norm}(\phi, t)$  since it leads to a corresponding decrease in the interaction length for second-harmonic generation.<sup>20</sup> It is important to point out, however, that while the changes in  $F_{norm}(\phi, t)$  play an important role in the behavior of  $S_{norm}(\phi, t)$ ,  $F_{norm}(\phi, t)$  never drops below the experimental noise and therefore cannot account for the vanishing of  $S_{norm}(\phi, t)$ . Thus,  $\chi^{(2)}$  must go to zero for  $S_{norm}(\phi, t)$  to reach zero.

### 3.2 The Response of $\chi^{(2)}$ to the Excitation

According to Eq. (5), dividing  $S_{norm}(\phi, t)$  (the filled circles in Fig. 3) by  $F_{norm}(\phi, t)$  (the open circles in Fig. 3) yields  $|\chi_{norm}^{(2)}(\phi, t)|^2$ . The resulting values of  $|\chi_{norm}^{(2)}(\phi)|^2$  appear in Fig. 4, shown at the same four pump-probe time delays as in Fig. 3. Figure 5 illustrates the time dependence of  $|\chi_{norm}^{(2)}(t)|^2$  for various excitation strengths. The results exhibit a range of behaviors, depending on the excitation strength. At pump fluences of 0.8 kJ/m<sup>2</sup> and higher,  $\chi^{(2)}$  goes to zero at a rate that increases with pump fluence: at 0.8 kJ/m<sup>2</sup> it takes about 2 ps to reach zero while at 1.5 kJ/m<sup>2</sup> it reaches zero within a pump-probe time delay of 130 fs. In contrast, at pump fluences less than or equal to 0.6 kJ/m<sup>2</sup>,  $\chi^{(2)}$  undergoes a partial decrease, but it does not reach zero. For pump fluences below 0.5 kJ/m<sup>2</sup>,  $\chi^{(2)}$  recovers to its initial value on a time scale of a few picoseconds.

Since  $\chi^{(2)}$  must go to zero for  $S_{norm}(\phi, t)$  to reach zero, the behavior of  $S_{norm}$  at fluences for which it drops to zero is similar to that of  $\chi^{(2)}$ . Thus, for high fluences the conclusions from previous second-harmonic generation experiments, in which changes in the dielectric constant are not explicitly taken into account, agree at least qualitatively with our  $\chi^{(2)}$  data.<sup>3-7</sup> In contrast, the behavior we observe at lower fluences has not been reported before. At fluences for which  $\chi^{(2)}$  does not reach zero, it is particularly important to account for the effects of changes in the dielectric constant on the measured second-harmonic signal. In fact, Fig. 6 shows that the  $S_{norm}$  measurements at a fluence of 0.4 kJ/m<sup>2</sup> are actually misleading if one assumes that the behavior of  $|\chi_{norm}^{(2)}|^2$  is given directly by the second-harmonic signal. Under this assumption, one would

conclude from Fig. 6 that at this fluence  $|\chi_{norm}^{(2)}|^2$  first rises and then drops below its initial value within a few picoseconds. However, our results show that at these fluences  $|\chi_{norm}^{(2)}|^2$  first *decreases* and then *recovers* to its initial value within a few picoseconds.

### 3.3 Long-Time Behavior: Reversible versus Irreversible Changes

Although we focused our attention on the behavior of  $\chi^{(2)}$  during the first 10 ps following the excitation, we also measured both the second-harmonic signal and the linear reflectivity at a time delay of a few seconds, after the material has reached its final state, shown in Fig. 7a, as well as the second-harmonic signal at a pump-probe time delay of 100 ps, shown in Fig. 7b. Figure 7a shows a sharp demarcation at 1.0 kJ/m<sup>2</sup> in the final state of  $\chi^{(2)}$ . For excitation strengths below 1.0 kJ/m<sup>2</sup>, both the second-harmonic signal and the linear reflectivity eventually return to their initial values.<sup>21</sup> However, neither the second-harmonic signal nor the linear reflectivity ever returns to its initial value for pump fluences above 1.0 kJ/m<sup>2</sup>. Thus, the changes induced in the material by the laser-pulse excitation are reversible if the pump fluence is below 1.0 kJ/m<sup>2</sup> but are irreversible if the pump fluence is above 1.0 kJ/m<sup>2</sup>. This conclusion is consistent with an examination of the sample through a microscope. By correlating pump pulse fluence with the size of damage spots on the sample measured through the microscope, we determined a threshold fluence for permanent damage of 1.0 kJ/m<sup>2</sup>, in agreement with the threshold for irreversible change evident in Fig. 7a. Figure 7b shows that for fluences greater than 0.6 kJ/m<sup>2</sup>, once  $S_{norm}$  vanishes, it remains zero for at least 100 ps. Thus, the recovery of  $\chi^{(2)}$  to its initial value at fluences between 0.6 and 1.0 kJ/m<sup>2</sup> occurs on a time scale which is orders of magnitude larger than the recovery times for fluences less than or equal to 0.5 kJ/m<sup>2</sup>.

## 4. Discussion

### 4.1 Three Regimes of Behavior

The results presented in the preceding section suggest three main regimes of behavior for  $\chi^{(2)}$  following laser-pulse excitation. In the low-fluence regime, below 0.5 kJ/m<sup>2</sup>,  $\chi^{(2)}$  exhibits a partial drop but recovers to its initial value within a few picoseconds. At medium fluences, from roughly 0.8 to 1.0 kJ/m<sup>2</sup>,  $\chi^{(2)}$  drops to zero on a time scale between a few hundred femtoseconds and a few picoseconds and remains zero for over 100 ps but eventually also recovers to its initial value. In the high-fluence regime, above 1.0 kJ/m<sup>2</sup>,  $\chi^{(2)}$  drops to zero within a few hundred femtoseconds and never recovers to its initial value. While a clear boundary at 1.0 kJ/m<sup>2</sup> separates the medium- and high-fluence regimes, no clear boundary separates the low- and medium-fluence regimes. Rather, the behavior gradually changes from low-fluence behavior to medium-fluence behavior between 0.5 kJ/m<sup>2</sup> and 0.8 kJ/m<sup>2</sup>.

The recovery time scale in the low-fluence range suggests that the behavior of  $\chi^{(2)}$  in this range is likely dominated by electronic effects. The laser-pulse excitation of electrons from the valence to the conduction band can directly affect  $\chi^{(2)}$  in a number of ways. Firstly, delocalized conduction electrons contribute little to  $\chi^{(2)}$  in GaAs compared to localized valence electrons.<sup>22-24</sup> Thus, we expect the excitation of a high density of electrons from the valence band to the conduction band by the pump pulse ( $> 10^{21}$  cm<sup>-3</sup>) to cause a drop in  $\chi^{(2)}$ . Secondly, the excited free carriers can also reduce the valence band contribution to  $\chi^{(2)}$  by bleaching resonant transitions due to phase-space filling. In addition, electronic screening of the ionic potential by the free carriers will modify the electronic band structure,<sup>25</sup> further affecting the valence band contribution to  $\chi^{(2)}$ . Because the magnitude of these electronic effects should depend on the free-carrier density, we expect  $\chi^{(2)}$  to return to its initial value as the excited free-carrier density relaxes through Auger recombination and diffusion following the excitation. The

picosecond time scale for the observed recovery of  $\chi^{(2)}$  in the low-fluence regime agrees with the time scale for Auger recombination and diffusion at free-carrier densities on the order of  $10^{21} \text{ cm}^{-3}$ .<sup>26, 27</sup> While the low-fluence recovery time is consistent with the expected behavior of the free carrier population, the source of the initial 1-ps delay between the excitation and the maximum change in  $\chi^{(2)}$  in this fluence regime is unclear. One possible cause for this delay is that in addition to its dependence on the free-carrier density,  $\chi^{(2)}$  may be strongly affected by the distribution of the free carriers in the energy bands. In this case, continued heating of the free carrier population by Auger recombination may counteract the effects of carrier density relaxation in the picosecond following the excitation and result in the observed behavior. Alternatively, it is possible that in this fluence regime the electronic excitation drives a slight structural distortion in the lattice, which is reversed when the carrier density decreases through recombination. If it is a temporary distortion of the lattice that is responsible for the partial drop in  $\chi^{(2)}$ , the picosecond delay in the drop would result from the timescale of the atomic motion.

Structural changes in the lattice most likely dominate the behavior of  $\chi^{(2)}$  in the medium- and high-fluence regimes. The time dependence of  $\chi^{(2)}$  in these two regimes is not consistent with electronic time scales. First of all, the recovery time for  $\chi^{(2)}$  is greater than 100 ps in the medium-fluence regime compared with electronic relaxation times of a few picoseconds in the low-fluence regime. Moreover, the drop in  $\chi^{(2)}$  to zero at fluences greater than  $0.6 \text{ kJ/m}^2$  cannot be accounted for by the roughly 10% valence band depopulation achieved by the pump pulse.<sup>24</sup> A structural change in the lattice, however, could lead to a vanishing of  $\chi^{(2)}$  on the observed time scales.<sup>28</sup> Recovery times for reversible structural changes should be comparable to lattice relaxation times, which are much greater than 100 ps.<sup>29</sup>

#### *4.2 Loss of Long-Range Order*

What does the vanishing of  $\chi^{(2)}$  imply about the structural changes in the lattice induced by the pump pulse? In the dipole approximation,  $\chi^{(2)} = 0$  in materials that have a center of inversion.<sup>15</sup> However, the loss of bulk, dipole  $\chi^{(2)}$  does not necessarily mean that the material has taken on a true center of inversion within each unit cell. A loss of long-range order on the scale of the wavelength of light is sufficient to cause such a drop in  $\chi^{(2)}$ . Experiments show that the degree of amorphization induced by low-dosage ion implantation ( $1 \times 10^{12} - 6 \times 10^{15} \text{ cm}^{-2}$  integrated flux of 80-keV  $\text{Te}^+$  and  $\text{S}^+$  ions) leads to a one-to-two order of magnitude drop in  $\chi^{(2)}$ .<sup>24</sup> The extent of ionic motion required for GaAs to lose long-range order is much smaller than that required for GaAs to take on a local center of inversion in each unit cell. Given the time scales involved in the data, a loss of long-range order is the most likely explanation for the observed drop in  $\chi^{(2)}$  to zero in the medium- and high-fluence regimes.

The data in Figs. 4, 5, and 7b then support the conclusion that, for pump fluences greater than  $0.6 \text{ kJ/m}^2$ , the pump pulse induces an instability in the covalent bonding of GaAs that leads to structural change in the lattice. The instability, which results from the excitation of a critical density of electrons from bonding valence states to antibonding conduction states,<sup>28, 30, 31</sup> occurs instantaneously with the generation of free carriers. Because the zincblende structure is no longer stable, the ions start to move away from their ground state positions. As the ions start to move, the material loses its long-range order, and  $\chi^{(2)}$  goes to zero. A stronger excitation results in a greater instability and, therefore, faster ionic motion and a faster drop in  $\chi^{(2)}$ . This interpretation of the data is consistent with the previous set of dielectric constant measurements mentioned in Section 1. The dielectric constant data, which indicate a collapse of the band gap caused by structural changes in the lattice, are fully discussed in other papers.<sup>8-11</sup>

In light of the above interpretation, the material exhibits a particularly interesting response to the excitation in the medium-fluence regime. As discussed in Section 3.3, pump fluences greater than the damage threshold of  $1.0 \text{ kJ/m}^2$  induce irreversible

changes in the material while the changes for pump fluences below this threshold are reversible. Thus, in the medium-fluence range the crystal *reversibly* undergoes sufficiently large structural changes to lose its long-range order. The ability to induce such large *transient* changes in a semiconductor's structural, electronic, and optical properties may be relevant for the development of optoelectronic switching devices.

## 5. Conclusion

We determined the behavior of  $\chi^{(2)}$  following femtosecond laser-pulse excitation from second-harmonic generation measurements by explicitly taking into account the effect of changes in the linear dielectric constant on the second-harmonic signal. We find that accounting for the changes in the linear dielectric constant is particularly important for pump fluences at which  $\chi^{(2)}$  does not vanish, because the changes in the second-harmonic signal at these fluences do not mirror the changes in  $\chi^{(2)}$ . The results show three regimes of behavior: (1) at low fluences, below 0.5 kJ/m<sup>2</sup>,  $\chi^{(2)}$  exhibits a partial drop and a recovery to its initial value within a few picoseconds, (2) at medium fluences, between 0.8 and 1.0 kJ/m<sup>2</sup>,  $\chi^{(2)}$  drops to zero but recovers to its initial value on a time scale greater than 100 ps, and (3) at high fluences, above 1.0 kJ/m<sup>2</sup>,  $\chi^{(2)}$  vanishes and never recovers to its initial value.

The structural changes that follow the laser-pulse excitation in the medium- and high-fluence regimes result from the destabilization of the covalent bonds by the excitation. The resulting ionic motion leads to a loss of long-range order, as indicated by the vanishing of  $\chi^{(2)}$ . This loss of long-range order is observed even at pump fluences below the damage threshold of 1.0 kJ/m<sup>2</sup>, a regime in which the induced changes are reversible.

## Acknowledgments

We appreciate many useful discussions with Professors N. Bloembergen, H. Ehrenreich, and M. Aziz. E. N. G. gratefully acknowledges a Fannie and John Hertz

Fellowship. This work was supported by contract numbers ONR N00014-89-J-1023 and NSF DMR 89-20490.

## References

- <sup>1</sup>C. V. Shank, R. Yen, and C. Hirlimann, Phys. Rev. Lett. **50**, 454 (1983).
- <sup>2</sup>M. C. Downer, R. L. Fork, and C. V. Shank, J. Opt. Soc. Am. B **2**, 595 (1985).
- <sup>3</sup>H. W. K. Tom, G. D. Aumiller, and C. H. Brito-Cruz, Phys. Rev. Lett. **60**, 1438 (1988).
- <sup>4</sup>T. Schröder, W. Rudolph, S. V. Govorkov, and I. L. Shumai, Appl. Phys. A **51**, 1438 (1990).
- <sup>5</sup>S. V. Govorkov, I. L. Shumay, W. Rudolph, and T. Schroeder, Opt. Lett. **16**, 1013 (1991).
- <sup>6</sup>K. Sokolowski-Tinten, H. Schulz, J. Bialkowski, and D. von der Linde, Appl Phys A **53**, 227 (1991).
- <sup>7</sup>P. N. Saeta, J. Wang, Y. Siegal, N. Bloembergen, and E. Mazur, Phys. Rev. Lett. **67**, 1023 (1991).
- <sup>8</sup>Y. Siegal, E. N. Glezer, and E. Mazur, Phys. Rev. B **49**, 16 (1994).
- <sup>9</sup>Y. Siegal, E. N. Glezer, L. Huang, and E. Mazur, in *Ultrafast Phenomena in Semiconductors*, edited by (SPIE, Bellingham, Washington, 1994).
- <sup>10</sup>Y. Siegal, E. N. Glezer, and E. Mazur, in *Femtosecond Chemistry*, edited by J. Manz and L. Wöste (Verlag Chemie, Berlin, 1994).
- <sup>11</sup>E. N. Glezer, Y. Siegal, L. Huang, and E. Mazur, Phys. Rev. B (Accepted for publication).
- <sup>12</sup>C. V. Shank, R. Yen, and C. Hirlimann, Phys. Rev. Lett. **51**, 900 (1983).
- <sup>13</sup>J. M. Liu, A. M. Malvezzi, and N. Bloembergen, in *Energy Beam-Solid Interactions and Transient Thermal Processing*, edited by D. K. Biegelsen, G. A. Rozgonyi and C. V. Shank (Materials Research Society, Pittsburgh, 1985).
- <sup>14</sup>S. A. Akhmanov, V. I. Emel'yanov, N. I. Koroteev, and V. N. Seminogov, Sov Phys Usp **28**, 1084 (1985).
- <sup>15</sup>Y. R. Shen, *The Principles of Nonlinear Optics*, (John Wiley & Sons, New York, 1984).
- <sup>16</sup>N. Bloembergen, and P. S. Pershan, Phys. Rev. **128**, 606 (1962).

- <sup>17</sup>In this case,  $\chi_{xyz}^{(2)}$  is the only element of the  $\chi^{(2)}$  tensor that contributes to second-harmonic generation.
- <sup>18</sup>Y. Siegal, Ph.D. Thesis, Harvard University, 1994.
- <sup>19</sup>J. K. Wang, Y. Siegal, C. Z. Lü, and E. Mazur, *Opt. Comm.* **91**, 77 (1992).
- <sup>20</sup>While a drop in the penetration depth certainly reduces the bulk dipole contribution to second-harmonic generation, its effect on the bulk quadrupole contribution is more complicated. The complication arises from the dependence of the quadrupole contribution to second-harmonic generation on the electric field gradient, which increases as the penetration depth decreases. More generally, Eq. (3), which gives the dependence of the nonlinear polarization on the dielectric constant, applies only to bulk dipole second-harmonic generation. An appropriate expression must be substituted for Eq. (3) to analyze second harmonic generation experiments on centrosymmetric materials such as Si where quadrupole contributions are important (such as in Ref. 3).
- <sup>21</sup>The noise apparent in the  $S_{norm}$  data at a pump-probe time delay of a few seconds (see Fig. 7a) stems from the photomultiplier tube we used to detect the second-harmonic signal. At other time delays, we collected enough data to improve the signal-to-noise ratio through averaging — the noise in the other  $S_{norm}$  data (see Figs. 3 and 7b) is comparable to the size of the data point symbol. However, we have fewer data points at the final-state time delay, limiting the reduction in noise achievable through averaging.
- <sup>22</sup>N. Bloembergen, R. K. Chang, S. S. Jha, and C. H. Lee, *Phys. Rev.* **174**, 813 (1968).
- <sup>23</sup>S. A. Akhmanov, N. I. Koroteev, G. A. Paitan, I. L. Shumay, M. V. Galjautdinov, I. B. Khaibullin, and E. I. Shtyrkov, *Opt Comm* **47**, 202 (1983).
- <sup>24</sup>S. A. Akhmanov, M. F. Galyautdinov, N. I. Koroteev, G. A. Paityan, I. B. Khaibullin, E. I. Shtyrkov, and I. L. Shumai, *Bull. Acad. Sci. USSR, Phys. Ser.* **49**, 86 (1985).
- <sup>25</sup>D. H. Kim, H. Ehrenreich, and E. Runge, *Solid State Commun.* **89**, 119 (1994).

<sup>26</sup>E. J. Yoffa, Phys. Rev. B **21**, 2415 (1980).

<sup>27</sup>H. M. van Driel, Phys Rev B **35**, 8166 (1987).

<sup>28</sup>P. Stampfli, and K. H. Bennemann, Phys. Rev. B **42**, 7163 (1990).

<sup>29</sup>F. Spaepen, in *Ultrafast Phenomena V*, edited by G. R. Fleming and A. E. Siegman (Springer-Verlag, Berlin, 1986).

<sup>30</sup>J. A. Van Vechten, R. Tsu, and F. W. Saris, Phys. Lett. **74A**, 422 (1979).

<sup>31</sup>R. Biswas, and V. Ambegoakar, Phys. Rev. B **26**, 1980 (1982).

## Figure Captions

- Fig. 1** Effect of a change in the dielectric constant of a material on the orientation of the electric field in the material relative to the crystallographic axes.
- Fig. 2** Probing geometry for second-harmonic generation measurements. A  $p$ -polarized probe beam is incident at  $45^\circ$  on a (100) GaAs sample in air. An ultraviolet mirror separates the reflected second-harmonic radiation, which is  $s$ -polarized, from the reflected fundamental radiation, which is still  $p$ -polarized. PMT: photomultiplier tube, PT: phototube, DBS: dichroic beamsplitter.
- Fig. 3** Normalized second-harmonic signal vs. pump fluence for four different pump-probe time delays. ●: measured second-harmonic signal,  $S_{norm}$  ; ○: second-harmonic signal calculated based solely on measured changes in dielectric constant,  $F_{norm}$  .
- Fig. 4** Square of the second-order susceptibility vs. pump fluence for four different pump-probe time delays.
- Fig. 5** Square of the second-order susceptibility vs. pump-probe time delay for various pump fluences. The curves are drawn to guide the eye. ○: 0.2 kJ/m<sup>2</sup>, ●: 0.4 kJ/m<sup>2</sup>, □: 0.6 kJ/m<sup>2</sup>, ■: 0.8 kJ/m<sup>2</sup>, ▲: 1.5 kJ/m<sup>2</sup>.
- Fig. 6** Normalized second-harmonic signal (□) and square of the second-order susceptibility (●) vs. pump-probe time delay at a pump fluence of 0.4 kJ/m<sup>2</sup>. The curves are drawn to guide the eye.

**Fig. 7** (a) Normalized second-harmonic signal (○) and normalized 45° reflectivity (●) at a time delay of a few seconds, after the material has reached its final state. (b) Normalized second-harmonic signal vs. pump fluence at a time delay of 100 ps.

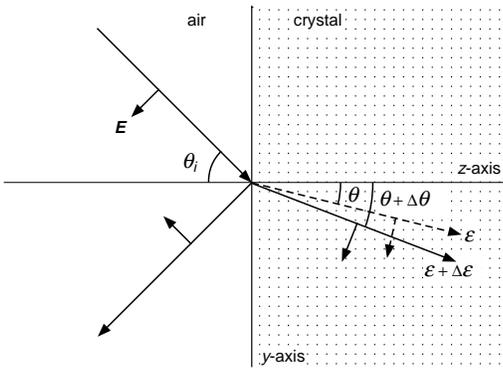


Fig. 1

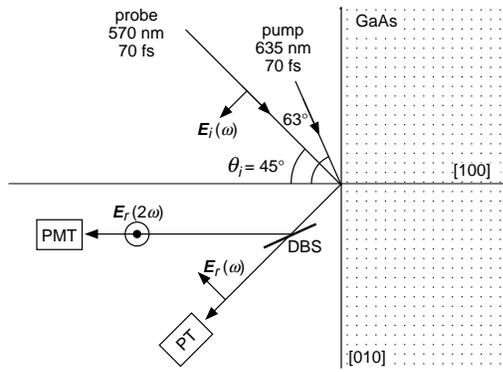


Fig. 2

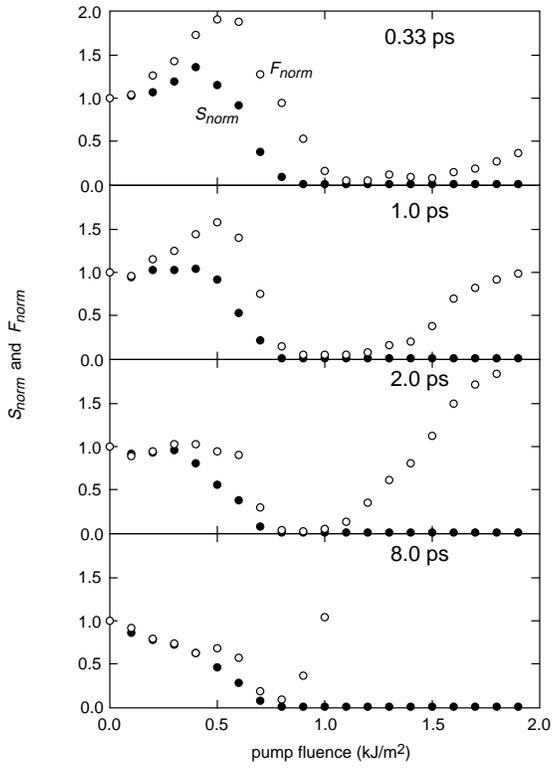


Fig. 3

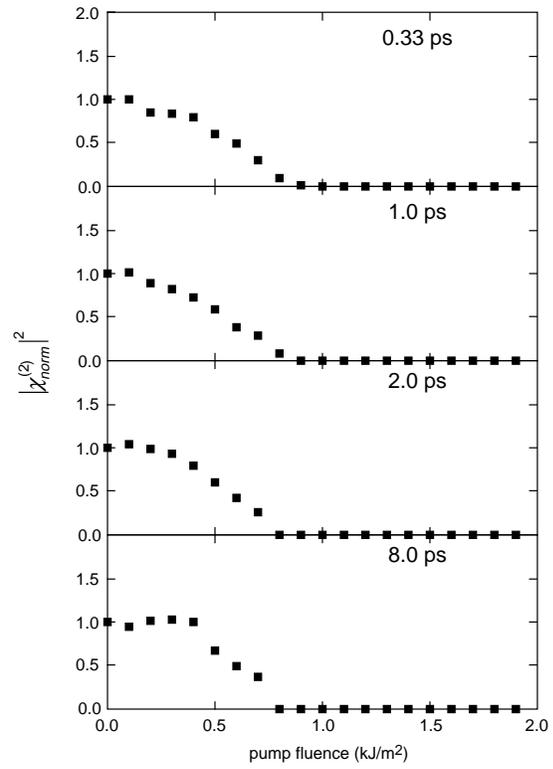


Fig. 4

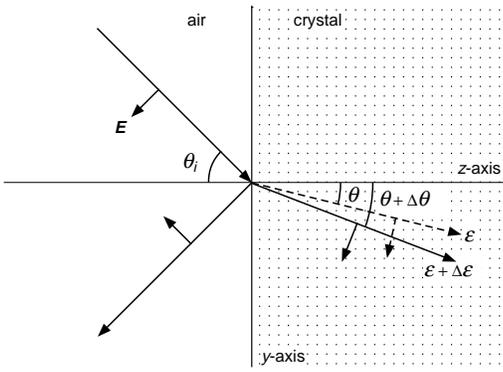


Fig. 1

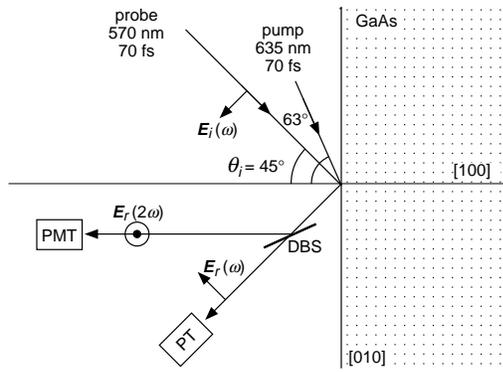


Fig. 2

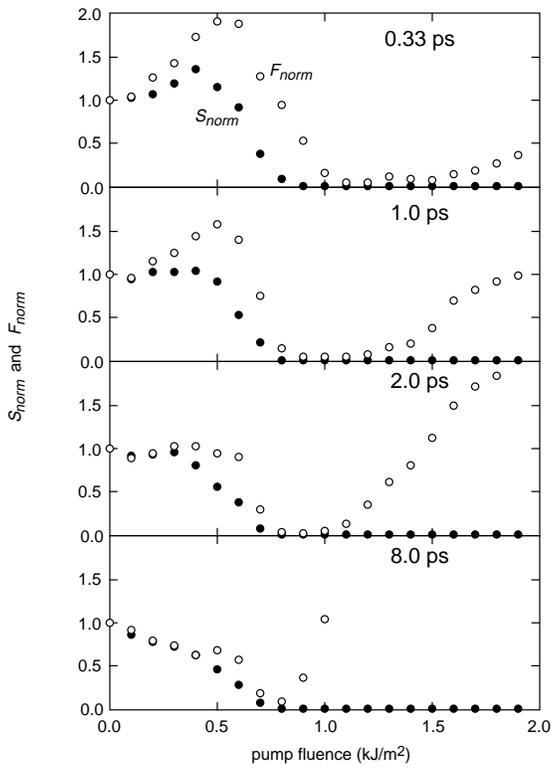


Fig. 3

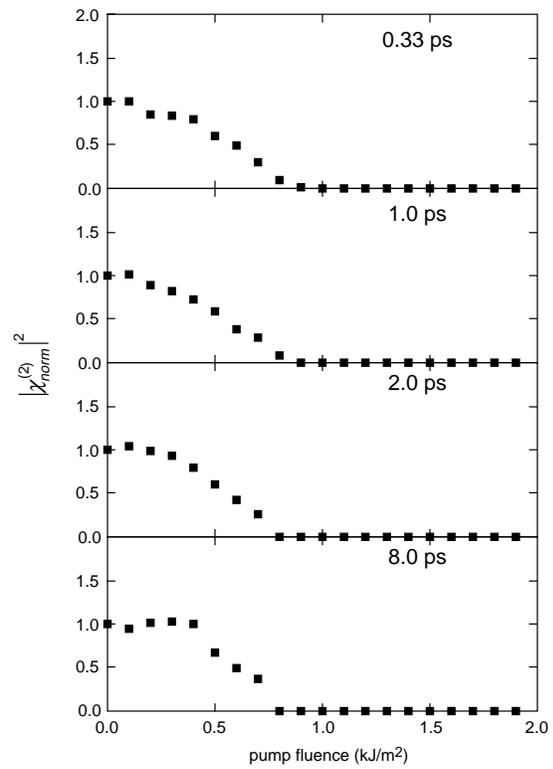


Fig. 4

# Chapter 5

## GaAs under ultrafast excitation: response of the dielectric function

### Introduction

Optical excitation of a semiconductor with photons of energy greater than the bandgap deposits its energy directly into the valence electrons. The electrons then transfer that excess energy to the ions by inducing lattice vibrations and thus heating the crystal. With ultrashort laser pulses the electronic excitation is complete before significant lattice heating occurs, thus creating a situation where the electrons and lattice are far from equilibrium. High intensity pulses can produce such high densities of excited electrons that the resulting charge redistribution destabilizes the bonding in the crystal. In the experiments described here we seek to understand the response of the electrons to the ultrafast excitation as well as the dynamics of the structural response of the lattice.

Time-resolved optical measurements provide an opportunity to observe ultrafast electron and lattice dynamics. Our previous results have shown that measuring fundamental optical properties, such as the dielectric constant<sup>1</sup> and second-order susceptibility<sup>2</sup> is a great improvement over ordinary reflectivity and second-harmonic measurements, and yields results qualitatively different from previous assumptions. We now move beyond single-frequency measurements using a new broadband spectroscopic technique to simultaneously monitor the response of the dielectric function  $\epsilon(\omega)$  of GaAs across the spectral range from 1.5 to 3.5 eV, with 100-fs temporal resolution.

### Background

The first studies of femtosecond laser-excited semiconductors using reflectivity and second-

harmonic generation showed evidence of rapid changes in the material (within a few hundred femtoseconds) following the excitation.<sup>4-9</sup> However, the nature of these material changes cannot be extracted from the results of these experiments. Interpretation of reflectivity and second-harmonic generation results is difficult because these quantities do not directly yield the behavior of intrinsic material properties. In particular, the reflectivity at a specific wavelength, polarization, and incident angle depends on both the real and imaginary parts of the dielectric constant at that wavelength. Furthermore, the measured second-harmonic radiation depends on the dielectric constant at both the fundamental and second-harmonic wavelengths as well as on the second-order susceptibility. The amount of information in the linear reflectivity and second-harmonic generation measurements, therefore, is not sufficient to uniquely determine the behavior of the linear or nonlinear optical material properties.

Our direct measurements of the dielectric constant in GaAs (at 2.2 and 4.4 eV)<sup>1</sup> have shown that contrary to previous assumptions<sup>9-11</sup>, the changes induced in the material by intense an excitation cannot be modeled using a Drude-model formalism. The Drude model is inadequate because it presumes that the changes in the dielectric constant are dominated by the free-carrier contribution to the optical susceptibility. Instead, we found that the optical changes are mainly due to changes in the bandstructure which alter the spectrum of interband transitions and thus the interband contribution to the susceptibility. The response is dominated by a drop in the average bonding-antibonding splitting, which is manifested by a decrease in frequency of the main absorption resonances in the GaAs dielectric function. Both the rate and extent of the drop in average splitting increase with pump fluence. At fluences below the damage threshold the drop is reversible.

Determining the dielectric constant also allowed us to extract the second-order susceptibility,

$\chi^{(2)}$ , from measurements of second-harmonic generation in GaAs<sup>2</sup>. The second-order susceptibility serves as a probe of symmetry, in particular as a measure of the long-range order in the lattice. We found that, at fluences above 60% of the damage threshold, the material loses its long-range order on a time scale that decreases from tens of picoseconds to 100 fs as the fluence is increased. Below the damage threshold,  $\chi^{(2)}$  eventually recovers to its original value, indicating that in the excitation range of 60–100% of the threshold, the loss of long-range order is reversible.

Several theoretical studies have addressed the electronic and structural response of semiconductors to intense ultrafast excitation. By calculating the effect on the bandstructure of a screened ionic potential and electron exchange correlation, it has been shown that the bandgap in GaAs decreases as the carrier density increases, dropping to zero when about 10% of the valence electrons are excited.<sup>12</sup> This result supports our observation that the optical changes are mainly due to the changes in the bandstructure rather than just the free-carrier contribution to the optical susceptibility. The electronic excitation also acts to destabilize the covalent bonding in the crystal and has been shown theoretically to result in rapid softening of the phonon modes to the point of lattice instability.<sup>13,14</sup> The resulting structural dynamics are currently being studied theoretically using molecular dynamics simulations.<sup>15</sup>

The electronic states and crystal structure are intimately connected to each other; for a given crystal structure, the electronic properties can be calculated, yet it is the electrons that dictate the equilibrium crystal structure in the first place. The electronic properties, *i.e.*, the bandstructure and corresponding eigenfunctions, determine the optical properties. An experimental goal is to measure a quantity that will provide a detailed comparison to theoretical models of possible electronic and structural dynamics in the material. The dielectric function,  $\epsilon(\omega)$ , is such a quantity, since it provides a full characterization of the linear optical properties. By measuring its evolution with

100-fs time resolution we obtain the most detailed view thus far of the electron and lattice dynamics during the induced phase transition.

## Experimental setup

We excite a (100) GaAs sample (Cr doped,  $\rho > 7 \times 10^7 \Omega \text{ cm}$ ), with a 70-fs, 1.9-eV (635-nm) laser pulse and probe with a weak broadband continuum pulse at a variable time delay. The probe pulse is focused more tightly than the pump to monitor a uniformly excited region.

To produce the broadband probe we use strong self-phase-modulation of a tightly focused 70-fs pulse in a transparent material. The ideal material should have a large third-order nonlinearity, small group velocity dispersion, and a high damage threshold. After comparing fused silica, optical glass (BK7), sapphire and several ionic crystals (NaCl, KCl,  $\text{CaF}_2$ ), we found that the broadest spectrum was produced in  $\text{CaF}_2$ . The setup for generating the broadband probe is shown in Fig.1 We focus a 20- $\mu\text{J}$ , 570-nm pulse in the center of a 2-mm-thick  $\text{CaF}_2$  crystal using a 50-mm lens. The white-light continuum is produced in a wide cone, which is collimated by an off-axis paraboloid mirror (rather than a lens) to avoid spectral dispersion. The white-light continuum is produced in a wide cone, which is collimated by an off-axis paraboloid mirror (rather than a lens) to avoid spectral dispersion.

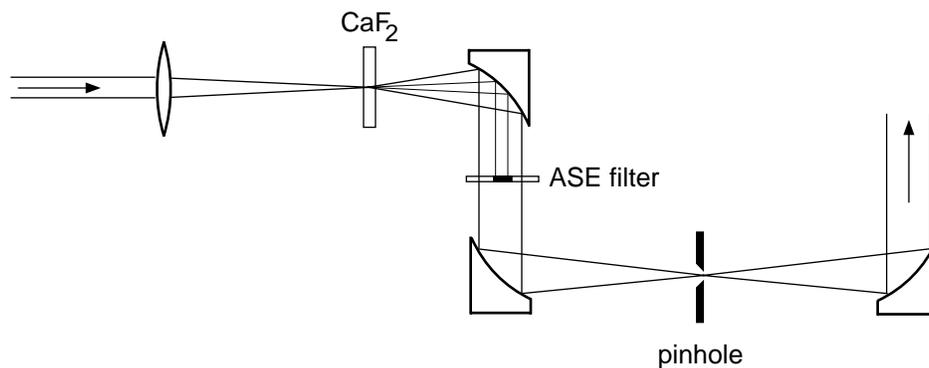


Fig. 1. Broadband probe generation and spatial filtering

There are several difficulties associated with working with the white-light probe beam. Most

importantly, the beam is not diffraction-limited due to the strong self-focusing in the  $\text{CaF}_2$ , and thus cannot be focused well. To circumvent this problem, we use a  $30\text{-}\mu\text{m}$  pinhole to select a small central portion of the beam, and then image the pinhole onto the sample. The small portion that is transmitted is sufficient to serve as a probe. Another problem is the presence of amplified spontaneous emission (ASE) produced in the amplifier. The ASE does not undergo self-phase modulation or self-focusing and thus remains in the central diffraction-limited part of the beam that is almost fully transmitted through the spatial filter. To solve this problem, we use an absorbing disk on a thin ( $180\text{-}\mu\text{m}$ ) fused silica substrate to block only the central part of the white-light beam which contains all of the ASE. The absorbing disk completely removes the ASE from the probe beam.

Figure 2 shows the white-light spectrum averaged over one hundred pulses. The vertical axis is the light intensity per CCD pixel in the spectrometer. Also shown is the spectrum after being filtered by a set of filters designed to flatten the probe spectrum (Schott BG1, BG24A, and a cell with a mixture of several dyes). Because the spectrum changes from pulse to pulse, we record both reference and reflected spectra on every shot.

The measurement setup shown schematically in Fig. 3. Fifteen percent of the probe is split off with a 1-mm-thick single-crystal sapphire disk (cut normal to its optic axis to prevent birefringence from rotating the polarization of the probe relative to the reference). The reference beam is guided around the sample and directed nearly colinearly to, but slightly above, the reflected probe beam. Meanwhile, the probe is tightly focused by an off-axis parabolic mirror onto the sample to an area inside the pumped region. The reflected light is collimated by another off-axis paraboloid, and sent towards the spectrometer nearly colinearly to the reference beam. Reflective optics are used to provide achromatic focusing, as well as to avoiding additional temporal dispersion. Both

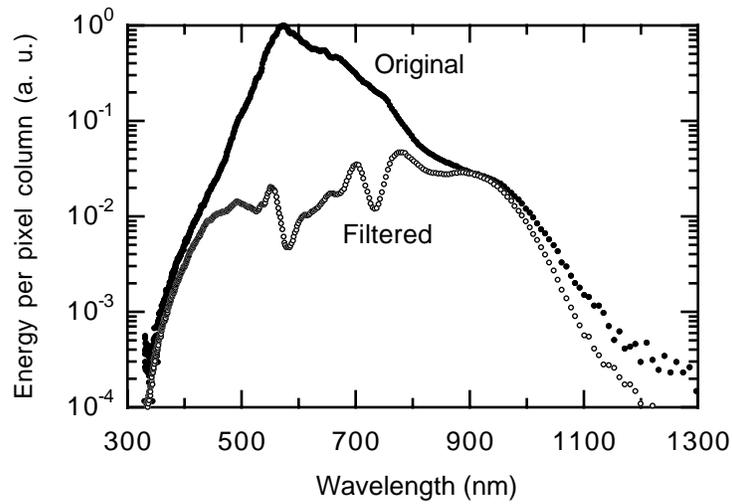


Fig. 2. Original and filtered spectra averaged over 100 pulses. The spectrum of a single pulse has very pronounced spectral variations, as can be seen later in Fig. 6.

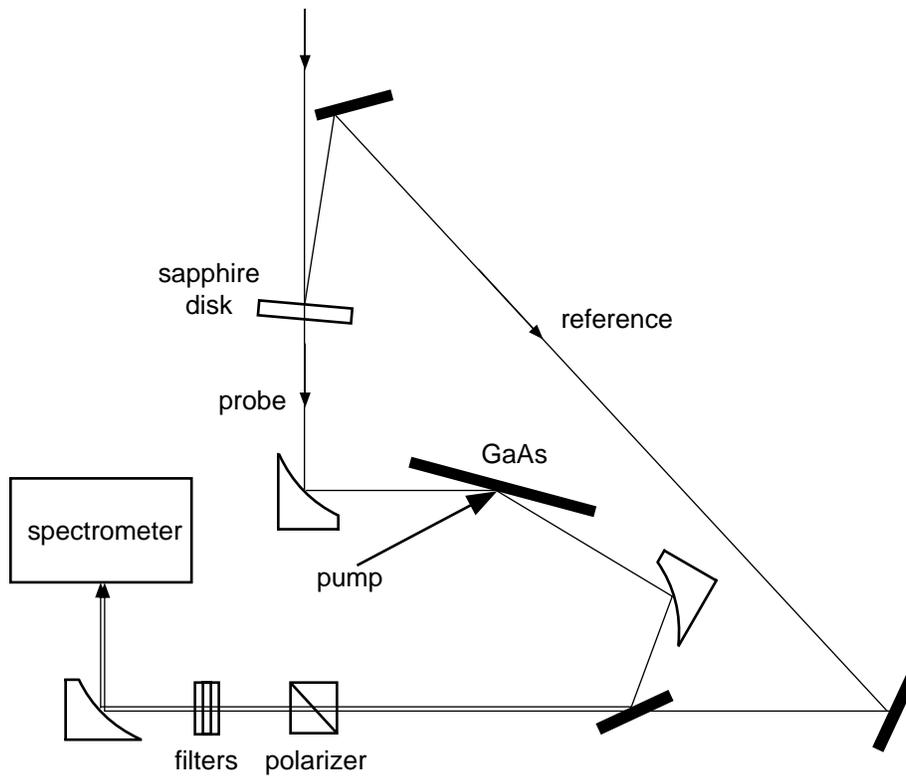


Fig. 3. Experimental setup.

beams pass through a Glan Taylor polarizer and the spectral filters. Finally the two beams are focused (slightly vertically displaced) onto the entrance slit of an imaging spectrometer, and the

two spectra are measured with a CCD camera at the output. Because the white light spans more than a factor of two in wavelength, The spectrometer uses a prism rather than a grating for spectral dispersion to avoid the overlap of higher diffraction orders.

Figure 4 shows the temporal dispersion of the white-light probe. The data points are obtained using sum-frequency generation of the probe with the pump pulse in a thin nonlinear crystal ( $150\text{-}\mu\text{m}$  thick BBO). The time delay is scanned to find the peak of the sum-frequency signal. To choose a particular spectral range band-pass filters are used in the sum-frequency beam rather than in the white-light probe itself since propagation through a filter affects its time of arrival. The angle of the nonlinear crystal is adjusted to assure phase matching at each wavelength point.

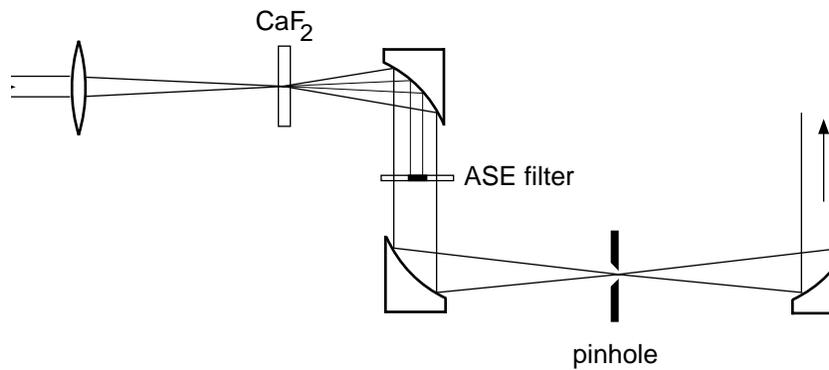


Fig. 4. Temporal dispersion of the spectrum of the white-light probe.

The dispersion shown in Fig. 4 is due to self-phase-modulation and also the propagation through the rest of the  $\text{CaF}_2$  disk, the  $180\text{-}\mu\text{m}$  fused silica substrate for the ASE filter, and the sapphire beam-splitter. Polarization and the spectral filtering are done after the probe is incident on the sample so as not to introduce additional dispersive elements before the time-resolved measurement. When dispersive elements are used after the sample they do not affect the time resolu-

tion of the experiment. Furthermore, even the dispersion acquired by the probe before it reaches the sample does not really limit the time resolution, but simply means that the pump-probe time delay is not the same for all spectral components. The difference in arrival time is taken into account after the time evolution has been measured, by appropriately shifting the ‘time-zero’ for each frequency based on the dispersion in Fig. 4.

## Measuring the dielectric function

To determine both the real and imaginary parts of the dielectric function of bulk GaAs, it is necessary to measure the reflectivity spectra at two different angles of incidence or at two different polarizations, and then numerically invert the Fresnel formulas at the two chosen angles and polarizations. The choice of angle and polarization is critical. This is apparent when one maps a regular grid of dielectric function values ( $\text{Re}(\epsilon)$ ,  $\text{Im}(\epsilon)$ ) to corresponding pairs of reflectivities, as shown in Fig. 5. The set of points in the rectangular grid in Fig. 5a covers the range of complex dielectric space from  $-20$  to  $20$  in  $\text{Re}(\epsilon)$ , and  $0$  to  $40$  in  $\text{Im}(\epsilon)$ . Figures 5b–d show this set of points mapped onto the plane of reflectivity pairs (on a logarithmic scale). The combination of p- and s-polarizations at a  $45^\circ$  angle of incidence (Fig. 5b) collapses the rectangular region onto a nearly one-dimensional curve, making it impossible to invert the mapping to uniquely determine the complex dielectric values from the reflectivity measurements. The combination of p-polarization at  $45^\circ$  with a different angle,  $0^\circ$ , produces a similar mapping (Fig. 5c). However, by going to higher angles of incidence with p-polarization, much more preferable mappings are obtained. In particular, best resolution is achieved when one reflectivity is measured for p-polarized light near the Brewster angle, and the other also p-polarized at a somewhat different, but still large angle. In our experiments, we use the combination of  $75$  and  $58$  degrees, shown in Fig. 5d. In the regions

where the points are densely clustered, a small uncertainty in reflectivity measurements will result in a large uncertainty in  $(\text{Re}(\epsilon), \text{Im}(\epsilon))$ ; conversely, in regions where the reflectivity pairs are widely spaced, the dielectric function can be determined with high precision.

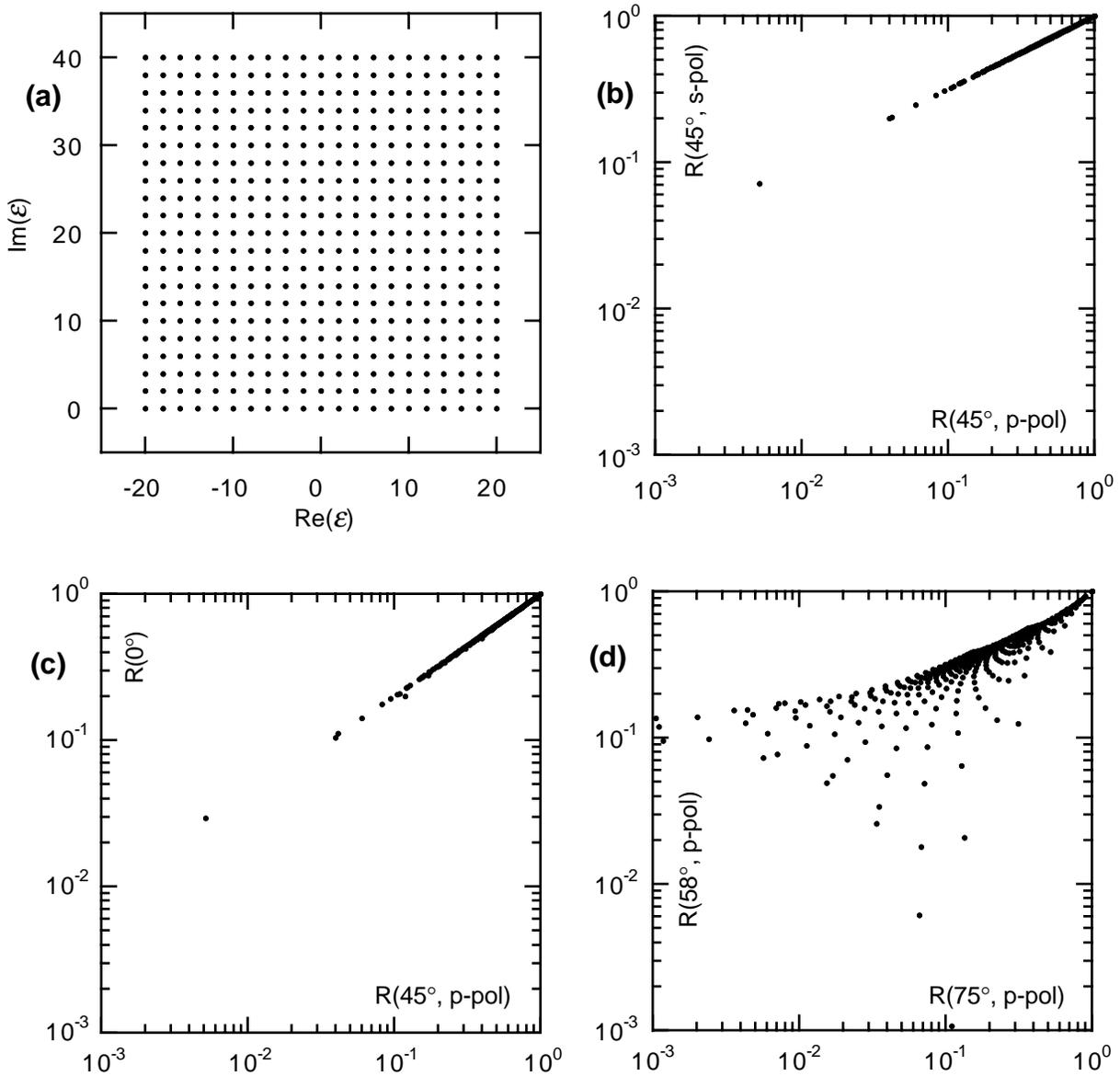


Fig. 5. Mapping of the complex dielectric-function plane to reflectivity pairs.

To determine the reflectivity spectrum at each of the incident angles, we measure a reference spectrum simultaneously with the reflected probe signal. An example of a single-shot measurement of both a reflected and a reference spectrum is shown in Fig. 6. (The spectral fluctuations

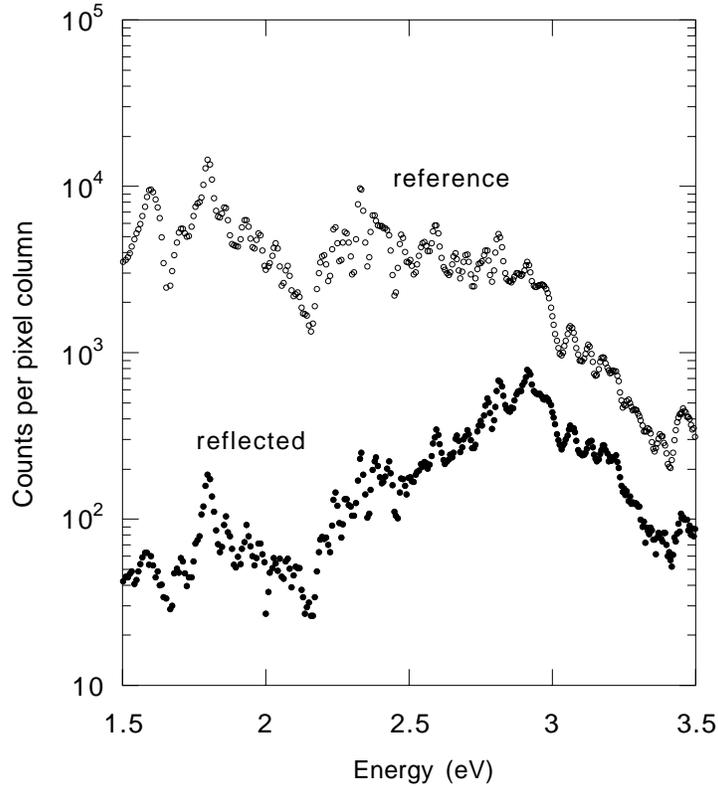


Fig. 6. Reflected and reference spectra from a single pulse.

seen in each of the signals are created in the self-phase-modulation process due to interference between equally spectrally-shifted components at various points on the pulse envelope.<sup>16)</sup> The available dynamic range in the measurements is over  $10^5$ , obtained by adding up the signal from a vertical column of several pixels on the CCD.

The reflectivity is given by the ratio of the reflected signal to the reference, with the beamsplitter ratio and differences in optical losses along the two paths taken into account. To do this, a calibration curve relating the measured ratio to the actual reflectivity was determined (for each of the two angle of incidence setups) by replacing the sample with a mirror of known reflectivity (Newport AL.2). This calibration allowed us to measure absolute reflectivity of the sample.

Because the GaAs surface oxidizes in air, we use a three-phase model (air-oxide-GaAs) in converting the reflectivity data to values of the dielectric constant. The conversion is done using

numerical optimization to determine the  $\epsilon(\omega)$  which best fits the pair of reflectivity values at each frequency. We calibrated the effective thickness of the oxide layer by measuring the reflectivity spectrum of GaAs (in the absence of excitation by a pump pulse) and comparing it with the calculated reflectivity spectrum based on the known values of  $\epsilon(\omega)$  for GaAs<sup>17</sup>, and  $\epsilon(\omega)=4$  for the oxide layer<sup>18</sup>, and using the effective oxide layer thickness as a fit parameter. This was done at a 75° incident angle, since the effect of the oxide layer is most pronounced near the Brewster angle. The oxide layer was found to be  $3.5 \pm 0.5$  nm thick, consistent with observations made using an atomic force microscope (AFM) measurement of an etched step. After the oxide layer thickness was determined and we had verified that the measured dielectric function values agreed with the known GaAs values, we took the unpumped GaAs reflectivity as a known reference and used it for all subsequent calibration in the time-resolved experiments instead of having to use the mirror.

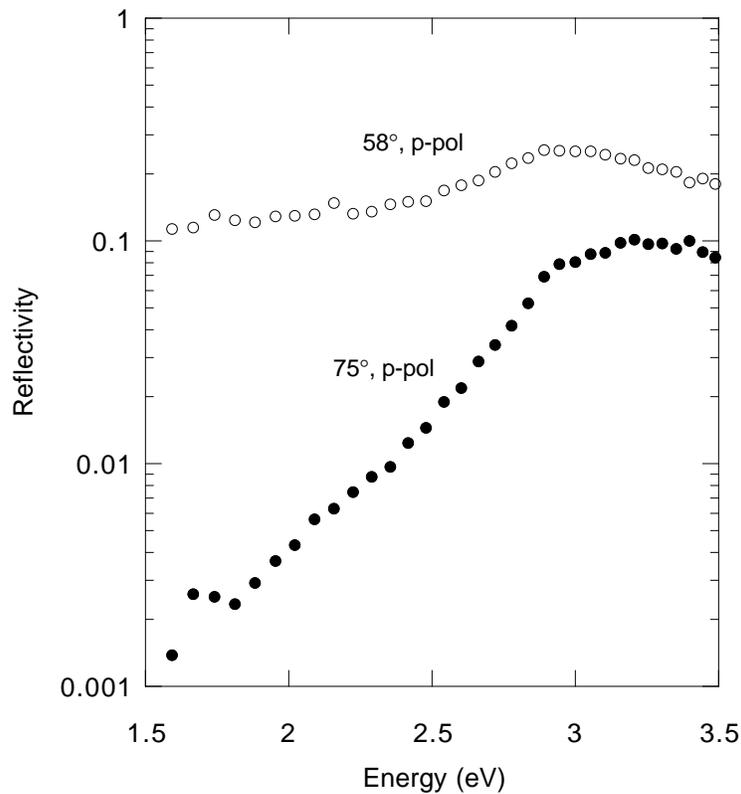


Fig. 7. Reflectivity of GaAs at 58 and 75 degrees, p-polarized.

To verify that the pump pulse is not contributing directly to the measured signal, we performed pumped experiments at ‘negative time delay’, i.e. with the probe pulse preceding the pump pulse. Figure 7 shows the reflectivity spectrum at  $75^\circ$  and  $58^\circ$ , averaged over several shots. Each data point represents an average over 10 pixel columns. A logarithmic scale is used to show the data because the dominant noise source scales with the reflectivity, and only at the very lowest values do background noise sources appear.

The experimentally obtained dielectric function at negative time delay is shown in Fig. 8, together with the known ground-state  $\epsilon(\omega)$ <sup>17</sup>. Uncertainty in the reflectivity measurements was assumed to be  $\pm 5\%$  (of the reflectivity values). The resulting uncertainty in  $\epsilon(\omega)$  depends on the resolution in the particular region of the complex dielectric plane, as demonstrated above in Fig 5. The pump pulse has no noticeable effect on the measurement. However, we note that without the probe, the pump alone does produce a small signal at 1.9 eV due to scattered pump light from the sample and at 1.4 eV due to the induced luminescence of the GaAs sample.

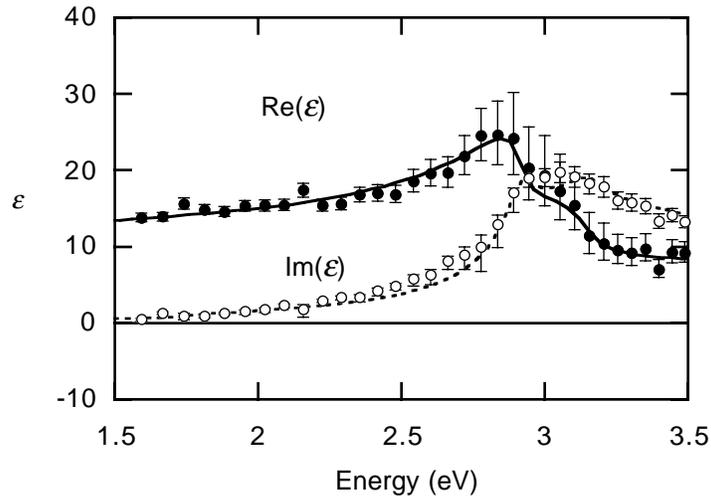


Fig. 8. Dielectric function at negative time delay.

## Results

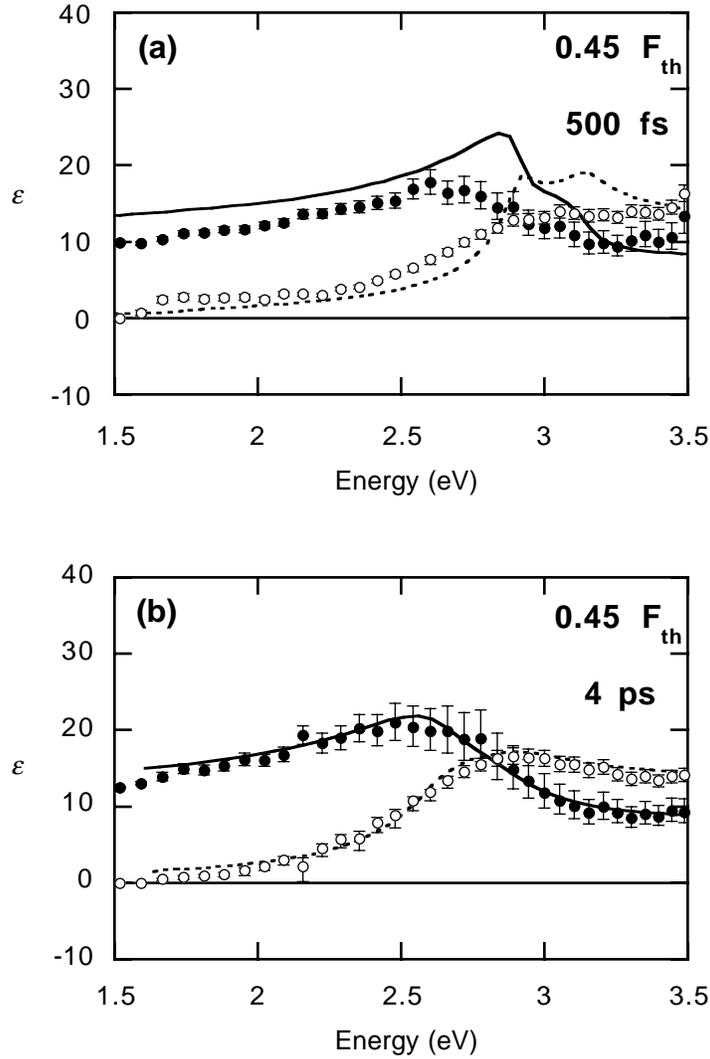


Fig. 9. Dielectric function ( $\bullet$  –  $\text{Re}[\epsilon(\omega)]$ ,  $\circ$  –  $\text{Im}[\epsilon(\omega)]$ ) for a pump fluence of  $0.45 F_{th}$  after 500 fs and 4 ps. The lines drawn on the 500-fs plot show the real (solid line) and imaginary (dashed line) parts of the dielectric function for unperturbed GaAs. The lines drawn on the 4-ps plot show the dielectric function for GaAs at 770 K.<sup>19</sup>

The data in Fig. 9 show  $\epsilon(\omega)$  at 500 fs and 4 ps after excitation by a pulse of 45% of the fluence threshold ( $0.45 F_{th}$ ) for observable permanent change. The threshold is  $1.0 \text{ kJ/m}^2$  ( $100 \text{ mJ/cm}^2$ ). The solid and dashed curves Fig. 9a correspond to the real and imaginary parts of the unperturbed  $\epsilon(\omega)$ . The 500-fs data display a drop in the real part and a broadening in the imaginary part of the

dielectric function which occur directly following the excitation. These initial changes are likely electronic in this pump fluence regime. By 4 ps  $\epsilon(\omega)$  has changed considerably, and agrees closely with ellipsometric measurements on GaAs at elevated temperature<sup>19</sup>, shown at 800 K by the solid and dashed curves in Fig. 9b, suggesting that the lattice has been heated and this process dominates the changes in  $\epsilon(\omega)$ . From 4 picoseconds on, we are able to fit the data in the ‘low-fluence’ regime, below about  $0.5 F_{th}$ , with the elevated temperature dielectric functions which have been measured up to 920 K.<sup>19</sup> The extracted lattice temperature is shown Fig. 10 as a function of the time delay for three different pump fluences, 0.2, 0.35, and  $0.45 F_{th}$ . The lattice temperature rises for about 10 ps following the excitation. As expected, the temperature rises faster and further with increasing fluence. In the first 2-3 picoseconds our data cannot be fitted with a dielectric function of heated GaAs, a finding that we attribute to the direct effect of the excited electrons on the band-structure.

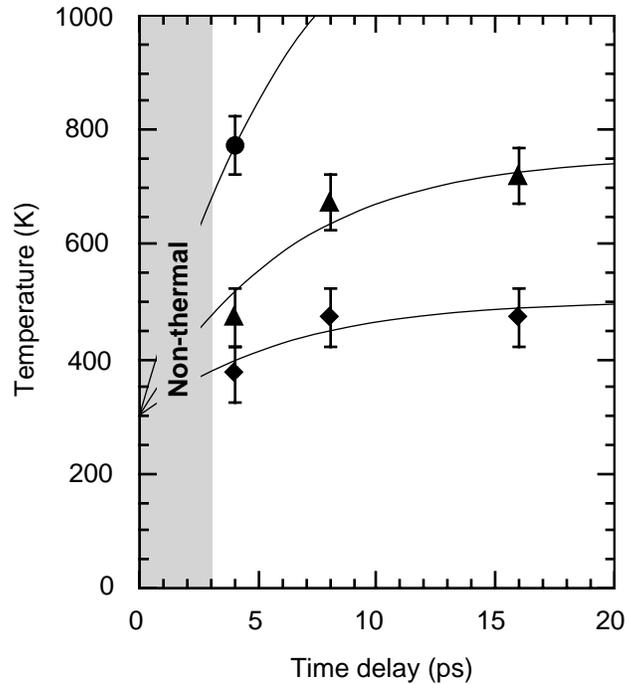


Fig. 10. Lattice heating in the low fluence range.

At a higher fluence of 70% of the threshold, the changes in  $\epsilon(\omega)$  are much more pronounced, as shown in Fig.11 at 500 fs and 4 ps. At 500 fs the imaginary part of  $\epsilon(\omega)$  is very broadly spread across the entire spectral range, and the real part has decreased and flattened (Fig. 11a). By 4 ps, the imaginary part is still broadly spread out and considerable larger, while the real part now

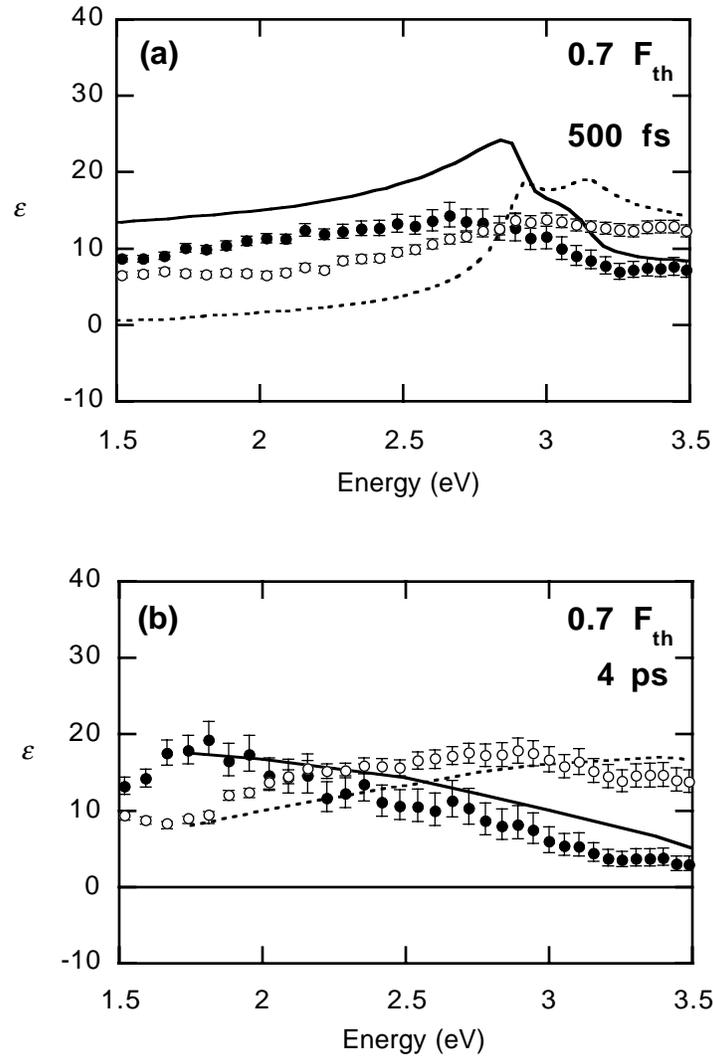


Fig. 11. Dielectric function ( $\bullet$  –  $\text{Re}[\epsilon(\omega)]$ ,  $\circ$  –  $\text{Im}[\epsilon(\omega)]$ ) for a pump fluence of  $0.7 F_{th}$  after 500 fs and 4 ps. The lines drawn on the 500-fs plot show the real (solid line) and imaginary (dashed line) parts of the dielectric function for unpumped GaAs. The lines drawn of the 4-ps plot show the dielectric function for amorphous GaAs.<sup>21</sup>

shows a negative slope in frequency across most of the spectral range. The dielectric function resembles that of amorphous GaAs<sup>21</sup>, shown by the curves, suggesting that the material has become disordered (Fig. 11b). A previous measurement<sup>2</sup> of the second-order optical susceptibility,  $\chi^{(2)}$ , of GaAs shows that at the same excitation fluence  $\chi^{(2)}$  drops to zero within a few picoseconds, indicating the loss of long-range order. Note that this occurs below the threshold for permanent change, so the disordering is reversible.

Lattice heating certainly occurs at  $0.7 F_{th}$ , but it can not account for the large change in the dielectric function at 4 ps. The heating probably contributes to the difference between the data at 4 ps and the dielectric function of amorphous GaAs at room temperature as shown in Fig. 11b. Additionally, although it is disordered, the material is unlikely to be exactly the same as in the amorphous state.

Above the damage threshold,  $\epsilon(\omega)$  changes dramatically. Fig. 12 shows  $\epsilon(\omega)$  at 500 fs and 4 ps after an excitation of 160% of the threshold (note the different scale). Within 500 fs much of the oscillator strength has moved from the large  $E_2$  resonance peak (initially centered at 4.75 eV) down to our spectral range. The real part is negative above 2.7 eV because most of the oscillator strength is at lower frequencies. By 4 ps the real part is negative above 1.8 eV, and possibly even below 1.8 eV. The data below 1.8 eV lie in a region of dielectric space which is not well-resolved in our reflectivity measurements and hence the uncertainty in  $\epsilon(\omega)$  is large. The metal-like dielectric function above 1.8 eV can be fitted with a Drude model, which gives a 12 eV plasma frequency and a 0.2 fs relaxation time. The large plasma frequency indicates that almost all of the original valence electrons are behaving as free carriers. Based on previous measurements at 4.4 eV,<sup>1</sup> under similar excitation,  $\epsilon(4.4 \text{ eV}) = -3 + 4i$ . The absence of any peaks in the imaginary part of  $\epsilon(\omega)$  from 1.8 to 3.5 eV and at 4.4 eV, indicates that the bonding-antibonding splitting has

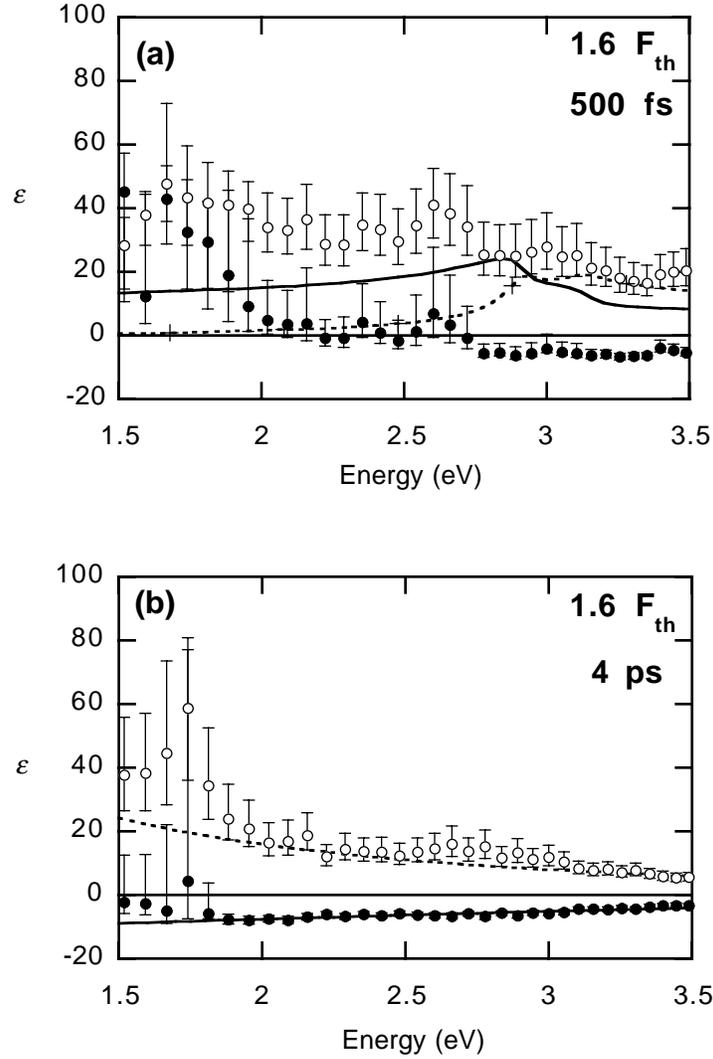


Fig. 12. Dielectric function ( $\bullet$  –  $\text{Re}[\epsilon(\omega)]$ ,  $\circ$  –  $\text{Im}[\epsilon(\omega)]$ ) for a pump fluence of  $1.6 F_{th}$  after 500-fs and 4-ps. The lines drawn on the 500-fs plot show the real (solid line) and imaginary (dashed line) parts of the dielectric function for unpumped GaAs. The lines drawn on the 4-ps plot show a Drude-model dielectric function with a plasma frequency of 12 eV and a relaxation time of 0.2 fs.

vanished. The changes in the dielectric function are much more drastic than those due to disordering and indicate changes in local bonding.

## Discussion

To understand the electron and lattice dynamics of GaAs after the 70-fs excitation, it is useful

to begin with the excitation process. Incident photons promote electrons from the occupied valence band to the empty conduction band, leaving holes behind. A carrier density on the order of  $10^{22} \text{ cm}^{-3}$ , or about 10% of the valence electrons, is created by the pump pulse at the damage threshold via linear and nonlinear absorption. Based on the joint density of states, we estimate that less than 0.01% of the valence electrons can be directly excited by the 1.9-eV pump pulse. Redistribution of the excited electrons within the  $\Gamma$ -valley during the pump pulse makes additional states available, but the phase space in the  $\Gamma$ -valley is limited to about  $10^{21} \text{ cm}^{-3}$ . However, several mechanisms allow a further increase in the carrier density. Scattering of the excited electrons into the L- and X- valleys during the excitation clears out the phase space for further transitions. The large density of states in the L-valley and its small energy difference from the  $\Gamma$ -valley make it a favored destination for the excited electrons. (A larger phase space is available to the holes within their  $\Gamma$ -valley.) Two-photon absorption makes possible transitions at 3.8-eV, where the joint density of states is much larger. Furthermore, sequential linear absorption of multiple photons can raise the carriers to much higher energies; this also clears out the phase space in the  $\Gamma$ -valley for further transitions, and the energetic electrons can create additional electron-hole pairs through impact ionization. Finally, as the excitation proceeds, the bandstructure itself is changing (as described below), making new regions of phase space accessible to direct transitions.

The excited free carriers directly affect the dielectric function in two ways. First, the presence of the free carriers changes the dielectric function by the contribution of free carrier absorption, as described by the Drude model. A decrease in the real part and an increase in the imaginary part are expected<sup>1</sup> due to a free carrier contribution. Although the observed changes at 500-fs time delay in both the  $0.45 F_{\text{th}}$  and  $0.7 F_{\text{th}}$  data (Figs. 9a and 11a) do show a decrease in the real part and an increase in the imaginary part below about 2.5 eV, the dominant cause of these changes is

not the free carrier contribution to the susceptibility. The strength of the changes at  $0.7 F_{\text{th}}$  would correspond to a plasma frequency of about 10 eV, and would indicate that a majority of valence electrons are contributing as free carriers — which is impossible since the total  $\epsilon(\omega)$  is not at all metal-like. Furthermore, the spectral shape of the changes (even below 2.5 eV) cannot be fitted with the Drude model even with both the carrier density and scattering time left as free parameters. Finally, above 2.5 eV the changes in  $\epsilon(\omega)$  do not even resemble the effect of a free-carrier Drude contribution. Instead, the broadening and shift to lower frequency of the absorption features seen in Fig. 9a, and more pronounced in Fig. 11a, are the consequence of the other effect of the free carriers, namely, the screening and bandgap renormalization in which many-body effects decrease the splitting between the valence and conduction bands<sup>12</sup>.

Since the average bonding-antibonding splitting increases with the strength of the ionic potential,<sup>22</sup> electronic screening reduces the average bonding-antibonding splitting. Note that electronic screening, through its modification of the band structure, changes the interband transition contribution to the dielectric constant and is therefore completely different from the direct intraband contribution of the free carriers to the dielectric constant through a Drude term. A recent calculation shows that when 10% of the valence electrons are excited to the conduction band, the direct gap at the X-point in the band structure of GaAs will decrease by roughly 2 eV due to electronic screening and many-body bandgap renormalization while the direct gap at the  $\Gamma$ -point changes only slightly.<sup>12</sup> Such an effect would lead to an immediate broadening and shift of the absorption peaks in the dielectric function as seen in Figs. 9a and 11a. These spectral measurements confirm our previous conclusion<sup>1</sup> that the initial response of the dielectric function to the excitation is dominated by changes in the electronic bandstructure and not by the optical susceptibility of the excited free carriers.

The free carriers relax through phonon emission, Auger recombination, radiative recombination and diffusion. Phonon emission and Auger recombination dominate during the first few ps after the excitation when the carriers are energetic and their density is high. Phonon emission provides the coupling that transfers energy from the electrons to the lattice. Auger recombination does not remove any energy from the electronic system, but rather reduces the carrier density while increasing the energy per carrier. The rapid decrease of carrier density by the Auger process leads to the diminishing role of direct electronic effects.

The carriers relax to the bottom of the bands on the picosecond timescale through phonon emission, causing the lattice becomes hotter. By 4 ps, the  $\epsilon(\omega)$  data matches the dielectric function a thermal state, suggesting that the initial electronic effect of the high density of carriers has largely subsided. Surprisingly, the lattice temperature continues to rise for a total of about 10 ps of continued heating. One possible explanation for this is that by 4 ps the carrier density has been significantly reduced by Auger recombination, but that a large fraction of the energy is still trapped in the electronic system. (The electrons are possibly even hotter than in the originally excited state.) As the carrier distribution cools, the lattice is further heated.

The lattice heating following femtosecond laser excitation is quite different from conventional thermal heating, and so it seems remarkable that our data agrees so well with the equilibrium measurements (Fig.9b). Ordinary heating involves two effects: the increase in vibrational motion of the atoms (the Debye-Waller effect), and the thermal expansion of the lattice. In the case of ultrafast laser-induced heating, lattice expansion cannot occur within several picoseconds, leaving the Debye-Waller effect as the only cause of the changes in  $\epsilon(\omega)$ . However, the most dominant contribution to the temperature dependence of the dielectric function of the GaAs is the Debye-Waller effect,<sup>20</sup> which is not related to volume changes. This explains the good agreements of our

data with measurements at elevated temperatures with conventional heating of GaAs, as seen in Fig. 9b.

In the medium fluence regime, represented by the data at  $0.7 F_{\text{th}}$ , the changes in  $\epsilon(\omega)$  are larger than and spectrally different from what can be extrapolated based on equilibrium measurements of heated GaAs. Although data are only available up to 920 K, the effect of heating is small relative to the changes observed in Fig. 11b. At the same time, it does not appear that the semiconductor is molten. The dielectric function of molten GaAs has not been measured because the Arsenic atoms diffuse out of the material, but based on other similar semiconductors such as Silicon,  $\epsilon(\omega)$  is metal-like in the molten state. The fact that the real part of  $\epsilon(\omega)$  in Fig. 11b is positive up to 3.5 eV indicates that much of the oscillator strength is still above 3.5 eV, in agreement with measurements at 4.4 eV,<sup>1</sup> and that the covalent bonds are still largely intact. These observations lead us to conclude that the measured  $\epsilon(\omega)$  corresponds to a disordered but yet not molten state. The dominant effect of the disordering on  $\epsilon(\omega)$  is the removal of the restriction of  $k$ -vector conservation in optical transitions. The disordering of the lattice results directly from the excitation — no signature of melting is observed prior the loss of crystalline order.

An insightful connection can be made to our previous measurements of  $\chi^{(2)}$ . With excitation below about  $0.5 F_{\text{th}}$ ,  $\chi^{(2)}$  recovers to its original value within 4 ps, whereas above  $0.6 F_{\text{th}}$   $\chi^{(2)}$  drops to zero within several picoseconds (and faster with increasing fluence) and does not recover for over 100 ps. Thus in the low fluence regime  $\chi^{(2)}$  indicates (after several picoseconds of carrier relaxation) that the long-range order in the crystal is preserved, whereas in the medium fluence regime the electronic excitation destabilizes the bonding enough to produce a disordered state. Interestingly, the disordered state does not have the metallic optical properties expected for a molten state. Rather it closely resembles the  $\epsilon(\omega)$  of solid amorphous GaAs. The bonding is still

largely covalent rather than metallic in character. The disordered state is not permanent: both the linear optical properties and second order susceptibility return to the original values on a timescale longer than 100 ps, indicating that in this fluence regime the semiconductor eventually returns to its original crystalline state.

In the high fluence regime the changes in  $\epsilon(\omega)$  are most drastic. To account for the time evolution of the shift in the resonance features of the dielectric constant, we must examine the effect of lattice structural change on the band structure. The electronic band structure is determined by the crystal structure. The semiconducting behavior of Group IV and III-V materials such as GaAs arises from the tetrahedrally-coordinated diamond or zincblende arrangement of the constituent atoms. If this arrangement is disturbed, the electronic properties will change accordingly. In general, deformation of the diamond or zincblende structure leads to a collapse of the band gap and a semiconductor-metal transition.<sup>22-25</sup> Just a 10% change in average bond length is enough to cause a semiconductor-metal transition.<sup>25</sup> Note that an ionic velocity as small as 25 m/s is already sufficient to achieve a 10% change in the GaAs bond length within 1 ps.

Because the covalent bonds of semiconductors like GaAs are stabilized by the valence electrons, excitation of a sufficient number of electrons from bonding valence states to antibonding conduction states can lead directly to lattice instability.<sup>13,14</sup> If the femtosecond pump pulse is intense enough to excite this critical density of electrons, the resulting instability in the lattice will cause the atoms to move towards a new minimum potential energy configuration. This deformation of the lattice begins immediately following the excitation but continues long after the incidence of the pump pulse. The change in the dielectric function accompanying the lattice deformation should therefore continue to progress in the picoseconds following the excitation, in agreement with the observed behavior.

## Conclusion

We determined the response of the dielectric function of GaAs to intense femtosecond excitation up to and above the threshold for permanent change. The measurements extend from 1.5 to 3.5 eV and provide the most detailed view thus far of the electron and lattice dynamics during the induced phase transition.

In the low fluence regime (up to about  $0.5 F_{\text{th}}$ ) the initial response is purely electronic in nature. The dominant effect is change in bandstructure rather than just absorption by free carriers: i.e. the largest effect of carriers on the optical properties is through their influence on the rest of the valence electrons rather than through their own direct contribution to  $\epsilon(\omega)$ . The changes in the dielectric function arise immediately after the excitation, consistent with the electronic nature of the effect. After several picoseconds, the effect of electronic excitation on the optical properties diminishes due to recombination and cooling of the electrons. The electronic cooling in turn heats the lattice, and this effect dominates the changes in  $\epsilon(\omega)$  after 4 ps. The observed  $\epsilon(\omega)$  shows excellent agreement with equilibrium measurements at elevated temperatures. The rate and extent of the heating increases with increasing fluence.

In the medium fluence regime ( $0.6\text{--}0.8 F_{\text{th}}$ ) the initial response is similar to the low fluence regime, but stronger due to the higher density of excited carriers. Four picoseconds after excitation at  $0.7 F_{\text{th}}$  the dielectric function resembles that of amorphous GaAs, indicating loss of order in the lattice. The disordered state is not molten, but instead closely resembles the  $\epsilon(\omega)$  of solid amorphous GaAs. The bonding is still largely covalent rather than metallic in character. The loss of order is reversible — the semiconductor eventually returns to its original crystalline state.

In the high fluence regime (above  $0.8 F_{\text{th}}$ ) by 500 fs we already see a very large downward shift in frequency of the major optical resonances in the semiconductor. This shift continues for several

picoseconds following the excitation and by 4 ps  $\epsilon(\omega)$  is metal-like above 1.8 eV. This downward progression suggests that structural changes in the lattice are likely responsible for these very drastic changes in the dielectric function. The structural changes are more extreme than just the disordering seen in the medium fluence regime and indicate changes in the lattice structure and the nature of the bonding in the semiconductor. In the range of 0.8–1.0  $F_{th}$  these changes are reversible.

## References

1. E.N. Glezer, Y. Siegal, L. Huang, and E. Mazur, Phys. Rev. B **51**, 6959 (1995).
2. E.N. Glezer, Y. Siegal, L. Huang, and E. Mazur, Phys. Rev. B **51**, 9589 (1995).
3. Y. Siegal, E.N. Glezer, L. Huang, and E. Mazur, Ann. Rev. of Mat. Sci. **25**, 223 (1995).
4. C. V. Shank, R. Yen, and C. Hirlimann, Phys. Rev. Lett. **51**, 900 (1983).
5. M. C. Downer, R. L. Fork, and C. V. Shank, J. Opt. Soc. Am. B **2**, 595 (1985).
6. H. W. K. Tom, G. D. Aumiller, and C. H. Brito-Cruz, Phys. Rev. Lett. **60**, 1438 (1988).
7. S. V. Govorkov, I. L. Shumay, W. Rudolph, and T. Schroeder, Opt. Lett. **16**, 1013 (1991).
8. K. Sokolowski-Tinten, H. Schulz, J. Bialkowski, and D. von der Linde, Appl Phys A **53**, 227 (1991).
9. P. N. Saeta, J. Wang, Y. Siegal, N. Bloembergen, and E. Mazur, Phys. Rev. Lett. **67**, 1023 (1991).
10. C. V. Shank, R. Yen, and C. Hirlimann, Phys. Rev. Lett. **50**, 454 (1983).
11. H. Kurz, and N. Bloembergen, in *Energy Beam-Solid Interactions and Transient Thermal Processing*, edited by D. K. Biegelsen, G. A. Rozgonyi and C. V. Shank (Materials Research Society, Pittsburgh, 1985).
12. D. H. Kim, H. Ehrenreich, and E. Runge, Solid State Commun. **89**, 119 (1994).
13. R. Biswas, and V. Ambegoakar, Phys. Rev. B **26**, 1980 (1982).
14. P. Stampfli, and K. H. Bennemann, Phys. Rev. B **42**, 7163 (1990).
15. Trey Graves, private communication.
16. Y. R. Shen, *The Principles of Nonlinear Optics* (John Wiley & Sons, New York, 1984).
17. E. D. Palik, in *Handbook of Optical Constants of Solids*, edited by E. D. Palik (Academic Press, Inc., New York, 1985).
18. R. F. Potter, in *Optical Properties of Solids*, edited by S. Nudelman and S. S. Mitra (Plenum, New York, 1969).
19. H. Yao, P.G. Snyder, and J.A. Woollam, preprint.
20. G. A. Samara, Phys. Rev. B **27**, 3494 (1983).
21. M. Erman et al., J. Appl. Phys. **56**, 2664 (1984).
22. W. A. Harrison, *Electronic Structure and the Properties of Solids: The Physics of the Chemical Bond* (Dover, New York, 1989).

- 23.V. M. Glazov, S. N. Chizhevskaya, and N. N. Glagoleva, *Liquid Semiconductors* (Plenum Press, New York, 1969).
- 24.W. Jank, and J. Hafner, *J. Non-Crystalline Solids* **114**, 16 (1989).
- 25.S. Froyen, and M. L. Cohen, *Phys. Rev. B* **28**, 3258 (1983).

## **Chapter 6**

### **3-D Optical Storage Inside Transparent Materials**

The following paper was to appear in Optics Letters.

# 3-D Optical Storage Inside Transparent Materials

E. N. Glezer, M. Milosavljevic, L. Huang, R. J. Finlay, T.-H. Her, J. P. Callan, and E. Mazur

*Division of Applied Sciences and Department of Physics, Harvard University, Cambridge, MA 02138*

We present a novel method for 3-D optical data storage that has submicron-size resolution, provides a large contrast in index of refraction, and is applicable to a wide range of transparent materials. Bits are recorded by focusing 100-fs laser pulses inside the material using a 0.65 NA objective. The laser pulse produces a submicron-diameter structurally altered region with high contrast in index of refraction. Binary information can be recorded by writing such bits in multiple planes, and read out with a microscope objective with a short depth of field. We demonstrate data storage and retrieval with 2- $\mu\text{m}$  in-plane bit spacing and 15- $\mu\text{m}$  inter-plane spacing (17 Gbits/cm<sup>3</sup>). Scanning electron microscopy and atomic force microscopy show structural changes confined to an area 200 nm in diameter.

Three-dimensional optical data storage offers the potential for very large recording capacity. In addition to extensive on-going research in volume holographic data storage, a number of recent papers report ‘point-like’ or ‘bit-wise’ binary 3-D optical storage in photopolymers<sup>1,2</sup> and photorefractive materials.<sup>3,4</sup> Using a nonlinear optical process in the medium, the optical interaction can be confined in all three dimensions to a micron-sized focal volume. Pioneering experiments demonstrated 3-D optical recording at a density of up to  $1.6 \times 10^{12}$  bits/cm<sup>3</sup> in a photopolymer using two-photon absorption and extremely tight focusing.<sup>2</sup> The stored information consisted of less than 1% changes in the local index of refraction, and was read out serially with a Nomarski differential interference contrast (DIC) laser microscope. Experiments in photorefractive materials<sup>3,4</sup> used linear absorption to record data at a density of up to  $4.2 \times 10^9$  bits/cm<sup>3</sup>. The authors suggested that the recording density could be increased using two-photon absorption. The stored data was read out with a phase-contrast microscope.<sup>3,4</sup>

Here we report a novel method for creating submicron-sized bits that have a large contrast in index of refraction and that can be read out with transmitted or scattered light under a standard microscope. The method can be used for permanent 3-D optical data storage in a wide range of materials including fused silica, fused quartz, sapphire, and various glasses and plastics, thus allowing for a storage medium that is mechanically, chemically and thermally very stable, and inexpensive. Unlike a photopolymer gel, there are no problems of distortion due to shrinkage and flow, and of isomerization due to ultraviolet light. And unlike photorefractive materials, the difficulties of fixing the recorded data are entirely avoided.

We tightly focus ultrashort laser pulses inside a transparent material to create localized structural changes, thereby altering the index of refraction. This process is used to record digital information in three dimensions by writing multiple planes as illustrated in Fig. 1. In our experiments, we translate the

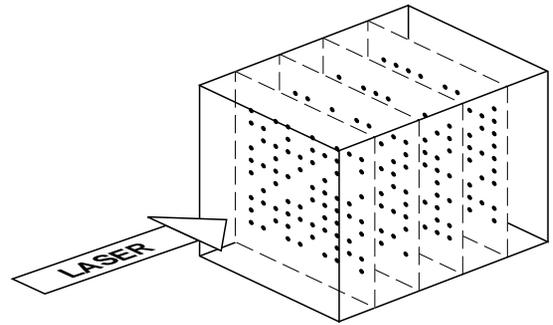


Fig. 1. Schematic diagram of 3-D optical data storage.

sample in the transverse plane, and move the focusing objective along the beam axis. Figure 2 shows an example of a random binary pattern stored inside fused silica, recorded using 0.5- $\mu\text{J}$ , 100-fs, 780-nm pulses from a regeneratively amplified Ti:Sapphire laser, focused by a 0.65 numerical aperture (NA) microscope objective. With this focusing, the threshold for observable structural change is 0.3  $\mu\text{J}$ . The pattern is read out in parallel using transmitted light in a microscope with a 0.95 NA objective. The spacing between adjacent bits is 2  $\mu\text{m}$ . The written spots can be viewed as dark or bright points depending on the focusing of the read-out objective. This change in brightness can be used as a contrast enhancing feature in a read-out system. During read-out, the depth discrimination provided by the short depth-of-field of the 0.95 NA objective is sufficient if adjacent layers are spaced by about 10  $\mu\text{m}$  or more. The longitudinal extent of the structurally altered regions is about 2.5  $\mu\text{m}$ . We recorded 10 layers spaced by 15  $\mu\text{m}$ , using a standard 0.65 NA refractive objective. Using an objective with a large working distance and an adjustment for aberrations caused by focusing into the material, it should be possible to record over 100 layers, spaced by 10  $\mu\text{m}$ .

The 1- $\mu\text{m}$  apparent size of the spots in Fig. 2 is at the resolution limit of the optical microscope. To obtain information about the lateral extent of the

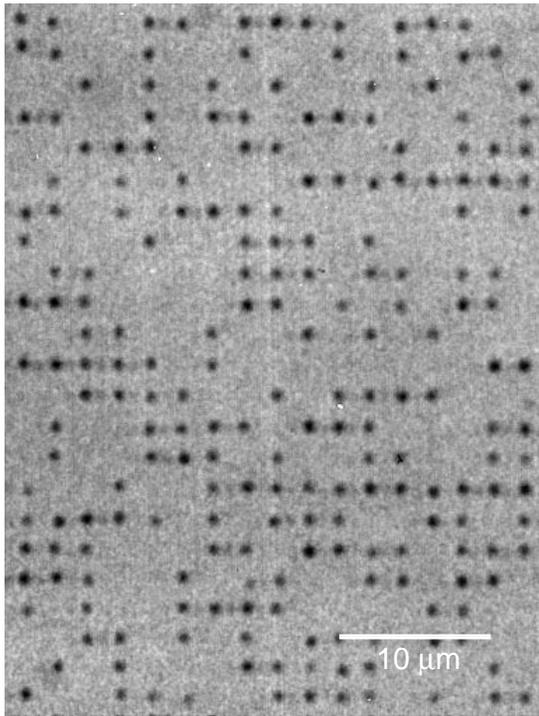


Fig. 2. Binary data pattern stored inside fused silica with 2- $\mu\text{m}$  bit spacing, photographed with an optical microscope using transmitted light.

structural changes in the material, we polished away the fused silica sample until the surface level reached the internally recorded structures, applied a 30-nm gold coat, and viewed the sample under a scanning electron microscope (SEM). Figure 3 shows such a SEM image of a 5 $\times$ 5- $\mu\text{m}$  regular array of spots recorded under conditions identical to those used in Fig. 2. The sample is tilted in the SEM to better show the morphology. The bright spots correspond to protrusions on the surface, while the dark spot (top row second from the right) corresponds to a cavity in the surface, as verified with an atomic force microscope (AFM). Whether a structurally altered region appears as a protrusion or a cavity depends on the level to which the material is polished. The inset in Fig. 3 shows a cavity at higher magnification. Both types of features have diameters of roughly 200 nm, implying a recording density limit of about  $10^{13}$  bits/cm<sup>3</sup> with a 0.65 NA objective.

It may be possible to write even smaller diameter and shorter length features using an objective with NA > 0.65. However, to read out data stored at densities significantly higher than what we have demonstrated requires using (serial) scanning techniques such as DIC laser microscopy.<sup>2</sup>

A likely mechanism for the creation of these structures is a ‘micro-explosion’ that occurs inside the material. Energy from the ultrashort laser pulse is coupled into the transparent material through a combination of multiphoton absorption and ava-

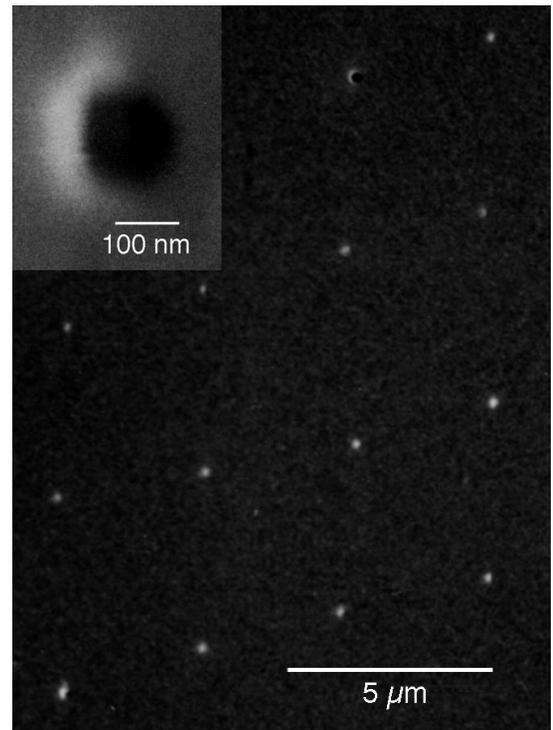


Fig. 3. A tilted SEM view of a polished cross-section through a regular array of bits with 5- $\mu\text{m}$  spacing. Inset shows a single bit at higher magnification.

lanche ionization.<sup>5-8</sup> Unlike ultrafast surface damage experiments,<sup>5-8</sup> however, ablation is not possible since the excited region is internal to the material. The ultrafast energy deposition creates very high temperatures and pressures inside the region; material is ejected from the center and forced into the surrounding volume, leading to the formation of a structure consisting of a void (or at least less dense material) surrounded by densified material. This mechanism is consistent with the SEM and AFM observations: the protrusions suggest the creation of denser, harder material, more resistant to the mechanical polishing; deeper polishing reveals a pit corresponding to a void which is created at the center of the micro-explosion.

The submicron bit size is not due to a simple thresholding effect which would occur with an excitation that is only a few percent above the threshold. The 0.5- $\mu\text{J}$  pulse energy used to make the patterns shown in Figs. 2 and 3 is well above the observed threshold of 0.3  $\mu\text{J}$ . Instead, the small size is likely caused by the nonlinearity of the absorption, which creates an excited region significantly smaller than the linear intensity distribution. Furthermore, self-focusing may be reducing the size of the beam waist. Finally, the dynamics of the micro-explosion may further confine the extent of the structurally altered region.

Surprisingly, the 0.5- $\mu\text{J}$  pulse energy used in these micro-explosions is *much less* than the energy

required to write bits in photosensitive materials:  $20\ \mu\text{J}$  per bit (2 mW for 10 ms, in 100-fs pulses at 100 MHz) was used for two-photon recording in a photopolymer with extremely tight focusing by a 1.4 NA objective<sup>2</sup>; linear absorption in a photorefractive<sup>4</sup> required  $60\ \mu\text{J}$  per bit (20 mW for 3 ms), also very tightly focused by a 1.0 NA oil-immersion objective lens. The threshold is higher in the experiments on photopolymers and photorefractives because the fraction of the incident laser energy that is absorbed in the focal volume is orders of magnitude smaller than in our experiments. Furthermore, while the change in index of refraction,  $\Delta n$ , in the photopolymer is  $8 \times 10^{-3}$ ,<sup>2</sup> and in the photorefractive material (LiNbO<sub>3</sub> crystal) only about  $10^{-6}$ ,<sup>4</sup> in our experiments the refractive index changes are large enough to be clearly visible under a standard microscope. Using diffraction from periodic arrays of bits (to be discussed in an upcoming publication) we estimate that  $\Delta n$  is in the range of 0.05–0.45, with the upper limit set by the silica-vacuum interface.

The structures produced by 100-fs pulses differ drastically from those produced by 200-ps and 10-ns pulses. With the longer pulses, the resulting structures are irregularly shaped, and cracks appear in the material even at energies only slightly above the threshold for structural change. Fig. 4 shows an example of the features produced by 200-ps pulses from the same laser system. The focusing conditions and beam profile are identical to those used in Fig. 2; the pulse energy is  $9\ \mu\text{J}$ , which is three times the observed threshold, and the point spacing is  $10\ \mu\text{m}$ . Note the large size of the spots, and clearly visible cracking produced with these longer pulses. With 10-ns pulses from a Nd:YAG laser, even larger, irregularly shaped spots are produced, and the cracking is more extensive.

With 100-fs pulses, the small, regular features described above are produced by pulses in a wide energy range of up to three times the  $0.3\text{-}\mu\text{J}$  threshold. Femtosecond pulses of even higher energy produce a ‘head and filament’ structure of  $20\text{--}40\ \mu\text{m}$  length that is visible to the unaided eye, and that can be used for internal marking and engraving 3-D patterns in a wide variety of transparent materials. We observed no cracking with the 100-fs pulses even up to pulse energies 100 times the threshold.

We have written information in a wide variety of transparent materials including fused silica, fused quartz, sapphire, BK7 optical glass, and plastic (acrylic). Surprisingly, the threshold for structural change is very similar in all of these materials — within a factor of 2 of the threshold of fused silica. A notable exception is diamond, where the threshold is at least 100 times greater. In all materials, and especially in diamond, the threshold for internal structural change is significantly higher than for surface ablation.

In conclusion, we have demonstrated a novel method for high density 3-D optical data storage using ultrashort laser pulses. The recording is done with ultrafast-laser-induced micro-explosions, pro-

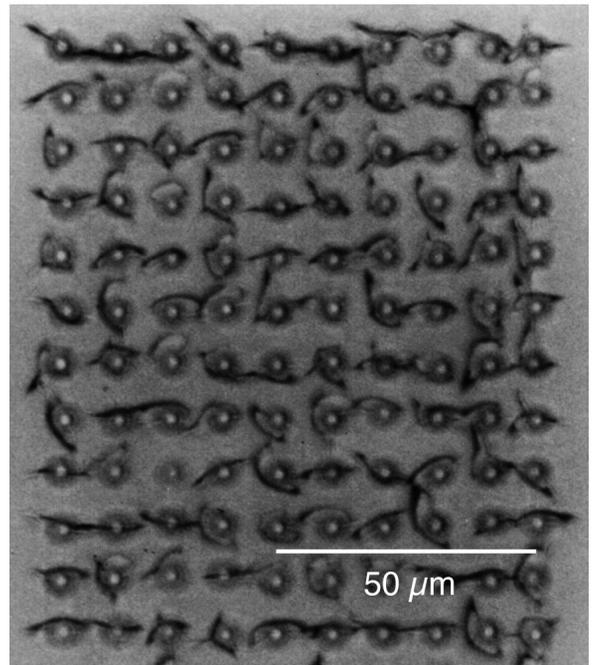


Fig. 4. Damage produced inside fused silica by 200-ps pulses. Optical microscope photograph using transmitted light. Note the difference in scale relative to Fig. 2; bits are spaced by  $10\ \mu\text{m}$ .

ducing localized submicron-diameter structures with high contrast in index of refraction. Parallel readout and depth discrimination are demonstrated with a standard microscope. Because a photosensitive medium is not required, a wide range of transparent materials can serve as the storage medium.

The method can also be used for engraving very fine-scale patterns inside transparent materials, without damaging or altering the surface. Other applications could include producing 3-D diffractive optical elements and periodic structures, patterning gratings in optical fibers, and creating materials with increased density and hardness.

We thank N. Bloembergen for insightful discussions. E.N.G. gratefully acknowledges financial support from the Hertz Foundation.

## References

1. D. A. Parthenopoulos and P. M. Rentzepis, *Science* **245**, 843 (1989).
2. J. H. Strickler and W. W. Webb, *Opt. Lett.* **16**, 1780 (1991); U.S. Patent #5,289,407 (1994).
3. Y. Kawata, H. Ueki, Y. Hashimoto, and S. Kawata, *Appl. Opt.* **34**, 4105 (1995).
4. H. Ueki, Y. Kawata, and S. Kawata, *Appl. Opt.* **35**, 2457 (1996).
5. J. Ihlemann, B. Wolff, and P. Simon, *Appl. Phys. A* **54**, 363 (1992).
6. D. Du, X. Liu, G. Korn, J. Squier, and G. Mourou, *Appl. Phys. Lett.* **64**, 3071 (1994).
7. B. C. Stuart, M. D. Feit, S. Herman, A. M. Rubenchik, B. W. Shore, and M. D. Perry, *Phys. Rev. Lett.* **74**, 2248 (1995); *J. Opt. Soc. Am. B*, **13**, 459 (1996).
8. D. von der Linde and H. Schüler, *J. Opt. Soc. Am. B*, **13**, 216 (1996).

## Chapter 7

# Ultrafast-laser driven micro-explosions inside transparent materials

### Introduction and background

We have recently discovered that submicron-diameter structures can be produced inside many transparent materials by tightly focused ultrashort laser pulses.<sup>1-3</sup> These structures or voxels have a high contrast in index of refraction and can be used for recording bits in multiple planes for high density 3-D optical data storage and other applications that involve writing very fine-scale patterns inside transparent materials.<sup>1-3</sup> In this paper we present the results of optical and structural examination of these submicron-diameter voxels and describe in more detail the physical processes involved in their formation.

### Laser damage in transparent materials: nanosecond pulses

Laser-induced damage in transparent materials has been a common phenomenon since its first observation in 1963, when transparent dielectrics were damaged by a pulsed ruby laser<sup>4</sup>. Damage can occur at the surface or in the bulk of the material. The fluence threshold for damage is always lower at the surface, so for unfocused laser beams damage usually first occurs at the surface. If, on the other hand, the beam is focused inside a transparent material by an external lens or by internal self-focusing, the damage can be confined entirely inside the bulk of the medium. The damaged region produced by focused nanosecond pulses is usually characterized by an irregular shape, extensive cracking, and strong optical scattering. With self-focused nanosecond pulses, long filaments are produced by a focus that moves back (opposite to the direction of light propagation) as the intensity increases during the pulse.<sup>5</sup>

The underlying goal in optical damage experiments has been to understand the cause of the damage so that materials can be perfected to withstand very high intensities. With nanosecond pulses, the damage threshold is usually determined by extrinsic factors such as impurities, crystal imperfections and, in the case of surface damage, by surface imperfections. Wide fluctuations in measured thresholds are typical, and damage is not necessarily initiated at the point of highest intensity.

### **Surface ablation: ultrashort vs. nanosecond pulses**

Recent experiments using picosecond and sub-picosecond laser pulses for surface ablation of transparent materials have shown that the effect of these pulses on the material is quite different from that of longer pulses.<sup>10-14</sup> Rather than depending on impurities or surface imperfections, the ablated area depends only on the beam profile. The energy deposition is initiated through a multi-photon absorption process (an intrinsic process), rather than depending on extrinsic effects for creating the initial electrons. Because the energy deposition is deterministic rather than statistical, there is very little shot-to-shot fluctuation in the fluence threshold for ablation.

Once the initial electrons are excited across the bandgap through multi-photon absorption, they acquire excess kinetic energy through sequential linear absorption and then scatter off other electrons exciting them across the bandgap by impact ionization. These new free electrons can then also undergo linear absorption, and this process continues as an avalanche. For pulses shorter than about 30 fs, a theoretical model suggests that multiphoton ionization alone is sufficient to produce the critical density of electrons for ablation.<sup>13</sup> Multiphoton absorption is dominant because of the high intensity, and because less time is available for a collisional avalanche during the sub-30-fs duration of the pulse.

With sub-picosecond pulses, the energy is deposited well before any ablation takes place.

Moreover, the electrons are ionized and heated faster than they can transfer their excess energy to the lattice. Thus a hot plasma is formed before the lattice is significantly heated. A comparison can be made to recent experiments on GaAs with 70-fs excitation, where a lattice heating time of about 10 ps has been observed below the damage threshold.<sup>15</sup> Above the threshold, changes in the lattice structure of GaAs are directly driven by the destabilization of the covalent bonding in less than 1 ps. This behavior is in contrast to excitation with pulses longer than a few picoseconds, where the laser pulse continues to deposit energy as the lattice heats up and possibly undergoes structural changes. Observations of the surface of fused silica after laser damage show that the morphology of the ablated surface also depends on the pulse duration: with pulses longer than about 20 ps, the damaged area is characterized by melting and boiling of the surface, whereas with pulses shorter than 10 ps, a shallow crater is produced where a thin layer of material has been removed by ablation.<sup>13</sup>

### **Bulk vs. surface**

An ultrashort pulse focused *inside* a transparent material deposits energy only in the focal volume, where the intensity is high enough for multiphoton absorption. The basic mechanisms for absorption — multiphoton and avalanche ionization — are the same in the bulk as at the surface. However, while surface states can contribute to absorption at the surface, in the bulk an electron must be excited across the entire energy bandgap (assuming the concentration of impurities is low). Additionally, in the bulk the role of self-focusing<sup>16</sup> needs to be considered, as well the defocusing effect of the plasma produced by the propagating pulse.

After the energy is deposited by the laser pulse, the atomic lattice begins to respond to the electronic excitation, but ablation is not possible because the excited volume is in the bulk of the material. The expansion that normally accompanies heating, melting, and vaporization is

restrained by the surrounding lattice which serves as an ideal confinement chamber, possibly producing tremendous pressures especially in extremely hard materials such as sapphire.

## **Description of experiments**

The experiments described in this paper are performed with a Kerr-lens-modelocked, regeneratively amplified Ti:Sapphire laser<sup>17</sup>. Chirped-pulse amplification<sup>18</sup> is used to avoid nonlinear optical effects in the amplifier. The 100-fs pulses from the laser oscillator are stretched to 200 ps, injected into an amplifier cavity, ejected after being amplified to 1 mJ in 25 roundtrips, and then recompressed to about 100 fs. A regenerative amplifier has the advantages of excellent mode quality directly out of the amplifier cavity, and very good pulse-to-pulse energy stability due to strong gain-saturation. After recompression, the amplified output is a 1-kHz train of 100-fs, 500- $\mu$ J pulses centered at a wavelength of 780 nm. Less than 0.1% of the laser pulse energy is needed to drive the micro-explosions.

To tightly focus the pulses inside the transparent material we use a 0.65 numerical aperture (NA) microscope objective. Refraction at the surface of the material causes spherical aberration, which adversely affects the focusing at depths greater than about 100  $\mu$ m. Deeper focusing can be achieved by using either an objective with adjustable spherical aberration, or one that is optimized for focusing to a specific depth together with a compensator plate of variable thickness.

In our 3-D data storage experiments,<sup>1-3</sup> we translate the sample in the transverse plane, and move the focusing objective along the beam axis with numerically controlled translation stages with 0.1- $\mu$ m precision. The bit pattern is controlled by a mechanical shutter. The multiple 2-D data planes were usually written in the back-to-front-plane order, but writing in the front-to-back order was also tested. No noticeable difference was observed between the two writing modes.

To determine the threshold for observable structural change, we wrote regularly spaced arrays

of voxels of decreasing pulse energy, all at the same depth of about  $100\ \mu\text{m}$ , and examined the area under a high-power microscope with up to 0.95 NA. Having a regularly spaced array of points made detection near the threshold of visible change much easier. Above the threshold, the voxels are visible either in transmission mode, or in scattered light using off-axis illumination.

Most of the results presented below were obtained in fused silica samples. We also examined the effects of tightly-focused ultrashort pulses in other materials, including several optical glasses, amorphous quartz, optical-grade sapphire, plastics (acrylics), and diamond.

For observation under a scanning electron microscope (SEM) and an atomic force microscope (AFM), we polished the material until the surface level reached the plane of voxels using diamond paste. Final polishing of the fused silica samples was done with Cerium Oxide. For SEM examination, the sample was coated with 30-nm of gold by sputtering.

To compare the effect of ultrafast excitation to the effect of longer laser pulses we used the 200-ps uncompressed output of the Ti:Sapphire amplifier, as well as 10-ns pulses from a Nd:YAG laser. The uncompressed Ti:Sapphire pulses are ideal for examining the effect of longer pulse-width, because the laser mode, wavelength, and focusing are identical to the 100-fs pulses.

## Results

Figure 1 shows an example of an array of voxels written in a binary data pattern inside fused silica, using  $0.5\text{-}\mu\text{J}$ , 100-fs, 780-nm pulses focused by a 0.65 numerical aperture (NA) microscope objective. Under these focusing conditions, the threshold for structural change is  $0.3\ \mu\text{J}$ . The spacing between adjacent bits is  $2\ \mu\text{m}$ . We recorded 10 layers spaced by  $15\ \mu\text{m}$ , using a standard 0.65 NA refractive objective. The image is read out using transmitted light in a microscope with a 0.95 NA objective. The written spots can be viewed as dark or bright points depending on the position of the read-out objective. During read-out, the depth discrimination provided by the short depth-

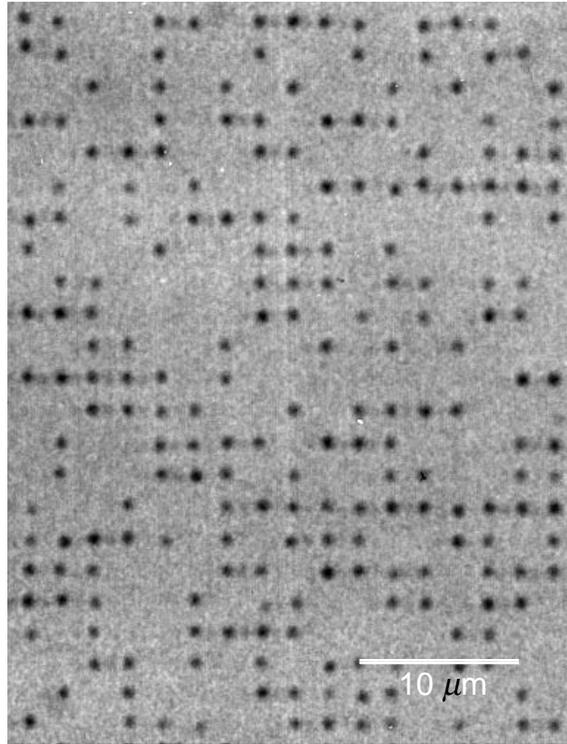


Fig. 1. Binary data pattern stored inside fused silica with 2- $\mu\text{m}$  bit spacing, photographed with an optical microscope using transmitted light.

of-field of the 0.95 NA objective is sufficient if adjacent layers are spaced by about 10  $\mu\text{m}$  or more. The longitudinal extent of the structurally altered regions is about 2.5  $\mu\text{m}$ .

With 100-fs pulses small, regularly-shaped voxels are produced by pulses in a wide energy range of up to three times the 0.3- $\mu\text{J}$  threshold for structural change in fused silica. Both the diameter and the longitudinal extent increase with energy. Pulses of even higher energy produce a head-and-filament structure of 20–40- $\mu\text{m}$  length that is visible to the unaided eye. We observed no cracking with the 100-fs pulses in fused silica even up to pulse energies 100 times the threshold.

We have written arrays of voxels in a wide variety of transparent materials including fused silica, fused quartz, sapphire, BK7 optical glass, and plastic (acrylic). Surprisingly, the threshold for structural change is very similar in all of these materials — within a factor of 2 of the threshold of fused silica. A notable exception is diamond, where the threshold is at least 100 times greater. In

all materials, and especially in diamond, the threshold for internal structural change is significantly higher than for surface ablation.

For even higher resolution than in Fig.1 we used a microscope with a 1.2 NA oil-immersion objective to examine the voxel arrays. Ordinarily, the voxels cannot be seen in reflection. However, under illumination with an annulus of light that is focused at an angle greater than the critical angle for total internal reflection the voxels can be seen clearly in reflected light against a dark background.

Figure 2 shows a 3-D projected view of a 2- $\mu\text{m}$  spaced voxel array, recorded under the same conditions as the pattern in Fig. 1 (0.5- $\mu\text{J}$ , 100-fs, 780-nm pulses; 0.65 NA focusing objective). The depth of the features corresponds to the brightness of the light reflected by the voxels. From this optical photograph we can see that the diameter of the voxels is less than 0.5  $\mu\text{m}$ .

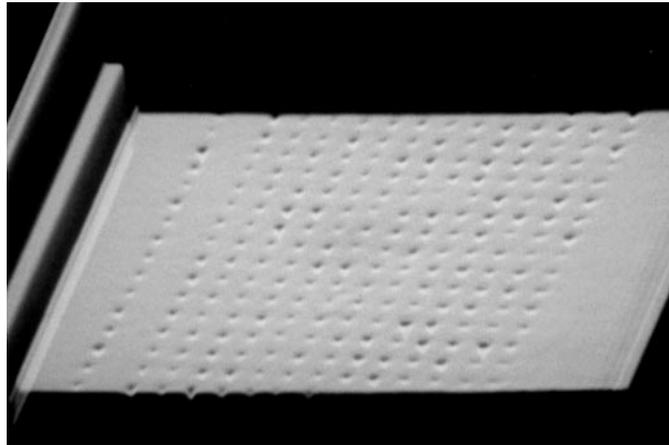


Fig. 2. 2- $\mu\text{m}$  spaced regular array of voxels inside fused silica, photographed in reflection using a 1.2 NA objective with above-critical-angle illumination, and shown in a 3-D projection.

The apparent size of the voxels in both Figs. 1 and 2 is at the resolution limit of the two optical microscopes. To obtain information about the extent of the structural changes in the material we polished away the fused silica sample until the surface level reached the internally recorded vox-

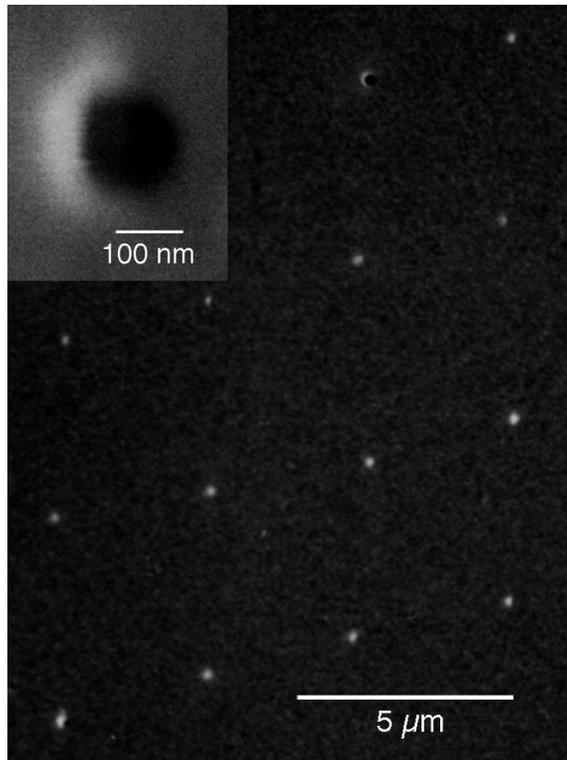


Fig. 3. A tilted SEM view of a polished cross-section through a regular array of voxels with  $5\text{-}\mu\text{m}$  spacing. Inset shows a single voxel at higher magnification.

els, and viewed the sample under a scanning electron microscope. Figure 3 shows a SEM image of a  $5\times 5\text{-}\mu\text{m}$  regular array of voxels recorded under conditions identical to those used in the patterns of Fig. 3. The sample is tilted in the SEM to better show the morphology. The bright spots correspond to protrusions or bumps on the surface, while the dark spot (top row second from the right) corresponds to a cavity or pit in the surface. Whether a structurally altered region appears as a bump or a pit depends on the level to which the material is polished. A typical pit is shown at higher magnification in the inset of Fig. 3. Both types of features have diameters of roughly 200 nm.

To examine the morphology of the voxels in more detail, we used an atomic force microscope (AFM). Figure 4 shows examples of depth profiles of the pits and bumps recorded by the AFM. The bumps are 7–10 nm high, while the pits are at least 35–40 nm deep. The pit depth may be

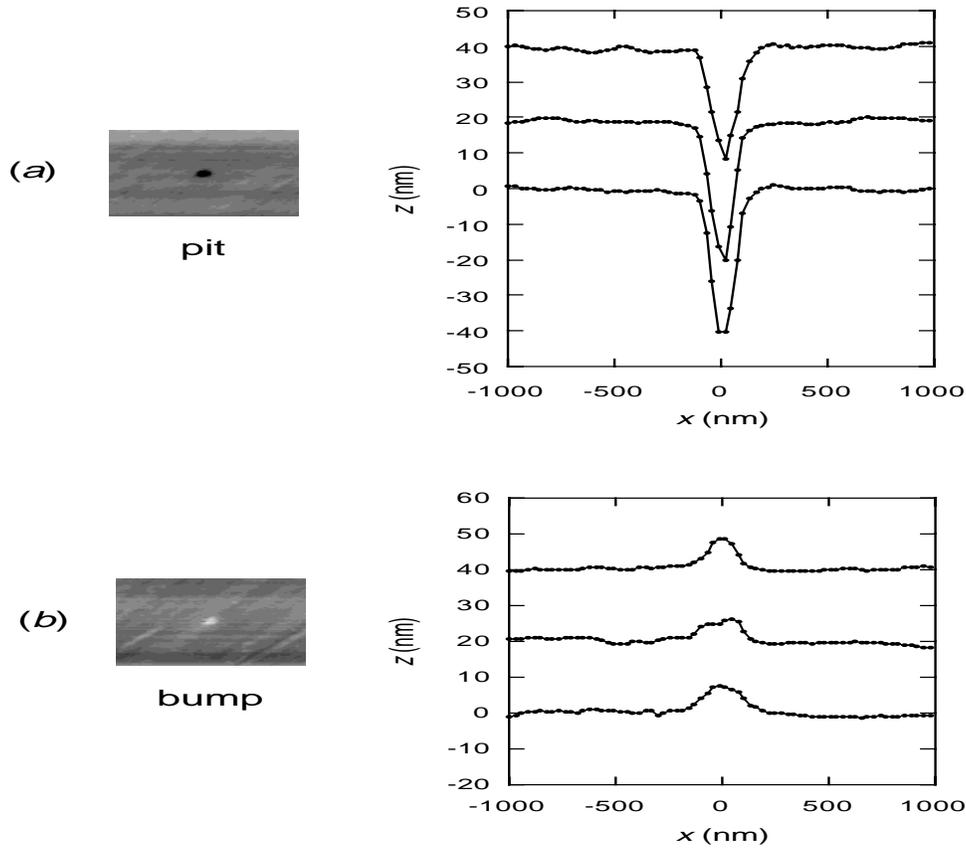


Fig. 4. AFM scans of three pits and three bumps. Also shown are AFM images of a pit and a bump.

greater than the measured value, because the measurement is limited by the size of the AFM tip (which is pyramidal with a sub-40-nm tip radius). The diameter of both the pits and the bumps is 200–250 nm.

Because of its small size, it is difficult to directly measure the optical properties of a single voxel. Instead, we used the diffraction pattern from periodic arrays of voxels to obtain quantitative information on the optical properties. Fig. 5 shows a diffraction pattern produced by a  $5 \times 5$ - $\mu\text{m}$  regular array of voxels written in plastic. A Helium-Neon laser beam was focused to illuminate a 20–30- $\mu\text{m}$  diameter area of the array, and the transmitted diffraction pattern was photographed. The spacing of the diffraction pattern is inversely proportional to the spacing of the array. Similar patterns were observed with voxel arrays in fused silica and sapphire. By compar-

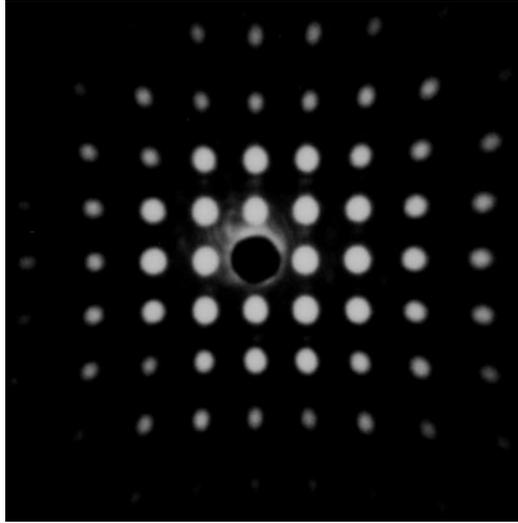


Fig. 5. Diffraction from a  $5\text{-}\mu\text{m}$  spaced array of voxels.

ing the intensities in the diffracted orders to the central peak we can estimate the change in the index of refraction inside the voxels (see Discussion section). In fused silica, the ratio of the diffracted light to the central peak was measured for  $3\times 3\text{-}\mu\text{m}$ ,  $2\times 2\text{-}\mu\text{m}$ , and  $1\times 1\text{-}\mu\text{m}$  spaced regular arrays, recorded under the same conditions as those shown in Figs. 1-4. To compare beams diffracted by the same angle in each of these arrays, we used the third, second, and first order peaks, respectively. The ratios of diffracted to central peak are  $2\times 10^{-4}$ ,  $6\times 10^{-4}$ , and  $6\times 10^{-3}$ , for the  $3\times 3\text{-}\mu\text{m}$ ,  $2\times 2\text{-}\mu\text{m}$ , and  $1\times 1\text{-}\mu\text{m}$  spaced regular arrays, respectively.

The voxels produced by 100-fs pulses differ drastically from the damaged regions produced by 200-ps and 10-ns pulses. With the longer pulses, the resulting structures are irregularly shaped, and cracks appear in the material even at energies only slightly above the threshold for structural change. An example of the damaged regions produced by 200-ps pulses from the same laser system is shown in Fig. 6. Fig. 6a shows a front view (seen along the direction of the writing beam), and Fig. 6b shows a side view with the laser beam incident from the left. The focusing conditions and beam profile during writing are identical to those used with the 100-fs pulses; the pulse

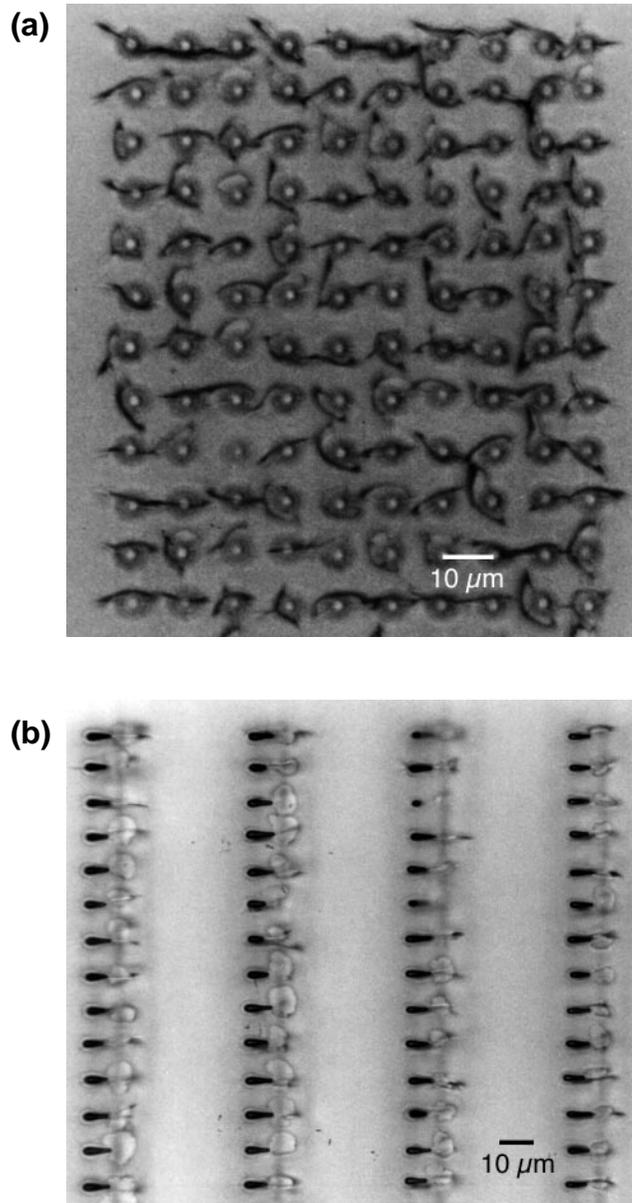


Fig. 6. Damage produced inside fused silica by 200-ps pulses. Optical microscope photograph using transmitted light: (a) front view, (b) side view. Note the differences in scale relative to Fig. 1.

energy is  $9 \mu\text{J}$ , which is three times the observed threshold with 200-ps pulses; the spacing is  $10 \mu\text{m}$ . Note the large size of the damaged regions, and the clearly visible cracking produced with these longer pulses. The side view shows interesting features in the damaged regions: a drop-shaped front end and an irregularly-shaped and cracked back end.

With 10-ns pulses from a Nd:YAG laser, even larger irregularly shaped damaged regions are produced, and the cracking is more extensive. Figure 7 shows a SEM view of a polished cross-section through a damaged region created inside fused silica by a 10-ns, 30- $\mu\text{J}$ , 532-nm wavelength pulse focused by a 0.25 NA objective. A cavity appears at the center of the damaged region, surrounded by radiating structural features extending over an area about 20- $\mu\text{m}$  in diameter. Note the dramatic contrast between the SEM views of this highly damaged region and the 200-nm diameter voxels of Fig. 3.

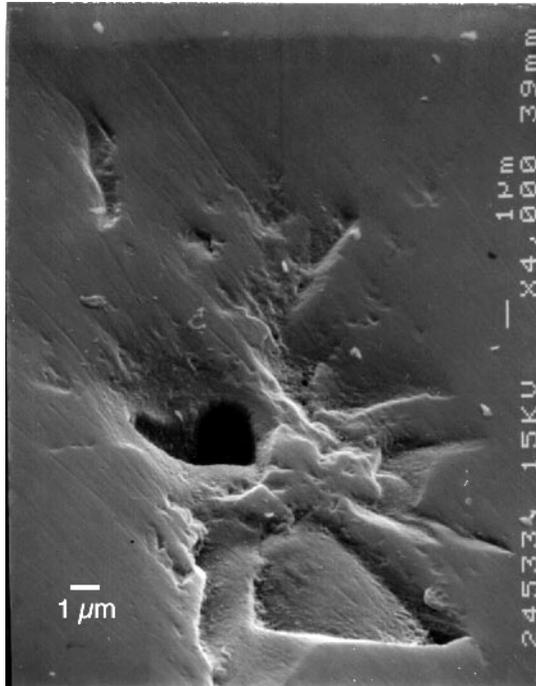


Fig. 7. A SEM view of damage produced inside fused silica by a 10-ns laser pulse.

All of the results presented so far involve examining the sample after the structural changes have occurred. We now turn to the electronic excitation that drives the structural changes. To observe the plasma created by the ultrafast excitation we carried out time-resolved imaging experiments. The excitation is created by a 100-fs pump pulse, and then imaged with a time-delayed 100-fs probe pulse of much lower fluence. The probe is incident perpendicular to the pump, illu-

minating the imaged region in a side view, as shown in Fig. 8. The transmitted probe light is imaged onto a CCD camera. Objectives with 0.20 NA are used for focusing and illumination, and a 0.65 NA objective is used for the imaging. The probe pulse is focused before reaching the excited region, providing a uniform illumination of the field of view. An image taken without a pump pulse is used for background subtraction and normalization.

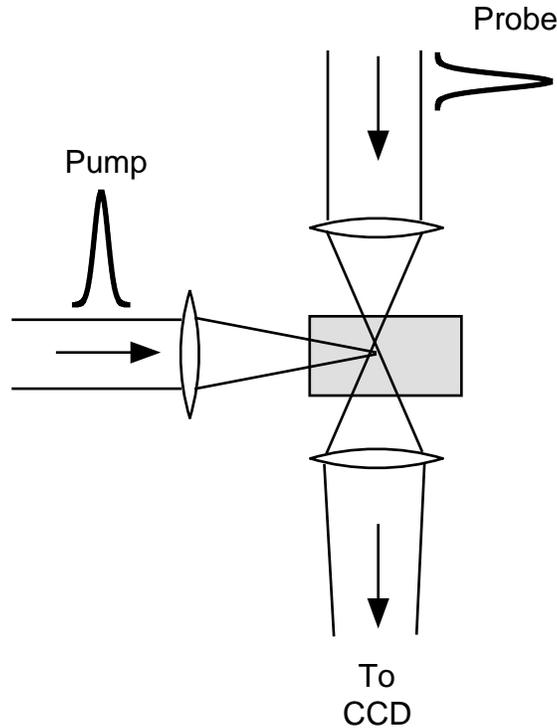


Fig. 8. Schematic diagram of the setup for time-resolved imaging.

Figure 9 shows the image of the pumped region 1 ps, 5 ps, 100 ps, and 1 second after the excitation by a pump pulse well above threshold. The horizontal length of the image is about  $200\ \mu\text{m}$ , and the pump pulse is incident from the left. The dark cone seen in the first three images is due to the probe light being absorbed in the plasma created by the pump pulse. No plasma is visible beyond the focal region of the pump pulse, suggesting that the pump pulse is strongly attenuated and does not propagate beyond the focal region. After one second, long after the plasma has

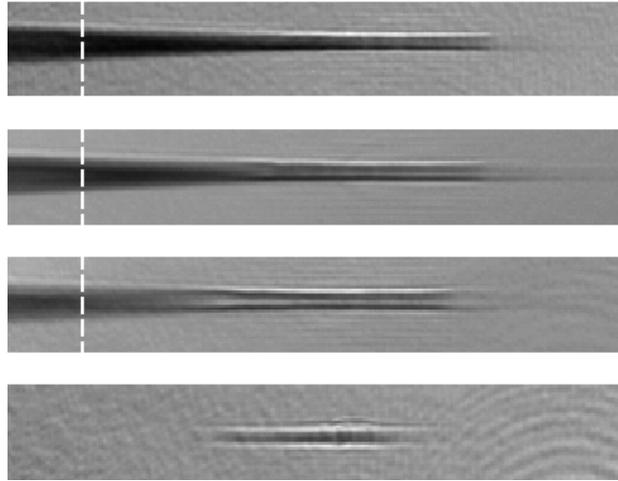


Fig. 9. Time-resolved images at 1 ps, 5 ps, 100 ps, and 1 second after excitation.

recombined and the material has cooled, a filament-like structurally altered region is visible where the pump was focused. No change is apparent in the larger cone region where plasma was visible.

Figure 10 shows three curves corresponding to vertical slices through the 1 ps, 5 ps, and 100 ps images of Fig. 9, as indicated by the dashed white lines in the images. The values give the fraction of light that is transmitted through the plasma. The absorption in the dark cone is initially about 45%, and decreases to about 35% after 100 ps. The plasma persists for over 100 ps in the region of the cone. It is difficult to determine the plasma density in the cone, since the Drude relaxation time,  $\tau$ , that characterizes the plasma, is not known. As a rough estimate, taking  $\tau$  to be in the range of 0.1–10 fs, gives a plasma density in the cone of about  $10^{19} - 10^{20} \text{ cm}^{-3}$ . At the focus, the density is probably even higher.

One interesting observation in the time-resolved images is that a high-density plasma is formed well outside the region where structural changes are apparent. This indicates that it is possible to create a plasma inside the material without causing structural change. Thus observing the presence of the plasma (*e.g.*, by monitoring the transmission of a probe pulse or the emission of the plasma) may not necessarily indicate that breakdown and permanent structural change take place.

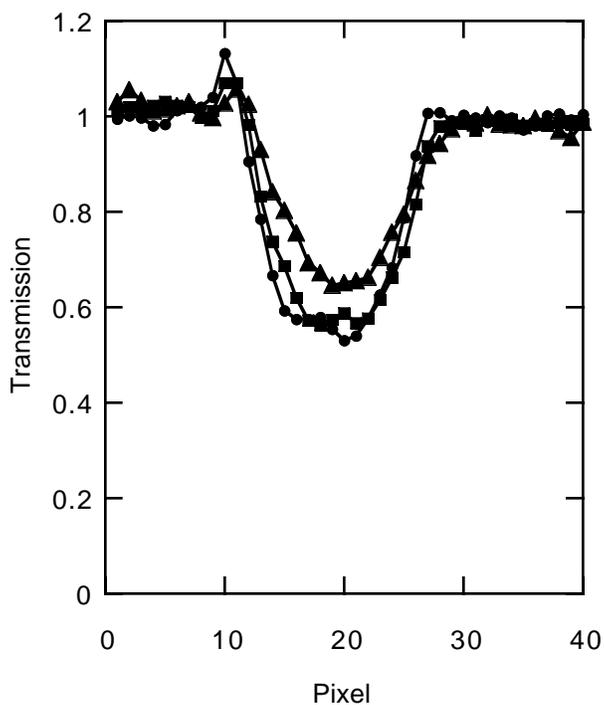


Fig. 10. Transmission of light in the plasma cone of Fig. 9 at 1 ps (●), 5 ps (■), and 100 ps (▲).

## Discussion

### Index of refraction in the voxels

We now analyze the results of the diffraction experiments to estimate the change in the index of refraction inside the voxels. We model each voxel as a cylindrical element with an index of refraction different from the bulk. The change in the index,  $\Delta n$ , affects the phase of the incident optical field as it is transmitted along the axis of the cylinder. The 2-D array of voxels forms a periodic diffractive structure, whose Fourier transform (squared) is measured in the diffracted pattern. For a given ratio of the diameter of the voxels to their spacing, and a known voxel length (about  $2.5 \mu\text{m}$ ), we can calculate the diffraction efficiency, defined as the ratio of the energy in the diffracted peak to the central one. An additional consideration is that the voxels are not all of the same length, which reduces the diffraction efficiency. In our calculation we assume that their

length varies  $\pm 50\%$ , and that the beam covers 10 periods of the array in each direction.

Figure 11 shows the calculated diffraction efficiency as a function of  $\Delta n$  for an array of voxels with a diameter equal to 10% of the spacing. The diffraction efficiency increases with  $\Delta n$  until the average phase difference in a voxel equals  $\pi$ . With further increase in  $\Delta n$  the diffraction efficiency oscillates, converging to a final value corresponding to a randomly phased diffractive structure. Also shown in Fig. 11 (as horizontal lines) are the measured diffraction efficiencies for 3, 2, and 1- $\mu\text{m}$  spaced arrays. If the effective optical diameter of the voxels is 300 nm, then the 3- $\mu\text{m}$  spacing should correspond to the condition used in the calculation that the diameter equals 10% of the spacing. Alternatively, if the diameter is 200 nm or 100 nm, then the calculation corresponds to the 2- $\mu\text{m}$  or 1- $\mu\text{m}$  spacing, respectively.

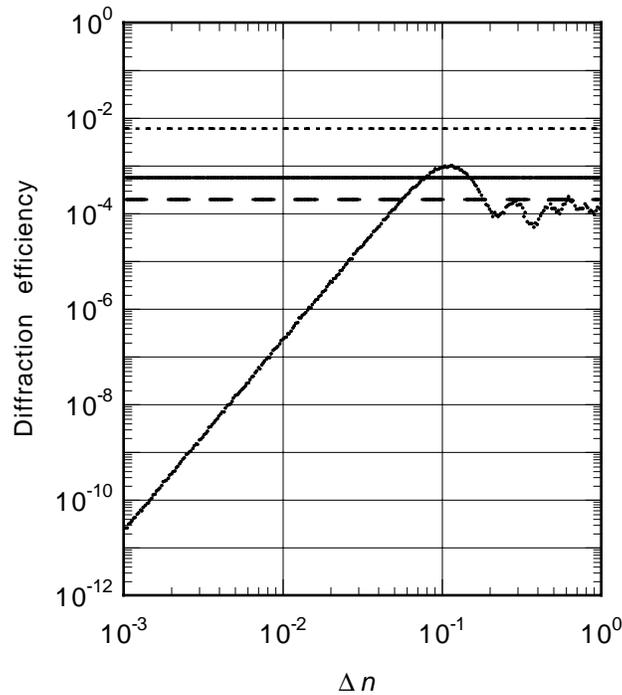


Fig. 11. Calculated diffraction efficiency of a voxel array as a function of  $\Delta n$  in the voxels. Assumptions: voxel diameter equal to 10% of spacing; 2.5- $\mu\text{m}$  voxel length with  $\pm 50\%$  variation; beam covers 10 periods of array. Horizontal lines are measured diffraction efficiencies of 3- $\mu\text{m}$  (dashed), 2- $\mu\text{m}$  (solid), and 1- $\mu\text{m}$  (dotted) spaced arrays.

The most important result of this analysis is that  $\Delta n$  is not small, *i.e.*, certainly not  $10^{-3}$  or even  $10^{-2}$ . Such a large change in index of refraction cannot be due to dislocations and slight structural changes in the material. For comparison, ultraviolet-light-induced compaction in Germanium-doped silica (used in fiber gratings) produces a  $\Delta n$  of  $10^{-4}$ – $10^{-3}$ .<sup>19</sup> (Although with special glass preparation by H<sub>2</sub> loading, changes as large as  $10^{-2}$  can be produced.<sup>19</sup>)

From Fig. 11 we can also see that the voxels are greater than 100-nm in diameter, because the 1- $\mu\text{m}$  spaced array diffracts light with higher efficiency ( $6\times 10^{-3}$ ) than is possible with 100-nm diameter voxels of any  $\Delta n$  according to the calculation. This is consistent with the SEM and AFM results. If the effective optical diameter of the voxels is 200 nm, then the diffraction efficiency from the 2- $\mu\text{m}$  spaced pattern ( $6\times 10^{-4}$ ) corresponds to the conditions used in the calculation. The diffraction efficiency from the 3- $\mu\text{m}$  spaced pattern ( $2\times 10^{-4}$ ) provides a comparison with the calculation if the effective optical diameter of the voxels is 300 nm. Considering the uncertainties both in the diameter of the voxels and in the variation in their length, we estimate that  $\Delta n$  is in the range of 0.05–0.45, where the upper limit is assumed to be the change across an interface between fused silica and vacuum. Additionally, as we noted above, while the voxels cannot ordinarily be seen in reflection, they are clearly seen in reflection (Fig. 2) by a 1.2 NA oil-immersion objective under illumination with an annulus of light above the total internal reflection angle of the silica-air interface. This provides further evidence that the change in index of refraction at the voxel interface is similar in magnitude to that at a silica-air interface.

Taken together with the structural examination of the polished samples, these results suggest that formation of a void by the micro-explosion is likely. Formation of a void must be accompanied by a densification of the surrounding region. This possibility is particularly intriguing in the case of sapphire, where no transition to a denser phase has ever been seen, even in shock experi-

ments with pressures of over 300 GPa (3 million atmospheres).<sup>20</sup>

### **Size of the voxels**

Why are the voxels so small? The FWHM diameter at the focus of our beam is estimated to be 0.9  $\mu\text{m}$ . This assumes that the input beam fills the 0.65 NA objective to the  $1/e^2$  intensity point (86% transmission). Any aberrations and imperfections in focusing can only increase the focal area. Yet the optical observations place an upper limit of 0.5- $\mu\text{m}$  on the diameter of the voxels, and the SEM and AFM observations indicate that the diameter of the structurally altered regions is only 200–250 nm.

The submicron diameter is not due to a simple thresholding effect which would occur with an excitation that is only a few percent above the threshold. The 0.5- $\mu\text{J}$  pulse energy used to make the voxels shown in Figs. 1–4 is well above the observed threshold of 0.3  $\mu\text{J}$ . Instead, the small size is likely caused by the nonlinearity of the absorption, which creates an excited region significantly smaller than the linear intensity distribution. Furthermore, self-focusing may be reducing the size of the beam waist. Finally, the dynamics of the micro-explosion may further confine the extent of the structurally altered region.

The effect of the nonlinearity of the absorption is straightforward. The characteristic diameter of the excited area scales as the inverse of the square root of the order of the nonlinearity. For example, if the absorption is proportional to  $I^5$ , then the diameter of the excited volume will be reduced to 45% of the diameter of the intensity distribution.

### **Self-focusing**

The effect of self-focusing is not as easy to predict. One might expect that under tight external focusing, self-focusing does not play much of a role. However, the intensity in the 100-fs

pulses is so high that self-focusing cannot be ruled out. To examine the possible effect of self-focusing on the tightly-focused beam, we model the beam propagation under our experimental focusing conditions, by numerically propagating a Gaussian beam in small steps through the medium. The intensity-dependent focusing is modeled by a progression of lenses of varying strength. Self focusing results from the intensity-dependent radial variation in the index of refraction,  $\Delta n = n_2 I$ , where  $n_2$  is proportional to the third-order susceptibility. The strength of the self-focusing lens is proportional to  $d^{-3}$ , where  $d$  is the beam diameter, because the intensity scales as  $d^{-2}$  and the curvature of the phase variation scales as  $d$ . We define a parameter  $s$  to characterize the strength of the self-focusing,

$$s = (1/\pi) n_2 P,$$

$P$  being the power in the beam. With this definition, the effective focal length of the self-focusing lens at any point along the beam is simply

$$f = w^3/s,$$

where  $w$  is the  $1/e^2$  radius of the beam. Figure 12 shows the beam profile (defined by its  $1/e^2$  intensity contour) for varying degrees of self-focusing:  $s = 0, 0.001, 0.0033,$  and  $0.01$ , expressed in units of  $\mu\text{m}^2$ . The beam propagates from left to right. Without self-focusing (dashed curve), the width of the beam waist is  $2w_0 = 1.6 \mu\text{m}$ , which corresponds to  $d_{\text{FWHM}} = 0.9 \mu\text{m}$ . As the power in the beam is increased, the beam waist moves to the left and its diameter decreases.

Figure 13 shows the beam waist diameter as a function of  $s$ . With sufficiently large  $s$ , self-focusing significantly reduces the beam waist, even under already tight focusing conditions. For example, at  $s = 0.01 \mu\text{m}^2$ ,  $d_{\text{FWHM}}$  is reduced to  $0.15 \mu\text{m}$ . To check if this is physically possible we can calculate the maximum  $\Delta n$  that would be produced at the peak intensity. The peak intensity is given by

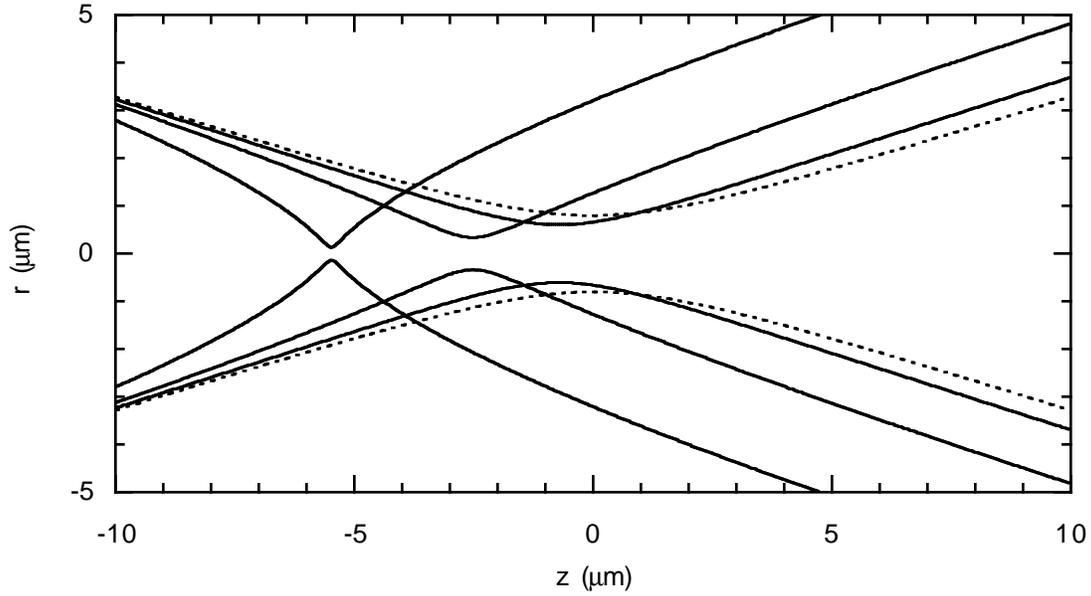


Fig. 12. Effect of self-focusing of the beam profile. Dashed curve: no self-focusing; solid curves:  $s = 0.001$ ,  $0.0033$ , and  $0.01$ .

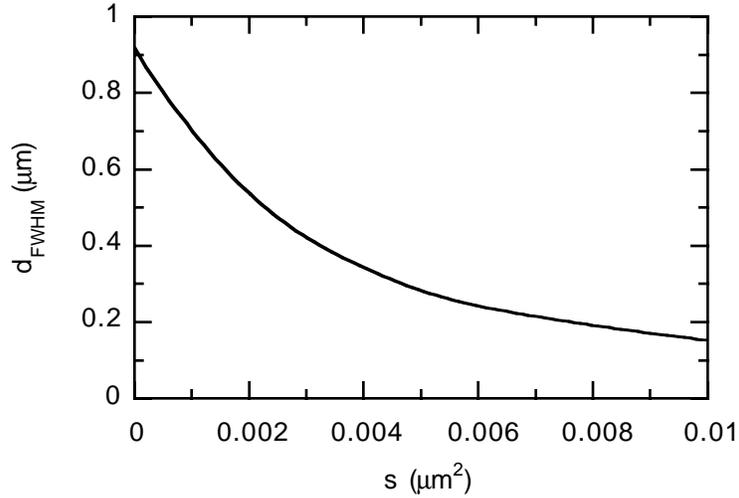


Fig. 13. Beam waist as a function of self-focusing strength.

$$I_0 = (2/\pi) (P/w_0^2),$$

and the corresponding maximum change in the refractive index is

$$\Delta n_{\text{MAX}} = n_2 I_0 = 2 (s/w_0^2).$$

With  $s = 0.01 \mu\text{m}^2$ , this results in  $\Delta n_{\text{MAX}} = 0.9$ , a very large change in the refractive index. Sev-

eral of the approximations in the model are likely to break down before such changes are produced. First, the linear increase of the change in index of refraction with intensity,  $\Delta n = n_2 I$ , probably saturates and may not be valid at such high intensity. Second, Gaussian beam propagation becomes less accurate in predicting the beam profile under extremely tight focusing. The underlying assumption of the slowly varying approximation becomes questionable when  $\Delta n$  changes considerably on the scale of a wavelength. Finally, the temporal profile of the pulse needs to be considered, including the effect of self-phase-modulation, which spectrally broadens the pulse. Despite these limitations the model shows that self-focusing can be significant even under tight external focusing conditions when  $s$  is large enough.

The magnitude of the self-focusing parameter can be estimated for our experimental conditions. For silica,  $n_2$  is about  $3 \times 10^{-16} \text{ cm}^2/\text{W}$  or  $3 \times 10^{-8} \text{ }\mu\text{m}^2/\text{W}$ .<sup>19</sup> A 100-fs, 0.5- $\mu\text{J}$  pulse corresponds to a power of  $5 \times 10^6 \text{ W}$ . Thus, for our experimental conditions  $s$  is about 0.05, or 5 times greater than even the largest value considered in Fig. 13. Clearly, self-focusing cannot be ignored.

### **Plasma: absorption and refraction**

Self-focusing, however, is not the only nonlinear effect that needs to be considered. Self-focusing results from the nonresonant part of the nonlinear susceptibility, is nearly instantaneous, and affects only the phase profile of the pulse. Higher-order processes that span the energy bandgap are resonant, and thus produce real populations in the conduction band. The plasma that is produced by multiphoton absorption contributes to both real and imaginary parts of the index of refraction, and hence can affect both the focusing and absorption of the rest of the pulse.

Which of the two effects of the plasma is dominant? It has been suggested that the defocusing caused by the plasma may be responsible for a much higher threshold observed in the bulk compared to the surface of transparent materials<sup>14</sup>. This observation would imply that the defocusing

occurs before much of the beam is absorbed and the expected intensities are never reached due to the defocusing effect of the plasma. In the following analysis we seek to investigate the dominant role of the plasma under the tight focusing conditions used our experiments, while making as few assumptions as possible about how the plasma is created.

The effect of the plasma is a contribution to the dielectric constant, given by the Drude model,

$$\varepsilon = \varepsilon_0 + \frac{i\omega_p^2\tau}{\omega(1-i\omega\tau)},$$

where  $\varepsilon_0$  is the dielectric constant of the material,  $\omega$  is the frequency of the light,  $\omega_p$  is the plasma frequency, and  $\tau$  is the scattering time. The square of the plasma frequency is proportional to the electron density,  $N$ . The scattering time is an important parameter which is difficult to estimate for this hot plasma and will be discussed below.

The refractive effect of the plasma produces a negative lens and results in a larger beam cross-section at focus. Figure 14 shows the cross-sectional area of the beam at focus as a function of the power in the pulse. The curve is a result of a numerical simulation with a Gaussian beam focused by a 0.65 NA lens, considering only refraction by the plasma (*i.e.*, assuming the decrease in the power of the beam is negligible), a scattering time  $\tau$  of 1 fs. The profile of the plasma is determined by the absorption, which is assumed to be a 5-photon process ( $N \sim I^5$ ) — the lowest order absorption process possible at the 780-nm wavelength. Although other absorption mechanisms may contribute, the results of the analysis are not critically dependent on the order of the nonlinearity. We define a critical power for defocusing,  $P_d$ , as the power at which the cross-sectional area doubles.

Absorption by the plasma reduces the beam intensity. Figure 15 shows the absorbed fraction of the light as a function of the power in the pulse. The curve is calculated by considering only the absorption by the plasma (*i.e.*, neglecting the refractive effect of the plasma) and integrating the

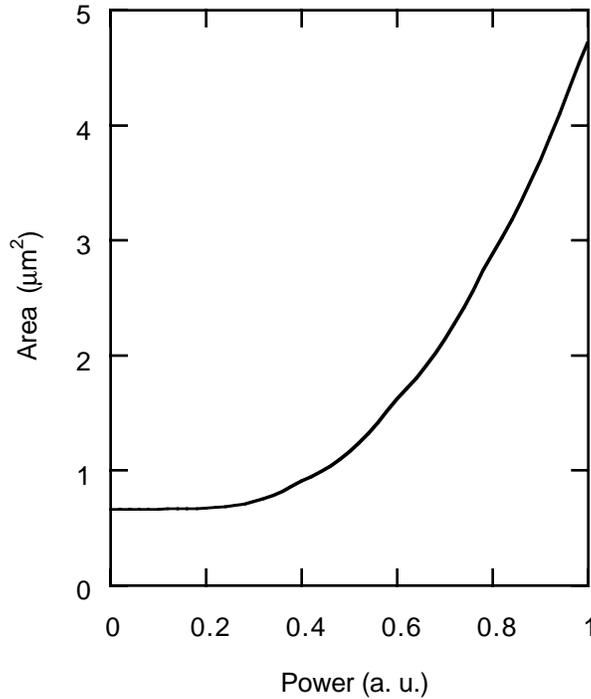


Fig. 14. Defocusing due to the effect of the plasma on the index of refraction

absorption along the beam axis. The focusing conditions and the units of power are the same as in Fig. 14. We define a critical power for the onset of absorption,  $P_{1\%}$ , as the power at which the absorption is 1%, and  $P_{50\%}$  as the power at which the absorption is 50%.

We now compare the critical power for defocusing with the critical power for absorption to see whether refraction or absorption by the plasma dominates. Whichever effect has the smaller critical power should dominate. Since both effects depend on the scattering time, which is not known, we calculate the effects over a wide range of  $\tau$ . Figure 16 shows the ratios  $P_{1\%}/P_d$  and  $P_{50\%}/P_d$  as a function of  $\tau$  over the range of 0.01-100 fs. The slope of the curves indicates that the longer the scattering time, the greater the effect of defocusing relative to absorption, (*i.e.*, the refraction is more important for a longer scattering time). Over the entire range, initial absorption (1%) occurs before any significant defocusing, even if the scattering time is very long (100 fs). If the scattering time is short (less than 0.3 fs), then most of the beam (over 50% on-axis) is

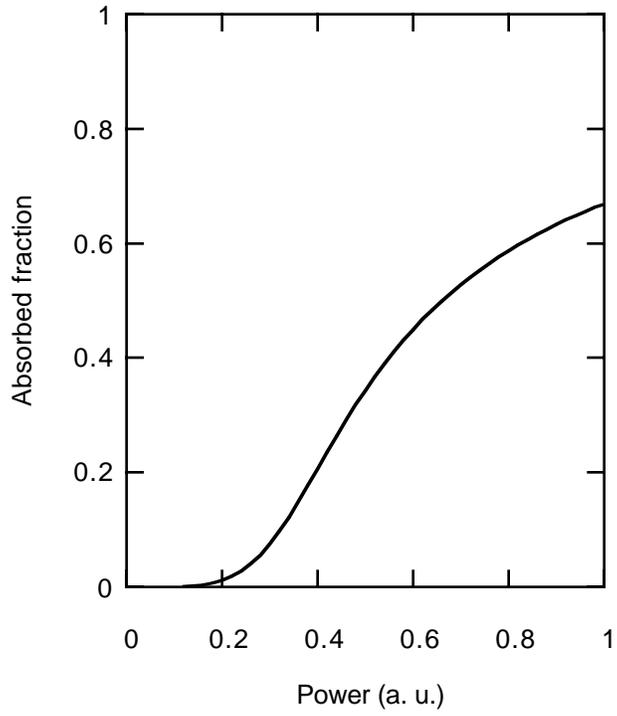


Fig. 15. Plasma absorption along the beam axis.

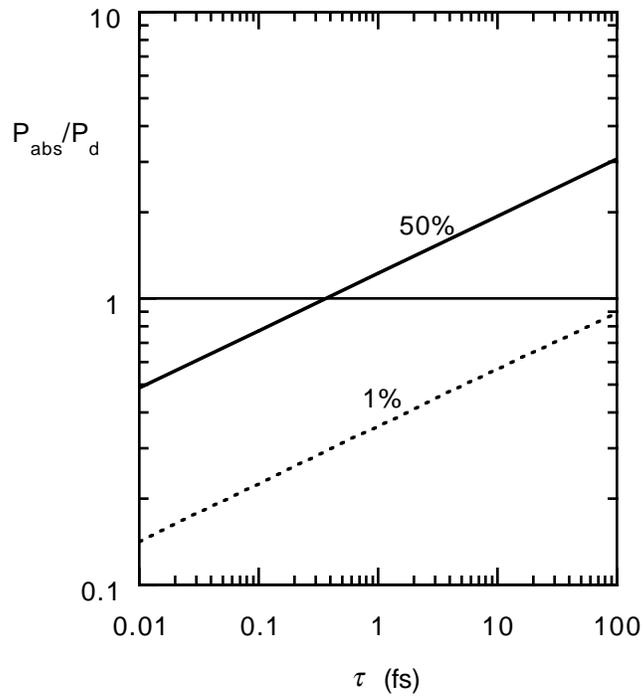


Fig. 16. Ratio of the critical power for defocusing to the critical power for absorption.

absorbed without defocusing playing any role.

The scattering time depends on the temperature and the density of a plasma. For metals,  $\tau$  is typically 1–10 fs at 77 K and 0.1–1 fs at 300 K.<sup>21</sup> Measuring  $\tau$  in the femtosecond-laser-created plasma would provide further insight into the relative roles of absorption and defocusing. For a more detailed model, the evolution of the cross-sectional profile of the pulse and the plasma needs to be considered in time and space (along the direction of propagation).

### **Ionization rate of the plasma**

By considering absorption alone and setting aside the effects of self-focusing for the moment, we can try to estimate the density of the plasma created by the pulse. The intensity at the focus of the 100-fs, 0.5- $\mu\text{J}$  pulse would be about 520 TW/cm<sup>2</sup>. A recent calculation<sup>13</sup> of multiphoton ionization (responsible for initiating an electron avalanche) on the surface of fused silica shows that a 100-fs, 12-TW/cm<sup>2</sup> pulse ionizes a plasma of 10<sup>18</sup> cm<sup>-3</sup> density, corresponding to a characteristic ionization rate of 10<sup>16</sup> cm<sup>-3</sup> fs<sup>-1</sup>. Assuming a five-photon absorption process, yields an ionization rate at an intensity of 520 TW/cm<sup>2</sup> of over 10<sup>24</sup> cm<sup>-3</sup> fs<sup>-1</sup>, enough to ionize all of the valence electrons within a fraction of a femtosecond. Self-focusing creates even higher intensities and thus only increases the ionization rate. With self-focusing, a 0.03- $\mu\text{J}$ , 100-fs pulse should be enough to fully ionize (10<sup>23</sup> cm<sup>-3</sup>) the electrons at the focus. Even without self-focusing, the energy threshold for full ionization would be 0.12  $\mu\text{J}$ . If the effect of the electron avalanche is included it will further lower the predicted threshold. However, we detect no plasma emission up to approximately the 0.3- $\mu\text{J}$  threshold for permanent structural change. This indicates that either the expected intensity is not attained, or the ionization rate is lower than predicted.

The unique conditions created by an ultrashort pulse focused tightly inside transparent materials push nonlinear optics to the extreme. Our estimate of the expected effects of self-focusing,

plasma defocusing, and plasma absorption, shows the limitations of extrapolations based on theories for lower intensities in explaining the energy threshold and voxel size. The created conditions may be beyond the limit of the perturbative approach of nonlinear optics. These experiments probe a unique regime of light propagation inside materials at intensities approaching a Petawatt per square centimeter, the electron ionization that accompanies it, and the material response to the extreme conditions created by the excitation.

### **Material response**

The lattice is heated by the plasma on a timescale of 10 ps or less. If the energy of a 0.5- $\mu\text{J}$  pulse is completely absorbed by a 1- $\mu\text{m}^3$  volume of fused silica (assuming a 0.75-J/gK heat capacity and 2.2-g/cm<sup>3</sup> density), the silica would be heated to an astounding temperature of 300,000 K. Since the volume remains constant in the confined geometry of the microexplosions, the solid values of the heat capacity are probably valid as rough approximations well above the ordinary melting and boiling temperatures because most of energy ordinarily absorbed in a phase transitions goes into a volume expansion.<sup>23</sup>

We have measured the (time-integrated) emission spectrum of the plasma, and found that the peak of the emission spectrum, i.e. the maximum in  $dI/d\nu$ , is at 2.1 eV. The frequency of the peak corresponds to a blackbody temperature of 8,500 K (although the measured spectrum is narrower than would be expected from a blackbody emission source). However, it is certainly possible that the temperature produced in the micro-explosion is much greater, but is not revealed in the emission spectrum. A fully ionized plasma can have a penetration depth as short as 10 nm. Such a plasma would act as an effective radiation blanket, shielding the radiation produced inside the micro-explosion. The light measured in the emission spectrum may be emanating from the surface of the plasma volume, and thus would not carry information about the internal temperature.

(A macroscopic example of this phenomenon is the sun, where the  $6 \times 10^3$  K blackbody spectrum reveals nothing about the  $10^7$  K internal temperature).

We have found that the emission spectrum is entirely independent of the strength of the excitation. By increasing the pulse energy we can raise the intensity of the emission by a factor of more than  $10^3$ , and yet the spectrum of the emission remains unchanged. Furthermore, excitation at a wavelength of 390-nm produces the same emission spectrum as the standard 780-nm pulse. These observations supports the possibility that the emission spectrum is only representative of the surface of the plasma, and that the inside may be much hotter.

The structural observations of the voxels (e.g. Figs. 3 and 4) indicate that a void is formed in the micro-explosion, surrounded by densified material. Such a structure could be the result of immense pressure forcing material from the center of the explosion outward. During cooling, the material apparently does not anneal, and the denser phase is frozen in. Further structural analysis is needed to characterize the densified material.

Very large pressures are needed to compress materials such as quartz and sapphire. Shock-wave experiments provide information about the compressibility of materials at extremely high pressures.<sup>24,25</sup> For example, in fused quartz, a pressure of 6.1 GPa is needed to produce a 10% volume reduction; a 50% reduction requires 35 GPa.<sup>23</sup> In sapphire, a 37-GPa pressure is needed for a 10% volume reduction (based on parameters measured in Ref.21). Extrapolation of these data suggest that a 50% reduction requires 560 GPa, or over 5 million atmospheres, which is beyond the highest pressures achieved in sapphire shock experiments (340 GPa).<sup>21</sup> Further experiments are needed to determine whether pressures of these magnitudes are indeed produced in the micro-explosions.

## Applications

The ultrashort-pulse driven micro-explosions described above provide a unique method for writing information inside transparent materials. Information-bearing patterns can be inscribed in the bulk of virtually any transparent solid object. The features that can be inscribed range from precise macroscopic patterns visible to the eye, down to micron and even sub-micron size structures which are only visible under a high-power microscope with proper illumination. With precise 2-axis or 3-axis translation, practically any image can be inscribed, including binary data, alphanumeric characters, identifying codes (visible or covert), and 3-D objects. The size and visibility of each voxel can be controlled by proper choice of focusing conditions and pulse energy.

A potentially significant application of this method is for high-density 3-D data storage, discussed in more detail in other publications.<sup>1-3</sup> By writing multiple 2-D planes of binary data, storage capacities on the order of 1 Tbit should be possible in a standard CD-size format. Special photosensitive materials are not required — ordinary glass or even plastic can be used as the storage medium, a medium that does not need curing or fixing, and is thermally, chemically, and optically stable and inexpensive. The information is stored permanently, and can only be erased by physically breaking the storage medium, or heating it to the melting or annealing temperature.

The ability to write 3-D objects with sub-micron precision may be useful for creating periodic structures (possibly including photonic bandgap crystals), diffractive optical elements and patterned grating in fibers. The process can also be used for fabricating regions of increased density and hardness in many transparent materials. The range of possible materials is not limited to those transparent to the eye, and includes semiconductors, since the transparency requirement is only that the laser pulse be of lower frequency than the bandgap of the material. The high temperatures and pressures produced in the micro-explosions may lead to novel methods for materials processing.

## Conclusion

Nonlinear absorption of a tightly focused ultrashort laser pulse provides a unique mechanism for rapidly depositing energy in a microscopic volume inside a transparent material. Extreme temperatures and pressures are produced inside the material, which results in a micro-explosion, forcing material out of the center and into the surrounding volume. In contrast to the large, irregular and cracked damage region produced by nanosecond pulses, the structurally altered region produced by ultrashort pulses is highly localized, regular, and shows no signs of cracking. The internal structures created with ultrashort pulses may find use in such applications as high-density 3-D data storage, sub-surface marking and engraving in transparent objects, fabrication of 3-D optical elements, and materials processing.

## References

1. E. N. Glezer, M. Milosavljevic, L. Huang, R. J. Finlay, T.-H. Her, J. P. Callan, and E. Mazur, to be published in *Opt. Lett.*;
2. E. N. Glezer, M. Milosavljevic, L. Huang, R. J. Finlay, T.-H. Her, J. P. Callan, and E. Mazur, in *Proc. Ultrafast Phenomena* (1996);
3. E. N. Glezer, M. Milosavljevic, L. Huang, R. J. Finlay, T.-H. Her, J. P. Callan, and E. Mazur, *OSA Technical Digest*, Vol. 12, *International Symposium on Optical Memory and Optical Data Storage*, Maui, Hawaii (1996).
4. P. D. Maker, R. W. Terhune, C. M. Savage, in *Proc. 3rd Int. Conf. Quantum Electronics*, edited by P. Grivet and N. Bloembergen, (Dunod, Paris 1964).
5. Y. R. Shen, *The Principles of Nonlinear Optics* (John Wiley & Sons, New York, 1984).
6. D. A. Parthenopoulos and P. M. Rentzepis, *Science* **245**, 843 (1989).
7. J. H. Strickler and W. W. Webb, *Opt. Lett.* **16**, 1780 (1991); U.S. Patent #5,289,407 (1994).
8. Y. Kawata, H. Ueki, Y. Hashimoto, and S. Kawata, *Appl. Opt.* **34**, 4105 (1995).
9. H. Ueki, Y. Kawata, and S. Kawata, *Appl. Opt.* **35**, 2457 (1996).
10. J. Ihlemann, B. Wolff, and P. Simon, *Appl. Phys. A* **54**, 363 (1992).
11. D. Du, X. Liu, G. Korn, J. Squier, and G. Mourou, *Appl. Phys. Lett.* **64**, 3071 (1994).
12. B. C. Stuart, M. D. Feit, S. Herman, A. M. Rubenchik, B. W. Shore, and M. D. Perry, *Phys. Rev. Lett.* **74**, 2248 (1995);
13. B. C. Stuart, M. D. Feit, S. Herman, A. M. Rubenchik, B. W. Shore, and M. D. Perry, *J. Opt. Soc. Am. B*, **13**, 459 (1996).
14. D. von der Linde and H. Schüler, *J. Opt. Soc. Am. B*, **13**, 216 (1996).
15. L. Huang, E. N. Glezer, J. P. Callan, and E. Mazur, *Proc. of Quant. Elect. and Laser Science*, Anaheim, CA (1996).

16. M. J. Soileau, W. E. Williams, N. Mansour, *Opt. Engineering* **28**, 1133, (1989).
17. S. Deliwala, Ph.D. Thesis, Harvard University, 1995.
18. D. Strickland and G. Mourou, *Opt. Comm.* **56**, 219 (1985).
19. T. Erdogan, OSA Technical Digest, Vol. 11, *Nonlinear Optics: Materials, Fundamentals and Applications*, Maui, Hawaii (1996).
20. J. S. Aitchison, Nonlinear Glasses, in Principles and Appl. of Nonl. Optical Mat.
21. D. Erskine, High-Pressure Science and Technology - 1993, Colorado Springs, CO. AIP Conference Proceedings **309**, 141 (1994).
22. N. W. Ashcroft , N. D. Mermin, *Solid State Physics* (Saunders College, Philadelphia, 1976).
23. R. W. Hopper and D. R. Uhlmann, *Journal of Applied Physics* **41**, 4023 (1970).
24. *LASL Shock Hugoniot Data*, edited by S. P. Marsh (University of Calif. Press, Berkeley, 1980).
25. I.-Y. S. Lee, J. R. Hill, D. D. Dlott, *J. Appl. Phys.* **75**, 4975 (1994).